ОБЗОРНАЯ СТАТЬЯ

УДК 681.324

ИНТЕГРИРОВАННЫЕ БОРТОВЫЕ ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ: ОБЗОР СОВРЕМЕННОГО СОСТОЯНИЯ И АНАЛИЗ ПЕРСПЕКТИВ РАЗВИТИЯ В АВИАЦИОННОМ ПРИБОРОСТРОЕНИИ

П.П. Парамонов, И.О. Жаринов

Парамонов Павел Павлович, доктор технических наук, профессор, генеральный директор Федерального государственного унитарного предприятия «Санкт-Петербургское опытно-конструкторское бюро «Электроавтоматика» им. П.А. Ефимова», профессор базовой кафедры машинного проектирования бортовой электронно-вычислительной аппаратуры НИУ ИТМО. Участник создания самолетов МиГ-23, МиГ-25, МиГ-27, МиГ-29, Су-25Т, Су-27, Су-80ГП, Ту-154, Ту-144, Ил-62, Ил-86 и др., а также вертолетов Ми-35, Ми-24К, Ка-50 («Черная акула»). Действительный член Академии навигации и управления движением, действительный член Петровской академии наук и искусств. Автор более 170 опубликованных работ, из которых 6 монографий, 30 учебно-методических работ, более 25 патентов. Почетный авиастроитель. Почетный работник высшего профессионального образования Российской Федерации. Заслуженный конструктор Российской Федерации. Отличник качества авиационной промышленности. Почетный гражданин Кировского района Санкт-Петербурга. Лауреат премии им. С.И. Мосина. Член редколлегии журнала «Научно-технический вестник информационных технологий, механики и оптики».

Жаринов Игорь Олегович, доктор технических наук, доцент, руководитель учебно-научного центра Федерального государственного унитарного предприятия «Санкт-Петербургское Опытно-конструкторское бюро «Электроавтоматика» им. П.А. Ефимова», заведующий базовой кафедрой машинного проектирования бортовой электронно-вычислительной аппаратуры НИУ ИТМО. Ведущий специалист в области теории и основ построения информационно-управляющего поля кабины пилотируемых летательных аппаратов, эргономической поддержки, разработки и эксплуатации авиационных систем. Участник создания бортовой аппаратуры для объектов Су-80ГП, Су-25СМ, Су-25УБМ и др. Специалист в области автоматизации проектирования бортового оборудования. Специалист в области компьютерной обработки электрофизиологических данных в медицинском приборостроении. Автор более 90 опубликованных работ, из которых 6 монографий, 15 учебных пособий, 6 отраслевых авторских свидетельств на программные продукты. Член диссертационного совета при НИУ ИТМО.

Приведен обзор современного состояния и анализ перспектив развития интегрированных бортовых вычислительных систем, применяемых в авиационном приборостроении. Основное внимание уделено проектам, выполненным в рамках концепции интегрированной модульной авионики. Подробно рассмотрены иерархические уровни проектирования модулей, крейтов (бортовых систем) и авиационных комплексов. Приведены примеры существующих отечественных и зарубежных разработок и их краткие технические характеристики, а также обширная библиография по тематике.

Ключевые слова: интегрированная модульная авионика, крейты, современное состояние, обзор.

Введение

Проблема проектирования перспективного бортового радиоэлектронного оборудования (БРЭО) относится сегодня к приоритетным направлениям развития авиационной промышленности Российской Федерации. Основные научные исследования в этой области осуществляются специалистами авиаприборостроения в рамках поиска новых концепций проектирования аппаратного [1–7] и программного [8, 9] обеспечения и в рамках внедрения новых технологий, материалов и элементной базы [10, 11] в образцы авиационной продукции. Поиск новых концепций проектирования аппаратного и программного обеспечения направлен на совершенствование структурной организации бортовых систем с целью удовлетворения непрерывно возрастающих требований:

– по обеспечению регулярности, безопасности и комфортабельности полетов летательных аппаратов;
– по обеспечению экономичности владения и удобства обслуживания летательных аппаратов (ЛА);
– по обеспечению заданной точности расчетов пилотажно-навигационных параметров;
– по обеспечению заданных надежностных характеристик приборного оборудования и т.д.

Внедрение новых технологий, материалов и элементной базы в образцы авиационной продукции направлено на изменение технико-экономических показателей БРЭО с целью придания бортовым системам новых качеств или с целью совершенствования количественных показателей существующей аппаратуры. На современном этапе разработчиками авионики достигнуты определенные успехи в решении частных задач проектирования отдельных компонентов за счет внедрения в разработки перспективной технологии функциональной, программной и аппаратной интеграции оборудования, однако проблема построения интегрированных вычислительных комплексов БРЭО остается нерешенной в полной мере до настоящего времени. Целью данной статьи является представление широкому кругу специалистов проблематики проектирования интегрированных вычислительных систем, составляющих вычислительное ядро перспективных комплексов БРЭО. Читателю будет представлен краткий обзор достигнутых сегодня разработчиками результатов и предложен для более подробного знакомства современный подход к проектированию перспективного БРЭО – подход на основе внедрения на борт ЛА структур интегрированной модульной авионики.

Обзор современных направлений проектирования вычислительных систем в авиационном приборостроении

Основные направления проектирования вычислительных систем в современном авиационном приборостроении реализуются в Российской Федерации в рамках следующих федеральных целевых государственных программ:

- «Развитие гражданской авиационной техники России на 2002–2010 г.г. и на период до 2015 г.»; выполняется в рамках постановления Правительства РФ от 15 октября 2001 года № 728: «О федеральной целевой программе развития авиационной техники России на 2002–2010 г.г. и на период до 2015 г.»;
- «Модернизация и развитие интегрированных комплексов и систем цифрового оборудования самолетов гражданской авиации России на период 2003–2010 г.г.» (АВИОНИКА–2010);
- «Развитие электронной компонентной базы и радиоэлектроники на 2008–2015 г.г.»; выполняется в рамках постановления Правительства РФ от 23 июня 2007 г. № 972-р. и др.

Соисполнителями программ выступают ОАО «ОКБ Сухого», ОАО «РСК «МиГ», ОАО «ММПП «Салют», ОАО «КнААПО», ОАО «РПКБ «Раменское», ФГУП «СПб ОКБ «Электроавтоматика» им. П.А. Ефимова», ОАО «Авиаприбор-Холдинг», ВНИИРА, ОАО «МНПК «Авионика», ОАО «Измеритель», ОАО «Научно-производственный комплекс «ЭЛАРА» им. Г.А. Ильенко», ФГУП «Уфимское приборостроительное производственное объединение», ОАО «ОКБ им. А.И. Микояна», ОАО «ОКБ им. А.С. Яковлева», ОАО «Туполев», ЦАГИ, КБ «Сигнал», ЛИ и ДБ им. М.М. Громова, ФГУП «НИИ аэронавигации», ОАО «Нижегородский авиастроительный завод «Сокол», ФГУП «Государственный Рязанский приборостроительный завод» и др. За рубежом перспективные исследования разработчиков в области авиаприборостроения осуществляются [12–16] под эгидой объединенного комитета НАТО по стандартизации архитектуры авиационных комплексов Allied Standard Avionics Architecture Council (ASAAC) в рамках программ создания новых концепций построения БРЭО: Digital Avionics Information System (DAIS), Pave Pillar / F22 (Raptor), Pave Pace / JSF (Joint Strike Fighter – F35), Module Avionics System Architecture (MASA), Integrated Sensor System (ISS), Standard Avionic Module (SAM), Advanced Avionics Architecture and Packing (AAAP) и др. Соисполнителями программ выступают компании Airbus, Bell Helicopter Textron, Boeing, General Motors, Rockwell Int., GTE Government Systems, American Airlines, United Airlines, Lockheed Martin, Tokyo Electric Power, European Aeronautic Defense and Space Company, European Association of Aerospace Construction и др.

Основным интернациональным результатом работы специалистов в области разработки архитектуры и компонентов БРЭО сегодня является утвержденная фирмой ARINC (Aeronautical Radio Inc., США) и развиваемая отечественными специалистами [17–27] концепция интеграции бортовой аппаратуры – концепция IMA (Integrated Modular Avionic), основы которой изложены в стандарте ARINC 651 «Design Guidance for Integrated Modular Avionic». Стандарт определил новое направление в совершенствовании авиационного приборостроения и наметил перспективы развития авиационной промышленности на десятилетия вперед.

ARINC 651 представляет собой согласованное и утвержденное законодательно в США мнение представителей авиакомпаний-эксплуатантов ЛА, фирм-производителей самолетов и авиационной техники США и ряда других стран по вопросу систематизации подхода к проектированию авионики. В Российской Федерации стандарт ARINC 651 носит информационный характер. Стандарт ARINC 651 явился основой серии документов ARINC, посвященных IMA (в частности, ARINC 624, 629, 636, 637, 650, 652, 653, 659 и др.). По мнению зарубежных специалистов [14–16], в перспективе аппаратура IMA должна заменить БРЭО, разработанное ранее в соответствии с требованиями стандартов 700-ой серии ARINC.

Концепция IMA охватывает следующие группы проектных задач авиаприборостроения [18]:

- цели создания и внедрения IMA в разработки БРЭО;
- используемые при разработке БРЭО технологии;

- реализация свойства отказоустойчивости бортового оборудования;
- принципы построения бортовой сети информационного обмена (бортовой сети данных);
- архитектура БРЭО;
- архитектура программного обеспечения БРЭО и его компонентов;
- сертификация разработок (аппаратного и программного обеспечения БРЭО);
- контролепригодность и ремонтопригодность БРЭО и объекта – ЛА – в целом;
- выбор источников и потребителей данных (абоненты бортовой локальной сети данных).

Концепция IMA предполагает разделение функциональных компонентов авиационной электроники на три иерархических уровня [28–32]:

- нижний уровень иерархии образуют унифицированные конструктивно-функциональные модули (КФМ) различного назначения, имеющие собственные вычислительные средства в компактном стандартизованном исполнении (например, по стандартам евромеханики с типоразмерами печатных плат 3U, 6U, 9U);
- средний уровень иерархии образуют мультипроцессорные вычислительные системы (крейты), создаваемые из модулей нижнего уровня и конструктивно выполненные в стандартизованном корпусе (в различных источниках литературы встречаются различные названия конструктива – крейт, стойка, шкаф, бокс, шасси);
- высший уровень иерархии представляет собой бортовую локальную вычислительную сеть на основе центрального сетевого интерфейса высокой пропускной способности, интегрирующую вычислительные средства крейтов среднего уровня.

Для организации бортового центрального сетевого интерфейса высокой пропускной способности сегодня специалистами авиаприборостроительных предприятий прорабатываются конструктивно-технологические решения по внедрению на авиационных и космических ЛА новых или уже существующих [33–38]:

- межсистемных интерфейсов Fiber Channel (ANSI X3T11); Scalable Coherent Interface (ANSI / IEEE Std 1596-1992); Myrinet; Gigabit Ethernet; ARINC 664; Asynchronous Transfer Mode; FireWire (IEEE 1394) и др.;
- внутрисистемных интерфейсов Scalable Coherent Interface – SCI (ANSI / IEEE Std 1596-1992); Fiber Channel – ANSI X3T11; VME; PCI Compact; SKY Channel Packet Bus; LVDS (TIA / EIA 644) и др.

Результаты сравнительного анализа различных принципов построения БРЭО [39–42] и основные ожидаемые технико-экономические преимущества, которые обеспечиваются введением в состав БРЭО аппаратуры, разработанной по концепции IMA, приведены в таблице.

<table>
<tr><th rowspan="2">Критерий или свойство</th><th colspan="3">Принцип построения архитектуры авиационного комплекса БРЭО</th></tr>
<tr><th>Независимый</th><th>Федеративный</th><th>Интегрированный (IMA)</th></tr>
<tr><td>Эволюционная совместимость разработок</td><td colspan="2">Используются специализированные проектные решения, разрабатываемые специально под каждое конкретное применение в составе летательного аппарата. Возможно также использование стандартизованных (готовых, заимствованных) проектных решений, если они функционально и технически совместимы с решаемыми комплексом БРЭО задачами.</td><td>Используется специализированная технология для модулей IMA, ориентированная на выполнение высоконадежных функции ЛА. Использование коммерческих технологий возможно только для реализации второстепенных функций ЛА.</td></tr>
<tr><td>Типы используемых бортовых сетевых решений (бортовая информационная сеть)</td><td colspan="3">Применяется вся номенклатура аналоговых и цифровых линий связи и шин передачи информации. Исторически первые поколения комплексов БРЭО базировались на аналоговых каналах связи. В современных комплексах БРЭО используются высокоскоростные цифровые каналы связи (в различных комплексах реализуются различные форматы протоколов передачи данных).</td></tr>
<tr><td rowspan="2">Отказоустойчивость авионики</td><td>Взаимное влияние отказа, возникшего в одном комплексе БРЭО, на системы другого комплекса исключается вследствие их независимости. Выявление отказа диагностируется путем использования встроенных средств аппаратно-программного контроля в отказавшем комплексе.</td><td>Имеет место распространение отказа в системе одного комплекса БРЭО на системы другого комплекса вследствие взаимного обмена данными. Диагностирование отказа реализуется автоматически средствами аппаратно-программного контроля обоих комплексов.</td><td>Системы, реализующие различные по категории важности бортовые задачи, обнаруживают и парируют отказы за счет использования процедур динамического назначения задач на имеющиеся в БРЭО работоспособные ресурсы. Для реализации отказоустойчивости применяются механизмы программного контроля и защиты информации.</td></tr>
<tr><td colspan="3">Избыточность аппаратных средств БРЭО обеспечивает защиту от случайных ошибок (сбоев) в измерениях. Избыточность программных средств БРЭО обеспечивает защиту от систематических ошибок и погрешностей вычислений.</td></tr>
</table>

Продолжение таблицы

<table>
<tr><th rowspan="2">Критерий или свойство</th><th colspan="3">Принцип построения архитектуры авиационного комплекса БРЭО</th></tr>
<tr><th>Независимый</th><th>Федеративный</th><th>Интегрированный (IMA)</th></tr>
<tr><td>Ремонтопригодность авионики</td><td colspan="2">Ремонтно-восстановительные работы проводятся на специализированных авиаремонтных заводах (АРЗ), где сосредоточены запасы и гарантийные технические аптечки исправной бортовой аппаратуры. При наличии нескольких АРЗ запасы аппаратуры должны находиться в каждом АРЗ в достаточном объеме, что увеличивает стоимость эксплуатации авионики. При использовании специализированных аппаратных средств на различных авиационных объектах ремонтопригодность БРЭО зависит от логистической поддержки обеспечения складов кондиционной продукцией.</td><td>Единая стандартизированная платформа авионики и единая номенклатура модулей IMA существенно упрощают техническое обслуживание объектов различного типа. В каждом АРЗ достаточно иметь ограниченный объем запаса унифицированных модулей IMA для обслуживания большого парка объектов.</td></tr>
<tr><td>Модернизация авионики</td><td>Модернизация аппаратных средств или прикладного программного обеспечения затрагивает только одну соответствующую систему. После изменения документации только в одной системе требуется повторная сертификация системы (летные испытания объекта).</td><td>Модернизация может производиться на уровне одной системы или одного комплекса БРЭО. После внесения изменений в документацию системы (комплекса) требуется проведение повторной сертификации разработки.</td><td>Модернизация модулей IMA не сказывается на прикладном программном обеспечении вследствие его независимости от аппаратной платформы. Изменениям подвергаются только встроенные средства аппаратно-программного контроля БРЭО с сохранением программного кода функциональных бортовых задач ЛА.</td></tr>
<tr><td rowspan="2">Потребность аппаратуры в системе охлаждения</td><td colspan="3">Каждая подсистема БРЭО способствует увеличению суммарного тепловыделения в составе объекта, в связи с чем требуется кондиционирование аппаратуры средствами охлаждения. Анализ объемов тепловыделения показывает, что для систем одного и того же функционального назначения комплексы IMA требуют заметно меньшего объема охлаждающего воздуха по сравнению с независимым или федеративным принципом построения авиационных комплексов.</td></tr>
<tr><td colspan="2">Для ослабления температурного режима работы электрорадиоэлементов наиболее важных систем БРЭО требуется постоянное применение средств охлаждения (встроенных или централизованных).</td><td>Так как крейты IMA интегрируют различную по критичности аппаратуру БРЭО, требуется постоянный обдув электрорадиоэлементов.</td></tr>
<tr><td>Время, стоимость и ответственность за разработку, сертификацию и комплексирование авионики</td><td colspan="2">Разработку и сопровождение средств авионики осуществляют несколько независимых фирм-производителей, каждая из которых обеспечивает последующее гарантийное и послегарантийное обслуживание своей аппаратуры. Применение стандартизованных функциональных элементов, включая применение коммерческих технологий, способствует переходу аппаратуры к принципам построения открытой архитектуры авионики и к унификации компонентов различных производителей.
Синергетические эффекты между различными поставщиками бортовых систем исключены: разработка, настройка, проверка и сертификация БРЭО выполняются каждым производителем независимо друг от друга, в связи с чем стоимость или время создания БРЭО в масштабах объекта не могут быть минимизированы.
Вопросы системной интеграции и комплексирования аппаратуры реализуются в составе объекта (или стенда отработки) после получения всех подсистем и находятся в компетенции главного (генерального) конструктора, ответственного за разработку объекта (летательного аппарата) в целом.</td><td>Производители БРЭО ориентированы на создание модулей IMA, совместимых с существующими авиационными стандартами на программном и аппаратном уровне.
Тестирование и сертификация каждого типа модуля IMA осуществляется производителем и не зависит от типа летательного аппарата. Порядок разработки и сертификации регламентируется соответствующими авиационными стандартами.
Системная интеграция и комплексирование БРЭО проводятся в едином технологическом цикле разработки на головном предприятии, поскольку прикладное программное обеспечение различных систем проектируется на базе единой аппаратно-программной платформы.</td></tr>
</table>

Продолжение таблицы

Критерий или свойство	Принцип построения архитектуры авиационного комплекса БРЭО		
	Независимый	Федеративный	Интегрированный (IMA)
Габаритные размеры, энергопотребление	Каждая система устанавливается в составе объекта индивидуально и является автономным потребителем бортовой электроэнергии соответствующей категории. Аппаратура первой категории – подсистемы БРЭО – является наиболее важной для выполнения летательным аппаратом бортовых задач (взлет, посадка).	Каждая система устанавливается в составе объекта индивидуально и является автономным потребителем бортовой электроэнергии. За счет использования принципа комплексной обработки информации объемы аппаратных средств БРЭО могут быть снижены.	Несколько различных систем объединены в составе единого конструктива. Используется принцип экономии габаритных размеров аппаратуры за счет максимального приближения источников и приемников бортовой информации друг к другу, реализуется свойство интеграции бортовой аппаратуры. На программном уровне регулируется объем задействованной аппаратуры БРЭО в зависимости от полетной ситуации.
Наращивание числа функциональных задач БРЭО, программирование БРЭО	Наращивание функциональных задач БРЭО реализуется путем изменения прикладного программного обеспечения и (или) введением дополнительных линий передачи информации между комплексами БРЭО. Доработка программного обеспечения, как правило, приводит к реализации свойства информационной избыточности БРЭО. Оба способа наращивания задач реализуются в пределах одной системы (комплекса). Могут изменяться масса, размеры, энергопотребление аппаратуры БРЭО. Если в состав БРЭО вводится дополнительно новая система, может потребоваться также введение новых специальных устройств (датчиков) – источников и приемников пилотажно-навигационной информации.	Наращивание функциональных задач БРЭО реализуется путем изменения прикладного программного обеспечения и (или) интеграцией с другими системами в единый комплекс. Доработка программ, как правило, приводит к реализации свойства информационной избыточности БРЭО. Второй способ затрагивает каналы передачи информации между системами, в связи с чем усложняется архитектура БРЭО. Может также потребоваться пересертификация БРЭО. Если в состав БРЭО вводится дополнительно новая система, может потребоваться также введение специальных устройств (датчиков) – источников и приемников пилотажно-навигационной информации.	Программные приложения, реализующие новые функции БРЭО, могут быть интегрированы в состав исходного кода бортовой задачи за счет использования модульного принципа построения программного обеспечения. При наличии в циклограмме бортовой задачи достаточного временного запаса интеграция новых фрагментов программного обеспечения реализуется без доработки уже существующих приложений и их программных кодов. Пересертификация существующих фрагментов программного обеспечения не требуется, сертифицируются только новые компоненты программного кода. Введение дополнительной функциональной бортовой задачи также может потребовать изменения топологии схемы связи БРЭО, прокладки дополнительных линий передачи информации в составе ЛА и пр.
Масса аппаратуры БРЭО	Масса аппаратуры БРЭО зависит от числа задействованных авиационных приборов и систем. Анализ различных принципов построения БРЭО для систем одного и того же функционального назначения показывает, что комплексы IMA нуждается в значительно меньшем количестве аппаратных средств, что уменьшает общую массу БРЭО по сравнению с независимым или федеративным принципом построения авиационных комплексов.		
	Соединение бортовой аппаратуры по принципу точка-точка обусловливает значительный вес бортовой кабельной сети за счет наличия прямых и обратных проводников каналов передачи информации.	Используются мультиплексированные линии связи между бортовыми источниками и приемниками информации, а также элементы высокой степени интеграции, за счет чего сокращается суммарный вес кабельной сети.	Используются высокоскоростные мультиплексированные линии связи, включая оптоволоконный канал передачи данных, между бортовыми источниками и приемниками информации, а также «системы на кристалле», за счет чего сокращается суммарный вес бортовой кабельной сети.

Продолжение таблицы

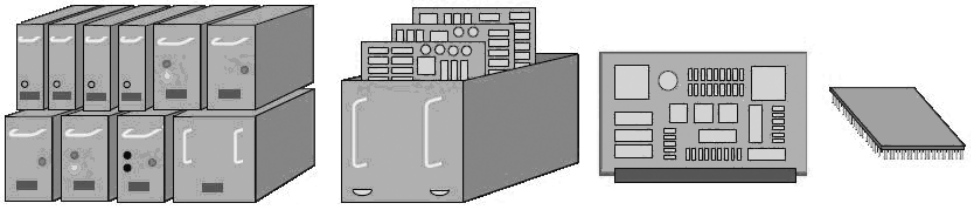
Критерий или свойство	Принцип построения архитектуры авиационного комплекса БРЭО		
	Независимый	Федеративный	Интегрированный (IMA)
			

Таблица. Сравнительные характеристики различных принципов построения авиационных комплексов

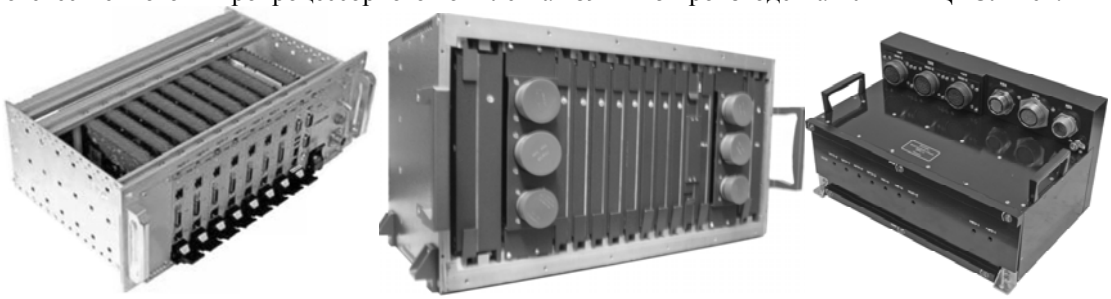
Обзор патентно-информационных ресурсов в области существующих разработок IMA

На фоне глобализации рынка авиационной продукции и услуг, расширения партнерских отношений российских предприятий с зарубежными компаниями и перспектив совместных работ в области реализации программ создания новых самолетов гражданской авиации представляется возможным использовать основные положения стандарта ARINC 651 в порядке методических рекомендаций для отечественных специалистов. Совместный опыт проектирования для авиастроительных предприятий Российской Федерации является, безусловно, положительным, поскольку позволяет интегрировать концептуальные российские идеи по разработке бортовой аппаратуры с существующим зарубежным научно-практическим заделом. Пример [43] такого международного сотрудничества – взаимодействие ФГУП «НИИ авиационного оборудования» с корпорацией Allied Signal (США), в результате которого был разработан интегрированный комплекс бортовой авионики для самолета-амфибии Бе-200.

Отечественные результаты разработок специалистов авиаприборостроительной отрасли в области структур IMA реализуются сегодня рядом предприятий Российской Федерации, в том числе ВНИИРА, ФГУП «СПб ОКБ «Электроавтоматика» им. П.А. Ефимова», МИЭА, НПО «Полет» и др., в частности, в рамках [18, 28, 42]:

– НИОКР «ИКБО-95МС»;
– НИОКР «Создание технологии проектирования и высокотехнологического производства аппаратно-программных компонентов информационно-вычислительной среды и периферийных средств для комплексов авиационного оборудования нового поколения на основе концепции IMA и необслуживаемого БРЭО»,

Специалистами проработаны конструкторско-технологические решения по разработке приемопередающих быстросменных КФМ для крейтов бортовой интегрированной радиотехнической системы, интегрированной системы радиосвязи и системы управления ЛА, КФМ для бортовой цифровой вычислительной системы. Среди примеров известных сегодня отечественных проектов бортового приборного оборудования, разработанного в соответствии с основными положениями концепции IMA, можно отметить (рис. 1, а) многопроцессорный вычислительный комплекс «МПВК» [30, 44] производительностью 25GFLOPS разработки ЗАО НПП «Авиационная и морская электроника». Комплекс «МПВК» выполнен на базе унифицированных одноплатных модулей-вычислителей серии CPC502 (до восьми модулей в едином специализированном конструктиве), изготавливаемых ЗАО НПФ «Доломант» по документации и под авторским надзором компании Fastwel в стандарте CompactPCI 3U. Комплекс «МПВК» по эксплуатационным свойствам соответствует гр. 2.1.2 ГОСТ РВ 20.39.304-98, ГОСТ РВ 20.39.305-98. В ближайшее время в рамках программы по импортозамещению элементной базы специалисты ЗАО НПП «Авиационная и морская электроника» в проекте «МПВК» предполагают переход на модули-вычислители на базе отечественного микропроцессорного комплекта 1892ВМ5 производства ГУП НПЦ «Элвис».



а б в

Рис. 1. Бортовые цифровые вычислительные системы структуры IMA: разработки ЗАО НПП «Авиационная и Морская Электроника» (а); разработки ФГУП «СПб ОКБ «Электроавтоматика» им. П. А. Ефимова» (б); разработки ОАО «РПКБ «Раменское» (в) (примеры)

Другой пример [28, 42] удачной разработки бортовой вычислительной системы в классе структур IMA – изделие «Крейт» (рис. 1, б), спроектированное и изготовленное в ФГУП «СПб ОКБ «Электроавтоматика» им. П.А. Ефимова» на базе модулей стандарта 6U в исполнении по гр.3.1 ГОСТ РВ 20.39.304-98, ГОСТ РВ 20.39.305-98. Вычислительная система «Крейт» является потребителем электроэнергии от бортовой сети постоянного и переменного тока с качеством и параметрами по ГОСТ 19705-89. В состав «Крейт» входят 10 модулей процессоров, обеспечивающих комплексную обработку пилотажно-навигационной информации и обмен по ГОСТ 18977-79 и РТМ 1495-75, ГОСТ Р 52070-2003, ARINC 664, Fiber Channel. Архитектура системы «Крейт» выполнена по двухконтурному типу в едином конструктиве и подробно рассмотрена в работах [45–47].

Практический интерес представляет также бортовая вычислительная система «БВС-1» [48], разработанная в ОАО «РПКБ «Раменское» в классе структур IMA и предназначенная для выполнения вычислительных и управляющих функций в составе информационно-управляющих систем перспективных комплексов бортового оборудования военной и гражданской авиации. Система «БВС-1» (рис. 1, в) обеспечивает резервированную полнодуплексную связь на основе последовательного высокоскоростного интерфейса коммутируемой архитектуры ARINC 664 между установленными в нее интеллектуальными электронными модулями (процессорные модули общего назначения, модули цифровой обработки сигналов и т.п.), а также связь с другими системами комплекса БРЭО в составе ЛА. Система «БВС-1» обеспечивает прием видеоинформации, формирование и наложение графических изображений и выдачу результирующей синтезированной видеоинформации на средства индикации информационно-управляющего поля кабины пилота ЛА. Процессорные модули и модули коммутатора сети выполнены в формате VPX-REDI (VITA 46, VITA 48) с типоразмером печатной платы 3U. На процессорные модули устанавливаются мезонинные модули графического контроллера или модуля массовой памяти. Суммарное быстродействие модулей-процессоров общего назначения (модуль выполнен на базе микропроцессора MPC8548 фирмы Freescale, тактовая частота работы 1–1,5 ГГц) – не менее 12000 MIPS. Объем оперативной памяти процессорных модулей: 3 Гбайт, объем постоянной памяти процессорных модулей: 1,5 Гбайт. Объем ПЗУ модуля массовой памяти: 8 Гбайт. Система «БВС-1» относится к классу приемников электроэнергии первой категории по ГОСТ 19705-89, раздел 3. Внешние интерфейсы системы «БВС-1» – ARINC 664 (Gigabit Ethernet 1000Base-SX), ARINC 818 (Fiber Channel 1x). Для информационного обмена с абонентами БРЭО используется многомодовый волоконно-оптический кабель, рабочая длина волны 850 нм.

Помимо указанных организаций, разработкой высокопроизводительных многопроцессорных вычислительных систем в классе структур IMA для авиационной промышленности в настоящее время в Российской Федерации занимаются такие известные предприятия, как ФГУП «Государственный научно-исследовательский институт авиационных систем» [49, 50], ОАО «Научно-конструкторское бюро вычислительных систем» (проект «Базис 5.0» [22]), ОАО «НИЦЭВТ» (работы в рамках НИР «Беркут», проект «Ястреб», проект «Ангара» [51]), ЗАО НТЦ «Модуль» (проект «Борт-М», проект «БПТС»), ОАО «НИИ ВК им. М.А. Карцева», ЗАО «МЦСТ» (проект «Эльбрус-90микро»), КБ «Корунд-М» (проект «Багет») и др. Ряд разработок, выполненных отечественными и зарубежными специалистами в области реконфигурируемых структур IMA, защищен патентами [47, 49, 50, 52, 53].

Для исследования и оптимизации проектных решений по отработке перспективных архитектур вычислительных систем, в частности, и авионики в целом учеными и техническими специалистами разрабатываются специализированные системы автоматизированного проектирования (САПР) [54–58]: GSSS (Generalized Structure Synthesis System) разработки Московского энергетического института (технического университета), САПР серии «САФАК» разработки ЦНИИ 30 МО РФ и др. Согласно системе классификации Флинна, большинство отечественных разработок в области построения многомодульных вычислительных систем относится к классу многопроцессорных систем с множественным потоком команд и множественным потоком данных.

Среди зарубежных проектов авиационной электроники [12–14, 43], выполненных в соответствии с основными положениями концепции IMA, заслуживают внимание комплексы ProLine 21, Primus Epic, Topdeck, включающие от 2 до 4 стандартных вычислительных крейтов IMA.

Интегрированный комплекс модульной авионики ProLine 21 разработки Rockwell Collins был впервые представлен в США в 1995 г. Комплекс предназначен для решения задач навигации и управления полетом гражданских транспортных самолетов бизнес-класса, используемых на местных воздушных авиалиниях, а также для современных вертолетов. Сегодня комплекс ProLine 21 устанавливается на реактивном самолете Premier I, на транспортных самолетах бизнес-класса Continental, Hawker 800XP, CitationJet моделей CJ1 и CJ2, на конвертоплане BA-609 и др. Внешний вид одного крейта IMA разработки Rockwell Collins и пример модуля IMA разработки National Aerospace Laboratory [15] показаны на рис. 2. Аналогичный по назначению комплекс БРЭО разработки Rockwell Collins для военных самолетов получил название Flight-2 и устанавливается сегодня на модернизируемых самолетах ВВС и ВМФ США: топливозаправщиках KC-135, транспортных самолетах C-130 и самолетах радиолокационного дозора P-3 Orion, а также на вертолетах Sikorsky моделей S-92, S-70.

Primus Epic – название бортового комплекса в классе структур IMA разработки компании Honeywell. Комплекс Primus Epic обладает открытой архитектурой, достаточной для масштабирования пилотажно-навигационного комплекса под любой тип гражданских ЛА – от маленьких турбовинтовых самолетов и вертолетов до реактивных лайнеров с числом перевозимых пассажиров более 100. Комплекс Primus Epic сегодня устанавливается на гражданских самолетах бизнес-класса Hawker Horizon, Hawker 450, Citation Sovereign, на вертолете AB139 и на реактивных самолетах семейства Fairchild–Dornier моделей 528JET / 728JET / 928JET, а также самолетах семейства Embraer моделей ERJ-170 / ERJ-190.

а б

Рис. 2. Крейт модульной авионики разработки Rockwell Collins (а) и модуль IMA разработки National Aerospace Laboratory (Голландия) (б)

Помимо бортовых авиационных комплексов ProLine 21 и Primus Epic, практическое применение получила разработка фирмы Thales Avionics – комплекс интегрированной модульной авионики Topdeck, установленный на военно-транспортных самолетах стран блока НАТО C-295, CN-235-300. В более совершенном исполнении известны также разработки бортовых комплексов структуры IMA по военным программам: Pave Pillar (комплекс БРЭО применяется в составе аппаратуры истребителя F-22A и боевого вертолета RAH-66); Pave Pace (комплекс БРЭО применяется в составе аппаратуры истребителя F-35).

Об актуальности проектирования новых комплексов БРЭО в классе структур IMA свидетельствует тот факт, что все электронное приборное оборудование нового пассажирского самолета Airbus A380 выполнено по концепции IMA (масса электронного оборудования в эквиваленте снижена на 40% при сохранении числа решаемых задач). С этой целью 33 проектных организации и научно-исследовательских института из 10 стран Европы под руководством Thales Avionics разработали проект бортовой электроники Victoria. БРЭО новейшего пассажирского самолета Boeing-787 «Dreamliner» (США) сегодня также разработано в классе структур IMA.

Авиационные комплексы структуры IMA имеют существенные преимущества перед своими предшественниками. Использование общих ресурсов позволяет уменьшить массогабаритные характеристики БРЭО, номенклатура однотипных модулей облегчает техническое обслуживание (эксплуатацию) объекта, уменьшается стоимость приобретения БРЭО, так как вместо разнотипных блоков, изготовляемых мелкими партиями, на заводах-изготовителях тиражируются модули одного и того же типа с ограниченной номенклатурой. Так, выигрыш от применения на практике бортового комплекса Primus Epic структуры IMA оценен [43] фирмой Honeywell по следующей системе показателей (в сравнении с аппаратурой, разработанной по стандартам 700-й серии ARINC):

- вес и геометрические размеры аппаратуры БРЭО снижены на 40%;
- надежность комплекса БРЭО увеличена на 100%;
- стоимость приобретения оборудования снижена на 30%;
- стоимость использования оборудования снижена на 50%;
- ремонтопригодность бортового оборудования повышена на 60%;
- протяженность электропроводки самолета снижена на 30%.

На самолете Hawker Horizon, где установлен комплекс Primus Epic, по сравнению, например, с Hawker 1000, где ранее устанавливался комплекс БРЭО предыдущего поколения SPZ-8000, достигнута экономия массы бортового оборудования на 45 кг, исключено 23 электронных блока, за счет чего надежность бортового комплекса увеличена вдвое.

Анализ достигнутых отечественными и зарубежными специалистами авиаприборостроения результатов позволяет выделить среди свойств многомодульных бортовых вычислительных средств, разрабатываемых по концепции IMA, следующие наиболее важные:

- мультипроцессорность, состоящая в том, что число модулей в системе не фиксировано заранее (относительно каждого вычислительного процесса оперативная память имеет два уровня: первый – наиболее быстрая по доступу внутренняя память модуля процессора, второй – память, размещаемая вне модуля, в том числе и в других модулях крейта);
- управляемая коммутация связей между элементами модулей, обеспечивающая создание динамиче-

ских вычислительных подсистем различной конфигурации (управление активностью процессоров в системных взаимодействиях);
– идентичность применяемых модулей относительно протоколов связи друг с другом, формата передаваемых данных, базового набора системы команд микропроцессоров и т.д. (при этом не исключается возможность наличия у модулей различных свойств на физическом уровне, например, из-за разной емкости оперативной и постоянной памяти, микропроцессоров различной производительности (неоднородные вычислительные среды) или специализированных сопроцессоров и др.).

Таким образом, структура многомодульной системы может быть различной в каждый момент времени. Сетевые коммутаторы могут содержать локальную память, способную запомнить несколько динамических конфигураций обеспечиваемых ими связей. Установка базовых конфигураций позволяет осуществить непосредственное статическое соединение между двумя или большим числом одновременно существующих в вычислительной системе потоков данных и абонентов.

Ключевое значение для построения многомодульных вычислительных систем структуры IMA имеет проблема обеспечения высокого уровня параллелелизма вычислений. Эффективно многомодульные вычислительные системы функционируют только при решении относительно малосвязных задач. При решении задач многоканальной обработки и задач, схожих с ними по организации вычислительного процесса, в которых задействуется большое число вычислительных модулей с интенсивным межмодульным информационным обменом, использование многомашинных вычислительных комплексов нецелесообразно. Как показывают результаты исследований специалистов ОАО «НИЦЭВТ» [59], полученные с использованием методики тестирования, разработанной Центром независимого межведомственного тестирования суперкомпьютерных систем ГУ РИНКЦЭ, на практике реальная производительность мультипроцессоров при решении многосвязных задач составляет лишь 5–10% от потенциальных возможностей вычислителей.

Структура многомодульных крейтов имеет аналогию со структурой нейронной сети человеческого мозга, пространство состояний которого и механизмы перехода от одного состояния к другому исследовались в работе [60]. Элементарные модули, обладающие собственными вычислительными средствами, могут интерпретироваться как нейроны, а переключатели – как аксоны (отростки нейронов). «Манипулирование» переключателем создает каждый раз новую конфигурацию. Таким образом, в многомодульном вычислителе реализуется свойство адаптации системы под алгоритм решаемой задачи и под наличие свободных (исправных) вычислительных ресурсов, что создает предпосылки для внедрения на борту ЛА реконфигурируемых вычислительных структур.

Модульный состав современных крейтов IMA, назначение КФМ

Унифицированные модули одного крейта IMA могут выполнять следующие бортовые функции [43]:
– автоматического и директорного управления полетом (автопилота);
– автоматического управления тягой двигателя (автомата тяги самолета);
– вычислительной системы самолетовождения;
– управления общесамолетными системами;
– вычислителя-генератора символов для системы индикации (информационно-управляющего поля);
– формирования аварийной, предупреждающей и уведомляющей сигнализации;
– предупреждения о критических режимах полета ЛА;
– предупреждения об опасном приближении ЛА к земле;
– контроля параметров взлета летательного аппарата;
– формирования и выдачи речевой и тональной звуковой сигнализации в кабине пилота;
– хранения и вывода на индикацию различной справочной информации (карты полета, радиоэлектронной информации от радиолокационной станции, контрольные перечни операций и др.), в том числе доступ во время полета к наземной справочной информации (метеорологические службы, службы планирования полетов, предприятия технического обслуживания и ремонта);
– радиосвязи с наземными службами;
– сбора информации для технического обслуживания, в том числе сбора и локализации отказов бортового оборудования комплекса БРЭО;
– концентрации данных от систем и измерительных датчиков ЛА для их использования другим приборным оборудованием комплекса БРЭО.

Для выполнения заданных функций крейты IMA комплектуются модулями различных типов. Наибольшее распространение на практике получил [47] следующий набор стандартных унифицированных КФМ:
– модуль вычислительный (МВ);
– модуль постоянной памяти (МПП);
– модуль графический (МГ);

– модуль ввода-вывода (МВВ);
– модуль напряжений (МН).

На рис. 3, а–д, представлены для примера фотографии модулей структуры IMA разработки ФГУП «СПб ОКБ «Электроавтоматика» им. П.А. Ефимова». Основные принципы построения модулей приведены в [47].

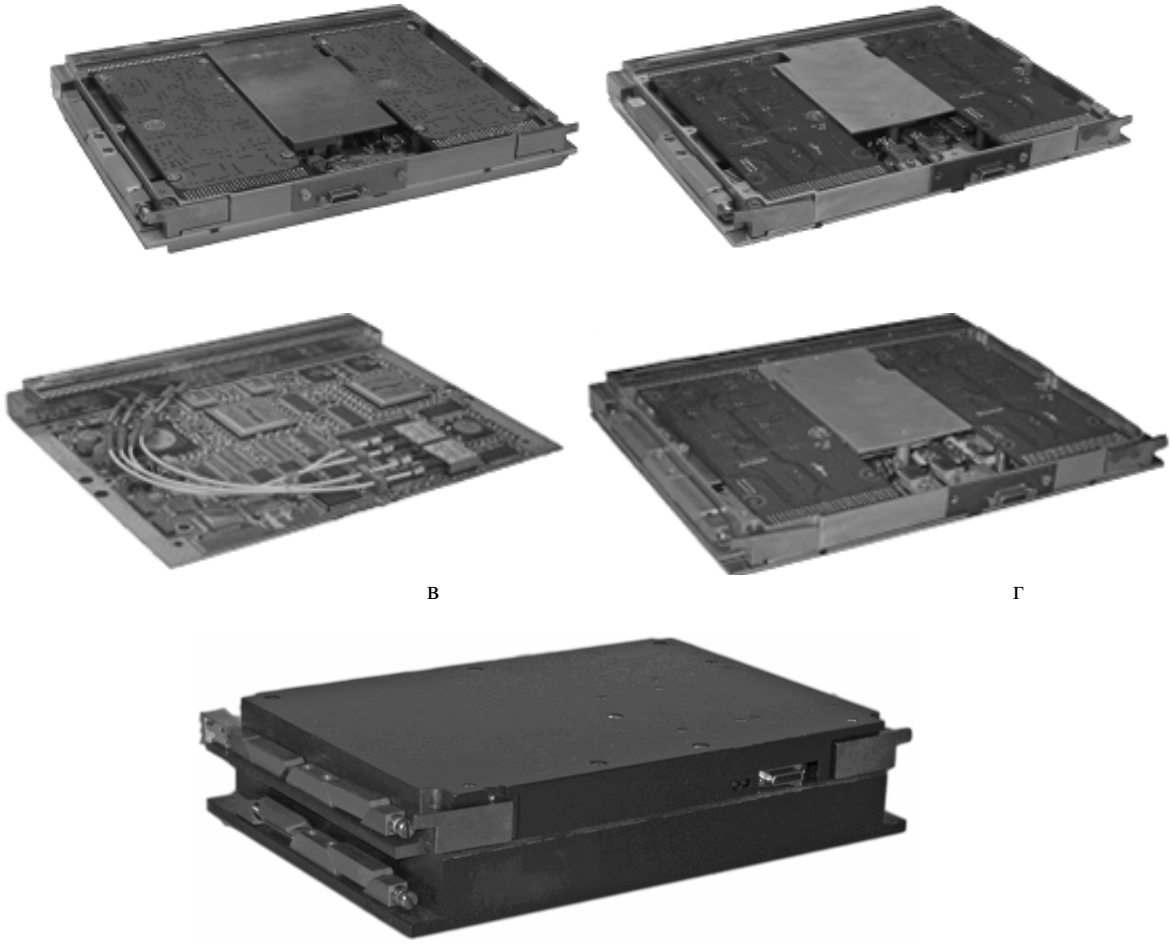


Рис. 3. Модули структуры IMA: модуль вычислительный (а); модуль постоянной памяти (б); модуль графический (в); модуль ввода-вывода (г); модуль напряжений (д)

МВ (рис. 3, а) представляет собой автономную бортовую вычислительную машину в одноплатном исполнении. Каждый из устанавливаемых в крейт МВ, а их может быть несколько (на практике известны случаи размещения в одном конструктиве до 10 вычислительных модулей), способен выполнять в режиме разделения времени несколько бортовых задач, причем задачи могут относиться к разным уровням критичности. Программное обеспечение для МВ имеет два уровня. Один из уровней представляет собой выполняемые задачи, другой уровень изолирует задачи от применяемой вычислительной платформы, обеспечивая их взаимонезависимость. В качестве примера применения модулей МВ следует отметить, что в составе одного крейта БРЭО Primus Epic самолета Hawker-4000 используется 2 аналогичных модуля-вычислителя (число крейтов – от 2 до 4 шт.), столько же модулей используется в составе крейта БРЭО AIMS самолета Boeing-777 (число крейтов – 2 шт.).

МПП (рис. 3, б) выполняет функцию электронной библиотеки, к которой при необходимости обращаются пилоты ЛА или электронные модули комплекса БРЭО, а также функцию записи (регистрации) в полете информации, используемой в процессе наземного технического обслуживания или ремонта. База данных экспертных систем также размещается в пределах МПП. В качестве примера применения модулей МПП следует отметить, что в составе одного крейта БРЭО Primus Epic самолета Hawker-4000 используется 1 аналогичный модуль (число крейтов в составе самолета – от 2 до 4 шт.), столько же модулей используется в составе крейта БРЭО AIMS самолета Boeing-777 (число крейтов в составе самолета – 2 шт.).

МГ (рис. 3, в) собирает пилотажно-навигационную информацию, подготавливает и передает на средства индикации информационно-управляющего поля сформированное изображение (информационный поток видеоинформации). МГ также получает видеоинформацию от измерительных датчиков ЛА,

например, от инфракрасной обзорной системы или радиолокатора. Полученное видеоизображение масштабируется, на него накладывается символьная информация (режим совмещения графической и видеоинформации), и подготовленное для индикации видеоизображение передается на многофункциональные бортовые индикаторы. В качестве примера применения МГ следует отметить, что в составе одного крейта БРЭО Primus Epic самолета Hawker-4000 используется 1 аналогичный модуль (число крейтов в составе самолета – от 2 до 4 шт.), столько же модулей используется в составе крейта БРЭО AIMS самолета Boeing-777 (число крейтов в составе самолета – 2 шт.).

МВВ (рис. 3, г) служит для приема той информации от самолетных систем и измерительных датчиков ЛА, которая поступает в комплекс БРЭО по бортовым интерфейсам различных видов, например, от систем, формирующих информацию по последовательному каналу связи или по мультиплексному каналу. МВВ преобразует поступающую информацию и передает ее во внутреннюю сеть крейта для использования модулями IMA. Наличие модулей ввода-вывода позволяет сосредоточить аппаратуру, необходимую для приема уникальных видов интерфейсов, в одном крейте; всем другим модулям комплекса БРЭО достаточно иметь в своем составе средства доступа к общей бортовой сети, в которую модуль ввода-вывода транслирует поступающую информацию. С целью уменьшения длины и массы электропроводки МВВ можно размещать и вне основного крейта – в непосредственной близости от датчиков сигналов в составе объекта, тогда конструктивно они выполняются в виде отдельного блока-концентратора. В качестве примера применения МВВ следует отметить, что в составе одного крейта БРЭО Primus Epic самолета Hawker-4000 используется 2 аналогичных модуля (число крейтов – от 2 до 4 шт.), 4 модуля используется в составе крейта БРЭО AIMS самолета Boeing-777 (число крейтов – 2 шт.).

МН (рис. 3, д) преобразует первичное напряжение бортовой сети (переменное ~115 В, 400 Гц, или постоянное +27 В) во вторичные напряжения, необходимые для низковольтового электропитания модулей крейта. Зачастую в МН реализуется двухканальная схема. Например, в комплексе ProLine 21 МН состоит из двух независимых формирователей вторичного питания, что существенно повышает надежность вычислительной системы в целом. Каждый из преобразователей способен обеспечивать электропитанием весь крейт, но в номинальном режиме эксплуатации, когда оба канала системы питания исправны, нагрузка распределяется между ними равномерно, вследствие чего преобразователи напряжения функционируют в облегченном электрическом и тепловом режиме. В качестве примера применения МН следует отметить, что в составе одного крейта БРЭО Primus Epic самолета Hawker-4000 используется 2 аналогичных модуля (число крейтов – от 2 до 4 шт.), столько же модулей используется в составе крейта БРЭО AIMS самолета Boeing-777 (число крейтов – 2 шт.).

Применение крейтов IMA в составе летательных аппаратов при разработке комплексов бортового радиоэлектронного оборудования

Успешное использование отечественными разработчиками методов структуризации позволило специалистам авиаприборостроительных предприятий [2, 3, 17, 21, 25, 26, 61] в последние годы сформировать обобщенную функциональную схему федеративных комплексов бортового оборудования (4-ое поколение БРЭО) практически для любого типа летательного аппарата. Рассматривая класс пилотируемых ЛА и ограничиваясь задачами самолетовождения, структуру бортовой авионики (техническое лицо комплекса БРЭО) сегодня представляют в виде многоконтурной системы информационного обмена – бортовой вычислительной сети, в которой:

- основной контур обеспечивает автоматизацию процессов, связанных с целевым назначением объекта (пилотаж/навигация ЛА), предусматривая отказоустойчивый детерминизм структуры БРЭО (за счет дублирования составляющих) для обеспечения заданных характеристик по его надежности;
- резервный контур локализует порядок применения специальных средств летательного аппарата на выполнение частных задач на основе обработки информации от оптико-электронных и радиолокационных датчиков;
- вспомогательный контур обеспечивает передачу экипажу пилотажно-навигационной информации на средства индикации информационно-управляющего поля, осуществляет информационное взаимодействие летчика и аппаратуры путем обработки и анализа состояния органов управления кабины пилота ЛА.

Основные свойства, которые обеспечиваются при такой федеративной организации БРЭО с системной ориентацией, следующие [29, 41]:

- БРЭО определяется как сложная территориально-распределенная система, состоящая из отдельных функциональных подсистем, связанных между собой стандартными информационными соединениями;
- разработка отдельных функциональных подсистем БРЭО осуществляется большей частью автономно различными проектными опытно-конструкторскими бюро, а последующее их комплексирование обеспечивает функциональную интеграцию БРЭО в целом на всех этапах его эксплуатации;
- процесс обработки информации распараллеливается во времени в неоднородных по своей организации и характеристикам вычислительных средах;

– разрабатываемое программное обеспечение БРЭО и его компонентов имеет модульную организацию;
– обеспечивается возможность сохранения статической структуры БРЭО в случае возникновения отказов оборудования;
– обеспечивается возможность модернизации БРЭО и наращивания числа функциональных подсистем без значительного изменения топологии физических соединений (схемы информационного обмена) на межсистемном (внутриобъектовом для ЛА) уровне.

Данный подход в течение длительного времени (до 4-го поколения БРЭО включительно) соответствовал логике развития бортовых систем, функционирование которых базировалось на разнородных физических принципах, а разработка которых в силу этого зачастую поручалась различным проектным организациям и различным главным конструкторам. Однако по мере усложнения функциональных задач и увеличения их количества федеративный принцип построения БРЭО в значительной степени исчерпал свои возможности.

Общепринятый подход [62–69] к проектированию перспективных авиационных комплексов состоит в том, что все подсистемы должны выбираться с позиции интеграции аппаратуры БРЭО в единое целое – «интеллектуальное вычислительное ядро». Доминирующая сегодня концепция интеграции базируется на оптимальной композиции всей совокупности бортовых систем (источников и приемников информации, средств индикации и обработки) на всех этапах проектирования, начиная с концептуальных (поисковых) исследований, эскизного, технического проектирования и заканчивая вопросами унификации и стандартизации основных конструктивно-функциональных модулей, в том числе серийно выпускаемых.

Особенности схем связи комплексов БРЭО для независимого, федеративного и интегрированного принципов построения комплексов БРЭО [39] приведены на рис. 4.

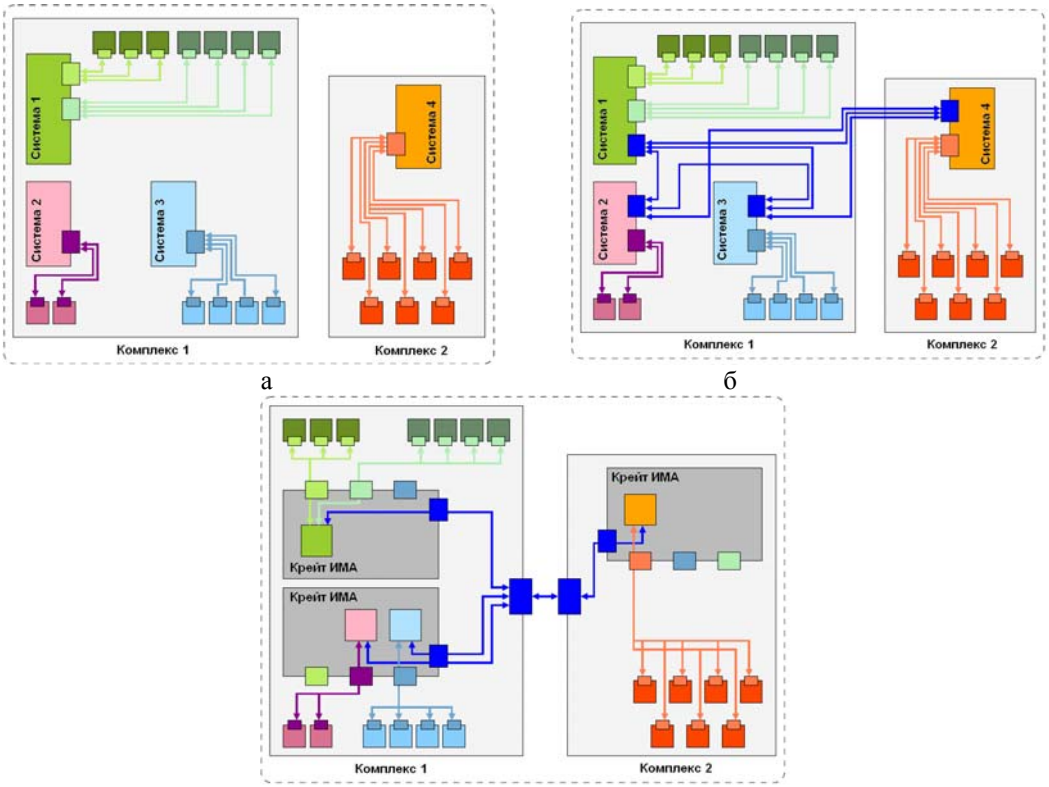


Рис. 4. Примеры построения структур авиационных комплексов: независимый принцип построения архитектуры БРЭО (а); федеративный принцип построения архитектуры БРЭО (б); интегрированный модульный принцип построения архитектуры БРЭО (в) (ИМА – интегрированная модульная авионика)

Авиационные комплексы БРЭО современных ЛА, находящиеся сегодня в эксплуатации, имеют системно-ориентированную детерминированную структурную организацию. Вычислительный ресурс такого БРЭО регулярно распределен между информационными каналами связи посредством организации отдельных подсистем (перераспределение задач между подсистемами на системном уровне не предусматривается), что не обеспечивает возможность интеграции бортового оборудования. Реализуется лишь резервирование (дублирование, реже троирование) наиболее важных функций, решаемых бортовыми вычислительными средствами. Анализ бортовых информационных потоков пилотажно-навигационных

данных свидетельствует [28, 67], что при жестком функциональном распределении задач между специализированными бортовыми вычислительными системами требуемые характеристики производительности федеративного БРЭО 4-го поколения не обеспечиваются даже в случае использования высокоскоростных каналов информационного обмена.

Для радикального повышения эффективности БРЭО разработчиками сегодня предпринимаются попытки его реализации и проектирования в классе мультимикропроцессорных структур IMA с программируемой архитектурой, допускающей динамическое перераспределение вычислительной мощности аппаратуры в зависимости от приоритета решаемых задач. Результирующая структура БРЭО на основе концепции IMA, разработанная специалистами авиаприборостроения [14], приведена на рис. 5.

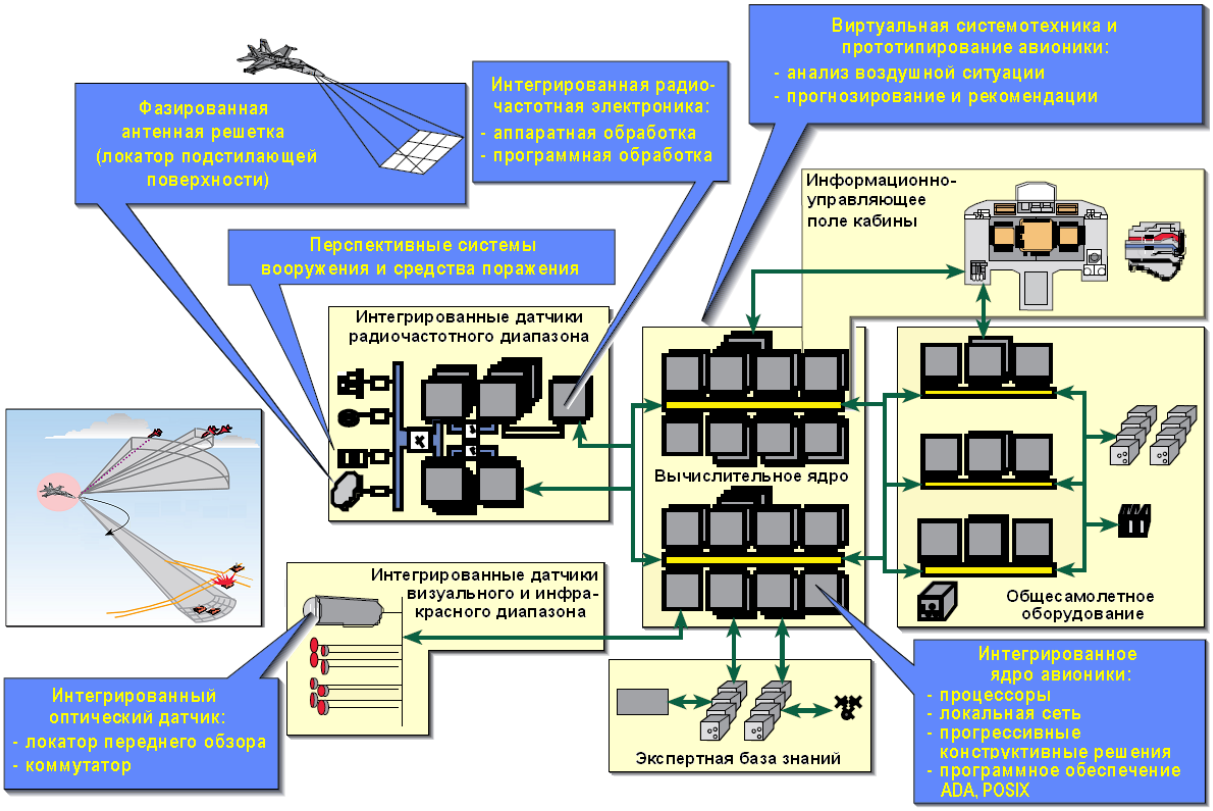


Рис. 5. Интегрированный комплекс БРЭО перспективного самолета, разработанный специалистами авиаприборостроительных предприятий в соответствии с концепцией IMA

Как следует из анализа рис. 5, интеграции в составе ЛА подлежат система измерительных датчиков радиочастотного диапазона, система датчиков визуального и инфракрасного диапазона частот, бортовая вычислительная система (вычислительное ядро комплекса авионики), информационно-управляющее поле кабины пилота летательного аппарата и общесамолетное оборудование.

Заключение

Подводя итоги аналитического обзора современного состояния и перспектив развития интегрированных бортовых вычислительных систем, применяемых в авиационном приборостроении, необходимо отметить, что техническую основу проектирования изделий структур IMA сегодня составляют:

– бортовая информационная сеть сбора и распределения в реальном масштабе времени поступающей от различных датчиков информации об окружающей ЛА обстановке (датчиков переднего и заднего обзора, аэрометрических датчиков и пр.), а также информации для устройств, представляющих собой аналоговые и цифровые средства измерения, средства анализа и преобразования данных;
– семейство унифицированных цифровых вычислительных модулей обработки и передачи данных, образующих в совокупности мультипроцессорный суперкомпьютер, конструктивно оформленный в виде стандартного крейта (число крейтов на борту зависит от числа функциональных задач и от производительности вычислительных средств конструктивно-функциональных модулей);
– математическое, алгоритмическое и программное обеспечение, включающее операционные системы реального времени, осуществляющее функции планирования и диспетчеризации вычислений и распределения вычислительных ресурсов (при назначении ресурсов выполнение задачи осуществляется вычислительными средствами конструктивно-функциональных модулей низшего уровня).

Учитывая, что полное внедрение методологии IMA в основы проектирования бортового радиоэлектронного оборудования – это проблема не сегодняшнего, а пока еще «завтрашнего» дня, большинство специалистов авиаприборостроительной отрасли [1, 2, 4, 5, 12, 17, 19, 22, 25–27, 29, 40, 43] склоняются к выводу, что первостепенными авиационными объектами проектирования в классе структур IMA должны стать:

– унифицированная интегрированная вычислительная система самолетовождения с открытой архитектурой;
– интегрированная система самолетовождения и индикации;
– интегрированная система самолетовождения, индикации и управления для различных самолетов гражданской и военной авиации.

Литература

1. Воробьев А.В. Научные основы создания отказоустойчивых интегрированных вычислительных комплексов систем управления летательными аппаратами: Автореф. дисс. … д-ра техн. наук: 05.13.05. – М.: МГИЭМ, 2010. – 32 с.
2. Джанджгава Г.И. Авионика пятого поколения: новые задачи – новая структура // Вестник авиации и космонавтики. – 2001. – № 5. – С. 8–10.
3. Джанджгава Г.И., Герасимов Г.И., Рогалев А.П., Шерман В.М., Сухоруков С.Я., Вершков В.В. Концепция создания интегрированных комплексов бортового оборудования летательных аппаратов нового поколения // Авиакосмическое приборостроение. – 2002. – № 6. – С. 38.
4. Евгенов А.В. Направления развития интегрированных комплексов бортового оборудования самолетов гражданской авиации // Авиакосмическое приборостроение. – 2003. – № 3. – С. 48–53.
5. Ефанов В.Н., Бодрунов С.Д. Открытые архитектуры в концепции авионики пятого поколения // Мир авионики. – 2004. – № 5. – С. 20–28.
6. Левин Д.Н., Пономаренко А.В., Сильвестров М.М. Концептуальный облик и особенности построения диалоговых моделирующих комплексов и действующего макета кабины для эргономического сопровождения разработки эргатического информационно-управляющего комплекса перспективного многофункционального маневренного самолета // Мехатроника, автоматизация, управление. – 2009. – № 12. – С. 52–59.
7. Михайлуца К.Т., Чернышев Е.Э., Шейнин Ю.Е. Основные архитектурные концепции информационно-вычислительной среды бортового оборудования перспективных летательных аппаратов // Современные технологии извлечения и обработки информации. – СПб: Радиоавионика, 2001. – 175 с.
8. Золотарев С.В. LynxOS-178 – сертифицированная ОСРВ для интегрированной модульной авионики // Мир компьютерной автоматизации. 2006. № 5 [Электронный ресурс]. – Режим доступа: http://www/rtsoft_training. ru/?p=600071, свободный. Яз. рус. (дата обращения 08.10.2009).
9. Паркинсон П., Киннан Л. Разработка критического по безопасности ПО для интегрированной модульной авионики [Электронный ресурс]. – Режим доступа: http://www.vxworks.ru/rus_safety-critical-sw-dev_wp-1107.pdf, свободный. Яз. рус. (дата обращения 27.09.2009).
10. Борисов Ю.И. Отечественная электронная промышленность и компонентная база. Перспективы развития // Электроника: НТБ. – 2006. – № 2. – С. 6–9.
11. Данилин Н.С., Белослудцев С.А. Проблемы применения современной индустриальной электронной компонентной базы иностранного производства в ракетно-космической технике // Современная электроника. – 2007. – № 7. – С. 8–12.
12. Колпаков К.М., Павлов А.М. Состояние и тенденции развития бортовых вычислительных систем перспективных летательных аппаратов. Аналитический обзор по материалам зарубежных источников / Под ред. Е.А. Федосова. – М.: ФГУП «НИИ АС», 2008. – Ч. 1. – 118 с.; Ч. 2. – 122 с.
13. Программа JSF и ее влияние на авионику боевых самолетов 5-го поколения. Аналитический обзор по материалам зарубежной печати / Под ред. Е.А. Федосова. – М.: ФГУП «НИИ АС», 2000. – 129 с.
14. Steidle Craig E. The Joint Strike Fighter Program // Johns Hopkins APL Technical Digest. – 1997. – V.18. – № 1. – P. 6–20.
15. Eveleens René L.C. Open Systems Integrated Modular Avionics – The Real Thing // Mission Systems Engineering. – Educational Notes RTO-EN-SCI-176, 2006. – Neuilly-sur-Seine, November, France: RTO. – Paper 2. – P. 2-1–2-22.
16. Stevens B., Price M. A Presentation to the Allied Standard Avionics Architecture Council Systems Working Group. – United Kingdom. 2005. July [Электронный ресурс]. – Режим доступа: http://assconline.co.uk/documents/BS_Slides-Systems_WG_200705.pdf, свободный. Яз. англ. (дата обращения 10.09.2010).
17. Анцев Г.В. Принципы построения бортовых информационно-управляющих систем высокоточного оружия нового поколения // Радиотехника. – 2001. – № 8. – С. 81–86.

18. Бабуров В.И., Пономаренко Б.В. Принципы интегрированной бортовой авионики. – СПб: Агентство «РДК-Принт», 2005. – 448 с.
19. Белый Ю.И. Построение комплексов БРЭО самолетов пятого поколения. Основные принципы // Аэрокосмическое приборостроение России в трудах специалистов корпорации «Аэрокосмическое оборудование» (1998–2003 гг.) / Под ред. С.Д. Бодрунова. – СПб: Изд-во ОАО «Пирометр», 2003. – Ч. 2. – С. 126–129.
20. Бражкин В.М., Герасимов Г.И. Развитие интегрированных комплексов бортового оборудования самолетов нового поколения // Мехатроника, автоматизация, управление. – 2008. – № 2. – С. 24–29.
21. Джанджгава Г.И., Герасимов Г.И., Петкевичус П.Ю. Развитие интегрированных комплексов бортового оборудования навигации, управления и наведения летательных аппаратов в разработках Раменского приборостроительного конструкторского бюро // Авиакосмическое приборостроение. – 2008. – № 2. – С. 3–8.
22. Итенберг И. Интегрированная модульная электроника – новая стратегия на рынке приборостроения // Новый оборонный заказ. Стратегии. – 2010. – № 5. – С. 64–65.
23. Кофман М.М., Парамонов П.П., Сабо Ю.И. Интеграция на основе системной отказоустойчивости — перспективный путь проектирования комплексов бортового оборудования и обеспечения безопасности полетов // Авиакосмическое приборостроение. – 2005. – № 8. – С. 25–31.
24. Кофман М.М., Суров Б.И., Сабо Ю.И., Суслов В.Д., Шек-Иовсепянц Р.А. Интеграция авионики — основное направление комплексирования бортового радиоэлектронного оборудования перспективных вертолетов и самолетов // Аэрокосмическое приборостроение России в трудах специалистов корпорации «Аэрокосмическое оборудование» (1998–2003 гг.) / Под ред. С.Д. Бодрунова. – СПб: Изд-во ОАО «Пирометр», 2003. – Ч. 2. – С. 143–153.
25. Парамонов П.П. Основы проектирования авионики. – Тула: Гриф и К., 2003. – 228 с.
26. Турчак А.А. Архитектура вычислительных систем для интегрированной модульной авионики перспективных летательных аппаратов // Радиотехника. – 2001. – № 8. – С. 87–95.
27. Федосов Е.А. IMA Russian Program – Overall presentation // European and Russian Joint Avionics Forum. – Moscow, 2009. –15 p.
28. Гатчин Ю.А., Жаринов И.О. Основы проектирования вычислительных систем интегрированной модульной авионики: монография. – М.: Машиностроение, 2010. – 224 с.
29. Павлов А.М. Принцип организации бортовых вычислительных систем перспективных летательных аппаратов // Мир компьютерной автоматизации. – 2001. – № 4 [Электронный ресурс]. – Режим доступа: www.mka.ru/?p= 41177, свободный. Яз рус. (дата обращения 25.04.2007).
30. Пятницких А. Бортовые компьютеры: варианты построения готовых систем // Современные технологии автоматизации. – 2008. – № 2. – С. 20–24.
31. Джонсон К., Леру П. Использование технологии объединения ресурсов для создания безопасных отказоустойчивых военных систем // Современные технологии автоматизации. – 2007. – № 4. – С. 72–76.
32. Буравлев А., Чельдиев М., Барыбин А., Костенко В., Тумакин Д., Петров Г. Масштабируемые мультипроцессорные вычислительные системы высокой производительности // Современные технологии автоматизации. – 2009. – № 3. – С. 72–82.
33. Беляков А.С., Путан Д.Б., Романов Д.А., Хахулин А.А. Интегрированная вычислительная среда малого космического аппарата // Авиакосмическое приборостроение. – 2008. – № 6. – С. 43–47.
34. Гобчанский О. Проблемы создания бортовых вычислительных комплексов малых космических аппаратов // Современные технологии автоматизации. – 2001. – № 4. – С. 28–34.
35. Горбачев С.В., Рождественский Д.А., Суворова Е.А., Шейнин Ю.Е. Масштабируемые архитектуры распределенных систем на технологии SpaceWire на базе платформы «МУЛЬТИКОР» // Вопросы радиоэлектроники. Сер. ОТ. – 2006. – Вып. 2. – С. 69–80.
36. Солохина Т.В., Петричкович Я.Я., Шейнин Ю.Е. Технология SpaceWire для параллельных систем и бортовых распределенных компьютеров // Электроника: НТБ. – 2006. – № 5. – С. 64–45.
37. Солохина Т.В., Петричкович Я.Я., Шейнин Ю.Е. Технология SpaceWire для параллельных систем и бортовых распределенных компьютеров // Электроника: НТБ. – 2007. – № 1. – С. 38–49.
38. Эрглис К.Э. Интерфейсы открытых систем. – М.: Горячая линия – Телеком, 2000. – 256 с.
39. Ott Aliki. System Testing In The Avionics Domain: Dissertation Zur Erlangung Des Grades Einer Doktorin Der Ingenieurwissenschaften. – Germany: Universität Bremen, 2007. – 434 p.
40. Федосеев Е.П., Колпаков К.М. Бортовые вычислительные системы сквозь призму авионики 21-го века // Мехатроника, автоматизация, управление. Приложение. – 2006. – № 3. – С. 17–22.
41. Колпаков К.М. История развития бортовых цифровых вычислительных машин в России // PCweek. 1999. №32. [Электронный ресурс]. – Режим доступа: www.pcweek.ru /year1999/№32/CP1251, свободный. Яз рус. (дата обращения 05.01.2006).
42. Жаринов И.О. Принципы построения и методы автоматизации проектирования вычислительных систем интегрированных комплексов бортового оборудования: Дисс. … д-ра техн. наук: 05.13.12. – СПб:

НИУ ИТМО. – 2011. – 298 с.
43. Кучерявый А.А. Бортовые информационные системы: Курс лекций / Под ред. В.А. Мишина и Г.И. Клюева. – 2-е изд. перераб. и доп. – Ульяновск: УлГТУ, 2004. – 504 с.
44. Севбо В., Орлов А., Лошаков А. Многопроцессорный вычислительный комплекс для задач «жесткого» реального времени // Современные технологии автоматизации. – 2007. – № 3. – С. 32–38.
45. Жаринов О.О., Видин Б.В., Шек-Иовсепянц Р.А. Принципы построения крейта бортовой многопроцессорной вычислительной системы для авионики пятого поколения // Научно-технический вестник СПбГУ ИТМО. – 2010. – № 4 (68). – С. 21–27.
46. Кузнецова О.А. Оценка надежности структурно избыточных комплексов авионики с учетом среднего времени между восстановлениями при отказах // Изв. вузов. Приборостроение. – 2012. – Т. 55. – № 3. – С. 65–69.
47. Богданов А.В., Васильев Г.А., Виноградов П.С., Егоров К.А., Зайченко А.Н., Ковернинский И.В., Петухов В.И., Романов А.Н., Смирнов Е.В., Уткин Б.В., Федосов Е.А., Шукалов А.В. Платформа интегрированной модульной авионики. Патент на полезную модель № 108868 U1 RU, МПК G06F 9/00, №2011121962/08. Заявл. 01.06.2011. Опубл. 27.09.2011.
48. Раменское приборостроительное конструкторское бюро. Бортовая вычислительная станция «БВС-1». [Электронный ресурс]. – Режим доступа: www.rpkb.ru, свободный. Яз рус. (дата обращения 05.01.2012).
49. Егоров К.А., Итенберг И.И., Ковернинский И.В., Тимченко А.П., Федосов Е.А., Чуянов Г.А. Платформа интегрированной модульной авионики. Патент № 2413280 C1 RU, МПК G06F 9/02. №2009127190/08. Заявл. 14.07.2009. Опубл. 27.02.2011. Бюл. № 6.
50. Егоров К.А., Итенберг И.И., Ковернинский И.В., Тимченко А.П., Федосов Е.А., Чуянов Г.А. Платформа интегрированной модульной авионики. Патент на полезную модель №88462 C1 RU, МПК G06F 9/00. № 2009127040/22. Заявл. 14.07.2009. Опубл. 10.11.2009.
51. Слуцкин А., Эйсымонт Л. Российский суперкомпьютер с глобально адресуемой памятью // Открытые системы. – 2007. – № 9 [Электронный ресурс]. – Режим доступа: http://www.osp.ru/os/2007/09/4569294, свободный. Яз. рус. (дата обращения 25.09.2009).
52. Герлих Х. Модульная система авионики самолета. Патент № 2413655 C2 RU, МПК B64C 19/00. № 2008123940/11. Заявл. 16.11.2006. Опубл. 10.03.2011. Бюл. № 7.
53. Абросимов О.В., Егоров К.А., Итенберг И.И., Ковальчученко А.Ф., Куликов Д.А., Сивцов С.А., Федосов Е.А., Чуянов Г.А. Базовая несущая конструкция платформы интегрированной модульной авионики. Патент на полезную модель № 106404 U1 RU, МПК G06F 1/16. № 2010129389/08. Заявл. 15.07.2010. Опубл. 10.07.2011.
54. Платунов В.С. Методология системных военно-научных исследований авиационных комплексов. – М.: «Дельта», 2005. – 344 с.
55. Топорков В.В. Модели распределенных вычислений. – М.: ФИЗМАТЛИТ, 2004. – 320 с.
56. Парамонов П.П., Гатчин Ю.А., Жаринов И.О., Жаринов О.О., Дейко М.С. Принципы построения отраслевой системы автоматизированного проектирования в авиационном приборостроении // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – № 6 (82). – С. 111–117.
57. Гатчин Ю.А., Жаринов И.О., Жаринов О.О. Архитектура программного обеспечения автоматизированного рабочего места разработчика бортового авиационного оборудования // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – № 2 (78). – С. 140–141.
58. Дейко М.С., Жаринов И.О. Применение симплекс-метода и метода искусственного базиса при проектировании бортового приборного оборудования // Научно-технический вестник информационных технологий, механики и оптики. – 2013. – № 1 (83). – С. 124–129.
59. Митрофанов В., Слуцкин А., Ларионов К., Эйсымонт Л. Направления развития отечественных высокопроизводительных систем // Открытые системы. 2003. № 5. [Электронный ресурс]. – Режим доступа: www.osp.ru/os/ 2003/05/183021.pdf, свободный. Яз. рус. (дата обращения 24.09.2009).
60. Жаринов И.О., Жаринов О.О. Статистический анализ информационных процессов с повторяющимися признаками формы в электрофизиологических исследованиях и медицинском приборостроении: монография. – СПб: «Нестор-история», 2011. – 151 с.
61. Джанджгава Г.И., Рогалев А.П., Бабиченко А.В., Сухоруков С.Я. Интегрированная динамически реконфигурируемая система комплексной обработки информации бортовых комплексов навигации, управления и наведения // Авиакосмическое приборостроение. – 2002. – № 6. – С. 8–14.
62. Афраймович Л.Г., Прилуцкий М.Х. Многоиндексные задачи распределения ресурсов в иерархических системах // Автоматика и телемеханика. – 2006. – № 6. – С. 194–205.
63. Богатырев В.А., Голубев И.Ю., Беззубов В.Ф. Организация межмашинного обмена в дублированных вычислительных комплексах // Изв. вузов. Приборостроение. – 2012. – Т. 55. – № 3. – С. 8–13.

64. Богатырев В.А. К распределению функциональных ресурсов в отказоустойчивых многомашинных вычислительных системах // Приборы и системы. Управление, контроль, диагностика. – 2001. – № 12. – С. 1–5.
65. Видин Б.В., Жаринов И.О., Жаринов О.О. Декомпозиционные методы в задачах распределения вычислительных ресурсов многомашинных комплексов бортовой авионики // Информационно-управляющие системы. – 2010. – № 1. – С. 2–5.
66. Копорский Н.С., Видин Б.В., Жаринов И.О. Организация вычислительного процесса в многомашинном бортовом вычислительном комплексе // Изв. вузов. Приборостроение. – 2006. – Т. 49. – № 6. – С. 41–50.
67. Парамонов П.П., Бобцов А.А., Видин Б.В., Жаринов И.О., Жаринов О.О., Сабо Ю.И., Шек-Иовсепянц Р.А. Проектирование систем бортового информационного обмена и их функциональных элементов: монография. – Тула: Гриф и К., 2010. – 208 с.
68. Парамонов П.П., Видин Б.В., Есин Ю.Ф., Жаринов И.О., Колесников Ю.Л., Кофман М.М., Сабо Ю.И., Шек-Иовсепянц РА. Теория и практика системного проектирования авионики: монография. – Тула: Гриф и К., 2010. – 365 с.
69. Сырямкин В.И., Шидловский В.С., Глушков Г.С., Лунев С.О., Бурмантов С.И. Интегрированные структурно-перестраиваемые корреляционно-экстремальные навигационные системы // Изв. вузов. Приборостроение. – 2012. – Т. 55. – № 2. – С. 51–55.

Парамонов Павел Павлович – ФГУП «Санкт-Петербургское опытно-конструкторское бюро «Электроавтоматика» им. П.А.Ефимова», доктор технических наук, профессор, генеральный директор, postmaster@elavt.spb.ru

Жаринов Игорь Олегович – ФГУП «Санкт-Петербургское опытно-конструкторское бюро «Электроавтоматика» им. П.А.Ефимова», доктор технических наук, доцент, руководитель учебно-научного центра, igor_rabota@pisem.net

1 ОПТИЧЕСКИЕ И ОПТИКО-ЭЛЕКТРОННЫЕ СИСТЕМЫ. ОПТИЧЕСКИЕ ТЕХНОЛОГИИ

УДК 681.7

ПРОБЛЕМЫ ПРОИЗВОДСТВА ВЫСОКОПРОЧНОГО ОПТИЧЕСКОГО ВОЛОКНА

А.Г. Коробейников, Ю.А. Гатчин, К.В. Дукельский, Е.В. Тер-Нерсесянц

Представлены результаты анализа основных технологических процессов получения высокопрочного оптического волокна. Рассмотрены основные факторы, влияющие на прочность и долговечность волоконных световодов. Дана рекомендация для изготовления высокопрочных длинномерных световодов MCVD-методом.

Ключевые слова: оптическое волокно, волоконные световоды, образование дефектов.

Введение

Все увеличивающееся применение волоконных световодов (ВС) в линиях связи, датчиках физических величин и волоконно-оптических приборах делает актуальной задачу повышения их прочности и долговечности. Существует ряд причин, приводящих к ухудшению работоспособности ВС, например, увеличение оптических потерь из-за микроизгибов при изменении температуры окружающей среды, рост потерь под действием ионизирующих излучений или из-за диффузии водорода в сердцевину световода в подводном кабеле [1–2]. Однако наиболее катастрофические последствия для волоконно-оптических систем и устройств происходят при разрушении световода в процессе эксплуатации. В связи с этим информация о возможности световодов разных типов выдерживать эксплуатационные нагрузки в течение периода эксплуатации всегда была критически важной и часто ограничивающей потенциальные применения в новых областях. По этой причине разработка новых и усовершенствование существующих методов получения оптического волокна с экстремально высокими прочностными показателями относятся к приоритетным задачам физики и химии полимеров.

Производство ВС на основе кварцевого стекла можно разделить на два основных этапа. Первый этап – это получение заготовки, представляющей собой стержень длиной порядка метра, а в диаметре около 10–80 мм. На втором этапе заготовки помещаются в плавильные печи, и из них тянут стеклянное волокно с нанесением защитного полимерного покрытия. При этом соотношение диаметров внутренних слоев в полученном волокне остается таким же, как в заготовке.

Существует три основных парофазных метода изготовления заготовок световодов:

1. осаждение из газовой фазы на внешнюю поверхность цилиндрической подложки (OVD);
2. осевое осаждение из газовой фазы на торце стержня (AVD);
3. модифицированное химическое осаждение из газовой фазы на внутреннюю поверхность кварцевой трубки (MCVD).

Целью настоящей работы является рассмотрение предлагаемых действий для повышения прочности оптического волокна при его производстве методом MCVD.

Анализ прочностных характеристик световодов

В настоящее время большинство ВС изготавливается из кварцевого стекла. Известно, что кварцевое стекло обладает очень высокой прочностью, около 15 ГПа [3]. Однако на больших длинах реализовать такой уровень прочности очень сложно [4]. Изоляция поверхности стекловолокна от воздействия окружающей среды нанесением полимерных оболочек методом MCVD обеспечивает сохранность прочностных свойств стекла, но не изолирует его от влаги. Этот метод является одним из наиболее ранних и простых способов изготовления заготовок ВС, заключающегося в осаждении стеклообразных слоев на внутреннюю поверхность кварцевой трубки. Он впервые был опубликован в работах Мак Чесни с соавторами в 1974 г. [5, 6]. Во многих странах к этому времени были освоены промышленные методы производства труб из кварцевого стекла. В связи с этим MCVD-метод изготовления ВС сразу получил широкое распространение во всем мире. Не прошло и десяти лет, как состояние разработок было доведено до использования световодов в волоконно-оптических линиях связи общей протяженностью в десятки тысяч километров.

В MCVD-методе изолированность от внешней среды реакционной зоны и линий подачи к ней реагентов гарантирует чистоту процесса. Малое содержание примесей обеспечивает высокую прозрачность синтезированного стекла. В то же время универсальность этой технологии создает благоприятные условия для разработки самых различных структур ВС. В этом случае MCVD-метод в сравнении со способом аксиального или наружного парофазного осаждения является наиболее приемлемым в части изготовления оптических волокон сложной конструкции, обладающих особыми свойствами.

Отечественные исследователи приводят экспериментальные данные, свидетельствующие об упрочняющем эффекте полимерного покрытия ВС, изготавливаемых MCVD-методом [7].

В зависимости от метода испытаний прочность ВС характеризуется как динамическая и статическая. Первая определяется величиной разрушающего напряжения, вторая – длительностью процесса разрушения под действием постоянного напряжения (долговечностью). Экспериментальные данные хорошо аппроксимируются уравнением [8]

$$\log \sigma_d = (1 + n)^{-1} \log v_\sigma + (1 + n)^{-1} \log k_d,$$

где σ – приложенное к световоду напряжение; v_σ – скорость нагружения; k_d – коэффициент, характеризующий кварцевое стекло; n – параметр влияния окружающей среды.

Размер дефекта r, расположенного на поверхности стекловолокна, определяет величину разрушающего напряжения σ (рис. 1) [9]:

$$\sigma = 0{,}474 \times 10^{-3}\, r^{-0{,}5} \text{ (ГПа)}. \quad (1)$$

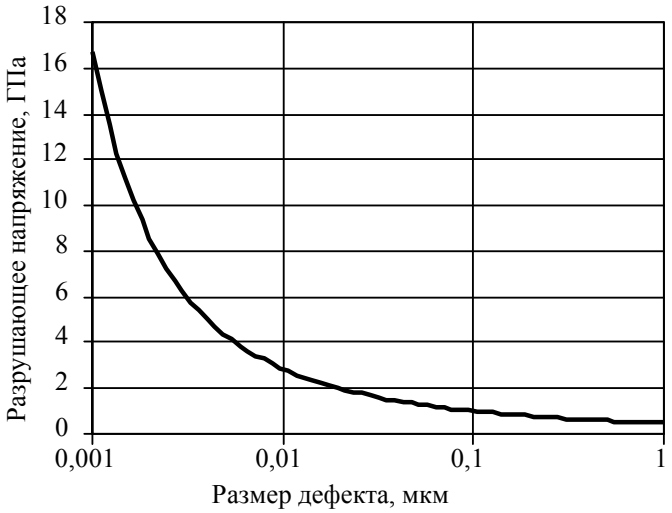


Рис. 1. Влияние размера дефекта на прочность световодов

На основании уравнения (1) можно оценить величину трещины, которая обусловливает верхний уровень прочности метровых отрезков ВС, равный 6 ГПа. Оказывается, что размер дефекта составляет около 6 нм. Эта величина в 37 раз больше длины структурного звена Si-O, равной 0,16 нм.

Прочность даже очень малых отрезков ВС (6–7 ГПа), измеренная методом изгиба, свидетельствует о гарантированном наличии трещин глубиной не менее 6 нм на поверхности стекловолокна диаметром 125 мкм и длиной около 1 мм [4]. Существует мнение [9], что истинная глубина начальных микротрещин на поверхности кварцевых ВС изменяется от 20 до 1000 Å.

Долговечность ВС τ зависит от величины приложенного к волокну напряжения σ и влажности окружающей среды [10]:

$$\ln \tau = \ln \tau_0 - \alpha \ln P + U_0/(RT) - \gamma\sigma/(RT),$$

где τ – длительность процесса разрушения; τ_0 – период атомных колебаний (10^{-13} с); α – порядок реакции гидролиза; γ – константа, характеризующая дефектность образца; P – давление паров воды; U_0 – энергия гидролитического разрыва связи Si-O; R – газовая постоянная; T – температура.

Сведения о величине α для кварцевого стекла противоречивы: рекомендованные значения порядка реакции гидролиза изменяются от 0,569 до 2,2 [11–15].

Для вероятностной оценки разрушения напряженного ВС широко используется модель «слабого звена» [16, 17]. Согласно этой модели, с увеличением длины волокна повышается частота появления более крупных дефектов. Вероятность F того, что прочность ВС длиной L окажется менее определенной величины напряжения σ с учетом статистики Вейбулла, определяется уравнением [18, 19]

$$\ln\ln [(1-F)^{-1}] = \ln L + m \ln\sigma = \ln L + \nu^{-1} \ln\sigma,$$

где $m \approx 1/\nu$; ν – коэффициент вариации.

Чем больше значение параметра m, тем в меньшей степени отличаются нижний и верхний уровень прочности. Экспериментальные данные для серии образцов, как правило, ложатся на прямую линию в координатах Вейбулла $\ln\sigma - \ln\ln (1-F)^{-1}$. Если статистика образования дефектов имеет разную природу, то экспериментальные данные низкопрочного состояния ложатся на искривленную линию (рис. 2) [20].

Срок службы ВС, находящегося под постоянной нагрузкой $\sigma_{раб}$, можно оценить его перемоткой при повышенном напряжении $\sigma_{пер}$ [21]:

$$\tau = B\, (\sigma_{пер})^{n-2} / (\sigma_{раб})^{n},$$

где n – параметр, характеризующий влияние влаги и равный 21,5±1,5; B – характеристика дефектности волокна, равная 10^{-3} ГПа^2·с [22].

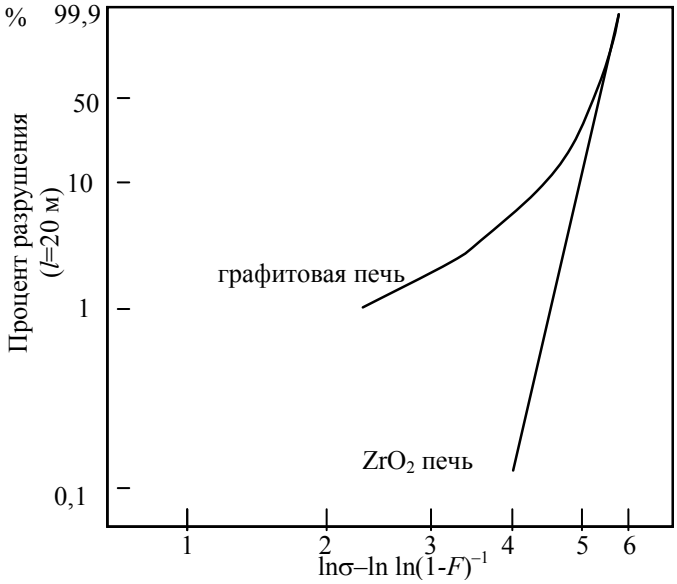


Рис. 2. Распределение вейбулловской вероятности прочных волокон, вытянутых в ZrO_2 печи, и низкопрочных, полученных в графитовой печи

Улучшение прочностных характеристик световодов

Разрушение ВС начинается, как правило, с дефекта, расположенного на поверхности волокна. В связи с этим качество стекломассы опорных кварцевых труб является определяющим фактором прочностных характеристик световодов, изготавливаемых MCVD-методом.

Экспериментально установлено, что трубы, изготавливаемые методом непрерывного формования из природного сырья, не могут обеспечить высокой прочности световодов, особенно если наплавление стекла происходит в среде водорода [23]. Трубы из синтетического кварцевого стекла намного дороже, но содержат существенно меньше микродефектов, благодаря чему обеспечивают высокопрочное состояние ВС [24].

При подготовке заготовки к вытягиванию волокна она в обязательном порядке проходит следующие операции [25]: промывку в чистом изопропиловом спирте, промывку в деионизованной воде, травление в растворе HF, промывку деионизованной водой и сушку, огненную полировку в пламени кислородно-водородной горелки при температуре порядка 2050ºС.

Прочность световодов существенно зависит от технологических параметров вытягивания волокна – температуры, натяжения и стерильности высокотемпературной зоны (запыленности). Изучению этого вопроса с целью оптимизации процесса вытягивания ВС посвящено несколько работ [20, 25–28].

В [28] анализируется состав частиц в высокотемпературной зоне печного пространства графитового нагревателя. Установлена взаимосвязь прочности ВС и размера этих частиц. В окислительных условиях индукционного нагрева легче создать стерильные условия, особенно когда нагреватель из диоксида циркония защищается диоксидом кремния [29–31]. С таким средством нагрева ВС получаются более прочными (рис. 2). С использованием кислородно-водородных горелок также можно обеспечить надлежащую чистоту высокотемпературной зоны, однако высокого уровня стабильности диаметра стекловолокна до настоящего времени достичь не удавалось [32]. Заготовки можно нагревать и CO_2-лазером, но работы в этом направлении не нашли должного продолжения [33–35].

Конструкция графитовой печи существенно влияет на прочностные свойства волокна. Так, введение внутрь нагревателя цилиндрических экранов, ограничивающих зону нагрева и обеспечивающих более надежную изоляцию от графитовых частиц, приводит к шестикратному увеличению нижнего уровня прочности, поднимая его с 80 до 500 кг/мм^2 [36].

С применением индукционного метода нагрева с использованием муфеля из стабилизированной двуокиси циркония получены лучшие по прочности длинномерные световоды. Вытянутый в таких условиях ВС длиной 8,5 км выдержал перемотку под напряжением 1,4 ГПа, а под напряжением 3,5 ГПа перемотан отрезок длиной 4 км [37]. Влияние дефектов на прочность стекловолокна можно ослабить за счет создания сжимающих напряжений в его поверхностном слое. Для этих целей на заготовку газофазным методом наносится тонкий слой легированного кварцевого стекла, обладающего более низкой вязкостью [37, 38], или слой с пониженным коэффициентом термического расширения [39].

Второй метод нашел применение в изготовлении ВС, поверхностный слой которых легирован диоксидом титана, понижающим коэффициент термического расширения кварцевого стекла [40, 41]. Нанесением тонкого металлического покрытия на ВС добиваются более чем двукратного повышения их прочности за счет изоляции стекла от влаги [42–44]. Однако повышенная жесткость металлической оболочки по сравнению с полимерным покрытием приводит к дополнительным оптическим потерям, особенно при температуре менее –10ºС [45].

Несмотря на прочностные достоинства такого метода защиты ВС, он не используется для изготовления длинномерных волокон из-за сложности обеспечения сплошности металлической оболочки на больших длинах. В производстве световодов для их защиты от влаги наиболее широкое распространение получили тонкие (менее 50 нм) углеродные слои, осаждаемые пиролитическим способом [46, 47]. При умеренных нагрузках долговечность световодов с углеродным покрытием превосходит долговечность волокон с полимерным покрытием в $10–10^5$ раз. Однако уровень предельной прочности у последних выше на 40–60%, что исключает использование углеродных покрытий для изготовления высокопрочных световодов [48].

Материалы полимерного покрытия, используемые для защиты стекловолокна от внешнего воздействия, подвергаются фильтрации с целью удаления абразивных частиц [49].

Заключение

Технологии получения волоконных световодов из соображений конкурентоспособности должны отличаться высокой производительностью, малозатратностью с максимально возможным использованием материалов отечественного производства. Разработка таких волоконных световодов диктует необходимость проведения комплексных исследований в следующих направлениях:

- изучение процессов по упрочнению кварцевых световодов и создание новых технических решений для повышения их прочностных свойств;
- модернизация газофазных методов изготовления опорных кварцевых труб и заготовок световодов;
- изучение свойств микроструктурированных оптических волокон;
- исследование влияния основных технологических факторов на характеристики световодов с последующей оптимизацией режимов их изготовления в условиях опытного производства.

Применение представленных результатов и рекомендаций позволят повысить прочность волоконных световодов, получаемых методом MCVD.

Литература

1. Коробейников А.Г., Дукельский К.В., Тер-Нерсесянц Е.В. Методы уменьшения оптических потерь в фотонно-кристаллическом оптическом волокне // Научно-технический вестник СПбГУ ИТМО. – 2010. – № 3 (67). – С. 5–11.
2. Коробейников А.Г., Гатчин Ю.А., Дукельский К.В., Ероньян М.А., Тер-Нерсесянц Е.В., Нестерова Н.А. Анализ совместимости фторсиликатных и боросиликатных слоев стекла для изготовления оптического волокна // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – № 3 (79). – С. 15–18.
3. Proctor B.A., Whitney I. and Johnson J.W. Strength of fused silica // Proc. Roy. Soc. London, ser. A. – 1967. – V. 297. – № 1451. – P. 534–557.
4. Давидович Н.М., Байкова Л.Г., Песина Т.И., Пух В.П., Радеева Е.И. Падение структурной прочности кварцевых волокон с полимерным покрытием под действием влажной среды // Физика и химия стекла. – 1990. – Т. 16. – № 4. – С. 566–570.
5. MacChesney J.B. Connor P.B., DiMarcello F.V., Simpson J.R., Lazay P.D. Preparation low loss optical fibers using simultaneous vapor phase deposition and fusion // Proc. 10th Int. Congr. Glass. – 1974. – P. 6.40–6.44.
6. MacChesney J.B. Connor P.B., Presby H.M. A new technique for preparation of low loss and graded index optical fibers // Proc. IEEE. – 1974. – V. 62. – № 9. – P. 1278–1279.
7. Бухтиарова Т.В., Дяченко А.А., Иноземцев В.П., Соколов А.В. Прочность и долговечность волоконно-оптических световодов // Итоги науки и техники. Связь. – Т. 8. Оптическая связь. – М.: ВИНИТИ, 1991. – С. 110–169.
8. Kalish D., Tariyal B.K. Static and dynamic fatigue of a polimer-coated fused silica optical fiber // J. Amer. Ceram. Soc. – 1978. – V. 61. – № 11–12. – P. 518–523.
9. Sakaguchi S., Nakahara M., Tajima Y. Drawing of high-strength long-length optical fiber // J. Non-Cryst. Solids. – 1984. – V. 64. – № 1–2. – P. 173–183.
10. Берштейн В.А. Механогидролитические процессы и прочность твердых тел. – Л.: Наука, Ленингр. отделение, 1987. – 317 с.
11. Hibino Y., Sakagucki S., Tajima J. Crack growth in silica glass under dynamic loading // J. Amer. Ceram. Soc. – 1984. – V. 67. – № 1. – P. 64.

12. Wiederhorn S.M., Bolz L.H. Stress corrosion and static fatigue of glass // J. Amer. Ceram. Soc. – 1970. – V. 53. – № 10. – P. 543.
13. Freiman S.W. Environmentally enhanced crack growth in glass // The Strength of Glass / ed. Kurkjian. – NY: Plenum press, 1985. – P. 197–215.
14. Duncan W.J., France P.W., Craig S.P. The effect of environment on the strength of optical fiber // The Strength of Glass / ed. Kurkjian. – NY: Plenum Press, 1985. – P. 309–326.
15. Богатырев В.А., Бубнов М.М., Румянцев С.Д., Семенов С.Л. Механическая надежность волоконных световодов // Труды ИОФАН. – М.: Наука, 1990. – Т. 23. – С. 66–93.
16. Kurkjian C.R. Statistics of the tensile strength of glass fibers for optical communication // Proc. 1st Intl. Otto Schott Coll. – Jena, DDR, 1978. – P. 379–387.
17. Богатырев В.А., Бубнов М.М., Дианов Е.М. Исследование механической прочности волоконных световодов для систем оптической связи // Квантовая электроника. – 1981. – Т. 8. – № 4. – С. 844–852.
18. Weibull W. A Statistical distribution function of wide applicability // J. Appl. Mech. – 1951. – V. 18. – № 9. – P. 293–297.
19. Леко В.К., Мазурин О.В. Свойства кварцевого стекла. – Л.: Наука, 1985. – 166 с.
20. Cohen M.I., Melliar-Smuth C.M. Recent advances in the fabrication of silica optical fibers // Int. Conf. Communic. – Seattle, Wash, 1980. – P. 55.1.1–55.1.7.
21. Богатырев В.А., Бубнов М.М., Вечканов Н.Н., Гурьянов А.Н., Семенов С.Л. Прочность стеклянных волоконных световодов большой длины // Труды ИОФАН. Волоконная оптика. – М.: Наука, 1987. – Т. 5. – С. 60–72.
22. Craig S.P., Duncan W.J., France P.W., Snodgas J.E. The strength and fatigue of large flaws in silica optical fiber // In Proc. of 8th ECOC. – Cannes, France, 1982. – P. 205–208.
23. Богатырев В.А., Бубнов М.М., Вечканова Н.Н. и др. Высокопрочные волоконные световоды, изготовленные методом химического осаждения из газовой фазы // Квантовая электроника. – 1982. – Т. 9. – № 7. – С. 1506–1509.
24. Шульц П. Производство оптических световодных волокон: Процессы и аппараты // Стеклообразное состояние. – Л., 1983. – С. 186–197.
25. Boniort J., Lboueq J., Bacle P. Improvement of optical fiber strength for submarine cables // SPJF. Optical fiber characteristics and standarts. – 1985. – P. 93–97.
26. Hanafusa H., Sakaguchi S., Hibino J. High-strength, long-length optical fibers // Rev. of the Elect. Commun. Laboratories. – 1985. – V. 33. – № 6. – P. 971–975.
27. Blyler L.L, DiMarcello F. V. Fiber drawing, coating and jacketing // Proc. IEEE. – 1980. – V. 68. – P. 1194–1198.
28. Sakaguchi S. Drawing of high-strength long-length optical fibers for submarine cables // J. Light wave techn. – 1984. – V. 2. – № 6. – P. 809–815.
29. Runk R. A zirconia induction furnace for drawing precision silica waveguides // Tecn. Digest Top. Meet. Opt. Fiber Comm. – Williamsburg, 1977. – P. TuB5-1.
30. DiMarcello F.V., Hart A.C., Williams J.C., Kurkjian C.R. High strength furnace-drawn optical fibers // Fiber Optics. Advances in Research and Development / ed. Bendow B. and Mitra S.S. – NY: Plenum, 1979. – P. 125–135.
31. Paec U.C., Schroeder C.M. Silica coated dual-tube zirconia induction furnace for high-strength fiber production // Electron. Lett. – 1986. – V. 22. – № 2. – P. 72–73.
32. Kaiser P., Hart A.C., Blyler L.L. Low-loss FEP-clad silica fibers // Appl. Opt. – 1975. – V. 14. – P. 156–162.
33. Jaeger R.E. Laser drawing of optical fibers // Am. Ceram. Soc. Bull. – 1976. – V. 55. – P. 270–273.
34. Paek U.C. Laser drawing of optical fibers // Appl. Opt. – 1974. – V. 13. – P. 1383–1386.
35. Белов А.В., Бубнов М.М., Гурьянов А.Н. и др. Вытяжка стеклянных волоконных световодов с помощью CO_2-лазера // Квантовая электроника. – 1978. – Т. 8. – С. 1169–1170.
36. Hiroshi Murata. Manufacturing of optical fibers in Japan // In Optical Fiber Communications. – Orlando, San Diego, New York, London, Toronto, Monreal, Sydney, Tokyo: Academic press, Inc., 1985. – V. 1. – P. 297–353.
37. Paek U.C., Kurkjian C.R. Calculation of cooling rate and induced stresses in drawing of optical fibers // J. Am. Ceram. Soc. – 1975. – V. 58. – P. 330.
38. Rongved L. Kurkjian C.R., Geyling F.T. Mechanical tempering of optical fibers // J. Non-Cryst. Solids. – 1980. – V. 42. – P. 579–584.
39. Oh S.M., Predieux P.H., Glavas X.G. Calculation of cooling rate in drawing of optical fibers // Opt. Letts. – 1982. – V. 7. – № 5. – P. 241–243.
40. Kurkjian C.R., Inniss D. Understanding mechanical properties of light guides: a commentary // Journal of SPIE. – 1991. – V. 30. – № 6. – P. 681–689.
41. Glaesemann G.S., Walter D.J. Method for obtaining long-length strength distributions for reliability prediction // Journal of SPIE. – 1991. – V. 30. – № 6. – P. 746–748.

42. Богатырев А.В., Бубнов М.М., Дианов Е.М., Румянцев С.Д., Семенов С.П. Прочность световодов в металлическом покрытии // Радиотехника. – 1988. – № 9. – С. 82–83.
43. Заявка Японии № 58-74543. Металлическое покрытие на волокне от 28.10.81, опубл. 06.05.83.
44. Богатырев А.В., Бубнов М.М., Румянцев С.Д., Семенов С.П. Механическая надежность волоконных световодов // Труды ИОФАН. – М.: Наука, 1990. – Т. 23. – С. 66–93.
45. Bogatyrjov V.A., Bubnov M.M., Dianov E.M., Makarenko A.Y., Rumyantsev S.D., Semjonov S.L., Sysoljatin A.A. High-strength hermetically tin-coated optical fibers // OFC (Optical fiber communic). – 1991. – P. 115.
46. DiMarcello F.V., Huff R.G., Lemaire P.J., Walker L.K. Hermetically sealed optical fibers. U.S. Pat. № 5000541, Mar. 19, 1991.
47. Lu K.T., Gleasemann G.S., Vandewoestine R.V., Kar G. Recent developments in hermetically coated optical fiber // J. Lightwave technol. – 1988. – V. 6. – № 2. – P. 240–241.
48. Ritter J.E. Service T.H., Jakus K. Predicted static fatigue behavior of specially coated optical glass fibers // J. Am. Ceram. Soc. – 1988. – V. 71. – № 11. – P. 988–992.
49. Kurkjian C.R., Krause J.T., Matthewson M.J. Strength and fatigue of silica optical fibers // J. Ligtwave Technology. – 1989. – V. 7. – № 9. – P. 1360–1370.

Коробейников Анатолий Григорьевич – Санкт-Петербургский филиал Института земного магнетизма, ионосферы и распространения радиоволн им. Н.В. Пушкова РАН, доктор технических наук, профессор, зам. директора по науке, Korobeynikov_A_G@mail.ru

Гатчин Юрий Арменакович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, профессор, зав. кафедрой, Gatchin@mail.ifmo.ru

Дукельский Константин Владимирович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат технических наук, доцент, KDukel@GOI.ru

Тер-Нерсесянц Егише Вавикович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат технических наук, доцент, volokno@goi.ru

УДК 535.314

ОПРЕДЕЛЕНИЕ ОПТИЧЕСКИХ ПОТЕРЬ ПРИ СТЫКОВКЕ СВЕТОВОДОВ С РАЗЛИЧНЫМ ДИАМЕТРОМ МОДОВОГО ПОЛЯ

А.В. Варламов, А.В. Куликов, В.Е. Стригалев, С.В. Варжель, С.М. Аксарин

Для минимизации оптических потерь при создании оптоволоконных систем проведено измерение углового распределения выходного излучения двулучепреломляющих оптических волокон с эллиптической напрягающей оболочкой с 4, 12, 16 и 18 мол.% GeO_2. Полученные данные позволяют прогнозировать потери при стыковке волокон с эллиптической напрягающей оболочкой с различной концентрацией GeO_2 между собой, а также со стандартными телекоммуникационными волокнами типа SMF-28, другими двулучепреломляющими световодами и элементами интегральной оптики.

Ключевые слова: диаметр модового поля, стыковка оптических волокон, двулучепреломление.

Введение

При создании распределенных волоконно-оптических фазовых интерферометрических датчиков (ФИД) на основе волоконных брэгговских решеток (ВБР) [1] необходимо подобрать оптические волокна (ОВ), обладающие, с одной стороны, достаточной фоторефрактивностью для записи в них решеток Брэгга [2], с другой стороны, приемлемыми для той или иной системы линейными оптическими потерями. Кроме того, для компоновки ФИД на основе ВБР необходимы данные по оптическим потерям при стыковке различных типов ОВ. Это позволит минимизировать суммарные оптические потери создаваемого измерительного комплекса. Оптические потери на стыковке ОВ можно определить, измерив диаметры модового поля световодов. В работе исследуются уникальные анизотропные одномодовые волоконные световоды (АОВС) с эллиптической напрягающей оболочкой (ESC), изготовленные по технологии [3, 4], обладающие высокой фоторефрактивностью для эффективной записи в них решеток Брэгга. Для увеличения фоторефрактивности исследуемых ОВ молярная концентрация диоксида германия (GeO_2) в сердцевине увеличена до 12, 16 и 18% по сравнению с 4% для «стандартного» АОВС с ESC. Актуальность работы состоит в том, что в литературе нет данных по измерению диаметра модового поля для исследуемых типов ОВ.

Зависимость оптических потерь на стыке световодов от диаметра модового поля

Величина потерь на стыке соединений определяется выражением [5]

$$D_{общ} = D_{ор} + D_{дм} + D_{ую} + D_{э} + D_{рпп} \text{ (дБ)}, \quad (1)$$

где $D_{ор}$ – потери из-за осевого рассогласования модовых полей; $D_{дм}$ – потери из-за разницы диаметров модовых полей; $D_{ую}$ – потери из-за погрешности угловой юстировки осей ОВ; $D_{э}$ – потери, обусловленные эллиптичностью модового поля; $D_{рпп}$ – потери из-за разницы показателей преломления.

В настоящей работе представлены результаты исследования оптических потерь из-за разницы диаметров модовых полей, поскольку они вносят определяющий вклад в суммарную величину потерь при стыковке волокон с различным содержанием GeO_2 в сердцевине. Основная часть оптического излучения, распространяющегося по волокну, локализована в его сердцевине, а остальная часть излучения распространяется в оболочке. По этой причине диаметр модового поля больше, чем физический диаметр сердцевины волокна. Диаметр модового поля является функцией длины волны источника, радиуса сердцевины волокна и профиля показателя преломления волокна. Таким образом, для оптических волокон с одинаковым физическим диаметром сердцевины диаметр модового поля может быть различным. Полагая, что поле моды в волокне аппроксимируется функцией Гаусса, потери на стыке одномодовых анизотропных волокон с различными диаметрами поля мод при условии согласования осей можно представить как [6]

$$D_{дм} = -10\log_{10}\left(\frac{2w_{0x}w'_{0x}}{w_{0x}^{2} + w'^{2}_{0x}} \cdot \frac{2w_{0y}w'_{0y}}{w_{0y}^{2} + w'^{2}_{0y}}\right), \quad (2)$$

где w_{0x} и w_{0y} – радиусы поля моды на уровне $1/e^2$ интенсивности оптического излучения на ортогональных осях для одного волокна; w'_{0x} и w'_{0y} – радиусы поля моды на уровне $1/e^2$ интенсивности оптического излучения на ортогональных осях для второго волокна.

С учетом малых отличий радиусов поля моды, измеренных для быстрой и медленной осей [7], формула (2) примет вид [6]

$$D_{дм} = -10\log_{10}\left(\frac{2w_{0}w'_{0}}{w_{0}^{2} + w'^{2}_{0}}\right)^{2}. \quad (3)$$

Определение диаметра модового поля световодов и оптических потерь при их стыковке

На рис. 1 представлена разработанная схема измерительной установки.

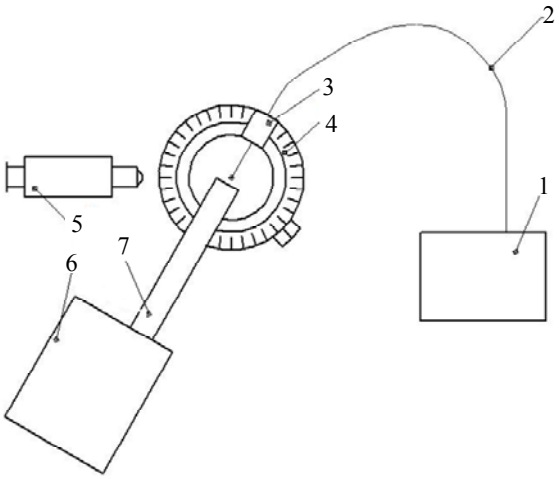


Рис. 1. Экспериментальная установка: 1 – источник оптического излучения; 2 – исследуемое волокно; 3 – волоконный вращатель; 4 – микроскоп с градуированным, поворачивающимся приборным столиком; 5 – микроскоп, перпендикулярный оси вращения поворотного столика; 6 – измеритель оптической мощности; 7 – экран

В качестве источника оптического излучения 1 применен лазерный диод с длиной волны излучения 1,55 мкм и мощностью 1 мВт. К источнику подсоединен исследуемый образец ОВ. Свободный конец световода 2 скалывается под углом 90º к оси волокна и крепится на волоконный вращатель 3, обеспечивающий вращение волокна вокруг своей оси. Вращатель закреплен на поворачивающемся приборном столике микроскопа. Микроскоп 4 используется для точной юстировки торца волокна и оси вращения столика. Перпендикулярно оси вращения поворотного столика расположен второй микроскоп 5, позволяющий установить оси двулучепреломляющих ОВ в одинаковое положение относительно оси вращения волокна. Одинаковое расположение осей позволяет не учитывать коэффициент поляризационной чувствительности измерителя оптической мощности и проводить измерения волокон в одинаковых условиях.

На расстоянии 1 см от торца волокна расположен измеритель оптической мощности 6 с экраном 7. Таким образом, в данной работе для определения диаметра модового поля световодов в дальнем поле перемещается конец волокна, а измеритель оптической мощности неподвижен, что позволяет достичь высокой точности измерений без использования компьютерного управления, как это предложено в работе [7], в которой используется установка с подвижным датчиком, управляемым компьютером.

Приборный столик микроскопа поворачивался с шагом 1º, при этом измерялась интенсивность оптического излучения. Экспериментальные графики зависимости нормированной интенсивности оптического излучения от угла для «быстрой» оси двулучепреломляющих волокон представлены на рис. 2. Исследования проведены на образцах АОВС с ESC, сердцевина которых легирована 4, 12, 16 и 18 мол.% GeO_2. Для сравнения исследованы ОВ широкого промышленного производства типа PANDA и SMF-28 [8].

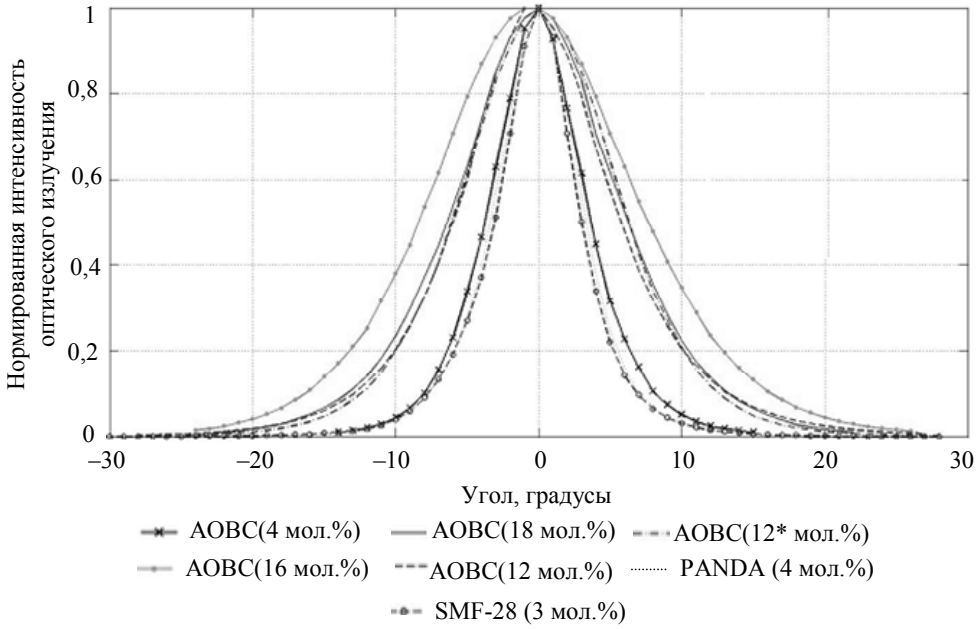


Рис. 2. Зависимость нормированной интенсивности излучения от угла; в скобках указана молярная концентрация GeO_2 в сердцевине ОВ

Тип волокна	АОВС					PANDA	SMF-28
Содержание GeO_2, мол.%	4	18	12*	16	12	4	3
Радиус модового поля, мкм	3,82	2,34	2,48	1,89	2,39	4,25	4,32

* – образец был изготовлен по модифицированной технологии MCVD (modified chemical vapor deposition) [9]

Таблица 1. Радиусы модовых полей исследованных типов ОВ, рассчитанные по формуле (4)

Стыкуемые волокна	**АОВС 4 мол.% GeO_2**	АОВС 18 мол.% GeO_2	АОВС 12 мол.% GeO_2 (*)	АОВС 16 мол.% GeO_2	АОВС 12 мол.% GeO_2	PANDA 4 мол.% GeO_2
АОВС 18 мол.% GeO_2	**1,01 (1,1–1,5)**					
АОВС 12 мол.% GeO_2 (*)	**0,79 (0,7–1,1)**	0,02				
АОВС 16 мол.% GeO_2	**1,99 (2,1–2,7)**	0,19	0,32			
АОВС 12 мол.% GeO_2	**0,93 (1,0–1,6)**	0,01	0,01	0,23		
PANDA 4 мол.% GeO_2	0,05	1,47	1,21	2,58	1,38	
SMF-28 3 мол.% GeO_2	0,07	1,54	1,27	2,67	1,45	0,01

* – образец был изготовлен по модифицированной технологии MCVD [9]

Таблица 2. Расчетные и измеренные (в скобках) потери при стыковке различных типов ОВ, дБ

Для исследованных образцов ОВ найдены телесные углы модового поля на уровне $1/e^2$, а также вычислены значения радиусов модового поля (табл. 1). Радиус модового поля вычислен по формуле [6]

$$w_0 = 2 \cdot \frac{\lambda}{\pi\theta_d} , \quad (4)$$

где w_0 – радиус модового поля; λ – длина волны; θ_d – телесный угол модового поля.

С учетом значений радиусов модового поля рассчитаны по формуле (3) теоретические потери при сварке различных типов ОВ (табл. 2). Измеренные значения потерь при стыковке приведены для соединений АОВС со «стандартным» содержанием GeO_2 4 мол.% с АОВС, сердцевина которых легирована 12, 16 и 18 мол.% GeO_2, предназначенных для записи ВБР. Как видно из таблицы, нижняя граница измеренных значений потерь близка к расчетной и зависит от степени легирования стыкуемых волокон GeO_2. Как и следовало ожидать, измеренные значения потерь несколько превышают расчетные, поскольку при расчете учтены лишь потери из-за различий диаметра модового поля $D_{дм}$, а иные виды потерь (1) не учтены.

Заключение

Для минимизации оптических потерь при создании оптоволоконных систем проведено измерение углового распределения выходного излучения анизотропных одномодовых волоконных световодов с эллиптической напрягающей оболочкой, сердцевина которых легирована 4, 12, 16 и 18 мол.% GeO_2. Полученные данные позволяют прогнозировать потери на стыковке оптических волокон между собой, а также со стандартными телекоммуникационными волокнами типа SMF-28, другими двулучепреломляющими световодами и элементами интегральной оптики. Измеренные значения потерь коррелируют с расчетными.

Литература

1. Варжель С.В., Стригалев В.Е. Метод устранения влияния сигнала помехи на чувствительность приема гидроакустической антенны на основе волоконных брэгговских решеток // Научно-технический вестник СПбГУ ИТМО. – 2010. – № 5 (69). – С. 5–8.
2. Варжель С.В., Куликов А.В., Асеев В.А., Брунов В.С., Калько В.Г., Артеев В.А. Запись узкополосных волоконных брэгговских отражателей одиночным импульсом эксимерного лазера методом фазовой маски // Научно-технический вестник СПбГУ ИТМО. – 2011. – Т. 75. – № 5. – С. 27–30.
3. Дукельский К.В., Ероньян М.А., Комаров А.В., Кондратьев Ю.Н., Ромашова Е.И., Серков М.М., Хохлов А.В. Тонкие анизотропные одномодовые волоконные световоды с эллиптической напрягающей оболочкой // Оптический журнал. – 2000. – Т. 57. – № 10. – С. 104–106.
4. Буреев С.В., Дукельский К.В., Ероньян М.А., Злобин П.А., Комаров А.В., Левит Л.Г., Страхов В.И., Хохлов А.В. Технология крупногабаритных заготовок анизотропных одномодовых световодов с эллиптической оболочкой // Оптический журнал. – 2007. – Т. 74. – № 4. – С. 85–87.
5. Каток В., Ковтун А., Руденко И. Оптоволокно: неразъемные соединения [Электронный ресурс]. – Режим доступа: http://citforum.ru/nets/hard/optonet/, свободный. Яз. рус. (дата обращения 02.11.2012).
6. Lefevre H. The Fiber-Optic Gyroscope. – London: Artech House, 1992. – 313 p.
7. Young M. Mode-field diameter of single-mode optical fiber by far-field scanning // Aplied Optics. – 1998. – V. 37. – № 24. – P. 5605–5619.
8. Mendez A., Morse T.F. Specialty Optical Fibers Handbook. – Burlington: Academic Press, 2007. – 250 p.
9. Варжель С.В., Ероньян М.А., Мешковский И.К., Стригалев В.Е. Минимизация оптических потерь одномодовых световодов с высокой концентрацией GeO_2 // Сборник трудов X международной конференции «Прикладная оптика-2012». – 2012. – С. 184–187.

Варламов Андрей Владимирович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, студент, b.a.b.92@bk.ru

Куликов Андрей Владимирович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, ассистент, a.kulikov86@gmail.com

Стригалев Владимир Евгеньевич – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат физ.-мат. наук, профессор, vstrglv@mail.ru

Варжель Сергей Владимирович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, vsv187@gmail.com

Аксарин Станислав Михайлович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, oskar2002@list.ru

УДК 535.399 (538.911)

ВЛИЯНИЕ ИОНОВ ЦЕРИЯ НА ПОГЛОЩЕНИЕ И ЛЮМИНЕСЦЕНЦИЮ МОЛЕКУЛЯРНЫХ КЛАСТЕРОВ СЕРЕБРА В СИЛИКАТНЫХ СТЕКЛАХ ПОСЛЕ ИОННОГО ОБМЕНА

И.А. Дёмичев, В.И. Егоров, Е.С. Постников, Е.М. Сгибнев, А.И. Сидоров, Т.А. Хрущева

Исследованы спектры люминесценции и поглощения силикатных стекол, содержащих серебро, введенное методом ионного обмена. Показано, что интенсивная люминесценция стекол возникает непосредственно после проведения ионного обмена, что связано с образованием нейтральных молекулярных кластеров серебра Ag_n. Показано, что при увеличении концентрации ионов церия в стекле происходит увеличение интенсивности люминесценции. Последующее ультрафиолетовое облучение стекол приводит к уменьшению интенсивности люминесценции. Термообработка при температуре 300–350ºС способствует усилению интенсивности люминесценции.

Ключевые слова: люминесценция, церий, ионный обмен, кластеры серебра.

Введение

Люминесцентные стекла, содержащие нейтральные атомы и молекулярные кластеры серебра Ag_n (n = 1–15), представляют практический интерес для создания люминофоров, спектральных преобразователей ультрафиолетового (УФ) излучения в солнечной энергетике, при разработке светодиодов белого свечения и люминесцентных волокон для волоконных датчиков и сенсоров. Известно, что нейтральные молекулярные кластеры Ag_n обладают интенсивной люминесценцией в видимой области спектра при возбуждении люминесценции излучением с длиной волны λ = 350–420 нм [1, 2]. До настоящего времени исследование люминесценции серебра в стеклах проводилось, в основном, в оксифторидных стеклах, в которых формирование нейтральных молекулярных кластеров Ag_n происходит непосредственно при варке стекла.

Фото-термо-рефрактивные (ФТР) стекла [3], содержащие серебро, фотосенсибилизатор – ионы церия Ce^{3+} и восстановитель – ионы сурьмы Sb^{5+}, являются удобной матрицей для синтеза и исследования оптических свойств атомарных и молекулярных ионов серебра, а также нейтральных молекулярных кластеров и наночастиц серебра. Изменение химического состава стекла, УФ облучение и термообработка (ТО) дают возможность управлять оптическими свойствами синтезированных кластеров и наночастиц. Так, исходно в ФТР стекле серебро находится в виде ионов и молекулярных ионов (Ag_n^{m+}), которые обладают очень слабой люминесценцией в видимой области спектра, а возбуждение их люминесценции происходит, в основном, УФ излучением. При УФ облучении в полосе поглощения ионов церия Ce^{3+} (λ = 305–315 нм) ионы церия отдают в стекло электроны, которые могут быть захвачены ионами серебра и ионами сурьмы. Это позволяет перевести часть ионов и молекулярных ионов серебра в нейтральное состояние. Последующая ТО при температуре выше температуры стеклования приводит к формированию нанокристаллов серебра, обладающих металлическими свойствами и плазмонным резонансом [4–6].

Метод ионного обмена (ИО) [7] позволяет сформировать вблизи поверхности стекла слой с высокой концентрацией серебра. При введении серебра в стекло происходит увеличение показателя преломления слоя стекла, содержащего серебро. При этом данный слой приобретает свойства градиентного оптического волновода [7]. Как показали наши предварительные эксперименты, часть серебра в этом слое находится не в виде ионов, а в виде нейтральных молекулярных кластеров, что обеспечивает интенсивную люминесценцию стекла в видимой области спектра при возбуждении люминесценции излучением с λ = 405 нм. Целью настоящей работы является исследование особенностей люминесценции молекулярных кластеров серебра в силикатных стеклах, после введения серебра методом ИО, а также влияния концентрации ионов церия, УФ облучения и ТО на люминесцентные свойства таких стекол.

Методика экспериментов

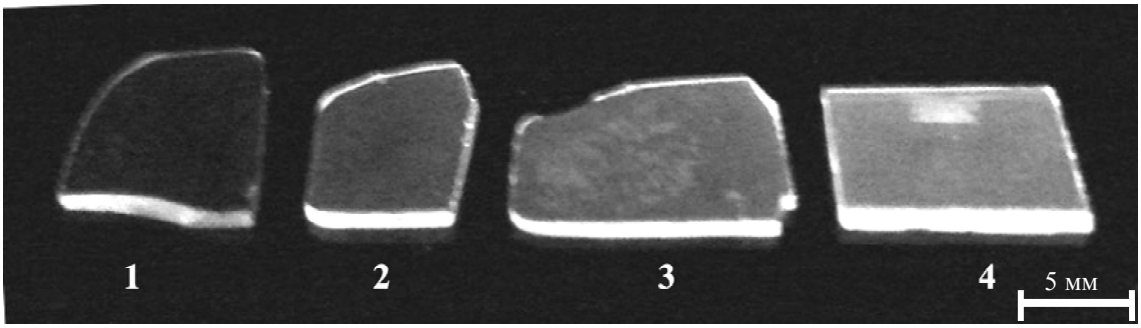
Для проведения исследований использовались силикатные стекла, изготовленные в НИУ ИТМО, и имеющие следующий основной состав: Na_2O-ZnO-Al_2O_3-SiO_2-NaF. Стекла содержали малые добавки фотосенсибилизатора – ионов церия в разной концентрации. Стекла подобного состава, содержащие ионы серебра, относятся к классу ФТР [3] и используются для записи объемных голограмм. Для синтеза стекол использовались особо чистые и химически чистые реактивы. Перед синтезом шихту стекла измельчали и гомогенизировали в шаровых мельницах (Pulverisette 5 Fritish). Стекла синтезировались в печах Gero в платиновых тиглях при температуре 1450°С с перемешиванием расплава платиновой мешалкой в воздушной атмосфере. Отжиг проводился при температуре 490°С с дальнейшим охлаждением по заданной программе (около 0,3 °С/мин). Температура стеклования была измерена на дифференциальном сканирующем калориметре STA6000 (Perkin–Elmer) и составляла 494°С. Образцы представляли собой полированные пластины толщиной 1,0–1,5 мм. Серебро в стекло вводилось методом ИО. Для этого стекло помещалось в расплав $AgNO_3$ (5%) + $NaNO_3$ (95%), находящийся в печи при температуре 315°С, на 15 мин. При ИО ионы Na^+ в приповерхностном слое стекла замещаются ионами Ag^+ [7]. Расчет пока-

зывает, что для указанного режима ионного обмена толщина слоя стекла, содержащего серебро, составляет 30 мкм.

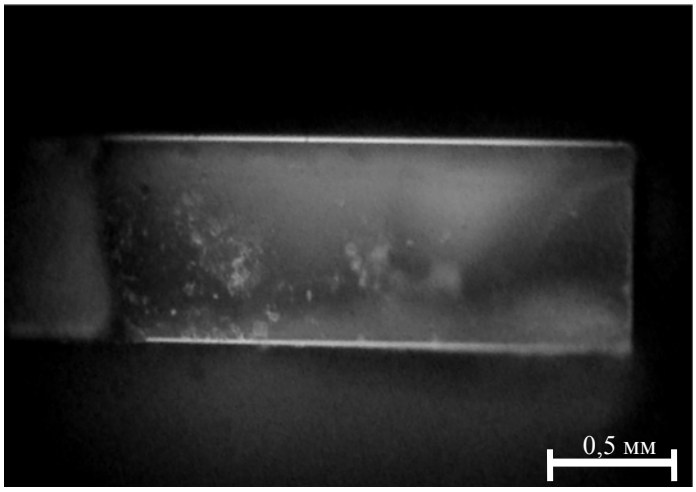
Для УФ облучения образцов использовалась ртутная лампа, имеющая полосы излучения в спектральном интервале 280–400 нм. УФ облучение проводилось при комнатной температуре в течение 20 мин. Расстояние от ртутной лампы до поверхности образца составляло 15 см. ТО образцов проводилась при температуре $t = 350°C$ в течение 2 часов в муфельной печи (Nabertherm) с программным управлением. Для исключения влияния продолжительных стадий нагрева и остывания печи образцы помещались в нагретую до нужной температуры печь и после термообработки извлекались из горячей печи. Измерения оптической плотности стекол проводились на спектрофотометре Cary500 (Varian) в спектральном интервале 300–800 нм с шагом 1 нм. Для измерения спектров люминесценции использовался волоконный спектрометр EPP2000-UVN-SR (StellarNet) с возбуждением люминесценции полупроводниковым лазером ($\lambda = 405$ нм). Регистрация спектров люминесценции и поглощения проводилась при комнатной температуре.

Экспериментальные результаты и обсуждение

На рис. 1, а, показана фотография люминесценции стекол с различной концентрацией ионов церия непосредственно после ИО. Фотография позволяет на качественном уровне оценить влияние ионов церия на люминесценцию стекол. Из рисунка видно, что стекло, не содержащее ионов церия, обладает относительно слабой люминесценцией. Увеличение концентрации ионов церия приводит к увеличению интенсивности люминесценции в видимой области спектра. На рис. 1, б, показана фотография люминесценции торца стекла после введения в него серебра методом ИО. Изображение получено с помощью оптического микроскопа. Из рис. 1 видно, что приповерхностные слои стекла, содержащие серебро, обладают интенсивной люминесценцией в видимой области спектра. Свечение объема стекла вызвано рассеяным излучением люминесценции, которое не было захвачено волноводными слоями. Наличие люминесценции в видимой области спектра при ее возбуждении излучением с длиной волны 405 нм указывает на формирование в стекле при ИО нейтральных молекулярных кластеров серебра. Известно (например, [8]), что люминесценция ионов и молекулярных ионов серебра возбуждается только УФ излучением с длиной волны менее 390 нм.



а



б

Рис. 1. Фотографии люминесценции: стекол с различной концентрацией оксида церия: 1 – 0 мол.%; 2 – 0,05 мол.%; 3 – 0,02 мол.%; 4 – 0,09 мол.%, с длиной волны возбуждения – 365 нм (а); торца стекла после введения в него серебра методом ИО, с длиной волны возбуждения – 405 нм (б)

На рис. 2 показаны спектры оптической плотности стекол после ИО для различной концентрации ионов церия. Из рисунка видно, что ИО приводит к увеличению оптической плотности образцов в УФ и видимой областях спектра. Увеличение поглощения в УФ области спектра вызвано присутствием в стекле как ионов и молекулярных ионов серебра, так и нейтральных атомов и нейтральных молекулярных кластеров серебра Ag_n ($n = 2–5$), имеющих полосы поглощения в данной области спектра [8–11]. Увеличение поглощения в спектральном интервале 400–800 нм вызвано присутствием в стекле нейтральных молекулярных кластеров серебра Ag_n ($n = 2–8$) [9–12]. Молекулярные кластеры серебра с $n > 5$ могут иметь различную геометрию [12], что вызывает смещение полос поглощения. Наложение полос поглощения кластеров различного типа приводит к появлению монотонной спектральной зависимости оптической плотности в интервале 400–800 нм.

Образование нейтральных атомов и нейтральных молекулярных кластеров в стекле при ИО связано с тем, что положительные ионы серебра в стекле в процессе ИО могут захватывать электроны с дефектов сетки стекла, например, с немостиковых ионов кислорода [7]. При дальнейшей термической диффузии атомов и ионов серебра в процессе ИО их взаимодействие приводит к образованию заряженных (Ag_n^{m+}) и нейтральных (Ag_n) молекулярных кластеров серебра. Увеличение концентрации ионов церия вызывает рост поглощения в видимой области спектра. Следовательно, присутствие ионов церия в стекле увеличивает эффективность ИО. Это приводит к увеличению концентрации ионного серебра в приповерхностном слое стекла и к повышению вероятности образования нейтральных молекулярных кластеров серебра. Причиной данного эффекта может быть увеличение неоднородностей (дефектов) в стекле при увеличении концентрации ионов церия. Увеличение концентрации дефектов приводит к увеличению подвижности ионов натрия и серебра, а также к росту концентрации ловушек для электронов, с которых эти электроны могут быть захвачены ионами серебра.

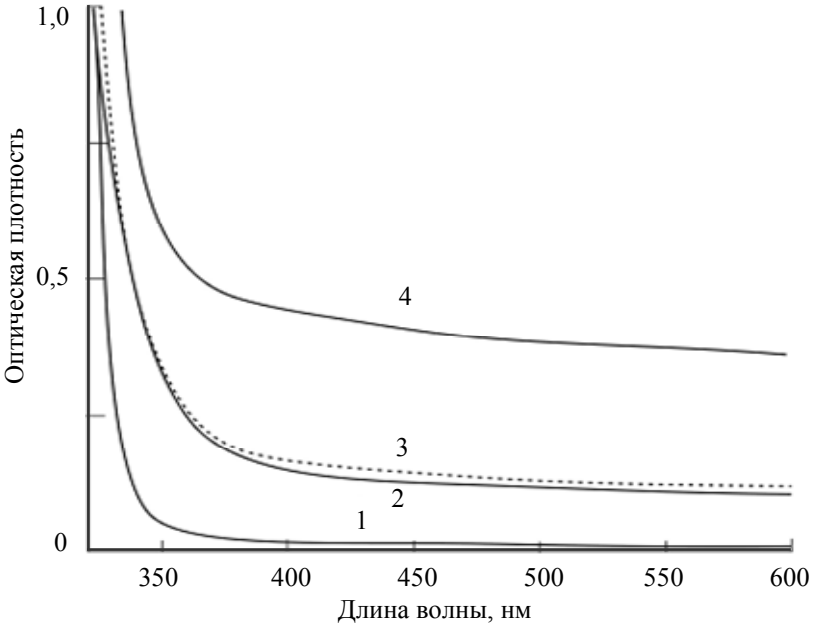


Рис. 2. Экспериментальные спектры оптической плотности стекол после ИО для различной концентрации оксида церия: кривая 1 – исходное стекло; кривая 2 – 0,02 мол.%; кривая 3 – 0,05 мол.%; кривая 4 – 0,09 мол.%

На рис. 3 представлены спектры люминесценции стекол с различной концентрацией ионов церия при возбуждении люминесценции излучением с $\lambda_{возб} = 405$ нм. Контрольные измерения, проведенные со стеклами, не содержащими серебра, показали отсутствие в них люминесценции. Из рисунка видно, что широкополосная люминесценция в спектральном интервале 500–750 нм возникает в стеклах непосредственно после ИО. Максимум полосы люминесценции приходится на длину волны 570 нм.

Большая спектральная ширина полосы люминесценции указывает на присутствие в стекле нескольких люминесцентных центров. Сравнение спектров люминесценции с данными, приведенными в работах [8, 10, 11], позволяет определить вклад различных молекулярных кластеров серебра в интегральную полосу люминесценции. Вклад в люминесценцию в спектральном интервале 450–500 нм вносят нейтральные молекулярные кластеры Ag_2 и Ag_4. В этом же спектральном интервале может происходить люминесценция атомарного серебра, однако она возбуждается только УФ излучением.

За люминесценцию в спектральном интервале 600–800 нм отвечают, в основном, молекулярные кластеры Ag_3. Увеличение концентрации ионов церия приводит к увеличению интенсивности люминесценции, что указывает на рост концентрации люминесцентных центров. Вставка в рис. 3 иллюстрирует

влияние предварительного УФ облучения на интенсивность люминесценции стекол после ИО. Из рисунка видно, что предварительное УФ облучение, проведенное непосредственно перед ИО, приводит к увеличению интенсивности люминесценции. Причиной эффекта является то, что при УФ облучении в полосе поглощения ионов церия происходит фотоионизация этих ионов. Свободные электроны, возникшие в процессе фотоионизации, могут быть захвачены дефектами стекла. Положительные ионы серебра, мигрирующие по стеклу при ИО, в свою очередь, захватывают эти электроны и переходят в нейтральное состояние. Это способствует увеличению концентрации нейтральных молекулярных кластеров серебра.

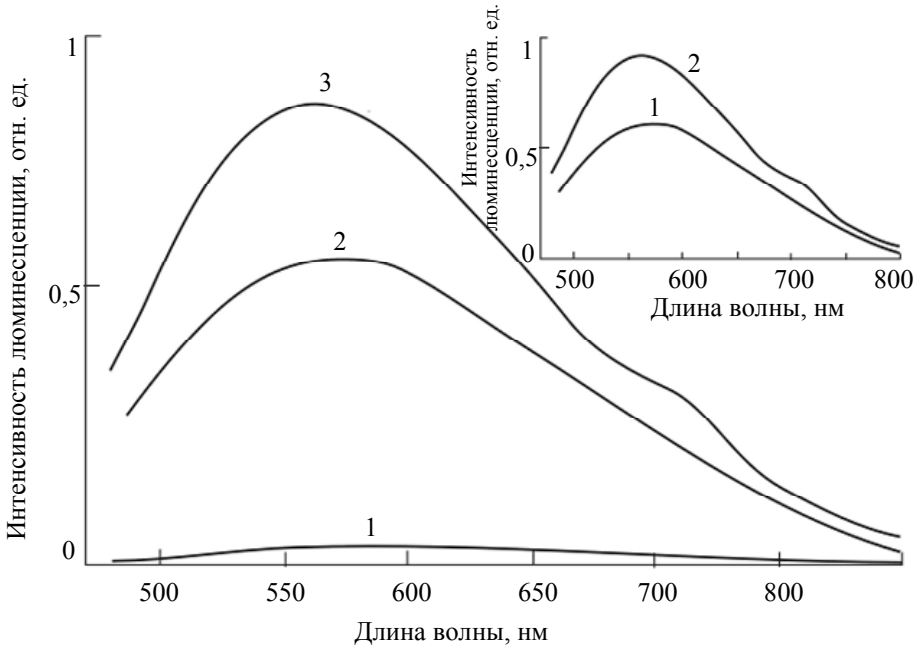


Рис. 3. Экспериментальные спектры люминесценции стекол после ИО для различной концентрации оксида церия: кривая 1 – 0,02 мол.%; кривая 2 – 0,05 мол.%; кривая 3 – 0,09 мол.%. На вставке – влияние предварительного УФ облучения: кривая 1 – без УФ облучения; кривая 2 – с УФ облучением. Длина волны возбуждения 405 нм

На рис. 4 показаны спектры оптической плотности стекол после ИО, УФ облучения и ТО. Из рисунка видно, что для всех образцов ТО приводит к появлению полосы поглощения в спектральном интервале 390–420 нм. Данная полоса поглощения связана с формированием в стекле наночастиц металлического серебра, имеющих плазмонный резонанс в указанной области спектра [4–6, 13].

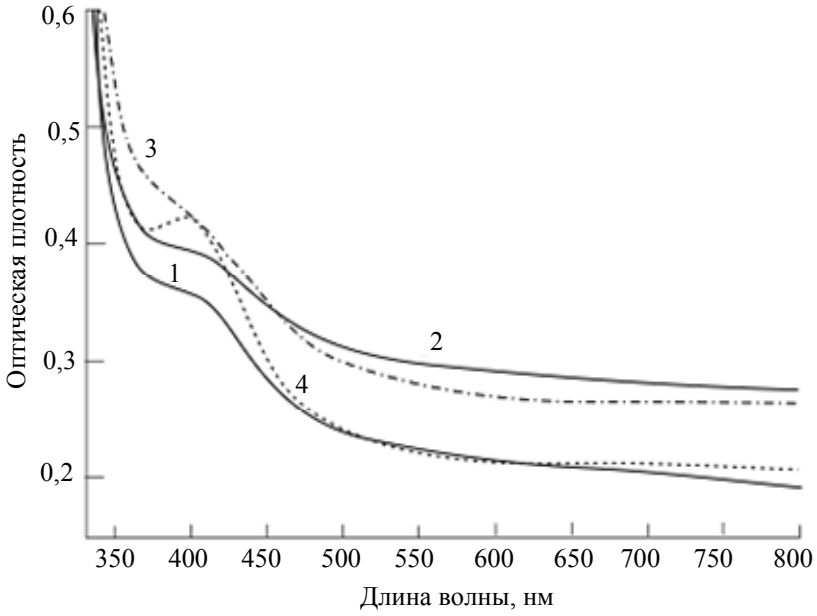


Рис. 4. Экспериментальные спектры оптической плотности стекол после ИО, УФ облучения и ТО для различной концентрации оксида церия: кривая 1 – 0,02 мол.%; кривая 2 – 0,05 мол.%; кривая 3 – 0,09 мол.%; кривая 4 – 0,05 мол.% с предварительным УФ облучением

Для образца, облученного УФ до ИО, плазмонная полоса поглощения имеет максимальную амплитуду. Это вызвано тем, что предварительное УФ облучение увеличивает концентрацию нейтральных молекулярных кластеров в стекле, которые являются центрами кристаллизации и роста наночастиц серебра.

На рис. 5 показаны спектры люминесценции стекол после ИО, УФ и ТО. Из рисунка видно, что и после ТО сохраняется описанная выше тенденция: рост концентрации ионов церия приводит к увеличению интенсивности люминесценции. Для сравнения на рисунке приведен спектр люминесценции ФТР стекла аналогичного состава, в которое серебро вводилось при варке стекла, после УФ облучения и ТО. Из сравнения кривых 3 и 4 на рис. 5 видно, что при вводе серебра в стекло методом ионного обмена полоса люминесценции уширяется, а ее максимум смещается в длинноволновую область спектра. Это указывает на возникновение различия в составе и концентрации нейтральных молекулярных кластеров серебра в зависимости от метода ввода серебра в стекло.

Вставка в рис. 5 иллюстрирует изменение интенсивности люминесценции стекла после УФ облучения и после ТО. Из рисунка видно, что УФ облучение приводит к некоторому уменьшению интенсивности люминесценции. Последующая ТО приводит к существенному увеличению интенсивности люминесценции. Причиной уменьшения интенсивности люминесценции после УФ облучения может быть то, что при УФ облучении происходит фотоионизация атомов и молекулярных кластеров серебра, имеющих большое количество полос поглощения в УФ области спектра. Это приводит к уменьшению концентрации таких молекулярных кластеров и уменьшению интенсивности люминесценции. В то же время свободные электроны, образовавшиеся в результате фотоионизации, могут быть захвачены заряженными молекулярными кластерами серебра, имеющими интенсивные полосы поглощения в видимой области спектра, но слабую люминесценцию при возбуждении излучением с $\lambda_{возб}$ = 405 нм.

Увеличение интенсивности люминесценции после ТО связано с тем, что ионы сурьмы, получившие дополнительные электроны при УФ облучении, отдают их в стекло при нагреве. Это приводит к переводу положительно заряженных молекулярных кластеров серебра в нейтральное состояние. Увеличение концентрации нейтральных молекулярных кластеров серебра проявляется в увеличении интенсивности люминесценции стекла в видимой области спектра. При этом наночастицы серебра также могут вносить вклад в люминесценцию стекла, так как они обладают люминесценцией в спектральном интервале 600–700 нм [14, 15].

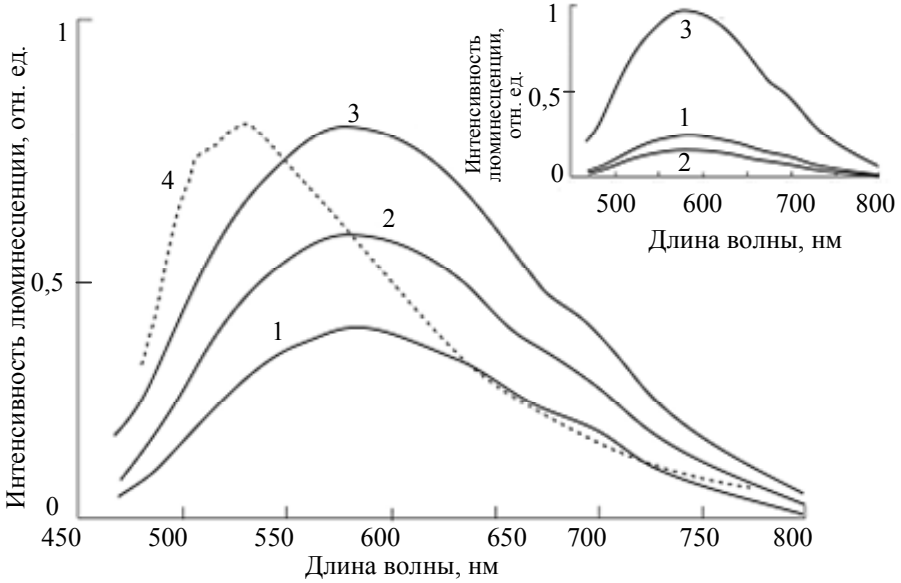


Рис. 5. Экспериментальные спектры люминесценции стекол после ИО, УФ облучения и ТО для различной концентрации оксида церия: кривая 1 – 0,02 мол.%; кривая 2 – 0,05 мол.%; кривая 3 – 0,09 мол.%; кривая 4 – фото-термо-рефрактивное стекло. На вставке: кривая 1 – после ИО; кривая 2 – после УФ облучения; кривая 3 – после ТО для концентрации оксида церия 0,09 мол.%. Длина волны возбуждения 405 нм

Заключение

Представленные результаты показывают, что при вводе серебра в силикатное стекло методом ионного обмена часть серебра формирует нейтральные молекулярные кластеры Ag_n (n = 2–6). Это приводит к увеличению поглощения стекла в видимой области спектра и появлению интенсивной широкополосной люминесценции в видимом диапазоне при возбуждении излучением с длиной волны 405 нм. Увеличение концентрации ионов церия в стекле приводит к увеличению поглощения в видимой области спектра и увеличению интенсивности люминесценции. Ультрафиолетовое облучение после ионного об-

мена приводит к уменьшению интенсивности люминесценции, что связано с переходом части нейтральных молекулярных кластеров серебра в заряженное состояние вследствие фотоионизации. Термообработка стекла при температуре ниже температуры стеклования существенно увеличивает интенсивность люминесценции. Полученные результаты могут быть использованы при разработке люминесцентных волноводных и волоконных датчиков и сенсоров.

Работа выполнена при поддержке РФФИ (грант № 12-02-31455), а также ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009–2013 годы (Соглашение № 14.132.21.1689, Минобрнауки РФ).

Литература

1. Kuznetsov A.S., Tikhomirov V.K., Moshchalkov V.V. Polarization memory of white luminescence of Ag nanoclusters dispersed in glass host // Optics Express. – 2012. – V. 20. – № 19. – P. 21576–21582.
2. Колобкова Е.В., Никоноров Н.В., Сидоров А.И., Шахвердов Т.А. Люминесценция молекулярных кластеров серебра в оксифторидных стеклах // Оптика и спектроскопия. – 2013. – Т.114. – С. 93–97.
3. Glebov L.B. Photochromic and photo-thermo-refractive glasses. Encyclopedia of smart materials / Ed. by M. Schwartz. – John Willey & Sons, 2002. – V. 2. – 770 p.
4. Kreibig U., Vollmer M. Optical properties of metal clusters. – Berlin: Springer-Verlag, 2010. – 532 p.
5. Nikonorov N.V., Sidorov A.I., Tsekhomskii V.A. Silver nanoparticles in oxide glasses: technologies and properties // Silver nanoparticles / Ed. by D.P. Perez. – Vukovar: In-Tech, Croatia, 2010. – P. 177–201.
6. Климов В.В. Наноплазмоника. – М.: ФИЗМАТЛИТ, 2009. – 480 с.
7. Tervonen A., West B.R., Honkanen S. Ion-exchanged glass waveguide technology: a review // Opt. Engineering. – 2011. – V. 50. – № 7. – P. 071107-1–071107-15.
8. Zheng W., Kurobori T. Assignments and optical properties of X-ray-induced colour centres in blue and orange radiophotoluminescent silver-activated glasses // Journal of Luminescence. – 2011. – V. 131. – № 1. – P. 36–40.
9. Ozin G.A., Huber H. Silver atoms and small silver clusters stabilized in zeolite Y: optical spectroscopy // Inorganic Chem. – 1978. – V. 17 – № 1. – P. 155–163.
10. Fedrigo S., Harbich W., Buttet J. Optical response of Ag2, Ag3, Au2, and Au3 in argon matrices // J. Chem. Phys. – 1993. – V. 99. № 8. – P. 5712–5717.
11. Felix C., Sieber C., Harbich W., Buttet J., Rabin I., Schulze W., Ertl G. Fluorescence and excitation spectra of Ag4 in an argon matrix // Chem. Phys. Lett. – 1999. – V. 313. – P. 105–109.
12. Lecoultre S., Rydlo A., Buttet J., Félix C., Gilb S., Harbich W. Ultraviolet-visible absorption of small silver clusters in neon: Agn (n = 1–9) // J. Chem. Phys. – 2011. – V. 134. – P. 184504–184510.
13. Колобкова Е.В., Никоноров Н.В., Асеев В.А. Влияние серебра на рост квантовых точек во фторофосфатных стеклах // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – № 5 (81). – С. 1–5.
14. Xu H., Suslick K.S. Sonochemical Synthesis of Highly Fluorescent Ag Nanoclusters // ACS Nano. – 2010. – V. 4. – P. 3209–3214.
15. Smith S.L., Nissamudeen K.M., Philip D., Gopchandran K.G. Studies on surface plasmon resonance and photoluminescence of silver nanoparticles // Spectrochim. Acta A. – 2008. – V. 71. – № 1. – P. 186–190.

Дёмичев Иван Алексеевич – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, demichevia@yandex.ru

Егоров Владимир Ильич – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, НИИ фотоники и информатики, инженер, egorovvl@gmail.com

Постников Евгений Сергеевич – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, НИИ фотоники и информатики, инженер, werrus@yandex.ru

Сгибнев Евгений Михайлович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, студент, sgibnevem@gmail.com

Сидоров Александр Иванович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, НИИ фотоники и информатики, профессор, aisidorov@qip.ru

Хрущева Татьяна Александровна – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, инженер; Санкт-Петербургский государственный политехнический университет, кандидат физ.-мат. наук, доцент, Khruta@qip.ru

УДК 535.016, 681.2.083

ПРИМЕНЕНИЕ ВРЕМЯПРОЛЕТНЫХ PMD-КАМЕР ДЛЯ ОПРЕДЕЛЕНИЯ ДАЛЬНОСТИ ДО ВОДНОЙ ПОВЕРХНОСТИ

Д.Ю. Крысин, А.В. Небылов

Представлены результаты, подтверждающие возможность применения времяпролетных PMD-камер для измерения малой дальности (0,2–5 м) до водной поверхности. Рассмотрены основные особенности технологии PMD. Приведено описание разработанной экспериментальной установки. Описан эксперимент, подтвердивший работоспособность PMD-камер при волнении водной поверхности в естественных условиях. Представлены экспериментальные 3D-изображения рельефа водной поверхности. Объясняется их специфический вид. Представлены графические результаты статистической обработки валидных пикселей последовательностей 3D-изображений. Рассмотрены безынерционные алгоритмы обработки 3D-изображений для получения оценки дальности до водной поверхности. Полученные результаты могут быть полезны при решении задач измерения уровня жидкости, определения параметров морского волнения и измерения высоты движения объекта над водной поверхностью.

Ключевые слова: времяпролетный метод, PMD-камера, 3D-изображение, измерение дальности, водная поверхность, экспериментальная установка, формирование изображений.

Введение

Проблема измерения малой дальности (0,2–5 м) до водной поверхности (ВП) является актуальной при решении задач измерения уровня жидкости, определения параметров морского волнения и измерения высоты движения объекта над водной поверхностью. Отсутствие соответствующих дальномеров, адекватных по цене и характеристикам, может стать серьезным препятствием при разработке перспективных систем управления подвижными объектами [1, 2].

В настоящее время активно развиваются методы технического зрения, предназначенные для бесконтактного измерения пространственного положения и формы объектов. Разнообразие и сфера применения этих методов непрерывно расширяются. Однако значительная их часть не может быть использована для измерения малых расстояний до ВП. Это объясняется особыми оптическими свойствами воды и ее поверхности. В работе [2] описывается метод, предназначенный для определения высоты и угловой ориентации объекта относительно невзволнованной ВП. В его основе лежит использование стереоскопической системы камер и свойства зеркального отражения ВП. В настоящей работе предлагается иной подход, позволяющий измерять дальность при волнении ВП, который базируется на использовании инновационных времяпролетных PMD-камер [3, 4].

Технология PMD

В основе технологии PMD (Photonic Mixer Device) лежит времяпролетный метод измерения расстояния [3]. Суть метода заключается в измерении времени, которое требуется световому сигналу для преодоления расстояния между камерой и отражающим свет объектом. Для этого времяпролетные PMD-камеры освещают сцену с помощью модулированного оптического излучения инфракрасного (ИК) диапазона с некоторой частотой модуляции f_{mod}. ИК-излучение отражается от объектов сцены и измеряется датчиком изображений камеры (светочувствительной матрицей). Далее специальная схема оценивает уровни корреляции и вычисляет фазовый сдвиг φ между излученным и принятым сигналами, который пропорционален дальности D до объекта:

$$D = \frac{c\varphi}{4\pi f_{mod}}, \qquad (1)$$

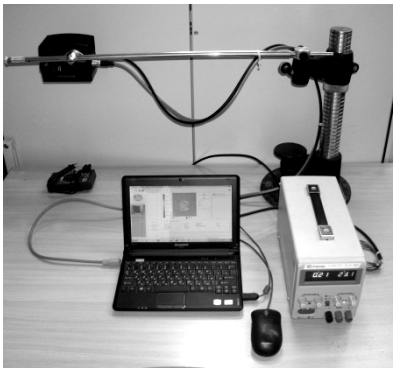
где c – скорость света. Сигнал с каждого элемента датчика изображений обрабатывается независимо. В результате получается изображение, являющееся трехмерной моделью поверхности объектов, находящихся в поле зрения камеры. Изображение представляет собой матрицу, количество элементов которой совпадает с количеством чувствительных элементов датчика изображений. Элементы матрицы содержат оценки дальности до соответствующих участков поверхности (в метрах). Первоначально дальность измеряется в радиальной системе координат с началом в центре камеры. Далее эти первичные данные могут быть переведены в декартову систему координат. В литературе можно встретить различные названия подобных изображений: карты дальности, карты глубин, 3D-изображения [4, 5]. В настоящей работе используется термин «3D-изображение».

Из приведенного описания следует, что PMD-камеры, в отличие от стереоскопических систем технического зрения, позволяют получать 3D-изображения, не прибегая к методам получения и обработки стереоскопических снимков. При этом, в отличие от лазерных сканеров, не используются движущиеся механические элементы, необходимые для развертки луча. Также не нужны сложные и требовательные к вычислительным ресурсам алгоритмы. Последние модели PMD-камер способны передавать управляющему устройству полностью готовые 3D-изображения с частотой до 90 Гц. Перечисленные факты обеспечивают возможность использования PMD-камер при построении систем, работающих практически в реальном времени.

Несмотря на очевидные достоинства, PMD-камеры имеют ряд ограничений. Из формулы (1) видно, что максимальная измеряемая дальность равна отношению скорости света к удвоенной частоте модуляции. При наличии в поле зрения камеры объектов, расстояние до которых больше, чем верхняя граница диапазона измерения, может возникнуть методическая погрешность, кратная максимальной измеряемой дальности. Систематическая погрешность измерения PMD-камер сравнительно мала [6]. Случайная составляющая может быть на порядок выше систематической и зависит от целого ряда внешних факторов. К ее увеличению приводят такие факторы, как малая амплитуда отраженного сигнала, высокая интенсивность внешнего освещения, движение объектов сцены, малое значение выдержки, использование вблизи другой времяпролетной камеры с той же частотой модуляции. Амплитуда отраженного сигнала в основном зависит от расстояния до объектов и их отражательной способности. Наличие в поле зрения камеры поверхностей с высоким коэффициентом отражения может привести к насыщению пикселей датчика и невозможности измерений. Множественные переотражения сигнала ухудшают точность, приводя к переоценке дальности. Несмотря на перечисленные ограничения, PMD-камеры являются одним из наиболее перспективных средств получения 3D-изображений.

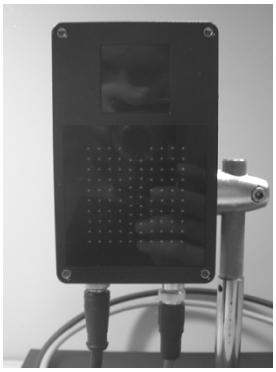
Экспериментальная установка

С целью исследования особенностей использования PMD-камер для измерения дальности до ВП была разработана установка (рис. 1, а), которая включает промышленную времяпролетную PMD-камеру O3D201 производства компании IFM (рис. 1, б, в), персональный компьютер (ПК), источник питания и штатив.



а

б



в

Рис. 1. Общий вид экспериментальной установки (а); виды PMD-камеры O3D201 (б, в)

PMD-камера O3D201 позволяет получать 3D-изображения и полутоновые изображения сцены (рис. 2), что дает возможность построения комплексных алгоритмов обработки информации. Наиболее важные с точки зрения эксперимента характеристики PMD-камеры O3D201 представлены в табл. 1.

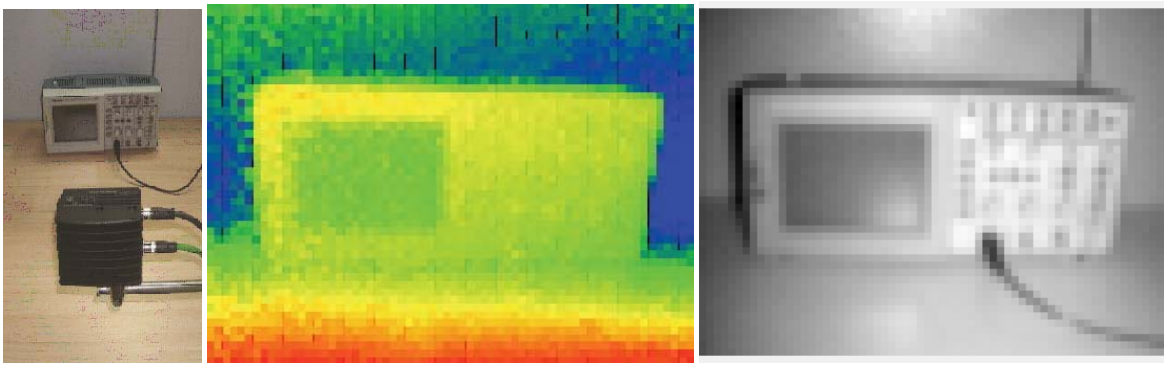
Параметры точности зависят от настроек камеры и ряда внешних факторов [7]. Типовые значения доверительных интервалов (±3σ), соответствующие измерению расстояния 4,5 м до поверхностей с различными коэффициентами отражения, имеют следующие значения:

– для белой поверхности (коэффициент отражения 0,9) – ±7 мм;
– для серой поверхности (коэффициент отражения 0,18) – ±20 мм;
– для черной поверхности (коэффициент отражения 0,06) – ±58 мм.

Тип датчика	PMD 3D chip
Тип подсветки	Массив ИК светоизлучающих диодов (850 нм)
Разрешение	50×64 пикселей
Углы обзора	30º×40º
Максимальная частота кадров	20 Гц
Интерфейс передачи данных	Ethernet 10Base-T/100Base-TX
Напряжение питания	24 В±10%

Таблица 1. Характеристики PMD-камеры O3D201

Для управления камерой с ПК используется приложение Efector pmd3d camera, доступное на сайте производителя [7]. Данное приложение обеспечивает настройку камеры, получение, визуализацию и сохранение 3D-изображений в памяти ПК.



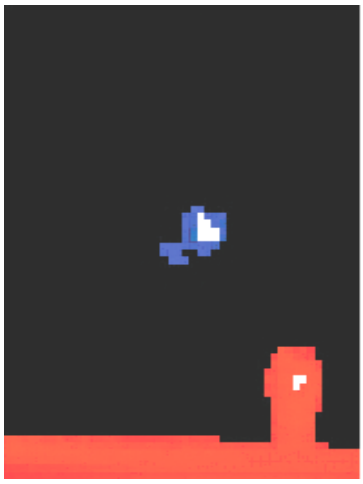
а б в

Рис. 2. PMD-камера O3D201 и сцена в лаборатории (а); 3D-изображение сцены (двумерная интерпретация – дальность закодирована в цвете) (б); полутоновое изображение сцены (в)

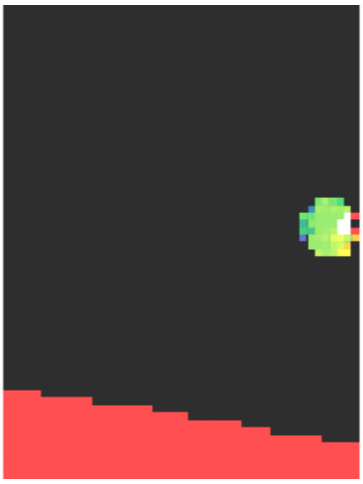
Эксперимент в естественных условиях

Авторами поставлен ряд экспериментов, направленных на определение возможности использования PMD-камер для измерения дальности до ВП и показавших положительные результаты. Опыты проводились как в лабораторных, так и в естественных условиях. Ниже приведено описание и результаты одного из экспериментов.

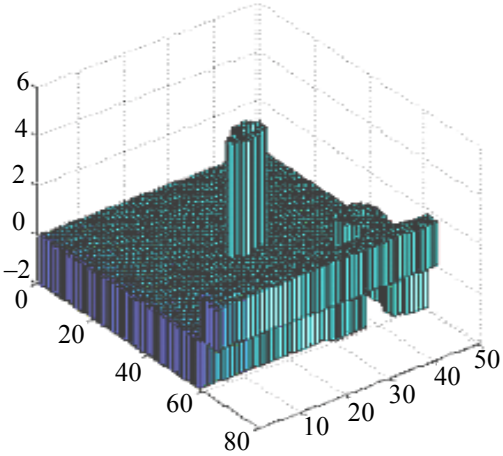
а



б



в



г

-2	-2	-2	-2	-2	-2	-2	-2	-2
-2	-2	-2	-2	-2	-2	-2	-2	-2
-2	-2	-2	-2	-2	-2	-2	-2	-2
-2	-2	-2	-2	-2	4,630	4,655	-2	-2
-2	-2	-2	-2	4,679	4,652	4,637	-2	-2
-2	-2	-2	-2	-2	4,620	4,642	4,699	-2
-2	-2	4,644	4,616	4,577	4,610	4,616	4,674	-2
-2	4,681	4,622	4,593	4,581	4,593	4,653	-2	-2
-2	4,659	4,605	-1	-1	4,591	4,610	-2	-2
-2	4,721	-1	-1	-1	4,593	4,609	-2	-2
-2	4,650	4,689	-1	-1	4,612	-2	-2	-2
-2	-2	4,625	4,714	4,684	4,642	-2	-2	-2
-2	-2	-2	-2	-2	-2	-2	-2	-2
-2	-2	-2	-2	-2	-2	-2	-2	-2
-2	-2	-2	-2	-2	-2	-2	-2	-2

д

Рис. 3. Натурный эксперимент: Санкт-Петербург, р. Фонтанка, Английский пешеходный мост (а); вид 3D-изображения ВП при горизонтальном положении камеры (б); вид 3D-изображения ВП при наклоне камеры (в); трехмерная визуализация 3D-изображения (г); фрагмент 3D-изображения (д)

Целью эксперимента было исследование особенностей функционирования PMD-камеры в естественных условиях при волнении ВП. Макет был установлен на мосту (рис. 3, а). Высота положения камеры относительно ВП была измерена с помощью отвеса и рулетки и составила 4,55 м. По оценке авторов, погрешность данного измерения не превышает 0,05 м. Во время эксперимента степень волнения ВП составляла 1 балл [8]. Установка была расположена со стороны моста ниже по течению, где водный поток является более возмущенным после прохождения вблизи опор моста.

Во время эксперимента были получены и записаны последовательности 3D-изображений ВП при различных углах наклона PMD-камеры. На рис. 3, б, в, показан вид этих изображений. На представленных изображениях дальность закодирована в цвете. Черные и белые участки изображений содержат невалидные пиксели, для которых не были получены оценки дальности. Черные пиксели не получили оценки дальности из-за недостаточной мощности отраженного сигнала, белые пиксели – из-за перехода в состояние насыщения вследствие повышенной интенсивности отраженного сигнала. В нижней части изображений зафиксирована конструкция моста, которая попала в поле зрения камеры из-за недостаточной длины выносной штанги. В процессе съемки использовалась выдержка (exposure time) 27,1 мс. Из рис. 3 видно, что полученные 3D-изображения ВП существенно отличаются от 3D-изображения, представленного на рис. 2. Принципиальное отличие заключается в том, что область валидных пикселей, содержащих оценки дальности, составляет всего несколько процентов от площади 3D-изображений ВП. При наклоне камеры эта область смещается к краям снимка.

На рис. 3, г, приведена трехмерная визуализация 3D-изображения, полученного с помощью камеры, находящейся в положении, близком к горизонтальному. На рис. 3, д, показан его фрагмент, включающий область валидных пикселей, содержащих оценки дальности до ВП в метрах. Невалидные пиксели имеют значения –1 и –2. Изображение не подвергалось обработке.

Формирование 3D-изображений ВП в PMD-камере

Полученные 3D-изображения не являются моделью ВП, находящейся в поле зрения камеры во время съемки. Это объясняется особыми оптическими свойствами воды и ее поверхности. Важной особенностью является то, что вода является оптически прозрачной средой, а на границе «воздух–вода» отражается только часть ИК излучения. При различных условиях может наблюдаться зеркальное, диффузное и смешанное отражение света. Так, представленные на рис. 2 предметы характеризуются диффузным отражением света.

Известно, что отражение света в значительной мере зависит от рельефа поверхности. Так, для возникновения зеркального отражения, в соответствии с критерием Релея [9], должно выполняться условие

$$\sigma < \frac{\lambda}{8\cos\theta}, \qquad (2)$$

где σ – среднеквадратичное значение шероховатости поверхности; λ – длина волны падающего электромагнитного излучения; θ – угол падения луча света.

Коэффициент зеркального отражения может быть вычислен помощью формулы Френеля [10]. Зависимость коэффициента отражения света от угла падения на ВП представлена на рис. 4, а. При вертикальном падении света коэффициент отражения составляет всего 0,02. Знание этой величины позволяет спрогнозировать необходимую мощность излучателя с учетом порога чувствительности датчика для обеспечения возможности измерений.

Если условие (2) не выполняется, возникает диффузное отражение. При этом пространственное распределение интенсивности отраженного света может быть предсказано только при наличии статистических данных о геометрии поверхности, которые, как правило, отсутствуют. Может также иметь место «смешанное» отражение, при котором проявляются свойства как зеркального, так и диффузного отражения.

Таким образом, характер отражения света сильно зависит от балльности волнения ВП. В естественных условиях при балльности 0–2 балла обычно преобладает зеркальное отражение света [11]. При более высокой балльности появляются волны с опрокидывающимися гребнями, что приводит к возникновению зон ВП с диффузным отражением света. Описанный авторами эксперимент с PMD-камерой проводился при волнении ВП порядка 1 балла. Следовательно, в эксперименте зеркальное отражение света преобладало на большей части ВП.

Расстояние между ИК излучателем и датчиком изображений в PMD-камере невелико по сравнению с измеряемой дальностью. ИК излучение отражается в направлении датчика только от ограниченной области ВП. Лучи света падают на нее под малыми углами (рис. 4, б). Это объясняет факт, что доля валидных пикселей составляет менее 10% (рис. 3), а наличие большого числа невалидных пикселей объясняется тем, что при съемке значительная часть ИК излучения от соответствующих участков ВП отражается не в направлении PMD-камеры.

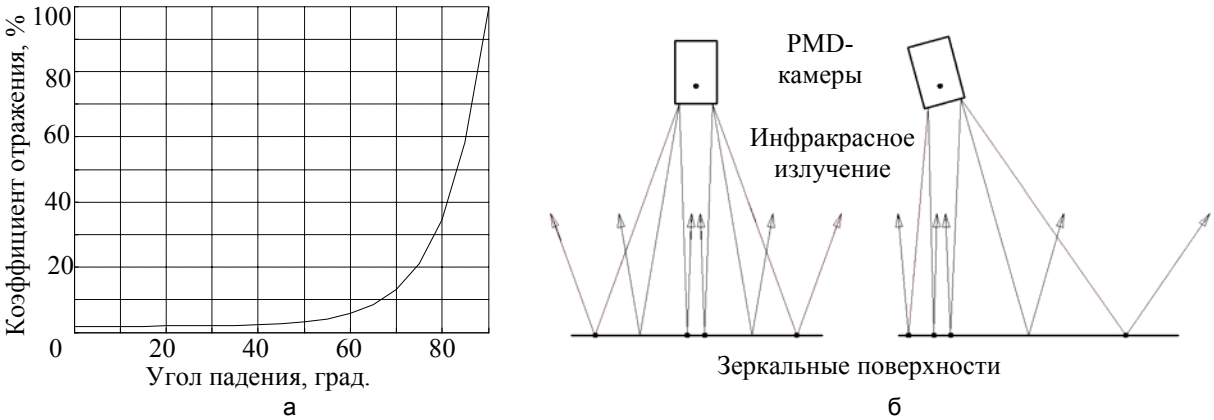


Рис. 4. Зависимость коэффициента отражения неполяризованного света от угла падения (а); ход лучей при зеркальном отражении света и различных положениях времяпролетной камеры (б)

Обработка 3D-изображений

Авторами проведена статистическая обработка полученных последовательностей 3D-изображений. Предварительно с 3D-изображений было удалено изображение конструкции моста. Обрабатывались только валидные пиксели. На рис. 5 представлены графические результаты. В левой части приведены результаты обработки массива из 20 изображений, полученных при положении камеры, близком к горизонтальному. Графики в правом столбце являются результатом обработки массива из 14 изображений, полученных при наклонном положении камеры (угол поворота вокруг продольной оси составлял 8–12º). Даже без применения специальных алгоритмов обработки видно, что значения валидных пикселей близки к измеряемому параметру 4,55 м. Область валидных пикселей стабильно присутствует на всех зарегистрированных 3D-изображениях.

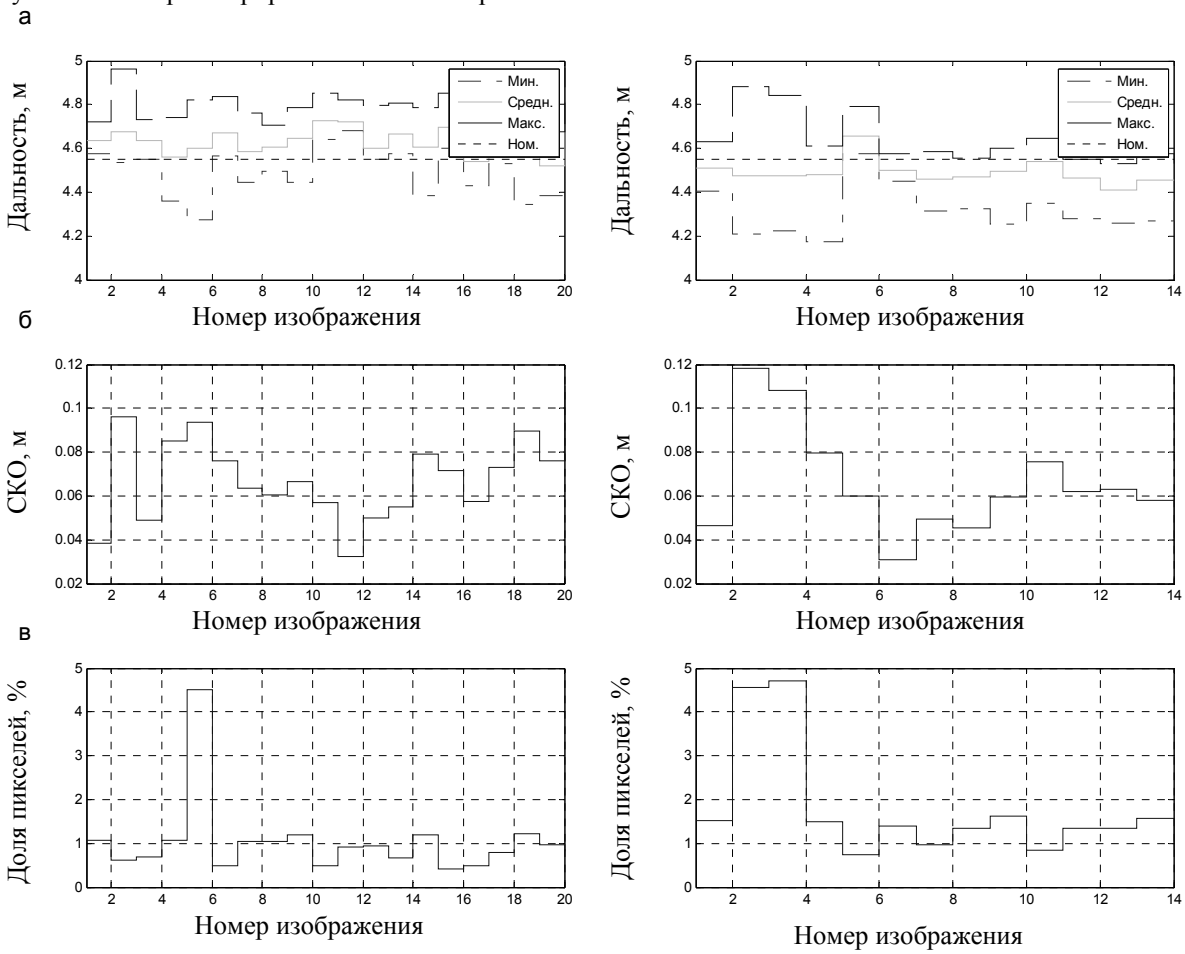


Рис. 5. Результаты обработки двух массивов 3D-изображений ВП: характеристики значений валидных пикселей (а); среднеквадратичные отклонения (СКО) значений валидных пикселей (б); доля валидных пикселей (в)

Безынерционные алгоритмы оценки дальности до ВП

Конечной целью обработки 3D-изображений ВП является получение оценки дальности до ВП. Под дальностью здесь понимается длина перпендикуляра, проведенного от камеры к плоскости, аппроксимирующей участок ВП, попадающий в поле зрения камеры. Ниже авторами предлагаются семь безынерционных алгоритмов, которые могут быть использованы для этой задачи. Алгоритмы протестированы на двух рассмотренных ранее массивах экспериментальных 3D-изображений (рис. 5), а результаты тестирования приведены в табл. 2.

Алгоритм 1. Заключается в непосредственном вычислении среднего значения всех валидных пикселей изображения. При съемке поверхности, находящейся на расстоянии 4,5 м от PMD-камеры, центральному пикселю 3D-изображения соответствует площадь поверхности около 59×59 мм [7]. Выше было показано, что в направлении камеры сигнал достаточной мощности отражается только от ограниченной области ВП. На рис. 3 соответствующая область валидных пикселей составляет порядка 8×8 пикселей. Следовательно, максимальное удаление пикселей от центра этой области составляет около 4 пикселей, что в реальности соответствует 0,25 м. Было сделано предположение, что эта область расположена симметрично относительно перпендикуляра, проведенного к плоскости, аппроксимирующей участок ВП. Полученные катеты (4,5 м и 0,25 м) прямоугольного треугольника дают гипотенузу порядка 4,51 м. Это значение отличается от длины перпендикуляра (первого катета) на доли процента. Исходя из этого, усреднение значений всех валидных пикселей может быть использовано для получения оценки дальности. Нужно отметить, что данный алгоритм приводит к завышенной оценке, так как в сумме используются результаты измерения расстояний, превышающих длину перпендикуляра.

Алгоритм 2. Заключается в непосредственном вычислении медианного значения всех валидных пикселей изображения. Использование этого подхода обычно приводит к уменьшению влияния импульсных помех [12, 13].

Алгоритм 3. Сложнее предыдущих и имеет следующие основные этапы: получение черно-белого изображения (1 – валидные пиксели; 0 – невалидные пиксели); поиск наибольшей по площади плоской фигуры, образованной валидными пикселями; определение ее центра тяжести; построение окружности вокруг центра тяжести; вычисление среднего значения валидных пикселей исходного 3D-изображения, лежащих внутри окружности.

Алгоритм 4. Отличается от алгоритма 3 последним этапом обработки валидных пикселей внутри окружности. Вместо среднего вычисляется медианное значение.

Алгоритм 5. Является развитием алгоритма 3. Содержит дополнительный этап, выполняемый перед процедурой поиска центра тяжести. На этом этапе реализуется морфологическая обработка черно-белого изображения [12]. Целью является компенсация смещения центра масс, возможного из-за насыщения пикселей.

Алгоритм 6. Является модернизацией алгоритма 4. По аналогии с алгоритмом 5 он содержит дополнительный этап морфологической обработки черно-белого изображения.

Алгоритм 7. Отличается от алгоритмов 3 и 4 тем, что вместо обработки совокупности пикселей вблизи центра тяжести в качестве оценки выбирается значение пикселя, лежащего в центре тяжести. Использование данного алгоритма затруднено из-за возможности попадания центра тяжести в область невалидных пикселей. По этой причине результаты работы алгоритма 7 в таблице не приводятся.

В табл. 2 для алгоритмов 3–6 представлены результаты для трех значений радиусов R_0 окружности, построенной вокруг центра тяжести (значения R_0 – в пикселях). Из таблицы видно, что результаты этих алгоритмов приближаются к результатам алгоритмов 1, 2 при увеличении радиуса окружности. Нужно отметить, что наиболее требовательные к вычислительным ресурсам алгоритмы 5, 6, дали наименьшие средние значения оценки дальности.

Номер алгоритма	Обработка массива 3D-изображений, полученного при горизонтальном положении камеры			Обработка массива 3D-изображений, полученного при наклонном положении камеры		
	R_0=1,0	R_0=1,5	R_0=2,0	R_0=1,0	R_0=1,5	R_0=2,0
1	4,635; 0,058			4,489; 0,057		
2	4,629; 0,056			4,493; 0,056		
3	4,615; 0,067	4,629; 0,062	4,631; 0,065	4,475; 0,049	4,482; 0,054	4,486; 0,055
4	4,609; 0,068	4,623; 0,062	4,625; 0,066	4,478; 0,053	4,485; 0,058	4,489; 0,058
5	4,615; 0,069	4,628; 0,064	4,629; 0,064	4,470; 0,053	4,480; 0,054	4,479; 0,060
6	4,609; 0,069	4,622; 0,063	4,622; 0,064	4,473; 0,056	4,484; 0,057	4,483; 0,060

Таблица 2. Результаты работы алгоритмов: Первое число – среднее значение оценки, м; второе число – СКО оценки, м

Заключение

Полученные результаты подтверждают возможность применения времяпролетных PMD-камер для измерения малой дальности до водной поверхности. Опыты, проведенные в естественных условиях, демонстрируют работоспособность метода при волнении водной поверхности. Оптические свойства воды объясняют специфический вид 3D-изображений ее поверхности. Рассмотренные безынерционные алгоритмы обработки могут быть использованы для вычисления оценки дальности до водной поверхности. Полученные таким способом оценки наклонной дальности и геометрической высоты движения над взволнованным морем полезны для реализации принципа комбинированного управления движением по ошибке и волновому возмущению [14].

Дальнейшие исследования будут направлены на изучение особенностей функционирования PMD-камер в более широком диапазоне условий внешней среды; разработку адаптивного программно-алгоритмического обеспечения; исследование параметров точности; испытания PMD-камер в динамике на водных средствах транспорта.

Литература

1. Небылов А.В. Измерение параметров полета вблизи морской поверхности. – СПб: СПбГААП, 1994. – 308 с.
2. Крысин Д.Ю. Метод определения пространственного положения объекта вблизи водной поверхности // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – № 6 (82). – С. 17–24.
3. Piatti D., Rinaudo F. SR-4000 and CamCube3.0 Time of light (ToF) Cameras: Tests and Comparison // Remote Sensing. – 2012. – Is. 4. – P. 1069–1089.
4. Kohoutek T.K. et al. Geo-referenced mapping using an airborne 3D time-of-flight camera // ISPRS Workshop Laser Scanning Archives. – 2011. – Volume XXXVIII-5/W12. – P. 151–155.
5. Мамаев И.Р. Управление и методы обработки информации при автоматизации извлечения объектов из контейнера с использованием PMD-технологии: Автореф. дисс. канд. техн. наук. – ГОУ ВПО Уфимский государственный авиационный технический университет, 2010. – 17 с.
6. Nitsche M. et al. Measuring streambed morphology using range imaging / A. Dittrich, K. Koll, J. Aberle, P Geisenhainer. (eds). – River Flow, 2010. – P. 1715–1722.
7. Электронная документация и программное обеспечение на Интернет-сайте компании IFM [Электронный ресурс]. – URL: http://www.ifm.com/ifmuk/web/dualis-download.htm, свободный (дата обращения 20.01.13).
8. Faltinsen O.M. Sea Loads on Ships and Offshore Structures. – Cambridge University Press, 1990. – 340 p.
9. Martin S. An introduction to ocean remote sensing. – Cambridge University Press, 2004. – 476 p.
10. Шрёдер Г., Трайбер Х. Техническая оптика. – М.: Техносфера, 2006. – 423 с.
11. Крысин Д.Ю., Рао В.П. Особенности применения систем технического зрения для измерения пространственного положения объекта вблизи водной поверхности // Научная сессия ГУАП: Сборник докладов: в 3-х частях. Ч. I. Технические науки. – СПб: ГУАП, 2012. – С. 23–25.
12. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2005. – 1072 с.
13. Зажигин Н.Н., Римских М.В., Шемплинер В.В., Юзликеев Я.В. Новые алгоритмы повышения точности восстановления искаженных изображений // Научно-технический вестник СПбГУ ИТМО. – 2006. – № 5 (28). – С. 49–54.
14. Небылов А.В. Гарантирование точности управления. – М.: Наука, 1998. – 304 с.

Крысин Дмитрий Юрьевич – Санкт-Петербургский государственный университет аэрокосмического приборостроения, ассистент, wtxt@ya.ru

Небылов Александр Владимирович – Санкт-Петербургский государственный университет аэрокосмического приборостроения, доктор технических наук, профессор, зав. кафедрой, nebylov@aanet.ru

2 ФОТОНИКА И ОПТОИНФОРМАТИКА

УДК 681.787:[517.977+517.443]

РЕКУРРЕНТНАЯ ОБРАБОТКА ДАННЫХ В СПЕКТРАЛЬНОЙ ОПТИЧЕСКОЙ КОГЕРЕНТНОЙ ТОМОГРАФИИ НА ОСНОВЕ ФИЛЬТРАЦИИ КАЛМАНА

М.А. Волынский, И.П. Гуров

Предложен метод динамической обработки данных в спектральной оптической когерентной томографии на основе фильтрации Калмана. Показана возможность адаптивного выбора необходимого количества спектральных отсчетов для обеспечения требуемого разрешения для конкретного исследуемого образца. Приведены примеры обработки экспериментальных данных, полученных при исследовании биологических объектов.

Ключевые слова: оптическая когерентная томография, спектральная интерферометрия, преобразование Фурье, фильтр Калмана.

Введение

Бесконтактный контроль внутренней микроструктуры объектов необходим для многих областей науки и современных технологий. Одно из перспективных направлений представляет оптическая когерентная томография (ОКТ), широко используемая в биомедицине [1–3]. Методы ОКТ основаны на принципах интерферометрии малой когерентности применительно к рассеивающим неоднородным средам и обеспечивают разрешение до долей микрометра при восстановлении томографических изображений, представляющих трехмерную внутреннюю микроструктуру неоднородных сред на глубине проникновения оптического излучения.

Как известно, методы ОКТ основываются на использовании корреляционного или спектрального интерферометра [1–5]. В корреляционном интерферометре осуществляют перемещение оптической системы относительно исследуемого объекта. При этом интерференционные полосы малой когерентности формируются в пределах длины когерентности излучения при интерференции части измерительной волны, отраженной от поверхности непрозрачного объекта или от слоя частично прозрачного неоднородного объекта, находящихся от светоделителя на расстоянии, равном оптической длине пути опорной волны.

В спектральном интерферометре оптическая длина пути опорной волны не равна оптической длине пути измерительной волны для всего диапазона высот рельефа непрозрачного объекта или исследуемой области частично прозрачного объекта. На выходе интерферометра размещен спектральный прибор, позволяющий определить составляющую отраженной измерительной волны для каждой из длин волн. При этом, используя преобразование Фурье-спектра, зарегистрированного приемником излучения, можно определить расстояние и степень отражения от каждого слоя [5]. Основное преимущество методов спектральной ОКТ состоит в отсутствии необходимости перемещения оптико-механических элементов, что обеспечивает более высокое быстродействие по сравнению с корреляционной ОКТ.

Вместо использования спектрального прибора, размещенного на выходе интерферометра, можно освещать интерферометр монохроматическим излучением с перестраиваемой последовательно во времени длиной волны. Преимущество метода спектральной ОКТ с перестраиваемой длиной волны состоит в возможности получения информации о микроструктуре объекта на всем наблюдаемом участке, а не только вдоль линии, соответствующей щели спектрального прибора.

В ОКТ с перестраиваемой длиной волны особенно важно обеспечить динамическую обработку получаемых данных, задавая шаг по длине волны и количество длин волн в зависимости от требований к разрешающей способности и быстродействию для конкретного исследуемого объекта.

В любом случае в спектральной ОКТ требуется выполнить преобразование Фурье для получения информации о внутренней микроструктуре объекта. При использовании преобразования Фурье возникают две проблемы. Во-первых, разрешающая способность по глубине среды зависит от количества спектральных отсчетов, однако возможность получения большого количества отсчетов (длин волн) в реальных системах ограничена шириной спектра излучения источника и разрешением по длинам волн. Ввиду сказанного, актуально минимизировать количество спектральных отсчетов при сохранении высокого аксиального разрешения. Во-вторых, поскольку преобразование Фурье является интегральным, для его реализации необходимо наличие полного набора отсчетов перед обработкой, что делает невозможной динамическую обработку данных непосредственно во время измерений при изменении длины волны и, как следствие, снижает быстродействие системы.

В настоящей работе предлагается использование алгоритма рекуррентной обработки данных в системе спектральной ОКТ с перестраиваемой длиной волны на основе фильтрации Калмана [6] для получения динамических оценок внутренней микроструктуры исследуемой среды на основе информации, содержащейся в сигнале спектральной интерференции.

Формирование сигналов спектральной интерференции

В спектральной интерферометрии регистрируют сигналы, пропорциональные значениям интенсивности (см., например, [5])

$$I(k) = G(k)\left| a_R \exp(-j2kr) + \int_0^L a(z)\exp\{-j2k[r + n_s(z)z]\}dz \right|^2, \quad (1)$$

где $G(k)$ – спектр источника излучения; a_R – амплитуда опорной волны; $j = \sqrt{-1}$; $k = 2\pi/\lambda$ – волновое число, определяемое длиной волны λ; $2r$ – оптическая длина пути опорной волны; $a(z)$ – амплитуда предметной волны, отраженной на глубине z в исследуемом образце в диапазоне глубин L; $n_s(z)$ – изменение показателя преломления по глубине среды. Полезная информация содержится в значениях $a(z)$, характеризующих степень отражения предметной волны по глубине исследуемой среды, и которые необходимо определить в результате обработки полученных значений $I(k)$.

В результате нормировки выражения (1) относительно спектра источника излучения и в предположении отсутствия дисперсии в среде (т.е. $n_s(z) = \text{const}$) интерферометрический сигнал в спектральной ОКТ определяется в области волновых чисел формулой

$$S(k) = \left| 1 + \int_0^\infty a(z)\exp(-j2kn_s z)dz \right|^2. \quad (2)$$

Обозначим как $A(k)$ интеграл под модулем в выражении (2). В результате разложения $A(k)$ на вещественную и мнимую части, $A(k) = x + jy$, получим:

$$S(k) = |1 + A(k)|^2 = |1 + x + jy|^2 = 1 + (x^2 + y^2) + (x + jy) + (x - jy) = 1 + |A(k)|^2 + A(k) + A^*(k). \quad (3)$$

Стандартный метод вычисления искомой амплитуды предметной волны $a(z)$ состоит в применении обратного преобразования Фурье к выделенной в выражении (3) информативной составляющей $A(k)$ сигнала $S(k)$, $k \geq 0$:

$$a(z) = \int_0^\infty A(k)\exp(j2kn_s z)dk. \quad (4)$$

Преобразование (4) является непараметрическим, поскольку в общем случае позволяет получать результаты независимо от вида (параметров) преобразуемой величины $A(k)$. Однако $A(k)$ можно рассматривать как набор отдельных гармонических составляющих, определяемых параметрами – частотой, амплитудой и начальной фазой. При этом задача сводится к параметрической идентификации гармонических составляющих и может быть решена с помощью процедуры линейной фильтрации Калмана [6]. Алгоритм дискретной фильтрации Калмана основан на рекуррентной процедуре предсказания значения сигнала на последующий шаг на основе информации, имеющейся на предыдущем шаге, с учетом известной параметрической модели сигнала и использовании ошибки предсказания для уточнения значений параметров на каждом шаге обработки.

Алгоритм обработки данных

Пусть имеется последовательность значений $A(n)$ и решается задача рекуррентной идентификации некоторой гармонической составляющей с фиксированной частотой f, содержащейся в этой последовательности. В рассматриваемом случае n обозначает номер в последовательности значений волнового числа k, задаваемых в спектральной ОКТ.

Искомая гармоническая составляющая определяется моделью $H(n) = u\cos(2\pi fn\Delta k)$, где u – амплитуда; Δk – шаг дискретизации в области волновых чисел, соответствующий шагу изменения длины волны. Рекуррентный алгоритм идентификации гармонической составляющей сводится к предсказанию ее амплитуды u для шага n:

$$\bar{u}(n) = \alpha u(n-1), \quad (5)$$

где коэффициент α задает линейный закон изменения амплитуды между соседними отсчетами (при постоянной амплитуде следует считать, что $\alpha = 1$), и ее апостериорной коррекции на основе невязки предсказания и наблюдения:

$$\tilde{u}(n) = \bar{u}(n) + P(n)[A(n) - C(n)\bar{u}(n)]. \quad (6)$$

В формуле (6) коэффициент перехода $C(n)$ является производной модели по искомому параметру, $C(n) = H'_u(n)$, коэффициент усиления $P(n) = R_{pr}C(n)[C^2(n)R_{pr} + R_n]^{-1}$, где R_{pr} – дисперсия ошибки априорной оценки оцениваемого параметра; R_n – дисперсия шума наблюдения [6].

Таким образом, алгоритм идентификации гармонической составляющей с заданной частотой сводится к вычислению коэффициента перехода $C(n)$ и коэффициента усиления $P(n)$. Если шаг дискретизации Δk принять равным единице, то алгоритм фильтрации (5)–(6) определяется следующим рекуррентным выражением:

$$\tilde{u}(n) = \overline{u}(n-1) + \frac{R_{pr}C(n)[A(n) - C(n)\overline{u}(n-1)]}{C^2(n)R_{pr} + R_n}, \tag{7}$$

где $C(n) = \cos(2\pi fn)$.

Значения R_{pr} и R_n выбираются исходя из априорной информации о сигнале и шуме. Поскольку отклонение начальной оценки искомого параметра от истинного значения амплитуды влияет лишь на скорость сходимости фильтра и не влияет на конечный результат, начальное значение оценки амплитуды может быть установлено в известной степени произвольно. Исследования показали, что наличие ненулевой начальной фазы также не влияет на результат оценивания амплитуды, а влияет только на скорость сходимости. Ввиду этого можно пренебречь дополнительной оценкой начальной фазы.

Примеры обработки сигналов в спектральной интерферометрии

Для проверки работы алгоритма (7) был использован модельный сигнал спектральной интерференции (рис. 1), соответствующий 500 значениям длины волны. Введем определение частоты как величины, обратной количеству отсчетов на одном периоде гармонической составляющей. Отметим, что при этом для данной последовательности минимальное значение частоты составляет 0,002, что соответствует одному периоду на всей последовательности. Сигнал на рис. 1 содержит две гармонические составляющие и описывается моделью

$$s(n) = \cos(2\pi f_1 n) + 0,5\cos(2\pi f_2 n) + w(n), \tag{8}$$

где $f_1 = 0,01$; $f_2 = 0,1$; $w(n)$ – белый гауссовский шум с дисперсией 0,01.

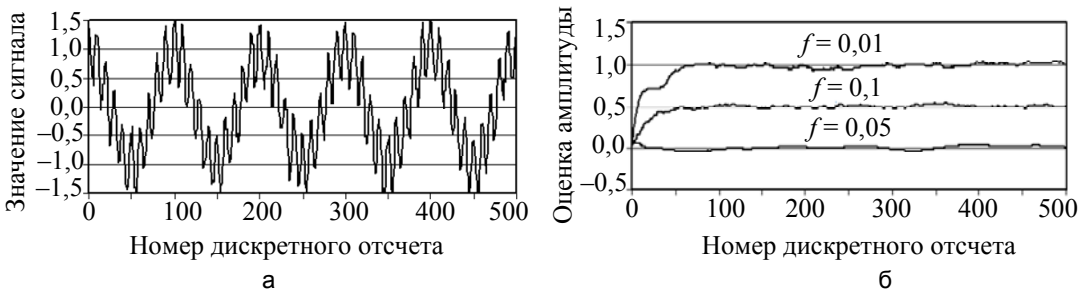


Рис. 1. Сигнал спектральной интерференции (а) и оценки амплитуд гармонических составляющих сигнала с частотами 0,01, 0,1 и 0,05 (б)

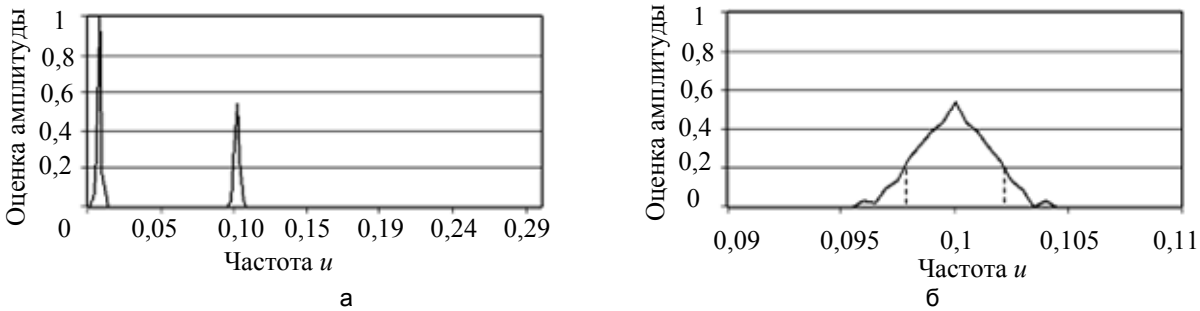


Рис. 2. Результат идентификации гармонических составляющих с частотами из интервала [0,002, 0,3] (а) и его увеличенный фрагмент в окрестности частоты 0,1 (б)

При идентификации гармонических составляющих сигнала (8) была введена дополнительная (гипотетическая) составляющая с частотой 0,05, отсутствующая в этом сигнале, для определения помехоустойчивости алгоритма идентификации в условиях воздействия шума. Видно, что после обработки 50 отсчетов (что соответствует половине периода гармонической составляющей с частотой f_1) оценки выходят на стационарные значения, равные 1; 0,5; 0 для частот 0,01; 0,1; 0,05 соответственно. Незначительные флуктуации оценок обусловлены наличием шума наблюдения с ненулевой дисперсией, что имеет место в реальных системах ОКТ. Ввиду указанного выше отсутствия влияния начальной фазы на результат идентификации гармонической составляющей анализируемый сигнал может иметь нецелое количество периодов на длине реализации. Как известно, если при преобразовании Фурье используется нецелое число периодов, то в таких случаях появляются артефакты (краевой эффект). При моделирова-

нии нами использовано целое число периодов для корректности сравнения характеристик предлагаемого метода с преобразованием Фурье.

На рис. 2 представлен спектр сигнала, показанного на рис. 1. Идентификация гармонических составляющих проводилась для частот в интервале [0,002, 0,3]. Увеличенный фрагмент спектра в окрестности частоты $f_2 = 0{,}1$ демонстрирует разрешающую способность метода.

Из рис. 2 видно, что ширина спектра выделенной гармонической составляющей с частотой $f_2 = 0{,}1$ на уровне 0,2 примерно равна удвоенному минимальному значению частоты (0,002), что соответствует разрешающей способности преобразования Фурье.

Выбор количества отсчетов

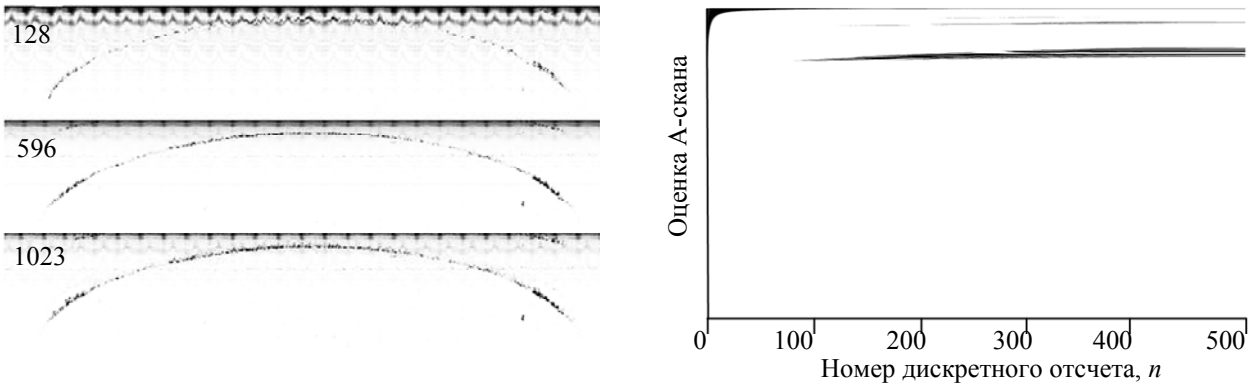
При практическом использовании систем спектральной ОКТ важно минимизировать количество отсчетов (длин волн) без снижения разрешения. При использовании дискретного преобразования Фурье при обработке сигналов в ОКТ требуемое количество отсчетов прямо пропорционально разрешающей способности по глубине. При исследовании произвольных объектов априорная информация о «необходимом» разрешении по глубине отсутствует. Быстрая сходимость метода рекуррентной обработки данных на основе линейного фильтра Калмана позволяет вводить критерий останова, т.е. ограничения количества отсчетов при достижении разрешения, достаточного для конкретного образца.

Можно ввести критерий останова на основе разности оценок амплитуды сигнала с заданной частотой на текущем n и предыдущем $(n-1)$ отсчетах, сравнивая эту разность с некоторым пороговым значением ε. Тогда критерий останова выражается выполнением условия

$$\tilde{u}(n) - \tilde{u}(n-1) < \varepsilon . \qquad (9)$$

Выбор порогового значения ε может осуществляться различными способами с учетом особенностей представления исходных данных. Если значения принадлежат множеству действительных чисел, то в качестве ε целесообразно выбирать среднеквадратическое отклонение (СКО) шума наблюдения. В системах ОКТ исходные данные, как правило, представляют 256 или 4096 градациями для 8- и 12-битного представления соответственно. В этом случае в качестве ε целесообразно выбирать минимальный шаг квантования по уровню.

На рис. 3 показан пример восстановления двумерного томографического изображения (В-скана) силиконовой модели кровеносного сосуда на 128, 596, 1023 шагах обработки и эволюция оценки одного из столбцов В-скана (А-скана) в процессе обработки. Критерий останова (9) выполняется на 312 отсчете, в то время как достижение аналогичного результата при использовании алгоритма обработки на основе быстрого преобразования Фурье требует 1024 отсчета. Таким образом, в данном примере требуемое количество отсчетов (длин волн) снижено примерно в три раза без снижения качества представления томограммы.



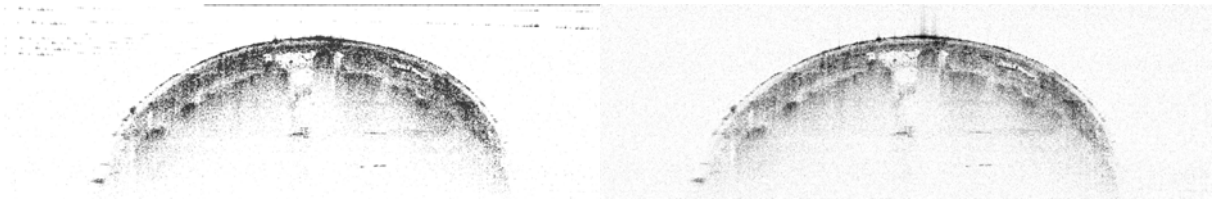
а б

Рис. 3. Оценка В-скана тест-объекта на 128, 596 и 1023 шагах обработки (а) и эволюция оценки А-скана при поступлении новых отсчетов сигнала (б). По вертикальной оси отложены инвертированные значения А-скана в зависимости от номера дискретного отсчета сигнала. Ширина В-сканов – 2 мм, глубина (размер А-скана) – 1,5 мм

Обработка биомедицинских данных

На рис. 4 показаны примеры В-скана личинки серой мясной мухи *Sarcophagidae*, полученные с помощью описанного в настоящей работе метода при использовании 98 отсчетов (длин волн) и с помощью метода на основе преобразования Фурье при 1024 отсчетах. Отличие соответствующих элементов изображений не превышает 12 градаций яркости из 255, т.е. погрешность составляет не более 5%. Отме-

тим, что результат, полученный с помощью предлагаемого в работе метода и не отличающийся от результата на основе преобразования Фурье, достигается к 356 отсчету.



а б

Рис. 4. В-сканы личинки серой мясной мухи *Sarcophagidae*, полученные с помощью описанного в настоящей работе метода (а) и с помощью метода на основе преобразования Фурье (б). Ширина В-сканов – 2 мм, глубина – 1 мм

Заключение

В системах спектральной оптической когерентной томографии с перестраиваемой длиной волны актуальна динамическая обработка данных, позволяющая сократить количество спектральных отсчетов (длин волн) при сохранении высокого разрешения. Традиционные методы обработки сигналов спектральной интерференции, основанные на преобразовании Фурье, не удовлетворяют этому требованию. Для повышения быстродействия систем спектральной оптической когерентной томографии в работе предложен алгоритм обработки данных на основе линейного фильтра Калмана, позволяющий проводить идентификацию гармонических составляющих с заданным набором частот и их вклад (амплитуду гармоник). Введенный критерий останова алгоритма позволяет использовать для восстановления полезной информации о внутренней микроструктуре исследуемого образца минимально необходимое количество спектральных отсчетов (длин волн) при обеспечении требуемого разрешения по глубине.

Следует отметить, что предложенный алгоритм нечувствителен к значениям начальной фазы в сигнале и не требует целого количества периодов сигнала, что является преимуществом по сравнению с преобразованием Фурье.

Описанный в работе алгоритм реализован в форме программного модуля [7].

Авторы выражают благодарность М.В. Волкову за полезное обсуждение особенностей программной реализации описанного в работе метода.

Работа выполнена при финансовой поддержке Министерства образования и науки Российской Федерации.

Литература

1. Tomlins P.H., Wang R.K. Theory, developments and applications of optical coherence tomography // J. Phys. D: Appl. Phys. – 2005. – V. 38. – P. 2519–2535.
2. Optical Coherence Tomography. Technology and Applications / W. Drexler, J.G. Fujimoto, eds. – Berlin-Heidelberg: Springer-Verlag, 2008. – 1346 p.
3. Гуров И.П., Волынский М.А., Жукова Е.В., Маргарянц Н.Б. Исследование растительных тканей методом оптической когерентной микроскопии // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – № 5 (81). – С. 42–47.
4. Волынский М.А., Воробьева Е.А., Гуров И.П., Маргарянц Н.Б. Бесконтактный контроль микрообъектов методами интерферометрии малой когерентности и оптической когерентной томографии // Изв. вузов. Приборостроение. – 2011. – Т. 54. – № 2. – С. 75–82.
5. Häusler G., Lindner M.V. «Coherence radar» and «Spectral radar» – new tools for dermatological diagnostics // J. Biomed. Opt. – 1998. – V. 3. – P. 21–31.
6. Kalman R.E. A new approach to linear filtering and prediction problems // Trans. ASME, J. Basic Eng. – 1960. – V. 82. – P. 35–45.
7. Волынский М.А. Программный модуль «Рекуррентная обработка данных в спектральной оптической когерентной томографии» // Свидетельство № 2012617881 от 31.08.2012.

Волынский Максим Александрович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат технических наук, доцент, maxim.volynsky@gmail.com

Гуров Игорь Петрович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, профессор, зав. кафедрой, gurov@mail.ifmo.ru

УДК 535.41: 681.7.014.3

АНАЛИЗ ПАРАМЕТРОВ ВОЛНОВОГО ПАКЕТА НА ОСНОВЕ ДИНАМИЧЕСКОГО ОЦЕНИВАНИЯ ФАЗЫ СПЕКТРАЛЬНЫХ СОСТАВЛЯЮЩИХ

Е.А. Воробьева, И.П. Гуров

Рассмотрен метод динамического анализа параметров волнового пакета, состоящий в выделении и оценивании амплитуд и фаз отдельных монохроматических составляющих волнового пакета. Показано, что определение положения максимума интенсивности волнового пакета можно осуществить по критерию равенства нулю начальных фаз монохроматических составляющих, оцениваемых при помощи дискретного нелинейного фильтра Калмана.

Ключевые слова: волновой пакет, спектральные составляющие, начальная фаза, фильтр Калмана.

Введение

При исследовании различных объектов оптическими методами используют разнообразные источники излучения. В ряде случаев излучение представляет собой совокупность монохроматических электромагнитных колебаний с различными частотами, образующих при распространении в пространстве волновой пакет. В зависимости от соотношения амплитуд и фаз составляющих волнового пакета возможно формирование распределений интенсивности различного вида. В частности, в области совпадения или малого отличия фаз формируется распределение интенсивности в форме импульса, положение максимума которого соответствует нулевому значению фазы всех монохроматических составляющих волнового пакета.

Свойства волнового пакета широко используются в оптической когерентной томографии (ОКТ) [1], поскольку идентификация положения максимума распределения интенсивности позволяет определить условие равенства оптической длины пути опорной и измерительной волн, тем самым выделить часть излучения, отраженного от отдельного слоя исследуемого частично прозрачного объекта, и в результате получить послойную томограмму внутренней микроструктуры объекта.

Определить положение максимума интенсивности можно различными методами, например, в результате детектирования огибающей распределения интенсивности [2] или нелинейной фильтрации распределения интенсивности [3]. Однако возможен и иной подход, предлагаемый в настоящей работе и состоящий в выделении и оценивании амплитуд и фаз отдельных монохроматических составляющих волнового пакета. При этом оказывается возможным определение положения максимума интенсивности по критерию равенства нулю начальных фаз монохроматических составляющих.

В работе рассматривается возможность фильтрации набора монохроматических составляющих в динамическом режиме с автоматическим определением положения максимума интенсивности импульса, представляющего волновой пакет.

Многочастотная модель интерферометрического сигнала для волнового пакета в системах ОКТ

Интерферометрический сигнал, характеризующий волновой пакет, можно представить в виде суммы гармонических составляющих с амплитудами a_i и частотами f_i

$$s(f_i, z) = \mu I(f_i, z) = a_i \cos\left(2\pi f_i z\right), \tag{1}$$

где μ – коэффициент преобразования интенсивности $I(f_i, z)$ в сигнал; z – независимая переменная. В случае огибающей гауссовой формы для интенсивности волнового пакета преобразование Фурье $G(f_i)$ огибающей также имеет гауссову форму:

$$a_i = G(f_i) = \exp\left(-\pi^2 \left(f_0 \pm \delta f_i\right)^2 \sigma^2\right), \tag{2}$$

где f_0 – среднее значение частоты, $f_0 = 1/\langle\lambda\rangle$, $\langle\lambda\rangle$ – среднее значение длины волны излучения; δf_i – отклонения частоты составляющих волнового пакета от среднего значения частоты.

При регистрации составляющие (1) на всех частотах суммируются, и для наблюдения доступен суммарный сигнал вида

$$s(z) = \sum_i a_i \cos(2\pi f_i z).$$

Поскольку в системах ОКТ разность хода опорной и измерительной волн меняется, в дискретном случае $z = k\Delta z$, где Δz – шаг дискретизации, регистрируемая последовательность отсчетов имеет вид

$$s(k) = \sum_i a_i \cos(2\pi f_i k \Delta z).$$

Для определения значения параметра σ в уравнении (2) рассмотрим спектр волнового пакета. Ширина спектра сигнала на уровне $1/e$

$$f_c = 1/\left(\pi\sigma\right). \tag{3}$$

Обозначим количество периодов косинусоиды, укладывающихся в интервале σ, как Q. С учетом (3) можно записать

$$Q = \pi\sigma / \lambda .$$

С другой стороны, интервал 2σ есть ширина спектра на уровне $1/e$ и примерно соответствует длине когерентности излучения $I_c = \langle\lambda\rangle^2 / \Delta\lambda$, где $\Delta\lambda$ – ширина спектра излучения. Нетрудно показать, что $\Delta\lambda / \langle\lambda\rangle = \Delta f / f_0$, откуда $\Delta\lambda = \langle\lambda\rangle \Delta f / f_0 = \langle\lambda\rangle^2 \Delta\lambda$. С учетом (3) имеем $\Delta f = 2 f_c = 2 / \pi\sigma$ и получаем $\Delta\lambda = 2\langle\lambda\rangle^2 / \pi\sigma$, откуда

$$\sigma = \frac{2}{\pi}\frac{\langle\lambda\rangle^2}{\Delta\lambda} = \frac{2\langle\lambda\rangle}{\pi}\frac{\langle\lambda\rangle}{\Delta\lambda}. \tag{4}$$

Поскольку $\langle\lambda\rangle$ в (4) определяется источником излучения, т.е. задается в известной степени произвольно, можно отнормировать (4), приняв первый сомножитель равным единице. Тогда параметр $\sigma^* = \langle\lambda\rangle / \Delta\lambda$ будет соответствовать относительной ширине спектра, т.е. числу, которое показывает, во сколько раз средняя длина волны больше, чем ширина спектра. Для типичных источников излучения, используемых с системах ОКТ, σ^* находится в интервале 10–50.

Восстановление параметров волнового пакета методом дискретной нелинейной фильтрации Калмана

Параметры интерферометрического сигнала с учетом отражения оптического излучения в образце можно описать в виде вектора

$$\mathbf{\Theta} = \left(\Phi_i(k), \varepsilon_i(k), f_i(k)\right)^T \tag{5}$$

где $\Phi_i(k)$ и $\varepsilon_i(k)$ – полная и начальная фазы, соответствующие частоте $f_i(k)$, k=0, 1, …, K – номер дискретного отчета в области независимой переменной; K – общее количество отсчетов; $i = 0, \pm 1, ..., \pm(L-1)/2$, $(2L+1)$ – количество частот в волновом пакете.

Оценку вектора параметров (5) в динамическом режиме можно получить с использованием фильтра Калмана [4]. При этом проводится обработка последовательности отсчетов, определяемой заданной моделью, которая выражается уравнением наблюдения, включающим влияние шума наблюдения. Эволюция вектора параметров характеризуется уравнением системы с матрицей перехода от предыдущего отсчета (состояния системы) к последующему отсчету.

В интерферометрической системе уравнение наблюдения может быть записано как

$$s(k) = \sum_i a_i \cos\left(\Phi_i\left(k, \varepsilon_i, f_i(k)\right)\right) + n(k),$$

где $n(k)$ – последовательность отсчетов шума наблюдения.

Матрица перехода $\mathbf{A}(k)$ для вектора параметров в уравнении системы имеет размерность $(6L+3)\times(6L+3)$ и определяется как

$$\mathbf{A}(k) = \begin{pmatrix} 1 & 0 & 2\pi\Delta z & ... & 0 \\ 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & & 0 \\ ... & & & ... & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Алгоритм дискретной нелинейной фильтрации Калмана, подробно описанный, например, в работе [5], с использованием рассмотренной выше многочастотной модели позволяет восстановить параметры, т.е. амплитуду и фазу составляющих интерферометрического сигнала. Однако при использовании модели с достаточно большим количеством частот L быстродействие алгоритма обработки заметно снижается. Для получения результатов, приведенных в работе, была использована модель сигнала с 30 частотами.

На рис. 1, а, представлен исходный смоделированный сигнал. Огибающая сигнала имеет два равных по амплитуде максимума в точках, которым соответствуют дискретные отсчеты сигнала 2980 и 6770. Несущая частота интерферометрического сигнала $f_0 = 0{,}015$, $\sigma^* = 15$.

Дискретный набор частот определяется как $f_i = f_0 \pm i\Delta f$, где $\Delta f = 3\cdot 10^{-5}$. В данном случае максимальное отклонение по частоте от несущей составляет $\Delta f \cdot (L/2) = 0{,}0005$. Таким образом, набор частот попадает в диапазон $\left(f_c^-, f_c^+\right)$, где граничные частоты равны 0,0145 и 0,0155 соответственно. На

рис. 1, б, представлен пример спектра огибающей сигнала. Пунктирными линиями обозначен диапазон, соответствующий интервалу $\left(f_c^-, f_c^+\right)$. На рис. 1, б, также отмечены первые 9 частот, которые использованы в модели интерферометрического сигнала в алгоритме фильтрации Калмана.

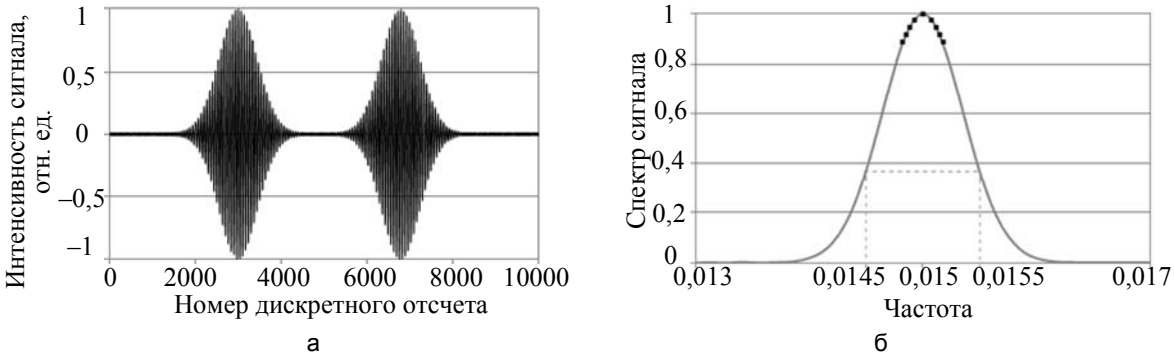


Рис. 1. Смоделированный сигнал (а) и спектр огибающей сигнала (б)

Покажем, что алгоритм дискретной нелинейной фильтрации Калмана позволяет восстановить параметры волнового пакета. На рис. 2, а, представлены восстановленные значения частот для трех сигналов, которым соответствует центральная частота (сплошная линия) и две боковые частоты $f_+ = f_0 + \Delta f$ и $f_- = f_0 - \Delta f$ (пунктирные линии). Как видно на рис. 2, а, значение частот мало корректируется в процессе фильтрации, поскольку было задано корректное начальное значение.

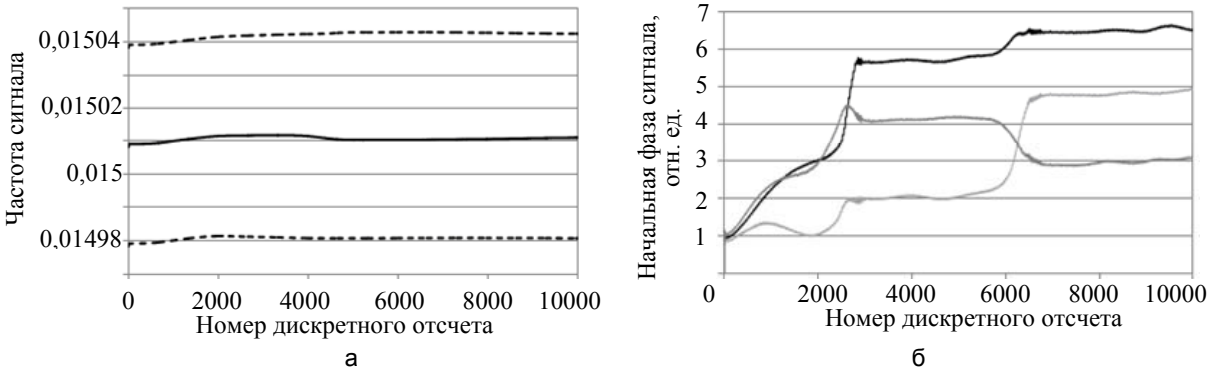


Рис. 2. Восстановленные значения частот (а) и начальных фаз для трех составляющих волнового пакета (б)

Наибольший интерес представляют восстановленные значения начальных фаз составляющих сигнала, поскольку на основе этих значений можно восстановить положения максимума огибающей сигнала, соответствующие нулевым фазам всех составляющих. На рис. 2, б, представлены восстановленные значения начальных фаз для сигналов, которым соответствуют частоты f_0, f_+ и f_-. Жирной линией показана начальная фаза сигнала с центральной частотой f_0, тонкими линиями – начальные фазы сигналов с боковыми частотами f_+ и f_-. Как видно на рис. 2, б, значение начальной фазы остается постоянным на интервале после появления очередного максимума огибающей и до появления следующего максимума. Важно отметить, что на представленном графике значение начальной фазы для сигнала, соответствующего центральной частоте, превышает значение 2π. Это связано с особенностями реализации алгоритма восстановления параметров сигнала. В данном случае начальная фаза сигнала есть остаток от деления на 2π восстановленного значения.

Для вычисления положений максимума огибающей рассмотрим три частоты f_0, f_+ и f_-. Значение сигнала можно представить в виде

$$s(k) = a_0 \cos(\varepsilon_0 + 2\pi f_0 k) + a_1 \cos(\varepsilon_1 + 2\pi f_+ k) + a_2 \cos(\varepsilon_2 + 2\pi f_- k)$$

где ε_0, ε_1 и ε_2 – начальные фазы сигналов с частотами f_0, f_+ и f_- соответственно. Значения a_0, a_1 и a_2 вычисляются по формуле (2).

При формировании волнового пакета положению максимума огибающей соответствует дискретный отсчет сигнала, в котором начальные фазы составляющих сигнала, формирующих волновой пакет, совпадают и равны нулю. Можно записать условия максимума огибающей волнового пакета

$$\varepsilon_0 + 2\pi f_0 k_0 = \varepsilon_1 + 2\pi f_+ k_0 + 2\pi n_1 = \varepsilon_2 + 2\pi f_- k_0 + 2\pi n_2, \qquad (6)$$

где n_1 и n_2 – целые числа; k_0 – положение максимума огибающей волнового пакета. В выражении (6) $2\pi n_1$ и $2\pi n_2$ – это фазы, которые «набегут» за счет разностей между f_0 и f_+, f_0 и f_- соответственно. Поскольку значение Δf задано, то можно записать

$$2\pi n_1 = 2\pi n_2 = 2\pi \Delta f k_0 ,$$

откуда

$$k_0 = \frac{n_1}{\Delta f} = \frac{n_2}{\Delta f} . \qquad (7)$$

На рис. 3, а, представлены восстановленные значения положения максимумов огибающей волнового пакета. Пунктирной линией обозначены истинные значения положения максимумов. Видно, что положение первого максимума алгоритм восстанавливает достаточно точно. Однако положение второго максимума определено не совсем точно, поскольку алгоритм учитывает влияние предыдущего максимума огибающей. Для устранения этого недостатка можно использовать переключаемую модель сигнала [3].

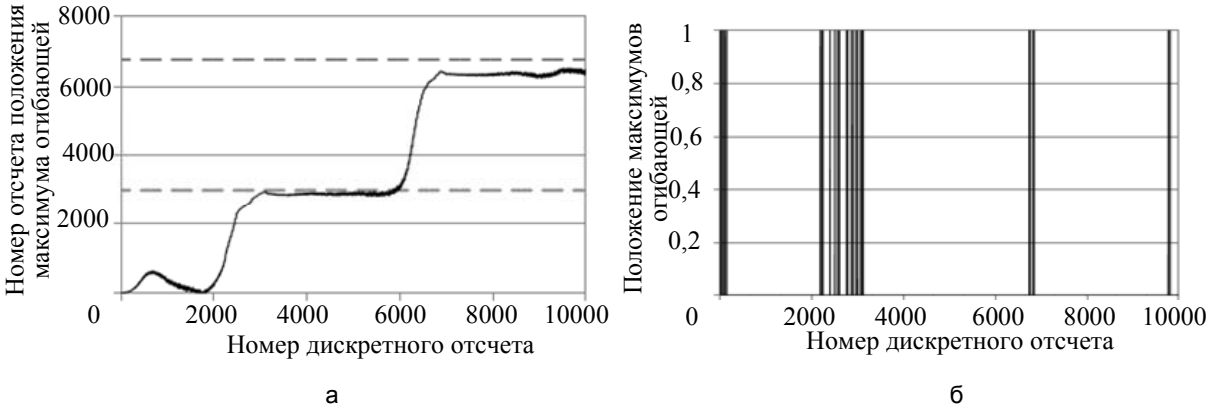


Рис. 3. Восстановленные значения положений максимумов огибающей сигнала, полученные с использованием уравнения (7) (а) и критерия равенства полной фазы сигналов 2π (б)

Другой способ оценки положения максимумов огибающей основан на анализе полных фаз составляющих сигнала, формирующих волновой пакет. Положению максимумов огибающей соответствуют дискретные отсчеты, для которых полная фаза сигналов кратна 2π. На рис. 3, б, представлены оценки максимума огибающей, полученные с использованием описанного критерия. Для трех сигналов, которым соответствуют частоты f_0, f_+ и f_-, были вычислены полные фазы. На графике дискретные отсчеты, для которых полные фазы трех сигналов кратны 2π, отмечены дельта-функциями. Алгоритм корректно оценил положение существующих максимумов огибающей волнового пакета, однако при этом возникают ложные максимумы. Ложный максимум в начале дискретной последовательности обусловлен начальной подстройкой дискретного нелинейного алгоритма фильтрации Калмана. Использование только трех частот для определения положения максимумов является причиной ложных результатов в остальных отсчетах дискретной последовательности. При увеличении количества частот надежность результатов заметно возрастает, однако это происходит за счет повышения вычислительной сложности.

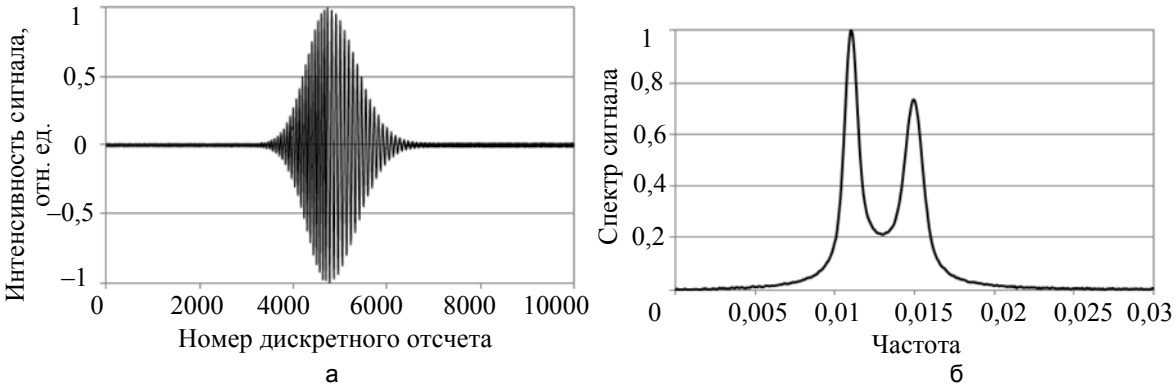


Рис. 4. Двухчастотный сигнал (а) и его спектр (б)

Рассмотрим волновой пакет с изменяющейся несущей частотой (рис. 4). Положению максимума огибающей волнового пакета соответствует дискретный отсчет 4729. Несущая частота волнового пакета до отсчета 4729 равна 0,015, после – 0,011. Рис. 4, б, иллюстрирует спектр сигнала.

Алгоритм дискретной нелинейной фильтрации Калмана с использованием многочастотной модели сигнала позволяет восстановить центральную частоту волнового пакета. На рис. 5 представлена оценка восстановленной центральной частоты волнового пакета. На интервале отсчетов от 0 до 4729 оценка значения частоты постепенно приближается к заданному значению 0,015. После отсчета 4729 алгоритм фиксирует изменение частоты. Восстановленное значение частоты постепенно корректируется до истинного значения 0,011.

Рис. 5. Восстановленные значения центральной частоты

Таким образом, алгоритм дискретной нелинейной фильтрации Калмана с использованием многочастотной модели позволяет восстановить параметры интерферометрических сигналов, формирующих волновой пакет, в том числе при локальном изменении среднего значения частоты сигнала.

Заключение

Проведенное рассмотрение особенностей формирования и анализа интерферометрических сигналов в системах оптической когерентной томографии на основе многочастотной модели сигнала показывает возможность получения динамических оценок параметров волнового пакета, в том числе при изменениях значения средней частоты сигнала волнового пакета. Алгоритм нелинейной фильтрации Калмана с использованием предложенной многочастотной модели интерферометрических сигналов позволил получить оценки начальной фазы частотных составляющих волнового пакета и определить положение максимумов огибающей по критерию равенства нулю начальных фаз составляющих и критерию равенства полной фазы целому числу фазовых циклов. При реализации рассмотренного алгоритма необходимо учитывать возрастание вычислительной сложности при увеличении числа учитываемых в модели сигнала частотных составляющих. Обработка большего числа составляющих обеспечивает повышение точности определения положения максимума огибающей волнового пакета.

Работа выполнена при финансовой поддержке Министерства образования и науки Российской Федерации.

Литература

1. Гуров И.П. Оптическая когерентная томография: принципы проблемы и перспективы // Проблемы когерентной и нелинейной оптики / Под ред. И.П. Гурова и С.А. Козлова. – СПб: СПбГУ ИТМО, 2004. – С. 6–30.
2. Воробьева Е.А., Гуров И.П., Петерсон М.В. Исследование метода обработки сигналов в системах оптической когерентной томографии с повышенным быстродействием // Изв. вузов. Приборостроение. – 2010. – № 3. – С. 65–73.
3. Gurov I., Volynsky M., Zakharov A. Evaluation of multilayer tissue in optical coherence tomography by the extended Kalman filtering method // Proc. SPIE. – 2007. – V. 6734. – P. 67341.
4. Kalman R.E. A new approach to linear filtering and prediction problems // Transactions of the ASME. – Journal of Basic Engineering, Series D. – 1960. – V. 82. – P. 34–45.
5. Гуров И.П., Захаров А.С. Анализ характеристик интерференционных полос методом нелинейной фильтрации Калмана // Оптика и спектроскопия. – 2004. – Т. 96. – № 2. – С. 210–216.

Воробьева Елена Александровна – ООО «Моторола-Мобилити», инженер-программист, lenavorobyeva@gmail.com

Гуров Игорь Петрович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, профессор, зав. кафедрой, gurov@mail.ifmo.ru

УДК 681.784.88, 535.36

ИДЕНТИФИКАЦИЯ АМИНОКИСЛОТ, ВХОДЯЩИХ В СОСТАВ КОЖИ ЧЕЛОВЕКА, С ПОМОЩЬЮ СПЕКТРОВ ДИАПАЗОНА 0,05–1,2 ТГЦ

А.А. Езерская, М.В. Цуркан, О.А. Смолянская, А.В. Капралова, Н.С. Балбекин, И.В. Романов, И.В. Прожеев, Е.А. Стрепитов, И.Ю. Просанов

Методом терагерцовой (ТГц) спектроскопии с разрешением во времени получены спектры кожи человека и таблетированных препаратов аминокислот в диапазоне 0,05–1,2 ТГц. Сделано сопоставление спектральных линий аминокислот для идентификации их в коже. Показано, что наиболее хорошо различимы в спектре кожи следующие аминокислоты: триптофан, цистеин, аланин, глицин, фенилаланин и гистидин. ТГц спектральные линии указанных аминокислот связаны с колебательными модами молекул, которые обусловлены химическими реакциями, процессами гидратации и конформных изменений молекул в коже человека. Определение спектральных линий аминокислот и других характерных молекул, участвующих в метаболических процессах кожи, необходимо для разработки ТГц диагностического прибора с максимальной чувствительностью и селективностью.

Ключевые слова: терагерцовая спектроскопия, диагностика, кожа человека, аминокислоты.

Введение

Исследования в области применения ТГц излучения в медицине направлены, в основном, на диагностику различных заболеваний человека и животных [1, 2]. Наиболее актуальными объектами диагностических исследований являются патологические новообразования кожи человека, а также динамика восстановления кожи после полученной травмы. В этом направлении получены положительные результаты, показывающие однозначную идентификацию пораженных областей кожи с помощью ТГц спектроскопии [3]. В том числе, большое количество экспериментальных работ по данной тематике было проведено в научно-образовательном центре фемтосекундной оптики и фемтотехнологий НИУ ИТМО. Однако до сих пор неизученным оставался вопрос идентификации различных биологических молекул, входящих в состав кожи человека, в ТГц спектрах. На данный момент имеются предположения о том, что отдельные пики ТГц спектров характеризуют поглощение воды, входящей в состав кожи [1, 4]. Также в литературе встречаются работы, в которых приводятся ТГц спектры отдельных биологических молекул [5, 6]. Однако общая картина изменения спектров ТГц излучения в результате взаимодействия с кожными покровами на данный момент отсутствует. Очевидно, что для разработки медицинских приложений необходимо заполнить эту нишу.

В ТГц диапазоне частот лежат линии поглощения, соответствующие вращательным движениям молекул, молекулярных групп сложных молекул. Также излучение оказывает влияние на молекулярные комплексы, структурированные за счет межмолекулярных взаимодействий – ван-дер-ваальсовых сил и водородных связей. Развитие ТГц спектроскопии позволяет получить новые сведения как о конформации молекул, так и о характеристиках межмолекулярных взаимодействий при различных условиях.

Особый интерес представляет оценка возможных конформационных изменений сложных биологических молекул при их взаимодействии или изменении условий и среды функционирования молекул. К настоящему времени хорошо известно, что функциональное состояние биологической молекулы зависит от ее конформационных особенностей [7–9]. В белке акцепторами ТГц излучения являются аминокислоты, поэтому представляет несомненный интерес изучение влияния излучения ТГц диапазона на состояние колебательно-вращательных мод индивидуальных аминокислот, в первую очередь таких, как характеристические белковые хромофоры – триптофан, тирозин и пр. Практически при любых физико-химических воздействиях на белок, включая и электромагнитное излучение, именно эти аминокислоты обнаруживают различия в поглощении излучения при стандартной УФ спектроскопии [7, 8].

В связи с этим представляется актуальным исследование влияния ТГц излучения на колебательно-вращательные частоты этих хромофоров в коже человека в ТГц диапазоне.

Материалы и методы

Терагерцовый рефлектометрический томограф. Исследование спектров отражения образцов в диапазоне 0,05–1,2 ТГц проводилось с помощью терагерцового рефлектометрического томографа. Его схема представлена на рис. 1. В основе работы устройства лежит принцип записи распределения электрической компоненты поля до и после отражения от образцов. Восстановление спектра ТГц излучения реализуется с помощью вычисления интеграла Фурье по записанным данным.

Луч с выхода FL-1 поступает на светоделитель пучка 1, где разделяется на два луча. Пробный луч распространяется по оптической схеме (система зеркал М, электрооптический кристалл 9, ахроматическая четвертьволновая пластина 10, призма Волластона 11) и поступает на вход балансного детектора 12. Луч накачки последовательно проходит оптико-механический модулятор 5, оптическую линию задержки 2 и поступает на вход полупроводникового кристалла InAs 3. Излучение с выхода 3 через систему двух параболических зеркал и делитель ТГц излучения 7 падает на исследуемый образец 6. Отраженное образцом излучение через параболическое зеркало 4 и делитель 7 падает на электрооптический кристалл CdTe 9, фо-

кусируясь линзой 8. Под действием электрической компоненты ТГц излучения изменяется плоскость поляризации электрооптического кристалла. Таким образом, электрическая компонента ТГц излучения с помощью фазосдвигающей пластинки 10 и призмы Волластона 11 детектируется балансным детектором 12 с синхронным усилением. Изменение времени задержки луча накачки приводит к сдвигу времени открытия электрооптического кристалла, что позволяет сканировать распределение электрической компоненты поля по времени. Данные измерений с синхронного усилителя заносятся в персональный компьютер с помощью встроенного аналого-цифрового преобразователя.

В таблице представлены основные характеристики ТГц рефлектометрического томографа.

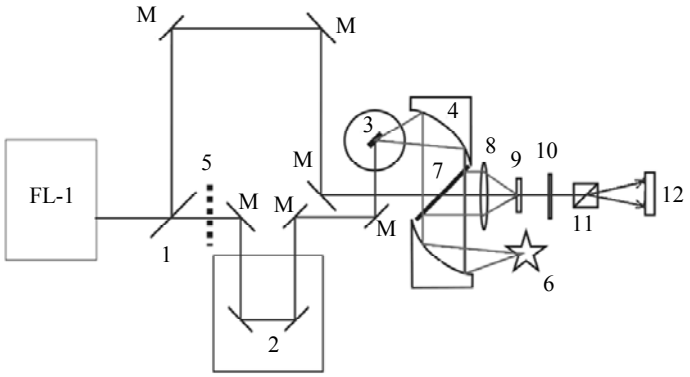


Рис. 1. Схема ТГц рефлектометрического томографа: FL-1 – лазер фемтосекундных импульсов на Yb:KYW; M – зеркала с $R \approx 100\%$ при падении светового пучка под 45°; 1 – светоделитель; 2 – оптическая линия задержки; 3 – полупроводниковый кристалл InAs; 4 – 45° параболические зеркала; 5 – оптико-механический модулятор; 6 – объект исследования; 7 – светоделитель ТГц излучения на основе пластины из высокоомного кремния; 8 – линза из ТРХ (поли-4-метилпентен-1) с f = 5 см; 9 – электрооптический кристалл CdTe; 10 – ахроматическая четвертьволновая пластина; 11 – призма Волластона; 12 – балансный детектор

Характеристика	Значение
Средняя мощность ТГц излучения, мкВт	30 ± 3
Мощность импульса, мВт	120 ± 10
Длительность импульса, пс	3
Энергия импульса, фДж	360 ± 30
Спектральный диапазон, ТГц	0,05–2,0
Отношение сигнал/шум в спектре	100
Частота повторения, МГц	75
Частота модуляции, Гц	433

Таблица. Характеристики ТГц рефлектометрического томографа

Методика измерений спектра пропускания кожи человека и отдельных аминокислот состоит из следующих этапов:

1. запускается ТГц рефлектометрический томограф, согласно порядку включения;
2. на персональном компьютере запускается программа управления томографом;
3. производится измерение амплитуды электрической компоненты поля в отсутствии измеряемого образца;
4. предварительно подготовленные образцы помещаются согласно схеме (рис. 1) в измерительную установку;
5. производится измерение амплитуды электрической компоненты поля, отразившегося от образца;
6. полученные данные заносятся в программу, реализующую алгоритм преобразования Фурье;
7. по полученным спектральным выборкам определяется частота отражения ТГц излучения в образцах.

Подготовка кожи для экспериментального исследования. Для получения спектра нормальной кожи облучался участок кожи руки человека с внешней стороны, представленный на рис. 2.

Подготовка аминокислот для экспериментального исследования. В качестве объектов были использованы таблетированные формы порошковых лиофиллизированных препаратов следующих аминокислот: аланин, глицин, гистидин, триптофан, фениланин, цистеин (фирма Sigma, США). Образцы для эксперимента готовились путем прессования порошков аминокислот в таблетки диаметром 3 мм и толщиной около 1,5 мм. Прессование производилось в вакууме в стальных оправках. Эксперименты проводились при комнатной температуре. Исследуемые образцы кожи и аминокислот устанавливались в перетяжку ТГц луча, диаметр которого составлял 1 мм. Время экспозиции образца под излучением составляло 5 минут. Спектр каждого образца записывался три раза. Результаты измерений по каждому образцу усреднялись.

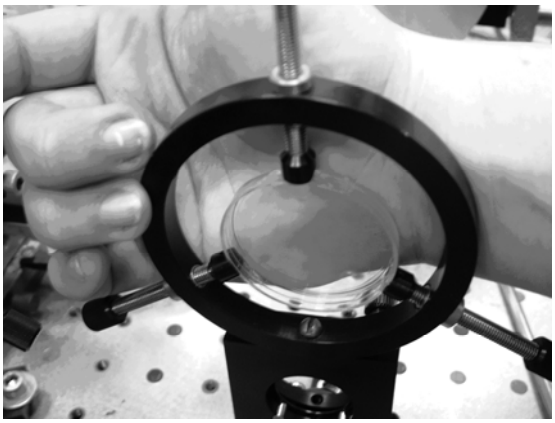


Рис. 2. Фотография руки женщины 21 года с участком облучаемой кожи

Результаты и их обсуждение

На рис. 3 представлен спектр отражения нормальной кожи человека.

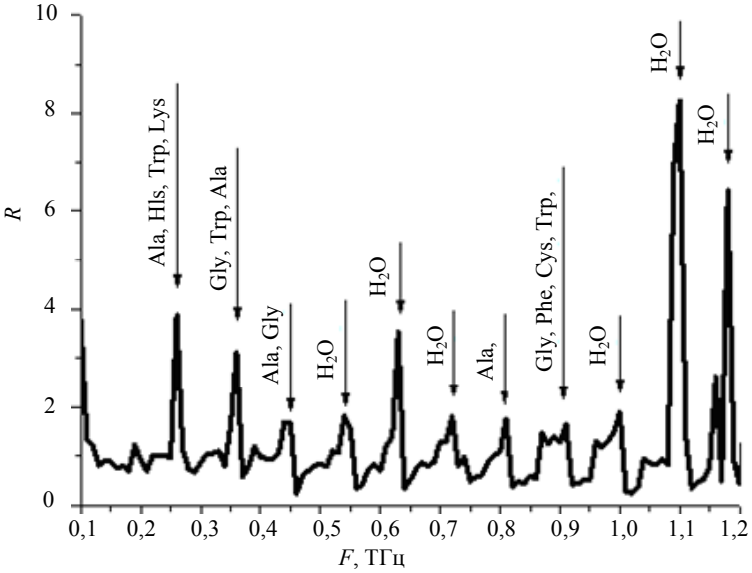


Рис. 3. ТГц спектр отражения участка нормальной кожи человека: *R* – ТГц отраженный сигнал, в условных единицах; *F* – частота. Ala – аланин; Cys – цистеин; His – гистидин; Gly – глицин; Trp – триптофан; Phe – фениланин; H_2O – вода

На спектре удалось выделить спектральные линии исследованных аминокислот. Так, рядом с линией отражения кожи на частоте 0,26 ТГц линия отражения имеется у аланина, гистидина, триптофана и лизина. Аминокислоты расположены в порядке уменьшения интенсивности пика отражения. Линия отражения кожи на частоте 0,36 ТГц совпала с линиями следующих аминокислот: глицин, триптофан и аланин. Отражение на частоте 0,82 ТГц совместилось с отражением аланина. Линии отражения кожи имеются в области частот 0,87; 0,89; 0,91 ТГц. Этим линиям сопоставлены следующие аминокислоты: глицин, фениланин, цистеин, аланин и триптофан. Указанные аминокислоты содержатся в больших количествах в коже и соединительной ткани, поэтому полезны для ее восстановления при повреждении [10]. Глицин замедляет дегенерацию мышечной ткани. Он является источником креатина – вещества, содержащегося в мышечной ткани и используемого при синтезе ДНК и РНК. Цистеин играет важную роль в процессах формирования тканей кожи. Он входит в состав альфа-кератина – основного белка ногтей, кожи и волос, способствует формированию коллагена и улучшает эластичность и текстуру кожи.

Линии отражения на частотах 0,55; 0,63; 0,73; 1 ТГц, а также 1,1 и 1,17 ТГц были отнесены к пикам, связанным с поглощением жидкой воды [11, 12] и водяного пара [13] в терагерцовой области частот. Субмиллиметровое, т.е. ТГц излучение поглощается парами воды и молекулярным кислородом [14]. С большой вероятностью можно считать, что в коже человека возможными фоторецепторами субмиллиметрового диапазона могут быть свободная вода и гидратные оболочки ионов и молекул, находящихся во внеклеточной жидкости, а также растворенный в ней молекулярный кислород. Изменение состояния воды и кислорода может вызвать изменение конформации примембранных белков.

Заключение

Получены ТГц спектры отражения нормальной кожи человека и некоторых аминокислот. Показано, что некоторые линии отражения кожи совпадают со следующими аминокислотами: триптофан, цистеин, аланин, глицин, фенилаланин, гистидин. Указанные аминокислоты содержатся в больших количествах в коже и соединительной ткани. ТГц спектральные линии указанных аминокислот связаны с колебательными модами молекул, которые обусловлены химическими реакциями, процессами гидратации и конформных изменений молекул в коже человека. Определение спектральных линий аминокислот и других характерных молекул, участвующих в метаболических процессах кожи, необходимо для разработки ТГц диагностического прибора с максимальной чувствительностью и селективностью.

Некоторые линии в ТГц спектрах кожи человека соотнесены с поглощением воды, находящейся во внеклеточной жидкости, а также с растворенным в ней молекулярным кислородом. Таким образом, изменение состояния воды и кислорода под влиянием ТГц излучения может вызвать изменение конформации примембранных белков в структуре кожи.

Полученные результаты актуальны для диагностики и терапии кожи человека, нормальной и с патологическими изменениями.

Работа поддержана Министерством образования и науки РФ (ГК № 14.B37.21.0248).

Литература

1. Езерская А.А., Романов И.В., Смолянская О.А., Грачев Я.В. Ранняя диагностика кариеса твердых тканей зубов с помощью лазерного излучения терагерцового диапазона // Научно-технический вестник СПбГУ ИТМО. – 2011. – № 6 (76). – С. 92–97.
2. Езерская А.А., Смолянская О.А., Парахуда С.Е., Грачев Я.В., Гончаренко А.О. Терагерцовые спектры пропускания и отражения катарактально измененных хрусталиков глаза человека // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – № 2 (78). – С. 33–36.
3. Joseph C.S., Yaroslavsky A.N., Al-Arashi M. Terahertz spectroscopy of intrinsic biomarkers for non-melanoma skin cancer // Proceedings of SPIE (SPIE, Bellingham, WA 2009). – V. 7215. – P. 72150I.
4. Laman N., Harsha S.S., Grischkowsky D., Melinger J.S. High-resolution waveguide thz spectroscopy of biological molecules // Biophysical Journal. – 2008. – V. 94. – P. 1010–1020.
5. Xu J., Plaxco K.W., Allen S.J. Probing the collective vibrational dynamics of a protein in liquid water by terahertz absorption spectroscopy // Protein Science. – 2006. – V. 15. – P. 1175–1181.
6. Казаринов К.Д. Биологические эффекты электромагнитного поля терагерцового диапазона // Электронная техника. Сер. 1: СВЧ-техника. – 2009. – Т. 503. – № 4. – С. 48–58.
7. Альбумин сыворотки крови в клинической медицине. – Кн. 2 / Под ред. Ю.А. Грызунова, Г.Е. Добрецова. – М.: ГЭОТАР, 1998. – 440 с.
8. Добрецов Г.Е., Миллер Ю.И. Биохимия и физико-химия сывороточного альбумина. Центры связывания органических молекул // Альбумин сыворотки крови в клинической медицине. – Кн. 1 / Под ред. Ю.А. Грызунова и Г.Е. Добрецова. – М.: ИРИУС, 1994. – С. 13–28.
9. Капралова А.В., Погодин А.С. Влияние терагерцового излучения различных диапазонов на конформацию молекул бычьего сывороточного альбумина // Вестник НГУ. Серия: Физика. – 2010. – Т. 5. – Вып. 4. – С. 182–185.
10. Аминокислоты [Электронный ресурс]. – Режим доступа: http://www.strong-atlet.net/main/dobavki/133-aminokisloty.html, свободный. Яз. рус. (дата обращения 17.10.2012).
11. Globus T., Anwar Mehdi, DeMaria Anthony J., Shur Michael S. Low-Terahertz spectroscopy of liquid water // Terahertz Physics, Devices and Systems II // Proc. of SPIE. – 2007. – V. 6772. – P. 67720S-1–67720S-11.
12. Анцыгин В.Д., Мамрашев А.А., Николаев Н.А., Потатуркин О.И. Малогабаритный терагерцовый спектрометр с использованием второй гармоники фемтосекундного волоконного лазера // Автометрия. – 2010. – Т. 46. – № 3. – С. 110–117.
13. Martin van Exter, Fattinger Ch., Grischkowsky D. Terahertz time-domain spectroscopy of water vapor // Optics Letters. – 1989. – V. 14. – № 20. – P. 1128–1130.
14. Айвазян Г.М. Распространение миллиметровых и субмиллиметровых волн в облаках. – Л.: Гидрометеоиздат, 1991. – 480 с.

Езерская Анна Александровна – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, студент, a.a.ezerskaya@gmail.com

Цуркан Мария Валерьевна – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, tsurkan.maria@yandex.ru

Смолянская Ольга Алексеевна – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат физ.-мат. наук, доцент, o_smolyanskaya@mail.ru
Капралова Ангелина Владимировна – Институт лазерной физики СО РАН, аспирант, kapralova@ngs.ru
Балбекин Николай Сергеевич – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, студент, godnikolay@yandex.ru
Романов Илья Владимирович – Томский государственный университет, аспирант, roil@elefot.tsu.ru
Прожеев Игорь Александрович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, студент, prozheev.iv@gmail.com
Стрепитов Евгений Александрович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, студент, e-strep@yandex.ru
Просанов Игорь Юрьевич – Институт химии твердого тела и механохимии СО РАН, доктор технических наук, ст. научный сотрудник, prosanov@mail.ru

УДК 53.086

ИССЛЕДОВАНИЕ МИКРОСТРУКТУРЫ ПОВЕРХНОСТНЫХ СЛОЕВ РАСТИТЕЛЬНОЙ ТКАНИ МЕТОДОМ ОПТИЧЕСКОЙ КОГЕРЕНТНОЙ МИКРОСКОПИИ

М.А. Волынский, И.П. Гуров, Е.В. Жукова, Н.Б. Маргарянц, Е.С. Рысева

Исследована микроструктура поверхностных слоев биологических тканей на примере изучения экзокарпия растительных плодов методом спектральной оптической когерентной микроскопии с перестраиваемой длиной волны в пределах 1305±75 нм и методом корреляционной оптической когерентной микроскопии при среднем значении длины волны 940 нм. Эксперименты выполнены для неповрежденных и дефектных структур, выполнен расчет профиля поверхности экзокарпия.
Ключевые слова: оптическая когерентная микроскопия, поверхность, 3D-изображение, экзокарпий.

Введение

Оптическая когерентная микроскопия – это современный метод неразрушающего контроля и исследования разнообразных сред, обладающих сложным пространственным строением. Оптические свойства таких микроструктур случайно-неоднородны, что затрудняет их диагностику традиционными оптическими методами. Изучение биотканей растительного происхождения методами оптической когерентной томографии (ОКТ) позволило исследовать структуру объектов на клеточном уровне, обнаружить характер видоизменения тканей в момент протекания важнейших физиологических процессов [1–3]. Прикладное значение исследований состоит в создании инновационных технических решений и диагностических приборов, обладающих высоким разрешением и быстродействием [4].

Оптические методы диагностики широко используются в растениеводстве и в плодоовощной отрасли для контроля качества продукции, создания диагностических комплексов наблюдения за процессами выращивания, хранения и сортификации овощных и фруктовых культур. Изучение плодов семейства яблоневых, которые относятся к группе приоритетных культур, привлекает особое внимание. Для контроля плодов широко используются спектрофотометрические и лазерные методы, люминесцентный анализ, а также методы компьютерного зрения, позволяющие одновременно диагностировать многие параметры развивающегося плода [5–7]. Широко используются методы отражения оптического излучения для неразрушающего контроля состояния поверхности плодовых культур [8].

Известно, что многие потребительские и органолептические свойства плодов яблок определяются состоянием внешнего слоя плода или его кожуры (экзокарпия). Экзокарпий – это многофункциональная биологическая ткань, имеющая слоистую структуру. Кожица плода выполняет защитную функцию от проникновения в плод болезнетворных микроорганизмов и механического повреждения мягких тканей, а также обеспечивает протекание выделительных процессов, сопровождающихся выходом на поверхность эфирных масел и влаги. Анализ состояния экзокарпия используется как один из критериев для определения сорта, сохранности и пищевой пригодности.

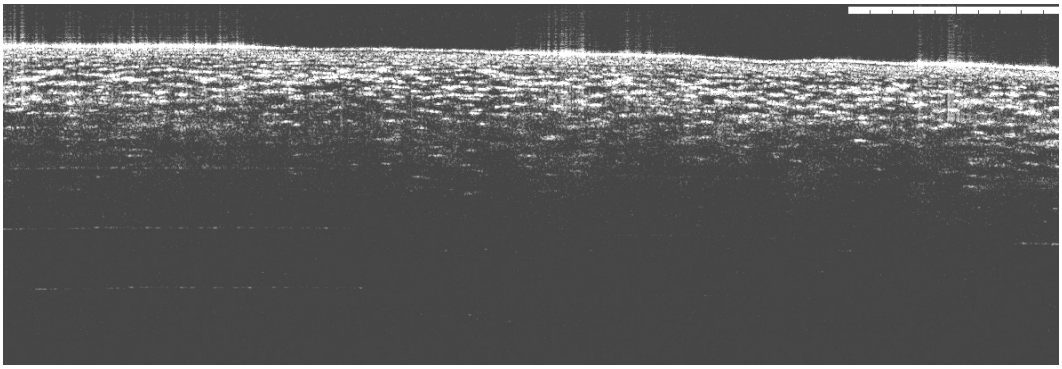
В настоящей работе были использованы спектральный и корреляционный методы оптической когерентной микроскопии для изучения экзокарпия плодов разных сортов.

Исследование микроструктуры ткани

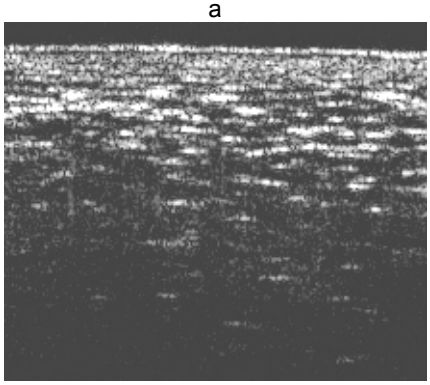
Для проведения экспериментальных исследований были выбраны сорта яблок Айдаред и Роял Гала, которые широко представлены на потребительском рынке, а также груши сорта Вильямс. Исследова-

ния микроструктуры поверхностного слоя выполнены на спектральном оптическом когерентном микроскопе (ОКМ), модель EX1301 OCT Microscope фирмы Michelson Diagnostics, Ltd (Великобритания).

Прибор оснащен источником излучения – лазером с перестраиваемой длиной волны. Максимум излучения находится на длине волны 1305 нм, диапазон перестройки по длинам волн составляет 150 нм. Мощность источника в области максимума излучения составляет 15 мВт. В процессе работы прибора можно выбирать ширину участка сканирования зондирующего излучения по поверхности изучаемого объекта в пределах 1–7 мм. Глубина сканирования составляет до 2 мм, при этом обеспечиваются условия для регистрации томограмм, так называемых В-сканов, представляющих собой изображение внутренней структуры объекта в плоскости, перпендикулярной поверхности, с разрешением по глубине около 10 мкм и в латеральной плоскости до 7,5 мкм. ОКМ обеспечивает возможность изучения микроструктуры биологических тканей. Принцип работы и возможности применения ОКМ для исследования оптических случайно-неоднородных сред и поверхностных слоев объектов органического и неорганического происхождения подробно рассмотрены в [9].



а

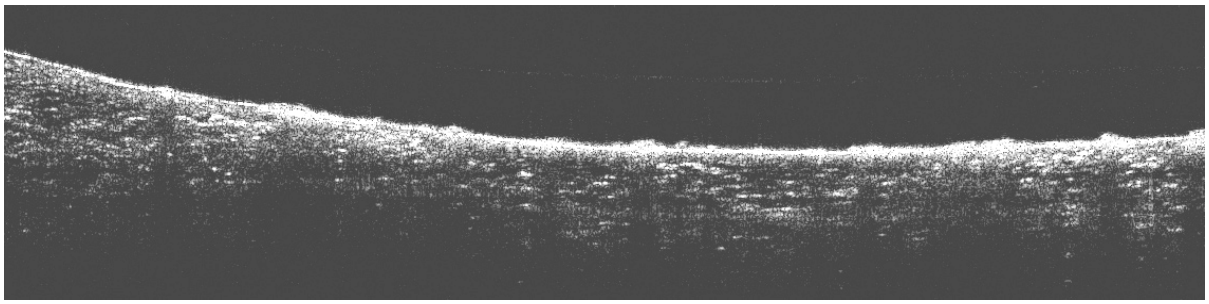


б

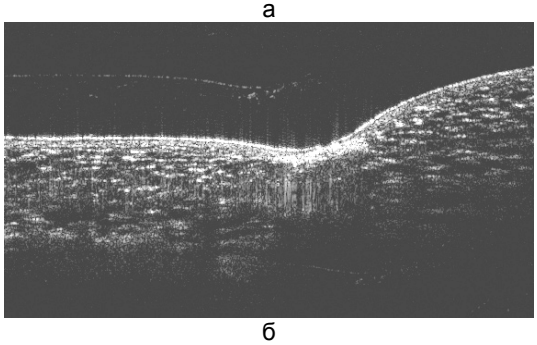
Рис. 1. В-сканы экзокарпия плода яблока сорта Айдаред: в области визуально здорового участка поверхности красного цвета (а), размер скана 5 мм×2 мм; экзокарпий и приповерхностная область мезокарпия (б), размер скана 0,91 мм×0,89 мм

На поверхности кожицы (экзокарпия) плода были выбраны несколько участков сканирования с целью сравнительного исследования тканей разной структуры и степени дефектности. На рис. 1, а, приведен В-скан экзокарпия и прилегающей к нему области для визуально здорового участка кожицы плода, который имел красную сплошную покровную окраску. Видно, что поверхность экзокарпия, обладая хорошими отражающими свойствами, все же пропускает излучение на глубину в пределах до 700 мкм. На томограмме по всей ширине сканирования просматривается ткань мякоти плода. Этот скан можно рассматривать как типичный для неповрежденных плодов яблок. Толщину экзокарпия можно оценить величиной 8–10 мкм. Под слоем экзокарпия, состоящего из слоев кутикулы и эпидермиса, находится слой мякоти (мезокарпий), на томограмме, приведенной на рис. 1, б, просматривается клеточное строение этого слоя.

Строение экзокарпия и прилежащих к нему тканей со структурными видоизменениями разного происхождения были также исследованы с помощью ОКМ. Так, на рис. 2, а, приведены томограммы для участков шероховатого уплотнения кожицы воронки яблока около плодоножки. Изменения кожицы плодов яблок в пределах воронки плодоножки используют в ряде случаев для определения сорта яблок. Видно, что на томограмме не просматривается микроструктура экзокарпия, для данной области характерно образование тканей с выраженной сильно рассеивающей структурой. На рис. 2, б, приведен пример дефекта в виде локального уплотнения ткани экзокарпия. На В-скане наблюдается как здоровая ткань мякоти плода, так и изменение структуры клеточной ткани в области дефекта. Уплотнение биотканей меняет условия визуализации слоев, лежащих под экзокарпием.



а



б

Рис. 2. В-сканы экзокарпия плода яблока сорта Айдаред: область уплотнения возле воронки у плодоножки (а), размер скана 4,67 мм×1,33 мм; в области дефекта на поверхности экзокарпия (б), размер скана 2,38 мм×1,54 мм

Для исследования был также выбран участок на поверхности экзокарпия, под которым просматривалась область потемневшей мякоти плода, томограмма приведена на рис. 3. В сравнении с томограммой здорового участка плода (рис. 1, а) видно, что процесс потемнения ткани сопровождается появлением локальных уплотнений клеточных структур под слоем экзокарпия. В данном случае эта стадия изменения микроструктуры соответствует начальному этапу необратимой деградации мякоти. Для сравнения были выполнены исследования микроструктуры приповерхностной к экзокарпию области для яблок сорта Роял Гала и груш сорта Вильямс. Результаты представлены соответственно на рис. 4, 5.

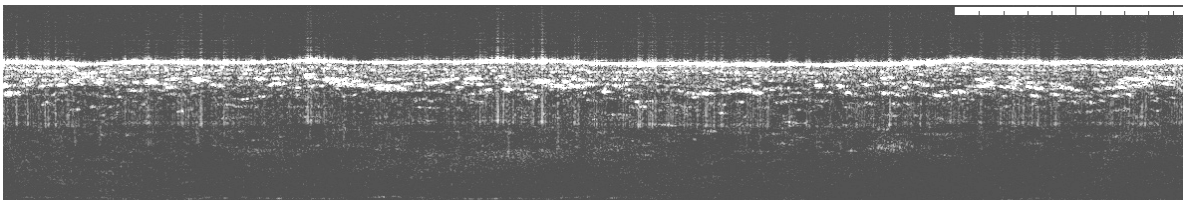
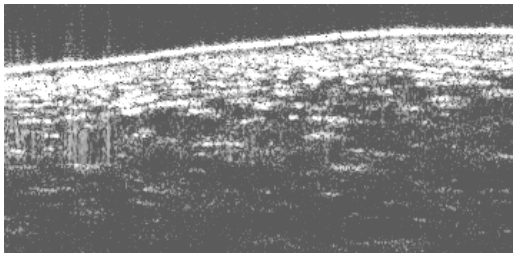
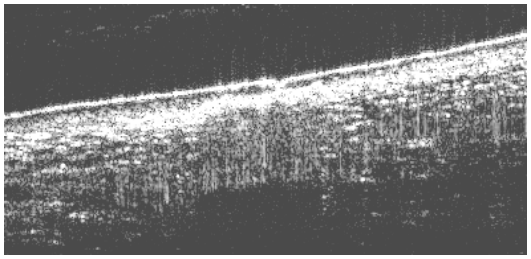


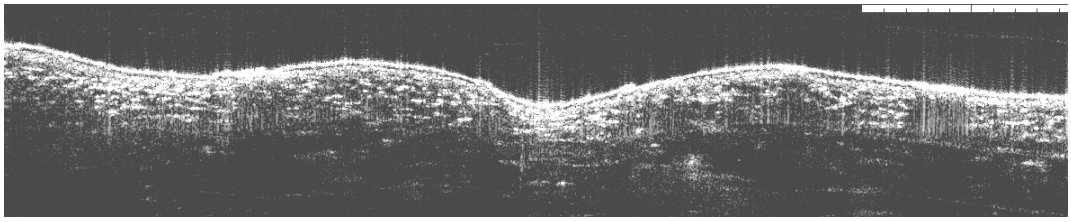
Рис. 3. В-скан поверхности плода яблока сорта Айдаред в области визуального потемнения под поверхностью экзокарпия, размер скана 4,88 мм×0,94 мм



а

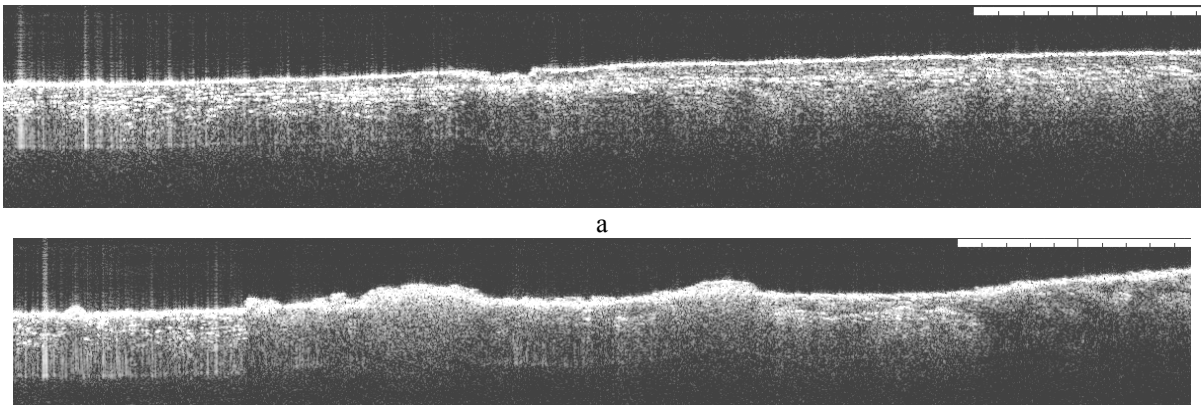


б



в

Рис. 4. В-сканы экзокарпия плода яблока сорта Роял Гала: здоровая поверхность экзокарпия (а), размер скана 1,47 мм×0,83 мм; в области уплотнения (б), размер скана 1,62 мм×0,91 мм; в области сморщивания кожицы плода яблока (в), размер скана 4,88 мм×0,94 мм



а

б

Рис. 5. В-сканы экзокарпия плода груши сорта Вильямс: визуально здоровая поверхность (а), размер скана 4,86 мм×0,95 мм; область локальных наростов (б), размер скана 4,90 мм×0,81 мм

На томограмме (рис. 4, б) видна область локального видоизменения клеточных структур под слоем эпидермиса, а характер изменения поверхности экзокарпия плода яблока, вызванный процессами усыхания, показан на рис. 4, в. Результаты изучения экзокарпия плода груши представлены двумя томограммами – для здорового и дефектного участков (рис. 5). Видно, что деградация биотканей сопровождается видоизменением в пределах значительной толщины. Толщины экзокарпия сорта яблок Роял Гала и груши сорта Вильямс можно оценить соответственно в 18 мкм и 9 мкм.

Исследование профиля поверхности экзокарпия

Поверхность одного плода яблок сорта Айдаред была дополнительно исследована методом корреляционной оптической когерентной микроскопии. Для этого был использован высокоразрешающий автоматический интерферометр малой когерентности, собранный на базе микроинтерферометра МИИ-4. В эксперименте источником излучения служил ИК диод (модель EDEN1LS3) мощностью 140 мВт, максимум излучения которого находится на длине волны 940 нм. Регистрация отраженного излучения проводилась в автоматическом режиме с использованием черно-белой видеокамеры (модель EVS VAC-135). Разрешение матрицы видеокамеры составляет 1280×1024 пикселов, область спектральной чувствительности перекрывает диапазон 0,4–1,1 мкм. Сочетая возможности микроскопа с увеличением 500$^{\times}$ и микроинтерферометра малой когерентности, прибор позволяет выполнить регистрацию томограммы (В-скана) в изучаемом сечении объекта с латеральным разрешением в 1 мкм и аксиальным разрешением менее 1 мкм. Принцип работы прибора и методика количественной обработки данных, полученных в ходе эксперимента с целью получения томографического изображения, подробно рассмотрены в [10].

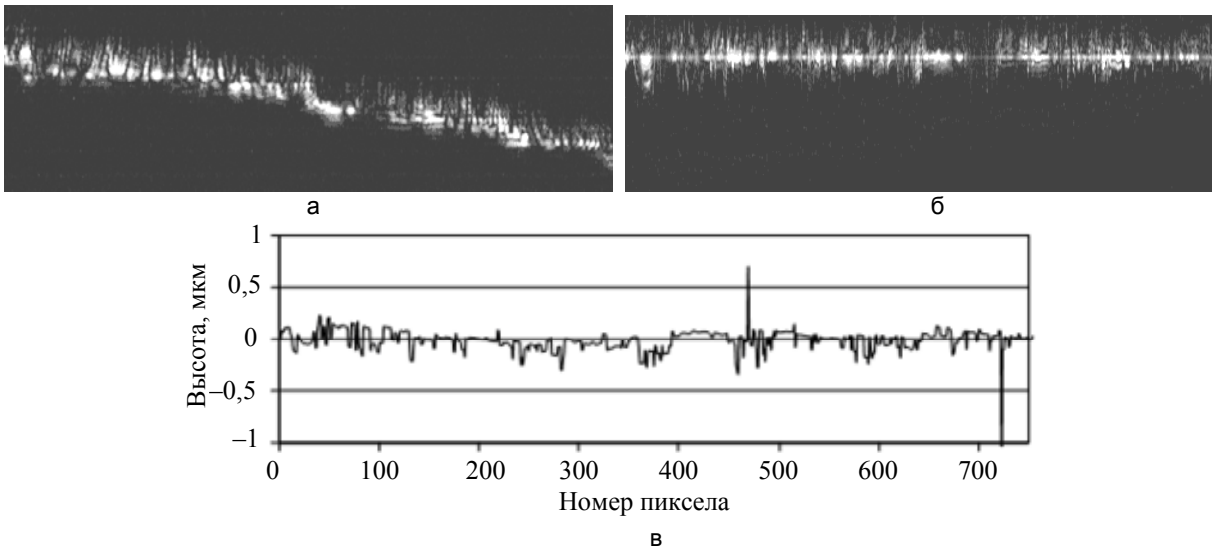


Рис. 6. В-скан экзокарпия плода яблока сорта Айдаред (а); В-скан после вычитания базовой линии (б); реконструированный профиль поверхности (в). Параметры сканирования: ширина 200 мкм, глубина 26 мкм

Для проверки воспроизводимости результатов измерений высот неровностей профиля экзокарпия и определения погрешности измерений были проведены два эксперимента. Первый эксперимент вклю-

чал регистрацию томограмм для различных участков на поверхности экзокарпия, во втором эксперименте была выполнена регистрация пяти томограмм одного и того же участка на поверхности плода яблока. Далее проведена обработка полученных данных для расчета рельефа поверхности экзокарпия.

На рис. 6, а, представлен В-скан экзокарпия плода яблока, полученный при однократном измерении в первом эксперименте. Для расчета высот микронеровностей профиля было необходимо выполнить процедуру вычитания базовой линии скана, результат представлен на рис. 6, б. Одномерный профиль микровысот исследуемой структуры был получен путем поиска положения пиксела с максимальной яркостью для каждого А-скана (столбца В-скана, представляющего собой распределение отражающей способности образца по его глубине). На рис. 6, в, показан реконструированный по результатам расчета профиль микронеровностей в зависимости от номера пиксела.

При анализе было рассчитано среднее квадратичное отклонение (СКО) высоты микронеровностей, величина которого для данного профиля (рис. 6, в) составила 0,69 мкм. Величину СКО следует рассматривать как основную характеристику микрорельефа экзокарпия для изучаемого образца. При измерении профиля других участков экзокарпия того же образца с субъективно сходным строением были получены значения СКО, равные 0,57; 0,47; 0,65 мкм, с погрешностью не более 0,038 мкм, что свидетельствует о возможности идентификации особенностей структуры экзокарпия на основе дополнительной статистической обработки В-сканов, что невозможно при субъективном анализе. Следует отметить, что указанная погрешность измерений не превышает методической погрешности, равной 0,8 мкм [10], поэтому достаточно одного измерения профиля на любом морфологически эквивалентном участке.

Во втором эксперименте для проверки воспроизводимости результатов измерений профиля поверхности были рассчитаны пять профилей поверхности экзокарпия в одном и том же месте поверхности плода. Для сохранения условий измерений образец непрерывно находился на столике микроинтерферометра и были приняты меры для минимизации случайных погрешностей, вызванных внешней вибрацией и смещением объекта, что позволило получить набор томограмм для одной и той же области экзокарпия. Для изучения воспроизводимости результатов измерений высот микрорельефа достаточно сравнения группы одномерных профилей исследуемой поверхности экзокарпия, поскольку количество рассматриваемых точек поверхности образца, равное 750, достаточно для надежных статистических оценок. Результаты вычисления среднего значения (математического ожидания) и СКО ошибки восстановления рельефа по пяти независимым измерениям показаны на рис. 7.

Из рис. 7, б, видно, что погрешность в среднем не превышает 0,6 мкм, в то время как методическая погрешность составляет 0,8 мкм [10], что позволяет судить о воспроизводимости результатов. Следует отметить, что при ширине В-скана в 200 мкм латеральное разрешение метода ограничено апертурой объектива и дифракционным пределом и составляет 1 мкм, поэтому для оценки погрешности измерения высот микрорельефа целесообразно выполнять усреднение значений по нескольким пикселам.

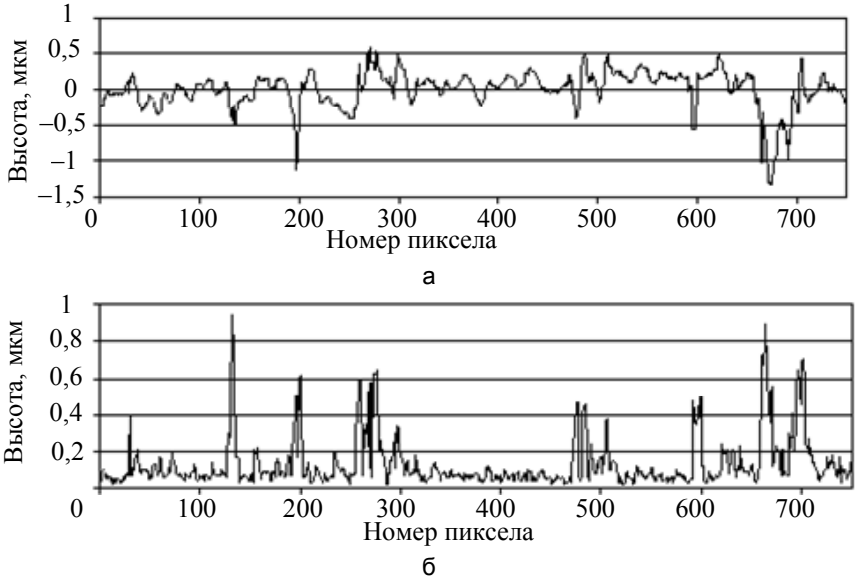


Рис. 7. Математическое ожидание (а) и дисперсия (б) оценки профиля поверхности экзокарпия для каждой точки поверхности

Заключение

Исследование экзокарпия плодов показало возможность применения метода спектральной оптической когерентной микроскопии для изучения микроструктуры поверхностных тканей таких объектов. Излучение с длиной волны 1305 нм проникает в ткани плода на глубину до 700 мкм, что позволяет визуализировать исследуемые области. На зарегистрированных В-сканах видны как слой экзокарпия, так и

структура прилежащих к нему тканей мезокарпия, обладающих довольно крупными клетками. Возможно наблюдение строения клеточной ткани и пустот, относящихся к межклеточному пространству. Уплотнение экзокарпия и деградация ткани мякоти под слоем эпидермиса сопровождаются локальными деформациями клеточных структур и приводят к изменению оптических свойств, увеличению рассеяния и отражения на образующихся микронеоднородностях. В работе были выполнены оценки толщин экзокарпия для сорта яблок Айдаред и Роял Гала, а также плода груши сорта Вильямс. Регистрация набора B-сканов позволяет выполнить реконструкцию исследуемой области в виде 3D-изображения [10] и последующий анализ состояния тканей с использованием методик сегментации и текстурного анализа. Такое представление данных широко применяется в материаловедении и промышленной диагностике и может быть перспективно для анализа состояния качества микроструктуры тканей плодов.

С целью количественного описания свойств поверхностного слоя экзокарпия выполнены профилометрические исследования методом оптической когерентной микроскопии. На примере изучения поверхности экзокарпия плодов яблока сорта Айдаред продемонстрированы возможности диагностики поверхностного микрорельефа, выполнены расчеты его высот, проведена оценка СКО. Эксперименты по проверке воспроизводимости результатов измерения микровысот показали, что СКО может быть рекомендовано в качестве характеристики экзокарпия.

Представленные результаты исследований и методика проведения анализа приповерхностной микроструктуры образцов могут быть полезны широкому кругу исследователей и рекомендованы для практического использования.

Работа выполнена при финансовой поддержке Министерства образования и науки Российской Федерации.

Литература

1. Сапожникова В.В., Каменский В.А., Куранов Р.В. Визуализация растительных тканей методом оптической томографии // Физиология растений. – 2003. – Т. 50. – № 2. – С. 316–320.
2. Hettinger J.W., Mattozzi M., Myer W.R., Williams M.E., Reeves A., Parsons R.L., Haskell R.C., Petersen D.C., Wang R., Medford J.I. Optical Coherence Microscopy. A technology for rapid, in vivo, non-destructive visualization of plants and plant cells // Plant Physiology. – 2000. – V. 123. – P. 3–15.
3. Кутис И.С., Сапожникова В.В., Куранов Р.В., Каменский В.А. Исследование методами оптической когерентной микроскопии и оптической когерентной томографии формо-функционального состояния тканей высших растений // Физиология растений. – 2005. – Т. 52. – № 4. – С. 635–640.
4. Shiina T., Kishiwaki D., Ito M., Honda T., Okamura Y. OCT system for plant measurement // Proc. of SPIE. – 2005. – V. 5948. – P. 59482D1–59482D9.
5. Родиков С.А. Методы и устройства определения зрелости яблок. – М.: ФИЗМАТЛИТ, 2009. – 216 с.
6. Kondo N. Machine vision based on optical properties of biomaterials for fruit grading system // Environ. Contr. Biol. – 2006. – V. 44. – № 3. – P. 151–159.
7. Бугадовская О.Н. Оптические методы диагностики зрелости и качества плодоовощной продукции // Вестник Мичуринского аграрного университета. – 2011. – № 2. – Ч. 2. – С. 83–92.
8. Solovchenko A.E. Relationships between chlorophyll and carotenoid pigments during on- and offtree ripening of apple fruit as revealed non-destructively with reflectance // Postharvest biology and Technology. – 2005. – V. 38. – P. 9–17.
9. Гуров И.П., Жукова Е.В., Маргарянц Н.Б. Исследование внутренней микроструктуры материалов методом оптической когерентной микроскопии с перестраиваемой длиной волны // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – № 3 (79). – С. 40–44.
10. Волынский М.А., Гуров И.П., Жукова Е.В. Рекуррентные алгоритмы обработки видеоинформации в системах оптической когерентной томографии // Оптический журнал. – 2012. – Т. 79. – № 11. – С. 28–35.

Волынский Максим Александрович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат технических наук, доцент, maxim.volynsky@gmail.com

Гуров Игорь Петрович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, профессор, зав. кафедрой, gurov@mail.ifmo.ru

Жукова Екатерина Владимировна – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат технических наук, доцент, zhukova-ifmo@mail.ru

Маргарянц Никита Борисович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, старший преподаватель, fosp@grv.ifmo.ru

Рысева Екатерина Сергеевна – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, ri.katerina@gmail.com

3 АНАЛИЗ И СИНТЕЗ СЛОЖНЫХ СИСТЕМ

УДК 62-50

ОБЕСПЕЧЕНИЕ ПАРАМЕТРИЧЕСКОЙ ε-ИНВАРИАНТНОСТИ ВЫХОДА СИСТЕМЫ ПРИ НЕОПРЕДЕЛЕННОСТИ ОБЪЕКТА УПРАВЛЕНИЯ

С.А. Александрова, А.А. Мусаев, О.В. Слита, А.В. Ушаков

Ставится задача обеспечения параметрической ε-инвариантности выхода непрерывной системы при неопределенности матрицы состояния исходного объекта. Задача решается для случая матрицы состояния объекта, заданной в сопровождающей форме, для которой оказываются невыполнимыми условия достижения абсолютной параметрической инвариантности. Результаты иллюстрируются примером.

Ключевые слова: неопределенность задания матрицы, сопровождающая форма, параметрическая ε-инвариантность, собственные значения и векторы.

Введение

Современные методы анализа и синтеза динамических систем позволяют решать задачи управления ими в условиях системных неопределенностей задания модели объекта [1–5]. Одной из постановочных версий управления в условиях неопределенности является задача обеспечения инвариантности выхода системы к неопределенности параметров модели исходного объекта [4, 5] или параметрической инвариантности.

В работе приводятся алгебраические условия достижимости абсолютной параметрической инвариантности [2, 4, 5] и делается переход к ситуации, когда полученные алгебраические условия оказываются невыполнимыми, что явилось мотивацией к постановке задачи обеспечения ε-инвариантности выхода системы с заданной оценкой величины ε.

Алгебраические условия обеспечения абсолютной параметрической инвариантности выхода системы

Рассматривается непрерывный объект управления (ОУ)

$$\dot{\mathbf{x}}(t) = (\mathbf{A} + \Delta\mathbf{A})\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)\,,\ \ \boldsymbol{x}(t)\big|_{t=0} = \boldsymbol{x}(0)\,,\ \ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)\,, \tag{1}$$

где $\mathbf{x} \in R^n, \mathbf{u} \in R^r, \mathbf{y} \in R^m$; $\mathbf{x}$, $\mathbf{u}$, $\mathbf{y}$ – векторы состояния, управления и выхода $\mathbf{A} \in R^{n\times n}, \mathbf{B} \in R^{n\times r}$, $\mathbf{C} \in R^{m\times n}$, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ – матрицы состояния, управления и выхода, при этом ($\mathbf{A}$, $\mathbf{B}$) и ($\mathbf{A}$, $\mathbf{C}$) образуют соответственно управляемую и наблюдаемую пары; $\Delta\mathbf{A} \in R^{n\times n}$ – неопределенность задания матрицы состояния.

Закон управления (ЗУ) объектом (1) формируется в виде

$$\mathbf{u}(t) = \mathbf{K}_g g(t) - \mathbf{K}\mathbf{x}(t)\,, \tag{2}$$

где $g(t)$ – задающее воздействие, матрица $\mathbf{K}$ с использованием метода модального управления [6] находится с помощью системы уравнений

$$\mathbf{M}\boldsymbol{\Lambda} - \mathbf{A}\mathbf{M} = -\mathbf{B}\mathbf{H}\,, \tag{3}$$
$$\mathbf{K} = \mathbf{H}\mathbf{M}^{-1}\,,$$

в которых наблюдаемая пара матриц $(\boldsymbol{\Lambda}, \mathbf{H})$ модальной модели задает желаемые динамические свойства системы (путем назначения структуры собственных чисел диагональной матрицы $\boldsymbol{\Lambda}$). Следует отметить, что если известны матрицы $\mathbf{A}, \mathbf{B}, \boldsymbol{\Lambda}, \mathbf{H}$, то уравнение (3) решается относительно матрицы $\mathbf{M}$, а если известны $\mathbf{A}, \mathbf{B}, \boldsymbol{\Lambda}, \mathbf{M}$, то уравнение (3) решается относительно матрицы $\mathbf{H}$ в форме

$$\mathbf{H} = \left(\mathbf{B}^T\mathbf{B}\right)^{-1}\mathbf{B}^T(\mathbf{A}\mathbf{M} - \mathbf{M}\boldsymbol{\Lambda})\,.$$

Замкнутая система, образованная ОУ (1) и ЗУ (2), записывается как

$$\dot{\mathbf{x}}(t) = \mathbf{F}\mathbf{x}(t) + \mathbf{G}g(t) + \Delta\mathbf{F}\mathbf{x}(t);\quad \mathbf{x}(0)\,; \tag{4}$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t);$$
$$\mathbf{F} = \mathbf{A} - \mathbf{B}\mathbf{K}\,,\ \mathbf{G} = \mathbf{B}\mathbf{K}_g\,, \Delta\mathbf{F} = \Delta\mathbf{A}\,.$$

ЗУ (2) должен обеспечивать желаемые показатели качества замкнутой системы при наличии в исходном объекте параметрической неопределенности, т.е. параметрическую инвариантность ее выхода. Для целей дальнейших исследований представим сигнальный компонент $\Delta\mathbf{F}\mathbf{x}(t)$ в декомпозированной форме:

$$\Delta\mathbf{F}\mathbf{x}(t)=\Delta\mathbf{A}\mathbf{x}(t)=\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}\begin{bmatrix}\Delta A_{11} & \Delta A_{12} & \dots & \Delta A_{1n}\end{bmatrix}\begin{bmatrix}x_1(t)\\x_2(t)\\\vdots\\x_n(t)\end{bmatrix}+\begin{bmatrix}0\\1\\\vdots\\0\end{bmatrix}\begin{bmatrix}\Delta A_{21} & \Delta A_{22} & \dots & \Delta A_{2n}\end{bmatrix}\begin{bmatrix}x_1(t)\\x_2(t)\\\vdots\\x_n(t)\end{bmatrix}+$$

$$+\dots+\begin{bmatrix}0\\0\\\vdots\\1\end{bmatrix}\begin{bmatrix}\Delta A_{n1} & \Delta A_{n2} & \dots & \Delta A_{nn}\end{bmatrix}\begin{bmatrix}x_1(t)\\x_2(t)\\\vdots\\x_n(t)\end{bmatrix}=\mathbf{D}\eta(t)\,. \qquad (5)$$

Вектор параметрического воздействия $\boldsymbol{\eta}(t)$ сформирован на правых мультипликативных компонентах элементов разложения $\Delta\mathbf{F}\mathbf{x}(t)$ в форме (5) в силу представления

$$\boldsymbol{\eta}(t)=col\left\{\eta_j=\begin{bmatrix}\Delta A_{j1} & \Delta A_{j2} & \dots & \Delta A_{jn}\end{bmatrix}\begin{bmatrix}x_1(t)\\x_2(t)\\\vdots\\x_n(t)\end{bmatrix}=(\Delta A)^j x(t)=h_j^T x(t)\right\}, \qquad (6)$$

где $(\Delta A)^j$ – j-ая строка матрицы $\Delta\mathbf{A}$. На левых сомножителях слагаемых этого выражения сформируем матрицу параметрического воздействия $\mathbf{D}=row\left\{\mathbf{D}_j=\left[0_{(j-1)\times 1};1;0_{(n-j)\times 1}\right];\, j=\overline{1,n}\right\}$.

С использованием (6) запишем уравнение (4) системы в виде

$$\dot{\xi}(t)=\mathbf{F}\mathbf{x}(t)+\mathbf{G}\mathbf{g}(t)+\mathbf{D}\boldsymbol{\eta}(t);\ \mathbf{y}(t)=\mathbf{C}\mathbf{x}(t)\,. \qquad (7)$$

Представление (7) системы (4) позволяет переформулировать задачу обеспечения абсолютной параметрической инвариантности $\boldsymbol{y}(t,\mathbf{F},g(t),\Delta\mathbf{A}\neq 0)=\boldsymbol{y}(t,\mathbf{F},g(t),\Delta\mathbf{A}\equiv 0)$ как задачу обеспечения сигнальной инвариантности [4, 7]

$$\mathbf{y}(t,\mathbf{F},g(t),\eta(t)\neq 0)=\mathbf{y}(t,\mathbf{F},g(t),\eta(t)\equiv 0)\,. \qquad (8)$$

Запишем выражение (8) в терминах преобразований Лапласа и передаточных функций (матриц)

$$Y(s,g(s),\eta(s)\neq 0)=\boldsymbol{\Phi}_{yg}(s)g(s)+\boldsymbol{\Phi}_{y\boldsymbol{\eta}}(s)\eta(s)=\boldsymbol{\Phi}_{yg}(s)g(s)\,, \qquad (9)$$

где $g(s)$ – Лапласов образ задающего воздействия $g(t)$; $\eta(s)$ – Лапласов образ «параметрического» воздействия $\boldsymbol{\eta}(t)$; $\boldsymbol{\Phi}_{yg}(s)$ – передаточная функция (матрица) отношения «задающее воздействие–выход системы»; $\boldsymbol{\Phi}_{y\eta}(s)$ – передаточная функция (матрица) отношения «параметрическое» воздействие–выход системы». Очевидно, что равенство (9) при $\eta(s)\neq 0$ выполняется, когда

$$\boldsymbol{\Phi}_{y\eta}(s)=0\,. \qquad (10)$$

Соотношение (10) представляет собой «сигнальный» аналог инвариантности выхода (ошибки) к неопределенностям задания матрицы состояния исходного объекта, которое выполняется при любых реализациях внешнего задающего воздействия $g(t)$. Приведем утверждение, одно из положений которого базируется на структуре собственных векторов [8] матрицы состояния проектируемой системы.

Утверждение 1. [3, 5]. Для того чтобы система (4) обладала абсолютной параметрической инвариантностью выхода к неопределенности задания матрицы исходного объекта или чтобы система (10) обладала сигнальной инвариантностью выхода относительно внешнего «параметрического» воздействия $\eta(t)$ в смысле условий (8), (9), т.е. чтобы передаточная функция (матрица) «параметрический» вход η – выход системы y » $\boldsymbol{\Phi}_{y\eta}(s)$ была бы нулевой и выполнялось равенство (10),

$$\boldsymbol{\Phi}_{y\eta}(s)=\mathbf{C}(s\mathbf{I}-\mathbf{F})^{-1}\mathbf{D}=\ \mathrm{row}\left\{\boldsymbol{\Phi}_{y\eta_j}=\mathbf{C}(s\mathbf{I}-\mathrm{F})^{-1}\mathbf{D}_j;\ j=\overline{1,p};1\le p\le n\right\}=0\,,$$

достаточно, чтобы

1. столбцы D_j матрицы $\mathbf{D}$ были собственными векторами матрицы $\mathbf{F}$;
2. столбцы D_j принадлежали ядру матрицы $\mathbf{C}$, т.е. чтобы выполнялось соотношение $\mathbf{C}D_j=0$.

Проблема параметрической ε-инвариантности

Проблема параметрической ε-инвариантности возникает в случае, когда невозможно достижение абсолютной параметрической инвариантности, т.е. когда не выполняется какое-либо из условий утверждения 1. Проиллюстрируем эту ситуацию [9] на примере объекта управления, заданного в сопровождающем управляемом базисе

$$\mathbf{A}=\begin{bmatrix}0 & 1 & 0 & \dots & 0\\ 0 & 0 & 1 & \dots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \dots & 1\\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-1}\end{bmatrix},\ \mathbf{B}=\begin{bmatrix}0\\0\\\vdots\\0\\b_n\end{bmatrix},\ \mathbf{C}=\begin{bmatrix}1 & 0 & 0 & \dots & 0\end{bmatrix}. \tag{11}$$

В этом случае любые системные неопределенности $\Delta\mathbf{A}$ возмущают только последнюю строчку матрицы состояния, и

$$\Delta\mathbf{A}=\begin{bmatrix}0 & 0 & 0 & \dots & 0\\ 0 & 0 & 0 & \dots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \dots & 0\\ \Delta a_0 & \Delta a_1 & \Delta a_2 & \dots & \Delta a_{n-1}\end{bmatrix}=\begin{bmatrix}0\\0\\\vdots\\0\\1\end{bmatrix}\cdot\begin{bmatrix}\Delta a_0 & \Delta a_1 & \Delta a_2 & \dots & \Delta a_{n-1}\end{bmatrix},$$

откуда для матрицы $\mathbf{D}$ получаем представление в виде матрицы-столбца

$$\mathbf{D}=\begin{bmatrix}0 & 0 & \cdots & 0 & 1\end{bmatrix}^T. \tag{12}$$

В силу канонической сопровождающей формы задания матрицы $\mathbf{A}$ и вида матрицы $\mathbf{B}$ (11) ЗУ (2) сохраняет каноническую сопровождающую форму матрицы состояния $\mathbf{F}$ системы [9], которая имеет собственные вектора ξ_i, формируемые по схеме Вандермонда:

$$\xi_i=\begin{bmatrix}1 & \lambda_i & \lambda_i^2 & \dots & \lambda_i^{n-1}\end{bmatrix}^T;\left(i=\overline{1,n}\right). \tag{13}$$

Как видно из структуры (13) собственных векторов, матрица-столбец $\mathbf{D}$ (12) не совпадает ни с одним из собственных векторов $\mathbf{F}$. Таким образом, при представлении матрицы состояния в форме (11) первое условие утверждения 1 выполнено не будет. Следовательно, абсолютная параметрическая инвариантность выхода для случая матрицы состояния, заданной в сопровождающей форме, недостижима. В этом случае следует перейти к обеспечению ε - инвариантности выхода проектируемой системы.

Оценка $\hat{\varepsilon}$ величины параметрической ε-инвариантности

Задачу формирования оценки $\hat{\varepsilon}$ величины параметрической ε-инвариантности решим в два этапа.

Этап 1. Модификация представления собственных векторов, построенных по схеме Вандермонда; им придается вид

$$\xi_i=\begin{bmatrix}\left(\lambda_i^{n-1}\right)^{-1} & \dots & \left(\lambda_i^2\right)^{-1} & \left(\lambda_i\right)^{-1} & 1\end{bmatrix}^T;\left(i=\overline{1,n}\right).$$

Этап 2. Представление матрицы-столбца $\mathbf{D}$ (12) в виде проекции на собственный вектор ξ_1 в форме $\alpha\xi_1=\mathbf{D}$, в которой коэффициент α ищется с помощью алгоритма Грама [10] из условия $\alpha=\dfrac{\left(\xi_1,\mathbf{D}\right)}{\left(\xi_1,\xi_1\right)}=\dfrac{\xi_1^T\mathbf{D}}{\xi_1^T\xi_1}=\dfrac{1}{1+\left(\lambda_1^{-1}\right)^2+\ldots+\left(\lambda_1^{-(n-1)}\right)^2}$.

Определим оценку $\hat{\varepsilon}$ на основе нормы невязки представления матрицы-столбца $\mathbf{D}$ его проекцией $\hat{\mathbf{D}}(\lambda_1)=\alpha\xi_1$ на вектор ξ_1 в форме

$$\hat{\varepsilon}(\lambda_1)=\left\{\left\|\mathbf{D}-\hat{\mathbf{D}}(\lambda_1)\right\|\Big/\left\|\mathbf{D}\right\|\right\}\cdot 100\%.$$

Пример

В качестве ОУ рассмотрим электропривод, исполнительный двигатель которого обладает механической характеристикой, содержащей восходящий участок [11]. Представление ОУ в виде (1) характеризуется матрицами $\mathbf{A}+\left[\Delta\mathbf{A}\right]=\begin{bmatrix}0 & 1\\ 0 & \left[-3,5;\,0,5\right]\end{bmatrix}$ с медианной составляющей $\mathbf{A}=\begin{bmatrix}0 & 1\\ 0 & -1,5\end{bmatrix}$ и интервальной $\left[\Delta\mathbf{A}\right]=\begin{bmatrix}0 & 0\\ 0 & \left[-2;2\right]\end{bmatrix}$, $\mathbf{B}=\begin{bmatrix}0 & 1\end{bmatrix}^T$, $\mathbf{C}=\begin{bmatrix}1 & 0\end{bmatrix}$.

Зададим требования к переходному процессу в виде его длительности $t_{\text{пп}}=2,4$ с и величины перерегулирования $\sigma=0\,\%$ в системе, в которую войдет объект. Синтезируем закон модального управле-

ния для номинального ОУ на основе биномиальной модальной модели, в которой $\lambda_1 = \lambda_2 = -2$. Тогда $\mathbf{K} = [5,25 \quad 5,875]$, $\mathbf{F} = \begin{bmatrix} 0 & 1 \\ -4 & -4 \end{bmatrix}$, $\mathbf{K}_g = 6,25$, $\mathbf{G} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$.

Результаты моделирования, представленные кривыми ошибки воспроизведения задающего воздействия $g(t) = t$ при медианном и двух угловых реализациях параметров ОУ ($\underline{e}(t)$, $\overline{e}(t)$ – значения ошибки $e(t)$, соответствующие угловым значениям интервальной матрицы состояния), показывают (рисунок, а), что выход системы не обладает абсолютной параметрической инвариантностью, а параметрическая ε-инвариантность характеризуется заметной величиной ε.

Синтезируем закон модального управления, задавая значения λ_1: $\lambda_1 = -10$, $\lambda_1 = -50$, $\lambda_1 = -100$.

Результаты моделирования представлены на рисунке, б ($\lambda_1 = -10$), рисунке, в ($\lambda_1 = -50$) и рисунке, г ($\lambda_1 = -100$), каждый из которых получен при медианных и двух угловых реализациях параметров исходного ОУ. Ни в одном из случаев не достигается абсолютная параметрическая инвариантность, но наблюдается заметное уменьшение величины ε параметрической инвариантности по мере уменьшения величины $\hat{\varepsilon}$ невязки аппроксимации вектора $\mathbf{D}$ собственным вектором, формируемым по схеме Вандермонда (таблица).

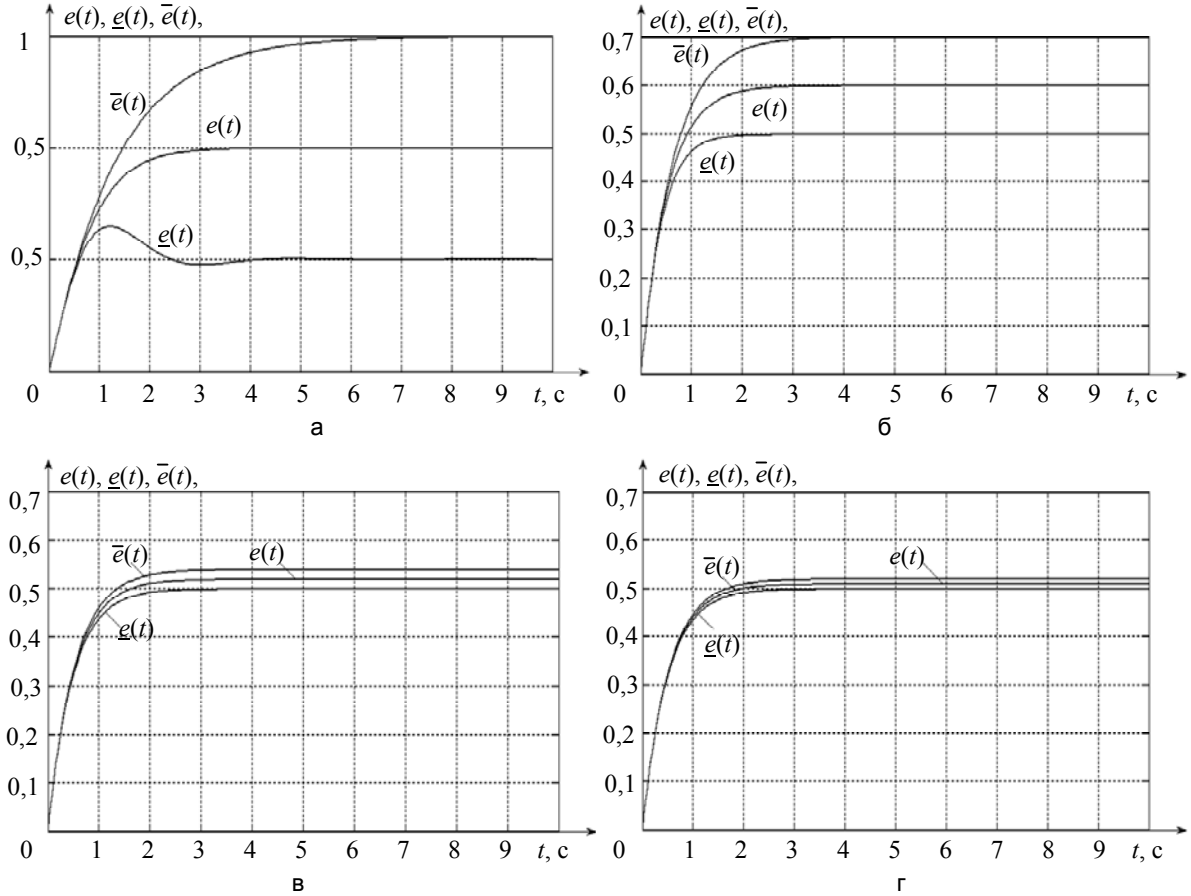


Рисунок. Ошибка по выходу для номинальной и угловых версий системы, спроектированной с помощью модального управления: $\lambda_1 = \lambda_2 = -2$ (а); $\lambda_1 = -10; \lambda_2 = -2$ (б); $\lambda_1 = -50; \lambda_2 = -2$ (в); $\lambda_1 = -100; \lambda_2 = -2$ (г)

Заключение

Показано, что для случая, когда матрица состояния исходного объекта задана в таком базисе, в котором недостижимо первое условие обеспечения абсолютной параметрической инвариантности, задача инвариантности выхода относительно неопределенности задания матрицы состояния объекта может быть решена в форме достижения параметрической ε-инвариантности с заданной ε.

Исследование выполнено при поддержке Министерства образования и науки Российской Федерации, соглашение 14. B37.21.0406 «Разработка многофункционального малогабаритного мультиротационного летательного аппарата».

	Матрица состояния спроектированной системы (медианная составляющая)	Собственный вектор	Величина $\hat{\varepsilon}$ невязки представления матрицы-столбца **D**	Матрица управления спроектированной системы	Величина ε параметрической инвариантности по ошибке
$\lambda_{1,2} = -2$	$\mathbf{F} = \begin{bmatrix} 0 & 1 \\ -4 & -4 \end{bmatrix}$	$\xi = \begin{bmatrix} -0,5 \\ 1 \end{bmatrix}$	$\hat{\varepsilon} = 50\%$	$\mathbf{G} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$	$\varepsilon = 50\%$
$\lambda_1 = -10$ $\lambda_2 = -2$	$\mathbf{F} = \begin{bmatrix} 0 & 1 \\ -20 & -12 \end{bmatrix}$	$\xi = \begin{bmatrix} -0,1 \\ 1 \end{bmatrix}$	$\hat{\varepsilon} = 10\%$	$\mathbf{G} = \begin{bmatrix} 0 \\ 20 \end{bmatrix}$	$\varepsilon = 16\%$
$\lambda_1 = -50$ $\lambda_2 = -2$	$\mathbf{F} = \begin{bmatrix} 0 & 1 \\ -100 & -52 \end{bmatrix}$	$\xi = \begin{bmatrix} -0,02 \\ 1 \end{bmatrix}$	$\hat{\varepsilon} = 5\%$	$\mathbf{G} = \begin{bmatrix} 0 \\ 120 \end{bmatrix}$	$\varepsilon = 3,85\%$
$\lambda_1 = -100$ $\lambda_2 = -2$	$\mathbf{F} = \begin{bmatrix} 0 & 1 \\ -200 & -102 \end{bmatrix}$	$\xi = \begin{bmatrix} -0,01 \\ 1 \end{bmatrix}$	$\hat{\varepsilon} = 1\%$	$\mathbf{G} = \begin{bmatrix} 0 \\ 200 \end{bmatrix}$	$\varepsilon = 2\%$

Таблица. Показатели параметрически ε-инвариантных систем

Литература

1. Буков В.Н., Бронников А.М. Условия инвариантности выхода линейных систем // Автоматика и телемеханика. – 2005. – № 2. – С. 23–35.
2. Никифоров В.О., Ушаков А.В. Управление в условиях неопределенности: чувствительность, адаптация, робастность. – СПб: СПб ГИТМО (ТУ), 2002. – 232 с.
3. Ackermann J. Robust control systems with uncertain physical parameters. – London: Springer-Verlag, 1993. – 406 p.
4. Слита О.В., Ушаков А.В. Обеспечение инвариантности выхода непрерывной системы относительно экзогенных сигнальных и эндогенных параметрических возмущений: алгебраический подход // Изв. РАН. Теория и системы управления. – 2008 – № 4. – С. 24–32.
5. Слита О.В., Ушаков А.В. Достаточные алгебраические условия параметрической инвариантности выхода линейной стационарной системы в первом приближении // Изв. РАН. Теория и системы управления. – 2010. – № 6. – С. 16–22.
6. Слита О.В., Ушаков А.В. Модальное управление: два способа реализации концепции подобия // Мехатроника, автоматизация, управление. – 2008. – № 9. – С. 7–13.
7. Дударенко Н.А., Полякова М.В., Ушаков А.В. Алгебраическая организация условий обобщенной синхронизируемости многоагрегатных динамических объектов // Научно-технический вестник СПбГУ ИТМО. – 2010. – № 2 (66). – С. 30–36.
8. Дударенко Н.А., Ушаков А.В. Структура собственных векторов матриц состояния многоканальных систем как вырождающий фактор // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – № 5 (81). – С. 52–58.
9. Слита О.В., Ушаков А.В. Модельное представление объекта управления в задаче параметрической инвариантности// Изв. вузов. Приборостроение. – 2006. – Т. 49. – № 1. – С. 14–20.
10. Гантмахер Ф.Р. Теория матриц. – М.: Наука, 1973. – 575 с.
11. Ковчин С.А., Сабинин Ю.А. Теория электропривода. – СПб: Энергоатомиздат, 1994. – 496 с.

Александрова Софья Александровна – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, студент, alexandrova_sophie@mail.ru

Мусаев Андрей Александрович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, студент, brein7@mail.ru

Слита Ольга Валерьевна – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат технических наук, доцент, o-slita@yandex.ru

Ушаков Анатолий Владимирович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, профессор, ushakov-AVG@yandex.ru

УДК 621.317.42

ЭЛЕКТРОМАГНИТНАЯ БЕЗОПАСНОСТЬ ЭЛЕКТРОТРАНСПОРТНЫХ СИСТЕМ: ОСНОВНЫЕ ИСТОЧНИКИ И ПАРАМЕТРЫ МАГНИТНЫХ ПОЛЕЙ[1]

Н.Г. Птицына, Ю.А. Копытенко, В.С. Исмагилов, А.Г. Коробейников

При эксплуатации транспортных средств, работающих на электротяге, генерируются магнитные поля, которые могут нарушить электромагнитную безопасность. Поэтому для ее обеспечения и электромагнитной совместимости необходимо знание характеристик и источников магнитных полей в электротранспорте. В работе проведено обобщение имеющейся информации о магнитных полях в электромобилях и сравнение их с результатами измерений, проведенными в других видах электрифицированного транспорта.

Ключевые слова: электромобиль, магнитное поле, электромагнитная безопасность, измерение магнитного поля.

Введение

Переход на электротранспорт является одной из наиболее актуальных мировых тенденций. Многие развитые высокотехнологичные страны имеют национальные программы по развитию экологичного электрического транспорта. Примером может служить программа Европейского Союза «Green Car Initiative», направленная на финансирование разработок в области создания электрического автотранспорта (с электрическими и с комбинированными, «гибридными» энергоустановками) и соответствующей инфраструктуры [1]. Общая сумма средств на данную программу достигает 1 млрд. евро. Россия также включилась во всеобщее движение по созданию автомобилей на электротяге.

Однако мировые производители электрического автомобилестроения столкнулись с серьезными проблемами обеспечения электромагнитной совместимости всех устройств на борту транспортной системы и электромагнитной безопасности пользователей электротранспорта [2–5]. В связи с этим тестирование, мониторинг и анализ магнитных полей (МП) в электротранспорте являются актуальной задачей.

Автомобили, работающие на электрической тяге, имеют силовые установки, датчики, устройства систем управления, информации и связи. Электрические токи, текущие через электродвигатель, цепи питания и батареи во время движения, генерируют МП в низкочастотных диапазонах (ультранизкочастотные (УНЧ), 0,001–10 Гц; крайне низкочастотные (КНЧ), 10–300 Гц). Более высокие гармоники электромагнитного поля в электромобиле генерируются разнообразными электронными устройствами на борту, информационными системами и системами связи. Например, в гибридных автомобилях наблюдаются магнитные импульсы до 5 кГц, которые генерируются во время переключения двигателя внутреннего сгорания и электрического режима. Кроме того, во всех типах автомобилей генерируется низкочастотное пульсирующее магнитное поле во время вращения стальных колесных дисков. Частота f этого поля определяется скоростью вращения колес, и обычно $f < 20$ Гц, но в спектре присутствуют также и гармоники с более высокой частотой [6].

Электромобиль является новой технологической системой, которая только сейчас выходит на широкий рынок. Производство и использование таких электромобилей пока не стало массовым, поэтому в мире и, в частности, в России отсутствуют достаточно обширные и детальные измерения МП в таких транспортных средствах. Работ о магнитных измерениях в гибридных автомобилях очень мало, а тестирование полей в полностью электрических автомобилях практически не проводится [7–9].

Поскольку в электромобилях, как и в других видах электротранспорта, МП генерируются токами, текущим по токовым системам (проводам и кабелям) транспортного средства, то можно полагать, что МП во всех транспортных системах, работающих на электрическом токе, будут иметь схожие параметры. Однако это предположение требует проверки. Целью настоящей работы является обобщение скудной информации о МП в электромобилях и сравнение их с результатами наших измерений, проведенными в других видах электрифицированного транспорта.

Магнитные поля в электромобиле

Vedholm и Hamerius [10] провели измерение магнитного поля (5–2000 Гц) в неподвижном гибридном электромобиле с включенным двигателем и кондиционером. Измерения проводились в 7 автомобилях в районе всех четырех сидений на уровне щиколоток, коленей, бедер, груди и головы. Так как автомобили были неподвижны, поля, генерированные вращением колес, отсутствовали. Магнитное поле, в котором находится водитель и пассажиры, было получено осреднением и представлено в таблице.

В гибридах, в которых батарея была размещена спереди (Авто 1–5), при отсутствии движения наблюдались маленькие поля. Более сильные поля наблюдались в автомобилях, в которых батарея была размещена сзади (Авто 6 и 7). В этих автомобилях батарея расположена под багажником или под задним

[1] Работа поддержана Министерством образования и науки РФ, государственный контракт № 11.519.11.1011.

сиденьем, и ток течет через весь автомобиль из передней части, от генератора до батареи. Такой большой токовый контур генерирует значительные МП. Максимальное поле 14 мкТл было отмечено у заднего правого сиденья (как правило, детского) на уровне ног.

Точка измерения	Авто 1	Авто 2	Авто 3	Авто 4	Авто 5	Авто 6	Авто 7
Левое переднее сиденье	0,12	0,11	0,15	0,22	0,14	2,6	3,2
Правое переднее сиденье	0,13	0,15	0,33	0,37	0,11	1,1	0,8
Левое заднее сиденье	0,06	0,04	0,03	0,03	0,06	2,4	4,0
Правое заднее сиденье	0,11	0,10	0,04	0,04	0,03	1,3	1,5

Таблица. Среднее магнитное поле (мкТл) в диапазоне частот (5–2000 Гц) в неподвижном гибридном электромобиле с включенным двигателем и кондиционером

Технический университет г. Биля по поручению Швейцарского федерального департамента здоровья провел измерения МП в двух гибридных автомобилях [3]. Измерения были сделаны в движении по городу и в лаборатории, где имитировались условия движения. Датчики МП помещались на переднее пассажирское сиденье, на пол у сиденья водителя, а также на место детского кресла (заднее правое сиденье). Колеса были сделаны из немагнитных материалов, чтобы исключить влияние МП, возникающего во время вращения ферромагнитных масс колес, на результаты измерений. Во время этого исследования установлено, что интенсивность МП постоянно менялась во время движения автомобиля и сильно зависела от способа ускорения и торможения. Наибольшая интенсивность наблюдается во время ускорения и торможения.

Во время движения МП на детском сиденье было в пределах 0,1–3 мкТл. МП в других измеренных точках имело примерно такой же уровень. Было получено, что в гибридных автомобилях генерируется смесь МП в частотном диапазоне 5–500 Гц. Надо заметить, что в данном исследовании измерительные приборы не фиксировали МП ниже 5 Гц.

В работе австралийских исследователей [8] приведены результаты измерений МП в разных точках гибридного автомобиля. Было найдено, что интенсивность поля находится в пределах 0–35 мГс (0–3,5 мкТл). Максимальные поля выше 10 мГс наблюдались на частоте 12 Гц. На рис. 1 показано МП, измеренное в задней части гибридного автомобиля во время движения в разном режиме (ускорение, торможение).

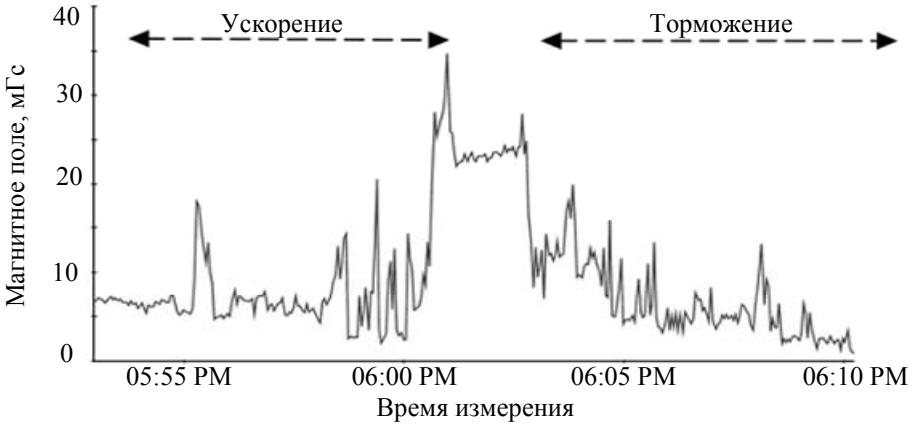


Рис. 1. Магнитные поля на левом заднем сидении гибридного автомобиля.

Группа исследователей из США исследовала МП в различных электромобилях производства США, а также на заправочной станции во время зарядки авто [11]. Эти автомобили использовали двигатели как постоянного, так и переменного тока. Измерения показали, как и в предыдущих случаях, что наибольшие поля генерировались во время максимального ускорения и рекуперативного торможения. Наиболее сильное поле было обнаружено на местах водителей. Во время зарядки электромобиля самые большие поля, равные 64,37 мГс, фиксировались на расстоянии 50 см от панели автоматического прерывателя тока. Результаты измерений установили, что в электромобилях генерируются низкочастотные МП в диапазоне частот 60–420 Гц. Авторы этого исследования заключают, что такие МП могут угрожать здоровью водителей и пассажиров, и поэтому их следует экранировать.

Достаточно детальное исследование МП в гибридном автомобиле Крайслер, ориентированное на экранирование и уменьшение этих полей, было проведено в совместном проекте армейской группы Army TACOM и Крайслер К° [12]. Поля измерялись индукционным магнитометром в широком диапазоне частот (0–50 кГц). В исследуемом автомобиле батарея находилась над задними колесами. При силе тока в 200 А максимальное поле ~1200 мГс (120 мкТл) наблюдалось в районе заднего сиденья, а пространственные градиенты доходили до 1000 мГс/м. Эти значения лежат в тех же пределах, что и в электропоездах.

Магнитные поля в электропоездах, метро, трамвае и троллейбусе

Авторы данной работы провели многочисленные измерения МП в различных видах электрического транспорта, как то: электролокомотивы и электрички, работающие как на постоянном, так и на переменном токе (16,34 Гц), метро, трамвай и троллейбус, работающие на постоянном токе. Результаты измерений и анализа частично представлены в [13–15]. Мониторинг полей, как правило, сопровождался фиксацией информации об условиях движения (ускорение, замедление, торможение, прохождение подстанций, встречных поездов или других ферромагнитных масс и т.д.), что позволило идентифицировать источники различных магнитных вариаций наблюдаемых в электротранспорте. Примеры такой идентификации показаны на рис. 2–4.

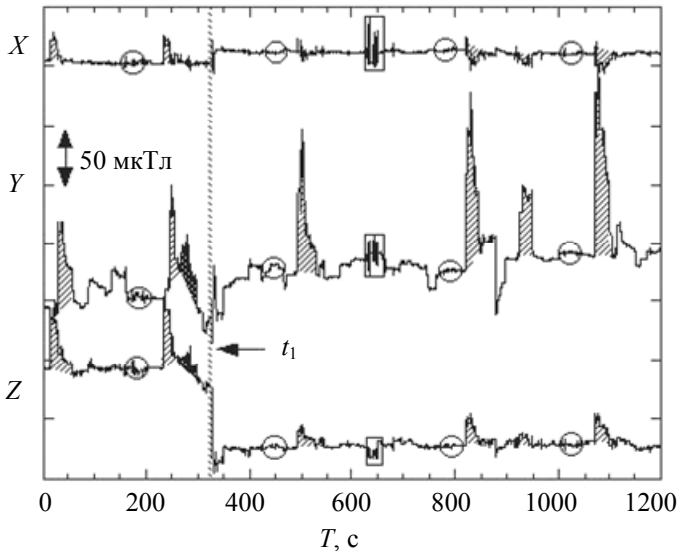


Рис. 2. УНЧ магнитное поле (компоненты *X*, *Y*, *Z*) в электричке, работающей на постоянном токе, в вагоне с моторами. Заштрихованная область – фаза ускорения, черная область – подстанция, прямоугольник – встречный поезд, круг и эллипс – фаза торможения. В момент t_1 датчик был помещен в точку непосредственно над электродвигателем. Резкое изменение поля в этот момент свидетельствует о больших пространственных градиентах

На рис. 2 представлена запись *X*-, *Y*- и *Z*-компонент МП, измеренного в пассажирском вагоне, т.е. в вагоне с электродвигателями. Вариации МП в направлении вдоль рельсов (*X*-компонента) пренебрежимо малы по сравнению с вариациями *Y*- и *Z*-компонент, перпендикулярных к рельсам. Наибольшие скачки поля происходят во время ускорения поезда, величина скачков в *Y*-компоненте на этом участке пути достигает 150 мкТл. Некоторое увеличение поля происходит также во время торможения. Встречные или стоящие вдоль пути поезда вызывают заметные вариации поля из-за влияния их ферромагнитных масс.

На рис. 3 приведена *Y*-компонента МП, измеренного в электровозе вблизи рабочего места машиниста. Видно, что максимальные значения вариаций здесь достигают 100 мкТл. Вариации, наблюдаемые в *X*- и *Z*-компонентах, имеют тот же вид, что и в *Y*-компоненте, но с гораздо меньшей амплитудой. Здесь при проведении измерений нулевые уровни датчиков соответствовали полям, измеренным на платформе до отправления поезда (т.е. обычному фону магнитной окружающей среды).

Данные, приведенные на рис. 3, позволяют оценить уровень постоянного МП внутри поезда. Видно, что величина поля внутри поезда обычно гораздо выше, чем на платформе, т.е. выше уровня магнитного фона окружающей среды. Однако иногда, например, при прохождении первой силовой подстанции значение поля становится ниже уровня окружающей среды (на рис. 3 для времени $T \approx 150$–200 с значение поля отрицательно).

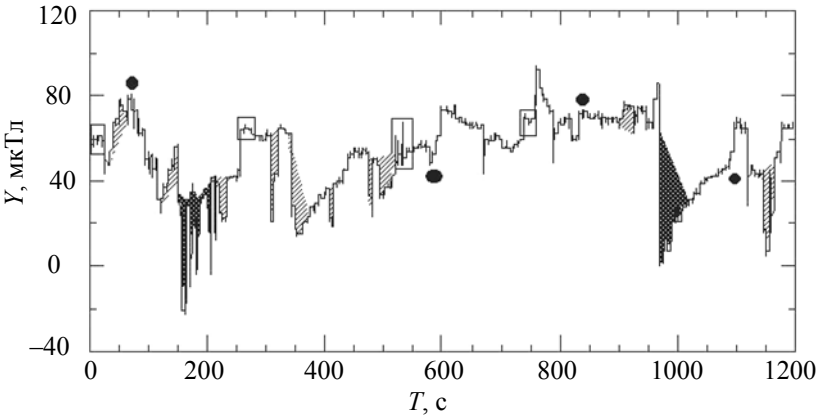


Рис. 3. *Y*-компонента УНЧ магнитного поля локомотива, работающего на постоянном токе, вблизи рабочего места машиниста (15 см от его головы). • – моменты нулевого тока. Остальные обозначения те же, что на рис. 2

Типичный пример записи МП в локомотивах, работающих на переменном токе (16,67 Гц) приведен на рис. 4. Измерения были выполнены на линии Цюрих-Берн. На рис. 4, а, показаны вариации трех компонент, зарегистрированные в течение 20 мин с частотой регистрации 200 Гц. Более детальная 20-секундная запись вариаций приведена на рис. 4, б. Этот рисунок показывает «движущуюся картину», получающуюся в результате непрерывных измерений МП. Оказалось, что для швейцарских электровозов, как и для российских поездов, можно связать различные особенности МП с рабочими режимами в различные моменты движения, как показано на рис. 4. Анализ показал, что на рабочем месте машиниста основным источником вариаций был ток в рельсовой цепи, который менялся в соответствии с обстоятельствами движения.

Рис. 2–4 демонстрируют, что основными источниками вариаций МП в электропоездах являются электрические токи в токовых цепях. Эти токи и, соответственно, МП уменьшаются / увеличиваются при смене режима движения (ускорение, торможение, прохождение подстанции). Кроме того, источником вариаций МП внутри является меняющаяся внешняя ферромагнитная обстановка (встречные поезда, мосты и т.д.).

Были проведены также измерения МП в трамвае ЛВК-86К [16]. Исследовалась зависимость амплитудных значений МП от режима эксплуатации вагона (стоянка, разгон, движение, торможение). Уровень УНЧ полей, возникающих в результате изменений тока в цепи согласно нуждам движения, менялся в пределах от нуля до нескольких сотен мкТл. Установлено, что в кабине водителя трамвая уровни индукции постоянного МП составили 3,4–98 мкТл, в салоне для пассажиров – 14–500 мкТл. В отдельных местах, вблизи отверстий в перегородке, отделяющей кабину водителя трамвая от салона для пассажиров, значения индукции достигали 1000 мкТл (вследствие возникающих там явлений дифракции и интерференции), что в 15 раз превышает постоянное поле Земли.

Источниками постоянных и переменных МП в трамвае являются токонесущие элементы, электрооборудование, печи электрообогрева. Высокие уровни МП регистрировались вблизи реостатного блока. Наибольшие значения индукции МП в салоне обнаружены в зоне расположения сидений пассажиров в начале вагона слева по линии расположения токоведущих элементов, где находятся места для пассажиров с детьми и инвалидов. В вагоне трамвая регистрировались значительные градиенты МП, являющиеся существенным биотропным параметром.

По той же методике [15, 16] были проведены измерения в двух типах метропоездов г. Санкт-Петербурга (вагон 81717 и автоматизированный вагон ЕМА). Как выяснилось, источниками МП в электропоездах метрополитена являются тяговые двигатели, пускотормозные сопротивления, групповые переключатели, кабели токоприемника, токонесущий провод. Значительные УНЧ вариации МП наблюдались в кабинах машинистов петербургского метро в горизонтальной Y-компоненте (перпендикулярно рельсам), где значения «от пика до пика» достигали 425 мкТл в кабине машиниста вагона 81717 и 350 мкТл в вагоне ЕМА. Таким образом, в этих местах МП в 7–14 раз превосходит величину естественного постоянного поля Земли (около 50 мкТл). В пассажирских салонах поле варьировало в пределах 75–300 мкТл. Анализ показал, что МП в поездах метрополитена и на платформах неоднородно по пространству.

Измерения МП в троллейбусе были проведены в Москве и Санкт-Петербурге. Исследовалась зависимость амплитудных значений МП от точки измерения и от режима эксплуатации троллейбуса (стоянка, разгон, движение, торможение). Точки измерений были выбраны вблизи источников наибольших колебаний индукции МП: в кабине водителя, в салоне на переднем пассажирском сидении, над тяговым электродвигателем, над компрессором и на задней площадке на трех высотах от пола: 0,5 м; 1,0 м; 1,4 или

1,7 м [17]. В каждой точке осуществлялось по 200 измерений в секунду каждой из трех компонент индукции МП в течение 3 мин. Результаты мониторинга записывались вместе с данными о режиме движения троллейбуса по маршруту.

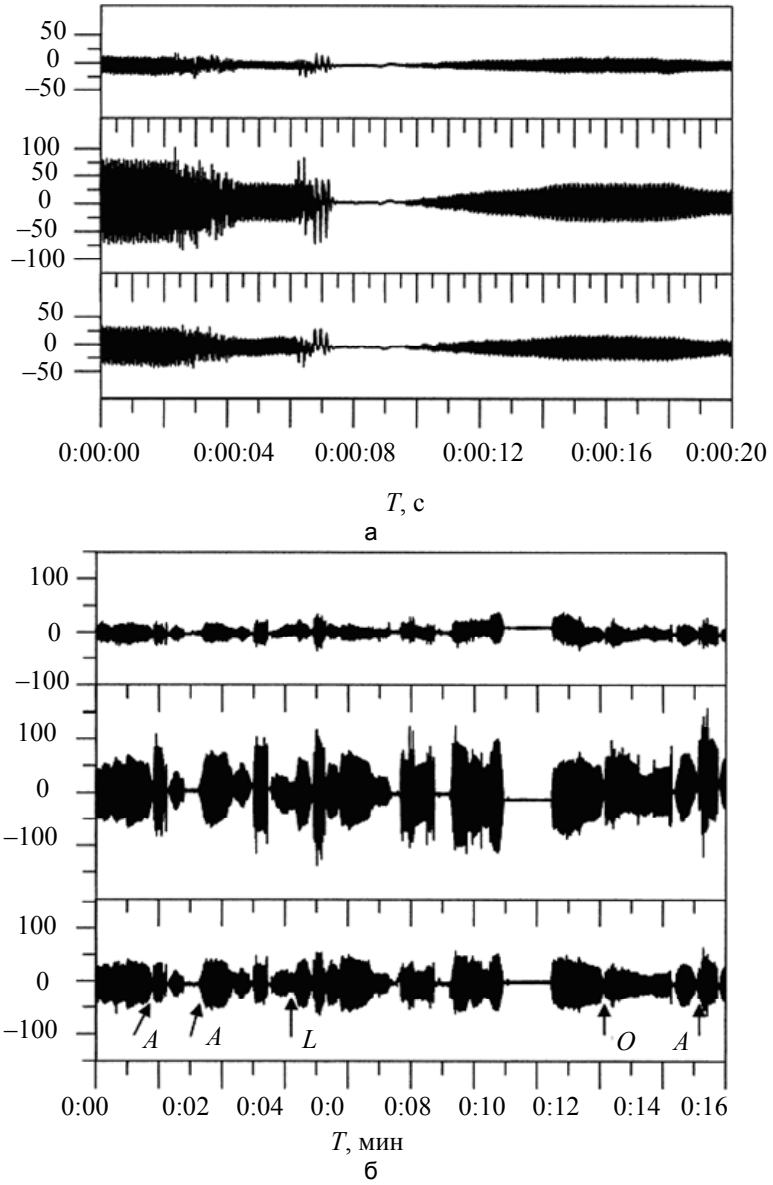


Рис. 4. Магнитное поле в швейцарском электролокомотиве, работающем на переменном токе (16,67 Гц): а – 20-минутная; б – 20-секундная запись. *A* – ускорение; *L* – низкая скорость; *O* – нулевой ток

В водительской кабине троллейбуса зафиксированы самые большие значения $H_{ср}$ = 93±45 мкТл за головой водителя в 20 см от проводов, по которым текут рабочие токи до 200 А от токоприемников к контакторной панели и тяговому электродвигателю. В этой точке имеется самый большой размах колебаний: $\Delta H = H_{макс} - H_{мин}$ = 20–36 = 168 мкТл. Наибольший размах зарегистрирован в горизонтальных компонентах: ΔX = 141 мкТл, ΔY = 172 мкТл, тогда как вертикальная компонента имеет меньший, но тоже значительный размах: ΔZ = 50 мкТл. Результаты измерений свидетельствуют, что МП в кабине троллейбуса не только резко меняется со временем, но и имеет пространственный градиент порядка 100 мкТл/м.

Аналогичное исследование было проведено в троллейбусах марки ЗИУ-9 и ЗИУ-10 в Санкт-Петербурге. Статистический анализ МП в кабине и пассажирском салоне петербургского троллейбуса показал, что в целом характер поля и его отклик на условия движения по маршруту имеют такой же характер, что и в московском троллейбусе. Основная мощность спектра поля в петербургском троллейбусе также лежит в области УНЧ частот (менее 10–15 Гц). Размах колебаний МП в петербургском троллейбусе, как и в Москве, зависит от режима движения и значительно увеличивается при движении троллейбуса в пробках при частых разгонах и торможениях.

Заключение

Сравнение измерений магнитного поля, проведенного в различных видах транспорта, показало, что эти поля кардинально отличаются от синусоидальных полей, которые генерируют линии передач (50 Гц или 60 Гц). Магнитные поля в электрическом транспорте, включая автомобиль, являются мультичастотными полями, которые быстро меняются во времени и пространстве автомобиля. Эти черты магнитного поля являются следствием суммирования многообразных источников магнитного поля на борту транспортного средства и меняющихся режимов движения (ускорение, торможение и т.д.). Однако большая часть магнитной энергии концентрируется в наиболее низких частотных диапазонах (квазистатическом и ультранизком, 0,001–10 Гц).

Ниже приведены максимальные уровни полей, встречающихся в различных видах электротранспорта, включая электромобили, в квазистатическом и ультранизкочастотном диапазоне [2, 4, 12, 15]:

- трамвай 500 мкТл;
- метро 450 мкТл;
- троллейбус 350 мкТл;
- электромобили/гибридные 140 мкТл;
- электропоезд 120 мкТл;
- электрокар 104 мкТл;
- легкий электробус 80 мкТл.

Видно, что наибольшие уровни полей встречаются в трамвае и метро. Измеренные к настоящему времени уровни полей в электрическом автомобиле сравнимы по порядку величин с магнитными полями, измеренными в электропоездах – в вагонах электричек и на рабочем месте машиниста электровоза. Эти магнитные поля значительно превосходят поля от линий передач, с которыми человек обычно сталкивается дома и на работе.

С учетом того факта, что водители и пассажиры электромобилей находятся в непосредственной близости от источников тока, очевидно, что для обеспечения их электромагнитной безопасности необходимо вести поиски разумных и недорогих способов уменьшения магнитных полей. Обычно инженерно-техническая защита строится на мероприятиях по ограничению эмиссионных параметров источников поля либо на базе экранирования поля. Это могут быть многослойные экраны, изготовленные из современных материалов на основе сплавов с аморфной и нанокристаллической структурой, которые не утяжеляют конструкцию авто и могут уменьшать низкочастотные магнитные поля, характерные для электротранспорта, в 5–10 раз [15].

Литература

1. European Green Car Initiative, 2008 [Электронный ресурс]. – Режим доступа: http://www.green-cars-initiative.eu/public/, свободный. Яз. англ. (дата обращения 05.01.2013).
2. Dietrich F.M., Jacobs W.L. Survey and assessment of electric and magnetic field (EMF) public exposure in the transportation environment. – US Department of Transportation, Federal Railroad Administration, Report № PB99-130908, 1999. – 52 p.
3. Swiss Federal Office of Public Health, Cars, hybrid cars. – Technical Report, 2009 [Электронный ресурс]. – Режим доступа: http://www.bag.admin.ch/themen/strahlung/ 00053/00673/02377/index.html?lang =en, свободный. Яз. англ. (дата обращения 05.01.2013).
4. Muc A.M. Electromagnetic Fields Associated with Transportation Systems, Radiation Health and Safety Consulting, Contract Report 4500016448, Air Health Effects Division, Healthy Environments and Consumer Safety Branch Health, Canada, 2002. – 56 p.
5. zur Nieden F., Scheier S., Frei S. Can Cross-Talk on PCBs Cause Permanent ESD-Damages // Proceedings of the 22th International Conference on Electromagnetic Compatiability. EMC Europe 2012. – Rome, Italy, 17–21 September. – 2012. – 73 p.
6. Milham S., Hatfield J., Richard T. Magnetic fields from steel-belted radial tyres: implications for epidemiologic studies // Bioelectromagnetics. – 1999. – V. 20. – № 7. – P. 440–445.
7. Electromagnetic Health Organization, EMF test of 2007 Toyota Prius Hybrid, Electromagn. Health, 2008 [Электронный ресурс]. – Режим доступа: http://electromagnetichealth.org/wp-content/uploads/2008/11/emf_test_of_2007_toyota_prius_hybrid-3.pdf свободный. Яз. англ. (дата обращения 05.01.2013).
8. Halgamuge M.N., Abeyrathne C.D., Mendis P. Measurements performed in electric trains – Comparison with ICNIRP Limit & Laboratory Experiments. Measurement and Analysis of Electromagnetic Fields from Trams, Trains and Hybrid Cars // Radiation Protection Dosimetry. – 2010. – V. 141. – № 3. – P. 255–268.
9. Птицына Н.Г., Исмагилов В.С., Копытенко Ю.А., Коробейников А.Г. Фазово-градиентный метод измерения магнитного поля в электромобиле // Научно-технический вестник информационных технологий, механики и оптики. – 2013. – № 1 (83). – С. 108–112.

10. Vedholm K., Hamnerius Y.K. Personal Exposure Resulting from Low Level Low Frequency Electromagnetic Fields in Automobiles // Second World Congress for Electricity and Magnetism in Medicine and Biology, June 8-13, Bologna, Italy, 1997. – Abstract F-9. – 445 p.
11. Berisha Shahin H., Karady George G., Raed Ahmad, Ray Hobbs, Donald Karner. Magnetic Field Generated From Different Electric Vehicles // Electric Transportation Application, DOI: 951934, 1995. – P. 1–5.
12. Snyder M. Magnetic Shielding for Electric Vechicles. Program Review. Contract DAAE07-93-C-R107. Army TACOM, Chrysler Corp. 1995. – 56 p.
13. Копытенко Ю.А., Виллорези Дж., Птицына Н.Г., Копытенко Е.А., Исмагилов В.С., Зайцев Д.Б., Воронов П.М., Тясто М.И., Юччи Н., Пфлюгер Д. Измерение магнитных полей электрического транспорта прибором, фиксирующим форму волны: российские и швейцарские поезда, метро, трамвай и троллейбус // Материалы второй международной конференции «Проблемы электромагнитной безопасности человека. Фундаментальные и прикладные исследования. Нормирование ЭМП: философия, критерии и гармонизация». – 20–24 сентября. – М., 1999. – С. 90.
14. Ptitsyna N.G., Kopytenko Y.A., Villoresi G., Pfluger D.H., Ismaguilov V., Iucci N., Kopytenko E.A., Zaitzev D.B., Voronov P.M., Tyasto M.I. Waveform Magnetic Field Survey in Russian DC- and Swiss AC-powered Trains: a Basis for Biologically Relevant Exposure Assessment // Bioelectromagnetics. – 2003. – V. 24. – № 8. – P. 546–556.
15. Птицына Н.Г., Виллорези Дж., Копытенко Ю.А. Тясто М.И. Магнитные поля на электротранспорте и экология человека. – СПб: Изд-во «Нестор-История», 2010. – 120 с.
16. Никитина В.Н., Ляшко Г.Г., Копытенко Ю.А. и др. Гигиеническая оценка электромагнитных полей в электропоездах и технологических зонах метрополитена // Медицина труда и промышленная экология. – 2002. – № 3. – С. 16–18.
17. Авенариус И.А., Лелюхин А.М., Копытенко Ю.А., Птицына Н.Г. Исследование низкочастотного магнитного поля в кабине и пассажирском салоне троллейбуса // Вестник МАДИ (ГТУ). – 2007. – № 3 (10). – С. 107–114.

Птицына Наталья Григорьевна – Санкт-Петербургский филиал Института земного магнетизма, ионосферы и распространения радиоволн им. Н.В. Пушкова РАН, кандидат физ.-мат. наук, ст. научный сотрудник, Nataliaptitsyna@yandex.ru

Копытенко Юрий Александрович – Санкт-Петербургский филиал Института земного магнетизма, ионосферы и распространения радиоволн им. Н.В. Пушкова РАН, доктор физ.-мат. наук, профессор, директор, office@izmiran.spb.ru

Исмагилов Валерий Сарварович – Санкт-Петербургский филиал Института земного магнетизма, ионосферы и распространения радиоволн им. Н.В. Пушкова РАН, кандидат физ.-мат. наук, ст. научный сотрудник, ученый секретарь, IVS@izmiran.spb.ru

Коробейников Анатолий Григорьевич – Санкт-Петербургский филиал Института земного магнетизма, ионосферы и распространения радиоволн им. Н.В. Пушкова РАН, доктор технических наук, профессор, зам. директора, Korobeynikov_A_G@mail.ru

УДК 621.396; 621.372

МЕТОД МАКСИМАЛЬНО ИНФОРМАТИВНОЙ ЗОНЫ ДЛЯ ПОСТРОЕНИЯ ВИРТУАЛЬНОЙ БАЗОВОЙ СТАНЦИИ В КИНЕМАТИЧЕСКИХ СХЕМАХ ГЕОДЕЗИЧЕСКИХ GPS-СЕТЕЙ

Р.А. Эминов, Х.Г. Асадов

Проанализирован существующий фактический материал по экспериментальной оценке погрешности позиционирования в VRS GPS-сетях, где мобильный приемник обеспечивается виртуальной базовой станцией. Предложен метод высокоинформативной зоны для устранения исходной неопределенности выбора базовых станций в целях построения минимальной GPS-сети, состоящей из трех базовых станций. Даны методические указания и рекомендации по использованию предложенного метода.

Ключевые слова: GPS-приемник, виртуальная базовая станция, кинематические схемы, позиционирование, геодезическая сеть.

Введение

Концепция сетевых кинематических систем GPS и GNSS возникла в середине 90-х г.г. и бурно развивается в настоящее время [1]. Обычные однобазовые кинематические системы реального времени (RTK) имели ограниченное предельное расстояние, равное 10–20 км. В настоящее время сетевые RTK (NRTK) системы, обладая высокой точностью, позволяют осуществить позиционирование при расстоянии между станциями 70–100 км.

Основными систематическими погрешностями, влияющими на точность мобильного приемника (МП) RTK-систем, являются многолучевость, атмосферные и эфемеридные погрешности [2]. Одним из перспективных путей дальнейшего повышения точности RTK-систем позиционирования можно считать метод создания виртуальной базовой станции (VRS), представляющей собой модель базовой станции (БС), находящейся в непосредственной близости МП. Использование метода VRS позволяет уменьшить погрешности позиционирования, возникающие из-за ошибок многолучевости, ионосферы и тропосферной задержки, а также эфемерид.

Как указывается в работе [3], концепция VRS позволяет осуществлять RTK-позиционирование с точностью 2 см в сетях БС с расстоянием между БС до 32 км.

Согласно [3], основными принципами создания VRS являются следующие:

- данные от сети базовых станций передаются в вычислительный центр;
- данные сети используются для вычисления моделей ионосферных, тропосферных и орбитальных погрешностей;
- неоднозначность оценки фазы несущего сигнала фиксируется с учетом базовых расстояний сети;
- фактические ошибки базовых расстояний определяются с сантиметровой точностью на основе измерения фиксированной фазы несущего сигнала;
- для предсказания суммарных погрешностей позиционирования мобильного приемника пользователя используется линейная комбинация моделей погрешностей;
- VRS создается в непосредственной близости от мобильного приемника;
- данные VRS передаются пользователю в стандартных форматах (RTCM).

Как будет показано далее, до сих пор открытым остается вопрос об обоснованном выборе места VRS в RTK-сети, а также вопрос о выборе БС для позиционирования мобильного приемника МП пользователя. До сих пор используются такие эвристические критерии, как «максимально близкие к мобильному приемнику БС» и «максимально близкое размещение VRS к мобильному приемнику».

Далее авторами рассматриваются вопросы замены эвристических критериев на информационные критерии.

Критический обзор существующих методов

Как сообщается в работе [4], в NRTK-системах расстояние между БС и мобильным приемником определяется в виде максимального радиуса круга, охватывающего зону, где сигнал соответствующей БС может быть использован эффективно совместно с посылками формата Международной радиотехнической комиссии морской службы (RTCM) (рис. 1).

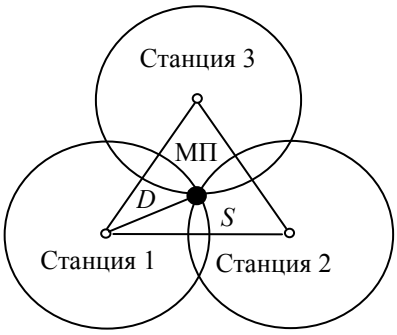


Рис. 1. Идеально симметричная конфигурация расположения МП и базовых станций [4]

В идеально симметричном случае имеет место следующее соотношение:

$D = 0,5774S$,

где D – расстояние между БС и МП пользователя; S – расстояние между БС. Согласно [4], погрешность позиционирования RTK GPS-сетей является линейно-возрастающей функцией D. Как сообщается в работе [5], в системе ROMPOS (Румыния) типа NRTK-системы высокая точность позиционирования достигается следующим образом. Сначала МП передает свои приблизительные данные о своей позиции в центр управления через Интернет. Центр управления на базе принятой информации выбирает БС, наиболее близкую к МП, в качестве основной. Далее первичная информация, полученная с этой БС, корректируется в виде геометрической позиции с учетом сетевой коррекции, соответствующей приблизительной позиции МП.

В настоящее время существуют такие конкурирующие технологии, как NRTK с VRS, NRTK с псевдо-базовой станцией (PRS), метод параметров коррекции по участку (FKR), а также концепция «хозяин–помощник» (MAC) и индивидуализированные коррекции такого же типа (i–Max) [6].

Согласно работе [7], оператор сети VRS должен выполнить следующие шаги:

1. определить атмосферные и орбитальные погрешности с сантиметровой точностью путем решения проблемы неоднозначности базовых расстояний в пределах сети;
2. моделировать позицию VRS, используя данные БС, расположенной наиболее близко к МП;
3. интерполировать погрешность сети в зоне расположения VRS, используя линейные или более сложные модели;
4. передать данные коррекции МП в реальном времени.

В отличие от концепции VRS NRTK, в методе MAC все данные коррекции передаются от сети к мобильному источнику в компактной форме с четким выделением различий коррекции для дисперсивных и недисперсивных полей применительно к каждой паре спутник–приемник. При этом мобильный приемник MAC GPS имеет возможность использовать более сложные интерполяционные алгоритмы для определения своей позиции, в то время как МП VRS вынужден «довериться» результатам вычисления его позиции в сервере сети.

Как сообщается в работе [8], для того чтобы использовать быстрое статическое или, в реальном времени, кинематическое (RTK) позиционирование с сантиметровой точностью на больших дистанциях, в некоторых частях Германии были развернуты сети опорных GPS-станций с шагом 30–50 км. В пределах этих сетей были смоделированы значения погрешностей, зависящих от расстояния, таких как погрешности из-за ионосферной и тропосферной рефракции и др., раздельно по спутникам с высоким временным разрешением. Такая процедура приводит к концепции VRS, выходная информация с которых фактически вырабатывается на компьютере, используя для этого данные реально существующих БС.

Как указывается в работе [8], одним из методов использования полного информационного содержания одновременно проводимых наблюдений на нескольких опорных станциях является их комбинирование таким способом, чтобы на VRS было сформировано оптимальное множество кодовых и фазовых измерений, пригодных для определения позиций мобильного приемника.

Согласно работе [9], для того чтобы результаты измерений фазы несущей частоты, полученные в БС сети, были трансформированы в смоделированный результат измерения фазы несущей частоты VRS, должен быть осуществлен ряд операций, к которым относятся следующие:

1. исключение неоднозначности, возникшей из-за двойного дифференцирования фазы несущей частоты;
2. вычисление моделей погрешностей, зависящих от расстояния:
   - ионосферная модель, основанная на линейной ионосферной комбинации;
   - геометрическая модель, содержащая тропосферные и орбитальные погрешности и основанная на линейной комбинации.

При этом обработка данных может быть осуществлена на центральном вычислителе, или может быть распределена между центральным вычислителем и пользователями.

Как указывается в работе [10], в сетевых RTK-системах позиционирования наиболее критической компонентой погрешности позиционирования является дифференциальная ионосферная остаточная погрешность между сетью опорных станций и мобильным приемником.

Согласно работе [11], концентрация электронов в ионосфере подвержена сильному суточному изменению. При этом, как сообщается в работе [12], после двойного дифференцирования тропосферные погрешности при межбазовых расстояниях 60–80 км могут превосходить ионосферные погрешности и достигать величины 2 см за 30-секундный временной интервал. Вместе с тем, вопрос выбора места создания VSR в мировой практике решается до сих пор по следующим эвристическим соображениям:

1. VSR должен быть расположен максимально близко к мобильному приемнику пользователя;
2. для позиционирования мобильного приемника должны быть выбраны те БС, которые максимально близки к мобильному приемнику.

Однако логика развития техники и технологии такова, что эвристические подходы к решению тех или других задач с появлением соответствующей теоретической базы должны быть заменены на научно обоснованные подходы к решению задач. Предпосылками для создания научной базы являются следующие хорошо известные положения техники GPS RTK-позиционирования.

1. Погрешность позиционирования базовых станций, интерполируемая в сети, имеет тренд увеличения с ростом расстояния от БС до мобильного приемника.
2. Погрешность позиционирования БС имеет как шумоподобные дисперсивные ионосферные составляющие, так и систематические (тропосферные, орбитальные и др.) составляющие [11, 12].
3. Доминирование шумоподобной ионосферной погрешности и ее дисперсивность (пространственная неоднородность) позволяют считать, что информация о погрешности позиционирования БС передается к пользователю с уменьшающимся по расстоянию отношением сигнал/шум.

Вышеуказанные положения позволяют ставить и решать задачу оптимального размещения VSR в пределах RTK GPS, с целью повышения информативности системы позиционирования.

Метод максимально информативной зоны для построения VSR

Как указывается в работе [1], при интерполировании погрешности VSR в пределах программного обеспечения сервера погрешности интерпретируются на основе данных окружающих БС с использованием метода интерполяции. При этом возможно также использовать метод экстраполяции (рис. 2).

При осуществлении интерполяции следует учесть свойство линейного увеличения погрешности от расстояния. При этом основными задачами сервера сети являются:

1. генерация данных для определения виртуальной позиции МП;
2. генерация RTCM-данных для виртуальной позиции МП.
3. передача RTCM-данных к МП.

Как нам представляется, интерполяционный метод определения погрешности координат МП в методическом плане исключает возможность осуществления высокоточного позиционирования, так как предполагает постоянное существование погрешности используемого метода – погрешности интерполирования. При этом в данном методе ничего не говорится о том, как следует поступать с погрешностью интерполяции и как извлечь из нее дополнительную информацию о позиции МП.

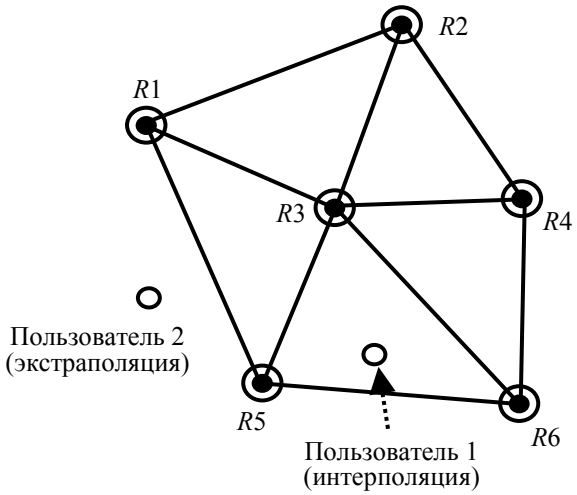


Рис. 2. Случаи необходимости осуществления интерполяции (Пользователь 1) и экстраполяции (Пользователь 2)

С учетом вышесказанного нами предлагается информационно-интерполяционный метод определения погрешности VRS-позиционирования, названный методом высокоинформативной зоны, суть которого заключается в следующем:

1. VRS располагается над МП, т.е. их места позиционирования совпадают;
2. осуществляется интерполирование погрешности и геометрически определяется интерполяционная оценка погрешности позиции VRS (МП);
3. определяется максимально информативная зона в поле расположения VRS (МП), в которой можно было бы принять максимальное количество информации при реализации определенного порядка передачи информации от БС о своей позиции с учетом заданных ограничительных условий;
4. для каждой позиции VRS (МП) в целях интерполирования выбирается такая тройка БС, максимальная информационная зона которых определяет точку размещения VRS (МП).

Получим основные базовые соотношения для предлагаемого метода. Допустим, что имеются шесть БС и МП с неточными координатами (рис. 3). Исходно выбираем станции БС1, БС3 и БС5. С учетом зашумленности сигнала каналов быстротечной составляющей ионосферной погрешности для передачи информации от БС к МП используется многократная посылка данных о собственных позициях.

Таким образом, переходя на непрерывную форму записи, количество информации, принятой в VRS, оценим как

$$F_1 = \int_0^{T_{\max}} \mathrm{M}_1 dT = \int_0^{T_{\max}} T \log_2 \left[\psi(l) + 1 \right] dT , \qquad (1)$$

где T – длительность передачи информации от БС к VRS с максимальным значением, равным $T_{\max}$; ψ – отношение сигнал/шум в переданной информации; l – расстояние между БС и VRS. $M_1 = T \log_2 \left[\psi(l) + 1 \right]$.

В первом приближении имеем

$$\psi(l) = \psi_0 + \psi' \cdot 1 , \qquad (2)$$

где ψ_0 – отношение сигнал/шум у базовых станций; $\psi' = \dfrac{d\psi}{dl}$.

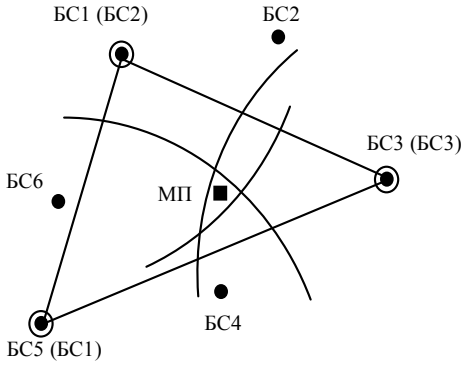


Рис. 3. Графическое пояснение предлагаемого метода высокоинформативной зоны: МП – мобильный приемник; БС1–БС6 – базовые станции исходной сети; (БС1)–(БС3) – базовые станции, используемые в треугольной сети

Введем на рассмотрение функцию связи между расстоянием l и длительностью T:

$$l = \varphi(T), \tag{3}$$

С учетом выражений (1)–(3) получаем

$$F_1 = \int_0^{T_{\max}} M_1 dT = \int_0^{T_{\max}} T \log_2\left[\psi_0 + \psi' \cdot \varphi(T) + 1\right] dT. \tag{4}$$

Введем ограничение на суммарную величину расстояний от МП до БС,

$$F_2 = \int_0^{T_{\max}} M_2 dT = \int_0^{T_{\max}} \varphi(T) dT = C_1, \tag{5}$$

где $C_1 = \text{const}$; $M_2 = \varphi(T)$.

Отметим, что в случае монотонного вида функции (3) выражение (5) можно интерпретировать как ограничение на суммарную длительность приема информации.

С учетом выражений (4) и (5) можно составить уравнение безусловной вариационной оптимизации:

$$F_3 = \int_0^{T_{\max}} T \log_2\left[\psi_0 + \psi' \cdot \varphi(T) + 1\right] dT + \lambda \int_0^{T_{\max}} \varphi(T) dT, \tag{6}$$

где λ – множитель Лагранжа.

Известно, что оптимальная функция $\varphi_{opt}(T)$ удовлетворяет условию Эйлера

$$\frac{d\left(M_1 + \lambda M_2\right)}{d\,\varphi(T)} = 0. \tag{7}$$

С учетом выражений (6) и (7) получаем

$$\frac{T \cdot \psi'}{(\ln 2)\left[\psi_0 + \psi' \cdot \varphi(T) + 1\right]} + \lambda = 0. \tag{8}$$

Из выражения (8) находим

$$\varphi(T) = -\left(\frac{\psi_0}{\psi'} + \frac{1}{\psi'} + \frac{T}{\lambda \cdot \ln 2}\right). \tag{9}$$

С учетом выражений (5) и (9) получим

$$-\int_0^{T_{\max}} \left(\frac{\psi_0}{\psi'} + \frac{1}{\psi'} + \frac{\mathrm{T}}{\lambda \cdot \ln 2}\right) dT = C_1, \tag{10}$$

откуда нетрудно вычислить

$$\lambda = -\frac{T_{\max}}{2 \cdot \ln 2\left[\dfrac{\psi_0}{\psi'} + \dfrac{1}{\psi'} + \dfrac{C_1}{T_m}\right]}. \tag{11}$$

С учетом выражений (8) и (11) получим

$$l = \varphi(T) = \frac{2T\left[\psi_0 + 1 + C_1\psi'\right]}{T_{\max} \cdot \psi'} - \frac{(\psi_0 + 1)}{\psi'}. \quad (12)$$

С учетом отрицательной величины ψ' примем $\psi' = -|\psi'|$. В этом случае уравнение (12) имеет следующий вид:

$$l = \varphi(T) = \frac{\psi_0 + 1}{|\psi'|} - \frac{2T\left[\psi_0 + 1 - C_1|\psi'|\right]}{T_{\max}|\psi'|}. \quad (13)$$

Таким образом, при известных значениях $\psi_0, |\psi'|, T_{\max}$, задаваясь линейной шкалой величин T, по выражению (13) можно вычислить значение l.

Далее по составленному множеству значений $\{T_i\}$, $i = \overline{1,3}$, в котором значения элементов множества подчиняются линейному закону изменения по индексу i, вычисляются соответствующие элементы l_i множества $\{l_i\}$, $i = \overline{1,3}$. После этого по вычисленным значениям l_i, $i = \overline{1,3}$ осуществляется геометрическое построение искомой высокоинформативной зоны, которая может содержать или не содержать МП.

Рассмотрим эти два случая раздельно.

1. Высокоинформативная зона содержит в себе позицию МП (рис. 4). В этом случае контуры построения высокоинформативной зоны, обозначенные как 1, 2, 3, уплотняются в сторону минимизации площади высокоинформативной зоны с целью уменьшения соответствующих времен $\{T_i\}$, $i = \overline{1,3}$, приема информации.
2. Если построенная высокоинформативная зона не содержит в себя МП, то возможны два варианта действий:
   1. VRS строится в высокоинформативной зоне, если не ожидается значительное удаление МП;
   2. осуществляется выбор тройки БС, обеспечивающей такое построение высокоинформативной зоны, в котором будет содержаться МП.

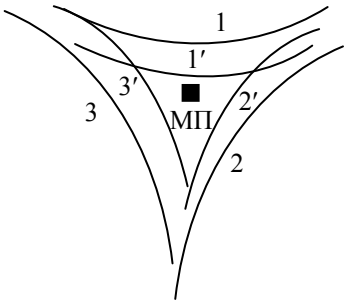


Рис. 4. Порядок уплотнения высокоинформативной зоны

Заключение

Таким образом, введенное в настоящей работе понятие максимальной информативной зоны позволяет устранить исходную неопределенность выбора минимальной сети БС для определения погрешности позиции мобильного приемника методом интерполяции.

Следует отметить, что в данном случае двухкритериальная концепция определения погрешности позиции мобильного приемника (определение по максимальной интерполяционной точности и по максимальному количеству информации) не подчиняется классической концепции многокритериальной оптимизации, так как здесь отсутствует антагонизм критериев, и выполнение требований по одному из критериев в пределе означает автоматическое выполнение требований другого критерия.

Основные выводы и положения проведенного исследования можно сформулировать следующим образом.

1. Систематизирован существующий фактический материал по экспериментальной оценке погрешности позиционирования в VRS GPS-сетях, где мобильный приемник обеспечивается виртуальной БС.
2. Предложен метод высокоинформативной зоны для устранения исходной неопределенности выбора БС для построения минимальной GPS-сети и дальнейшего позиционирования мобильного приемника пользователя.
3. Даны методические указания и рекомендации по использованию предложенного метода.

Литература

1. Landau H., Vollath U., Chen X. Virtual reference station systems // Journal of Global Positioning Systems. – 2002. – V. 1. – № 2. – P. 137–143.
2. Vollath U., Landau H., Chen X., Doucet K., Pagels Ch. Network RTK versus single base RTK understanding the error characteristics // Proceedings of ION GPS 2002. – 24–27 September. – Portland, OR, 2002. – P. 2774–2781.
3. Vollath U., Deking A., Landau H., Pagels Ch. Long-range RTK positioning using virtual reference stations. – P. 470–474 [Электронный ресурс]. – Режим доступа: http//www.ucalgary.ca/engo_webdocs/SpecialPublications/Kis%2001/PDF/0802/PDF, свободный. Яз. англ. (дата обращения 30.10.2012).
4. Feng Y., Wang J. GPS RTK performance characteristics and analysis // Journal of Global Positioning Systems. – 2008. – V. 7. – № 1. – P. 137–143.
5. Badesku G., Stefan O., Badesku R., Ortelecan M., Veres S.I. Positioning system GPS and RTK VRS type, using the Internet as a base a network of multiple stations // FIG Working Week 2011, Bridging the Gap between Cultures Marrakesh, Morocco. – 18–22 May. – 2011. – P. 1–11.
6. Takac F., Zelzer O. The relationship between network RTK solutions MAC, VRS, PRS, FKP and i-MAX // Proceedings of ION GNSS 2008. – Savannah, GA. – P. 348–355.
7. Janssen V. A comparison of the VRS and MAC principles for network RTK // International Global Navigation Satellite Systems Society. – IGNSS Symposiun 2009. Holiday Inn Surfers Paradise, Old, Australia, 1–3 December, 2009. – P. 1–15 [Электронный ресурс]. – Режим доступа: http//www.ecite.utas.edu.au/60284/1/2009_Janssen_IGNSS2009_proceedings_version.pdf, свободный. Яз. англ. (дата обращения 25.10.2012).
8. Wanninger L. The performance of virtual reference stations in active Geodetic GPS-networks under solar maximum conditions // Proc. of ION GPS'99. – Nashville TN, 1999. – P. 1419–1427.
9. Wanninger L. Improved ambiguity resolution by regional differential modeling of the ionosphere // Proc. of ION GPS'95. – 1995. – P. 55–62.
10. Landau H., Vollath U., Chen X. Virtual reference stations versus broadcast solutions in network RTK – advantages and limitations, GNSS, Graz, Austia, April, 2003. – P. 1–15 [Электронный ресурс]. – Режим доступа: http://www.gootechmabna.com/catalouges/Papers/3sencore/15.pdf, свободный. Яз. англ. (дата обращения 21.11.2012).
11. Wu S., Zhang K., Silcock D. Magnitudes and temporal variations of the tropospheric and ionospheric errors in GPSnet // Journal of Global Positioning Systems. – 2010. – V. 9. – № 1. – P. 61–67.
12. Wu S., Zhang K., Yuan Y., Wu F. Spatio-temporal characteristics of the ionospheric TEC variation for GPSnet-based real-time positioning in Victoria // Journal of Global Positioning Systems. – 2006. – V. 5. – № 1–2. – P. 52–57.

Эминов Рамиз Ахмет оглы – Азербайджанская государственная нефтяная академия (г. Баку), кандидат технических наук, доцент, eminovramiz@mail.ru

Асадов Хикмет Гамид оглы – НИИ аэрокосмической информатики (г. Баку), доктор технических наук, доцент, начальник отдела, asadzade@rambler.ru

УДК 681.532.65

АЛГОРИТМЫ КОМПЕНСАЦИИ ПУЛЬСАЦИЙ МОМЕНТА ПРЕЦИЗИОННОГО ЭЛЕКТРОПРИВОДА НА БАЗЕ СИНХРОННОЙ МАШИНЫ С ПОСТОЯННЫМИ МАГНИТАМИ

В.С. Томасов, С.Ю. Ловлин, А.В. Егоров

Рассмотрена модель электропривода, учитывающая нелинейности синхронной машины с постоянными магнитами: зубцовый момент и момент высших гармоник потокосцепления ротора. Предложены алгоритмы идентификации параметров синхронной машины с постоянными магнитами и компенсации влияния пульсаций зубцового момента и момента гармоник. Данные алгоритмы увеличивают эффективность прецизионного электропривода за счет повышения точности позиционирования. Изложенные в работе решения имеют экспериментальное подтверждение.

Ключевые слова: прецизионный электропривод, синхронная машина с постоянными магнитами, пульсации момента, зубцовый момент, момент гармоник.

Введение

К числу актуальных проблем современной электромеханики и силовой электроники относится проектирование систем управления прецизионными сервоприводами для оптико-механических систем и, в частности, для высокоточных комплексов позиционирования и слежения [1–6]. Проектирование серво-

привода систем наведения и слежения таких комплексов требует решения широкого круга задач, связанных с необходимостью обеспечения высокоточного позиционирования, вращения исполнительного органа с инфранизкими угловыми скоростями при значительных величинах момента сопротивления на валу [1]. Высокое качество наведения должно обеспечиваться при вращающихся массах от нескольких десятков килограммов до нескольких десятков тонн при наличии возмущающих неравномерных моментов от сил вязкого и сухого трения, в том числе в подшипниковых узлах, кабельном переходе, переменных (зубцовых) моментов электродвигателя, ветровых и динамических нагрузок, включая момент дисбаланса оптической оси. Обеспечение таких точностей возможно лишь при использовании безредукторных следящих электроприводов осей на основе вентильных двигателей [1–6].

В прецизионных электроприводах чаще всего используются синхронные машины с возбуждением от постоянных магнитов (СМПМ). СМПМ характеризуются большим отношением вращающего момента к моменту инерции ротора, определяющим предельное быстродействие машины, а также низкими потерями в меди за счет отсутствия отдельных обмоток возбуждения и токов намагничивания. Высокая эффективность СМПМ позволяет использовать полностью закрытые конструкции с естественным охлаждением. Использование редкоземельных постоянных магнитов позволяет обеспечить высокую плотность магнитного потока в воздушном зазоре, что упрощает конструкцию двигателей высокой удельной мощности.

Одной из проблем приводов на базе СМПМ является неравномерность вращающего момента. Момент двигателя меняется периодически в зависимости от положения ротора во время его вращения. Результирующий мгновенный момент СМПМ состоит из двух составляющих: постоянного, или полезного, момента и пульсирующего момента. Существуют два источника пульсаций момента. Первый источник – момент, вызванный несинусоидальным распределением плотности магнитного потока в воздушном зазоре (момент гармоник). Второй – зубцовый момент, появляющийся в результате неравномерного распределения магнитной проницаемости статора.

Один из известных способов минимизации пульсаций момента электрической машины – системы управления с таблицами поиска (lookup table) [7–9]. Таблица поиска – это структура данных, обычно массив или ассоциативный массив, используемая с целью замены процедуры вычисления на операцию простого поиска [9]. Таблицы поиска для компенсации пульсаций момента составляются по результатам эксперимента или серии экспериментов. В процессе эксперимента фиксируются амплитудные значения тока в зависимости от угла поворота ротора при движении с постоянной скоростью. В этом случае таблица поиска будет содержать информацию о необходимых значениях тока для компенсации всех возмущений. Но некоторые из возмущений зависят от направления движения, а также от величин и знаков токов фаз СМПМ, или же со временем могут изменяться. Все это нельзя учесть при составлении таблицы поиска (в общем случае моменты трения при движении в положительную сторону и в отрицательную сторону не равны по модулю при движении), и поэтому она может содержать меняющиеся во времени составляющие момента возмущения. При изменении этих составляющих эффективность таблицы может снизиться вплоть до появления дополнительных возмущений.

Пульсации момента гармоник и зубцового момента являются свойством СМПМ и не меняются со временем. Рассмотренный ниже алгоритм рассчитан только на компенсацию этих пульсаций, что является его преимуществом по сравнению с таблицей поиска по причине вышеуказанных проблем.

Пульсации вращающего момента

Зубцовые пульсации вращающего момента (зубцовый момент) возникают в результате взаимодействия между постоянными магнитами, установленными на роторе, и зубцами статора. Существует множество методов для уменьшения зубцового момента, таких как скос пазов и (или) магнитов; изменение формы магнитов; введение вспомогательных пазов или зубцов; оптимизация соотношения полюсной дуги к полюсному шагу магнитов; использование дробного соотношения числа пазов к числу полюсов; и т.д. [10, 11]. Несмотря на то, что для определения результирующего зубцового момента сейчас широко используется метод конечных элементов, аналитические методы остаются полезным средством для быстрого расчета его формы [11, 12].

Один из таких методов – метод суперпозиции, используемый для получения результирующей формы зубцового момента на базе рассчитанного зубцового момента для одного зубца статора [13]. Этот момент может быть получен либо методом конечных элементов, либо путем разложения в ряд Фурье:

$$M_{sc} = \sum_{i=1,2,3...}^{\infty} M_{sc.i} \sin(2p \cdot i \cdot \alpha),$$

где M_{sc} – зубцовый момент от одного паза статора; $M_{sc.i}$ – амплитуда i-ой гармоники; $2p$ – число полюсов; α – геометрический угол поворота ротора.

Результирующий зубцовый момент для пазов статора может быть получен аналитически, учитывая угловой сдвиг пазов и применяя метод суперпозиции [13]:

$$M_{cog} = \sum_{n=1,2,3...}^{\infty} M_{sc.i} N_s \cdot \sin(N_c n\alpha) . \quad (1)$$

где N_c – наименьшее общее кратное $2p$ и N_s. Анализируя формулу (1), можно сделать вывод, что порядок основной гармоники зубцового момента равен наименьшему общему кратному N_c числа полюсов ротора и пазов статора. Модель, полученная аналитическим методом, позволяет анализировать форму и гармонический состав зубцового момента.

При сбалансированных синусоидальных токах в распределенных трехфазных обмотках статора формируется синусоидальная магнитодвижущая сила (МДС) в воздушном зазоре. Синусоидальное распределение потока ротора достигается за счет изменения формы магнитов и регулирования их направления намагничивания. На практике идеальное синусоидальное распределение недостижимо, и противо-ЭДС содержит несколько высших гармоник, что приводит к устойчивым пульсациям момента:

$$e_a(t) = -c_e\Omega(t)\left(\sin(p\alpha(t)) + \sum_{i=1,2,3...}^{\infty} K_i \sin((2i+1)p\alpha(t))\right),$$

$$e_b(t) = -c_e\Omega(t)\left(\sin\left(p\alpha(t) + \frac{2\pi}{3}\right) + \sum_{i=1,2,3...}^{\infty} K_i \sin\left((2i+1)p\alpha(t) + \frac{2\pi}{3}\right)\right),$$

$$e_c(t) = -c_e\Omega(t)\left(\sin\left(p\alpha(t) - \frac{2\pi}{3}\right) + \sum_{i=1,2,3...}^{\infty} K_i \sin\left((2i+1)p\alpha(t) - \frac{2\pi}{3}\right)\right),$$

где c_e – конструктивная постоянная противо-ЭДС; p – количество пар полюсов; K_i – амплитуда i-ой гармоники противо-ЭДС относительно первой. При синусоидальных токах в обмотках статора,

$$i_a(t) = I_m \sin(p\alpha(t) + \psi),\ i_b(t) = I_m \sin\left(p\alpha(t) + \psi + \frac{2\pi}{3}\right),$$
$$i_c(t) = I_m \sin\left(p\alpha(t) + \psi - \frac{2\pi}{3}\right), \quad \textbf{(2)}$$

где ψ – угол сдвига фаз, I_m – амплитуда фазных токов, основной момент принимает вид [5]

$$M_{ME1}(t) = \frac{3}{2} c_e I_m \cos(\psi) . \quad (3)$$

Момент, обусловленный взаимодействием высших гармоник магнитного поля ротора с токами в фазах статора, равен

$$M_{MEh}(t) = \frac{3}{2} c_e I_m \sum_{i=1,2,3...}^{\infty} K_i \cos(2pi\alpha - \psi) . \quad (4)$$

Следует отметить, что гармонический состав момента гармоник включает в себя и гармоники зубцового момента. Это связано с тем, что высшие гармоники МДС вызывают и гармоники ЭДС (соответственно момент гармоник M_{MEh}), и зубцовый момент M_{cog}. Кроме того, момент гармоник имеет магнитоэлектрическую природу и зависит от тока в обмотках статора, в то время как зубцовый момент возникает за счет изменения энергии постоянных магнитов [14]. По этой причине алгоритмы компенсации их влияния должны быть разными, и идентифицировать оба момента необходимо раздельно, что усложняется однообразием их гармонического состава.

Синтез алгоритма компенсации влияния нелинейностей электрической машины

Модель момента гармоник $M_{MEh}(\alpha(t))$ (4) соответствует синусоидальным токам статора (2), поэтому для синтеза алгоритма компенсации на базе моделей (4) и (1) необходимо замкнуть контур тока. Контур тока настраивается на апериодическое звено первого порядка с постоянной времени T_T [15]. В результате электромагнитный момент двигателя принимает вид

$$M(t) = \frac{3}{2} c_e I_m \cos(\psi) + \frac{3}{2} c_e I_m \sum_{i=1,2,3...}^{\infty} K_i \cos(2pi\alpha - \psi) + \sum_{n=1,2,3...}^{\infty} M_{sc.i} N_s \cdot \sin(N_c n\alpha) . \quad (5)$$

Тогда добавка к заданию амплитуды тока I_m, благодаря которой формула (5) преобразовывается в (3), имеет следующий вид:

$$I_m^* = \frac{-\frac{2}{3} c_e^{-1} \sum_{n=1,2,3...}^{\infty} M_{sc.i} N_s \cdot \sin(N_c n\alpha) - I_m \sum_{i=1,2,3...}^{\infty} K_i \cos(2pi\alpha - \psi)}{\cos(\psi) + \sum_{i=1,2,3...}^{\infty} K_i \cos(2pi\alpha - \psi)} . \quad (6)$$

Уравнение (6) не учитывает динамические особенности контура тока, поэтому всегда будет присутствовать ошибка компенсации пульсаций зубцового момента и момента гармоник. Компенсация, учитывающая динамику контура тока, имеет вид

$$I_m^* = \left(-\frac{2}{3} c_e^{-1} \sum_{n=1,2,3...}^{\infty} M_{sc.i} N_s \cdot (\sin(N_c n\alpha) + T_T N_c n\Omega(t) \sin(N_c n\alpha)) - \right.$$

$$\left. -I_m \sum_{i=1,2,3...}^{\infty} K_i (\cos(2pi\alpha - \psi) - 2pi\Omega(t) T_T \sin(2pi\alpha - \psi)) \right) \times \qquad \textbf{(7)}$$

$$\times \left(\cos(\psi) + \sum_{i=1,2,3...}^{\infty} K_i (\cos(2pi\alpha - \psi) - 2pi\Omega(t) T_T \sin(2pi\alpha - \psi)) \right)^{-1}.$$

Таким образом, при известных параметрах $M_{sc.i}$, N_s, K_i, p, c_e компенсация (7) позволяет минимизировать пульсации электромагнитного момента, не учитывая влияние изменения тока в фазе. В случае, когда динамическими особенностями контура тока можно пренебречь, стоит пользоваться компенсацией (6), так как для ее вычисления не используется скорость вращения привода, и сама формула проще, т.е. займет меньше времени при вычислении на микроконтроллере.

Идентификация параметров моделей момента гармоник и зубцового момента

Расчет параметров модели момента гармоник (4) и зубцового момента (1) требует точного моделирования электромагнитного поля в воздушном зазоре электрической машины при таких известных параметрах машины, как скос пазов статора, высота воздушного зазора, расположение обмоток статора, высота магнитов и т.п. [16–19]. Такие данные очень редко сообщаются разработчиками электрических машин, так как составляют интеллектуальную собственность. Кроме того, расчет электромагнитных процессов в воздушном зазоре электрической машины – весьма трудоемкая задача, в отличие от определения параметров моделей (4) и (1) экспериментальным путем. Идентификация параметров моделей момента гармоник (4) и зубцового момента (1) осуществляется при постоянной скорости, так как динамический момент привода равен нулю, и основной момент принимает вид

$$M_{ME1}(t) = -M_{MEh}(\alpha(t)) - M_{cog}(\alpha(t)) + M_{\text{тр}}(\alpha(t)) + M_{\text{каб}}(\alpha(t)) + M_{\text{дисб}}(\alpha(t)),$$

где $M_{\text{тр}}$ – момент трения; $M_{\text{каб}}$ – момент кабельного перехода (который при постоянной скорости вращения можно считать линейным); $M_{\text{дисб}}$ – момент дисбаланса оптической оси.

Измеряемый угол по датчику положения ротора $\alpha_{\text{ДП}}(t)$ отличается от электрического угла φ:

$$\theta = p\alpha_{\text{ДП}}(t) - p\alpha(t),$$

где θ – угол расхождения нуля датчика положения ротора с нулем электрического угла. Тогда (2) принимает вид

$$i_a(t) = I_m \sin(p\alpha_{\text{ДП}}(t) - \theta + \theta^*),\ i_b(t) = I_m \sin\left(p\alpha_{\text{ДП}}(t) - \theta + \theta^* + \frac{2\pi}{3} \right),$$

$$i_c(t) = I_m \sin\left(p\alpha_{\text{ДП}}(t) - \theta + \theta^* - \frac{2\pi}{3} \right),$$

где θ^* – задаваемый параметр, позволяющий регулировать угол между магнитным потоком ротора и магнитным потоком статора.

При вращении в положительную и отрицательную сторону с постоянной скоростью получают зависимости $I_{m.1}^+(\alpha_{\text{ДП}})$ и $I_{m.1}^-(\alpha_{\text{ДП}})$ для θ_1, $I_{m.2}^+(\alpha_{\text{ДП}})$ и $I_{m.2}^-(\alpha_{\text{ДП}})$ для θ_2. Этим данным соответствует следующее уравнение:

$$(I_{m.1}^+(\alpha_{\text{ДП}}) - I_{m.1}^-(\alpha_{\text{ДП}}))\cos(\theta^* - \theta_1) - (I_{m.2}^+(\alpha_{\text{ДП}}) - I_{m.2}^-(\alpha_{\text{ДП}}))\cos(\theta^* - \theta_2) =$$

$$= -(I_{m.1}^+(\alpha_{\text{ДП}}) - I_{m.1}^-(\alpha_{\text{ДП}})) \times \sum_{i=1,2,3...}^{\infty} K_i \cos(2pi\alpha_{\text{ДП}} - \theta^* + \theta_2) + \qquad (8)$$

$$+(I_{m.2}^+(\alpha_{\text{ДП}}) - I_{m.2}^-(\alpha_{\text{ДП}})) \times \sum_{i=1,2,3...}^{\infty} K_i \cos(2pi\alpha_{\text{ДП}} - \theta^* + \theta_2).$$

Параметры модели момента гармоник K_i определяются из уравнения (8) методом наименьших квадратов (МНК). Затем для определения параметров зубцового момента $M_{sc.i}$ составляется следующее уравнение:

$$\frac{3}{2}c_e(I_m^+(\alpha_{ДП})+I_m^-(\alpha_{ДП}))\cos(\theta^*-\theta)=-\frac{3}{2}c_e(I_m^+(\alpha_{ДП})+I_m^-(\alpha_{ДП}))\times$$

$$\times\sum_{i=1,2,3...}^{\infty}K_i\cos(2pi\alpha_{ДП}-\theta^*+\theta)-2\sum_{n=1,2,3...}^{\infty}M_{sc.i}N_s\sin(N_c n\alpha_{ДП}-\theta^*)+f(\alpha_{ДП}),$$

причем

$$f(\alpha)=(M_{тр}^+(\alpha_{ДП})-M_{тр}^-(\alpha_{ДП}))+2M_{каб}(\alpha_{ДП})+2M_{дисб}(\alpha_{ДП}).$$

При принятых допущениях относительно линейности $M_{каб}$ и при несовпадении периода $M_{дисб}$ с периодом гармоник M_{cog}, параметры $M_{sc.i}$ достаточно точно определяются с помощью МНК.

Эксперимент

Оценка эффективности алгоритма компенсации проводилась путем сравнения двух трехконтурных систем управления с контурами тока, скорости и положения, настроенных с учетом и без учета нелинейностей исполнительной электрической машины. Для проведения эксперимента использован уникальный стенд, спроектированный и изготовленный в ОАО НПК «СПП» для проведения экспериментальных научно-исследовательских работ в НИУ ИТМО, полностью имитирующий поведение азимутальной оси опорно-поворотного устройства. Стенд состоит из двухмассового механизма и позволяет изменять коэффициент жесткости между первой и второй массами, а также момент инерции второй массы. Для измерения угла поворота оси стенд оснащен прецизионным датчиком фирмы «Renishaw». На стенде установлен трехфазный синхронный двигатель типа RSM-P-36-275*25 BS производства фирмы ООО «Рухсервомотор». Согласно паспортным данным, СМПМ имеет следующие параметры: коэффициент противо-ЭДС $c_e=6,2$ В·с/рад; активное сопротивление фазы $R=1,2$ Ом; электромагнитная постоянная времени $T_e=8$ мс; количество пар полюсов $p=24$; количество пазов $N_s=36$. Напряжение в звене постоянного тока – 48 В.

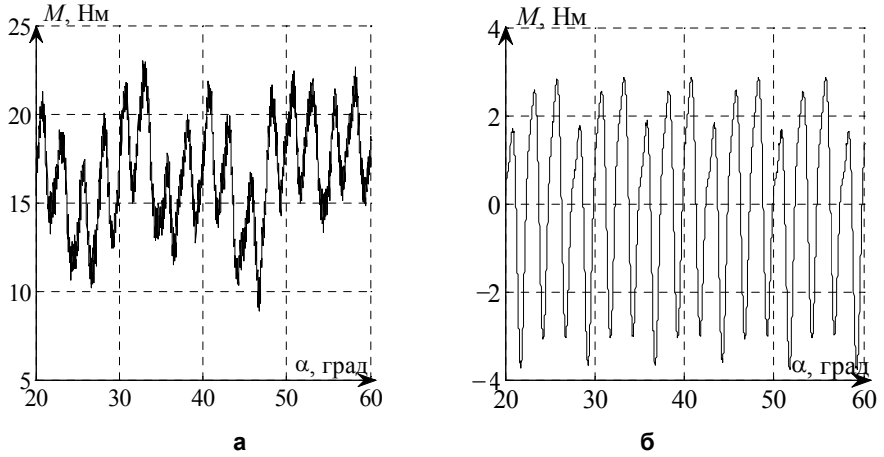


Рис. 1. График основного момента при движении с постоянной скоростью (а) и график результатов идентификации зубцового момента и момента гармоник СМПМ (б)

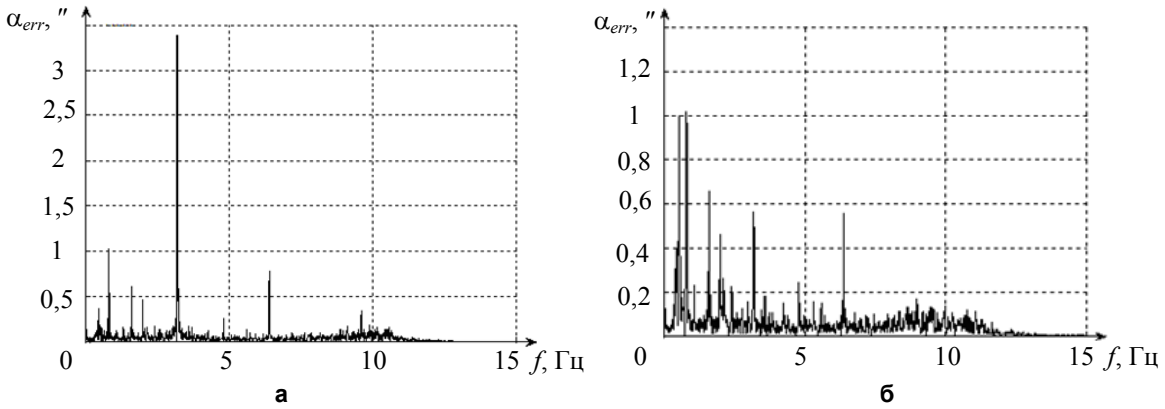


Рис. 2. Спектральный состав ошибки трехконтурной системы управления, настроенной без (а) и с учетом (б) влияния нелинейностей исполнительной электрической машины

Идентификация параметров моделей момента гармоник и зубцового момента проводилась согласно методике, описанной выше. На вход трехконтурной системы управления подавалось линейно-возрастающее задающее воздействие, соответствующее скорости вращения $1\,\text{град/с}$. При этом среднеквадратичное отклонение (СКО) ошибки по углу составляло $1{,}3''$. Результаты идентификации изображены на рис. 1.

Спектральный состав ошибки α_{err} при отработке линейно изменяющегося сигнала задания со скоростью $8\,\text{град/с}$ системами управления с компенсацией и без компенсации влияния зубцового момента и момента гармоник изображен на рис. 2.

В таблице приведены значения СКО ошибок слежения при отработке системами задающих сигналов, изменяющихся с постоянной скоростью. Очевидно, что использование алгоритма компенсации позволяет существенно повысить точность слежения прецизионного электропривода.

Скорость слежения, град/с	СКО трехконтурной системы управления, угл. секунды	СКО трехконтурной системы управления с алгоритмом компенсации влияния нелинейностей СМПМ, угл. секунды
1	1,5	0,7
8	5,6	2,9

Таблица. Результаты применения алгоритма компенсации на испытательном стенде

Заключение

Предложена модель электропривода, учитывающая нелинейности синхронной машины с постоянными магнитами. Разработаны алгоритмы идентификации параметров моделей зубцового момента и момента гармоник и компенсации влияния нелинейностей синхронной машины с постоянными магнитами. Разработанные алгоритмы позволили повысить точность сопровождения космических аппаратов при движении оптической оси с постоянной скоростью. Рассмотренные в работе алгоритмы идентификации и компенсации были использованы при разработке цифровых сервоприводов для ОПУ-799, ОПУ-834 проектов «Прицел» и «Моренос», выполненных по заказам ОАО НПК «СПП».

Литература

1. Васильев В.Н., Томасов В.С., Шаргородский В.Д., Садовников М.А. Состояние и перспективы развития прецизионных электроприводов комплексов высокоточных наблюдений // Изв. вузов. Приборостроение. – 2008. – № 6. – С. 5–11.
2. Томасов В.С., Денисов К.М., Толмачев В.А. Следящие электроприводы систем наведения оптико-механических комплексов нового поколения. Проблемы и достижения // Тр. V междунар. (XVI Всеросс.) конференции по автоматизированному электроприводу. – 2007. – С. 175–177.
3. Томасов В.С., Овчинников И.Е., Егоров А.В. Энергоподсистема большого алтайского телескопа траекторных измерений // Известия тульского государственного университета. – Тула: Изд-во ТулГУ, 2010. – Вып. 3. – Ч. 3. – С. 216–222.
4. Балковой А.П., Цаценкин В.К. Прецизионный электропривод с вентильными двигателями – М.: Издательский дом МЭИ, 2010. – 328 с.
5. Овчинников И.Е. Вентильные электрические двигатели и привод на их основе (малая и средняя мощность) // Курс лекций – СПб: Корона-Принт, 2010. – 336 с.
6. Сабинин Ю.А. Позиционные и следящие электромеханические системы: Учебное пособие для вузов. – СПб: Энергоатомиздат, 2001. – 208 с.
7. Campbell-Kelly Martin, Croarken Mary, Robson Eleanor. The History of Mathematical Tables From Sumer to Spreadsheets. – 1-st ed. – New York, USA: Oxford University Press. – 2003. – 372 p.
8. Iqbal Husain. Minimization of Torque Ripple in SRM Drives // IEEE Transactions on industrial electronics. – 2002. – V. 49. – № 1. – P. 28–39.
9. Wu A.P., Chapman P.L. Cancellation of Torque Ripple Due to Nonlinearities of Permanent Magnet Synchronous Machine Drives // Proc. IEEE Power Electronics Specialist Conf. – 2003. – P. 256–261.
10. Jahns T.M., Soong W.L. Pulsating torque minimization techniques for permanent magnet ac motor drives: a review // IEEE Trans. Power Electronics. – 1996. – V. 43. – P. 321–330.
11. Zhu Z.Q., Howe D. Influence of design parameters on cogging torque in permanent magnet machines // IEEE Trans. on Energy Conversion. – 2000. – V.15. – № 4. – P. 407–412.
12. Goto M., Kobayashi K. An analysis of the cogging torque of a DC motor and a new technique of reducing the cogging torque // Electrical Engineering in Japan. – 1983. – V. 103. – № 5. – P. 113–120.
13. Zhu Z.Q., Ruangsinchaiwanich S., Howe D. Synthesis of Cogging Torque Waveform from Analysis of a Single Stator Slot // IEEE International Conference on Electric Machines and Drives. – 2005. – P. 125–130.

14. Holtz J., Springob L. Identification and compensation of torque ripple in high-precision permanent magnet motor drive // IEEE Transactions on Industrial Electronics. – 1996. – V. 43. – № 2. – P. 309–320.
15. Ловлин С.Ю., Тушев С.А. Информационная подсистема цифрового электросилового привода с компенсацией пульсаций момента вентильного двигателя // Сборник тезисов докладов конференции молодых ученых. – СПб: СПбГУ ИТМО, 2011. – Вып. 2. – С. 250–251.
16. Jingqiu Qiao, William Cai. Calculation and Error Analysis of Electromagnetic Torque for a Wheel Permanent-Magnet Motor // IEEE Transactions on Industry Applications. – 2006. – V. 42. – № 5. – P. 1155–1161.
17. Arash Kiyoumarsi. Prediction of torque pulsations in brushless permanent–magnet motors using improved analytical technique // Journal of Electrical Engineering. – 2010. – V. 61. – № 1. – P. 37–43.
18. Steinbrink J. Analytical Determination of the Cogging Torque in Brushless Motors Excited by Permanent Magnets // IEEE International. – 2007. – V. 1. – P. 172–177.
19. Favre E., Cardoletti L., Jufer M. Permanent-Magnet Synchronous Motors: A Comprehensive Approach to Cogging Torque Suppression // IEEE Transactions on industry applications. – 1993. – V. 29. – № 6. – P. 1141–1149.

Томасов Валентин Сергеевич – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат технических наук, доцент, зав. кафедрой, tomasov@ets.ifmo.ru

Ловлин Сергей Юрьевич – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, младший научный сотрудник, seri-l@yandex.ru

Егоров Алексей Вадимович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, инженер-исследователь, alexeykey@rambler.ru

4 КОМПЬЮТЕРНЫЕ СИСТЕМЫ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

УДК 681.3

РАСПРЕДЕЛЕНИЕ ФРАГМЕНТИРОВАННЫХ КАДРОВ ПО РЕЗЕРВИРОВАННЫМ КАНАЛАМ ВЫЧИСЛИТЕЛЬНОЙ СЕТИ

А.В. Осипов

Предложен алгоритм распределения кадров по резервированным гетерогенным каналам, учитывающий физические особенности каналов. Приведена имитационная модель работы алгоритма на базе MATLAB/Simulink, подтверждающая преимущества предлагаемого алгоритма по сравнению с алгоритмами RoundRobin и RateBalance.
Ключевые слова: резервирование каналов, канальный уровень.

Введение

Потребность в построении высокопроизводительных вычислительных сетей в условиях ограниченных ресурсов послужила одной из причин развития технологий агрегации, которые используют существующие физические каналы передачи данных для построения канала с большей пропускной способностью [1, 2]. Агрегирование каналов позволяет не только повысить суммарную пропускную способность сети [3], но и более эффективно использовать существующие каналы передачи данных [4]. Одновременное использование нескольких каналов связано с вариантностью организации межмашинного обмена. Реализация эффективного обмена при объединении в единую резервированную среду нескольких каналов требует исследования методов распределения кадров по физическим каналам и оценки их эффективности [5–10].

Алгоритмы распределения кадров по резервированным каналам

Эффективность использования пропускной способности магистралей при объединении на канальном уровне в значительной степени определяется алгоритмом распределения кадров по физическим каналам. В существующих технологиях агрегирования применяется алгоритм RoundRobin, циклически распределяющий поток кадров по имеющимся магистралям. Однако алгоритм RoundRobin, обеспечивая балансировку нагрузки, может не сохранять порядок следования кадров. Согласно алгоритму RoundRobin, каждый следующий кадр передается через очередную магистраль независимо от времени доставки предыдущего кадра. Доля инверсий в потоке принятых кадров может быть низка только при агрегировании абсолютно идентичных физических каналов. Функция исправления инверсий (нарушений порядка доставки кадров) может быть реализована на верхних уровнях сетевой модели, однако, ввиду ограниченности размера буферов передатчика и приемника, процесс исправления инверсий ведет к резкому снижению эффективности использования гетерогенных каналов, так как максимальная пропускная способность в каждом канале ограничена минимальным из значений пропускных способностей всех каналов.

В работе [11] для распределения кадров по физическим каналам передачи данных предложен адаптивный алгоритм RateBalance. В основу алгоритма RateBalance положен следующий принцип: кадр передается по тому каналу, который доставит его раньше. Для каждого физического канала i хранится время t_i окончания передачи кадров, поставленных в очередь отправки канала. Алгоритм выполняет распределение кадров по N доступным каналам, основываясь на оценке времени доставки кадра. Номер канала i для отправки кадра размером s выбирается, исходя из соотношения $i = \min_j f(s, j)$, где $f(s, j) = t_j + s / b_j$, b_j – пропускная способность канала j. При этом время t_i освобождения канала i корректируется по формуле $t_i := t_i + s / b_i$ [11].

Для каналов с различной производительностью пропускная способность агрегированного канала, использующего RateBalance, заметно выше. При этом RateBalance дает хорошее приближение к пиковой производительности (равной сумме производительностей физических каналов, составляющих агрегированный канал). Алгоритм RateBalance лучше сохраняет исходный порядок кадров, соседние кадры входной последовательности находятся рядом и в выходной. Соответственно, сдвиг скользящего окна будет выполняться непрерывно и без задержек, в отличие от алгоритма RoundRobin.

Существующие алгоритмы распределения кадров по каналам не определяют оптимальный вариант фрагментации пакетов. В настоящей работе предложен алгоритм распределения кадров по агрегированным каналам, осуществляющий оптимальную фрагментацию пакетов.

Оптимальная фрагментация и распределение кадров по каналам

Для повышения эффективности использования пропускных способностей физических каналов предлагается осуществлять оптимальную фрагментацию каждого пакета на кадры, причем длины кадров,

пересылаемых по физическим каналам, должны соответствовать оптимальному размеру для каждого физического канала. Таким образом, пакет может быть разбит на кадры разной длины.

Предлагаемый алгоритм DPB (Delivery probability balancing) учитывает следующие параметры:

– параметры пакетов, поступающих на отправку:
  – длина пакета (бит) – l;
– параметры используемых физических каналов:
  – средняя величина задержки начала отправки кадра (с) – t_d;
  – вероятность возникновения битовой ошибки – *BER*;
  – скорость передачи данных (бит/с) – br;
  – оптимальная длина кадра (бит) – l_o;
  – объем служебной информации, присоединяемой к кадру (бит) – l_h.

Средняя величина задержки начала отправки кадра определяется используемым протоколом передачи данных в физическом канале и битовой скоростью в канале. Физический канал может обладать нулевой задержкой, например, в случае использования системы Aloha со случайным доступом. В случае использования механизма CSMA/CA при обнаружении занятого состояния канала происходит задержка повторной попытки отправки на случайное число слотов ожидания из интервала [0, $2^{BE} - 1$], где –*BE* (Backoff exponent) – показатель степени границы интервала количества слотов ожидания. Алгоритм DPB при распределении данных по физическим каналам учитывает среднюю величину суммарной временной задержки физического канала перед отправкой кадра.

Вероятность возникновения битовой ошибки (*BER* – bit error rate) в физическом канале определяется мощностью сигнала на приемнике (которая зависит от мощности передатчика), спектральной плотностью шумов и используемой модуляцией. Если допустить, что спектральная плотность шумов и функциональная зависимость мощности на приемнике от мощности передатчика остаются неизменными в процессе функционирования сети, то значение *BER* можно определить расчетным путем даже при изменении скорости, мощности и типа модуляции на передатчике в процессе функционирования сети.

В зависимости от особенностей вычислительной сети (производительность узлов, наличие или отсутствие необходимости снижения энергопотребления) используемые в алгоритме значения параметров физических каналов и оптимальной фрагментации могут быть определены на этапе создания/настройки сети или же определяться динамически перед отправкой каждого пакета, либо при изменении физических параметров канала (например, битовой скорости). Рассмотрена оценка эффективности алгоритма распределения для стационарного режима работы сети, при котором интенсивность запросов и средняя длина пакетов не меняется во времени.

Задача определения оптимальной длины кадра для одного физического канала, учитывающая физические параметры канала, рассмотрена в [12]. Определение оптимальной длины кадров может быть выполнено однократно при проектировании/настройке сети, либо, если узлы сети обладают достаточными вычислительными и энергетическими ресурсами, может выполняться по расписанию или перед каждой отправкой пакета.

Распределитель кадров по алгоритму DPB схематично представлен на рис. 1.

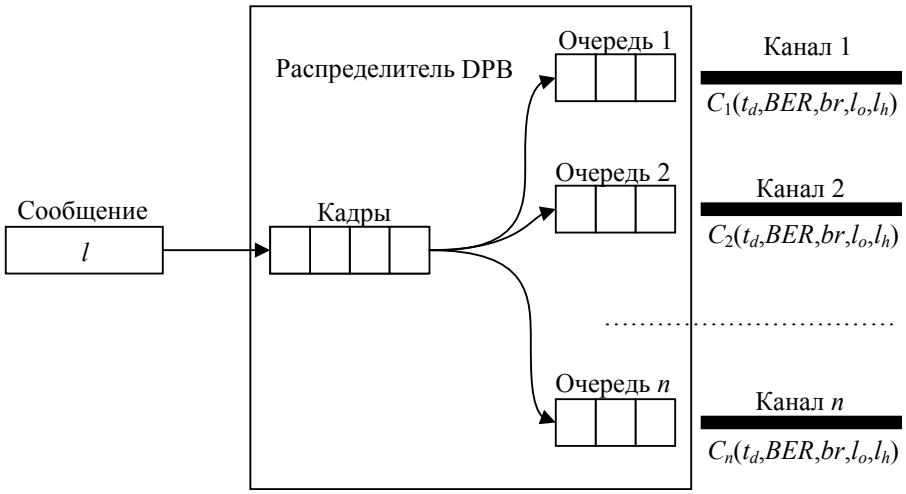


Рис. 1. Распределитель DPB

Задача оптимального распределения кадров по нескольким физическим каналам сводится к минимизации функции максимального времени отправки всех кадров из одной очереди:

$$F = \max\left(\left(\frac{l_{o1}+l_{h1}}{br_1}+t_{d1}\right)\cdot k_1, \left(\frac{l_{o2}+l_{h2}}{br_2}+t_{d2}\right)\cdot k_2, \ldots, \left(\frac{l_{on}+l_{hn}}{br_n}+t_{dn}\right)\cdot k_n\right),$$

где n – число агрегированных физических каналов (шт); l_{on} – оптимальная длина кадра для физического канала n (бит); l_{hn} – объем служебной информации, присоединяемой к кадру в физическом канале n (бит); br_n – битовая скорость передачи в физическом канале n (бит/с); t_{dn} – средняя величина задержки начала отправки кадра в физическом канале n (с); k_n – число кадров исходного пакета, определяемых к отправке через физический канал n (шт).

Для алгоритмов RoundRobin, RateBalance и DPB построены имитационные модели с использованием библиотеки SimEvents пакета MATLAB/Simulink. Имитационная модель представлена на рис. 2.

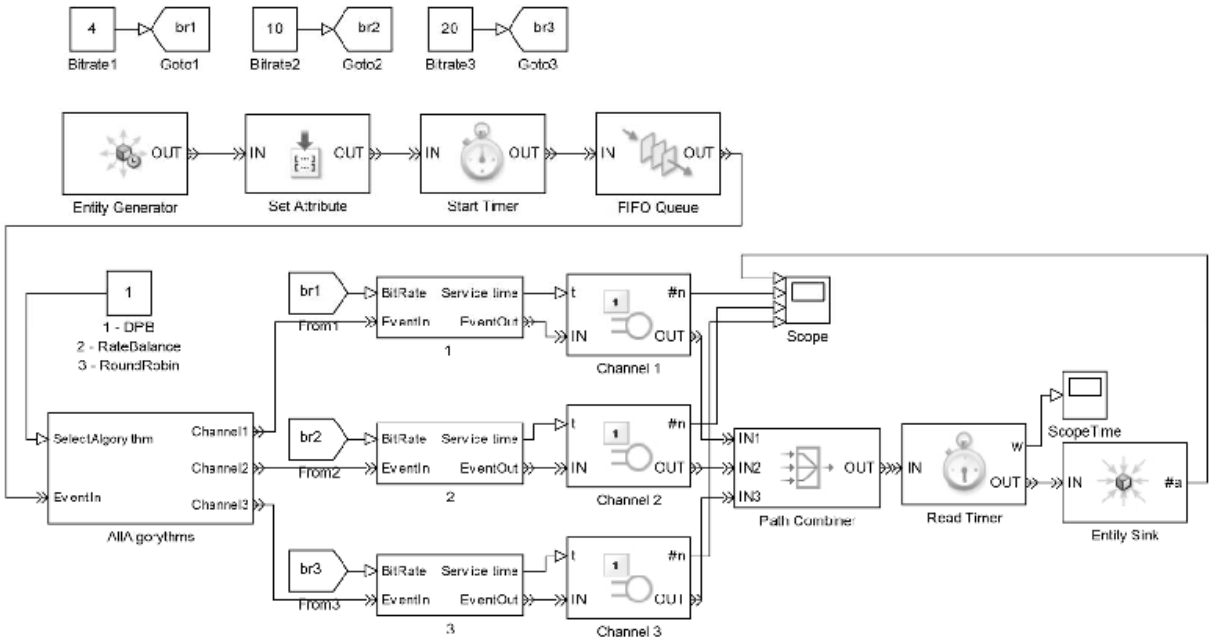


Рис. 2. Имитационная модель сети с тремя параллельными каналами

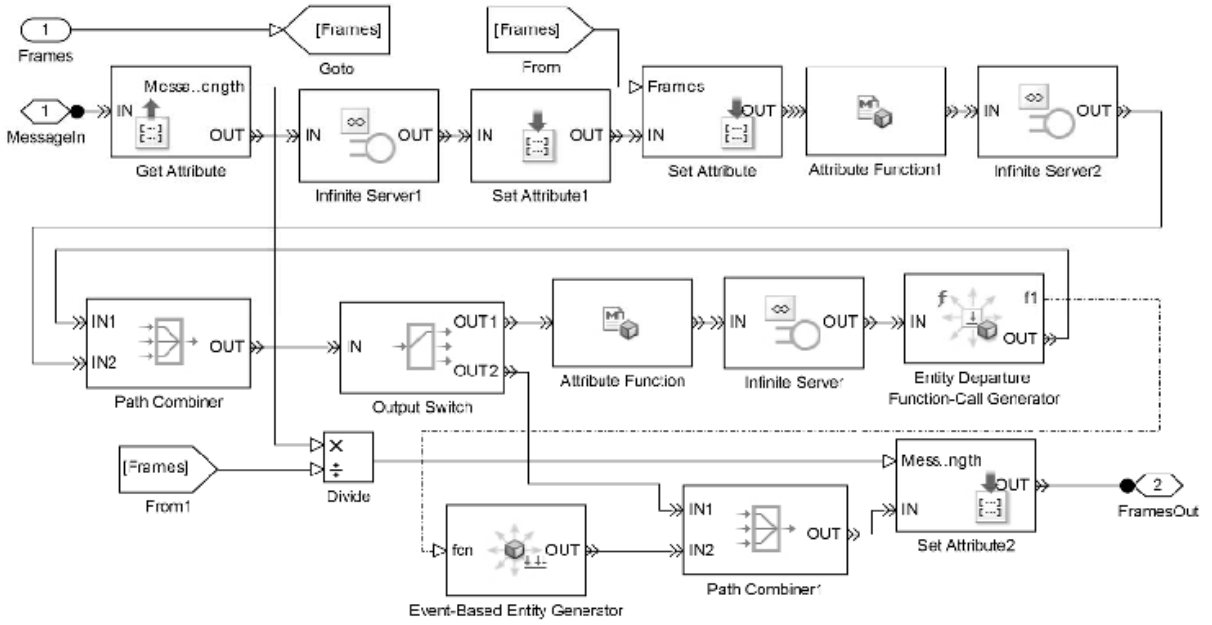


Рис. 3. Процесс фрагментации, реализованный в Simulink

Процесс фрагментации, осуществляемый алгоритмом DPB, представлен на рис. 3. Модель описывает вариант организации системы, в которой оптимальная длина кадра при заданной загрузке для каж-

дого канала определяется по априорно известным данным или по статистическим данным, собранным в процессе эксплуатации. Следовательно, при стационарности процессов передачи, когда интенсивность запросов и средняя длина пакетов не меняется во времени, на поиск оптимального варианта фрагментации отправляемого пакета время не тратится.

Адекватность модели проверена со следующими исходными значениями.

1. При установлении пропускных способностях второго и третьего канала, близких к нулю, передача данных происходит только по первому каналу. Все три алгоритма показывают одинаковое среднее время пребывания требования в системе.
2. При установлении равных пропускных способностей для всех каналов алгоритмы RoundRobin и RateBalance показывают одинаковое среднее время пребывания требования в системе, а алгоритм DPB – втрое меньшее (так как алгоритм осуществляет фрагментацию каждого пакета на кадры, и кадры отправляются по трем каналам одновременно).

Произведено имитационное моделирование вычислительной сети, функционирующей по трем параллельным каналам передачи данных с пропускными способностями соответственно 4 Мбит/с, 10 Мбит/с и 20 Мбит/с. Сеть осуществляет передачу тестовых пакетов длиной 120 кбит. В ходе первого этапа имитационного моделирования измеряется число отправленных пакетов N за время работы сети t, равное 0,1 с. Результаты измерений приведены на рис. 4.

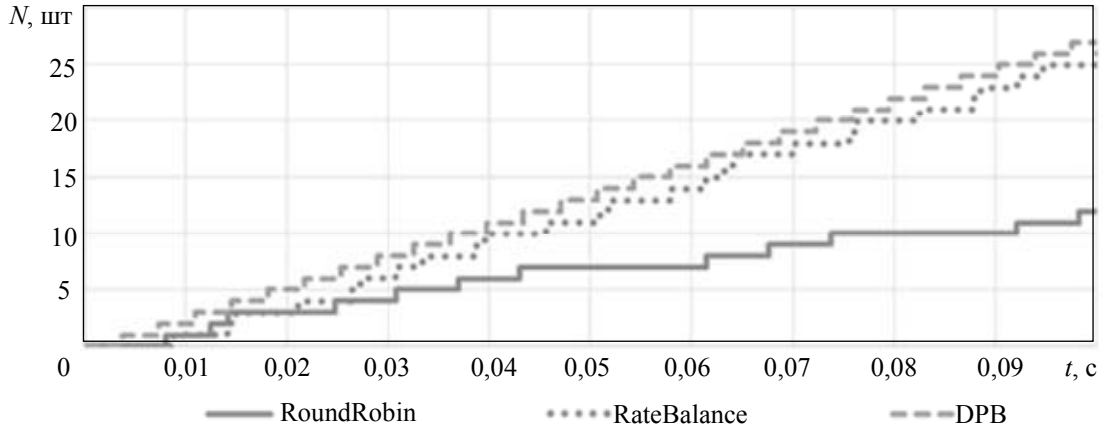


Рис. 4. Сравнение числа отправленных пакетов

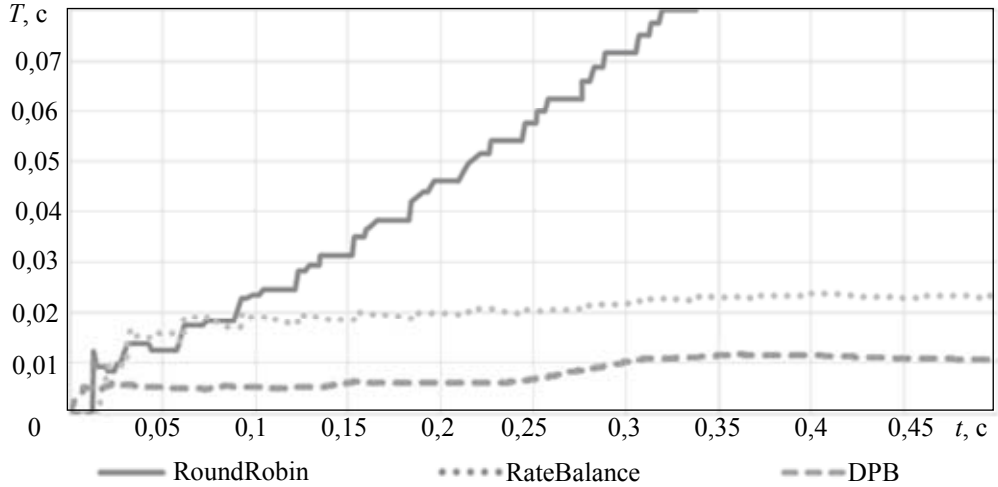


Рис. 5. Сравнение среднего времени пребывания требования в системе

Результаты имитационного моделирования показывают, что исследуемая вычислительная сеть, использующая алгоритм RoundRobin, обслуживает вдвое меньшее число заявок на отправление пакетов по сравнению с алгоритмами RateBalance и DPB. В ходе второго этапа имитационного моделирования за время работы сети t, равное 0,5 с, измеряется среднее время пребывания требования в системе T. Как видно из рис. 5, в процессе функционирования сети с алгоритмом RoundRobin среднее время пребывания требования в системе постоянно увеличивается. Это объясняется нарушением условия эргодичности системы, при котором общая интенсивность обслуживания системы должна превосходить интенсивности входного потока требований.

Эффективность использования пропускной способности гетерогенных каналов (для значений пропускной способности 4 Мбит/с, 10 Мбит/с, 20 Мбит/с) алгоритмом RoundRobin слишком низка, чтобы обслужить имеющийся входящий поток требований. Алгоритм RateBalance позволяет сохранить условие эргодичности системы, однако среднее время пребывания требования в системе вдвое больше по сравнению с алгоритмом DPB.

Заключение

Предложен алгоритм распределения кадров по резервированным каналам, производящий оптимальную фрагментацию пакета перед отправкой, что позволяет сократить среднее время пребывания требования в системе. Построена имитационная модель стационарного процесса передачи в вычислительной сети для алгоритмов RoundRobin, RateBalance и DPB, подтверждающая преимущества алгоритма DPB в случае стационарности процесса передачи.

Литература

1. Олифер В.Г., Олифер Н.А. Компьютерные сети. Принципы, технологии, протоколы. – СПб: Питер, 2010. – 944 с.
2. Таненбаум Э. Компьютерные сети. – 5-е изд. – СПб: Питер, 2012. – 960 с.
3. Tunnel provisioning with link aggregation: United States Patent Application Publication, Inventor: Ronen Solomon. – Pub. Date: Oct. 27, 2011. – Pub. No.: US 2011/0261829 A1.
4. McLean I., Christensen K. Reducing Energy Use: A Dual-Channel Link // IEEE Communications Letters. – 2012. – V. 16. – № 3. – P. 411–413.
5. Богатырев В.А. К управлению взаимосвязью при резервировании неполносвязных магистралей локальной сети // Электронное моделирование. – 1998. – № 4. – С. 28–32.
6. Богатырев В.А. Комбинаторный метод оценки отказоустойчивости многомагистрального канала // Методы менеджмента качества. – 2000. – № 4. – С. 30–35.
7. Богатырев В.А. Оценка надежности резервированного канала с межмагистральной ретрансляцией кадров // Автоматизация и современные технологии. – 2000. – № 4. – С. 2–4.
8. Богатырев В.А. Оценка вероятности полной связности локальных сетей при неполнодоступности резервированных магистралей // Электронное моделирование. – 1999. – № 5. – С. 102–112.
9. Богатырев В.А., Евлахова А.В., Котельникова Е.Ю., Богатырев С.В., Осипов А.В. Организация межмашинного обмена при резервировании магистралей // Научно-технический вестник СПбГУ ИТМО. – 2011. – № 2 (72) – С. 171.
10. Осипов А.В., Богатырев В.А. Варианты объединения разнотипных каналов вычислительной сети // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – № 2 (78). – С. 145.
11. Поляков А.Ю. Адаптивный подход к распределению информационных блоков по каналам передачи данных // Электросвязь. – 2009. – № 6. – С. 32–35.
12. Осипов А.В., Фролова Н.И. Задача оптимизации при сегментации пакетов в беспроводных сетях // Сборник трудов молодых ученых и сотрудников кафедры ВТ. – СПб: СПбГУ ИТМО, 2012. – Вып. 3. – С. 44–46.

Осипов Андрей Владимирович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, osipov-andrey@mail.ru

УДК 519.872

АППРОКСИМАЦИЯ ВЕРОЯТНОСТНЫХ РАСПРЕДЕЛЕНИЙ В МОДЕЛЯХ МАССОВОГО ОБСЛУЖИВАНИЯ

Т.И. Алиев

Для вероятностных распределений с коэффициентами вариации, отличными от единицы, получены математические зависимости для аппроксимации по двум заданным моментам распределения с использованием мультиэкспоненциальных распределений. Для аппроксимации распределений с коэффициентами вариации меньше единицы предлагается использовать гипоэкспоненциальное распределение, которое, в отличие от распределения Эрланга, имеющего дискретные значения коэффициента вариации, позволяет формировать случайные величины с любым значением коэффициента вариации в интервале (0; 1).

Ключевые слова: вероятностное распределение, коэффициент вариации, аппроксимация, гипоэкспоненциальное распределение, гиперэкспоненциальное распределение, распределение Эрланга.

Введение

В качестве моделей вычислительных систем и сетей широкое применение находят математические модели массового обслуживания, позволяющие проводить анализ эффективности их функционирования и

решать задачи системотехнического проектирования [1, 2]. При этом наибольшее распространение получили так называемые экспоненциальные модели, в которых протекающие в исследуемых системах процессы описываются случайными величинами, распределенными по экспоненциальному закону, обладающему свойством отсутствия последействия. Это замечательное свойство экспоненциального распределения используется при построении моделей в терминах марковских процессов, представляющих собой особый класс случайных процессов, развитие которых не зависит от предыстории процесса [3]. Благодаря этому для многих моделей массового обслуживания удается достаточно просто получить конечные результаты, в том числе в явном виде и в аналитической форме, для расчета характеристик исследуемой системы. Исходя из этого, часто при исследовании систем, в которых временные процессы отличаются от экспоненциальных, стремятся свести эти процессы к экспоненциальному представлению.

Для экспоненциального закона распределения случайных величин, определенных в области положительных значений, коэффициент вариации, описывающий разброс значений случайной величины, равен единице. Если реальные временны*е* интервалы имеют значения коэффициента вариации, отличающиеся от единицы, использование экспоненциального распределения может привести к значительным погрешностям конечных результатов. В этих случаях в качестве аппроксимирующих функций законов распределений могут использоваться мультиэкспоненциальные распределения, представляющие собой композицию экспоненциальных распределений, а именно: распределение Эрланга, когда коэффициент вариации временного интервала меньше единицы, $0 < \nu < 1$, и гиперэкспоненциальное распределение, когда коэффициент вариации временного интервала больше единицы, $\nu > 1$ [3]. При этом аппроксимация реального распределения в простейшем случае может выполняться по двум первым моментам распределения – математическому ожиданию и коэффициенту вариации.

Аппроксимация распределений с коэффициентом вариации $0 < \nu < 1$

Положим, что математическое ожидание и коэффициент вариации некоторой случайной величины τ, определенной в положительной области действительных чисел, соответственно равны t и ν, причем $0 < \nu < 1$. Для аппроксимации закона распределения такой случайной величины в теории массового обслуживания обычно используют распределение Эрланга [3]. Случайная величина, распределенная по закону Эрланга *k*-го порядка E_k, представляет собой сумму *k* экспоненциально распределенных случайных величин (фаз) с одинаковым математическим ожиданием $M[\tau]$.

Математическое ожидание и коэффициент вариации случайной величины, распределенной по закону Эрланга *k*-го порядка, равны соотвтственно

$$M_{E_k} = kM[\tau]; \qquad \nu_{E_k} = \frac{1}{\sqrt{k}},$$

где $k = 1, 2, ...$ – параметр распределения Эрланга, принимающий целочисленные значения.

Тогда для заданных реальных (измеренных) значений математического ожидания t и коэффициента вариации $\nu\,(0 < \nu < 1)$ некоторой случайной величины τ, определенной в положительной области действительных чисел, параметры аппроксимирующего распределения Эрланга будут определяться следующим образом:

$$k = \left]\frac{1}{\nu^2}\right[; \qquad M[\tau] = \frac{t}{k},$$

где]*x*[означает ближайшее целое, большее *x*. Нетрудно убедиться, что распределение Эрланга позволяет аппроксимировать только те реальные распределения, коэффициенты вариации которых имеют следующие значения: 0,707 (при *k* = 2); 0,577 (при *k* = 3); 0,5 (при *k* = 4) и т.д.

Для аппроксимации распределений с любым значением коэффициента вариации, находящимся в интервале (0; 1), рассмотрим многофазное распределение с разными значениями математического ожидания экспоненциальных распределений в разных фазах: $t_i = M_i[\tau] \quad (i = \overline{1,k})$, которое будем называть гипоэкспоненциальным распределением.

Проанализируем свойства гипоэкспоненциального распределения на примере двухфазного распределения, математическое ожидание и второй начальный момент которого будут равны $M_{h_2} = t_1 + t_2$; $\alpha^{(2)}_{h_2} = 2(t_1^2 + t_1 t_2 + t_2^2)$, откуда коэффициент вариации $\nu_{h_2} = \sqrt{t_1^2 + t_2^2}\,/(t_1 + t_2)$.

На рис. 1 показана зависимость коэффициента вариации гипоэкспоненциального распределения 2-го порядка от отношения t_1 / t_2 параметров экспоненциальных составляющих. Как видно из графика, коэффициент вариации гипоэкспоненциального распределения изменяется в пределах от 1 до 0,7, точнее, до значения 0,707, т.е. коэффициента вариации распределения Эрланга 2-го порядка, когда параметры экспоненциальных составляющих равны между собой, $t_1 = t_2$.

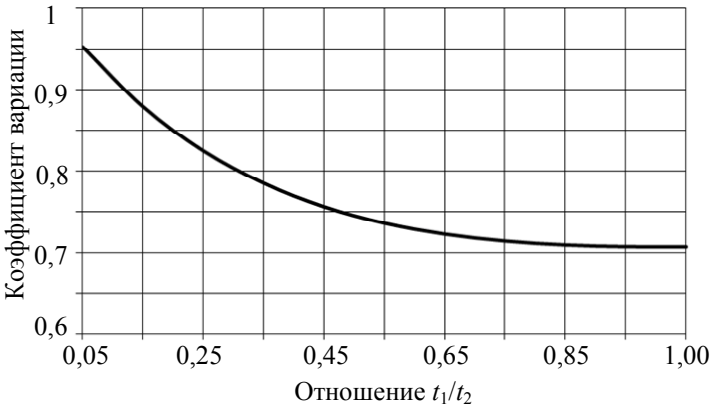


Рис. 1. Зависимость коэффициента вариации от отношения t_1/t_2

Очевидно, что для увеличения интервала изменения коэффициента вариации гипоэкспоненциального распределения необходимо вместо двухфазного использовать многофазное (k-фазное, $k > 2$) представление. Математическое ожидание и коэффициент вариации гипоэкспоненциального распределения k-го порядка будут равны

$$M_{h_k} = \sum_{i=1}^{k} t_i; \qquad \nu_{h_k} = \sqrt{\sum_{i=1}^{k} t_i^2} \Big/ \sum_{i=1}^{k} t_i \,. \tag{1}$$

Коэффициент вариации гипоэкспоненциального распределения k-го порядка лежит в интервале $(1/\sqrt{k}; 1)$, причем с увеличением порядка k левая граница этого интервала приближается к нулю ($k = 2, 3, \ldots$).

Рассмотрим теперь задачу аппроксимации реального распределения с коэффициентом вариации $0 < \nu < 1$ гипоэкспоненциальным распределением k-го порядка. Положим, что известны математическое ожидание t и коэффициент вариации ν (причем $0 < \nu < 1$) некоторой случайной величины τ, определенной в положительной области действительных чисел. Для простоты без потери общности положим, что аппроксимирующее гипоэкспоненциальное распределение k-го порядка содержит только два типа экспоненциальных фаз: k_1 фаз с параметром $\alpha_1 = 1/t_1$ и $k_2 = k - k_1$ фаз с параметром $\alpha_2 = 1/t_2$, где t_1 и t_2 – математические ожидания экспоненциально распределенных случайных величин в фазах первого и второго типов соответственно. Тогда из (1) следует, что математическое ожидание и коэффициент вариации будут равны

$$M_{h_k} = k_1 t_1 + k_2 t_2; \quad \nu_{h_k} = \sqrt{k_1 t_1^2 + k_2 t_2^2} \Big/ (k_1 t_1 + k_2 t_2),$$

причем $k_1 + k_2 = k$.

Таким образом, для аппроксимации по двум моментам необходимо, чтобы выполнялись следующие два условия:

$$\left.\begin{aligned} & k_1 t_1 + k_2 t_2 = t; \\ & \frac{\sqrt{k_1 t_1^2 + k_2 t_2^2}}{k_1 t_1 + k_2 t_2} = \nu \end{aligned}\right\},$$

где t и ν – соответственно математическое ожидание и коэффициент вариации аппроксимируемого (реального) распределения. Полагая, что значения k_1 и k_2 известны, решим полученную систему уравнений относительно неизвестных t_1 и t_2. Из первого уравнения системы следует, что

$$t_2 = \frac{t - k_1 t_1}{k_2}\,. \tag{2}$$

Подставляя это выражение во второе уравнение, после некоторых преобразований получим квадратное уравнение с одним неизвестным t_1:

$$k_1 (k_1 + k_2) t_1^2 - 2 k_1 t\, t_1 + (t^2 - k_2 \nu^2 t^2) = 0\,.$$

Решая это квадратное уравнение, получим:

$$t_1 = \frac{t}{k}\left[1 \pm \sqrt{\frac{k_2}{k_1}\left(k\nu^2 - 1\right)}\right], \tag{3}$$

где $k = k_1 + k_2$.

В качестве решения могут использоваться оба корня квадратного уравнения.

Для того чтобы в (3) под знаком квадратного корня была неотрицательная величина, необходимо выполнение условия

$$k \geq \frac{1}{\nu^2}, \tag{4}$$

которое определяет минимальное количество фаз в аппроксимирующем гипоэкспоненциальном распределении, т.е. порядок распределения. Для того чтобы второе решение со знаком минус перед квадратным корнем давало $t_1 \geq 0$, дополнительно необходимо выполнение условия

$$k_2 \leq \frac{1}{\nu^2}. \tag{5}$$

Объединив условия (4) и (5), окончательно получим вполне очевидное условие

$$k_2 \leq \frac{1}{\nu^2} \leq k. \tag{6}$$

Подставим теперь (3) в (2) и найдем t_2:

$$t_2 = \frac{t}{k}\left[1 \mp \sqrt{\frac{k_1}{k_2}\left(k\nu^2 - 1\right)}\right],$$

для которого получим условие, аналогичное условию (6):

$$k_1 \leq \frac{1}{\nu^2} \leq k.$$

Окончательно имеем следующие выражения для аппроксимации гипоэкспоненциальным распределением *k*-го порядка законов распределения случайных величин с коэффициентом вариации $0 < \nu < 1$:

$$k \geq \frac{1}{\nu^2}; \tag{7}$$

$$\left.\begin{aligned} t_1 &= \frac{t}{k}\left[1 + \sqrt{\frac{k_2}{k_1}\left(k\nu^2 - 1\right)}\right] \\ t_2 &= \frac{t}{k}\left[1 - \sqrt{\frac{k_1}{k_2}\left(k\nu^2 - 1\right)}\right] \end{aligned}\right\} \text{ или } \left.\begin{aligned} t_1 &= \frac{t}{k}\left[1 - \sqrt{\frac{k_2}{k_1}\left(k\nu^2 - 1\right)}\right] \\ t_2 &= \frac{t}{k}\left[1 + \sqrt{\frac{k_1}{k_2}\left(k\nu^2 - 1\right)}\right] \end{aligned}\right\}. \tag{8}$$

Алгоритм аппроксимации реального распределения с коэффициентом вариации $0 < \nu < 1$ гипоэкспоненциальным распределением *k*-го порядка при заданных значениях математического ожидания t и коэффициента вариации ν случайной величины τ, определенной в положительной области действительных чисел, формулируется следующим образом:

1. на основе выражения (7) по заданному значению коэффициента вариации ν определяется минимально необходимое число экспоненциальных фаз k в аппроксимирующем распределении как ближайшее большее целое по отношению к $1/\nu^2$;
2. выбирается значение $k_1 < k$ и рассчитывается $k_2 = k - k_1$;
3. на основе (8) рассчитываются значения t_1 и t_2.

Результаты аппроксимации на основе выражений (7) и (8) при разных значениях k_1 и $k_2 = k - k_1$ различаются значениями третьего и более высоких моментов распределения, учет которых при аппроксимации реальных распределений не вызывает принципиальных трудностей, но сопровождается более громоздкими математическими выкладками. К тому же, во многих случаях влияние этих моментов на конечные результаты исследований оказывается незначительным.

Пример 1. Пусть математическое ожидание и коэффициент вариации аппроксимируемого выражения соответственно равны $t = 10$ и $\nu = 0,4$.

В соответствии с изложенным выше алгоритмом:

1. минимально необходимое число экспоненциальных фаз k в аппроксимирующем распределении $k = 7$ ($k \geq 1/0,16 = 6,25$);
2. выберем значение $k_1 = 3$, тогда $k_2 = 7 - 3 = 4$;
3. на основе (8) рассчитываются значения $t_1 = 2$ и $t_2 = 1$.

Таким образом, в качестве аппроксимирующего распределения выбираем гипоэкспоненциальное распределение 7-го порядка, в котором 3 экспоненциальные фазы имеют математическое ожидание, равное 2, и 4 фазы – математическое ожидание, равное 1.

Аппроксимация распределений с коэффициентом вариации $\nu > 1$

Положим, что математическое ожидание и коэффициент вариации некоторой случайной величины τ, определенной в положительной области действительных чисел, соответственно равны t и ν, причем $\nu > 1$. В этом случае для аппроксимации закона распределения в теории массового обслуживания используют гиперэкспоненциальное распределение [3], представляющее собой композицию экспоненциальных распределений.

Случайная величина, распределенная по гиперэкспоненциальному закону, представляет собой совокупность случайных величин (фаз), распределенных по r разным экспоненциальным законам, причем появление случайной величины, принадлежащей *i*-ой фазе, происходит с вероятностью

$$q_i \ (i = \overline{1, r}), \quad \sum_{i=1}^{r} q_i = 1 .$$

В простейшем случае гиперэкспоненциальное распределение может быть представлено в виде двух экспоненциальных распределений (рис. 2).

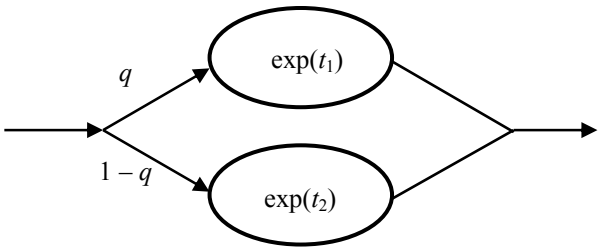


Рис. 2. Двухфазное гиперэкспоненциальное распределение

Параметрами такого распределения являются: t_1 и t_2 – математические ожидания экспоненциальных распределений; q – вероятность формирования случайной величины по первой экспоненте. Полученное таким образом распределение является трехпараметрическим. Это означает, что аппроксимация может выполняться по трем числовым моментам. Выбор значений параметров t_1, t_2 и q гиперэкспоненциального распределения по двум моментам предполагает наличие некоторого произвола.

Таким образом, задача аппроксимации гиперэкспоненциальным распределением сводится к определению значений параметров t_1, t_2 и q в зависимости от известных значений математического ожидания t и коэффициента вариации ν аппроксимируемого закона распределения случайной величины τ.

Математическое ожидание и второй начальный момент гиперэкспоненциального распределения соответственно равны

$$t = q\,t_1 + (1-q)\,t_2; \tag{9}$$

$$t^{(2)} = 2[q t_1^2 + (1-q)\,t_2^2].$$

Тогда коэффициент вариации гиперэкспоненциального распределения

$$\nu^2 = \frac{2[q\,t_1^2 + (1-q)\,t_2^2]}{t^2} - 1 ,$$

откуда

$$2\,[q t_1^2 + (1-q)\,t_2^2] = t^2\,(1+\nu^2). \tag{10}$$

Из (9) имеем:

$$t_2 = \frac{t - q\,t_1}{1-q} . \tag{11}$$

Подставив последнее выражение в (10), после некоторых преобразований получим квадратное уравнение

$$2q\,t_1^2 - 4q\ t\,t_1 + [1 + q - (1-q)\nu^2\,]t^2 = 0 .$$

Решая это квадратное уравнение относительно t_1, получим:

$$t_1 = t\left[1 \pm \sqrt{\frac{1-q}{2q}(\nu^2 - 1)}\right]. \tag{12}$$

Для того чтобы гарантировать $t_1 > 0$, в качестве решения выберем корень уравнения со знаком плюс перед знаком радикала:

$$t_1 = \left[1 + \sqrt{\frac{1-q}{2q}(\nu^2 - 1)}\right] t . \tag{13}$$

Подставим (13) в (11) и найдем t_2 :

$$t_2 = \left[1 - \sqrt{\frac{q}{2(1-q)}(\nu^2 - 1)}\right] t \,. \qquad (14)$$

Потребуем, чтобы $t_2 \geq 0$, т.е. $\frac{q}{2(1-q)}(\nu^2 - 1) \leq 1$. Отсюда

$$q \leq \frac{2}{1+\nu^2} \,. \qquad (15)$$

Выражения (13)–(15) можно использовать для аппроксимации любого закона распределения с коэффициентом вариации $\nu > 1$ двухфазным гиперэкспоненциальным распределением, для чего достаточно выбрать значение вероятности q из условия (15) и рассчитать значения t_1 и t_2 в соответствии с (13) и (14). В частном случае, когда $q = \frac{2}{1+\nu^2}$, с использованием (13) и (14) получим:

$$t_1 = \frac{\nu^2 + 1}{2} t; \quad t_2 = 0 \,. \qquad (16)$$

Последние выражения соответствуют однофазному представлению гиперэкспоненциального распределения, которое является частным случаем так называемого распределения Кокса.

Пример 2. Пусть $\nu = 3$, тогда в соответствии с (15) $q \leq 0,2$.

1. Выберем $q = 0,1$, тогда в соответствии с (13) и (14) $t_1 = 7\,t$; $t_2 = \frac{1}{3} t$.
2. Выберем $q = 0,2$, тогда в соответствии с (16) $t_1 = 5t$; $t_2 = 0$.

Заметим, что полученные для t_1 и t_2 выражения (13) и (14) симметричны. Можно показать, что если выбрать в качестве решения квадратного уравнения (12) второй корень со знаком минус перед знаком радикала и потребовать, чтобы выражение в квадратных скобках не было отрицательным, то получим:

$$t_1 = t\left[1 - \sqrt{\frac{1-q}{2q}(\nu^2 - 1)}\right]; \qquad t_2 = t\left[1 + \sqrt{\frac{q}{2(1-q)}(\nu^2 - 1)}\right],$$

а условие для выбора значения q примет вид $q \geq \frac{\nu^2 - 1}{\nu^2 + 1}$, что равносильно перестановке экспоненциальных фаз (рис. 2) гиперэкспоненциального распределения.

Заключение

Полученные соотношения позволяют выполнить на основе заданных двух моментов вероятностного распределения точный расчет параметров гипо- и гиперэкспоненциальных распределений, используемых для аппроксимации распределений случайных величин, определенных в положительной области действительных чисел и имеющих отличный от единицы коэффициент вариации. Благодаря этому процессы, протекающие, например, в неэкспоненциальных сетях массового обслуживания и системах с приоритетами, могут быть представлены в терминах марковского процесса. Использование гипоэкспоненциального распределения вместо распределения Эрланга позволяет выполнить аппроксимацию распределений с любым значением коэффициента вариации в интервале (0; 1), что повышает точность результатов расчета характеристик функционирования таких систем. Кроме того, предлагаемая аппроксимация по двум заданным моментам вероятностного распределения может использоваться для формирования случайных величин, распределенных по произвольному закону, при проведении имитационных экспериментов.

Литература

1. Алиев Т.И., Муравьева-Витковская Л.А. Приоритетные стратегии управления трафиком в мультисервисных компьютерных сетях // Изв. вузов. Приборостроение. – 2011. – Т. 54. – № 6. – С. 44–49.
2. Алиев Т.И. Задачи синтеза систем с потерями // Изв. вузов. Приборостроение. – 2012. – Т. 55. – № 10. – С. 57–63.
3. Клейнрок Л. Теория массового обслуживания: Пер. с англ. – М.: Машиностроение, 1979. – 432 с.

Алиев Тауфик Измайлович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, профессор, зав. кафедрой, :aliev@d1.ifmo.ru

УДК 004.75

ОРГАНИЗАЦИЯ ОТКАЗОУСТОЙЧИВОЙ РЕПЛИКАЦИИ ПОТОКОВЫХ ДАННЫХ ПОДСИСТЕМЫ ТРАНСПОРТА СООБЩЕНИЙ В АРХИТЕКТУРЕ CORBA ПРИ ПОМОЩИ ТЕХНОЛОГИИ ZEROMQ

Ф.А. Козлов

Обсуждаются особенности построения отказоустойчивой распределенной системы, главным функционалом которой является репликация потоковых данных. Представлены концепции реализации данной системы с использованием технологии ZeroMQ. Рассматриваются основные преимущества технологии ZeroMQ перед другими аналогами в создании данного вида распределенных систем.

Ключевые слова: распределенная система, репликация, кластер, брокер, ZeroMQ.

Введение

CORBA (Common Object Request Broker Architecture) – технологический стандарт написания распределенных приложений. Технология CORBA создана для поддержки разработки и развертывания сложных объектно-ориентированных прикладных систем. В настоящее время используется множество систем CORBA, и их разработка и стандартизация развиваются по мере того, как растет число установленных систем. Примером таких систем является системы электронного документооборота (СЭД).

Технологический стандарт CORBA позволяет отдельным сервисам, написанным на разных языках программирования, взаимодействовать друг с другом. Взаимодействие реализуется с помощью распределенных объектов [1]. В СЭД, основанных на CORBA, используются подсистемы транспорта сообщений (ПТС), являющиеся надстройкой над реализацией CORBA и предоставляющие интерфейс для обмена данными между серверными и клиентскими модулями системы. В компаниях, состоящих из нескольких географически удаленных офисов, возникает необходимость создания единой СЭД. Единая СЭД – это распределенная система, позволяющая производить выборочный обмен информацией между локальными СЭД.

Распределенная система – это информационная система, состоящая из независимых элементов, находящихся на удалении друг от друга. Одно из главных требований к распределенной системе – ее единство в представлении пользователя. Применение технологии CORBA в распределенной СЭД затрудняет ее дальнейшую поддержку, так как возникает необходимость синхронизации версий каждого сервера СЭД при изменении интерфейса одного из распределенных объектов [2]. Решением данной задачи является создание распределенной системы ПТС СЭД. Обмен данных в ПТС основан на передаче сообщений в событиях. Такое решение позволяет производить передачу сообщений с локального сервера СЭД в распределенную систему, создавая необходимые фильтры внешних сообщений.

Главной функцией распределенной системы ПТС СЭД является передача сообщения, полученного одной ПТС от локального модуля, всем ПТС, входящим в распределенную систему. Таким образом, производится репликация данных, поступающих на ПТС, в распределенной системе. Основой репликации является передача данных от подчиненного узла системы к главному узлу, который передает полученное сообщение на все подчиненные узлы. Недостатком такой схемы является ее низкая надежность. При выходе из строя главного узла распределенная система полностью теряет свою работоспособность. Повышение надежности распределенной системы возможно путем создания на ее основе отказоустойчивого кластера.

Целью настоящей работы является разработка метода построения отказоустойчивого кластера, производящего выборочную репликацию данных ПТС СЭД.

Анализ технологий передачи данных

Для обеспечения связи локальной ПТС СЭД с распределенной системой необходимо создать специальный модуль – брокер, предоставляющий функционал для взаимодействия с узлами кластера. Существуют готовые решения, позволяющие производить обмен данных и репликацию в распределенных системах. Наиболее распространенные технологии в данной области – RabbitMQ и ActiveMQ. Обе технологии используют протокол AMPQ [3] и предоставляют готовый брокер для обмена сообщениями между клиентами и сервером. К недостаткам готовых брокеров при организации распределенных СЭД относятся избыточный функционал, большой объем занимаемого дискового пространства и увеличение пакета передаваемых данных путем наложения дополнительных системных протоколов. Эти недостатки влияют на скорость передачи данных в распределенной системе. Высокая скорость передачи потоковых данных является одним из определяющих требований к распределенной СЭД.

Среди технологий обмена сообщениями выделяется низкоуровневая библиотека ZeroMQ, которая предоставляет интерфейс для обмена сообщениями через системные сокеты. Проведенные измерения времени передачи пакетов данных при использовании технологий ZeroMQ, RabbitMQ и ActiveMQ (табл. 1) показали, что технология ZeroMQ по скоростным характеристикам опережает свои аналоги. Измерения производились на ОС Ubuntu 12.04 с 6 потоками Pentium i7 и 2 ГБ оперативной памяти. Тестируемые технологии производили обмен пакетами размером 10 Б по связи «запрос–ответ». Относитель-

ная погрешность полученных результатов не превышает 1%. Средние скорости передачи пакетов составили: ZeroMQ – 27270 Б/с, RabbitMQ – 7000 Б/с, AtciveMQ – 230 Б/с. Высокая скорость передачи сообщений в сокетах ZeroMQ позволяет использовать данную технологию при создании брокера высокопроизводительной распределенной системы.

Количество пакетов	Время передачи пакетов, мс			
	RabbitMQ (Python)	ActiveMQ (Python)	ZeroMQ (Python)	ZeroMQ (C++)
100	–	4376	–	–
1000	1414	42137,66	352,33	379
10000	14694,66	430036,66	3332	3592,33
100000	140387,33	–	32465,33	36238,33

Таблица 1. Время передачи пакетов данных между клиентом и сервером

Анализ подсистемы транспорта сообщений

ПТС управляет каналами, по которым передаются сообщения. Модуль, получивший доступ к каналу, имеет возможность обмена сообщениями с подписанными на этот канал модулями СЭД. Сообщения в канале могут располагаться как в памяти самого ПТС, так и в базе данных. При сбое ПТС сохраненные в базе сообщения восстанавливаются при перезапуске системы.

Каналы ПТС разделяются на стабильные и нестабильные – по способу хранения сообщений в памяти, в также на глобальные и локальные – по области видимости канала. Для создания распределенной системы необходимо организовать потоковую передачу всех сообщений, поступающих в глобальные каналы ПТС, на локальный брокер, являющийся узлом кластера.

Архитектура распределенной системы

Так как в распределенной системе СЭД отсутствует условие обеспечения строгой идентичности всех серверов, в кластере используется асинхронная репликация данных. Организация отношений между узлами распределенной системы производится по схеме «главный–подчиненный» [4]. Предусматривается наличие в системе одного главного сервера (ГС), с которым соединены все остальные подчиненные сервера (ПС). Такой подход позволяет избежать лишних связей между узлами системы.

При использовании схемы «главный–подчиненный» необходимо обеспечить возможность автоматического возобновления репликации при сбое ГС в кратчайшие сроки. Для решения данной задачи создается кластер по схеме «Ротация резерва» (Rotating Standby) [5]. Эта схема предполагает наличие резервного сервера (РС), который при сбое ГС занимает его место. Место РС в такой ситуации занимает один из активных ПС. Главным преимуществом данной схемы является автоматическое возобновление работы кластера после сбоя и низкий промежуток простоя системы [6].

Брокер потоковых данных

Для реализации описанной выше архитектуры необходимо создать модуль брокера, предоставляющий определенный функционал в зависимости от ролей узла в кластере (рис. 1).

Рис. 1. Диаграмма вариантов использования системы

ПС обеспечивает асинхронный обмен сообщениями между локальным ПТС и ГС кластера. При этом ПС проходит регистрацию на ГС и обрабатывает системные сообщения кластера. Задачей ГС является прием и рассылка сообщений с удаленных ПТС. ГС производит выбор и наблюдение за РС, а при его сбое производит поиск нового РС среди ПС, а также производит регистрацию всех серверов кластера для дальнейшей рассылки системных сообщений. РС наблюдает за состоянием ГС. При сбое ГС РС переходит в состояние ГС и делает рассылку всем ПС системных сообщений о переподключении. Наблюдение за состояниями серверов производится с помощью технологии Heartbeat [7].

Для реализации асинхронного взаимодействия как на уровне функционала, предоставляемого брокером (одновременная работа ГС и ПС на одном узле), так и на уровне функционала сервера (асинхронный прием и отправка сообщений на ПС) необходимо создать систему потоков. Первый уровень потоков запускает функционал, предоставляемый ГС, ПС или РС. Второй уровень потоков обеспечивает асинхронную работу внутри каждого сервера. Для реализации потоков брокера в исследовании была использована технология Boost Thread [8].

Каждый поток брокера оперирует определенным набором сокетов для обмена данными. Технология ZeroMQ предоставляет расширения системных сокетов для решения определенных задач. Основными типами связи сокетов в ZeroMQ являются:

– «Публикация–подписка» (PUB–SUB);
– «Запрос–ответ» (REQ–REP);
– «Потоковая передача» (PUSH–PULL);
– «Уникальная пара» (PAIR).

«Публикация–подписка» является связью, где на один публикующий сервер подписывается множество серверов. Это односторонняя связь. «Публикация–подписка» используется в брокере для рассылки сообщений для ПТС и системных сообщений кластера на ГС, а также для наблюдения за состояниями серверов.

Связь «Запрос–ответ» представляет собой классическую синхронную модель взаимодействия клиента и сервера. Клиент производит запрос необходимой информации на сервер и получает ответ. В брокере такая связь необходима для выбора РС и регистрации серверов в кластере.

Связи «Потоковая передача» и «Уникальная пара» являются односторонними и предназначены для потоковой передачи данных без блокировки сокета. «Потоковая передача» используется в брокере для сбора ГС сообщений с удаленных ПТС. «Уникальная пара» отличается от «Потоковой передачи» тем, что в первой присутствуют только один передающий и один принимающий сокет [9]. Связь используется в брокере для обмена сообщениями между ПТС и брокером.

Структура ПС, РС и ГС состоит из главного и прикладных потоков. Потоки производят управление и обработку результатов на определенных сокетах. Каждый сокет брокера имеет один тип и выполняет одну функцию.

Система устроена так, что брокер перенаправляет полученные с кластера сообщения на вход локального ПТС (рис. 2). Таким образом, в случае задержек при обработке полученных сообщений на ПТС новые сообщения с кластера будут вставать в очередь обработки на ПТС, а не на локальном брокере. Независимо от работоспособности локального ПТС, брокер будет способен в полной мере поддерживать работоспособность всего кластера.

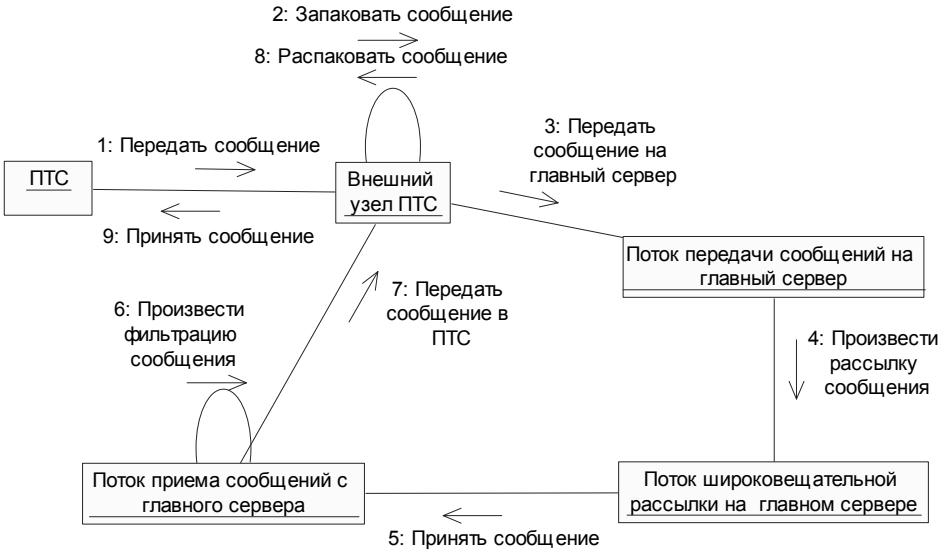


Рис. 2. Диаграмма коммуникации потоков системы при репликации данных в кластере

Способность технологии ZeroMQ производить обмен данных через сокеты на внутрипроцессном уровне позволяет управлять потоками в брокере с минимальной затратой ресурсов на время передачи сообщения.

Заключение

Разработанный брокер позволяет объединять удаленные подсистемы транспорта сообщений в отказоустойчивый кластер и производить их репликацию. Были проведены измерения времени репликации данных одной подсистемы транспорта сообщений на все подсистемы транспорта сообщений распределенной системы (табл. 2). Из полученных результатов следует, что брокер обладает высокой скоростью передачи данных. При увеличении узлов системы происходит дополнительная нагрузка на ГС, что снижает производительность брокера.

Недостатками брокера является отсутствие гарантии доставки сообщения и низкий уровень безопасности канала. Разработанная архитектура предполагает использование брокера в закрытых и защищенных сетях, в которых скорость передачи сообщения преобладает над надежностью его передачи. Такими системами являются подсистемы транспорта сообщений систем электронного документооборота, системы мгновенных сообщений и системы записи логов. Данные системы обладают высокой частотой передачи данных и требуют от распределенной системы высокой пропускной способности.

Количество пакетов	Время передачи пакетов для определенного количества узлов в кластере, мс			
	2	3	4	5
10000	121	171	219	301
100000	1091	1570	2091	2668
1000000	10926	14804	20481	26903

Таблица 2. Время репликации данных в кластере

Разработанная архитектура позволяет создавать брокеры с расширенным функционалом. В брокере возможна организация передачи данных по защищенному соединению [10] и реализация восстановления и повторной передачи недоставленных сообщений на ГС. Для оптимизации передачи сообщений в брокере возможна фильтрация определенных каналов подсистемы транспорта сообщений. Упомянутые расширения предусматривают дополнительную нагрузку и вводятся в зависимости от требований к системе.

Литература

1. Таненбаум Э., ван Стеен М. Распределенные системы. Принципы и парадигмы. – СПб: Питер, 2003. – 877 с.
2. Zinky J., Bakken D., Schantz R. Architectural Support for Quality of Service for CORBA Objects // Theory and Practice of Object Systems. – 1997. – V. 3. – P. 55–73.
3. Vinoski S. Advanced Message Queuing Protocol // IEEE Internet Computing. – 2006. – V. 10. – №. 6 – P. 87–89.
4. Hercule K., Eugene M., Paulin B., Joel L. Study of the Master-Slave replication in a distributed database // IJCSI International Journal of Computer Science Issues. – 2001. – V. 8. – P. 319–326.
5. Маркарян К. Построение кластера высокой надежности под управлением GNU/Linux [Электронный ресурс]. – Режим доступа: http://www.ibm.com/developerworks/ru/library/l-Cluster_Linux_1/, свободный. Яз. рус. (дата обращения 20.01.2013).
6. Шубинский И.Б. Структурное резервирование в информационных системах. Предельные оценки // Надежность – 2012. – № 1. – С. 118–125.
7. Fei-fei Li, Xiang-zhan Yu, Gang Wu. Design and Implementation of High Availability Distributed System Based on Multi-level Heartbeat Protocol // CASE '09 Proceedings of the 2009 IITA International Conference on Control, Automation and Systems Engineering. – 2009. – P. 83–87.
8. Williams A., Botet Escriba V. The Boost C++ Libraries BoostBook Documentation Subset. Chapter 30. Thread 4.0.0 [Электронный ресурс]. – Режим доступа: http://www.boost.org/doc/libs/1_53_0/doc/html/thread.html, свободный. Яз. англ. (дата обращения 15.01.2013).
9. Hintjens P., ZeroMQ: The Guide [Электронный ресурс]. – Режим доступа: http://zguide.zeromq.org/, свободный. Яз. англ. (дата обращения 15.01.2013).
10. Tariq M., Koldehofe B., Altaweel A., Rothermel K. Providing basic security mechanisms in broker-less publish/subscribe systems // Conference: Distributed Event-Based Systems (DEBS). – NY, USA, 2010. – P. 38–49.

Козлов Федор Алексеевич – ОАО «Промышленные информационные системы», программист, аспирант, kozlovfedor@gmail.com

5 ЛАЗЕРНЫЕ И БИОМЕДИЦИНСКИЕ ТЕХНОЛОГИИ

УДК 611.16:[681.787+681.784.8]

ИЗГОТОВЛЕНИЕ ТКАНЕИМИТИРУЮЩИХ ФАНТОМОВ И КАПИЛЛЯРОВ И ИХ ИССЛЕДОВАНИЕ МЕТОДОМ ОПТИЧЕСКОЙ КОГЕРЕНТНОЙ ТОМОГРАФИИ

А.В. Быков, М.В. Волков, М.А. Волынский, И.П. Гуров, М. Киннунен, Н.Б. Маргарянц, А.П. Попов

Рассмотрены методики изготовления тканеимитирующих фантомов и капилляров из поливинилхлорид-пластизоля и силикона для их использования в качестве тест-объектов в оптической когерентной томографии и эмуляции кожи и капилляров. Даны сравнительные характеристики этих материалов и рекомендации по их использованию. Показаны примеры визуализации фантомов методом оптической когерентной томографии. Показана возможность использования информации В-сканов для определения показателя преломления.

Ключевые слова: тканеимитирующие фантомы, капилляры, поливинилхлорид-пластизоль, силикон, оптическая когерентная томография, рефрактометрия.

Введение

Для исследования различных методов бесконтактной диагностики в медицине и биологии распространено использование тест-объектов, которые имитируют живые системы [1–8]. В англоязычной литературе такие объекты принято называть фантомами (phantoms).

Для развития систем оптической когерентной томографии (ОКТ) как систем бесконтактного контроля биологических микрообъектов [9, 10] изготовление фантомов также является важной задачей. Во-первых, возможность изготовления фантомов с заданными оптическими и геометрическими характеристиками позволяет использовать фантомы в качестве «стандартных» объектов в ОКТ для метрологической поверки приборов [8]. Во-вторых, в ОКТ фантомы в качестве эмуляторов кожи и капилляров интересны ввиду того, что часто нет возможности апробации новых методов диагностики на живых системах. Дополнительным преимуществом использования фантомов является их статичность. Впервые использование фантомов для ОКТ предложено в работе [11].

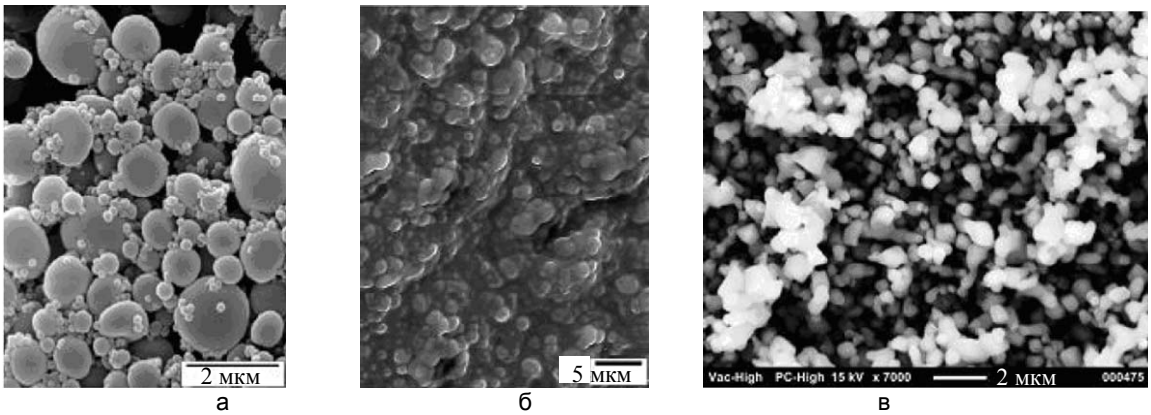
Актуальной задачей является выбор материала и методики изготовления фантомов. Анализ ранних публикаций [1–7], посвященных изготовлению и исследованию фантомов, позволяет выделить следующие материалы: интралипид, полиакриламидный гель, альбумин, агар, фибрин, эпоксидная смола, силикон. Работа с некоторыми из указанных материалов сложна. Например, интралипид представляет собой жировую эмульсию, не предусматривающую твердое состояние. Альбумин и агар быстро деградируют из-за роста грибков. Фибрин сложен в получении, поскольку представляет собой белок, образующийся из фибриногена плазмы крови в печени под воздействием фермента тромбина. Эпоксидная смола и силикон – похожие по свойствам материалы, удобные для изготовления капилляров. В работе [11] в качестве материала для изготовления фантомов предложен поливинилхлорид-пластизоль (ПВХ-пластизоль).

В настоящей работе рассмотрены особенности изготовления и исследования фантомов из ПВХ и силикона с указанием преимуществ и недостатков использования этих материалов на основе экспериментального опыта, накопленного в ходе совместных работ НИУ ИТМО и Университета Оулу (Финляндия).

Планарные фантомы из ПВХ

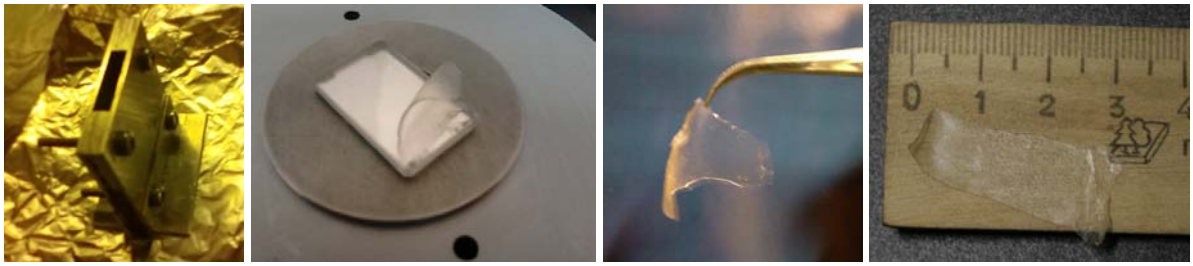
ПВХ-пластизоль представляет собой золь с частицами пластика, которые при термическом воздействии соединяются. Путем добавления в пластизоль наночастиц с известными размерами и известной концентрацией можно управлять рассеянием среды, которое может быть рассчитано с использованием теории Ми. В качестве наночастиц обычно используются TiO_2, ZnO или Fe_3O_4. Так, на рис. 1 показаны изображения ПВХ-пластизоля и наночастиц оксида титана, полученные с помощью электронного микроскопа, а также примеры полученных фантомов с различными свойствами.

Существуют два способа изготовления фантомов из ПВХ. Первый предполагает использование кюветы, в которую наливается пластизоль до термической обработки. Пример кюветы показан на рис. 2, а. Преимуществом данного способа является возможность контролировать геометрические параметры фантомов и изготавливать фантомы толщиной до 5 мм. Недостаток состоит в необходимости обеспечения герметичности кюветы во избежание вытекания пластизоля до его затвердевания. Второй способ заключается в поливе пластизоля на термостойкую поверхность с последующей термообработкой. Преимуществом такого метода является возможность изготовления тонких структур (толщиной порядка 200–300 мкм), а недостаток состоит в невозможности контроля толщины фантома. Пример фантома, изготовленного бескюветным способом, показан на рис. 2, б–г.



а б в

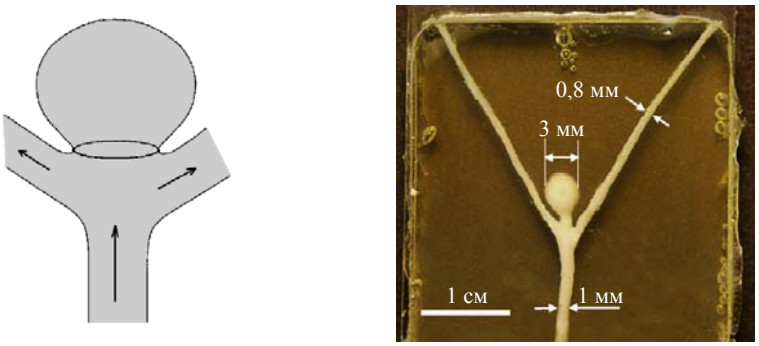
Рис. 1. ПВХ-пластизоль до (а) и после (б) термической обработки, наночастицы TiO_2 (в) со средним диаметром 650 нм



а б в г

Рис. 2. Кювета для изготовления фантома с заданными геометрическими характеристиками (а) и фантом, изготовленный бескюветным способом (б–г)

В качестве дополнительного преимущества фантомов, изготовленных с использованием кюветы, можно отметить возможность эмуляции капиллярной структуры путем закрепления в кювете проволоки с последующим ее удалением. На рис. 3 показан пример капиллярной сети в планарном фантоме, моделирующей аневризму [12].



а б

Рис. 3. Гипотетическая модель аневризмы (а) и ее реализация (б) в прозрачном фантоме. Сосуды заполнены 4% раствором интралипида

Следует отметить, что процесс изготовления фантомов из ПВХ трудоемок, требует специализированного оборудования и лабораторных условий с необходимым классом чистоты. Указанный факт является недостатком использования ПВХ-пластизоля в качестве материала для изготовления фантомов.

Цилиндрические фантомы из силикона

Другим материалом, составляющим альтернативу ПВХ-пластизолю, является силикон. В отличие от ПВХ, для работы с силиконом не требуется соблюдения специальных условий и наличия дорогостоящего оборудования. Силиконовым фантомам необходимо придать требуемую форму и оставить на сутки для отверждения на воздухе. Главный недостаток силиконовых фантомов состоит в сложности обеспечения равномерного распределения наночастиц в вязком силиконе, что делает невозможным управление оптическими свойствами таких фантомов. Практика показала, что из силикона наиболее удобно изготавливать фантомы цилиндрической формы, которые могут эмулировать капилляры. Отверстие внутри такого фантома может быть сделано путем введения проволоки или нити в процессе формирования фанто-

ма с последующим ее удалением после отверждения фантома. На рис. 4 показаны примеры таких капилляроподобных фантомов.

Рис. 4. Капилляроподобные фантомы из силикона

В таблице приведены сравнительные сведения о фантомах из ПВХ-пластизоля и силикона.

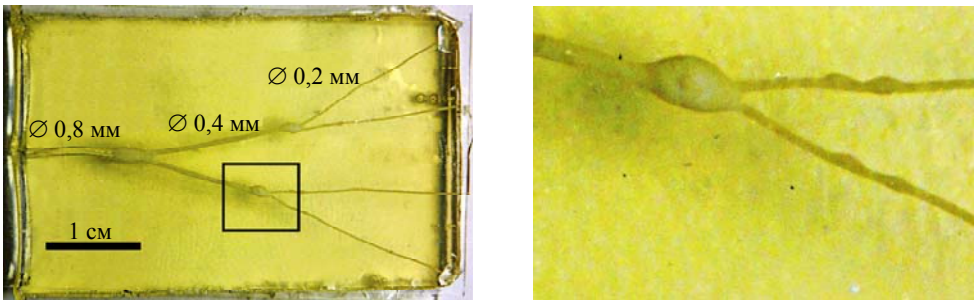
Характеристика	Фантомы из ПВХ-пластизоля		Фантомы из силикона
	С кюветой	Без кюветы	
Планарная структура	+	+	+/–
Цилиндрическая структура	+/–	–	+
Возможность добавления наночастиц	+	+	–
Возможность контроля толщины	+	–	+/–
Возможность эмуляции капилляров	+	–	+
Простота изготовления	–	+/–	+

Таблица. Сравнение фантомов, изготовленных различными способами

Исследование тканеимитирующих фантомов и капилляров методом ОКТ

В качестве моделей биологических объектов использовались фантомы с оптическими свойствами, близкими к реальным биотканям, и обладающие капиллярными сетями. Фантомы были изготовлены на основе ПВХ-пластизоля с рассеивающими частицами TiO_2. Их оптические свойства (коэффициенты рассеяния и поглощения, коэффициент анизотропии) были оценены путем спектрофотометрических измерений. Капиллярная структура была получена с помощью вложенной медной проволоки различных диаметров (рис. 5).

Для исследования капиллярной системы, показанной на рис. 5, использовался спектральный оптический когерентный томограф (ИПФ РАН, Россия) со следующими характеристиками: центральная длина волны 910 нм; ширина спектра источника 49 нм; мощность 0,6 мВт; разрешение по глубине (в воздухе) 10 мкм; диапазон сканирования (по вертикальной оси) 1,25 мм.

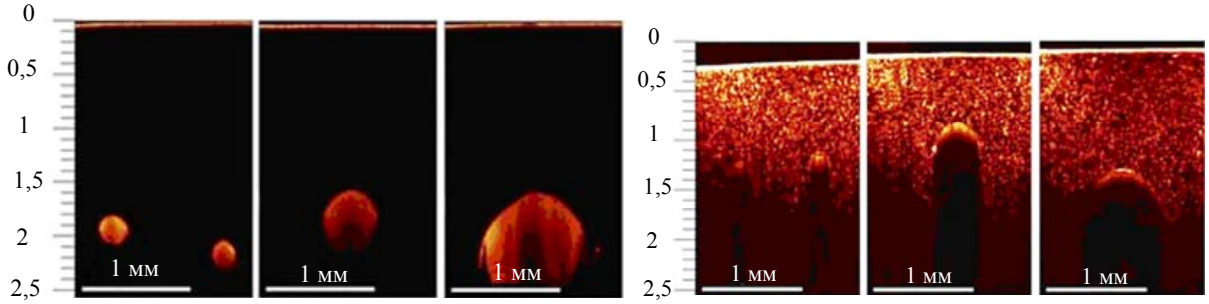


а б

Рис. 5. Изготовленная капиллярная система в прозрачном слое фантома (а) и увеличенный фрагмент области, выделенной прямоугольником (б). Знак Ø – диаметр поперечного сечения капилляра

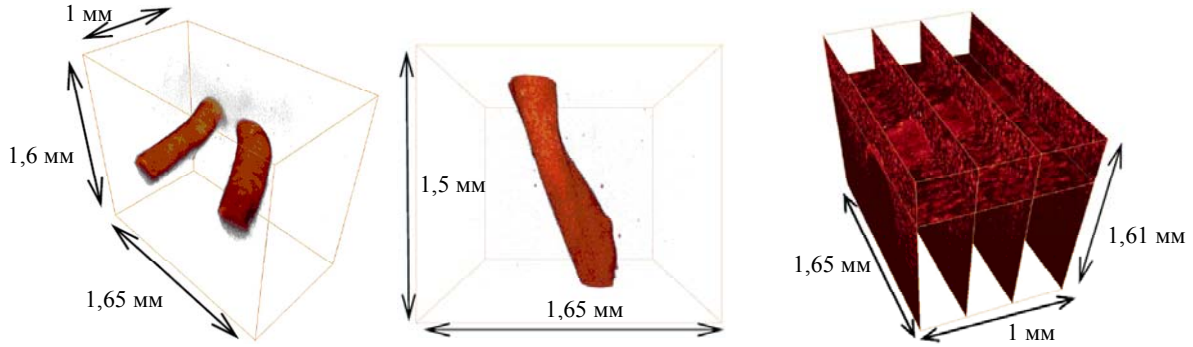
На рис. 6 представлены томограммы капилляров в прозрачном фантомном слое, заполненных 4% раствором интралипида, и капилляров, заполненных цельной кровью, вложенных в рассеивающий фантомный слой. С помощью методов ОКТ было определено, что глубина вложения (оптическая глубина) тонких сосудов (Ø 0,2 мм) находится в диапазоне 1,4–2 мм, для сосудов среднего размера (Ø 0,4 мм) глубина вложения составляет 0,9–1,6 мм, для толстых сосудов (Ø 0,8 мм) она составляет 1,3–1,6 мм.

На рис. 7 приведены трехмерные ОКТ-изображения тонких сосудов (Ø 0,2 мм) в прозрачном фантомном слое, заполненных 4% интралипидом, и сосудов толщиной 0,4 мм, заполненных цельной кровью в рассеивающем фантомном слое. Случай на рис. 7, б, соответствует патологически деформированному сосуду.



а б

Рис. 6. Томограммы капилляров в прозрачном фантомном слое, заполненных 4% интралипидом (а), и капилляров, заполненных цельной кровью и вложенных в рассеивающий фантомный слой (б). Концентрация рассеивающих частиц TiO_2 составляет 1,5 мг/мл. По вертикальной оси отложена оптическая длина пути (в мм)

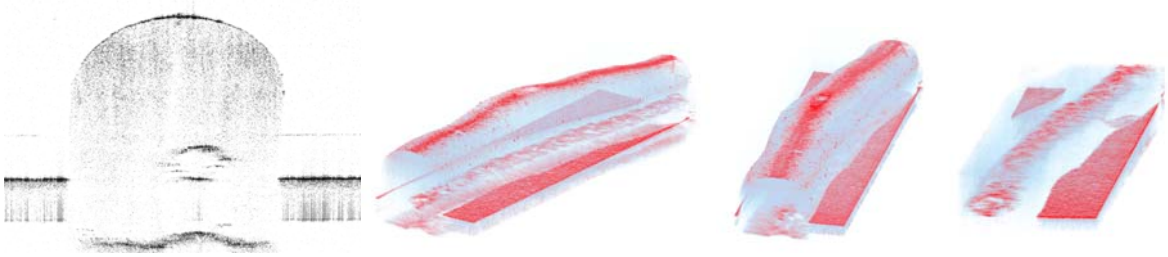


а б в

Рис. 7. Трехмерные ОКТ-изображения сосудов толщиной 0,2 мм в прозрачном фантомном слое, заполненных 4% интралипидом: два близко расположенных сосуда (а), патологически деформированный сосуд (б) и сосуд толщиной 0,4 мм в рассеивающем фантоме (в). Концентрация рассеивающих частиц TiO_2 составляет 1,5 мг/мл

Фантомы из силикона также могут быть визуализированы методом ОКТ. Для их исследования был использован оптический когерентный микроскоп с перестраиваемой длиной волны (модель EX-1301, Michelson Diagnostics Ltd., Великобритания) с центральной длиной волны 1305 нм, диапазоном перестройки длины волны 150 нм, мощностью излучения на выходе прибора порядка 3 мВт, разрешающей способностью по глубине 10 мкм и глубиной сканирования 2 мм.

На рис. 8 показаны результаты визуализации методом ОКТ силиконового капилляра толщиной около 1 мм и диаметром внутреннего отверстия около 200 мкм, что соответствует размерам кровеносных и лимфатических сосудов человека.



а б в г

Рис. 8. В-скан капилляра из силикона (а) и его трехмерные ОКТ-изображения: поверхность и внутренний канал (б), поверхность (в), внутренний канал (г). Ширина В-скана – 2 мм, глубина 2 мм. Длина участка капилляра на трехмерных изображениях – 1 см

Рефрактометрия с помощью ОКТ

Одним из преимуществ метода ОКТ применительно к исследованию фантомов является возможность его использования как косвенного метода определения эффективного показателя преломления и анализа его изменчивости по площади образца.

На рис. 9, а, показан пример В-скана края фантома на подложке (предметном стекле). Видно, что за счет не равного единице показателя преломления материала фантома его нижняя граница «утоплена»

в стекло. Отношение оптической длины пути вдоль каждого столбца В-скана к геометрической длине пути позволяет для каждого пиксела в горизонтальном направлении вычислить эффективный показатель преломления. На рис. 9, б, показано распределение показателя преломления вдоль В-скана.

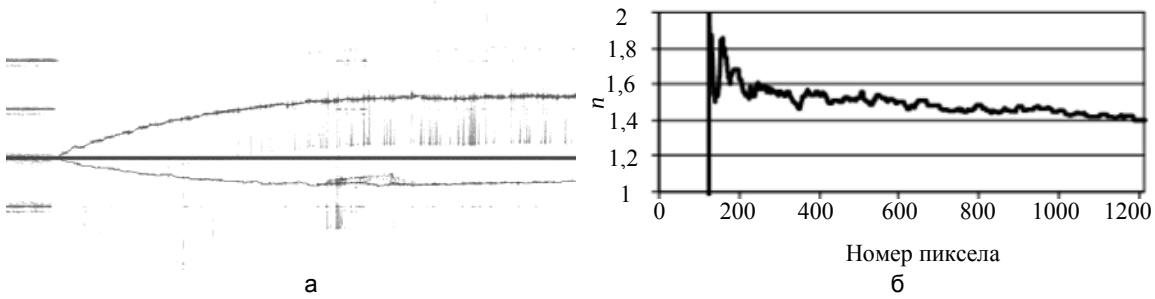


Рис. 9. В-скан края фантома на подложке (а) и распределение эффективного показателя преломления (*n*) вдоль В-скана (б). Ширина В-скана – 5 мм, глубина – 2 мм

Аналогичная методика позволяет получить оценку среднего эффективного показателя преломления для всего фантома. Использование метода ОКТ позволяет определять различия показателя преломления с погрешностью не более долей процента. Такая точность достаточна при исследовании свойств биотканей.

Заключение

Тканеимитирующие фантомы и капилляры являются интересным объектом для моделирования живых систем с возможностью их исследования методом ОКТ. Наиболее перспективными материалами представляются ПВХ-пластизоль и силикон. В настоящей работе проанализированы преимущества и недостатки использования этих материалов. На основе результатов исследований можно сделать вывод о том, что материал следует выбирать в зависимости от вида изготавливаемого фантома: для изготовления планарных структур необходимо использовать ПВХ, а для цилиндрических (эмуляция капилляров) – силикон. При визуализации капилляров целесообразно использовать интралипид или кровь в качестве контрастирующего агента. Дополнительным преимуществом использования метода ОКТ является возможность определения эффективного показателя преломления для каждой точки в плоскости образца.

Работа выполнена при финансовой поддержке Министерства образования и науки Российской Федерации.

Литература

1. Bini M., Ignesti A., Millanta L., Olmi R., Rubino N., Vanni R. The polyacrylamide as a phantom material for electromagnetic hyperthermia studies // IEEE Trans. Biomed. Eng. – 1984. – V. 31. – P. 317–322.
2. Surowiec A., Shrivastava P., Astrahan M., Petrovick Z. Utilization of a multilayer polyacrylamide phantom for evaluation of hyperthermia applicators // Int. J. Hyperthermia. – 1992. – V. 8. – P. 795–807.
3. Royston D., Poston R., Prahl S. Optical properties of scattering and absorbing materials used in the development of optical phantoms at 1064 nm // J. Biomed. Opt. – 1996. – V. 1. – P. 110–116.
4. Iizuka M., Sherar M., Vitkin I. Optical phantom materials for near infrared laser photocoagulation studies // Lasers Surg. Med. – 1999. – V. 25. – P. 159–169.
5. Kharine A., Manohar S., Seeton R., Kolkman R., Bolt R., Steenbergen W., de Mul F. Poly(vinylalcohol) gels for use as tissue phantoms in photoacoustic mammography // Phys. Med. Biol. – 2003. – V. 48. – P. 357–370.
6. Iizuka M., Sherar M., Vitkin I. Optical phantom materials for near infrared laser photocoagulation studies // Lasers Surg. Med. – 1999. – V. 25. – P. 159–169.
7. De Korte C., Cespedes E., van der Steen A., Norder B., te Nijenhuis K. Elastic and acoustic properties of vessel mimicking material for elasticity imaging // Ultrason. Imaging. – 1997. – V. 19. – P. 112–126.
8. Tomlins P., Woolliams P., Tedaldi M., Beaumont A., Hart C. Measurement of the 3D point-spread function in an OCT imaging system // Proc. SPIE. – 2008. – V. 6847. – P. 68472Q-1.
9. Волынский М.А., Гуров И.П., Воробьева Е.А., Маргарянц Н.Б. Бесконтактный контроль микрообъектов методами интерферометрии малой когерентности и оптической когерентной томографии // Изв. Вузов. Приборостроение. – 2011. – Т. 54. – № 2. – С. 7–14.
10. Гуров И.П., Волынский М.А., Жукова Е.В., Маргарянц Н.Б. Исследование растительных тканей методом оптической когерентной микроскопии // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – № 5 (81). – С. 42–47.

11. Bykov A.V., Popov A.P, Kinnunen M., Prykäri T., Priezzhev A.V., Myllylä R. Skin phantoms with realistic vessel structure for OCT measurements // Proc. SPIE. – 2010. – V. 7376. – P. 73760F-1–73760F-6.
12. Bykov A.V., Popov A.P., Priezzhev A.V., Myllylä R. Multilayer tissue phantoms with embedded capillary system for OCT and DOCT imaging // Proc. SPIE. – 2011. – V. 8091. – P. 80911R-1–80911R-6.

Быков Александр Викторович – Университет Оулу, кандидат физ.-мат. наук, исследователь, bykov@ee.oulu.fi

Волков Михаил Владимирович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат технических наук, доцент, ph-m.volkov@yandex.ru

Волынский Максим Александрович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат технических наук, доцент, maxim.volynsky@gmail.com

Гуров Игорь Петрович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, профессор, зав. кафедрой, gurov@mail.ifmo.ru

Киннунен Матти – Университет Оулу, Финляндия, D.Sc. (Tech.), исследователь, matti.kinnunen@ee.oulu.fi

Маргарянц Никита Борисович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, старший преподаватель, fosp@grv.ifmo.ru

Попов Алексей Петрович – Университет Оулу, кандидат физ.-мат. наук, исследователь, bykov@ee.oulu.fi

УДК 53.086

ФОРМИРОВАНИЕ ГИПЕРСПЕКТРАЛЬНЫХ ДАННЫХ О МИКРООБЪЕКТАХ С ВЫСОКОЙ СТЕПЕНЬЮ ПРОСТРАНСТВЕННОЙ НЕРАВНОМЕРНОСТИ РАСПРЕДЕЛЕНИЯ ОПТИЧЕСКОЙ ПЛОТНОСТИ

И.П. Гуров, А.И. Лопатин, А.В. Мельников

Предложен метод формирования гиперспектральных данных о микрообъектах, основанный на освещении наблюдаемого образца в микроскопе последовательно на различных длинах волн и подборе времени накопления фотоэлектрического заряда матричного фотоприемника для каждого участка изображения при изменении длины волны, что позволяет получать информацию без искажений на участках с малым и сильным отражением и сформировать результирующее изображение с улучшенным контрастом.

Ключевые слова: гиперспектральные приборы, микроскопия, спектрофотометрия.

Введение

При наблюдении различных объектов методами микроскопии степень отражения излучения на отдельных локальных участках зависит от свойств конкретного объекта. В случае равномерного распределения оптической плотности микрообъекта на наблюдаемом участке при формировании гиперспектральных данных, когда объект освещается последовательно на различных длинах волн, необходимо учитывать средний уровень облученности на каждой длине волны, приходящийся на все элементы видеоматрицы, чтобы в зависимости от этого уровня выполнить необходимую процедуру подстройки времени накопления видеоматрицы (выдержки) или яркости источника излучения [1].

При наличии у микрообъекта высокой пространственной неравномерности распределения оптической плотности может возникнуть ситуация, при которой отдельные участки изображения микрообъекта могут быть передержаны, а другие недодержаны. В этой ситуации невозможно сформировать гиперспектральные данные по всему полю изображения микрообъекта без потери полезной информации. Подстройка времени накопления фотоэлектрического заряда видеоматрицы или яркости источника излучения не могут устранить этот недостаток, поскольку приводят к изменению интенсивности всех участков изображения микрообъекта, тогда как для компенсации пространственной неравномерности оптической плотности микрообъекта необходимо подстроить интенсивность только на определенном локальном участке для каждой длины волны.

Управление уровнем облученности отдельных участков может быть выполнено различными методами. Один из известных методов заключается в размещении перед видеоматрицей нейтрального светофильтра (управляемого транспаранта) с изменяемым локальным коэффициентом пропускания, величина которого может подстраиваться для каждого участка светофильтра. Однако применение данного метода заметно усложняет конструкцию гиперспектрального микроскопа, требует использования дополнительных дорогостоящих узлов и специализированного программного обеспечения для управления параметрами светофильтра.

В работах [2, 3] рассмотрены методы компенсации неравномерного распределения облученности объекта, которые основаны на получении набора изображений, записанных с различными значениями выдержки. Полученный набор изображений позволяет синтезировать результирующее изображение без участков, интенсивность которых выходит за пороговые значения. Процедура синтеза изображения заключается в последовательном определении наиболее подходящего значения выдержки при регистрации каждого участка изображения. В результате реализуется регистрация отдельных участков изображения с различными значениями выдержки и компенсируется неравномерное распределение облученности объекта.

В настоящей работе предложен метод компенсации неравномерности пространственного распределения оптической плотности наблюдаемого микрообъекта в гиперспектральной микроскопии при записи набора видеокадров с различной выдержкой для каждой длины волны освещения и формирования гиперспектральных данных без потери полезной информации.

Формирование гиперспектральных данных

При наблюдении микрообъекта гиперспектральные данные получают в форме набора изображений, записанных с различными значениями выдержки, на каждой длине волны λ. Для каждого участка изображения микрообъекта подбирается изображение из полученного набора, на котором интенсивность не выходит за пределы пороговых значений. Зависимость значения выдержки, которое является оптимальным для регистрации участка изображения микрообъекта, от дискретных координат x и y обозначим как $t_{exp}(\lambda, x, y)$.

Отдельный элемент видеоматрицы позволяет регистрировать энергетическую экспозицию, которая является произведением облученности соответствующего элемента видеоматрицы и времени накопления электрического заряда (выдержки) [4]:

$$H(\lambda, x, y) = E(\lambda, x, y)t_{exp},$$

где $H(\lambda, x, y)$ – экспозиция; $E(\lambda, x, y)$ – облученность отдельного элемента видеоматрицы; t_{exp} – время накопления фотоэлектрического заряда.

Для каждого элемента изображения рассчитывается отношение значений экспозиции, полученных при регистрации изображения микрообъекта, и базовой линии [5], соответствующей априорно известным значениям спектра источника излучения. Если изображение микрообъекта и базовая линия получены с разными значениями выдержки, то полученное значение необходимо умножить на поправочный коэффициент, равный отношению значений выдержки:

$$T(\lambda, x, y) = \frac{H(\lambda, x, y)}{H_0(\lambda, x, y)} \times \frac{t_{0exp}}{t_{exp}}, \quad (1)$$

где $H_0(\lambda, x, y)$ – экспозиция при регистрации базовой линии; t_{exp} – значение выдержки при регистрации изображения микрообъекта; t_{0exp} – значение выдержки при регистрации базовой линии.

Поскольку видеоматрица при отсутствии падающего на нее излучения регистрирует уровень темнового тока, который необходимо учитывать при расчете [6], соотношение (1) принимает вид

$$T(\lambda, x, y) = \frac{H(\lambda, x, y) - H_{dc}(x, y)}{H_0(\lambda, x, y) - H_{dc}(x, y)} \times \frac{t_{0\exp}}{t_{exp}}, \quad (2)$$

где $H_{dc}(x, y)$ – уровень темнового тока элемента видеоматрицы с дискретными координатами x и y.

Если диапазона значений выдержки недостаточно, чтобы полностью компенсировать неравномерность пространственного распределения оптической плотности микрообъекта, то осуществляется подстройка яркости источника излучения при формировании набора изображений на каждой длине волны, например, изменением цветовой температуры источника [1]. При регистрации изображений, записанных с разной яркостью источника излучения в зависимости от длины волны, необходимо учитывать дополнительный поправочный коэффициент, равный отношению яркостей источника излучения при записи изображений микрообъекта и базовой линии. В этом случае выражение (2) имеет вид

$$T(\lambda, x, y) = \frac{H(\lambda, x, y) - H_{dc}(x, y)}{H_0(\lambda, x, y) - H_{dc}(x, y)} \times \frac{t_{0exp}}{t_{exp}} \times \frac{L_0(\lambda)}{L(\lambda)},$$

где $L(\lambda)$ – яркость источника излучения при записи изображения микрообъекта; $L_0(\lambda)$ – яркость источника при записи базовой линии. Результаты расчета поправочного коэффициента для галогенной лампы представлены в работе [1].

Подбор оптимального значения выдержки для каждого элемента изображения микрообъекта проводится таким образом, чтобы интенсивность находилась наиболее близко к середине динамического

диапазона видеоматрицы. На рис. 1 показан график зависимости интенсивности в отдельной точке изображения от значения выдержки видеоматрицы.

Рис. 1. Зависимость интенсивности в отдельной точке изображения от значения выдержки

В данном примере при значении выдержки, равном 1 мс, интенсивность наиболее близко расположена к середине динамического диапазона видеоматрицы.

Экспериментальные результаты

Для апробации предложенного метода формирования гиперспектральных данных о микрообъектах с высокой степенью пространственной неравномерности распределения оптической плотности был создан экспериментальный образец микроскопа-гиперспектрофотометра, блок-схема которого показана на рис. 2.

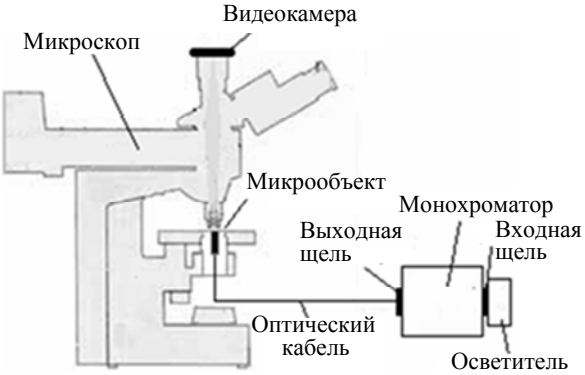


Рис. 2. Блок-схема экспериментального образца микроскопа-гиперспектрофотометра

Микроскоп-гиперспектрофотометр построен на основе микроскопа Микмед-6 (вариант 7), освещаемого на различных длинах волн с помощью монохроматора МДР 206 с дифракционной решеткой 1200 штр/мм и осветителем с галогенной лампой. Для регистрации изображений использована черно-белая видеокамера VAC-135 форматом 1280×1024 элементов. Для формирования изображений микрообъектов в плоскости видеокамеры использовался адаптер с линейным полем зрения 1/2″ и увеличением $1^{\times}$. Микрообъективы имели увеличение и числовые апертуры $10^{\times}$/0,25 и $40^{\times}$/0,65 соответственно. Экспериментальный образец микроскопа-гиперспектрофотометра формирует гиперспектральные данные при последовательном изменении длины волны излучения, освещающего микрообъект.

На рис. 3 показан пример набора изображений микрообъекта *Ulterus quer mit Larven*, сформированных с помощью микрообъектива с увеличением $40^{\times}$ при освещении микрообъекта на длине волны 550 нм и зарегистрированных с различным временем выдержки. Из рис. 3 видно, что для отображения структуры микрообъекта без потери информации по всему полю зрения некоторые участки изображения должны быть получены при малых выдержках, а другие участки изображения – при больших выдержках.

На рис. 4 показаны два изображения микрообъекта, полученные с микрообъективом $40^{\times}$ при двух значениях выдержки, 0,5 мс и 2,0 мс, на длине волны 600 нм. С учетом изображения базовой линии по формуле (1) были рассчитаны коэффициенты пропускания для локального участка 1 (T_1 = 19,0%) с использованием изображения, полученного с выдержкой 0,5 мс, а для локального участка 2 (T_2 = 1,4%), приходящемся на оптически более плотную область микрообъекта, – с использованием второго изображения, полученного с выдержкой 2,0 мс. Из рис. 4 видно, что детали микроструктуры на локальном участке 1 ясно различимы на рис. 4, а, тогда как изображение на рис. 4, б, содержит информацию о структу-

ре микрообъекта на локальном участке 2. Таким образом, совместное использование набора изображений, сформированных и зарегистрированных в соответствии с предлагаемым методом, позволяет получить полную информацию о структуре микрообъекта по всему полю зрения в различных длинах волн.

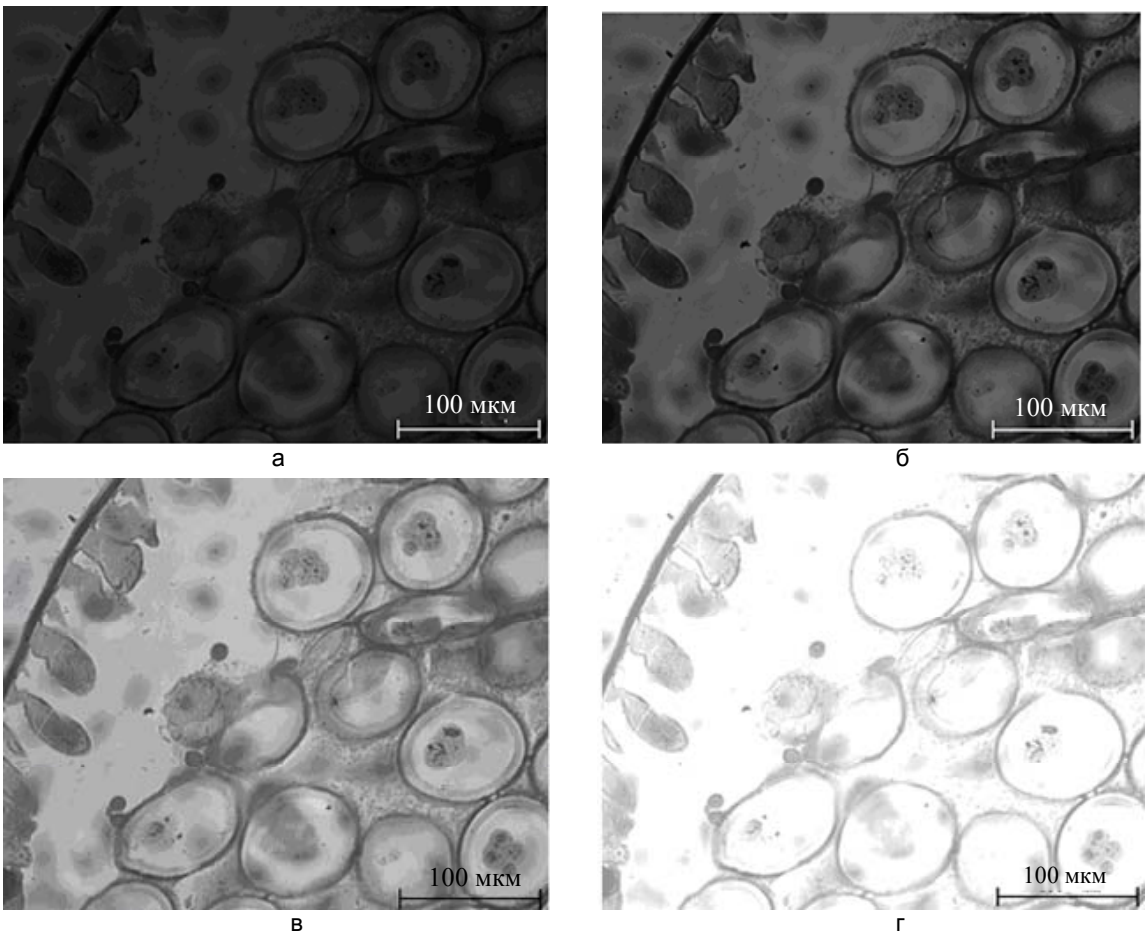


Рис. 3. Набор изображений, полученный на длине волны 550 нм с микрообъективом 40$^{\times}$ и выдержкой 0,2 мс (а); 0,8 мс (б); 1,4 мс (в); 2,0 мс (г)

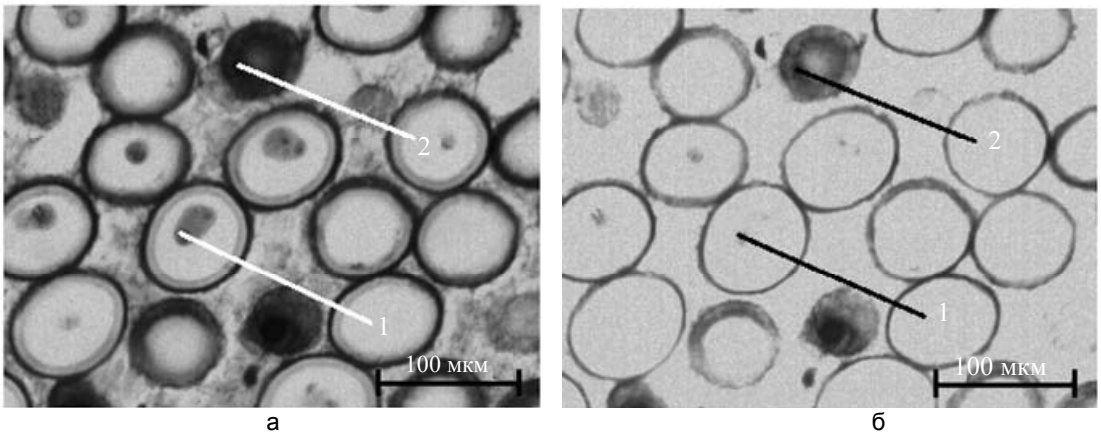


Рис. 4. Изображения, полученные на длине волны 600 нм с микрообъективом 40$^{\times}$ и выдержкой 0,5 мс (а); 2,0 мс (б). Коэффициенты пропускания локальных участков 1 и 2 равны T_1 = 19,0% и T_2 = 1,4%

Заключение

Экспериментальная апробация предложенного метода формирования гиперспектральных данных о микрообъектах с высокой степенью пространственной неравномерности распределения оптической плотности показала возможность его применения для исследования спектральных свойств локальных особенностей медицинских и биологических препаратов.

Метод показал высокую эффективность при регистрации спектра пропускания различных участков микроструктуры клеток, который имеет большой диапазон изменения оптической плотности. Разработанные алгоритмы и компьютерные программы позволяют использовать микроскоп-гиперспектрофотометр во многих областях биомедицинских исследований.

Работа выполнена при поддержке Министерства образования и науки Российской Федерации.

Литература

1. Гуров И.П., Лопатин А.И., Мельников А.В. Метод компенсации спектральной неоднородности источника излучения подстройкой его цветовой температуры для гиперспектральных приложений в микроскопии // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – № 5 (81). – С. 38–42.
2. Nayar S.K., Branzoi V. Adaptive dynamic range imaging: optical control of pixel exposure over space and time // Proc. the Ninth IEEE International Conference on Computer Vision (ICCV). – 2003. – V. 2. – P. 1168–1175.
3. Nuske S., Roberts J., Wyeth G. Extending the dynamic range of robotic vision // Proc. IEEE Int. Conference on Robotics and Automation. – 2006. – P. 162–167.
4. Понс Ж., Форсайт Д. Компьютерное зрение: Пер. с англ. – М.: Вильямс, 2004. – 928 с.
5. Белашенков Н.Р., Гуров И.П., Лопатин А.И., Мельников А.В. Микроскоп-спектрофотометр с матричным фотоприемником // Научно-технический вестник СПбГУ ИТМО. – 2007. – № 9 (43). – С. 260–265.
6. Бобровников Л.З. Электроника. – СПб: Питер, 2004. – 560 с.

Гуров Игорь Петрович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, профессор, зав. кафедрой, gurov@mail.ifmo.ru

Лопатин Александр Иосифович – ОАО «ЛОМО», кандидат физ.-мат. наук, главный оптик, ailopatin@gmail.com

Мельников Алексей Викторович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, инженер-исследователь, melnikov.alexey@gmail.com

УДК 53.05: 519.219: 519.714.3

ИДЕНТИФИКАЦИЯ ПАРАМЕТРОВ ВНУТРЕННЕЙ МИКРОСТРУКТУРЫ БИОТКАНЕЙ С ИСПОЛЬЗОВАНИЕМ ФОРМАЛИЗМА СТОХАСТИЧЕСКИХ РАЗНОСТНЫХ УРАВНЕНИЙ

Е.А. Воробьева, И.П. Гуров, А.Х. Киракозов

Разработан метод моделирования многослойных случайно-неоднородных сред на основе использования математического аппарата стохастических разностных уравнений. Предложен метод описания случайной границы среды. Приведены примеры представления многослойных случайно-неоднородных сред с различными параметрами их внутренней микроструктуры.

Ключевые слова: микроструктура, стохастические дифференциальные уравнения, многослойные среды.

Введение

Определение внутренней микроструктуры случайно-неоднородных сред неразрушающими оптическими методами представляет важное направление научных исследований и высоких технологий. Перспективным методом исследований является оптическая когерентная томография (ОКТ), обеспечивающая наиболее высокое разрешение при неразрушающем контроле микрообъектов [1–3].

Количественный анализ свойств микроструктуры объектов возможен при учете физических особенностей взаимодействия оптического излучения с веществом и использовании адекватного математического описания микроструктуры исследуемых объектов. Существуют различные подходы к исследованию случайно-неоднородных сред, например, метод диаграмм, метод интегралов по траекториям, метод уравнения переноса и др. [4–5]. Метод диаграмм связан с использованием аппарата квантовой теории поля, а именно, с построением диаграмм Фейнмана [6].

Метод интегралов по траекториям, впервые предложенный Р. Фейнманом, получил широкое распространение [6–8]. Сущность метода заключается в том, что излучение рассматривается как поток фотонов, проходящих через среду по всевозможным траекториям, рассеиваясь на неоднородностях [9]. Интегрирование по всем траекториям (суммирование вкладов по всем траекториям) позволяет описывать распространение света в случайно-неоднородной среде [4]. Виды реализации метода интегралов по траекториям можно разделить на два класса: аналитические и стохастические. Модели рассеяния на броуновских частицах и на потоках частиц составляют основу аналитических методов [9–10], в которых ис-

пользуются некоторые приближения для решения задачи. Метод Монте-Карло является стохастическим методом моделирования рассеяния в случайно-неоднородной среде [11–12].

Одним из наиболее плодотворных методов исследования случайно-неоднородных сред является метод уравнения переноса [4, 13]. Поскольку на сегодняшний день не найдено решение уравнения переноса, корректно описывающее рассеяние всех порядков, стохастическое описание является, по существу, единственным подходом, позволяющим предсказывать результаты экспериментов в случаях, когда важную роль играют как рассеяние низких порядков, так и многократное рассеяние.

Модель случайно-неоднородной среды можно построить также на основе формализма стохастических дифференциальных уравнений [14]. При этом важно учитывать априорную информацию об исследуемой среде, например, о наличии слоистой структуры с неровными границами и т.п.

В настоящей работе рассматриваются особенности описания случайно-неоднородных сред на основе формализма стохастических дифференциальных уравнений первого порядка (в форме уравнения Ланжевена) и демонстрируется возможность определения параметров микроструктуры исследуемой среды.

Стохастические дифференциальные уравнения для описания случайно-неоднородных сред

Рассмотрение свойств случайно-неоднородной среды дает основание для использования стохастических математических моделей для моделирования ее внутренней микроструктуры. Рассмотрим стохастическое дифференциальное уравнение Ланжевена

$$\frac{d\theta}{dt} = -\alpha\theta + \alpha n(t)\,, \tag{1}$$

где $n(t)$ – случайный гауссовский процесс с нулевым средним значением и равномерной спектральной плотностью мощности (белый гауссов шум); α – константа.

Уравнение (1) представляет собой альтернативное описание случайных реализаций величины $\theta(t)$ при эволюции плотности вероятности $p(\theta,t)$, определяемой уравнением Фоккера–Планка. В свою очередь, уравнение Фоккера–Планка широко используется для описания многих процессов в физике и химии [15], а также при изучении процессов самоорганизации в сложных системах, включая биологические системы [16]. Кроме этого, известно, что решение уравнения (1) обладает свойством фрактальности, которое характерно для разнообразных процессов в естественной природе. Таким образом, уравнение вида (1) применимо для описания пространственного распределения степени отражения излучения в случайно-неоднородных средах биологической природы ввиду клеточного механизма формирования биотканей. При этом распределение коэффициента отражения в среде по глубине определяется уравнением (1) в форме

$$\frac{d\theta}{dz} = -\alpha\theta + \alpha n(z)\,, \tag{2}$$

где z – координата по глубине среды.

Уравнение (2) удобно для использования ввиду простоты при моделировании и возможности варьировать спектральные свойства получаемых реализаций $\theta(z)$. За счет выбора подходящего значения α можно установить характерный масштаб неоднородностей: чем меньше величина этого параметра, тем более крупным является характерный масштаб неоднородностей моделируемой среды.

Большинство биотканей имеет слоистую структуру [17, 18]. По этой причине для описания реальных биотканей уравнение (2) требуется преобразовать к нестационарному виду, когда $\alpha = \alpha(\theta, z)$ или формирующий шум $n(z)$ характеризуется изменяющимися параметрами – переменной дисперсией или шириной спектра, т.е. является «окрашенным» шумом.

В методах ОКТ, как отмечалось выше, получают значение коэффициента (однократного) отражения по глубине среды вдоль координаты z (так называемые А-сканы), совокупность которых составляет томограмму (В-скан). Поскольку томограмма является двумерным представлением микроструктуры в сечении исследуемого объекта, при описании распределения коэффициента отражения среды нужно учитывать коррелированность характеристик в соседних А-сканах.

Дискретная двумерная модель томограммы показана на рис. 1. Модель представляет собой двумерную сетку, состоящую из $N \times M$ ячеек, при этом каждой ячейке соответствует инверсное значение коэффициента отражения: чем оно больше, тем ярче выглядит ячейка на рисунке. Значение коэффициента отражения в ячейке с координатами (i, j) обозначим как $\theta(i, j)$.

Используя уравнение (2), можно записать стохастическое разностное уравнение для описания изменения коэффициента отражения внутри среды с учетом коррелированности значений в соседних точках в виде

$$\theta(i, j) = (1-a)\left(a\theta(i-1, j\right) + b\theta(i+1, j) + c\theta(i, j+1) + d\theta(i, j-1) + w(i, j)\,, \tag{3}$$

где a, b, c, d – коэффициенты, сумма которых должна быть равна единице. Таким образом, получаем $N \times M$ уравнений и $N \times M$ неизвестных:

$$\begin{pmatrix} 1 & c & b & 0 & . \\ a & 1 & c & b & . \\ 0 & a & 1 & c & . \\ 0 & 0 & a & 1 & . \\ . & . & . & . & . \end{pmatrix} \theta = w, \qquad (4)$$

где θ – столбец неизвестных $\theta(i, j)$; w – столбец случайных величин $w(i, j)$. При решении системы уравнений (4) полученные значения будут полностью удовлетворять уравнению (3).

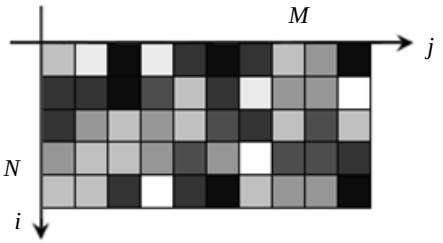
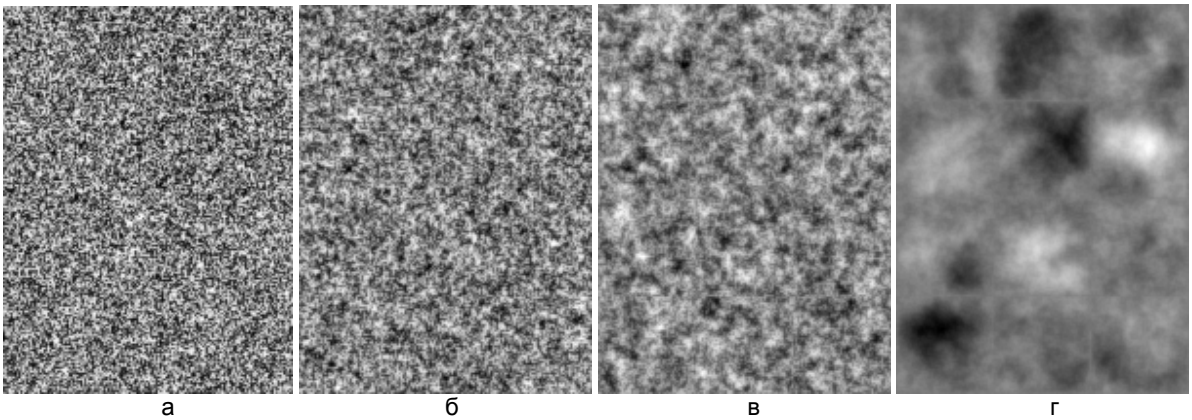


Рис. 1. Дискретная двумерная модель В-скана

Рис. 2 иллюстрирует пространственные распределения коэффициента отражения в среде при a=b=c=d=0,25 и различных значениях α. Значение параметра α влияет на скорость изменения характеристик среды, что позволяет определить среды с различными масштабами локальных неоднородностей.



а б в г

Рис. 2. Примеры пространственного распределения коэффициента отражения в среде для α=0,5 (а); α=0,2 (б); α=0,05 (в); α=0,01 (г) (размер 150×200 пикселей)

Заметим, что параметр α определяет граничную частоту диапазона частот спектральной плотности мощности случайного процесса, определяемого уравнением (2) (см. [19]), что позволяет интерпретировать свойства случайно-неоднородных сред с позиций пространственно-частотного подхода.

Описание границы случайно-неоднородной среды

В описанном выше алгоритме предполагается, что моделируемая среда имеет ровную границу. Однако получаемые ОКТ-изображения реальных биологических сред практически никогда не имеют ровной внешней границы.

Для описания случайной границы целесообразно использовать решение уравнение Ланжевена (2), в котором в качестве случайной реализации θ рассматривается начальная координата границы x_0. Рассмотрим алгоритм определения случайной границы более подробно.

Для формирования случайной границы необходимо задать функцию $x_0(j)$, которая для каждого j-го столбца определяет номер ячейки, относящейся к границе среды для этого столбца [14]. Для уравнения Ланжевена (2) известно точное решение: если $\theta_0 = \theta(z = 0)$, то

$$\theta(z) = \theta_0 e^{-\alpha z} + \alpha \int_0^z n(\xi) e^{-\alpha(z-\xi)} d\xi . \qquad (5)$$

Перепишем уравнение (5) для координаты границы z_j:

$$\psi(z_j) = \psi(\alpha, \sigma, z_{j-1}, \Delta z) = z_{j-1} e^{-\alpha z} + \alpha \int_0^z n(\xi) e^{-\alpha(z-\xi)} d\xi \,, \tag{6}$$

где функция $\psi(z_j)$ представляет решение уравнения (1) при заданных параметрах (α, σ), σ – среднеквадратичное отклонение, характеризующее формирующий гауссов шум $n(z)$.

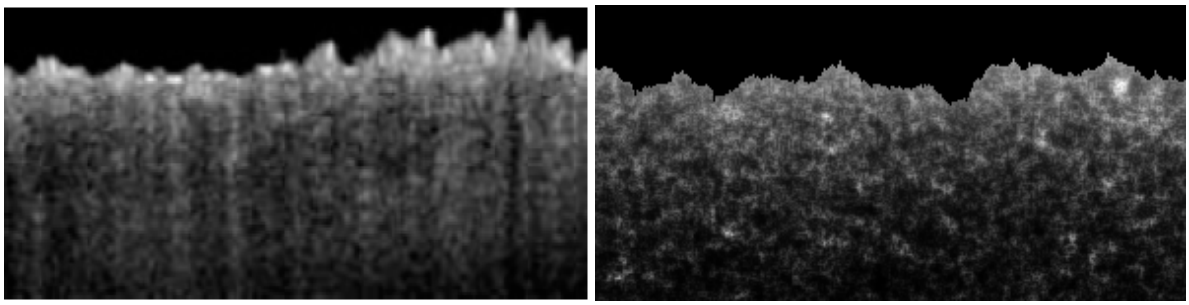
Алгоритм вычислений состоит из двух этапов. Сначала выбираются некоторые значения для параметров $\alpha, \sigma, z_0, \Delta z$, и для каждого *j*-го столбца по формуле (6) рекурсивно вычисляются значения

$z_j = \psi'(\alpha, \sigma, z_{j-1}, \Delta z),\ j = 1, ..., M$.

Затем функция $z_0(j)$ вычисляется по следующей формуле [14]:

$$z_0(j) = \mathrm{Int}\left(1 + \left(\max\left\{z_0(j)\right\} - 1\right) \frac{z_j - \min_j z_j}{\max_j z_j - \min_j z_j}\right), \tag{7}$$

где $\mathrm{Int}(\cdot)$ обозначает функцию округления до ближайшего целого числа, $\max\left\{z_0(j)\right\}$ – заданное максимальное значение функции $z_0(j)$. Формула (7), по существу, представляет операцию масштабирования значений z_j на отрезок $\left[1,\ \max\left\{z_0(j)\right\}\right]$.



а б

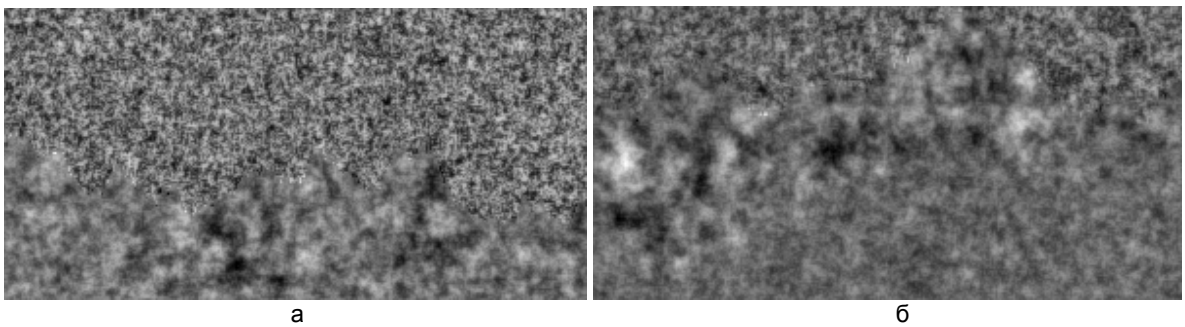
Рис. 3. Томограмма свиной кожи (а) и результат синтеза В-скана для среды со случайной верхней границей для α=0,1 (б) (размер 300×150 пикселей)

На рис. 3, б, показан результат моделирования случайно-неоднородной среды со случайной верхней границей. Сравнение рис. 3, а, б, демонстрирует соответствие реальной томограммы биоткани и синтезированной томограммы.

Описание структуры многослойной случайно-неоднородной среды

Реальные биологические ткани в большинстве случаев состоят из нескольких слоев, отличающихся свойствами их микроструктуры. Для многослойной среды необходимо определить параметры для каждого отдельного слоя и для границ слоев [14].

На рис. 4 представлены примеры синтезированных многослойных случайно-неоднородных сред. Случайная граница между слоями задается так же, как и внешняя граница среды.



а б

Рис. 4. Фрагменты представления двухслойной (а) и трехслойной (б) случайно-неоднородных сред (размер 300×150 пикселей)

На рис. 4, а, показан фрагмент двухслойной среды при значениях параметров для обоих слоев *a*=*b*=*c*=*d*=0,25 и значениях α=0,2 для верхнего слоя, α=0,1 для нижнего слоя. На рис. 4, б, представлен фрагмент трехслойной среды с параметрами α, равными 0,1; 0,01; 0,05 для верхнего, среднего и нижнего слоев соответственно.

Рассмотренный алгоритм позволяет компактно описать внутреннюю микроструктуру случайно-неоднородной среды с произвольным числом слоев при различной степени выраженности границ между слоями. Представленные среды на рис. 4 имеют ярко выраженную границу между слоями. Для моделирования слабо выраженной границы было реализовано усреднение коэффициентов α_i и α_j на границе *i*-го и *j*-го слоев.

Алгоритм вычислений иллюстрируется блок-схемой на рис. 5. В качестве входных данных алгоритм получает описание каждого из слоев случайно-неоднородной среды и параметры для определения случайных границ. На первом этапе вычислений осуществляется инициализация параметров среды. Для каждого слоя формируется случайная граница и выполняется ее сглаживание при необходимости.

Второй этап вычислений заключается в разбиении среды на отдельные равные части. Время выполнения вычислений значений в каждой точке среды и требуемая память имеют квадратичную зависимость от размера обрабатываемой области, что ограничивает допустимые размеры области расчета.

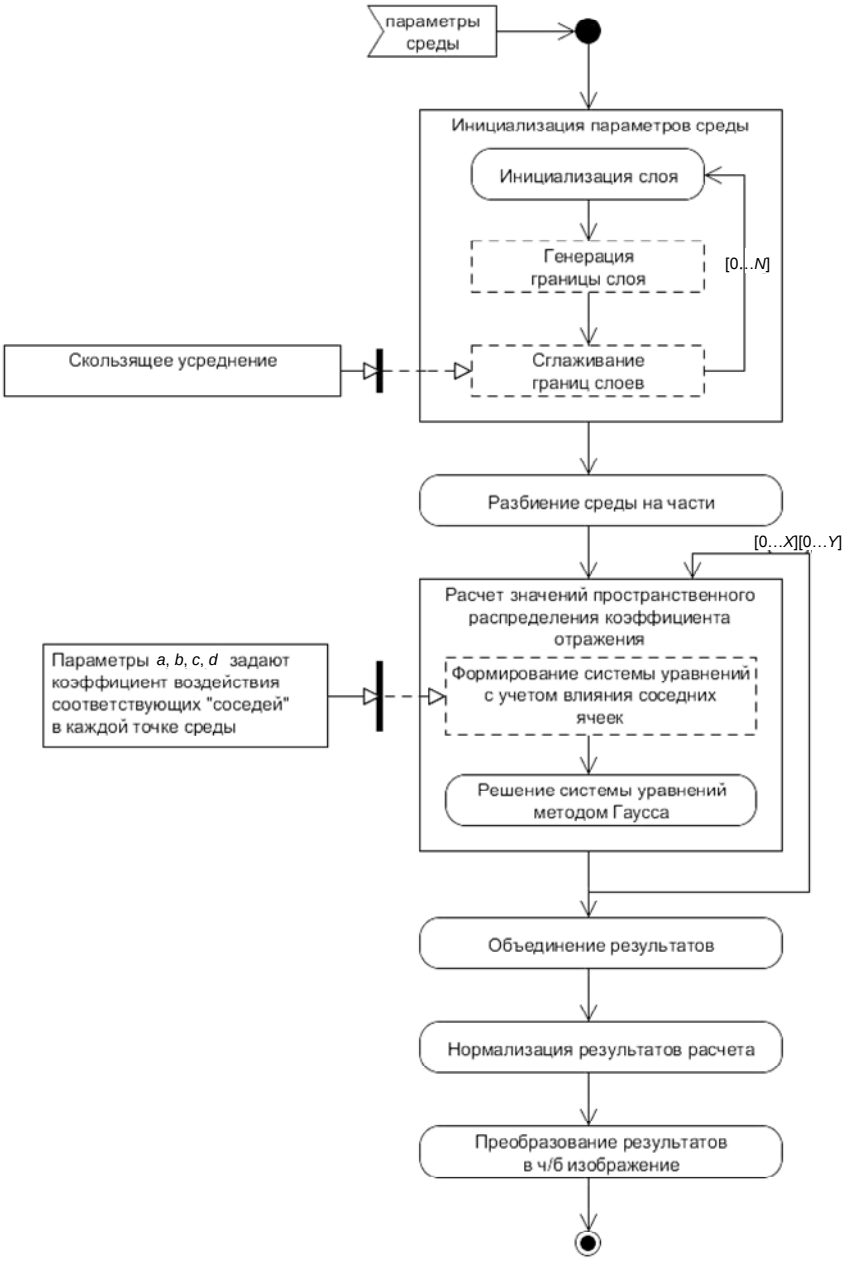


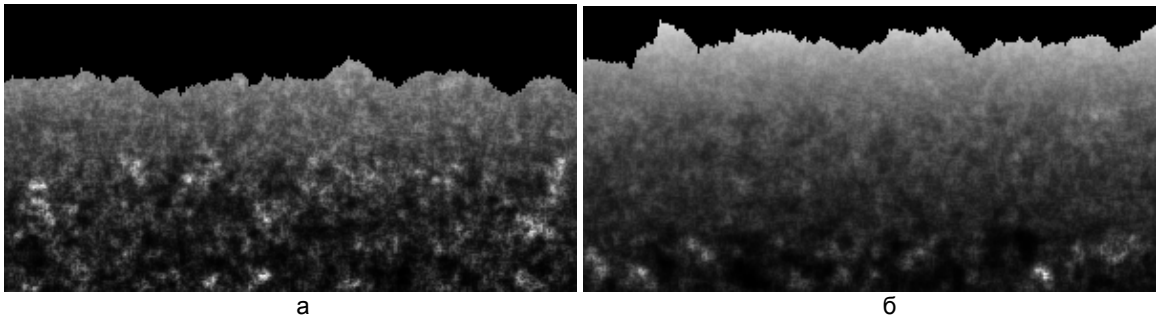
Рис. 5. Блок-схема представления многослойной случайно-неоднородной среды со случайными границами

Третий этап вычислений состоит в формировании системы уравнений с учетом влияния соседних ячеек с весовыми коэффициентами a, b, c, d и ее решении методом Гаусса. Этот этап выполняется отдельно для всех частей среды, и затем результаты вычислений объединяются в форме матрицы.

Завершающие шаги алгоритма состоят в нормализации полученных результатов и преобразовании в черно-белое изображение. Нормализация проводится для удобства преобразования полученных значений в полутоновый диапазон.

На рис. 6 представлены примеры многослойных сред со слабо выраженными границами между слоями, полученные с помощью описанного выше алгоритма вычислений.

Представленная на рис. 6, а, среда состоит из двух слоев, параметр α для верхнего слоя равен 0,13, а для нижнего 0,08. Среда, изображенная на рис. 6, б, состоит из трех слоев, для которых параметр α равен 0,85; 0,9; 0,95 для верхнего, среднего и нижнего слоев соответственно.



а б

Рис. 6. Смоделированные В-сканы для двухслойной (а) и трехслойной (б) случайно-неоднородных сред со слабо выраженными случайными границами между слоями (размер 300×150 пикселей)

При описании сложных многослойных сред требуется задавать в общем случае четыре «переключаемых» параметра для каждого слоя и для каждой границы слоев. Число параметров снижается при описании преимущественно однородных сред.

Заключение

Разработанный подход к определению микроструктуры биотканей на основе математического аппарата стохастических разностных уравнений обладает высокой гибкостью в зависимости от выбора параметров модели, что позволяет адекватно представить случайно-неоднородные среды с различной микроструктурой. Предложенный метод позволяет определять случайно-неоднородную среду ограниченным числом параметров для описания слоев и их границ. Приведены примеры описания однослойных и многослойных случайно-неоднородных сред со случайными границами, иллюстрирующие адекватность предложенного представления реальным средам.

Разработанный подход обеспечивает возможность создания тестовых моделей (виртуальных неоднородных сред) для верификации алгоритмов обработки информации в оптической когерентной томографии.

Работа выполнена при финансовой поддержке Министерства образования и науки Российской Федерации.

Литература

1. Гуров И.П. Оптическая когерентная томография: принципы проблемы и перспективы // Проблемы когерентной и нелинейной оптики / Под ред. И.П. Гурова и С.А. Козлова. – СПб: СПбГУ ИТМО, 2004. – С. 6–30.
2. Alarousu E., Gurov I., Kalinina N., Karpets A., Margariants N., Myllyla R., Prykari T., Vorobeva E. Full-field high-resolving optical coherence tomography system for evaluating paper materials // Advanced Laser Technologies 2007 // Proc. SPIE. – 2007. – V. 7022. – P. 7022-12.
3. Gurov I., Karpets A., Margariants N., Vorobeva E. Full-field high-speed optical coherence tomography system for evaluating multilayer and random tissues // O3A: Optics for Arts, Architecture, and Archaeology Proc. SPIE. – 2007. – V. 6618. – P. 6618-07.
4. Скипетров С.Е. Диффузионно-волновая спектроскопия в средах с пространственно неоднородной динамикой рассеивателей: Дисс. … канд. физ.-мат. наук: 01.04.21. – М., 1998. – 153 с.
5. Воробьева Е.А., Гуров И.П. Модели распространения и рассеяния оптического излучения в случайно-неоднородных средах // Проблемы когерентной и нелинейной оптики. Сборник научных статей / Под ред. И.П. Гурова, С.А. Козлова. – СПб: СПб ГИТМО (ТУ), 2006. – С. 82–98.
6. Фейнман Р., Лейтон Р., Сэндс М. Фейнмановские лекции по физике. Том 8. Квантовая механика – I // М.: Мир, 1966. – 267 с.
7. Фейнман Р., Хибс А. Квантовая механика и интегралы по траекториям. – М.: Мир, 1968. – 382 с.

8. Path integrals and their applications in quantum, statistical, and solid state physics / Eds. G.J. Papadopoulos, J.T. Devreese. – NY: Plenum press, 1978. – 515 p.
9. Maret G., Wolf P.E. Multiple light scattering from disordered media. The effect of brownian motion of scatterers // Z. Phys. B. – 1987. – V. 65. – № 4. – P. 409–413.
10. Wu X-L., Pine D.J., Chaikin P.M., Huang J.S., Weitz D.A. Diffusing-wave spectroscopy in a shear flow // J. Opt. Soc. Am. B. – 1990. – V. 7. – № 1. – P. 15–20.
11. Ярославский И.В., Тучин В.В. Распространение света в многослойных рассеивающих средах. Моделирование методом Монте-Карло // Оптика и спектроскопия. – 1992. – Т. 72. – № 4. – С. 934–939.
12. Feng S., Zeng F. Monte Carlo simulations of photon migration path distributions in multiple scattering media // Proc. SPIE. – 1991. – V. 1888. – P. 78–89.
13. Кейз К., Цвайфель П. Линейная теория переноса. – М.: Мир, 1972. – 385 с.
14. Воробьева Е.А., Киракозов А.Х. Идентификация стохастических моделей случайно-неоднородных сред в оптической когерентной томографии // Проблемы когерентной и нелинейной оптики. Сборник научных статей / Под ред. И.П. Гурова, С.А. Козлова. – СПб: СПб ГИТМО (ТУ), 2008. – С. 120–129.
15. Ван-Камепен Н.Г. Стохастические процессы в физике и химии. – М.: Высшая школа, 1990. – 376 с.
16. Хакен Г. Информация и самоорганизация. Макроскопический подход к сложным системам. – М.: Мир, 1991. – 240 с.
17. Руководство по оптической когерентной томографии / Под ред. Н.Д. Гладковой, Н.М. Шаховой и А.М. Сергеева. – М.: ФИЗМАТЛИТ, 2007. – 296 с.
18. Уманец А.В. Анализ видов тестовых образцов материалов в оптической когерентной томографии // Проблемы когерентной и нелинейной оптики. Сборник научных статей / Под ред. И.П. Гурова, С.А. Козлова. – СПб: СПб ГИТМО (ТУ), 2008. – С. 130–136.
19. Ахманов С.А., Дьяков Ю.А., Чиркин А.С. Введение в статистическую радиофизику и оптику. – М.: Наука, 1981. – 640 с.

Воробьева Елена Александровна – ООО «Моторола-Мобилити», инженер-программист, lenavorobyeva@gmail.com

Гуров Игорь Петрович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, профессор, зав. кафедрой, gurov@mail.ifmo.ru

Киракозов Александр Христофорович – ООО «Яндекс», руководитель группы, hristoforich@yandex.ru

6 МАТЕРИАЛОВЕДЕНИЕ И НАНОТЕХНОЛОГИИ

УДК 544.313.2.031

ОПРЕДЕЛЕНИЕ ХИМИЧЕСКИХ И ФАЗОВЫХ ПРЕВРАЩЕНИЙ ПРИ ИМПУЛЬСНОМ ЛАЗЕРНОМ ОБЛУЧЕНИИ МНОГОКОМПОНЕНТНЫХ СПЛАВОВ НА ВОЗДУХЕ МЕТОДАМИ ХИМИЧЕСКОЙ ТЕРМОДИНАМИКИ

В.П. Вейко, А.А. Слободов, Г.В. Одинцова

Предложен расчетный термодинамический метод определения фазово-химического состава поверхности металлов и сплавов, формируемых при лазерном воздействии в атмосфере, в зависимости от их компонентного состава, условий лазерного облучения и состава атмосферы. На примере лазерного нагревания на воздухе легированной стали показано, что из множества различных возможных реакций взаимодействия железа, никеля или хрома с компонентами воздуха (кислородом, азотом, углеродом, их соединениями, атмосферной влагой и др.) реализуются строго определенные. В первую очередь, это процессы окисления металлов с образованием оксидной пленки, фазовый и химический состав которой определяется температурой и временем нагревания. Результаты расчетов экспериментально подтверждены методом энерго-дисперсионной рентгеновской спектроскопии.

Ключевые слова: лазерные технологии, физическая химия, тонкие пленки, лазерное облучение металлов, термодинамика, расчет.

Введение

Известно, что при нагревании металлов и сплавов на воздухе за счет химического взаимодействия с атмосферой на их поверхности образуются тонкие цветные интерференционные пленки [1]. При лазерном воздействии этот эффект может быть локализован в пределах зоны облучения, что позволяет, например, создать технологию цветной лазерной маркировки [2–3]. Главной проблемой, препятствующей внедрению этой технологии в промышленность, является сложность воспроизведения цветов при смене облучаемых материалов, а также параметров и условий воздействия. Для реализации возможности «лазерного управления» интерференционными цветами требуется детальная информация о физико-химических процессах на поверхности, в том числе, определение типов протекающих реакций, а также химического и фазового состава полученных пленок.

Заметим, что задача выявления фазово-химических изменений поверхности при лазерном облучении выходит на первый план и во многих других приложениях – например, в методе лазерного формирования дифракционных оптических элементов на пленках хрома [4], при лазерном структурировании поверхности кремния для повышения его поглощательной способности с целью интенсификации теплового воздействия лазерного излучения [5] и пр.

Однако экспериментальные методы исследования, такие как рамановская спектроскопия, электронная или рентгеновская спектроскопия соответствующих вторичных излучений при электронной бомбардировке объекта, не только дороги и трудоемки, но и не всегда приводят к желаемому результату. Расшифровка рамановских спектров и идентификация химических соединений по ним, а также идентификация химических соединений по данным электронной или рентгеновской спектроскопии, особенно в случае лазерного облучения многокомпонентных сплавов на воздухе, – это весьма трудоемкая задача [6, 7].

Качественный анализ ситуации

Теоретическое рассмотрение этих процессов было проведено в ряде работ для гетерогенных процессов, когда их лимитирующей стадией является диффузия (например, [8]). В случае лазерного окисления хрома и некоторых других металлов (никель, алюминий и др.), когда образуется плотная оксидная пленка (на основе Cr_2O_3, NiO, Al_2O_3), пленка предохраняет нижележащие слои от непосредственного взаимодействия с атмосферой. Тогда, как известно, диффузия через слой окисла действительно является лимитирующей стадией гетерогенного процесса окисления, и эта теория достаточно хорошо описывает эксперименты при милли-, микро- и наносекундных длительностях лазерного воздействия. При более коротких воздействиях лимитирующей стадией процесса может стать диффузия атомов кислорода из газовой среды к поверхности. Однако, как можно показать, это ограничение становится существенным при переходе к еще более коротким длительностям воздействия (~10 нс и менее).

Если образующийся оксид – дефектный, пористый или недостаточно плотный и не предохраняет лежащие под ним слои металла от дальнейшего окисления, складывается иная ситуация. В этом случае процесс протекает в кинетической области, т.е. лимитирующей стадией, определяющей общую скорость гетерогенного окисления, является чисто химическая кинетика – электронные явления, которые гораздо менее инерционны, чем диффузионные, а эффективная скорость реакции существенно зависит от температуры согласно уравнению Аррениуса. При воздействии лазерных импульсов с длительностью ~100 нс

и частотами следования ~ 10^4–10^5 Гц на сплавы железа с углеродом, т.е. стали, как раз реализуется ситуация пористых окислов, в которых окисление не лимитируется процессами диффузии в слое окисла и в воздухе.

Анализ протекающих при нагревании металлов и сплавов фазовых и химических превращений на их поверхности и условий формирования интерференционных пленок позволяет заключить, что интегрально они, как правило, подчиняются законам термодинамики. Современный же уровень развития теории и методов химической термодинамики и практика их применения в физико-химических исследованиях позволяют решать задачи определения искомых продуктов взаимодействия для разнообразных, в том числе и достаточно сложных многокомпонентных гетерогенных систем и процессов.

При лазерных воздействиях областью возможного приложения термодинамических методов могут быть непрерывные или импульсно-частотные процессы. Конечно, точность определения выхода реакции при этом может ограничиваться временными и пространственными колебаниями температуры в реакционном пространстве. Наиболее благоприятными с этой точки зрения являются режимы с большой частотой следования импульсов и большим пространственным перекрытием облучаемых зон на поверхности, когда распределение температур можно считать квазистационарным.

Это обстоятельство должно существенно отодвигать границу нестационарности процессов лазерного нагревания в область кратковременных воздействий из-за того, что собственные времена химических взаимодействий весьма малы. Таким образом, рассматриваемая ситуация представляется весьма интересной для проверки применимости методов химической термодинамики и кинетики при импульсных лазерных воздействиях.

Цель настоящего исследования заключается в том, чтобы показать возможность и разработать методику применения термодинамических методов в процессах лазерной термохимии, а также подтвердить результаты расчета экспериментом.

Термодинамический метод исследования химических аспектов воздействия лазерного излучения на поверхность металлов и их сплавов. Исходные теоретические положения

Как отмечалось, корректным и потенциально весьма эффективным подходом к выявлению возможных и реализуемых взаимодействий и их количественному описанию является термодинамический метод. При изобарно-изотермических условиях протекающие процессы целесообразно описывать с помощью энергии Гиббса системы: $G \equiv G_{P,T}(x)$, где x – искомый фазово-химический состав системы [9].

Как известно [10, 11], физико-химические процессы в системе протекают в направлении уменьшения ее соответствующего термодинамического потенциала, в данном случае энергии Гиббса $G_{P,T}(x)$. При этом конечное, равновесное состояние соответствует достижению минимума этого потенциала, $G_{P,T}(x) \to \min\limits_{x}$, к нахождению условий реализации которого и сводится, по существу, рассматриваемая задача.

Требование, обеспечивающее корректность применения термодинамического подхода, заключается в том, чтобы скорости всех протекающих самопроизвольных процессов были существенно выше скорости (вынужденной, задаваемой внешним воздействием) изменения состояния системы.

Суть используемого расчета заключается в следующем. Рассчитывается энергия Гиббса G и ее изменение ΔG как для каждого из компонентов-участников термодинамического процесса (реакции), так и для системы в целом – обычно через изменения энтальпии H и энтропии S (ΔH и ΔS) [12]. В частности, для изотермического процесса имеем: $\Delta G_T = \Delta H_T - T\Delta S_T$, где T – неизменная температура окружающей среды; индекс «P» давления здесь опущен.

Тепловые эффекты процессов зависят, в первую очередь, от температуры, при которой они происходят, поэтому для их вычисления при произвольных температурах необходимо знать температурные зависимости стандартных изобарных теплоемкостей реагирующих веществ $C_j^0(P_0,t)$, а также температуры T_j и тепловые эффекты (энтальпии) фазовых (структурных, агрегатных) превращений $\Delta H_j^0(P_0,T_j)$.

Тогда расчет стандартной энергии Гиббса $\Delta G_{P,T}^0$ химической реакции (т.е. разность суммарных энергий Гиббса продуктов и реагентов) осуществляется следующим образом [13]:

$$\Delta G_T = \Delta_f H^0\left(P_0,T_0\right) - TS^0\left(P_0,T_0\right) + \int_{P_0}^{P} V^0\left(p,T\right)dp -$$

$$-\sum_{j=1}^{k}\left[T\int_{T_{j-1}}^{T_j}\frac{C_j^0\left(P_0,t\right)}{t}dt - \int_{T_{j-1}}^{T_j} C_j^0\left(P_0,t\right)dt + \Delta H_j^0\left(P_0,T_j\right)\left(\frac{T}{T_j}-1\right)\right], \quad (1)$$

где $\Delta_f H^0(P_0,T_0)$, $S_0(P_0,T_0)$ – соответственно энтальпия образования и энтропия вещества при стандартных (P_0 = 1 бар, T_0 = 298,15 К) условиях; $V^0(P,T)$ – барическая зависимость стандартного мольного объема на интервале $[P_0,P]$ при температуре *T*. Поскольку условием самопроизвольности изобарно-изотермического процесса является уменьшение энергии Гиббса системы в ходе процесса, $\Delta G^0_{P,T} < 0$, то о разрешенности реакции в пределах соответствующего температурного диапазона свидетельствует отрицательный знак изменения энергии Гиббса.

Помимо термической нестационарности процесса превращений и образования соответствующего продукта, имеется еще одна особенность – кинетическая (временная), обусловленная следующим: даже если процесс и продукт (т.е. искомое равновесное состояние исследуемой системы) для заданных условий определены термодинамически корректно, то это еще не дает ответ на вопрос, пройдет ли за время лазерного воздействия данный процесс до конца и, тем самым, успеет ли образоваться соответствующий продукт взаимодействия. Таким образом, встает задача определения времени, затрачиваемого на достижение искомого равновесного (или квазиравновесного) состояния, необходимого для получения ожидаемого продукта взаимодействия. Исходя из этого, целесообразно привлечение опытных данных в целях сопоставления и сравнительного анализа с расчетными данными по химическому и фазовому составу пленок, образующихся в результате нагрева на поверхности образца.

Термодинамический расчет процесса взаимодействия стали с атмосферой при лазерном нагревании

Процесс взаимодействия лазерного излучения с поверхностью нержавеющей стали с энергетической и химической точек зрения порождает комплекс фазовых и химических превращений. Если процесс лазерной обработки осуществляется на воздухе, то на поверхности нержавеющей стали марки 12Х18Н10Т (содержащей 71% Fe; 0,12% C; 18% Cr; 10% Ni, а также незначительное количество легирующих элементов) наиболее вероятно протекание реакций взаимодействия основных компонентов стали и воздуха между собой. Определим возможность протекания этих реакций путем расчета энергии Гиббса $\Delta G^0_{P,T}$ по выражению (1). Исходные термодинамические данные для расчета взяты из [13].

На рис. 1 приведены примеры рассчитанных зависимостей энергии Гиббса образования реакций от температуры, полученные на основе уравнения энергии Гиббса химической реакции (1). Для каждой реакции образования расчет энергии Гиббса осуществлялся в температурном диапазоне 298–1000 К.

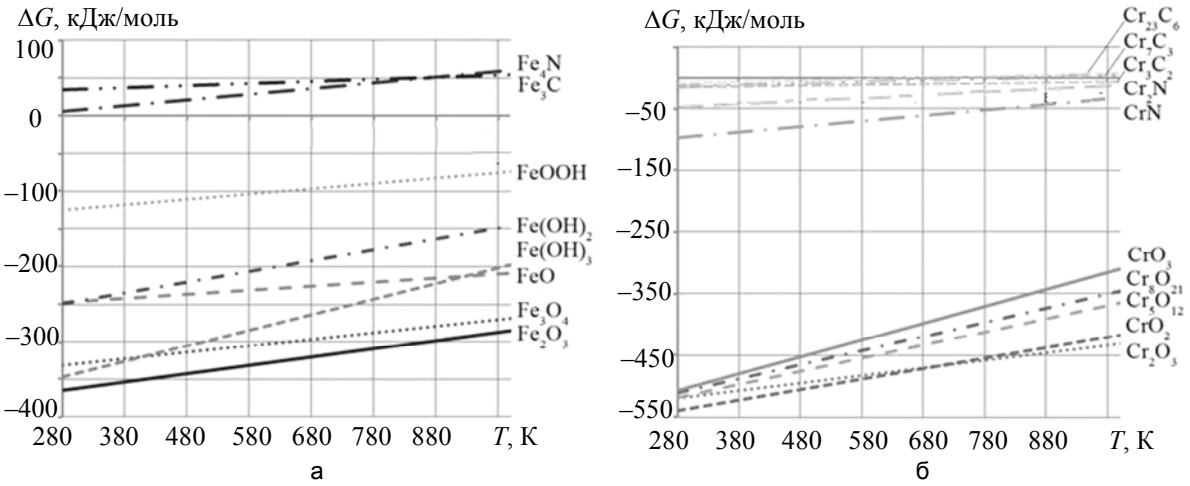


Рис. 1. Зависимость энергии Гиббса образования оксидов, гидроксидов, карбидов и нитридов компонентов стали от температуры: железо (а), хром (б)

Положительные значения энергии Гиббса ($\Delta G > 0$) на графиках рис. 1 свидетельствуют о «запрещенности» соответствующих реакций, в то время как отрицательные – об их «разрешенности», при этом чем меньше энергия Гиббса, тем более вероятна соответствующая реакция.

Экспериментальные результаты

Процессы лазерного воздействия изучались на образцах нержавеющей стали марки 12Х18Н10Т в виде шлифованных пластин толщиной 2 мм. Облучение образцов проводилось путем экспонирования поверхности размерами 5×5 см сканирующим пучком волоконного лазера на длине волны λ = 1,07 мкм со средней мощностью 1–10 Вт, длительностью импульсов τ = 100 нс, частотой следования *f* = (10–100) кГц при плотности мощности $q = (0{,}1–1{,}8)\cdot 10^8$ Вт/см^2. Целью облучения было формирование на поверхности

цветных пленок, которые могут быть использованы для технологии цветной лазерной маркировки металлических поверхностей. На рис. 2 представлены примеры изображений, полученных при помощи данной технологии. Для формирования элемента изображения применялось построчное сканирование поверхности образца пучком лазерного излучения со скоростью $V = (1–250)$ мм/с и шагом строк сканирования $N = (0–50)$ мкм. Облучение проводилось в нормальных лабораторных условиях в воздушной атмосфере.

а б в г

Рис. 2. Примеры изображений (а–г), полученных методом цветной лазерной маркировки поверхности нержавеющей стали марки 12Х18Н10Т толщиной 2 мм

Для определения элементного состава полученных пленок использовался метод энерго-дисперсионной рентгеновской спектроскопии, реализованный на базе растрового электронного микроскопа Inspect FEICompany. Данные таблицы показывают массовую долю каждого компонента из смеси ($\omega(X)$) для необработанной поверхности нержавеющей стали и на цветных пленках после лазерного нагрева до определенной температуры.

Температура образцов в процессе облучения определялась с помощью инфракрасной камеры FLIR Titanium 520 M, время интегрирования – от 3 мкс, максимальная частота обновления полных кадров – от 380 Гц до 20 кГц, точность измерений ±1%, коэффициент серости 0,87±0,13 (получен путем калибровки измерений ИК-камеры с помощью термопар). Измеренные значения температур, при которых формировались соответствующие цветные пленки, указаны в таблице.

Цвет стали после лазерного нагрева	Температура образца при лазерном нагреве, К	$\omega(X)$, %						
		Fe	Cr	Ni	O	Ti	Si	Mn
исходный состав (до лазерного воздействия)	293	70,76	17,53	8,44		0,49	0,89	1,89
Голубой	604–609	64,13	15,26	7,74	10,3	0,41	0,75	1,41
Красный	658–667	59,71	13,3	7,49	17,78	0,52	0,4	0,8
Зеленый	694–717	59,57	12,04	7,64	18,89	0,46	0,48	0,92

Таблица. Результаты рентгеновской энерго-дисперсионной спектроскопии поверхности

В связи с многокомпонентностью стали определение точного фазового состава образовавшейся пленки этим методом невозможно, однако химический состав поверхностных соединений может быть уверенно выявлен. В результате энерго-дисперсионого анализа в составе полученных в результате лазерного воздействия соединений из атмосферных газов обнаружен только кислород, таким образом, можно считать доказанным образование на поверхности легированной стали лишь оксидов металлов. Видно также, что различным цветам соответствует различное содержание кислорода и при увеличении температуры нагрева содержание кислорода в пленке увеличивается.

Анализ результатов расчета и их обсуждение

Из графиков (рис. 1) видно, что, согласно расчету, наиболее вероятным является образование соединений железа и хрома с кислородом, так как величина энергии Гиббса для них во всем диапазоне температур значительно меньше, чем для соединений с азотом, углеродом или водой. Это подтверждается и результатами энерго-дисперсионной рентгеновской спектроскопии, где из компонентов атмосферы обнаружен только кислород. Согласно результатам расчетов, при нагревании поверхности нержавеющей стали до температуры 700 К в атмосфере образуется многокомпонентная пленка со следующим фазовым составом: Fe_2O_3, CrO_2, NiO. При температурах 700–1000 К CrO_2 частично восстанавливается до Cr_2O_3, фазовый состав пленки становится другим: Fe_2O_3, Cr_2O_3, NiO. Данные энерго-дисперсионного анализа (процентный состав химических элементов) и экспериментальные данные лазерного воздействия [14, 15] и традиционного нагрева [16] подтверждают полученные теоретические результаты.

При недостатке кислорода, когда вероятность образования FeO и NiO велика, наиболее вероятно образование двойных оксидов: при температурах ниже 1000 К это $FeCr_2O_4$. Так как процентное содержание железа в стали больше, чем остальных элементов, то его избыток будет связываться в простой оксид Fe_2O_3. Экспериментальные данные [17] подтверждают возможность образования двойного оксида $FeCr_2O_4$ на внутреннем слое (подложка–пленка) с увеличением количества проходов и образования простых оксидов железа на внешнем слое (пленка–воздух).

Таким образом, сопоставление результатов расчета с экспериментом показывает их хорошее качественное совпадение. Кроме того, расчет термодинамическим методом позволяет получить важную дополнительную информацию о химических реакциях при лазерном нагревании среды в химически активной атмосфере.

Заключение

В результате проведенных расчетов и исследований можно сделать выводы, что методы химической термодинамики с учетом кинетических ограничений применимы для определения состава «продуктов» при лазерных воздействиях на металлы и сплавы в химически активной атмосфере для импульсно-частотных процессов. Есть все основания полагать, что такой подход можно использовать и при собственно кратковременных импульсных воздействиях вплоть до пико- и наносекундных (например, [18]).

В работе установлен фазово-химический состав пленок, полученных при взаимодействии лазерного излучения с поверхностью стали в атмосфере, который определяется температурой и скоростью нагревания поверхности. С помощью предложенного метода возможно определение продуктов взаимодействия с участием других легирующих элементов, входящих в состав стали. Приведенные экспериментальные данные подтверждают полученный фазово-химический состав пленок.

Авторы выражают благодарность А.И. Денисюку и Т.Б. Пришибилович за определение состава образцов; аспирантам кафедры ЛТиЭП НИУ ИТМО А.Д. Кочетову и В.А. Кочетовой за помощь в проведении тепловизионных исследований.

Работа выполнена при поддержке государственного контракта № 14.B37.21.0144, государственного контракта № 11.519.11.4017, гранта Президента РФ для ведущей научной школы 619.2012.2, гранта РФФИ 12–02–01194а.

Литература

1. Арзуов М.И., Барчуков А.И., Бункин Ф.В., Кириченко Н.А., Конов В.И., Лукьянчук Б.С. Особенности лазерного нагрева окисляющихся металлов в воздухе при наклонном падении излучения // Квантовая электроника. – 1979. – Т. 6. – № 3. – С. 466–472.
2. Odintsova G.V., Gorny S.G., Veiko V.P., Otkeeva A.V. Laser induced multicolor image formation on metal surfaces // Proc. of SPIE. – 2011. – V. 7996. – P. 799605-1–7.
3. Вейко В.П., Горный С.Г., Одинцова Г.В., Патров М.И., Юдин К.В. Формирование многоцветного изображения при лазерном окислении металлов // Изв. вузов. Приборостроение. – 2011. – Т. 54. – № 2. – С. 47–53.
4. Veiko V.P., Shakhno E.A., Poleshcuk A.G., Korolkov V.P., Matyzhonok V. Local laser oxidation of thin metal films: ultra-resolution in theory and in practice // J. of Laser Micro/Nanoengineering. – 2008. – V. 3. – P. 201–205.
5. Бонч-Бруевич А.М., Либенсон М.Н., Макин В.С., Пудков Д.С. О совместном действии импульсного и непрерывного оптического излучения на металлы // Письма в ЖТФ. – 1977. – Т. 3. – № 5. – С. 193–197.
6. Veiko V.P., Baranov A.V., Bogdanov K.V., Fedorov A.V., Yarchuk M.V., Ivanov A.I., Berwick K. Laser-induced thermo-oxidation of thin chromium films: micro-Raman characterization // J. of Raman spectrosc. – 2011. – V. 42. – № 9. – P. 1780–1783.
7. Горелик С.С., Расторгуев Л.Н., Скаков Ю.А. Рентгенографический и электронооптический анализ. – М.: Металлургия, 1970. – 366 с.

8. Либенсон М.Н. Лазерно-индуцированные оптические и термические процессы в конденсированных средах и их взаимное влияние. – М.: Наука, 2007. – 424 с.
9. Гиббс Д.В. Термодинамика. – М.: Наука, 1982. – 584 с.
10. Пригожин И., Дефэй Р. Химическая термодинамика. – Новосибирск: Наука, 1966. – 509 с.
11. Суворов С.А., Слободов А.А., Мусевич В.А., Иксанов Ф.Р. Термодинамический анализ химических и фазовых равновесий в системе MgO-C-Al-Al_2OC-$Al_2B_4C_7$-Al_4O_4C-Al_2O_3-Al_4C_3-воздух // Новые огнеупоры. – 2007. – № 10. – С. 44–50.
12. Слободов А.А., Кремнев Д.В., Мищенко А.В., Качер Е.Б., Наймушин А.Б. Возможности и эффективность термодинамического физико-химического моделирования химико-технологических систем и процессов // Известия СПбГТИ(ТУ). – 2009. – № 5 (31). – С. 26–31.
13. Yokokawa H., Fujishige M., Ujiie S., Dokiya M. CTC: Chemical thermodynamic computation system // J. Nat. Chem. Lab. Ind. – 1988. – V. 83. – P. 1–122.
14. Козаков А.Т., Яресько С.И. Исследование методом оже-спектроскопии состава поверхности многокомпонентных сплавов при импульсном лазерном воздействии // Физика и химия обработки материалов. – 2010. – № 3. – С. 67–73.
15. Li Z.L., Zheng H.Y., The K.M., Liu Y.C., Lim G.C., Seng H.L., Yakovlev N.L. Analysis of oxide formation induced by UV laser coloration of stainless steel // Appl. Surf. Sci. – 1999. – V. 256. – P. 1582–1588.
16. Xu J., Bai X., He F., Fan Y. The growth mechanism of the colored film on the stainless steel studied by 18O tracing and nuclear reaction analysis // Nucl. Instr. and Meth. in Phys. Res. B. – 1999. – V. 149. – P. 147–152.
17. Jervisa T.R., Williamsonb D.L., Hirvonena J.P. Characterization of the surface oxide formed by excimer laser surface processing of AISI 304 stainless steel // Mater. Lett. – 1990. – V. 9. – P. 379–383.
18. Veiko V.P., Yarchuk M.V., Ivanov A.I. Diffusionless Oxidation and Structure Modification of Thin Cr Films by the Action of Ultrashort Laser Pulses // Laser Physics. – 2012. – V. 22. – №. 8. – P. 1310–1316.

Вейко Вадим Павлович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, профессор, зав. кафедрой, veiko@lastech.ifmo.ru

Слободов Александр Арсеньевич – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор химических наук, профессор, aslobd@gmail.com

Одинцова Галина Васильевна – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, ogv-r@yandex.ru

УДК 535:771.36

ФОТОГРАФИЧЕСКИЕ ТЕХНОЛОГИИ НА ОСНОВЕ ЖИДКИХ КРИСТАЛЛОВ[1]

М.Г. Томилин

Регистрирующие среды на основе жидких кристаллов описаны в концепции двухступенной фотографии. Предложенный формализм позволяет описать регистрацию излучений и распределения физических полей на поверхности на основе развитых методов классической фотографии.

Ключевые слова: регистрирующие среды, жидкие кристаллы, излучения, распределения физических полей, поверхность, двухступенный процесс.

Введение

Гигантская оптическая анизотропия жидких кристаллов (ЖК) и их высокая лабильность, проявляемая при внешних воздействиях, позволяет рассматривать их как уникальные регистрирующие среды (РС), осуществляющие запись как световых полей, так и полей различной физической природы – электрических, магнитных, акустических, температурных и сил межмолекулярных взаимодействий. Наибольшей чувствительностью к воздействию полей обладают нематические жидкие кристаллы (НЖК), не имеющие, в отличие от холестерических и смектических материалов, надмолекулярной структуры.

Предпосылкой для визуализации полей является исходная упорядоченность молекул НЖК, которая может быть ими нарушена, вызывая деформацию слоя НЖК. Взаимодействие этих полей с регистрирующими слоями на основе НЖК может происходить или дистанционно при воздействии излучений видимого, ИК, СВЧ диапазонов и акустических полей, или при контакте слоя НЖК с изучаемой поверхно-

[1] Статья подготовлена по материалам приглашенного доклада, прочитанного на международной конференции The Latin America optics & photonics в Сан Себастьяне, Бразилия, 10–13 ноября 2012 [1]

стью, на которой имеется распределение физических полей. Индуцируемые полями деформации в НЖК регистрируются в поляризованном свете на просвет или отражение. Регистрацию деформаций слоя НЖК можно осуществлять как в реальном времени, так и с памятью. Для расширения возможностей регистрации можно использовать холестерические, смектические и полимерные ЖК (ПЖК).

Особенность регистрирующих сред на ЖК состоит в том, что близкие по структуре органические вещества могут использоваться как для дистанционной записи излучений, так и для контактной регистрации распределений физических полей на поверхности материалов и изделий. Более 80000 синтезированных жидкокристаллических соединений позволяют оптимизировать состав и свойства РС. Принципы детектирования излучений и физических полей описываются хорошо разработанной теорией деформации ЖК [2].

Высокая лабильность НЖК в мезофазе, обеспечиваемая текучестью и малой вязкостью, определяет их максимальную чувствительность к внешним воздействиям, делает естественным их использование в реальном времени, но затрудняет фиксирование ими статических изображений по эксплуатационным соображениям. Обращение с изображениями в среде, находящейся в мезофазе, вызывает трудности при длительном хранении (стекание, промокание и др.). Для практических случаев актуально решение противоречивой задачи: детектирование внешних воздействий в мезофазе ЖК и сохранение записанной информации в твердой фазе ЖК.

Цель настоящей работы состоит в развитии двухступенных методов регистрации изображений внешних физических полей с эффектом памяти, при которых регистрация внешних воздействий осуществляется в мезофазе, а эксплуатация и хранение – в твердокристаллическом состоянии.

Дистанционные методы записи на ЖК излучений различных длин волн

Методы регистрации излучений на ЖК можно отнести к несеребряным фотографическим процессам. Различают регистрацию, осуществляемую ЖК-композитами, например, на основе холестериков (рис. 1), смектиков, полимерных ЖК (рис. 2 и 3), и структурами типа «фотопроводник–ЖК» (рис. 4).

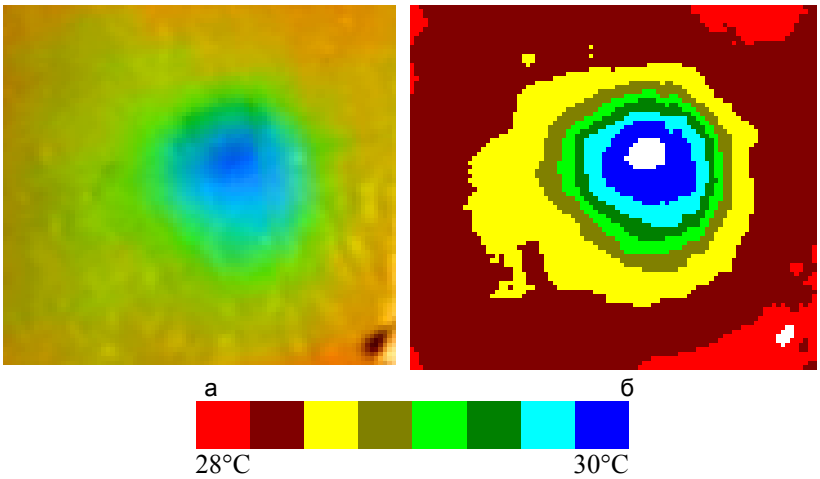


Рис. 1. Визуализация структуры пучка в сечении одномодового лазера диаметром 3 мм в ИК-диапазоне (а) и распределение температуры по сечению (изотермы) (б). Внизу показана цветотемпературная шкала [3]

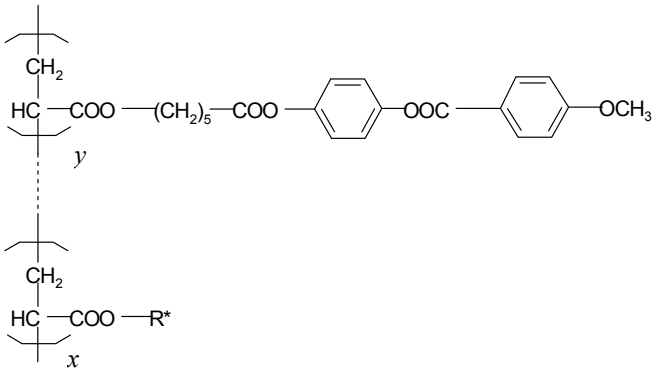


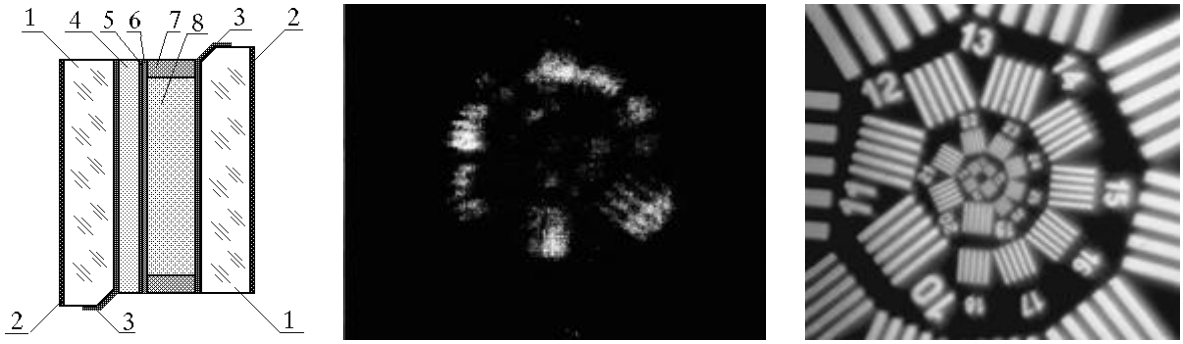
Рис. 2. Пример ПЖК со структурой хиральных ЖК-сополимеров с нематогенными и хиральными фотохромными группами [4]

ЖК может быть капсулирован в полимерную пленку для защиты от внешних воздействий [2]. Регистрация оптических изображений в видимом диапазоне длин волн может осуществляться на светочувствительном ПЖК, например, со структурой хиральных ЖК-сополимеров с нематогенными и хиральными фотохромными группами [4] (рис. 2). Примеры регистрации негативных и позитивных оптических изображений на ПЖК приведены на рис. 3.

а б

Рис. 3. Примеры негативного изображения с памятью (а) и позитивного изображения с памятью (б) на ПЖК, «вписанные» в формат 24×36 мм

Структуры «фотопроводник–ЖК» используются в системах динамической голографии и устройствах адаптивной оптики для коррекции фазовых искажений фронта световой волны в реальном времени. На рис. 4 показан пример коррекции изображения, осуществляемой такой структурой.



а б в

Рис. 4. (а) – структуры «фотопроводник–ЖК» отражательного типа: 1 – стеклянные подложки; 2 – просветляющие покрытия; 3 – прозрачные электроды; 4 – фотопроводник; 5 – светоблокирующий слой; 6 – зеркало; 7 – прокладки; 8 – слой ЖК. Изображения штриховой миры размером 24×36 мм до (б) и после (в) коррекции, создаваемые зеркалом телескопа [2]

Контактные методы визуализации распределения физических полей на поверхности материалов и изделий

Визуализация распределения физических полей на поверхности материалов и изделий является уникальной особенностью РС на основе НЖК, отличающей их от известных РС на других принципах. Основная идея применения НЖК в качестве РС состоит в том, что при деформации они выполняют функцию оптической реплики с большой величиной оптической анизотропии, осуществляющей фазовую задержку проходящего света. Слои НЖК, нанесенные на изучаемую поверхность, изучаются в поляризованном оптическом микроскопе (ПОМ). Распределение и величина локальных фазовых задержек в поле зрения ПОМ соответствуют распределению и интенсивности исследуемого поля на поверхности объекта. Задание деформации осуществляется в сверхтонком приповерхностном слое НЖК, который транслирует двумерную локальную ориентацию в толщу слоя. Граница «поверхность–НЖК» выполняет функцию командного слоя, а происходящие в ней явления имеют наноразмерную шкалу. Для накапливания фазовой задержки, достаточной для визуализации деформаций (а значит, и изучаемых полей), используют слои НЖК толщиной порядка 1 мкм.

Теория деформаций слоя НЖК под действием внешних полей описана в работе [2]. Принцип контактного метода визуализации физических полей показан на рис. 5. Наибольший практический интерес вызывает возможность визуализации структурных неоднородностей на поверхности и распределения невидимых физических полей. При выведении из оптической схемы слоя НЖК, поляризатора или анализатора наблюдатель теряет возможность регистрировать фазовую задержку. Это наглядно выявляет тот вклад в изображение, который привносит деформированный слой НЖК.

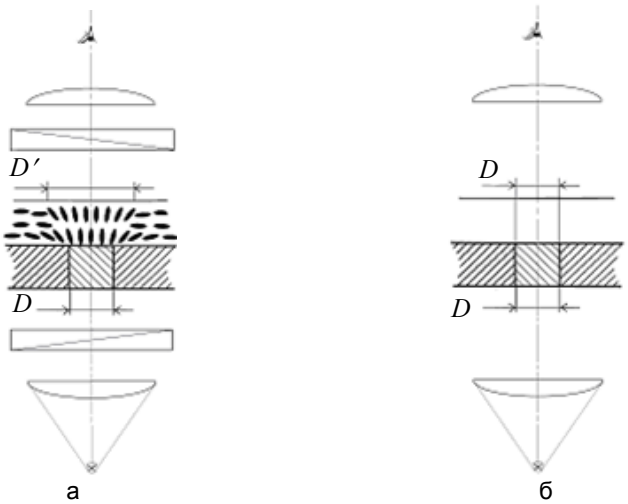


Рис. 5. (а) – в поляризованном свете наблюдаются деформации слоя НЖК *D′*, вызванные дефектом; (б) – в неполяризованном свете невидимый структурный дефект *D* наблюдается сквозь прозрачный слой НЖК, деформации в нем не видны, и поэтому дефект на поверхности не виден

Недостатки сохранения изображения в мезофазе можно устранить путем двухступенной регистрации изображения физических полей или дефектов на поверхности: на первой стадии реализуются условия формирования деформаций в текучем слое мезофазы НЖК, а на второй стадии осуществляется переход мезофазы в твердое состояние. Возможны, по крайней мере, два способа фиксирования изображения в твердой фазе.

В первом способе используется термотропные свойства НЖК-материала, имеющего температуру существования мезофазы выше комнатной температуры. На первой стадии НЖК нагреванием переводится в мезофазу, где он обладает малой вязкостью, что позволяет ему деформироваться под действием слабых полей поверхности [5]. На второй стадии НЖК охлаждается до комнатной температуры, переходя в твердокристаллическую фазу, но сохраняя при этом деформированную структуру в течение длительного времени. Описанный способ обладает реверсивностью. Твердокристаллическая фаза нагреванием может быть вновь переведена в мезофазу, в которой производится перезапись изображения поля, после чего мезофаза охлаждением переводится в твердокристаллическое состояние. Возможность такой записи была подтверждена авторами экспериментально при регистрации слабых магнитных полей [6]. На рис. 6, а, показана визуализированная структура магнитных доменов на поверхности сплава на основе Co.

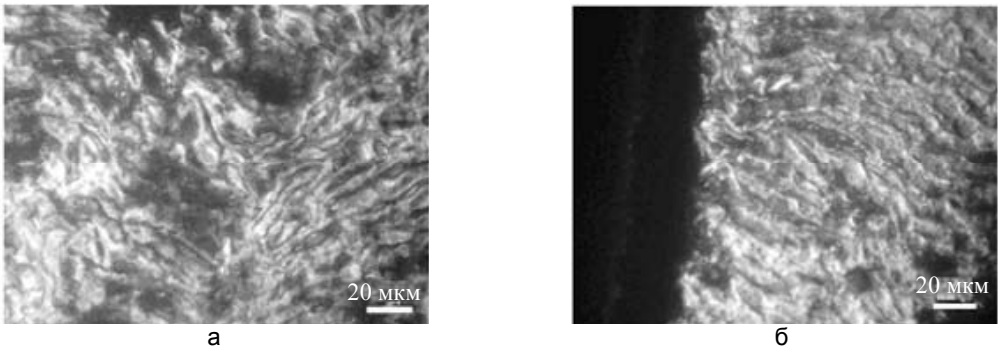


Рис. 6. Два примера (а) и (б) изображений магнитных доменов (темные области на фоне зерен), полученных с помощью смектического высокотемпературного ЖК на основе $L_3Tb(C_{12}H_{25}SO_4)_3$ [5]

После охлаждения до кристаллического состояния (рис. 6, б) материал сохраняет без видимых изменений изображения структуры доменов в твердой фазе.

Во втором способе используется фиксирование картины деформации слоя НЖК-методом полимеризации. Для этого в состав НЖК добавляется примерно 5% весового содержания полимера, что практически не изменяет его свойств.

Деформация слоя НЖК под действием распределения поверхностных полей осуществляется при комнатной температуре, после чего деформированная структура подвергается воздействию УФ излучения, вызывающего полимеризацию слоя (рис. 7). Способ был разработан для стабилизации исходной поверхностной и объемной ориентации молекул НЖК [7] и предложен авторами для фиксирования деформированной структуры.

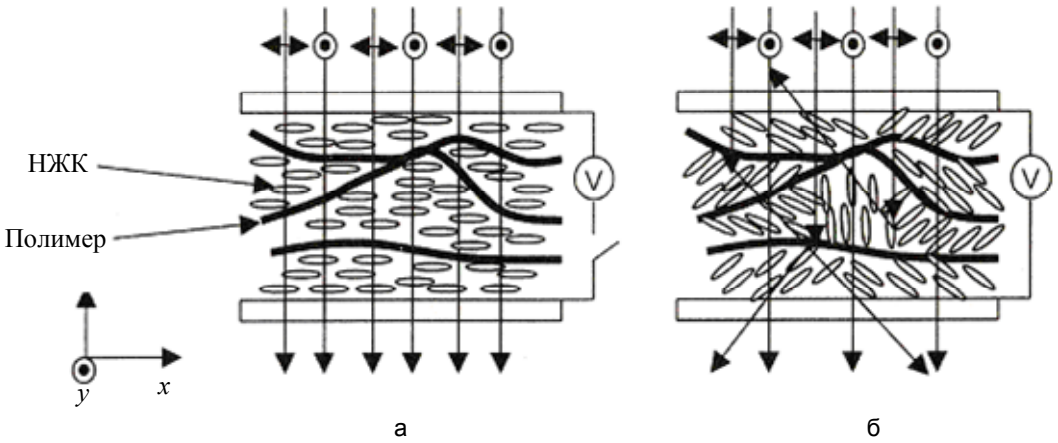


Рис. 7. Фиксирование деформированного слоя НЖК за счет его полимеризации при воздействии УФ излучения: структура смеси НЖК с полимером до внешнего воздействия и УФ облучения (а); структура после внешнего воздействия и УФ облучения (б) [7]

Разработанные методы показали возможность их уникального применения в широких областях физики, химии, биологии, медицины и высоких технологий, а предложенный подход позволяет использовать для описания РС на ЖК такие характеристики из арсенала классической фотографии, как функция рассеяния линии, информационная емкость, частотно-контрастная характеристика и др. [8].

Заключение

Процессы регистрации различных физических полей, включая оптические излучения различных длин волн, на основе использования жидких кристаллов впервые рассмотрены с позиций двуступенных технологий несеребряной фотографии. Показано, что противоречивую задачу их использования в качестве регистрирующих сред можно осуществить детектированием внешних воздействий в мезофазе и сохранением записанной информации в твердой фазе жидких кристаллов.

Литература

1. Tomilin M.G. Photographic technologies based on liquid crystals // Conference program and abstracts. LAOP 2012. – Mersedias Beech. – 13–15 November. – P. 22.
2. Томилин М.Г., Невская Г.Е. Фотоника жидких кристаллов. – СПб: Изд-во Политехн. ун-та, 2011. – 742 с.
3. Жаркова Г.М., Коврижина В.М., Хачатурян В.М. Полимерно-жидкокристаллические материалы для оптической диагностики температурных полей и их применение в аэрофизическом эксперименте // Современные оптические методы исследования потоков: Коллективная монография / Под редакцией Б.С. Ринкевичюса. – М.: Оверлей, 2011. – 360 с. (С. 275 – 290).
4. Шибаев В.П. Полимерные кентавры // Природа. – 2012. – № 6. – С. 12–24.
5. Беляев В.В. Вязкость нематических жидких кристаллов. – М.: ФИЗМАТЛИТ, 2002. – 224 с.
6. Томилин М.Г., Галяметдинов Ю.Г., Кузнецов П.А. Применение парамагнитных ЖК с гигантской магнитной анизотропией для изучения структуры доменов в магнитных материалах // Оптический журнал. – 2005. – Т. 74. – № 9. – С. 32–36.
7. Wu Sh.-T., Yang D.-K. Reflective LC displays. – J. Willey & Sons, Ltd. – 2001. – 335 p.
8. Джеймс Т.Х. Теория фотографического процесса. – Л.: Химия, 1980. – 672 с.

Томилин Максим Георгиевич – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, профессор, mgtomilin@mail.ru

7 СИСТЕМЫ АВТОМАТИЗИРОВАННОГО ПРОЕКТИРОВАНИЯ

УДК 681.3

АЛГОРИТМ РАЗБИЕНИЯ СВЕРХБОЛЬШИХ ИНТЕГРАЛЬНЫХ СХЕМ С АДАПТИВНЫМ УПРАВЛЯЮЩИМ ПАРАМЕТРОМ

П.Н. Филиппенко

Рассматривается задача разбиения сверхбольших интегральных схем. В качестве математической модели описания интегральных схем выбран граф. Предложена модификация муравьиного алгоритма оптимизации, используемая для решения задачи разбиения графа. Муравьиный алгоритм – это оптимизационный метод, основанный на принципах самоорганизации и других полезных особенностях поведения муравьев. Предлагается методика, основанная на муравьином алгоритме, с улучшенным методом начального распределения и динамической корректировкой управляющих поиском параметров. Экспериментальные исследования и сравнение с известными алгоритмами показали, что предложенная методика обеспечивает улучшение решений задачи разбиения сверхбольших интегральных схем.
Ключевые слова: разбиение, муравьиный алгоритм, управляющий параметр.

Введение

В настоящее время сверхбольшие интегральные схемы (СБИС) могут состоять из миллионов транзисторов. Учитывая огромное количество элементов, из которых состоит схема, ее топология не может быть спроектирована целиком. Это обусловлено ограничениями возможностей вычислительных мощностей компьютеров и недостаточными объемами памяти для единовременной обработки всей схемы [1]. В связи с этим необходимо разбиение схемы с целью получения блоков меньших размеров, работа с которыми возможна и эффективна.

Для математической описания СБИС используем аппарат теории графов. Пусть $G=(V,E)$ – неориентированный граф, где $V=\{v_i \mid i=1,2,...,n\}$ – множество вершин графа; $E=\{e_u \mid u=1,2,...,n\}$ – множество ребер графа [2]. Для решения задачи разбиения воспользуемся методом перехода от схемы к графу $G=(V,E)$, при котором элементам схемы соответствуют вершины графа V, а электрические цепи представляются ребрами E. В таком случае для решения задачи необходимо разбить V на k непустых частей, $V=V_1\cup V_2\cup..\cup V_k, \forall t\neq j, V_t\cap V_j=\varnothing$. Критерием оптимизации выступает суммарное число ребер F, находящихся между частями разбиения $V_1, V_2,..., V_k$. Цель оптимизации – минимизация F.

Муравьиный алгоритм – это оптимизационный метод, основанный на принципах самоорганизации и других полезных особенностях поведения муравьев [3]. Поскольку в основе муравьиного алгоритма лежит моделирование перемещения муравьев по заданным путям, этот подход является эффективным способом поиска решений оптимизационных задач [4], допускающих представление в виде графовой модели [2]. Рассмотрим обобщенно традиционный муравьиный алгоритм. Он представляет собой подход к решению комбинаторных задач оптимизации, в частности, задачи нахождения оптимального разбиения множества вершин V графа G на части в соответствии с какой-либо функцией качества [2]. Функцией качества может выступать величина минимального количества ребер, попадающих в разрез между кусками графа разбиения [5]. Муравьиный алгоритм рассматривает компоненты решения C и таблицу феромонов T и построен на принципах самоорганизации и коллективной адаптации [2, 4, 5]. Муравей является агентом, производящим строительство решения C. Решение находится последовательным объединением вершин в подмножества разбиения [6]. Во время избирательного движения агента-муравья от вершины к вершине он принимает решения о добавлении их в соответствующий блок разбиения [5, 7]. Каждая вершина характеризуется значением феромона, отложенного при инициализации или на предыдущих итерациях работы алгоритма агентами, уже проходившими через эту вершину [6]. Вероятность перемещения к вершине и добавления ее в соответствующий блок разбиения определяется агентом и зависит от количества находящегося на вершине феромона и управляющей информации, заложенной в системе изначально [8, 9]. Вероятность вхождения в формируемый кусок для каждой вершины $v_i\in V$ определяется [10] как

$$\Pr(v_i)\equiv\frac{\tau_i^{\alpha}\eta(v_i)^{\beta}}{\sum\limits_{v_j\in V}\tau_j^{\alpha}\eta(v_j)^{\beta}}, \qquad (1)$$

где $\tau_i\in T$ означает, что феромон соответствует вершине v_i, а $\eta(v_i)^{\beta}$ представляет собой дополнительные априорные знания или управляющую информацию, с помощью которой муравей

оценивает привлекательность вершины v_i [7]. Параметры $\alpha \geq 0$ и $\beta \geq 0$ определяют соотношение между информаций, заключенной в феромонах, и управляющей информацией соответственно [6].

В итоговой части муравьиного подхода осуществляется обновление феромона, и удачные решения помечаются более чем остальные, что повышает вероятность их использования в решениях, найденных на следующих итерациях. Также в этой процедуре феромоны частично испаряются с течением времени [2]. Это делается во избежание застревания в локальных оптимумах во время поиска колониями новых решений.

Приведем способ решения задачи оптимизации с использованием муравьиного алгоритма [5–10]:

1. собрать решение-кандидата (кусок разбиения) из множества вершин на основе вероятностей, руководствуясь величиной оставленного феромона на вершине и управляющей информацией.
2. обновить модель феромонов, основываясь на качестве решения-кандидата.

Повторять первые два шага, пока результат не будет удовлетворять критериям либо не будет достигнут лимит итераций работы алгоритма [6].

Предложенный метод начального распределения муравьев

В начале работы муравьиного алгоритма через некоторые пути проходит множество муравьев, но есть пути, которые не посещаются. Локальная эвристика, основанная на видимости, поощряет муравьев выбрать вершины, находящиеся ближе. В некоторых вершинах может быть много муравьев, в то время как в других их не будет вообще. Поскольку количество феромона на каждом пути первоначально идентично, муравей, главным образом, использует расстояние между двумя вершинами как эвристический фактор, когда выбирает направление к следующей вершине [10]. В то время как муравьи находятся в относительно большем количестве в определенной вершине, плотность феромона на близлежащих ребрах пути будет усилена из-за относительно большего числа муравьев, путешествующих вдоль пути. Этот путь не обязательно является самым коротким путем, но при этом локально оптимальное решение найдено, и муравьи создают ситуацию попадания алгоритма в локальную яму.

Чтобы решить эту проблему, алгоритм строится таким образом, чтобы муравьи были распределены равномерно, т.е. размещает m муравьев в n вершинах и удостоверяется, что по крайней мере один муравей находится в каждой вершине. Таким образом, производительность поиска и вероятность получения лучшего результата увеличены.

Стратегия изменения управляющего параметра

В муравьином алгоритме управляющая информация важна для генерации высококачественных маршрутов. Поскольку ценность следов феромона не дает достаточной информации в ранней стадии накопления опыта системой и не может направлять муравьев к построению хороших маршрутов, для управляющих параметров могут быть установлены значения большей величины [6, 9]. С другой стороны, на более поздних стадиях управляющему параметру могут быть более полезны небольшие величины значений, потому что следы феромонов к этому моменту накапливают достаточно информации для требуемого поведения, и управляющая информация может вводить в заблуждение, сбивая с правильного пути поиска и направляя к локальным оптимумам. Управляющий параметр устанавливается в качестве константы в традиционном муравьином алгоритме [5, 7]. Этот параметр подбирают экспериментальным путем или аналитически [8]. В настоящей работе в качестве управляющего параметра выступает параметр β (1). Если ценность управляющего параметра является большой, муравей будет пытаться найти другой маршрут (отличный от уже найденного, имеющего большую концентрацию феромона и вызывающего преждевременную сходимость). Очевидно, что маленькая величина управляющего параметра может привести к плохой работе на ранней стадии накопления опыта муравьями. Однако небольшие значения ценности управляющего параметра могут быть эффективны в случае достаточно продолжительного поиска [10, 11]. Исходя из этого, используется интуитивный адаптивный управляющий параметр для муравьиного алгоритма. Предлагается способ назначения адаптивного управляющего параметра, приводящий к улучшению производительности поиска.

Когда муравьиный алгоритм начинает работу, он располагает одинаковой информацией о каждом из маршрутов, в этот момент информационная энтропия системы максимальна. С увеличением количества феромона на путях энтропия будет постепенно уменьшаться. Если в текущий момент энтропией не управлять, она, в конечном счете, уменьшится до нулевого значения. Феромон при этом будет сосредоточен только на одном пути и максимален. В такой ситуации будет получено ошибочное решение, приводящее к преждевременной сходимости [10, 12]. Для того чтобы преодолеть преждевременную сходимость, вызывающую дефекты решения сложных комбинаторных оптимизационных задач с помощью алгоритма, основанного на симуляции муравьиной колонии, необходима его модификация. Предлагается муравьиный алгоритм, основанный на анализе информационной энтропии и использующий значения управляющего параметра, задающегося энтропией. Каждый след – дискретная случайная переменная в матрице распределения феромонов. Энтропия случайной переменной определена как

$$E(X) = -\sum_{i=1}^{r} P_i \log P_i \,, \tag{2}$$

где P_i представляет вероятность возникновения каждого следа в матрице распределения феромонов. Для симметрии n вершин матрицы феромонов: $n(n-1)/2$ отмеченных следов феромона, $r = n(n-1)/2$. Максимальная энтропия задается как

$$E_{\max} = -\sum_{i=1}^{r} P_i \log P_i = -\sum_{i=1}^{r} \frac{1}{r} \log \frac{1}{r} = \log r \,. \tag{3}$$

Предлагается использовать величину энтропии в качестве показателя, определяющего степень накопления информации, содержащейся в феромонах на всех участках пути. Управляющий параметр обновляется в соответствии со следующим правилом:

$$\beta = \begin{cases} 5 & A < \mathrm{E}' \le 1 \\ 4 & B < E' \le A \\ 3 & C < E' \le B \\ 2 & 0 < E' \le C \end{cases}, \text{когда} \tag{4}$$

$$E' = 1 - \frac{E_{\max} - E}{E_{\max}} \,, \tag{5}$$

где E' является величиной энтропии для текущей матрицы распределения феромонов и рассчитывается по формуле (3), а A, B, C – границы, обусловленные величиной накопления феромонов. Значение границы A установлено в пределах 0,81 ~ 0,89, границы B – 0,74 ~ 0,54, а границу C решено сделать эвристически основанной на величине B.

Рисунок. Упрощенная схема модифицированного муравьиного алгоритма

Предложенный алгоритм объединяет в себе метод начального распределения положения муравьев, повышающий качество поиска на начальных этапах работы, и стратегию изменения управляющего параметра, предотвращающую преждевременную сходимость. Работу предложенного модифицированного алгоритма можно условно описать следующими шагами.

Шаг 1. Ввод исходных данных и инициализация параметров алгоритма.

Шаг 2. Расчет максимальной энтропии по формуле (4).

Шаг 3. Размещение муравьев в стартовых позициях согласно предложенной стратегии распределения.

Шаг 4. Поиск решений муравьиным алгоритмом, описанным выше.

Шаг 5. Перерасчет системы, переопределение и запоминание лучшего результата, полученного в результате работы алгоритма.

Шаг 6. Возможно, улучшение результатов с помощью локального поиска, осуществляемого другим алгоритмом на базе решения, полученного на текущем этапе работы муравьиного алгоритма.

Шаг 7. Проверка выполнения необходимого числа итераций алгоритма. Если не выполнено, то переход к шагу 4, если выполнено, то переход к шагу 8.

Шаг 8. Вычисление значений величин энтропии на ребрах графа по формуле (2).

Шаг 9. Расчет нового управляющего параметра поиска по формуле (5). Для выполнения следующей итерации алгоритма переход к шагу 4.

Шаг 10. Формирование окончательного решения.

Шаг. 11. Конец работы алгоритма.

Упрощенная блок-схема алгоритма приведена на рисунке.

Алгоритм разбиения блоков был реализован на языке C++. При создании программы использовался пакет Visual Studio 2008 для 32-разрядных операционных систем семейства Windows, тестирование и эксперименты проводились на IBM-совместимом компьютере с процессором Intel(R) Duo Core(TM)Duo 2Ghz и оперативной памятью 512 MB DIMM PC-3200.

Эффективность разработанного алгоритма разбиения оценивалась путем проведения серии экспериментов на специальных схемах (бенчмарках) PrimGA1. Выполнялось сравнение с муравьиными алгоритмами [6, 12], не использующими предложенные в настоящей работе модификации. В качестве параметров, используемых для сравнения, выступали σ – стабильность алгоритма и κ – коэффициент, определяющий степень улучшения решений.

Стабильность алгоритма определяется как число лучших значений целевой функции на определенном интервале итераций исследуемого алгоритма. Чем больше значение σ, тем более нестабилен алгоритм. Стабильность σ вычисляется следующим образом:

$$\sigma = \left| \frac{f_{\text{ср}} - f}{f_{\text{ср}}} \right|,$$

где $f_{\text{ср}}$ – усредненное значение лучшей целевой функции на определенном интервале итераций; f – значение лучшей целевой функции исследуемого алгоритма.

*К*оэффициент κ, определяющий степень улучшения решений, вычисляется по формуле

$$\kappa = \left(100 \frac{f}{f'} \right) - 100 \text{ [\%]},$$

где f' – значение лучшей целевой функции алгоритма, с которым сравнивается исследуемый алгоритм; f – значение лучшей целевой функции исследуемого алгоритма.

В результате сравнения с другими муравьиными алгоритмами [6, 12] удалось установить, что представленный в работе модифицированный муравьиный алгоритм находит оптимальные решения за меньшее число итераций и эффективней в среднем на 6%.

Заключение

Наилучшие результаты муравьиные алгоритмы показывают для задач с большими размерностями областей поиска, к которым относится задача разбиения на блоки. Алгоритмические усовершенствования простого алгоритма приводят к улучшению качества решений. Поскольку от выбора управляющих параметров сильно зависит поведение муравьиных алгоритмов, именно к этой проблеме обращено наибольшее внимание. Предложенный алгоритм объединяет в себе метод начального распределения положения муравьев, повышающий качество поиска на начальных этапах работы, и стратегию изменения управляющего параметра, предотвращающую преждевременную сходимость.

В качестве направления для развития дальнейших исследований можно отметить распараллеливание алгоритма с целью улучшения получаемых решений [13]. Взаимодействия, происходящие на уровне самоорганизации внутри колонии, следует дополнить взаимообменом информацией с другими обособ-

ленными колониями, сосуществующими в едином пространстве поиска. При этом просчет взаимодействий внутри каждой колонии целесообразно реализовывать на отдельных вычислительных машинах, а взаимодействие на межколониальном уровне – через сеть. В случае продолжения работы в указанном направлении анализ способа выбора управляющих параметров способен повысить качество полученных решений.

Литература

1. Балюк Л.В., Курейчик В.В., Сороколетов П.В. Перспективная технология интегрированного поиска в САПР // Известия ЮФУ. Технические науки. Тематический выпуск «Интеллектуальные САПР». – 2007. – № 2 (77). – С. 18–25.
2. Dorigo M., Stützle T. Ant Colony Optimization. – London: A Bradford, 2004. – 305 p.
3. Курейчик В.М., Кажаров А.А. Муравьиные алгоритмы для решения транспортных задач // Известия РАН. Теория и системы управления. – 2010. – № 1. – С. 32–45.
4. Гладков Л.А., Курейчик В.В., Курейчик В.М. Биоинспирированные алгоритмы в оптимизации. – М.: ФИЗМАТЛИТ, 2009. – 384 с.
5. Курейчик В.М., Кажаров А.А. О некоторых модификациях муравьиного алгоритма // Известия ЮФУ. Технические науки. Тематический выпуск «Интеллектуальные САПР». – 2008.– № 4 (81). – С 7–12.
6. Лебедев О.Б. Муравьиный алгоритм разбиения // Известия ЮФУ. Технические науки. Тематический выпуск «Интеллектуальные САПР». – 2009. – № 12 (101). – С. 53–58.
7. Курейчик В.В., Курейчик В.М., Родзин С.И. Концепция эволюционных вычислений, инспирированных природными системами // Известия ЮФУ. Технические науки. Тематический выпуск «Интеллектуальные САПР». – 2009. – № 4 (93). – С. 16–24.
8. Штовба С.Д. Муравьиные алгоритмы // Математика в приложениях. – 2004. – № 4 (4). – С. 70–75.
9. Лебедев Б.К., Лебедев О.Б. Эволюционные процедуры разбиения СБИС // Известия ЮФУ. Технические науки. – 2008. – № 4 (81). – С. 20–26.
10. Hlaing Z., Khine M. An Ant Colony Optimization Algorithm for Solving Traveling Salesman Problem // International Proceedings of Computer Science and Information Technology. – 2011. – V. 16. – P. 54–59.
11. Hinne M., Marchiori. E. Cutting Graphs Using Competing Ant Colonies and an Edge Clustering Heuristic // Proceedings of the 11th European conference on Evolutionary computation in combinatorial optimization. – 2011. – V. 6622. – P. 60–71.
12. Randall M., Lewis A. Modifications and additions to ant colony optimization to solve the set partitioning problem // Proceedings of the 2010 6th IEEE International Conference on e-Science Workshops. – 2010. – P. 110–115.
13. Филиппенко П.Н., Шашелов А.А., Сеитова С.В. Построение систем, обрабатывающих большие вычисления: проблемы и тенденции // Известия ЮФУ. Технические науки. Тематический выпуск «Интеллектуальные САПР». – 2010. – № 12 (113). – С. 196–203.

Филиппенко Петр Николаевич – Южный федеральный университет, аспирант, peterpain@mail.ru

8 МЕТОДЫ И СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ

УДК 004.056.53

ДЕРЕВЬЯ УЩЕРБОВ КАК МОДЕЛЬ ОЦЕНКИ РИСКОВ ПОТЕРЬ ДОСТУПНОСТИ ПОСЛЕ ПРОВЕДЕНИЯ ИЗМЕНЕНИЙ В ФИНАНСОВЫХ ИНФОРМАЦИОННЫХ СИСТЕМАХ

С.А. Арустамов, М.Г. Генин

Рассматривается метод прогнозирования и анализа рисков потерь доступности критических информационных систем с помощью построения деревьев ущербов. Приводится описание метода построения дерева ущербов для каждого из возможных событий, которые могут привести к потере доступности. Приводится пример, иллюстрирующий использование метода, на основании которого можно выбрать наиболее рациональную стратегию тестирования программного обеспечения.

Ключевые слова: информационные системы, оценки, анализ, доступность, риск потери доступности, ущерб, дерево ущербов, стратегия тестирования ПО.

Введение

В работе рассматриваются риски потери доступности программного обеспечения (ПО) при проведении обновлений финансовых информационных систем [1, 2]. Как правило, финансовые информационные системы имеют жесткие требования по доступности. Под доступностью в работе будем понимать возможность системы выполнять заявленные функции их легальными пользователями.

Обычно риск потери доступности [3] определяют как пару, состоящую из вероятности наступления того или иного события, приведшего к потере доступности после проведения обновления, и ущерба от наступления этого события. В настоящей работе предлагается ввести расширенное понятие ущерба в предположении, что он носит вероятностный характер. Например, утрата части клиентской базы и связанная с этим потеря доли рынка в результате сбоя ПО может и не произойти. Все зависит от множества объективных и субъективных факторов, таких как продолжительность неработоспособности системы в результате сбоя, время сбоя, наличие клиентов, которые в момент сбоя работали с системой, и т.д. Исходя из этого, можно говорить как о величине ущерба, так и о вероятности наступления ущерба при условии, что сбой произошел. Основываясь на этих предположениях, определим риск потери доступности следующим образом:

$$R_i = \{P_i, L_i(P_{Li})\}, \quad (1)$$

где R_i – риск наступления i-го события; P_i – вероятность наступления i-го события; $L_i(P_{Li})$ – величина ущерба от i-го события, наступающая с вероятностью P_{Li}, при этом ущерб будем рассматривать как дискретную случайную величину (ДСВ).

Будем считать, что на момент начала исследования текущая версия системы до обновления не содержит ошибок, т.е. риски потери доступности для нее отсутствуют. Предметом исследования является обнаружение и предотвращение в максимально короткие сроки проблем, которые могут возникнуть в результате обновления системы. Для решения данной задачи предлагается следующий подход. Для каждого из выявленных возможных негативных событий (сбоев), которые могут иметь место в результате обновления[1], проводится оценка риска потери доступности на основе анализа возможных ущербов от этого события. События с более высоким риском признаются более опасными. На основе этого строится стратегия тестирования обновления, по которой приоритет отдается тестированию компонентов ПО с высокими рисками. Для количественной оценки риска потери доступности в работе предлагается метод построения дерева ущерба для каждого из выявленных негативных событий. Вопросы технологии тестирования ПО не являются предметом данной работы.

Построение дерева ущербов для *i*-го события

Продолжим дальше развивать понятие ущерба. Обратим внимание на то, что ущерб L_i , возникший в результате события i, может быть, вообще говоря, составным, т.е. может быть разложен на несколько независимых составляющих:

$$L_i = \{L_{ij}\}, j = \{1, 2, \ldots, N\},$$

[1] Выявление проведенных обновлений автоматически отслеживается программными агентами, установленными на критических узлах системы. Сведения о проведенных изменениях передаются на централизованный сервер.

где L_{ij} – j-я составляющая ущерба L_i, возникшего в результате i-го события. В самом деле, в результате сбоя системы могут иметь место одновременно и независимо друг от друга несколько негативных последствий. Например, в результате неполадок новой версии СУБД перестали быть доступными независимые друг от друга функции (такие, например, как ввод документов и подготовка отчетности). Как следствие, возникает несколько независимых друг от друга ущербов. Продолжая цепочку рассуждений дальше, по аналогии с (1), можно заметить, что у каждой j-й составляющей ущерба L_{ij} может быть своя вероятность возникновения P_{Lji}, которая не зависит от других составляющих. Отталкиваясь от приведенного выше примера, можно предположить, что ущерб в результате невозможности ввода документов никак не связан с ущербом из-за проблем модуля подготовки отчетности. Вероятность возникновения первого связана с возможным недовольством клиентов или руководителей, а вторая – с возможными штрафными санкциями со стороны надзорных органов. В результате получаем следующее выражение для риска потери доступности:

$$R_i=\{P_i,\ \{L_{i1}(P_{Li1}),\ L_{i2}(P_{Li2}),\ L_{i3}(P_{Li3}),\ \ldots,\ L_{iN}(P_{LiN})\}\},\qquad(2)$$

где N – количество независимых составляющих ущерба от i-го события.

Рассмотрим теперь ситуацию, когда составляющие ущерба являются зависимыми друг от друга, т.е. когда одна составляющая ущерба является следствием другой. Это возможно в случае, когда в результате неполадок в работе системы часть клиентов переходит к другому поставщику услуг, в результате чего уменьшается доля контролируемого рынка, и т.д. Возникает ситуация, когда один ущерб может вызвать другой, и т.д., т.е. возникает цепочка вложенных ущербов. Каждый из ущербов в цепочке имеет свою собственную вероятность возникновения, но при условии, что реализовались все предыдущие ущербы из цепочки. В данном примере уменьшение доли рынка возможно только при переходе части клиентов к другому поставщику, а переход на другую платформу возможен, в свою очередь, только как следствие неполадок в работе системы. В этом случае риск потери доступности для зависимых случайных величин можно определить следующим образом:

$$R_i=\{P_i,\ \{L_{i1}(P_{Li1})\},\ L_{i2}(P_{Li2}|_{Li1}),\ \ldots,\ L_{iN}(P_{LiN}|_{LiN-1\ldots Li3Li2Li1})\}\},\qquad(3)$$

где N – количество зависимых ущербов от i-го события в данной цепочке.

Возможно существование ситуаций, когда цепочка зависимых ущербов завершается несколькими независимыми друг от друга ущербами, каждый из которых, в свою очередь, может быть началом новой цепочки зависимых ущербов, т.е. присутствует комбинация зависимых и независимых ДСВ.

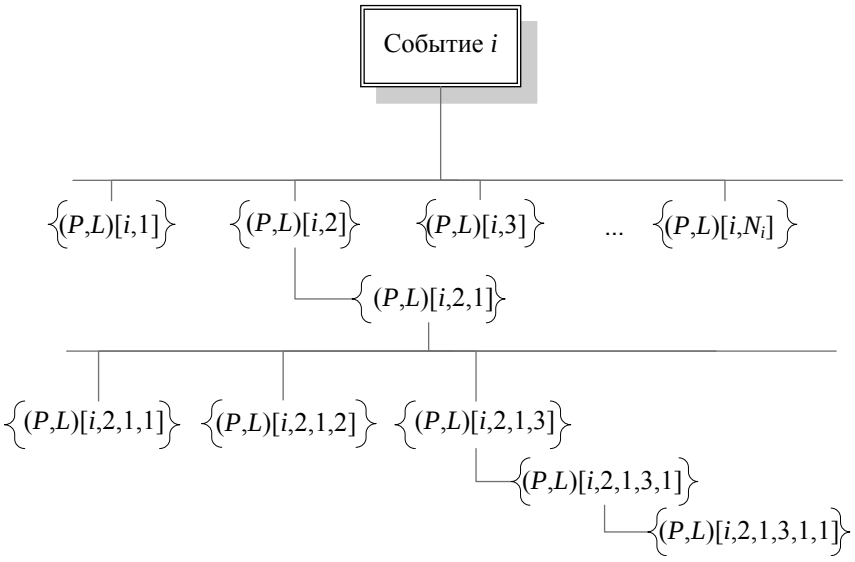


Рис. 1. Графическое изображение дерева ущербов, порождаемого i-м событием

В общем случае можно сказать, что в результате события i может иметь место набор независимых ущербов, некоторые из которых могут, в свою очередь, породить цепочку зависимых. Каждая из порожденных цепочек зависимых ущербов может, в свою очередь, заканчиваться несколькими независимыми ущербами, которые тоже могут порождать свои цепочки, и т.д. Формально это можно описать различными вложенными комбинациями формул (2) и (3), однако представить общий случай в виде одной формулы сложно, так как выражение будет слишком громоздким. Вместо этого предлагается вариант графического представления совокупности ущербов в виде дерева.

Перед тем, как начать построение дерева, перейдем к более удобному, на наш взгляд, способу обозначения пары «вероятность–ущерб». Вместо нижних индексов будем использовать индексы, заключенные в квадратные скобки. Такие обозначения используются, например, для описания многомерных массивов. Индексы в скобках будем указывать сразу для пары (P, L), а не для каждой из составляющих, так

как индексы для P и для L из одной пары всегда совпадают. Например, запись $L_{i1}(P_{Ii1})$ преобразуется в $(P, L)[i, 1]$. Тогда запись (3) для независимых ущербов будет выглядеть следующим образом:

$R_i=\{P_i, \{(P, L)[i, 1], (P, L)[i, 2], (P, L)[i, 3], \ldots, (P, L)[i, N_i]\}\}$,

где N_i – количество независимых составляющих ущерба от i-го события.

Для обозначения ущерба, подчиненного предыдущему, мы будем использовать индекс следующего уровня. Например, ущерб, который подчинен ущербу $(P, L)[i, 2]$, мы будем обозначать как $(P, L)[i, 2, 1]$, следующий за ним в цепочке ущерб будет уже $(P, L)[i, 2, 1, 1]$, и т.д. Запись (3) для зависимых ущербов в новой форме будет такой:

$R_i=\{P_i, \{(P, L)[i, 1], (P, L)[i, 1, 1], (P, L)[i, 1, 1, 1] ,\ldots, (P, L)[i, 1,\ldots,1]\}\}$.

Например, если у узла $(P, L)[i, 1, 1]$ будет не один подчиненный ему ущерб, а N_i подчиненных ущербов, независимых друг от друга, то будем обозначать их как $(P, L)[i,1,1,1]$, $(P, L)[i, 1, 1, 2]$, $(P, L)[i, 1, 1, 3]$, $\ldots$, $(P, L)[i, 1, 1, N_i]$.

Используя данный подход к обозначению ущербов разных уровней (узлов дерева), мы можем для наглядности перейти к модели ущербов в виде дерева (рис. 1).

Свертка дерева ущерба для *i*-го события. Ожидаемое значение ущерба

Выше было рассмотрено разложение ущерба из формулы (1) на составляющие с помощью построения дерева ущербов. Тем не менее, общий возможный ущерб, указанный в формуле (1), наряду с его составными частями, также имеет смысл как отдельная числовая величина, а разложение его на составляющие в виде дерева ущерба – это лишь вариант представления общего ущерба. Следовательно, должен существовать способ преобразования данных из дерева ущербов в величину общего возможного ущерба из формулы (1). Назовем эту величину ожидаемым значением ущерба L_i^{est}.

Для ее вычисления будем применять подход, основанный на расчете математического ожидания случайной величины. Вначале отдельно рассмотрим варианты расчета математического ожидания для независимых и зависимых ущербов, а затем их комбинацию для расчета ожидаемого значения ущерба для всего дерева.

Для независимых событий (горизонтальный уровень дерева) средневзвешенное значение ДСВ (математическое ожидание) вычисляется как сумма произведений вероятности и величины каждой из возможных комбинаций ущербов, разделенная на сумму вероятностей комбинаций этих ущербов. Так как перечисляются все возможные комбинации независимых ущербов одного горизонтального уровня, то их совокупность можно считать полной группой событий, поэтому сумма вероятностей всех возможных комбинаций, т.е. знаменатель формулы, будет равен единице.

Рассмотрим дерево k при гипотезе, что число его независимых ущербов горизонтального уровня равно 2. Для простоты записи обозначим вероятность j-ого ущерба P^k_j, а его величину L^k_j (j = 1,2). Тогда ожидаемое значение ущерба для дерева k будет следующим:

$$L_k^{est} = \frac{P_1^k \cdot (1-P_2^k)\cdot L_1^k + P_2^k \cdot (1-P_1^k)\cdot L_2^k + P_1^k \cdot P_2^k \cdot (L_1^k + L_2^k) + (1-P_1^k)\cdot(1-P_2^k)\cdot L^k(0)}{P_1^k \cdot (1-P_2^k) + P_2^k \cdot (1-P_1^k) + P_1^k \cdot P_2^k + (1-P_1^k)\cdot(1-P_2^k)}, \quad (4)$$

где $L^k(0) = 0$ – нулевая величина ущерба, соответствующая исходу, при котором он отсутствует.

Так как $L^k(0) = 0$, а знаменатель, образуя сумму вероятностей полной группы событий (ущербов), равен единице, то формула (4) примет следующий вид:

$$L_k^{est} = P_1^k \cdot (1-P_2^k)\cdot L_1^k + P_2^k \cdot (1-P_1^k)\cdot L_2^k + P_1^k \cdot P_2^k \cdot (L_1^k + L_2^k). \quad (5)$$

После аналогичных рассуждений и преобразований для дерева m с гипотезой о трех независимых ущербах формула расчета ожидаемого ущерба примет вид

$$\begin{aligned} L_m^{est} &= P_1^m \cdot (1-P_2^m)\cdot(1-P_3^m)\cdot L_1^m + P_2^m \cdot (1-P_1^m)\cdot(1-P_3^m)\cdot L_2^m + P_3^m \cdot (1-P_1^m)\cdot(1-P_2^m)\cdot L_3^m + \\ &+ P_1^m \cdot P_2^m \cdot (1-P_3^m)\cdot(L_1^m + L_2^m) + P_2^m \cdot P_3^m \cdot (1-P_1^m)\cdot(L_2^m + L_3^m) + P_1^m \cdot P_3^m \cdot (1-P_2^m)\cdot(L_1^m + L_3^m) + \\ &+ P_1^m \cdot P_2^m \cdot P_3^m \cdot (L_1^m + L_2^m + L_3^m). \end{aligned} \quad (6)$$

В общем случае с $N > 3$ независимыми составляющими ущерба от i-го события формула для вычисления L_i^{est} будет очень громоздкой, поэтому в работе она не приведена. Кроме того, на практике крайне редко удается выделить более трех независимых существенных ущербов, порождаемых одним узлом дерева.

При суммировании зависимых ущербов одной ветви дерева будем использовать тот же подход, что и для (5)–(6), перебирая все возможные комбинации ущербов для ветви дерева, но учитывая при этом, что реализация каждого следующего ущерба из данной ветви возможна при условии, что все предыдущие ущербы реализовались. В отличие от независимых ущербов одного горизонтального уровня, совокупность зависимых ущербов вертикальной ветви не образует полную группу событий, так как не все возможные комбинации реализуются. Ущерб d не может реализоваться в случае, если не реализовались ущербы 1, 2, …, (d–1).

Рассмотрим дерево q при гипотезе, что оно состоит из одной ветви, содержащей три последовательных ущерба ($d = 3$). В этом случае ожидаемое значение ущерба примет вид

$$L_q^{est} = \frac{(1-P_1^q)\cdot L^q(0) + P_1^q\cdot(1-P_2^q)\cdot(1-P_3^q)\cdot L_1^q + P_1^q\cdot P_2^q\cdot(1-P_3^q)\cdot(L_1^q+L_2^q) + P_1^q\cdot P_2^q\cdot P_3^q\cdot(L_1^q+L_2^q+L_3^q)}{(1-P_1^q)+P_1^q\cdot(1-P_2^q)\cdot(1-P_3^q)+P_1^q\cdot P_2^q\cdot(1-P_3^q)+P_1^q\cdot P_2^q\cdot P_3^q}, \quad (7)$$

где $L^q(0)$ – нулевой ущерб. Так как $L^q(0) = 0$, то формула (7) примет следующий вид:

$$L_q^{est} = \frac{P_1^q\cdot(1-P_2^q)\cdot(1-P_3^q)\cdot L_1^q + P_1^q\cdot P_2^q\cdot(1-P_3^q)\cdot(L_1^q+L_2^q) + P_1^q\cdot P_2^q\cdot P_3^q\cdot(L_1^q+L_2^q+L_3^q)}{(1-P_1^q)+P_1^q\cdot(1-P_2^q)\cdot(1-P_3^q)+P_1^q\cdot P_2^q\cdot(1-P_3^q)+P_1^q\cdot P_2^q\cdot P_3^q}. \quad (8)$$

Поскольку совокупность зависимых ущербов вертикальной ветви не образует полную группу событий, знаменатель формулы (7) будет всегда меньше единицы, поэтому его придется учитывать при расчете ожидаемого значения ущерба для зависимых событий ветви дерева.

С использованием этой методики можно рассчитать ожидаемое значение ущерба для всего дерева. Вначале, используя (8), будем проводить свертку от концов самых нижних ветвей до нижних горизонтальных уровней. Затем, в соответствии с (5)–(6), будем проводить свертку самих этих горизонтальных уровней. Повторяя эти две операции, мы постепенно свернем все дерево, двигаясь снизу вверх от самых длинных ветвей к вершине. На рис. 2 показан пример процесса поэтапной свертки дерева ущербов.

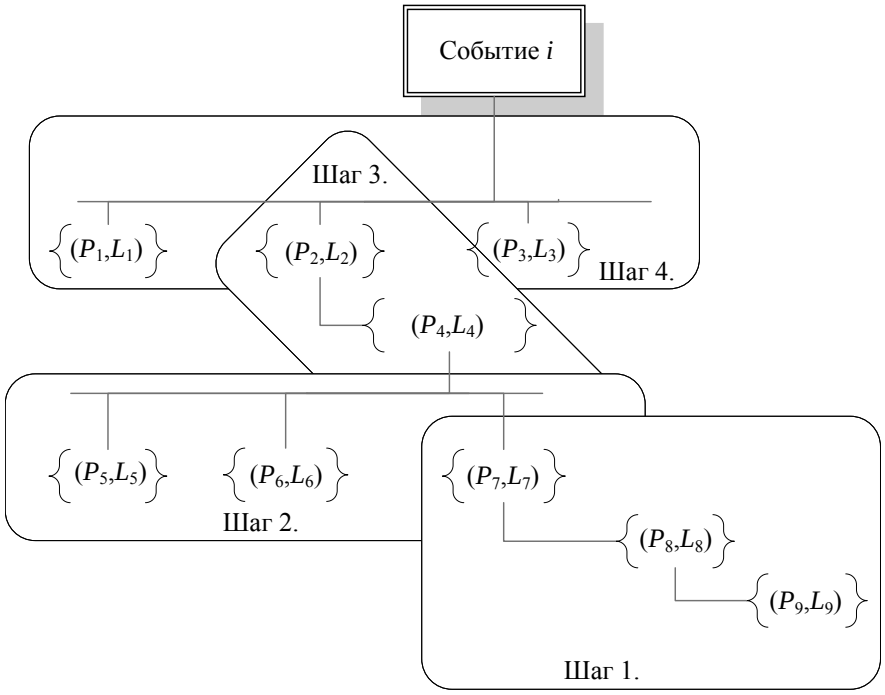


Рис. 2. Пример пошагового алгоритма свертки для дерева ущербов

Ниже приведена последовательность действий по расчету ожидаемого значения ущерба для данного дерева в соответствии с шагами, показанными на рис. 2.

Шаг 1.

$$L^{est1} = L_7 + \frac{P_8\cdot(1-P_9)\cdot L_8 + P_8\cdot P_9\cdot(L_8+L_9)}{(1-P_8)+P_8\cdot(1-P_9)+P_8\cdot P_9}.$$

Так как узел $\{(P_7, L_7)\}$ является не только источником зависимых ущербов $\{(P_8, L_8)\}$ и $\{(P_9, L_9)\}$, но также принадлежит горизонтальному уровню с независимыми ущербами $\{(P_5, L_5)\}$ и $\{(P_6, L_6)\}$, его вероятность P_7 будет учтена на следующем шаге. Аналогичным образом будем действовать ниже для узлов $\{(P_4, L_4)\}$ и $\{(P_2, L_2)\}$.

Шаг 2.

$$L^{est2} = L_4 + P_5\cdot(1-P_6)\cdot(1-P_7)\cdot L_5 + (1-P_5)\cdot(1-P_7)\cdot P_6\cdot L_6 + (1-P_5)\cdot(1-P_6)\cdot P_7\cdot L^{est1} +$$
$$+P_5\cdot P_6\cdot(1-P_7)\cdot(L_5+L_6) + P_5\cdot P_7\cdot(1-P_6)\cdot(L_5+L^{est1}) + P_6\cdot P_7\cdot(1-P_5)\cdot(L_6+L^{est1}) +$$
$$+P_7\cdot P_8\cdot P_9\cdot(L_5+L_6+L^{est1}).$$

Шаг 3.

$$L^{est3} = L_2 + \frac{P_4\cdot L^{est2}}{(1-P_4)+P_4}.$$

Шаг 4.

$$L^{est} = P_1 \cdot (1-P_2)\cdot(1-P_3)\cdot L_1 + (1-P_1)\cdot P_2 \cdot (1-P_3)\cdot L^{est3} + (1-P_1)\cdot(1-P_2)\cdot P_3 \cdot L +$$
$$+P_1 \cdot P_2 \cdot (1-P_3)\cdot(L_1 + L^{est3}) + P_1 \cdot P_3 \cdot (1-P_2)\cdot(L_1 + L_3) + P_2 \cdot P_3 \cdot (1-P_1)\cdot(L^{est3} + L_3) +$$
$$+P_1 \cdot P_2 \cdot P_3 \cdot (L_1 + L^{est3} + L_3).$$

Использование ожидаемого значения ущерба для того или иного события после свертки всего дерева ущербов позволяет получить некоторую обобщенную оценку величины возможного ущерба от данного события. Эту оценку можно будет использовать в дальнейшем при анализе риска потери доступности при возникновении данного события. Кроме того, для любого события *r*, связанного с потерей доступности ПО, наряду с ожидаемым значением ущерба L_r^{est} логично ввести понятие максимального значения ущерба $L_r^{\max}$, которое рассчитывается при условии, что вероятность любого из составляющих ущербов события *r* равна единице.

Градации ущербов и вероятностей

Для выполнения операций сложения и сравнения величин дискретных вероятностных ущербов необходимо, чтобы их величины были одной размерности. В то же время природа ущерба может быть самой разной (финансовой, имиджевой, временной и т.д.), поэтому свести имиджевые, временные потери или санкции надзорных органов (отбор лицензии, запрет на выполнение определенных видов деятельности и т.д.) к некоторой заранее рассчитанной денежной величине может быть крайне затруднительно. При невозможности расчета абсолютных величин ущербов в денежном выражении предлагается перейти от абсолютных значений величин ущерба к некоторым условным значениям, тем самым упростив данную модель.

Представим ущерб в некоторых условных единицах, соответствующих степени его существенности для нас, например, очень низкий, низкий, средний, высокий и очень высокий. Для того чтобы с величинами ущербов можно было выполнять операции сложения и умножения, поставим в соответствие величинам ущербов числовой ряд {1, 2, 3, 4, 5} так, что числу 5 соответствует максимальный ущерб (табл. 1).

Очень низкий (ОН)	Низкий (Н)	Средний (С)	Высокий (В)	Очень высокий (ОВ)
1	**2**	**3**	**4**	**5**

Таблица 1. Градации величин ущербов

Перейдем ко второй компоненте ущерба – вероятности. Поступим так же, как поступали при оценке величины ущерба: введем градации вероятности. Будем использовать три уровня вероятности ущерба: низкую (Н), среднюю (С), высокую (В), поставив им в соответствие числовой ряд {0,3; 0,6; 0,9}[2] (табл. 2).

Низкий (Н)	Средний (С)	Высокий (В)
0,3	**0,6**	**0,9**

Таблица 2. Градации вероятностей ущербов

После введения градаций величин ущербов и их вероятностей можно перейти к оценке ущерба в целом как пары «вероятность–величина», используя механизм расчета расчетного значения ущерба, описанный в предыдущем разделе.

Ущерб и продолжительность потери доступности

При построении дерева ущербов представляется целесообразным учитывать возможное изменение вероятности возникновения или величины того или иного ущерба от продолжительности потери доступности системы. Приведем некоторые примеры в подтверждение этого.

1. Вероятность вынесения штрафных санкций за задержку при сдаче отчетности равна единице вне зависимости от времени задержки. Величина штрафа зависит от длительности задержки: к основной сумме добавляются пени. В данном случае от времени зависит только величина ущерба.
2. Вероятность перехода того или иного клиента на другую систему зависит от времени неработоспособности. Величина ущерба при уходе одного клиента не зависит от времени и является фиксированной величиной.

[2] Предложенное авторами разбиение представляется удобным с точки зрения программной реализации. Можно использовать и другую шкалу градаций величин ущербов и вероятностей, которую можно настраивать в зависимости от сложности эксплуатируемой инфраструктуры.

Перейдем теперь к описанию этих ситуаций в модели дерева ущербов. С точки зрения вероятности целесообразно рассмотреть две ситуации:

1. сохранение во времени вероятности, указанной на момент, когда ущерб может возникнуть (В–В–…–В);
2. рост вероятности ущерба при простое системы (Н–...–Н–С–…–С–В–…–В).

С точки зрения оценки эволюции величины ущерба во времени ситуация аналогична: величина ущерба либо такая же, как была указана на начальный момент времени существования данного ущерба, либо возрастает по мере простоя системы[3].

Для некоторых типов ущербов, таких как пени за просрочку штрафа, зависимость вероятности и величины ущерба от времени можно точно указать, зная величину штрафа и пени за просрочку. Для оценки других видов ущербов придется выдвигать предположения на основе экспертных оценок.

Таким образом, анализируя возможный рост вероятности и величины ущерба для каждого из ущербов в построенном дереве, можно перейти от статического изображения дерева ущербов к его динамическому изображению, указав возможную эволюцию во времени каждого из ущербов дерева. Стоит отметить, что как ожидаемое значение ущерба L_i^{est}, так и максимальный ущерб $L_i^{\max}$ тоже будут зависеть от продолжительности простоя системы.

Описание тестов предложенной методики

Для исследования эффективности предложенной методики использовалась тестовая ИТ-инфраструктура финансового-кредитного учреждения, на которой была установлена обновленная система дистанционного банковского обслуживания и проведены изменения инфраструктуры, необходимые для ее функционирования. Предметом исследования были выявление и предотвращение в максимально короткие сроки возможных проблем, которые могли возникнуть в результате изменений компонентов финансовой информационной системы.

Обновление затрагивало следующие компоненты системы: административный модуль (АМ), модуль рублевых платежей (РП), а также модуль обмена с внешними подсистемами (ОВП).

На основе анализа возможных проблем при обновлении каждого из указанных компонентов был получен список возможных негативных событий (уязвимостей), связанный с обновлениями этих компонентов (табл. 3).

Модуль	Уязвимость
АМ	Некорректная обработка прав существующих пользователей
АМ	Ошибки при заведении новых пользователей/изменении прав существующих пользователей
РП	Ошибки при заведении клиентом платежей нового формата ускоренной доставки
РП	Ошибки при заведении клиентом платежей старых форматов
ОВП	Ошибки при обработке платежей нового формата ускоренной доставки на стороне банка
ОВП	Ошибки при обработке платежей старых форматов на стороне банка

Таблица 3. Список возможных уязвимостей

При построении дерева ущербов для каждого из выявленных событий использовались условные, а не абсолютные значения ущербов. В результате анализа деревьев ущербов, построенных для случая возможной неработоспособности каждого из перечисленных компонентов, были получены значения ожидаемых L_i^{est} и максимальных $L_i^{\max}$ ущербов (табл. 4).

На основе полученных данных были сделаны следующие выводы. Значения обоих параметров L_i^{est} и $L_i^{\max}$ однозначно указывают на события с максимальными возможными ущербами, на предотвращение которых надо обратить особое внимание при тестировании: (1), (6), (4) (см. табл. 4). Вместе с тем, для событий с меньшими возможными ущербами параметр L_i (свертка всех ущербов для дерева данного события) дает более точную оценку и позволяет различать события с одинаковыми параметрами $L_i^{\max}$ – простая сумма всех ущербов для дерева (2), (3) (см. табл. 4). Несмотря на то, что эти различия незначительны, тестировать систему с целью предотвращения обоих событий нужно практически с одним и тем же приоритетом.

В целом на основании проведенного анализа можно сделать вывод о том, что в данном случае расчета параметра $L_i^{\max}$ (простой суммы ущербов по каждому из событий) вполне достаточно для проведения сравнительной оценки рискованности выявленных событий и выбора стратегии тестирования. Тем не менее, в связи с многообразием ситуаций, связанных с расчетом ущербов, порождаемых негативными

[3] Ситуации, при которых вероятность или величина ущерба снижаются по мере простоя системы, авторами не рассматриваются, так как не имеют практического смысла.

событиями, целесообразно анализировать как общую сумму ущербов $L_i^{\max}$, так и свертку величин ущербов по вероятности L_i^{est}, так как учет вероятностей ущербов дает дополнительную информацию для выбора стратегии тестирования.

Модуль	Уязвимость	Возможное время неработоспособности					
		1 час		4 часа		8 часов	
		L_i^{est}	$L_i^{\max}$	L_i^{est}	$L_i^{\max}$	L_i^{est}	$L_i^{\max}$
АМ	Некорректная обработка прав существующих пользователей (1)	3,51	6	8,79	12	18,49	29
АМ	Ошибки при заведении новых пользователей/ изменении прав существующих пользователей (2)	0,3	1	0,78	2	7,53	10
РП	Ошибки при заведении клиентом платежей нового формата ускоренной доставки (3)	0,3	1	0,63	2	6,26	10
РП	Ошибки при заведении клиентом платежей старых форматов (4)	3,32	6	8,78	12	12,2	15
ОВП	Ошибки при обработке платежей нового формата ускоренной доставки на стороне банка (5)	0,3	1	0,78	2	3,51	6
ОВП	Ошибки при обработке платежей старых форматов на стороне банка (6)	6,41	12	10,98	15	18,43	25

Таблица 4. Оценка ожидаемых и максимальных ущербов для выявленных уязвимостей

Заключение

В работе был рассмотрен метод прогнозирования риска потери доступности финансовых информационных систем на основе построения деревьев ущербов для выявленных возможных негативных событий. На основе полученных данных созданы предпосылки для выбора рациональной стратегии тестирования программного обеспечения за счет детального учета информации, необходимой для принятия обоснованных решений. Проведены эксперименты, результаты которых подтверждают эффективность предложенной методики.

Литература

1. Арустамов С.А., Генин М.Г. Минимизация рисков потери доступности программного обеспечения после установки обновлений или изменения функциональности // Научно-технический вестник СПбГУ ИТМО. – 2011. – № 3 (73). – С. 111–116.
2. Генин М.Г. Методика проведения обновлений, снижающая риски доступности информационных ресурсов. Информатика, моделирование, автоматизация проектирования // Сборник научных трудов под ред. Н.Н. Войта. – Ульяновск: УлГТУ, 2011. – С. 125–137.
3. Арустамов С.А., Генин М.Г. Методы снижения рисков потерь доступности программного обеспечения критических информационных систем // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – № 4 (80). – С. 131–136.

Арустамов Сергей Аркадьевич – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, профессор, sergey.arustamov@gmail.com

Генин Михаил Геннадьевич – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, gmg@rambler.ru

9 ЭКОНОМИКА И ФИНАНСЫ. МЕНЕДЖМЕНТ

УДК 51.7, 519.86, 519.68

ПРИМЕНЕНИЕ ТЕОРЕТИЧЕСКОГО ЦЕНООБРАЗОВАНИЯ ОПЦИОНОВ В АВТОМАТИЧЕСКИХ ТОРГОВЫХ СТРАТЕГИЯХ

А.В. Торопов, Д.В. Иванов, Ю.А. Шполянский

Аналитическими и численными методами реализованы модели теоретического ценообразования европейских, американских, азиатских и барьерных опционов. Для случая многократных выплат дискретных дивидендов применен оригинальный вычислительный подход. На основе этих моделей разработаны автоматические торговые стратегии маркет-мейкинга, хеджирования, управления рисками и торговли комбинациями опционов, успешно используемые крупными банками и трейдинговыми фирмами.

Ключевые слова: теоретическое ценообразование, опцион, торговая стратегия, алгоритмическая торговля.

Введение

Алгоритмическая торговля (algorithmic trading) – процесс совершения сделок на финансовых рынках в автоматическом режиме с помощью специализированных компьютерных программ, реализующих различные торговые стратегии – алгоритмы принятия решений о совершении торговых операций. Такие программы называют автоматическими торговыми стратегиями.

Роль алгоритмической торговли в финансовых организациях неуклонно растет с каждым годом. В последнее десятилетие особенно популярной становится высокочастотная торговля (high-frequency trading). Этот вид алгоритмической торговли представляет собой многократное открытие и закрытие позиций по одним и тем же финансовым инструментам в течение торгового дня. Высокочастотная торговля получает активное развитие, в том числе благодаря своему положительному влиянию на рынок ценных бумаг: повышает ликвидность, привлекает малых инвесторов и обеспечивает лучшую цену при больших объемах сделки [1–3].

Для выполнения автоматических торговых стратегий необходима обширная программно-техническая инфраструктура – система для алгоритмической торговли, решающая такие задачи, как получение котировок с удаленных электронных рынков; создание и управление заказами; работа с портфелями ценных бумаг; обеспечение интерактивного графического взаимодействия с пользователем в условиях большого объема разнотипной информации и т.д. На современных биржах торгуются сотни тысяч финансовых инструментов, цены изменяются тысячи раз в секунду, а заключение сделок занимает микросекунды. Именно поэтому важнейшим требованием к системам алгоритмической торговли со стороны финансовых организаций является максимальное быстродействие на всех уровнях. Примером такой системы является коммерческий программный комплекс компании Tbricks AB [4].

Система Tbricks представляет собой распределенную систему с клиент-серверной архитектурой, разделенную на отдельные сервисы. Ее специфическими особенностями являются широкие функциональные возможности, покрывающие требования различных финансовых организаций; хорошая масштабируемость; эффективная работа на любом количестве серверов (имеется успешный опыт использования от 1 до 20 серверов с 8–24 CPU каждый); автоматическое перераспределение нагрузки (load-balancing).

Для реализации финансовой бизнес-логики в торговых стратегиях необходим инструментарий, позволяющий в автоматическом режиме принимать решения о посылке/изменении заказа на рынок или о заключении сделки. Один из широко распространенных подходов – использование теоретического ценообразования (theoretical pricing). Задачей настоящей работы было применение теоретического ценообразования для разработки автоматических торговых стратегий. На основе современных моделей теоретической оценки финансовых инструментов авторами созданы стратегии торговли опционами, маркет-мейкинга, хеджирования, управления рисками. Стратегии реализованы в виде плагинов для коммерческой системы алгоритмической торговли Tbricks.

Теоретическое ценообразование

Финансовые инструменты можно разделить на две большие группы – базовые активы и производные инструменты, или деривативы (derivatives). Цена производных инструментов зависит от цены и свойств базового актива. Базовыми активами могут быть акции, финансовые индексы или другие деривативы (например, фьючерсы). Одним из самых распространенных производных финансовых инструментов является опцион – соглашение на право купить (колл-опцион) или продать (пут-опцион) базовый инструмент по определенной цене (цена исполнения) и в определенное время (срок истечения) [5]. На современных биржах торгуется множество видов опционов, различающиеся деталями соглашения. Наиболее распространенными являются европейские и американские опционы [5]. Кроме этого, достаточно

популярны экзотические (exotic) опционы. Авторами рассмотрены азиатские и барьерные опционы [5–8], а также турбо варранты (turbo warrant) [9].

Основной задачей теоретического ценообразования является вычисление справедливой стоимости [6], которая определяет рациональную, свободную от арбитражных возможностей оценку потенциальной рыночной стоимости финансового инструмента, принимающую во внимание объективные рыночные факторы. Известны различные подходы к вычислению справедливой стоимости, приводящие к аналитическим или численным методам расчета [5–9].

Наиболее известной является модель Блэка–Шоулза–Мертона, которая описывает зависимость цены опциона $V(S,t)$ от цены базового актива S (спот-цена, от spot price) и времени t в виде дифференциального уравнения в частных производных (ДУЧП) [5, 6]:

$$\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}+rS\frac{\partial V}{\partial S}-rV=0\,, \qquad (1)$$

где r – безрисковая процентная ставка; σ – волатильность.

Кроме справедливой стоимости, важную роль в теоретическом ценообразовании играют ее частные производные [5, 6], например: «дельта» – $\Delta=\partial V/\partial S$; «гамма» – $\Gamma=\partial^2 V/\partial S^2$; «тета» – $\theta=\partial V/\partial t$; «вега» – $\nu=\partial V/\partial\sigma$ и др., показывающие меру чувствительности к изменению различных рыночных параметров. В финансовой математике их называют греческими производными, так как они обозначаются буквами греческого алфавита.

При некоторых видах начальных и граничных условий, связанных со свойствами опционов, уравнение (1) имеет аналитические решения. Нами использованы аналитические формулы для оценки справедливой стоимости европейских [5, 6, 8] и азиатских [7] опционов, а также турбо варрантов [9]. Эти формулы корректны при отсутствии выплат дискретных дивидендов по базовому активу.

В общем случае найти точное решение не удается, и на первый план выходят численные методы [5, 6, 10, 11]. Для американских опционов использовалась биномиальная модель [10]. Для остальных видов опционов при наличии дискретных дивидендов с датами выплат $t=t_{d,k}$, $k=1..K, K\geq 1$ был реализован оригинальный подход, схема которого показана на рис. 1. Временной промежуток I от даты исполнения T до даты выплаты последнего дивиденда $t_{d,K}$ (в сторону уменьшения t) покрывается аналитической формулой для цены опциона без дивидендов. Далее на участке II от $t_{d,K}$ до текущего момента $t=0$ применяется схема Кранка–Николсона численного решения ДУЧП (1) [6].

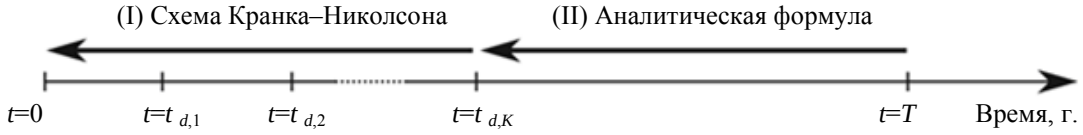


Рис. 1. Метод расчета цен опционов с многократными дискретными дивидендами

К достоинствам этого метода можно отнести сокращение размера расчетной области, уменьшение погрешности за счет применения точных аналитических решений на участке I и использования гладкого терминального условия при $t=t_{d,K}$ на участке II. Подобраны параметры сетки, обеспечивающие необходимую точность (относительная ошибка менее 10^{-4} по сравнению с аналитическим решением при обнулении дивидендов) при высокой скорости расчета (менее 200 мкс): 200 шагов сетки по цене базового актива и 50 шагов по времени на один год. Анализ для случая азиатских опционов приведен в работе [12]. Детальному описанию применения метода для других видов опционов планируется посвятить отдельную работу.

Перечисленные методы интегрированы нами в систему высокочастотной алгоритмической торговли компании Tbricks AB [4]. На основе полученной функциональности разработан ряд торговых стратегий, речь о которых пойдет в следующих разделах.

Маркет-мейкинг

Маркет-мейкер (market-maker) – один из важнейших участников биржевой торговли, берущий на себя риски приобретения и хранения ценных бумаг для организации их продаж. Маркет-мейкер заключает с биржей контракт, согласно которому он обязан держать активные заказы на покупку и продажу опционов (квотировать, от to quote) в течение всего торгового времени. Такие соглашения оговаривают группы квотируемых опционов, максимальную разницу между ценой покупки и продажи (спрэд, от spread) и другие условия. От биржи маркет-мейкер получает либо денежную компенсацию, либо уменьшенную комиссию на заключение сделок, что дает ему возможность вести торговую деятельность с меньшими издержками. Маркет-мейкеры повышают ликвидность квотируемых инструментов и обеспечивают цену, близкую к лучшей на рынке, при больших объемах сделок (увеличивают глубину рынка).

Главным вопросом при создании стратегии маркет-мейкера становится выбор цен для котировок, так как торговля по необоснованным ценам может принести серьезные убытки.

Естественное решение здесь – базирование котировок на теоретической стоимости, так как это позволяет правильно застраховать свои риски путем хеджирования сделок (см. далее раздел «Хеджирование»). Алгоритм, реализованный нами, проиллюстрирован на рис. 2. Базовый инструмент торгуется на рынке по ценам покупки S_b и продажи S_a. Соответствующие теоретические цены покупки V_b и продажи V_a производного инструмента раздвигаются в меньшую и большую сторону до значений M_b и M_a соответственно, чтобы обеспечить задаваемый маркет-мейкером спрэд.

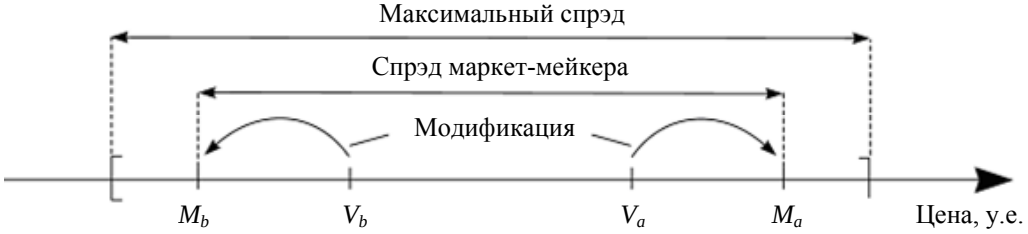


Рис. 2. Алгоритм выбора цен для квотирования

Важным требованием при реализации такого подхода является высокая скорость расчета теоретической стоимости при изменении цены базового актива. Задержки в расчете приводят к квотированию инструментов по устаревшим ценам и соответственно могут принести убытки маркет-мейкеру. На современных финансовых рынках на одном активе может базироваться до тысячи опционов с разными ценами и сроками исполнения. Наша реализация обеспечивает пересчет квот на все опционы, базированные на одном активе, за 300–500 мкс, что, по отзывам финансовых организаций, соответствует современным требованиям к системам высокочастотной торговли.

Теоретические значения предварительно рассчитываются для некоторого диапазона вокруг текущей цены базового актива и кэшируются в памяти. Это позволяет достичь указанной выше скорости пересчета квот вне зависимости от видов опционов и наличия дивидендных выплат. При движении цены диапазон модифицируется в соответствии с ценой базового актива. Чтобы избежать постоянного изменения границ диапазона в связи с малыми колебаниями цены на рынке, был применен механизм реагирования только на существенные движения цены.

Хеджирование

Хеджирование (hedge) – это стратегия страхования рисков, связанных с изменением рыночных параметров, например, движением цены базового актива (дельта-хеджирование) или изменением волатильности (вега-хеджирование).

Рассмотрим опцион на покупку акции, дельта которой в настоящий момент составляет $\Delta_0 = \partial V / \partial S\big|_{S=S_0}$, где S_0 – текущая цена акции. Сформируем портфель Π ценных бумаг, продав один колл-опцион и купив Δ_0 единиц базового актива:

$$\Pi = \begin{cases} -1 \text{ колл-опцион} \\ +\Delta_0 \text{ базового актива} \end{cases}.$$

Справедливая стоимость P портфеля Π как функция цены актива S имеет вид $P(S) = -V(S) + \Delta_0 \cdot S$.

Разложим эту функцию в ряд Тейлора вблизи цены S_0:

$$P(S) = P(S_0) + \frac{\partial P}{\partial S}\bigg|_{S=S_0} \cdot (S - S_0) + \frac{1}{2} \cdot \frac{\partial^2 P}{\partial S^2}\bigg|_{S=S_0} \cdot (S - S_0)^2 + O\left[(S - S_0)^3\right].$$

Для нашего портфеля при $S = S_0$ имеем:

$$\Delta_P = \frac{\partial P}{\partial S}\bigg|_{S=S_0} = -\frac{\partial V}{\partial S}\bigg|_{S=S_0} + \Delta_0 = 0, \qquad \Gamma_P = \frac{\partial^2 P}{\partial S^2}\bigg|_{S=S_0} = -\frac{\partial^2 V}{\partial S^2}\bigg|_{S=S_0} = -\Gamma_0.$$

Видно, что при малых отклонениях цены S относительно S_0 цена портфеля в линейном приближении остается неизменной. Такой портфель называют дельта-нейтральным. В квадратичном приближении стоимость портфеля изменяется пропорционально $\Gamma_P = -\Gamma_0$, т.е. при $\Gamma_P > 0$ стоимость P возрастает независимо от знака $(S - S_0)$, что выгодно для инвестора.

При существенных отклонениях цены S портфель перестает быть дельта-нейтральным, так как $\Delta = \partial V / \partial S$ становится заметно отличным от Δ_0. Чтобы восстановить свойство дельта-нейтральности, необходимо изменить позицию по базовому активу. Такой процесс называют динамическим хеджированием. Важно отметить, что слишком частая балансировка невыгодна из-за затрат на транзакционную стоимость сделок. Эти затраты учитываются при определении оптимальной частоты корректировки портфеля.

Если Γ_P велика по абсолютному значению, то Δ_P существенно изменяется даже при относительно небольших колебаниях цены S (так как $\Gamma_P = \partial \Delta_P / \partial S$), что приводит к необходимости частой балансировки портфеля. В связи с этим для инвестора предпочтительнее иметь Γ_P положительной, но ограниченной по величине. Обеспечить требуемое значение Γ_P позволяют дельта-нейтральные комбинации, которые состоят из актива и набора опционов, базированных на нем. Дельта такой комбинации приблизительно равна нулю, а гамма имеет известное значение. Покупая/продавая на рынке дельта-нейтральные комбинации, инвестор добивается желаемого значения Γ_P.

В системе Tbricks эту логику реализуют специальные стратегии: хеджирование, создание дельта-нейтральных комбинаций, балансировка портфеля.

Управление рисками

Важной задачей для инвестора является оценка потенциальной прибыли/убытков портфеля ценных бумаг. Эту задачу относят к управлению рисками. Используя методы теоретического ценообразования, можно прогнозировать, как будет меняться теоретическая стоимость портфеля в зависимости от рыночных параметров, например, цен базовых активов и их волатильностей, календарного времени и т.д.

Справедливая стоимость портфеля ценных бумаг *FMR* (fair market result) вычисляется по формуле

$$FMR = \sum_j \left(P_j \cdot V_j - I_j \right),$$

где индекс j пробегает по всем позициям портфеля, V_j – справедливая стоимость инструмента в позиции, P_j – объем позиции, а I_j – количество средств, инвестированных в позицию. На рис. 3 для примера приведены зависимости *FMR*-портфеля, аналогичного портфелю Π из раздела «Хеджирование», от цены базового актива (рис. 3, а) в диапазоне $[S_0 - 10\% S_0;\ S_0 + 10\% S_0]$, где S_0 − текущая цена, и волатильности (рис. 3, б) в диапазоне $[\sigma_0 - 15\%;\ \sigma_0 + 15\%]$, где σ_0 – текущее значение волатильности в процентах.

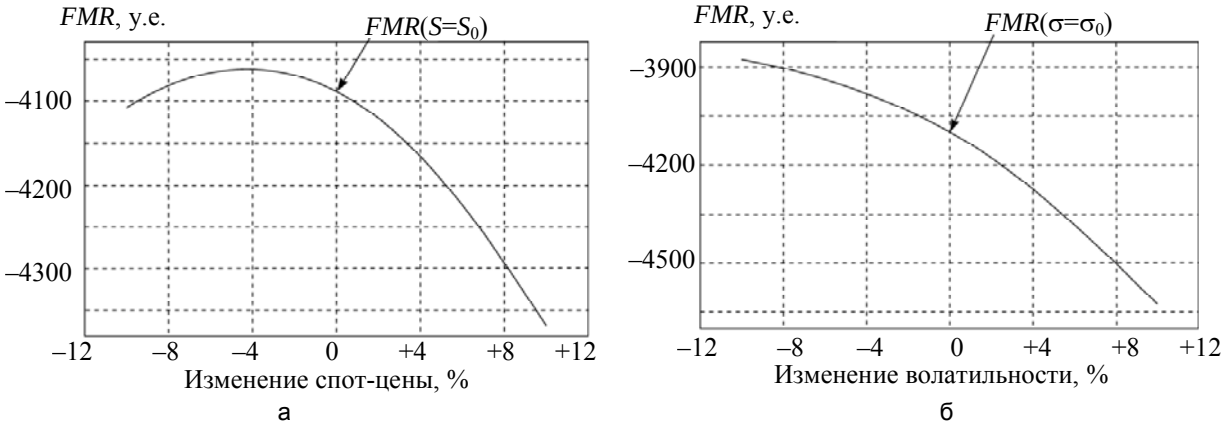


Рис. 3. Зависимости справедливой стоимости портфеля от спот-цены (а) и волатильности (б)

Существенной проблемой при построении таких зависимостей является большой объем вычислений. Портфели крупных трейдинговых компаний состоят из десятков и даже сотен тысяч позиций, а для построения детальных графиков необходимо рассчитать значения функции *FMR* для каждой позиции в большом числе точек. Трудоемкость создания многомерных прогнозов еще увеличивается на порядки. В системе Tbricks с помощью различного рода оптимизаций технического характера нам удалось обеспечить построение графиков функции *FMR* для больших портфелей в пределах 20–30 с. Такой результат позволяет многократно оценивать риски в течение торгового дня.

Торговля комбинациями опционов

Комбинации опционов являются отдельным видом финансовых инструментов, над которыми можно проводить торговые операции. Простейший пример таких инструментов – комбинации стрэддл (straddle) и стрэнгл (strangle). Комбинация стрэддл состоит из колл- и пут-опционов с одинаковым сро-

ком истечения и ценой исполнения $X = S_0$, где S_0 – цена базового актива в момент покупки комбинации. Комбинация стрэнгл состоит из пут- и колл-опционов с одинаковым сроком истечения, но разными ценами исполнения ($X_P < S_0 < X_C$). На рис. 4 изображены диаграммы доходности комбинаций стрэддл (рис. 4, а) и стрэнгл (рис. 4, б). Пунктирной линией показан доход в момент истечения опционов ($t = T$), сплошной – до истечения опционов ($t < T$).

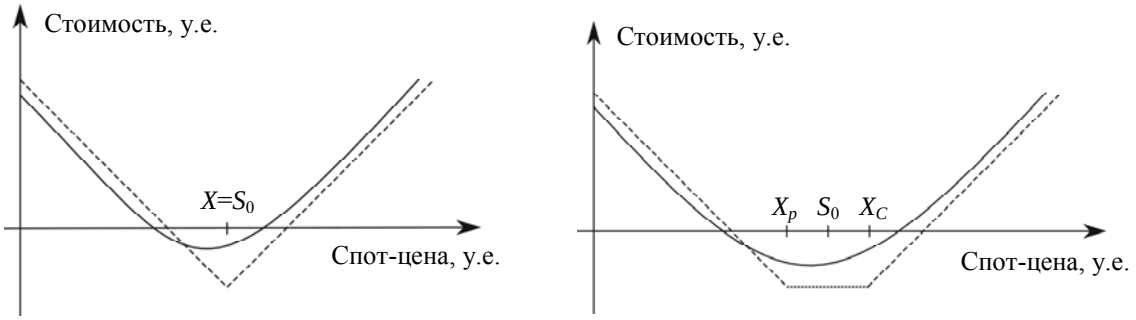


Рис. 4. Диаграммы доходности комбинации стрэддл (а) и стрэнгл (б). Пунктирная линия – доход в момент $t = T$, сплошная – доход в момент $t < T$

Из рис. 4 видно, что стоимость комбинаций имеет минимум вблизи цены S_0, при которой она приобретается. Это означает, что при движении цены базового актива относительно минимума в любую сторону справедливая стоимость комбинации возрастает. В этом случае можно ожидать роста рыночной стоимости комбинации, что выгодно для инвестора. Такой анализ, основанный на теоретическом ценообразовании, можно применять и к более сложным комбинациям.

В системе Tbricks при участии авторов была реализована функциональность, позволяющая формировать любые комбинации опционов, осуществлять торговые операции над ними на рынке, рассчитывать их теоретическую стоимость и анализировать ее поведение в зависимости от различных параметров.

Заключение

В работе аналитическими и численными методами реализованы различные модели теоретического ценообразования европейских, американских, азиатских, барьерных опционов и турбо варрантов. Подобраны параметры алгоритмов, обеспечивающие необходимую точность расчета (относительная погрешность менее 10^{-4}) при наименьших затратах вычислительных ресурсов (менее 200 мкс на расчет с дискретными дивидендами). Создан вариант коммерческой библиотеки для вычисления теоретических цен. На ее основе разработан ряд автоматических торговый стратегий:

– маркет-мейкинг;
– хеджирование;
– управление рисками;
– торговля комбинациями опционов.

Стратегии интегрированы в систему высокочастотной алгоритмической торговли Tbricks и успешно применяются клиентами – крупными европейскими банками, трейдинговыми фирмами и хедж-фондами.

Авторы выражают признательность коллегам J. Hansbo, Е.Ю. Шевкоплясу, К.Н. Романову за плодотворное сотрудничество по тематике работы и полезные замечания.

Литература

1. Gomber P., Arndt B., Lutat M., Uhle T. High-Frequency Trading // Goethe University Press. – 2011. – P. 1–82.
2. Hendershott T., Jones C., Menkveld A. Does Algorithmic Trading Improve Liquidity? // The Journal of finance. – 2010. – V. 66. – P. 1–33.
3. Chlistalla M. High-frequency trading // Deutsche Bank Research. – 2011. – P. 1–19.
4. Tbricks AB. Official website [Электронный ресурс]. – Режим доступа: http://www.tbricks.com, свободный. Яз. англ. (дата обращения 10.11.1012).
5. Hull J.C. Options, futures and other derivatives. – 7-th ed. – Prentice-Hall International, Inc, 2009. – 814 p.
6. Wilmott P. Paul Willmott on Quantitative Finance. – 2-d ed. – John Wiley & Sons, Ltd., 2006. – V. 1–3. – 1500 p.
7. Curran M. Valuing Asian and portfolio options by conditioning on the geometric mean price // Management Science. – 1994. – V. 40. – № 12. – P. 1705–1711.
8. Haug E.G. The complete guide to option pricing formulas. – 2-d ed. – McGraw-Hill, 2007. – 536 p.
9. Eriksson J. Explicit Pricing Formulas for Turbo Warrants // submitted to Risk Magazine. – 2005. – P. 1–6.

10. Vellekoop M.H., Nieuwenhuis J.W. Efficient pricing of derivatives on assets with discrete dividends // Applied Mathematical Finance. – 2006. – V. 13. – № 3. – P. 265–284.
11. Ikonen S., Toivanen J. Pricing American options using LU decomposition // Applied Mathematical Sciences. – 2007. – V. 1. – № 51. – P. 2529–2551.
12. Косяков М.С., Пономарев М.В., Иванов Д.В., Шполянский Ю.А. Методы расчета цен опционов азиатского типа с дискретными дивидендами вне периода усреднения // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – № 5 (81). – С. 143–147.

Торопов Александр Владимирович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант; Tbricks AB, инженер-программист; tav@tbricks.com, toropov@rain.ifmo.ru

Иванов Дмитрий Владимирович – Tbricks AB, математик, dm.vl.ivanov@gmail.com

Шполянский Юрий Александрович – Tbricks AB, ведущий математик; Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доцент; доктор физ.-мат. наук, доцент; shpolyan@mail.ru

10 НОВЫЕ ТЕХНОЛОГИИ В ОБРАЗОВАНИИ

УДК 004.9

ПЛАНИРОВАНИЕ И ОЦЕНИВАНИЕ ОЖИДАЕМЫХ РЕЗУЛЬТАТОВ ОСВОЕНИЯ КОМПЕТЕНЦИЙ ФГОС ВПО

В.Н. Васильев, Л.С. Лисицына

Предложен подход планирования и оценивания результатов освоения компетенций ФГОС ВПО в контексте развития информационно-образовательной среды вуза, основанный на структурировании содержания дисциплины под формирование результатов освоения составляющих компетенций стандарта и принципе междисциплинарного обучения. Проанализирована связь результатов освоения компетенций ФГОС ВПО с видами и формами занятий студентов, а также с видами и формами их контроля в дисциплине. Рассмотрены методики разработки оценочных средств для текущего, рубежного и промежуточного контроля ожидаемых результатов освоения компетенций ФГОС ВПО. Приведены примеры использования разработанных методик.

Ключевые слова: ожидаемые результаты освоения компетенций ФГОС ВПО, компетентностная модель выпускника, методика разработки фонда оценочных средств дисциплины.

Введение

С переходом на федеральные государственные стандарты высшего профессионального образования (ФГОС ВПО) вузы России приступили в 2011 г. к реализации компетентностно-ориентированных основных образовательных программ (ООП) для уровневой подготовки выпускников (бакалавр, магистр). Отличительной особенностью ФГОС ВПО является то, что в них впервые государство устанавливает квалификационные требования к выпускникам вузов в виде общекультурных и профессиональных компетенций (ОК и ПК), перечень которых в различных ФГОС ВПО составляет 30–60 формулировок. Все компетенции ФГОС ВПО являются рамочными, т.е. позволяют различным выпускающим кафедрам вузов проектировать компетентностные модели выпускников своих ООП сообразно состоянию и тенденциям развития научных школ вузов, потребностей потенциальных работодателей не только в своем регионе, но и в целом по России и в мире. Это обстоятельство открывает широкие возможности для разработчиков ООП, конкурентоспособность которых определяется спросом на рынке образовательных услуг.

Многие вузы перестраивают сейчас свои информационно-образовательные среды в соответствии с требованиями ФГОС ВПО. Анализ основных задач развития информационно-образовательной среды вуза в связи с переходом на ФГОС ВПО был дан авторами в работе [1], в которой подробно изложен сетевой подход к автоматизации планирования ожидаемых результатов освоения компетенций ФГОС ВПО и управления на их основе компетентностными моделями выпускников ООП; в соответствии с требованиями ФГОС ВПО такие модели должны ежегодно пересматриваться вузом совместно с потенциальными работодателями выпускников ООП.

Настоящая работа посвящена проблеме оценивания ожидаемых результатов освоения компетенций ФГОС ВПО. Основной акцент в работе сделан на связь конкретных результатов освоения компетенций ФГОС ВПО и оценочных средств, главное предназначение которых – контроль учебных достижений студентов, ожидаемых в дисциплине ООП.

Компетентностная модель выпускника ООП и уровни ее представления

Компетентностная модель выпускника ООП, реализующей ФГОС ВПО, в процессе ее разработки и модернизации описывается на следующих уровнях представления [2]:

1. составные компетенции ФГОС ВПО с установленными минимальными (ожидаемыми) уровнями формирования их компонентов;
2. составляющие компетенции ФГОС ВПО в дисциплинах ООП;
3. результаты освоения (компоненты) компетенций ФГОС ВПО (конкретные знания, умения (владения), личностные качества выпускников).

Принцип междисциплинарного обучения при разработке составляющих компетенций ФГОС ВПО в дисциплине ООП

Среди компетенций ФГОС ВПО из задания на разработку рабочей программы дисциплины ООП [3] вначале следует выделить основные, дополнительные и сопутствующие компетенции стандарта. Основные компетенции непосредственно связаны с содержанием предметной области изучения дисциплины, а дополнительные компетенции, как правило, являются компетенциями из других предметных областей обучения (из других дисциплин), но частично формируются на содержании данной дисциплины. В этом случае в дисциплине планируются междисциплинарные результаты освоения компетенций ФГОС ВПО. Это означает, что на едином содержании отдельного раздела дисциплины можно формировать ре-

зультаты освоения различных компетенций ФГОС ВПО. При этом ведущей компетенцией, как правило, выступает основная, а ведомой – дополнительная компетенция ФГОС ВПО. Результаты освоения сопутствующих компетенций ФГОС ВПО, например, навыки работы в команде, ведения дискуссий и т.п., формируются в дисциплине за счет применяемых технологий обучения и преподавания

Пример 1. Разработка составляющих компетенций в дисциплине «Дискретная математика». «Дискретная математика» является вариативной дисциплиной математического и естественнонаучного цикла дисциплин в составе ООП по подготовке бакалавров направления 230400 – «Информационные системы и технологии» и служит основой для формирования результатов освоения следующих компетенций ФГОС ВПО.

Индекс компетенции	Компетенция ФГОС ВПО в дисциплине	Ожидаемый уровень освоения компетенции*
ОК-6	Владение широкой общей подготовкой (базовыми знаниями) для решения практических задач в области информационных систем и технологий	З2-У2-СЛ1
ПК-12	Способность разрабатывать средства реализации информационных технологий (методические, информационные, математические, алгоритмические, технические и программные)	З1-У1-СЛ1

* – установлены в соответствии с действующим в НИУ ИТМО тарификатором результатов обучения [3]

Таблица 1. Компетенции ФГОС ВПО и ожидаемые уровни их формирования в дисциплине «Дискретная математика»

В связи с этим целью дисциплины «Дискретная математика» (табл. 1) является формирование результатов освоения следующих составляющих компетенций выпускника ООП, разработанных на основе принципа междисциплинарности обучения (здесь компетенция ОК-6 является основной, а компетенция ПК-12 – дополнительной):

- способность применять теорию множеств для разработки математических и алгоритмических средств информационных технологий;
- способность применять теорию графов для разработки математических средств информационных технологий;
- способность применять методы решения оптимизационных задач на графах для разработки алгоритмических средств информационных технологий;
- способность применять теорию сетей Петри для разработки математических средств информационных технологий;
- способность применять транспортные сети для разработки математических средств информационных технологий;
- способность применять гиперграфы для разработки математических средств информационных технологий.

Формированию результатов освоения каждой составляющей компетенции будет соответствовать отдельный раздел в содержании данной дисциплины [3].

Компоненты для формирования компетенций ФГОС ВПО

Формирование компетенций осуществляется через освоение знаний, приобретение навыков и развитие личностных качеств (ответственности, дисциплинированности, организованности и т.п.) в ходе учебного процесса. Исходя из этого, любая компетенция ФГОС ВПО является результатом синтеза (соединения на практике) в процессе обучения следующих трех взаимосвязанных компонентов.

1. Знаниевый компонент (знания), характеризующий знания обучающегося как основу для осуществляемой деятельности.
2. Функциональный компонент (умения и навыки), характеризующий умения обучающегося выполнять определенные действия и навыки владения определенными инструментами, технологиями, оборудованием и т.п. в осуществляемой деятельности.
3. Социально-личностный компонент, опирающийся на личностные качества обучающегося и характеризующий его отношение к осуществляемой деятельности (обучению).

Формулировки компонентов для формирования компетенций ФГОС ВПО в дисциплине должны содержать конкретные и диагностируемые в ходе контроля знания, умения, навыки и личностные качества.

Пример 2. Разработка компонентов для формирования компетенций ФГОС ВПО в дисциплине «Дискретная математика». Для формирования способности применять теорию множеств для разработки математических и алгоритмических средств информационных технологий (пример 1) сформулированы следующие знания, умения и навыки, ожидаемые у каждого студента на соответствующем уровне тарификатора результатов обучения (компетенций) [3].

Знания для формирования ОК-6:
1. правила построения четких и нечетких множеств;
2. правила построений и вычислений математических выражений на основе теории множеств;
3. правила определения расстояний между множествами в линейной метрике (расстояние Хемминга) и в метрике Евклида;
4. правила определения меры нечеткости и индексов нечеткости множеств;
5. правила построения прямых и обратных отношений на множествах, композиций на их основе;
6. правила построения функций (отображений) на множествах;
7. свойства специальных бинарных отношений.

Знания для формирования ПК-12:
1. представление о прямых и косвенных способах построения функции принадлежности нечеткому множеству;
2. проблемы разработки алгоритмов обработки нечеткой информации.

Умения для формирования ОК-6:
1. построить нечеткие множества и подмножества;
2. определять расстояния между множествами;
3. определять меры нечеткости и индексы нечеткости множеств;
4. строить бинарные (прямые, обратные) отношения на четких и нечетких множествах;
5. находить композицию бинарных отношений на четких и нечетких множествах;
6. определять свойства заданного специального бинарного отношения;
7. определять классы эквивалентности для множеств;
8. строить диаграммы Хассе для упорядоченных множеств;
9. определять функции на множествах.

Умения для формирования ПК-12:
1. разрабатывать простейшие алгоритмы обработки нечеткой информации.

Навыки для формирования ОК-6:
1. находить результат вычисления математических выражений на основе теории множеств;
2. сравнивать множества по нечеткости.

Принципы разработки фондов оценочных средств в дисциплинах ООП

Фонд оценочных средств (ОС) дисциплины – набор контрольно-измерительных материалов, необходимый преподавателю для контроля (фиксации, измерения и оценки) заявленных им в дисциплине результатов освоения компетенций ООП у каждого студента. Набор может включать в себя различные виды ОС, которые зависят от планируемых результатов освоения компетенций ФГОС ВПО в дисциплине (тесты, индивидуальные задания для контрольных, домашних, лабораторных и курсовых работ, типовые экзаменационные теоретические вопросы и практические задачи с наборами конкретных вариантов заданий к ним), а также набор измерителей для оценки результатов освоения компетенций ФГОС ВПО у студентов, например, критериев для оценки тестов, отчетов по различным работам и т.п. Фонд ОС дисциплины по составу должен удовлетворять принципу полного покрытия ожидаемых знаний, умений и навыков (ЗУН) и личностных качеств обучаемых, контролируемых в процессе изучения дисциплины.

Связь результатов освоения компетенций ФГОС ВПО с видами и формами занятий при изучении дисциплины

В дисциплинах ООП планируются два типа занятий – аудиторные и внеаудиторные занятия. К аудиторным занятиям относятся лекционные и практические занятия, а к внеаудиторным – самостоятельная работа студентов (СРС). *Лекционные занятия* проводятся в основном в форме лекций и мастер-классов. Лекция – это аудиторное занятие по передаче учебной информации от преподавателя к студентам, направленная в основном на приобретение студентами новых теоретических и фактических знаний. В отличие от обычной лекции, передача учебной информации в мастер-классе проводится с использованием компьютерных и других технических средств. *Практические занятия* проводятся в форме практикумов, лабораторных работ, семинаров, коллоквиумов и т.д.; все практические занятия направлены на приобретение навыков, а в ряде случаев и на приобретение новых теоретических и фактических знаний. Практикум – это практическое занятие под руководством преподавателя, в ходе которого выполняется решение конкретных задач (изучение и применение на практике конкретных методов, разработка алгоритмов, расчеты характеристик и моделирование объектов и т.п.). Лабораторная работа – практическое занятие студента под руководством преподавателя, связанное с использованием учебного, научного или производственного оборудования (конкретных приборов, устройств и т.п.), в котором обязательным этапом является проведение эксперимента с целью получения новых фактических знаний и практических умений. Семинар, коллоквиум – это практическое занятие под руководством преподавателя, в ходе которого выполняется систематизация теоретических и фактических знаний в определенном контексте (подготов-

ка и презентация материала по определенной теме, обсуждение ее, формулирование выводов и заключения). *Самостоятельная работа студентов* (СРС) – внеаудиторное занятие студентов, в ходе которого студенты могут приобретать новые теоретические и фактические знания, умения и навыки, а также демонстрировать их на практике. СРС может проводиться под управлением преподавателя, например, при изучении дополнительного теоретического материала студенты могут воспользоваться консультацией преподавателя, в том числе и дистанционной. Одной из форм управляемой СРС является работа над курсовым проектом. Курсовое проектирование – познавательная деятельность студента, связанная с выполнением проекта технического объекта, системы, прибора, технологии и др. (удовлетворяющего заданным требованиям при определенных ограничениях), направленная в основном на приобретение новых фактических знаний и навыков в синтезе (соединении) на практике. По этой причине курсовой проект является одним из основных испытаний сформированности компетенций ФГОС ВПО у студента. Другим вариантом такого испытания могут быть комплексные домашние задания, выполняемые также в рамках СРС. На рис. 1 показана связь видов и форм занятий студентов и формируемых в них результатов обучения дисциплины.

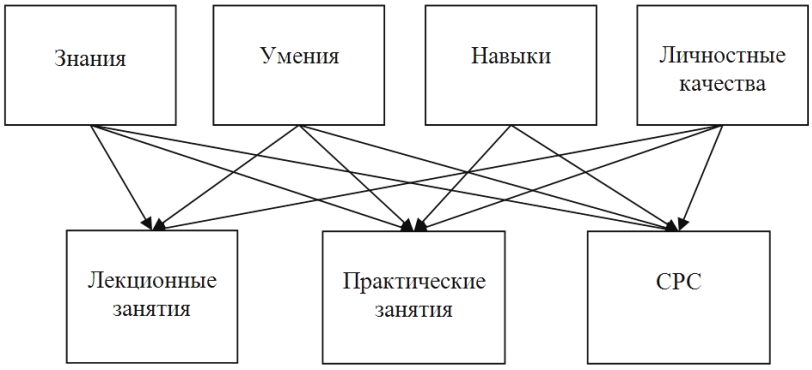


Рис. 1. Связь между формируемыми результатами освоения компетенций ФГОС ВПО и видами занятий в дисциплине

Связь результатов освоения компетенций ФГОС ВПО с видами и формами контроля в дисциплине

В дисциплине осуществляется текущий, рубежный и промежуточный контроль запланированных результатов освоения компетенций ФГОС ВПО.

Для проведения *текущего контроля* рекомендуется проводить опросы, в том числе в виде компьютерных тестов. При разработке тестовых заданий рекомендуется, в первую очередь, проверять умение применять знание на практике, например, переводить числовые данные из одной системы счисления в другую, кодировать числа, символьные данные и изображения и т.п. Если соответствующее задание будет выполнено студентом неверно, рекомендуется задавать дополнительный вопрос, который должен установить, знает ли студент соответствующее содержание, например, правило перевода чисел из одной системы счисления, соответствующий способ кодирования данных и т.п. Если в данном разделе запланированы практические (лабораторные) работы, то контроль знаний и навыков (многократно закрепленных умений с применением конкретного инструмента, метода, оборудования и т.д.) может осуществляться при сдаче и защите отчетов по лабораторным работам, домашних заданий и т.п. Оценка личностных качеств студентов, проявленных в ходе текущего контроля, проводится по критериям, установленным как дополнительное условие (сдача тестов, лабораторных и контрольных работ и т.п. в установленный срок; умение работать в команде, умение вести диалог и т.п.).

Для проведения *рубежного контроля* рекомендуется проводить контрольные работы, цель которых – проверка ЗУН, запланированных в разделах дисциплины для данного модуля. При этом оценка их сформированности должна быть подтверждена на практике в ограниченное время. Задания для контрольных работ должны содержать теоретические вопросы, для ответа на которые от студента потребуется, например, обоснование выбора способа (подхода) решения поставленной задачи; поиск и обоснование ошибок, допущенных в решении поставленной задачи известным методом; аргументированное обоснование того метода, который был применен для решения поставленной задачи; и т.п. Практические задания для рубежного контроля должны быть направлены на проверку сформированности результатов составляющих компетенций ФГОС ВПО.

Промежуточный контроль проводится в форме экзаменов и зачетов, цель которых – проверка сформированности результатов освоения компетенций ФГОС ВПО в данной дисциплине. Экзаменационные и зачетные задания делятся на две группы – теоретические вопросы и практические задачи. Целью теоретических вопросов является проверка способности студентов применять полученные знания в нестандартной ситуации, когда постановка конкретной проблемы требует применения полученных знаний и умений в их синтезе (соединении). При этом студенты должны демонстрировать в ограниченное время

умение анализировать проблему и предлагать способы ее решений. Практические задания должны позволять студенту демонстрировать полученные навыки в контексте компетенций ФГОС ВПО. На рис. 2 показана связь видов и форм контроля и оцениваемых в них результатов освоения компетенций ФГОС ВПО в дисциплине.

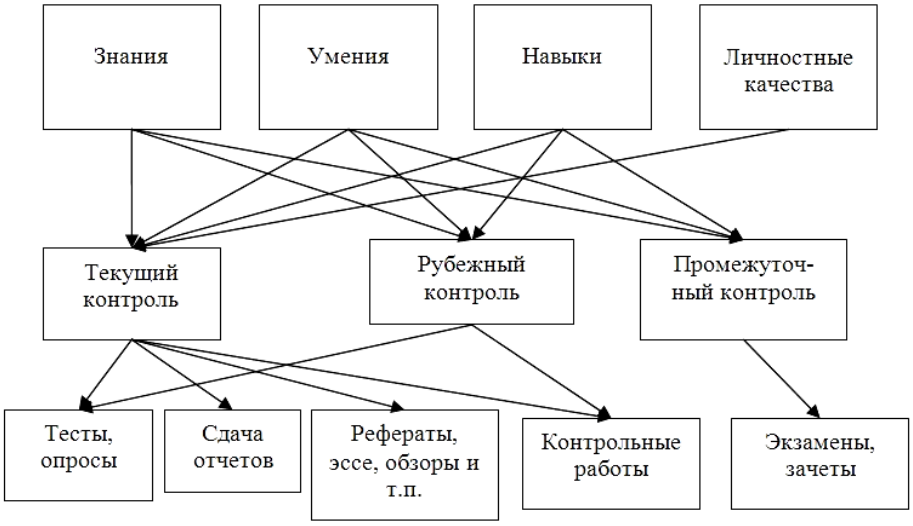


Рис. 2. Связь между оцениваемыми в них результатами освоения компетенций ФГОС ВПО и видами, формами контроля в дисциплине

Методика разработки фонда оценочных средств для текущего контроля

Структурированное содержание дисциплины (каждая составляющая компетенция ФГОС ВПО соответствует отдельному разделу дисциплины) определяет виды учебных занятий в зависимости от состава заявленных результатов освоения компетенций ФГОС ВПО (рис. 1). Формы их текущего контроля (рис. 2) преподаватель выбирает, исходя из отобранного содержания дисциплины, необходимого и достаточного для формирования запланированных результатов. Методика разработки фонда оценочных средств (ОС) для текущего контроля в разделе дисциплины состоит в следующем.

1. Запланировать формы и виды ОС для контроля результатов освоения компетенций, формируемых в разделе дисциплины.
2. Установить соответствие каждого вида ОС конкретным результатам освоения компетенций, формируемых в разделе дисциплины.
3. Разработать план проверки этих результатов.
4. Разработать по плану проверки конкретные ОС.

Пример 3. Разработка форм и ОС текущего контроля для формирования компетенций ФГОС ВПО в дисциплине «Дискретная математика». Для текущего контроля запланированных результатов освоения компетенций в дисциплине (пример 1) были разработаны формы контроля, представленные в табл. 2.

Номер раздела	Наименование раздела	Формы контроля		
		Знаний и умений	Навыков	Личностных качеств
1	Основы теории множеств	Компьютерный тест № 1	Сдача отчетов по лабораторным работам	Сдача в установленный срок
2	Основы теории графов	Компьютерный тест № 2	Сдача отчетов по лабораторным работам	Сдача в установленный срок
3	Оптимизация на графах	Компьютерный тест № 3	Сдача отчетов по лабораторным работам	Сдача в установленный срок
4	Основы теории сетей Петри	Компьютерный тест № 4	Сдача отчетов по лабораторным работам	Сдача в установленный срок
5	Транспортные сети	Сдача отчетов по лабораторным работам		Сдача в установленный срок
6	Основы теории гиперграфов	Компьютерный тест № 5		Сдача в установленный срок

Таблица 2. Формы текущего контроля в дисциплине «Дискретная математика»

Рассмотрим далее на примере разработки ОС в форме компьютерного компетентностно-ориентированного теста № 3 (табл. 2) для текущего контроля знаний и умений, запланированных для формирования способности применять методы решения оптимизационных задач на графах для разработки алгоритмических средств информационных технологий (пример 1). В этом разделе запланированы знания типовых оптимизационных задач на графах и умение определять их оценки на практике на простейших графах. Покажем отличие содержательного и компетентностного подходов к контролю на примере задачи поиска наибольших пустых подграфов в графе, для оценки которой используется число внутренней устойчивости графа.

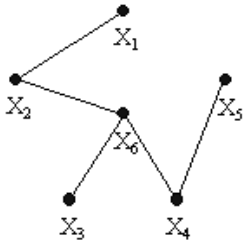
Тестовое задание закрытого типа для проверки знания содержания может выглядеть следующим образом.

Число внутренней устойчивости графа определяет

☐	число всех пустых подграфов.
☐	число всех полных подграфов.
☐	число вершин в наибольшем полном подграфе.
☐	число вершин в наибольшем пустом подграфе.

При компетентностном подходе требуется проверка знания на практике. Для этого можно предложить следующую цепочку из двух заданий. В цепочке второе задание предъявляется студенту после правильного ответа на первое задание.

Определить число внутренней устойчивости графа $\alpha_0(G)$.



Ответ:_____

Ответ правильный. Уточните, какое множество вершин соответствует найденному значению $\alpha_0(G)$?

☐	{X1, X3, X5}
☐	{X1, X2, X3}
☐	{X2, X6, X4}

Методика разработки фонда оценочных средств для рубежного и промежуточного контроля

Для подготовки оценочных средств рубежного и промежуточного контроля (рис. 2) разработчики рабочих программ дисциплин составляют типовые теоретические вопросы и типовые практические задачи для проверки запланированных конкретных результатов освоения составляющих компетенций ФГОС ВПО в дисциплине. Основной принцип разработки типовых вопросов и задач к разделу дисциплины – полное покрытие тех знаний, умений и навыков, которые ожидаются после завершения его изучения. В общем случае преподаватель может предложить не один, а несколько типовых вопросов и типовых задач. Каждый типовой теоретический вопрос должен проверять умение применять знание на практике.

Пример 4. Разработка типовых теоретических вопросов для рубежного и промежуточного контроля в дисциплине «Дискретная математика». Для формирования способности применять теорию множеств для разработки математических и алгоритмических средств информационных технологий было запланировано знание правил построения четких и нечетких множеств и умение строить четкие и нечеткие подмножества (пример 2). Для их проверки можно использовать следующие типовые вопросы (для каждого типового вопроса в фонде ОС подбираются конкретные задания).

1. В заданном универсальном множестве построить и изобразить графически нечеткое множество **A**, описанное с помощью характеристического свойства его элементов. Обосновать ответ.
2. В заданном универсальном множестве построено нечеткое множество **A** путем перечисления его элементов и степеней их принадлежности. Построить другое нечеткое множество **B**, которое будет подмножеством для **A**. Обосновать ответ.

3. В заданном универсальном множестве построено нечеткое множество **A** путем перечисления его элементов и степеней их принадлежности. Построить другое нечеткое множество **B**, которое будет более (менее) нечетким, чем множество **A**. Обосновать ответ.
4. В заданном универсальном множестве было построено нечеткое множество **A** путем перечисления его элементов и степеней их принадлежности. Студент построил другое нечеткое множество **B**, которое должно быть более (менее) нечетким, чем множество **A**. Проанализируйте ответ студента и аргументировано объясните возможные ошибки.

Пример 5. ОС в дисциплине «Дискретная математика» (практические задачи). Типовая практическая задача – «Задано множество **X**, на котором построено бинарное отношение R. Перечислите все элементы множества R, укажите степени их принадлежности. Определите свойства этого отношения. Обоснуйте свой ответ».

Варианты оценочных средств к данному заданию:

1. **X** = {10, 15, 19, 21}; R ={(x,y): (x,y)∈X*X; |x–y| приблизительно равно 10};
2. **X** = {10, 15, 19, 21}; R ={(x,y): (x,y)∈X*X; |x–y| приблизительно равно 5};
3. **X** = {2, 5, 9, 11}; R ={(x,y): (x,y)∈X*X; (x+y) намного меньше, чем x*y};
4. **X** = {2, 5, 9, 11}; R ={(x,y): (x,y)∈X*X; (x+y) немного меньше, чем x*y};
5. **X** = {2, 5, 9, 11}; R ={(x,y): (x,y)∈X*X; (x/y) намного меньше, чем x}.

Здесь знак «*» обозначает прямое произведение множества X самого на себя.

Заключение

Через два года вузы будут поставлены перед проблемой массового приема в магистратуру выпускников бакалаврских ООП. Уже есть мнение о том, что для этого необходимо проводить отбор лучших выпускников бакалаврских ООП на основе единого экзамена (бакалаврского минимума) в рамках направления подготовки. Каким должен быть такой экзамен? В данной работе мы изложили результаты наших исследований, которые, на наш взгляд, могут стать методической базой для разработки экзаменационных работ бакалаврского минимума, цель которого – дифференциация знаний, умений и навыков, сформированных в базовой части ООП данного направления подготовки в любом вузе России.

В работе предложены методики планирования результатов освоения компетенций ФГОС ВПО и разработки оценочных средств для их текущего, рубежного и промежуточного контроля на примере дисциплины «Дискретная математика». Методики разработаны в рамках подхода, основанного на структурировании необходимого и достаточного содержания дисциплины под формирование результатов освоения составляющих компетенций ФГОС ВПО в дисциплине в соответствии с принципом междисциплинарного обучения.

Литература

1. Васильев В.Н., Лисицына Л.С. Основные направления развития информационно-образовательной среды вуза в связи с переходом на ФГОС ВПО // Компьютерные инструменты в образовании – 2012. – № 4. – С. 3–9.
2. Пирская А.С. Методика оценивания компетенций выпускника // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – № 1 (77). – С. 135–141.
3. Лисицына Л.С., Лямин А.В., Шехонин А.А. Разработка рабочих программ дисциплин (модулей) в составе основных образовательных программ, реализующих ФГОС ВПО. Методическое пособие. – СПб: СПбГУ ИТМО, 2011.– 63 с.

Васильев Владимир Николаевич – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, профессор, ректор, rector@mail.ifmo.ru

Лисицына Любовь Сергеевна – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, доцент, зав. кафедрой, lisizina@mail.ifmo.ru

УДК 159.9; 004.81

ИНТЕЛЛЕКТНЫЕ СИМБИОНТЫ В ЭРГАТИЧЕСКИХ СИСТЕМАХ

С.Ф. Сергеев

Показано, что искусственный и естественный интеллекты связаны со сложностью организации наделенными ими систем в процессе их жизнедеятельности (функционирования) в среде. Разработана модель симбиотических объединений систем естественного и искусственного интеллектов, включенных в организованные и организуемые технические среды. Показано, что искусственные среды влияют на формирование и проявление интеллектуальных способностей погруженных в них пользователей в силу аутопоэтического характера возникающих межсистемных отношений. Введено понятие диффузного интеллекта как синергетического взаимообъединения искусственного и естественного интеллектов в организованной среде.

Ключевые слова: интеллект, искусственный интеллект, гибридный интеллект, диффузный интеллект, сложность, иммерсивные среды, коммуникация.

Введение

Информационные технологии предоставляют разработчикам эргатических систем широкий спектр средств, повышающих их интеллектуальность, формируя отношения человека-оператора с технической системой, подобные отношениям, возникающим в процессе естественной социальной коммуникации [1]. Вместе с тем отметим, что взаимодействие человека с искусственными информационными средами отличается от его взаимодействия с естественными средами в силу структурно-функциональной дополнительности искусственных сред с когнитивными структурами человека, что ведет к процессам непрерывной модификации психических функций оператора [2]. Возникает проблема симбиоза между системами разной природы – биологическими, наделенными механизмами психического отражения и активного целеполагания, и техническими, реализующими технологии искусственного интеллекта.

Наиболее часто эффективность функционирования сложных систем связывают с понятием «системный гибридный интеллект» [3], которое получает новые смыслы при включении человека в сложноорганизованные среды. Техногенные среды ведут к новым формам интеграции человека с машиной, формируя гибридные и симбиотические формы естественного интеллекта, отличающиеся от классического интеллекта, возникшего в процессе жизнедеятельности человека в изолированных культурных средах и естественной среде жизнедеятельности. Это ведет к постановке задачи направленной модификации интеллектуальной сферы человека-оператора и повышения его эффективности в процессе взаимодействия с компонентами искусственного интеллекта эргатической системы. Цель настоящей работы – рассмотрение возможных вариантов симбиотического объединения человека-оператора с техническими системами и средами, обладающими свойствами систем с искусственным интеллектом.

Интеллект в системно-психологических представлениях

Поиски оснований естественного интеллекта, а в более широком плане – человеческого ума, возможности создания их искусственных аналогов приобретают особое значение в связи с интенсивным развитием техники и технологий и пониманием ограниченности индивидуального человеческого разума [4, 5]. Развитие информационных технологий питает наблюдаемый в последнее десятилетие бум в области искусственного интеллекта.

Существование различий между людьми по умению решать задачи одного класса неоспоримо. Это позволило выдвинуть предположение о существовании в человеке особых ментальных структур, умственных способностей, отраженных в понятии «интеллект». Именно на этом предположении был основан самый известный классический тест для оценки коэффициента интеллекта (IQ, intelligence quotient) – шкала Д. Векслера, которая по настоящее время служит эталоном для оценки человеческого ума в его рациональных и культурно-лингвистических формах [6].

В современной интерпретации, предложенной в работах М.А. Холодной, интеллект – это форма организации когнитивного опыта человека, представленного в виде «накопленных» в ходе онтогенеза понятийных психических структур, степень сформированности которых определяет структурные характеристики субъективного пространства интеллектуального отражения. Основное назначение интеллекта по М.А. Холодной – построение особого рода репрезентаций происходящего, связанных с воспроизводством объективного знания о мире [7]. Отметим, что эта задача решается посредством создания и использования когнитивных инструментов [8, 9], которые должны быть достаточно удобны и эффективны для пользователя [10].

Будучи психологическими конструктами, проявляющимися при решении тестовых задач, показатели естественного интеллекта, несмотря на кажущуюся их «объективность», с трудом могут быть применены в технике при оценке систем искусственного интеллекта. Естественный интеллект имеет совсем другую природу, нежели «интеллектуальные возможности», заложенные в сложные технические системы их разработчиками.

В классических, главным образом, факторных моделях интеллекта рассматривается его адаптивная функция по отношению к окружающей и внутренней средам. Определение интеллекта как некоторой

способности, обусловливающей успешность адаптации индивида к новым условиям, является наиболее общим местом в моделях, использующих методологию классической научной рациональности [11].

В.Н. Дружининым сформулирован принцип, согласно которому трудность задачи определяет уровень интеллекта, необходимого для ее правильного решения [12, с. 14–25]. Фактор, обеспечивающий переработку сложной информации и детерминирующий индивидуальную продуктивность, В.Н. Дружинин называет «индивидуальным когнитивным ресурсом». Следствием этих положений является принцип: чем труднее задача, тем более высокий уровень развития интеллекта требуется для ее правильного решения. Вместе с тем заметим, что понятие «трудность задачи» не является синонимом понятия «сложность задачи». Трудность задачи связана с ее субъективным восприятием, интерпретацией и оценкой субъектом, который в процессе решения может как недооценивать, так и переоценивать объективную сложность задачи.

В классическом подходе среда рассматривается главным образом как отдельный фактор, влияющий на формирование и проявления интеллекта, а не как его неотъемлемая часть. Эта отделенность интеллекта от среды создает иллюзии его независимости и локализации в субстрате организма человека.

При изучении интеллекта никто не обращает внимания на его активный, организующий характер при формулировании задач, их оптимизацию под когнитивные средства и возможности, имеющиеся у конкретного обладателя интеллекта. Естественный интеллект активен и избирателен по своей природе. Он упрощает среду деятельности до уровня, позволяющего ее операционализировать и тем самым активно преобразовывать в нужном направлении.

В модели интеллекта Д. Гилфорда показаны возможности человека по универсальному структурированию среды. Выделено 120 интеллектуальных способностей (позднее их число было увеличено до 150), каждая из которых связана с определенным классом задач, представленных в координатах содержание–операция–продукт [13].

Р. Мейли выдвинул гипотезу о том, что структуру интеллекта составляют четыре фактора – доступной сложности, пластичности, целостности, беглости [14]. Экспериментальная проверка этой гипотезы показала, что теоретически выделенные факторы являются инвариантными и постоянно устанавливаются в структуре, начиная с 6-летнего возраста. На основе проведенных исследований Р. Мейли предложил «компонентную модель» интеллекта, в которой факторы трактовались как условия (компоненты) индивидуальных различий в выполнении разных интеллектуальных актов. Они могут относиться как к индивиду, так и к окружающей среде. Следовательно, структура интеллекта, включающая эти факторы, включает в себя структуру взаимодействия индивида со средой.

В теории множественного интеллекта Говарда Гарднера выделено девять типов интеллекта: лингвистический, логико-математический, визуально-пространственный, телесно-кинестетический, музыкальный, натуралистический, межличностный, внутриличностный и экзистенциальный [15]. Все виды интеллектов по Гарднеру равноценны, каждый представляет собой особый способ взаимодействия человека с окружающей действительностью. Очевидно, что каждый интеллект в данной модели эффективен только по отношению к специализированной среде, отражающей соответствующие формы культуры.

В практике инженерного проектирования сложных систем данные подходы имеют ограниченное применение, так как получаемые в соответствии с ними модели не универсальны и работают только в узком диапазоне условий.

Интеллект и сложность среды

Мы наделяем систему высоким интеллектом только в случае, если она проявляет способности к решению сложных задач, не имеющих заранее готовых алгоритмов решения. Предполагается, что для решения простых задач интеллект не требуется. В связи с этим интеллект оказывается связанным с миром задач, различающихся между собой сложностью решения, и может интерпретироваться с позиций концепций сложности.

Выделяют «онтологическую» и «семиотическую» простоту/сложность. Под первой понимают сложность материального, физического мира, а под второй – оценку знаковых систем. Понятие интеллекта в большей мере связано с семиотической сложностью, хотя мир задач обусловлен и онтологической сложностью. Отношения между этими видами сложности и определяют средово-системный континуум, в котором порождается понятие «интеллект». Таким образом, традиционное понимание интеллекта как суммы локально принадлежащих когнитивному аппарату человека способностей (инструментальных возможностей искусственной системы в случае искусственного интеллекта) должно быть заменено на интегральные свойства, обусловленные средой и действующей (живущей) в ней системой. Интеллект определяется в значительной мере степенью освоенности конкретной среды системой, действующей в ней, степенью воплощенности в нее.

Большинство исследователей сложного считают, что представления о сложном могут быть сформулированы в следующих концептах: множественность; динамическое разнообразие; нелинейность; неравномерность; сложность самоорганизующихся систем. Эти же категории применимы и к понятию «ин-

теллект». Множественность рассматривается как многокомпонентность. Она относится к описанию сложных систем, непрерывно эволюционирующих и изменяющихся. Сложная система в этой парадигме предстает как процесс бесчисленного усложнения ее сущностей, возникновения новых элементов как элементов для новых творений. Интеллект может быть представлен как свойство, воплощенное в сложную динамическую систему, позволяющее достигать результат, определяемый условиями задачи.

Множественность в концепциях динамического разнообразия дополняется качественной характеристикой – разнообразием. Разнообразие связано с асинхронным существованием в среде динамически существующих и сосуществующих систем и их распадающихся элементов, которые могут образовывать, в свою очередь, новые системные сущности.

Концепт неравномерности отражает принципиальную неравномерность распределения в пространстве одновременно существующих различных форм материи (энергии, вещества), ведущую к локальной самоорганизации и возникновению новых систем. Постулируется холистический характер мира, который разделяется на элементы только в процессе работы механизмов человеческого интеллекта и сознания.

Концепция сложности самоорганизующихся систем отражает непрерывную динамику мира во всех ее принципиально непознаваемых количественно-качественных проявлениях, а интеллект в ней является эмерджентным свойством сложной системы, позволяющим последней эффективно решать задачи активного (формирование искусственной среды, созидающая деятельность) и пассивного (адаптация, приспособление к среде) существования в мире.

Синергетический базис искусственного интеллекта

Е.Н. Князева обобщает существующие взгляды на сложность и сложные системы, формулируя и детализируя характерные свойства сложных систем в рамках синергетической парадигмы:

- сложность есть множество элементов системы, соединенных нетривиальным образом оригинальными связями друг с другом. Сложность есть динамическая сеть элементов, соединенных по определенным правилам;
- сложность есть внутреннее разнообразие системы, разнообразие ее элементов или подсистем, которое делает ее гибкой, способной изменять свое поведение в зависимости от меняющейся ситуации;
- сложность есть многоуровневость системы (существует архитектура сложности);
- сложные системы являются системами открытыми, т.е. обменивающимися веществом, энергией и (или) информацией с окружающей средой. Границы сложной системы порой трудно определить (видение ее границ зависит от позиции наблюдателя);
- сложные системы – это такие системы, в которых возникают эмерджентные феномены (явления, свойства), которые не могут быть «вычитаны» из анализа поведения отдельных элементов;
- сложные системы имеют память, для них характерно явление гистерезиса, при смене режима функционирования процессы возобновляются по старым следам (прежним руслам);
- сложные системы регулируются петлями обратной связи: отрицательной, обеспечивающей восстановление равновесия, возврат к прежнему состоянию, и положительной, ответственной за быстрый, самоподстегивающийся рост, в ходе которого расцветает сложность [16, с. 77–78].

Очевидно, что приведенные Е.Н. Князевой дефиниции сложных систем применимы и к интеллектуальным системам, так как в этом случае интеллект является функцией от сложности самоорганизующейся системы.

Далее в концепции Е.Н. Князевой делается ряд существенных дополнений к классическим взглядам на проблему сложных систем. Во-первых, сложные системы, по ее мнению, являются системами операционально закрытыми. Система одновременно является открытой и замкнутой по отношению к окружающей среде. Операциональная замкнутость означает селективность системы, наличие границы, упорядочивающей отношения системы со средой и окружающими системами. Система и среда проявляют взаимную активность. Среда меняет систему, но и система активно видоизменяет окружающую среду, вступая в коэволюцию с нею.

Сложность в современных синергетических концепциях является возникающим и исчезающим феноменом, циклически порождаемым в циклах самоорганизации. Многие ее аспекты перекликаются с понятием интеллекта. Можно предположить, что интеллектуальные функции отражают некоторые общие принципы самоорганизации аутопоэтических систем, в частности, их способность активно изменять сложность среды [17]. Отсюда следует, что все аутопоэтические системы, в том числе и образуемые социальными коммуникациями, обладают воплощенными в них интеллектуальными способностями.

Симбиотические интеллектные системы и среды

Интеллектуальные эргатические системы представляют собой системы организованной сложности (таблица). Будучи организованными человеком, они содержат в себе в скрытой форме часть функций,

присущих системам, наделенным естественным интеллектом. Это искусственный интеллект, воплощенный в искусственную среду.

Тип интеллектного образования	Отношения между компонентами	Центр активности и управления, механизм	Отношения со средой деятельности, границы
естественный интеллект	самоорганизация, аутопоэтическая система	сознание, эго-система человека	активное преобразование мира, границы динамично изменяются
искусственный интеллект	фиксированная или переменная программно-аппаратная структура	программа, алгоритм в структурированной или структурируемой среде	реализация алгоритма, ситуативное управление, границы фиксированные
гибридный интеллект	симбиоз, адаптация организованного и аутопоэтического компонентов к среде, объединения на макроуровне с приоритетом сознания	человек в структурированной среде	взаимная адаптация естественного и искусственного интеллектов, границы переменные
диффузный интеллект	селективные связи на всех уровнях аутопоэтически организованной и организуемой сред и человека	возникает в организованной среде	синергетическое объединение, границы формируются под задачу

Таблица. Формы интеллектных образований

Таким образом, мы приходим к понятию симбиотического диффузного интеллекта, который представляет собой систему, включающую естественный и искусственный интеллекты в их синергетическом взаимодействии в сложноорганизованной искусственной среде. Чем выше уровень искусственного интеллекта, распределенного в среде эргатической системы, тем более эффективен при прочих равных условиях интеллект человека. Свойства симбиотического диффузного интеллекта зависят от способности естественного интеллекта пользователя освоить интерфейс и интеллектуальную компоненту среды. Отметим отличие понятия «диффузный интеллект» от понятия «гибридный интеллект», применяемого в сложных эргатических системах. Гибридный интеллект отражает свойство сложной системы, возникающее в процессе функционирования в ней естественного интеллекта, а диффузный интеллект возникает как синергетическое свойство в сложноорганизованной среде в процессе ее когнитивной коэволюции с системой естественного интеллекта. Гибридный интеллект основан на представлениях об адаптации систем, а диффузный – о взаимодействии и взаимной ориентации.

Эффективное взаимодействие естественного и искусственного интеллектов усиливает возможности человеческого интеллекта в искусственной среде, но только в определенных, предоставляемых технологиями границах.

Меняется парадигма интеллекта с его понимания как локально существующего в отдельной когнитивной системе феномена на способность человека обеспечивать эффективные межсистемные взаимодействия, ведущие к получению требуемого результата (решения задачи) в сложноорганизованных интеллектуальных средах.

Интеллект эргатических систем с искусственным интеллектом

Любая интеллектуальная эргатическая система представляет собой, в известной мере, автономную систему, осуществляющую с помощью оператора деятельность по достижению заданной цели в конкретной среде. Чем больше функций во взаимодействиях со средой без присутствия оператора может выполнять система, тем более сложной и интеллектуальной она кажется наблюдателю. Интеллект искусственной технической системы возникает и проявляется в условиях организованной сложности, в отличие от живых систем, активно организующих среду. Система искусственного интеллекта эргатической системы является системой организованной сложности, в которой взаимодействие элементов и функциональных областей строятся на неслучайных отношениях, заданных проектировщиком в некоторых контекстах, которые непрерывно меняются, следуя логике эволюции задачи, вызываемой оператором.

Функциональная сложность эргатической системы может быть оценена по аналогии с когнитивной сложностью человека по количеству и типам решаемых задач. Ее сложность (и интеллектуальность) может определяться в привязке к конкретным решаемым задачам, к сложности программных средств, используемых в управляющем компьютере системы, с объемом и степенью дифференцированности кон-

тактов системы со средой и оператором. В качестве критерия сложности (интеллектуальности) может использоваться длина выполняемого алгоритма и количество циклов, позволяющих его реализовать.

Перспективным методом определения интеллектуальности эргатической системы является метод, учитывающий свойства культурного поля проектного коллектива, участвующего при создании системы. Чем шире спектр специалистов и более высока междисциплинарная нагрузка, тем система сложнее и интеллектуальнее.

Появление эргатических систем с искусственным интеллектом требует рассматривать вопросы взаимодействия естественного интеллекта оператора системы с искусственным интеллектом, воплощенным в технической части эргатической системы.

Можно дать ряд общих определений интеллекта и интеллектных симбионтов, действующих в эргатической системе и возникающих в процессе объединения искусственного и естественного интеллектов и среды деятельности.

1. Интеллект есть форма активной самоорганизации сложной системы, вовлекающая погруженного в среду пользователя в созидающие изменения.
2. Интеллект связан со средой как механизм ее организации, обеспечивающий процессы самоорганизации системы им наделенной.
3. Интеллект распределен в континууме «система–среда» и воплощен в циклах самоорганизации системы, действующей в среде.
4. Естественный интеллект представляет собой организующую сложность в организуемой среде, а искусственный интеллект – организованную сложность в организованной среде.
5. Гибридный и диффузный интеллекты представляют собой симбионты, включающие организующую и организованную сложности систем в их синергетическом и симбиотическом взаимодействиях как инструмент достижения цели актором в организованной и организуемой средах.
6. Интеллект отражает результаты селекции и применения самоорганизующейся системой эффективных способов достижения цели в организованной среде.

Заключение

Интеллектные симбионты являются проявлением различных форм межсистемных объединений различной материальной природы, в которых проявляется активная форма преобразования и структурирования среды.

Диффузный интеллект в организованной среде эргатической системы является формой самоорганизации когнитивной системы человека, погруженного в сложноорганизованную информационно-коммуникативную среду, в которой взаимодействуют формы искусственного и естественного интеллектов, находящиеся в симбиотическом взаимодействии, что ведет к появлению эффективных способов структурирования содержания искусственной среды и проявляется в эффективном достижении целей системы.

Гибридные формы интеллектных симбионтов представляют собой объединения, не нарушающие системную целостность составляющих их симбионтов. При этом совместное функционирование симбионтов порождает новые, полезные для гибридной системы интеллектуальные свойства.

При проектировании сложных эргатических систем важно правильно выбрать методы и технологии их интеллектуализации, которые, в основном, зависят от среды деятельности. Попытка полной автономизации и автоматизации системы связана со степенью организованности среды. При работе в условиях организованных сред целесообразно использование методов распределенного искусственного интеллекта, а в слабоструктурированных средах – включение и усиление возможностей естественного интеллекта.

При создании сложных эргатических систем необходимо учитывать эмерджентные свойства, возникающие вследствие сложной организации среды – эффекты интеллектуализации, кооперативных и гибридных форм интеллектов и их симбионтов. Включение в сложные техногенные среды человека также связано с эффектами модификации его личности и когнитивных систем, что ведет к возникновению техно-психических симбиозов. Например, память человека заменяется памятью технической системы, функции внимания и поиска информации – работой поисковых систем и т.д.

Материалы настоящего исследования могут быть использованы при создании систем интерфейсов сложных эргатических систем, включающих симбиотические объединения аутопоэтического типа, возникающие в распределенных эргатических системах с социальными формами управления и взаимодействия с операторами.

Литература

1. Сергеев С.Ф. Методология проектирования тренажеров с иммерсивными обучающими средами // Научно-технический вестник СПбГУ ИТМО. – 2011. – № 1 (71). – С. 109–114.

2. Сергеев С.Ф. Методология создания мехатронных систем с искусственным интеллектом // Известия ТулГУ. Технические науки: В 3 частях. – Ч. 1. – Тула: Изд-во ТулГУ, 2011. – Вып. 5. – С. 245–249.
3. Венда В.Ф. Системы гибридного интеллекта: Эволюция, психология, информатика. – М.: Машиностроение, 1990. – 448 с.
4. Сергеев С.Ф. Психологические основания проблемы искусственного интеллекта // Мехатроника, автоматизация, управление. – 2011. – № 7. – С. 2–6.
5. Сергеев С.Ф. На пути от биоорганизации к киберорганизации: человек в тени искусственного интеллекта // Естественный и искусственный интеллект: методологические и социальные проблемы / Под ред. Д.И. Дубровского и В.А. Лекторского. – М.: Канон+, РООИ Реабилитация, 2011. – С. 48–59.
6. Wechsler D. Wechsler Adult Intelligence Scale. – New York: The Psychological Corporation, 1955.
7. Холодная М.А. Когнитивные стили. О природе индивидуального ума: Учебное пособие. – М.: ПЕР СЭ, 2002. – 304 с.
8. Сергеев С.Ф. Инструменты обучающей среды: интеллект и когнитивные стили // Школьные технологии. – 2010. – № 4. – С. 43–51.
9. Сергеев С.Ф. Инструменты обучающей среды: стили обучения // Школьные технологии. – 2010. – № 5. – С. 19–27.
10. Сергеев С.Ф. Когнитивная педагогика: пользовательские свойства инструментов познания // Школьные технологии. – 2011. – № 2. – С. 35–41.
11. Степин В.С. Классика, неклассика, посттнеклассика: критерии различения // Постнеклассика: философия, наука, культура. – СПб: Издательский дом Міръ, 2009. – С. 249–295.
12. Дружинин В.Н. Психология общих способностей – СПб: Питер, 1999. – 368 с.
13. Гилфорд Д. Три стороны интеллекта // Психология мышления. – М.: Прогресс, 1965. – С. 434–437.
14. Meili R. Analytischer Intelligenztest (AIT) / Verlag H. Huber. – Bern, Stuttgart, Toronto, 1971. – 119 p.
15. Гарднер Г. Структура разума. Теория множественного интеллекта. – М.: Вильямс, 2007. – 512 с.
16. Князева Е.Н. Темпоральная архитектура сложности // Синергетическая парадигма. Синергетика инновационной сложности. – М.: Прогресс-Традиция, 2011. – С. 66–86.
17. Сергеев С.Ф., Заплаткин Ю.Ю., Захаревич М.А. Механизмы редукции в самоорганизующихся мирах циклических коммуникаций // Проблемы управления и моделирования в сложных системах: Труды XIV Международной конференции (19–22 июня 2012 г., Самара, Россия). – Самара: Самарский научный центр РАН, 2012. – С. 131–137.

Сергеев Сергей Федорович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор психологических наук, профессор, ssfpost@mail.ru

УДК 14.01.85

ИСПОЛЬЗОВАНИЕ ИНФОРМАЦИОННО-КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ В ЭЛЕКТРОННОМ ОБУЧЕНИИ ИНОСТРАННЫМ ЯЗЫКАМ

И.Б. Тампель, Е.В. Краснова, Е.А. Панова, К.Е. Левин, О.С. Петрова

Рассматриваются способы применения в системе электронного обучения иностранным языкам технологий автоматического распознавания и синтеза речи для задач отработки произношения на сегментном и супрасегментном уровнях, тренировки коммуникативных навыков, поверки словарного запаса обучаемого, тренировки навыка понимания на слух, а также для создания системы голосовой навигации. Несмотря на некоторые ограничения, такое применение данных технологий эффективно как для упрощения задач реализации процесса обучения, так и для повышения удобства использования системы.

Ключевые слова: электронное обучение, иностранные языки, ИКТ, речевые технологии, распознавание речи, синтез речи.

Введение

Спрос на изучение иностранных языков в настоящее время, несомненно, велик. Однако из-за стремительно ускоряющегося ритма жизни человеку все меньше времени удается уделять обучению в целом и изучению иностранных языков, в частности. В связи с этим в последнее время резко возросла необходимость в модернизации процесса обучения, которое должно быть доступным, мобильным и удобным. В последние годы Министерство образования и науки Российской Федерации всерьез озаботилось вопросом повышения уровня доступности образования и с этой целью приступило к выполнению мероприятий по созданию электронных образовательных ресурсов [1].

Электронные образовательные ресурсы создаются для широкого круга пользователей. Возможность дистанционного использования электронных ресурсов делает обучение мобильным: с помощью

персонального компьютера, смартфона или планшетного компьютера с установленным устройством доступа к вычислительной сети пользователь может самостоятельно обучаться в интерактивном непрерывном режиме. Также имеется возможность учесть предпочтения пользователя и его индивидуальные успехи и, таким образом, автоматически адаптировать курс обучения к конкретному пользователю.

Внедрение информационно-коммуникационных технологий (ИКТ) в электронные образовательные ресурсы, в частности, технологий автоматического синтеза (TTS – Text-to-Speech) и распознавания (ASR – Automatic speech recognition) речи, которым посвящена данная работа, позволяет сделать системы электронного обучения максимально удобными в использовании (например, благодаря голосовой навигации внутри курса). Что касается систем электронного обучения иностранным языкам, внедрение инновационных речевых технологий открывает также методические возможности. Очень важно отметить и социальную направленность подобных разработок: например, внедрение технологий TTS и ASR в системы электронного обучения делает эти системы доступными для людей со слабым зрением или с ограниченной возможностью движения (нарушением опорно-двигательного аппарата), т.е. обеспечивает людям с ограниченными возможностями взаимодействие с компьютером.

Цель настоящей работы – описать, каким образом с помощью технологий автоматического распознавания и синтеза речи решаются некоторые задачи системы электронного обучения иностранным языкам.

Создание программного комплекса «Лингвистический тренажер»

В настоящее время осуществляется внедрение ИКТ в систему электронного обучения иностранным языкам «Лингвистический тренажер», которая создается ООО «ЦРТ» (Центр речевых технологий). Лингвистический тренажер (ЛТ) представляет собой программный комплекс с элементами искусственного интеллекта, предназначенный для освоения навыков понимания на слух и беглого говорения на иностранных языках (предусмотрено обучение девяти языкам) с помощью компьютера, в том числе различных мобильных устройств, на основе вопросно-ответных коммуникаций.

Программная архитектура ЛТ предусматривает следующие компоненты:

- система обучения, в которой содержатся как учебные материалы (теоретический материал и упражнения), так и информация об обучаемом (о результатах обучения, об истории взаимодействия пользователя с системой и т.д.) и модуль расчетов; на основе этих данных создается экспертная система, которая позволяет модифицировать учебный курс в зависимости от результатов обучения пользователя, т.е. адаптировать курс к конкретному пользователю;
- служебные функции, обеспечивающие надежное функционирование системы в целом;
- пользовательский интерфейс, которому при разработке ЛТ уделяется особое внимание: необходимо создать максимально удобный интерфейс, так как программа создается для широкого круга пользователей [2].

Применение ИКТ в «Лингвистическом тренажере»

Использование речевых технологий в системах электронного обучения иностранным языкам (обучение произношению, чтению вслух, освоение некоторых коммуникативных навыков) началось относительно недавно [3], поэтому, а также ввиду ряда технических ограничений [4], в существующих сервисах полноценного их применения практически не наблюдается.

В ЛТ предусмотрено освоение навыков понимания на слух и говорения на девяти языках. В разработанной технологии электронного обучения иностранным языкам предполагается использование систем ASR и TTS для русской речи, таким образом, они применимы в той части ЛТ, в которой происходит обучение русскому языку, а также в интерфейсе для русскоговорящих пользователей.

Использование технологий ASR и TTS позволяет решить задачи голосовой навигации по сервису и некоторые методические задачи, а именно, отработка произношения на сегментном и супрасегментном уровнях, тренировка коммуникативных навыков, поверка словарного запаса обучаемого, тренировка навыка понимания на слух.

Использование технологии ASR

Существуют разные типы автоматического распознавания речи, все они применимы для решения разных функциональных и методических задач ЛТ.

Для голосовой навигации по сервису ЛТ, которая необходима как для мобильного использования сервиса, так и для использования его людьми с ограниченными физическими возможностями, применяется технология распознавания команд. Распознавание команд – это разновидность распознавания дискретной речи, т.е. изолированно произнесенных слов. Суть данной технологии состоит в том, что система реагирует на изолированно произнесенные слова или словосочетания из небольшого заранее заданного словаря. Применимость распознавания речи данного типа ограничивается заданным словарем, в то же время за счет этого ограничения точность распознавания достаточно высока [5].

В методических целях технология распознавания команд применима, например, для устной тренировки и проверки словарного запаса обучаемого: для этого также задается определенный словарь, состоящий из слов, которые пользователю необходимо выучить, чтобы успешно пройти урок.

Основная же область применения ASR в целях обучения иностранному языку – тренировка произношения (CAPT – Computer Assisted Pronunciation Training), на ней следует остановиться подробнее. Тренировка фонетики иностранного языка подобным способом очень востребована и удобна, поскольку пользователь может сам определять время и количество тренировок, отслеживать свои результаты, повторять пройденный материал столько раз, сколько ему необходимо, что невозможно реализовать при аудиторном обучении.

Главное отличие систем ASR, используемых для обучения иностранному языку, от стандартных систем – это распознавание акцентной речи. Системы, предназначенные для распознавания речи носителей языка, могут быть лишь частично применимы в целях обучения иностранному языку, и то лишь на более высоком уровне обучения, тогда как тренировка фонетики, в основном, требуется на начальном этапе. Более того, такие системы не будут делать различий между похожими звуками изучаемого и родного языка (например, между английским придыхательным апикально-альвеолярным /t/ и русским зубным дорсальным /т/), что не позволит разработать эффективную обратную связь для коррекции произношения. При обучении произношению на иностранном языке могут возникать и другие трудности: наличие в изучаемом языке звуков, которые отсутствуют в родном языке (например, межзубные английские /ð/ и /θ/, отсутствующие в русском языке), а также различное контекстное употребление похожих звуков (например, испанские звонкие смычные /b/, /d/, /g/ не употребляются в интервокальной позиции, а русские /б/, /д/, /г/ – в абсолютном конце (в позиции перед паузой)).

Таким образом, одним из наиболее перспективных направлений разработки систем распознавания акцентной речи является добавление информации о родном языке пользователя, что даст представление об ошибках, которые он, скорее всего, сделает. На ошибки конкретных пользователей могут быть натренированы как акустические, так и языковые модели распознавания речи (подробнее об акустических и языковых моделях см. далее) [6]. Акустические модели осуществляют распознавание ошибок на фонетическом уровне (см. приведенный пример о разнице между русским /т/ и английским /t/), языковые модели осуществляют распознавание ошибок на фонематическом уровне, например, замена носителем русского языка английского /θ/ на /t/ в слове «three» приведет к смысловой ошибке. Поскольку фонематические ошибки являются более грубыми, то при разработке системы ошибок носителей определенного языка, им должно быть уделено больше всего внимания. Наиболее эффективным материалом для отработки подобных ошибок являются упражнения на минимальные пары, подобранные специально с учетом трудностей произношения носителей определенного языка (например, при изучении английского языка трудности произношения для русскоязычных и испаноязычных пользователей будут отличаться). Как показали исследования [7], можно ранжировать носителей разных языков по количеству ошибок, которые они могут сделать в конкретном изучаемом языке, а также выявлять универсальные ошибки и ошибки, характерные для определенных языков и для определенных языковых групп. Например, при изучении произношения голландского языка меньше всего трудностей наблюдалось у англичан и носителей скандинавских языков (носителей типологически близких языков), чем у носителей славянских, романских и восточных языков [7].

Тренировка произношения на супрасегментном уровне включает в себя отработку как словесного ударения (уровень слова), так и интонации (уровень фразы). Для распознавания интонации иностранного языка требуется выделение параметров, существенно отличающихся от сегментного уровня: частоты основного тона (ЧОТ) – основного параметра в большинстве языков, а также интенсивности и длительности супрасегментных единиц. Если на сегментном уровне и уровне словесного ударения произношение можно охарактеризовать как «правильное» или «неправильное», то на интонационном уровне высказывание лучше характеризовать как «корректное» и «некорректное» [8], поскольку в данном случае нет четкого определения параметра «правильности» произношения. Исключением в данном случае являются тональные языки, в которых тон неразрывно связан со смыслом (например, китайский язык). В естественном общении критерием «корректности» является достижение коммуникативной цели, в системе распознавания подобным критерием являются значения диапазонов ЧОТ, интенсивности и длительности супрасегментных единиц. В плане описания интонационной системы наиболее удобным представляется подход С. Оде [9] при описании интонации английского, голландского и русского языков. Данный подход был разработан в автосегментной фонологии и является наглядным и удобным, особенно с точки зрения сопоставления интонационных систем разных языков.

Для отработки обучаемым коммуникативных навыков технология ASR также применима: она может использоваться при создании упражнений, представляющих собой вопросно-ответные коммуникации. Например, упражнение представляет собой диалог, где пользователю необходимо устно давать ответы на определенные вопросы, при этом в зависимости от того, что ответит обучаемый, следующий вопрос пользователю автоматически выбирается из нескольких вариантов (вопросно-ответные коммуника-

ции, представляющие собой граф). Для такого типа упражнений применима технология распознавания ключевых слов: исходя из распознанных ключевых слов, ЛТ выбирает следующий вопрос. Данная технология заключается в распознавании отдельных участков слитной речи, при этом системе распознавания не требуется полностью преобразовывать в текст произнесенную речь, достаточно распознать лишь определенные слова, на основе которых в ЛТ может строиться дальнейшая траектория взаимодействия с обучаемым [5, 10].

Нагляднее всего можно показать такой принцип работы системы на примере отработки альтернативных вопросов, ответы на которые являются полными (рисунок).

Рисунок. Пример использования технологии распознавания ключевых слов в учебном упражнении

Однако для того, чтобы в подобных и других упражнениях обеспечить полноценный автоматический анализ произнесенного пользователем текста (в том числе с точки зрения наличия/отсутствия грамматических ошибок), технологии распознавания ключевых слов недостаточно, необходимо использовать технологию распознавания слитной речи. Распознанный таким образом текст далее анализируется парсерами, т.е. в этом случае возможна отработка обучаемым правил морфологии и синтаксиса изучаемого языка [11].

Распознавание слитной речи осуществляется на большом словаре (LVCSR – Large vocabulary continuous speech recognition) и основано на акустических и языковых моделях. Акустические модели – это шаблоны звуков речи, с которыми сопоставляется входной речевой сигнал. Наибольшее распространение в системах распознавания с большим словарем получило представление акустических моделей, соответствующих акустическим событиям, например, аллофонам (монофонам, дифонам или трифонам), с помощью моделей скрытого марковского процесса [12–14]. Языковая модель – обязательная часть систем распознавания слитной речи. Не любая последовательность слов является предложением, между словами есть грамматические и семантические связи. Языковая модель позволяет узнать, какие последовательности слов в языке более вероятны, а какие – менее. Информация от языковой модели либо учитывается одновременно с информацией от акустической модели (каждая со своим весом), либо включается на втором этапе. Использование языковой модели помогает сократить пространство поиска и снять неоднозначность при выборе из нескольких близких по стоимости акустических гипотез (для русского языка, например, помогает правильно распознать слово в нужном падеже). Процесс подготовки лингвистических моделей, как правило, автоматизирован и основан на статистической обработке большой выборки текстового материала.

Качество распознавания слитной речи обеспечивается модулем оценки качества сигнала, а также модулем определения речевой активности, который позволяет выделить в звуковом потоке неречевые участки и участки, содержащие речь. Из анализа исключаются участки пауз, что позволяет сократить время обработки данных, повысить достоверность распознавания, а также реализовать адаптивные алгоритмы распознавания, настраивающиеся на определенный уровень помех и искажений [15, 16]. Таким образом, система распознавания слитной речи работает достаточно эффективно и вполне применима для методических задач ЛТ.

Однако для полноценного и эффективного использования распознавания слитной речи для отработки коммуникативных навыков обучаемого также необходима настройка на акцентную речь [17]. Это условие относится к системам ASR не только для русской речи.

Эффективность работы ASR в применении к обучению иностранному языку заключается не только в высокой надежности распознавания, но и в наглядной интерпретации результата распознавания для пользователя. В данном случае необходимо учитывать, для какого круга пользователя разрабатывается система обучения языку. Для широкой аудитории обратная связь должна быть представлена в максимально простой и наглядной форме, результат распознавания может выдаваться как «правильный»/«неправильный», либо «корректный»/«некорректный» (использование системы баллов и системы «правильности» в процентах представляется неэффективным), например, могут даваться указания по основным трудностям, характерным для пользователей данного языка. Если система обучения языку разрабатывается для студентов языковых специальностей, в качестве обратной связи могут предоставляться графики спектра мощности, ЧОТ и т.д. Не только преподаватели, но и сами студенты могут сравнивать свои результаты с эталонными произнесениями носителей изучаемого языка. Исследования показывают [1, 4, 6], что подобная обратная связь эффективна в обучении.

Отдельной задачей является адаптация системы ASR в целях обучения иностранному языку для людей с речевыми нарушениями [18]. На основе опыта ООО «ЦРТ» по исследованию акустических свойств речевых нарушений в дальнейшем предполагается разработка системы распознавания дефектной речи.

Использование технологии TTS

Для голосовой навигации по ЛТ применима также технология синтеза речи. Озвучка навигационных сообщений синтезом речи не только обеспечивает удобство в использовании сервиса слабовидящими, но и исключает необходимость трудоемкой записи дикторской речи. Технология синтеза русской речи основана на методе Unit Selection (разновидность компилятивного синтеза), в котором из большой речевой базы могут выбираться элементы различной величины (аллофоны, последовательности аллофонов, предложения) [19–21]. В настоящее время эта технология позволяет достичь максимальной естественности синтезированной речи. Как пишет в своей книге «Text-to-Speech Synthesis» один из основоположников технологии Unit Selection Пол Тэйлор [19], в то время как по разборчивости синтезируемой речи эта технология сравнима с другими (также существует формантный, артикуляторный синтез, синтез на основании статистических моделей), по естественности речи она значительно их превосходит.

Принцип технологии TTS основан на том, что из большой речевой базы по полученной на ранних этапах работы синтезатора фонетико-просодической последовательности выбираются наиболее подходящие элементы. Качество синтеза напрямую зависит от размера и качества используемой речевой базы (необходимо использовать полную, сбалансированную и корректно размеченную базу).

Технология TTS также может быть использована для методических целей, например, для упражнений, представляющих собой диктовку на иностранном языке, которые необходимы для освоения навыков понимания на слух иностранной речи. В целом эффективность диктовки с помощью TTS можно оспорить, так как автоматически синтезированная русская речь на данный момент не обладает полностью естественным звучанием и отличима от реальной речи. С этой точки зрения для подобных целей было бы правильным использовать запись речи диктора, чтобы исключить возможные искажения, которые могут негативно повлиять на корректное восприятие иностранной речи обучаемым. Однако, как показали исследования, диктант посредством синтезированной речи является более эффективным на первых порах обучения, поскольку синтезированная речь имеет большую разборчивость, чем речь человека [22].

Для целей отработки произношения, например, упражнений на имитацию звуков, отдельных слов и фраз, может быть использована технология макрокомпилятивного синтеза, который обеспечивает еще более высокую естественность звучания и не допускает малейших искажений звукового сигнала. Макрокомпилятивный синтез гарантирует стыковку фрагментов непрерывной речи не на уровне аллофонов, а на уровне отдельных слов или даже фраз. Однако данный тип синтеза удобен для обеспечения идеального произнесения только отдельных, часто повторяемых фрагментов речи.

Применение TTS в электронной системе обучения иностранным языкам позволяет добавлять в ресурс новые учебные материалы, которые необходимо озвучить, и при этом снимает задачу записи речи диктора.

Использование TTS в ЛТ в методических целях планируется в перспективе.

Заключение

Информационно-коммуникационные технологии в настоящее время достигли очень высокого уровня, что позволяет использовать их в самых разнообразных сферах. Эффективным является применение информационно-коммуникационных технологий в образовании, в первую очередь – в электронных образовательных ресурсах.

Разработки в области речевых технологий можно и необходимо внедрять в системы электронного обучения, в частности – в системы обучения иностранным языкам. ООО «ЦРТ» реализует эту идею в разрабатываемом при финансовой поддержке Министерства образования и науки Российской Федерации программном комплексе «Лингвистический тренажер». Так как сервис предназначен для освоения уст-

ных коммуникативных навыков, использование в нем новейших речевых технологий логично. В работе были описаны способы решения некоторых задач лингвистического тренажера с помощью технологий автоматического распознавания и синтеза речи. Все технологии, о которых шла речь, разработаны ООО «ЦРТ».

Эффективность применения речевых технологий в лингвистическом тренажере заключается не только в упрощении задач реализации процесса обучения (применение технологий непосредственно к учебному материалу), но и в повышении удобства использования лингвистического тренажера как для широкого круга пользователей, так и для людей с ограниченными физическими возможностями (использование технологий для навигации по сервису).

В настоящее время имеются некоторые ограничения, которые не позволяют использовать системы ASR и TTS в полной мере, однако стремительное развитие информационно-коммуникационных технологий позволяет прогнозировать скорое преодоление этих ограничений. В целом же применение речевых технологий в обучении значительно повышает доступность образования, на что в данный момент нацелены мероприятия, реализуемые государством.

Литература

1. Материалы к выступлениям Министра образования и науки РФ [Электронный ресурс]. – Режим доступа: http://минобрнауки.рф/%D0%BF%D1%80%D0%B5%D1%81%D1%81-%D1%86%D0%B5%D0%BD%D1%-82%D1%80?events_sections=5, свободный. Яз. рус. (дата обращения 06.12.2012).
2. Киселев В.В., Елисеева О.Е., Ковалёнок Ю.И., Хитрова Ю.Н. Архитектура лингвистического тренажера для экспресс-освоения навыков общения на иностранном языке // Научно-технический вестник информационных технологий, механики и оптики. – 2013. – № 1 (83). – С. 166–170.
3. Egan K.B. Speaking: A critical skill and challenge // CALICO Journal. – 1996. – V. 16. – № 3. – P. 277–293.
4. Dodigovic M. Artificial Intelligence in Second Language Learning: Raising Error Awareness. – Toronto, 2005. – 304 p.
5. Официальный сайт ООО «ЦРТ»: Распознавание речи [Электронный ресурс]. – Режим доступа: http://www.speechpro.ru/technologies/recognition, свободный. Яз. рус. (дата обращения 06.12.2012).
6. Stanley Th., Hacioglu K. Improving L1-Specific Phonological Error Diagnosis in Computer Assisted Pronunciation Training. Proceeding of Interspeech-2012 [Электронный ресурс]. – Режим доступа: http://interspeech2012.org/accepted-abstract.html?id=555, свободный. Яз. англ. (дата обращения 12.02.2013).
7. Neri A., Cucchiarini C., Strik H. Segmental errors in Dutch as a second language: How to establish priorities for CAPT // Proceeding of the InSTIL/ICALL Symposium. – 2004. – P. 13–16.
8. Arias J.P., Yoma N.B., Vivanco H. Automatic intonation assessment for computer aided language learning // Speech Communication 523. – 2010. – P. 254–267.
9. Оде С. Заметки о понятии тональный акцент на примере русского языка // Проблемы фонетики V. – 2007. – С. 237–249.
10. Хитров М.В. Распознавание речи в системах контроля персонала // Директор по безопасности. – 2012. – № 7 – С. 38–41.
11. Cucchiarini C., Neri A., Strik H. Oral proficiency training in Dutch L2: The contribution of ASR-based corrective feedback // Speech Communication. – 2009. – V. 51. – № 10. – P. 853–863.
12. Young S.J. A review of large-vocabulary continuous speech recognition // IEEE Signal Processing Mag. – 1996. – V. 13. – № 5. – P. 45–57.
13. Jelinek F. Statistical Methods for Speech Recognition. – MA.: Cambridge, 1998. – 188 p.
14. Рабинер Л.Р. Скрытые марковские модели и их применение в избранных приложениях при распознавании // ТИИЭР. – 1989. – Т. 77. – № 2. – С. 86–120.
15. T. Gandhi, V.K. Antiwal, A.K. Jain Evaluation of Smoothed Language Models. – 2011. – V. 2. – P. 487–491.
16. Wilson I. Using Praat and Moodle for Teaching Segmental and Suprasegmental Pronunciation // Proc. of the 3rd International WorldCALL Conference: Using Technologies for Language Learning. – 2008. – P. 112–115.
17. Levow G., Olsen M.B. Modeling the language assessment process and result: Proposed architecture for an automatic oral proficiency assessment. Computer mediated language assessment and evaluation in natural language processing // Proceedings of a symposium sponsored by the Association for Computational Linguistics and International Association of Language Learning Technology. – 1999. – P. 24–31.
18. Ruiter M.B., Beijer L., Cucchiarini C., Krahmer E., Rietveld T.C.M., Strik H., Van Hamme H. Human language technology and communicative disabilities: requirements and possibilities for the future // Language Resources and Evaluation. – 2012. – V. 46. – № 1. – P. 143–151.
19. Taylor P. Text-to-Speech Synthesis. – MA.: Cambridge, 2009. – 626 p.

20. Black A.W., Hunt A.J. Unit Selection in a Concatenative Speech Synthesis Using a Large Speech Database // Proc. of ICASSP 96. – 1996. – V. 1. – P. 373–376.
21. Vepa J. Join Cost for Unit Selection Speech Synthesis. Phd thesis. – University of Edinburgh, 2004. – 241 p.
22. Pellegrini T., Costa A., Trancoso I. Less errors with TTS? A dictation experiment with foreign language learners. Proc. of Interspeech–2012 [Электронный ресурс]. – Режим доступа: http://interspeech2012.org/accepted-abstract.html?id=1105, свободный. Яз. англ. (дата обращения 12.02.2013).

Тампель Иван Борисович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат технических наук, тьютор, tampel@speechpro.com

Краснова Екатерина Викторовна – ООО «ЦРТ», младший научный сотрудник, krasnova@speechpro.com

Панова Екатерина Александровна – ООО «ЦРТ», научный сотрудник, panova@speechpro.com

Левин Кирилл Евгеньевич – ООО «ЦРТ», кандидат технических наук, руководитель отдела, levin@speechpro.com

Петрова Ольга Сергеевна – ООО «ЦРТ», менеджер проектов, Petrova-o@speechpro.com

КРАТКИЕ СООБЩЕНИЯ

УДК 004.822

ПОДХОД К КОЛЛЕКТИВНОЙ РАЗРАБОТКЕ ОНТОЛОГИЙ

А.В. Невидимов, И.А. Бессмертный

Рассмотрен подход к коллективной разработке онтологий на основе идеологии систем контроля версий с центральным хранилищем. Предложены решения для проблем обеспечения обратной совместимости онтологии, а также для решения конфликтов вносимых экспертами правок.

Ключевые слова: онтологии, совместная разработка, системы контроля версий.

При разработке информационных систем все более актуальными становятся семантические технологии [Л]. Структурирование знаний в виде онтологий позволяет уменьшать трудозатраты на обработку информации и принятие решений практически во всех предметных областях. Для получения максимально качественной онтологии той или иной предметной области разумным представляется задействовать не одного, а нескольких экспертов. В этой связи возникает задача обеспечения распределенной коллективной работы над онтологиями, и ее решение сопровождается следующими сложностями:

– разные эксперты могут иметь разные точки зрения на предметную область, что приводит к различиям в онтологиях этой предметной области, созданных разными экспертами. Это приводит к необходимости обеспечить слияние онтологий;
– параллельно с разработкой онтологии может разрабатываться программное обеспечение, использующее эту онтологию в качестве источника знаний. Прикладные программы должны использовать именно ту версию разрабатываемой онтологии, с которой они совместимы.

В настоящей работе предлагается подход, основанный на идеологии систем контроля версий с центральным репозиторием, таких как Git, Mercurial и SVN. Приведем список ключевых правок онтологии в таком подходе с их описанием и особенностями.

– Добавление новых триплетов вида <Сущность–субъект, Предикат, Сущность–объект> не нарушает обратной совместимости онтологии с ее предыдущей версией, так как все множество фактов, которые могли быть получены из старой версии онтологии, могут быть получены и из новой. Важным является также необходимость определять случаи параллельного добавления разными экспертами схожих сущностей для исключения дублирования данных в онтологии.
– Удаление старых триплетов в некоторых случаях может привести к нарушению обратной совместимости. Факт наличия или отсутствия обратной совместимости предлагается определять путем попытки логического вывода всех удаленных триплетов из новой онтологии, и если хотя бы один триплет не может быть выведен, то считается выявленным нарушение обратной совместимости. Удаление триплетов может также привести к нарушению валидности правок, вносимых другим экспертом.
– Удаление элементов из множества сущностей или множества отношений всегда приводит к нарушению обратной совместимости и может привести к нарушению валидности правок, вносимых другим экспертом.

В работе предлагается алгоритм коллективной работы над онтологиями, позволяющий избежать возникновения конфликтов и противоречий:

– Перед внесением в основную ветку онтологии набора правок проверить на совместимость с текущим состоянием онтологии;
– Если результирующая онтология несовместима с исходной, сформировать конфликтный список и предложить автору правок внести соответствующие изменения;
– Если автор вводит новую сущность, выполнить поиск в онтологии схожих сущностей с помощью вычисления расстояния Левенштейна между их названиями или на основе данных ресурса WordNet о синонимичности понятий. Сформировать список схожих понятий и вернуть автору для корректировки правок;
– Проверить обратную совместимость новой версии онтологии. Если в данной ветви онтологии есть нарушения обратной совместимости, опубликовать новый URL, являющийся точкой доступа SPARQL-endpoint, что обеспечит прикладным программам возможность пользоваться последней версией онтологии, с которой они совместимы.

Предложенный алгоритм в настоящее время проходит апробацию в проекте создания онтологии для корпоративной информационной системы, обеспечивающей работу торговых площадок.

[Л]. Бессмертный И.А. Семантическая паутина и искусственный интеллект // Научно-технический вестник СПбГУ ИТМО. – 2009. – № 6 (64). – С. 77–83.

Невидимов Александр Валентинович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, студент, alex.nevidimov@gmail.com

Бессмертный Игорь Александрович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат технических наук, доцент, igor_bessmertny@hotmail.com

УДК 004.896

АВТОНОМНАЯ НАВИГАЦИЯ МОБИЛЬНОГО РОБОТА НА ОСНОВЕ УЛЬТРАЗВУКОВОГО ДАТЧИКА ИЗМЕРЕНИЯ РАССТОЯНИЙ

М.О. Костишин, И.О. Жаринов, В.Д. Суслов

Рассматривается задача автономного управления мобильным роботом на основе единичного ультразвукового датчика измерения расстояний.

Ключевые слова: навигация, управление движением.

Решение задачи управления в системе автономной навигации мобильной робототехнической системы имеет важное практическое значение. Робототехнические системы применяются при чрезвычайных ситуациях для анализа экологической обстановки в зонах химико-биологического или радиационного загрязнения, при разборе завалов после землетрясений, для формирования рельефа местности после селевых сходов и оползней в зонах бедствия и т.д. Для автономного движения робототехническая система должна обладать системой навигации, на основании результатов измерений которой предотвращаются столкновения робота с препятствиями. Сегодня существуют различные автономные системы навигации, применяемые в робототехнических системах, решающих задачи специального назначения и бытовые хозяйственные задачи (например, роботы-пылесосы) [Л]. Реализуемые в системе навигации бытовых робототехнических систем принципы движения основаны на хаотичном перемещении устройства по обслуживаемой территории. Большинство таких устройств не оснащено измерительными датчиками и интеллектуальной системой обработки, парирующими столкновение робота с препятствиями.

Для организации автономного движения мобильного робота авторами предлагается принцип, основанный на анализе результатов измерений ультразвуковым датчиком расстояний от робота (датчика) до впереди стоящего препятствия как в статическом положении, так и в процессе движения робота. Наличие препятствия в направлении движения робота определяется по времени прихода отраженного радиоимпульса, излученного и принятого датчиком в ультразвуковом диапазоне частот. Расчет дальности осуществляется по формуле $x = ct/2$, где t – время прихода отраженного радиоимпульса с момента излучения; c – скорость распространения ультразвуковой волны в воздушном пространстве, $c = 340$ м/с.

В процессе движения робота датчик и измерительная система непрерывно осуществляют измерение расстояния до впереди стоящего препятствия. При достижении минимально допустимого порогового значения (300 мм) от робота до препятствия робот останавливается и производит замер расстояния слева и справа. Если расстояние слева оказывается больше, чем справа, робот движется в направлении наибольшего запаса движения (налево). При отсутствии беспрепятственных направлений движения осуществляется остановка робота и движение робота в обратную сторону. Измерение расстояния слева и справа от робота сопровождается вращением по кругу единичного датчика.

Результатом работы является алгоритм автономного управления мобильным роботом с использованием единичного ультразвукового датчика. Алгоритм реализован на экспериментальном макете робототехнической системы СМАРТБОТ, приведенном на рисунке. Программное обеспечение работы информационно-измерительной системы робота написано на языке С. Управление роботом осуществляется микроконтроллером Atmega-2560. Управление двигателями колесной базы реализовано на основе четырех двигателей постоянного тока. Круговое движение ультразвукового датчика осуществляется сервоприводом FS5106B. В качестве датчика измерения расстояния использовался ультразвуковой датчик ParallaxPing.

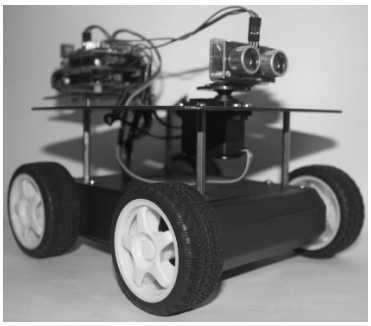


Рисунок. Мобильный робот СМАРТБОТ

Данный подход к реализации автономной системы навигации робототехнической системы является перспективным в области автомобилестроения. Применение серии датчиков, установленных по периметру робота, позволяет панорамно оценивать окружающее пространство вокруг робота с целью предотвращения столкновения робота с препятствием или с другими роботами при организации упорядоченного движения в транспортном потоке. Представленный принцип автономной навигации мобильного робота может быть использован при разработке системы автопилота для автомобильного транспорта.

Л. Принцип действия iRobot Roomba (фрагмент). Порядок автоматической уборки [Электронный ресурс]. – Режим доступа: http://irobot-home.ru/about-irobot/principle-of-operation-irobot-roomba/, свободный. Яз. рус. (дата обращения 15.01.2013).

Костишин Максим Олегович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, job.max@me.com
Жаринов Игорь Олегович – ФГУП «СПб ОКБ «Электроавтоматика» им. П.А. Ефимова», доктор технических наук, доцент, руководитель учебно-научного центра, igor_rabota@pisem.net
Суслов Владимир Дмитриевич – ФГУП «СПб ОКБ «Электроавтоматика» им. П.А. Ефимова», руководитель экспертного совета, postmaster@elavt.spb.ru

УДК 681.324

ПРИНЦИПЫ ОРГАНИЗАЦИИ АРХИТЕКТУРЫ ПЕРСПЕКТИВНЫХ БОРТОВЫХ ЦИФРОВЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ В АВИОНИКЕ

Е.В. Книга, И.О. Жаринов, А.В. Богданов, П.С. Виноградов

Рассматриваются основные принципы построения перспективных бортовых цифровых вычислительных систем в авиационном приборостроении. Приводится структурная схема перспективной бортовой вычислительной системы.
Ключевые слова: интегрированная модульная авионика, вычислительные системы, архитектура.

Бортовые цифровые вычислительные системы (БЦВС) современного летательного аппарата (ЛА) представляют собой сложные технические объекты проектирования, решающие в составе ЛА различные функциональные задачи: определение пилотажно-навигационных параметров в режиме взлета, горизонтального полета, посадки ЛА; слежение за техническим состоянием бортового оборудования; координация работы всех бортовых подсистем; сбор, хранение, обработка и выдача пилоту на средства индикации объективной информации, получаемой как от информационно-измерительной системы ЛА, так и от органов управления информационно-управляющего поля кабины пилота.

Работы по проектированию архитектуры БЦВС для перспективных типов ЛА в настоящее время проводятся рядом предприятий Российской Федерации: ФГУП «СПб ОКБ «Электроавтоматика» им. П.А. Ефимова» (Санкт-Петербург), ВНИИРА (Санкт-Петербург), МИЭА (г. Москва), НПО «Полет» (Нижний Новгород), ФГУП «НИИ авиационных систем» (г. Москва), ОАО «РПКБ «Раменское» (г. Москва) и др. Достигнутые в настоящее время разработчиками результаты защищены патентами [1–5]: RU 2413280 C1; RU 2413655 C2; RU 108868 U1; RU 106404 U1; RU 88462 U1 и др.

Общими техническими решениями в предлагаемых разработчиками архитектурах перспективных БЦВС являются 4 принципа – открытости архитектуры, модульности, унификации и стандартизации, функциональной и аппаратной интеграции. Однако есть и отличия, влияющие на эффективность применения различных вариантов архитектур БЦВС в решении прикладных бортовых задач.

Существенным для показателей качества БЦВС является тип примененного в них внутрисистемного интерфейса. В проектах известных сегодня отечественных образцов БЦВС [4–6] в качестве внутрисистемного интерфейса используются интерфейсы ARINC664 (Gigabit Ethernet 1000Base-SX, AFDX), CompactPCI (PICMG 2.0, D3.0), PCI Express, RapidIO, VME64x и др.

Построение БЦВС на основе последовательных внутрисистемных интерфейсов типа ARINC664 с высокой скоростью передачи данных требует включения в состав БЦВС специализированных устройств сопряжения (контроллеров), обеспечивающих взаимодействие быстродействующих компонентов БЦВС с менее производительными бортовыми интерфейсами. Построение БЦВС на основе параллельных внутрисистемных интерфейсов типа CompactPCI, PCI Express, RapidIO (LP-LVDS), VME64x с большим количеством проводников во внутрисистемном интерфейсе сегодня не может обеспечивать высокую отказоустойчивость БЦВС при работе интерфейса в гигагерцовом диапазоне частот и, следовательно, отказобезопасность работы БЦВС в целом.

Таким образом, тип внутрисистемного интерфейса является определяющим в выборе архитектуры БЦВС. Очевидно, перспективная архитектура БЦВС базируется сегодня на сетевых технологиях с применением высокоскоростных последовательных внутрисистемных интерфейсов, допускающих коммутацию электрических межмодульных соединений и, следовательно, возможность построения динамически реконфигурируемых вычислительных структур. Одним из путей практической реализации перспективных БЦВС является построение БЦВС с применением технологии коммутируемых высокоскоростных интерфейсов SpaceWire. Структурная схема перспективной БЦВС, разработанной в ФГУП «СПб ОКБ «Электроавтоматика» им. П.А.Ефимова» [2, 7], представлена на рисунке. БЦВС построена на основе унифицированных конструктивно-функциональных модулей (КФМ). В качестве КФМ выступают разные по назначению модули:

– вычислительные модули, производящие сложные расчеты для управления полетом ЛА;
– модули ввода-вывода, обеспечивающие функции обмена информацией по последовательным каналам (ПК), по мультиплексным каналам обмена (МКИО), обмен разовыми командами (РК);

– графические модули, обрабатывающие изображение для его вывода на средства бортовой индикации, например, по интерфейсу ARINC 818 (Fibre Channel 1x);
– модули массовой памяти, предназначенные для хранения в своем постоянном запоминающем устройстве функционального программного обеспечения и занесения его при инициализации БЦВС в оперативное запоминающее устройство модулей – модуль вычислительный (МВ);
– модули электропитания, обеспечивающие преобразование напряжения бортовой резервированной сети во вторичные напряжения, необходимые для электропитания КФМ.

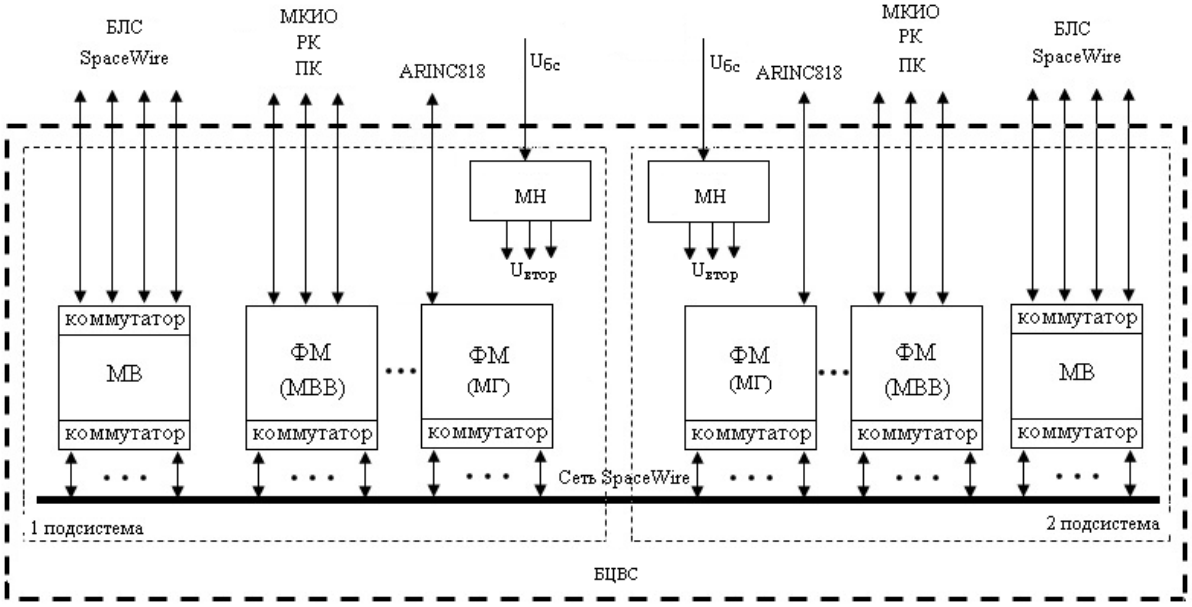


Рисунок. Структурная схема БЦВС разработки ФГУП «СПб ОКБ «Электроавтоматика» им. П.А. Ефимова: ФМ – функциональный модуль; МВВ – модуль ввода–вывода; МГ – модуль графический; МН – модуль электропитания; БЛС – бортовая локальная сеть; $U_{бс}$ – напряжение бортовой сети; $U_{втор}$ – вторичные напряжения

Как следует из рисунка, БЦВС содержит две подсистемы, состоящие из модулей МВ, МН и ФМ различного назначения. Дублирование подсистем обеспечивает уровень резервирования бортового оборудования, достаточный для достижения требуемой по техническому заданию наработки на отказ БЦВС. Каждый КФМ содержит неблокирующий коммутатор интерфейса SpaceWire, выполненный на элементной базе комплекта «Мультиборт», разработанного компанией ОАО НПЦ «Элвис».

Преимуществами использования в составе БЦВС в качестве внутрисистемного интерфейса каналов SpaceWire являются:
– повышение надежности работы внутренней сети обмена данными благодаря возможности полного резервирования каналов связи за счет соединения КФМ БЦВС по топологии «двойная звезда»;
– достижение максимального уровня пропускной способности внутрисистемного интерфейса в границах соединенных компонентов по топологии «полносвязная сеть»;
– обеспечение программно управляемого исполнения функциональных задач авионики за счет соединения компонентов БЦВС по топологии «полносвязная сеть».

1. Абросимов О.В., Егоров К.А., Итенберг И.И., Ковальчученко А.Ф., Куликов Д.А., Сивцов С.А., Федосов Е.А., Чуянов Г.А. Базовая несущая конструкция платформы интегрированной модульной авионики. Патент на полезную модель №106404 U1 RU, МПК G06F 1/16. №2010129389/08. Заявл. 15.07.2010. Опубл. 10.07.2011.
2. Богданов А.В., Васильев Г.А., Виноградов П.С., Егоров К.А., Зайченко А.Н., Ковернинский И.В., Петухов В.И., Романов А.Н., Смирнов Е.В., Уткин Б.В., Федосов Е.А., Шукалов А.В. Платформа интегрированной модульной авионики. Патент на полезную модель №108868 U1 RU, МПК G06F 9/00, №2011121962/08. Заявл. 01.06.2011. Опубл. 27.09.2011.
3. Герлих Х. Модульная система авионики самолета. Патент №2413655 C2 RU, МПК B64C 19/00. №2008123940/11. Заявл. 16.11.2006. Опубл. 10.03.2011. Бюл. № 7.
4. Егоров К.А., Итенберг И.И., Ковернинский И.В., Тимченко А.П., Федосов Е.А., Чуянов Г.А. Платформа интегрированной модульной авионики. Патент №2413280 C1 RU, МПК G06F 9/02. №2009127190/08. Заявл. 14.07.2009. Опубл. 27.02.2011. Бюл. № 6.
5. Егоров К.А., Итенберг И.И., Ковернинский И.В., Тимченко А.П., Федосов Е.А., Чуянов Г.А. Платформа интегрированной модульной авионики. Патент на полезную модель №88462 C1 RU, МПК G06F 9/00. №2009127040/22. Заявл. 14.07.2009. Опубл. 10.11.2009.

6. Севбо В., Орлов А., Лошаков А. Многопроцессорный вычислительный комплекс для задач «жесткого» реального времени // Современные технологии автоматизации. – 2007. – № 3. – С. 32–38.
7. Жаринов О.О., Видин Б.В., Шек-Иовсепянц Р.А. Принципы построения крейта бортовой многопроцессорной вычислительной системы для авионики пятого поколения // Научно-технический вестник СПбГУ ИТМО. – 2010. – № 4 (68). – С. 21–27.

Книга Екатерина Викторовна – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, ekovinskaya@gmail.com
Жаринов Игорь Олегович – ФГУП «СПб ОКБ «Электроавтоматика» им. П.А. Ефимова», доктор технических наук, руководитель учебно-научного центра, igor_rabota@pisem.net
Богданов Андрей Викторович – ФГУП «СПб ОКБ «Электроавтоматика» им. П.А. Ефимова», начальник отдела, andrew_bogdanov@rambler.ru
Виноградов Петр Сергеевич – ФГУП «СПб ОКБ «Электроавтоматика» им. П.А. Ефимова», зам. начальника отдела, vinogradov@elavt.spb.ru

УДК 623.64:623.74

РЕАЛИЗАЦИЯ СТРУКТУРЫ ДАННЫХ, ИСПОЛЬЗУЕМЫХ ПРИ ФОРМИРОВАНИИ ИНДИКАЦИОННОГО КАДРА В БОРТОВЫХ СИСТЕМАХ КАРТОГРАФИЧЕСКОЙ ИНФОРМАЦИИ

П.П. Парамонов, П.В. Коновалов, И.О. Жаринов, Ю.А. Кирсанова, С.Б. Уткин

Рассматривается структура данных, используемых при построении индикационного кадра, выводимого на экран бортовой системы картографической информации.
Ключевые слова: навигация, картография, бортовые системы.

Наличие на борту самолета системы, предоставляющей экипажу доступ к геопространственным данным, позволяет повысить эффективность управления и точность навигации летательного аппарата [1]. Человеку привычно восприятие информации в виде образов, поэтому картографическая информация, представленная на средствах бортовой индикации в форме структурированного графического изображения, позволяет также уменьшить утомляемость летного состава и увеличить безопасность перелетов. Пример отображения картографической информации на средствах бортовой индикации представлен на рис. 1.

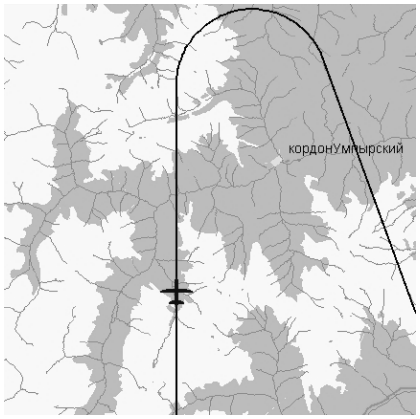


Рис. 1. Фрагмент отображаемой на средствах индикации картографической информации с отмеченным маршрутом полета

Визуализация картографической информации является ресурсоемким процессом [2]. Несмотря на непрерывный рост производительности вычислительных средств, на данный момент внедрение геопространственных данных для решения задач навигации и пилотирования в полной мере на борту летательных аппаратов не реализовано. В этой связи актуальной является задача разработки специализированной структуры данных, используемой в программном обеспечении формирования и вывода геопространственных данных на средства бортовой индикации.

Геопространственные данные представляют собой цифровой массив данных, включающий метаданные и картографические модели. Метаданные содержат справочную пространственную информацию. Картографические модели включают:

– модель местности, представляющую собой векторизованные данные об объектах местности;
– модель рельефа, содержащую данные о неровностях земной поверхности в матричном виде;
– модель обстановки, содержащую векторизованные данные об объектах навигационной и тактической обстановки в районе полетов.

Метаданные содержат следующую инфорацию: идентификационную информацию *array_id*; контрольную информацию *array_size*; данные *scale_index* об источнике картографической информации, используемой для построения моделей; параметры картографической проекции *plane_arguments*[]; пара-

метры системы координат *plane_X*, *plane_Y*, *pixel_size* и адресную информацию *data_adress*[]. Структура метаданных в синтаксисе языка С имеет вид

{ *array_id*, *array_size*, *scale_index*, *plane_arguments*[], *plane_X*, *plane_Y*, *pixel_size*, *data_adress*[] }.

Модель местности включает в себя классификационные данные, метрические данные, семантические данные. Классификационные данные включают в себя указатели на свойства объектов, которые определяют их отношение к одному из следующих элементов картографической информации: гидрография и гидротехнические сооружения; дорожная сеть и дорожные сооружения; населенные пункты; растительный покров и грунты; рельеф суши; промышленные объекты; социально-культурные объекты; границы, ограждения. Структура модели местности в синтаксисе языка С имеет вид

{ *model_id*, *model_size*, *metric_descriptors*[], *semantic_descriptors*[], *placement_descriptors*[] },

где *model_id*, *model_size* – справочные данные о модели; *metric_descriptors*[] – таблица адресов, указывающих на место хранения метрической информации; *semantic_descriptors*[] – таблица адресов, указывающих на место хранения семантических данных; *placement_descriptors*[] – таблица адресов, указывающих на информацию о расположении объектов модели местности.

Метрические данные представляют собой координаты объектов. В зависимости от форм и размеров объектов в масштабе представления геопространственных данных различают следующие виды объектов местности:

- площадной объект, метрика которого включает последовательность координат, описывающих положение границ объекта;
- линейный объект, метрика которого включает последовательность координат, описывающих положение осевой линии объекта;
- векторный объект, метрика которого включает координаты двух точек, первая из которых указывает на местоположение центра объекта, а вторая – на точку конца вектора, проведенного из центра и определяющего ориентацию объекта относительно сторон света;
- точечный объект, метрика которого включает координаты, указывающие на местоположение центра объекта в отображаемом фрагменте карты.

Структура метрических данных в синтаксисе языка С имеет вид

{ *object_id*, *object_key*, *visualization_parameters*[], *object_color*, *object_primitives*[] },

где *object_id*, *object_key* – идентификационная информация о метрическом объекте; *visualization_parameters*[] – параметры визуализации объекта (атрибуты отображения); *object_color* – цвет, используемый для отображения объекта на средствах индикации; *object_primitives*[] – графические примитивы (линии, знаки, дуги, окружности, специальные символы и пр. [3]), используемые для построения объекта.

Семантические данные представляют собой данные, специфичные для отдельных типов объектов (названия объектов, высоты препятствий, значения магнитных склонений поля Земли и т.п.).

Модель рельефа относится к пространственным моделям, представляемым в растровом виде, и содержит массив данных высот рельефа в зоне полета летательного аппарата. Данные по высотам рельефа представляют собой набор значений, характеризующих максимальные превышения высот поверхности земли над определенным уровнем в пределах заданной области полета летательного аппарата. Модель рельефа в синтаксисе языка С имеет вид

{ *model_id*, *model_size*, *relief_descriptors*[] },

где *model_id*, *model_size* – идентификационная и справочная информация; *relief_descriptors*[] – таблица адресов, указывающих на место хранения данных о рельефе местности в массиве геопространственных данных. *Модель обстановки* имеет структуру, аналогичную структуре модели местности, и отличается от нее только типом хранимой информации (навигационная и тактическая обстановка в зоне полета).

Для обработки геопространственной информации в бортовых системах картографической информации используется условная система координат и послойное представление моделей [4]. Послойное представление определяет способ хранения геопространственных данных, при котором в пределах каждого слоя отображается только часть общей информации. Принцип послойного представления данных показан на рис. 2. Из анализа рис. 2 следует, что Слой 1 содержит часть семантических данных модели местности, сообщающих пилоту названия населенных пунктов в зоне полета. Слой 2 содержит метрические данные модели местности, представляющие карту мира в заданном масштабе. Полученный в результате совмещения слоев информационный кадр дает представление о взаимном расположении геоинформационных объектов. Такое представление дает возможность при изменении тактической обстановки перерисовывать только отдельные слои карты, снижая вычислительные затраты на обработку геопространственной информации при выводе.

Для отображения геопространственных данных на основе введенных структур данных используются следующие принципы формирования изображений на бортовых средствах индикации:

- поточное чтение метрических данных с селекцией видов отображаемых в каждом слое данных;

– формирование списка данных, подлежащих отображению, на основе анализа семантической информации объектов, попадающих в кадр индикации и масштаба представления геопространственных данных;
– идентификация объектов по заданным координатам, определение свойств объектов и их семантических параметров;
– определение метрических данных, семантических параметров и свойств объекта по заданной идентификационной информации.

Описанная структура данных реализована в составе вычислительного модуля, входящего в состав бортовой системы картографической информации, разработанной в ФГУП «СПб ОКБ «Электроавтоматика» им. П.А. Ефимова».

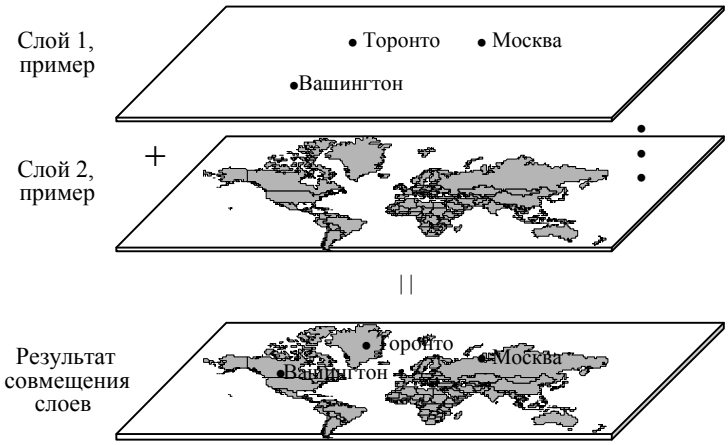


Рис. 2. Послойные формирования описания картографической информации

1. Жаринов И.О., Жаринов О.О. Бортовые системы картографической информации. Принципы построения геоинформационных ресурсов: Учебное пособие. – СПб: СПбГУ ИТМО, 2008. – 40 с.
2. Парамонов П.П., Ильченко Ю.А., Жаринов И.О. Теория и практика статистического анализа картографических изображений в системах навигации пилотируемых летательных аппаратов // Датчики и системы. – 2001. – № 8. – С. 15–19.
3. Парамонов П.П., Ильченко Ю.А., Жаринов И.О., Тарасов П.Ю. Структурный анализ и синтез графических изображений на экранах современных средств бортовой индикации на плоских жидкокристаллических панелях // Авиакосмическое приборостроение. – 2004. – № 5. – С. 50–57.
4. Копорский Н.С., Видин Б.В., Жаринов И.О. Система бортовой картографической информации пилотируемых летательных аппаратов. Основные принципы построения // Сб. трудов 10-й международной конференции «Теория и технология программирования и защиты информации». – СПб: СПбГУ ИТМО, 2006. – С. 18–23.

Парамонов Павел Павлович – ФГУП «СПб ОКБ «Электроавтоматика» им. П.А. Ефимова», доктор технических наук, профессор, генеральный директор, postmaster@elavt.spb.ru
Коновалов Павел Викторович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, olkesomewhere@gmail.com
Жаринов Игорь Олегович – ФГУП «СПб ОКБ «Электроавтоматика» им. П.А. Ефимова», доктор технических наук, доцент, руководитель учебно-научного центра, igor_rabota@pisem.net
Кирсанова Юлия Александровна – ФГУП «СПб ОКБ «Электроавтоматика» им. П.А. Ефимова», начальник отдела, juliakirsan@yandex.ru
Уткин Сергей Борисович – ФГУП «СПб ОКБ «Электроавтоматика» им. П.А. Ефимова», начальник сектора, ser33duck@mail.ru

УДК 004.89, 004.942, 338.24

МЕТОДЫ И КОГНИТИВНЫЕ ТЕХНОЛОГИИ ИНФОРМАЦИОННОГО ОБЕСПЕЧЕНИЯ УПРАВЛЕНИЯ РИСКОУСТОЙЧИВЫМ РЕГИОНАЛЬНЫМ РАЗВИТИЕМ

А.В. Маслобоев

Впервые исследована глобальная безопасность развития региональных социально-экономических систем, подверженных влиянию множества внешних и внутренних факторов. Разработан комплекс методов и когнитивных технологий виртуализации процессов управления рискоустойчивым региональным развитием.
Ключевые слова: когнитивные технологии, виртуализация, управление, безопасность, региональное развитие.

Современное геополитическое и социально-экономическое положение Российской Федерации (РФ) обусловливает ряд проблем внешнего и внутреннего характера, касающихся обеспечения безопасности в различных сферах и уровнях организации социума. Особую значимость эти проблемы приобретают для арктических регионов, превращающихся в одну из основных баз будущего развития страны. На

сегодняшний день Арктическая зона РФ является объектом сферы национальных интересов ведущих мировых держав, что ослабляет позиции присутствия РФ в Арктике, владеющей значительными ее территориями, и формирует вектор угроз национальным интересам РФ в этом районе – геополитическим, социально-экономическим, оборонным, демографическим и экологическим. Повышение интереса к Российской Арктике обусловливает высокую актуальность темы защиты интересов РФ в Арктической зоне и выводит задачу обеспечения глобальной безопасности развития Арктических регионов России на передний план, позиционируя ее как самостоятельную проблему, требующую научной проработки. Решение данной задачи затрудняется необходимостью интеграции, обработки и анализа большого объема разноплановой информации для различных ведомств, а также согласованности информационного взаимодействия соответствующих структур безопасности в едином информационном поле. Анализ мер, осуществляемых РФ по развитию сферы информационных технологий для задач обеспечения безопасности в Российской Арктике, свидетельствует о том, что их эффективность существенно снижается отсутствием единого информационного пространства Арктической зоны РФ [1], на базе которого представляется возможным формирование целостной информационной инфраструктуры для комплексного решения задач управления глобальной безопасностью арктических регионов.

Целью исследований является разработка и анализ когнитивных моделей, методов и технологий, ориентированных на создание сетецентрических проблемно-ориентированных мультиагентных виртуальных пространств, обеспечивающих комплексную информационно-аналитическую поддержку и повышение эффективности согласованной деятельности субъектов региональной безопасности на всех этапах решения задач управления рискоустойчивым развитием и функционированием социально-экономических систем (СЭС) регионального уровня, а также синтеза допустимых траекторий безопасного развития региона с учетом его внутренней динамики и внешних факторов.

Современный этап развития компьютерных наук обозначил новый мэйнстрим в области создания технологий построения интеллектуальных систем поддержки принятия решений в сфере прогнозирования и стратегического планирования развития сложных динамических систем – когнитивные информационные технологии. Когнитивные информационные технологии, в частности, технология мультиагентных систем, являются достаточно эффективным инструментом для решения широкого круга трудноформализуемых задач в различных предметных областях, связанных с управлением организационными и техническими системами, характеризующимися динамичностью, разнообразием и сложностью входящих в их состав компонентов, а также огромными объемами потоков циркулирующей разноплановой информации и высокими требованиями к времени ее обработки. Таким образом, когнитивный подход и реализующие его технологии могут найти применение для решения задач управления безопасностью развития региональных СЭС и их компонентов на основе агентно-ориентированной виртуализации процессов управления данными системами. Агентно-ориентированная виртуализация и создание проблемно-ориентированных мультиагентных виртуальных пространств [2] – это новый виток развития методологии когнитивного управления сложными системами. Конвергенция когнитивных, мультиагентных и веб-технологий обеспечивает базис для получения качественно новых технологических решений в области разработки средств информационно-аналитической поддержки управления безопасностью функционирования сложных динамических систем и процессов.

Управление региональной безопасностью на основе мультиагентных виртуальных пространств позволяет проиграть и оценить возможные варианты диффузии ситуации на треке развития региона на агентных имитационных моделях в многомерном пространстве критериев для последующего синтеза траекторий безопасного функционирования региональных СЭС, адекватных стратегиям регионального развития. В таком ключе концепция мультиагентной виртуализации может быть расширена на решение широкого спектра задач, связанных с управлением региональным развитием.

В ходе исследований сформулирована и решена научно-техническая проблема управления глобальной безопасностью развития СЭС регионального масштаба в Арктической зоне РФ. Для решения этой проблемы разработаны методы и средства виртуализации проблемно-ориентированной деятельности субъектов регионального управления, созданы когнитивные модели и технологии формирования проблемно-ориентированных мультиагентных виртуальных пространств для задач управления глобальной безопасностью. Формализован и реализован когнитивный подход к решению задач информационно-аналитической поддержки управления региональной безопасностью, основанный на интеграции методов концептуального, системно-динамического и мультиагентного моделирования сложных динамических систем, а также метода и соответствующей когнитивной мультиагентной технологии виртуализации процессов управления безопасностью. Реализация предложенного подхода обеспечила возможность адаптивного моделирования целенаправленной деятельности субъектов безопасности и формирования проблемно-ориентированных виртуальных пространств как ситуационно-коалиционных мультиагентных систем управления безопасностью, взаимодействующих в синтезируемой единой информационно-аналитической среде.

Впервые исследована глобальная безопасность развития региональных СЭС, подверженных влиянию множества внешних и внутренних факторов. Развиты понятийный аппарат управления глобальной безопасностью, сетецентрический подход к организации управления в проблемно-ориентированных рас-

пределенных информационных системах, методы и технологии мультиагентной виртуализации и концептуального синтеза динамических моделей сложных систем. На основе интеграции полученных результатов разработаны новые когнитивные модели и технологии, обеспечивающие возможность формализации и численного решения задач синтеза траекторий безопасного развития региональных систем в многомерном пространстве критериев и позволяющие проводить выработку рекомендаций по обеспечению региональной безопасности в меняющихся условиях. В итоге предложен комплекс методов и технологий виртуализации процессов управления безопасностью функционирования региональных СЭС, дающий возможность формирования расширяемой многофункциональной сетецентрической информационно-аналитической среды поддержки управления региональной безопасностью. Полученные теоретические и методологические результаты представляют собой новый когнитивный подход к построению региональных распределенных информационных систем на основе сетецентрических мультиагентных виртуальных пространств, интегрированных в глобальную информационную инфраструктуру.

Практическая значимость результатов исследования заключается в разработке методологических основ, конкретных методов и технологий управления глобальной безопасностью, реализация которых обеспечивает формирование расширяемой, многофункциональной информационной инфраструктуры, ориентированной на поддержку стабильного функционирования и развития региональных СЭС в условиях меняющейся внутренней и внешней среды в краткосрочной и долгосрочной перспективах. На основе полученных в работе практических результатов сформирована мультиагентная виртуальная сетецентрическая информационно-аналитическая среда поддержки управления безопасностью развития Мурманской области – МИАС «Безопасный виртуальный регион», представляющая собой иерархическое виртуальное пространство региона как интеграционную площадку для проблемно-ориентированных ситуационно-коалиционных мультиагентных систем управления региональной безопасностью, мониторинга социально-экономического развития и кризисного реагирования.

Специфическими особенностями синтезируемого виртуального пространства региона как целостной информационно-аналитической среды поддержки управления региональной безопасностью с унифицированной точкой доступа на основе веб-технологий [2] являются мультиагентная реализация, сетецентрическая организация, децентрализация функций управления, интеграция разнородных информационных и исполнительных ресурсов, синергетические свойства (самоорганизация компонентов на основе коалиционных взаимодействий когнитивных агентов).

Применение мультиагентного подхода для задач управления рискоустойчивым региональным развитием позволяет создать адекватную информационно-аналитическую среду поддержки управления региональной безопасностью, учитывая распределенность, динамичность и структурную сложность региональных компонентов. При таком подходе представляется возможность реализовать виртуализацию функций управления отдельными составляющими региональной безопасности за счет делегирования их когнитивным агентам, а на основе проблемно-ориентированных коалиционных взаимодействий агентов, можно будет обеспечить эффективное функционирование самоорганизующегося виртуального пространства региона и его компонентов, а также поддержание допустимого уровня безопасности развития на перспективу. Мультиагентные виртуальные пространства регионов – следующий этап на пути создания виртуальных тренажеров для чиновников, менеджеров, системных аналитиков, военных, предназначенных для интеллектуальной поддержки принятия решений по управлению региональными СЭС на основе когнитивных мультиагентных моделей и технологий, а также для автоматизированного синтеза адаптивных траекторий рискоустойчивого регионального развития с учетом динамики социально-экономической среды.

Полученные результаты использовались при разработке «Стратегии социально-экономического развития Мурманской области до 2025 года», а также при реализации «Стратегии развития Арктической зоны Российской Федерации и обеспечения национальной безопасности на период до 2020 года» на территории Мурманской области. Разработанные модели, методы и программные средства использованы при разработке технологической платформы для реализации системы освещения социально-экономической и промышленно-экологической обстановки в Арктических регионах РФ.

Исследования проводились при поддержке РФФИ (грант № 12-07-00138 «Разработка когнитивных моделей и методов формирования интегрированной информационной среды поддержки управления безопасностью Арктических регионов России»).

1. Маслобоев А.В., Путилов В.А. Информационно-аналитическая поддержка управления безопасностью развития Арктических регионов России: задачи, методы, технологии // Научно-технический вестник СПбГУ ИТМО. – 2011. – № 3 (73). – С. 143–145.
2. Маслобоев А.В., Путилов В.А. Интегрированная информационно-аналитическая среда поддержки управления региональной безопасностью: этапы и технологии реализации // Труды Института системного анализа РАН. – 2012. – Т. 62. – № 3. – С. 61–73.

Маслобоев Андрей Владимирович – Институт информатики и математического моделирования технологических процессов Кольского научного центра РАН, ст. научный сотрудник; Кольский филиал Петрозаводского государственного университета, заведующий кафедрой; кандидат технических наук, доцент, masloboev@iimm.kolasc.net.ru

УДК 004.056.53

ПАРАЛЛЕЛЬНЫЙ АЛГОРИТМ ОБУЧЕНИЯ СТРУКТУРЫ БАЙЕСОВСКОЙ СЕТИ

С.А. Арустамов, В.Ю. Дайнеко

Представлена реализация масштабируемого параллельного алгоритма обучения структуры байесовской сети. Произведен сравнительный анализ последовательного и параллельного алгоритмов.

Ключевые слова: обучение структуры байесовской сети, вычисления на графических процессорах.

В основе алгоритмов обучения структуры байесовской сети (БС) лежат два подхода: ограничение условной независимости (в этом случае сеть может оказаться неориентированной, частично обученной) и оценка меры качества сети (оптимизация локального поиска с полным перебором всех комбинаций, сопровождающаяся экспоненциальным ростом вычислений с увеличением числа узлов сети). Известный гибридный алгоритм обучения Max-Min Hill-Climbing (MMHC) включает оба этих подхода и состоит в следующем [Л]. Вначале применяется алгоритм ограничения условной независимости Max-Min Parents and Children (MMPC), который на основе данных обучения восстанавливает «скелет» БС. Далее в MMHC используется алгоритм оптимизации меры качества сети Hill-Climbing (HC). Тем не менее, применение гибридного алгоритма MMHC в работе с крупными сетями не решает проблему минимизации времени обучения. Цель представляемой авторами работы – уменьшение времени обучения гибридного алгоритма MMHC.

Предлагается использование платформы параллельных вычислений на графических процессорах (ГП) Compute Unified Device Architecture (CUDA), которая позволяет добиться значительного вычислительного ускорения. Архитектура CUDA для ГП поколения Geforce 600 позволяет объединить 4 видеокарты с двуядерными ГП, что позволяет одновременно работать с 12 288 потоковыми процессорами. Основные подходы для распараллеливания алгоритма обучения MMHC заключаются в следующих оптимизациях для работы с архитектурой CUDA.

1. Разделение обрабатываемых данных на подмножества и вычисление промежуточных данных для каждого подмножества отдельными потоковыми процессорами, работающими одновременно.
2. Использование механизма синхронизации доступа к общей памяти для каждого из потоков.
3. Масштабирование размеров заданий пулом потоков для функции ядра в зависимости от выбранного количества работающих потоков в блоках вычисления и используемой памяти для их работы.
4. Сведение к минимуму использования ветвлений внутри функций ядра.

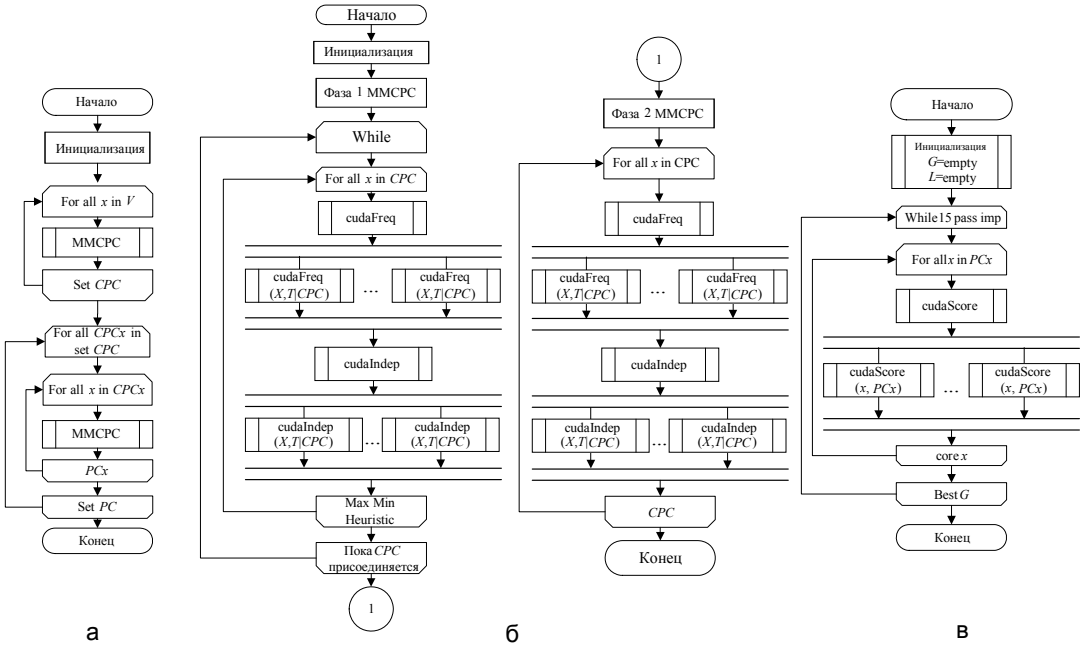


Рисунок. Блок-схема алгоритма MMPC (а); блок-схемы параллельных алгоритмов Max-Min Candidate Parents and Children (MMCPC) (б) и HC (в)

Разработаны следующие функции ядра архитектуры CUDA на языке C: cudaFreq – вычисление частоты появления всех комбинаций принимаемых значений между тестируемыми переменными из обучающего набора данных D с N экспериментами; cudaIndep – тест на независимость двух переменных X и Y при данной переменной Z; cudaScore – вычисление меры качества сети для целевого узла T от его потомков. Блок-схемы применяемых алгоритмов представлены на рисунке.

Для сравнения производительности разработанного параллельного алгоритма MMHC с последовательной реализацией использовался персональный компьютер Intel Core i5 2500 3,5 ГГц с оперативной памятью 4 ГБ; видеокарта Nvidia Geforce 9600GT с 64 потоковыми процессорами; операционная система

Ubuntu 12.04; набор разработчика CUDA Toolkit 4.0. Использовался тип данных float. Тестирование производилось пятью разными входными данными размером выборок от 1024 до 16384 записей. Размер БС – 14 узлов. В результате эксперимента для размера выборки данных 16384 время работы параллельного алгоритма сократилось до 36 раз и составило 1,46 с по сравнению с последовательным алгоритмом MMHC.

Разработанный параллельный алгоритм обучения структуры БС позволяет значительно сократить время обучения и масштабируется в зависимости от доступности потоковых процессоров.

[Л]. Tsamardinos I., Brown L.E., Aliferis C.F. The max-min hill-climbing bayesian network structure learning algorithm // Springer Science, Machine learning. – 2006. – V. 65. – № 1. – P. 31–78.

Арустамов Сергей Аркадьевич – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, профессор, sergey.arustamov@gmail.com

Дайнеко Вячеслав Юрьевич – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, daynekovy@yandex.ru

SUMMARY

REVIEW ARTICLE

INTEGRATED ON-BOARD COMPUTING SYSTEMS: PRESENT SITUATION REVIEW AND DEVELOPMENT PROSPECTS ANALYSIS IN THE AVIATION INSTRUMENT-MAKING INDUSTRY

P. Paramonov, I. Zharinov

The article deals with present situation review and analysis of development prospects for integrated on-board computing systems, used in the aviation instrument-making industry. The main attention is paid to the projects carried out in the framework of an integrated modular avionics. Hierarchical levels of module design, crates (on-board systems) and aviation complexes are considered in detail. Examples of the existing products of our country and from abroad and their brief technical characteristics are given and voluminous bibliography on the subject matter as well.

Keywords: integrated modular avionics, crates, present situation, review.

Pavel Paramonov – Saint Petersburg Scientific Design Bureau "Electroavtomatica" n.a. P. A. Efimov, General director, D.Sc., Professor, postmaster@elavt.spb.ru

Igor Zharinov – Saint Petersburg Scientific Design Bureau "Electroavtomatica" n.a. P. A. Efimov, Head of learning scientific center, D.Sc., Associate professor, igor_rabota@pisem.net

OPTICAL AND OPTICAL ELECTRONIC SYSTEMS. OPTICAL TECHNOLOGIES

ASPECTS OF HIGH-STRENGTH OPTICAL FIBER MANUFACTURE

A. Korobeynikov, Yu. Gatchin, K. Dukel'skiy, E. Ter-Nersesyantz

Analysis results of the basic technological processes for high-strength optical fiber manufacture are presented. Major factors influencing the durability of fiber optical paths are considered. Manufacturing recommendation for high-strength long-length optical paths by MCVD method is given.

Keywords: optical fiber, fiber optical waveguides, defects formation.

Anatoly Korobeynikov – Pushkov Institute of Terrestrial Magnetism, Ionosphere and Radio Wave Propagation of the Russian Academy of Sciences St. Petersburg Branch (SPbF IZMIRAN), Deputy director for science, D.Sc., Professor, Korobeynikov_A_G@mail.ru

Yuri Gatchin – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Department head, D.Sc., Professor, gatchin@mail.ifmo.ru

Konstantin Dukel'skiy – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Associate professor, PhD, KDukel@GOI.ru

Egiche Ter-Nersesyantz – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Associate professor, PhD, volokno@goi.ru

DETERMINATION OF OPTICAL LOSSES AT FIBERS JOINING WITH DIFFERENT MODE-FIELD DIAMETER

A. Varlamov, A. Kulikov, V. Strigalev, S. Varzhel', S. Aksarin

For optical losses reducing at fiber optic systems creation the angular distribution of the output radiation of birefringent optical fibers with elliptical stress cladding with 4, 12, 16 and 18 mol. % GeO_2 is measured. Obtained data make it possible to predict losses at joining of optical fibers with elliptical stress cladding and different concentrations of GeO_2 with each other, as well as standard telecom fibers, such as SMF-28, other birefringent optical fibers and integrated optics components.

Keywords: mode-field diameter, joining of optical fibers, birefringence.

Andrey Varlamov – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, student, b.a.b.92@bk.ru

Andrey Kulikov – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Assistant, a.kulikov86@gmail.com

Vladimir Strigalev – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Professor, PhD, Associate professor, vstrglv@mail.ru

Sergey Varzhel' – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Postgraduate, vsv187@gmail.com

Stanislav Aksarin – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Postgraduate, oskar2002@list.ru

CERIUM IONS INFLUENCE ON A LUMINESCENCE AND ABSORPTION OF MOLECULAR SILVER CLUSTERS IN SILICATE GLASSES AFTER ION EXCHANGE

I. Dyomichev, V. Egorov, E. Postnikov, E. Sgibnev, A. Sidorov, T. Khrushcheva

Luminescence and absorption spectrums of silicate glasses containing silver, embedded by ion exchange, were investigated. Intensive luminescence of such glasses was shown to appear right after ion exchange that can be

associated with neutral molecular clusters of silver Ag_n formation. The increase of cerium ions concentration was also shown to cause the increase of luminescence intensity of the samples. The following UV-treatment decreases luminescence intensity. Thermal treatment under 300-350 ^{0}C temperature increases the luminescence intensity.
Keywords: luminescence, cerium, ion exchange, silver clusters.
Ivan Dyomichev – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, postgraduate, demichevia@yandex.ru
Vladimir Egorov – S&R Institute of Nanophotonics and Optical Informatics of Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, engineer, egorovvl@gmail.com
Evgeniy Postnikov - S&R Institute of Nanophotonics and Optical Informatics of Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, engineer, werrus@yandex.ru
Evgeniy Sgibnev – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, student, sgibnevem@gmail.com
Alexander Sidorov – S&R Institute of Nanophotonics and Optical Informatics of Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Professor, D.Sc., Professor, aisidorov@qip.ru
Tatiana Khrushcheva – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, engineer; Saint Petersburg State Polytechnic University, Associate professor, PhD, Khruta@qip.ru

APPLICATION OF TIME-OF-FLIGHT PMD-CAMERAS FOR DISTANCE MEASUREMENT TO WATER SURFACE

D. Krysin, A. Nebylov

The article deals with results that confirm the possibility of time-of-flight PMD-cameras application for short-distance measuring (0.2-5 m) to water surface. Main features of the PMD-technology are considered. Developed experimental setup is presented. An experiment, confirming the efficiency of PMD-cameras under natural conditions is described. Experimental 3D-images of water surface shape are presented with corresponding explanation. Graphical results of statistical processing of valid pixels of 3D-images are presented. 3D-image processing algorithms are proposed for estimate of distance. The results may be useful in solving the problems of liquid level measurement, determining the parameters of sea waves and measuring the height of the movement of object above the water surface.
Keywords: time-of-flight method, PMD-camera, 3D-image, distance measurement, water surface, experimental setup, image formation.
Dmitriy Krysin – Saint Petersburg State University of Aerospace Instrumentation, assistant, wtxt@ya.ru
Alexander Nebylov – Saint Petersburg State University of Aerospace Instrumentation, Department head, D.Sc., Professor, nebylov@aanet.ru

PHOTONICS AND OPTICAL INFORMATICS

RECURRENT DATA PROCESSING IN SPECTRAL OPTICAL COHERENCE TOMOGRAPHY BASED ON KALMAN FILTERING

M. Volynsky, I. Gurov

The method of dynamic data processing in spectral optical coherence tomography based on Kalman filtering is suggested. Adaptive choice possibility of necessary number of spectral samples to ensure required resolution for a particular investigated sample is shown. Examples of experimental data processing obtained at biological objects investigating are presented.
Keywords: optical coherence tomography, spectral interferometry, Fourier transform, Kalman filter.
Maxim Volynsky – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Associate professor, PhD, maxim.volynsky@gmail.com
Igor Gurov – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Department head, D.Sc., Professor, gurov@mail.ifmo.ru

ANALYSIS OF THE WAVE PACKET PARAMETERS BASED ON DYNAMIC EVALUATION OF THE SPECTRAL COMPONENTS PHASE

E. Vorob'yeva, I. Gurov

The method of dynamic analysis for the wave packet parameters is considered. The method involves selection and evaluation of the amplitudes and phases of the separate monochromatic components of a wave packet. Positions of the intensity maximum of the wave packet can be determined using initial phases of monochromatic components estimated by nonlinear discrete Kalman filter.
Keywords: wave packet, spectral components, initial phase, Kalman filter.
Elena Vorob'yeva – LLC «Motorola Mobility», software developer, lenavorobyeva@gmail.com
Igor Gurov – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Department head, D.Sc., Professor, gurov@mail.ifmo.ru

AMINO ACIDS IDENTIFICATION IN THE HUMAN SKIN BY SPECTRA APPLICATION OF 0,05-1,2 THZ RANGE

A. Ezerskaya, M. Tsurkan, O. Smolyanskaya, A. Kapralova, N. Balbekin, I. Romanov, I. Prozheev, E. Strepitov, I. Prosanov

Spectra of human skin and tableted preparations of amino acids in 0,05 ÷1,2 THz range are obtained by terahertz spectroscopy method with time resolution. Comparison of the amino acids spectral lines is made for their identification in the skin. It is shown that the following amino acids are the most clearly visible in the skin spectrum: tryptophan, cysteine, alanine, glycine, phenylalanine and histidine. Terahertz spectral lines of the mentioned amino acids are associated with molecules oscillating modes, caused by chemical reactions, hydration processes and molecules conformal changes in the human skin. Determination of amino acids spectral lines and other specific molecules, involved in the metabolic processes of the skin, is necessary for the development of terahertz diagnostic tool with maximum sensitivity and selectivity.

Keywords: terahertz spectroscopy, diagnostics, human skin, amino acids.

Anna Ezerskaya – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, student, a.a.ezerskaya@gmail.com

Maria Tsurkan – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, postgraduate, tsurkan.maria@yandex.ru

Olga Smolyanskaya – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Associate professor, PhD, o_smolyanskaya@mail.ru

Angelina Kapralova – Institute of Laser Physics SB RAS, postgraduate, kapralova@ngs.ru

Nicolay Balbekin – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, student, godnikolay@yandex.ru

Il'ya Romanov – Tomsk State University, postgraduate, roil@elefot.tsu.ru

Igor Prozheev – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, student, prozheev.iv@gmail.com

Evgeniy Strepitov – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, student, e-strep@yandex.ru

Igor Prosanov – Institute of Solid State Chemistry and Mechanochemistry SB RAS, senior staff scientist, D.Sc., prosanov@mail.ru

STUDY OF SURFACE LAYERS MICROSTRUCTURE FOR PLANT TISSUE BY OPTICAL COHERENCE MICROSCOPY

M. Volynsky, I. Gurov, E. Zhukova, N. Margaryants, E. Ryseva

The surface layers microstructure of biological tissue on the example of plant fruit exocarp was investigated by spectral optical coherence microscopy with tunable wavelength in the 1305 ± 75 nm range and by the method of correlated optical coherence microscopy at the mean wavelength value equal to 940 nm. The experiments were performed for the intact and defect structures, and calculation of the surface profile of exocarp was performed.

Keywords: optical coherence microscopy, surface, 3D-image, exocarp.

Maxim Volynsky – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Associate professor, PhD, maxim.volynsky@gmail.com

Igor Gurov – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Department head, D.Sc., Professor, gurov@mail.ifmo.ru

Ekaterina Zhukova – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Associate professor, PhD, Associate professor, zhukova-ifmo@mail.ru

Nikita Margaryants – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, senior lecturer, fosp@grv.ifmo.ru

Ekaterina Ryseva – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, postgraduate, ri.katerina@gmail.com

ANALYSIS AND SYNTHESIS OF COMPLEX SYSTEMS

PARAMETRICAL ε-INVARIANCE SUPPORT OF THE SYSTEM OUTPUT UNDER THE PLANT UNCERTAINTY

S. Alexandrova, A. Musaev, O. Slita, A. Ushakov

The article deals with the problem of parametrical ε-invariance support of continuous system output. The problem is solved for the plant state matrix given in attending form with impracticable conditions of absolute parametric invariance. Results are illustrated with an example.

Keywords: specification uncertainty of a matrix, attending form, parametrical ε-invariance, eigenvalues and eigenvectors.

Sophia Alexandrova – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, student, alexandrova_sophie@mail.ru

Andrey Musaev – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, student, brein7@mail.ru
Olga Slita – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Associate professor, PhD, Associate professor, o-slita@yandex.ru
Anatoliy Ushakov – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Professor, D.Sc., Professor, ushakov-AVG@yandex.ru

ELECTROMAGNETIC SAFETY OF ELECTRIC TRANSPORT SYSTEMS: MAIN SOURCES AND PARAMETERS OF MAGNETIC FIELDS

N. Ptitsyna, Yu. Kopytenko, V. Ismagilov, A. Korobeynikov

Magnetic fields produced by electric drive vehicles may break electromagnetic safety. For electromagnetic safety and electromagnetic compatibility knowledge about characteristics and sources of magnetic fields in the electric transport is necessary. The article deals with analysis of available data about magnetic fields in electric cars and comparison with results of our measurements carried out in the other types of electrified transport systems.

Keywords: electric vehicle, magnetic field, electromagnetic safety, magnetic field measurement.

Natalia Ptitsyna – Pushkov Institute of Terrestrial Magnetism, Ionosphere and Radio Wave Propagation of the Russian Academy of Sciences Saint Petersburg Branch (SPbF IZMIRAN), senior scientific researcher, PhD, nataliaptitsyna@ya.ru
Yuri Kopytenko – Pushkov Institute of Terrestrial Magnetism, Ionosphere and Radio Wave Propagation of the Russian Academy of Sciences Saint Petersburg Branch (SPbF IZMIRAN), Director, D.Sc., Professor, office@izmiran.spb.ru
Valery Ismagilov – Pushkov Institute of Terrestrial Magnetism, Ionosphere and Radio Wave Propagation of the Russian Academy of Sciences Saint Petersburg Branch (SPbF IZMIRAN), scientific secretary, PhD, senior scientific researcher, IVS@izmiran.spb.ru
Anatoly Korobeynikov – Pushkov Institute of Terrestrial Magnetism, Ionosphere and Radio Wave Propagation of the Russian Academy of Sciences Saint Petersburg Branch (SPbF IZMIRAN), Deputy director for science, D.Sc., Professor, Korobeynikov_A_G@mail.ru

METHOD OF MAXIMAL INFORMATIVE ZONE FOR VIRTUAL REFERENCE STATION DEVELOPMENT IN KINEMATIC SYSTEMS OF GPS NETWORKS

R. Eminov, H. Asadov

The existing actual material on experimental assessment of positioning error in VRS GPS networks is analyzed where the mobile receiver is provided with virtual reference station. The method of highly informative zone is suggested for removal of initial uncertainty in reference station selection with the aim to develop minimal GPS network consisting of three reference stations. Methodical recommendations and directions are given for the suggested method application.

Keywords: GPS receiver, virtual reference station, kinematic schemes, positioning, geodesic network.

Ramiz Eminov – Azerbaijan State Oil Academy, Associate professor, PhD, Associate professor, eminovramiz@mail.ru
Hikmet Asadov – S&R Institute of Aerospace Informatics, Section head, D.Sc., Associate professor, asadzade@rambler.ru

COMPENSATION ALGORITHMS OF TORQUE RIPPLE FOR PRECISION ELECTRIC DRIVE WITH PERMANENT-MAGNET SYNCHRONOUS MOTOR

V. Tomasov, S. Lovlin, A. Egorov

The article deals with electric drive model considering the non-linearity of permanent-magnet synchronous motor such as cogging torque and field harmonic torque. Permanent-magnet synchronous motor parameters identification algorithms and cogging torque compensation and field harmonic torque compensation are proposed. These algorithms increase the efficiency of precision electric drive by magnifying the accuracy of positioning. The results have experimental confirmation.

Keywords: precision electric drive, permanent-magnet synchronous motor, torque ripple, cogging torque, field harmonic torque.

Valentin Tomasov – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Department head, PhD, Associate professor, tomasov@ets.ifmo.ru
Sergey Lovlin – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, postgraduate, seri-l@yandex.ru
Alexey Egorov – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, engineer- researcher, alexeykey@rambler.ru

COMPUTER SYSTEMS AND INFORMATION TECHNOLOGIES

DISTRIBUTION OF FRAGMENTED FRAMES OVER REDUNDANT CHANNELS OF COMPUTER NETWORK

A. Osipov

Algorithm of frames distribution over redundant heterogeneous channels, taking into account the physical characteristics of channels is proposed. Simulation model of the algorithm based on MATLAB / Simulink confirms the advantages of the algorithm as compared to RoundRobin and RateBalace algorithms.

Keywords: channels redundancy, channel level.

Andrey Osipov – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, postgraduate, osipov-andrey@mail.ru

APPROXIMATION OF PROBABILITY DISTRIBUTIONS IN QUEUEING MODELS

T. Aliev

For probability distributions with variation coefficient, not equal to unity, mathematical dependences for approximating distributions on the basis of first two moments are derived by making use of multi exponential distributions. It is proposed to approximate distributions with coefficient of variation less than unity by using hypoexponential distribution, which makes it possible to generate random variables with coefficient of variation, taking any value in a range (0; 1), as opposed to Erlang distribution, having only discrete values of coefficient of variation.

Keywords: probability distribution, coefficient of variation, approximation, hypoexponential distribution, hyperexponential distribution, Erlang distribution.

Taufik Aliev – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Department head, D.Sc., Professor, aliev@d1.ifmo.ru

DEVELOPMENT OF TRANSPORT SUBSYSTEM STREAMING DATA REPLICATION CLUSTER IN CORBA-SYSTEM WITH ZEROMQ TECHNOLOGY

F. Kozlov

The article deals with the peculiarities of distributed cluster system creation with streaming data replication. Ways of replication cluster implementation in CORBA-systems with ZeroMq technology are presented. Major advantages of ZeroMQ technology over similar technologies are considered in this type distributed systems creation.

Keywords: distributed system, replication, cluster, broker, ZeroMQ.

Fedor Kozlov – JSC “Industrial Information Systems”, programmer, kozlovfedor@gmail.com

LASER AND BIOMEDICAL TECHNOLOGIES

FABRICATION OF TISSUE-SIMULATIVE PHANTOMS AND CAPILLARIES AND THEIR INVESTIGATION BY OPTICAL COHERENCE TOMOGRAPHY TECHNIQUES

A. Bykov, M. Volkov, M. Volynsky, I. Gurov, M. Kinnunen, N. Margaryants, A. Popov

Methods of tissue-simulative phantoms and capillaries fabrication from PVC-plastisol and silicone for application as test-objects in optical coherence tomography (OCT) and skin and capillary emulation are considered. Comparison characteristics of these materials and recommendations for their application are given. Examples of phantoms visualization by optical coherence tomography method are given. Possibility of information using from B-scans for refractive index evaluation is shown.

Keywords: tissue-simulative phantoms, capillaries, PVC-plastisol, silicone, optical coherence tomography, refractometry.

Alexander Bykov – University of Oulu, Finland, researcher, PhD, bykov@ee.oulu.fi
Mikhail Volkov – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Associate professor, PhD, ph-m.volkov@yandex.ru
Maxim Volynsky – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Associate professor, PhD, maxim.volynsky@gmail.com
Igor Gurov – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Department head, D.Sc., Professor, gurov@mail.ifmo.ru
Matti Kinnunen – University of Oulu, Finland, researcher, D.Sc (Tech), matti.kinnunen@ee.oulu.fi
Nikita Margaryants – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, senior lecturer, fosp@grv.ifmo.ru
Alexey Popov – University of Oulu, Finland, researcher, PhD, popov@ee.oulu.fi

FORMATION OF HYPERSPECTRAL DATA ON MICROSCOPIC OBJECTS WITH HIGH DEGREE OF SPATIAL NON-UNIFORMITY OF OPTICAL DENSITY DISTRIBUTION

I. Gurov, A. Lopatin, A. Melnikov

A method of hyperspectral data formation on micro objects is suggested. The method is based on an observed sample illuminating in microscope sequentially at different wavelengths and on accumulation time tuning of photo electric charge of matrix photo detector for each image area under wavelength change. This makes it possible to obtain undistorted information at the areas with weak reflection as well as strong reflection and provides formation of resulting image with enhanced contrast.

Keywords: hyperspectral instruments, microscopy, spectrophotometry.

Igor Gurov – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Department head, D.Sc., Professor, gurov@mail.ifmo.ru

Alexander Lopatin – "LOMO", PLC, Chief optician, PhD, ailopatin@gmail.com

Alexey Melnikov – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, engineer, melnikov.alexey@gmail.com

PARAMETRIC IDENTIFICATION OF THE INTERNAL MICROSTRUCTURE OF BIOLOGICAL TISSUES BY STOCHASTIC DIFFERENCE EQUATIONS FORMALISM

E. Vorob'yeva, I. Gurov, A. Kirakozov

Modeling method for multilayer random heterogeneous tissues is developed based on mathematical tools of stochastic difference equations. A method for describing the random boundary of tissue is suggested. Examples of multilayer random heterogeneous tissues with different parameters of their internal microstructure are presented.

Keywords: microstructure, stochastic differential equations, multilayer tissues.

Elena Vorob'yeva – «Motorola Mobility» LLC, software developer, lenavorobyeva@gmail.com

Igor Gurov – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Department head, D.Sc., Professor, gurov@mail.ifmo.ru

Alexander Kirakozov – «Yandex» LLC, team leader, hristoforich@yandex.ru

MATERIALS TECHNOLOGY AND NANOTECHNOLOGIES

ANALYSIS OF CHEMICAL AND PHASE TRANSITIONS OF MULTICOMPONENT ALLOYS UNDER PULSED LASER EXPOSURE IN THE AIR BY METHODS OF CHEMICAL THERMODYNAMICS

V. Veiko, A. Slobodov, G. Odintsova

The article deals with computational thermodynamic method for determination of phase chemical composition of metal alloys surface formed under laser action in the atmosphere, depending on its volume components, conditions of laser exposure and atmosphere composition. By giving an example of laser heating of complex alloy (alloyed steel) in the air it is demonstrated that from a set of various possible reactions of interaction between iron, nickel or chrome with air components (oxygen, nitrogen, carbon, their compounds, atmospheric moisture, etc.), only strictly defined reactions are realized. Primarily, these are metal oxidation processes with the formation of an oxide film, whose phase and chemical composition is determined by temperature and heating time. Calculation data are confirmed by the experimental data provided by energy-dispersive X-ray spectroscopy.

Keywords: laser technologies, physical chemistry, thin films, laser exposure of metals, thermodynamics, calculation.

Vadim Veiko – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Department head, D.Sc., Professor, veiko@lastech.ifmo.ru

Aleksander Slobodov – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Professor, D.Sc., Professor, aslobd@gmail.com

Galina Odintsova – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, postgraduate, ogv-r@yandex.ru

PHOTOGRAPHIC TECHNOLOGIES BASED ON LIQUID CRYSTALS

M. Tomilin

Recording mediums on the base of liquid crystals (LC) are described in conception of two-step photography. Suggested formalism opens the possibilities to describe the radiation recording and distribution of physical fields on the surface using developed methods of traditional photography.

Keywords: recording mediums, liquid crystals, radiation, distribution of physical fields, surface, two- step photography.

Maksim Tomilin – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Professor, D.S., Professor, mgtomilin@mail.ru

COMPUTER-AIDED DESIGN SYSTEMS

VLSI PARTITIONING ALGORITHM WITH ADAPTIVE CONTROL PARAMETER

P. Filippenko

The article deals with the problem of very large-scale integration circuit partitioning**.** A graph is selected as a mathematical model describing integrated circuit**.** Modification of ant colony optimization algorithm is presented, which is used to solve graph partitioning problem. Ant colony optimization algorithm is an optimization method based on the principles of self-organization and other useful features of the ants' behavior. The proposed search system is based on ant colony optimization algorithm with the improved method of the initial distribution and dynamic adjustment of the control search parameters. The experimental results and performance comparison show that the proposed method of very large-scale integration circuit partitioning provides the better search performance over other well known algorithms.

Keywords: partitioning, ant colony algorithm, control parameter.

Petr Filippenko – Southern Federal University, postgraduate, peterpain@mail.ru

INFORMATION SECURITY METHODS AND SYSTEMS

TREES OF DAMAGES AS A MODEL OF RISKS ASSESSMENT FOR AVAILABILITY LOSSES AFTER CHANGES IN FINANCIAL INFORMATION SYSTEMS

S. Arustamov, M. Genin

The article deals with the methodology for risks assessment of availability losses in financial information systems after changes made in them by using trees of damages. A description of damages tree generation for each identified possible event is presented that potentially can lead to the system availability loss. An example is given, illustrating the methodology application that gives the possibility to choose the optimal software testing strategy.

Keywords: information systems, estimation, analysis, availability, risk of availability loss, damage, tree of damages, software testing strategy.

Sergei Arustamov – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Professor, D.Sc., Professor, sergey.arustamov@gmail.com

Mikhail Genin – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, postgraduate, gmg@rambler.ru

ECONOMICS AND FINANCES. MANAGEMENT

APPLICATION OF THEORETICAL OPTIONS PRICING TO ALGORITHMIC TRADING STRATEGIES

A. Toropov, D. Ivanov, Yu. Shpolyanskiy

We implemented analytical and numerical models of theoretical pricing for European, American, Asian and barrier options. An original computational approach is applied to the case of multiple payments of discrete dividends. The models are in the ground of the developed automatic strategies for algorithmic trading including those for market-making, hedging, risk management, and trading of combinations. These strategies are used successfully by large banks and trading firms.

Keywords: theoretical pricing, option, trading strategy, algorithmic trading.

Alexander Toropov – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, postgraduate; Tbricks AB, engineer-programmer, tav@tbricks.com, toropov@rain.ifmo.ru

Dmitry Ivanov – Tbricks AB, mathematician, dm.vl.ivanov@gmail.com

Yuri Shpolyanskiy – Tbricks AB, leading mathematician; Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Associate professor, D.Sc., Associate professor, shpolyan@mail.ru

NEW TECHNOLOGIES IN EDUCATION

PLANNING AND ESTIMATION OF EXPECTED COMPETENCES LEARNING OUTCOMES FOR FSES HPE

V. Vasilev, L. Lisitsyna

The article deals with approach for planning and estimation of competences learning outcomes for Federal State Educational Standard of Higher Professional Education (FSES HPE) in the framework of information and educational environment development of the higher education institution, based on the discipline content structuring for formation of the standard competences learning results and the principle of interdisciplinary training. Connection of FSES HPE competences learning outcomes with types and forms of students' trainings, types and forms of their control in a discipline is analyzed. Techniques of estimation means development for the current, module and intermediate control of expected competences learning outcomes for FSES HPE are considered. Examples of the developed techniques application are given.

Keywords: FSES HPE expected competences learning outcomes, graduating student competence-based model, development techniques of the discipline estimation facilities fund.
Vladimir Vasilev – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Rector, D.Sc., Professor, rector@mail.ifmo.ru
Lyubov' Lisitsyna – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Department head, D.Sc., Associate professor, lisizina@mail.ifmo.ru

SYMBIONTS OF INTELLECTS IN ERGATIC SYSTEMS

S. Sergeev

The article deals with the link between the natural (or artificial) intelligence and the environment system complexity. Symbiotic integration model is developed that includes both natural and artificial intelligence in organized technical environments. It is shown that the artificial environments tend to influence the intelligence abilities development and manifestation because of autopoetics type of intersystem links. The term "diffusive intelligence" is introduced meaning the integration of natural and artificial intelligence in structured environment.
Keywords: intelligence, artificial intelligence, hybrid intelligence, diffusive intelligence, complexity, immersive environments, communication.
Sergei Sergeev – Saint Petersburg State University of Information Technologies, Mechanics and Optics, Professor, D.Sc., ssfpost@mail.ru

APPLICATION OF INFORMATION AND COMMUNICATION TECHNOLOGIES IN COMPUTER AIDED LANGUAGE LEARNING

I. Tampel, E. Krasnova, E. Panova, K. Levin, O. Petrova

The article deals with the various ways of application for automatic speech recognition, Text-to-Speech technology, pronunciation and communication skills training, vocabulary check of the taught person, audition skills training in computer aided language learning (CALL-system). In spite of some constraints such technologies application is effective both for education problems simplification and for comfort growth of the system application.
Keywords: e-learning, foreign languages, ICT, speech technologies, automatic speech recognition (ASR), text-to-speech (TTS).
Ivan Tampel – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, tutor, PhD, senior research scientist, tampel@speechpro.com
Ekaterina Krasnova – "Speech Technologies Center", Ltd, junior research scientist, krasnova@speechpro.com
Ekaterina Panova – "Speech Technologies Center", Ltd, research scientist panova@speechpro.com
Kirill Levin – "Speech Technologies Center", Ltd, Head of speech recognition department, PhD, levin@speechpro.com
Olga Petrova – "Speech Technologies Center", Ltd, R&D Project manager, Petrova-o@speechpro.com

BRIEF REPORTS

CO-OPERATIVE ONTOLOGY DEVELOPMENT APPROACH

A. Nevidimov, I. Bessmertny

The paper deals with an approach to co-operative ontology development based on the ideology of version control systems with a central repository. Some problem solutions of ontology backward compatibility and resolution of conflicts between updates made by experts are proposed.
Keywords: ontologies, co-operative development, version control systems.
Alexander Nevidimov – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, postgraduate, alex.nevidimov@gmail.com
Igor Bessmertny – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Associate Professor, PhD, Associate Professor, igor_bessmertny@hotmail.com

AUTONOMOUS NAVIGATION OF MOBILE ROBOT BASED ON ULTRASONIC SENSOR FOR DISTANCE MEASUREMENT

M. Kostishin, I. Zharinov, V Suslov

The problem of autonomous mobile robot control based on a single ultrasonic distance measurement sensor is considered.
Keywords: navigation, control motion.
Maksim Kostishin – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, postgraduate, job.max@me.com
Igor Zharinov – Saint Petersburg Scientific Design Bureau "Electroavtomatica" n.a. P. A. Efimov, Head of learning scientific center, D.Sc., Associate professor, igor_rabota@pisem.net
Vladimir Suslov – Saint Petersburg Scientific Design Bureau "Electroavtomatica" n.a. P. A. Efimov, Chief of expert council, postmaster@elavt.spb.ru

RULES OF ARCHITECTURE DESIGN FOR ADVANCED ONBOARD DIGITAL COMPUTER SYSTEMS IN AVIONICS

E. Kniga, I. Zharinov, A. Bogdanov, P. Vinogradov

The article deals with basic design rules of advanced onboard digital computer systems in avionics. A block diagram of advanced onboard computer system is given.

Keywords: integrated modular avionics, digital computing systems, architecture.

Ekaterina Kniga – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, postgraduate, ekovinskaya@gmail.com

Igor Zharinov – Saint Petersburg Scientific Design Bureau "Electroavtomatica" n.a. P. A. Efimov, Head of learning scientific center, D.Sc., Associate professor, igor_rabota@pisem.net

Andrew Bogdanov – Saint Petersburg Scientific Design Bureau "Electroavtomatica" n.a. P. A. Efimov, Chief of science-research department, andrew_bogdanov@rambler.ru

Peter Vinogradov – Saint Petersburg Scientific Design Bureau "Electroavtomatica" n.a. P. A. Efimov, Deputy chief of science-research centre, vinogradov@elavt.spb.ru

IMPLEMENTATION OF DATA STRUCTURE IN INDICATION FRAME FORMATION FOR ONBOARD GEOSPATIAL DATA COMPUTER SYSTEMS

P. Paramonov, P. Konovalov, I. Zharinov, Yu. Kirsanova, S. Utkin

Data structure used in indication frame formation displaying in onboard navigation system is considered.

Keywords: navigation, cartography, onboard systems.

Pavel Paramonov – Saint Petersburg Scientific Design Bureau "Electroavtomatica" n.a. P. A. Efimov, General director, D.Sc., Professor, postmaster@elavt.spb.ru

Pavel Konovalov – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, postgraduate, olkesomewhere@gmail.com

Igor Zharinov – Saint Petersburg Scientific Design Bureau "Electroavtomatica" n.a. P. A. Efimov, Head of learning scientific center, D.Sc., Associate professor, igor_rabota@pisem.net

Yulia Kirsanova – Saint Petersburg Scientific Design Bureau "Electroavtomatica" n.a. P. A. Efimov, Chief of science-research department, juliakirsan@yandex.ru

Sergei Utkin – Saint Petersburg Scientific Design Bureau "Electroavtomatica" n.a. P. A. Efimov, Chief of science-research sector – main specialist, ser33duck@mail.ru

METHODS AND COGNITIVE TECHNOLOGIES FOR MANAGEMENT INFORMATION SUPPORT OF RISK-SUSTAINABLE REGIONAL DEVELOPMENT

A. Masloboev

The global security of regional socio-economic systems development liable to impact by multiple external and internal factors has been investigated for the first time. The battery of methods and cognitive technologies for management processes agent-based virtualization of risk-sustainable regional development has been designed.

Keywords: cognitive technologies, virtualization, management, security, regional development.

Andrei Masloboev – Institute for Informatics and Mathematical Modeling of Technological Processes Kola Science Center of the Russian Academy of Sciences, senior research scientist; Kola Branch of Petrozavodsk State University, Department head, PhD, Associate professor, masloboev@iimm.kolasc.net.ru

PARALLEL ALGORITHM FOR BAYESIAN NETWORK STRUCTURE LEARNING

S. Arustamov, V. Dayneko

The article deals with implementation of a scalable parallel algorithm for structure learning of Bayesian network. Comparative analysis of sequential and parallel algorithms is done.

Keywords: structure learning of Bayesian network, computing on GPUs.

Sergei Arustamov – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Professor, D.Sc., Professor, sergey.arustamov@gmail.com

Vyacheslav Dayneko – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, postgraduate, daynekovy@yandex.ru