НАУЧНО ТЕХНИЧЕСКИЙ ВЕСТНИК информационных технологий, механики и оптики май-июнь 2013 выпуск 3 (85)



ОБЗОРНАЯ СТАТЬЯ

Современное состояние и перспективы развития зарубежных тепловизионных систем	Тарасов В.В., Якушенков Ю.Г.	1
ОПТИЧЕСКИЕ И ОПТИКО-ЭЛЕКТРОННЫЕ СИСТЕМЫ. ОПТИЧЕСКИЕ ТЕХН	ЮЛОГИИ	
Способ коррекции кривизны поля в широкоугольных объективах	Безруков В.А., Карпова Г.В.	14
Способ углового селективного регулирования направленного светопропускания	Закируллин Р.С.	17
Восстановление непрерывных спектров адаптивным способом вычислительных	Сизиков В.С., Кривых А.В.	22
экспериментов с регуляризацией		•
Метод решения задачи динамического выбора маршрутов и назначения длин волн	Агеев Д.В., Переверзев А.А.	29
в сетях WDM с учетом явления четырехволнового смешивания		26
Прогнозирование качества изооражении космических объектов	Алтухов А.И., I нусарев Н.В., Коршунов Д.С.	30
ФОТОНИКА И ОПТОИНФОРМАТИКА		12
Расчет времени когерентности фемтосекундного спектрального суперконтинуума	Мельник М.В., Цыпкин А.Н.	42
в кварцевом оптическом волокне		
Коллимация и фокусировка параксиального волнового пакета, получаемого при дифракции	Дрозоов А.А., Козлов С.А.	40
в дальнеи зоне исходно однопериоднои волны		
АНАЛИЗ И СИНТЕЗ СЛОЖНЫХ СИСТЕМ		
Робастное субоптимальное управление боковым движением летательного аппарата	Фуртат И.Б.	51
в режиме захода на посадку		
Исследование колебательности процессов в апериодических непрерывных системах,	Акунов Т.А., Дударенко Н.А., Полинова Н.А.,	33
порождаемои фактором кратности сооственных чисел	Ушаков А.В.	<i>(</i> 1
Проектирование и технологическая подготовка сетей станции	I ришенцев А.Ю., Коробеиников А.Г.	61
вертикального зондирования ионосферы		
КОМПЬЮТЕРНЫЕ СИСТЕМЫ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ		
Надежность резервированного вычислительного комплекса	Богатырев В.А., Алексанков С.М., Демидов Д.В.,	67
при ограниченном восстановлении	Беззубов В.Ф.	
Анализ эффективности применения методов хеширования в файловых системах,	Иванов Е.Ю., Косяков М.С.	72
работающих в пользовательском режиме		
Оптимизация распределения запросов между кластерами отказоустоичивои	Богатырев В.А., Богатырев А.В., Голубев И.Ю.,	//
вычислительной системы	Богатырев С.В.	0.7
Алгоритм и программа поиска и исследования м-матриц	Балонин Ю.Н., Сергеев М.Б.	82
Агентное имитационное моделирование оизнес-процессов в нотации еЕРС	Ланцев Е.А., Доррер М.Г.	00
тодель планирования частотно-временного ресурса в нисходящем канале связи	Гаркуша С.Б., Басиленко Ю.А.	92
ЛАЗЕРНЫЕ СИСТЕМЫ И ЛАЗЕРНЫЕ ТЕХНОЛОГИИ		00
исследование характеристик излучения моноимпульсного Ег: у п-лазера	Назаров В.В., Хлопонин Л.В., Храмов В.Ю.,	99
в условиях пространственной неоднородности коэффициента усиления	Феооров Н.А.	102
моделирование неустоичивого резонатора моноимпульсного Nd: YAG-лазера	Дуоинкин И.Н., Пазаров Б.Б., Алопонин Л.Б., Уранов Р.Ю	103
с градиентным выходным зеркалом	лрамов Б.Ю.	
лазерные и биомедицинские технологии		100
Герагерцовые спектры отражения кожи с дерматитами	Смолянская О.А., Езерская А.А., Прожеев И.В.,	108
и морфологическими изменениями		112
Моделирование воздушного потока в каналах нерегулярной формы	Воронин А.А., Лукьянов Г.П., Перонов Р.В. Удинов С.И. Корициници А.Ю. Снотобаний Р.И.	113
Менорина измерения и расчета характеристик дыхания	Хинков С.И., Кормилицын А.Ю., Скоруоский В.И.	110
Измерение параметров дыхания дагчиком воздушных потоков	Кормилицын А.Ю., Линков С.И., Скоруоскии Б.И.	122
МАТЕРИАЛОВЕДЕНИЕ И НАНОТЕХНОЛОГИИ		120
Идентификация центров люминесценции в стеклах с ионами меди и хлора	Бабкина А.Н., Никоноров Н.В., Сидоров А.И.,	129
И	Шахвероов І.А., Ширинев П.С.	122
изменение поглощательной спосооности кремния при фемтосекундном	1 ук И.В., Шаноыоина І.Д., Яковлев Е.Б.	133
лазерном воздеиствии		127
механизм микроструктурирования системы SIO ₂ /SI при оолучении	Скворцов А.М., Луинь Конг Ту, Лалецкии Р.А.	13/
сканирующим пучком импульсного волоконного лазера	Warnes M.D. Hoginica D.D.	112
получение наноструктурированных пленок Ат ₂ O ₃ методом электрохимического	Жуков м.д., Левичев Д.Д.	145
технологии приборостроения. организация производства		1.47
Влияние шероховатости поверхности на усилие распрессовки прессовых соединении	Иванов А.Ю. Леонов Д.Б.	14/
методы и системы защиты информации		151
методика анализа архитектуры систем защиты информации на основе типовых элементов	Сухопаров М.Е., Соловьев И.Н., Леоеоев И.С.,	131
	Комаров И.И.	
НОВЫЕ ТЕХНОЛОГИИ В ОБРАЗОВАНИИ	<i>K</i>	
Автоматизация составления заданий по математике методом вариации параметров	Костюченко Е.В.	133
КРАТКИЕ СООБЩЕНИЯ		
Фотодеструкция наночастиц серебра в фото-термо-рефрактивных стеклах	Игнатьев Д.А., Игнатьев А.И., Никоноров Н.В.	158
Оптико-электронный комплекс для экспресс-анализа обогатимости руд твердых полезных	Алёхин А.А.	159
ископаемых оптическим методом		1
Организация библиотек в Altium Designer	Кузнецова О.В., Романова Е.Б.	160
Применение гиперсписков в разработке веб-портала	Бессмертный И.А., Пономарева Е.П.	162
SUMMARY		164



Издание Санкт-Петербургского национального исследовательского университета информационных технологий, механики и оптики

ГЛАВНЫЙ РЕДАКТОР..... доктор технических наук, профессор В.О. Никифоров РЕДАКЦИОННАЯ КОЛЛЕГИЯ доктор технических наук, профессор А.А. Бобцов доктор технических наук А.В. Бухановский доктор технических наук, профессор В.А. Валетов доктор физико-математических наук, старший научный сотрудник Т.А. Вартанян доктор технических наук М.А. Ган доктор технических наук, профессор Ю.А. Гатчин кандидат технических наук, старший научный сотрудник Н.Ф. Гусарова доктор технических наук, профессор А.В. Демин кандидат технических наук, доцент Н.С. Кармановский (заместитель главного редактора) доктор физико-математических наук, профессор Ю.Л. Колесников доктор физико-математических наук, профессор С.А. Козлов доктор технических наук, профессор А.Г. Коробейников доктор технических наук, профессор В.В. Курейчик доктор технических наук, доцент Л.С. Лисицына кандидат технических наук, доцент В.Г. Мельников доктор технических наук, профессор Ю.И. Нечаев доктор технических наук, профессор Н.В. Никоноров доктор технических наук, профессор А.А. Ожиганов доктор технических наук, профессор П.П. Парамонов доктор физико-математических наук, старший научный сотрудник Е.Ю. Перлин доктор технических наук, профессор И.Г. Сидоркина доктор технических наук О.А. Степанов доктор технических наук, профессор В.Л. Ткалич доктор технических наук, профессор А.А. Шалыто доктор технических наук, профессор Ю.Г. Якушенков

Ответственный секретарь Г.О. Артемова Редактор, верстка Н.Ф. Гусарова Редактор Л.Н. Точилина Перевод Н.Г. Лещикова Графика М.В. Герасимова

Адрес: 197101, Санкт-Петербург, Кронверкский пр., 49, НИУ ИТМО, комн. 330

Телефон / факс (812) 233 45 51 http: //ntv.ifmo.ru E-mail:karmanov@mail.ifmo.ru



ПРАВИЛА ПРЕДСТАВЛЕНИЯ РУКОПИСЕЙ

ОБЩИЕ ПОЛОЖЕНИЯ

Редакция принимает к рассмотрению оригинальные (нигде ранее не опубликованные, за исключением материалов научных конференций) работы по тематике журнала, представляющие новые научные результаты, полученные лично авторами статьи. Публикуются также обзоры, характеризующие современное состояние актуальных направлений развития науки и техники, сообщения о научных конференциях, материалы научных дискуссий и рецензии на новые книги.

Тематика журнала включает разделы, отражающие достижения науки, техники и технологии в области современного приборостроения по направлениям: оптика, оптические системы и технологии, оптическое материаловедение и нанотехнологии, информационные и компьютерные системы и технологии, методы и системы защиты информации, системный анализ, системы автоматического управления, мехатроника, системы автоматизированного проектирования, микроэлектроника, биомедицинские технологии. В качестве дополнительных разделов журнал также публикует работы по экономике и финансам.

Материал статьи должен быть изложен в соответствии со следующей структурой.

В аннотации, рассчитанной на самый широкий круг читателей, необходимо кратко (объемом до 100 слов) изложить научное содержание статьи.

Во введении необходимо представить содержательную постановку рассматриваемого вопроса, краткий анализ известных из научной литературы решений, критику их недостатков и преимущества (особенности) предлагаемого подхода.

В основном тексте статьи должна быть представлена строгая постановка решаемой задачи, изложены и обстоятельно разъяснены (доказаны) полученные утверждения и выводы, приведены результаты экспериментальных исследований или математического моделирования, иллюстрирующие сделанные утверждения. Основной текст статьи может быть разбит на отдельные разделы: «Постановка задачи», «Основной результат», «Результаты моделирования» и т.п.

В заключении необходимо кратко сформулировать основные результаты, прокомментировать их и, если возможно, указать направления дальнейших исследований.

Все статьи проходят процедуру рецензирования. После ее завершения автору направляется экземпляр рукописи с замечаниями рецензента и научного редактора для доработки. При опубликовании статьи указываются даты ее поступления и последующих переработок.

ОСНОВНЫЕ ТРЕБОВАНИЯ К ОФОРМЛЕНИЮ СТАТЕЙ

Объем статьи, включая иллюстрации, таблицы и список литературы, не должен превышать 6 страниц.

Оригинал-макет статьи выполняется в редакторе Word for Windows. Шрифт Times New Roman, размер шрифта 12 pt, межстрочный интервал одинарный, поля сверху 20 мм, остальные поля 25 мм. Абзацный отступ 10 мм.

Аннотация печатается шрифтом Times New Roman 10 обычный, межстрочный интервал одинарный, выравнивается по ширине страницы. В аннотации не допускаются ссылки на цитированную литературу и громоздкие формулы.

Название раздела печатается шрифтом Times New Roman 12 полужирный, межстрочный интервал одинарный, отделяется от предыдущего раздела полуторной пустой строкой, а от последующего текста одной пустой строкой.

Название подраздела печатается шрифтом Times New Roman 12 полужирный, межстрочный интервал одинарный, пустыми строками не отделяется.

Формулы набираются в редакторе Microsoft Equation Ed. 3.0. Параметры стиля: текста — Times New Roman, прямой; функции — Times New Roman, прямой; переменной — Times New Roman, наклонный; греческих букв — Symbol, прямой; символов — Symbol, прямой; матрицы-вектора — Times New Roman, прямой, полужирный; чисел — Times New Roman, прямой. Размеры: обычный — 12 рt, крупный индекс — 10 рt, мелкий индекс — 8 рt, крупный символ — 16 рt, мелкий символ - 12 pt. Отдельные переменные в тексте допускается набирать в текстовом режиме с соблюдением требований, указанных для формул. Разделительный знак десятичных дробей — запятая.

При оформлении рисунков и графиков используется шрифт Times New Roman. Стиль написания переменных должен соответствовать требованиям к основному тексту.

Используются только единицы измерения, соответствующие системе СИ.

Список использованной литературы оформляется в соответствии с ГОСТ 7.0.5-2008. Нумерация позиций — согласно очередности ссылок в тексте.

Плата с аспирантов за публикацию рукописей не взимается. Более подробную информацию смотрите на сайте http://ntv.ifmo.ru

.....

Подписано к печати 15.04.2013 Тираж 350 экз. Заказ № 33(85)

.....

Отпечатано в учреждении «Университетские телекоммуникации» Адрес: 197101, Санкт-Петербург, Кронверкский пр., 49

.....

Журнал зарегистрирован Федеральной службой по надзору в сфере связи,

информационных технологий и массовых коммуникаций, свидетельство ПИ № ФС77-47243 от 10.11.2011 г. Подписной индекс 47 197 по каталогу агентства Роспечать

(Газеты. Журналы)

REDOBAT CIENTIFIC AND TECHNICAL OF INFORMATION TECHNOLOGIES, MECHANICS AND OPTICS CAHKT NUMBER 3 (85) MAY-JUNE 2013 ISSN 2226-1494 NAUČNO-TEHNIČESKIJ VESTNIK INFORMACIONNYH TEHNOLOGIJ, MEHANIKI I OPTIKI **REVIEW ARTICLE** V.Tarasov, Y. Yakushenkov 1 Modern state and development perspectives of foreign infrared imagers **OPTICAL AND OPTICAL ELECTRONIC SYSTEMS. OPTICAL TECHNOLOGIES** Correction method of field curvature in wide-angle objective lens V. Bezrukov, G. Karpova 14 Method of angular selective regulation for directional light transmission R. Zakirullin 17 Continuous spectra restoration by an adaptive way of computing experiments with regularization V. Sizikov, A. Krivykh 22 29 Solution method of dynamic routing and wave length assignment problem for WDM networks D. Ageyev, A. Pereverzev with taking into account four-wave mixing effect A. Altukhov, N. Gnusarev, D. Korshunov 36 Image quality forecasting for space objects PHOTONICS AND OPTICAL INFORMATICS Coherence time calculation of femtosecond spectral supercontinuum in quartz optical fiber M. Melnick, A. Tsypkin 42 Collimation and focusing of paraxial wave packet obtained by diffraction in the far area A. Drozdov, S. Kozlov 46 of initial single-cycle wave ANALYSIS AND SYNTHESIS OF COMPLEX SYSTEMS I Furtat 51 Robust suboptimal control of aircraft lateral movement at landing conditions Process oscillativity study in aperiodic continuous systems, generated by eigenvalues T. Akunov, N. Dudarenko, N. Polinova, A. Ushakov 55 multiplication factor Design and engineering background for station networks of vertical ionosphere sounding A. Grishentcev, A. Korobeynikov 61 COMPUTER SYSTEMS AND INFORMATION TECHNOLOGIES Reliability of redundant computing complex with recovery limitations V. Bogatyrev, S. Aleksankov, D. Demidov, V. Bezzubov 67 Efficiency analysis of hashing methods for file systems in user mode E. Ivanov, M. Kosyakov 72 77 Queries distribution optimization between clusters of fault-tolerant computing system V. Bogatyrev, A. Bogatyrev, I. Golubev, S. Bogatyrev The algorithm and program of M-matrices search and study Yu. Balonin, M. Sergeev 82 Agent-based approach to business processes simulation in eEPC notation Ye. Lantsev, M. Dorrer 86 Planning model of time-frequency resource in the downlink of LTE technology S. Garkusha, Yu. Vasylenko 92 LASER SYSTEMS AND LASER TECHNOLOGIES 99 Radiation characteristics research of the Q-switched Er:YLF laser under conditions V. Nazarov, L. Khloponin, V. Khramov, N. Fedorov of the gain spatial inhomogeneity Numerical simulation of Q-switched Nd: YAG laser with unstable resonator and output I. Dubinkin, V. Nazarov, L. Khloponin, V. Khramov 103 variable reflectivity mirror LASER AND BIOMEDICAL TECHNOLOGIES Terahertz reflectance spectra of skin dermatitis and morphological changes O. Smolyanskaya, A. Ezerskaya, I. Prozheev, 108 E. Strepitov A. Voronin, G. Luk'yanov, R. Neronov Computational modeling of airflow in nonregular shaped channels 113 S. Hankov, A. Kormilitsin, V. Skorubskiy The method for measurement and calculation of breath characteristics 118 Breath parameters measurement by air flow sensor A. Kormilitsin, S. Khankov, V. Skorubskiy 122 MATERIALS TECHNOLOGY AND NANOTECHNOLOGIES Identification of luminescent centres in glass with copper and chlorine ions A. Babkina, N. Nikonorov, A. Sidorov, 129 T. Shakhverdov, P. Shirshnev I. Guk, G. Shandybina, Ye. Yakovlev Modification of silicon absorptivity under femtosecond laser pulse 133 Microstructuring mechanism of SiO,/Si system under irradiation A. Skvortsov, Huynh Cong Tu, R. Khaletskiy 137 by a scanning pulsed fiber laser beam Al₂O₃ nanostructured films creation by method of electrochemical anodizing M. Zhukov, V. Levichev 143 INSTRUMENT-MAKING TECHNIQUE. ORGANIZATION OF PRODUCTION Effect of surface roughness on the strength of press connections A. Ivanov, D. Leonov 147 INFORMATION SECURITY METHODS AND SYSTEMS M. Sukhoparov, I. Solovyev, I. Lebedev, I. Komarov Analysis method of information security systems architecture based on standard elements 151 NEW TECHNOLOGIES IN EDUCATION Automated preparing of mathematical tasks by the parameter variation method 155 E. Kostvuchenko BRIEF REPORTS Photo destruction of silver nanoparticles in photo-thermo-refractive glass D. Ignatiev, A. Ignatiev, N. Nikonorov 158 Optical-electronic system for express analysis of ore dressability for mineral raw materials A. Alekhin 159 by optical method Organization of libraries in Altium Designer O. Kuznetsova, E. Romanova 160 Hyper-lists applying to web-portal development I. Bessmertny, E. Ponomareva 162

SUMMARY

164

УДК 621.384 СОВРЕМЕННОЕ СОСТОЯНИЕ И ПЕРСПЕКТИВЫ РАЗВИТИЯ ЗАРУБЕЖНЫХ ТЕПЛОВИЗИОННЫХ СИСТЕМ В.В. Тарасов, Ю.Г. Якушенков

Виктор Васильевич Тарасов – доктор технических наук, профессор, генеральный директор ОАО «ЦНИИ «Циклон». Хорошо известны его работы в области создания высокочувствительных матричных приемников излучения на основе микроболометрических чувствительных слоев, структур с множественными квантовыми ямами и сверхрешеток типа II. Пионерскими для нашей страны являются работы В.В. Тарасова по созданию плоскостных дисплеев на основе органических светодиодов. Одновременно В.В. Тарасов ведет преподавательскую деятельность, с 2000 г. являясь профессором кафедры оптико-электронных приборов МИИГАиК. Им опубликованы 3 монографии. Он автор и соавтор 38 научных статей, 29 авторских свидетельств и патентов, 7 учебно-методических пособий.



Юрий Григорьевич Якушенков – доктор технических наук, профессор, Заслуженный деятель науки и техники РСФСР. С 1972 г. и по настоящее время является заведующим кафедрой оптико-электронных приборов МИИГАиК. С 1985 г. по 2010 г. одновременно с заведованием кафедрой являлся деканом факультета оптического приборостроения МИИГАиК. Научные труды Ю.Г. Якушенкова составляют 21 книгу, более 120 статей, 27 авторских свидетельств на изобретения и патентов. За последние 3 года им опубликованы два учебника, одна монография и 10 статей в реферируемых журналах. Его учебник для оптических специальностей вузов «Теория и расчет оптико-электронных приборов» выдержал шесть изданий, награжден Дипломом почета ВДНХ, переведен на английский язык, удостоен в 2001 г. премии Правительства РФ в области науки и техники. Он является членом президиума Учебно-методического объединения вузов России по образованию в области приборостроения и оптотехники. Член редакционной коллегии журнала «Научнотехнический вестник информационных технологий, механики и оптики»

Излагаются краткие сведения о программах развития и современном состоянии разработок зарубежных тепловизионных систем и их элементной базы, в первую очередь, оптических систем и фотоприемных устройств. Особое внимание уделяется тепловизионным системам 3-го поколения, работающим в двух и более спектральных диапазонах и использующим матричные фотоприемные устройста. Приводятся примеры разработок тепловизионных систем военного, гражданского и «двойного» назначения. Рассматриваются перспективы развития тепловизионных систем в ближайшие годы.

Ключевые слова: тепловизионная система, оптическая система, фотоприемное устройство.

Введение

Тепловизионная техника и технологии – одно из самых быстро развивающихся направлений оптикоэлектронного приборостроения. На всех представительных мировых форумах, конференциях, выставках оптико-электронных систем (ОЭС), например, ежегодных конференциях «Оптика для обороны и безопасности», проводимых Международным обществом по оптической технике (SPIE), или на ежегодных Международных форумах «Optics-Expo», организуемых Всероссийским выставочным центром, тепловизионные системы (TBC) занимают очень большую часть в программах и экспозициях [1].

Большинство ТВС работают в одном, двух и значительно реже – в нескольких диапазонах инфракрасной (ИК) области оптического спектра. В принятых за рубежом обозначениях этими диапазонами с соответствующими граничными длинами волн являются:

- NWIR (Near Wave Infrared) ближневолновый (0,7–0,9 мкм);
- SWIR (Short Wave Infrared) коротковолновый (0,9–3,0 мкм);
- MWIR (Middle Wave Infrared) средневолновый (3–5 мкм);
- LWIR (Long Wave Infrared) длинноволновый (8–14 или 8–12 мкм).

Показательно, что со времени выхода в свет публикаций [2–4] прошло совсем немного времени, однако ведущие зарубежные компании заметно увеличили номенклатуру устройств, применяемых в TBC, а также и самих систем.

Наиболее распространенными фотоприемными устройствами (ФПУ) в ТВС продолжают оставаться ФПУ на базе тройных соединений КРТ (кадмий–ртуть–теллур); СКЯ (квантово-размерные структуры – структуры с квантовыми ямами); СР-II (сверхрешетки типа II); МБ (микроболометры). Большое число фирм в 25 странах мира (США, Канада, Великобритания, Франция, Германия, Россия, Израиль, Италия, Индия и др.) занято созданием и совершенствованием ТВС самого различного применения, как военного, так и гражданского [IIN – 2013, № 2, С. 1].

Объем мирового рынка ТВС гражданского применения в 2011 г. вырос на 18% по сравнению с предыдущим годом и составил 3,2 млрд долларов США (\$), хотя рынок ТВС «двойного» (гражданского и военного) применения в 2012 г. увеличился всего на 2% из-за сокращения в США таких крупных программ, как ТWS (прицельные системы) и DVE-FOS (системы для ночного вождения). Можно отметить, что объем продаж ТВС в 2012 г., близкий к 7 млрд \$, несколько уменьшился по сравнению с 2011 г. из-за уменьшения военной составляющей бюджета США и ряда европейских стран. Однако в странах Ближнего Востока и Азии он сохранился и будет расти в ближайшие 5 лет [IIN – 2013, № 2, С. 2].

Ожидается, что к 2017 г. рынок продаж ТВС вырастет на 11%, в первую очередь, за счет развития следующих направлений:

- ТВС военного назначения для наземных, воздушных, космических, ракетных и военно-морских комплексов;
- ТВС гражданского и «двойного» назначения, применяемые для управления транспортными средствами, в охранных и антитеррористических системах, в радиометрии и термографии и др.;
- ТВС, работающие активно-пассивным методом и использующие ФПУ на базе InGaAs и МБ;
- разработка более совершенных охлаждаемых матричных ФПУ на базе КРТ, СКЯ и СР-II, в частности, работающих при температурах охлаждения около 150±20 К и обозначаемых как НОТ (High Operating Temperature) -ΦПУ;
- разработка более совершенных неохлаждаемых матричных ФПУ (МБ).

Таким образом, можно сделать вывод о продолжающемся совершенствовании и расширении круга применения ТВС. Ниже будут показаны основные направления развития этих систем и их элементной базы. В конце статьи будут приведены некоторые примеры ТВС военного, гражданского и «двойного» назначения.

Некоторые программы развития тепловизионных систем

Непрерывно растет число программ развития TBC и их основных узлов. Так, Директорат вооружений научно-исследовательской лаборатории Военно-воздушных сил США (AFRL/RW) и Агентство противоракетной обороны (MDA) составили список наиболее актуальных областей исследований из 18 наименований, в число которых входят:

- распознавание автономных целей;
- методы и аппаратура для моделирования работы систем защиты от стратегического и тактического ракетного оружия [IIN – 2013, № 1, С. 5];
- совершенствование имитаторов наблюдаемых сцен и разработка на их основе испытательной аппаратуры, с помощью которой можно будет воспроизвести характеристики обнаруживаемых с помощью ОЭС и ТВС целей для различной фоноцелевой обстановки, а также предсказать степень совершенства разрабатываемых систем для имитируемых условий эксплуатации;
- исследование узлов ТВС, работающих активным методом (лидаров или ладаров) лазеров, приемных систем, систем формирования лазерного излучения, сканирующих систем, дальномеров, систем обработки и распознавания принимаемых сигналов и изображений.

Агентство по перспективным научно-исследовательским разработкам DARPA (США) в 2013 г. создало рассчитанную на 45 месяцев программу создания облегченных (менее 1 кг), дешевых, экономичных и малогабаритных ТВС на базе матричных ФПУ, устанавливаемых на стрелковом оружии и шлемах и предназначенных для значительного повышения характеристик систем обнаружения угроз, распознавания и идентификации целей в дневных и ночных условиях, а также в сложной метеорологической обстановке (программа PIXNET – Pixel Network for Dynamic Vizualization) [IIN – 2013, № 2, С. 1]. Предусматривается, что нашлемная аппаратура может работать в двух или нескольких спектральных оптических диапазонах (в видимом, NWIR-, SWIR-, MWIR- и LWIR-диапазонах). Системы питания должны обеспечить непрерывную работу в течение четырех или восьми часов с длительностью рабочих циклов в 25%, 50% и 100% от этого времени. Наименьший формат нашлемных ФПУ должен составлять 1280×1024 пикселов (далее размерность опускается), для прицелов стрелкового оружия – 640×480, а частота кадров – 60 Гц.

Программа DARPA рассматривает отдельно возможность создания двухдиапазонных ФПУ с соотношениями 1:1 размеров пикселов, чувствительных в SWIR- и LWIR-диапазонах, при формате 640×512 и 4:1 при форматах 1280×1024 и 2048×1536. В состав разрабатываемой аппаратуры должны входить следующие основные блоки: ОЭС визуализации (камера), беспроводная связь со смартфоном, система объединения получаемых изображений и дисплей. Предполагается, что основу канала, работающего в SWIR-диапазоне (до 1,5–1,7 мкм), могут составить ФПУ на базе InGaAs в сочетании с термоэлектрическими охлаждающими устройствами и системой коррекции неоднородности отдельных пикселов ФПУ.

Очень актуальной остается проблема обнаружения минных полей, которая включена в программу научных и технологических задач Агентства по развитию науки и технологий США (Science and Technology Broad Agency) [IIN – 2012, № 6, С. 12].

DARPA поддерживает программу производства дешевых (стоимостью менее 500 \$) ТВС на базе неохлаждаемых МБ-ФПУ (Low Cost Thermal Imager Manufacturing – LCTI-M). В соответствии с этой программой предусматривается развитие производства дешевой моллированной (изготавливаемой путем горячего прессования) оптики, освоенного фирмой LPT [IIN – 2012, № 7, С. 4], а компания RVS активно работает над созданием МБ формата 1280×1024 с размерами пикселов 12 мкм (далее будем обозначать формат и размеры пикселов как 1280×1024/12) и камер на их основе с массой менее 25 г, объемом не более 3,5 см³ и потребляемой мощностью менее 0,5 Вт [IIN – 2012, № 1, С. 3].

Одновременно DARPA продолжает развивать ранее объявленные программы, например, MANTIS (Multispectral Adaptive Networked Tactical Imaging Systems) [2, 5] применительно к многодиапазонным нашлемной аппаратуре и прицельным устройствам, работающим в широком спектральном диапазоне, от видимого до длинноволнового ИК, и осуществляющим адаптивное объединение цифровых изображений [IIN – 2013, № 2, С. 1].

Принятая в армии США программа создания ряда (линейки) усовершенствованных прицелов (Advanced Weapon Sight Technology – AWST) рассчитана на 28 месяцев [IIN – 2013, № 6, С. 1]. Она предусматривает использование крупноформатных ФПУ с размерами пикселов не более 17 мкм, работающих как в SWIR-диапазоне при низком уровне облученности наблюдаемой сцены, так и в LWIRдиапазоне. Прицелы должны работать в комплексе с встроенным лазерным дальномером, «баллистическим» компьютером и автоматически перестраиваемой прицельной сеткой. Беспроводная видеосистема крепится на шлеме солдата. Нашлемные очки ночного видения осуществляют объединение (смешение) изображений, формируемых электронно-оптическим преобразователем 2-го поколения [6] и тепловизионным МБ-ФПУ 320×240/17 [IIN – 2012, № 6, С. 1]. В рамках программы AWST компания RVS разрабатывает крупноформатные InGaAs-ФПУ форматом 1280×1024/12,5.

Директорат ночного видения и электронных датчиков (NVESD) армии США сформулировал требования к следующему поколению ФПУ для использования их в составе TBC, предназначенных для наземных систем переднего обзора (Next Gen FLIR B-Kit). Эти системы должны работать в MWIR- и LWIR-диапазонах. Разрабатываемые ФПУ будут иметь формат 860×480/12, а затем 1280×720/12, и размещаться в дьюаровском охлаждающем устройстве с переменной полевой диафрагмой. В состав TBC должны входить блоки контроля электроники и обработки изображений. Диафрагменное число объектива может изменяться от 2 до 5 при соответствующем изменении фокусного расстояния.

В США принята программа создания пограничных интегрированных комплексов, размещаемых на специальных вышках – Integrated Fixed Towers (IFT). В состав комплексов входят и TBC [IIN – 2013, N_{2} 3, C. 9].

Технический центр стратегического командования обороны армии США по противодействию ракетно-космических угрозам (U.S. Army Space and Missile Defense Command /Army Forces Strategic Command – OSASMDC/ARSTRATTC) заинтересован в исследованиях, разработке и реализации новых инновационных технологий аппаратного и программного обеспечения комплексов для защиты от всех типов ракет. В [IIN – 2013, № 2, С. 10] приводится большой список разработок, интересующих этот центр, в частности:

- следящие ОЭС;
- органические и полимерные оптические материалы;
- характеристики целей в оптическом диапазоне спектра;
- технологии изготовления оптических и фотоэлектрических компонентов;
- дистанционное зондирование;
- тепловизионные системы и технологии;
- многодиапазонные (многоспектральные) ФПУ;
- обработка сигналов и изображений;
- моделирование и имитация ОЭС;
- облегченные оптические компоненты;
- охлаждающие устройства;

- неохлаждаемые ФПУ;
- совершенствование способов молекулярно-лучевой эпитаксии для изготовления многодиапазонных матричных ФПУ и ряд других [IIN 2012, № 10, С. 10].

Агентство MDA обратилось к аккредитованным колледжам и университетам или институтам, дающим высшее образование, с призывом проводить исследования и разработки перспективных систем противоракетной обороны (программы MSTAR – Missile Defense Agency and Technology Advanced Research). К числу основных направлений программы, помимо перечисленных выше, входят также лазеры и интегрированные ИК комплексы. Контракты по MSTAR составляют 500 тыс. \$ на 2 года и 700 тыс. \$ на 3 года [IIN – 2012, № 8, С. 4].

В США реализуется общирная программа создания ТВС высокого разрешения в MWIRдиапазоне, предназначенных для обнаружения запусков ракет и космических летательных аппаратов. Задачи, сформулированные в этой программе, содержат:

- технологию выращивания пленочных приемников методами эпитаксии;
- технологию изготовления фотодиодных ФПУ форматов 2048×2048/30 с граничной длиной волны спектральной характеристики 5 мкм и рабочей температурой 130 К; неоднородность пленок и пикселов не должна превышать 1–2%;
- разработку методов и аппаратуры радиометрических испытаний ТВС, включающих определение квантовой эффективности, эквивалентной шуму облученности, эквивалентной шуму разности температур (ЭШРТ), темновых токов, величины и неоднородности чувствительности, вклада схемы считывания и накопления зарядов (СНСЗ) в шум ФПУ.

Реализация программы рассчитана на 33 месяца [IIN – 2012, № 7, С. 7].

На 2013 г. намечены летные испытания ТВС с ФПУ 3-го поколения формата 1280×720 пикселов, разработанных компанией RVS в порядке реализации программы Common Sensor Payload – GPS по развитию беспилотных летательных аппаратов (БПЛА) и армейского разведывательного вертолета ARH. Эта компания в 2007 г. выиграла тендер стоимостью 800 млн \$ со сроком окончания работ в 2016 г. [IIN – 2012, № 1, С. 4].

Через NVESD правительственные организации и армия США собирают информацию о путях создания перспективных ТВС НОТ-типа для наблюдения и разведки обширных площадей (5×5 км) в MWIRдиапазоне с частотой выдачи информации 1–5 Гц. При этом наибольший интерес представляют следующие вопросы:

- Какие минимальные размеры пикселов ФПУ могут быть достигнуты в ближайшие три года?
- Как совершенствовать традиционную технологию гибридизации фоточувствительных матриц с очень малыми размерами пикселов и схем накопления и считывания зарядов?
- Как измерить и промоделировать функцию передачи модуляции, перекрестные связи, квантовую эффективность и коэффициенты заполнения ФПУ с очень малыми размерами пикселов?
- Какой наибольший формат может быть достигнут для ФПУ на одном чипе?
- Как оптимально сочетать требования по увеличению частоты кадров и по уменьшению себестоимости крупноформатных ФПУ с малыми размерами пикселов?

Ответы на эти вопросы должны помочь реализовать к 2015 г. крупноформатное ФПУ (4000×4000/5-8) на базе эпитаксиальных технологий [IIN – 2012, № 8, С. 14].

Долгосрочная программа развития, выдвинутая командованием сил специального назначения США (Special Operation Forces), предусматривает развитие таких технологий, как:

- создание камуфляжа обмундирования для отдельного солдата в широком спектральном диапазоне;
- разработка датчиков высокого разрешения для обнаружения и слежения за целями в видимом и ИК диапазонах;
- разработка адаптивных и всепогодных систем ночного видения;
- создание необнаруживаемых электромагнитных систем связи [IIN 2012, № 6, С. 11].

Федеральное агентство Германии по оборонным технологиям и их обеспечению (Federal Agency for Defense Technology and Procurement – BWB) развивает программу создания тепловизионных прицелов различных видов для разведки и обнаружения целей на больших расстояниях, сочетаемых с безопасным для глаза лазерным дальномером, цифровым магнитным компасом, цифровой системой контроля огня и беспроводной линией связи.

В Канаде реализуется программа внедрения ТВС в транспортные средства, прежде всего воздушные (Entranced Flight Vision Systems – EFVS), которая предусматривает использование в плохих погодных условиях нашлемных ТВС на базе МБ фирмы Max-Viz. Общая стоимость таких систем составляет около 100 тыс. \$ [IIN – 2012, № 4, С. 4].

Современное состояние ОЭС и ТВС 3-го поколения

К настоящему времени практически реализованы значения параметров ОЭС и ТВС 3-го поколения, которые 10 лет назад рассматривались как перспективные [5, 7], а именно:

- формат ФПУ 1280×720;
- частота кадров 60 Гц;
- частота просмотра внутреннего кадрового окна 480 Гц;
- диафрагменные числа объективов К=2,5 для широкого углового поля 3,85°×2,2° и К=6 для узкого углового поля 1,6°×0,9°;
- размер пиксела матричного приемника излучения 20 мкм;
- рабочие спектральные диапазоны 3,4–5,0 и 7,8–10,5 мкм;
- ЭШРТ при температуре фона 300 К 35 мК (при К=6), 25 мК (при К=2,5);
- динамический диапазон 78 дБ;
- диапазон температур наблюдаемой сцены 227–360 К;
- рабочая температура (температура охлаждения) ФПУ более 80 К.

Рассмотрим примеры такой реализации применительно к основным узлам современных ОЭС и ТВС – их оптическим системам и ФПУ.

Оптические материалы и оптические системы

Основная тенденция развития оптических систем TBC просматривается в стремлении создавать системы, способные работать в широком спектральном диапазоне, что необходимо для разработки двухи многодиапазонных ОЭС и TBC без существенного усложнения их конструкции [5]. При этом большое внимание уделяется снижению массы и габаритов оптических систем, а также их стоимости. Для систем, работающих в различных оптических диапазонах, важно достижение малого сдвига плоскости изображений при переходе от одного диапазона к другому.

Продолжают привлекать большое внимание вопросы атермализации оптических систем [8], в частности, возможность применять пассивную атермализацию путем использования линз из халькогенидных и других оптических материалов с оправами из алюминия, что особенно важно для многодиапазонных линзовых и зеркальных систем [IIN – 2012, № 3, С. 3].

Компания LPT продвигает на рынок оптические детали и компоненты, изготавливаемые путем моллирования (горячего прессования) из халькогенидных стекол GeAsSe и GaSbSe, позволяющих улучшить атермализацию объективов TBC, поскольку показатели преломления этих стекол мало зависят от температуры. Компоненты, изготовленные из таких стекол, имеют меньшую массу, нежели распространенные германиевые компоненты. При изготовлении асферических и гибридных линз компания пользуется алмазным точением [IIN – 2013, № 2, С. 4].

Фирма Murata Manufacturing (Япония) развила удешевленный технологический процесс моллирования германиевых и кремневых линз при температуре 400°С и 730°С соответственно. Изготовленные по этой технологии линзы обладают почти такой же функцией передачи модуляции, что и линзы, изготавливаемые традиционным методом полирования [IIN – 2012, № 2, С. 2].

Для 3-го поколения ТВС вождения в ночных условиях разрабатываются объективы из халькогенидных линз (компания Umicore) и из кристаллического германия (компания Ophir) [IIN – 2012, № 5, С. 3].

Компания InView Technology (США) продолжает исследования технологии изготовления микрозеркальных матриц большого формата – до 1024×768, используемых в качестве модуляторов потоков, собираемых на отдельные пикселы InGaAs-ФПУ, работающих в коротковолновом ИК диапазоне [IIN – 2012, № 12, С. 6].

Не ослабевает интерес к панорамным оптическим системам [2, 9–11], хотя большинство из публикаций последнего времени посвящено системам с оптико-механическим сканированием, а не панорамным объективам, работающим в статическом режиме. Так, компания L-3 CE сообщает о создании TBC с панорамным полем обзора (360°) по одной из координат с вращающейся камерой на базе InSb-ФПУ формата 1280×1024. В поле 360° формируются 24 отдельных кадра, строящихся в единой фокальной плоскости (плоскости изображения сцены). Частота кадров составляет 2,5 Гц [IIN – 2013, № 2, С. 8].

В ИК модуле «Attica», разработанном компанией Cassidian Optronics Gmbh (прежнее название – Carl Zeiss Optronics), для перископа командира танка «Leopard 2», используется матричное КРТ-ФПУ 384×288/25. Применение в этом модуле 2×2-микросканирования [6] позволило увеличить эффективный формат до 768×576 [IIN – 2012, № 12, С. 2].

Фотоприемные устройства

ФПУ на базе тройных соединений КРТ. Усилия многих разработчиков КРТ-ФПУ в последние годы направлены на увеличение их пространственного разрешения, удешевления за счет замены материала подложек – GaAs вместо CdZnTe, что позволяет увеличить площади подложек более чем в два раза, а также на создание более совершенных двух- и трехдиапазонных ФПУ [12–16].

Для MWIR-диапазона выпускаются КРТ-ФПУ форматов 640×512/20, 1280×1024/12 (компания Barbara Focalplane), 1280×720/15 (Raytheon).

Компания SG разрабатывает двухдиапазонные (для MWIR и LWIR-диапазонов) КРТ-ФПУ форматов 640×512/20, 640×512/24, 860×480/20. Для производства КРТ-ФПУ, работающих при температуре охлаждения T_0 более 220 К (MWIR-диапазон) и имеющих формат 1024×768/16, используется эпитаксия из металлоорганической вакуумной фазы. Последней разработкой SG явилось КРТ-ФПУ для MWIRдиапазона формата 1920×1280/12, в котором пикселы схемы считывания и накопления расположены вдоль одной из коротких сторон мозаики фоточувствительных пикселов, что заметно увеличивает коэффициент заполнения ФПУ [IIN – 2012, № 12, С. 1].

Совместно с фирмами Caeleste и Easics (Бельгия) SG по заказу Европейского космического агентства развивает производство крупноформатных КРТ-ФПУ, охлаждаемых до 77 К и чувствительных в NWIR- и SWIR-диапазонах [IIN – 2012, № 6, С. 5].

ФПУ на базе структур с квантовыми ямами (СКЯ-ФПУ). Круг применений СКЯ-ФПУ расширяется благодаря их относительно низкой стоимости и высокой стабильности и однородности параметров, даже при меньшей, чем у КРТ-ФПУ, квантовой эффективности [17]. Эти ФПУ непрерывно совершенствуются. Так, компания ТRT заметно улучшила параметры двухдиапазонных СКЯ-ФПУ, работающих на длинах волн 4,6 и 8,6 мкм [IIN – 2013, № 3, С. 4]. Новые ФПУ выполняются в гибридном исполнении. Они используют систему охлаждения фирмы Sofradir. В СНСЗ применяются емкостные трансимпедансные усилители с точным контролем смещения. Одни и те же усилители используются в МWIR- и LWIR-диапазонах. Совершенствование СНСЗ заключается в использовании двух выходов (по 10 мегапикселов на каждый), увеличении частоты кадров до 100 Гц, снижении числа шумовых электронов до 228 в MWIR- и до 1100 в LWIR-диапазонах. Величина ЭШРТ в обоих диапазонах – менее 50 мК, время накопления – 7 мс, процент годных пикселов – 99,5, рабочая температура – 70 К.

ФПУ на InGaAs. Эти ФПУ в последнее время привлекают к себе большое внимание, поскольку они работают в SWIR-диапазоне, широко используемом для создания ТВС активно-пассивного типа и систем «двойного» назначения. Появилось много сообщений о разработке крупноформатных матричных ФПУ на лавинных InGaAs-фотодиодах [18, 19]. Такие ФПУ могут обеспечить λ_{rp} =2,2 мкм и полосу пропускания частот до 25 ГГц [IIN – 2012, № 12, С. 5].

Компания RVS разработала InGaAs-ФПУ для работы в видимом, NWIR- и SWIR-диапазонах формата 1280×1024/20 [IIN – 2012, № 1, С. 2].

Матричные ФПУ на лавинных InGaAs-фотодиодах производятся компанией SBC. Они предлагаются для работы в составе ладаров автономных БПЛА и других робототехнических комплексов. Частота кадров этих устройств составляет 5–6 Гц, поле обзора – 60°×30°. При работе на дальности 20 м их разрешение в плоскости сцены составляет 45 см на один пиксел. Эта же компания для работы в 3D-пространстве предлагает матрицы на лавинных InGaAs-диодах, работающих в режиме счета одиночных фотонов [IIN – 2012, № 2, С. 4].

Компания Xenics (Бельгия) запустила в производство неохлаждаемые InGaAs-ФПУ, работающие в спектральном диапазоне 0,85–1,6 мкм и имеющие очень малое время накопления зарядов – менее 80 нс. Эти ФПУ сочетаются с системой обработки и коррекции изображений в реальном масштабе времени. Частота кадров составляет 28 Гц. При охлаждении до 77 К такие ФПУ формата 640×512/20 имеют очень низкий уровень шума – до 5 электронов в секунду, однако частота кадров при этом резко уменьшается – до 2,5 Гц [IIN – 2013, № 3, С. 2].

Компанией FLIRS создана ТВС Таи SWIR для работы в SWIR-диапазоне с форматом InGaAs-ФПУ 640×512/15. Потребляемая Таи SWIR-мощность – менее 4 Вт, масса – менее 130 г, габаритные размеры – 3,8×3,8×4,8 см [IIN – 2013, № 3, С. 3].

ФПУ на основе сверхрешеток типа II (СР). Известные преимущества ФПУ на СР перед КРТ-ФПУ и СКЯ-ФПУ, а именно, лучшая однородность, более простая технология изготовления, возможность поглощать падающее по нормали излучение, уменьшенная скорость Оже-рекомбинации, возможность более просто изменять граничную длину волны, особенно сильно стимулируют развитие этих устройств [2].

В последние годы появился ряд сообщений о разработке двух- и многодиапазонных СР-ФПУ. Так, в NU были созданы фотодиодные InAs/GaSb-ФПУ, имеющие квантовую эффективность η =50% в MWIRи LWIR-диапазонах с λ_{rp} =4,8–5,0 и 11,0 мкм соответственно. Формат ФПУ для MWIR-диапазона составил 640×512/15 [IIN – 2012, № 8, С. 4]. При комнатной температуре такие ФПУ имеют удельную обнаружительную способность $D^*=1,7\cdot10^{10}$ Вт⁻¹·см·Гц^{1/2}. При T₀=150 К они имеют $D^*=10^{13}$ Вт⁻¹·см·Гц^{1/2} [IIN – 2012, № 4, С. 2].

В NU разработаны двухдиапазонные СР-ФПУ формата 1000×1000, работающие при T₀=68 К (MWIR- и 81 К (LWIR-диапазоне) с ЭШРТ 19 мК и 27 мК соответственно, при диафрагменном числе объектива *К*=2 и времени накопления 0,13 мс при η=81% [IIN – 2013, № 3, С. 8].

Компания IRC выпускает TBC для научных исследований на базе СР-ФПУ формата 640×512/20, которая может работать в широком спектральном диапазоне 2,5–11,5 мкм или отдельно в MWIR- или LWIR-диапазонах, выделяемых сменными оптическими фильтрами. Частота кадров может достигать 475 Гц. Стоимость камеры – около 200 тыс. \$ [IIN – 2012, № 9, С. 4].

В [IIN – 2012, № 2, С. 3] сообщается о разработке институтом KRISS (Республика Корея) двухдиапазонного (MWIR- и LWIR-диапазоны) ФПУ на базе СР (InAs/InGaSb) формата 320×256, имеющего ЭШРТ, равную 100 мК при температуре 77 К. Совместно с UNM (University of New Mexico) KRISS разработал для MWIR-диапазона ФПУ формата 320×256, имеющее ЭШРТ 23,8 мК при диафрагменном числе объектива *K*=4 и температуре охлаждения 77 К.

Исследователи из JPL достигли увеличения граничной длины волны спектральной характеристики CP-ФПУ на InAsSb до 6 мкм за счет использования квантовых точек в InSb при $T_0=225$ K [IIN – 2013, N_{2} 3, C. 1]. В большинстве известных ФПУ на InSb и KPT, работающих при $T_0=150$ K, эта длина волны ограничивается примерно 4,2 мкм.

Матричные ФПУ, работающие при температурах охлаждения выше 150±20 К (НОТ-ФПУ). Эти ФПУ в последние годы привлекают большое внимание, поскольку увеличенная T_0 позволяет значительно уменьшить габариты, массу, энергопотребление и стоимость ТВС, а достаточно хорошо отработанная технология изготовления ФПУ на базе материалов группы A_3B_5 и возможность производить подложки приемников больших размеров обеспечивают им преимущества перед КРТ-ФПУ, работающими в том же MWIR-диапазоне [19]. Очень важно, что охлаждающие устройства НОТ-ФПУ по сравнению с криогенными устройствами, создающими T_0 менее 90 К, имеют существенно больший срок безотказной наработки. Это также снижает стоимость всей ТВС и упрощает ее конструкцию. Как пример, в [IIN – 2012, № 4, С. 1] указывается, что объем охлаждающей системы ТВС на базе НОТ-ФПУ, разработанной компанией L-3 CE, составляет 290 см³, а для ТВС на базе InSb – 1700 см³. Ожидается, что совершенствование охлаждающих устройств, подобно тем, что работают совместно с НОТ-ФПУ, приведет к заметному расширению рынка ТВС гражданского и «двойного» назначения в ближайшие 5 лет [IIN – 2012, № 6, С. 7].

Сегодня НОТ-ФПУ разрабатываются на базе ряда материалов – InAs/GaSb (CP), InAsSb (фотодиоды, HgCdTe (КРТ) и InSb.

Компания LMSBF выпустила на рынок гражданской продукции НОТ-ФПУ для MWIR-диапазона (3,0–4,2 мкм), работающее в сочетании с системой охлаждения по циклу Стирлинга (T_o до 135 K), имеющей срок наработки более 20000 ч. Система потребляет около 5 Вт. В ФПУ используются материалы n-типа (InAs и барьеры между слоями AlAsSb или InAsSb). Формат ФПУ – 1280×1024/12. Компания NS создала работающую в диапазоне 3,6–4,9 мкм ИК камеру с массой менее 1,5 кг на базе этого ФПУ. В другой камере той же компании используется InSb-ФПУ формата 640×512/15. Последняя имеет ЭШРТ порядка 30 мК при T_o=95 К. Масса этой камеры – 335 г [IIN – 2012, № 2, С. 2].

Компания DRST выпускает ИК камеру Zafiro 640 с НОТ-ФПУ на КРТ формата 640×480/12, работающую в MWIR-диапазоне при T₀=120 К. Масса камеры – менее 1,2 кг [IIN – 2013, № 3, С. 5]. Эта же компания предполагает использовать КРТ-ФПУ формата 640×512/12, работающую при T₀=120 К, создаваемой линейным криохолодильником типа Стирлинга со сроком службы 17000 ч, для разработки ручного целеуказателя с массой менее 0,9 кг.

Неохлаждаемые ФПУ. Неохлаждаемые и сравнительно дешевые ФПУ на базе МБ продолжают заметно совершенствоваться в последнее десятилетие. Достигнутое уменьшение размеров пикселов МБ привело к увеличению формата матриц при сохранении высокой чувствительности [20–22]. Сегодня в ряде компаний разработаны крупноформатные МБ-ФПУ с размером пикселов менее 25 мкм и ЭШРТ менее 30–50 мК, с преобразованием «аналог–цифра» внутри чипа ФПУ, значительно увеличенным динамическим диапазоном и заметным уменьшением тепловой постоянной времени. Например, компания RVS создала такие МБ-ФПУ формата 2048×1536/17.

Создание МБ с размерами пикселов 17 мкм позволило достичь для ТВС с МБ-ФПУ бо́льших дальностей распознавания целей, чем для ТВС 2-го поколения, в которых используются охлаждаемые ФПУ и оптико-механическое сканирование [IIN – 2013, № 3, С. 6].

Сообщается, что компания DRST вслед за выпуском в 2011 г. миниатюрной TBC Tamarisk 320 на VO_x -MБ формата $320 \times 240/17$ выходит в 2013 г. на рынок с миниатюрной TBC Tamarisk 640 формата $640 \times 480/17$. Габариты этой TBC – $46 \times 40 \times 31$ мм, масса – менее 60 г, энергопотребление – менее 1,5 Вт. Камера может работать с частотой кадров 30 или 9 Гц. Ее ЭШРТ составляет менее 50 мК при диафрагменном числе объектива K=1. Она снабжена сменными объективами с фокусными расстояниями от 14,25 до 65 мм и угловыми полями от 44° до 9,6°, соответственно [IIN – 2012, № 5, С. 4].

Компания UI разработала ряд МБ-ФПУ форматов 160×120/17, 384×288/17, 640×480/17 и 1024×768/17 с ЭШРТ порядка 40 мК при постоянной времени менее 10 мс. Энергопотребление этих ФПУ снижено до 60 мВт по сравнению с 110 мВт у выпускавшихся ранее МБ-ФПУ формата 384×288/25 [IIN – 2012, № 8, С. 7]. Компания СЕА/Leti (Франция) по четырехлетнему контракту объемом 26 млн евро разрабатывает ФПУ с такими же форматами, но размерами пикселов 12 мкм [IIN – 2012, № 2, С. 2].

В самое последнее время появилось сообщение о том, что наряду с МБ на традиционных материалах (VO_x и α -Si) создаются первые образцы МБ формата 640×480/12 на VNiO_x, у которого температурный коэффициент сопротивления в два раза больше, чем у VO_x.

Наряду с МБ продолжают развиваться и некоторые другие типы неохлаждаемых ФПУ. Так, компания TSB продолжает совершенствовать неохлаждаемые ФПУ на основе КД-диодов (кремний на диэлектрике) [2, 23]. Сообщается о создании таких ФПУ формата 320×240/22 с СНСЗ, помещенной на той же подложке, что и фоточувствительные пикселы. В СНСЗ применена двойная коррелированная выборка в сочетании с дифференциальным усилителем вместо одного транзисторного усилителя и аналогоцифрового преобразователя, что уменьшило зависимость выходного сигнала от температуры.

При разработке ФПУ очень большого формата создание столь же крупноформатных СНСЗ становится отдельной проблемой. Работы по ее решению ведутся рядом компаний. Можно отметить в качестве примера СНСЗ форматов 640×512/25 и 1280×1024/12,5, предложенные компанией МТ для гибридных InGaAs-ФПУ, а также другие СНСЗ для ФПУ, работающих в MWIR- и LWIR-диапазонах.

Продолжаются работы по созданию неохлаждаемых гибридных систем ночного видения [6], в которых изображение, создаваемое на фотокатоде электронно-оптического преобразователя, через микроканальную пластину поступает на ПЗС-приемник [IIN – 2012, № 6, С. 3].

Новые физические принципы работы ФПУ. Компания Tanner Research (США) предложила МБ нового типа (наноболометр), в котором поглощение излучения происходит в металлических наночастицах, работающих как оптические резонаторы. Резонансные частоты зависят от материала частиц, их размеров и формы. Падающее на них ИК излучение вызывает изменение электрического поля вокруг наночастиц. Этот принцип был реализован на примере одноэлементного болометра для длины излучения 1,5 мкм [24].

В [24] приводятся сведения о разработке ФПУ на основе PbSe, полученного осаждением из паровой фазы (VDP). Формат опытных образцов ФПУ достигал 32×32, а частота кадров – более 2 кГц. Достоинствами таких ФПУ являются высокое быстродействие, отсутствие охлаждения, совместимость с технологией кремниевых комплементарных структур металл–оксид–полупроводник (КМОП), однородность пикселов, невысокая стоимость. В перспективе разработчики ожидают увеличения формата, уменьшения размеров пикселов, доведения частоты кадров до 20 кГц и выше.

Исследования, проводимые в компании RTI International (США), посвящены разработке дешевых ФПУ для SWIR-диапазона, основанных на гетеропереходах в фуллерене PbS-C60. Опытные одноэлементные образцы обладали $\eta=20\%$ в диапазоне 0,4–1,4 мкм. Величина *D** составила 5·10¹¹ Bt⁻¹·см ·Гц^{1/2}, а плотность темнового тока при комнатной температуре – 130 нА/см². Ожидается, что такие ФПУ могут иметь $D^*=10^{13}$ Bt⁻¹·см ·Гц^{1/2} и работать в спектральном диапазоне 0,25–2,2 мкм. Формат их может составить 1000×1000, а стоить они могут около 1000 \$ [IIN – 2012, № 6, С. 4].

Некоторые примеры разрабатываемых сегодня ТВС

Двухдиапазонные ТВС активно-пассивного типа

Продолжается совершенствование систем активно-пассивного типа для военных и гражданских применений, в которых активный канал включает генератор излучения (обычно – лазер), облучающий сцену (обнаруживаемые или наблюдаемые объекты) и работающий, как правило, в NWIR-, SWIR- или LWIR-диапазонах. В пассивном канале обычно принимается собственное излучение объектов в MWIR-или LWIR-диапазонах [2, 12]. Часто пассивный канал используется для предварительного обнаружения целей, а активный, обладающий принципиально более высоким пространственным разрешением, – для распознавания и идентификации целей. Кроме того, активный канал используется для локации объектов, что позволяет получать трехмерную информацию в «смотрящем» режиме, т.е. без механического сканирования. Хорошо известными системами такого типа являются лазерные локаторы (ладары).

Совсем недавно появилось сообщение о разработке в университете NU ФПУ для активного канала этих систем на базе InAs/GaSb-CP-II формата 320×256 [IIN – 2013, № 2, С. 5]. При диафрагменном числе K=2,3 и времени накопления 10 мс для рабочей температуры 140 К в MWIR-диапазоне была получена ЭШРТ в 49 мК. Для температуры 110 К ЭШРТ уменьшилась до 13 мК при времени накопления 30 мс. Чувствительность SWIR-канала ограничивалась шумами системы считывания при температуре 150 К.

Системы военного назначения

Среди ТВС самого различного военного назначения можно выделить несколько групп, привлекающих наибольшее внимание потребителей и разработчиков. К ним можно отнести:

- системы разведки, обнаружения, целеуказания и прицеливания, устанавливаемые как на легком стрелковом оружии, так и на стационарных или передвижных наземных или морских носителях (бронетранспортеры, танки, суда и др.);
- системы воздушного и космического базирования (самолетные, вертолетные, на БПЛА и др.), решающие те же задачи разведки, прицеливания, контроля огня и т.д.;
- наземные, воздушные и космические навигационные системы и системы вождения (управления) транспортных средств;
- системы защиты от поражающих средств противника.

Сюда же можно отнести системы обнаружения различных химических соединений, радиационной угрозы, противопожарные, охранные и правоохранительные TBC, которые являются во многих случаях системами «двойного» назначения.

При рассмотрении в начале этого обзора программ развития TBC, а затем их основных узлов и требований к ним указывались современные пути их практической реализации и приводились параметры некоторых из них.

Системы наземного базирования. Всепогодные прицелы круглосуточного действия, устанавливаемые на стрелковом оружии, продолжают создаваться многими зарубежными производителями. Эти прицелы служат также для разведки и контроля огня. Они сочетаются с нашлемными очками ночного видения. Одной из основных задач, стоящих перед их разработчиками, является уменьшение времени распознавания цели, что должно на 50% повысить безопасность и выживаемость солдата-оператора [IIN – 2013, № 3, С. 7].

Примером систем, входящих в снаряжение солдата будущего, являются ТВС, разрабатываемые компанией RM по заданию бундесвера. Начатая в 2012 г. разработка включает в себя создание нескольких блоков. В их число входит система разведки и прицеливания на больших расстояниях, предназначенная для командира группы или снайпера. Эта ТВС строится на базе КРТ-ФПУ формата 640×512 и работает в MWIR-диапазоне в сочетании с лазерным дальномером, цифровым магнитным компасом, системой глобального позиционирования GPS и беспроводной линией связи [IIN – 2013, № 3, С. 14].

Правительство США развивает трехлетнюю программу создания ИК бинокля со следующими параметрами:

- угловое поле: широкое не менее 12°×9°, узкое максимально 4°×3°;
- минимальный формат ФПУ на InSb 640×480;
- два рабочих спектральных диапазона 3–5 и 8–12 мкм.

Бинокль должен входить в комплекс, в составе которого предусмотрены безопасный для глаза лазерный дальномер (измеряемая дальность до 4 км), система глобального позиционирования GPS, цифровой магнитный компас с разрешением 1°. Габариты всего комплекса – не более 29×30×12 см, масса – до 6 кг, энергопитание – от элементов со сроком службы более 3 ч [IIN – 2012, № 9, С. 9].

Другой разработкой, поддерживаемой армией и корпусом морской пехоты США, является создание облегченного бинокля для работы в видимой области (датчик изображения формата 700×400 пикселов с переменным увеличением 5[×]) и в LWIR-диапазоне (МБ формата 320×240 с переменным увеличением 2[×]). В состав бинокля массой менее 3 кг должны входить лазерный дальномер (дальность действия более 2,5 км), GPS-система и цифровой компас. Бинокль должен работать при температуре окружающей среды –25°C–50°C [IIN – 2012, № 6, С. 11].

Компания Sagem поставила французской армии 150 многофункциональных всепогодных ИК биноклей дальнего действия нового поколения. Они обеспечивают видение днем и ночью и предназначены для обнаружения противника, точного целеуказания, разведки и дальнометрирования. Бинокль имеет встроенный приемник GPS и беспроводную линию связи [25].

Целям разведки, распознавания и идентификации целей, а также определения их координат служит работающая в MWIR-диапазоне TBC переменного увеличения Sofue XF, созданная компанией Thales. Формат охлаждаемого ФПУ составляет 384×288 [IIN – 2013, № 3, С. 15].

Новые средства камуфляжа. Наряду с совершенствованием ТВС обнаружения, распознавания и идентификации различных целей создаются новые средства противодействия этим системам. Так, компания SSZCTL (Швейцария) развивает технологии камуфляжа обмундирования солдат за счет снижения излучательной способности поверхности одежды путем использования металлотканых материалов, подобных шерсти, и специальных покрытий. Использование этих материалов снижает заметность объекта более чем на 60%. Эффективные наблюдаемые температуры людей, одетых в такое обмундирование, могут снижаться до 12–21°C [IIN – 2012, № 6, С. 2].

Системы воздушного и космического базирования. Как отмечалось выше, ТВС продолжают рассматриваться как одно из средств решения многих задач, стоящих перед военно-воздушной и ракетно-космической техникой, а также перед системами разведки и противодействия ракетным угрозам. Космические ОЭС и ТВС разведки играли и играют важную роль при обнаружении запусков ракет, для чего преимущественно используется SWIR-диапазон. Для отслеживания первых этапов полета баллистических ракет (на высотной и заатмосферной стадиях полета) оптимальным считается MWIR-диапазон, а для слежения за «холодными» отделяющимися блоками – LWIR- и сверхдальний ИК диапазоны.

Примером одной из ОЭС, служащих для навигации, обзора окружающего пространства, обнаружения ракетных угроз и огневых средств противника может служить обзорная ТВС для самолетаистребителя F-35 LII, базирующаяся на шести отдельных крупноформатных InSb-ФПУ (640×480 или 1000×1000 фирмы L-3 CE), просматривающих поле обзора 360° в MWIR-диапазоне [IIN – 2012, № 9, C. 4; IIN – 2013, № 3, C. 13].

Та же компания в партнерстве с SG создает первую из восьми ИК систем подавления ракетных средств вероятного противника, развиваемых по заданию армии США. В частности, разрабатывается система защиты вертолетов от портативных носимых средств поражения на базе НОТ-КРТ-ФПУ формата 384×288, работающая в MWIR-диапазоне. Компании RVS и Boeing разработали неохлаждаемое МБ-ФПУ формата 640×480 для использования в ТВС ракетных следящих комплексах [IIN – 2012, № 9, С. 5].

Компания L-3 IOS (США) заключила контракт стоимостью 8,9 млн \$ на разработку ИК датчика, работающего в MWIR-диапазоне. Датчик устанавливается на БПЛА, предназначенном для разведки и имеющем размеры 3,4×1,0×4,3 м. Высоты полета могут достигать 4,5 км, а скорость – 230 км/ч. Датчик обладает пространственным разрешением, достаточным для обнаружения живой силы и военной техни-ки [IIN – 2012, № 12, С. 7].

Для работы в составе БПЛА компания ООЅ (США) разработала ИК камеру, в которой предусмотрены системы коррекции неоднородности отдельных пикселов, компенсации неработающих пикселов, автоматического контроля экспозиции, адаптивного контроля динамического диапазона выходных сигналов. В камере используется стандартная схема считывания и накопления зарядов ROIC фирмы Indigo/FLIR [IIN – 2012, № 11, С. 12].

Компания AV (США) получила контракт на 16,5 млн \$ для разработки к концу июня 2013 г. ТВС для БПЛА RQ-11B Raven, запускаемого с рук. В датчике используются МБ компании FLIRS 640×480/17, а также лазерный целеуказатель, совмещенный с камерой видимого диапазона [IIN – 2012, № 11, С. 8].

Системы вождения автомобилей и военной техники. Фирма Autoliv предлагает в 2013 г. двухдиапазонную систему ночного видения для вождения автомашин нового класса Mercedec-Benz. Видение в LWIR-диапазоне осуществляется с помощью МБ на окиси ванадия 320×240/17, разработанного FLIRS. Для активного канала, работающего в условиях естественной ночной освещенности, используется NWIR-диапазон. Пешеходы на пути будут обнаруживаться на расстоянии до 160 м, а крупные животные – до 100 м. В 2014 г. компании ВМW и Audi предполагают оснастить свои новые автомашины системами ночного видения, работающими только в LWIR-диапазоне. Предполагается, что цена такой системы составит около 2500 \$ [IIN – 2013, № 1, С. 2].

Для вождения военной техники постоянно разрабатываются все новые TBC. Так, компания SG предлагает крупноформатные ФПУ 1920×1080/12. Компания DRS также производит ряд систем для вождения, а именно:

- систему DVE Wide, состоящую из трех широкопольных подсистем с угловыми полями 40°×30° и ФПУ формата 640×480 и обеспечивающую поле обзора 107°×30° (вместо существующего стандартного поля 40°×30°);
- систему DVE Ultra Wide, состоящую из двух систем переднего обзора с угловыми полями 40°×30°;
- систему DVE Ultra Wide с объективом типа «рыбий глаз», имеющим угловое поле 114°×98° [IIN 2012, № 11, С. 3].

Примеры ТВС «двойного» и гражданского применения

В литературе содержится много примеров эффективного использования TBC «двойного» и гражданского применения для решения важных задач в самых различных областях [2, 5, 6, 26–36]. Можно привести еще несколько, сведения о которых появились в 2012–2013 г.г.

Медицина. Для оптической когерентной томографии компания UTC Aerospace Systems (США) предлагает новую цифровую камеру, работающую в SWIR-диапазоне (SD-OCT), в которой применены линейные ФПУ на базе InGaAs из 2048 пикселов с размерами 10 мкм. Камера может отсканировать более 76000 линий в секунду. Предполагается использовать камеру в офтальмологии, а также в волоконно-оптических спектрометрах [IIN – 2013, № 11, С. 6].

Для офтальмологических применений компания Bay Spec (США) предлагает ИКС, работающую на длине волны 1,31 мкм и позволяющую наблюдать внутриглазные полости, включая сетчатку, макулу, глазные нервы с разрешением 10 мкм. В предшествующих системах исследования и измерения проводились на длине волны 1,05 мкм [IIN – 2013, № 11, С. 6].

Компания Xenics (Бельгия) встроила линейные InGaAs-ФПУ в высокоскоростную сканирующую камеру компании Lynx. Линейка из 2048 пикселов может осуществлять сканирование с частотой до 40 кГц. Производятся также камеры с линейными ФПУ, состоящими из 512 пикселов размерами 25 мкм и 1024 пикселов размерами 12,5 мкм. Эти камеры предназначены для оптической когерентной томографии и систем машинного зрения [IIN – 2013, № 11, С. 6].

Газоанализаторы. Компания Chemring Detection Systems (США) получила заказ на создание более чем 100 ИК датчиков для обнаружения на удалении в атмосфере до 5 км облачных химических, радиационных и биологических образований естественного или искусственного происхождения. Датчики будут обнаруживать их в пассивном режиме в спектральном диапазоне 7–14 мкм [IIN – 2013, № 3, С. 16].

Для обнаружения и определения состава и концентрации различных газов рядом фирм разрабатываются газоанализаторы на базе ИК интерферометров Фабри–Перо. Так, фирма Denso (Япония) разработала такой интерферометр, работающий в MWIR-диапазоне и служащий для обнаружения газов C₂H₅OH, CO₂, CO, CH₃ и др. [IIN – 2012, № 9, С. 1].

Газы и токсические химические вещества обнаруживаются с помощью двухдиапазонной (видимый и длинноволновый ИК каналы) камеры, разработанной фирмой Mine Safety Appliances США. В ИК канале используется МБ-ФПУ компании FLIRS 640×480/17. Стоимость камеры составляет 60–80 тыс. \$ [IIN – 2012, № 5, С. 1].

Компания Telops (Канада) предлагает TBC Hyper-Cam для обнаружения и идентификации газов, а также для их визуализации в реальном масштабе времени, в которой применен ИК Фурье-спектрометр высокого разрешения с матричным МБ-ФПУ формата 320×256. Стоимость прибора – около 500 тыс. \$ [IIN – 2012, № 8, С. 5].

Гиперспектральная система визуализации, основанная на Фурье-спектрометре, в котором применено КРТ-ФПУ формата 256×256, чувствительное в спектральном диапазоне 7–11 мкм, разработана фирмой ВО (США). Система имеет спектральное разрешение 0,7⁻¹ см и обнаруживает ряд газов на расстояниях до 10 км [IIN – 2012, № 5, С. 5].

Компания TCL (Япония) применила волоконно-оптический кабель, лазерную подсветку на длине волны 9,5 мкм и термопару для создания ТВС обнаружения паров алкоголя (этанола) у водителей автотранспорта [IIN – 2012, № 9, С. 1].

Радиометрия, пирометрия и термография. Японское агентство по обороне (Japan Defense Agency) проводило с помощью ИК термографов измерения температур на АЭС «Фукусима» [IIN – 2012, № 9, С. 7]. Для измерения температур в диапазоне 700°С–1800 °С компания Ортіs GmbH (Германия) предлагает пирометр, работающий без охлаждения ФПУ в спектральном диапазоне 0–1,1 мкм [IIN – 2012, № 2, С. 4].

Новые ИК термометры (62 Мах), разработанные компанией Fluke (США), по цене (около 100 \$) близки к существующим приборам. Они имеют водо- и пылезащищенную конструкцию и могут измерять температуры от 30°C до 500°C или от 30°C до 600°C с погрешностью ±1,5% на расстояниях, в 10–12 раз превышающих размер контролируемого участка [IIN – 2012, № 6, С. 7].

Термографическая камера R300, построенная на базе MБ 320×240/23,5, предлагается компанией NEC Avio Infrared (Япония). ЭШРТ камеры – менее 3 мК. Она может применяться для измерения температур до 500°С (или до 2000°С) с погрешностями от 1°С± 1% до 2°С±2%. Камера имеет объектив переменного увеличения (зум-объектив), а также панорамную систему обзора [IIN – 2012, № 3, С. 12].

Продолжают развиваться ТВС для обнаружения очагов пожаров и контроля распространения возгораний. Требования к пространственному разрешению таких систем повышаются: сегодня нужны системы с разрешением очагов возгораний на местности порядка нескольких сантиметров, а погрешность определения их координат не должна превышать 10 м. Предполагается, что такие системы должны быть двухдиапазонными (MWIR- и LWIR-диапазоны) [2; IIN – 2013, № 3, С. 8].

Заключение

Несмотря на небольшой спад производства в 2012 г., развитие тепловизионных систем различного назначения неуклонно продолжается. Ожидается, что к 2017 г. рынок продаж тепловизионных систем вырастет на 11%. Растет число областей применения тепловизионных систем, заметно улучшается их пространственное, энергетическое и временное разрешение. Ведущие компании-разработчики тепловизионных систем наибольшее внимание уделяют двухдиапазонным тепловизионным системам с матричными фотоприемными устройствами, работающим в двух диапазонах ИК спектра.

Литература

- Якушенков Ю.Г. Международные форумы «Оптические приборы и технологии OPTICS-EXPO» // Фотоника. – 2012. – № 4 (34). – С. 30–33.
- Тарасов В.В., Торшина И.П., Якушенков Ю.Г. Инфракрасные системы 3-го поколения. М.: Логос, 2011. – 240 с.
- 3. Якушенков Ю.Г. Тенденции развития малогабаритных инфракрасных систем 3-го поколения, работающих активно-пассивным методом // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – № 3 (79). – С. 11–14.
- 4. Тарасов В.В., Якушенков Ю.Г. Тенденции совершенствования элементной базы инфракрасных систем 3-го поколения // Изв. вузов. Приборостроение. 2012. Т. 55. № 5. С. 56–63.
- 5. Тарасов В.В., Якушенков Ю.Г. Двух- и многодиапазонные оптико-электронные системы с матричными приемниками излучения. – М.: Университетская книга; Логос, 2007. – 192 с.
- 6. Тарасов В.В., Якушенков Ю.Г. Инфракрасные системы «смотрящего» типа. М.: Логос, 2004. 444 с.
- Horn S., Norton P., Cincotta T. et al. Challenges for third-generation cooled imagers // SPIE Proc. 2003. V. 5074. – P. 44–51.
- 8. Якушенков Ю.Г. Теория и расчет оптико-электронных приборов. М.: Логос, 2011. 568 с.

- 9. Bjork C., Wan W. Mid-wave infrared (MWIR) panoramic sensor for various applications // SPIE Proc. 2010. V. 7660. P. 76600B-1...9.
- Nichols J.M., Waterman J.R. Performance characteristics of a submarine panoramic infrared sensor // SPIE Proc. – 2010. – V. 7660. – P. 766005-1...9.
- 11. Fontanella J.-C., Delacourt D., Klein Y. ARTEMIS: first naval staring IRST in service // SPIE Proc. 2010. V. 7660. P. 766006-1...11.
- Тарасов В.В., Якушенков Ю.Г. Некоторые тенденции совершенствования оптико-электронных систем визуализации изображений 3-го поколения и их элементной базы (по материалам зарубежной печати) // Сб. Оптико-электронные системы визуализации и обработки оптических изображений. М.: ЦНИИ «Циклон», 2011. Вып. 3. С. 66–89.
- 13. Abbott P, Pillans L., Knowles P., McEwen R.K. Advances in dual-band IRFPAs made from HgCdTe grown by MOVPE // SPIE Proc. 2010. V. 7660. P. 766035-1...11.
- 14. Melkonian L., Bangs J., Elizondo L. et al. Performance of MWIR and SWIR HgCdTe-based focal plane arrays at high operating temperatures // SPIE Proc. 2010. V. 7660. P. 76602W-1...11.
- 15. McKeag W., Veeder T., Wang J. et al. New development in HgCdTe APDS and LADAR receivers // SPIE Proc. 2011. V. 8012. P. 801230-1...14.
- de Borniol E., Castelen P., Guellec F. et al. A 320×256 HgCdTe avalanche photodiode focal plane array for passive and active 2D and 3D imaging // SPIE Proc. – 2011. – V. 8012. – P. 801232-1...7.
- 17. Breiter R., Wendler J, Lutz H. et al. High operating temperature IR-modules with reduced pitch for SWaP sensitive applications // SPIE Proc. 2011. V. 8012. P. 80122V-1...14.
- MacDougal M., Hood A., Geske J. et al. InGaAs focal plane arrays for low light level SWIR imaging // SPIE Proc. – 2011. – V. 8012. – P. 801221-1...10.
- Dixon P., Hess C.D., Chuan L. et al. Dual-band technology on indium gallium arsenide focal plane arrays // SPIE Proc. – 2011. – V. 8012. – P. 80121V-1...7.
- 20. Mounier E. Technical and market trends for microbolometers for thermography and night vision // SPIE Proc. 2011. V. 8012. P. 80121U-1...6.
- 21. Fritze J., Munzberg M. The new megapixel thermal imager family // SPIE Proc. 2011. V. 8012. P. 801205-1...8.
- 22. Li C., Skidmore G.D., Han C.J. Uncooled VOx infrared sensor development and application // SPIE Proc. 2011. V. 8012. P. 80121N-1...8.
- Ohnakado T., Ueno M., Ohta Y. et al. Novel readout circuit architecture realizing TEC-less oferation for SOI diode uncooled IRFPA // SPIE Proc. – 2009. – V. 7298. – P. 7298OV-1...10.
- 24. Vergara G., Linazes-Herrero R., Gutierrez-Alvarez et al. VDP PbSe technology fills the existing gap in uncooled, low-cost and fast IR imagers // SPIE Proc. 2011. V. 8012. P. 80121Q-1...11.
- 25. Военно-промышленный курьер. 2013. № 3 (471). С. 12.
- Linares-Herrero R., Archilla-Prat V., Montojo-Supervielle M.T. et al. High-speed IR monitoring of a turbojet engine gas flow using an uncooled MWIR imaging sensor // SPIE Proc. - 2011. - V. 8013. - P. 8013OP-1...13.
- 27. Santa Cruz G.A., Gonzalez S.R., Dagrosa A. et al. Dynamic infrared imaging for biological and medical applications in boron neutron capture therapy // SPIE Proc. 2011. V. 8013. P. 801307-1...18.
- Kauppinen T., Siikanen S. Impruvement of energy efficiency the use of thermography and and airtightness test in verification of thermal performance of school buildings // SPIE Proc. – 2011. – V. 8013. – P. 8013O9-1...8.
- 29. Naranjo E., Baliga S., Park J. et al. IR gas cloud imaging in oil and gas applications: immunity to false stimuli // SPIE Proc. 2011. V. 8013. P. 8013OB-1...10.
- Bison P., Marinetti S., Cuogo G. et al. Corrosion detection on pipelines by IR thermography // SPIE Proc. 2011. – V. 8013. – P. 8013OF-1...10.
- 31. Vela B.M., Hsien S.-J., Palomares J.B.D.G. Comparative analysis of pulse and active thermography for investigating solder joint geometry prediction // SPIE Proc. 2011. V. 8013. P. 8013OH-1...20.
- 32. Tarin M. Pulse and lock-in IR NDT in complex structures // SPIE Proc. 2011. V. 8013. P. 801312-1...12.
- 33. Alicandro C.J., DeMarco R.W. 1024x768 XGA Uncooled Camera Core achieves new levels of performance in a small package // SPIE Proc. 2011. V. 8012. P. 80121H-1...6.
- Cramer K.E., Winfree P. Fixed Eigenvector Analysis of Thermographic NDE Data // SPIE Proc. 2011. V. 8013. – P. 8013OT-1...11.
- 35. Koulas C.E. Extracting wildfire characteristics using hyperspectral, LiDAR and thermal IR remote sensing systems // SPIE Proc. 2009. V. 7298. P. 72983Q-1...11.
- 36. Mouroulis P. Compact infrared spectrometers // SPIE Proc. 2009. V. 7298. P. 729803-1...10.

Перечень некоторых сокращений и аббревиатур, используемых в настоящем обзоре

- Издания: IIN Infrared Imaging News, SPIE Proc. SPIE Proceedings.
- Организации: AFRL/RW Air Force Research Laboratory, Munitions Directorate (США); DARPA Defense Research Projects Agency (США); KRISS Korea Research Institute of Standarts & Science; MDA Missile Defense Agency (США); NVESD Night Vision & Electronic Sensors Directorate (США).
- Компании: AV AeroViroment Inc. (США); BO Bruker Optics (США); DRST DRS Technologies (США); FLIRS FLIR Systems (США); IRC IRCamera (США); JPL Jet Propulsion Lab. (США); LMSBF Lockheed Martin Santa Barbara Focalplane (США); LPT LightPath Technologies (США); L3-CE L-3 Cincinnati Electronics (США); MT Mikro-Tasarim (Турция); NS Nova Sensors (США); NU Northwestern University (США); OOS Optimum Optical Systems (США); RM Reinmetall (Германия); RVS Raytheon Vision Systems (США); SG Selex Galileo Великобритания); SBC (Spectrolab, а Boeing Company (США); TCL Toyota Central R&D Labs (Япония); TRT Thales Research & Technology (Франция); TSB Toshiba (Япония); UI Ulis (Франция); UM Umicore (США).

Тарасов Виктор Васильевич	_	Московский государственный университет геодезии и картографии, доктор
		технических наук, профессор; ОАО «ЦНИИ «Циклон», генеральный дирек-
		тор, cyclone@asvt.ru
Якушенков Юрий Григорьевич	_	Московский государственный университет геодезии и картографии, доктор

Якушенков Юрий Григорьевич – Московский государственный университет геодезии и картографии, доктор технических наук, профессор, заведующий кафедрой, yakush@miigaik.ru

ОПТИЧЕСКИЕ И ОПТИКО-ЭЛЕКТРОННЫЕ СИСТЕМЫ. ОПТИЧЕСКИЕ ТЕХНОЛОГИИ

УДК 535.317.6 СПОСОБ КОРРЕКЦИИ КРИВИЗНЫ ПОЛЯ В ШИРОКОУГОЛЬНЫХ ОБЪЕКТИВАХ

В.А. Безруков, Г.В. Карпова

Предложен один из способов коррекции кривизны поля в широкоугольных объективах, приведена методика его реализации. Для решения задачи синтеза коррекционных компонентов использованы корректирующие элементы с определенными по величине отношениями продольных увеличений вдоль главного луча. Конструктивные параметры этих элементов математически связаны с их реально вносимыми аберрациями, в частности, такой аберрацией, как астигматизм.

Ключевые слова: широкоугольные объективы, кривизна поля, астигматизм, корректирующие линзы.

Введение

Занимаясь вопросом исправления кривизны поля, нельзя не отметить, что эта аберрация наиболее трудно поддается коррекции в широкоугольных объективах. По этой причине целесообразно рассмотреть один из способов коррекции кривизны поля в таких объективах, который основывается на преобразовании астигматизма от предшествующей части исходной оптической системы в необходимое изменение кривизны поля после корректирующего компонента.

Решение задачи определения конструктивных параметров анастигматических линз для бесконечно удаленного предмета и для положения предмета на конечном расстоянии было дано проф. М.М. Русиновым [1]. Это условие не позволяет синтезировать коррекционные компоненты с вносимым наперед заданным астигматизмом. В настоящей работе предлагается метод, который основывается на решении более общей задачи, а именно – расчета конструктивных параметров корректирующих линз конечной толщины с наперед заданным астигматизмом при определенных по величине отношениях продольных увеличений вдоль главного луча. Работа является дальнейшим развитием метода синтеза оптических систем профессора М.М. Русинова.

Метод синтеза коррекционного компонента на основе условия получения наперед заданного астигматизма линз

Как известно [1], рост меридиональной кривизны изображения по отношению к росту сагитталь-

ной кривизны в оптической системе зависит от величины отношения $\frac{Q_m}{Q_s}$, где Q_m и Q_s – продольные

увеличения в меридиональной и сагиттальной плоскостях в точках предмета и изображения на главном луче. При конечном положении предмета это отношение для оптической системы (коррекционного компонента) будет иметь вид

$$\frac{Q_m}{Q_s} = \frac{1 + \frac{\Delta' t - \Delta' F}{z_s'}}{1 + \frac{\Delta t - \Delta F}{z_s}},\tag{1}$$

где Δt , ΔF , Δt и ΔF – астигматизм вдоль главного луча в пространстве изображений, в задней фокальной плоскости, в пространстве предметов, в передней фокальной плоскости соответственно; z_s и z'_s – отрезки вдоль главного луча от фокальной точки до предмета и изображения в сагиттальной плоскости.

Рассматривая выражение (1) применительно к одиночной линзе (элементу), можно сделать вывод, что изменение величины и знака астигматизма Δt элемента, устанавливаемого после корректируемой системы, будет приводить к изменению величины отношения продольных увеличений. В свою очередь, астигматизм Δt линзы конечной толщины связан с основными ее параметрами вдоль главного луча выражением [2], которое при отрезках $t_{1,m}$ и $t_{1,s}$ от точек преломления главного луча на первой поверхности линзы до соответствующих точек предмета имеет вид

$$\frac{1}{t_{2F,s}} = \frac{\cos^{2} \varepsilon_{2}}{\frac{\cos^{2} \varepsilon_{1}}{t_{1,m}} + \frac{1}{t'_{1F,s}}} - \tilde{d} - \frac{\cos^{2} \varepsilon_{2}'}{n_{r_{2}}\Delta' t + \frac{1}{\frac{1}{\frac{1}{t_{1,s}} - \tilde{d}}} - \frac{1}{t'_{2F,s}}},$$
(2)

где $n_{r1} = \frac{n_1}{n_2}$ и $n_{r2} = \frac{n_2}{n_3}$ – относительные показатели преломления на первой и второй поверхностях.

Из выражения (2) следует, что при одних и тех же значениях отрезков $t'_{1F,s}$, и $t_{2F,s}$ углы преломления ε'_1 и падения ε_2 главного луча соответственно на первой и второй поверхностях могут иметь либо одинаковые, либо разные знаки. В соответствии с терминологией проф. М.М. Русинова линзы, имеющие одинаковые знаки углов, будем называть корректирующими линзами первого рода, а линзы, имеющие разные знаки, линзами второго рода.

В основном, силовые компоненты широкоугольных объективов обладают отрицательной кривизной поля, поэтому для ее коррекции используются обычно отрицательные линзы. Метод синтеза силовых компонентов широкоугольных объективов приведен в работе [3]. Исследование отрицательных корректирующих линз первого и второго рода при конечном положении предмета, удовлетворяющих условию (2), и положительном значении $\Delta't$ показали:

- для линз обоего рода отношение продольных увеличений, большее единицы, определяется условием Δ't >> Δt, приводящим к тому, что рост меридиональной кривизны изображения опережает рост сагиттальной в направлении положительных значений;
- для линз обоего рода отношение продольных увеличений, меньшее единицы, определяется условием Δ't << Δt, которое приводит к росту меридиональной кривизны изображения в сторону отрицательных значений, а сагиттальной – в сторону положительных значений. Таким образом, наблюдается движение той и другой кривизны навстречу друг другу.

Создавая в исходной оптической системе положительный астигматизм, будем иметь более быстрое изменение меридиональной кривизны изображения по отношению к изменению сагиттальной кривизны. Устанавливая в пространстве изображений оптической системы отрицательную корректирующую линзу первого или второго рода, получаем возможность увеличить изменение меридиональной кривизны изображения или уравнять это изменение с изменением сагиттальной кривизны в зависимости от величины отношения (1) корректирующей линзы.

Следовательно, использование обобщенного условия получения наперед заданного астигматизма линз конечной толщины (2) и результатов исследования корректирующих линз первого и второго рода позволяет реализовать на практике предлагаемый способ.

Синтез коррекционного компонента должен осуществляться по главному лучу, относительно которого ведется компоновка исходной оптической системы по методу проф. М.М. Русинова [1]. При этом число элементов, составляющих компонент, а также их род зависит от величин и знаков меридиональной и сагиттальной кривизны изображения предшествующей им оптической системы. Функция элемента состоит в преобразовании астигматизма от предшествующей ему оптической системы в заданное изменение кривизны поля после этого элемента. При этом конструктивные параметры элемента, обеспечивающие требуемое по величине отношение продольных увеличений и связь с предшествующим базовым компонентом оптической системы, определяются формулами, вытекающими из соотношения (2), которые приведены в работе [2].

Как показал предложенный метод, при синтезе коррекционных компонентов наиболее предпочтительно использовать корректирующие линзы второго рода с большим выносом входного зрачка, имею-

щие отношение продольных увеличений больше единицы, $\frac{Q_m}{Q_s} > 1$, и корректирующие линзы первого

рода с прямой ориентировкой первой преломляющей поверхности, т.е. поверхности, расположенной вогнутостью к входному зрачку, с небольшим выносом входного зрачка и отношением продольных увели-

чений, меньшим единицы, $\frac{Q_m}{Q_s} < 1$.

Результаты синтеза коррекционных компонентов

В качестве примера, подтверждающего правильность выбранных исходных позиций и эффективность способа, были синтезированы два коррекционных компонента (рис. 1, 2). График изменения астигматизма системы (рис. 2), в которой осуществлялась коррекция кривизны поверхности изображения, соответствует угловому полю $2\omega = 141^{\circ}$. Астигматизм вдоль главного луча при $\omega = -70^{\circ}30'$ составляет величину $\Delta t = 1,17$ мм. Как следует из характера изменения меридиональной и сагиттальной кривизны поля исходной оптической системы (рис. 1), коррекционный компонент должен состоять как минимум из двух отрицательных линз.

Первый элемент должен влиять на изменение кривизны таким образом, чтобы меридиональная кривизна перешла в область положительных значений, а сагиттальная кривизна находилась в области небольших отрицательных или положительных значений, т.е. это должен быть элемент второго рода с отношением продольных увеличений, большим единицы. Исходя из этого, первым элементом для обоих

вариантов явилась линза второго рода с выносом входного зрачка $s_p = -105,96$ мм, величиной предмета y = 15,51 мм и отношением продольных увеличений, большим единицы, а $Q_m = 2,2191$, $Q_s = 1,8798$ при $\Delta' t = 3,2738$ мм. Конструктивные параметры имеют следующие значения: $r_1 = -29,50026$ мм; $r_2 = -209,01318$ мм; $d_1 = 3,001$ мм; $n_2 = 1,7849$. Косая толщина линзы $\tilde{d} = 6,0709$ мм.



Рис. 1. Коррекционный компонент: изменение астигматизма по полю после первой линзы (а); изменение астигматизма по полю после второй линзы, т.е. после всей системы (б)



Рис. 2. Коррекционный компонент: изменение астигматизма системы, в которой осуществлялась коррекция кривизны поля (а); изменение астигматизма по полю после второй линзы, т.е. после всей системы (б)

Второй элемент должен сохранить направление изменения сагиттальной кривизны, а изменение меридиональной кривизны должно быть направлено в сторону отрицательных значений, причем этот элемент должен быть первого рода с отношением продольных увеличений, меньшим единицы. По этой причине в качестве второго элемента в обоих случаях была использована корректирующая линза первого рода с прямой ориентировкой ее первой преломляющей поверхности и выносом входного зрачка $s_p = -37,05$ мм, величиной предмета y = -21,292 мм и отношением продольных увеличений, меньшим единицы. Для первого варианта (рис. 1) продольные увеличения составляют $Q_m = 0,8772$, $Q_s = 1,2871$ при $\Delta't = 0,5187$ мм, а конструктивные параметры имеют следующие значения: $r_3 = -54,99945$ мм; $r_4 = \infty$; $d_3 = 2,99$ мм; $n_4 = 1,805998$. Косая толщина линзы $\tilde{d} = 6,2734$ мм.

Для второго варианта (рис. 2) продольные увеличений составляют $Q_m = 0,8091$, $Q_s = 1,3060$ при $\Delta t = 0$, а конструктивные параметры имеют следующие значения: $r_3 = -54,99945$ мм; $r_4 = 735,81714$ мм; $d_3 = 3,069$ мм; $n_4 = 1,805998$. Косая толщина линзы $\tilde{d} = 6,6399$ мм. Воздушный промежуток между линзами в обоих вариантах $d_2 = 5,7$ мм.

Заключение

В основном широкоугольные объективы для получения хорошего качества изображения рассчитаны с использованием геометрического виньетирования широких наклонных пучков, что отрицательно сказывается на светораспределении в пространстве изображений. Предложенный метод синтеза силовых и коррекционных компонентов широкоугольных объективов позволяет получить высокое качество изображения в пределах всего углового поля без использования геометрического виньетирования широких наклонных пучков, что определяет более равномерное светораспределение в пространстве изображений [4]. Выполненный расчет коррекционных компонентов подтвердил правильность выбранных исходных позиций и эффективность метода.

Литература

- 1. Русинов М.М. Техническая оптика. Л.: Машиностроение, 1979. 488 с.
- 2. Безруков В.А. Обобщенное условие получения заранее заданного астигматизма линз // Оптикомеханическая промышленность. – 1986. – № 1. – С. 16–18.
- 3. Безруков В.А., Карпова Г.В. Синтез силовых компонентов широкоугольных объективов // Оптический журнал. – 2012. – № 5. – С. 32–34.
- 4. Зацепина М.Е. Расчет светосильного (1:1,2) киносъемочного объектива с асферическими поверхностями // Научно-технический вестник СПбГУ ИТМО. – 2011. – № 5 (75). – С. 147–148.

Безруков Вячеслав Алексеевич	-	Санкт-Петербургский национальный исследовательский университет
		информационных технологий, механики и оптики, кандидат техниче-
		ских наук, доцент, bezrukov@mail.ifmo.ru
Карпова Галина Васильевна	-	Санкт-Петербургский национальный исследовательский университет
		информационных технологий, механики и оптики, кандидат техниче-
		ских наук, доцент, Karpova3101@mail.ru

УДК 535.31, 681.7.06 СПОСОБ УГЛОВОГО СЕЛЕКТИВНОГО РЕГУЛИРОВАНИЯ НАПРАВЛЕННОГО СВЕТОПРОПУСКАНИЯ Р.С. Закируллин

Представлен способ углового селективного регулирования направленного светопропускания при движении источника света относительно остекленного объекта. Оптический фильтр с решеточными слоями на двух поверхностях стеклянной подложки обеспечивает угловое регулирование за счет относительного расположения решеток. Обе решетки состоят из поглощающих, рассеивающих или отражающих полос, чередующихся с полосами необработанной поверхности стекла фильтра. Приведен графоаналитический метод расчета угловых характеристик светопропускания фильтра плоскопараллельной формы. При заранее известной траектории источника света характеристики светопропускания можно предварительно адаптировать к изменению углов падения лучей. Результаты графоаналитического расчета подтверждены экспериментально. Рассмотрены возможности применения способа регулирования в архитектурном остеклении, оптических и светотехнических устройствах.

Ключевые слова: направленное светопропускание, оптический фильтр, решетка с чередующимися полосами, графоаналитический расчет, угловое селективное регулирование.

Введение

Регулирование направленного светопропускания в зависимости от угла падения лучей имеет практическое значение в случаях, когда источник света и (или) остекленный объект движутся друг относительно друга. При изменении угла падения солнечных лучей на окно проходящее в помещение излучение контролируется для достижения оптимальных условий освещения и инсоляции. Угловая зависимость светопропускания учитывается при производстве оптических устройств разного назначения. При увеличении угла падения коэффициенты отражения и поглощения возрастают в соответствии с формулами Френеля и законом Бугера–Ламберта, и светопропускание самопроизвольно уменьшается. Дополнительное регулирование необходимо для получения оптимального светопропускания в тех или иных угловых диапазонах.

В оконных конструкциях со смарт-стеклами регулирование осуществляется дискретно в двух режимах – при пропускании электрического тока через электрохромные [1], жидкокристаллические [2] слои или слои с мелкодисперсными частицами стекло пропускает направленное излучение, при выключении тока проходящее излучение рассеивается. Для регулирования непосредственно по углам падения лучей применяются жалюзи и другие устройства перераспределения светового потока [3].

Определение направления (угла падения) излучения по максимуму интенсивности используется как способ ориентировки оптической системы [4]. Существует много методов исправления различных видов геометрической и хроматической аберрации в оптических устройствах [5]. В светотехнических и оптических приборах для распределения мощности световых потоков по разным направлениям распространены ступенчатые френелевские линзы [6]. Для коррекции зрения применяются очки с многофокальными линзами и позонным распределением показателя преломления. Солнцезащитные очковые лин-

зы с «градиентным окрашиванием» имеют переменные характеристики светопропускания вследствие постепенного изменения цвета и (или) насыщенности окраски поверхности.

Настоящая работа посвящена дальнейшему развитию способа углового селективного регулирования направленного светопропускания, описанного в материалах заявок на изобретение, поданных в США [7] и в Российской Федерации [8]. Исследуется оптический фильтр нового типа, предназначенный для регулирования светопропускания без дополнительных устройств перераспределения световых потоков. Поставлена задача определения взаимосвязи между оптическими и геометрическими параметрами фильтра и угловыми характеристиками его светопропускания. Задача решается расчетным и экспериментальным путями. Обсуждаются перспективы и пути применения способа в архитектурном остеклении, в оптических и светотехнических устройствах.

Оптический фильтр и метод расчета характеристик светопропускания

Фильтр состоит из плоскопараллельной стеклянной подложки, на обеих поверхностях которой формируются тонкослойные решеточные слои с чередующимися полосами. На поверхность стекла через определенное расстояние наносятся поглощающие, отражающие или рассеивающие полосы, остальные полосы чистой поверхности стекла пропускают излучение направленно. Угловая селективность регулирования обеспечивается за счет взаимного расположения пропускающих полос на входной и выходной поверхностях – при изменении углов падения меняется доля излучения, проходящего через обе решетки фильтра.

На рис. 1 приведена схема для графоаналитического расчета плоскопараллельного решеточного фильтра с чередующимися пропускающими и поглощающими полосами (изображены тонкими и толстыми линиями соответственно). Указаны границы направленного светопропускания через входную решетку фильтра при падении пучков 1 и 2. Пучок 1 падает под характеристическим углом фильтра – углом падения Θ луча, проходящего через центры пропускающей полосы входной решетки и поглощающий полосы выходной решетки. Этот угол определяет сдвиг решеток фильтра друг относительно друга. Для расчета приняты следующие геометрические и оптические параметры: ширины пропускающих и поглощающих полос $t_1 = 3$ мм и $t_2 = 1$ мм на входной решетке, $t_3 = 2,5$ мм и $t_4 = 1,5$ мм на выходной решетке, показатель преломления n = 1,5 и толщина стекла s = 4 мм. Шаги полос (суммарные ширины двух соседних чередующихся полос) на входной и выходной решетках одинаковы: $t_1 + t_2 = t_3 + t_4 = 4$ мм.



Рис. 1. Схема плоскопараллельного решеточного фильтра

Через входную решетку при любом угле падения проходит одна и та же часть падающего излучения, равная отношению ширины пропускающей полосы к величине шага полос. Через выходную решетку излучение проходит в зависимости от угла падения, так как указанные на рисунке границы направленного светопропускания передвигаются относительно чередующихся полос выходной решетки. Таким образом, интенсивность проходящего через весь фильтр излучения имеет угловую зависимость. Графоаналитический расчет устанавливает взаимосвязь между оптическими и геометрическими параметрами фильтра и угловыми характеристиками его светопропускания.

Коэффициент светопропускания т плоскопараллельного фильтра рассчитывается по формуле

$$\tau = \frac{h}{t_1 + t_2},\tag{1}$$

где *h* – ширина светопропускания (общая ширина той части пропускающих полос выходной решетки, через которую преломленные лучи проходят направленно в пределах одного шага полос). На рис. 1 указана ширина h_2 для пучка 2. Ширина светопропускания рассчитывается на основе функции смещения lпреломленного луча [9] – расстояния, на которое при данном угле падения Θ преломленный под углом Θ_n луч смещается на выходной поверхности относительно непреломленного луча при нормальном угле падения (на рис. 1 указаны смещения l_1 и l_2 для лучей 1 и 2):

$$l = \frac{s\sin\Theta}{\sqrt{n^2 - \sin^2\Theta}} \,. \tag{2}$$

При анализе рис. 1 для диапазона углов падения от 0° до 14,22° получена расчетная формула ширины светопропускания $h = 0, 5t_1 - 0, 5t_4 + l_c - l$, где l_c – смещение преломленного луча при характеристическом угле фильтра. Ширина светопропускания уменьшается в соответствии с этой формулой до угла падения 14.22°, при котором самый нижний луч, прошедший направленно через входную поверхность, после преломления точно попадает на нижний край поглощающей полосы выходной поверхности. При этом экстремальном угле выполняется равенство $l = -0.5t_1 + 0.5t_4 + l_c$. Точное значение 14,22° экстремального угла падения определено путем подстановки вычисленного значения смещения в формулу, полученную из формулы (2):

$$\Theta = \arcsin\left(\frac{nl}{\sqrt{s^2 + l^2}}\right).$$

При дальнейшем анализе рис. 1 выделены еще три диапазона с одинаковым характером изменения ширины светопропускания (диапазоны разделены экстремальными углами, для которых указаны равенства для точного определения их значений):

- от 14,22° до 45,55° ($l = 0,5t_1 - 0,5t_4 + l_c$) ширина светопропускания постоянна, $h = t_1 - t_4$;

- от 45,55° до 68,53° $(l = -0,5t_1 + t_3 + 0,5t_4 + l_c)$ ширина светопропускания увеличивается, $h = 0,5t_1 - 0,5t_4 - l_c + l;$

– от 68,53° до 90° ширина светопропускания постоянна, $h = t_3$.

По формуле (1) рассчитаны коэффициенты светопропускания для углов падения от 0° до 90° с подстановкой значений ширины светопропускания, полученных по расчетным формулам для каждого из четырех диапазонов регулирования. По результатам графоаналитического расчета на рис. 2 построена линия зависимости коэффициента светопропускания от угла падения (угловая характеристика светопропускания) для данного фильтра.



Рис. 2. Угловые характеристики светопропускания решеточного фильтра

На рис. 2 представлена также расчетная характеристика светопропускания, скорректированная с учетом угловой зависимости отражения и поглощения для наиболее интересного с практической точки зрения диапазона углов падения от 0° до 60°. Формула для коррекции коэффициента пропускания для фильтра плоскопараллельной формы получена из формул Френеля и закона Бугера-Ламберта:

$$\tau = (1-\rho)^2 \exp\left(-\alpha s \sqrt{1+\sin^2 \Theta/n^2-\sin^2 \Theta}\right),$$

где р – суммарный амплитудный коэффициент отражения с учетом перпендикулярно и параллельно поляризованных составляющих падающей световой волны, α – натуральный коэффициент поглощения стекла.

Эксперименты для фильтра с принятыми оптическими и геометрическими параметрами проведены с использованием расширенного коллимированного пучка лазерного излучения с длиной волны 532 нм. Скорректированная и экспериментальная характеристики фильтра имеют хорошую согласованность – среднее отклонение составляет 2,3%. Приведены также характеристики пропускания дополнительных образцов фильтра с входной и выходной решетками. Эти линии горизонтальны, что подтверждает независимость светопропускания от угла падения при применении только одной решетки. При приближении к 60° светопропускание этих образцов начинает падать из-за увеличения коэффициента отражения.

Графоаналитический расчет позволяет найти необходимые параметры решеток фильтра для получения требуемых в конкретном случае угловых характеристик светопропускания. При регулировании проходящего через окно солнечного излучения данный способ обеспечит оптимальные характеристики при любом азимуте ориентации окна по сторонам света, так как, в отличие от жалюзи с горизонтальными или вертикальными ламелями, чередующиеся полосы решеток фильтра можно располагать под любым углом.

Особенности применения способа регулирования светопропускания

На рис. З изображена линза со сферическими поверхностями с центрами в точках O_1 и O_2 и радиусами R_1 и R_2 , на поверхности которой нанесены чередующиеся поглощающие слои в форме колец (выделены толстыми линиями). Размеры и расположение колец на обеих поверхностях подобраны для обеспечения максимального пропускания параллельно падающих лучей при их соосном с осью линзы падении. При любом другом угле падения (штриховые линии) светопропускание уменьшается из-за попадания части преломленных лучей на поглощающие кольцевые полосы второй (выходной) решетки линзыфильтра. По максимуму светопропускания можно определить угол падения лучей и ориентировать ось оптической системы по направлению к источнику света. Возможно использование криволинейных полос на поверхностях солнцезащитных очковых линз для ограничения попадания на сетчатку глаза лучей вредного, например, ультрафиолетового диапазона при определенных диапазонах углов падения лучей.



Рис. 3. Схема светопропускания сферической линзы с чередующимися кольцами



Рис. 4. Схема прохождения лучей через призму с чередующимися полосами (a); схема отражения лучей под заранее заданным углом (б)



Рис. 5. Схема исправления сферической аберрации

На рис. 4 приведены примеры получения различных светоцветовых эффектов. При падении параллельного пучка белого света на призму с чередующимися полосами за счет дисперсии достигается выборочное пропускание лучей определенного диапазона по длине волны (рис. 4, а). По схеме, представленной на рис. 4, б, обеспечивается видимость рисунка или надписи, нанесенных на зеркальную полосу внутренней поверхности остекленного объекта, под определенным и заранее заданным углом наблюдения. На внешнюю поверхность прикрепляется прозрачная пластина с чередующимися полосами, пропускающая падающий свет только под определенным углом (диапазоном углов). Для различения отраженного изображения от наблюдаемого напрямую пластина окрашивается, т.е. является светофильтром. Возможно применение данной схемы для защиты от подделок.

Для перераспределения направлений лучей, проходящих через остекленный объект, используются толстослойные чередующиеся полосы с разными показателями преломления. На рис. 5 изображена схема исправления сферической аберрации линзы с кольцевыми чередующимися полосами и уменьшающимися к краям линзы показателями преломления ($n_3 < n_2$). Показатель преломления стекла линзы n_1 ($n_2 < n_1$). На верхней половине разреза линзы показано прохождение лучей 1, 2 и 3 без участия полос. С удалением падающих лучей от оси аберрация усиливается – точка F_3 более отдалена от точки F_2 фокусирования параксиальных лучей 1 по сравнению с точкой F_2 (луч 0 проходит по оси линзы). Показатели преломления n_2 и n_3 и толщины полос подбираются так, чтобы лучи 2 и 3 также собирались в точке F_1 фокусировки параксиальных лучей (показано на нижней половине линзы). Подобную схему можно применить в очках при коррекции зрения для позонного распределения показателя преломления (вместо многофокальных линз), в светораспределителях осветительных приборов для изменения направлений световых потоков (вместо френелевских линз).

Рассмотренный способ обеспечивает селективное регулирование интенсивности и направления проходящего излучения в зависимости от углов падения с предварительной адаптацией к их изменению при известной траектории источника света или его расположении относительно остекленного объекта. Возможность ступенчатого распределения оптических и геометрических параметров чередующихся полос решеток позволяет достичь позонного регулирования светопропускания, в том числе с изменением спектра излучения.

Заключение

Графоаналитическим методом рассчитана зависимость коэффициента светопропускания от угла падения для плоскопараллельного фильтра. Установлена связь между оптическими и геометрическими параметрами чередующихся полос решеток фильтра и угловыми характеристиками его светопропускания. Расчетные и экспериментальные данные в диапазоне углов падения от 0° до 60° отличаются в среднем на 2,3%.

Рассмотрены особенности применения способа регулирования в архитектурном остеклении, в оптических и светотехнических устройствах. Дальнейшие исследования направлены на разработку технологии для обеспечения углового регулирования светопропускания окон без дополнительных средств перераспределения световых потоков.

Литература

- Andersson A.M., Granqvist C.G., Stevens J.R. Electrochromic Li_xWO₃/poymer laminate/Li_yV₂O₅ device: toward an all-solid-state smart window // Appl. Opt. – 1989. – V. 28. – P. 3295–3302.
- Sueda K., Tsubakimoto K., Miyanaga N., Nakatsuka M. Control of spatial polarization by use of a liquid crystal with an optically treated alignment layer and its application to beam apodization // Appl. Opt. – 2005. – V. 44. – P. 3752–3758.

- Andersen M., Rubin M., Powles R., Scartezzini J.-L. Bi-directional transmission properties of Venetian blinds: experimental assessment compared to ray-tracing calculations // Solar Energy. – 2005. – V. 78. – № 2. – P. 187–198.
- Антонов И.Н., Закируллин Р.С., Малков А.И., Пожар М.С., Руссов В.М., Чакак А.А. Устройство для измерения распределения плотности энергии лазерного излучения. Авт. свид. СССР, кл. G01J 5/02.1988.
- 5. Гуриков В.А. Эрнст Аббе. М.: Наука, 1985. 228 с.
- 6. Трембач В.В. Световые приборы: Учеб. для вузов по спец. «Светотехника и источники света». М.: Высшая школа, 1990. 463 с.
- Zakirullin R.S. Expedient of regulation of the directional gear transmission of light. Заявка США, кл. G02B 5/22.2011 [Электронный ресурс]. – Режим доступа: http://www.wipo.int/patentscope/search/en/search.jsf, свободный. Яз. англ. (дата обращения 04.03.2013).
- Закируллин Р.С. Способ регулирования направленного светопропускания. Заявка РФ, кл. G02B 5/20.2012 [Электронный ресурс]. – Режим доступа: http://www1.fips.ru/fips_servl/fips_servlet, свободный. Яз. рус. (дата обращения 04.03.2013).
- 9. Закируллин Р.С. Селективное регулирование направленного светопропускания по углам падения лучей // ЖТФ. 2012. Т. 82. № 10. С. 134–136.

Закируллин Рустам Сабирович – Оренбургский государственный университет, кандидат технических наук, доцент, rustam.zakirullin@gmail.com

УДК 535.338.1+519.642.3+519.6 ВОССТАНОВЛЕНИЕ НЕПРЕРЫВНЫХ СПЕКТРОВ АДАПТИВНЫМ СПОСОБОМ ВЫЧИСЛИТЕЛЬНЫХ ЭКСПЕРИМЕНТОВ С РЕГУЛЯРИЗАЦИЕЙ В.С. Сизиков, А.В. Кривых

Рассмотрена обратная задача спектроскопии – восстановление непрерывных спектров путем математической обработки измеренных спектров, искаженных аппаратной функцией спектрометра и помехами. Задача сводится к решению интегрального уравнения Фредгольма I рода. Задача его решения некорректна, поэтому для получения устойчивого решения используется метод регуляризации Тихонова. При этом применен адаптивный способ вычислительных экспериментов, согласно которому, наряду с исходным спектром P, обрабатывается модельный спектр Q с задаваемым истинным спектром z и моделируемым измеренным спектром u с учетом дополнительной (априорной) информации об истинном спектре P. Это позволяет выбрать параметр регуляризации α . Предложенная методика может быть использована для повышения разрешающей способности спектрометра. Приведены численные иллюстрации.

Ключевые слова: непрерывный спектр, обратная задача спектроскопии, интегральное уравнение, метод регуляризации Тихонова, способ вычислительных экспериментов, повышение разрешающей способности спектрометра.

Введение

Измеренный спектрометром (например, интерферометром Фабри–Перо) спектр $u(\lambda)$ (где λ – длина волны) обычно отличается от истинного спектра $z(\lambda)$ [1–8]. Это проявляется, во-первых, в большей сглаженности спектра $u(\lambda)$ по сравнению с $z(\lambda)$, а именно, в спектре $u(\lambda)$ не разрешены близкие линии, сглажена тонкая структура спектральной линии, что является результатом воздействия аппаратной функции спектрального прибора [1–9]. Во-вторых, это проявляется в зашумленности спектра $u(\lambda)$, а именно, слабые линии «тонут» в шуме, что является результатом погрешностей измерений [1–3], а также воздействия среды, через которую проходит излучение [10].

Дадим следующее определение аппаратной функции (A Φ) [3, 6–8] (ср. [9, С. 32, 704]): аппаратной функцией $K(\lambda, \lambda')$ спектрометра называется его реакция (в виде измеренной интенсивности) на дискретную линию единичной интенсивности и длины волны λ' при настройке спектрометра на длину волны λ .

Форма аппаратной функции (ширина и т.д.) может заметно меняться с изменением длины волны настройки $\lambda \in [\lambda_{\min}, \lambda_{\max}]$, где $[\lambda_{\min}, \lambda_{\max}]$ – диапазон длин волн изучаемой части спектра. Обычно с увеличением λ АФ становится шире, что характерно для широкополосной спектрометрии, например, изучения спектра звезды во всем видимом диапазоне. Если же АФ практически не изменяется при изменении λ , то АФ является разностной (инвариантной): $K(\lambda, \lambda') = K(\lambda - \lambda')$, что имеет место, например, при изучении тонкой структуры отдельной линии [3, 6, 8], когда диапазон [$\lambda_{\min}, \lambda_{\max}$] мал.

На рис. 1 в качестве примера приведен смоделированный непрерывный измеренный спектр $u(\lambda)$, сглаженный аппаратной функцией спектрометра $K(\lambda, \lambda')$, а также зашумленный (и дискретизированный) измеренный спектр $\tilde{u}(\lambda) = u(\lambda) + \delta u$ (где $\delta u - шум$) и АФ спектрометра, причем, поскольку в дан-

ном примере $K(\lambda, \lambda') - \phi$ ункция неразностная, то приведено два ее «сечения» (подробности примера см. дальше). В принципе похожий вид может иметь непрерывный узкополосный спектр [6, С. 200], например, сверхтонкая структура отдельной линии, обусловленная магнитными или электрическими полями (эффект Зеемана или Штарка), а также тепловым уширением (эффект Доплера) [10], однако в этом случае диапазон [$\lambda_{\min}, \lambda_{\max}$] мал, а АФ – разностная: $K(\lambda, \lambda') = K(\lambda - \lambda')$.



Рис. 1. *u*(λ) – измеренный без шума спектр; $\tilde{u}(\lambda)$ – измеренный зашумленный и дискретизированный спектр; $K(\lambda^*, \lambda')$ – AΦ при некоторой длине волны настройки $\lambda = \lambda^*$; [$\lambda_{\min}, \lambda_{\max}$] – широкий диапазон длин волн

Как будет видно далее, в примере на рис. 1 в измеренном спектре $u(\lambda)$ (тем более, в зашумленном спектре $\tilde{u}(\lambda)$) не разрешены близкие линии и не выявлены слабые, причем этот эффект тем сильнее, чем шире АФ $K(\lambda, \lambda')$ (а также чем выше уровень шумов), другими словами, чем меньше разрешающая способность спектрометра [1, 9].

В данной работе ставится известная обратная задача спектроскопии – задача восстановления истинного спектра $z(\lambda)$ по измеренному спектру $\tilde{u}(\lambda)$ и аппаратной функции $K(\lambda,\lambda')$ [1–8, 11–15]. Данная задача описывается интегральным уравнением (см. дальше), задача решения которого некорректна, поэтому его обычно решают методом регуляризации Тихонова. При этом важным является вопрос о выборе параметра регуляризации α . В данной работе предлагается новый адаптивный способ (вычислительных экспериментов) для выбора параметра α .

Математическая формулировка обратной задачи спектроскопии

Рассмотрим случай непрерывного спектра, обычно характерного для веществ с повышенной плотностью (расплавленный жидкий металл, плазма и т.д.). Измеренная интенсивность $u(\lambda)$ при настройке спектрометра на длину волны λ равна сумме (интегралу) по всем истинным интенсивностям $z(\lambda)$ с весовой функцией *K*:

$$u(\lambda) = \int_{a}^{b} z(\lambda') K(\lambda, \lambda') d\lambda',$$

где $a = \lambda_{\min}$, $b = \lambda_{\max}$, откуда, варьируя значение λ (т.е. выполняя сканирование по спектру) и учитывая зашумленность спектра $u(\lambda)$, получим:

$$\int_{a}^{b} K(\lambda,\lambda') z(\lambda') d\lambda' = \tilde{u}(\lambda) , \quad c \le \lambda \le d ,$$
(1)

где [c,d] – пределы изменения λ (обычно более широкие, чем [a,b]).

В соотношении (1) известны (измерены или заданы) $\tilde{u}(\lambda)$, $K(\lambda, \lambda')$, $a, b, c, d, a z(\lambda')$ является искомым истинным спектром. Соотношение (1) есть интегральное уравнение Фредгольма I рода, причем $K(\lambda,\lambda')$ является ядром уравнения, $\tilde{u}(\lambda)$ – правой частью, а $z(\lambda')$ – искомой функцией. Если $K(\lambda,\lambda') = K(\lambda - \lambda')$, то

$$\int_{0}^{\infty} K(\lambda - \lambda') z(\lambda') d\lambda' = \tilde{u}(\lambda), \quad 0 \le \lambda < \infty.$$
⁽²⁾

Соотношение (2) есть интегральное уравнение Фредгольма I рода типа свертки на полуоси. Решение уравнения (1) или (2) дает возможность, в принципе, восстановить истинный спектр $z(\lambda)$. Однако задача решения уравнений (1) и (2) является некорректной (существенно неустойчивой) [2–4, 6, 8, 16]: если решать уравнение (1), например, методом квадратур, а уравнение (2) – методом преобразования Фурье (инверсной фильтрации), то в качестве решения получим так называемую «пилу» [3, 6, 8] – крайне неустойчивое решение. По этой причине для устойчивого решения этих уравнений необходимо применение таких методов, как регуляризация Тихонова [2–4, 6–8, 11–16], параметрическая фильтрация Винера [3, 6, 8, 16], итеративная регуляризация Фридмана [6, 8, 16] и др.

При обработке спектра в широком диапазоне длин волн следует учитывать изменение формы $A\Phi K(\lambda,\lambda')$ с изменением длины волны настройки λ . При обработке же спектра в узкой полосе следует использовать уравнение Фредгольма I рода с разностным ядром (ср. (2)):

$$\int_{a}^{b} K(\lambda - \lambda') z(\lambda') d\lambda' = \tilde{u}(\lambda), \quad c \le \lambda \le d.$$
(3)

Задача решения уравнений (1)–(3) связана с задачей редукции к идеальному спектральному прибору [1–4, 9, 17] – с одним из вариантов редукционной проблемы Рэлея [3, 6, 8, 13]. Успешное решение задачи редукции позволит путем математической обработки результатов измерений повысить разрешающую способность спектрального прибора. В настоящей статье воспользуемся методом регуляризации Тихонова. Что касается других устойчивых методов (фильтрации Винера, итеративных методов и др.), то они изложены в различных публикациях ([3, 6, 8, 16] и др.) и также могут быть применены для устойчивого восстановления спектров.

Краткая формулировка метода регуляризации Тихонова

Запишем уравнение (1) в виде

$$Az \equiv \int_{a}^{b} K(\lambda,\lambda') z(\lambda') d\lambda' = \tilde{u}(\lambda), \quad c \le \lambda \le d , \qquad (4)$$

где *А* – оператор, соответствующий ядру *К*. Метод регуляризации Тихонова сводится к решению интегрального уравнения Фредгольма II рода

$$\alpha z_{\alpha}(t) + \int_{a}^{b} B(t,\lambda') z_{\alpha}(\lambda') d\lambda' = U(t), \quad a \le t \le b,$$
(5)

где $\alpha > 0$ – параметр регуляризации, а новое ядро и новая правая часть равны

$$B(t,\lambda') = B(\lambda',t) = \int_{c}^{d} K(\lambda,t) K(\lambda,\lambda') d\lambda, \qquad U(t) = \int_{c}^{d} K(\lambda,t) \tilde{u}(\lambda) d\lambda$$

В таком варианте уравнение (5) обычно решается методом квадратур [3, 4, 6, 8, 16]. Если же рассматривать уравнение (2) или (3), то его решение методом регуляризации Тихонова будет включать преобразование Фурье и α -регуляризацию (подробности см. в [3, 4, 6, 8, 11–16, 18–21]).

При этом важным является вопрос о выборе параметра регуляризации α и об учете дополнительной (априорной) информации относительно искомого спектра $z(\lambda)$. Существует ряд способов выбора параметра регуляризации α : способ невязки, обобщенный принцип невязки, метод перекрестной значимости, локальный регуляризующий алгоритм, способ подбора и др. [3, 4, 6, 8, 13, 16, 18–21].

Способ вычислительных экспериментов

В данной работе получает дальнейшее развитие способ вычислительных экспериментов для выбора параметра регуляризации α (другие его названия – способ псевдообратного оператора, способ эталонных, или модельных примеров, способ моделирования) [3, 6, 8, 16, 22, 23]. Данный способ учитывает дополнительную (априорную) информацию об искомом спектре (оценку количества спектральных линий, их параметров и т.д.) и поэтому является интерактивным и адаптивным способом.

Кратко изложим способ вычислительных экспериментов.

Рассмотрим операторное уравнение I рода: Az = u (ср. (4)). Полагаем, что вместо точных u и K известны \tilde{u} и \tilde{K} такие, что $\|\tilde{u}-u\| \le \delta$, $\|\tilde{A}-A\| \le \xi$, где δ и ξ – верхние оценки погрешностей по норме правой части u и ядра K. При использовании метода регуляризации Тихонова решается операторное

уравнение $\alpha z_{\alpha} + \tilde{A}^T \tilde{A} z_{\alpha} = \tilde{A}^T \tilde{u}$ (ср. (5)), где A^T – транспонированный оператор. Обозначим $\Delta z_{\alpha} = z_{\alpha} - z$ – погрешность регуляризованного решения z_{α} , а z – точное решение (нормальное псевдорешение [16, 20, 21]). В работах [16, 22, 23] получена следующая оценка относительной погрешности регуляризованного решения по норме:

$$\frac{|\Delta z_{\alpha}||}{||z||} \leq \varepsilon(\alpha), \qquad (6)$$

где

$$\varepsilon(\alpha) = \frac{\left\| \tilde{A} \right\| \eta}{2\sqrt{\alpha}} + \frac{p\alpha}{p\alpha + 1}.$$
(7)

Здесь $\eta = \delta_{\text{отн}} + \xi_{\text{отн}}$, причем $\delta_{\text{отн}} = \delta / ||u||$ и $\xi_{\text{отн}} = \xi / ||A||$ – относительные погрешности исходных данных; $p = ||\tilde{A}^+||^2$, A^+ – псевдообратный оператор: $A^+u = z$ [20, С. 184]. Функция $\varepsilon(\alpha)$ является верхней огибающей для истинной относительной погрешности

$$\sigma_{\rm oth}(\alpha) = \frac{\|\Delta z_{\alpha}\|}{\|z\|}.$$
(8)

В работах [16, 22] показано, что функция $\varepsilon(\alpha)$ имеет (единственный) минимум при условии $p \cdot (||\tilde{A}^+||\eta)^2 < 27/16 \approx 1,69$ или $||\tilde{A}^+||\cdot||\tilde{A}||\eta < 3\sqrt{3}/4 \approx 1,30$. Согласно соотношениям (6)–(8), оценка относительной погрешности $||\Delta z_{\alpha}||/||z||$ регуляризованного решения z_{α} зависит от \tilde{A} и η (точнее, от произведения $||\tilde{A}||\eta$). По этой причине, если решается несколько задач (другими словами, обрабатывается несколько спектров) с одинаковыми \tilde{A} и η , то для них оценки погрешности

$$\sigma_{\text{OTH}}(\alpha) = \frac{\|\Delta z_{\alpha}\|}{\|z\|} \leq \frac{\|A\| \eta}{2\sqrt{\alpha}} + \frac{p\alpha}{p\alpha + 1}$$

будут одинаковыми.

Отсюда следует, что при решении некоторого исходного примера P (т.е. при обработке исходного спектра \tilde{u}_p) с неизвестным решением (спектром) z_p можно использовать результаты решения другого, модельного, примера Q с известным (заданным) точным решением (спектром) z_Q , причем с такими же \tilde{A} и η , что и в примере P. При этом при решении примера Q можно рассчитать функцию $\sigma_{\text{отн}}(\alpha)_Q = ||\Delta z_{\alpha Q}||/||z_Q||$ и по ней найти $\alpha_{\text{опт}Q}$ – оптимальное значение α , при котором $\sigma_{\text{отн}}(\alpha)_Q = \min_{\alpha}$. Это значение $\alpha_{\text{опт}Q}$ может быть использовано при решении исходного примера (спектра) P. При этом необходимо также определить $p = ||\tilde{A}^+||^2$. Оценка p может быть получена путем подбора такого значения p, при котором огибающая кривая $\varepsilon(\alpha)$ касается набора кривых $\sigma_{\text{отн}}(\alpha)_Q$ (см. рис. 3).

Добавим, что для повышения эффективности изложенного способа модельный пример Q (или несколько примеров) должен содержать дополнительную информацию об исходном примере (спектре) P, а именно, оценку количества спектральных линий (максимумов) в искомом спектре z_P , соотношений их интенсивностей и значений их длин волн. Данную оценку должен делать опытный спектроскопист. Использование такой информации в модельном примере Q позволит более удачно выбрать параметр α . Данный способ следует считать адаптивным и интерактивным способом.

Численная иллюстрация

В рамках системы программирования MATLAB 7 был разработан пакет программ для восстановления истинных непрерывных спектров $z(\lambda)$ путем численного решения интегрального уравнения Фредгольма I рода методом регуляризации Тихонова с использованием способа вычислительных экспериментов.

Сначала был рассмотрен первый пример (рис. 1) – оригинал *P*, у которого известен зашумленный измеренный спектр $\tilde{u}(\lambda)$ на равномерной сетке узлов $\lambda = \lambda_{\min}, \lambda_{\min} + h, ..., \lambda_{\max}$, где $\lambda_{\min} = 450$ нм; $\lambda_{\max} = 650$ нм; $h = \Delta \lambda = \text{const} = 1$ нм – шаг дискретизации; $n = (\lambda_{\max} - \lambda_{\min})/h = 200$ – число шагов дискретизации по λ . Известна также аппаратная функция – дифракционная АФ Рэлея (ср. [1, 4, 5]) вида

$$K(\lambda,\lambda') = \frac{1}{\gamma(\lambda)} \operatorname{sinc}^2\left(\frac{\lambda-\lambda'}{\gamma(\lambda)}\right) = \frac{1}{\gamma(\lambda)} \left\{\frac{\sin\left[\pi(\lambda-\lambda')/\gamma(\lambda)\right]}{\pi(\lambda-\lambda')/\gamma(\lambda)}\right\}^2,\tag{9}$$

где $\gamma(\lambda)$ – полуширина АФ по уровню 0, равная приблизительно ширине АФ по уровню 0,5, которую мы положили равной $\gamma(\lambda) = 8\lambda/\lambda_{min}$ нм. Спектр полагается широкополосным (от фиолетового до красного), поэтому ширина АФ непостоянна, а именно, $\gamma(\lambda_{min}) = 8$ нм, а $\gamma(\lambda_{max}) = 11,55$ нм, т.е. $\gamma(\lambda_{max})/\gamma(\lambda_{min}) = 1,44$. При этом истинный спектр $z(\lambda)$ в примере *P* неизвестен.

Из рис. 1 видно, что измеренный спектр $u(\lambda)$ имеет довольно сложную структуру, а именно, содержит шесть явных флуктуаций, две из которых (при $\lambda \approx 525$ нм и $\lambda \approx 620$ нм), скорее всего, состоят каждая из двух линий, но они не разрешились из-за того, что АФ имеет немалую ширину и, тем самым, ограничивает разрешающую способность спектрометра. Кроме того, есть намек на то, что при $\lambda \approx 507$ нм и $\lambda \approx 543$ нм имеются еще две слабые линии. Таким образом, все указывает на то, что на самом деле в спектре имеются не менее восьми спектральных линий. В связи с этим в качестве второго (модельного или эталонного) примера Q был составлен близкий к оригиналу P пример, истинный спектр $z_0(\lambda)$ которого состоит из 9 спектральных линий в виде гауссиан:

$$\begin{aligned} z_{\varrho}(\lambda) &= 2,0 \exp\left\{-\left[(\lambda - 486)/10\right]^{2}\right\} + 0,4 \exp\left\{-\left[(\lambda - 512)/5\right]^{2}\right\} + \\ &+ 8,5 \exp\left\{-\left[(\lambda - 522)/2\right]^{2}\right\} + 9,2 \exp\left\{-\left[(\lambda - 530)/2\right]^{2}\right\} + \\ &+ 0,5 \exp\left\{-\left[(\lambda - 542)/5\right]^{2}\right\} + 8,2 \exp\left\{-\left[(\lambda - 566)/6\right]^{2}\right\} + \\ &+ 2,5 \exp\left\{-\left[(\lambda - 592)/4\right]^{2}\right\} + 4,5 \exp\left\{-\left[(\lambda - 614)/7\right]^{2}\right\} + \\ &+ 3,0 \exp\left\{-\left[(\lambda - 626)/5\right]^{2}\right\}. \end{aligned}$$

Измеренный спектр $u_Q(\lambda)$ в примере Q был рассчитан согласно выражению

 $u_{\mathcal{Q}}(\lambda) = \int_{a}^{b} K(\lambda, \lambda') z_{\mathcal{Q}}(\lambda') d\lambda', \quad c \leq \lambda \leq d ,$

численно. При этом a = 460, b = 640, c = 450, d = 650 нм.

Погрешности измеренной $u_p(\lambda)$ были оценены примерно в 1%, что соответствует среднеквадратическому отклонению СКО $\approx 0,02$. По этой причине к значениям $u_Q(\lambda)$ были добавлены случайные нормальные погрешности с СКО от 0,01 до 0,025, что соответствует $\delta_{\text{отн}} \approx 0,5 - 1,25\%$, поскольку значение $\delta_{\text{отн}}$ в исходном примере *P* известно неточно. АФ спектрометра в примере *Q* была взята в виде (9), причем (поскольку АФ известна также неточно) $\gamma(\lambda)$ было взято равным $\gamma(\lambda) = (8 + \xi)\lambda/\lambda_{\min}$, где $\xi = 0 - 0,3$, что соответствует $\xi_{\text{отн}} \approx 0 - 3\%$.

Далее модельный пример Q был решен методом квадратур с регуляризацией Тихонова с помощью разработанной m-функции Tikh.m [6, C. 207] для ряда значений параметра регуляризации α , и была построена зависимость относительной погрешности регуляризованного решения $z_{\alpha}(\lambda)$ по отношению к точному решению $z(\lambda)$ (см. (8)):

$$\sigma_{\text{oth}}(\alpha) = \frac{\left\| z_{\alpha}(\lambda) - z(\lambda) \right\|}{\left\| z(\lambda) \right\|}.$$

На рис. 2 представлены зависимости $\sigma_{\text{отн}}(\alpha)$ для ряда погрешностей $\delta_{\text{отн}}$ и $\xi_{\text{отн}}$. На рис. 2 представлена также огибающая $\varepsilon(\alpha)$ (см. (7)), при построении которой было положено $\eta = 10^{-2}$ и $\|\tilde{A}\| = \| a \|_{L_2} / \| z \|_{L_2} = 0.82$. Для ряда значений p от 100 до 270 были рассчитаны кривые $\varepsilon(\alpha)$ (рис. 2). Было выбрано то значение p, при котором одна из кривых касается набора кривых $\sigma_{\text{отн}}(\alpha)$, а именно, p = 100. Этому соответствует значение параметра регуляризации $\alpha = 10^{-3}$. Из рис. 2 видно, что, несмотря на разброс кривых $\sigma_{\text{отн}}(\alpha)$ и $\varepsilon(\alpha)$, значения p и, как следствие, α определяются уверенно.

При значении $\alpha = 10^{-3}$, выбранном с помощью решения модельного примера Q как вспомогательного восстановлен спектр в исходном примере P (рис. 3). Как видно из рис. 3, в примере P разрешились близкие линии и восстановились слабые линии, правда, на краях спектра проявился эффект Гиббса, однако в слабой форме (на уровне погрешностей метода). Аналогичные результаты получены для других, весьма различных, непрерывных спектров [3, 6–8, 14, 22, 23], т.е. изложенная в работе методика вычислительных экспериментов может быть использована для широкого класса спектров (с близкими линиями, со слабыми линиями, узкими и широкими линиями и т.д.).

2

460

470

480

490

500

510

520

530

540



Рис. 3. 1 – истинный спектр $z_p(\lambda)$; 2 – измеренный спектр $\tilde{u}(\lambda)$; 3 – восстановленный спектр $z_{\alpha P}(\lambda)$ Заключение

550

560

570

580

590

600

610

620

λ, λ', нм

630

640

Практическое использование изложенной методики позволит повысить разрешающую способность спектрометра. Спектральный прибор может быть соединен с компьютером или со спецпроцессором с заложенным в него математическим и программным обеспечением, реализующим методы и численные алгоритмы решения обратной задачи спектроскопии. В результате такого комплексирования (соединения прибора с компьютером) можно разрешить близкие и выделить слабые линии спектров излучения (или поглощения), а именно, в физике – спектров газов, жидкостей, металлов, плазмы; в астрофизике – спектров звезд, планет, галактик, туманностей, комет; в металлургии – спектров расплавленных металлов в домнах; в геофизике – спектров залежей руд, минералов, нефти, газа и т.д.

Работа выполнена при поддержке РФФИ (грант № 13-08-00442).

Литература

- 1. Раутиан С.Г. Реальные спектральные приборы // Успехи физических наук. 1958. Т. 66. Вып. 3. С. 475–517.
- 2. Кочиков И.В., Курамшина Г.М., Пентин Ю.А., Ягола А.Г. Обратные задачи колебательной спектроскопии. – М.: Изд-во МГУ, 1993. – 204 с.
- 3. Сизиков В.С. Математические методы обработки результатов измерений. СПб: Политехника, 2001. 240 с.
- 4. Старков В.Н. Конструктивные методы вычислительной физики в задачах интерпретации. Киев: Наукова думка, 2002. 264 с.
- 5. Fleckl T., Jäger H., Obernberger I. Experimental verification of gas spectra calculated for high temperatures using the HITRAN/HITEMP database // J. Phys. D: Appl. Phys. 2002. V. 35. P. 3138-3144.
- 6. Сизиков В.С. Обратные прикладные задачи и MatLab. СПб: Лань, 2011. 256 с.
- 7. Сизиков В.С., Кривых А.В. Применение способа эталонных примеров при решении обратной задачи спектроскопии методом регуляризации // Изв. вузов. Приборостроение. 2011. Т. 54. № 9. С. 44–51.
- 8. Сизиков В.С. Интегральные уравнения и MatLab в задачах томографии, иконики и спектроскопии. Saarbrücken: LAP, 2011. 252 с.
- 9. Физический энциклопедический словарь / Гл. ред. А.М. Прохоров. М.: Сов. энциклопедия, 1984. 944 с.
- 10. Ландсберг Г.С. Оптика: Учебное пособие для вузов. 6-е изд. М.: ФИЗМАТЛИТ, 2003. 848 с.
- 11. Преображенский Н.Г., Седельников А.И. Оптимизация спектроскопических измерений на основе методов регуляризации // Журнал прикладной спектроскопии. 1981. Т. 35. Вып. 4. С. 592–599.
- 12. Брагинская Т.Г., Клюбин В.В. Решение обратной задачи спектроскопии оптического смещения методом регуляризации Тихонова. Препринт № 855. – Л.: ЛИЯФ, 1983. – 60 с.
- Глазов М.В., Болохова Т.А. Решение редукционной проблемы Рэлея с использованием различных модификаций метода регуляризации // Оптика и спектроскопия. – 1989. – Т. 67. – Вып. 3. – С. 533– 537.
- 14. Кривых А.В., Сизиков В.С. Комплексированное восстановление непрерывных спектров с использованием псевдообратной матрицы // XLI Неделя науки СПбГПУ: материалы научно-практической конференции с международным участием. Ч. XIII. – СПб: Изд-во Политехн. ун-та, 2012. – С. 240–242.
- 15. Кривых А.В., Сизиков В.С. Обработка дискретных спектров с помощью алгоритма интегральной аппроксимации // Научно-технический вестник СПбГУ ИТМО. 2011. № 5 (75). С. 14–18.
- 16. Верлань А.Ф., Сизиков В.С. Интегральные уравнения: методы, алгоритмы, программы. Киев: Наукова думка, 1986. 544 с.
- Краулиня Э.К., Лиепа С.Я., Пикалов В.В., Скудра А.Я. К проблеме исследования атомной сенсибилизированной флуоресценции по контурам спектральных линий // Некорректные обратные задачи атомной физики / Сб. статей под ред. Н.Г. Преображенского. – Новосибирск: ИТПМ, 1976. – 133 с.
- 18. Воскобойников Ю.Е., Преображенский Н.Г., Седельников А.И. Математическая обработка эксперимента в молекулярной газодинамике. – Новосибирск: Наука, 1984. – 240 с.
- 19. Воскобойников Ю.Е., Мухина И.Н. Локальный регуляризирующий алгоритм восстановления контрастных сигналов и изображений // Автометрия. – 2000. – № 3. – С 45–53.
- 20. Морозов В.А. Регулярные методы решения некорректно поставленных задач. М.: Наука, 1987. 240 с.
- 21. Тихонов А.Н., Гончарский А.В., Степанов В.В., Ягола А.Г. Численные методы решения некорректных задач. М.: Наука, 1990. 232 с.
- 22. Сизиков В.С. Обобщенный метод редукции измерений // Электронное моделирование. 1991. Т. 13. № 4. С. 7–14.
- 23. Верлань А.Ф., Сизиков В.С., Мосенцова Л.В. Метод вычислительных экспериментов для решения интегральных уравнений в обратной задачи спектроскопии // Электронное моделирование. 2011. Т. 33. № 2. С. 3–12.

Сизиков Валерий Сергеевич	-	Санкт-Петербургский национальный исследовательский университет	
		информационных технологий, механики и оптики, доктор технических	
		наук, профессор, sizikov2000@mail.ru	
Кривых Александр Владимирович	зых Александр Владимирович – Санкт-Петербургский национальный исследовательский универ		
		информационных технологий, механики и оптики, аспирант,	
		krivykh1987@mail.ru	

УДК 621.391 МЕТОД РЕШЕНИЯ ЗАДАЧИ ДИНАМИЧЕСКОГО ВЫБОРА МАРШРУТОВ И НАЗНАЧЕНИЯ ДЛИН ВОЛН В СЕТЯХ WDM С УЧЕТОМ ЯВЛЕНИЯ ЧЕТЫРЕХВОЛНОВОГО СМЕШИВАНИЯ

Д.В. Агеев, А.А. Переверзев

Решается задача динамического выбора маршрутов и назначения длин волн в сетях WDM. Задача является актуальной для управления сетью WDM при обслуживании запросов на установку кратковременных соединений и передачу избыточного трафика. Предлагаемый метод решения задачи является эвристическим и усовершенствует ранее известный метод за счет учета влияния четырехволнового смешивания и использования новой метрики на этапе поиска потенциальных для использования маршрутов. Это позволило уменьшить вероятность блокировки вызовов при установке соединений в среднем на 13% и значение *Q*-фактора на 0,812.

Ключевые слова: маршрут, RWA, световой путь, проектирование, четырехволновое смешивание, вероятность блокировки.

Введение

Постоянно растущая потребность в увеличении пропускной способности современных телекоммуникационных систем при развертывании транспортных сетей приводит к необходимости использования технологии спектрального уплотнения каналов (Wavelength-division multiplexing, WDM). Передача информации в WDM-сетях производится вдоль световых путей. Под световым путем будем понимать последовательность оптических каналов, которые используются для передачи потока от источника к получателю с использованием одной и той же длины волны. Указанное ограничение на использование одной длины волны вдоль всего пути может привести к уменьшению количества установленных соединений и неэффективному использованию имеющихся длин волн. Таким образом, задача выбора маршрутов и назначения длин волн в WDM-сетях является важной [1].

Рассмотрим задачу: необходимо выбрать маршрут между парой источник-получатель и назначить ему длину волны таким образом, чтобы обеспечить максимум количества устанавливаемых соединений. Эта задача известна как маршрутизация и распределение длин волн (Routing and Wavelength Assignment, RWA) [2–4]. В зависимости от периодичности решения задачи RWA можно выделить статическое и динамическое распределение.

Задача статического распределения решается в основном на этапе проектирования телекоммуникационной сети, а наиболее часто используемыми критериями оптимальности являются максимум количества установленных соединений или минимум вероятности блокировки.

Задача динамического распределения возникает при управлении потоками в телекоммуникационной сети. Важным фактором при выборе метода решения является его вычислительная сложность и близость получаемого результата к оптимальному.

При построении WDM-сетей возникает нелинейное явление четырехволнового смешивания (ЧВС), что повышает уровень битовых ошибок (Bit Error Rate, BER). Таким образом, при выборе маршрутов для установки соединения необходимо учитывать явление ЧВС.

Одним из перспективных направлений развития современных сетей доступа является совместное использование беспроводных и оптических технологий. Преимущество таких сетей заключается в возможности использования лучших черт обеих технологий. Для беспроводных технологий – это мобильность, высокая скорость развертывания. Оптические технологии позволяют обойти такие недостатки беспроводных сетей, как ограниченность частотного ресурса, проблемы электромагнитной совместимости и другие явления, связанные с особенностями распространения радиоволн.

Использование оптических сетей в качестве транспортной среды позволяет обеспечить передачу радиосигналов между базовыми станциями непосредственно на высокой частоте. Данная технология известна в международной литературе как RoF (Radio over Fiber) [5, 6]. При больших размерностях сетей доступа на основе технологии RoF (более сотни базовых станций) возникает явление всплесков трафика на небольшом промежутке времени. Это связано с увеличением количества ресурсов, запрашиваемых мобильными абонентами в некоторых сотах. В связи с этим возникает необходимость в использовании новых или перераспределении уже выделенных частотных ресурсов. Эта задача получила название MSCA (maximal service channel allocation) [7]. При ее решении производится распределение имеющихся свободных ресурсов сети WDM на непродолжительный промежуток времени, что требует решения задачи динамического выбора маршрутов и назначения длин волн (Dynamic Routing and Wavelength Assignment, DRWA).

Для решения задачи DRWA используются эвристические методы решения, которые базируются на фиксированной, фиксированной альтернативной и адаптивной маршрутизациях.

При использовании фиксированной маршрутизации для установки соединения между заданной парой источник–получатель используется заранее определенный маршрут. Как правило, этот маршрут является кратчайшим путем. Достоинством такого подхода является отсутствие потребности в сборе информации о состоянии сети. К его недостаткам можно отнести то, что при невозможности установки соединения по фиксированному маршруту запрос на установку соединения блокируется.

В случае фиксированной альтернативной маршрутизации для установки соединения используется список фиксированных кратчайших маршрутов. Вначале алгоритм производит установку соединения по первому маршруту в списке, если на маршруте нет свободных длин волн, то установка соединения происходит по следующему маршруту в списке. Недостаток такого подхода аналогичен предыдущему: если во всех маршрутах, содержащихся в списке, отсутствуют свободные длины волн, то запрос блокируется.

Адаптивная маршрутизация находит маршрут для заданного запроса установки соединения динамически, однако для этого необходимы мониторинг и учет текущего состояния сети. Преимуществом данного подхода является гибкость и низкий уровень вероятности блокировки по сравнению с методами, которые приведены выше. Таким образом, наиболее целесообразным является использование адаптивной маршрутизации [8].

Изложенный в работах [9, 10] метод адаптивной маршрутизации наименее загруженного маршрута (Least Load Routing, LLR) является одним из представителей распределенной модели управления. Основной идеей LLR является выбор маршрута, который включает в себя наибольшее количество свободных длин волн.

Метод адаптивной маршрутизации фиксированного наименее загруженного маршрута (Fixed Paths Least Congestion routing, FPLC) базируется на выборе фиксированного маршрута с наименьшим количеством задействованных вдоль него длин волн [11]. Результаты показали, что FPLC-алгоритм может значительно улучшить производительность по сравнению с другими фиксированными альтернативными алгоритмами маршрутизации. Однако, если в телекоммуникационной сети находится оптический конвектор, методы LLR и FPLC работают неэффективно, потому что при выборе маршрута они учитывают только распределение свободных длин волн, а длина маршрута не учитывается.

В основе метода, который базируется на расчете весовой функции и выборе наименее загруженного маршрута (Weighted Least-Congestion Routing First-Fit, WLCR-FF) [12], лежит выбор маршрута с наибольшим весом, рассчитываемым как

$$W(R) = \frac{F(R)}{\sqrt{h(R)}}$$

где W(R) – взвешенный коэффициент для маршрута R; F(R) – количество свободных длин волн на R - ом маршруте; h(R) – количество узлов на маршруте R. Данный метод учитывает не только пропускную способность маршрутов, но и их длину.

Большинство известных методов адаптивной маршрутизации выбирает маршруты из заранее заданного множества. Как правило, альтернативные пути для каждой пары источник-получатель, которые включены в это множество, являются реберно-непересекающимися кратчайшими путями [13]. При установке соединения маршрут передачи выбирается из заданного множества, базируясь на значении весовой функции для этого пути. При отсутствии свободных длин волн вдоль заранее заданных путей алгоритм пытается найти новый путь, содержащий свободные длины волн, и если такой путь найти не удается, то вызов блокируется.

Одним из представителей описанного выше класса методов адаптивной маршрутизации, который характеризуется достаточно высокой производительностью и высокой степенью близости получаемого решения к оптимальному, является метод динамической маршрутизации длин волн (Dynamic Wavelength Routing, DWR) [14]. Метод DWR предусматривает применение на первом этапе алгоритма выбора путей, базирующегося на использовании наименее загруженного пути с наименьшей степенью узлов (Least Congestion with Least Nodal-degree Routing algorithm, LCLNR) и алгоритма динамической маршрутизации длин волн с двух концов (Dynamic two-end wavelength routing algorithm, DTWR). Алгоритм маршрутизации с двух концов требует больших затрат времени. Кроме того, при расчете весов маршрутов не учитывается явление четырехволнового смешивания, что приводит к установке соединений, не удовлетворяющих требованиям по допустимой вероятности битовой ошибки.

Для устранения указанных недостатков требуется проведение дополнительных исследований, разработка новых и модернизация существующих методов решения задачи динамического выбора маршрутов и назначения длин волн в сетях WDM. Таким образом, целью данной работы является снижение вероятности блокировки вызовов при установке соединений и уменьшении вероятности битовой ошибки (BER) вследствие увеличения значения *Q*-фактора.

Математическая постановка задачи распределения длин волн в транспортной сети WDM

При решении задачи маршрутизации топологию сети будем описывать связным графом G(N,L), где $N = \{n_i\}$ – множество узлов сети; $L = \{l_{gu}\}$ – множество каналов связи оптической сети, где $l_{gu} = (n_g, n_u)$. Математическая постановка задачи приведена ниже.

Пусть n_s – узел источника при передаче трафика; n_d – узел получателя при передаче трафика; N_s – соседние узлы узла источника; N_d – соседние узлы узла получателя; W – множество длин волн, использование которых допустимо в рамках используемой технологии оптической сети; $K_{sd} = \{K_{sd}^i\}$ – множество маршрутов между узлами n_s и n_d .

Введем следующие обозначения: $h_{sd}^i = \|K_{sd}^i\| - 1$ – количество узлов на *i*-ом маршруте от $n_s \kappa n_d$; W_{sd}^i – множество свободных длин волн на *i*-ом маршруте; ω_{sd}^i – количество свободных длин волн на маршруте K_{sd}^i , где $\omega_{sd}^i = \|W_{sd}^i\|$; x_{sd}^{ih} – установка соединения с использованием λ -ой длины волны по маршруту K_{sd}^i ; ω_{gu} – количество свободных длин волн в l_{gu} -оптическом канале связи; W_{gu} – множество свободных длин волны по маршруту K_{sd}^i ; ω_{gu} – количество свободных длин волн в l_{gu} -оптическом канале связи; W_{gu} – множество свободных длин волн в l_{gu} -оптическом канале связи; M_{gu} – множество свободных длин волн в l_{gu} -оптическом канале связи; $deg(n_i)$ – степень вершины n_i в графе G(N,L); Ω_{sd}^i – количество длин волн, доступных для использования вдоль маршрута K_{sd}^i ; $\pi_{sd} \in K_{sd}$ – выбранный маршрут при установке соединения между парой (n_s, n_d) ; $O(\pi_{sd})$ – функция, определяющая целесообразность выбора маршрута π_{sd} .

Решение задачи сводится к необходимости найти:

- π_{sd} - маршрут передачи оптического сигнала между узлами (n_s, n_d) ;

- λ_{sd} – длину волны, используемую вдоль маршрута π_{sd} .

Критерий оптимальности – минимум вероятности блокировки:

$$Z_{Blocking} = \frac{\left|K_{sd}\right| - \sum_{i \in K} K_{sd}^{i}}{\left|K_{sd}\right|} \to \min.$$

Метод DWR базируется на использовании двух алгоритмов – LCLNR и DTWR.

Суть алгоритма LCLNR заключается в том, что при выборе светового пути между узлами выбирается наименее загруженный световой путь, тем самым уменьшая вероятность блокировки. Список *k*кратчайших путей для каждой пары (*s*, *d*) вычисляется с помощью модифицированного алгоритма Дейкстры [15]. Выбор маршрута из списка *k*-кратчайших путей осуществляется на основе коэффициента целесообразности, вычисляемого как

$$O(\pi_{sd}) = \frac{w}{h}$$

Данное выражение учитывает пропускную способность маршрута и количество узлов, через которые проходит маршрут. После этого будет выбран маршрут с максимальным значением коэффициента целесообразности. Если существует несколько маршрутов, то будет выбран маршрут с минимальной суммой степеней узлов вдоль маршрута,

$$\min\sum_{i=n_s}^{n_d} \deg(n_i)$$

Если после выполнения данного условия число маршрутов более 1, то маршрут будет выбран случайным образом. Псевдокод алгоритма LCLNR приведен ниже:

Входные данные:

граф сети G(N,L) и множество предоопределенных K кратчайших маршрутов, $K_{sd}^i, i \in [1,k]$ для соединений $e \in \{s,d\}$ и $\pi_{sd} \in K_{sd}^i, i \in [1,k]$

Вычисление функции $O(K_{sd}^i)$

 $O(K_{sd}^i) = w_i / h_i, \forall i \in [1, |K|]$

if $O(K^i_{sd}) = 0, \forall i \in [1, \left|K\right|]$ then блокировка вызова на установку соединения; return;

else
$$\pi_{sd} = \max_{i \in \{1..k\}} O(K_{sd}^i), \pi_{sd} \cup Temp;$$

end if
if $|Temp| \ge 2$, then вычисляется $\pi_{sd} = \min_{p \in [1,P]} \sum_{n_i \in P_{Temp_j}} \deg(n_i), P = |K_{sd}|$.
if $P \ge 2$ then маршрут выбирается случайным образом из множества Temp
end if
end if

Далее происходит установка соединения по выбранному маршруту и назначается длина волны $x_{sd}^{i\lambda}$ с помощью использования метода first-fit [3].

Если не удается установить соединение с помощью алгоритма LCLNR, то применяется алгоритм DTWR. Алгоритм DTWR рассматривает три сценария.

- Сценарий А. Отсутствуют свободные длины волн в оптических каналах связи от узлов источников и соседних узлов.
- Сценарий В. Есть некоторое количество свободных длин волн в оптических каналах связи от узлов источников и соседних узлов.
- Сценарий С. Есть свободные длины волн между узлами источниками смежными, но отсутствуют свободные длины волн на промежуточных узлах на световом маршруте.

Алгоритм DTWR работает только при сценариях **В** и **С**. Первый шаг алгоритма DTWR – определение сценария. Если сценарий – не **A**, то выполняется удаление ребер из графа, которые не содержат свободных длин волн. После этого выполняется поиск кратчайшего маршрута с помощью алгоритма Дейкстры. На последнем шаге происходит передача этого маршрута на рассмотрение алгоритму LCLNR. Псевдокод алгоритма DTWR приведен ниже.

while
$$\pi_{sd} = \emptyset$$
;

do вычислить
$$\omega_{n_s n_v}$$
, $\omega_{n_d n_c}$, $W_{n_s n_v}$ и $W_{n_d n_c}$

if ($W_{n_sn_v}=0$) или ($W_{n_dn_c}=0$) then вызов блокируется; return;

else if
$$\left(\bigcup_{i,\,n_i\in N_s}W_{si}\right)\cap\left(\bigcup_{i,\,n_i\in N_d}W_{id}\right)\neq\emptyset$$
 then

вызов на запрос установки соединения блокируется; return;

else удаляются ребра с графа G(N,L), где $\omega_{n_sn_v}=\omega_{n_dn_c}=0$; генерируется граф G'(N,L')

1 paų 0 (11, 1

end if

Вычисляется множество путей $K_{n,n_d}^{\prime i}$ с помощью алгоритма Дейкстры;

if $K_{n_sn_d}^{\prime i}=\varnothing$ then $K_{sd}^i=\varnothing$, вызов на установку соединения блокируется return:

else $K^i_{sd} \leftarrow K^{\prime i}_{n,n_d}$, переход к алгоритму LCLNR

end if

end while.

При исследовании метода решения задачи DRWA были обнаружены следующие недостатки:

- при использовании алгоритма DTWR возникают большие временные затраты на этапе проверки условий сценариев. Этот фактор является критичным при решении задачи DRWA;
- алгоритмом LNCLR при выборе светового маршрута не учитывается такое нелинейное искажение в волоконно-оптических сетях, как четырехволновое смешивание (ЧВС);
- метод назначения длин волн first-fit неэффективен при наличии ЧВС [16].

Разработка метода устранения приведенных недостатков

Последствием ЧВС является появление побочных сигналов, в том числе на длинах волн, соответствующих другим рабочим каналам, что может привести к росту ошибок и ухудшению эффективности системы WDM. Для учета влияния данного явления при выборе маршрута была использована [17] следующая формула расчета мощности помехи ЧВС:

$$P_{ijk}(f_i,f_j,f_k) = \frac{\eta}{9} D^2 \gamma^2 P_i P_j P_k e^{-\alpha L} \left\{ \frac{(1-e^{-\alpha L})^2}{\alpha^2} \right\},$$

где P_i , P_j , P_k – мощности входных канальных сигналов на частотах f_i , f_j , f_k соответственно; D – коэффициент вырожденности; α – коэффициент затухания оптического волокна; L – длина отрезка оптического волокна.

Коэффициент нелинейности на длине волны λ рассчитывается по формуле

$$\gamma = \frac{2 \pi n_2}{\lambda A_{3 \phi \phi}}$$

где n_2 – коэффициент нелинейности показателя преломления; $A_{_{3\phi\phi}}$ – эффективная площадь оптического волокна.
Эффективность η ЧВС описывается следующим выражением:

$$\eta = \frac{\alpha^2}{\alpha^2 + \Delta\beta^2} \left| 1 + \frac{4e^{-\alpha L} \sin^2\left(\frac{\Delta\beta L}{2}\right)}{(1 - e^{-\alpha L})} \right|$$

Коэффициент фазового согласования $\Lambda\beta$ зависит от хроматической дисперсии $D_c(\lambda)$ и описывается следующим выражением:

$$\Lambda\beta = \frac{2\pi\lambda_k^2}{c}\Delta f_{ik}\Delta f_{jk}\left[D_c(\lambda) + \frac{\lambda_k^2}{2c}(\Delta f_{ik} + \Delta f_{jk})\frac{dD_c(\lambda)}{d\lambda}\right]$$

где интервал между каналами $\Lambda f_{ik} = |f_i - f_k|$, $\Lambda f_{jk} = |f_j - f_k|$; c – скорость света.

Мощность помехи ЧВС на частоте f_m равна сумме мощностей всех комбинационных продуктов:

$$P_{\rm YBC}(f_m) = \sum_{i=1}^N \sum_{j=1}^N P_{ijk}(f_i, f_j, f_k) ,$$

где N-количество каналов.

Проведем расчет мощности усиленного спонтанного излучения (ASE) с помощью выражения $P_{ase} = 2n_{sp}(G-1)hf_m\Delta f_0$,

где n_{sp} – коэффициент спонтанной эмиссии усилителя; h – постоянная Планка; Λf_0 – ширина полосы пропускания оптического фильтра демультиплексора WDM; G – коэффициент усиления усилителя.

Допустим, что все участки имеют одинаковую длину, следовательно, мощность усиленного спонтанного излучения на входе приемного оптического модуля (ПрОМ) равна сумме соответствующих мощностей на выходе всех усилителей:

$$P_{ase \sum} = P_{ase} (N_{yc} + 1) ,$$

где N_{yc} – количество усилителей. Тогда мощность ЧВС на входе ПрОМ равна

$$P_{\text{YBC}\sum} = P_{\text{YBC}} G(N_{\text{yc}} + 1)$$

На выходе фотоприемника оптический шум ЧВС и ASE соответственно формируют электрический сигнал с мощностями

$$\begin{split} P_{e^{\text{UBC}}\sum} &= 2z^2 P_{\text{BX}} \frac{P_{\text{UBC}}\sum}{8}, \\ P_{ease\sum} &= 4z^2 P_{\text{BX}} P_{ease\sum} \cdot \frac{\Delta f_e}{\Delta f_0} \end{split}$$

где Δf_e – полоса пропускания электрического усилителя ПрОМ.

Чувствительность фотоприемника равна

$$z=\frac{\eta e}{hf_m},$$

где η – квантовая эффективность фотодетектора; е – заряд электрона.

Q-фактор и связанная с ним вероятность ошибки рассчитываются по формулам

$$Q \approx \frac{P_{\text{BX}}}{\sqrt{P_{ease}\sum + P_{e^{\text{HBC}}\sum}}},$$

$$P_{\text{out}} = \frac{1}{\sqrt{2\pi}} \int_{Q}^{\infty} e^{-\frac{x^2}{2}} dx.$$
(1)

где $P_{e^{\text{чBC}}}$ – суммарная мощность ЧВС, $P_{\text{вх}}$ – входная мощность; $P_{ease\sum}$ – суммарная мощность ASE. Полученное выражение (1) используется при расчете целевой функции на этапе выбора светового маршрута алгоритмом LCLNR следующим образом:

$$O(\pi_{sd}) = \frac{w \cdot FMP(R)}{h}$$

где FMP(R) – коэффициент, который уменьшает значение $O(\pi_{sd})$, если вероятность ошибки для маршрута R превысила допустимое значение.

На этапе расчета коэффициента целесообразности выбора маршрута для установки соединения происходит сравнение рассчитанного значения *Q*-фактора на каждом оптическом канале, через который проходит рассматриваемый маршрут, с допустимым значением *Q*-фактора. Если значение *Q*-фактора превышает допустимую норму, то значение целевой функции уменьшается.

Как показали исследования, использование алгоритма DTWR приводит к большим временным затратам, поэтому в работе предлагается вместо него использовать алгоритм Дейкстры для поиска кратчайшего маршрута между заданной парой источник–получатель с использованием следующей метрики:

$$\ln(P(K_{sd}^{i})) = \ln(1 - \frac{W_{sn}}{W}) + \dots + \ln_{m}(1 - \frac{W_{kd}}{W}),$$

где w_{sn} – количество занятых длин волн в оптическом канале между узлами *s* и *n*. Данная метрика характеризует вероятность успешной установки соединения, так как величина $P(K_{sd}^i)$ пропорциональна вероятности не блокирования маршрута.

Методика исследования эффективности предложенных методов базируется на имитационном моделировании. На вход имитационной модели поступал простейший поток заявок на установку соединения по световому пути. В дальнейшем производился выбор маршрута и назначение ему длины волны. В случае успешной установки соединения в модели соответствующие ресурсы помечались как занятые. В имитационной модели также моделировалась длительность соединения, после которого занятые соединением сетевые ресурсы освобождались. При проведении эксперимента использовалась топология оптической сети с 20 узлами, ранг каждого узла в сети – не меньше 3. Эксперимент проводился при различных интенсивностях поступления запросов на установку соединения, интервалы поступления запросов и время удержания ресурсов являлись случайными величинами, распределенными по экспоненциальному закону. Величина загруженности сети изменялась в пределах (0,2–0,9).

Для назначения длин волн световому маршруту в предложенном методе использовался метод random, а в базовом – first-fit. В результате проведения эксперимента были получены следующие результаты, показанные на рис. 1.



Рис. 1. Зависимость вероятности блокировки от интенсивности трафика для базового и модифицированного методов алгоритмов

В ходе анализа полученных результатов (рис. 1) установлено, что наибольший выигрыш вероятности блокировки (19%) соответствует интенсивности трафика 1,5. При интенсивности трафика 1 вероятность блокировки снижается на 14%, при 0,8 – на 8%. В эксперименте с интенсивностью трафика 0,4 из-за низкой вероятности блокировки не было зафиксировано ни одного заблокированного вызова.

На рис. 2 показаны значения *Q*-фактора базового и модифицированного методов при различных интенсивностях поступления запросов на установку соединения.

Анализируя полученные результаты, можно увидеть, что выигрыш *Q*-фактора растет с уменьшением интенсивности поступления запросов на установку соединения.



Рис. 2. Зависимость значений *Q*-фактора от интенсивности трафика для базового и модифицированного методов алгоритмов

При совместном анализе рисунков 1, 2 можно сделать вывод, что предложенные изменения в методе адаптивной маршрутизации по сравнению с базовым методом позволяют снизить вероятность блокировки и вероятность битовых ошибок (которая является функцией Q-фактора) и требуют меньшего числа вычислительных итераций. Наибольший выигрыш по уменьшению вероятности блокировки имеет место при средней интенсивности поступления вызовов. В случае низкой интенсивности трафика быстрее освобождаются ресурсы сети, и необходимость использования алгоритма DTWR уменьшается. Наибольший выигрыш для *Q*-фактора, напротив, получается при низкой интенсивности поступления вызовов.

Заключение

В работе проведен анализ методов решения задачи динамической маршрутизации и назначения длин волн (DRWA). Показано, что наиболее эффективным является метод адаптивной маршрутизации, как обеспечивающий гибкость и низкий уровень вероятности блокировки. В работе был рассмотрен метод адаптивной маршрутизации на основе использования алгоритмов LCLNR и DTWR. Показано, что использование алгоритма DTWR приводит к большим временным затратам на стадии проверки условий сценария, что является критичным при решении задачи DRWA. Не учитываются нелинейные искажения в волоконно-оптических сетях, такие как ЧВС, при выборе светового маршрута с помощью алгоритма LNCLR. Метод назначения длин волн first-fit неэффективен в случае наличия в каналах нелинейных явлений.

Для устранения данных недостатков предложена модификация алгоритма: вместо использования алгоритма DTWR предложен алгоритм Дейкстры с использованием метрики, которая основывается на пропускной способности маршрутов; введены математические выражения для учета ЧВС при расчете вероятности выбора светового маршрута; использован метод random при назначении длин волн световым маршрутам вместо first-fit.

Проведен эксперимент с использованием средств имитационного моделирования на искусственно синтезируемой топологии сети. При моделировании решалась задача RWA при различных интенсивностях поступлений вызовов на установку соединения. В ходе анализа полученных результатов было выявлено, что модифицированный метод адаптивной маршрутизации имеет наибольшую эффективность при интенсивности в 1,5 единицы, при этом выигрыш составляет 19%, тогда как среднее значение выигрыша составляет 13%. Наибольшее значение *Q*-фактора наблюдалось при самой низкой интенсивности трафика (0,4), так как в этом случае задействовано наименьшее число длин волн в оптических каналах. Средний выигрыш *Q*-фактора составил 0,812.

Литература

- Gangxiang S., Rodney S.T. Translucent Optical Networks: The Way Forward // IEEE Communications Magazine. – 2007. – V. 45. – P. 48–54.
- 2. Ramesh G., Sundaravadivelu S. Reliable Routing and Wavelength Assignment for Optical WDM Networks // European Journal of Scientific Research. – 2010. – V. 48. – № 1. – P. 85–96.
- Leonardi E., Mellia M., Marsan M.A. Algorithms for the logical topology design in WDM all-optical networks // Optical Networks. – 2000. – V. 1. – P. 35–46.

- 4. Mokhtar A., Azizoglu M. Adaptive wavelength routing in all-optical networks // IEEE/ACM Transactions on Networking. 1998. V. 6. № 2. P. 197–206.
- 5. Jianjun Yu. Radio-over-optical-fiber networks: introduction to the feature issue // Journal of Optical Networking. 2009. V. 8 (5). P. 481–488.
- 6. Gomes P.H., Fonseca da N.L.S., Branquinho O.C. Optimization of the use of Radio Resource of Radio-Over-Fiber Access Networks // Global Telecommunications Conference (GLOBECOM 2010). – 2010. – P. 1–5.
- Mirosław Klinkowski, Marek Jaworski, Davide Careglio. Channel Allocation in Dense Wavelength Division Multiplexing Radio-over-Fiber Networks // 12th International Conference on Transparent Optical Networks. - 2010. – P. 1–5.
- 8. Zang H., Jue J. P., Mukherjee B. A review of routing and wavelength assignment approaches for wavelength-routed optical networks // Optical Network Magazine. 2000. V. 1. № 1. P. 47–60.
- Karasan E., Ayanoglu E. Effects of wavelength routing and selection algorithms on wavelength conversion gain in WDM optical networks // IEEE/ACM Transactions on Networking. – 1998. – V. 6. – № 2. – P. 186– 196.
- 10. Birman A. Computing Approximate Blocking Probabilities for a Class of All-Optical Networks // IEEE Journal on Selected Areas in Communications. 1996. V. 14. № 5. P. 852–857.
- 11. Li L., Somani A.K. Dynamic Wavelength Routing Using Congestion and Neighborhood Information // IEEE/ACM Transactions on networking. 1999. V. 7. № 5. P. 779–786.
- Xiaowen Chu, Bo Li, Zhensheng Zhang. A Dynamic RWA Algorithm in a Wavelength-Routed All-Optical Network with Wavelength Converters // INFOCOM. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies. – 2003. – V. 3. – P. 1795–1804.
- 13. Yates J.M., Rumsewicz M.P. Wavelength converters in dynamically-reconfigurable WDM networks // IEEE Communications Surveys & Tutorials. 1999. V. 2. № 2. P. 2–15.
- Kungmang Lo, Daryoush Habibi, Quoc Viet Phung, Hoang Nghia Nguyen. Dynamic Wavelength routing in all optical meshnetwork // Proceedings of Asia Pacific Conference on Communications. – 2005. – P. 178–182.
- 15. Bhandari R. Survivable Networks: Algorithms for Diverse Routing. Kluwer Academic Publishers, 1999. 200 p.
- 16. Агеев Д.В., Ковальчук В.К., Переверзев А.А. Планирование распределения длин волн при проектировании транспортной сети DWDM // Восточно-европейский журнал передовых технологий. 2011. № 5/3 (53). С. 25–29.
- 17. Педяш В.В., Решетников О.С. Оптимізація потужності лінійного сигналу системи DWDM // Цифрові Технології. 2009. № 5. С. 27–33.

Агеев Дмитрий Владимирович	-	Харьковский национальный университет радиоэлектроники, доктор
		технических наук, профессор, dm@ageyev.in.ua
Переверзев Александр Анатольевич	_	Харьковский национальный университет радиоэлектроники, аспирант,
		pereverzev_aa@mail.ru

УДК 528.88

ПРОГНОЗИРОВАНИЕ КАЧЕСТВА ИЗОБРАЖЕНИЙ КОСМИЧЕСКИХ ОБЪЕКТОВ

А.И. Алтухов, Н.В. Гнусарев, Д.С. Коршунов

Описывается подход к прогнозированию качества изображений космических объектов, который может использоваться при планировании рабочих программ бортовых оптико-электронных систем спутников дистанционного зондирования. В основе предлагаемого авторами подхода лежит оценивание влияния на передаточные характеристики оптико-электронных систем наблюдения космических комплексов баллистических условий орбитальной съемки в задачах прогнозирования разрешающей способности и линейного разрешения на объекте.

Ключевые слова: оптико-электронные системы, разрешающая способность, линейное разрешение на объекте, сдвиг оптического изображения.

Введение

В настоящее время численность неуправляемых объектов искусственного происхождения, т.е. так называемого космического мусора, в околоземном космическом пространстве достигла такой величины, что становится неразумно не считаться с реальной опасностью повреждения дорогостоящей космической техники в результате столкновений с техногенными орбитальными частицами и даже более крупными фрагментами [1]. В этих условиях особо актуальна задача контроля технического состояния спутников с целью своевременного обнаружения неисправностей и принятия решений по их устранению, что обеспе-

чивает целостность ключевых звеньев отечественной космической инфраструктуры и ее эффективное функционирование [2]. Решение данной задачи возможно с использованием детальных изображений космических аппаратов (КА), полученных в ходе орбитальной инспекции. Примером такой инспекции может служить применение французского спутника дистанционного зондирования Земли (ДЗЗ) «Pleiades-1» для съемки аварийного спутника ДЗЗ «ENVISAT», принадлежащего Европейскому космическому агентству «ESA». Многократная уникальная съемка производилась в панхроматическом и многоспектральном режимах (рис. 1) с расстояния около 100–120 км, во время пролета спутника «Pleiades-1» под аппаратом «ENVISAT». В соответствии с программой съемки «Pleiades-1» был развернут телескопом от Земли и сопровождал аварийный спутник в течение восьми минут [3].

Полученные в ходе инспекции снимки позволили оценить состояние обшивки «ENVISAT», а также выявить изменения в его пространственной ориентации.

Результаты моделирования движения КА [4–7] с целью анализа пространственно-временных характеристик съемки объектов бортовыми оптико-электронными системами космических средств позволили сделать вывод о том, что возможность получения изображений, пригодных для интерпретации, значительно ограничивается влиянием на их качество баллистических условий орбитальной съемки. Наблюдения показали, что наиболее качественная съемка возможна, когда наблюдающий и наблюдаемый спутники находятся на орбитах с близкими параметрами. Например, спутник «Pleiades-1» имеет круговую солнечносинхронную орбиту высотой 697/699 км, что примерно на 70 км ниже высоты орбиты «ENVISAT» (766/768 км), а угол между плоскостями орбит составляет всего 6°. В случае, когда движение КА выполняется по орбитам, лежащим в разных плоскостях, изображения объектов получаются размытыми и смазанными. Причиной искажений является сдвиг изображения фокальной плоскости объектива бортовой оптико-электронной системы, возникающий вследствие высоких относительных скоростей движения КА в момент съемки, рост которого происходит по мере сокращения дальности между спутниками.



Рис. 1. Снимки аварийного спутника «ENVISAT»

Таким образом, ограничения, накладываемые баллистическими условиями орбитальной съемки на качество изображений, ведут к снижению периодичности контроля состояния КА и снижению возможности применения средств ДЗЗ для орбитальной инспекции в целом. С целью повышения качества изображений авторы предлагают оценивать влияние сдвига, вызванного баллистическими условиями съемки, на передаточные характеристики бортовых оптико-электронных систем в задаче прогнозирования линейного разрешения на объекте. Это позволит найти оптимальные пространственно-временные характеристики съемки объектов, обеспечивающие устранение искажений вследствие высоких относительных скоростей и допустимую дистанцию съемки для получения необходимого линейного разрешения. Новизна предлагаемого авторами подхода заключается в том, что поиск оптимальных условий съемки выполняется как для спутников, имеющих схожие параметры орбит, так и для спутников, орбиты которых лежат в разных плоскостях.

Модель линейного разрешения на объекте

В различных условиях съемки оптико-электронная система формирует изображения различного качества с точки зрения передачи структуры и формы объекта. Для прогнозирования качества изображений авторами предложена модель линейного разрешения на объекте, которая позволяет оценить передаточные свойства бортовой оптико-электронной системы и рассчитать показатели качества изображений, получаемых этой системой в заданных условиях съемки.

Под качеством понимается свойство изображения, характеризующее способность нести в себе сведения о геометрических и фотометрических характеристиках и параметрах объектов. Показатель качества изображений – величина, служащая конкретным индикатором этого свойства [8, 9]. Оценивание качества изображений космических объектов выполняется по таким показателям, как разрешающая способность и линейное разрешение на объекте. Под разрешающей способностью понимается максимальная пространственная частота периодической решетки, штрихи которой визуально различимы на изображении, полученном оптико-электронной системой, при использовании в качестве объекта стандартной миры заданного контраста [10]. Под линейным разрешением на объекте понимается половина периода предельно разрешаемого поля миры, приведенного к объекту наблюдения [8, 9].



Рис. 2. Звенья оптико-электронной системы

В предложенной модели основными звеньями оптико-электронной системы (рис. 2) являются объектив, матрица прибора с зарядовой связью (МПЗС) и сдвиг изображения. Оценивание передаточных свойств оптико-электронной системы выполняется на основе ее частотно-контрастной характеристики (ЧКХ), которая показывает, как изменяется контраст изображения периодической решетки с изменением пространственной частоты [8, 9].

При формировании изображения объектив оптико-электронной системы оказывает размывающее воздействие, причинами которого являются дифракция световой волны на ограничивающих ее диафрагмах и остаточные аберрации. Передаточные свойства объектива в частотной области хорошо описываются ЧКХ идеального дефокусированного объектива, которая находится по формуле

$$T_{o}(\mathbf{v}) = \left[\sin\left(\pi \mathbf{v}\Delta'\left(\frac{1}{k}-\lambda \mathbf{v}\right)\right)\right] / \left(\frac{\pi \mathbf{v}\Delta'}{k}\right)$$

где Δ' – дефокусировка объектива; k – диафрагменное число. Определение передаточных свойств МПЗС выполняется на основе ЧКХ диффузионного расплывания заряда, геометрической ЧКХ дискретного переноса и технологических потерь приемника лучистой энергии, а также ЧКХ фазы МПЗС. В частности, ЧКХ диффузионного расплывания заряда рассчитывается по формуле

$$T_{drz}(\mathbf{v}) = \left(1 - \frac{\exp(-\alpha_s W)}{1 + \alpha_s L}\right) / \left(1 - \frac{\exp(-\alpha_s W)}{1 + \alpha_s L_o}\right),$$

где W – глубина обедненного слоя; L_o – диффузионная длина носителей; α_s – коэффициент поглощения в кремнии на средней длине волны. Геометрическая ЧКХ дискретного переноса и технологических потерь МПЗС в режиме временной задержки накопления находится по формуле:

$$T_{gpd}(\mathbf{v}) = \left(\sin(a)/a\right) \left(\sin(b)/b\right) \exp\left[-b_p \left(4\mathbf{v}d\right)^2\right]$$

где *a* – коэффициент геометрического размера элемента МПЗС; *b* – коэффициент дискретного переноса заряда; *b_p* – относительная величина технологических потерь МПЗС; *d* – размер элемента МПЗС. Расчет ЧКХ фазы приемника лучистой энергии выполняется по формуле:

$$T_f(\mathbf{v}) = \cos(\pi \mathbf{v} d)^{\mathbf{v} d}$$

Графические зависимости, полученные на основе экспериментальных данных о передаточных свойствах объектива и МПЗС, приведены на рис. 3.

Результирующая ЧКХ матрицы ПЗС рассчитывается по формуле

$$T_{pzs}\left(\mathbf{v}\right) = T_{drz}\left(\mathbf{v}\right)T_{gpd}\left(\mathbf{v}\right)T_{f}\left(\mathbf{v}\right).$$

Орбитальная съемка в условиях высоких относительных скоростей движения спутников приводит к возникновению сдвига изображения, построенного объективом оптико-электронной системы и, как следствие, к снижению качества снимка. Оценить величину сдвига изображения можно по формуле

$$T_{sd}\left(\nu\right) = \left[\frac{\sin\pi H_{sd}\eta_{0}\nu}{\pi H_{sd}\eta_{0}\nu}\frac{\sin\pi H_{sd}\left(1-\eta_{0}\right)\nu}{\pi H_{sd}\left(1-\eta_{0}\right)\nu}\right],$$

где η_0 – оптический коэффициент полезного действия фотозатвора; H_{sd} – сдвиг оптического изображения, который рассчитывается по формуле

$$H_{sd} = T_{ex} V_{im},$$

где T_{ex} – время экспонирования; V_{im} – скорость движения оптического изображения в фокальной плоскости. Результирующая ЧКХ оптико-электронной системы равна произведению ЧКХ ее звеньев и рассчитывается как



В качестве показателя для оценивания качества изображения, характеризующего способность оптико-электронной системы передавать структуру и форму мелких деталей объекта, используется разрешающая способность, которая может быть определена как пространственная частота, для которой выполняется равенство

$$cT_{os}\left(\nu\right)=C_{p}\left(\nu\right),$$

где *с* – контраст, при котором находится разрешающая способность; $T_{os}(v)$ – результирующая ЧКХ оптико-электронной системы; $C_p(v)$ – пороговый контраст на изображении объекта. В свою очередь, пороговый контраст, визуально воспринимаемый на изображении объекта, рассчитывается по формуле

 $C_{p}(v) = (2, 3q_{p} / g) (D_{za} + D_{st})^{0.5}$, где q_{p} – пороговое отношение сигнал/шум; g – градиент светосигнальной характеристики; D_{za} – дисперсия шума зрительного анализатора; D_{st} – дисперсия шума системы тракта. Вторым показателем качества изображений служит линейное разрешение на объекте, характеризующее линейный размер минимального элемента различимого на изображении объекта и рассчитываемое по формуле

$$L_p = \frac{dk_m}{2R_c f'}$$

где *d* – дальность съемки; *f* ′ – фокусное расстояние; *k*_m – коэффициент формы объекта.

Влияние баллистических условий съемки на качество изображений

Результаты экспериментальных расчетов показывают, что параметры орбит спутников значительно влияют на скорость движения оптического изображения и, как следствие, на качество получаемых снимков. Например, в случаях, когда движение спутников выполнялось по круговым орбитам с высотами 650–750 км, а угол между плоскостями орбит не превышал 10°, максимальное значение скорости движения оптического изображения достигало 3–9 мм/с. В случаях, когда угол между плоскостями орбит спутников менялся от 10° до 180°, скорость движения оптического изображения достигала значений 40– 130 мм/с. Подобные изменения скорости движения оптического изображения существенно влияют на ЧКХ сдвига изображения и, как следствие, на результирующую ЧКХ всей оптико-электронной системы. Как видно из рис. 4, чем ниже скорость движения оптического изображения и, следовательно, меньше величина сдвига, тем выше частота штрихов, которые могут быть переданы оптико-электронной системой при формировании изображения периодической решетки.

На рис. 5 представлены графические зависимости, наглядно демонстрирующие, как изменяется разрешающая способность оптико-электронной системы при различных ЧКХ. Из рис. 5 видно, что, чем выше скорость движения оптического изображения и, следовательно, больше величина сдвига, тем ниже контраст объекта на изображении, построенном оптико-электронной системой, и ниже разрешающая способность. При выполнении экспериментальных расчетов также было установлено, что максимальный сдвиг изображения соответствует минимальной дальности съемки. На рис. 6 представлены графические зависимости линейного разрешения на объекте от дискретной совокупности пространственно-временных положений спутников на орбитах в моменты съемки, где точка 60 по оси абсцисс соответствует минимальной дистанции наблюдения, равной 120 км.



Рис. 6. Линейное разрешение на объекте

Из рис. 6 видно, что минимальная дальность съемки соответствует наименьшему значению линейного разрешения на объекте и наилучшему качеству изображения только тогда, когда параметры орбит наблюдаемого и наблюдающего спутников имеют схожие параметры (кривые 1, 2 и 3). В случае, когда орбиты спутников лежат в разных плоскостях, из-за сдвига оптического изображения снимки получаются худшего качества, чем при съемке с больших дистанций (кривые 4 и 5). Таким образом, прогнозирование качества изображений позволяет определить оптимальные пространственно-временные параметры съемки для получения снимков, пригодных для интерпретации.

Заключение

В работе рассмотрена модель линейного разрешения на объекте, которая позволяет прогнозировать качество изображений космических объектов, получаемых оптико-электронными системами спут-

ников дистанционного зондирования. Достоинством данной модели является то, что оценивание передаточных свойств оптико-электронных систем спутников и расчет показателей качества итоговых изображений осуществляется с учетом баллистических условий орбитальной съемки, а именно – с учетом относительных скоростей движения космического аппарата. Модель линейного разрешения может использоваться на этапе планирования программы работы бортовых оптико-электронных систем, что обеспечит экономию ресурса орбитальных комплексов и повысит периодичность контроля технического состояния спутников.

Литература

- 1. Иванов В.Л., Меньшиков В.А., Пчелинцев Л.А., Лебедев В.В. Космический мусор: Проблема и пути ее решения. М.: Патриот, 2002. 360 с.
- 2. Суханов С.А., Шаргородский В.Д., Шилин В.Д. Система контроля космического пространства. М., 2011. 206 с.
- Кучейко А.А. Уникальное применение спутника ДЗЗ орбитальная инспекция [Электронный ресурс]. – Режим доступа: http://www.scanex.ru/ru/news/ News_Preview.asp, свободный. Яз. рус. (дата обращения 28.02.2013).
- Иванов Н.М., Лысенко Л.Н. Баллистика и навигация космических аппаратов: Учебник для вузов. М.: Дрофа. 2004. – 544 с.
- 5. Гнусарев Н.В. Геодезическое и баллистическое обеспечение космических систем дистанционного зондирования Земли. СПб: ВКА имени А.Ф. Можайского, 2008. 220 с.
- 6. Разыграев А.П. Основы управления полетом космических аппаратов. М: Машиностроение, 1990. 440 с.
- Мамон П.А., Кульвиц А.В. Теория полета космических аппаратов. Курс лекций. СПб: ВКА имени А.Ф. Можайского, 2007. – 160 с.
- Мельканович А.Ф. Фотографические средства и их эксплуатация. М.: Министерство обороны, 1992. – 440 с.
- 9. Авдеев С.П. Анализ и синтез оптико-электронных приборов. СПб: Правда, 2000. 680 с.
- 10. ГОСТ 23935-79. Аэрофотоаппаратура и аэрофотографирование. Термины и определения Введ. 01.01.1981. М.: Гос. комитет СССР по стандартам. 24 с.

Алтухов Александр Иванович Военно-космическая академия имени А.Ф. Можайского, кандидат технических наук, зав. кафедрой, ааi_51@mail.ru Гнусарев Николай Васильевич Военно-космическая академия имени А.Ф. Можайского, кандидат технических наук, доцент, gnusarev2006@rambler.ru Коршунов Денис Сергеевич Военно-космическая академия имени А.Ф. Можайского, соискатель, Korshunov.Denis@rambler.ru

2

ФОТОНИКА И ОПТОИНФОРМАТИКА

УДК 535.3

РАСЧЕТ ВРЕМЕНИ КОГЕРЕНТНОСТИ ФЕМТОСЕКУНДНОГО СПЕКТРАЛЬНОГО СУПЕРКОНТИНУУМА В КВАРЦЕВОМ ОПТИЧЕСКОМ ВОЛОКНЕ М.В. Мельник, А.Н. Цыпкин

Показано, что при генерации фемтосекундного спектрального суперконтинуума в кварцевом волокне в областях нормальной, нулевой и аномальной дисперсий групповых скоростей с увеличением центральной длины волны фемтосекундного лазерного импульса на входе время когерентности излучения со сверхшироким спектром значительно уменьшается, однако в области нулевой групповой дисперсии кварцевого волокна существует «скачок» времени когерентности. Например, для длительности исходного импульса 40 фс и пиковой интенсивности 10¹³ Вт/см² при генерации спектрального суперконтинуума в кварцевом волокне на длине волны исходного излучения 800 нм время когерентности равно 22 фс, при увеличении длины волны время когерентности уменьшается до 4 фс (1180 нм). В области нулевой групповой дисперсии (1260 нм) время когерентности резко увеличивается до 20 фс, после чего опять уменьшается, достигая минимума в 4 фс (1560 нм).

Ключевые слова: фемтосекундный спектральный суперконтинуум, время когерентности излучения.

Введение

Генерация суперконтинуума в оптических волокнах – хорошо изученный процесс и может включать различные нелинейные эффекты, такие как самомодуляция и кросс-модуляция, солитонные эффекты, комбинационное рассеяние, модуляционная неустойчивость и четырехволновые смешения. Большинство спектральных суперконтинуумов генерируются в области аномальной дисперсии групповой скорости, в данном случае в процессе уширения спектра доминируют солитонные эффекты [1–4]. Эти эффекты приводят к тому, что спектр становится чувствительным к импульсу, что, в свою очередь, приводит к низкой согласованности между соседствующими импульсами, а следовательно, и к низкой когерентности.

Избавиться от этого можно двумя способами – либо тщательно выбирать входные параметры [4], либо генерировать спектральный суперконтинуум в области с нормальной дисперсией групповых скоростей, где солитоны формироваться не могут, а генерация фемтосекундного спектрального суперконтинуума происходит в основном из-за фазовой самомодуляции (ФСМ) [5–7]. ФСМ является внутренним детерминированным процессом, который сохраняет согласованность входных импульсов. По этой причине спектральный суперконтинуум, сгенерированный в этом режиме, обладает высокой спектральной согласованностью и стабильностью. Однако в этом случае требуется большая интенсивность лазерных импульсов.

В настоящее время вопрос когерентности считается почти полностью проработанным. Однако работ по исследованию времени когерентности обнаружено не было, а знание о времени когерентности спектрального суперконтинуума может быть использовано, например, для расчета максимального разрешения оптической когерентной томографии [8–11].

В настоящей работе выявлены оптимальные условия генерации высококогерентного фемтосекундного спектрального суперконтинуума в оптическом кварцевом волокне на основе определения времени когерентности исследуемого излучения и зависимости времени когерентности от длины волны. Например, при увеличении длительности входного импульса время когерентности увеличивается. В области нормальной групповой дисперсии существует максимум времени когерентности на длине волны 800 нм, после чего величина времени когерентности значительно уменьшается при смещении к области аномальной групповой дисперсии. Однако в области нулевой групповой дисперсии (~1260 нм для оптического кварцевого волокна) существует скачок увеличения времени когерентности, после чего величина времени когерентности продолжает убывать и достигает своего минимума в области аномальной групповой дисперсии оптической среды. Например, для длительности исходного импульса 40 фс на длине волны 800 нм время когерентности равно 22 фс, при увеличении длины волны время когерентности падает до 4 фс на длине волны 1180 нм. В области нулевой групповой дисперсии время когерентности резко увеличивается до 20 фс, после чего опять уменьшается, достигая минимума в 4 фс на длине волны 1560 нм.

Математическая модель самовоздействия фемтосекундных световых импульсов в диэлектрической среде

Распространение интенсивного светового импульса в волноведущей диэлектрической среде с нормальной групповой дисперсией и нерезонансной нелинейностью может быть описано следующим уравнением [12, 13]:

$$\frac{\partial E}{\partial z} - a \frac{\partial^3 E}{\partial \tau^3} + b \int_{-\infty}^{\tau} E d\tau' + g E^2 \frac{\partial E}{\partial \tau} = 0, \qquad (1)$$

где E – электрическое поле светового импульса; z – пространственная координата, вдоль которой распространяется импульс; $\tau = t - \frac{N_0}{c} z$ – время в движущейся с импульсом системе координат, t – время, c – скорость света в вакууме; параметры N_0 , a и b описывают линейный показатель преломления среды и его дисперсию:

$$n(\omega) = N_0 + ac\omega^2 - c\frac{b}{\omega^2},$$
(2)

параметр *g* характеризует безынерционную кубическую по полю нелинейность поляризационного отклика среды. Уравнение (1) описывает самовоздействие светового импульса, в том числе при его высокой интенсивности, и генерацию спектрального суперконтинуума [12]. Теоретический анализ этого уравнения будем проводить для случаев, когда спектр интенсивных импульсов находится в области нормальной, нулевой и аномальной групповой дисперсии среды. Метод решения данного уравнения приведен в работе [13]. Например, для распространенного в лазерной технике кварцевого стекла зависимость (2) при N_0 = 1,45, $a = 4,04 \cdot 10^{-42} \text{ с}^3/\text{м}, b = 0$ и коэффициенте нелинейного показателя преломления n_2 = 2,9 $\cdot 10^{-16}$ см²/Вт описывает дисперсию линейного показателя преломления стекла с точностью до третьего знака после запятой в диапазоне от 550 до 1100 нм [6]. Центральная длина волны $\lambda_0 = 2\pi c/\omega_0$ = 1260 нм спектра входного импульса соответствует длине волны нулевой групповой дисперсии кварцевого волокна. Импульс на входе в

среду будем полагать гауссовым вида $E = E_0 \cdot e^{-\left(\frac{t}{\tau_p}\right)^2} \cdot \sin(\omega_0 t)$, где E_0 – максимальная амплитуда электрического поля излучения; τ_p – его длительность; ω_0 – центральная частота спектра.

Зависимость времени когерентности от длины волны для импульсов разных длительностей

Для определения времени когерентности рассмотрим взаимодействие светового фемтосекундного импульса с самим собой, сдвинутым на временную задержку. Зная временную структуру поля на выходе из оптической среды при генерации спектрального суперконтинуума, можно определить интерференционный сигнал по следующей формуле:

$$I(\Delta \tau) = \int_{-\infty}^{+\infty} (E(t) + E(t + \Delta \tau))^2 dt , \qquad (3)$$

где $I(\Delta \tau)$ – зависимость интенсивности интерференционного поля от временной задержки; E(t), $E(t + \Delta \tau)$ – временная структура поля на выходе из оптической среды и то же поле, сдвинутое во времени; $\Delta \tau$ – временной сдвиг между интерферирующими импульсами.

На рис. 1–3 представлены результаты расчета интерференционного сигнала для временной структуры на выходе из кварцевого оптического волокна длиной 10 мм при генерации спектрального суперконтинуума для исходного импульса с длительностью $\tau_p = 40$ фс, пиковой интенсивностью $I = 10^{13}$ BT/см² и центральной длиной волны излучения $\lambda_0 = 800$ нм (рис. 1, область нормальной групповой дисперсии кварцевого оптического волокна), $\lambda_0 = 1260$ нм (рис. 2, область нулевой групповой дисперсии кварцевого оптического волокна), $\lambda_0 = 1560$ нм (рис. 3, область аномальной групповой дисперсии кварцевого оптического волокна), определенные по формуле (3).







Рис. 2. Интерференционный сигнал для временной структуры на выходе из кварцевого оптического волокна длиной 10 мм при генерации спектрального суперконтинуума для исходного импульса длительностью т_ρ = 40 фс, пиковой интенсивностью *I* = 10¹³ Вт/см² и центральной длиной волны излучения λ₀ = 1260 нм



Рис. 3. Интерференционный сигнал для временной структуры на выходе из кварцевого оптического волокна длиной 10 мм при генерации спектрального суперконтинуума для исходного импульса длительностью т_ρ = 40 фс, пиковой интенсивностью *I* = 10¹³ Вт/см² и центральной длиной волны излучения λ₀ = 1560 нм

Из рис. 1–3 видно, что при смещении центральной длины волны в область аномальной групповой дисперсии оптической среды время когерентности импульса со сверхшироким спектром значительно уменьшается.





На рис. 4 приведен пример графического определения времени когерентности фемтосекундного спектрального суперконтинуума. В качестве огибающей $U(\Delta \tau)$ используется экспоненциальная функция (кривая 2), накладываемая на функцию интерференционного сигнала (кривая 1) для временной структу-

ры на выходе из кварцевого оптического волокна, которую можно описать следующей формулой: $({}^{t})^{2}$

 $U(\Delta \tau) = e^{(\tau_{\text{ког}})}$, где $\tau_{\text{ког}}$ – время когерентности фемтосекундного спектрального суперконтинуума.

Определяя время когерентности таким методом для различных длин волн, получим зависимость времени когерентности от центральной длины фемтосекундного лазерного импульса на входе, представленную на рис. 5.





Заключение

Методами численного моделирования рассчитаны времена когерентности излучения спектральных суперконтинуумов, генерируемых в различных областях групповой дисперсии кварцевого оптического волокна. Проанализирована зависимость времени когерентности от центральной длины волны фемтосекундного лазера. Показано, что при увеличении центральной длины волны время когерентности значительно уменьшается, однако, в области нулевой групповой дисперсии кварцевого стекла существует скачок увеличения времени когерентности. Например, для длительности исходного импульса 40 фс на длине волны 800 нм время когерентности равно 22 фс, при увеличении длины волны время когерентности падает до 4 фс на длине волны 1180 нм. В области нулевой групповой дисперсии время когерентности резко увеличивается до 20 фс, после чего опять уменьшается, достигая минимума в 4 фс на длине волны 1560 нм. Следовательно, использование излучения спектрального суперконтинуума, генерируемого в области нормальной групповой дисперсии, является целесообразным для использования его, например, в оптической когерентной томографии, системах передачи информации, из-за высокого значения времени когерентности и квазилинейной частотной модуляции временной структуры.

Работа поддержана грантами ГК № 16.740.11.0459, ГК № соглашения 14.132.21.1392 и РФФИ № 12-02-31633.

Литература

- Dudley J.M., Coen S. Supercontinuum generation in photonic crystal fiber // Rev. Mod. Phys. 2006. V. 78. – № 4. – P. 1135–1184.
- 2. Islam M.N., Sucha G., Bar-Joseph I., Wegener M., Gordon J.P., Chemla D.S. Femtosecond distributed soliton spectrum in fibers // J. Opt. Soc. Am. B. 1989. V. 6. № 6. P. 1149–1158.
- Herrmann J., Griebner U., Zhavoronkov N., Husakou A., Nickel D., Knight J.C., Wadsworth W.J., Russell P. St. J., Korn G. Experimental evidence for supercontinuum generation by fission of higher-order solitons in photonic fibers // Phys. Rev. Lett. – 2002. – V. 88. – № 17. – P. 173901 (4 pages).
- Corwin K.L., Newbury N.R., Dudley J.M., Coen S., Diddams S.A., Washburn B.R., Weber K., Windeler R.S. Fundamental amplitude noise limitations to supercontinuum spectra generated in a microstructured fiber // Appl. Phys. B. – 2003. – V. 77. – № 2–3. – P. 269–277.
- 5. Агравал Г. Нелинейная волоконная оптика. М.: Мир, 1996. 324 с.
- 6. Беспалов В.Г., Козлов С.А., Сутягин А.Н., Шполянский Ю.А. Сверхуширение спектра интенсивных фемтосекундных лазерных импульсов и их временное сжатие до одного колебания светового поля // Оптический журнал. 1998. Т. 65. № 10. С. 85–88.
- 7. Heidt A.M. Pulse preserving flat-top supercontinuum generation in all-normal dispersion photonic crystal fibers // J. Opt. Soc. Am. B. 2010. V. 27. № 3. P. 550–559.
- Povazay B., Bizheva K., Unterhuber A., Hermann B., Sattmann H., Fercher A.F., Drexler W., Apolonski A., Wadsworth W.J., Knight J.C., Russell P.S., Vetterlein M., Scherzer E. Submicrometer axial resolution optical coherence tomography // Opt. Lett. – 2002. – V. 27. – № 20. – P. 1800–1802.

- 9. Humbert G., Wadsworth W., Leon-Saval S., Knight J., Birks T., Russell P.St.J., Lederer M., Kopf D., Wiesauer K., Breuer E., Stifter D. Supercontinuum generation system for optical coherence tomography based on tapered photonic crystal fiber // Optics Express. – 2006. – V. 14. – № 4. – P. 1596–1603.
- 10. Hartl I., Li X.D., Chudoba C., Ghanta R.K., Ko T.H., Fujimoto J.G., Ranka J.K., Windeler R.S. Ultrahighresolution optical coherence tomography using continuum generation in an air–silica microstructure optical fiber // Optics Letters. – 2001. – V. 26. – № 9. – P. 608–610.
- 11. Гуров И.П., Киракозов А.Х. Анализ методов обработки интерферометрических данных в спектральной оптической когерентной томографии // Научно-технический вестник информационных технологий, механики и оптики. 2012. № 2 (78). С. 21–24.
- Козлов С.А., Сазонов С.В. Нелинейное распространение импульсов длительностью в несколько колебаний светового поля в диэлектрических средах // Журнал экспериментальной и теоретической физики. – 1997. – Т. 111. – В. 2. – С. 404–418.
- 13. Шполянский Ю.А. Сценарии развития фемтосекундного спектрального суперконтинуума // Проблемы когерентной и нелинейной оптики / Под ред. И.П. Гурова и С.А. Козлова. СПб, 2000. С. 136–153.

Мельник Максим Владимирович	-	Санкт-Петербургский национальный исследовательский университет					
		информационных технологий, механики и оптики, студент,					
		maxim.melnick@gmail.com					
Цыпкин Антон Николаевич	_	Санкт-Петербургский национальный исследовательский университет					
		информационных технологий, механики и оптики, аспирант,					
		tsypkinan@mail.ru					

УДК 535.3, 535.42 КОЛЛИМАЦИЯ И ФОКУСИРОВКА ПАРАКСИАЛЬНОГО ВОЛНОВОГО ПАКЕТА, ПОЛУЧАЕМОГО ПРИ ДИФРАКЦИИ В ДАЛЬНЕЙ ЗОНЕ ИСХОДНО ОДНОПЕРИОДНОЙ ВОЛНЫ А.А. Дроздов, С.А. Козлов

Показано, что при коллимации параксиального волнового пакета, получаемого при дифракции в дальней зоне исходно однопериодной волны, формируется «Ж-образная» полуторапериодная пространственно-временная структура, длительность которой пропорциональна расстоянию от оси пучка, а центральная ее часть движется быстрее периферийной. С увеличением пройденного коллимированным излучением расстояния от коллимирующего до фокусирующего зеркала волновой пакет в фокусе последнего меняется от полуторапериодной волны до однопериодной и снова до полуторапериодной, максимум спектра которой на оси пучка смещается сначала в область высоких частот, а затем обратно в область низких частот. Однопериодная волна в фокусе зеркала получается идентичной исходной для коллимирующего и фокусирующего зеркал с равными фокусными расстояниями при удвоенном фокусном расстоянии между ними.

Ключевые слова: однопериодные волны, параксиальные волновые пакеты, дальняя зона дифракции, коллимация и фокусировка излучения.

Введение

С развитием фемтосекундной лазерной техники и микроэлектроники стало возможным получать предельно короткие по числу колебаний однопериодные электромагнитные волны со спектром в диапазоне 0,1–10 ТГц [1]. Такое терагерцовое излучение находит широкое применение в системах обнаружения наркотических и взрывчатых веществ, других скрываемых объектов за преградами, непрозрачными, например, в оптическом спектральном диапазоне, для медицинской диагностики и в других приложениях [2]. Важной задачей является оптимизация систем испускающих, коллимирующих и фокусирующих подобные однопериодные электромагнитные волны. Закономерности дифракции волн, испущенных источниками однопериодного излучения, к настоящему времени уже хорошо изучены (см. [3] и обзор в ней). В дальней зоне они становятся полуторапериодными. Однако работ по целенаправленному теоретическому изучению особенностей их дальнейшей динамики после коллимирования и последующей фокусировки не было. В настоящей работе теоретически изучены особенности коллимации и фокусировки сложной пространственно-временной полевой структуры, которая получается при дифракции в дальней зоне (зоне дифракции Фраунгофера) параксиального волнового пакета, испущенного исходно однопериодным источником. Показано, что коллимированное излучение имеет «Ж-образную» полуторапериодную структуру, центральная часть которой движется быстрее периферийной. Показано, что фокусирующее зеркало, поставленное сразу за коллимирующим, в своем фокусе полуторапериодное поле излучения не меняет. С увеличением пройденного коллимированным излучением расстояния от коллимирующего до фокусирующего зеркала волновой пакет в фокусе последнего меняется от полуторапериодной волны до однопериодной и снова до полуторапериодной, максимум спектра которой на оси пучка смещается сначала в область высоких частот, а затем обратно в область низких частот. При равных фокусных расстояниях коллимирующего и фокусирующего зеркал однопериодная волна в фокусе последнего получается идентичной исходной однопериодной волне (которая испускается источником, а затем трансформируется при дифракции, коллимации и фокусировке) с точностью до фазового сдвига в π при расстоянии между зеркалами, равном их удвоенному фокусному расстоянию.

Математическая модель дифракционной динамики поля

Будем рассматривать электрическое поле, испускаемое источником, в виде однопериодной волны с гауссовым поперечным распределением вида [3]

$$E(x, y, t) = E_0 \exp\left(-\frac{x^2 + y^2}{\rho^2}\right) \frac{t}{\tau} \exp\left(-\frac{t^2}{\tau^2}\right),\tag{1}$$

где E_0 – амплитуда входного поля, а ρ и τ – исходные поперечный размер волнового пакета и его длительность. Поле (1) проиллюстрировано на рис. 1, а. Светло-серым участкам плоскостных изображений соответствуют максимальные положительные значения поля, а темно-серым – максимальные отрицательные. На иллюстрации $\rho/\lambda_0 = 10$, где $\lambda_0 = \sqrt{2\pi}c\tau$ [4] – длина волны излучения, соответствующая максимуму спектра исходного волнового пакета; *с* – скорость света в вакууме.

Параксиальная однонаправленная динамика электрического поля *E* в прозрачной оптической среде, дисперсией которой можно пренебречь (например, в воздухе), может быть описана уравнением [5, 6]

$$\frac{\partial E}{\partial z} + \frac{N_0}{c} \cdot \frac{\partial E}{\partial t} = \frac{c}{2N_0} \Delta_{\perp} \int_{-\infty}^{t} Edt', \qquad (2)$$

где z – направление, вдоль которого распространяется волновой пакет; N_0 – линейный показатель пре-

ломления среды;
$$\Delta_{\perp} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$
 – поперечный лапласиан; *t* – время.



Рис. 1. Пространственно-временная структура распределения электрического поля на входе в среду (1) (а), а также ее структура в дальней зоне дифракции [3] (б)

В дальней зоне дифракции (при $z \gg \frac{\pi \rho^2 N_0}{\{\lambda\}_{\min}}$, где $\{\lambda\}_{\min}$ – минимальная длина волны из диапазона

длин волн, в котором находится практическая часть энергии исходного излучения [3]), пространственновременное распределение поля из исходного (1) принимает вид [3]

$$E(t', x, y, z) = E_0 A^3(x, y, z) \frac{T(z)}{\tau} \left[1 - 2 \left(A(x, y, z) \frac{t'}{\tau} \right)^2 \right] \exp\left(- \left(A(x, y, z) \frac{t'}{\tau} \right)^2 \right),$$
(3)

где $A(x, y, z) = \left(1 + \left(\frac{2T(z)}{\tau}\right)^2 \frac{x^2 + y^2}{\rho^2}\right)^2$, $T(z) = \frac{\rho^2 N_0}{2cz}$ и «запаздывающее», в том числе вследствие кри-

визны волнового фронта, время $t' = t - \frac{N_0}{c} \left(z + \frac{x^2 + y^2}{2z} \right)$. Поле (3) иллюстрировано на рис. 1, б. Полагали

 $z = z_0 = 2 \frac{\pi \rho^2 N_0}{\lambda_0}$. Исходя из этого, например, при характерных для терагерцовых импульсов начальных параметрах $\rho = 3$ мм и $\lambda_0 = 0,3$ мм, поле на рис. 1, б, иллюстрировано для расстояния от источника

 $z_0 = 20$ см. Как видно из рисунка, волна исходного вида (1) в дальней зоне дифракции становится полуторапериодной по всему волновому пакету. Длинноволновая часть излучения на периферии пучка распространяется медленнее, чем коротковолновая на оси пучка.

Рассмотрим теперь, как волновой пакет (3), получаемый при дифракции в дальней зоне исходно однопериодной волны (1), будет эволюционировать после коллимирующего зеркала с фокусным расстоянием $f_{\text{кол.}}$.

Динамика коллимированного параксиального волнового пакета

Поставим в дальней зоне дифракции излучения на расстоянии z_0 от источника коллимирующее зеркало с фокусным расстоянием $f_{\text{кол.}} = z_0$ и функцией отражения $R(x, y) = \exp(-ik(x^2 + y^2)/2f_{\text{кол.}})$ [7], где $k = N_0 \omega/c$. Тогда поле волнового пакета сразу за коллимирующим зеркалом примет вид

$$E_{_{\text{KOI.}}}\left(t'', x, y, z_{_{0}}\right) = \frac{N_{_{0}}\rho^{^{2}}}{2cz_{_{0}}}E_{_{0}}\frac{\tau^{^{2}}}{\tau^{^{3}}_{_{\text{KOI.}}}}\left(1 - 2\frac{t''^{^{2}}}{\tau^{^{2}}_{_{\text{KOI.}}}}\right)\exp\left(-\frac{t''^{^{2}}}{\tau^{^{2}}_{_{\text{KOI.}}}}\right),\tag{4}$$

где $t'' = t - \frac{N_0 z_0}{c}$; $\tau_{\text{кол.}}^2 = \tau^2 + \left(\frac{\rho N_0}{c}\right)^2 \frac{x^2 + y^2}{z_0^2}$. Это поле иллюстрировано на рис. 2, а. Как следует из фор-

мулы (4) и видно из рисунка, сразу за зеркалом формируется «Ж-образная» полуторапериодная пространственно-временная полевая структура, длительность которой $\tau_{\text{кол.}}$ увеличивается с расстоянием от оси пучка.

Будем теперь рассматривать выражение (4) как новое граничное условие динамики коллимированного поля излучения. На рис. 2, б–д, приведены результаты численного расчета по уравнению (2) с граничным условием (4) эволюции пространственно-временной полевой структуры волнового пакета на разных расстояниях от коллимирующего зеркала в оптической среде. Из рисунка видно, что из-за дифракции центральная часть «Ж-образного» волнового пакета движется быстрее периферийной.



Рис. 2. Пространственно-временная структура распределения электрического поля коллимированной волны (4) на расстоянии $z_0 = 0$ от коллимирующего зеркала (а) и ее дальнейшая эволюция в оптической среде на расстояниях: $f_{\text{кол.}}$ (б); $2f_{\text{кол.}}$ (в); $3f_{\text{кол.}}$ (г); $6f_{\text{кол.}}$ (д)

Теперь рассмотрим особенности фокусировки коллимированного излучения, представленного на рис. 2, фокусирующим зеркалом с фокусным расстоянием $f_{\phi o \kappa} = f_{\kappa o \pi}$.

Динамика сфокусированного волнового пакета

На рис. 3 представлены пространственно-временные распределения электрического поля излучения в фокусе зеркала в зависимости от пройденного коллимированным волновым пакетом расстояния от коллимирующего до фокусирующего зеркала. Из рис. 3, а, видно, что при фокусировке полуторапериодной «Ж-образной» полевой структуры, полученной сразу за коллимирующим зеркалом, в фокусе зеркала число колебаний полученного волнового пакета не меняется. Из рис. 3, б–д, видно, что с увеличением пройденного коллимированным излучением расстояния от коллимирующего до фокусирующего зеркала волновой пакет в фокусе последнего меняется от полуторапериодной волны до однопериодной и снова до полуторапериодной.



Рис. 3. Пространственно-временные структуры распределения электрического поля излучения в фокусе зеркала при фокусировке волнового пакета, полученного сразу за коллимирующим зеркалом (а), и на расстояниях до него $f_{\phi o \kappa}$ (б); $2f_{\phi o \kappa}$. (в); $3f_{\phi o \kappa}$. (г); $6f_{\phi o \kappa}$. (д)



Рис. 4. Нормированные профили электрического поля (слева) и квадраты модулей спектральной плотности (справа) на оси пучка в фокусе (жирные линии) при фокусировке волнового пакета, полученного сразу за коллимирующим зеркалом (а), и на расстояниях от него: $f_{\phi o \kappa}$ (б); $2f_{\phi o \kappa}$. (в); $3f_{\phi o \kappa}$.



На рис. 4 дополнительно проиллюстрированы временные профили электрических полей (нормированные на максимум напряженности электрического поля источника E_0) и квадраты модулей спектров (нормированные на максимум квадрата модуля спектральной плотности источника $|G_0|^2$) в фокусе на оси волнового пакета. На рисунке частота $\Omega = \omega/\omega_{max}$ нормирована на частоту $\omega_{max} = \sqrt{2}/\tau$ [4, 8], соответствующую максимум квадрата модуля спектра исходной однопериодной волны источника. Из рисунка видно, что при расстоянии между коллимирующим и фокусирующим зеркалом, равном их двойному фокусному расстоянию, поле с точностью до фазового сдвига в π и его спектр совпадают с полем и спек-

тром однопериодной волны, исходно испущенной источником излучения. При меньших или больших расстояниях между зеркалами сфокусированная волна становится при малых отклонениях слабо выраженной, а при больших – ярко выраженной полуторапериодной со спектром, смещенным в высокочастотную область. Максимумы поля и спектра при этом растут, что обусловлено более глубокой фокусировкой излучения (см. также рис. 3).

Заключение

В работе теоретически изучены особенности коллимации и фокусировки сложной пространственно-временной полевой структуры, получаемой при дифракции в дальней зоне параксиального исходно однопериодного волнового пакета. Показано, что при коллимации такого пакета формируется «Жобразная» полуторапериодная волна, длительность которой пропорциональна расстоянию от оси пучка, а центральная ее часть движется быстрее периферийной. С увеличением пройденного коллимированным излучением расстояния от коллимирующего до фокусирующего зеркала волновой пакет в фокусе последнего меняется от полуторапериодной волны до однопериодной и снова до полуторапериодной, максимум спектра которой на оси пучка смещается сначала в область высоких частот, а затем обратно в область низких частот.

При равных фокусных расстояниях коллимирующего и фокусирующего зеркал однопериодная волна в фокусе последнего получается идентичной исходной однопериодной волне источника с точностью до фазового сдвига в π при расстоянии между зеркалами, равном их удвоенному фокусному расстоянию.

Работа поддержана грантом РФФИ 11-02-01346-а и ГК № 14.512.11.0020.

Литература

- 1. Беспалов В.Г., Городецкий А.А., Денисюк И.Ю., Козлов С.А., Крылов В.Н., Лукомский Г.В., Петров Н.В., Путилин С.Э. Методы генерации сверхширокополосных ТГц импульсов фемтосекундными лазерами // Оптический журнал. – 2008. – Т. 75. – № 10. – С. 34–41.
- Lee Y.S. Principles of Terahertz Science and Technology. Corvalis: Springer Science+Business Media, 2009. – 347 p.
- 3. Ezerskaya A.A., Ivanov D.V., Kozlov S.A., Kivshar Yu.S. Spectral Approach in the Analysis of Pulsed Terahertz Radiation // J. Infrared Milli. Terahz. Waves. 2012. V. 33. № 9. P. 926–942.
- Drozdov A.A., Kozlov S.A., Sukhorukov A.A., Kivshar Yu.S. Self-phase modulation and frequency generation with few-cycle optical pulses in nonlinear dispersive media // Phys. Rev. A. 2012. V. 86. № 5. P. 053822–10.
- 5. Беленов Э.М., Назаркин А.В. О некоторых решениях уравнений нелинейной оптики без приближения медленно меняющихся амплитуд и фаз // Письма в ЖЭТФ. 1990. Т. 51. № 5. С. 252–255.
- 6. Козлов С.А., Сазонов С.В. Нелинейное распространение импульсов длительностью в несколько колебаний светового поля в диэлектрических средах // ЖЭТФ. – 1997. – Т. 111. – № 2. – С. 404–418.
- 7. Папулис А. Теория систем и преобразований в оптике. М.: Мир, 1971. 495 с.
- 8. Дроздов А.А., Козлов С.А. Фазовая самомодуляция однопериодных оптических волн // Научнотехнический вестник СПбГУ ИТМО. – 2011. – № 2 (72). – С. 99–105.

Дроздов Аркадий Анатольевич	-	Санкт-Петербургский национальный исследовательский университет ин-				
		формационных технологий, механики и оптики, аспирант,				
		arkadiy.drozdov@gmail.com				
Козлов Сергей Аркадьевич	-	Санкт-Петербургский национальный исследовательский университет ин-				
		формационных технологий, механики и оптики, доктор физмат. наук,				
		профессор, декан, заведующий кафедрой, kozlov@mail.ifmo.ru				

3

АНАЛИЗ И СИНТЕЗ СЛОЖНЫХ СИСТЕМ

УДК 519.7 РОБАСТНОЕ СУБОПТИМАЛЬНОЕ УПРАВЛЕНИЕ БОКОВЫМ ДВИЖЕНИЕМ ЛЕТАТЕЛЬНОГО АППАРАТА В РЕЖИМЕ ЗАХОДА НА ПОСАДКУ¹ И.Б. Фуртат

Рассматривается робастное субоптимальное управление боковым движением летательного аппарата в режиме захода на посадку. Разработан алгоритм, обеспечивающий субминимизацию интегрального критерия качества в условиях параметрических возмущений, действующих на модель летательного аппарата. Приведены результаты моделирования, иллюстрирующие эффективность используемого алгоритма.

Ключевые слова: летательный аппарат, робастное управление, оптимальное управление.

Введение

В настоящее время в классе задач адаптивного и робастного управления существует большое количество решений. В основной части из них [1–3] целью управления является выполнение предельного целевого неравенства. Однако на практике, как правило, дополнительно предъявляются требования к обеспечению качества переходных процессов в системе управления.

В условиях определенности модели объекта на сегодняшний день в классической теории управления разработано большое количество методов и подходов [4–7]. Задача оптимального и субоптимального управления объектами, подверженными воздействию неизвестных внешних возмущений, рассмотрена в [8]. В работе [9] подробно рассматривается задача адаптивно-оптимального и субоптимального управления, где для параметрически неопределенных объектов предложены модификации метода динамического программирования, принципа максимума Понтрягина, вариационного метода и т.д.

Одним из наиболее распространенных подходов оптимального управления неопределенными объектами является метод H_{∞} -оптимизации [5, 10]. Однако применение методов H_{∞} -оптимизации приводят к регуляторам с высоким динамическим порядком, которые к тому же могут быть нереализуемыми. Как отмечено в [5], задача H_{∞} -оптимизации существенно упрощается, если для объекта управления можно построить сверхстабилизирующий регулятор. Но класс таких объектов слишком узок. Метод H_{∞} -оптимизации был применен в [10] для управления продольным движением самолета. В [10] также приведены результаты использования H_2 регулятора и анизотропного регулятора для управления самолетом.

Наиболее близкой по идее к настоящей работе является работа [11]. В [11] предполагалось, что неизвестные матрицы линейного объекта управления можно представить в виде суммы номинальных и неизвестных матриц. Схема управления включает в себя два контура: основной контур, предназначенный для оптимального управления номинальным объектом, и контур параллельной обратной связи, компенсирующей функциональную и параметрическую неопределенность объекта. Теоретические результаты были применены к управлению боковым движением самолета. В [12] с использованием подходов [3] и [11] решена задача субоптимального управления нестационарными динамическими объектами в условиях параметрических и внешних возмущений. Полученный в [12] алгоритм прост и имеет невысокий динамический порядок. Результат [12] был распространен для управления многосвязными объектами в [13].

В настоящей работе предлагается робастное субоптимальное управление боковым движением летательного аппарата в режиме захода на посадку. Динамическая модель объекта подвержена действию параметрических возмущений. Целью управления является субминимизация интегрального критерия качества с бесконечной верхней границей. Формирование управляющего сигнала и критерия качества осуществляется аналогично [11]. Компенсация неопределенностей основана на методе вспомогательного контура [3], развитого для решения задач субоптимального управления в [12, 13]. Приведены результаты численного моделирования, иллюстрирующие эффективность предложенной схемы.

Постановка задачи

Рассмотрим модель бокового движения летательного аппарата в режиме захода на посадку: $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t)$,

где $\mathbf{x}(t) = \left[\Delta z(t), \Delta \psi(t), \Delta \gamma(t), \Delta \omega_x(t)\right]^T$, $u(t) = \delta_{\mathfrak{H}}(t)$ – вектор состояния и сигнал управления по линеари-

(1)

¹ Исследование выполнено при поддержке Министерства образования и науки Российской Федерации, соглашение № 14.В37.21.0421 «Разработка автономной бортовой системы навигации и управления многофункциональными мультиротационными летательными аппаратами», а также в рамках ФЦП «Научные и научно-педагогические кадры инновационной России на 2009–2013 г.г.», соглашение № 8846, и гранта РФФИ № 12-01-31354.

зованной математической модели движения летательного аппарата без учета скольжения, $\Delta z(t)$ – величина бокового отклонения центра масс летательного аппарата от продольной оси взлетно-посадочной полосы, $\Delta \psi(t)$ – угол между продольной осью взлетно-посадочной полосы и горизонтальной проекцией вектора скорости летательного аппарата, $\Delta \gamma(t)$ – изменение угла крена летательного аппарата, $\Delta \omega_x(t)$ – изменение угловой скорости вращения летательного аппарата относительно его продольной оси, $\delta_{\Im}(t)$ – отклонение элеронов от балансировочного положения; $\mathbf{A}=\mathbf{A}_{\mathrm{H}}+\Delta\mathbf{A}$; $\mathbf{B}=\mathbf{B}_{\mathrm{H}}+\Delta\mathbf{B}$. Индексом «Н» будем обозначать номинальные матрицы, т.е. матрицы с известными элементами. Так, \mathbf{A}_{H} и вектор \mathbf{B}_{H} линеаризованной модели при скорости летательного аппарата 85 м/с характеризуются следующими числовыми значениями [14]:

$$\mathbf{A}_{\mathrm{H}} = \begin{bmatrix} 0 & 85 & 0 & 0 \\ 0 & 0 & 0, 12 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -2 \end{bmatrix}, \quad \mathbf{B}_{\mathrm{H}} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 3, 4 \end{bmatrix}.$$
(2)

Возмущения всех элементов матрицы A_H и коэффициентов вектора B_H должны быть согласованы между собой. Исходя из этого, параметрические возмущения зададим в виде [11]

где ρ – произвольное действительное число, 9 – произвольное положительное действительное число. Целью управления является синтез алгоритмической структуры управляющего устройства, обес-

$$J = \int_{0}^{\infty} \left[\mathbf{x}^{T}(t) \mathbf{Q} \mathbf{x}(t) + r u_{0}^{2}(t) \right] dt , \qquad (4)$$

где весовая матрица и коэффициент целевого функционала выбраны в виде [11]

печивающей субминимизацию функционального критерия качества

$$\mathbf{Q} = \begin{bmatrix} 6, 25 \cdot 10^{-6} & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 31 \end{bmatrix}, \ r = 93,$$
(5)

 $u_0(t)$ – сигнал оптимального управления. Аналогичная цель управления формируется в [11] для управления неопределенными объектами. Отличие данной работы состоит в формировании сигнала управления $u_k(t)$, необходимого для компенсации параметрических возмущений.

Метод решения

Принимая во внимание (2) и (3), преобразуем уравнение (1) к виду $\dot{\mathbf{x}}(t) = \mathbf{A}_{\mathrm{H}} \mathbf{x}(t) + \mathbf{B}_{\mathrm{H}} u(t) + \mathbf{B}_{\mathrm{H}} \phi(x, u)$,

(6)

где $\phi(\mathbf{x}, u) = \begin{bmatrix} 0 & 0 & \rho \end{bmatrix} \mathbf{x}(t) + \vartheta u(t)$. Следуя [11], представим сигнал управления u(t) в виде суммы $u(t) = u_0(t) + u_k(t)$

и выделим в (6) номинальный объект управления

$$\mathbf{x}(t) = \mathbf{A}_{\mathrm{H}} \mathbf{x}(t) + \mathbf{B}_{\mathrm{H}} u_0(t) .$$
⁽⁷⁾

Решим для (7) задачу оптимального управления с минимизацией критерия качества (4). Согласно [4–7], для объекта (7) стационарный оптимальный закон управления определится в виде

 $u_0(t) = -\mathbf{K}_0 \mathbf{x}(t)$, (8) где $\mathbf{K}_0 = r^{-1} \mathbf{B}_{\mathrm{H}}^T \mathbf{H}$, матрица $\mathbf{H} = \mathbf{H}^T > 0$ является решением матричного алгебраического уравнения Лурье–

Риккати $\mathbf{A}_{\mathrm{H}}^{T}\mathbf{H} + \mathbf{H}\mathbf{A}_{\mathrm{H}} - \mathbf{H}\mathbf{B}_{\mathrm{H}}\mathbf{R}^{-1}\mathbf{B}_{\mathrm{H}}^{T}\mathbf{H} = -\mathbf{Q}$. Для матриц вида (2) и (5) матрицы **H** и **K**₀ определятся в виде

0,0001	0,0647	0,0225	0,0071		0,0003	
0,0647	77,2913	31,7923	10,2230	V	0,3737	
0,0225	31,7923	26,6652	8,8020	, n =	0,3218	•
0,0071	10,2230	8,8020	9,4033		0,3438	
	0,0001 0,0647 0,0225 0,0071	0,00010,06470,064777,29130,022531,79230,007110,2230	0,00010,06470,02250,064777,291331,79230,022531,792326,66520,007110,22308,8020	0,0001 0,0647 0,0225 0,0071 0,0647 77,2913 31,7923 10,2230 0,0225 31,7923 26,6652 8,8020 0,0071 10,2230 8,8020 9,4033	$\begin{bmatrix} 0,0001 & 0,0647 & 0,0225 & 0,0071 \\ 0,0647 & 77,2913 & 31,7923 & 10,2230 \\ 0,0225 & 31,7923 & 26,6652 & 8,8020 \\ 0,0071 & 10,2230 & 8,8020 & 9,4033 \end{bmatrix}, \mathbf{K} =$	$\begin{bmatrix} 0,0001 & 0,0647 & 0,0225 & 0,0071 \\ 0,0647 & 77,2913 & 31,7923 & 10,2230 \\ 0,0225 & 31,7923 & 26,6652 & 8,8020 \\ 0,0071 & 10,2230 & 8,8020 & 9,4033 \end{bmatrix}, \mathbf{K} = \begin{bmatrix} 0,0003 \\ 0,3737 \\ 0,3218 \\ 0,3438 \end{bmatrix}$

С учетом оптимального управления (8) перепишем уравнение (6) в виде

$$\dot{\mathbf{x}}(t) = \mathbf{A}_0 \mathbf{x}(t) + \mathbf{B}_{\mathrm{H}} u_k(t) + \mathbf{B}_{\mathrm{H}} \varphi_1(\mathbf{x}, u) ,$$

где

$$\mathbf{A}_{0} = \mathbf{A}_{\mathrm{H}} - \mathbf{B}_{\mathrm{H}} \mathbf{K}_{0} = \begin{bmatrix} 0 & 85 & 0 & 0 \\ 0 & 0 & 0,12 & 0 \\ 0 & 0 & 0 & 1 \\ -0,0009 & -1,2707 & -1,0941 & -3,1688 \end{bmatrix}, \ \varphi_{1}(\mathbf{x}, u) = \varphi(\mathbf{x}, u) + \mathbf{K}_{0} \mathbf{x}(t) \,.$$

Согласно [3], для выделения возмущений введем вспомогательный контур в виде [12, 13] $\dot{\mathbf{x}}_{v}(t) = \mathbf{A}_{0}\mathbf{x}_{v}(t) + \mathbf{B}_{H}u_{k}(t), \ \mathbf{x}_{v}(0) = \mathbf{x}(0),$ (10)

где $\mathbf{x}_{v}(t) \in \mathbb{R}^{4}$ – вектор состояния. Принимая во внимание (9) и (10), рассмотрим уравнение для рассогласования $\xi(t) = \mathbf{x}(t) - \mathbf{x}_{v}(t)$:

$$\dot{\boldsymbol{\xi}}(t) = \mathbf{A}_0 \,\boldsymbol{\xi}(t) + \mathbf{B}_{\mathrm{H}} \boldsymbol{\varphi}_1(\mathbf{x}, u) \,. \tag{11}$$

Перепишем (11) в виде системы дифференциальных уравнений

$$\dot{\xi}_{1} = 85\xi_{2},
\dot{\xi}_{2} = 0,12\xi_{3},
\dot{\xi}_{3} = \xi_{4},$$
(12)

 $\dot{\xi}_4 = \begin{bmatrix} -0,0009 & -1,2707 & -1,0941 & -3,1688 \end{bmatrix}^T \xi + 3,4\varphi_1(x,u).$

Из четвертого уравнения (12) сигнал $\phi_1(x, u)$ можно определить как

$$\varphi_1(\mathbf{x}, u) = 3, 4^{-1} \left(\dot{\xi}_4(t) - \begin{bmatrix} -0,0009 & -1,2707 & -1,0941 & -3,1688 \end{bmatrix}^T \xi(t) \right).$$

Для оценки сигнала $\dot{\xi}_4(t)$ воспользуемся наблюдателем

$$\dot{\bar{\xi}}_4(t) = \frac{p}{\mu p + 1} \xi_4(t) , \qquad (13)$$

где $\mu > 0$ – достаточно малое число. Тогда из (9) видно, что для компенсации $\phi_1(x, u)$ закон управления $u_k(t)$ можно сформировать в виде

$$u_k(t) = -3, 4^{-1} \left(\frac{\dot{\xi}}{\xi_4}(t) - \begin{bmatrix} -0,0009 & -1,2707 & -1,0941 & -3,1688 \end{bmatrix}^T \xi(t) \right).$$
(14)

В результате сформирована система управления боковым движением летательного аппарата в режиме захода на посадку, представленная уравнениями (8), (10), (13), (14). Доказательство работоспособности системы управления (8), (10), (13), (14) в общем виде рассмотрено в [12], поэтому здесь не приводится.

Рассмотрим возмущения в (1), аналогичные [11]. Пусть в (3) $\rho = -4$ и $\vartheta = 1,7$. Последнее означает, что эффективность элеронов уменьшилась вдвое. Начальное отклонение $\Delta z(0) = 400$ м.



Рис. 1. Графики переходных процессов бокового отклонения центра масс летательного аппарата

Положим в (13) $\mu = 0,01$. На рис. 1 показан переходный процесс по величине бокового отклонения центра масс летательного аппарата от продольной оси взлетно-посадочной полосы $\Delta z(t)$ для возмущенной системы (1) и $\Delta z_{\rm H}(t)$ для номинального объекта (7), на рис. 2 – переходный процесс по отклонению

(9)

элеронов от балансировочного положения $\delta_{\Im}(t)$ для (1) и $\delta_{\Im onr}(t)$, где $\delta_{\Im onr}(t)$ – оптимальное управление для (7). Еще раз отметим, что при моделировании на возмущенную систему (1) подана сумма сигналов (8) и (14), на номинальный объект (7) подан только сигнал (8).



Рис. 2. Графики расходов рулей

Рис. 2 показывает, что из-за наличия параметрической неопределенности увеличивается расход рулей. По результатам моделирования также можно сказать, что для неопределенного объекта (1) значение интегрального критерия качества (4) равно 18,377. Для номинального объекта (7) значение критерия качества равно 18,998. Абсолютная погрешность равна 0,621. Другими словами, предложенный алгоритм робастен к наличию параметрических возмущений. Сравнительный анализ полученных результатов и данных [11, 15] показал, что предложенная в работе система управления проста в расчете и реализации, а погрешность при минимизации интегрального критерия качества может быть уменьшена за счет уменьшения параметра µ в (13).

Заключение

В работе предложен синтез робастного субоптимального закона управления боковым движением летательного аппарата в режиме захода на посадку. Полагается, что модель летательного аппарата описывается параметрически неопределенным линейным дифференциальным уравнением четвертого порядка. Для оптимального управления неопределенным объектом, в соответствии с [11], предложено закон регулирования рассматривать как сумму сигнала оптимального управления номинальным объектом и сигнала компенсации неопределенностей в исходном объекте. В свою очередь, формирование сигнала компенсации основано на идее подхода [3]. Численное моделирование показало хорошую робастность по отношению к параметрическим возмущениям.

Литература

- 1. Мирошник И.В., Никифоров В.О., Фрадков А.Л. Нелинейное и адаптивное управление сложными динамическими системами. СПб: Наука, 2000. 549 с.
- 2. Фуртат И.Б. Алгоритм субынвариантного управления по выходу линейным структурно неопределенным динамическим объектом // Изв. вузов. Приборостроение. 2011. № 9 (54). С. 22–27.
- 3. Цыкунов А.М. Алгоритмы робастного управления с компенсацией ограниченных возмущений // Автоматика и телемеханика. 2007. № 7. С. 103–115.
- 4. Афанасьев В.Н., Колмановский В.Б., Носов В.Р. Математическая теория конструирования систем управления. М.: Высшая школа, 2003. 614 с.
- 5. Поляк Б.Т., Щербаков П.С. Робастная устойчивость и управление. М.: Наука, 2002. 303 с.
- Методы классической и современной теории автоматического управления. Теория оптимизации автоматического управления / Под ред. К.А. Пупкова и Н.Д. Егупова. М.: Издательство МГТУ им. Н.Э. Баумана, 2004. Т. 4.– 744 с.
- 7. Воронов А.А., Ким Д.П., Лохин В.М. и др. Теория автоматического управления. Теория нелинейных и специальных систем автоматического управления. М.: Высшая школа, 1986. Ч. 2. 504 с.
- 8. Сиван Р., Квакернаак Х. Линейные оптимальные системы управления. М.: Мир, 1977. 653 с.
- 9. Тертычный-Даури В.Ю. Адаптивная механика. М.: Факториал Пресс, 2003. 464 с.
- Методы классической и современной теории автоматического управления. Синтез регуляторов систем автоматического управления / Под ред. К.А. Пупкова и Н.Д. Егупова М.: Издательство МГТУ им. Н.Э. Баумана. 2004. Т. 3. 616 с.

- 11. Буков В.Н. Вложение систем. Аналитический подход к анализу и синтезу матричных систем. Калуга: Изд-во научной литературы Н.Ф. Бочкаревой, 2006. – 720 с.
- 12. Фуртат И.Б. Робастное субоптимальное управление линейными нестационарными объектами по выходу // Мехатроника, автоматизация, управление. – 2009. – № 7. – С. 7–12.
- 13. Фуртат И.Б. Децентрализованное субоптимального управление по быстродействию многосвязными объектами с запаздыванием по состоянию в системах АСУ // Промышленные АСУ и контроллеры. 2010. № 3. С. 16–19.
- 14. Летов А.М. Динамика полета и управление. М.: Наука, 1969. 360 с.
- 15. Клебан В.О., Шалыто А.А. Разработка системы управления малоразмерным вертолетом // Научнотехнический вестник СПбГУ ИТМО. – 2011. – № 2 (72). – С. 12–15.
- *Фуртат Игорь Борисович* Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат технических наук, доцент; Институт проблем машиноведения РАН, ст. научный сотрудник, cainenash@mail.ru

УДК 62.50: 681.5.01 ИССЛЕДОВАНИЕ КОЛЕБАТЕЛЬНОСТИ ПРОЦЕССОВ В АПЕРИОДИЧЕСКИХ НЕПРЕРЫВНЫХ СИСТЕМАХ, ПОРОЖДАЕМОЙ ФАКТОРОМ КРАТНОСТИ СОБСТВЕННЫХ ЧИСЕЛ

Т.А. Акунов, Н.А. Дударенко, Н.А. Полинова, А.В. Ушаков

Рассматривается устойчивая апериодическая непрерывная система, матрица состояния которой обладает вещественным спектром кратных собственных чисел, кратность которых равна размерности ее вектора состояния. Показывается, что если модуль собственного числа меньше единицы, то в свободном движении системы по норме вектора состояния обнаруживается колебательность, проявляющаяся в наличии начального выброса, сменяющегося монотонным движением к состоянию покоя. Установлено, что величина выброса тем больше, чем меньше по модулю собственное число и больше его кратность.

Ключевые слова: вещественные собственные числа, кратность, свободное движение, норма, выброс.

Введение. Постановка задачи

Ставится задача исследования свободного движения устойчивой линейной непрерывной многомерной динамической системы по норме вектора ее состояния с целью изучения влияния на это поведение кратности собственных чисел ее матрицы состояния и значения их модуля. В настоящей работе поставленная задача решается для случая вещественных кратных собственных чисел. Более того, предполагается, что кратность собственного числа равна размерности вектора состояния. Как будет показано, приходится констатировать системное явление, состоящее в том, что в апериодической системе (системе с вещественными собственными числами) при кратности собственных чисел больше единицы и значениях модуля собственных чисел меньших единицы возникает возможность появления заметных выбросов нормы вектора состояния в свободном движении. Обнаруживается, что величина выброса растет с уменьшением модуля собственных чисел и с увеличением их кратности. Более того, появляется возможность «обменивать» модуль собственных чисел на их кратность в классе систем с фиксированным значением выброса. Первоначально задача решается для случая представления матрицы состояния в канонической жордановой форме [1], затем исследования переносятся на произвольный случай.

Аналитическое исследование свободного движения непрерывной многомерной апериодической системы для случая кратных вещественных собственных значений ее матрицы состояния

Рассмотрим линейную гурвицеву непрерывную многомерную динамическую систему, задаваемую [2, 3] в векторно-матричной форме

$$\dot{\mathbf{x}}(t) = \mathbf{F}\mathbf{x}(t), \mathbf{x}(t)\Big|_{t=0} = \mathbf{x}(0), \qquad (1)$$

где $\mathbf{x}(0), \mathbf{x}(t)$ – вектора соответственно начального и текущего состояний системы; \mathbf{F} – ее матрица состояния; $\mathbf{x}(0), \mathbf{x}(t) \in \mathbb{R}^n; \mathbf{F} \in \mathbb{R}^{n \times n}$. Матрица системы \mathbf{F} , заданная в произвольном базисе, такова, что ее характеристический полином $D(\lambda)$ имеет представление

$$D(\lambda) = \det(\lambda \mathbf{I} - \mathbf{F}) = \left\{ (\lambda - \alpha)^n = \lambda^n + \sum_{i=1}^n (-1)^i C_n^i \alpha^i \lambda^{n-i}; \alpha : Jm(\alpha) = 0 \right\}.$$
 (2)

Такая ситуация может возникнуть, когда при синтезе методами модального управления [3] матрица состояния F системы задается во фробениусовой форме, сопровождающей характеристический полином $D(\lambda)$, в котором C_n^i – число сочетаний из n по i. Дополним условие (2) наличия в алгебраическом спектре собственных чисел $\sigma\{\mathbf{F}\} = \{\lambda_i = \arg[\det(\lambda \mathbf{I} - \mathbf{F}) = 0]: \lambda_i = \alpha; i = \overline{1, n}\}$ матрицы \mathbf{F} единственного вещественного элемента кратности $\mu = n = \dim(x)$ условием, накладываемым на дефект характеристической матрицы ($\lambda \mathbf{I} - \mathbf{F}$) [1] матрицы \mathbf{F} , который должен принимать единичное значение.

Тогда [1] каноническая форма матрицы, построенная на спектре σ {**F**} собственных чисел матрицы **F**, будет представлять собой (*n*×*n*)-клетку Жордана *J*(α) вида

$$\mathbf{J}(\alpha) = \begin{bmatrix} \alpha & 1 & 0 & \dots & 0 \\ 0 & \alpha & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \alpha \end{bmatrix}.$$
 (3)

Следует заметить, что матрица в форме Жордана $J(\alpha)$ порождает автономную динамическую систему вида (1), задаваемую в жордановом каноническом базисе,

$$\dot{\tilde{\mathbf{x}}}(t) = \mathbf{J}(\alpha) \,\tilde{\mathbf{x}}(t), \,\tilde{\mathbf{x}}(t)\Big|_{t=0} = \tilde{\mathbf{x}}(0), \tag{4}$$

в которой вектор $\tilde{\mathbf{x}}$ и матрица $\mathbf{J}(\alpha)$ состояния связаны с вектором \mathbf{x} и соответственно с матрицей \mathbf{F} состояния исходной системы (1) векторно-матричными соотношениями

$$\mathbf{x} = \mathbf{S}\tilde{\mathbf{x}}, \, \mathbf{S}\mathbf{J}(\alpha) = \mathbf{F}\mathbf{S} \,. \tag{5}$$

В (5) S – $(n \times n)$ -матрица неособого преобразования подобия, допускающая представление матрицы F в форме

$$\mathbf{F} = \mathbf{S}\mathbf{J}(\alpha)\mathbf{S}^{-1}.$$
 (6)

В свою очередь, жорданова матрица $J(\alpha)$ в силу (3) может быть представлена в аддитивно декомпозированном виде

$$\mathbf{J}(\alpha) = diag\left\{\lambda_i = \alpha; i = \overline{\mathbf{I}, n}\right\} + \mathbf{J}(0) = \alpha \mathbf{I} + \mathbf{J}(0),$$
(7)

где J(0) – нильпотентная матрица [1] индекса v = n.

Теперь поставим задачу исследования свободного движения системы (4) по вектору ее состояния в скаляризованной форме. Решение системы (4) $\tilde{\mathbf{x}}(t) = \tilde{\mathbf{x}}(t, \tilde{\mathbf{x}}(0))$ имеет [1–3] вид

$$\tilde{\mathbf{x}}(t) = \tilde{\mathbf{x}}(t, \tilde{\mathbf{x}}(0)) = \exp\{\mathbf{J}(\alpha)t\}\tilde{\mathbf{x}}(0).$$
(8)

Скаляризацию векторного процесса (8) осуществим на основе использования согласованных [1] векторных и матричных норм, в результате чего на основе (7) получим цепочку соотношений

$$\left\|\tilde{\mathbf{x}}(t)\right\| = \left\|\exp\left\{\mathbf{J}(\alpha)t\right\}\tilde{\mathbf{x}}(0)\right\| \le \left\|\exp\left\{\mathbf{J}(\alpha)t\right\}\right\| \cdot \left\|\tilde{\mathbf{x}}(0)\right\| = e^{\alpha t} \left\|\exp\left\{\mathbf{J}(0)t\right\}\right\| \cdot \left\|\tilde{\mathbf{x}}(0)\right\|.$$
(9)

В (9) компонент $\exp \{ \mathbf{J}(0) t \}$ мультипликативной цепи элементов имеет [1–3] при $\mu = n$ представление

$$\exp\left\{\mathbf{J}(0)t\right\} = \exp\left\{\begin{bmatrix} 0 & 1 & 0 & \dots & 0\\ 0 & 0 & 1 & \dots & 0\\ \vdots & \vdots & \vdots & \dots & \vdots\\ 0 & 0 & 0 & \dots & 1\\ 0 & 0 & 0 & \dots & 0 \end{bmatrix} t\right\} = \begin{bmatrix} 1 & t & (2)^{-1}t^{2} & \dots & \left[(\mu-1)!\right]^{-1}t^{\mu-1}\\ 0 & 1 & t & \dots & \left[(\mu-2)!\right]^{-1}t^{\mu-2}\\ \vdots & \vdots & \vdots & \dots & \vdots\\ 0 & 0 & 0 & \dots & t\\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} \right].$$
(10)

Из (10) видно, что столбцовая норма $\|\exp\{\mathbf{J}(0)t\}\|_{1}$, определяемая последним столбцом матричной экспоненты $\exp\{\mathbf{J}(0)t\}$, ее строчная норма $\|\exp\{\mathbf{J}(0)t\}\|_{\infty}$, определяемая первой строкой экспоненты, и оценка спектральной нормы $\|\exp\{\mathbf{J}(0)t\}\|_{\infty}$, задаваемая [4] мажорирующим неравенством

$$\left\|\exp\left\{\mathbf{J}(0)t\right\}\right\|_{2} \leq \left\{\left\|\exp\left\{\mathbf{J}(0)t\right\}\right\|_{1} \cdot \left\|\exp\left\{\mathbf{J}(0)t\right\}\right\|_{\infty}\right\}^{1/2}$$

совпадают и определяются выражением

$$\left\|\exp\left\{\mathbf{J}(0)t\right\}\right\|_{p} = 1 + t + (1/2)t^{2} + \ldots + (1/(\mu - 1)!)t^{\mu - 1} = \sum_{k=0}^{\mu - 1} (1/k!)t^{k}, (p = 1, 2, \infty).$$

Таким образом, норма матричной экспоненты $\|\exp\{\mathbf{J}(\alpha)t\}\|$ удовлетворяет соотношению

$$\left\| \exp\{\mathbf{J}(\alpha)t\} \right\| = e^{\alpha t} \sum_{k=0}^{\mu-1} (1/(k!))t^k .$$
(11)

Для дальнейших исследований выделим такое $\tilde{\mathbf{x}}(0) = \arg\{\|\tilde{\mathbf{x}}(0)\| = 1\}$, для которого выполняется точное равенство

$$\|\tilde{\mathbf{x}}(t)\| = \|\exp\{\mathbf{J}(\alpha)t\}\| \cdot \|\tilde{\mathbf{x}}(0)\|_{\|\tilde{\mathbf{x}}(0)\|=1} = e^{\alpha t} \sum_{k=0}^{\mu-1} (1/(k!))t^k.$$
(12)

Теперь при фиксированной кратности $\mu = n$ собственного числа $\lambda = \alpha$ поставим задачу оценки знака скорости изменения нормы $\|\tilde{\mathbf{x}}(t)\|$ в момент t = 0 как синдрома характера развития процессов в системе в функции от областей значений $\lambda = \alpha$. Дифференцированием по времени выражения (12) получим:

$$\frac{d}{dt} \|\tilde{\boldsymbol{x}}(t)\| = \frac{d}{dt} \left\{ e^{\alpha t} \sum_{k=0}^{\mu-1} (1/(k!)) t^k \right\} \bigg|_{t=0} = \left\{ \alpha e^{\alpha t} \sum_{k=0}^{\mu-1} (1/(k!)) t^k + e^{\alpha t} \sum_{k=0}^{\mu-2} (1/(k!)) t^k \right\} \bigg|_{t=0} = \alpha + 1.$$
(13)

Соотношение (13) позволяет сепарировать процессы по их качеству в системе (4) с матрицей состояния в виде жордановой клетки полной размерности по норме вектора состояния в функции от значения кратного собственного числа $\lambda = \alpha$. Ясно, что при любом отрицательном значении $\lambda = \alpha$ и при любой его кратности процессы в системе (4) являются сходящимися, потому что мультипликативный член $e^{\alpha t}$ в выражении (12) для $\|\tilde{x}(t)\|$ имеет бесконечное число элементов разложения по степеням t, в то

время как член $\sum_{k=0}^{\mu-1} (1/(k!)) t^k$ – конечное. Следовательно, всегда найдется такой момент времени $t = t^*$, с

которого начинает проявляться доминирование экспоненциального сомножителя $e^{\alpha t}$. Теперь рассмотрим следующие ситуации.

Ситуация 1. $\alpha < 0, |\alpha| > 1, \left\{ \frac{d}{dt} \| \tilde{\mathbf{x}}(t) \| \right\} \Big|_{t=0} < 0$, процесс $\| \tilde{\mathbf{x}}(t) \|$ сходится к нулю и мажорируется экс-

понентой в форме $\left\| \tilde{\mathbf{x}}(t) \right\| \le e^{(\alpha+1)t} \left\| \tilde{\mathbf{x}}(0) \right\|$.

Ситуация 2. $\alpha = -1, \left\{ \frac{d}{dt} \| \tilde{\mathbf{x}}(t) \| \right\} \Big|_{t=0} = 0$, начальная скорость нулевая, но при t > 0 в силу (13) уста-

навливается отрицательная скорость, определяемая выражением

$$\frac{d}{dt} \left\| \tilde{\mathbf{x}}(t) \right\| = \left\{ \alpha e^{\alpha t} \sum_{k=0}^{\mu-1} \left(\frac{1}{k!} \right) t^k + e^{\alpha t} \sum_{k=0}^{\mu-2} \left(\frac{1}{k!} \right) t^k \right\} \bigg|_{\alpha=-1} = -\left(\frac{1}{(\mu-1)!} \right) e^{-t} t^{(\mu-1)}.$$
(14)

Скорость изменения нормы $\|\tilde{\mathbf{x}}(t)\|$ на траекториях системы характеризуется экстремумом, наблюдаемым в момент t_m , определяемым в силу (14) соотношениями

$$t_m = \arg\left\{\frac{d^2}{dt^2} \|\tilde{\mathbf{x}}(t)\| = 0\right\} = \arg\left\{\frac{d}{dt} \left(e^{-t} t^{(\mu-1)}\right) = 0\right\} = \mu - 1,$$

при этом скорость изменения нормы $\|\tilde{\mathbf{x}}(t)\|$, будучи отрицательной, равна величине

$$\max\left(\frac{d}{dt}\left\|\tilde{\mathbf{x}}(t)\right\|\right) = -\frac{(\mu-1)^{(\mu-1)}}{(\mu-1)!}e^{-(\mu-1)!}$$

Процесс $\|\tilde{\mathbf{x}}(t)\|$ сходится к нулю в силу представления (12). Процесс мажорируется экспоненциальной функцией, т.е. выполняется неравенство

 $\left\|\tilde{\mathbf{x}}(t)\right\| \leq \rho e^{\gamma t} \left\|\tilde{\mathbf{x}}(0)\right\|,$

в котором параметры (р, ү) определяются из условия

$$(\rho,\gamma) = \arg \left\{ \min_{\rho,\gamma} \left\| \| \tilde{\mathbf{x}}(t) \| - \rho e^{\gamma t} \| \tilde{\mathbf{x}}(0) \| \| \& \left(\frac{d}{dt} \left(\rho e^{\gamma t} \| \tilde{\mathbf{x}}(0) \| \right) \right) \right|_{t=(\mu-1)} = -\frac{(\mu-1)^{(\mu-1)}}{(\mu-1)!} e^{-(\mu-1)} \& \rho \ge 1 \right\}.$$

Ситуация 3 (предмет статьи). $\alpha < 0, |\alpha| < 1, \left\{ \frac{d}{dt} \| \tilde{\mathbf{x}}(t) \| \right\} \Big|_{t=0} > 0$. Процесс $\| \tilde{\mathbf{x}}(t) \|$ на начальном от-

резке времени расходится, достигая максимума в момент t_M , определяемый соотношениями

$$t_{M} = \arg\left\{\frac{d}{dt}\left\|\tilde{\mathbf{x}}(t)\right\| = 0\right\} = \arg\left\{\left(1+\alpha\right)\sum_{k=0}^{\mu-2} (1/k!)t^{k} + \alpha\left(1/(\mu-1)!\right)t^{(\mu-1)} = 0\right\},\tag{15}$$

а далее сходится к нулю. Таким образом, процесс $\|\tilde{\mathbf{x}}(t)\|$ на траекториях свободного движения апериодической системы обнаруживает выброс, численно определяемый величиной $\alpha (\alpha < 0, |\alpha| < 1)$ кратного собственного числа и значением μ его кратности. Очевидно свойство процесса $\|\tilde{\mathbf{x}}(t)\|$: чем меньше величина $|\alpha| < 1$ и чем больше его кратность μ , тем больше величина его выброса над уровнем $\|\tilde{\mathbf{x}}(t)\|$. Для иллюстрации этого результата произведем вычисление момента t_M с помощью (15) и выброса в кривой $\|\tilde{\mathbf{x}}(t)\|$ апериодической системы для момента $t = t_M$ в силу соотношения (12) для различных значений $\alpha:(\alpha < 0, |\alpha| < 1)$ и кратностей μ . Результаты вычислений приведены в табл. 1, 2.

μ	2	3	4	5	10
α			t_M		
-0,2	4	8,9	13,9	18,8	43,8
-0,02	49	99	149	199	449

μ	2	3	4	5	10	
α	$\max_{t} \left(\left\ \tilde{\mathbf{x}}(t) \right\ \right) = \left\ \tilde{\mathbf{x}}(t_{M}) \right\ $					
-0,2	2,25	8,35	34,7	151,6	3,32·10 ⁵	
-0,02	18,8	690,4	$2,86 \cdot 10^4$	$1,25 \cdot 10^{6}$	$2,72 \cdot 10^{14}$	

Таблица 1. Значения моментов выброса в кривой $\| ilde{x}(t)\|$

Таблица 2. Значения выбросов $\max_{\mathbf{x}} \left(\| \tilde{\mathbf{x}}(t) \| \right) = \| \tilde{\mathbf{x}}(t_M) \|$ кривой $\| \tilde{\mathbf{x}}(t) \|$

Вернемся теперь к исходной системе (1) с матрицей состояния F, заданной в произвольном базисе. По аналогии с (8) с использованием (6) можно записать:

$$\mathbf{x}(t) = \mathbf{x}(t, \mathbf{x}(0)) = \exp\{\mathbf{F}t\} \mathbf{x}(0) = \mathbf{S} \exp\{\mathbf{J}(\alpha)t\} \mathbf{S}^{-1}\mathbf{x}(0).$$
(16)

Если в (16) перейти к скаляризованным векторным процессам по норме вектора состояния системы (1), то получим, с использованием (11), цепочку соотношений

$$\|\mathbf{x}(t)\| = \|\mathbf{S}\exp\{\mathbf{J}(\alpha)t\}\mathbf{S}^{-1}\mathbf{x}(0)\| \le \|\mathbf{S}\| \cdot \|\exp\{\mathbf{J}\alpha t\}\| \cdot \|\mathbf{S}^{-1}\| \cdot \|\mathbf{x}(0)\| = c\{\mathbf{S}\} e^{\alpha t} \sum_{k=0}^{\mu-1} (1/(k!))t^k \|\mathbf{x}(0)\|,$$

где $c\{\mathbf{S}\} = \|\mathbf{S}\| \cdot \|\mathbf{S}^{-1}\|$ – число обусловленности матрицы **S**, удовлетворяющее [4] условию $1 \le c\{\mathbf{S}\} < \infty$. Значения $\|\mathbf{\tilde{x}}(t)\|$ будут в $c\{\mathbf{S}\}$ раз превышать значения $\|\mathbf{\tilde{x}}(t)\|$, сохраняя ту же зависимость от модуля $|\alpha|$ собственного числа $\lambda = \alpha$ и его кратности μ .

Компьютерное исследование свободного движения непрерывной многомерной апериодической системы для случая кратных вещественных собственных чисел ее матрицы состояния

Компьютерное исследование процессов по норме $\|\tilde{\mathbf{x}}(t)\|$ как функции собственного числа $\lambda = \alpha$ и его кратности $\mu = n$ проводилось в соответствии с соотношением $\|\tilde{\mathbf{x}}(t)\| = \|\exp\{\mathbf{J}(\alpha)t\}\tilde{\mathbf{x}}(0)\| \le \|\exp\{\mathbf{J}(\alpha)t\}\|\cdot\|\tilde{\mathbf{x}}(0)\|$ по его мажорирующей части в модельной среде пакета MATLAB. Результаты моделирования процессов в форме $\|\tilde{\mathbf{x}}(t)\|$ для единого набора кратностей $\mu = n = 2; 3; 5$ и 10 и значений $\lambda = \alpha = -2; -0, 2$ и -0, 02 представлены на рисунках.

На рис. 1 приведены кривые для случая $\lambda = \alpha = -2$. Процессы $\|\tilde{\mathbf{x}}(t)\|$ сходятся монотонно без выбросов (см. ситуацию 1). Кривая 1 соответствует случаю $\mu = n = 2$, а кривая $4 - \mu = n = 10$.



Рис. 1. Кривые процессов $\|\tilde{\mathbf{x}}(t)\|$ при $\lambda = \alpha = -2$; и $\mu = n = 2$; 3; 5;10 (кривые 1–4 соответственно) На рис. 2 приведены 4 кривые для случая $\lambda = \alpha = -0, 2$. Процессы $\|\tilde{\mathbf{x}}(t)\|$ обнаруживают выбросы, нарастающие с увеличением $\mu = n$ (см. ситуацию 3).



Рис. 2. Кривые процессов $\|\tilde{\mathbf{x}}(t)\|$ при $\lambda = \alpha = -0, 2$ и $\mu = n = 2$ (а); 3(б); 5(в) и 10(г)



Рис. 3. Кривые процессов $\|\tilde{\mathbf{x}}(t)\|$ при $\lambda = \alpha = -0,02$ и $\mu = n = 2(a); 3(6); 5(в)$ и 10(г)



Рис. 4. Кривые постоянных значений $\max(\|\tilde{\mathbf{x}}(t)\|) = \|\tilde{\mathbf{x}}(t_M)\| = \text{const}$

На рис. 3 приведены 4 кривые для случая $\lambda = \alpha = -0,02$. Процессы $\|\tilde{\mathbf{x}}(t)\|$ имеют заметные выбросы, нарастающие с увеличением $\mu = n$ (см. ситуацию 3).

На рис. 4 приведены кривые постоянных значений $\max_{t} \left(\| \tilde{\mathbf{x}}(t) \| \right) = \| \tilde{\mathbf{x}}(t_M) \| = \text{const на плоскости}$ «µ – λ», иллюстрирующие возможность «обмена» кратности на значение кратного собственного числа врешаемой задаче.

На рис. 5 приведены кривые при $\lambda = \alpha = -0, 2$ для случая системы (1), в которой матрица F задана в сопровождающей строчной форме, и системы (4). Процессы $\|\tilde{\mathbf{x}}(t)\|$ имеют характер кривых $\|\tilde{\mathbf{x}}(t)\|$, но в каждый момент в $c\{\mathbf{S}\}$ раз превышают их значения.

Завершая рассмотрение влияния фактора кратности собственных чисел на качество процессов в апериодических системах, следует отметить, что если спектр собственных чисел матрицы **F** имеет несколько кратных чисел σ {**F**} = { $\lambda_i = \alpha_j : i = \overline{1, \mu_j}; j = \overline{1, q}; \sum_{j=1}^{q} \mu_j = n$ }, то каноническое представление **F** в жордановой форме будет содержать q жордановых клеток размерности ($\mu_j \times \mu_j$) каждая. Для такого случая соотношение (9) принимает вид

$$\|\tilde{\mathbf{x}}(t)\| = \|diag\{\exp\{\mathbf{J}(\alpha_{j})t\}; j = \overline{\mathbf{I},q}\}\tilde{\mathbf{x}}(0)\| \le e^{\overline{\alpha}t} \|\exp\{\mathbf{J}_{(\overline{\mu}\times\overline{\mu})}(0)t\}\| \cdot \|\tilde{\mathbf{x}}(0)\|,$$

rge $\overline{\alpha} = \max_{j}\{\alpha_{j}: \alpha_{j} < 0 \& |\alpha_{j}| < 1; j = \overline{\mathbf{I},q}\}; \overline{\mu} = \max_{j}\{\mu_{j}; j = \overline{\mathbf{I},q}\}.$



Заключение

Установлено, что кратность собственных чисел матриц состояния устойчивых апериодических непрерывных систем, как и структура их собственных векторов [5], оказывается важным системным фактором, наделяющим динамические процессы в системе весьма специфическими свойствами, которые могут приводить к нежелательным последствиям разрушительного характера. Чтобы не допустить обнаруженного эффекта кратности собственных чисел при синтезе методами модального управления [3], матрицу состояния **F** системы следует наделить спектром собственных чисел, не содержащих кратные элементы.

Работа подготовлена при поддержке проекта 14.В37.21.0406 «Разработка многофункционального малогабаритного мультиротационного летательного аппарата».

Литература

- 1. Гантмахер Ф.Р. Теория матриц. М.: Наука, 1973. 575 с.
- 2. Андреев Ю.Н. Управление конечномерными линейными объектами. М.: Наука, 1976. 424 с.
- Дударенко Н.А., Слита О.В., Ушаков А.В. Математические основы современной теории управления: аппарат метода пространства состояний: Учебное пособие / Под ред. А.В. Ушакова – СПб: СПбГУ ИТМО, 2008. – 323 с.
- 4. Голуб Дж., Ван Лоун Ч. Матричные вычисления: Пер. с англ. М.: Мир, 1999. 548 с.
- 5. 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.

Акунов Таалайбек Абакирович	 Санкт-Петербургский национальный исследовательский университет
	информационных технологий, механики и оптики, докторант,
Tudanawa Hamazua Azawandanawa	เล่นแบงใน/เกิดที่เห็น Court Поторбиргоний, และแอนอส แม่นี้ และสอสอกอาจาร อหมนี้ มนแคงกอนสอส
дуоаренко паталия Алексаноровна	 Санкт-петероургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат техниче- ских наук, доцент, dudarenko@yandex.ru
Полинова Нина Александровна	 Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, студент, polinova_nina@mail.ru
Ушаков Анатолий Владимирович	 Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, профессор, ushakov-AVG@yandex.ru

УДК 004.045; 004.428; 004.942 ПРОЕКТИРОВАНИЕ И ТЕХНОЛОГИЧЕСКАЯ ПОДГОТОВКА СЕТЕЙ СТАНЦИЙ ВЕРТИКАЛЬНОГО ЗОНДИРОВАНИЯ ИОНОСФЕРЫ А.Ю. Гришенцев, А.Г. Коробейников

Проведен анализ структуры сети станций вертикального зондирования ионосферы. Рассмотрены особенности разработки и принцип построения комплексов программ автоматизированной обработки, анализа и хранения данных зондирования ионосферы. Построена концептуальная модель комплексной системы управления базами данных. В практику исследования ионосферы внедрено соответствующее программное обеспечение. Показаны результаты применения предложенных авторами алгоритмов и программ автоматизированной обработки и анализа данных вертикального зондирования ионосферы.

Ключевые слова: проектирование, разработка, программное обеспечение, исследования, вертикальное зондирование ионосферы.

Введение

Ионосфера – это часть верхней атмосферы, где плотность свободных электронов достаточна, чтобы оказывать значительное влияние на распространение радиоволн. Отсюда следует, что исследование структуры ионосферы важно как для понимания физики протекающих в ней процессов, так и для решения разнообразных радиофизических задач, связанных с распространением радиоволн. Для практических нужд человечества изучение ионосферы связано, например, с эффективностью функционирования спутниковых систем радиосвязи и координатно-временного обеспечения, с разработкой радаров с синтезированной апертурой, использующих все более низкие частоты радиоволн и способных осуществлять глубинное зондирование земных грунтов.

Анализ состояния ионосферы производят различными методами, одним из которых является вертикальное зондирование ионосферы. Этот метод реализуется на станциях вертикального зондирования ионосферы (СВЗИ).

В настоящее время происходит модернизация сети станций СВЗИ, находящихся на балансе Института земного магнетизма, ионосферы и распространения радиоволн им. Н.В. Пушкова РАН (ИЗМИРАН, Россия), Государственного научного центра «Арктический и антарктический научноисследовательский институт» (ААНИИ, Россия). На этих сетях СВЗИ производятся достаточно длительные приборные наблюдения ионосферы, позволяющие делать выводы об идущих в ионосфере динамических процессах. В настоящее время на территории Российской Федерации функционируют не менее 20 СВЗИ (в СССР было 40), образующие сеть и постоянно передающие данные в один из мировых центров данных (МЦД). Для более глубокого понимания происходящих в ионосфере процессов необходимо использование архивных данных, собранных СВЗИ. Это позволит, например, проводить синтез и анализ математических (в том числе статистических) моделей явлений в ионосфере и показывать их связи с другими процессами природного и антропогенного происхождения [1–7]. Актуальность таких исследований, включающих непрерывный мониторинг, моделирование, сбор, обработку, анализ и хранение данных, была показана в ряде работ [1–3].

Современные методы радиозондирования ионосферы можно разделить на три класса: вертикальное зондирование, наклонное зондирование, спутниковое радиопросвечивание. При этом вертикальное зондирование ионосферы (ВЗИ) является исторически наиболее применяемым приборным методом исследования, а данные ВЗИ на сегодняшний день накоплены за наиболее длительный период.

С появлением современных вычислительных и коммуникационных средств и технологий возникла необходимость концептуальной модернизации сети СВЗИ, включающей:

- разработку программного обеспечения (ПО) и баз данных;
- разработку аппаратного обеспечения;
- разработку средств коммуникаций и инфраструктуры;
- разработку и исследование математических моделей;
- внедрение принципов интероперабельности;
- разработку методов и средств защиты информации;
- развертывание, тестирование и настройку оборудования и ПО;
- системный анализ проектных решений в целом.

В настоящей работе проведен анализ структуры сети СВЗИ и рассмотрены вопросы разработки ПО и баз данных применительно к проектированию и технологической подготовке сети СВЗИ.

Постановка задачи

Для полноценного использования современных аппаратных средств, реализации методов передачи, обработки, анализа и хранения данных требуется разработка специального ПО. В большинстве случаев потребность разработки ПО связана с поступлением нового импортного технологического оборудования и с включением России в международную сеть ионосферных наблюдений. Последнее обусловлено постоянно изменяющимся требованиями международного координационного центра. Таким образом, разработка специального ПО, позволяющего проводить автоматизированную обработку, анализ и хранение данных отечественной сети СВЗИ, является актуальной.

Сеть станций вертикального зондирования

Рассматриваемые в данной работе СВЗИ расположены в высокоширотной области, что определяет особенности регистрируемых данных, а следовательно, и методов их обработки и анализа. Кроме того, практически все СВЗИ оборудованы дополнительными средствами мониторинга, позволяющими проводить покомпонентные наблюдения за магнитным полем Земли в географической точке расположения СВЗИ. В связи с этим такие станции наблюдения часто называют комплексными магнито-ионосферными обсерваториями (КМИО).

Ценность приборных наблюдений заключается в их непрерывности, что позволяет отслеживать динамику процессов в ионосфере, наблюдать скорость развития, длительность и периодичность возмущений. Практика показала, что периодичность зондирования должна составлять не более нескольких минут. Результатом получения серии зондирующих импульсов является получение так называемой ионограммы. Обычный объем файла одной ионограммы измеряется десятками килобайт. Таким образом, в сети КМИО за месяц накапливаются существенные объемы информации, которые требуют обработки, анализа, систематизации, передачи и хранения. Хранение и обработка полученных результатов решаются как непосредственно на КМИО, так и в центре сбора данных (центр данных).

Операции, производимые на КМИО:

- развертывание, настройка и эксплуатационное обслуживание зондов и сопутствующего оборудования;
- регистрация ионограмм;
- предварительная обработка;
- сопровождение архива;
- организация взаимодействия с центром данных.

Операции, производимые в центре данных:

- централизованная разработка методологии исследований, включающая концептуальные технические решения, разработку новых и модернизацию имеющихся аппаратно-программных средств, методов обработки, анализа и хранения данных;
- развертывание, настройка и эксплуатационное обслуживание вычислительных ресурсов и сопутствующего оборудования;
- организация информационного взаимодействия с отдельными КМИО;
- организация и выполнение централизованной систематической обработки и анализа данных;
- организация взаимодействия с другими исследовательскими центрами, в том числе МЦД. На рис. 1 приведена организационная структура сети КМИО.



Рис. 1. Организационная структура сети КМИО

Спецификой построения сети КМИО, удовлетворяющей современным требованиям к центрам данным, является выполнение операций доставки данных в режиме реального времени и интероперабельность, в основе которых лежит унификация форматов данных и протоколов обмена. Доступность сети Интернет на КМИО позволяет применить сетевые протоколы обмена в качестве основного средства коммуникации. Ионограммы предоставляются в двух форматах – графическом (PNG) и бинарном. Бинарный формат имеет меньший объем и предпочтительнее для передачи, хранения и обработки. Кроме того, бинарные файлы содержат некоторую служебную информацию, которая утрачивается при преобразовании в графический формат.

СВЗИ оборудованы приемо-передающими анализаторами спектра отраженных от ионосферы сигналов (зондами) отечественного производства АИС-М, СП-3-М, БАЗИС-М и САDI канадского производства. Индекс «М» в обозначении отечественного анализатора означает, что он имеет цифровой интерфейс, позволяющий производить настройку режимов зонда, а также автоматически получать и сохранять данные в электронном виде в ЭВМ [8].

Протоколы и форматы данных

В последнее время на отечественных СВЗИ значительное распространение получили ионозонды Canadian Advenced Digital Ionosonde (CADI) производства Scientific Instrumentation Limited (SIL, Канада) (рис. 2). Такие зонды используют достаточно хорошо документированный класс форматов ионограмм:

- md1: диапазон высот до 510 км, одна или более фиксированная частота;
- md2: диапазон высот до 510 км, базовый набор частот;
- md3: то же, что md1, но диапазон высот до 1020 км;
- md4: то же, что md2, но диапазон высот до 1020 км.

В таблице приведены некоторые сравнительные характеристики зонда АИС-М (анализатор ионосферного спектра модифицированный) и САDI [9].

Характеристика	АИС-М	CADI (2009)
Год производства	1957 (модификация 2002 г.)	2009
Диапазон частот (от), МГц	1–20	1–30
Мощность передатчика, кВт	2,5	0,6
Диапазон высот, км	90–1500	90-1020
Максимальная частота повторения импульсов, Гц	40	256
Максимальная длительность импульсов, мкс	70	40

Таблица. Некоторые характеристики ионосферных зондов АИС-М и CADI



Рис. 2. Системный модульный блок ионозонда CADI(а), внешняя антенна(б), дневная ионограмма, город Троицк (55,5° с.ш., 37,5° в.д.) в спокойных геофизических условиях (в). По горизонтали – частота в МГц

На сегодняшний день часть CB3И оборудована зондами отечественного производства АИС-М (или подобными). Учитывая длительный, многолетний период наблюдения с использованием теперь уже частично устаревших отечественных анализаторов, даже при их повсеместной замене на более новые, сохранение полученных данных в существенно варьирующихся современных форматах является актуальной задачей.

Дорогостоящее ПО, дополнительно предоставляемое с оборудованием CADI, содержит модули управления зондами и обработки ионограмм. Однако поставщики импортного оборудования предоставляют его с «закрытым кодом», что является существенной проблемой для его интеграции в разрабатываемую сеть КМИО. По этой причине для полноценного использования данных ионозондов необходимо повторно разрабатывать специальные методы, алгоритмы и реализовывать их в виде специального ПО. Следует отметить, что подобная проблема (закрытости, отсутствия некоторой документации, сокрытия инженерных и научных решений) достаточно характерна для случаев использования импортных комплектующих в отечественных сложных информационных системах, особенно если они разработаны зарубежными оборонными предприятиями и содержат недокументированные возможности, недоступные отечественным пользователям.

Обеспечение центров данных

Центр данных является логическим центром сети КМИО и выполняет функции управления работой отдельных КМИО, сбора данных и предварительной обработки результатов (на местах), последующей обработки, анализа и архивирования данных. Кроме того, сотрудники центра данных решают задачи методологического уровня, относящиеся к разработке новых методов интерпретации данных, выявлению взаимосвязей процессов в ионосфере с другими природными и антропогенными факторами. Центр данных также осуществляет взаимодействие с внешним окружением (рис. 1).

Разработанный авторами комплекс систем управления базами данных (КСУБД) с применением реляционной модели построен с учетом следующих принципов:

- соответствие таблиц базы данных реальным физическим объектам;
- достаточное количество полей и индексов в таблицах базы данных для оперативного извлечения информации;
- устойчивость к программным и аппаратным сбоям и ошибкам;
- быстродействие, достигаемое за счет оптимизации запросов, структуры таблиц и форматов хранимой информации.

В структуре базы данных выделены следующие кластеры (рис. 3):

- исходные данные;
- обработка и анализ;
- заключения, публикации, комментарии;
- прогнозы.

Рассмотрим назначение каждого кластера в отдельности.

Исходные данные. Кластер содержит в себе исходные данные, получаемые со СВЗИ, а также результаты первичной обработки, производимой «на местах».

Обработка и анализ. Кластер содержит результаты обработки исходных данных при помощи ЭВМ, а также человеком.

Заключения, публикации, комментарии. Кластер содержит систематизированные, возможно, обобщенные, результаты анализа, собственную базу знаний, отчетов, технической документации и других материалов, связанных с предметной областью.

Прогнозы. Кластер является результатом работы математических моделей (интегрированных с базами данных) и позволяет производить упорядоченное хранение данных прогнозов, оценку адекватности прогнозирования и протоколирование рекомендуемых корректировок для математических моделей.



Рис. 3. Организационно-логическая модель базы данных

Рассмотренная организационно-логическая модель позволяет производить независимое хранение данных, являющихся результатами различных этапов исследовательской деятельности [1, 2, 10, 11], интегрировать математические методы обработки, анализа и прогнозирования непосредственно в систему. В свою очередь, такая интеграция и непрерывный мониторинг конкурирующих математических моделей с оценкой адекватности их расчетов позволяют вести непрерывную доработку и модернизацию. Еще одним положительным фактором интеграции документации в виде заключений, публикаций и комментариев является возможность использования базы знаний в качестве учебной для молодых специалистов или желающих пройти дополнительную подготовку. Применение кластерной структуры, основанной на объектной модели разработанного программного обеспечения, делает модель КСУБД легко модернизируемой.

Практическая реализации КСУБД осуществлена на основе реляционной СУБД MySQL. СУБД MySQL является кроссплатформенной и в рассматриваемой работе использовалась на базе операционной системы семейства Linux. MySQL имеет интерфейсы для транслируемых языков, таких как PHP, Perl, и для компилируемых в исполняемые модули C, C++. СУБД MySQL является хорошо зарекомендовавшим себя на индустриальном рынке информационных технологий продуктом, позволяет работать (в версиях MySQL выше 3.22) с таблицами до 2⁶³ байт, поддерживает механизм транзакций и псевдотранзакций на основе блокировок, обеспечивает многопользовательскую работу.

Для дополнительной разгрузки СУБД, способствующей улучшению скоростных характеристик выполнения запросов и повышению надежности, предусмотрено уменьшение объемов таблиц MySQL путем разделения данных на табличные и файловые. К табличным данным отнесем настройки оборудования, время и особенности регистрации ионограммы, «компактные» результаты обработки и др. К файловым данным отнесем ионограммы, файлы изображений, некоторые виды документации, например, формата Portable Document Format (PDF) и др. Такой подход позволяет в несколько раз сократить размеры таблиц, следовательно, значительно снизить нагрузку на СУБД и ускорить ее работу.

В виде отдельного программного модуля и программных библиотек на языке С++ авторами разработан комплекс «SkySpectrum» [1, 11] позволяющий производить автоматизированную обработку. преобразование форматов данных и визуализацию ионограмм [12]. В ходе исследований и экспертного анализа программного комплекса на базе обработки данных ионосферной станции Воейково (Санкт-Петербург) за 2002-2003 г.г. выявлено, что разработанные методы автоматизированной фильтрации и анализа ионограмм позволяют определить значение критической частоты f_0 с погрешностью до 16% в 68% случаев, 23% ионограмм были отсеяны на уровне фильтрации как не поддающиеся машинному анализу из-за высокого уровня шумов, в 9% СВЗИ критические частоты были определены с погрешностью более 16%. С учетом географического расположения СВЗИ и особенностей высокоширотной ионосферы данные результаты можно оценивать как положительные, что позволяет внедрять разработанные методы и программы в практические исследования. Следует отметить, что не существует отечественных программ, аналогичных разработанным авторами и предназначенных для автоматизированного анализа ионограмм, зарубежные разработчики подобных программ не раскрывают исходный код, что практически исключает встраивание зарубежного ПО в отечественные системы. В этом смысле разработанные авторами методы и алгоритмы представляют особую ценность для развития отечественных средств автоматизированного анализа данных СВЗИ.

Заключение

Предложенная организационно-логическая модель комплекса систем управления базами данных позволяет производить построение центров данных (рис. 1) мониторинга окружающей среды с применением реализованных авторами программных средств автоматизированной обработки и анализа.

Авторами разработаны системы:

- автоматизированной обработки данных вертикального зондирования ионосферы (применяется в СПбФ ИЗМИРАН);
- автоматизированного сбора и хранения данных вертикального зондирования ионосферы (используется в ААНИИ).

Разработаны предметно-ориентированные программные приложения и программные библиотеки, две из которых зарегистрированы в Федеральной службе по интеллектуальной собственности, патентам и товарным знакам Российской Федерации [10, 11]. Авторами получен акт о внедрении компьютерной программы «SkySpectrum» от 03.07.2012 г. [11] в научно-исследовательскую практику СПбФ ИЗМИРАН.

Литература

- 1. Гришенцев А.Ю., Коробейников А.Г. Обратная задача радиочастотного зондирования ионосферы // Журнал радиоэлектроники. – 2010. – № 10 [Электронный ресурс]. – Режим доступа: http://jre.cplire.ru/jre/oct10/6/text.html, свободный. Яз. рус. (дата обращения 10.02.2013).
- 2. Гришенцев А.Ю., Коробейников А.Г. Разработка модели распределения плотности токов при возбуждении ионосферы высокочастотным облучением // Изв. вузов. Приборостроение. 2010. Т. 53. № 12. С. 41–47.
- 3. Гришенцев А.Ю., Коробейников А.Г. Разработка модели решения обратной задачи вертикального зондирования ионосферы // Научно-технический вестник СПбГУ ИТМО. – 2011. – № 2 (27). – С. 21–26.
- Смирнов В.М. Интерпретация ионосферных возмущений в период слабых землетрясений // Электронный журнал «Исследовано в России» [Электронный ресурс]. – Режим доступа: http://zhurnal.ape.relarn.ru/articles/2003/012.pdf, свободный. Яз. рус. (дата обращения 10.02.2013).
- 5. Липеровский В.А., Похотелов О.А., Шалимов С.Л. Ионосферные предвестники землетрясений. М.: Наука, 1992. 304 с.
- 6. Яковлев О.И., Якубов В.П., Урядов В.П., Павельев А.Г. Распространение радиоволн. М.: ЛЕНАНД, 2009. 496 с.
- 7. Яковлев О.И., Павельев А.Г., Матюгов С.С. Спутниковый мониторинг Земли: Радиозатменный мониторинг атмосферы и ионосферы. М.: Либроком, 2010. 208 с.
- 8. Ким Ю.В. Модернизация ионосферной станции «АИС». Отчет по теме № 690. ИЗМИРАН / Руководитель темы Ю.В. Ким. Троицк, 2002. 24 с.
- 9. Canadian advanced digital ionosonde. System manuals / Scientific instrumentation limited. 2233 Hanselman Avenue Saskatoon, SK S7L 67A, Canada, 2009. 22 p.
- 10. Гришенцев А.Ю., Муромцев Д.И. Система управления данными наблюдений солнечно-земной физики «МІ» // Регистрация программы для ЭВМ от 21.07.2011 г. – № 2011615714.
- 11. Гришенцев А.Ю., Коробейников А.Г. Программа обработки и анализа данных ионосферного спектрографа АИС-М «SkySpectrum» // Регистрация программы для ЭВМ от 28.09.2011 г. – № 2011617569.
- Piggott W.R., Rawer K. URSI handbook of Ionogram Interpritation and Reduction. INAG (Ionospheric Network Advisory Group) WORLD DATA CENTER A. National Academy of Sciences, 2101 Constitution Avenue, N.W. Washington, D.C., U.S.A., 20418. – Second edition, november, 1972. – 145 p.

Гришенцев Алексей Юрьевич – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат технических наук, доцент, tigerpost@ya.ru

Коробейников Анатолий Григорьевич – Санкт-Петербургский филиал Института земного магнетизма, ионосферы и распространения радиоволн им. Н.В.Пушкова РАН, доктор технических наук, профессор, зам. генерального директора; Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, профессор; Korobeynikov_A_G@mail.ru

4

КОМПЬЮТЕРНЫЕ СИСТЕМЫ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

УДК 681.3 НАДЕЖНОСТЬ РЕЗЕРВИРОВАННОГО ВЫЧИСЛИТЕЛЬНОГО КОМПЛЕКСА ПРИ ОГРАНИЧЕННОМ ВОССТАНОВЛЕНИИ В.А. Богатырев, С.М. Алексанков, Д.В. Демидов, В.Ф. Беззубов

Предложены марковские модели надежности дублированных вычислительных комплексов, позволяющие учесть влияние на нестационарный коэффициент готовности ограниченных возможностей восстановления, вызванных недопустимостью потери данных и прерываний вычислительного процесса во время восстановления. Ключевые слова: надежность, отказоустойчивость, доступность, дублированный вычислительный комплекс, резервирование, нестационарный коэффициент готовности.

Введение

Высокая надежность, отказоустойчивость и безопасность систем критического применения, в частности, функционирующих в реальном времени, реализуется за счет резервировании основных подсистем, в том числе средств хранения, обработки и передачи данных.

Безопасность систем критического применения достигается при минимизации возможностей возникновения опасных отказов и последствий от их возникновения. Функциональная безопасность систем связана с минимизацией частоты и последствий отказов в выполнении отдельных функций системы, в том числе связанных с безопасностью [1, 2]. Обеспечение функциональной безопасности в значительной мере определяется информационной безопасностью и, прежде всего, связано с минимизацией риска потери данных (в том числе формируемых в ходе вычислительного процесса) при отказах системы, вызванных внешними или внутренними причинами, в том числе в результате злонамеренных воздействий.

В системах критического применения используемые вычислительные узлы, как правило, строятся на основе дублированных или троированных вычислительных комплексов. Возможности сохранения данных после отказов дублированных вычислительных комплексов (ДВК) во многом определяются организацией средств комплексирования и доступа к памяти.

Для достижения высокой отказоустойчивости, функциональной и информационной безопасности при сохранении устойчивости вычислительного процесса средства комплексирования ДВК должны обеспечивать доступность информационных ресурсов (памяти) обоих полукомплексов, даже при отказах процессоров одного из них, тем самым сохраняя возможности функционирования (или данных) в режиме деградации. Формирование работоспособной структуры в ряде случаев требует реконфигурации с целью использования в вычислительном процессе сохраненных ресурсов обоих полукомплексов. В настоящее время достаточно хорошо исследованы модели надежности дублированных систем [3–5], в том числе восстанавливаемых дублированных систем при различных дисциплинах ограниченного восстановления, под которым понимается восстановление при образовании очереди на восстановление отказавших узлов.

Потенциал обеспечения надежности управляющих вычислительных систем во многом определяется особенностями прикладных процессов, в ряде случаев ограничивающих возможности восстановления системы после отказов.

В связи с высоким вниманием к анализу безопасности и функциональной надежности систем критического применения, в том числе систем, критичных к нарушениям доступности и целостности информационных ресурсов, представляется актуальным исследование их надежности. Модели надежности таких систем должны учитывать срывы вычислительного процесса и (или) потерю стратегически важной информации, потенциально приводящие к катастрофическим последствиям, после которых восстановление вычислительной системы теряет смысл. Модели надежности указанного класса систем, помимо оценки традиционных показателей надежности, должны находить вероятности опасных состояний, при которых возможны срыв критичных вычислительных (управляющих) процессов и невосполнимая потеря критически важных данных.

В работе предложены марковские модели надежности восстанавливаемых ДВК, отличающиеся возможностью учета ограниченного восстановления, при невосполнимой потере накопленных данных (результатов вычислений) или при недопустимости прерываний вычислительного (управляющего) процесса во время процесса восстановления системы.

Задачи исследования

В работе решается задача оценки надежности ДВК с учетом ограниченного восстановления после отказов. Типовая структура дублированного вычислительного комплекса, представленная на рис. 1, состоит из двух полукомплексов, каждый из которых содержит процессор (ЦП) и память (М). Доступ ЦП к ресурсам сопряженного полукомплекса обеспечивается средствами комплексирования (адаптером со-пряжения – АС) [6–7].



Рис. 1. Структура дублированного вычислительного комплекса

Средства комплексирования должны обеспечивать доступ к памяти сопряженного полукомплекса с целью сохранения информационных ресурсов (результатов решения прикладных задач) даже при отказах одного из дублированных ЦП. Таким образом, для информационно безопасного, отказоустойчивого функционирования комплекса АС должен обеспечивать прямой доступ к памяти двух полукомплексов.

Надежность ДВК в определенной мере зависит от ограничений восстановления системы после отказов, определяемых особенностями прикладных процессов [8, 9].

Рассмотрим ограничения восстановления комплекса. Оно нереализуемо в случаях, когда:

- потеря данных недопустима, а перерывы вычислительного процесса во время восстановления допустимы (S1);
- потеря данных и перерывы вычислительного процесса во время восстановления недопустимы (S2);
- потеря данных и перерывы вычислительного процесса во время восстановления допустимы (S3).

Потеря данных происходит, когда при информационной зависимости прикладных функциональных задач, решаемых вычислительной системой, отказывает память в двух полукомплексах, что приводит к невозможности выполнения задач, т.е. к отказу системы, при котором ее восстановление невозможно.

Перерывы вычислительного процесса во время восстановления недопустимы, например, для управляющих систем, работающих в реальном времени, при этом будем считать, что для поддержки вычислений достаточно работоспособности хотя бы одного ЦП при доступности для него памяти хотя бы в одном полукомплексе. При отказе одного ЦП доступ исправного ЦП к памяти другого (сопряженного) полукомплекса должен поддерживаться средствами комплексирования по прямому доступу.

Случай ограничения восстановления, когда потеря данных недопустима, а перерывы вычислительного процесса во время восстановления допустимы, не рассматривается, так как при отказе памяти в двух полукомплексах вычислительный процесс не реализуем, и при потере данных в обоих полукомплексах их восстановление считается невозможным.

Модели надежности с учетом ограничений восстановления

При построении марковской модели [3] надежности восстанавливаемого ДВК с учетом отмеченных ограничений восстановления (S1–S3) будем считать, что в каждый момент времени восстанавливается только один узел. Будем считать известными интенсивности отказа процессора – λ_0 , памяти – λ_1 , адаптера (средств комплексирования) – λ_2 , а интенсивность восстановления любого узла равна μ . При достижении состояния, при котором вычислительный процесс в соответствии с перечисленными условиями (S1–S3) восстановить не удается, считается, что достигнуто невосстанавливаемое состояние (состояния полного отказа системы), когда выполнение потока прикладных задач невозможно.

Марковские модели надежности ДВК для различных случаев ограниченного восстановления после отказов S1–S3 приведены на рис. 2–4, на которых выделены состояния:

- работоспособные, при которых вычисления реализуемы;
- отказа, при которых восстановление системы возможно;
- отказа, при которых восстановление системы невозможно.

На графах переходов (рис. 2–4) вершины, представляющие состояния системы, пронумерованы (A0-A17 для рис. 2; B0-B17 для рис. 3; C0-C14 для рис. 4). В поле вершин графов представлены изображения, поясняющие состояния ДВК, в соответствии с рис. 1, при этом отказавшие узлы перечеркнуты. Переходы (ребра) графов, обозначенные сплошными линиями, соответствуют отказам, а пунктирные – восстановлениям. На графе для всех переходов указаны соответствующие им интенсивности отказов λ_0 , λ_1 , λ_2 и восстановлений μ .

Для варианта ограниченного восстановления *S*1 признаком невосстанавливаемого состояния является отказ блоков памяти двух полукомплексов (состояния *A*4, *A*5, *A*7, *A*8, *A*10, *A*11, *A*12, *A*15, *A*16, *A*17 неработоспособны, среди них состояния *A*5, *A*8, *A*11, *A*16, *A*17 – невосстанавливаемые, неработоспособные состояния на рис. 2–4 затемнены).
Для варианта ограниченного восстановления S2 невосстанавливаемое состояние возникает при отказе двух блоков памяти или двух процессоров. Для варианта S2, при исправности AC, способность выполнения вычислительного процесса во время восстановления сохраняется, если исправен хотя бы один ЦП и один узел памяти, а при отказе AC – если одновременно исправны ЦП и модуль памяти одного полукомплекса. Для систем, соответствующих рис. 3, все неработоспособные состояния (B4, B5, B7, B8, B10, B11, B15, B16, B17) являются невосстанавливаемыми.

Для варианта восстановления S3 все состояния являются восстанавливаемыми (к неработоспособным состояниям относятся C4, C5, C7, C8, C12, C13, C14), т.е. ограничений по восстановлению, обусловленных определенными сочетаниями отказов узлов, нет.



Рис. 2. Граф переходов для марковской модели ДВК при ограничении восстановления S1

Для вычисления вероятности нахождения системы в работоспособном состоянии по графам, представленным на рис. 2–4, в соответствии с известными правилами [3] составляется система дифференциальных уравнений, решение которой позволяет найти вероятности всех состояний ДВК, в том числе работоспособных.

Надежность ДВК охарактеризуем нестационарным коэффициентом готовности K(t), который определяется как вероятность того, что в заданный момент времени *t* система находится в одном из работоспособных состояний. Коэффициент K(t) вычисляется при суммировании вероятностей всех работоспособных состояний. Заметим, что нестационарный коэффициент готовности (функция готовности) K(t)при возможностях восстановления, соответствующих S3 (когда все состояния допускают восстановления), с увеличением *t* стремится к стационарному коэффициенту готовности. При вариантах ограничений восстановления S1 и S2 (когда есть состояния, для которых восстановление системы после отказов не производится) с увеличением *t* вероятность работоспособного состояния системы стремится к нулю.

С учетом сформулированных условий работоспособности состояний ДВК для вариантов *S*1–*S*3 ограничений восстанавливаемости системы нестационарные коэффициенты готовности вычисляются как

$$\begin{split} K_{S1}(t) &= p_{A0}(t) + p_{A1}(t) + p_{A2}(t) + p_{A3}(t) + p_{A6}(t) + p_{A9}(t) + p_{A13}(t) + p_{A14}(t), \\ K_{S2}(t) &= p_{B0}(t) + p_{B1}(t) + p_{B2}(t) + p_{B3}(t) + p_{B6}(t) + p_{B9}(t) + p_{B13}(t) + p_{B14}(t), \\ K_{S3}(t) &= p_{C0}(t) + p_{C1}(t) + p_{C2}(t) + p_{C3}(t) + p_{C6}(t) + p_{C9}(t) + p_{C10}(t) + p_{C11}(t), \end{split}$$

где $p_{A0}(t), p_{A1}(t), ..., p_{A14}(t), p_{B0}(t), p_{B1}(t), ..., p_{B14}(t), p_{C0}(t), p_{C1}(t), ..., p_{C11}(t)$ соответствуют вероятностям состояний A0, A1, ..., A14 системы S1 (рис. 2), состояний B0, B1, ..., B14 системы S2 (рис. 3) и состояний C0, C1, ..., C14 системы S3 (рис. 4).



Рис. 3. Граф переходов для марковской модели ДВК при ограничении восстановления S2



Рис. 4. Граф переходов для марковской модели ДВК при ограничении восстановления S3

Вероятность отказа для систем S1–S4 вычисляется как $1 - K_{S1}, 1 - K_{S2}, 1 - K_{S3}$.

Предложенные модели позволяют вычислить вероятности опасных состояний. Так, вероятность потери информационных ресурсов для систем S1 определяется как

 $K_{aS1}(t) = p_{A5}(t) + p_{A8}(t) + p_{A11}(t) + p_{A16}(t) + p_{A17}(t),$

а вероятность неопасного (восстанавливаемого) отказа – как

$$K_{bS1}(t) = p_{A5}(t) + p_{A8}(t) + p_{A11}(t) + p_{A16}(t) + p_{A17}(t) + p_{A7}(t) + p_{A10}(t) + p_{A12}(t) + p_{A15}(t) .$$

Для системы S2 вероятность опасных состояний вычисляется как $1 - K_{s_2}$.

Результаты расчетов

Результаты расчета нестационарного коэффициента готовности K(t) для рассматриваемых вариантов S1-S3 ограниченного восстановления дублированных комплексов после отказов представлены на рис. 5. Расчеты проведены для случая $\lambda_0 = 0.8 \cdot 10^{-4}$, $\lambda_1 = 0.5 \cdot 10^{-4}$, $\lambda_2 = 0.3 \cdot 10^{-4}$ 1/ч и $\mu = 1$ 1/ч. Кривые 1–3 соответствуют вариантам S1-S3 ограничений восстанавливаемости после отказов.

Расчеты выполнены в системе компьютерной математики Mathcad-15 путем решении системы дифференциальных уравнений, составленных по графам на рис. 2–4 по методу Рунге–Кутты.



Рис. 5. Нестационарный коэффициент готовности для вариантов ограниченного восстановления дублированных комплексов после отказов. Кривые 1–3 соответствуют вариантам ограничений восстанавливаемости S1–S3

Расчеты подтверждают существенность влияния ограничений восстанавливаемости после отказов на надежность вычислительных комплексов. Таким образом, при оценке надежности и безопасности критически важных систем рекомендуется использование предлагаемых моделей, позволяющих учесть особенности ограниченного восстановления при невосполнимой потере информационных ресурсов и при прерывании критически важных вычислительных процессов при восстановлении.

Современные информационные управляющие системы, как правило, представляют собой распределенные системы, в которых множество узлов обработки и хранения данных объединяются в кластеры, связанные через многоуровневую сеть. В системах критического применения в качестве базовых узлов могут использоваться дублированные вычислительные комплексы, при оценке надежности и безопасности которых могут применяться предлагаемые в настоящей работе модели как части системы [10–14].

Заключение

Предложены марковские модели дублированных вычислительных комплексов, позволяющие учесть влияние на надежность ограниченности возможностей восстановления при недопустимости потери данных и прерываний вычислительного процесса во время восстановления.

Представленные модели надежности могут быть использованы при прогнозировании надежности и выборе вариантов построения резервированных вычислительных комплексов защищенных систем критического применения, в том числе работающих в распределенных системах реального времени.

Работа выполнена на кафедре вычислительной техники НИУ ИТМО в рамках НИР «Разработка методов и средств системотехнического проектирования информационных и управляющих вычислительных систем распределенной архитектуры».

Литература

- ГОСТ Р МЭК 61508-2-2007. Функциональная безопасность систем электрических, электронных, программируемых электронных, связанных с безопасностью. Часть 2. Требования к системам. – Введ. 01.09.2008. – М.: Стандартинформ, 2008. – 68 с.
- ГОСТ Р МЭК 61508-4-2007. Функциональная безопасность систем электрических, электронных, программируемых электронных, связанных с безопасностью. Часть 4. Термины и определения. – Введ. 01.06.2008. – М.: Стандартинформ, 2008. – 27 с.
- 3. Половко А.М., Гуров С.В. Основы теории надежности. СПб: БХВ-Петербург, 2006. 704 с.
- Shunji Osaki, Toshihiko Nishio. Reliability Evaluation of Some Fault-Tolerant Computer Architectures. Springer-Verlag, Berlin Heidelberg, New York, 1980. –129 p.
- 5. Sorin D. Fault Tolerant Computer Architecture. Morgan & Claypool, 2009. 116 p.
- Богатырев В.А., Башкова С.А., Беззубов В.Ф., Полякова А.В., Котельникова Е.Ю., Голубев И.Ю. Надежность дублированных вычислительных комплексов // Научно-технический вестник СПбГУ ИТМО. – 2011. – № 6 (76). – С. 74–78.

- Богатырев В.А., Демидов Д.В., Алексанков С.М. Оценка надежности дублированных комплексов с учетом контроля // Materiali VII mezinarodni vedecko-prakticka konference Aktyalni vymozenosti vedy. – 2011. Чехия, Прага 27.06.2011–05.07.2011. – Прага: Education and science, 2011. – С. 57–58.
- Bogatyrev V.A. Exchange of Duplicated Computing Complexes in Fault Tolerant Systems // Automatic Control and Computer Sciences. 2011. V. 46. № 5. P. 268–276.
- 9. Богатырев В.А., Бибиков С.В. Оценка функциональной безопасности систем, связанных с безопасностью // Технико-технологические проблемы сервиса. – 2011. – № 4. – С. 45–47.
- 10. Богатырев В.А. Надежность вариантов размещения функциональных ресурсов в однородных вычислительных сетях // Электронное моделирование. – 1997. – № 3. – С. 21–27.
- 11. Богатырев В.А., Богатырев С.В. Объединение резервированных серверов в кластеры высоконадежной компьютерной системы // Информационные технологии. 2009. № 6. С. 41–47.
- 12. Богатырев В.А. Надежность функционально-распределенных резервированных структур с иерархической конфигурацией узлов // Изв. вузов. Приборостроение. 2000. № 4. С. 67–70.
- 13. Богатырев В.А. Надежность отказоустойчивых вычислительных систем реального времени, компонуемых из многофункциональных модулей // Информационные технологии. 2000. № 10. С. 11–16.
- 14. Богатырев В.А. Отказоустойчивые кластеры дублированных вычислительных комплексов // Информационные технологии. 2012. № 1. С. 9–15.

Богатырев Владимир Анатольевич	- Санкт-Петербургский национальный исследовательский университет
	информационных технологий, механики и оптики, доктор технических
	наук, профессор, Vladimir.bogatyrev@gmail.com
Алексанков Сергей Михайлович	- Санкт-Петербургский национальный исследовательский университет
-	информационных технологий, механики и оптики, студент, alek-
Демиоов Даниил Валентинович	– Санкт-петероургский национальный исследовательский университет
	информационных технологий, механики и оптики, студент,
	daniil.demidov@gmail.com
Беззубов Владимир Федорович	- Санкт-Петербургский национальный исследовательский университет
	информационных технологий, механики и оптики, аспирант,
	bezzubov_vf@mail.ru

УДК 004.451 АНАЛИЗ ЭФФЕКТИВНОСТИ ПРИМЕНЕНИЯ МЕТОДОВ ХЕШИРОВАНИЯ В ФАЙЛОВЫХ СИСТЕМАХ, РАБОТАЮЩИХ В ПОЛЬЗОВАТЕЛЬСКОМ РЕЖИМЕ

Е.Ю. Иванов, М.С. Косяков

Рассмотрены особенности протоколов взаимодействия виртуальной файловой системы и файловой системы, их влияние на производительность микроядерных операционных систем. На примере файловой системы ext2, реализованной для операционной системы MINIX 3, установлено, что время идентификации файлового объекта в микроядерных операционных системах может увеличиваться до 26 раз по сравнению с монолитными системами. Показано, что при использовании рекомендованных в работе методов хеширования можно добиться производительности рассматриваемой части стека ввода–вывода, практически не уступающей производительности монолитных операционных систем. Ключевые слова: файловая система, операционная система, микроядерная ОС, методы хеширования, протоколы

Ключевые слова: файловая система, операционная система, микроядерная ОС, методы хеширования, протоколь взаимодействия ВФС и ФС, MINIX 3, PUFFS.

Введение

Несмотря на достижения современных исследований в области надежности операционных систем (OC) и верификации программного обеспечения, отказы OC, вызванные аппаратными и программными ошибками, по-прежнему приводят к снижению доступности прикладных сервисов, в том числе в системах стратегического назначения [1]. По некоторым оценкам, в OC на каждую тысячу строк кода приходится до шести ошибок [2, 3]; размер же современных OC превышает 6 000 000 строк кода. Согласно [4], до 85% ошибок находится в коде драйверов устройств и файловых систем (ФС), причем в случае монолитных ядер OC любая из них потенциально может привести к краху всей системы. Считается, что переход к модульным микроядерным OC, в которых основной функционал реализован в виде системных процессов, работающих в непривилегированном пользовательском режиме, позволяет решить данную проблему [5].

Отказоустойчивость микроядерных OC основана на принципе минимальных привилегий и изоляции ошибок в системных компонентах (модулях), в частности, в драйверах и ФС: при отказе одного из компонентов ядро OC продолжает функционирование и в состоянии перезапустить сбойный системный процесс. Время восстановления одного компонента значительно меньше, чем время восстановления всей системы, благодаря чему повышается коэффициент готовности. Кроме того, такая структурная организация ОС позволяет увеличить время наработки на отказ, так как с точки зрения прикладного сервиса отказ драйвера или ФС с последующим перезапуском соответствующего компонента в большинстве случаев может остаться вообще незамеченным.

Функционирование драйверов и ФС в виде отдельных процессов пользовательского режима приводит к потере производительности микроядерных ОС по сравнению с монолитными ОС [6]. Одной из причин снижения производительности является отсутствие возможности совместного использования служебных структур данных различными компонентами ОС. В микроядерных ОС системные процессы вынуждены явным образом обмениваться данными, используя средства ядра ОС, и осуществлять преобразование данных из одного формата в другой при их передаче согласно специальным протоколам взаимодействия. В частности, подобные накладные расходы появляются в микроядерных ОС из-за выделения виртуальной файловой системы (ВФС) и файловой системы в независимые процессы. При этом эффективность механизма их взаимодействия становится основным фактором, определяющим производительность этой части стека ввода–вывода в сравнении с монолитными ОС.

С целью увеличения производительности микроядерных ОС в данной работе проанализирована эффективность применения различных методов хеширования в ФС, работающих в пользовательском режиме, и рассмотрена целесообразность предоставления ВФС дополнительной информации о структурах данных, поддерживаемых ФС, как, например, в протоколе PUFFS взаимодействия ВФС и ФС для ОС NetBSD [7]. Для этого в микроядерной ОС MINIX 3 авторами реализована ФС ext2, что позволило, кроме прочего, провести сравнение производительности рассматриваемой части стека ввода-вывода в микроядерных ОС с аналогичными решениями для монолитных систем, а именно с ФС ext2, работающей в ОС Linux [8]. Показано, что при использовании рекомендованных в работе методов хеширования можно добиться производительности рассматриваемой части стека ввода-вывода микроядерных ОС, практически не уступающей производительности монолитных ОС.

Структурная организация стека ввода-вывода

На рис. 1 показана структурная организация стека ввода-вывода в монолитных и микроядерных ОС. При обращении к файлу пользовательский процесс выполняет системный вызов, передавая ОС дескриптор файла *fd* (*file descriptor*). Для монолитных ОС система переводится в режим ядра, а управление передается ВФС, которая по *fd* находит объект *vnode* (*virtual node*), содержащий высокоуровневое описание файла. Среди прочего, *vnode* предоставляет указатель на структуру данных ФС *istruct* (*inode struct*), описывающую сам файловый объект. ВФС использует этот указатель на *istruct* при вызове соответствующего обработчика запроса ФС. Таким образом, ВФС идентифицирует запрашиваемый файловый объект посредством прямого доступа к *istruct*. Данное описание применимо к Linux [8] и к большинству других UNIX-подобных ОС.

В микроядерных ОС ВФС и ФС представляют собой отдельные независимые процессы, работающие в пользовательском режиме. Как и в монолитной системе, ВФС использует *fd* для поиска соответствующего *vnode*. Затем, согласно протоколу взаимодействия ВФС и ФС, ВФС осуществляет запрос к ФС, используя некоторый идентификатор файла *id*. Использование в качестве *id* непосредственно адреса *istruct* является нежелательным, так как в этом случае нарушается принцип изоляции, и ФС не сможет перемещать *istruct* в памяти между вызовами со стороны ВФС. Тем не менее, некоторые протоколы взаимодействия ВФС и ФС, например, протокол PUFFS, реализованный для ОС NetBSD, в целях повышения производительности позволяют ВФС хранить и передавать адреса структур данных *istruct*, поддерживаемых ФС [7].

Другие протоколы, такие как протокол взаимодействия ВФС и ФС в ОС MINIX 3, в качестве *id* используют числовой идентификатор *inum*, получаемый от ФС при создании или открытии файла [9]. Использование идентификатора *inum*, отличного от адреса *istruct*, приводит к необходимости хранения в ФС таблицы соответствий *inum* адресам *istruct* и поиска по ней при идентификации файлового объекта. Так как эта операция используется при каждом обращении к ФС, ее неэффективная реализация может сказаться на производительности файловой системы. Стоит отметить, что данная проблема отличается от поиска *vnode* по значению файлового дескриптора *fd* в ВФС. В последнем случае количество хранимых *fd*, среди которых осуществляется поиск, соответствует числу файлов, открытых одним вызывающим процессом, т.е. относительно невелико. По этой причине операция поиска *vnode* по *fd* практически не оказывает влияния на время обслуживания запроса в ВФС. Поиск же адреса *istruct* по *inum* осуществляется среди всех идентификаторов файлов, открытых всеми процессами. Кроме того, ФС обычно кэширует сами структуры *istruct* на случай повторного обращения к недавно закрытым файлам, что приводит к дополнительному увеличению пространства поиска и, как следствие, времени выполнения запроса.

В связи с этим в работе проводится анализ эффективности применения различных методов хеширования при реализации таблицы соответствий *inum* адресам структур данных *istruct*. Измеряемой характеристикой является среднее время выполнения эталонных задач, представляющих собой интегральные нагрузки, широко используемые для оценки производительности стека ввода-вывода OC [10]. В качестве характеристики производительности используется величина, обратная времени выполнения, что соответствует числу задач, исполненных в единицу времени.



Рис. 1. Структурная организация стека ввода-вывода в монолитных и микроядерных ОС

Параметры проведения экспериментов

Эксперименты проводились на реализованной в работе файловой системе ext2 для микроядерной OC MINIX 3, функционирующей на персональном компьютере с процессором Intel Pentium 4 2.80 ГГц, оперативной памятью DDR 512 МБ и диском Seagate 40 ГБ 7200 грт. Рассматривались следующие типы эталонных пользовательских нагрузок:

- «Сборка». Сборка ОС MINIX 3 из каталога /usr/src, включая ядро, драйверы, ФС и библиотеки. Полное время выполнения этой задачи T составляло около 40 мин.
- «Копирование». Копирование дерева каталогов с исходными файлами /usr/src на целевую файловую систему ext2. Общий объем данных составлял порядка 170 МБ, число файлов 24 694 (1 620 каталогов), средняя глубина каталогов равнялась четырем. Время *T* равнялось примерно 20 с.
- «Статистика». Вызов системной команды stat для всех файлов из каталога /usr/src в случайном порядке. В отличие от нагрузки «Копирование», в результате которой происходят последовательные обращения к диску, в нагрузке «Статистика» обращения к блокам диска осуществляются в произвольном порядке. Время T равнялось примерно 50 с.

Медленные дисковые операции значительно более сильно влияют на производительность стека ввода-вывода, чем вычислительные операции внутри ФС, что затрудняет измерение времени т поиска адреса *istruct* по числовому идентификатору файла *inum*. Кроме того, при учете времени выполнения дисковых операций резко увеличивается длительность проведения экспериментов. В связи с этим авторами, в том числе, анализировалась работа указанной части стека ввода-вывода изолированно без учета дисковых операций. Функционирование такой части осуществлялось в специальном программном окружении, имитирующем остальные компоненты ФС. Для этого случая нагрузка на исследуемую часть стека ввода-вывода создавалась путем воспроизведения записанных ранее операций ФС, сформированных при выполнении перечисленных выше пользовательских задач.

С целью повышения точности измерений функция Hyper-Threading процессора была отключена. Процессу, исполняющему код исследуемой части стека ввода-вывода, был присвоен наивысший абсолютный приоритет, что исключало прерывание его работы для предоставления процессорного времени другим процессам. Подкачка страниц виртуальной памяти была запрещена. В качестве *inum* и адреса *istruct* в протоколе взаимодействия ВФС и ФС в ОС MINIX 3 используются целые 32-разрядные числа, поэтому их обработка на 32-битном процессоре не требует дополнительных накладных расходов.

Анализ эффективности методов хеширования

Результаты экспериментов показали, что при линейном поиске адреса *istruct* по *inum* и при ограничении максимального числа *n* хранимых структур *istruct* значениями до 512 время идентификации файлового объекта т увеличивается до 26 раз по сравнению со временем, наблюдаемым при прямой адресации, используемой в монолитных системах. Для увеличения производительности рассматриваемой части стека ввода–вывода рассмотрена эффективность применения различных методов хеширования при реализации таблицы соответствий *inum* адресам структур данных *istruct*. Производительность хеш-таблиц напрямую зависит от выбора хеш-функции, метода разрешения коллизий и размера хеш-таблицы. В работе использованы эталонные реализации хеш-таблиц на основе открытой адресации с квадратичной последовательностью проб (Google dense) и двойным хешированием (GNU libiberty). Для оценки эффективности хеш-таблиц с разрешением коллизий при помощи цепочек использована собственная реализация, позволяющая, в отличие от других, полностью контролировать все параметры вычисления. Например, в популярной реализации *hash_map*, входящей в GNU GCC, размер хеш-таблицы динамически изменяется в зависимости от коэффициента заполнения α. Кроме того, результаты экспериментов показали, что собственная реализация работает быстрее, чем *hash_map* при аналогичных параметрах.

Для отображения множества возможных значений *inum* в *m*-элементное множество хеш-значений, где *m* – размер хеш-таблицы, используются следующие хеш-функции *h*:

- модульное хеширование «MOD»: *h*(*inum*) = *inum mod m*. Применяется в GNU libiberty и *hash_map*;

– модульное хеширование «AND» для случаев, когда *m* является степенью двойки: *h(inum) = inum and* (*m* – 1). Данная функция проста в вычислительном плане и дает распределение, схожее с распределениями модульных хеш-функций общего вида [11]. Используется в Google dense и в авторской реализации хеш-таблицы.

Собственная реализация хеш-таблицы, кроме хеш-функции «AND», поддерживает хеш-функцию «MOD», мультипликативное и универсальное хеширование, популярную хеш-функцию «MurmurHash» [12]. Наши результаты показали, что, несмотря на большую вычислительную сложность, перечисленные функции не дают заметного преимущества в распределении хеш-значений по сравнению с хеш-функцией «AND».

В таблице приведено среднее время т идентификации файлового объекта при реализации различных типов файловых нагрузок для случая n = 512. Размер хеш-таблиц для Google dense и реализации авторов m = 2048, для libiberty m = 2039. Разница между наблюдаемыми минимальными и максимальными значениями т составляла не более 7% от представленных средних значений.

	Таблицы соответствий inum адресам istruct							
Тип нагрузки	Google dense	libiberty	Авторская реализация	Прямая адресация	Линейный поиск			
«Сборка»	354	410	214	165	4315			
«Копирование»	21	24	11	9	237			
«Статистика»	10	13	6	5	104			

Таблица. Среднее время идентификации файлового объекта т, мс

Согласно полученным результатам, для случая работы с файлами разрешение коллизий на основе цепочек работает быстрее, чем открытая адресация. Преимущество метода «цепочек» сохраняется также при изменении размеров хеш-таблиц. Это объясняется высокой интенсивностью открытия и закрытия файлов при выполнении нагрузок, что требует частого удаления элементов из таблицы соответствий. При использовании открытой адресации удаленные элементы продолжают храниться в хеш-таблице, тем самым увеличивая время поиска. Более низкая производительность открытой адресации с двойным хешированием по сравнению с квадратичной последовательностью проб указывает на то, что время, затрачиваемое на вычисление дополнительных хеш-значений, превышает потери времени на дополнительных пробах [13]. Кроме того, видно, что при использовании авторской реализации хеш-таблицы с разрешением коллизий при помощи цепочек время идентификации файлового объекта лишь немногим превосходит соответствующее время, наблюдаемое при прямой адресации, используемой в монолитных системах. В этой связи в дальнейшем в работе рассматривается именно авторская реализация хеш-таблицы.

На рис. 2 показана зависимость времени т от размера *m* хеш-таблицы для нагрузки «Сборка» при ограничении на число хранимых *istruct* n = 512. Графики, представленные пунктиром, получены при кэшировании последнего найденного сопоставления *inum* адресу *istruct*, что позволяет реже выполнять поиск в таблице соответствий. При значениях $m \ge 4096$, соответствующих коэффициенту заполнения $\alpha = n / m \le 0.25$, значение т становится практически идентичным результатам, полученным при использовании прямой адресации. Эксперименты показали, что данная зависимость справедлива и для других значений *n* при $\alpha \le 0.25$. Кроме того, при таких значениях коэффициента заполнения α выигрыш от кэширования результатов последней операции поиска становится совсем незначительным. Перечисленные утверждения остаются верными и для других файловых нагрузок, так как идентификация файлового объекта необходима при выполнении любой файловой операции.

Результаты измерения полного времени T выполнения пользовательских задач с учетом дисковых операций показали, что при небольших $n \le 1024$ значения $T^{\Pi\Pi}$, получаемые для случая линейного поиска адреса *istruct* по *inum* со средним временем O(n), практически совпадают со значениями T^{XT} для случая применения хеш-таблиц со средним временем поиска O(1). Это объясняется тем, что медленные диско-

вые операции оказывают значительно большее влияние на время T, чем вычислительные операции внутри ΦC . Однако при увеличении числа n хранимых в памяти *istruct* отношение времен $T^{\Pi\Pi}/T^{XT}$ выполнения пользовательских задач для этих случаев также возрастает и достигает двух раз уже при $n \approx 4096$.



Рис. 2. Зависимость среднего времени идентификации файлового объекта т от размера хеш-таблицы *m* на примере нагрузки «Сборка»

Заключение

В работе рассмотрены особенности протоколов взаимодействия ВФС и ФС и их влияние на производительность микроядерных ОС, обладающих повышенной надежностью в сравнении с монолитными системами. Для детального изучения этой части стека ввода–вывода была разработана реализация файловой системы ext2, функционирующая в микроядерной ОС MINIX 3. С ее помощью было показано, что протоколы взаимодействия ВФС и ФС, использующие числовой идентификатор *inum* для поиска файловой структуры *istruct*, описывающей запрашиваемый файловый объект, привносят значительные накладные расходы. А именно, время идентификации файлового объекта увеличивается до 26 раз по сравнению с монолитными системами, что особенно заметно при рассмотрении работы ФС без учета дисковых операций, например, при доступе к блокам файла, находящимся в кэше ФС.

В качестве возможного пути решения этой проблемы в работе рассматривается подход, реализованный в протоколе PUFFS взаимодействия ВФС и ФС для ОС NetBSD, позволяющий ВФС хранить адреса структур данных, поддерживаемых ФС, в частности адреса *istruct*. К сожалению, данный подход нарушает принцип изоляции компонентов микроядерных ОС и ведет к существенной зависимости функционирования ВФС от работы ФС, тем самым понижая надежность ОС.

В этой связи в работе предлагается использовать альтернативный путь увеличения производительности рассматриваемой части стека ввода-вывода, позволяющий сохранить изоляцию ВФС от ФС и основанный на применении хеш-таблиц в ФС для поиска адресов *istruct* по *inum*. Результаты измерений времени идентификации файлового объекта при использовании различных методов хеширования показали допустимость применения простой в вычислительном плане модульной хеш-функции при размере хеш-таблицы $m = 2^x$. Было установлено, что для рассматриваемого случая работы с файлами разрешение коллизий при помощи цепочек работает быстрее, чем метод открытой адресации. Показано, что использование указанных методов хеширования позволяет достичь сопоставимых, а при увеличении размера хеш-таблицы – практически идентичных значений времени идентификации файловых объектов для микроядерных ОС в сравнении с монолитными системами. Это позволяет авторам сделать вывод о нецелесообразности применения подхода, реализованного в протоколе PUFFS.

Анализ работы всего стека ввода-вывода микроядерной ОС MINIX 3 с учетом дисковых операций показал, что при ограничении максимального числа хранимых структур *istruct* значениями до 1024 добавление хеш-таблиц не привносит заметного увеличения производительности. Это связано с тем, что медленные дисковые операции значительно сильнее влияют на производительность стека ввода-вывода, чем вычислительные операции внутри файловой системы. Однако с ростом числа хранимых в памяти *istruct*, вызванным как увеличением максимального числа открытых файлов, так и кэшированием самих структур *istruct* на случай повторного обращения к недавно закрытым файлам, использование хештаблицы становится необходимым.

Реализация ext2 для MINIX 3 выполнена при поддержке свободного университета Амстердама (VU University Amsterdam) и входит в состав операционной системы MINIX, начиная с версии 3.1.8. Авторы выражают признательность своему коллеге, научному программисту университета VU T. Veerman за плодотворное сотрудничество по тематике работы и полезные замечания.

Литература

- 1. Титов А.В. Методика оценки надежности встроенных программных средств при редких отказах // Изв. вузов. Приборостроение. 2010. Т. 53. № 10. С. 38–41.
- 2. Ostrand T.J., Weyuker E.J. The distribution of faults in a large industrial software system // SIGSOFT Softw. Eng. Notes. - 2002. - V. 27. - № 4. - P. 55-64.
- 3. Ostrand T.J., Weyuker E.J., Bell R.M. Where the bugs are // SIGSOFT Softw. Eng. Notes. 2004. V. 29. № 4. P. 86–96.
- Chou A., Yang J., Chelf B. et al. An empirical study of operating systems errors // SIGOPS Oper. Syst. Rev. - 2001. - V. 35. - № 5. - P. 73-88.
- Herder J.N., Bos H., Gras B. et al. Construction of a Highly Dependable Operating System // Proceedings of the Sixth European Dependable Computing Conference. EDCC '06. Washington, DC, USA: IEEE Computer Society. – 2006. – P. 3–12.
- Leslie B., Chubb P., Fitzroy-dale N. et al. User-level Device Drivers: Achieved Performance // Journal of Computer Science and Technology. – 2005. – V. 20. – P. 654–664.
- Kantee A. puffs Pass-to-Userspace Framework File System // Proceedings of the AsiaBSDCon. 2007. P. 29–42.
- 8. Лав Р. Ядро Linux. Описание процесса разработки: Пер. с англ. 3-е изд. М.: Вильямс, 2013. 496 с.
- 9. The VFS-FS protocol. MINIX 3 Developers Guide [Электронный ресурс]. Режим доступа: http://wiki.minix3.org/en/DevelopersGuide/VfsFsProtocol, своб. Яз. англ. (дата обращения: 02.03.2013).
- 10. Traeger A., Zadok E., Joukov N., Wright C.P. A nine year study of file system and storage benchmarking // ACM Transactions on Storage (TOS). 2008. V. 4. № 2. P. 1–56.
- 11. Седжвик Р. Фундаментальные алгоритмы на С++. Части 1-4: Анализ/Структуры данных/Сортировка/Поиск: Пер. с англ. Киев: Издательство «ДиаСофт», 2001. 688 с.
- 12. Appleby A. Murmurhash v3 [Электронный ресурс]. Режим доступа: http://sites.google.com/site/murmurhash/, свободный. Яз. англ. (дата обращения 16.03.2013).
- 13. Кнут Д.Э. Искусство программирования. Т. 3. Сортировка и поиск: Пер. с англ. 2-е изд. М.: Вильямс, 2000. 832 с.

Иванов Евгений Юрьевич

Косяков Михаил Сергеевич

 Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, студент i@eivanov.com
 Санкт-Петербургский национальный исследовательский университет информаци-

онных технологий, механики и оптики, кандидат технических наук, доцент, mkosyakov@gmail.com

УДК 004.738

ОПТИМИЗАЦИЯ РАСПРЕДЕЛЕНИЯ ЗАПРОСОВ МЕЖДУ КЛАСТЕРАМИ ОТКАЗОУСТОЙЧИВОЙ ВЫЧИСЛИТЕЛЬНОЙ СИСТЕМЫ В.А. Богатырев, А.В. Богатырев, И.Ю. Голубев, С.В. Богатырев

Предложена оценка надежности распределенных вычислительных систем, предусматривающих перераспределение запросов при изменениях потоков запросов, отказах и отключениях узлов системы, объединяемых в совокупность кластеров. Предложена и решена задача оптимизации процесса перераспределения запросов между кластерами с учетом его влияния на задержки обслуживания и надежность системы.

Ключевые слова: оптимизация, надежность, перераспределение запросов, кластер, отказоустойчивость.

Введение

Повышение отказоустойчивости, надежности и производительности распределенных вычислительных систем, объединяющих в единую систему множество отдельных кластеров [1–3], достигается в результате динамического перераспределения запросов [4–7] между ними с учетом изменений загруженности кластеров, отказов и временных отключений их узлов.

В распределенной инфраструктуре [1–3], консолидирующей множество ресурсов, объединенных в кластеры, перераспределение запросов (нагрузки) может осуществляться между узлами как одного, так и различных кластеров, соединенных через сеть. При перераспределении запросов между кластерами увеличиваются издержки на взаимосвязь через сеть, но возрастают возможности балансировки загрузки и адаптации к отказам и отключениям узлов, что обусловливает актуальность оптимизации процесса распределения запросов [8, 9].

Задача оптимизации системы

Объектом исследования является распределенная вычислительная система (рис. 1), включающая *М* локальных кластеров и общедоступный кластер, объединяющий *m* серверов.



Рис. 1. Структура распределенной системы

В результате перераспределения запросов от локальных кластеров в общедоступный кластер обеспечивается сбалансированность нагрузки узлов системы и устойчивость системы к отказам и перегрузкам серверов локальных кластеров. Перераспределение запросов от некоторого локального кластера, содержащего в исходном (до отказов) состоянии *n* серверов, в общедоступный кластер осуществляется через *N* резервированных коммутационных узлов (маршрутизаторов или коммутаторов) [9].

При оптимизации структуры определяется число (кратность резервирования) серверов в локальных кластерах n и в общем кластере m, а также число коммутационных узлов N, обеспечивающие наибольшую надежность системы P при заданных ограничениях на стоимость построения системы s. При оценке надежности системы, в отличие от [9], где условие работоспособности системы сформулировано как требование сохранения в каждой подсистеме хотя бы одного узла, в предлагаемой работе учитываются нижние ограничения на число узлов в подсистемах, при которых не возникают перегрузки соответствующих кластеров.

При оптимизации процесса распределения запросов с учетом возможности отказов и отключений узлов общедоступного кластера будем считать заданными средние времена выполнения запросов в серверах кластеров и в коммутационных узлах v_0, v_1 , их интенсивности отказов λ_0, λ_1 , и восстановлений μ_0, μ_1 . Будем считать известными вероятности *r* нахождения во включенном состоянии серверов общедоступного кластера. Оптимизация проводится при заданной интенсивности потока запросов λ , поступающего в локальный кластер и при необходимости перераспределяемого через сеть в общедоступный кластер, на который от других кластеров системы через сеть дополнительно направляется поток запросов с интенсивностью $\Lambda = \beta\lambda$.

При оптимизации структуры будем считать стоимости серверов локальных и общедоступного кластера, а также стоимость коммутационных узлов соответственно равными c_0, c_1, c_2 .

В результате оптимизации процесса распределения потока запросов, поступающего в локальный кластер, ищется их доля, перераспределяемая через сеть в общедоступный кластер, при которой минимизируется среднее время пребывания запросов *T*.

Отличие предлагаемой задачи оптимизации распределения запросов от [9] заключается в учете возможностей отказов, восстановлений и отключений серверов общедоступного кластера в процессе функционирования. Учет возможности отключения серверов общедоступного кластера обусловлен тем, что предоставляемые им услуги по обслуживанию внешних для него запросов могут проводиться в фоновом режиме и поэтому могут отбрасываться при высокой нагрузке серверов, при решении важных для владельца кластера (сервера) задач, при профилактическом обслуживании или временных отключениях узлов по другим причинам.

Оценка надежности системы

Определим вероятность работоспособности системы для локального кластера из n серверов с учетом возможности использования в качестве резерва ресурсов m серверов общедоступного кластера, связь с которым обеспечивается через N коммутационных узлов.

Предположим, что пропускная способность каждого коммутационного узла достаточна, чтобы не ограничивать возможности перераспределения запросов, т.е. если исправен хотя бы один коммутационный узел, то запросы могут перераспределяться в общедоступный кластер, но для реализации такого пе-

рераспределения в локальном кластере должен быть исправен хотя бы один вычислительный узел. С учетом этих условий вероятность работоспособности системы составляет

$$P = (1 - P_1) \left[\sum_{i=a}^{n} C_n^i p_0^i (1 - p_0)^{n-i} \right] + P_1 \sum_{j=b}^{n+m} \left(C_{m+n}^j - d_j C_m^j \right) p_0^j (1 - p_0)^{n+m-j},$$
(1)

где $d_j = 1$, если $j \le m$, иначе j = 0; $P_1 = \sum_{i=1}^{N} C_N^i p_1^i (1 - p_1)^{N-i}$ – вероятность исправности коммутационной

подсистемы, при этом из соображений отсутствия перегрузки кластеров значения *a* и *b* определяются как ближайшие целые, бо́льшие λv_0 и $\lambda (1+\beta) v_0$.

Надежность узлов определим по коэффициентам готовности, вычисляемым для серверов и коммутационных узлов соответственно как [10, 11]

 $p_0 = \mu_0 / (\lambda_0 + \mu_0); p_1 = \mu_1 / (\lambda_1 + \mu_1).$

Формула (1) не учитывает возможность случайных временных отключений серверов общедоступного кластера, с учетом доступности серверов с вероятностью *r* имеем

$$P = (1 - P_1) \left[\sum_{i=a}^n C_n^i p_0^i (1 - p_0)^{n-i} \right] + P_1 \left[\sum_{i=1}^n C_n^i p_0^i (1 - p_0)^{n-i} \right] \sum_{j=b}^m C_m^j p_2^j (1 - p_2)^{m-j},$$

где $p_2 = rp_0$. Для систем критического применения, не допускающих наличие узлов, отказ которых может вызвать отказ системы, в качестве базовых средств вычислений используются резервированные вычислительные комплексы [12]. Простейшая структура дублированного вычислительного комплекса (ДВК), скомплектованная из двух связанных через адаптер сопряжения (АС) полукомплексов, включающих процессоры (П) и модули памяти (М), представлена на рис. 2, а. Модель надежности ДВК, допускающего возможность совместной работы процессора и модуля памяти разных полукомплексов, сводится к хорошо изученной в теории надежности модели мостиковой схемы [10, 11], приведенной на рис. 2, б.

Надежность (коэффициент готовности) ДВК, в соответствии с моделью по рис. 2, б, вычисляется как $p_0 = p_a (1 - (1 - p_p)^2) (1 - (1 - p_M)^2) + (1 - p_a) (1 - (1 - p_p p_M)^2)$,

где при заданных интенсивностях отказов $\lambda_p, \lambda_M, \lambda_a$ и восстановлений μ_p, μ_M, μ_a процессора, памяти и

адаптера сопряжения соответственно имеем $p_p = \mu_p / (\lambda_p + \mu_p), p_M = \mu_M / (\lambda_M + \mu_M), p_a = \mu_a / (\lambda_a + \mu_a).$

В случае невозможности совместной работы процессоров и модулей памяти разных полукомплексов надежность ДВК вычислим как

 $p_0 = (1 - (1 - p_p p_M)^2).$



Рис. 2. Структура (а) и модель надежности (б) ДВК

Для ДВК с ограниченным восстановлением (одновременный ремонт нескольких узлов невозможен) коэффициент готовности определяется как сумма вероятностей работоспособных состояний, для нахождения которых процесс отказов и восстановлений представляется марковским процессом, при этом составляется граф переходов и уравнения Чепмена–Колмогорова, в результате решения которых и определяются искомые вероятности. При оценке вероятностей работоспособных состояний и коэффициента готовности ДВК по рис. 2, а, могут использоваться результаты, полученные в [13].

Оптимизация структуры

При оптимизации структуры рассматриваемой вычислительной системы ищется число серверов *n* в локальных кластерах, число серверов *m* в общедоступном кластере и кратность резервирования *N* коммутационных узлов, обеспечивающие максимум надежности системы, $P = \max_{m,n,N=1}^{n} P(m, n, N, g, \lambda)$, при огра-

ничении стоимости *s* ее реализации $(Mc_0n + c_1N + c_2m) \le s$, и условия стационарности функционирования узлов (отсутствия перегрузки узлов).

1

Поиск максимума *P* может основаться на переборе, реализуемом с использованием средств системы компьютерной математики Matchcad-15.

Целью оптимизации структуры может быть минимизация среднего времени пребывания запросов в системе [14] при ограничении средств *s* на ее построение, $T = \min_{m,n,N,g} T(m,n,N,g,\lambda)$, при этом среднее время пребывания запросов в системе вычисляется [9] как

$$T = g\left(\frac{v_0}{1 - \frac{g\lambda v_0}{n}}\right) + (1 - g)\left(\frac{2v_1}{1 - \frac{((1 - g) + \beta)2\lambda v_1}{N}} + \frac{v_0}{1 - \frac{((1 - g) + \beta)\lambda v_0}{m}}\right),$$
(2)

где (1 - g) – средняя доля запросов, перераспределяемых через сеть от локального кластера в общедоступный. При поиске оптимального *g* необходимо учитывать условие стационарного режима функционирования узлов (условие отсутствия перегрузки узлов) [9]:

$$\left(\frac{g\lambda v_0}{n} < 1\right) \land \left(\frac{\left(\left(1-g\right)+\beta\right)2\lambda v_1}{N} < 1\right) \land \left(\frac{\left(\left(1-g\right)+\beta\right)\lambda v_0}{m} < 1\right).$$
(3)

При необходимости оптимизация может быть проведена по мультипликативному критерию $r(m,n,N,g,\lambda) = \max_{n} (P(m,n,N)/T(m,n,N,g,\lambda)).$

Оптимизация процесса перераспределения запросов

При заданной структуре системы (сформированной при рассмотренной выше структурной оптимизации) проведем оптимизацию процесса распределения запросов с учетом возможности отказов и отключений исправных узлов общедоступного кластера с вероятностью (1 - r). Оптимизация проводится при заданной средней интенсивности потока запросов λ , поступающего в локальный кластер и при необходимости перераспределяемого через сеть в общедоступный кластер.



Рис. 3. Оптимальная доля запросов, перераспределяемых через сеть

В результате оптимизации процесса распределения потока запросов, поступающего в локальный кластер, ищется их доля, перераспределяемая через сеть в общедоступный кластер, при которой минимизируется среднее время пребывания запросов T. $T = \min T(m, n, N, g, \lambda)$, где при модернизации (2) и (3) имеем

$$T = g \left(\frac{v_0}{1 - \frac{g\lambda v_0}{n}} \right) + (1 - g) \left(\frac{2v_1}{1 - \frac{((1 - g) + \beta)2\lambda v_1}{N_c}} + \frac{v_0}{1 - \frac{((1 - g) + \beta)\lambda v_0}{m_c}} \right)$$
$$\left(\frac{g\lambda v_0}{n} < 1 \right) \land \left(\frac{((1 - g) + \beta)2\lambda v_1}{N_c} < 1 \right) \land \left(\frac{((1 - g) + \beta)\lambda v_0}{m_c} < 1 \right)$$

при математических ожиданиях числа коммутационных узлов N_c и доступных исправных серверов общедоступного кластера, вычисляемых как

$$N_{c} = \sum_{i=1}^{N} i C_{N}^{i} p_{1}^{i} (1-p_{1})^{N-i}, \ m_{c} = \sum_{j=1}^{m} j C_{m}^{j} p_{2}^{j} (1-p_{2})^{m-j},$$
$$\left(\frac{g\lambda v_{0}}{n} < 1\right) \wedge \left(\frac{\left((1-g) + \beta\right) 2\lambda v_{1}}{N_{c}} < 1\right) \wedge \left(\frac{\left((1-g) + \beta\right) \lambda v_{0}}{m_{c}} < 1\right).$$

Для примера проведем оптимизацию процесса распределения запросов при n = 8 шт., N = 5 шт., m = 23 шт.; $v_0 = 10$ с, $v_1 = 1$ с, r = 0.8; $\lambda_0 = \lambda_2 = 10^{-4}$ 1/ч, $\lambda_1 = 0.5 \cdot 10^{-4}$ 1/ч; $\mu_0 = \mu_1 = \mu_2 = 1$ 1/ч. Результаты поиска оптимальной доли (1 - g), распределяемых через сеть в общедоступный кластер запросов, в зависимости от интенсивности входного потока запросов λ 1/с представлены на рис. 3 при $\beta = 0.5$ и $\beta = 1$. Рост доли неперераспределяемых запросов g при незначительной интенсивности λ потока запросов объясняется влиянием дополнительных задержек при передаче запросов через сеть, а при значительной интенсивности λ – перегрузкой общедоступного кластера.

Заключение

Поставлены и решены задачи оптимизации структуры вычислительной системы и процесса перераспределения через сеть потока запросов от локальных кластеров в общедоступный кластер с учетом возможностей отказов, восстановлений и отключений серверов общедоступного кластера. Перераспределение запросов реализуется с целью минимизации среднего времени пребывания запросов при адаптации системы к отказам узлов и изменениям потока запросов.

Предложены модели надежности и массового обслуживания вычислительных систем динамического перераспределения запросов (нагрузки) между кластерами, которые могут быть использованы при оценке надежности и выборе рациональных вариантов организации перераспределения запросов в системах с объединением вычислительных ресурсов в локальные и общедоступные кластеры, связанные через сеть.

Работа выполнена на кафедре вычислительной техники НИУ ИТМО в рамках НИР «Разработка методов и средств системотехнического проектирования информационных и управляющих вычислительных систем распределенной архитектуры».

Литература

- 1. Таненбаум Э., Ван Стеен М. Распределенные системы. Принципы и парадигмы. СПб: Питер. 2003. 877 с.
- 2. Clark T. The New Data Center. New technologies are radically reshaping the data center. Brocade Bookshelf. San Jose, 2010. – 156 p.
- 3. Кармановский Н.С., Гатчин Ю.А., Терентьев А.О., Федоров Д.Ю., Беккер М.Я. Информационная безопасность при облачных вычислениях: проблемы и перспективы // Научно-технический вестник СПбГУ ИТМО. 2011. № 1 (71). С. 97–102.
- 4. Богатырев В.А. К повышению надежности вычислительных систем на основе динамического распределения функций // Изв. вузов. Приборостроение. 1981. № 8. С. 62–65.
- 5. Богатырев В.А. Распределение заданий в многомашинных вычислительных системах // Изв. вузов. Приборостроение. 1986. № 5. С. 43–47.
- 6. Богатырев В.А. Надежность функционально-распределенных резервированных структур с иерархической конфигурацией узлов // Изв. вузов. Приборостроение. 2000. № 4. С. 67–70.
- 7. Богатырев В.А. Надежность вычислительных систем с функциональной реконфигурацией на основе перераспределения задач // Информационные технологии. 2001. № 7. С. 22–27.
- 8. Богатырев В.А., Богатырев С.В. Объединение резервированных серверов в кластеры высоконадежной компьютерной системы // Информационные технологии. 2009. № 6. С. 41–47.
- Bogatyrev V.A., Golubev I.Y., Bogatyrev S.V. Optimization and the Process of Task Distribution between Computer System Clusters // Automatic Control and Computer Sciences. – 2012. – V. 46. – № 3. – P. 103– 111.
- 10. Гуров С.В., Половко А.М. Основы теории надежности. СПб: БХВ-Петербург, 2006. 704 с.
- 11. Черкесов Г.Н. Надежность аппаратно-программных комплексов. СПб: Питер, 2005. 479 с.
- 12. Bogatyrev V.A. Exchange of Duplicated Computing Complexes in Fault tolerant Systems // Automatic Control and Computer Sciences. 2011. V. 46. № 5. P. 268–276.
- Богатырев В.А., Башкова С.А., Беззубов В.Ф., Голубев И.Ю., Котельникова Е.Ю., Полякова А.В. Надежность дублированных вычислительных комплексов // Научно-технический вестник СПбГУ ИТМО. – 2011. – № 6. – С. 74–78.
- 14. Алиев Т.И. Основы моделирования дискретных систем. СПб: СПбГУ ИТМО, 2009. 363 с.

Богатырев Владимир Анатольевич	-	Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор техниче-
		ских наук, профессор, Vladimir.bogatyrev@gmail.com
Богатырев Анатолий Владимирович	-	Санкт-Петербургский национальный исследовательский университет
		информационных технологий, механики и оптики, аспирант,
		gangleon@gmail.com Vladimir.bogatyrev@gmail.com
Голубев Иван Юрьевич	_	Санкт-Петербургский национальный исследовательский университет
		информационных технологий, механики и оптики, аспирант,
		www.golubev@mail.ru
Богатырев Станислав Владимирович	_	ООО «Айти Хаус», главный инженер; Санкт-Петербургский нацио-
		нальный исследовательский университет информационных техноло-
		гий, механики и оптики, аспирант, realloc@gmail.com
		Vladimir.bogatyrev@gmail.com

УДК 004.421 АЛГОРИТМ И ПРОГРАММА ПОИСКА И ИССЛЕДОВАНИЯ М-МАТРИЦ Ю.Н. Балонин, М.Б. Сергеев

Рассматриваются алгоритм и программный комплекс поиска и исследования матриц ортогональных базисов – минимаксных матриц (М-матриц). Приведена схема алгоритма, даны комментарии к блокам расчета, пояснен интерфейс программного комплекса MMatrix, разработанного с участием авторов статьи. Результатом работы универсального алгоритма являются матрицы Адамара, матрицы Белевича (С-матрицы, conference matrices) и матрицы, дополняющие указанные и близкие к ним по свойствам четных и нечетных порядков, в частности, матрица 22-го порядка, для которого С-матрицы не существует. Приведены примеры портретов найденных альтернативных матриц 255-го и 257-го порядков, отвечающих последовательностям чисел Мерсенна и Ферма. Пояснен новый путь получения матриц Адамара, отличный от ранее известных переборных процедур и процедур, опирающихся на вычисление символов Лагранжа, имеющий теоретическое и прикладное значения.

Ключевые слова: ортогональные матрицы, М-матрицы, матрицы Адамара, матрицы Мерсенна, матрицы Ферма, численные методы, алгоритм вычислений.

Введение

Для применения в процедурах построения помехоустойчивых и защитных кодов, в маскировании информации необходимы оригинальные ортогональные базисы [1, 2], поиск которых возможен в исследовательской программной среде. В настоящей работе описывается программный комплекс MMatrix, разработанный с участием авторов специально для этой цели и зарегистрированный под названием «Программа поиска М-матриц» [3]. Научная концепция построения матриц ортогональных базисов, обладающих экстремальными минимаксными свойствами, рассмотрена в работах [4–7].

В классе ортогональных матриц особое место занимают матрицы спектрального преобразования Фурье, определенные над полем вещественных или комплексных чисел. Деление весьма условно, поскольку, так же как и в случае жордановых матриц собственных значений и собственных векторов, существуют комплексная и вещественная формы, связанные между собой. При переходе к вещественной форме столбцы, порожденные значениями четных и нечетных базисных функций и разнесенные в комплексной форме по отдельным составляющим, размещаются в одной матрице рядом. В итоге утрачивается одно из свойств комплексных матриц: значения синусов и косинусов кратных частот трактуются теперь уже не как вещественная и мнимая составляющие, вместе образующие катеты прямоугольного треугольника с гипотенузой единичной длины, а как самостоятельные значения. Иными словами, модуль каждого элемента в вещественной форме зависит от периода дискретизации и не равен, в общем случае, единице. Тем не менее, частные формы вещественных ортогональных матриц с единичными нормами элементов возможны. Согласно гипотезе Адамара, порядок этих матриц кратен 4.

С появлением вычислительных средств преимущества, которые предполагают вычисления со столь просто устроенными матрицами, обеспечили серьезный интерес к исследованию матриц Адамара – именно так они были названы. К последовательности матриц порядков 2, 4, 8, 16, 32 и т.п., порождаемых процедурой удвоения порядка от 1, предложенной еще Сильвестром, Адамар добавил еще две стартовые матрицы пропущенных порядков 12 и 20. Процесс нахождения таких матриц плохо поддается формализации, компьютерный поиск отчасти подтвердил гипотезу Адамара, но порядок матрицы, подлежащий проверке, не превышает пока тысячи. Дальнейший прогресс в этом направлении возможен при включении в область исследований матриц с элементами нескольких значений, в том числе и пропущенных нечетных порядков включительно.

Минимаксные ортогональные матрицы (М-матрицы)

В работе [7] введено определение уровней матрицы, которым соответствуют значения ее элементов. Введение уровней позволяет, во-первых, классифицировать матрицы, во-вторых – представлять их графические портреты.

Определение 1. Матрица Адамара – квадратная двухуровневая матрица \mathbf{H}_n порядка *n*, кратного 4, состоящая из элементов из множества $\{1, -1\}$, столбцы которой ортогональны:

 $\mathbf{H}_{n}^{T}\mathbf{H}_{n}=n\mathbf{I},$

где I – единичная матрица.

Определение 2. Матрица Белевича (C-matrix, conference-matrix) – квадратная трехуровневая матрица C_n порядка *n*, кратного 2, с нулевой диагональю и остальными элементами из множества $\{1, -1\}$, обладающая свойством

 $\mathbf{C}_n^T \mathbf{C}_n = (n-1) \mathbf{I}.$

На классе ортогональных матриц заданной размерности и та, и другая матрицы (после ортонормирования) экстремальны по весьма простому критерию – они минимальны по значению максимума среди абсолютных значений их элементов. В этом смысле их можно называть минимаксными ортогональными матрицами или частными случаями таких матриц для четных порядков.

Определение 3. М-матрица или обобщенная матрица Адамара в строгом смысле – это матрица, максимум абсолютных значений элементов (*m*-норма) которой минимален на классе ортогональных матриц (после ортонормирования строк или векторов) заданного четного или нечетного порядков *n*.

Определение 4. М-матрица или обобщенная матрица Адамара в менее строгом смысле – это матрица, максимум абсолютных значений элементов (*m*-норма) которой локально-минимален на классе ортогональных матриц (после ортонормирования строк или векторов) заданного четного или нечетного порядков *n*.

Важными частными случаями М-матриц являются бинарные (адамаровы) матрицы, тринарные (С-матрицы), а также матрицы, содержащие, в общем случае, более трех уровней.

Алгоритм поиска М-матриц

Алгоритм поиска **М**-матриц ориентирован на достижение глобального (или локального, при снижении требований) минимума максимума абсолютных значений ортогональной матрицы, т.е. минимальной *m*-нормы.

В качестве начального приближения используется кососимметрическая теплицева матрица **A** оператора гильбертова преобразования с единичной диагональю, внедиагональные элементы вычисляются как функция разности индексов $a_{ij} = M/(i-j)$, где M – масштабный множитель. Перед использованием столбцы матрицы **A** нормируются. Ниже описаны этапы итерации процесса поиска **M**-матриц, представленной на схеме на рис. 1.



Рис. 1. Схема реализации алгоритма поиска **М**-матриц: 1 – перестановка столбцов; 2 – ограничение норм; 3 – ортогонализация Грамма-Шмидта; 4 – уменьшение величины сжатия *р*

Блок 1. Перестановка столбцов итерируемой матрицы так, чтобы первым стал столбец с максимальным по абсолютной величине элементом, вторым и последующими – тот, который менее всего уступает по этому показателю предыдущему столбцу.

Блок 2. Ограничение норм элементов матрицы насыщением: абсолютные значения всех элементов должны быть понижены до границы насыщения, s = pm, p < 1, m – текущее значение максимального по абсолютной величине элемента матрицы (*m*-норма).

Блок 3. Ортогонализация сжатой матрицы по методу Грама–Шмидта. Перестановка столбцов создает эффективное зацепление за максимально измененный в желаемом смысле вектор, ортогонализация не меняет его направление, не восстанавливает, как это может быть в противном случае.

Блок 4. Уменьшение величины сжатия пересчетом порога p = bp - b + 1, b < 1 (обычно 0,995). Переход к блоку 1 для совершения следующей итерации.

Практика показала, что количество требуемых итераций – около тысячи. Полезно осуществить несколько запусков, варьируя стартовое значение для *p*, пока финальная *m*-норма не перестанет меняться. Выбор начального приближения при реализации предложенного алгоритма определен следующим. Ортогональная единичная матрица I не годится в качестве начального условия, поскольку ортогонализация всего лишь восстановит ее. Случайные матрицы создают ненужные проблемы для воспроизведения эксперимента. Стартовые симметричные матрицы тоже мало подходят: эволюция к оптимуму связана с прохождением неоптимальных локальных стадий, навязывание симметричной структуры, желаемой для итоговой матрицы, существенно снижает вариативность поиска в начале.

Линейный оператор гильбертова преобразования, связанный с матрицей **A**, представляет собой фазовращатель, дифференцирующий тригонометрические функции (сигналы). Он тесно связан с оператором Фурье, используемым также в теории комплексных матриц Адамара. Годятся и иные модификаторы единичной матрицы, необходимая для поиска модификация вектор-столбцов **A** управляется всего одним параметром M. Финальная ортогональная матрица формируется на выходе блока 3.

Таким образом, на каждом цикле итерируемая матрица ортогонализуется и нормируется. Далее матрица сжимается по *m*-норме, например, применением функции насыщения для абсолютных значений элементов вектор-столбцов. Ограничение абсолютных значений приводит к потере ортогональности векторов, восстанавливаемой на следующей итерации.

Этот процесс иллюстрируется рис. 1. Последовательные сжатия и ортогонализации (растяжение) образуют пульсации, в процессе которых ортогональный базис поворачивается, стремясь занять наиболее компактное положение, при котором проекции базисных векторов, т.е. максимум абсолютных значений элементов столбцов, будут принимать минимальное значение.

Амплитуда пульсаций регулируется коэффициентом насыщения p, это значение, согласно заложенной итерационной формуле, стремится к единице. Работа алгоритма зависит всего лишь от двух параметров: коэффициента раствора вектор-столбцов M < 1 матрицы Гильберта **A** и начального значения p, имеющего в программном комплексе обозначение P (P < 1). Оба этих параметра, M и P, вынесены в отдельные окна интерфейса программного комплекса для варьирования при принятии решения о продолжении циклов итераций или рестарте с новой матрицей **A**. Образуется эффективная двухпараметрическая процедура поиска **M**-матриц, которая способна находить при смене параметров M и P не только абсолютные, но и локальные оптимумы по заданному критерию m.

Интерфейс программного комплекса MMatrix

Интерфейс программного комплекса MMatrix представлен на рис. 2. Верхний ряд составляют кнопки генерации начальных стартовых матриц на основе известных заранее аналитических форм в виде матриц Гильберта (Hilbert) и, в дополнение к ней, Адамара (Hadamard), Белевича (Belevitch), Мерсенна (Mersenne) [4], Ферма (Fermat) [6] и других.

В окнах ниже фиксируется графический портрет матрицы, где интенсивность серого цвета пикселя свидетельствует об уровне ее элемента (слева), и выводится профиль уровней элементов найденных матриц – гистограмма (справа).

Имеются окна для задания общего количества итераций, промежуточных точек вывода, значений шага процедуры сжатия элементов, начального значения коэффициента сжатия *P* и других. Программно реализуется генерация некоторых рекомендуемых значений параметров двухпараметрической процедуры оптимизации для успешного поиска оптимальных и регулярных локально-оптимальных структур.

На панель интерфейса выведены кнопки процедур перестановки, сортировки, инверсии значений элементов уровней. Встроенный матричный калькулятор позволяет вычислить произведение Кронекера с целью поиска матриц высокого порядка на основе матриц более низких порядков.

Имеется окно ввода ранее найденной матрицы для продолжения итераций с нею. Интерфейс предусматривает введение символьных обозначений уровней, что позволяет применить процедуры поиска уравнений их связи. Программный комплекс обеспечивает возможность фиксировать найденные результаты в текстовом файле и в виде изображений матриц.

На рис. 3 приведены две М-матрицы матричных фракталов Эйлера (Мерсенна) 255-го и Ферма 257-го порядков, найденные с использованием описываемого программного комплекса MMatrix. Это левая (рис. 3, а) и правая (рис. 3, б) ветви основной последовательности матриц порядков, равных степени двойки.

Заключение

Достоинство реализованного в программном комплексе MMatrix итерационного алгоритма состоит в том, что результатом его работы может являться и локальный оптимум, который интересен сам по себе – именно так обнаруживаются малоуровневые субоптимальные регулярные структуры.

Рассматриваемый алгоритм обладает уникальными качествами, поскольку, хотя основа его построена на работе с вещественными числами, результатом вычислений являются целочисленные решения системы квадратичных уравнений, следующих из условия ортогональности матрицы. В результате его работы единообразно находятся, например, и М-матрицы 12-го и 20-го порядков, исторически полученные Адамаром совершенно иным способом, и матрицы Белевича порядков 6, 10, 14, 18 и т.д., имеющие тоже вполне самостоятельную историю их получения.

Известно, что проблема Адамара открыта, т.е. общего алгоритма поиска этих матриц нет. В таких обстоятельствах независимый путь получения матриц Адамара, отличный от переборных процедур, имеет большое прикладное значение. В теории матриц Белевича есть порядки, для которых матриц такого вида не существует. Это немедленно поднимает вопрос об альтернативных решениях, которые могут быть получены и исследованы при помощи такого комплекса. Так, например, было получено минимаксное решение для 22-го порядка [5]. Вес матриц Адамара и Белевича в теории информации внушителен, что поднимает интерес к их замещениям, если матрицы отсутствуют.

■ MN	latrix	-	-	-	-	-		-	-		l	-		x
Hada	amard	Mersen	ne	Belevito	h	Fermat		Jacobsth)-matrix	Fourier	•	Hilbe	ert
9	\mathbb{N}	2	Ņ	23	S									
	ŭ					X								
	Q.	Ň	5	Š.										
Š		3	È											
				¥8	Ś									
	<u>S</u>	8		86		50								
🗆 AI	bs_C	F	R	m-no	rm 0	.138461		✓ Abs	С	F m*	sqrt(n)	1.116	312	
N	65		P (0.75	М	0.1		Optim	nal	Text				-
lter	2000	Vie	ew	100	dP	0.005		Regu	lar]			Print	
к	0.3		5	0.3	W	2.5				St	art		austian	
	st.	sort	[satm		sats		Get					quation	5
0.138 0.123 0.123 0.123 0.123	8461 0. 3795 0. 3795 -0. 3795 -0. 3795 -0.	123795 (138461 -(111538 (111538 -(111538 -(111538 -().123).111).138).111).111	3795 0.12 1538 -0.11 3461 -0.11 1538 0.13 1538 -0.11	3795 1538 1538 8461 1538	0.123795 -0.111538 -0.111538 -0.111538 -0.111538 0.138461	0.1 -0.1 0.1 0.1 0.1	23795 0.1 11538 0.1 38461 -0.1 38461 0.1 38461 0.1	23795 38461 11538 38461 38461	0.123795 0.138461 0.138461 -0.111538 0.138461	0.123795 0.13846 0.13846 0.13846 0.13846 -0.111538	5 0.12 -0.11 0.13 0.13 0.13	23795 11538 38461 - 38461 38461	
0.123	3795-0. 3795 0. 3795 0. 3795 0.	138461 -(138461 (138461 (138461 ().138).111).138).138	3461 0.13 1538 0.13 3461 -0.11 3461 0.13	8461 8461 1538 8461	0.138461 0.138461 0.138461 -0.111538	-0.1 -0.1 -0.1	11538 0.1 11538 0.1 11538 -0.1 11538 -0.1	38461 11538 11538	-0.111538 -0.111538 0.138461 -0.111538	-0.111538 -0.111538 -0.111538 0.13846	3 -0.1 3 0.13 3 0.13 1 0.13	38461 - 38461 38461 38461	
•													Þ	
														Ν
ab	out	inv. ro	w	inv. co		swap rows	;][swap cols	s	Inv				
sym	metry	sort		abc		cba		trans		sum	sortosc	0	.001	1
A	A'A	MS		MR		MS1		MR1		MS2	MR2		+ X	2
										_		_		

Рис. 2. Интерфейс программного комплекса MMatrix



Рис. 3. Примеры портретов найденных матриц Мерсенна (размер 255×255 пикселей) (а) и Ферма (размер 257×257 пикселей) (б) 255-го и 257-го порядков соответственно

Литература

- 1. Аршинов М.Н., Садовский Л.Е. Коды и математика М.: Наука, 1983. 144 с.
- 2. Блэйхут Р. Быстрые алгоритмы цифровой обработки сигналов. М.: Мир, 1989. 448 с.
- 3. Сергеев М.Б., Балонин Н.А., Балонин Ю.Н. Программа поиска М-матриц. Свидетельство о государственной регистрации программы для ЭВМ № 2012614356 от 16 мая 2012 г.
- 4. Балонин Н.А., Сергеев М.Б., Мироновский Л.А. Вычисление матриц Адамара–Мерсенна // Информационно-управляющие системы. 2012. № 5. С. 92–94.
- 5. Балонин Ю.Н., Сергеев М.Б. М-матрица 22-го порядка // Информационно-управляющие системы. 2011. № 5. С. 87–90.
- 6. Балонин Н.А., Сергеев М.Б., Мироновский Л.А. Вычисление матриц Адамара–Ферма // Информационно-управляющие системы. – 2012. – № 6 (61). – С. 90–93.
- 7. Балонин Н.А., Сергеев М.Б. О двух способах построения матриц Адамара–Эйлера // Информационноуправляющие системы. – 2013. – № 1 (62). – С. 7–10.

Балонин Юрий Николаевич	_	Санкт-Петербургский государственный университет аэрокосмического прибо-
		ростроения (ГУАП), программист, tomaball@mail.ru
Сергеев Михаил Борисович	-	Санкт-Петербургский государственный университет аэрокосмического прибо-
		ростроения (ГУАП), доктор технических наук, профессор, зав. кафедрой; НИИ
		информационно-управляющих систем НИУ ИТМО, директор, mbse@mail.ru

УДК 65.012.122 АГЕНТНОЕ ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ БИЗНЕС-ПРОЦЕССОВ В НОТАЦИИ еЕРС Е.А. Ланцев, М.Г. Доррер

Предложен и разработан теоретико-множественный аппарат описания моделей бизнес-процессов микроэкономической системы в методологии ARIS eEPC, а также агентных имитационных моделей AnyLogic. Разработан метод трансляции модели в нотации eEPC в агентную имитационную модель AnyLogic. На основе разработанного метода проведен эксперимент по трансляции модели в нотации eEPC в агентную имитационную модель AnyLogic. Ключевые слова: агентное имитационное моделирование, eEPC, ARIS, AnyLogic.

Введение

Для анализа бизнес-процессов могут применяться непосредственно программные продукты для имитационного моделирования – Arena, AnyLogic, GPSS и др. Но использование специализированных систем имитационного моделирования для анализа деятельности предприятий и организаций сопряжено со значительной сложностью представления и согласования разработанных моделей с экспертами предметной области – менеджерами, экономистами, технологами. Для них более традиционным является представление деятельности предприятия в виде графических моделей бизнес-процессов в формате IDEF0/IDEF3, BPMN, EPC.

Существующие разработки в области трансляции между формальным описанием бизнеспроцессов и его имитационной моделью в основном нацелены на получение в конечном итоге дискретно-событийной имитационной модели. В качестве целевой имитационной модели может использоваться как модель в формате существующей имитационной системы (Arena, GPSS, CPN Tools), так и формат для системы имитационного моделирования собственной разработки. Агентные модели в этой области применяются мало, в качестве имитационной системы используются собственные разработки [1–3].

В отличие от существующих разработок, в настоящей работе реализован более общий подход, теоретически применимый и к другим событийным моделям бизнес-процессов.

Методы и материалы исследования

ARIS eEPC – расширенная нотация описания цепочки процесса, управляемого событиями, которая предназначена для детального описания бизнес-процесса и отражает логику его выполнения [4]. Модель бизнес-процесса представляет собой граф, состоящий из вершин – объектов модели, а также дуг (направленная связь) и ребер (ненаправленная связь) – связей между ними, которые не только соединяют объекты модели, но и служат для определения отношения между ними.

Модель в нотации eEPC (extended Event Driven Process Chain) определим как [5]:

 $G = \{X, V\},$

где X – непустое множество объектов модели (вершин графа); V – непустое множество дуг V_n и ребер V_r , таких, что $V = V_n \cup V_r$.

Объекты модели (вершины графа):

 $X = \{S, F, D, I, P, R\},\$

где S – множество объектов типа «событие»; F – множество объектов типа «функция»; D – множество объектов типа «документ» и «информационный носитель»; I – множество объектов типа «Должность», «Организационная единица»; P – множество объектов типа «Продукт/услуга»; R – множество объектов типа «правило» XOR, OR, AND, задающих логику ветвления/слияния бизнес-процесса.

Множество объектов типа «событие» S состоит из множеств $S_s \cup S_f \cup S_m$. Здесь S_s – непустое множество начальных событий, определяемое как $S_s = \{s_i | s_i \in S: \neg \exists s_j \in S: j \leq i\}; S_f$ – непустое множество конечных событий, определяемое как $S_f = \{s_i | s_i \in S: \neg \exists s_j \in S: j \leq i\}; S_m$ – множество событий внутри модели, определяемое как $S_m = \{s_i | s_i \in S: \exists s_j \in S: j \leq i\}; S_m$ – множество событий внутри модели, определяемое как $S_m = \{s_i | s_i \in S: \exists s_j \in S: j < i\},$ где $s_i - i$ -ое событие модели.

Множество правил ветвления R состоит из элементов множества $R_{xor} \cup R_{or} \cup R_{and}$, где R_{xor} – множество правил ветвления типа «исключающее ИЛИ» (XOR); R_{or} – множество правил ветвления типа «ИЛИ» (OR); R_{and} – множество правил ветвления типа «ИЛИ» (OR); R_{and} – множество правил ветвления типа «ИЛИ» (AND).

Множество объектов типа I состоит из элементов $I_{pos} \cup I_{org}$, где I_{pos} – множество объектов типа «Должность»; I_{org} – множество объектов типа «Организационная единица».

Множество объектов типа D состоит из элементов $D_{doc} \cup D_{inf}$, где D_{doc} – множество объектов типа «Документ», D_{inf} – множество объектов типа «Информационный объект».

Множество дуг V_n определяется как

 $V_{n} = \{ V_{sf}, V_{fs}, V_{sr}, V_{rs}, V_{fr}, V_{rf}, V_{fp}, V_{pf} \},\$

где $V_{sf} = \{(s_i, f_j) \mid s_i \in S, f_j \in F\}$ – множество дуг, имеющих направление от события к функции; $V_{fs} = \{(f_i, s_j) \mid f_i \in F, s_j \in S\}$ – множество дуг, имеющих направление от функции к событию; $V_{sr} = \{(s_i, r_j) \mid s_i \in S, r_j \in R\}$ – множество дуг, имеющих направление от события к объекту типа «правило ветвления»; $V_{rs} = \{(r_i, s_j) \mid r_i \in R, s_j \in S\}$ – множество дуг, имеющих направление от объекта типа «правило ветвления»; $V_{rs} = \{(r_i, s_j) \mid r_i \in R, s_j \in S\}$ – множество дуг, имеющих направление от объекта типа «правило ветвления»; $V_{rf} = \{(f_i, r_j) \mid f_i \in F, r_j \in R\}$ – множество дуг, имеющих направление от функции к событию типа «правило ветвления»; $V_{rf} = \{(r_i, f_j) \mid r_i \in R, f_j \in F\}$ – множество дуг, имеющих направление от объекта типа «правило ветвления»; $V_{rf} = \{(r_i, f_j) \mid r_i \in R, f_j \in F\}$ – множество дуг, имеющих направление от объекта типа «правило ветвления»; $V_{rf} = \{(r_i, f_j) \mid r_i \in R, f_j \in F\}$ – множество дуг, имеющих направление от объекта типа «правило ветвления»; $V_{rf} = \{(r_i, f_j) \mid r_i \in R, f_j \in F\}$ – множество дуг, имеющих направление от объекта типа «правило ветвления»; $V_{rf} = \{(r_i, f_j) \mid r_i \in R, f_j \in F\}$ – множество дуг, имеющих направление от объекта типа «правило ветвления»; $V_{rf} = \{(r_i, f_j) \mid r_i \in R, f_j \in F\}$ – множество дуг, имеющих направление от объекта типа «правило ветвления»; $V_{pf} = \{(p_i, f_j) \mid p_i \in P, f_j \in F\}$ – множество дуг, имеющих направление от функции к продукту/услуге; $V_{pf} = \{(p_i, f_j) \mid p_i \in P, f_j \in F\}$ – множество дуг, имеющих направление от продукта/услуги к функции.

Множество ребер V_r определяется как

 $V_r = \{V_{fl}, V_{fd}\},$

где $V_{fi} = \{(i_i, f_j) \mid i_i \in I, f_j \in F\}$ – множество ребер, соединяющих функции и должности; $V_{fd} = \{(f_i, d_j) \mid f_i \in F, d_i \in D\}$ – множество ребер, соединяющих функции и документы.

Агентная модель в AnyLogic peanusyercя с помощью базового объекта – активного объекта [6]. В общем виде агентная имитационная модель AnyLogic, основанная на событийных моделях eEPC, состоит из активных объектов (агентов), имеющих каждый свои правила поведения (задаются стейтчартом – диаграммой состояний) и взаимодействующих с внешней средой и друг с другом посредством пересылки сообщений (специальный программный класс на языке Java).Определим агентную модель AnyLogic как *ABM* = {*E*, *A*},

где *E* – множество объектов, составляющих окружение агента; *A* – множество агентов (активных объектов).

Множество объектов Е, составляющих окружение агента, определим как

 $E = \{Tm, Var, Ds, Msg\},\$

где *Tm* – множество таймеров; *Var* – множество переменных; *Ds* – множество таблиц для сохранения статистических данных о работе модели и синхронизации; *Msg* – множество сообщений для взаимодействия агентов со средой и друг другом.

Множество агентов А определим как

 $A = \{Sch, Par, Var, Fn\},\$

где Sch – непустое множество стейтчартов; Par – множество параметров агента; Fn – множество функций (специальный элемент AnyLogic для записи функций на языке Java) агента.

Множество стейтчартов Sch определим как

 $Sch = \{Bc, St, Trn\},\$

где *Bc* – непустое множество указателей начального состояния; *St* – непустое множество состояний; *Trn* – множество переходов.

Перевод графических элементов модели в нотации еЕРС в агентную имитационную модель AnyLogic производится согласно табл. 1.

Рассмотрим подробнее метод преобразования модели в нотации еЕРС в агентную имитационную модель AnyLogic, в процессе преобразований будем использовать соответствия, приведенные в табл. 1. Агент реализуется с помощью базового объекта AnyLogic – активного объекта. В качестве прототипа для класса агента используются организационные единицы еЕРС (должность, организационная единица и другие), т.е. исполнители – ресурсы, необходимые для выполнения бизнес-функции, но не изменяемые ею (например, людские ресурсы, оборудование, и т.п.). В имитационной модели создаются классы активных объектов, и далее в модели используются любое число экземпляров класса.

Коммуникация между агентами осуществляется с помощью особого класса – сообщения (объект «сообщение» реализован через класс на языке Java), с его помощью агенту передается информация о состоянии, в которое ему нужно перейти. Сообщения могут поступать из внешней среды (например, по сигналу таймера) или от других агентов. У каждого из агентов имеется своя очередь из сообщений, которая обрабатывается, когда агент освобождается (дисциплина очереди – FIFO). Начальные события модели еЕРС в агентной модели моделируются с помощью таймеров, которые генерируют заявки на обслуживание по заданному расписанию. Также сами агенты генерируют сообщения друг для друга.

Поведение агента задается стейтчартом. Простые состояния стейтчарта соответствуют функциям еЕРС. Стейтчарт агента состоит из одного гиперсостояния и нескольких простых состояний. В гиперсостоянии есть только одно состояние – ожидание, в котором агент находится в самом начале моделирования, а также переходит после каждого простого состояния. Из гиперсостояния в какое-либо из простых состояний стейтчарт переходит, следуя последовательности смены состояний, либо по сигналу из внешней среды, либо по сигналу от других агентов.

Вероятностная логика правил ветвления eEPC «ИЛИ» и «исключающего ИЛИ» (OR и XOR) реализуется в методе (под методом здесь подразумевается соответствующий термин объектноориентированного программирования) перехода «Action» на языке Java при смене состояний (простого состояния на гиперсостояние). Таким образом, моделируется принятие агентом решения – в какое следующее состояние он должен перейти далее или же он должен перейти в режим ожидания и отправить сообщение другому агенту для обработки.

Параллельное выполнение модели функций еЕРС в имитационной модели также реализуется в методе перехода «Action» на языке Java при смене состояний рассылкой сообщений необходимым агентам. Синхронизация правила ветвления «И» (AND) происходит при помощи внутренних таблиц (DataSet), в которые записывается состояние выполненных функций или произошедших событий.

Использование материальных ресурсов (например, количество определенных материалов на складе) и информационных ресурсов (например, количество созданных/обработанных документов) в имитационной модели фиксируется в таблицах (массивах) или же в простых переменных (вещественные, целые) при помощи кода на языке Java.

Получившаяся в результате преобразований по предложенному методу агентная имитационная модель является «заготовкой» для дальнейшего более детального моделирования – сосредоточения непосредственно на агентах, их поведении и коммуникациях.

Эксперимент по трансляции в агентную имитационную модель

Для примера трансляции по предложенному методу рассмотрим модель бизнес-процесса «Формирование заказа на поставку товара» в нотации еЕРС (рис. 1).

Вначале строим модель в AnyLogic на основе предложенного метода. В представленной модели еЕРС введены три организационные единицы: «Служба безопасности», «Отдел выписки» и «Диспетчер», соответственно они будут являться в имитационной модели агентами. Затем, после получения структуры агентной имитационной модели, определим численные характеристики объектов модели, необходимые для проведения имитационного моделирования: показатели для функций еЕРС (время выполнения, количество единиц используемых ресурсов); количество исполнителей (организационные единицы) и соответственно экземпляров агентов, вероятности наступления событий в правилах ветвления, в данном случае вероятности для «исключающего ИЛИ» (XOR).

Объект модели	Графическое	Соответствующий	Описание элемента AnyLogic
eEPC	обозначение	элемент AnyLogic	
Функция (Function)		Состояние (State)	Простое состояние стейтчарта
$F \in X \in G$		$St \in Sch \in A \in ABM$	(диаграммы состояний).
	Function	state8	
		: [: : : : : :] :	
Начальное событие		Таймер (Timer)	Наступление начальных
(Event)		$Tm \in E \in ABM$	событий определяется
$S_s \in S \in X \in G$	Event	Chtimer1	таймером по заданному
		. 79	аналитиком расписанию.
Событие		Переход (Transition)	Переход из гиперсостояния
(Event)		$Trn \in Sch \in A \in ABM$	стейтчарта в простое
$S_m, S_f \in S \in X \in G$		}►	состояние.
, ,			
	Материальны	ые и информационные	ресурсы
Документ,		Переменная	Переменные используются для
Информационный	Description	(Variable)	моделирования изменяющихся
носитель	Document	$Var \in E \in ABM$	характеристик и для хранения
(Document)		Ovariable	результатов моделирования.
$D \in X \in G$,		ресурсов происходит
(Droduct/Service)			в состояниях стейтнарта
(FIOUUCI/Service)	Product/Service		в состояниях стентнарта,
$T \in A \in G$			Java.
Исполнители – рес	урсы, необходимые для	выполнения бизнес-ф	ункции, но не изменяемые ею
Должность		Класс активного	Поведение активного объекта
(Position)	Position	объекта	описывается с помощью стейт-
$I_{pos} \in I \in X \in G$	FOSILION	(Active Object)	чарта. В имитационной модели
		$A \in ABM$	создаются классы активных
Организационная		· agent* · · ·	объектов и далее в модели
единица			используется любое число
(Organizational unit)	Organizational unit		экземпляров класса.
$I_{org} \in I \in X \in G$			
	<u> </u>		
Правило ИЛИ		Переход	При переходе из простого
(OR rule)		(Transition)	состояния стейтчарта в
$R_{or} \in R \in X \in G$		$Trn \in Sch \in A \in ABM$	гиперсостояние, в методе
Правило		Ì ►	перехода «Action» на языке
исключаюшего	(\times)		Java программируется логика
ИЛИ			принятие агентом решения.
(XOR rule)			
$R_{xor} \in R \in X \in G$			
Правило И			
(AND rule) P = P = V = C			
$\pi_{and} \in \pi \in X \in G$			

Таблица 1. Описание соответствия объектов модели нотации еЕРС и элементов агентной имитационной модели на языке AnyLogic





Фрагменты созданной агентной имитационной модели AnyLogic на основе модели eEPC представлены на рис. 2, 3. На рис. 2, а, показано дерево объектов модели, на рис. 2, б, отображается состав объектов, вложенных в класс «Main». Среда (окружение) моделируется классом «Main» (переменные, таймер, таблицы), в который вложены активные объекты – агенты (sB_depart, v_depart, dispetcher). Поскольку активные объекты технически являются классами на языке Java, то не представляется возможным давать им имена на русском языке, и для пояснений приходится ограничиваться текстовыми комментариями на самой модели. В данном примере, чтобы не перегружать модель комментариями, не были сделаны комментарии к переменным (например, переменная «cancelOrders» хранит количество отмененных заказов, переменная «tr all» хранит количество сообщений, сгенерированных таймером, и т.д.). Далее на рис. 3 представлен пример создания агента «Диспетчер» и его стейтчарта на основе объектов модели еЕРС (табл. 1).

Для верификации полученной агентной имитационной модели AnyLogic, проведено сравнение ее выходных данных с выходными данными имитационной модели встроенного в ARIS ToolSet имитатора ARIS Simulation. Сопоставив выходные данные экспериментов по загруженности исполнителей (табл. 2), а также по выполненным/отмененным заказам (табл. 3), полученные для одной и той же исходной модели еЕРС, сделали вывод, что имитационная модель ARIS Simulation и агентная имитационная модель AnyLogic показывают сравнимые результаты моделирования при одинаковых исходных данных.



а

Рис. 2. Активные объекты (а) и их окружение (б) агентной имитационной модели AnyLogic



Рис. 3. Пример создания стейтчарта (диаграммы состояний) для программирования поведения агента «Диспетчер»

	ARIS Simulation, %	AnyLogic, %
Отдел выписки	97,90	98,40
Служба безопасности	88,30	83,80
Диспетчер	55,60	58,40

Таблица 2. Сводные усредненные данные по загруженности исполнителей, %

	ARIS Simulation	AnyLogic
Выполнено заказов	18	18
Отменено	1	3

Таблица 3. Сводные усредненные данные по выполненным/отмененным заказам

Заключение

Установлена возможность взаимной связи элементов модели в нотации eEPC и агентной модели AnyLogic, paspaботан метод трансляции. Проведенный эксперимент по трансляции модели из нотации eEPC в агентную имитационную модель показал совпадение выходных данных по ключевым характеристикам исследуемых моделей в системах ARIS Simulation и AnyLogic. Дальнейшим развитием работы является автоматизация получения агентной имитационной модели на основе репозитария моделей в системе ARIS Toolset, что позволит еще больше сократить время проектирования имитационных моделей, а также упростит процесс их дальнейшей эксплуатации.

Литература

- 1. Тихонов С.В. Моделирование бизнес-процессов с использованием аппарата теории систем массового обслуживания // Сборник трудов II конференции «II школа-семинар молодых ученых «Управление большими системами». Т. 2. Воронеж: Научная книга, 2007. С. 5–59.
- 2. Валькман Ю.Р., Муцаковская К.В. Определение базиса для построения мультиагентной системы имитационного моделирования бизнес-процессов в нотации ВРММ // Збірник наукових праць. – Інституту проблем моделювання в енергетиці. – 2012. – № 62. – С. 101–108.
- 3. Жевнерчук Д.В. Методика построения имитационных моделей с применением больших массивов данных на основе интеграции IDEF3, OLAP, GPSS технологий. Ижевск, 2006. 146 с.
- 4. Шеер А.В. Бизнес-процессы. Основные понятия. Теория. Методы. М.: Весть-МетаТехнология, 1999. 182 с.
- Доррер М.Г., Ланцев Е.А., Шаргаева А.И. Событийный анализ бизнес процессов в нотации еЕРС // Тр. XV Международной ЭМ конференции по эвентологической математике и смежным вопросам / СФУ, НИИППБ, КГТЭИ. – Красноярск, 2011. – С. 89–96.
- 6. Карпов Ю.Г. Имитационное моделирование систем. Введение в моделирование с AnyLogic 5. СПб: БХВ-Петербург, 2006. 400 с.

Ланцев Евгений Алексеевич – Сибирский государственный технологический университет, аспирант, red_snow@list.ru

Доррер Михаил Георгиевич – Сибирский государственный технологический университет, кандидат технических наук, доцент, mdorrer@mail.ru

УДК 621.391 МОДЕЛЬ ПЛАНИРОВАНИЯ ЧАСТОТНО-ВРЕМЕННОГО РЕСУРСА В НИСХОДЯЩЕМ КАНАЛЕ СВЯЗИ ТЕХНОЛОГИИ LTE С.В. Гаркуша, Ю.А. Василенко

Приведены результаты разработки математической модели планирования частотно-временного ресурса нисходящего канала связи технологии LTE. Предложенная модель направлена на обеспечение гарантированного качества обслуживания пользователей беспроводной сети путем выделения пользовательским станциям требуемых скоростей передачи. Проведен сравнительный анализ предложенной модели с существующими методами с точки зрения обеспечения общей производительности нисходящего канала связи, степени балансировки пропускной способности, а также вероятности выделения пользовательских станциям требуемых скоростей передачи. Показано, что предложенная модель в условиях высоких требований к скорости передачи пользовательских станций, по сравнению с известными методами, позволяет на 5–20% повысить степень балансировки пропускной способности инсходящего канала связи, а также на 40–100% повысить вероятность выделения пользовательским станциям требуемой скорости передачи. **Ключевые слова:** LTE, частотно-временной ресурс, ресурсный блок, блок планирования, математическая модель, требуемая скорость передачи.

Введение

В технологии LTE (Long-Term Evolution), разработанной 3GPP (3rd Generation Partnership Project), одним из эффективных путей повышения производительности и улучшения основных показателей каче-

ства обслуживания (Quality of Service, QoS) является усовершенствование сетевых протоколов и механизмов, отвечающих за планирование доступных сетевых ресурсов. К подобного рода ресурсам, прежде всего, относятся временной ресурс – OFDM-символы (Orthogonal frequency-division multiplexing) и частотный ресурс – частотные поднесущие. Под OFDM-символом понимается промежуток времени, на протяжении которого амплитуда и фаза модулируемых поднесущих неизменна. Необходимо заметить, что решение задачи о выделении радиоресурсов в первую очередь основывается на требованиях к QoS и может быть возложено на систему управления радиоресурсом (Radio Resource Management, RRM), а именно на планировщик (scheduler). В технологии LTE, как и в HSDPA или WiMAX, механизмы планирования ресурсов нисходящего канала связи (DownLink) не определены стандартом, оставляя право выбора за производителями оборудования базовых станций (evolved NodeB, eNodeB) [1–3].

Результатом решения задачи распределения частотного и временного ресурсов должно быть закрепление блоков планирования (Scheduling Block, SB) за пользовательскими станциями (User Equipment, UE) в нисходящем канале связи одного кадра. Блок планирования является наименьшим структурным элементом, выделяемым одной пользовательской станции и формируемым двумя соседними ресурсными блоками (Resource Block, RB) на одинаковых поднесущих (subcarrier) [4].

В работе были проанализированы методы распределения частотного и временного ресурсов, использующие алгоритм Round Robin Scheduler [3, 5, 6]. Использование алгоритма Round Robin Scheduler предполагает выделение в равной степени частотно-временного ресурса для каждой UE. В базовой реализации Round Robin Scheduler выделяет каждой UE временной интервал, в рамках которого эта UE получает монопольный доступ к каналу. При этом скорость передачи данных UE ограничивается фактическим значением отношения сигнал/шум (signal to noise Ratio, SNR). Также были проанализированы методы [5, 6], использующие алгоритм Max C/I Ratio, который предоставляет частотно-временной ресурс пользовательской станции с самыми лучшими значениями SNR, не обеспечивая справедливость распределения этого ресурса между UE. Кроме того, в работе проанализированы методы распределения частотного и временного ресурсов, использующие алгоритм пропорционального справедливого распределения обслуживания (Proportional Fair Scheduling) [5–7]. Проведенный анализ показал, что алгоритм Proportional Fair Scheduling способствует UE, которая имеет высокое значение SNR, одновременно обеспечивая достаточное количество частотных и временных ресурсов UE с худшим значением SNR.

Анализ показал, что для предоставления доступа к частотно-временному ресурсу технологии LTE наиболее приемлемым был бы механизм, включающий в себя особенности Round Robin и Max C/I Ratio алгоритмов. Выбор алгоритма зависит от категории и величины нагрузки. Правильный выбор алгоритма для предоставления доступа особенно важен при большой нагрузке. Использование алгоритма Proportional Fair Scheduling направлено на применение для интерактивного «best effort» класса данных во избежание ситуации, при которой некоторые UE никогда не получат доступ к частотно-временному ресурсу. Указанный класс обслуживания (Class of Service, CoS) обеспечивает доставку данных UE по мере возможностей без гарантий скорости передачи данных. Повышение качества обслуживания при планировании частотновременного ресурса каждой UE должно быть направлено на обеспечине гарантированной скорости передачи с возможностью доступа к дополнительной (не гарантируемой) полосе пропускания. Однако ни один из проанализированных механизмов не способен обеспечить подобного CoS. В результате проведенного анализа принято решение о необходимости разработки математической модели планирования частотновременного ресурса в нисходящем канале связи технологии LTE, сформулированной как задача распределения SB для обеспечения гарантированной скорости передачи UE.

Математическая модель распределения ресурсных блоков в нисходящем канале

Предлагаемая математическая модель направлена на применение в беспроводных сетях технологии LTE, использующей временное и частотное разделение каналов. При разработке математической модели учитывается тот факт, что наименьшей структурной единицей радиоресурса, которой можно управлять при решении задачи планирования, является SB [1].

В технологии LTE предложено три вида распределения ресурсов. Разрабатываемая модель направлена на использование нулевого вида (Resource Allocation Type 0), предполагающего объединение RB в так называемые группы ресурсных блоков (resource block groups, RBG), которые выделяются UE. При этом количество RB, входящих в состав одной RBG (p), зависит от используемой ширины частотного канала. Если деление количества RB на параметр p не дает целочисленного значения, то крайняя RBG будет иметь размер, меньший p [8].

С целью учета числа подкадров, выделенных для передачи информации в нисходящем канале связи [1, 2], в математическую модель введено понятие матрицы конфигураций нисходящего канала – прямоугольной матрицы с количеством строк, соответствующим количеству конфигураций кадра (L), и с количеством столбцов, соответствующим количеству подкадров (K) в кадре, т.е.

 $H = \left\| h_{l,k} \right\|, \, (\, l = \overline{0,L\!-\!1} \, ; \, \, k = \overline{0,K\!-\!1} \,),$

[1, если k-й подкадр при l-й конфигурации используется

где $h_{l,k} = \begin{cases} для передачи информации в нисходящем канале связи; \end{cases}$

0, в противном случае.

В ходе решения задачи распределения блоков планирования в рамках предлагаемой модели необходимо обеспечить расчет булевой управляющей переменной ($x_{k,m}^n$), определяющей порядок распределения блоков планирования:

$$x_{k,m}^{n} = \begin{cases} 1, \text{ если } m$$
-й блок планирования на k-м подкадре выделен n-й UE;

где $m = \overline{0, M-1}$; $k = \overline{0, K-1}$; $n = \overline{1, N}$; M – число SB, формируемых на протяжении передачи одного подкадра; N – количество UE.

При расчете искомых переменных $x_{k,m}^n$ необходимо выполнить ряд важных условий-ограничений.

1. Условие закрепления *m* -го SB нисходящего канала связи на протяжении передачи *k* -го подкадра не более чем за одной UE:

$$\sum_{n=1}^{N} x_{k,m}^{n} \le 1 \quad (k = \overline{0, K-1}; \ m = \overline{0, M-1}).$$
(2)

2. Условие выделения UE блоков планирования только нисходящего канала:

$$\frac{\sum_{m=0}^{M-1}\sum_{n=1}^{N} x_{k,m}^{n}}{M} \le h_{l,k} \ (k = \overline{0, K-1}; l - используемая конфигурация кадра).$$
(3)

3. Условие закрепления за *n*-й UE количества SB, обеспечивающего необходимую скорость передачи в нисходящем канале связи при используемой схеме модуляции и кодирования (Modulation and Coding Scheme, MCS):

$$\sum_{m=\frac{M}{2}-3}^{\frac{M}{2}+2} R_{SB}^{0} x_{0,m}^{n} + \sum_{m=\frac{M}{2}-3}^{\frac{M}{2}+2} R_{SB}^{5} x_{5,m}^{n} + \sum_{m=0}^{\frac{M}{2}-4} R_{SB}^{0-9} x_{0,m}^{n} + \sum_{m=\frac{M}{2}+3}^{M-1} R_{SB}^{0-9} x_{0,m}^{n} + \sum_{m=0}^{\frac{M}{2}-4} R_{SB}^{0-9} x_{5,m}^{n} + \sum_{m=0}^{\frac{M}{2}-4} R_{SB}^{0-9} x_{0,m}^{n} + \sum_{m=0}^{M-1} R_{SB}^{0-9} x_{5,m}^{n} + \sum_{m=0}^{M-1} \sum_{k=0}^{M-1} R_{SB}^{0-9} h_{l,k} x_{k,m}^{n} + \sum_{m=0}^{M-1} \sum_{k=0}^{K-1} R_{SB}^{0-9} h_{l,k} x_{k,m}^{n} \geq R_{SB}^{n-9} h_{l,k} x_{k,m}^{n} \geq R_{SB}^{n} + \sum_{m=0}^{M-1} \sum_{k=0}^{M-1} R_{SB}^{0-9} h_{l,k} x_{k,m}^{n} = R_{SB}^{0-9} h_{l,k} x_{k,m}^{n} \geq R_{SB}^{n} + \sum_{m=0}^{M-1} \sum_{k=0}^{M-1} R_{SB}^{0-9} h_{l,k} x_{k,m}^{n} = R_{SB}^{0-9} h_{l,k} x_{k,m}^{n} \geq R_{SB}^{n} + \sum_{m=0}^{M-1} \sum_{k=0}^{M-1} R_{SB}^{0-9} h_{l,k} x_{k,m}^{n} = R_{SB}^{0-9} h_{l,k} x_{k,m}^{n} \geq R_{SB}^{n} + \sum_{m=0}^{M-1} \sum_{k=0}^{M-1} R_{SB}^{0-9} h_{l,k} x_{k,m}^{n} = R_{SB}^{0-9} h_{l,k} x_{k,m}^{n} \geq R_{SB}^{n} + \sum_{m=0}^{M-1} \sum_{k=0}^{M-1} R_{SB}^{0-9} h_{l,k} x_{k,m}^{n} = R_{SB}^{0-9} h_{l,k} x_{k,m}^{n} \geq R_{SB}^{n} + \sum_{m=0}^{M-1} \sum_{k=0}^{M-1} R_{SB}^{0-9} h_{l,k} x_{k,m}^{n} = R_{SB}^{0-9} h_{l,k} x_{k,m}^{n} \geq R_{SB}^{n} + \sum_{m=0}^{M-1} \sum_{k=0}^{M-1} R_{SB}^{0-9} h_{l,k} x_{k,m}^{n} = R_{SB}^{0-9} h_{l,k} x_{k,m}^{n} \geq R_{SB}^{n} + \sum_{m=0}^{M-1} \sum_{k=0}^{M-1} R_{SB}^{0-9} h_{l,k} x_{k,m}^{n} = R_{SB}^{0-9} h_{l,k} x_{k,m}^{n} = R_{SB}^{0-9} h_{l,k} x_{k,m}^{n} \geq R_{SB}^{n} + \sum_{m=0}^{M-1} \sum_{k=0}^{M-1} R_{SB}^{0-9} h_{l,k} x_{k,m}^{n} = R_{SB}^{0-9} h_{l,k} x$$

при $n = \overline{1, N}$, l – используемая конфигурация кадра. В выражении (4) использованы следующие переменные:

- $R_{SB}^{0} = \left(\left(N_{symb}^{RB} N_{SF}^{RB} - N_{PDCCH} - N_{PSS,SSS} - N_{PBCH} \right) K_{s} - N_{pilot} N_{SF}^{RB} \right) R_{c}^{n,m} k_{b}^{n,m} / KT_{SF}$ – пропускная способность SB [9], формируемых на нулевом подкадре и передающих сигналы первичной синхрони-

ность SB [9], формируемых на нулевом подкадре и передающих сигналы первичной синхронизации (Primary Synchronization Signal, PSS), вторичной синхронизации (Secondary Synchronization Signal, SSS) и широковещательный сигнал (Physical Broadcast Channel, PBCH), где $N_{PSS,SSS} = 2$, $N_{PBCH} = 4$ OFDM-символов соответственно; N_{symb}^{RB} – количество OFDM-символов, формирующих один ресурсный блок; $N_{SF}^{RB} = 2$ – количество RB, формируемых на одинаковых поднесущих и выделяемых UE на протяжении передачи одного подкадра; $N_{PDCCH} = 3$ – количество OFDM-символов в каждом подкадре, выделенных для передачи сигнала управления (Physical Downlink Control Channel, PDCCH); $N_{pilot} = 4$ – ресурсные элементы (Recourse Element, RE), выделенные для передачи пилотных сигналов в одном RB; K_s – число поднесущих для передачи данных в одном RB и одном SB; $R_c^{n,m}$ – скорость кода, используемого при кодировании сигнала *n*-й UE на поднесущих *m*-го SB; $k_b^{n,m}$ – битовая загрузка OFDM-символа *n*-й UE на поднесущих *m*-го SB; $T_{SF} = 1$ мс – время передачи одного подкадра;

- $R_{SB}^{5} = \left(\left(N_{symb}^{RB} N_{SF}^{RB} N_{PDCCH} N_{PSS,SSS} \right) K_{s} N_{pilot} N_{SF}^{RB} \right) R_{c}^{n,m} k_{b}^{n,m} / KT_{SF}$ пропускная способность SB [9], формируемых на пятом подкадре и передающих сигналы PSS и SSS;
- $R_{SB}^{0-9} = \left(\left(N_{symb}^{RB} N_{SF}^{RB} N_{PDCCH} \right) K_s N_{pilot} N_{SF}^{RB} \right) R_c^{n,m} k_b^{n,m} / KT_{SF}$ пропускная способность SB [9], формируемых на всех подкадрах и содержащих из служебной информации только PDCCH.
- *R*^{*n*}_{трб} требуемая скорость передачи *n*-й UE.

4. Условие объединения RB в RGB размеров, удовлетворяющих ширине используемого частотного канала:

$$x_{k,m}^{n} = \frac{1}{p} \sum_{z = \left\lfloor \frac{m}{p} \right\rfloor^{p}}^{\left\lfloor \frac{M}{p} \right\rfloor + 1} x_{k,z}^{n} \ (n = \overline{1, N}; \ k = \overline{0, K-1}; \ m = \overline{0, \left\lfloor \frac{M}{p} \right\rfloor p - 1});$$
(5)

$$x_{k,m}^{n} = \frac{1}{p} \sum_{z=\left|\frac{M}{p}\right|p}^{M} x_{k,z}^{n} \ (n = \overline{1,N}; \ k = \overline{0,K-1}; \ m = \boxed{\left|\frac{M}{p}\right|p - 1,M} \), \tag{6}$$

где $\lfloor \rfloor$ – обозначение округления числа до ближайшего целого в меньшую сторону. Условие (6) используется для объединения блоков планирования крайней группы ресурсных блоков, размеры которой могут быть меньше *p*.

Расчет искомых переменных (1) в соответствии с условиями-ограничениями (2)–(6) целесообразно осуществлять в ходе решения оптимизационной задачи, обеспечивая минимум или максимум предварительно выбранного критерия качества решения задачи распределения SB. Задача распределения SB может быть решена с использованием критерия оптимальности, направленного на максимизацию общей производительности нисходящего канала связи. Таким образом, критерий оптимальности примет следующий вид:

$$\max \sum_{n=1}^{N} \left(\sum_{\substack{m=\frac{M}{2}-3}}^{\frac{M}{2}+2} R_{SB}^{0} x_{0,m}^{n} + \sum_{\substack{m=\frac{M}{2}-3}}^{\frac{M}{2}+2} R_{SB}^{5} x_{5,m}^{n} + \sum_{m=0}^{\frac{M}{2}-4} R_{SB}^{0-9} x_{0,m}^{n} + \sum_{\substack{m=\frac{M}{2}+3}}^{M-1} R_{SB}^{0-9} x_{0,m}^{n} + \sum_{\substack{m=\frac{M}{2}+3}}^{M-1} R_{SB}^{0-9} x_{5,m}^{n} + \sum_{\substack{m=\frac{M}{2}+3}}^{M-1} R_{SB}^{0-9} x_{0,m}^{n} + \sum_{\substack{m=\frac{M}{2}+3}}^{M-1} R_{SB}^{0-9} x_{0,m}^{n} + \sum_{\substack{m=\frac{M}{2}+3}}^{M-1} R_{SB}^{0-9} x_{0,m}^{n} + \sum_{\substack{m=\frac{M}{2}+3}}^{M-1} R_{SB}^{0-9} x_{5,m}^{n} + \sum_{\substack{m=0}}^{M-1} \sum_{\substack{k=0}}^{M-1} R_{SB}^{0-9} h_{l,k} x_{k,m}^{n} + \sum_{\substack{m=0}}^{M-1} \sum_{\substack{k=0}}^{M-1} R_{SB}^{0-9} h_{l,k} x_{k,m}^{n} \right),$$
(7)

при учете условий-ограничений (2)–(6). Использование критерия оптимальности (7) совместно с условиями-ограничениями (2)–(6) направлено на обеспечение гарантированной скорости передачи с возможностью доступа к дополнительной (не гарантируемой) полосе пропускания, что в CoS соответствует подклассу В.

Сформулированная задача с математической точки зрения является задачей смешанного целочисленного нелинейного программирования – MINLP (Mixed Integer NonLinear Programming). Переменная, используемая в критерии оптимальности (7), является целочисленной, ограничения на искомые переменные (2)–(4) носят линейный, а ограничения (5) и (6) – нелинейный характер.

Анализ решений задачи распределения блоков планирования

С целью анализа решений по планированию ресурсных блоков в нисходящем канале связи, получаемых с использованием известных методов, а также предложенной модели (1)–(7), рассмотрим пример, в котором для распределения блоков планирования в нисходящем канале связи были использованы следующие исходные данные: количество UE N = 5; число SB, формируемых на протяжении передачи одного временного слота, M = 15; количество OFDM-символов, формирующих один ресурсный блок, $N_{symb}^{RB} = 7$; значения эффективности MCS-сигналов UE для различных ресурсных блоков – в соответствии с рис. 1; вид разделения каналов – TDD; количество подкадров, используемых для передачи информации в нисходящем канале связи, K = 4 (конфигурация кадра – № 1 [1, 2]); количество SB, входящих в состав одной RBG, p = 2.

Решение сформулированной в работе оптимизационной задачи было получено с использованием системы MATLAB R2012b. При этом была задействована программа minlpAssign пакета оптимизации TOMLAB. Всем пользовательским станциям были установлены одинаковые требуемые скорости передачи, которые принимали значения $R_{rp6}^n = 0 - 0.26$ Мбит/с.

Как показали результаты моделирования, общая производительность нисходящего канала связи при использовании известных методов на протяжении всего интервала измерения не изменялась и составила для метода Round Robin 0,9622 Мбит/с, для метода Proportional Fair – 1,2377 Мбит/с, а для метода Мах C/I Ratio – 1,4192 Мбит/с. Общая производительность нисходящего канала связи при использовании предложенной модели (1)–(7) на участке $R_{rp6}^n = 0-0,15$ Мбит/с имела максимальное значение, соответствующее методу Мах C/I Ratio, и составляла 1,4192 Мбит/с. На интервале $R_{rp6}^n = 0,15-0,26$ Мбит/с общая производительность на 3%, до значения 1,3641 Мбит/с.



На рис. 2, а, приведены результаты моделирования, отображающие динамику изменения степени балансировки пропускной способности нисходящего канала между UE. Степень балансировки пропускной способности определялась в соответствии с выражением [10]

$$F^{i} = 1 - \left(\max_{n} R_{n}^{i} - \min_{n} R_{n}^{i} \right) / \sum_{n=1}^{N} R_{n}^{i} ,$$

где R_n^i – скорость передачи, выделенная *n*-й UE на *i*-м интервале измерения, $n = \overline{1, N}$. Как показали результаты моделирования (рис. 2, а), степень балансировки пропускной способности нисходящего канала связи при использовании известных методов на протяжении всего интервала измерения не изменялась и составила для метода Round Robin 0,9421, для метода Proportional Fair – 0,9163, а для метода Мах С/I Ratio – 0,8214. Степень балансировки пропускной способности нисходящего канала связи при использовании предложенной модели (1)–(7) на участке $R_{rp6}^n = 0-0,15$ Мбит/с имела минимальное значение, соответствующее методу Мах С/I Ratio, и составляла 0,8214. На интервале $R_{rp6}^n = 0,15-0,26$ Мбит/с степень балансировки пропускной способности увеличилась до 0,9859.





На рис. 2, б, приведены результаты расчета вероятности выполнения требований по скорости передачи, выделяемой всем UE. Вероятность выполнения требований по скорости передачи на *i* -м интервале измерения определялась в соответствии с выражением

$$P^{i} = \sum_{n=1}^{N} Q_{n}^{i} / N$$

где $\sum_{n=1}^{N} Q_{n}^{i}$ – количество UE, которым выделена требуемая скорость передачи на *i* -м интервале измере-

ния, т.е.

$$Q_n^i = \begin{cases} 0, \text{ если } R_n^i < R_{\text{трб}}^i; \\ 1, \text{ если } R_n^i \ge R_{\text{трб}}^i. \end{cases}$$

Как показали результаты моделирования, при изменении требуемой скорости передачи от 0,15 Мбит/с и выше все рассмотренные известные методы не обеспечивают требуемого значения скорости передачи для всех пользовательских станций. Выполнение требования по скорости передачи известными методами обеспечивается только при невысоких значениях R_{rp6}^n , а именно при $R_{rp6}^n = 0 - 0,15$ Мбит/с. Использование предложенной модели (1)–(7) обеспечивает выделение требуемой скорости передачи UE на всем интервале измерения $R_{rp6}^n = 0 - 0,26$ Мбит/с.

Заключение

Установлено, что одной из основных задач в беспроводной сети, функционирующей с использованием технологии LTE, является задача обеспечения требуемого качества обслуживания, в состав которой входит выделение пользовательским станциям сети необходимой скорости передачи в нисходящем канале связи. Также установлено, что обеспечение требуемой скорости передачи в технологии LTE может быть достигнуто путем решения задачи распределения блоков планирования в нисходящем канале связи. В связи с этим проанализированы существующие механизмы распределения блоков планирования между пользовательскими станциями в нисходящем канале связи беспроводной сети, функционирующей с использованием технологии LTE.

На основании выявленных недостатков известных решений предложена математическая модель, представленная рядом линейных и нелинейных условий-ограничений. Новизна модели состоит в формулировке задачи распределения блоков планирования как задачи перераспределения доступной пропускной способности нисходящего канала технологии LTE для передачи информации в направлении пользовательских станций при учете их территориальной удаленности (вида системы модуляции и кодирования).

Сравнительный анализ показал, что в условиях высоких требований к скорости передачи пользовательских станций использование модели (1)–(7), по сравнению с известными методами, позволяет на 5–20% повысить степень балансировки пропускной способности нисходящего канала связи (рис. 2, а), а также на 40–100% повысить вероятность выделения пользовательским станциям требуемой скорости передачи (рис. 2, б). При этом производительность нисходящего канала связи, в условиях высоких требований к скорости передачи, на 3% меньше производительности, получаемой с использованием метода Max C/I Ratio, и на 10–42% больше производительности с использованием методов Round Robin и Proportional Fair.

Использование предложенной модели направлено на обеспечение каждой пользовательской станции гарантированной скорости передачи в нисходящем канале связи с возможностью доступа к дополнительной (негарантируемой) полосе пропускания. Это в классах сервиса соответствует подклассу В CoS, тогда как другие методы не гарантируют выделения требуемой скорости передачи, что в классах сервиса CoS соответствует подклассу А.

Литература

- 3GPP TS 36.211. 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation (Release 11). – Valbonne, France: Sophia Antipolis, 2012. – 108 p.
- 2. Ghosh A., Zhang J., Muhamed R., Andrews J. Cr. Fundamentals of LTE. Prentice Hall, USA, 2010. 464 p.
- Iosif O., Banica I. On the Analysis of Packet Scheduling in Downlink 3GPP LTE System // The Fourth International Conference on Communication Theory, Reliability, and Quality of Service (CTRQ 2011). – 2011. – P. 99–102.
- Dahlman E., Parkvall S., Skold J., Beming P. 3G Evolution HSPA and LTE for Mobile Broadband. Academic Press, 2008. 608 p.
- Kawser M.T., Farid H.M.A.B., Hasin A.R., Sadik A.M.J., Razu I.K. Performance Comparison between Round Robin and Proportional Fair Scheduling Methods for LTE // International Journal of Information and Electronics Engineering. – 2012. – V. 2. – № 5. – P. 678–681.
- Galaviz G., Covarrubias D.H., Andrade A.G., Villarreal S. A resource block organization strategy for scheduling in carrier aggregated systems // EURASIP Journal on Wireless Communications and Networking. – 2012. – P. 107–124.

- Girici T., Zhu C., Agre J.R., Ephremides A. Proportional Fair Scheduling Algorithm in OFDMA-BasedWireless Systems with QoS Constraints // Journal of communications and networks. – 2010. – V. 12. – № 1. – P. 30–42.
- 3GPP TS 36.213. 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures (Release 11). – Valbonne, France: Sophia Antipolis, 2012. – 104 p.
- 9. Ермолаев В.Т. Флаксман А.Г. Теоретические основы обработки сигналов в системах мобильной радиосвязи. – Нижний Новгород: НГУ им. И.Н. Лобачевского, 2010. – 107 с.
- Adibah M.R.H., Kumbesan S., Riyaj B., Leijia W. Modeling and simulation of packet scheduling in the downlink long term evolution system // Proceedings of 15th Asia-Pacific Conference on Communications (APCC 2009). – 2009. – P. 68–71.

Гаркуша Сергей Владимирович

- Харьковский национальный университет радиоэлектроники, кандидат технических наук, докторант; Полтавский университет экономики и торговли, доцент, sv.garkusha@mail.ru
- Василенко Юрий Александрович
- Филиал ЗАО «Киевстар Дж.Эс.ЭМ», г. Харьков, зам. директора, tks2006@ukr.net



ЛАЗЕРНЫЕ СИСТЕМЫ И ЛАЗЕРНЫЕ ТЕХНОЛОГИИ

УДК 621.373.826 ИССЛЕДОВАНИЕ ХАРАКТЕРИСТИК ИЗЛУЧЕНИЯ МОНОИМПУЛЬСНОГО ER:YLF-ЛАЗЕРА В УСЛОВИЯХ ПРОСТРАНСТВЕННОЙ НЕОДНОРОДНОСТИ КОЭФФИЦИЕНТА УСИЛЕНИЯ

В.В. Назаров, Л.В. Хлопонин, В.Ю. Храмов, Н.А. Федоров

Разработана модель твердотельного лазера с продольной диодной накачкой, позволяющая исследовать характеристики излучения Er:YLF-лазера в условиях пространственной неоднородности коэффициента усиления. Исследована зависимость энергетических и пространственных характеристик излучения от пространственных параметров пучка накачки, а также от концентрации активатора в активном элементе. Приведена зависимость энергии моноимпульса от соотношения радиуса основной моды и радиуса пучка накачки.

Ключевые слова: устойчивый резонатор, продольная диодная накачка, неоднородность распределения коэффициента усиления, перетяжка пучка, параметр качества пучка.

Введение

В последнее время благодаря разработке эффективных мощных диодов возрос интерес к лазерам с диодной накачкой [1]. Лазеры с продольной диодной накачкой обеспечивают высокую эффективность накачки и более высокие пространственные характеристики излучения генерации в сравнении с лазерами с поперечной накачкой, что делает их перспективными для медицинских приложений. Одним из наиболее важных факторов, определяющих эффективность лазерной генерации, а также выходную мощность, является проблема пространственного перекрытия излучения накачки и пучка генерации, что определяется модовым составом генерируемого излучения. Влияние размеров пучков накачки и лазерного излучения на величину порога и эффективность лазерной генерации исследовано в ряде работ [2–4], в большинстве которых рассматривается усредненный по объему активного элемента размер пучка накачки. Несмотря на то, что данный подход дает возможность получить простые аналитические зависимости для проведения оптимизации, влияние таких параметров, как расходимость и положение перетяжки пучка накачки в объеме активного элемента, на энергетические и пространственные характеристики лазерной генерации не может быть учтено в рамках таких простых моделей. В ряде работ определение интеграла перекрытия пучков накачки и излучения генерации рассматривается при помощи численных методик [5–6], в которых учитывается генерация исключительно вблизи порога.

В настоящей работе приведены результаты исследования моноимпульсного Er:YLF-лазера с учетом радиальной и продольной неоднородностей коэффициента усиления активной среды.

Взаимодействие излучения накачки с активным элементом

Для учета влияния пространственной неоднородности коэффициента усиления на энергетические и пространственные характеристики излучения была разработана математическая модель накачки, каустика пучка накачки в котором подчиняется линейному закону (рис. 1). Такая форма каустики позволяет достаточно простым способом промоделировать накачку с произвольным значением параметра качества пучка M^2 .



Рис. 1. Модель взаимодействия излучения пучка накачки с активной средой

В модели рассматривался активный элемент лазера со следующими геометрическими параметрами: длина активной среды $L_{ae} = 50$ мм, радиус активного элемента $R_{ae} = 1$ мм. Накачка осуществлялась гауссовым пучком с заданным радиусом на входе в активную среду w_p , радиусом в перетяжке w_0 и положением самой перетяжки z_d . Радиус гауссова пучка определялся по уровню e^{-2} от максимального значения интенсивности (рис. 2).



Рис. 2. Определение параметров гауссова пучка

В рамках модели активная среда разбивалась на слои N_{layer} толщиной dz, в которых по отдельности решалась задача взаимодействия импульса излучения накачки с резонансно поглощающей средой. Параметры излучения накачки (интенсивность излучения и радиус пучка накачки), полученные после прохождения первого слоя, использовались в качестве исходных данных при расчете взаимодействия со вторым слоем, и т.д. Математически подобная задача сводится к решению нелинейной системы уравнений в частных производных для функций I(z,t) и U(z,t):

$$\frac{\partial I_p(z,t)}{\partial z} = -(\sigma_{abs} \cdot N_1 - \sigma_{em} \cdot N_2) \cdot I_p - b \cdot I_p$$
$$\frac{\partial N_2(z,t)}{\partial t} = (\sigma_{abs} \cdot N_1 - \sigma_{em} \cdot N_2) \cdot I_p - \frac{N_2}{T_1} - 2 \cdot \gamma \cdot N_2^2,$$
$$N_2 = N_s - N_1$$

где I – интенсивность излучения накачки; N_1, N_2 – населенность нижнего и верхнего уровней активной среды; N_s – концентрация активных центров; $\sigma_{abs}, \sigma_{em}$ – сечения поглощения и вынужденного излучения для излучения накачки; b – коэффициент нерезонансных потерь; T_1 – время жизни верхнего уровня; γ – коэффициент ап-конверсии.

Длительность импульса накачки т в расчетах составляла 2 мс; параметры активной среды, используемые для вычислений, представлены в таблице, в которой приведены также коэффициенты апконверсии по данным [7].

Длина волны накачки λ_p , мкм		0,97
Сечения эмиссии и поглощения на длине волны накачки, см ² σ_{em}		$0,34 \cdot 10^{-20}$
		$1,1.10^{-20}$
Время жизни верхнего лазерного уровня T_1 , с	$4 \cdot 10^{-3}$	
Модельные коэффициенты ап-конверсии ү, см ³ /с	1.10^{-17} [6]	
Коэффициент нерезонансных потерь b , см $^{-1}$	$1 \cdot 10^{-3}$	
Длина волны генерации λ_{gen} , мкм	2,81	

Таблица. Параметры кристалла Er:YLF при T = 300 K

Исследование энергетических характеристик моноимпульсного режима работы Er:YLF-лазера

Для расчета пространственно-временной структуры поля используется модель лазерного устойчивого резонатора, позволяющая исследовать процесс формирования поля лазерного излучения с учетом дифракции и усиления активной среды [8]. Моделирование проводилось для следующих геометрических параметров резонатора лазера: радиус апертуры a = 1 мм; база резонатора $L_{base} = 0,1$ м; глухое зеркало – плоское (кривизна глухого зеркала $R_1 = 0$ м⁻¹), коэффициент отражения выходного зеркала R0 = 0,95; число Френеля F = 3,6; кривизна выходного зеркала R_2 была изменяемым параметром с целью управления радиусом основной моды резонатора лазера.

В ходе моделирования была построена зависимость энергии Е моноимпульса от положения перетяжки пучка накачки в объеме активного элемента (рис. 3). Зависимости были получены для следующих параметров пучка накачки и кривизны выходного зеркала: радиус пучка накачки $w_p = 0,8$ мм, радиус перетяжки $w_0 = 0,6$ мм, кривизна выходного зеркала $R_2 = 0,6$ м⁻¹.

Как видно из представленных графиков, для концентрации активатора больше 1% смещение перетяжки вглубь активного элемента приводит к существенному увеличению энергии моноимпульса, в то же время при концентрации 0,3% увеличение энергии несущественно. При увеличении концентрации больше 1% оптимальное положение перетяжки пучка накачки, при котором наблюдается максимальная энергия при данной концентрации, начинает смещаться к левому краю активного элемента за счет того, что при больших концентрациях область усиления ограничена первыми слоями активного элемента.



Рис. 3. Зависимость энергии моноимпульса от положения перетяжки пучка накачки при концентрации активатора 0,3%, 1%, 1,5% и 2% (кривые 1–4 соответственно)



Рис. 4. Зависимость энергии моноимпульса от концентрации активатора при кривизне выходного зеркала 0,4 м⁻¹, 0,6 м⁻¹ и 0,9 м⁻¹ (кривые 1–3 соответственно)

В ходе дальнейшего моделирования была исследована зависимость энергии моноимпульса от концентрации активатора при фокусировке пучка ближе к выходному торцу активного элемента ($z_d = 50$ мм) и радиусах пучка накачки на входе в активный элемент и в плоскости перетяжки 0,8 мм и 0,6 мм соответственно (рис. 4). Моделирование проводилось при различных значениях кривизны выходного зеркала R_2 .

При указанных геометрических параметрах пучка накачки максимальная энергия моноимпульса была достигнута для концентрации активатора 1,5%. Для данной концентрации характерна наиболее рав-

номерная прокачка активного элемента вдоль продольной оси, что приводит к усилению излучения во всех слоях. Также была определена оптимальная кривизна выходного зеркала, которая составила 0,6 м⁻¹. Такая конфигурация резонатора позволяет получить максимальную энергию моноимпульса для разных концентраций и параметр качества пучка M^2 , близкий к единице. Уменьшение кривизны выходного зеркала приводит к уменьшению энергии моноимпульса при неизменном качестве пучка, а увеличение – также к уменьшению энергии, но при существенном снижении качества пучка ($M^2 > 2$) за счет генерации дополнительных поперечных мод.

Для более тщательного исследования влияния кривизны выходного зеркала (радиуса основной моды резонатора) на энергию моноимпульса была построена зависимость энергии моноимпульса от соотношения радиуса основной моды w_{00} и радиуса пучка накачки для различных значений концентраций активатора (рис. 5). Моделирование проводилось при следующих параметрах пучка накачки: $z_d = 50$ мм, $w_p = 0.8$ мм, $w_0 = 0.6$ мм.





Как видно из полученных зависимостей, при увеличении концентрации активатора (C > 1,5%) оптимальное соотношение радиуса основной моды и радиуса пучка накачки, при котором достигается максимум энергии моноимпульса, смещается в сторону увеличения радиуса основной моды. Данный эффект можно объяснить тем, что при больших концентрациях активатора усиление излучения генерации происходит за счет более эффективного снятия накопленной инверсии на периферии активного элемента.

Заключение

Разработана математическая модель твердотельного моноимпульсного Er:YLF-лазера с продольной диодной накачкой, которая позволяет исследовать влияние пространственной неоднородности коэффициента усиления на пространственные и энергетические параметры излучения генерации. Разработанная модель накачки позволяет достаточно простым способом описать пучок накачки с произвольным значением M^2 .

С помощью разработанной модели были исследованы энергетические зависимости моноимпульса лазера от положения перетяжки пучка накачки в активном элементе. Найдена оптимальная концентрация активатора и оптимальное соотношение радиусов основной моды и пучка накачки для достижения максимальной энергии моноимпульса. Представлена зависимость энергии моноимпульса от соотношения радиуса основной моды и радиуса пучка накачки для различных значений активатора.

Работа выполнена в рамках НИР № 411406 (номер государственной регистрации 01 201261810 от 23.05.2012), финансируемой из централизованных средств НИУ ИТМО.

Литература

- 1. Иночкин М.В., Назаров В.В., Сачков Д.Ю., Сидорова О.П., Хлопонин Л.В., Храмов В.Ю. Особенности многочастотной генерации излучения эрбиевых лазеров // Изв. вузов. Приборостроение. – 2012. – Т. 55. – № 5. – С. 79–85.
- Risk W.P. Modeling of longitudinally pumped solid state lasers exhibiting reabsorption losses // J. Opt. Soc. Am. B. – 1988. – V. 5. – P. 1412–1423.

- 3. Fan T.Y., Sanchez A. Pump source requirements for end pumped lasers // IEEE J. of Quantum Electronics (QE). 1990. V. 26. № 2. P. 311-316.
- 4. Laporta P., Brussard M. Design criteria for mode size optimization in diode pumped solid state lasers // IEEE J. of Quantum Electronics (QE). 1991. V. 27. № 10. P. 2319–2326.
- Pfistner C., Albers P., Weber H.P. Influence of spatial mode matching in end-pumped solid state lasers // Appl. Phys. B. – 1992. – V. 54. – № 1. – P. 83–88.
- 6. Chen Y.F., Liao T.S., Kao C.F. et al. Optimization of fiber coupled laser diode end pumped lasers // IEEE J. of Quantum Electronics (QE). 1996. V. 32. № 11. P. 2010–2016.
- 7. Ткачук А.М., Разумова И.К., Мирзаева А.А., Малышев А.В., Гапонцев В.П. Up-конверсия и заселение возбужденных уровней иона эрбия в кристаллах LiY1-хErxF4 (x=0.003-1) при непрерывной накачке излучением InGaAs-лазерных диодов // Оптика и спектроскопия. 2002. Т. 92. № 1. С. 73–88.
- 8. Назаров В.В., Хлопонин Л.В., Храмов В.Ю., Федоров Н.А. Исследования пространственных характеристик излучения твердотельных лазеров в условиях продольной диодной накачки // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – № 4 (80). – С. 34–39.

Назаров Вячеслав Валерьевич	-	Санкт-Петербургский национальный исследовательский университет информационных технологий механики и оптики, кандидат технических
Хлопонин Леонид Викторович	_	наук, ст. научный сотрудник, lab255@grv.ifmo.ru Санкт-Петербургский национальный исследовательский университет
		информационных технологий, механики и оптики, кандидат технических наук, ст. научный сотрудник, l_khloponin@yahoo.com
Храмов Валерий Юрьевич	_	Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических
		наук, профессор, зав. кафедрой, khramov@grv.ifmo.ru
Федоров Никита Анатольевич	_	Санкт-Петербургский национальный исследовательский университет
-		информационных технологий, механики и оптики, студент, nikifedorov@mail.ru

УДК 621.373.826

МОДЕЛИРОВАНИЕ НЕУСТОЙЧИВОГО РЕЗОНАТОРА МОНОИМПУЛЬСНОГО Nd:YAG-ЛАЗЕРА С ГРАДИЕНТНЫМ ВЫХОДНЫМ ЗЕРКАЛОМ И.Н. Дубинкин, В.В. Назаров, Л.В. Хлопонин, В.Ю. Храмов

Представлен метод моделирования генерации лазерного излучения в неустойчивом радиально-симметричном резонаторе с выходным градиентным зеркалом. Приведены результаты исследований влияния параметров градиентного зеркала на пространственные и энергетические характеристики лазерного излучения. Проведено моделирование лазерной генерации в режимах активной и пассивной модуляции добротности. Проведен сравнительный анализ пространственных и энергетических характеристик излучения для указанных режимов.

Ключевые слова: моделирование, неустойчивый резонатор, градиентное зеркало, модуляция добротности.

Введение

Среди разрабатываемых в настоящее время лазеров твердотельные лазеры, в частности Nd:YAGлазеры, занимают одно из лидирующих мест по объему производства. Мощные импульсные Nd:YAGлазеры могут использоваться в различных технологических процессах благодаря компактности и надежности конструкции, стабильности параметров и возможности применения оптических волокон для доставки излучения до объекта воздействия. Они нашли широкое применение для обработки материалов, в медицине, научных исследованиях, лазерных системах с преобразованием частоты излучения.

Проблеме распространения лазерного излучения в открытых неустойчивых резонаторах, начиная с работы Сигмана [1], посвящено большое количество теоретических и экспериментальных работ, основные результаты которых подытожены в известных монографиях [2, 3]. Математические модели и основные численные методы решения указанной проблемы изложены в [4–8]. Однако предыдущие исследования [9, 10], имеющие непосредственное отношение к тематике настоящей работы, как правило, не учитывали такие факторы, как динамика поля лазерного излучения внутри резонатора, дифракционные эффекты, влияние усиления излучения в активной среде.

Математическая модель моноимпульсного лазера

Одним из наиболее эффективных методов решения задачи о лазерной генерации в случае радиально симметричного резонатора является метод, основанный на быстром преобразовании Ганкеля [1]. Кратко рассмотрим разработанный авторами алгоритм, реализующий математическую модель, основанную на указанном методе.

Первый шаг алгоритма заключался в том, что на основе быстрого преобразования Ганкеля был реализован метод Фокса и Ли для пустого резонатора, содержащего тонкую линзу, эквивалентную тепловой линзе активного элемента, а также градиентное выходное зеркало. В результате было получено распределение комплексной амплитуды поля, соответствующее суперпозиции собственных мод резонатора, что позволило определить потери для исследуемого резонатора. Указанное распределение задает стартовое «шумовое» поле, которое будет усиливаться в активной среде в процессе лазерной генерации.

На втором шаге был реализован метод Фокса и Ли для резонатора, содержащего активную среду и насыщающийся поглотитель. Начальным распределением поля здесь служило распределение, полученное на предыдущем шаге для пустого резонатора. Для моделирования процесса взаимодействия излучения внутри резонатора с активной средой и насыщающимся поглотителем были использованы уравнения Франца–Нодвика [6]. Стоит отметить, что на данном шаге критерием прекращения вычислений являлось снижение интенсивности излучения на несколько порядков относительно пикового значения.

В результате указанных процедур с шагом дискретизации по времени, равным времени обхода резонатора, определялось распределение интенсивности поля на выходном зеркале резонатора в ходе генерации. Полученные данные использовались для определения энергетических и пространственных характеристик импульса излучения лазерной генерации с учетом характеристик выходного зеркала резонатора.

Основными выражениями при реализации метода Фокса и Ли были уравнения, полученные в работах [1, 4–6] и преобразованные в соответствии с рассматриваемой задачей. Эти уравнения можно записать в следующем виде:

$$\begin{aligned} f_n &= r_0 \exp(\alpha n) \cdot E_1 \left(r_0 \exp(\alpha n) \right) \cdot \exp\left[-\frac{i\pi}{\lambda B} A r_0^2 \exp(2\alpha n) \right]; \\ g_{n+m} &= \alpha \frac{2\pi}{\lambda B} r_0 \rho_0 \exp(\alpha (n+m)) \cdot J_0 \left(\frac{2\pi}{\lambda B} r_0 \rho_0 \exp(\alpha (n+m)) \right); \\ \psi &= IFFT \left[IFFT \left(f_n \right) \cdot FFT \left(g_{n+m} \right) \right]; \\ E_2 \left(\rho_0 \cdot \exp(\alpha m) \right) &= i \frac{\psi(\rho_0 \exp(\alpha m))}{\rho_0 \exp(\alpha m)} \cdot \exp\left[-\frac{i\pi}{\lambda B} D \rho_0^2 \exp(2\alpha m) \right], \end{aligned}$$
(1)

где n, m – целочисленные индексы, соответствующие узлам расчетной сетки; A, B, D – элементы лучевой матрицы оптической системы; E_1 – начальное распределение комплексной амплитуды поля на опорной плоскости; E_2 – искомое распределение комплексной амплитуды поля; λ – длина волны излучения; J_0 – функция Бесселя нулевого порядка; r_0, ρ_0, α – числовые параметры, определяемые в соответствии с [4]; *FFT*, *IFFT* – операторы прямого и обратного быстрого преобразования Фурье соответственно.

Основными выражениями при моделировании процесса взаимодействия лазерного излучения с активной средой и насыщающимся поглотителем были уравнения, полученные в [7] и модифицированные в соответствии с решаемой задачей:

$$\begin{cases} U_i^{(AE,SH)} = \frac{U_i^{(AE,SH)} \cdot \exp\left[-\sigma^{(AE,SH)} U_i^{(AE,SH)} L^{(AE,SH)}\right]}{\exp\left[2\sigma^{(AE,SH)} \frac{c}{n^{(AE,SH)}} I_i \tau\right] + \exp\left[-\sigma^{(AE,SH)} U_i^{(AE,SH)} L^{(AE,SH)}\right] - 1};\\ I_i = \frac{I_i}{1 - \left(1 - \exp\left[-\sigma^{(AE,SH)} U_i^{(AE,SH)} L^{(AE,SH)}\right]\right) \cdot \exp\left[-2\sigma^{(AE,SH)} \frac{c}{n^{(AE,SH)}} I_i \tau\right]}, \end{cases}$$
(2)

где *i* – целочисленные индексы, соответствующие узлам расчетной сетки; $U^{(AE,SH)}$ – разность населенностей рабочих уровней в активной среде или насыщающемся поглотителе; $\sigma^{(AE,SH)}$ – сечение рабочего перехода в активной среде или насыщающемся поглотителе; $L^{(AE,SH)}$ – характерный размер активной среды или насыщающегося поглотителя; $n^{(AE,SH)}$ – показатель преломления активной среды или насыщающегося поглотителя; τ – длительность импульса излучения постоянной интенсивности на входе в среду; I_i – интенсивность излучения в *i*-ой точке расчетной сетки; *c* – скорость света в вакууме.

Заметим, что уравнения (2) являются частным решением задачи о взаимодействии импульса излучения с двухуровневой средой и применимы только в случае, если со средой взаимодействует импульс постоянной во времени интенсивности. По этой причине в рамках данной модели усиливаемый импульс представлялся последовательностью постоянных во времени элементарных импульсов.

Для определения пространственных характеристик импульса излучения лазерной генерации был проведен анализ распределения плотности энергии в фокальной плоскости модельной тонкой линзы, размещенной на выходе лазерного резонатора. С помощью выражений (1) по уровню 90% энергии опре-
делялся диаметр лазерного пучка вблизи фокуса указанной линзы, что, в свою очередь, позволяло легко оценить угловую расходимость излучения.

Результаты исследования характеристик лазерного излучения

На основе разработанного алгоритма было проведено численное исследование генерации излучения в неустойчивом резонаторе, схема которого показана на рис. 1. С целью упрощения задачи оптимизации геометрические параметры резонатора были зафиксированы.



Рис. 1. Геометрия исследуемого резонатора: АЭ – активный элемент; М1, М2 – зеркала резонатора; А – блок пассивной модуляции добротности; Б – блок активной модуляции добротности; Д – диафрагма; *F* – тепловая линза. Оптическая сила тепловой линзы 0,5 м⁻¹. Все линейные величины приведены в мм

Для данной геометрии коэффициент увеличения резонатора составляет 1,6, а эквивалентное число Френеля равно 7,5. В качестве насыщающегося поглотителя в режиме пассивной модуляции добротности был использован кристалл GSGG:Cr:Mg.

Профиль коэффициента отражения выходного градиентного зеркала определяется выражением

$$R = R_0 \exp\left(-2\frac{r^n}{\omega_m^n}\right),\,$$

где R_0 – максимальный коэффициент отражения; r – радиальная координата, ω_m – радиус (характерный размер) градиентного зеркала; n – показатель степени супергауссова профиля коэффициента отражения выходного зеркала.



Рис. 2. Зависимость энергии импульса излучения от показателя степени супергауссова профиля коэффициента отражения выходного зеркала: 1 – активная модуляция добротности; 2 – пассивная модуляция добротности с использованием насыщающегося поглотителя с начальным пропусканием 60%. Параметры градиентного зеркала: максимальный коэффициент отражения – 0,25; радиус градиентного зеркала – 1,35 мм

Параметры градиентного зеркала				
Максимальный коэффициент отражения	0,25–0,3			
Радиус (характерный размер) градиентного зеркала, мм	1,2–1,5			
Показатель степени супергауссова профиля	6–8			
коэффициента отражения выходного зеркала				
Параметры насыщающегося поглотителя				
Начальный коэффициент поглощения, см ⁻¹	1,0–1,5			
Характерный линейный размер, см	0,1–0,3			
Концентрация активатора, 10 ¹⁸ см ⁻³	5,5–8,5			

Таблица 1. Характеристики насыщающегося поглотителя и градиентного зеркала



Рис. 3. Поперечное распределение интенсивности в лазерном пучке: 1 – активная модуляция добротности; 2 – пассивная модуляция добротности с использованием насыщающегося поглотителя с начальным пропусканием 80%; 3 – пассивная модуляция добротности с использованием насыщающегося поглотителя с начальным пропусканием 60%. Параметры градиентного зеркала: минимальный коэффициент отражения – 0,25; радиус градиентного зеркала – 1,35 мм; показатель степени супергауссова профиля коэффициента отражения выходного зеркала – 6

Результаты оптимизации параметров градиентного зеркала и насыщающегося поглотителя с целью достижения более однородного поперечного распределения интенсивности в импульсе лазерного излучения и повышения энергии импульса излучения в режиме пассивной модуляции добротности приведены в табл. 1. Зависимость энергии импульса излучения от показателя степени супергауссова профиля коэффициента отражения выходного зеркала показана на рис. 2. Дальнейшие результаты приведены для указанных в табл. 1 значений параметров поглотителя и зеркала.

Поперечное распределение плотности мощности в импульсе лазерного излучения для случаев активной модуляции добротности и пассивной модуляции добротности с использованием насыщающихся поглотителей с различным начальным пропусканием показано на рис. 3. Результаты приведены для момента достижения пиковой мощности в импульсе генерации. Из приведенных на рис. 3 данных очевидно, что в режиме пассивной модуляции добротности повышается однородность лазерного пучка.

Оценка неоднородности поперечного распределения интенсивности лазерного пучка Δ проводилась при помощи выражения

$$\Delta = \frac{I_{\max} - I_0}{I_{\max}}$$

где I_0 – минимальное значение интенсивности в приосевой области; I_{max} – максимальное значение интенсивности лазерного пучка.

Основные характеристики излучения для различных режимов модуляции приведены в табл. 2. Следует отметить, что для оценки значения коэффициента усиления слабого сигнала в активной среде, исходя из энергии накачки, были использованы выражения из работы [10].

Параметр	Значение			
Параметр	Активная модуляция	Пассивная модуляция		
Энергия импульса генерации при энергии импульса накачки 50 Дж, мДж	1026	814		
Длительность импульса излучения по полувысоте, нс	30	50		
Угловая расходимость излучения, мрад	1,8	1,5		
Угловая расходимость излучения, мрад				
Неоднородность по поперечному сечению, %	18	11		

Таблица 2. Характеристики излучения лазерной генерации

Заключение

В работе приведено описание разработанной математической модели, позволяющей исследовать генерацию лазерного излучения в неустойчивом резонаторе с выходным градиентным зеркалом. Была проведена оптимизация параметров градиентного зеркала с целью достижения более однородного поперечного распределения плотности мощности излучения и максимальной энергии импульса генерации. Представлены результаты численного моделирования параметров лазерного излучения для режимов активной и пассивной модуляции добротности. Сравнительный анализ указанных режимов показал, что использование метода пассивной модуляции добротности позволяет повысить однородность поперечного распределения интенсивности в импульсе лазерного излучения и снизить угловую расходимость излучения. Проведен анализ влияния параметров градиентного зеркала и насыщающегося поглотителя на характеристики лазерного излучения.

Литература

- 1. Siegman A.E., Miller H.Y. Unstable optical resonator loss calculations using the prony method // Applied Optics. 1970. V. 9. № 12. P. 2729–2763.
- 2. Ананьев Ю.А. Оптические резонаторы и проблемы расходимости лазерного излучения. М.: Наука, 1979. 328 с.
- 3. Ананьев Ю.А. Оптические резонаторы и лазерные пучки. М.: Наука, 1990. 264 с.
- 4. Murphy W.D., Bernabe M.L. Numerical procedures for solving nonsymmetrical eigenvalue problems associated with optical resonators // Applied Optics. 1978. V. 17. № 15. P. 2358–2365.
- 5. Карамзин Ю.Н., Конев Ю.Б. Численные методы исследования работы неустойчивых телескопических резонаторов с учетом дифракции и эффекта насыщения в активной среде. Препринт ИПМ им. М.В. Келдыша АН СССР, 1974. № 57. 14 с.
- 6. Елкин Н.Н., Напартович А.П. Прикладная оптика лазеров. М.: ЦНИИатоминформ, 1989. 183 с.
- Lee M. Frantz, John S. Nodvik. Theory of Pulse Propagation in a Laser Amplifier // J. Appl. Phys. 1963. V. 34. – P. 2346–2349.
- 8. Аладов А.В., Беззубик В.В., Белашенков Н.Р., Карасев В.Б., Назаров В.В. и др. Применение зеркал с переменным коэффициентом отражения в компактных твердотельных лазерных системах // Изв. вузов. Приборостроение. 1998. Т. 41. № 3. С. 53–56.
- Morin M. Graded reflectivity mirror unstable laser resonators // Optical and Quantum Electronics. 1997. V. 29. – № 8. – P. 819–866.
- 10. Зверев Г.М., Голяев Ю.Д., Шалаев Е.А., Шокин А.А. Лазеры на алюмоиттриевом гранате с неодимом. М.: Радио и связь, 1985. 144 с.

Дубинкин Илья Николаевич	-	Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, студент, ilva dubinkin@mail.ru
Назаров Вячеслав Валерьевич	_	Санкт-Петербургский национальный исследовательский университет
		информационных технологий, механики и оптики, кандидат технических наук, ст. научный сотрудник, lab255@grv.ifmo.ru
Хлопонин Леонид Викторович	_	Санкт-Петербургский национальный исследовательский университет
		информационных технологий, механики и оптики, кандидат технических
		наук, ст. научный сотрудник, l_khloponin@yahoo.com
Храмов Валерий Юрьевич	_	Санкт-Петербургский национальный исследовательский университет
		информационных технологий, механики и оптики, доктор технических
		наук, профессор, зав. кафедрой, khramov@grv.ifmo.ru

6

ЛАЗЕРНЫЕ И БИОМЕДИЦИНСКИЕ ТЕХНОЛОГИИ

УДК 681.784.88, 535.36 ТЕРАГЕРЦОВЫЕ СПЕКТРЫ ОТРАЖЕНИЯ КОЖИ С ДЕРМАТИТАМИ И МОРФОЛОГИЧЕСКИМИ ИЗМЕНЕНИЯМИ О.А. Смолянская, А.А. Езерская, И.В. Прожеев, Е.А. Стрепитов

Проведены работы по поиску возможности диагностики дерматитов и морфологических изменений кожи человека с помощью излучения терагерцовой области частот (0,05–2,0 ТГц). Особенности различных типов заболеваний кожи человека *in vivo* проявляются во всем исследуемом диапазоне частот, особенно в колебательной области (1,5–2,0 ТГц). Обусловлены они обратным рассеянием на новообразованиях кожи в ее верхних слоях. В терагерцовых спектрах отражения хорошо различимы спектральные линии различных дерматитов, пигментных пятен, гематом. Терагерцовое излучение хорошо проникает через перевязочные средства. При этом за одно сканирование, длящееся около одной минуты, обрабатываются спектры не только перевязочного материала, но и различных слоев кожи. Ключевые слова: терагерцовая диагностика, дерматиты, морфологические изменения кожи.

Введение

В настоящее время дерматологи, онкологи и пластические хирурги осуществляют диагностику кожных заболеваний человека различными методами и приборами в зависимости от тяжести заболевания. К ним относятся дерматоскоп [1, 2], анализ крови (аллергены, антитела, антигены) [3], анализ соскоба (полимеразная цепная реакция, микроскопия), цитология [4], гистология (биопсия) [5], микробиология (посев на флору с определением чувствительности к антибиотикам и противогрибковым препаратам) [4], люминесцентная лампа Вуда [6]. Это относительно простые способы диагностики. В более тяжелых случаях, когда, например, требуется распознать опухоль кожи, применяют более сложную и дорогостоящую аппаратуру [7]. Это оптическая когерентная томография [8], ультразвуковая и магнитно-резонансная томография [9]. Однако современные диагностические методы не могут определить, является ли определенное пятно опухолью, и, если да, то доброкачественная она или злокачественная. Они имеют ограниченные возможности оценки начального размера повреждения кожи и диагностики процесса заживления. Приборы опасны, так как подвергают человека воздействию вредного для здоровья ионизирующего излучения. Они имеют большое число ограничений в использовании, а также высокую стоимость. Некоторые ограничения связаны со сложностью передачи излучения через перевязочные средства и воспаленные участки кожи человека.

Разрабатываемые в настоящее время технологии, использующие терагерцовое (ТГц) излучение в биомедицинской томографии и спектроскопии, выгодно отличаются от современных методов исследования. Основное преимущество ТГц диагностики патологий по сравнению с существующими методами диагностики – это возможность за одно сканирование осуществить обработку спектральных характеристик биоткани, т.е. распознать патологию и одновременно осуществить визуализацию ее внутренней структуры, т.е. рассчитать площадь патологии.

В настоящей работе исследовалась возможность бесконтактной диагностики кожи человека, пораженной дерматитами и морфологическими изменениями, неионизирующим ТГц излучением диапазона 0,05–2,0 ТГц. При этом оценивалась возможность осуществлять диагностику через перевязочные средства. Появление в ТГц спектре дерматита отражения, отличного от спектра нормальной кожи, по нашим оценкам, свидетельствует о наличии воспалительного процесса. Происходит обратное рассеяние ТГц сигнала от гематомы – скопления крови при повреждениях тканей. Характерные размеры гематомы сопоставимы с частотой ТГц сигнала.

Материалы и методы

ТГц рефлектометрический спектрометр. Исследование спектров отражения кожи в диапазоне 0,05–2,0 ТГц проводилось с помощью ТГц рефлектометрического спектрометра. Фотография общей схемы установки представлена на рис. 1. Описание работы спектрометра для диагностики заболеваний кожи человека приведено в работе [10]. Отношение сигнал/шум не хуже 600, средняя мощность не менее 30 мкВт, спектральный диапазон 0,02–2,0 ТГц, пространственное разрешение 300 мкм, длительность импульса 3 пс, спектральное разрешение рефлектометрического спектрометра составляет 5 ГГц, погрешность определения – 2,5%. В основе работы устройства лежит принцип записи распределения электрической компоненты поля до и после отражения от кожи. Восстановление спектра ТГц излучения реализуется с помощью вычисления интеграла Фурье по записанным данным.

Испытуемый плотно прижимает пораженную дерматитом область кожи к центральной части экрана (рис. 2). В качестве экрана используется зафиксированная в строго вертикальном положении плос-

копараллельная пластина, прозрачная в используемом диапазоне излучения, при условии строгой горизонтальности оптической оси. Таким образом, излучение падает на экран, затем на кожу. Тем самым обеспечивается ортогональность плоскости кожи к падающему излучению.



Рис. 1. Схема терагерцового рефлектометрического спектрометра



Рис. 2. Тестовая диагностика дерматита кожи: с лейкопластырем (а); с медицинским бинтом (б)

Подготовка кожи, пораженной дерматитом, для диагностики. Объектами диагностики являлись пораженные дерматитами участки кожи человека *in vivo*. Проводилась диагностика различных типов дерматитов, таких как папиллома, гематома, химический ожог, крапивница, а также пигментные пятна. На рис. 3 представлены фотографии заболеваний, которые диагностировались на лабораторном макете ТГц спектрометра.

В некоторых экспериментах на участок кожи с дерматитом был наложен влагостойкий воздухопроницаемый бактерицидный лейкопластырь фирмы SANITAPLAST (Турция, производитель Бетасан), изготовленный из тонкой полимерной перфорированной ленты, покрытой гипоаллергенным кремом. Сорбционная подушечка выполнена из 100%-ной вискозной марли.

В экспериментах для каждого образца производилось по 5 измерений. Статистическая обработка результатов производилась в предположении нормальности распределения Стьюдента с доверительной вероятностью 0,95.

Порядок проведения измерений тестовой диагностики кожи

- 1. Запустить лабораторный макет ТГц рефлектометрического спектометра согласно порядку включения.
- 2. На персональном компьютере запустить программу для управления лабораторным макетом ТГц рефлектометрического спектометра.

- 3. Согласно описанию применения программы для управления лабораторным макетом ТГц рефлектометрического спектометра произвести измерение сигнала-референса ТГц спектра отражения.
- 4. Пациенту плотно прижать участок кожи, на котором выявлено заболевание, к экрану. Процедуру можно выполнить с нанесенным на кожу лекарственным средством и перевязочным материалом (лейкопластырь, бинт).
- 5. Произвести сканирование ТГц сигнала. Процедура длится около двух минут.
- 6. Выделить на графике ТГц сигнала области, отраженные от перевязочного средства (если есть), а также области слоев кожи.
- 7. Обработать сигнал, полученный от пораженной кожи, с целью получения частотной зависимости ТГц отражения диагностируемой области кожи.
- 8. Оценить характер диагностируемого заболевания кожи и (или) степень его заживления по амплитуде и характерным частотам полученного ТГц спектра отражения.





Рис. 3. Фотографии дерматитов на руке: заживший химический ожог кожи на запястье руки (женщины 34 года) (а); дерматит кожи (крапивница) на внешней стороне руки (женщины 21 года) (б); папиллома кожи на запястье руки (женщины 34 года) (в); гематома кожи на внутренней стороне сгиба руки (женщины 34 года) (г); пигментное пятно кожи внешней стороны руки (женщины 21 года) (д)

Результаты и их обсуждение

Рис. 4 демонстрирует возможность проведения ТГц диагностики кожи перебинтованной руки или при приклеивании на кожу бактерицидного лейкопластыря.



Рис. 4. ТГц сигнал, отраженный от кожи перебинтованной руки человека (серый) и с приклеенным на нее бактерицидным лейкопластырем (черный)



Рис. 5. ТГц спектры отражения участка кожи с крапивницей (1), гематомой (2), нормального участка кожи (3), участка кожи, поврежденной химическим ожогом (4), пигментным пятном (5) и папилломой (6)

Первый сигнал отражения появляется через 7 пс. Он несет информацию об экране, к которому прижат диагностируемый участок кожи. Второй сигнал наблюдается через 16 пс от начала сканирования. Этот сигнал характеризует появление бинта или лейкопластыря. Третий сигнал обнаруживается через 22–25 пс после начала сканирования. Он позволяет получить ТГц спектр отражения кожи. В случае если на кожу не приклеен лейкопластырь, то второй сигнал наблюдаться не будет. При этом сигнал от кожи

будет прослеживаться через 24 пс после начала сканирования. Разница отражений лейкопластыря и участка кожи в 6–9 пс позволяет хорошо их различать невооруженным глазом.

На рис. 5 приводится сравнение ТГц спектров различных дерматитов и морфологически измененной кожи человека с нормальной кожей. Важно отметить характерные пики отражения кожи, которые характеризуют определенное заболевание и отличают его от другого заболевания. Крапивница (1) имеет отличительный пик отражения на 1,1 и 1,42 ТГц; гематома (2) – на 1,62 ТГц; химический ожог (4) – на 1,1 и 1,84 ТГц; пигментное пятно (5) – на 1,43, 1,62 и 1,69 ТГц, папиллома (6) – на 1,62 и 1,69 ТГц. При этом ТГц спектр отражения неповрежденной кожи (3) не имеет особенностей в диапазоне 0,05–2 ТГц.

Заключение

Проведены работы по поиску возможности диагностики дерматитов и морфологических изменений кожи человека с помощью излучения ТГц области частот 0,05–2,0 ТГц. Особенности различных типов заболеваний кожи человека *in vivo* проявляются во всем исследуемом диапазоне частот, особенно в колебательной области (1,5–2,0 ТГц). Обусловлены они обратным рассеянием на новообразованиях кожи в ее верхних слоях. В ТГц спектрах отражения хорошо различимы спектральные линии различных дерматитов, пигментных пятен, гематом. ТГц излучение хорошо проникает через перевязочные средства. При этом за одно сканирование, длящееся около одной минуты, обрабатывается спектр не только перевязочного материала, но и различных слоев кожи.

Таким образом, показана перспективность создания метода бесконтактной диагностики кожи человека в ТГц области частот. Доступность открытых участков кожи позволят проводить исследования *in vivo* на человеке и набирать статистику. Легко будет доставить излучение до пораженного участка, не требуя сложной конструкции «оптической руки».

Работа поддержана Министерством образования и науки Российской Федерации (ГК № 14.512.11.0020).

Литература

- 1. Mark G. Lebwohl, Warren R. Heymann, John Berth, Jan Coulson. Treatment of skin disease: comprehensive therapeutic strategies. 3-d ed. Hardcover: Elsevier, 2010. 833 p.
- 2. US 2012/0008838 (A1) Guyon Isabelle [US]; Barnhill Stephen [US]. System and metod for remote melanoma screening, pub. 21.12.2010.
- 3. Данилова Л.А. Анализы крови и мочи. СПб: Салит-Медкнига, 2003. 128 с.
- 4. Глик Б., Пастернак Дж. Молекулярная биотехнология. Принципы и применение: Пер. с англ. М.: Мир, 2002. 589 с.
- Kiernan J.A. Histological and Histochemical Methods: Theory and Practice. 4-th ed. Bloxham, UK: Scion, 2008. – 481 p.
- Методы клинических лабораторных исследований / Под ред. проф. В.С. Камышникова. 4-е изд. М.: МЕДпресс-информ, 2011. – 752 с.
- 7. Давыдов М.И., Ганцев Ш.Х. Атлас по онкологии. М.: МИА, 2007. 416 с.
- 8. Tuchin Valery V. Handbook of Coherent-Domain, Biomedical Diagnostics, Environmental Monitoring, and Materials Science. 2-d ed. NY: Springer Science+Business Media, 2013. 1332 p.
- 9. Мэнсфилд П. Быстрая магнитно-резонансная томография // Успехи физических наук. 2005. Т. 175. № 10. С. 1044–1052.
- 10. Езерская А.А., Цуркан М.В., Смолянская О.А. Идентификация аминокислот, входящих в состав кожи человека, с помощью спектров диапазона 0,05–1,2 ТГц // Научно-технический вестник информационных технологий, механики и оптики. 2013. № 2 (84). С. 50–54.

Смолянская Ольга Алексеевна	-	Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат физмат.
Езерская Анна Александровна	_	наук, доцент, o_smolyanskaya@mail.ru Санкт-Петербургский национальный исследовательский университет
		информационных технологий, механики и оптики, студент, a.a.ezerskaya@gmail.com
Прожеев Игорь Владимирович	-	Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, студент, prozbeev iv@gmail.com
Стрепитов Евгений Александрович	-	Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, студент, e- stree@vandex_ru

УДК 532.542.4:533.6.011.32:612.215.41 МОДЕЛИРОВАНИЕ ВОЗДУШНОГО ПОТОКА В КАНАЛАХ НЕРЕГУЛЯРНОЙ ФОРМЫ

А.А. Воронин, Г.Н. Лукьянов, Р.В. Неронов

Проанализированы основные подходы к компьютерному моделированию движения воздуха в носовых каналах человека при дыхании. Описаны различные модели турбулентности, которые могут быть применены для расчета скорости и давления воздуха. Проиллюстрированы результаты экспериментального измерения параметров воздушного потока на входе в носовую полость. Приведена геометрическая модель носовой полости, полученная из снимков компьютерной томографии, а также результаты моделирования движения воздуха в рамках данной модели. Приведены полученные поля скоростей и температур во время вдоха и выдоха.

Ключевые слова: численное моделирование воздушных потоков, компьютерная томография, носовая полость, турбулентность.

Введение

Проблема моделирования движения газа в каналах нерегулярной формы остро стоит в различных областях науки, в том числе и биомедицинских исследованиях. Так, осуществление операций по корректированию формы внутренних носовых перегородок, лечению различных заболеваний верхних дыхательных путей требует высокой квалификации оперирующего специалиста в сочетании с максимальным количеством информации об особенностях строения дыхательных путей пациента. Нередко индивидуальные различия в строении носовых каналов и пазух становятся причиной врачебных ошибок, которые могут привести к серьезным осложнениям, вплоть до летального исхода в случае повреждения прилегающих отделов головного мозга. Одним из возможных путей решения указанной проблемы может стать создание неинвазивной методики, позволяющей оперирующему специалисту получать данные об индивидуальных особенностях строения внутренних носовых каналов пациента, планировать траектории надрезов и анализировать результат планируемых действий до начала операции. Достижение таких результатов возможно при условии наличия виртуальной геометрической модели верхних дыхательных путей и возможностей модели раимие.

В последние годы несколькими группами исследователей был предпринят ряд попыток построения оптимальной геометрической модели внутренних носовых каналов человека с последующим моделированием движения воздуха при дыхании. Первые такие работы связаны с результатами, полученными группой исследователей из университета Пенсильвании (США) в 1993 году [1]. Исходная геометрическая модель верхних дыхательных путей была реконструирована по данным компьютерной томографии, а моделирование движения воздуха осуществлялось для правой части носовой полости для вдоха и выдоха при различных фиксированных значениях массового расхода. Авторы исследования пришли к выводу, что движение воздуха в общем случае носит турбулентный характер, однако, ввиду незначительного изменения профиля скоростей потока с изменением массового расхода, в невозбужденном состоянии спокойного дыхания (нормированный массовый расход воздуха примерно равен 150 мл/с) численное моделирование на основе моделей для ламинарного режима движения приводит к удовлетворительным результатам и хорошей корреляции с экспериментальными данными.

В последующие годы было разработано множество физических моделей движения потока воздуха при дыхании, основанных на геометрических данных, полученных по результатам компьютерной томографии [2–5]. При этом в более ранних работах осуществлялся расчет исключительно для стационарного режима движения воздуха [2, 5], в рамках которого были получены глобальные поля распределения скоростей и давлений потока без учета временных пульсаций данных динамических величин. Рядом исследователей [3, 4] были предприняты попытки описания движения воздуха при дыхании с помощью моделей для турбулентного течения жидкости: моделей k– ε и Спаларта–Альмараса [3], а также k– ω модели Уилкокса [4]. Данные физические модели турбулентности позволяют осуществлять моделирование движения крупномасштабных турбулентных вихревых структур, но, вместе с тем, не дают возможности описать пульсации мгновенных значений пространственных компонент скорости, а также давления внутри потока, так как в своей основе содержат значения осредненных по Рейнольдсу параметров уравнений Навье–Стокса для движения жидкости.

В настоящем исследовании основное внимание было уделено разработке подробной нестационарной физической модели движения воздуха при дыхании, наиболее адекватно описывающей перемещение характерных для турбулентного потока вихревых структур, а также построению качественной сетки конечных элементов, позволяющей реализовать наибольшее доступное пространственное разрешение с учетом имеющихся в наличии вычислительных ресурсов.

Экспериментальное исследование характеристик потока воздуха при дыхании

Первым этапом исследования воздушного потока в носовой полости человека при дыхании стало измерение характеристик воздуха на входе в носовые каналы во время вдоха-выдоха, которое было осу-

ществлено при помощи многоканального ринологического прибора, разработанного на кафедре компьютерной теплофизики и энергофизического мониторинга НИУ ИТМО [6, 7]. Диагностический элемент данного прибора представляет собой клипсу с закрепленными датчиками температуры миниатюрных размеров (рис. 1), устанавливаемую на носовой перегородке на входе в носовую полость. Датчики позволяют фиксировать изменение температуры воздуха в процессе дыхания без искажения естественных характеристик воздушной струи.



а







Рис. 2. Пульсации температуры воздуха на входе в носовую полость человека при дыхании

Результаты измерений пульсаций температуры на входе в носовую полость приведены на рис. 2. На основании полученных экспериментальных данных была разработана методика диагностики заболеваний верхних дыхательных путей [6, 8, 9].

Обработка результатов таких измерений позволяет определить наличие патологических отклонений в функционировании дыхательной системы человека, однако не дает возможности выявить особенности строения носовых каналов, связанные с диагностируемым заболеванием, что существенно ограничивает возможность применения данного метода при планировании оперативного вмешательства. В связи с данной проблемой следующим этапом исследования стала разработка подробной геометрической модели внутренних носовых каналов и последующее моделирование движения воздуха в рамках верхних дыхательных путей.

Построение геометрической модели внутренних носовых каналов

Реконструирование геометрической модели верхних дыхательных путей осуществлялось по данным компьютерной томографии пациентов отоларингологического отделения медицинского центра «Адмиралтейские верфи» (Санкт-Петербург). Такие данные представляют собой набор изображений поперечных срезов черепной коробки человека, расположенных с шагом 0,625 мм. При помощи специализированного программного обеспечения Mercury Amira из изображений срезов были выделены части, соответствующие воздуху, находящемуся внутри черепной коробки. После обобщения заданных параметров сегментации на весь набор срезов и их последующего наложения, по данным томографии было восстановлено трехмерное изображение внутренних носовых каналов (рис. 3).

В рамках геометрической модели (рис. 3) при помощи программных пакетов Altair Hypermesh и Ansys Icem CFD были построены сетки конечных элементов: нерегулярная треугольная поверхностная сетка (800 000 элементов), нерегулярная тетраэдрическая объемная сетка (9 000 000 элементов с максимальным размером грани 0,3 мм, что соответствует доступной разрешающей способности томографа).





Методы расчета турбулентного движения воздушного потока

Для описания движения жидкости как сплошной среды используется система уравнений Навье– Стокса, включающая в себя уравнение неразрывности (1), а также уравнения движения для трех компонент скорости потока (2).

$$\frac{\partial \rho}{\partial \tau} + \frac{\partial \rho u_j}{\partial x_j} = 0 , \qquad (1)$$

$$\frac{\partial \rho u_i}{\partial \tau} + \frac{\partial \rho u_j u_i}{\partial x_j} = -\frac{\partial P}{\partial x_i} + \frac{\partial \tau_{ij}}{\partial x_j}, \qquad (2)$$

где u_i , u_j (i, j = 1, 2, 3) – проекции мгновенной скорости потока; x_i , x_j (i, j = 1, 2, 3) – пространственные координаты; τ – время; ρ – плотность воздуха; P – мгновенное значение давления в точке потока.

Уравнения (1) и (2) содержат переменные u_i , u_j для мгновенных скоростей потока, изменение

которых в пространстве и времени при турбулентном режиме течения имеет настолько сложный характер, что не поддается точному расчету. По этой причине все используемые в настоящее время математические модели турбулентности предполагают внесение в систему уравнений (1) и (2) ряда упрощений, наиболее распространенным из которых является так называемое осреднение по Рейнольдсу, состоящее в замене мгновенных значений скоростей и давлений суммами их осредненных и пульсационных значений. Существуют различные модификации таких осредненных уравнений, которые были положены в основу математических моделей турбулентности – так называемые k– ϵ и k– ω модели [10], гибридная модель Ментера [11] и многие другие. Принципиальным недостатком указанных моделей является исходное осреднение пульсирующих величин (скорости, давления) по всему частотному спектру колебаний, что исключает возможность моделирования мелкомасштабных турбулентных вихрей, посредством которых происходит диссипация кинетической энергии потока.

Указанный недостаток может быть исключен посредством использования методов LES (Large Eddy Simulation – моделирование больших вихрей) и DES (Detached Eddy Simulation – моделирование отсоединенных вихрей), получивших широкое распространение в последние годы. В основе данных методов лежит прямой численный расчет искомых параметров турбулентного течения в рамках масштабов, заданных специальным фильтром. Разрешение масштабов, меньших размера фильтра, происходит по-

средством приближенного моделирования. В роли фильтра, как правило, используется расчетная сетка конечных элементов. Таким образом, часть частотного спектра турбулентных колебаний искомых величин рассчитывается точно, при помощи уравнений (1) и (2). Вихри наименьших размеров, точный расчет которых не представляется технически возможным, моделируются через осредненные значения скоростей и давлений потока.

В настоящем исследовании было осуществлено моделирование движения воздуха при помощи математической модели DES. Масштабы, меньшие размеров расчетной сетки, моделировались при помощи гибридной модели Ментера [11].

Результаты моделирования

Моделирование движения воздуха в рамках геометрической модели (рис. 3) было произведено при помощи программного пакета Ansys Fluent.

Граничные условия для расчета: p = 0, T = 293 K – на входе в носовую полость; $u_1 = u_2 = u_3 = 0$, T = 310 K – на стенке; $P = 50 \cos(\pi(0,869\tau - 0,47828)) + 50$ – давление в носоглотке при вдохе; $P = 50 \sin(\pi(1,17647\tau + 0,5)) - 50$ – давление в носоглотке при выдохе; T = 310 K – температура в носоглотке.

Был осуществлен нестационарный расчет для полного цикла вдох-выдох длительностью 4 с с шагом 0,1 с. Для расчета каждого временного шага использовалось 50 итераций. Результаты расчета приведены на рис. 4, 5. Рис. 4 иллюстрирует распределение скоростей потока воздуха при дыхании (во время вдоха). Максимальные значения мгновенной скорости в каждый момент времени наблюдаются на входе в носовую полость как при вдохе, так и при выдохе. При этом пространственное распределение скоростей потока носит неравномерный характер, претерпевая колебательные изменения с течением времени.



Рис. 4. Поле скоростей потока воздуха при дыхании: линии тока (а); контурный график (б)

Изображение линий тока вдоль продольной плоскости носовых каналов (рис. 4, а) демонстрирует ярко выраженную вихревую природу потока воздуха. Наиболее сложные траектории движения отмечены в области клиновидной пазухи, а также нижнего носового хода. Полученная турбулентная картина движения воздуха полностью согласуется с основной функцией, выполняемой верхними дыхательными путями – интенсификацией процессов тепло- и массообмена, обеспечивающих достаточный прогрев, фильтрацию и увлажнение воздуха, поступающего в дыхательную систему из окружающей среды.

Рис. 5 иллюстрирует распределение температуры воздуха, поступающего в носовую полость человека во время вдоха. Характер рассчитанного поля температур позволяет сделать вывод, что к моменту достижения околоносовых пазух, воздух оказывается прогретым фактически до уровня температуры тела. Полученные в результате численного моделирования значения температуры, скорости и давления воздуха в носовой полости в процессе дыхания соответствуют данным, полученным при помощи ринологического прибора (рис. 1) и описанным, в частности, в работе [6].



Рис. 5. Поле температур воздуха при дыхании (вдох)

Заключение

К основным результатам настоящего исследования следует отнести построение подробной геометрической модели верхних дыхательных путей на основе снимков компьютерной томографии, а также последующее моделирование движения воздуха при дыхании. Расчет параметров турбулентного течения был произведен на основе математической модели отсоединенных вихрей, что позволило получить достаточно полную картину течения, включающую в себя перемещение турбулентных вихрей в пространстве и их изменение во времени. Полученные результаты расчета (рис. 4, 5) представляют несомненный практический интерес также и для медицинских работников, так как позволяют проанализировать параметры проходимости носовых ходов и вентилируемости околоносовых пазух.

В настоящее время авторы продолжают работу над обработкой расчетных данных моделирования, а также созданием твердотельной модели внутренних носовых каналов, на базе которой планируется дальнейшее экспериментальное исследование параметров турбулентного движения воздуха.

Литература

- Hahn I., Scherer PW., Mozell M.M. Velocity profiles measured for airflow through a large-scale model of the human nasal cavity // J. Appl. Physiology. – 1993. – V. 75 (5). – P. 2273–2287.
- Tang H., Tu J.Y., Li H.F., Au-Hijleh B., Xue C.C., Li C.G. Dynamic Analysis of Airflow Features in a 3D Real-Anatomical Geometry of the Human Nasal Cavity // 15th Australasian Fluid Mechanics Conference. – Sydney, Australia, 13–17 December 2004. – Paper AFMC00174. – P. 1–3.
- Zhao K., Dalton P., Yang G.C., Scherer P.W. Numerical Modeling of Turbulent and Laminar Airflow and Odorant Transport during Sniffing in the Human and Rat Nose // Chemical Senses Advance Access. – 2006. – V. 31. – № 2. – P. 107–118.
- Wen J., Inthavong K., Tian Z.F., Tu J.Y., Xue C.L., Li C.G. Airflow Patterns in Both Sides of a Realistic Human Nasal Cavity for Laminar and Turbulent Conditions // 16th Australasian Fluid Mechanics Conference. – Crown Plaza, Gold Coast, Australia, 2–7 December 2007. – P. 68–70.
- Zhao K., Cowart B.J., Rawson N.E., Scherer P.W., Clock K.T. Nasal airflow and odorant transport modeling in patients with chronic rhinosimusitis // Thomas Jefferson University Department of Otolaryngology. – Head and Neck Surgery Faculty Papers. – 2007. – Paper 10.
- 6. Рассадина А.А. Измерения и анализ флуктуаций температуры, скорости и давления в каналах нерегулярной формы: Автореф. дис. ... канд. техн. наук. СПб: СПбГУ ИТМО, 2007. 18 с.
- 7. Лукьянов Г.Н., Рассадина А.А. Применение миниатюрных термисторов в качестве термоанемометров // Научно-технический вестник СПбГУ ИТМО. – 2005. – № 2 (18). – С. 68–72.
- 8. Лукьянов Г.Н., Рассадина А.А., Усачев В.И. Определение состояния человека по характеристикам его дыхания // Научно-технический вестник СПбГУ ИТМО. 2005. № 2 (18). С. 73–77.
- 9. Пилипенко Н.В., Казарцев Я.В. Оптимальное планирование эксперимента при идентификации процессов теплообмена сенсоров теплового потока // Изв. вуз. Приборостроение. – 2011. – Т. 54. – № 5. – С. 88–93.

10. Wilcox D.C. Turbulence Modeling for CFD. - DCW Industries, Inc., 1993. - 460 p.

11. Menter F.R. Zonal Two Equation $k - \omega$ Turbulence Models for Aerodynamic Flows // AIAA Paper. – 1993. – V. 93 (2906). – 21 p.

Воронин Алексей Анатольевич	_	Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, ale vor@rambler.ru
Лукьянов Геннадий Николаевич	_	Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, профессор, зав. кафедрой, Gen-lukjanow@yandex.ru
Неронов Роман Витальевич	-	Военно-медицинская академия им. С.М. Кирова, кандидат медицинских наук, преподаватель, nrvspb@mail.ru

УДК 681.02

МЕТОДИКА ИЗМЕРЕНИЯ И РАСЧЕТА ХАРАКТЕРИСТИК ДЫХАНИЯ С.И. Ханков, А.Ю. Кормилицын, В.И. Скорубский

Рассматривается теплофизическая модель энергетического обмена воздушных потоков вдоха и выдоха, которая использует свойства турбулентности. Показано, что на основе этой модели могут быть построены датчики измерения и методы расчета параметров потоков, характеризующих объем легких и скорость воздушных потоков. Изменения параметров потока относительно нормальных могут быть связаны с нарушениями работы легких и сердечнососудистой системы. Результаты исследований модели легли в основу конструирования датчика параметров вихревых потоков дыхания и алгоритмов расчета параметров в портативных приборах функциональной диагностики. Ключевые слова: дыхание, мониторинг состояния, турбулентность, аускультация, рекуперация, конвективный теплообмен.

Введение

История измерений параметров дыхания [1] показывает актуальность продолжения этих исследований с целью организации оперативных измерений и диагностики на основе современных компьютерных технологий.

Легкие – воздушный насос (вентилятор) низкого давления. Движение воздуха происходит в результате перепада давления между внутренним объемом легких и атмосферой. В настоящее время получены новые данные о влиянии состояния сердечно-сосудистой системы на дыхание, появились технические возможности его оценки и автоматической обработки получаемых данных с датчиков параметров дыхания. В клинической медицине стандартным способом оценки вентиляции легких является спирометрия. Измеряется объемная скорость воздушного потока и изменения объема [1]. Датчики параметров воздушных потоков могут иметь разные физические принципы детектирования [1, 2]:

- 1. Манометрический метод прямое измерение давления воздушного потока при вдохе и выдохе [3].
- 2. В электромагнитных вихревых расходомерах жидкость, движущаяся в постоянном магнитном поле, создает ЭДС, частота которой прямо пропорциональна частоте вихреобразования.
- Ультразвуковой метод вихри усиливаются и достигают своего развития ниже по потоку, где происходит их детектирование. Из анализа амплитудно-модулированного ультразвукового сигнала определяется величина объемного расхода [4].
- 4. Емкостной датчик регистрирует изменение емкости за счет деформации чувствительного элемента.
- 5. Метод изгибных напряжений пьезосенсор регистрирует совокупность тепловых и механических воздействий от вихревых потоков.
- 6. Термальный регистрируется динамика изменения температуры с изменением во времени энтальпии воздушного потока, в результате регистрируются вихревые колебания воздушного потока.

Большинство приборов спирометрии построены с использованием датчиков типов 1–5 и применяются как средства измерения при клинической аттестации, а также для проведения кратковременных медицинских проб. Однако эти приборы не пригодны для длительного мониторинга состояния пациента в реальной жизнедеятельности, что необходимо для оперативного контроля состояния, лечения и повышения надежности измерения. Портативность оборудования измерения и анализа параметров дыхания на основе компьютерных средств управления позволяет проводить оперативные измерения в реальном времени и принимать решения на основе быстрой диагностики.

В настоящее время интерес представляют термальные методы, для которых хорошо изучены принципы возникновения вихревых потоков в дыхании и их влияние на измеряемые параметры, связанные с теплообменом. Качество систем автоматической спирометрии определяется комплексом средств и методов, в том числе – качеством датчиков параметров вихревых потоков, алгоритмами анализа сигналов и диагностики, производительностью и ресурсами памяти компьютеров. Термальные методы измерений поддерживаются теорией теплофизических расчетов, одна из ведущих школ в этой области представлена монографией [5].

Целью исследований, представленных в работе, является обоснование методов расчета параметров дыхания на основе измерений датчиками, регистрирующими энергетику теплообмена в реальном времени.

Турбулентность вихревых потоков в легких

С учетом геометрической формы легких показано, что потоки в верхних путях легких являются турбулентными (вихревыми) [1]. Свойства турбулентности для выявления и измерения параметров потоков исследуются как прямыми клиническими измерениями, так и оперативными, косвенными, в частности, акустическими приборами.

Акустические явления в легких широко используются для врачебного обследования пациента (методы аускультации). Аускультация легких производится в определенных точках на поверхности грудной клетки. При дыхании выслушивают везикулярные и бронхиальные шумы. Везикулярные шумы выслушивают над легочной тканью, а бронхиальные – над трахеей и крупными бронхами. В классическом описании нормальные и патологические шумы определяют диаграммами (спектограммы).

Установлено [2], что нормальные звуки генерируются турбулентным потоком в воздушных путях, громкость (энергия колебаний) пропорциональна скорости потока. Аускультация отображает не только процесс генерации звука, но и процессы резонанса и поглощения между воздушными путями и датчиком звуков, что используется для диагностики по изменению спектра сигнала. Регистрация акустических явлений непосредственно в воздушном потоке передает информацию о спектре и энергии процесса генерации звука. Спектр определяется ритмом работы сердца, а энергия – теплообменом воздушного потока.

Вихри образуются на границах тела обтекания (канала на вдохе и выдохе) [2]. При этом частота образования вихрей прямо пропорциональна скорости потока:

 $f = \operatorname{Sh}(v/d),$

(1)

где *f* – частота образования вихрей Кармана; Sh – число Струхаля; *v* – скорость потока среды; *d* – ширина тела обтекания.

Начало турбулентного режима течения определяется критерием Рейнольдса

$$\operatorname{Re} = \frac{ud}{v},$$
(2)

где *и* – скорость среды, м/с; v – коэффициент кинематической вязкости; *d* – диаметр канала.

При Re > 10000 устанавливается турбулентный режим, для ламинарного режима Re < 2300, в промежутке имеет место переходный режим.

Параметры турбулентных потоков, определяющие энергообменные процессы при дыхании или процессы массообмена, являются необходимым фактором жизнедеятельности и важнейшими показателями изменений состояния сердечно-сосудистой системы. При этом оцениваются отклонения параметров процессов массо- и теплообмена от среднестатистических значений, принимаемых в качестве нормы в клинических исследованиях.

Например, принимая диаметр трахеи 1,5 см, объемный расход $W = 5 \text{ л} = 5 \cdot 10^{-3} \text{ м}^3$ (измеряется датчиком потока) и кинематическую вязкость $v = 16 \cdot 10^{-6} \text{ м}^2/\text{с}$, получим значение критерия Рейнольдса Re = 26500. В этом случае имеет место турбулентный поток, и при значении критерия Струхаля Sh = 0,2 при скорости потока u = 28,3 м/с, измеряемой датчиком, частота колебаний вихревых потоков определяется из соотношений (1), (2):

$$f = \operatorname{Sh} \frac{u}{d} = 380$$
 Гц.

Доступные внешнему мониторингу интегральные массо-энергетические параметры воздушных потоков – температура воздушного потока на вдохе T_0 и выдохе T_w ; объемный или массовый расход или скорость воздушного потока M. Выделяемая энергия в выдыхаемом потоке или мощность теплового потока P связаны между собой соотношениями [5]

$$P = cME(T_w - T_0); \quad E = 1 - \exp(-\varphi); \quad \varphi = \frac{\alpha S}{cM},$$
(3)

где *E* – коэффициент недорекуперации; φ – показатель недорекуперации; α – коэффициент теплоотдачи; *S* – площадь поверхности теплообмена.

Зависимости справедливы при условии постоянства температуры дыхательных путей. Из этого следует, что:

1. при больших значениях *S* в формуле (3) имеем $\phi > 4, E \rightarrow 1$ и

 $P = cM(T_w - T_0),$

происходит полная рекуперация, температура стенки трахеи при вдохе охлаждается до температуры внешнего воздуха;

2. при малых значениях *S* и ϕ в формуле (3) *E* $\rightarrow \phi$ и

$$E \approx \frac{\alpha S}{cM}, \ P = \alpha S \left(T_w - T_0 \right), \tag{5}$$

(4)

т.е. имеет место недорекуперация, и необходим расчет коэффициента теплоотдачи α, характеризующего интенсивность конвективного теплообмена.

Определение параметров воздушного потока и интенсивности теплообмена

Расчет коэффициента теплоотдачи проводится с использованием уравнений, содержащих безразмерные критерии подобия. Критерий Рейнольдса (2) определяет скорость потока жидкости или газа (в рассматриваемом случае – воздуха) и может быть выражен через объемный или массовый расход, плотность и (для канала круглого сечения) –диаметр d [5]:

$$\operatorname{Re} = \frac{ud}{v} = \frac{4M}{\pi d\gamma v} = \frac{4W}{\pi d\nu},$$
(6)

где
 $\gamma-$ плотность воздуха;
 W-объемный расход воздуха.

В зависимости от величины критерия Рейнольдса вычисляется значение критерия Нуссельта (Nu) [4], который определяет величину коэффициента теплоотдачи α в (5)

$$\alpha = \mathrm{Nu}\lambda_f/d_s$$

где λ – коэффициент теплопроводности воздуха. В частности, для ламинарного режима (Re< 2300) справедлива следующая формула вычисления критерия Нуссельта [6]

$$\overline{\mathrm{Nu}}_{f} = 0,15\sqrt[3]{\mathrm{Re}}_{f} \cdot \mathrm{Pr}_{f}^{0,43} \cdot \mathrm{Gr}_{f}^{0,1};$$
(7)

где Pr – критерий Прандтля; Gr – критерий Грасгофа.

Индекс *f* обозначает, что соответствующие коэффициенты определяются при температуре внешнего воздуха, а *w* – при температуре охлаждаемой или нагреваемой поверхности. Критерий Прандтля вычисляется по формуле

$$\Pr = \frac{v}{a} = \frac{v c \gamma}{\lambda}$$

Критерий Грасгофа определяется по формуле

$$\operatorname{Gr} = \beta g \frac{d^3}{v^2} (T_w - T_f),$$

где β – коэффициент объемного расширения воздуха; *g* – ускорение силы тяжести; *d* – внутренний диаметр канала.

(8)

(9)

Для переходного режима (Re = 2300-10000) из формулы (7) следует:

$$\overline{Nu}_f = k \cdot Pr_f^{0,43} = 0,958 k$$

а для турбулентного режима (Re больше 10000)

 $\overline{Nu_f} = 0,023 \cdot Re_f^{0,8} \cdot Pr_f^{0,43}$.

Длина теплообменника (трахеи), при которой реализуется условие полной рекуперации, определяется как

$$L \ge 4 \frac{CM}{\alpha \pi d}; M = \gamma W.$$

Таким образом, для расчета теплоотдачи Р:

- 1. вычисляется значение критерия Рейнольдса (6);
- определяется коэффициент теплоотдачи α с использованием критерия Нуссельта по одной из формул (8) или (9) в зависимости от величины критерия Рейнольдса;
- 3. вычисляется коэффициент конвективного массообмена β (м/с)

$$\beta = \frac{\alpha}{c\gamma};$$

4. вычисляется интенсивность конденсации или испарения

$$j = \frac{\beta}{RT} (P_s - P_w), \, \kappa \Gamma / (\mathbf{M}^2 \cdot \mathbf{c}),$$

где R = 461 Дж/кг; T – термодинамическая температура; P_s – парциальное давление пара над конденсирующей или испаряющей поверхностью; P_w – давление насыщенных паров, соответствующее температуре конденсирующей или испаряющей поверхности. В случае положительного значения j происходит конденсация с выделением тепловой энергии, а при отрицательном значении – испарение с поглощением энергии;

5. по коэффициенту теплоотдачи α можно определять величину поглощаемой или отдаваемой тепловой мощности.

Расчет энергии воздушного потока

Мощность воздушного потока P рассчитывается на основе тепловых измерений параметров потока датчиком и складывается из мощности, определяемой теплосодержанием Pt и кинетической энергией потока Pk:

P = Pt + Pk.

Мощность теплосодержания определяется энтальпией газа и вычисляется как

$$Pt = Mi; i = c \Delta T; \Delta T = T_{\Pi} - T_{c}; M = \gamma W;$$

где M – массовый расход; i – удельная энтальпия воздуха; c – удельная теплоемкость воздуха; ΔT – температурный напор потока; T_n – температура потока; T_c – температура окружающей среды; γ – плотность воздуха; W – объем расхода. В свою очередь, мощность теплосодержания вычисляется как $Pk = MV^2/2$; V = W/S, (11)

где *V* – скорость потока; *S* – проходное сечение в канале потока. Таким образом, можно получить выражение для общей мощности в виде

 $P = M(i + W^2/S).$

Отсюда следует, что теплосодержание потока растет пропорционально объемному расходу, а кинетическая энергия пропорциональна третьей степени объемного расхода.

Например, подставив в формулы (10) и (11) справочные значения ($c = 10^3$ Дж/кг·К; $\gamma = 1,16$ кг/м³; $\Delta T = 16$ К; $W = 5 \cdot 10^{-3}$ м³; $S = 1,8 \cdot 10^{-4}$ м² и V = 28,3 м/с), получим:

Pt = 93 BT; Pk = 2,32 BT; Pt/Pk = 40.

Таким образом, энергия в выдыхаемом воздушном потоке в 40 раз больше кинетической энергии потока. Тогда амплитуда пульсаций, связанная с воздействием на датчик импульсов отдельных вихрей, будет мала по сравнению с общей амплитудой сигнала. Для выделения этих пульсаций необходимо использовать схемотехнические и алгоритмические средства определения параметров вихревых компонентов сигнала.

Заключение

В работе предложена методика расчета энергии теплообмена в потоке воздуха с использованием параметров дыхания (скорость потока, температура на входе и выходе, объем расхода), измеряемых датчиком воздушного потока. В статье [7] приводятся характеристики разработанного для измерений датчика параметров воздушного потока при вдохе и выдохе и результаты исследования параметров датчика на реальных объектах.

Литература

- 1. Джорж Уэст. Физиология дыхания. Основы. М.: Мир. 1988. 200 с.
- 2. Стерлягов А.А. Роль биосистемы легких в теплообмене организма // Математическая морфология: электронный математический и медико-биологический журнал. 1998. Т. 3 № 1. С. 158–165.
- ГОСТ 8.563.1-97. Измерение расхода и количества жидкости и газов методом переменного перепада давления. – Введ. 01.10.1999. – Минск: Изд-во стандартов, 1997. – 64 с.
- ГОСТ 30457-97. Акустика. Определение уровней звуковой мощности источников шума на основе интенсивности звука. Измерение в дискретных точках. Технический метод. – Введ. 01.01.1999. – М.: Изд-во стандартов, 1997. – 14 с.
- 5. Дульнев Г.Н., Парфенов В.Г., Сигалов А.В. Методы расчета теплового режима приоров. М.: Радио и связь, 1990. 312 с.
- 6. Habibullo I. Abdussamatov, Sergey I. Khankov, Yevgeniy V. Lapovok. The Thermal Inertia Characteristics of the System Ocean-Atmosphere // Journal of Geographic Information System. 2012. V. 4. № 5. Р. 479–482 [Электронный ресурс]. Режим доступа: http://www.SciRP.org/journal/jgis, свободн.
- 7. Кормилицын А.Ю., Ханков С.И., Скорубский В.И. Измерение параметров дыхания датчиком воздушных потоков // Научно-технический вестник информационных технологий, механики и оптики. 2013. № 3. С. 122–129.

Ханков Сергей Иванович Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, вед. научный сотрудник, Leva0007@rambler.ru ООО «Инкарт», технический директор, Alex_kkk@incart.ru Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат технических наук, доцент; Государственный политехничекский университет, доцент; vskorubski@yandex.ru

(10)

УДК 681.02 ИЗМЕРЕНИЕ ПАРАМЕТРОВ ДЫХАНИЯ ДАТЧИКОМ ВОЗДУШНЫХ ПОТОКОВ

А.Ю. Кормилицын, С.И. Ханков, В.И. Скорубский

Рассматриваются задачи анализа и исследования параметров нового датчика воздушных потоков. Приведены результаты экспериментальных исследований датчика и сравнения его параметров с существующими в клинической практике датчиками. Сравнение выполняется по ряду параметров (частотное распределение спектральной плотности сигнала, реакция датчика на акустические воздействия, линейность и инерционность датчика; реакция датчика на энергию потока с регулируемой температурой; влияние жесткости конструкции).

Ключевые слова: спирометрические измерения, амплитудно-частотные характеристики, линейность и инерционность, чувствительность.

Введение

Спирометрические (акустические) измерения широко применяются для оценки вентиляции легких. В частности, измеряется объемная скорость воздушного потока и изменение объема [1]. Недостатками существующих спирометрических датчиков являются:

- низкая чувствительность акустических датчиков;
- узкий диапазон амплитудно-частотной характеристики (АЧХ) измеряемого акустического сигнала;
- конструктивные ограничения при оперативном применении.

Сконструирован датчик характеристик вихревых потоков дыхания с акустическим принципом действия и косвенным воздействием акустической волны. В датчике используется воздействие вихревых потоков на эффекты, свойственные пленке поляризованного поливинилиденфторида (ПВДФ) [2]. Пленка обладает пьезоэффектом в широком частотном диапазоне от 0,01 Гц до 100 кГц и формирует электрический потенциал при разных видах энергетического воздействия. Используется также пироэлектрический эффект.

Электрический потенциал формируется при изменении условий теплообмена (например, при изменении тепловых потоков, воздействующих на поверхность конвективными турбулентными течениями или внешним тепловым излучением). В технических характеристиках пленки [2] приводятся коэффициенты механо-электрического и температурно-электрического преобразования.

Датчик не реагирует на постоянную составляющую тепловых потоков, а регистрирует только приращение параметров, когда динамика их изменения соответствует частоте рабочего диапазона частот датчика. Учитывая это обстоятельство, можно путем конструктивных изменений датчика изменить или ограничить его частотный диапазон, выбрать чувствительность датчика в заданных диапазонах частот или создать условия для демпфирования сигнала. Это позволяет управлять формой сигнала, сглаживая определенные частоты колебаний и настраивая датчик на рабочий режим. Способом демпфирования может быть выбрано изменение массы чувствительного элемента с одновременной частичной изоляцией его поверхности от воздействующих внешних потоков. Такая частичная изоляция за счет нанесения на рабочую поверхность пленки определенной толщины создает дополнительное тепловое сопротивление между рабочей поверхностью датчика и источниками внешних потоков. В результате увеличивается общая теплоемкость датчика, что ведет к изменению его термоинерционных свойств. Процессом можно управлять, выбирая толщину пленки.

В режиме регистрации механических колебаний важно иметь возможность регулировать частотный диапазон регистрируемых внешних механических воздействий. При мониторинге вихревых колебаний можно управлять характеристиками принимаемого сигнала, сглаживая и усредняя механические пульсации, создаваемые воздействием вихревых потоков. Регулировка может быть выполнена при выборе подложки с определенной массой и размещением на рабочей поверхности датчика слоя пленки. При этом изменится собственная частота колебания пластины и реакция на внешние механические колебания без потери энергии.

Таким образом, можно регулировать конфигурацию регистрируемого сигнала, приводя его к виду, отражающему наиболее важные закономерности периодически изменяющихся параметров дыхания.

Датчик воздушных потоков позволяет измерить мгновенную скорость воздушного потока, объем воздуха при вдохе и выдохе в заданном интервале, изменение температуры воздуха в интервалах вдохавыдоха, а также индицировать ЭКГ при вдохе, выдохе и задержке дыхания.

Акустические и тепловые пьезоэффекты, используемые в датчике, преобразуются в электрические сигналы, для которых могут быть использованы средства цифровой обработки сигналов для получения значений необходимых параметров дыхания:

- спектрограмма импульса с разрешением до уровня регистрации пульсаций вихревых компонент воздушного потока;
- распределение интенсивности частот колебаний воздушных потоков дыхания;
- суммарная мощность теплообмена воздушного потока;

- кинетическая энергия воздушных потоков.
- На реальных объектах выполнялись клинические исследования с применением сконструированного датчика. Исследовались следующие параметры:
- 1. частотное распределение спектральной плотности сигнала;
- 2. линейность и инерционность датчика;
- 3. реакция датчика на акустические воздействия;
- 4. реакция датчика на энергию потока с регулируемой температурой;
- 5. влияние жесткости конструкции датчика при использовании изолирующих материалов на АЧХ.

По каждой характеристике проведены эксперименты и выделены сигналы, демонстрирующие измеряемые параметры. При косвенных измерениях выполнены различные методы численных преобразований для получения необходимых прямых значений.

Целью выполненных исследований являлось создание новых методов и средств измерения параметров воздушных потоков дыхания, необходимых для развития функциональной диагностики в медицине, повышения надежности соответствующих алгоритмических методов диагностики и лечения, в частности, важнейших из них, связанных с сердечно-сосудистой системой.

Частотное распределение спектральной плотности сигнала

На временной диаграмме (рис. 1) приведены формы сигналов теплового потока W(t), получаемых с датчика. Сигналы на рис. 1 (температура потока – рис. 1, а, скорость потока – рис. 1, б) отражают особенности дыхания человека, трудно поддающиеся механическому моделированию. На экспериментальном датчике потоков получены в фазе выдоха высокочастотные колебания (рис. 1, в), частота и амплитуда которых изменяется в соответствии с изменением скорости потока.

Для выделения вихревых компонентов сигнала вычисляется распределение спектральной плотности сигнала. На рис. 2 приведена зависимость спектральной плотности от частоты для осциллограммы, представленной на рис. 1.

Из сигнала, записанного в режиме свободного дыхания, выделены интервалы выдоха. При этом от начальной фазы к конечной скорость потока уменьшалась, что в соответствии с расчетами в [3, 4] должно приводить к уменьшению частоты образования вихрей.



температура (а); скорость потока (б); высокочастотные колебания (в)



Рис. 2. Спектральная плотность сигнала

При скорости потока менее 2 м/с режим течения потока изменяется от турбулентного к переходному, и образование вихрей прекращается. Спектральная характеристика с частотой выше 259 Гц размыта, поэтому нельзя выделить единственную доминирующую частоту. Предполагаем, что информативным является интервал девиации частот, характерных для спокойного дыхания, с турбулентным режимом конвективного потока.

Частота вихрей изменяется пропорционально скорости потока, которая не является постоянной величиной на всем периоде выдоха. При изменении аэродинамического сопротивления дыхательного канала скорость потока изменяется. Новой скорости будет соответствовать своя характеристика частотного распределения спектральной плотности принимаемого сигнала. С увеличением аэродинамического сопротивления спектр будет смещаться в сторону меньших частот. Это смещение можно использовать как индикатор возникновения преграды воздушному потоку.

Характеристики линейности и инерционности чувствительного элемента датчика

Для калибровки чувствительности использовался твердотельный лазер Nd:YAG, обладающий высокой стабильностью энергии излучения и длительности импульса. Параметры использованного лазера: длина волны излучения лазера 1064 нм, диаметр луча с размерами светового пятна 6×6 мм, длительность импульса по половине амплитуды 3 нс. Характер изменения энергии в импульсе после ослабления излучения фильтрами по результатам измерения специализированным оптическим измерителем мощности иллюстрируется рис. 3. В качестве чувствительного элемента использовалась пленка ПВДФ.







Рис. 4. Реакция чувствительного элемента (в, г) на поглощенное излучение лазера при энергии воздействия 3,5 мДж (а, в) и 5,5 мДж (б, г), обусловленное выравниванием температуры поверхности

Измерения проводились с помощью осциллографа Tektronix с полосой 200 МГц. Параметры входа 1 МОм и 20 пФ. Осциллограф Tektronix имеет функцию записи файлов, что позволяет сохранить значительно больше точек сигнала, чем представлено на экране, что необходимо для дальнейшего анализа методами цифровой обработки. На рис. 3 представлены отклики чувствительного элемента на поглощаемую энергию лазерного излучения. Графики демонстрируют линейность датчика (пропорциональность энергии воздействия 3,5 мДж и 5,5 мДж) и небольшую инерционность (задержка на 3 порядка меньше длительности вдоха-выдоха) относительно теплового воздействия. На рис. 4 в более узком диапазоне (4 мс) приведена реакция на фотоэффект пленки ПВФД. Графики также подтверждают линейность и невысокую инерционность датчика (уровень сигнала устанавливается за 3 мс при частоте дыхания 0,3 Гц).

Для синхронизации осциллографических измерений в эксперименте использовался кремниевый фотодиод. Поскольку амплитуда сигнала от лазерного импульса значительно превосходит амплитуду сигнала от излучения лампы накачки, для момента синхронизации был выбран передний фронт сигнала от лазерного импульса. В более продолжительном интервале (через 10 мкс после воздействия лазером) наступает процесс полной тепловой релаксации, сопровождающийся незначительными по амплитуде собственными колебаниями, которые могут быть устранены конструктивно.

Сравнение сигналов датчиков

Спирометрический метод исследования дыхания направлен на выявление амплитудных и временных характеристик сигнала в процессе измерения объемного расхода или скорости потока при выполнении определенных дыхательных тестов. Критерии соответствия выбираются по наиболее информативным характеристикам. В данном случае используется критерий совпадения формы исследуемых сигналов с выбранным эталоном. Эталонный сигнал считывается метрологическим измерителем объемного расхода.

В эксперименте на грудной клетке были размещены лента плетизмографа и канал ЭКГ с регистрацией биполярной реографии [5]. Со всех датчиков, в том числе и с датчика воздушного потока, проводилась синхронная запись сигналов. Для сравнения сигналы разной физической природы представлены в нормированном виде на рис. 5.

Визуально можно выделить наиболее близкие по форме сигналы плетизмограммы (рис. 5, кривая 3), реограммы (рис. 5, кривая 4) и датчика вихревых потоков (рис. 5, кривая 5). При этом следует учитывать следующие особенности измерения:

- 1. эталонный сигнал может представлять собой набор значений в некотором конечном коридоре (диапазоне) или набор реализаций;
- 2. данные, поступающие с датчиков, частично зашумлены;

- 3. анализ может проводиться как по одной выборке периода дыхания, так и по множеству выборок. При оценке по нескольким периодам сигнала процесс классифицируется как нестационарный;
- 4. выбирается пороговый уровень сигналов с датчика, который может иметь доверительный интервал значений.



Рис. 5. Фрагмент исходных нормированных сигналов с косвенных датчиков дыхания: 1 – сигналы ЭКГ; 2 – изменение объема воздуха в цикле дыхания; 3 – сигнал плетизмограммы; 4 – сигнал реограммы; 5 – сигнал с датчика вихревых потоков; 6 – сигнал с датчика давления (с амплитудой от 0 до 1); 7 – сигнал с датчика температуры (с амплитудой от –4,4 до –5,4 мкВ)

Чувствительность датчика определяется как коэффициент преобразования единицы физической величины в электрический сигнал (напряжение) и является характеристикой, привязанной к физическому принципу, на котором построен датчик, однако единица измерения может быть не связана с измеряемой косвенно величиной. Чтобы обеспечить сравнение сигнала с датчика с эталонным сигналом датчика объема, необходимо выполнить предварительную обработку сигнала [5]. В программу исследования достоверности косвенных методов входят следующие параметры:

- амплитудные искажения относительного объемного расхода фаз вдоха и выдоха;
- амплитудные искажения скорости потока;
- ошибка определения времени привязки к началу фаз вдоха и выдоха, сопоставляемых с показаниями тахометра;
- чувствительность датчика по отношению к относительному объемному расходу;

– АЧХ.

Исследуемый интервал делится на участки, соответствующие псевдоциклическому процессу вдоха-выдоха. Производится накопление данных и построение гистограммы распределения из N циклов вдох-выдох. Количество циклов определяется равенством площадей по уровню 0,5. Таким образом, требуется не более 100 циклов вдохов-выдохов.

Алгоритм расчета сравнительных характеристик применяется отдельно для оценки искажения форм объемного расхода на вдохе и выдохе, а также профиля скорости потока во времени. Динамические характеристики процесса определяет АЧХ, которая строится на основании анализа технической документации или на основании измерений по стандартной методике. Данные, полученные в результате исследований характеристик линейности и инерционности, можно распространить на частотные характеристики акустической чувствительности датчика, так как внешняя энергия всегда вызывает деформацию (непосредственно механическую или температурную) и пъезоэффект.

Датчик можно аттестовать как микрофон с нелинейностью АЧХ в пределах от +10% до -30% в диапазоне 20 Гц -1 кГц. Линейность характеристики в диапазоне 0–100 Па не хуже 7%. В акустических измерениях за опорную величину уровня звукового давления (SPL), принимают давление 10^{-5} Па, формирующее звук, соответствующий порогу слухового ощущения. Тогда динамический диапазон датчика составляет не менее 110 дБ [6].

Акустические сигналы и скорости воздушного потока проявляются в изменении давления. Воздушные потоки образуются за счет скорости потока и сопротивления дыхательных путей, акустические сигналы формируются при образовании продольных волн в результате наличия элементов, способных создавать механические колебания. При этом акустические компоненты сигнала значительно превосходят по энергии вихревые потоки, а также перекрывают их в частотной области. Появление значительных акустических сигналов приводит к невозможности регистрации вихревых потоков.

Исследование датчика в потоках с регулируемой температурой

В процессе дыхания человека в каждый момент времени присутствуют компоненты суммарной энергии потока, обусловленные его теплосодержанием и его кинетической энергией. Энергии разделимы в модельном эксперименте с использованием установки имитации дыхания с регулируемой температурой потока, которая регистрировалась датчиком температуры, а энергия и скорость потока – датчиком воздушного давления. На рис. 6 представлены осциллограммы на выходах датчиков при температуре воздушного потока из установки 20°С относительно окружающей среды, а на рис. 7 – аналогичные осциллограммы при температуре 0°С.

Регистрируемая датчиком энергия потока, нагретого установкой имитации дыхания, более чем на порядок выше потока без теплосодержания. При этом датчик воздушного потока демонстрирует высокую чувствительность к механическим колебаниям по сравнению с приемником воздушного давления.



Рис. 6. Осциллограммы при температуре теплового потока 20°С: изменение температуры (а); скорость потока воздуха (б); вихревой поток (в)



Рис. 7. Осциллограммы при температуре теплового потока 0°С: изменение температуры (а); скорость потока воздуха (б); вихревой поток (в)

Заключение

Реализация спирометрического метода измерения параметров в предлагаемом датчике потоков вдоха и выдоха позволяет выполнить измерение комплекса параметров сигналов и более надежную диагностику состояния пациента. Сравнение с известными и распространенными приборами измерения параметров (таблица) подтверждает эффективность предлагаемого датчика и методики измерений. Дальнейшие работы в этой области связаны с совершенствованием алгоритмической базы медицинской диагностики и совершенствованием датчика до уровня серийных образцов, доступных для мониторинга.

	Типы датчиков				
Параметры	Датчик вих- ревых пото- ков	Плетизмограф	Реограф	Датчик воз- душного давления	Датчик тем- пературы
Вид измеряемого сигна-	Энергия	Реактивное со-	Реактивное со-	Давление	Температура
ла	потока	противление	противление		
Диапазон сигнала	0-200 Вт	0-20 Ом	0-2 кОм	±300 Па	0–40°C
Диапазон параметра при	2-40 Вт	0,05–0,2 Ом	0,2–4 Ом	±50 Па	20–37°C
вдохе-выдохе					
Чувствительность к	0,16	125	0,2	2,43	1,1
параметру	B/BT	мВ/Ом	мВ/Ом	мкВ/Па	мВ/°С
Коэффициент искаже- ния формы объемного расхода на вдохе	0,0750	0,0201	0,0481	_	-
Коэффициент искаже- ния формы объемного расхода на выдохе	0,0813	0,0214	0,0505	_	_
Коэффициент искаже- ния формы скорости потока	0,0971	_	_	_	_
Ошибка привязки по времени фазы вдоха	150 мс	40 мс	150 мс	10 мс	_
Ошибка привязки по времени фазы выдоха	200 мс	40 мс	150 мс	10 мс	250 мс
АЧХ	0-100 кГц	0-5 Гц	0-5 Гц	0-3 кГц	-
Чувствительность объ- емного расхода	7 B	25 мВ	0,8 мВ	0,1 мВ	_
Акустическая чувстви- тельность	18 мВ/Па	_	_	2,43 мкВ/Па	_

Таблица. Сравнение характеристик датчиков дыхания на основе различных физических эффектов с учетом результатов, полученных в работе.

Литература

- 1. Смирнов И.В., Старшов А.М. Функциональная диагностика. ЭКГ, реография, спирография. Серия: Профессиональная медицина. М.: Эксмо, 2008. 224 с.
- 2. Техническое описание производителя. «Measurement Specialties Ltd» [Электронный ресурс]. Режим доступа: http://www.meas-spec.com/product/t_product.aspx?id=2488#, свободный. Яз. англ. (Дата обращения 02.03.2013).
- 3. Стерлягов А.А. Роль биосистемы легких в теплообмене организма // Математическая морфология: электронный математический и медико-биологический журнал. 1998. Т. 3 № 1. С. 158–165.
- 4. Трифонов Е.В. Спирограмма объем потока время. Пневмапсихосоматология человека. Русскоангло-русская энциклопедия. – 15-е изд. – 2012. – 104 с.
- 5. Трифонов Е.В. Объем смеси газов дыхательных путей. Пневмапсихосоматология человека. Русскоангло-русская энциклопедия. – 15-е изд. – 2012. – 346 с.

Кормилицын Александр Юрьевич	_	ООО «Инкарт», технический директор, Alex_kkk@incart.ru
Ханков Сергей Иванович	-	Санкт-Петербургский национальный исследовательский университет
		информационных технологий, механики и оптики, доктор технических
		наук, вед. научный сотрудник, Leva0007@rambler.ru
Скорубский Владимир Иванович	-	Санкт-Петербургский национальный исследовательский университет
		информационных технологий, механики и оптики, кандидат техниче-
		ских наук, доцент; Государственный политехничекский университет,
		доцент; vskorubski@yandex.ru

МАТЕРИАЛОВЕДЕНИЕ И НАНОТЕХНОЛОГИИ

УДК 535.399

ИДЕНТИФИКАЦИЯ ЦЕНТРОВ ЛЮМИНЕСЦЕНЦИИ В СТЕКЛАХ С ИОНАМИ МЕДИ И ХЛОРА

А.Н. Бабкина, Н.В. Никоноров, А.И. Сидоров, Т.А. Шахвердов, П.С. Ширшнев

Представлены результаты исследования спектров люминесценции и спектров возбуждения люминесценции калиевоалюмоборатных стекол, содержащих ионы меди и хлора. Показано, что до термообработки люминесценция стекол связана с молекулярными кластерами Cu_n (n < 10). Широкополосная люминесценция в видимой области спектра, возникающая после термообработки, вызвана появлением в матрице стекла молекулярных кластеров (CuCl)_n и (Cu₂O)_n.

Ключевые слова: люминесценция, молекулярный кластер, стекло, хлорид меди, ионы меди, ионы хлора.

Введение

Стекла с нанокристаллами CuCl обладают ярко выраженными нелинейно-оптическими свойствами [1, 2], в ряде случаев – фотохромизмом [3], и являются удобными объектами для исследования оптических и физических свойств нанокристаллов CuCl. Нанокристаллы CuCl обладают экситонной люминесценцией, которая наблюдается в ультрафиолетовой области спектра при криогенных температурах [4–6]. В то же время, как показали наши предварительные эксперименты, в стеклах, содержащих ионы меди и хлора, после термообработки возникает интенсивная широкополосная люминесценция в видимой области спектра при комнатной температуре. Причем эта люминесценция сохраняется и после продолжительной термообработки стекол, приводящей к формированию в них нанокристаллов CuCl. В связи с этим возникает вопрос о природе этих люминесцентных центров. Предварительные эксперименты показали, что люминесценция возникает как в силикатных стеклах, так и в калиево-алюмоборатных (КАБ) стеклах, содержащих ионы меди и хлора. Однако в силикатных стеклах с нанокристаллами CuCl наблюдается фотохромизм [3], проявляющийся в обратимом увеличении поглощения при облучении коротковолновым излучением. Это затрудняет проведение люминесцентных измерений. КАБ стекла с нанокристаллами CuCl не являются фотохромными [7], поэтому они были выбраны в качестве объекта исследований.

Целью настоящей работы было определение типов центров люминесценции в КАБ стеклах с ионами меди как до термообработки, так и после термообработок при разных режимах, в том числе после термообработки, приводящей к формированию в стекле нанокристаллов CuCl.

Методика экспериментов

КАБ стекла были синтезированы в НИУ ИТМО и имели следующий состав: $K_2O-Al_2O_3-B_2O_3-Cu_2O$ -NaCl. NaCl при синтезе создает мягкие восстановительные условия, что способствует сохранению ионов меди в одновалентном состоянии. Синтез проводился при температуре 1300°C. Образцы представляли собой полированные бесцветные стеклянные пластины толщиной 5 мм. Для измерения спектров поглощения были изготовлены пластины толщиной 200 мкм. Отсутствие у образцов окраски указывает на отсутствие или малое количество в них ионов двухвалентной меди Cu²⁺, дающих зеленую окраску. Термообработка образцов проводилась в муфельных печах (Nabertherm) с программным управлением при температуре 300°C и 380°C в течение 20 ч.

Измерения оптической плотности фото-термо-рефрактивных стекол проводились на спектрофотометре Cary500 (Varian) в спектральном интервале 200–1000 нм с шагом 1 нм. Для измерения спектров люминесценции и спектров возбуждения люминесценции использовался спектрофлуориметр MPF-44A (PerkinElmer). Измерения спектров люминесценции и поглощения проводились при комнатной температуре.

Обсуждение результатов

На рис. 1 показаны спектры люминесценции образцов до термообработки и после термообработок в указанных выше режимах. Из рисунка видно, что до термообработки люминесценция сосредоточена, в основном, в синей области спектра (кривая 1 на рис. 1). Авторы работ [8–10] связывают люминесценцию в этой области спектра с нейтральными молекулярными кластерами меди Cu_n (n < 10). Наличие структуры у этой полосы люминесценции свидетельствует о вкладе в люминесценцию молекулярных кластеров с разным количеством атомов, входящих в состав этих кластеров. Длинноволновый хвост полосы люминесценции может быть связан с люминесценцией ионов Cu⁺ и димеров Cu⁺–Cu⁺[11]. Термообработка при $t = 300^{\circ}$ С приводит к исчезновению полосы люминесценции в спектральном интервале 440–480 нм и появлению интенсивной широкополосной люминесценции в спектральном интервале 480–680 нм с максимумом на длине волны 580 нм (кривая 2 на рис. 1). При термообработке при $t = 380^{\circ}$ С в стекле происходит формирование нанокристаллов CuCl, на что указывает длинноволновое смещение полосы поглощения КАБ стекла и появление на спектре поглощения экситонной полосы поглощения нанокристаллов CuCl (см. вставку в рис. 1). Однако появление в стекле нанокристаллов CuCl слабо влияет на спектр его люминесценции (кривая 3 на рис. 1). Это указывает на то, что нанокристаллы CuCl при комнатной температуре не вносят вклад в данную полосу люминесценции.





На рис. 2 показаны спектры возбуждения люминесценции для различных длин волн люминесценции КАБ стекла после термообработки при $t = 380^{\circ}$ С в течение 20 ч. Из рисунка видно, что в КАБ стеклах с медью присутствуют две ярко выраженные полосы возбуждения люминесценции: одна – с максимумом на длине волны 280 нм, вторая – с максимумом на длине волны 370 нм. Первая полоса возбуждения люминесценции соответствует люминесценции, в основном, в спектральном интервале 400–500 нм, а вторая полоса возбуждения люминесценции — в спектральном интервале 540–650 нм. Наличие двух полос возбуждения люминесценции позволяет предположить появление после термообработки двух типов центров люминесценции в КАБ стеклах, содержащих медь.

Как было отмечено выше, в исходном стекле присутствуют молекулярные кластеры меди Cu_n. При термообработке при $t = 300^{\circ}$ C в этих кластерах могут происходить реакции с ионами хлора Cl⁻ и кислорода O⁻. В результате этого в стекле будет происходить трансформация молекулярных кластеров меди Cu_n в молекулярные кластеры (CuCl)_n и (Cu₂O)_n. Для проверки этого предположения были измерены спектры люминесценции силикатного стекла, содержащего медь, но не содержащего хлор. Ионы меди вводились в стекло методом ионного обмена [12] при $t = 500^{\circ}$ C. После ионного обмена стекло не приобретало зеленую окраску. Это указывает на то, что после ионного обмена медь находится в стекле в одновалентном состоянии.

На рис. 3 показаны спектры возбуждения и люминесценции силикатного стекла, содержащего ионы одновалентной меди. Из рисунка видно, что в данном случае наблюдается одна полоса возбуждения люминесценции с максимумом на $\lambda = 385$ нм и одна полоса люминесценции с максимумом на $\lambda = 570$ нм. Как было показано в работе [13], подобными характеристиками люминесценции в стекле обладает Cu₂O. Сравнение рис. 3 и рис. 2 позволяет сопоставить длинноволновую ($\lambda = 385$ нм) полосу возбуждения люминесценции в КАБ стекле с полосой возбуждения люминесценции Cu₂O в силикатном стекле. В этом случае полоса люминесценции в спектральном интервале 550–650 нм в КАБ стекле может быть связана только с молекулярными кластерами (Cu₂O)_n. Небольшое отличие в спектральном положении максимумов возбуждения и люминесценции для стекол двух типов может быть вызвано как различием в размерах молекулярных кластеров (Cu₂O)_n, так и различием в их окружении. В этом случае за коротковолновую ($\lambda = 385$ нм) полосу возбуждения люминесценции и полосу люминесценции в спектральном интервале 450–500 нм в КАБ стекле могут отвечать молекулярные кластеры (CuCl)_n.



Рис. 2. Спектры возбуждения люминесценции КАБ стекла после термообработки при *t* = 380°С в течение 20 ч. Цифры у кривых – длина волны люминесценции в нанометрах



Рис. 3. Спектр возбуждения при длине волны люминесценции 570 нм (а); спектр люминесценции при длине волны возбуждения 385 нм (б) силикатного стекла, содержащего ионы одновалентной меди

При термообработке стекла при $t = 380^{\circ}$ С термическая диффузия ионов меди и хлора приводит к увеличению размера молекулярных кластеров (CuCl)_n. При достижении ими критического размера, когда начинают проявляться кристаллические свойства, их люминесценция в видимой области спектра исчезает. Однако при этом в стекле происходит параллельный процесс – формирование новых молекулярных кластеров (CuCl)_n, которые после термообработки обеспечивают люминесценцию стекла в спектральном интервале 400–500 нм.

Заключение

Представленные результаты показывают, что калиево-алюмоборатные стекла, содержащие медь, обладают яркой люминесценцией в видимой области спектра. В исходном стекле, не прошедшем термообработку, люминесценция в спектральном интервале 440–480 нм вызвана присутствием в стекле молекулярных кластеров меди Cu_n. Термообработка при $t \ge 300^{\circ}$ C приводит к реакциям этих молекулярных кластеров с ионами хлора и кислорода и их трансформации в кластеры вида (CuCl)_n и (Cu₂O)_n, обеспечивающие люминесценцию в спектральных интервалах 450–500 нм и 550–600 нм.

Работа выполнена при поддержке гранта РФФИ «Мой первый грант» на 2012–2013 годы (мол_а 12-02-31896) и ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009–2013 годы (Соглашение № 14.В37.21.0169 Минобрнауки РФ).

Литература

- 1. Никоноров Н.В., Сидоров А.И., Цехомский В.А. Низкопороговый нелинейно-оптический отклик фотохромных стекол с нанокристаллами хлорида меди // Оптический журнал. – 2008. – Т. 75. – № 12. – С. 61–65.
- 2. Ким А.А., Никоноров Н.В., Сидоров А.И., Цехомский В.А., Ширшнев П.С. Нелинейно-оптические эффекты в стеклах с нанокристаллами хлорида меди // Письма в ЖТФ. 2011. Т. 37. № 9. С. 22–28.
- Dotsenko A.V., Glebov L.B., Tsekomskii V.A. Physics and chemistry of photochromic glasses. N.W.: CRC Press, 1998. – 190 p.
- 4. Edamatsu K., Oohata G., Shimizu R., Itoh T. Generation of ultraviolet entangled photons in a semiconductor // Nature. - 2004. - V. 431. - P. 167-170.
- 5. Oda M., Shen M.Y., Saito M., Goto T. Photobrightening of CuBr nanocrystals in PMMA // J. of Lumin. 2000. V. 87–89. P. 469–471.
- 6. Park S., Jeen G., Kim H., Kim I., Jang K. Thermal properties of CuCl quantum dots embedded in a aluminoborosilicate glass matrix // J. of the Korean Phys. Soc. – 2000. – V. 37. – № 3. – P. 309–312.
- 7. Никоноров Н.В., Цехомский В.А., Ширшнев П.С. // Научно-технический вестник СПбГУ ИТМО. 2011. № 5 (75). С. 23–26.
- 8. Vazquez-Vazquez C., Banobre-Lopez M., Mitra A., Lopez-Quintela M.A., Rivas J. Synthesis of Small Atomic Copper Clusters in Microemulsions // Langmuir. 2009. V. 25. № 14. P. 8208–8216.
- Wei W., Lu Y., Chen W., Chen S. Wentao Wei, Yizhong Lu, Wei Chen and Shaowei Chen. One-Pot Synthesis, Photoluminescence, and Electrocatalytic Properties of Subnanometer-Sized Copper Clusters // J. Am. Chem. Soc. 2011. V. 133. P. 2060–2063.
- 10. Vilar-Vidal N., Blanco M.C., Lopez-Quintela M.A., Rivas J., Serra C. Electrochemical synthesis of very stable photoluminescent copper clusters // J. Phys. Chem. C. 2010. V. 114. P. 15924–15930.
- Srikumar T., Kityk I.V., SrinivasaRao Ch., Gandhi Y., Piasecki M., Bragiel P., Ravi Kumar V., Veeraiah N. Photostimulated optical effects and some related features of CuO mixed Li 20–Nb 2O 5–ZrO 2–SiO 2 glass ceramics // Ceramics Intern. – 2011. – V. 37. – P. 2763–2779.
- Tervonen A., West B.R., Honkanen S. Ion-exchanged glass waveguide technology: a review // Opt. Engineering. 2011. V. 50. P. 071107.
- Gurin V.S., Alexeenko A.A., Kaparikha A.V. Fabrication and optical features of the sol-gel derived silica glasses doped with copper oxide nanoparticles and europium // Materials Lett. – 2011. – V. 65. – P. 2442– 2444.

Бабкина Анастасия Николаевна	-	Санкт-Петербургский национальный исследовательский университет
		информационных технологий, механики и оптики, студент,
		babkinauha@ya.ru
Никоноров Николай Валентинович	-	Санкт-Петербургский национальный исследовательский университет
		информационных технологий, механики и оптики, доктор физмат.
		наук, профессор, зав. кафедрой, Nikonorov@oi.ifmo.ru
Сидоров Александр Иванович	_	Санкт-Петербургский национальный исследовательский университет
		информационных технологий, механики и оптики, доктор физмат.
		наук, профессор, ст. научный сотрудник, aisidorov@qip.ru
Шахвердов Теймур Азимович	_	Санкт-Петербургский национальный исследовательский университет
1 01		информационных технологий, механики и оптики, кандидат физмат.
		наук, вед. инженер, teimur@list.ru
Ширшнев Павел Сергеевич		Санкт-Петербургский национальный исследовательский университет
1 1		информационных технологий, механики и оптики, студент,
		pavel.shirshnev@gmail.com

УДК 621.373.535 ИЗМЕНЕНИЕ ПОГЛОЩАТЕЛЬНОЙ СПОСОБНОСТИ КРЕМНИЯ ПРИ ФЕМТОСЕКУНДНОМ ЛАЗЕРНОМ ВОЗДЕЙСТВИИ И.В. Гук, Г.Д. Шандыбина, Е.Б. Яковлев

Представлены результаты теоретического моделирования совместного влияния нелинейностей поглощательной способности и коэффициента поглощения на пространственно-временное распределение электронно-дырочной плазмы и на динамику удельного поглощенного потока в кремнии в течение действия фемтосекундного лазерного импульса. Показано, что основной вклад в изменение удельного поглощенного потока вносит увеличение поглощения горячим электронным газом. Полученные результаты сопоставлены с известными представлениями о поляритонном механизме, которые используются для интерпретации фемтосекундного лазерного микроструктурирования кремния. Показана необходимость учета динамики поглощательной способности при оценках режимов ультракороткой лазерной обработки полупроводников.

Ключевые слова: поглощательная способность, скин-эффект, электронно-дырочная плазма, фемтосекундное микроструктурирование поверхности.

Введение

В последнее время не ослабевает практический интерес к явлениям, протекающим на поверхности полупроводников под действием ультракоротких импульсов. В частности, модификация оптических, электрических и структурных свойств поверхности кремния может быть использована при разработке светоизлучающих устройств, в том числе лазеров, для получения фотонных структур, эффективных элементов солнечных батарей, лазерной спектроскопии [1–4] и т.д.

Фундаментальные исследования особенностей взаимодействия ультракоротких лазерных импульсов с различными средами показали, что во время фемтосекундного лазерного импульса существенно изменяются оптические свойства твердого тела без изменения химического состава и морфологии поверхности. Все эти изменения происходят при нагревании решетки, которое начинается после окончания импульса, и становятся заметными после воздействия большого количества импульсов лазерного излучения. Трудности экспериментального исследования динамики процессов, протекающих в течение воздействующего ультракороткого импульса, способствовали усилению роли математического моделирования при их исследовании. В большинстве моделей, описывающих процесс интенсивного фотовозбуждения и нагрева полупроводников и диэлектриков, основное внимание уделяется изменению в процессе облучения дифференциальной оптической характеристики – коэффициента поглощения, а изменение интегральной оптической характеристики – поглощательной способности (ПС) – обычно не учитывается [5]. Возникает необходимость детального анализа механизмов и динамики изменения ПС для определения вклада этих изменений в общую картину процессов фемтосекундного воздействия лазерного излучения на полупроводники.

В работе исследуется совместное влияние нелинейностей ПС и коэффициента поглощения на процесс фемтосекундного двухфотонного возбуждения кремния с учетом внешних фото-, термоэффектов для уточнения пространственно-временного распределения электронно-дырочной плазмы и удельной плотности поглощенного потока. Численное моделирование проводится для пластин монокристаллического кремния толщиной 300 мкм, облучаемых импульсами длительностью 80 фс, генерируемых хромфорстеритовой лазерной системой ($\lambda = 1,25$ мкм) при плотности энергии линейно поляризованного излучения 0,5–2 Дж/см².

Известно, что результат фемтосекундного структурирования поверхности кремния зависит от характера распределения оптических свойств полупроводника по глубине [6]. Это позволяет судить о процессах, развивающихся в фотовозбужденном полупроводнике, по типу сформировавшихся в результате лазерного воздействия поверхностных структур. Полученные результаты анализируются в соответствии с экспериментальными данными «поляритонного» фемтосекундного микроструктурирования поверхности кремния [7].

Приближение слабоаномального высокочастотного скин-эффекта

Теория скин-эффекта успешно описывает ПС металлов. Согласно ее выводам, на величину ПС металлов влияет степень локальности отклика электронов проводимости в поле электромагнитной волны, частотная зависимость A_{ω} , зависимость от температуры, обусловленная увеличением частоты электронфононных столкновений $A_T \sim T$, и рассеяние электронов на поверхности A_s [8]:

 $A = A_{\omega} + A_T + A_s.$

(1)

При переходе в область воздействия ультракоротких лазерных импульсов изменяется сам механизм ПС: времена электрон-фононных столкновений становятся больше длительности импульса, и на ПС начинает влиять рост температуры электронного газа, а поверхностное поглощение реализуется в условиях сильных эмиссионных потоков электронов с поверхности. Используя выражения для плазменной частоты электронов (ω_p), частоты электрон-электронных столкновений, скорости (v_F) и энергии Ферми (ε_F)

$$\omega_p = \sqrt{\frac{4\pi n q_e^2}{m_e}}, \ \gamma_e = \frac{9\omega_p^4}{64\pi^2 n v_F^3} \left[\left(\frac{\hbar\omega}{2\pi\varepsilon_F}\right)^2 + \left(\frac{k_B T_e}{\varepsilon_F}\right)^2 \right], \ v_F^2 = 2\varepsilon_F / m_e, \ \varepsilon_F = \frac{\hbar^2}{2m_e} (3\pi^2 n)^{2/3},$$

где m_e – эффективная масса, примерно равная массе свободного электрона; q_e – заряд электрона; γ_e – частота столкновений электронов, определяемая частотой электрон-электронных столкновений; $\omega = 1,5 \cdot 10^{15} \text{ c}^{-1}$ – частота лазерного излучения; k_B – постоянная Больцмана; n – концентрация электронов, в приближении слабоаномального высокочастотного скин-эффекта можно получить следующее выражение для ПС

$$A = b \left(\frac{\hbar\omega}{2\pi}\right)^2 + b \left(k_B T_e\right)^2 + p \frac{3v_F}{4c_0},$$

rge $b = \frac{q_e^3 m_e^{7/2}}{\hbar^7 \pi^{19/6} 3^{1/3} n^{11/6}}.$ (2)

В выражении (2) в соответствии с выражением (1) первые два слагаемых определяют объемное поглощение (частотную и температурную компоненты ПС), а третье слагаемое характеризует дополнительное поглощение, обусловленное диффузным отражением электронов от поверхности в глубину материала: p – феноменологический параметр, $0 ; <math>c_0$ – скорость света.

В полупроводниках в результате многочисленных предыдущих исследований сверхбыстрой электронной динамики установлено образование плотной (более 10^{21} см⁻³) электронно-дырочной плазмы, поэтому для качественного анализа механизма и динамики ПС кремния можно применить классическую теорию скин-эффекта.

Результаты теоретического моделирования

Для численных расчетов авторами используется модель двухфотонного фотовозбуждения кремния, в которой уравнения теплопроводности двухтемпературной модели дополнены уравнением концентрации неравновесных носителей n(z,t), граничными условиями для потоков двухфотонной фотоэмиссии и термоэмиссии и выражением дифференциального закона Бугера для плотности поглощенного потока фотонов J(z,t) с куполообразной формой временного распределения интенсивности лазерного излучения (ось z направлена в глубину полупроводника) [9]. В процессе фотовозбуждения с ростом концентрации электронно-дырочной плазмы изменяется коэффициент поглощения $\alpha(z,t)$, представляющий собой сумму коэффициента двухфотонного поглощения внутреннего фотоэффекта, коэффициента двухфотонного поглощения внешнего фотоэффекта и коэффициента поглощения на свободных электронах.

В первом приближении предполагается, что ПС полупроводника A является величиной постоянной. Согласно формулам Френеля, для $\lambda = 1,25$ мкм при комнатной температуре и толщине пластины кремния 300 мкм A = 0,2. Численная модель позволяет проследить динамику электронно-дырочной плазмы при куполообразной форме лазерного импульса и определить распределение концентрации возбужденных носителей в кремнии в глубину. Уже в первые фемтосекунды лазерного импульса происходит металлизация полупроводника; максимум концентрации возбужденных носителей ввиду эмиссионных явлений на поверхности заглубляется – в приповерхностном слое образуется оптически неоднородная структура.

Вводя в дифференциальный закон Бугера ПС, зависящую от концентрации неравновесных носителей согласно (2) и интегрируя результат по *z*, можно проследить вклад каждого из слагаемых в ПС. Результаты численного расчета представлены на рис. 1, где на фоне профиля лазерного импульса показана динамика каждого из компонент ПС. Видно, что поверхностная компонента дает слабый вклад в общую ПС, но может влиять на поглощение в начале импульса, частотная компонента также изменяется слабо, и только температурная компонента ПС стремительно возрастает. Действительно, по мере металлизации кремния растут плазменная частота, частота электрон-электронных столкновений и скорость Ферми электронов, что приводит к некоторому уменьшению частотной компоненты и росту поверхностной составляющей ПС. Интенсивный рост температурной компоненты ПС обусловлен ее квадратичной зависимостью от температуры электронного газа, которая в течение лазерного импульса может достигать значений, превышающих 10⁴ К.

Проведенный качественный анализ показал, что в фотовозбужденном полупроводнике ПС может сильно изменяться благодаря температурной составляющей.

С физической точки зрения определяющей характеристикой фемтосекундного нагрева полупроводника является удельный поглощенный поток $\alpha(z,t)J(z,t)$. На рис. 2 представлена динамика удельного поглощенного потока фотонов, проинтегрированного по глубине относительно падающего потока,

$$Q(t) = \frac{\int_0^\infty \alpha(z,t) J(z,t) dz}{J_0(t)}$$

Рост Q в начале импульса обусловлен ростом коэффициента поглощения α и соответственно уменьшением объема, в котором поглощается поток. Далее значение α стабилизируется и при постоянной ПС Q не изменяется. Учет изменяющейся ПС приводит к стремительному росту удельного поглощенного потока во второй половине импульса.



Рис. 1. Изменение поверхностной (точечная линия), частотной (штрих-пунктирная линия), температурной (сплошная линия) компонент ПС кремния во время действия лазерного импульса. Пунктирной линией показана временная форма лазерного импульса; плотность энергии 1 Дж/см²



Рис. 2. Динамика удельного поглощенного потока фотонов (в относительных единицах) при постоянной ПС *A* = 0,2 (пунктирная линия) и при изменяющейся ПС (сплошная линия)



Рис. 3. Динамика концентрации носителей в фотовозбужденном кремнии (в максимуме по глубине) (а) и динамика движения координаты максимума $z_{\rm max}$ концентрации (б) при изменяющейся ПС (сплошная линия), пунктирная линия для A = 0,2

Учет зависимости ПС от температуры электронного газа существенно влияет на динамику распределения концентрации электронно-дырочной плазмы. На рис. 3, а, видно, что концентрация неравновесных носителей в течение всего лазерного импульса (более $5 \cdot 10^{21}$ см⁻³) превышает значение, при котором знак диэлектрической проницаемости в кремнии становится отрицательным, и полупроводник остается в металлизированном состоянии практически весь импульс. В глубине координата максимума концентрации неравновесных носителей с учетом изменяющейся ПС сдвигается ближе к поверхности (рис. 3, б).

Сравнение с экспериментом по фемтосекундному микроструктурированию кремния

Рассмотрим полученные результаты в соответствии с экспериментальными данными «поляритонного» фемтосекундного микроструктурирования поверхности кремния [7]. При облучении полированных пластин монокристаллического кремния лазерными импульсами длительностью 80 фс ($\lambda = 1,25$ мкм) при плотности энергии линейно поляризованного излучения 0,5–1 Дж/см² наблюдалось формирование поверхностных периодических структур, ориентированных перпендикулярно вектору поляризации лазерного излучения и связанных с возбуждением и интерференцией поверхностных поляритонов с падающей электромагнитной волной. Известно, что возбуждение плазмон-поляритонов возможно, если действительная часть диэлектрической проницаемости активной среды отрицательна, что характерно для металлов. При воздействии фемтосекундного излучения переход в металлоподобное состояние поверхности кремния обусловлен интенсивным фотовозбуждением. Высокая концентрация неравновесных носителей может изменить знак диэлектрической проницаемости, создавая условия для возбуждения поверхностных плазмон-поляритонов (ППП).

Согласно результатам расчета по численной модели, учитывающей зависимость ПС полупроводника от температуры электронного газа, уровень высокой концентрации электронно-дырочной плазмы, обеспечивающий металлизацию полупроводника вблизи поверхности, поддерживается в течение всего импульса, а максимум значения концентрации неравновесных носителей приближается к поверхности. Такое распределение обеспечивает условия для возбуждения ППП, лучшие по сравнению с численными результатами, полученными при расчете с постоянной ПС.



Рис. 4. Электронное изображение поверхности кремния, облученного лазерными фемтосекундными импульсами. Стрелкой показано направление вектора поляризации лазерного излучения

Фемтосекундное микроструктурирование кремния, имеющее «поляритонную» природу, наблюдают многие исследователи (см., например, [10–11]). На рис. 4 показано типичное электронное изображение поверхности кремния после фемтосекундного лазерного облучения 1200 импульсами ($\lambda = 1,25$ мкм, $\tau = 80$ фс). Ясно видны микроструктуры, они ориентированы перпендикулярно вектору поляризации излучения. Период структур – около 1,2 мкм (сравним с длиной волны лазера). Ориентация структур следует за вектором поляризации, и их вращение не зависит от ориентации кристаллографических осей образца. Формирование рельефа зависит от времени экспозиции. Глубина рельефа увеличивается с числом лазерных импульсов. В облученной области обнаружено присутствие окислов кремния. Подобные структуры типичны для поляритонного механизма микроструктурирования.

Также следует обратить внимание на результат, имеющий большое значение для ультракороткой лазерной обработки полупроводников: удельный поглощенный поток растет по мере действия лазерного импульса не только за счет уменьшения глубины, на которой происходит поглощение излучения [5], но и за счет роста температурной компоненты поглощательной способности.

Заключение

Из результатов проведенного анализа следует, что основной механизм изменения поглощательной способности при фемтосекундных воздействиях на полупроводники обусловлен квадратичной зависимостью поглощательной способности от температуры электронного газа. Дополнение существующей модели двухфотонного фотовозбуждения кремния с учетом внешней эмиссии динамикой поглощательной способности позволило получить пространственно-временное распределение концентрации неравновесных носителей, обеспечивающее более благоприятные условия для возбуждения поверхностных плазмон-поляритонов, что соответствует типу микроструктур, наблюдаемых экспериментально. Изменение поглощательной способности в процессе фемтосекундного фотовозбуждения кремния оказывает существенное влияние на динамику удельного поглощенного потока, т.е. его нужно учитывать при оценках режимов ультракороткой лазерной обработки полупроводников.

Авторы благодарны С.В. Заботнову за предоставленную фотографию микроструктурированной поверхности кремния, а также Г.А. Марциновскому за полезные обсуждения.

Работа выполнена при поддержке грантов РФФИ 12-02-01194-а, 13-02-00033-а, 13-02-00971-а и ведущей научной школы НШ-619.2012.2.

Литература

- Gattass R., Mazur. E. Femtosecond laser micromachining in transparent materials // Nat. Phot. 2008. V. 2. – P. 219–225.
- Juodkazis S., Mizeikis V., Misawa1 H. Three-dimensional microfabrication of materials by femtosecond lasers for photonics applications // J. App. Phys. – 2009. – V. 106. – 051101 (14 pages).
- Bellini N., Vishnubhatla K.C., Bragheri F., Ferrara L., Minzioni P., Ramponi R., Cristiani I., Osellame R. Femtosecond laser fabricated monolithic chip for optical trapping and stretching of single cells // Opt. Exp. – 2010. – V. 18. – P. 4679–4688.
- Sibbett W., Lagatsky A.A., Brown C.T.A. The development and application of femtosecond laser systems // Opt. Exp. – 2012. – V. 20. –№ 7. – P. 6989–7001.
- 5. Gamaly E.G. The physics of ultra-short laser interaction with solids at non-relativistic intensities // Phys. Rep. 2011. V. 508. P. 91-243.
- Марциновский Г.А., Смирнов Д.С., Шандыбина Г.Д., Заботнов С.В., Головань Л.А., Тимошенко В.Ю., Кашкаров П.К. Ультракороткие возбуждения поверхностных поляритонов и волноводных мод в полупроводниках // Оптика и спектроскопия. – 2008. – Т. 105. – № 1. – С. 75–81.
- 7. Остапенко И.А., Заботнов С.В., Шандыбина Г.Д., Головань Л.А., Червяков А.В., Рябчиков Ю.В., Яковлев В.В., Тимошенко В.Ю., Кашкаров П.К. Микро- и наноструктурирование поверхности кристаллического кремния под действием фемтосекундных лазерных импульсов // Известия РАН. Сер. физическая. – 2006. – Т. 70. – № 9. – С. 1315–1317.
- 8. Либенсон М.Н. Лазерно-индуцированные оптические и термические процессы в конденсированных средах и их взаимное влияние. СПб: Наука, 2007. 423 с.
- Dyukin R.V., Martsinovskiy G.A., Sergaeva O.N., Shandybina G.D., Svirina V.V., Yakovlev E.B. Interaction of Femtosecond Laser Pulses with Solids: Electron/Phonon/Plasmon Dynamics // Laser Pulses – Theory, Technology and Applications / Ed. by Igor Peshko. – Rijeka: InTech, 2012. – P. 197–218.
- Golosov E.V., Ionin A.A., Kolobov Y.R., Kudryashov S.I., Ligachev A.E., Makarov S.V., Novoselov Y.N., Seleznev L.V., Sinitsyn D.V. Topological evolution of self-induced silicon nanogratings during prolonged femtosecond laser irradiation // Applied phisics A. Material science & processing. – 2011. – V. 104. – P. 701–705.
- 11. Han Y., Qu S. The ripples and nanoparticles on silicon irradiated by femtosecond laser // Chemical Physics Letters. 2010. V. 495. P. 241-244.

Гук Игорь Владимирович	-	Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, icorruk@ymail.com
Шандыбина Галина Дмитриевна	-	санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат физмат. наук, доцент, corchand@gmail.com
Яковлев Евгений Борисович	_	Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, профессор, yak@lastech.ifmo.ru

УДК 621.315.592 МЕХАНИЗМ МИКРОСТРУКТУРИРОВАНИЯ СИСТЕМЫ SiO₂/Si ПРИ ОБЛУЧЕНИИ СКАНИРУЮЩИМ ПУЧКОМ ИМПУЛЬСНОГО ВОЛОКОННОГО ЛАЗЕРА

А.М. Скворцов, Хуинь Конг Ту, Р.А. Халецкий

Рассматривается механизм микроструктурирования системы SiO₂/Si при облучении импульсным волоконным лазером ИЛИ-1-50 с применением режима сканирования лазерного пучка. Приведены результаты микроструктурирования при такой мощности излучения, когда происходит локальное плавление – кристаллизация кремния, но не разрушается пленка SiO₂. Показано, что при малых дозах облучения (облучение в сходящихся лучах лазера) плавление кремния локализуется на поверхности кремния на структурных дефектах, в частности, в местах выхода на поверхность дислокаций, генерируемых лазерным облучением. Показано, что с ростом дозы облучения вначале происходит увеличение площади областей расплава, затем – смыкание этих областей и образование протяженных узких «каналов» расплава. При этом пленка SiO₂ приобретает морфологию структурированной поверхности кремния.

Ключевые слова: лазерное облучение, плавление кремния, кристаллизация, структурные дефекты, дислокации, микроструктурирование, сканирование.

Введение

На основе кремния изготавливается до 95% всех видов полупроводниковых устройств, в том числе изделий интегральной микроэлектроники. В настоящее время традиционная микроэлектроника вплот-

ную приблизилась к своим физическим пределам, в частности, в отношении размеров активных элементов. Так, например, в микропроцессорах, выпускаемых ведущими компаниями-производителями микросхем по 32 нм-технологии, достигнуты следующие размеры областей затворной композиции МОПтранзисторов: длина затвора 32 нм, длина канала 25 нм, толщина подзатворного слоя диоксида кремния 1,2 нм. В связи с этим, наряду с сохранением тенденции дальнейшего снижения размеров элементов при традиционной технологии ультрабольших интегральных схем (УБИС), проводятся интенсивные исследования по созданию нанокомпозитных материалов и разработке принципиально новых методов конструирования изделий наноэлектроники [1–3]. К таким новым нанокомпозитам можно отнести тонкие пленки диоксида кремния со встроенными в них нанокристаллами кремния [4].

С начала нового тысячелетия расширился круг исследований, направленных на создание устройств оптоэлектроники на основе кремния (кремниевой фотоники). Это обусловлено открытием в 1990 году Л. Кэнхемом интенсивной фотолюминесценции в пористом кремнии [5]. Оказалось, что могут быть получены излучатели фотонов в широком спектральном диапазоне в зависимости от размеров кристаллических наночастиц кремния в слое пористого кремния. Эти слои, по существу, представляют собой структуры из окисленных нанокристаллов и нанокристаллических нитей кремния. Позднее была показана возможность люминесценции в других структурах, содержащих нанокристаллы кремния в пленках диоксида кремния, выращенных на кремниевых подложках, имеющих высокую плотность нанокристаллов как для создания устройств вычислительной техники, так и для элементов кремниевой фотоники. Авторами уже было показано [9], что для формирования таких нанокомпозитов может быть применено лазерное микроструктурирование термически окисленных кремниевых пластин (системы SiO₂/Si).

В настоящей работе приведены результаты исследования процессов микроструктурирования системы SiO₂/Si путем сканирования поверхности системы излучением иттербиевого волоконного импульсного лазера при дозах облучения от начала микроплавления кремния (граничная доза) до мощностей, при которых плавление кремния не сопровождается разрушением пленки диоксида кремния.

Методика эксперимента

В качестве экспериментальных образцов использовались пластины монокристаллического кремния КДБ-10, ориентированные в кристаллографической плоскости (111). На пластинах методом термического окисления во влажном кислороде был выращен слой SiO₂ толщиной 150 нм.

Для облучения системы SiO₂/Si использовался импульсный иттербиевый волоконный лазер ИЛИ-1-50 с длиной волны $\lambda = 1062$ нм и случайной поляризацией. Номинальная выходная мощность лазера 50 Вт, длительность импульсов 120 нс, номинальная энергия в импульсе при максимальной выходной энергии 1,0 мДж. Облучение проводилось в сходящихся лучах [10] с частотой следования импульсов 50 кГц, что обеспечивает необходимую зону облучения и равномерное распределение мощности по площади облучения по сравнению со сфокусированным пучком. Диаметр пятна облучения подложки в нашем случае составлял около 120 мкм.

В процессе облучения лазерный луч перемещался с переменной скоростью по прямой линии при помощи двух зеркал гальванометрического сканатора. Начальная и конечная скорости – 100 мм/с, в средней точке луч останавливался на 1 с (рис. 1). Длина полосы облучения составила 10 мм, ширина – около 120 мкм. Такой режим обеспечивал плавное изменение дозы облучения по линии сканирования, что позволило отследить характер микроструктурирования системы SiO₂/Si в зависимости от дозы облучения. Контроль морфологии облученных образцов проводился на оптическом микроскопе Axio Imager A1m фирмы Carl Zeiss, укомплектованном цифровой видеокамерой высокого разрешения.

Доза облучения рассчитывалась по формуле

$$D = P_{\text{имп}} \cdot t \cdot f \cdot \tau / S ,$$

где $P_{имп}$ – мощность одиночного импульса, Вт; t – время воздействия, с; f – частота повторения импульсов, Гц; τ – длительность лазерного импульса, с; S – площадь пятна лазерного облучения.

Обсуждение полученных результатов

Предварительный выбор режима облучения и сканирования должен был обеспечить переход из граничной области мощности, где отсутствует плавление кремния, к появлению точечного плавления кремния (отдельных субмикронных областей). Таким начальным режимом облучения был выбран ранее указанный режим облучения: энергия импульсов 1,0 мДж, начальная и конечная скорости сканирования 100 мм/с. На рис. 2 приведены микрофотографии участков линии сканирования на той части, где последовательно происходит рост дозы облучения. Следует заметить, что после перемещения лазерного луча по линии сканирования в результате остывания ранее облученной части подложки происходит рекристаллизованные области кремния и пленка SiO₂ сохраняют морфологию расплавленных областей.



Рис. 1. Экспериментальная установка (а), схема процесса сканирования (б), график изменения скорости лазерного луча по линии сканирования (в), доза облучения в пересчете на скорость движения луча (г)

Как следует из рис. 2, а, при дозе облучения $D = 2,3 \text{ Дж/см}^2$ на поверхности кремния появляются области плавления кремния субмикронных размеров, и только отдельные, наиболее крупные области достигают размеров в насколько десятых долей микрометра в диаметре. Причиной их появления могут являться структурные дефекты исходного материала подложки, а также дефекты, появляющиеся в результате воздействия лазерного излучения на поверхность кремния [11, 12].

Повышение дозы облучения приводит к увеличению размеров площадей расплавленных областей без появления новых центров микроплавления (рис. 2, б). Размеры расплавленных областей здесь вырастают до 0,2–1,0 мкм. Отдельные области имеют диаметр 2–2,4 мкм. Начиная с дозы, близкой к 3,14 Дж/см², наряду с ростом площадей расплавленного кремния появляются новые центры плавления (рис. 2, в). Общая площадь расплавленных областей на этом участке полосы сканирования составляет порядка 10%. Контуры границ рекристаллизованных областей начинают терять форму кругов. Дальнейший рост дозы облучения приводит к изменению формы кристаллизованных областей кремния, которая приближается к форме равносторонних треугольников (рис. 2, г), характерных для выхода плоскостей скольжения на поверхность кремниевой пластины, ориентированной в главной кристаллографической плоскости (111). Как известно [13], форму равносторонних треугольников на поверхности кремния приобретают ямки анизотропного химического травления в местах выхода дислокаций на поверхность, имеющую ориентацию (111). В нашем случае плавление кремния, подобно химическому травлению, возникало в локальных областях поверхность кремниевой подложки генерируемых лазером дислокаций.

С ростом дозы облучения возрастает температура области облучения, увеличивается число дислокаций, которые выходят на поверхность, увеличиваются площади расплавленных областей. Наконец, в результате дальнейшего роста дозы облучения ($D = 7,02 \text{ Дж/см}^2$) происходит слияние локальных областей расплава, и в средней части полосы сканирования образуются узкие «каналы», т.е. области расплавленного кремния длиной до 20 мкм и шириной 5–7 мкм (рис. 3). Видно также, что параллельно этим цен-

тральным «каналам», ближе к периферии полосы сканирования, формируются более короткие «каналы» длиной до 5–8 мкм. Как видно из микрофотографии, все эти «каналы» вытянуты в направлении движения сканирующего луча. К области с полосами расплава с обеих сторон примыкают участки с высокой плотностью локальных областей «анизотропного плавления» (областей треугольной формы), обусловленных действием механизма генерации дислокаций и их выходом на поверхность подложки. Ширина этих участков составляет 10–15 мкм, и к ним, в свою очередь, примыкают участки, в которых, вследствие малых плотностей мощности, только зарождаются области локального плавления.





Рис. 2. Микрофотографии участков линии сканирования, полученных при различных дозах облучения: $D = 2,3 \, \text{Дж/см}^2$ соответственно зоне 5 (X = -5 мм) на рис. 1, б, г (а); $D = 2,68 \, \text{Дж/см}^2$ соответственно зоне 4 (X = -4 мм) на рис. 1, б, г (б); $D = 3,14 \, \text{Дж/см}^2$ соответственно зоне 3 (X = -3 мм) на рис. 1, б, г (в); $D = 3,98 \, \text{Дж/см}^2$ соответственно зоне 2 (X = -2 мм) на рис. 1, б, г (г)

Во время остановки луча на одну секунду (X=0) доза в области облучения составила 2,3 · 10³ Дж/см² (рис. 1, г). При этом в облученной области произошло разрушение пленки SiO₂ и испарение поверхностного слоя кремния. После того, как луч лазера вновь начал движение с нарастающей скоростью, по ходу движения сформировалась такая же топология областей, что и в начале сканирования, но в обратной последовательности. Вначале формируется область с «каналами» (аналогичная рис. 3), затем последовательно области, аналогичные областям рис. 2, г, в, б, а.

Следует отметить, что пленка SiO₂ полностью повторяет конфигурацию поверхности рекристаллизованных областей кремниевой подложки. Это объясняется тем, что при температуре плавления кремния (1423°C) пленка SiO₂ становится пластичной (температура плавления пленки SiO₂ равна 1700°C).

Чтобы отследить развитие механизма микроструктурирования системы SiO₂/Si при дальнейшем увеличении дозы облучения, лазерный луч был сфокусирован до диаметра 80 мкм, мощность импульса была увеличена до максимальной (50 Вт). На рис. 4 представлены микрофотографии, полученные при сканировании сфокусированного лазерного луча при разной мощности лазерного излучения. Использовался следующий режим облучения подложки: энергия импульсов 1,0 мДж, перемещение лазерного луча – равномерное со скоростью 70 мм/с, шаг сканирования 250 мкм, диаметр сфокусированного пучка 80 мкм. Площадь сканирования составила 5×5 мм².


Рис. 3. Микрофотография участка линии сканирования при дозе облучения $D = 7,02 \text{ Дж/см}^2$

На микрофотографиях (рис. 4) показаны типичные участки двух областей сканирования, полученных на одной кремниевой подложке при различных плотностях мощности. Верхние и нижние полосы получены движением лазерного луча слева направо, а средние – справа налево. Следует обратить внимание на то, что вид полос, полученных при противоположных направлениях сканирования луча, различается. Однако с ростом плотности мощности эта разница становится менее выраженной. По-видимому, отмеченная разница связана с некоторыми погрешностями настройки системы сканирования.



Рис. 4. Микрофотографии участков областей сканирования лазерным пучком при дозах облучения *D* = 7,69 Дж/см² (а) и *D* = 8,35 Дж/см² (б)

Как видно из рис. 4, а, при дозе облучения $D = 7,69 \, \text{Дж/см}^2$ наблюдаются полосы плавления кремния шириной порядка 80 мкм по всей длине линии сканирования пучка, т.е. ширина полос фактически совпадает с диаметром пятна облучения на подложке. При дозе $D = 8,35 \, \text{Дж/см}^2$ (рис. 4, б) ширина полосы расплавленного кремния уже достигает 130 мкм и несколько превышает половину шага сканирования. Дальнейшее увеличение плотности мощности приводит к увеличению ширины полосы плавления кремния и высоты рельефа микроструктурированной поверхности полосы. Это связано с тем, что плавление кремния начинается в центре полосы, куда попадает максимальное число фотонов лазерного облучения, причем здесь наблюдается наибольшая глубина области плавления. Дальше от центра образуются «каналы» плавления, которые доходят до края облучения, смыкаются, приобретают направление к центру полосы и образуют сплошную полосу плавления кремния. Зарождение областей плавления здесь, повидимому, происходит по тому же механизму, что и при облучении сходящимся пучком. Плавление кремния начинается на поверхности кремния в местах выхода дислокаций; с ростом дозы облучения увеличиваются скорость генерации дислокаций и локальные площади расплавленного кремния, которые, соединяясь, образуют «каналы расплава», а те, в свою очередь, сливаясь, образуют сплошную полосу расплавленного кремния. Далее, вследствие высокой теплопроводности кремния и большой энергии фононов, получаемой от фотонов лазерного излучения, происходит передача энергии «фононным ветром» на смежные, необлученные области подложки. В результате происходит плавление кремния и в прилежащих к полосе областях, не подвергшихся прямому лазерному облучению. Отсюда – полученные экспериментальные факты: во-первых, ширина полосы расплавленного кремния больше диаметра сканирующего лазерного пятна; во-вторых, с увеличением дозы облучения растет разница между диаметром пятна облучения и шириной полосы расплавленного кремния.

Морфология поверхности системы SiO₂/Si в полосах плавления-кристаллизации кремния может быть объяснена следующим образом. Пленка SiO₂ в области плавления кремния (температура плавления кремния равна 1423°C) становится пластичной. Температура плавления SiO₂, как уже указывалось, равна 1700°C. Наибольшей пластичностью по сравнению с областями, прилегающими к твердой подложке, пленка обладает в центральной части полосы. Под действием давления потока фотонов лазерного излучения на поверхность системы SiO₂/Si пленка SiO₂ прогибается на разную величину, которая связана как с разной пластичностью, так и с неравномерным распределением плотности фотонов в пучке. В результате под давлением пленки расплав кремния оттесняется от середины, а движение прогиба вместе с движением лазерного луча перемещает расплав по направлению сканирования луча. После смещения луча и кристаллизации охлаждаемой области системы твердеющая пленка SiO₂ приобретает форму кристаллизованного кремния. После стравливания пленки SiO₂ в растворе плавиковой кислоты видно, что микротопология поверхности кристаллизованного кремния имеет ту же форму, что и с пленкой.

Вышеизложенный механизм действует до тех пор, пока не произойдет смыкание и частичное перекрытие полос сканирования. Дальнейшее увеличение плотности мощности облучения приводит к началу разрушения пленки SiO₂. Как правило, разрушение пленки совпадает с началом абляции кремния. Следует заметить, что еще до смыкания полос сканирования в облученных областях системы SiO₂/Si появляются линии скольжения, пересекающиеся под углом 60° , что свидетельствует о появлении пластической деформации во всей облученной области.

Заключение

Основным механизмом, приводящим к возникновению локальных микрообластей расплавленного кремния под воздействием лазерного облучения с большой частотой следования импульсов, является дислокационный механизм. Именно подпитка кристаллической решетки кремния энергией фотонов с частотой 50 кГц приводит к интенсивной генерации собственных точечных дефектов (вакансий и междоузельных атомов кремния). В результате происходит зарождение и рост дислокаций в кремнии вблизи границы SiO₂/Si [10]. В областях выхода дислокаций на поверхность и начинается локальное плавление кремния.

Увеличение дозы облучения приводит к вовлечению в процесс дефектообразования все большего количества атомов кристаллической решетки, росту числа дислокаций и числа центров плавления, увеличению площадей расплавленных областей. Начиная с дозы облучения 7,02 Дж/см², происходит соединение отдельных расплавленных областей в протяженные (образование «каналов»), и далее, с увеличением дозы облучения, формируется сплошная расплавленная полоса. Морфология поверхности микроструктурированной поверхности системы SiO₂/Si связана с особенностью взаимодействия расплава кремния, пластичной пленки SiO₂, границей твердой и жидкой фазы кремниевой подложки и упругим воздействием пучка фотонов на систему SiO₂/Si.

Авторы признательны сотрудникам кафедры лазерных технологий и экологического приборостроения (зав. кафедрой профессор В.П. Вейко) за обсуждение материалов работы и профессору Е.Б. Яковлеву за предоставление фотографии экспериментальной установки. Мы благодарим доцента Н.С. Кармановского за ценные советы по оформлению статьи

Литература

- 1. Tiwari S., Rana F., Chan K., Hanafi H., Wei C., Buchanan D. Volatile and nonvolatile memories in silicon with nano-crystal storage // IEEE Int. Electron Devices Meeting Tech. Dig. 1995. P. 521–524.
- Dimitrakis P. et al. Silicon nanocrystal memory devices obtained by ultra-low-energy ion-beam synthesis // Solid-State Electronics. – 2004. – V. 48. – P.1511–1517.
- Gonzalez-Varona O., Ga rrido B., Cheylan S., Perez-Rodriguez A., Cuadras A., Morante J.R. Control of tunnel oxide thickness in Si-nanocrystal array memories obtained by ion implantation and its impact in writing speed and volatility // Appl. Phys. Lett. – 2003. – V. 82. – № 13. – P. 2151–2153.
- 4. Скворцов А.М., Фам Куанг Тунг. Структура нанокластеров кремния в системе кремний диоксид кремния // Изв. вузов. Приборостроение. 2009. Т. 52. № 3. С. 69–73.
- Canham L.T. Silicon quantum wire array fabrication by electrochemical and chemical dissolution of wafers // Appl. Phys. Lett. – 1990. – V. 57. – № 10. – P. 1046–1048.
- Rinnert H., Vergnat M., Burneau A. Evidence of light-emitting amorphous silicon clusters confined in a silicon oxide matrix // J. Appl. Phys. – 2001. – V. 89. – № 1. – P. 237–243.

- Yang Yang, Ling Xu, Fei Yang, Wenqiang Liu, Jun Xu, Zhongyuan Ma, Kunji Chen. Enhanced visible photoluminescence from nc-Si/SiO_x films deposited by electron beam evaporation // Journal of Non-Crystalline Solids. 2010. V. 356. № 50–51. P. 2790–2793.
- 8. Костюк А.Б., Белов А.И., Жаворонков И.Ю. и др. Светоизлучающие ионно-синтезированные структуры на основе нанокристаллов кремния в оксидных матрицах // Вестник Нижегородского университета им. Н.И. Лобачевского. 2010. № 5 (2). С. 264–270.
- Skvortsov A.M., Veiko V.P., Sokolov V.I., Pham Qung Tung, Khaletsky R.A. Laser modification of thermal oxide films on silicon // International conference «Fundamentals of laser assisted micro- and nanotechnologies» (FLAMN-10): Abstracts. – St. Petersburg: Russia, 2010. – P. 103.
- 10. Скворцов А.М., Вейко В.П., Хуинь Конг Ту. Применение импульсного волоконного лазера для микроструктурирования системы SiO₂/Si // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – № 5 (81). – С. 128–133.
- Banishev A.F., Golubev V.S., Kolmnev Y.U. Generation and accumulation of dislocations on the silicon surface ander the fction of pulse-periodic emission from a YAG:Nd laser // Technical Phisics. – 2001. – V. 46. – № 8. – P. 962–967.
- 12. Банишев А.Ф., Новикова Л.В. Образование обратимых и необратимых структурных дефектов на поверхности кремния под действием лазерного импульса // Физика и химия обработки материалов. – 1992. – № 4. – С. 55–59.
- 13. Скворцов А.М., Жарова Ю.А., Ткалич В.Л. Микроструктурирование поверхности монокристаллов кремния в электронике // Изв. вузов. Приборостроение. 2006. Т. 49. № 1. С. 60–65.

Скворцов Альберт Матвеевич
 Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, профессор, a-skvortsov@yandex.ru
 Хуинь Конг Ту
 Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, Picochip912@yahoo.com
 Халецкий Роман Александрович
 Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, Picochip912@yahoo.com
 Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат технических наук, доцент, halecky@yandex.ru

УДК 539.232; 542.06; 546-1 ПОЛУЧЕНИЕ НАНОСТРУКТУРИРОВАННЫХ ПЛЕНОК Al₂O₃ МЕТОДОМ ЭЛЕКТРОХИМИЧЕСКОГО АНОДИРОВАНИЯ М.В. Жуков, В.В. Левичев

Методом сканирующей зондовой микроскопии исследованы поверхности оксидных пленок алюминия. Изучены электрические параметры процесса анодирования различных образцов алюминия для получения структурированного окисла. Произведено исследование топографии поверхностной структуры оксидных пленок на сканирующем зондовом микроскопе NTegra Aura.

Ключевые слова: оксид алюминия, массив структурированных пор, тонкие пленки, наноразмерные поры, анодирование алюминия.

Введение

В последнее время в области нанотехнологий и прикладной химии особый интерес представляют работы, направленные на создание наноразмерных массивов пористых структур на поверхности образцов алюминия для дальнейшего применения полученных образований в практических целях [1]. Полученный структурированный оксид, как правило, исследуется методами электронной (СЭМ) [2] и сканирующей зондовой (СЗМ) [3] микроскопии.

Известно, что структура слоя оксида алюминия представляет собой набор гексагональных ячеек (рис. 1) с размерами 50–500 нм. Оксидная пленка обладает защитными свойствами, такими, как высокая коррозионная стойкость, прочность и твердость [4]. На геометрические параметры ячеек при этом наибольшее влияние оказывают состав электролита, электрические режимы формирования, температура, время и шероховатость поверхности образца.

В процессе анодирования образуется два вида оксидных пленок. Первая пленка, находящаяся непосредственно на алюминии, носит название барьерной. Она возникает практически моментально на чистом алюминии при комнатной температуре и имеет аморфную структуру. Такая пленка растет в нейтральных растворах, в которых оксид алюминия трудно растворим (борная кислота, борат аммония и тетраборат аммония в этиленгликоле). Вторая пленка, образующаяся поверх барьерной, называется пористой и представляет как раз оксидный слой гексагональных ячеек (рис. 1). Она растет в кислых электролитах, в которых оксид может не только осаждаться, но и растворяться (разбавленная серная, фосфорная, хромовая, щавелевая кислоты и смеси неорганических и органических кислот) [5].

В электролитической ячейке алюминий служит анодом, а катодом может служить пластина из свинца, нержавеющей стали, никеля или алюминия. При замыкании цепи на электроде, подключенном к отрицательному полюсу источника (катоде), происходит реакция электрохимического восстановления, т.е. присоединение электронов к атомам и молекулам материала катода:

 $2Al = 2Al^{3+} + 6e^{-}$.

Взаимодействие электронов с ионами водорода приводит к образованию нейтральных молекул водорода. На аноде происходит реакция окисления за счет отдачи электронов в электролит [5]:

 $6H^+ + 6e^- = 3H_2$.

Суммарную реакцию при этом можно записать как [6]

 $2\mathrm{Al} + 3\mathrm{H}_2\mathrm{O} = \mathrm{Al}_2\mathrm{O}_3 + 3\mathrm{H}_2\,.$

Во многих работах для образования оксида алюминия применяют значительную длительность анодирования поверхности (до 4 часов) [7].



Рис. 1. Структура оксида алюминия

Цель настоящей работы – исследование возможности применения электрохимического анодирования для получения пористых пленок оксида алюминия в растворе серной кислоты и возможности использования получаемой структуры в качестве основы для создания наноструктурированных материалов.

Описание эксперимента

Процесс получения экспериментальных образцов состоял из нескольких этапов:

- пробоподготовка образца перед экспериментом;
- анодирование алюминия;
- очистка образца от остатков электролита и продуктов реакции;
- контроль получившегося окисла методом C3M.

На полученных C3M-изображениях светлые области скана соответствуют более высоким точкам (рис. 2, а), а темные – более низким относительно среднего уровня, равного половине максимального перепада высот исследуемой области. Разрешающая способность данного метода зависит от размеров кончика зонда (около 10 нм). В данной работе использовался C3M типа NTegra Aura производства компании NT-MDT.



Рис. 2. СЗМ-изображение рельефа поверхности алюминия до анодирования (а) и после анодирования (б)

Писки из	Образина химинески		Пленка алюминия,
диски из	чистого алюминия	Листовой алюминий	напыленная на
дюралюминия			слюдяную подложку
15,8	17,1	8,5	5,3

Характеристики исходных образцов для получения анодного окисла представлены в таблице.

Таблица. Значения R_a исходных образцов, нм

Пробоподготовка дюралюминиевых дисков и образцов чистого алюминия состояла в механической шлифовке абразивной пастой ГОИ № 2 для обеспечения ровной поверхности. Листовой алюминий и тонкие пленки механической обработке не подвергались, так как обладали зеркально гладкой поверхностью.

Для контроля шероховатости проведено C3M-исследование поверхности алюминиевых образцов. С помощью специальной программы обработки C3M-изображений Image Analysis проведен расчет средней арифметической шероховатости R_a [8], основанный на вычислении средних перепадов высот по поверхности заданного размера. По результатам измерений (таблица) видно, что наименьшей средней арифметической шероховатостью обладают образцы, представляющие собой напыленную на слюде пленку алюминия. По этой причине дальнейшие исследования проводились именно с этими образцами. Образцы представляли собой подложку слюды круглой формы диаметром около 10 мм. В качестве подслоя использовалась тонкая (10 нм) пленка титана. На подложку с подслоем напылялась пленка алюминия толщиной 500 нм.

В качестве электролита использовался 20% раствор серной кислоты при температуре около 0°С. Перед окислением образцы очищались ацетоном или изопропиловым спиртом. Контроль процесса окисления осуществлялся по величине тока.

При выборе величины тока при анодировании учитывался тот факт, что при низком уровне тока пленка оксида не успевает расти, так как процесс стравливания оксида протекает значительно быстрее процесса его роста, а при достаточно большом значении уровня тока происходит полное зарастание пор.

В результате предварительных экспериментов для данного образца была выбрана величина тока $I \sim 0.1$ А. Время анодирования составило около 15–20 с. На изображении (рис. 3, а) представлена поверхность полученного оксида алюминия с порами диаметром около 150 нм и глубиной 30–40 нм. При этом хорошо видно, что поры равномерно расположены по всей поверхности пленки алюминия.



Рис. 3. Увеличенное СЗМ-изображение рельефа поверхности анодированного алюминия (а) и U-образный профиль сечения одной поры (б)

Обсуждение результатов

Следует отметить, что особое влияние на получаемую структуру имеет исходная степень шероховатости вещества, так как на крупных дефектах поверхности происходит разнонаправленный рост окисла алюминия. В связи с этим наиболее приемлемыми для создания структурированных оксидных пленок оказались образцы с тонкой пленкой алюминия, напыленной на слюду. Диаметр полученных пор составил около 150 нм при глубине 30–40 нм.

Выбор оптимальных значений силы тока существенно зависел от интенсивности газообразования на поверхности электродов. При интенсивном газообразовании на поверхности электрода происходила неравномерная реакция окисления, что отрицательно сказывалось на качестве образующегося окисла. Для дальнейшего исследования формирования окисла на поверхности тонких пленок можно использо-

вать микроскоп токов ионной проводимости, использующих зонд в виде стеклянного микрокапилляра с проводящим сердечником [9].

Заключение

Изучение различных исходных образцов алюминия показало, что для задач наноструктурирования поверхности алюминия наиболее подходят свеженапыленные пленки алюминия (500 нм) на слюдяную подложку с тонким (10 нм) адгезионным подслоем титана (шероховатость поверхности образцов $R_a \sim 5$ нм). Показана принципиальная возможность получения методом электрохимического анодирования поверхности алюминия упорядоченной пористой структуры оксида алюминия, обладающей равномерно распределенными по поверхности наноразмерными порами размером около 150 нм и U-образным профилем глубиной 30–40 нм.

Литература

- 1. Напольский К.С. Синтез пространственно упорядоченных металл-оксидных нанокомпозитов на основе пористого Al₂O₃ / Под ред. А.Р. Кауля. М.: Московский государственный университет им. М.В. Ломоносова, 2008. 211 с.
- 2. Росляков И.В., Напольский К.С., Елисеев А.А., Лукашин А.В., Чернышов Д.Ю., Григорьев С.В. Синтез магнитных наночастиц с контролируемой анизотропией функциональных свойств в матрице из пористого оксида алюминия // Российские нанотехнологии. 2009. Т. 4. № 3–4. С. 69–72.
- 3. Белов А.Н., Гаврилов С.А., Шевяков В.И. Особенности получения наноструктурированного анодного оксида алюминия // Российские нанотехнологии. 2006. Т. 1. № 1–2. С. 223–227.
- 4. Аверьянов Е.Е. Справочник по анодированию. М.: Машиностроение, 1988. 224 с.
- 5. Андруцкая О.М. Процесс анодирования алюминия // Промышленная окраска. 2007. Т. 3. С. 2.
- Jinsub Choi. Fabrication of monodomain porous alumina using nanoimprint lithography and its application. Dissertation. – 2004. – 103 p.
- 7. Лукащук Т.С., Ларин В.И., Пшеничная С.В. Формирование наноструктурированных анодных оксидов алюминия в щавелевой кислоте // Вестник Харьковского национального университета. 2010. № 932. Вып. 19 (42). С. 112–118.
- ISO 4287. Geometrical Product Specifications (GPS) Surface texture: Profile method Terms, definitions and surface texture parameters. – 01.04.1997. – Geneva, International Organization for Standardization. – 35 p.
- 9. Стовпяга А.В., Французов Г.С. Изготовление и аттестация зондов из стеклянных микрокапилляров для сканирующего зондового микроскопа // Научно-технический вестник СПбГУ ИТМО. 2011. № 1 (71). С. 93–96.

 Жуков Михаил Валерьевич
 Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, студент, cloudjyk@yandex.ru, zerocomb@bk.ru

 Левичев Вадим Вячеславович
 Санкт-Петербургский национальный исследовательский университет инфор

Левичев Вадим Вячеславович – Санкт-Петероургскии национальный исследовательский университет информационных технологий, механики и оптики, кандидат физ.-мат. наук, доцент, levichev_vadim@mail.ru

8

ТЕХНОЛОГИЯ ПРИБОРОСТРОЕНИЯ. ОРГАНИЗАЦИЯ ПРОИЗВОДСТВА

УДК 621.81.004.17:620.191.355.001.5 ВЛИЯНИЕ ШЕРОХОВАТОСТИ ПОВЕРХНОСТИ НА УСИЛИЕ РАСПРЕССОВКИ ПРЕССОВЫХ СОЕДИНЕНИЙ А.Ю. Иванов, Д.Б. Леонов

Приводятся результаты экспериментального исследования влияния шероховатости поверхности на усилие распрессовки прессовых соединений с использованием непараметрических критериев оценки и контроля микрогеометрии поверхностей. Разработана методика, позволяющая определить лучшую из возможных в конкретных производственных условиях шероховатость поверхности, обеспечивающую необходимое усилие распрессовки соединений.

Ключевые слова: шероховатость поверхности, непараметрический метод, усилие распрессовки прессовых соединений.

Введение

Изделия приборостроения и машиностроения представляют совокупность взаимосвязанных между собой деталей. Отклонения размеров, формы и расположения поверхностей даже в одной из деталей вызывают отклонения формы или отклонения в расположении других деталей сборочной единицы. Эти отклонения в сумме оказывают определенное воздействие на эксплуатационные показатели изделий, что обусловливает необходимость оптимизации всех факторов, которые могут повлиять на их качество.

По способу сопряжения деталей в прессовых посадках различают соединения, получаемые:

- путем приложения осевой силы, обеспечивающей взаимное смещение сопрягаемых деталей в заданное положение – продольная запрессовка;
- за счет температурных деформаций одной из деталей (или обеих), вводимых без натяга одна в другую
 поперечная запрессовка [1].

В обоих случаях необходимая прочность соединения достигается созданием при сборке монтажного натяга, который определяет величину контактного давления, возникающего на сопрягаемых поверхностях. Эксплуатационные показатели прессовых соединений зависят от многих конструкторскотехнологических факторов: макро- и микрогеометрии сопрягаемых поверхностей, отклонения их формы, например, от цилиндричности, фактического натяга между сопрягаемыми деталями и т.д.

На сегодняшний день имеется ряд работ, связанных с определением и расчетом параметров прессовых соединений [1–4], но в них шероховатость поверхности рассматривается с точки зрения параметрических критериев. Известно, что при продольной запрессовке часть микронеровностей деформируется или срезается, снижая напряженно-деформированное состояние соединения. При создании соединений поперечной запрессовкой такого явления не происходит, поэтому логично предположить, что шероховатость сопрягаемых поверхностей в них существенно влияет на усилие распрессовки таких соединений. Целью работы является экспериментальная проверка этого предположения с использованием непараметрических критериев оценки и контроля микрогеометрии поверхностей (в качестве которых используются графические изображения функций плотности распределения ординат или тангенсов углов наклона профилей и т.п.) [5–7], позволяющих более информативно описать исследуемую поверхность.

Объект исследования

На рис. 1 изображено соединение, полученное поперечной запрессовкой за счет температурных деформаций одной из деталей.



Рис. 1. Соединение, полученное поперечной запрессовкой: номинальный диаметр сопряжения *D* = 20 мм; наружный диаметр втулки *d* = 30 мм; длина сопряжения *I* = 20 мм

Детали изготовлены из стали 40. Модули упругости материала втулки и вала $E_1 = E_2 = 206 \ \Gamma \Pi a$; предел текучести стали $\sigma_T = 313 \ M\Pi a$; коэффициент трения при продольном смещении деталей $f_1 = 0,14$;

 μ_1 , μ_2 – коэффициенты Пуассона (для стали $\mu \approx 0,3$); по условиям эксплуатации соединение нагружено осевой силой $F = 6,65 \cdot 10^3$ Н. Расчет прочности соединения проводился по классическим формулам [8]. У посадок с натягом неподвижность сопрягаемых деталей под действием нагрузок обеспечивается силами трения, возникающими при упругой деформации деталей, создаваемой натягом. Минимальный допускаемый натяг определяется исходя из возможных наибольших сил, действующих на сопряжение:

$$N_{\min}^{F} = \frac{F}{\pi l f_{1}} \left(\frac{C_{1}}{E_{1}} + \frac{C_{2}}{E_{2}}\right) = 0,19 \cdot 10^{-3} \text{ mm},$$

где С₁, С₂ – вспомогательные коэффициенты, определяемые по формулам (1) и (2)

$$C_{1} = \frac{1 + \left(\frac{D}{d}\right)^{2}}{1 - \left(\frac{D}{d}\right)^{2}} + \mu_{1} = 2,87,$$

$$C_{2} = \frac{1 + \left(\frac{d_{1}}{D}\right)^{2}}{1 - \left(\frac{d_{1}}{D}\right)^{2}} - \mu_{2} = 0,7,$$
(1)
(2)

где d_1 – диаметр осевого отверстия, $d_1 = 0$. Максимальное давление P_{max} определяется из условия прочности сопрягаемых деталей. В качестве P_{max} берется меньшее из допустимых значений давлений – $P_{\text{доп}}$, которые рассчитываются по следующим формулам:

- для втулки: $P_{\text{доп}} = 0,58 \,\sigma_{\text{T}} \left[1 - \left(\frac{D}{d_2} \right)^2 \right] = 101,6 \text{ МПа};$

- для вала:
$$P_{\text{доп}} = 0,58 \,\sigma_{\text{T}} \left[1 - \left(\frac{d_1}{D} \right)^2 \right] = 181,5 \,\text{МПа.}$$

Так как максимальный натяг рассчитывается из условий прочности деталей, определяем величину максимально допустимого натяга для данного сопряжения:

$$N_{\max}^{P} = P_{\max} D\left(\frac{C_1}{E_1} + \frac{C_2}{E_2}\right) = 0,036$$
 MM.

С учетом полученных результатов определяем допуски прессового соединения по ГОСТ [9] и выбираем посадку:

$$N_{\max} = 0,035 \text{ mm} < N_{\max}^{P} = 0,036 \text{ mm} ; \ N_{\max} = 0,001 \text{ mm} > N_{\min}^{P} = 0,19 \cdot 10^{-3} \text{ mm} ; \ \varnothing 20 \frac{H7 (^{+0,021})}{p6 (^{+0,022}_{+0,022})}.$$

Определяем температуру нагрева охватывающей детали, необходимую для поперечной запрессовки [10]:

$$t = \frac{\Delta}{d\alpha} + t_0 = \frac{0,036}{20 \cdot 12, 4 \cdot 10^{-6}} + 20 = 170^{\circ} \text{C} ,$$

где Δ – максимальный натяг в соединении, мм; α – коэффициент линейного расширения материала охватывающей детали, $\alpha = 12,4 \cdot 10^{-6} \, {}^{\circ}\mathrm{C}^{-1}$; t_0 – температура помещения, °С.

Проведенные теоретические расчеты показывают, в каких пределах должен находиться действительный натяг в соединении при проведении экспериментальных исследований.

Экспериментальные результаты и обсуждение

Для экспериментального исследования влияния шероховатости поверхности на усилие распрессовки прессового соединения из материала сталь 40, изготовлено 9 образцов, которые разделены на 3 группы. Образцы в этих группах существенно отличаются только шероховатостью их функциональных поверхностей. Чтобы создать разный исходный микрорельеф, на каждом образце (группе образцов) используются различные методы и режимы обработки. При этом технология обработки каждого образца и режимы обработки представлены в табл. 1. Все изготовленные образцы подвергаются распрессовке соединения с фиксацией приложенного усилия (табл. 2). Распрессовка соединения проводилась на разрывной машине RAUENSTEIN ZDM 10/91.

Из табл. 2 видно, что при одинаковом значении действительного натяга в прессовых соединениях с разной шероховатостью поверхности меняется усилие распрессовки (образцы 3, 5, 8).

Учитывая, что шероховатость поверхности влияет на усилие распрессовки прессовых соединений, появляется возможность оптимизировать это влияние. Задача оптимизации шероховатости не решается с использованием стандартных параметров шероховатости [11]. Такую возможность обеспечивают непараметрические критерии оценки и контроля шероховатости, позволяющие установить лучшую микрогеометрию этих поверхностей. В качестве критериев принято использовать графические изображения различных функций, таких как плотности распределения ординат и тангенсов углов наклона профилей, функций распределения ординат и тангенсов углов наклона профилей, опорных кривых профиля (кривых Аббота). На рис. 2 показаны кривые плотностей распределения тангенсов углов наклона профилей испытуемых поверхностей.

Номер группы образцов	Вид обработки	Режимы обработки	R_z , мкм	
(вал и втулка)			Вала	Втулки
1	Строгание	<i>S</i> = 0,7 мм/об, <i>n</i> = 355 об/мин	17,64	15,04
2	Шлифование	n = 180 об/мин, $V = 35$ м/сек	12,69	10,05
3	Полирование	<i>n</i> = 1400 об/мин	2,1	2,65

Таблица 1. Виды и режимы обработки образцов: S – подача; n – обороты; V – скорость резания; R_z – высота неровностей профиля по десяти точкам (параметрический критерий оценки шероховатости поверхности)

Номера образцоя	В	Усилие распрессовки, Н	Действительный натяг, мм
	1	$10,9 \cdot 10^3$	0,023
1 группа	2	$11,8\cdot10^{3}$	0,024
	3	$13,3 \cdot 10^3$	0,028
2 группа	4	$15,4\cdot 10^{3}$	0,025
	5	$17,6\cdot10^{3}$	0,028
	6	$16,2\cdot 10^3$	0,027
3 группа	7	$22 \cdot 10^{3}$	0,026
	8	$26,6\cdot10^3$	0,028
	9	$26 \cdot 10^3$	0,028

Таблица 2. Результаты эксперимента





Согласно данной методике, получены графики непараметрических критериев, шероховатости поверхностей втулок и валов из табл. 2, где номер графика соответствует номеру образца. По графикам можно судить, как меняется шероховатость поверхности в результате применения разных видов и режимов обработки, и при достаточном для статистической обработки количестве экспериментов можно определить лучшую из возможных в данных условиях (эталонную) микрогеометрию сопрягаемых поверхностей, обеспечивающую наибольшее усилие распрессовки прессовых соединений (рис. 3).

Средняя кривая показывает эталонное значение, а две других – допуски на отклонение этой величины. Эти кривые, принятые в качестве эталонных, относятся к поверхностям образцов, которые, при прочих равных условиях, показали наилучшие результаты эксперимента. Это также означает, что, если соответствующие непараметрические критерии шероховатости поверхностей серийных деталей будут соответствовать эталонным (рис. 3), то после описанной выше запрессовки требуемое усилие распрессовки будет гарантировано. Вероятно, существуют шероховатости поверхностей, которые могут обеспечить, при прочих равных условиях, большее усилие распрессовки, но в данной работе была определена лучшая микрогеометрия, обеспечиваемая условиями производства конкретного предприятия.



Рис. 3. «Эталонный» график «Плотность распределения тангенсов углов наклона профиля» с указанием поля допусков на отклонение для валов (а), втулок (б). Средняя кривая – эталонное значение

Заключение

Практическая значимость исследования состоит в разработке методики определения влияния шероховатости поверхности на усилие распрессовки прессовых соединений и нахождения лучшей из возможных микрогеометрий в данных производственных условиях. Можно утверждать, что разработанная методика позволяет найти лучшую из возможных микрогеометрию для любого предприятия и не имеет значения, каким образом будут получены эталонные кривые при обработке поверхностей – такие поверхности будут гарантировать ожидаемые свойства. Но в условиях серийного производства данного предприятия для надежности использованы параметры, давшие лучшие результаты, полученные в ходе эксперимента.

Литература

- Буканова И.С., Роговой В.М. О прочности прессового соединения втулка-корпус // Материалы конференция студентов, аспирантов и профессорско-преподавательского состава, посвященная 60-летию Алтайского ГТУ им. И.И.Ползунова. – Ч. 11. Инженерно-физический факультет. – Барнаул: изд-во АлтГТУ, 2002. – С. 37.
- Буканова И.С., Татаркин Е.Ю., Ятло И.И. Технологическое обеспечение повышенной несущей способности неподвижных соединений типа «корпус-втулка» // Ползуновский вестник. – 2009. – № 1–2. – С. 183–187.
- 3. Воячек И.И. Расчет прочности соединений с натягом, собранных поперечным методом // Изв. вузов. Машиностроение. 1996. № 4–6. С. 15–25.
- 4. Григорьева О.А. Технологическое обеспечение прочности профильных неподвижных соединений упругопластическим деформированием элементов соединения: Автореф. дис. ... канд. техн. наук. Омск, 2004. 20 с.
- 5. Мусалимов В.М., Валетов В.А. Динамика фрикционного взаимодействия. СПб: СПбГУ ИТМО, 2006. 191 с.
- 6. Валетов В.А., Иванов А.Ю. Непараметрический подход к оценке качества изделий // Металлообработка. – 2010. – № 6. – С. 55–59.
- 7. Валетов В.А. Оптимизация микрогеометрии поверхностей деталей в приборостроении. Л.: ЛИТМО, 1989. 100 с.
- 8. Анухин В.И. Допуски и посадки. Выбор и расчет, указание на чертежах: Учебное пособие. 2-е изд., перераб. и доп. СПб: Изд-во СПбГТУ, 2001. 219 с.
- ГОСТ 25347-82. Основные нормы взаимозаменяемости. Единая система допусков и посадок. Поля допусков и рекумендуемые посадки. – Введ. 01.01.78. – М.: Изд-во стандартов, 1983. – 53 с.
- 10. Морозов И.М., Шамин В.Ю. Основы технологии сборки в машиностроении: Учебное пособие. 2-е изд., перераб. и доп. Челябинск: ЮУрГУ, 2006. 72 с.
- 11. Иванов А.Ю., Леонов Д.Б. Технологические методы обеспечения качества изделия // Научнотехнический вестник СПбГУ ИТМО. – 2011. – № 5 (75). – С. 111–113.

Иванов Андрей Юрьевич	_	Санкт-Петербургский национальный исследовательский университет инфор-
		мационных технологий, механики и оптики, кандидат технических наук, до-
		цент, ivaanur72@mail.ru
Леонов Димилян Божидаров	_	BM3 AO Сопот, Болгария, зам. директора, аспирант, dimilqn@mail.ru

9

МЕТОДЫ И СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ

УДК 004.056 МЕТОДИКА АНАЛИЗА АРХИТЕКТУРЫ СИСТЕМ ЗАЩИТЫ ИНФОРМАЦИИ НА ОСНОВЕ ТИПОВЫХ ЭЛЕМЕНТОВ М.Е. Сухопаров, И.Н. Соловьев, И.С. Лебедев, И.И. Комаров

Предложен метод анализа архитектуры системы защиты информации на основе типовых элементов. Подход основан на использовании вероятностных характеристик типовых элементов защиты, учитывает поток событий для одного типового элемента системы защиты информации, что позволяет производить сравнительный анализ и обоснование выбора архитектуры на ранних этапах жизненного цикла. Приведен пример анализа архитектуры системы. Ключевые слова: анализ архитектур СЗИ, выбор архитектуры, оценка уязвимости.

Введение

Повсеместное использование глобальных вычислительных сетей, систем распределенных вычислений, облачных технологий, появление элементов «открытого контура» в системах управления, где необходимо обеспечение конфиденциальности, целостности и доступности циркулирующих данных, обусловливает важность задач защиты информационно-телекоммуникационных объектов.

Развитие современных технологий проектирования и возрастающие возможности средств передачи информации дают возможность реализовывать различные топологии информационных систем, оптимизированные для решения предметно ориентированных задач, что требует прогнозной оценки показателей качества обеспечения информационной безопасности на различных стадиях жизненного цикла [1].

Общий подход к построению защиты информационных объектов состоит в том, чтобы для каждой угрозы внедрить решение, которое снизит вероятность ее реализации до приемлемого уровня. Особенностью современных средств обеспечения информационной безопасности является многоуровневый принцип защиты, где злоумышленнику для осуществления угроз конфиденциальности, целостности и доступности необходимо преодолеть разные элементы [2].

Постановка задачи

Широкое распространение типовых средств защиты информации известных производителей ведет к увеличению активности по поиску существующих в них уязвимостей [3]. Если одно из таких средств недостаточно снижает риски, то внедряются несколько решений разных разработчиков.



Рис. 1. Структура защищаемого объекта

Например, для информационно-телекоммуникационных систем, имеющих доступ в Интернет, устанавливаются различные антивирусы на шлюз, почтовый сервер и рабочие станции. Удаленные и мо-

бильные рабочие места, имеющие доступ к информационным ресурсам информационнотелекоммуникационной системы (ИТКС), оснащаются средствами защиты, рекомендованными администраторами политики безопасности. Таким образом, в системе появляется регламентированный набор средств защиты. В связи с этим возникает ряд задач, связанных с оценкой заданных показателей защищенности объектов.

Рассмотрим систему защиты распределенных ресурсов *S*, где для каждого защищаемого объекта $o_i \in O$ подсистема защиты может состоять из набора элементов системы защиты информации (СЗИ) из $m_i \in M$. В качестве примера можно привести систему, препятствующую утечке, например, текстовых документов. Преодоление элементов СЗИ, обрабатывающих текстовую информацию, может быть связано с уязвимостями, возникающими вследствие несовершенства алгоритмов обработки текстовой информации на морфологическом, синтаксическом и семантическом уровнях, с применением средств примитивного кодирования, использованием псевдографики в отправляемых и принимаемых сообщениях, систем программирования, содержащихся в обычных текстовых редакторах и офисных приложениях. Исходя из этого, необходимо учитывать широкий спектр возможностей, имеющихся у потенциального нарушителя, и использовать разнообразные элементы защиты на разных уровнях.

На рис. 1 представлен защищаемый информационно-технический объект. Особенностью преодоления защиты является возможность не только прямого воздействия на последовательные уровни защиты элементов СЗИ, но поиск уязвимостей в каждом элементе защиты по отдельности.

Предлагаемый подход

Допустим, что в ИТКС осуществляются процессы информационного противоборства, одной из составляющих которого являются атаки на лингвистический модуль определения данных, содержащих конфиденциальную информацию центрального сервера управления DLP-системы [4]. Исследуя систему, возможно определить уязвимости, связанные с несовершенством алгоритмов морфологического, синтаксического, семантического уровней обработки естественного языка, которые с определенной вероятностью позволят преодолеть средства защиты, осуществляющие идентификацию информации сообщений [5]. Например, для лингвистического модуля определения данных использование методов примитивного кодирования дает возможность обойти подсистемы поиска по точному совпадению слов с учетом морфологии, анализа регулярных выражений, но должно быть обнаружено при идентификации цифровых отпечатков или подсистемой поведенческого анализа. При посылке различных видов потенциально опасных сообщений в зависимости от свойств информации могут быть преодолены одни типовые элементы и в то же время обнаружена угроза другими [6].

Обозначим λ – количество преодолений типового элемента СЗИ в единицу времени *t*. Для упрощения модели будем считать, что для одного типового элемента СЗИ поток событий обладает свойствами стационарности, отсутствия последействия и ординарности. Событие преодоления одного элемента защиты зависит только от длины временного промежутка, в течение которого приходится обрабатывать текстовые сообщения. Событие преодоления одного элемента защиты не зависит в любом промежутке времени от того, появлялись события в прошлом или нет. Появление более одного события за малый промежуток времени практически невозможно. Вероятность того, что произойдет преодоление *n* типовых элементов СЗИ, будет определяться как

$$p(n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$

Одной из задач построения СЗИ является получение количественных оценок, позволяющих производить сравнение и выбор элементов систем, а также выносить обоснованное решение об использовании той или иной архитектуры.

Допустим, что система состоит из типовых элементов. Для определения архитектуры необходимо определить их требуемое количество.

Введем порог l < N, показывающий предельно допустимый уровень преодоленных элементов. Вероятность того, что в течение времени *t* не менее *l* элементов СЗИ будут преодолены, равна

$$p(n \ge l) = 1 - \sum_{n=1}^{l} \frac{(\lambda t)^n}{n!} e^{-\lambda t} .$$

На рис. 2 приведены зависимости для λ =1 при значениях *n* и *l*, равных (12 и 8), (10 и 7), (15 и 10) соответственно.

На рис. 3 приведены зависимости для различных λ при значениях *n* и *l*, равных (12 и 8), (10 и 7), (15 и 10) соответственно и периода времени t = 1.



Рис. 2. Вероятность преодоления защиты в течение периода времени t



Рис. 3. Вероятность преодоления защиты при различных интенсивностях

С другой стороны, данный подход позволяет сравнивать архитектуры СЗИ. Поясним подход на примере. Пусть для последнего рассмотренного случая необходимо выбрать структурную схему, позволяющую потенциально достигать наибольшие показатели защищенности.

Допустим, что необходимо обосновать создание последовательности из элементов СЗИ по первому или второму типу в зависимости от статистических показателей, характеризующих преодоление. Для этого составляем отношение

$$\frac{p(n \ge l_1)}{p(n \ge l_1)} = \frac{1 - \sum_{n=1}^{l_1} \frac{(\lambda t)^n}{n!} e^{-\lambda t}}{1 - \sum_{n=1}^{l_2} \frac{(\lambda t)^n}{n!} e^{-\lambda t}}$$

На рис. 4 показаны отношения вероятностей непреодоления защиты различных архитектур.



Рис. 4. Отношения вероятностей непреодоления защиты различных архитектур

Анализ зависимости показывает предпочтение той или иной архитектуры в зависимости от интенсивности отказов элементов СЗИ.

Заключение

Широкое распространение типовых средств защиты информации известных производителей, их постоянный анализ с целью преодоления со стороны субъектов информационного противоборства обусловливает необходимость развития прогнозных методов, основанных на технических характеристиках отдельных элементов системы защиты информации.

В работе продемонстрирована возможность применения метода анализа архитектуры систем защиты информации на основе типовых элементов для решения типовой задачи анализа и оценки заданных показателей защищенности объектов, зависящих от выбора элементов системы защиты информации.

Предложенный метод обеспечивает возмоность формальной оценки преимуществ разных архитектур, определяющих свойства системы защиты информации, получить прогнозные результаты расчетов для оценки системы защиты информации, выделить наиболее уязвимые элементы системы защиты информации и определить направления повышения защищенности системы.

Работа выполнена в рамках НИР № 12360.

Литература

- 1. Гвоздев А.В., Зикратов И.А., Лебедев И.С., Лапшин С.В., Соловьев И.Н. Прогнозная оценка защищенности архитектур программного обеспечения // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – № 4 (80). – С. 126–130.
- 2. Зикратов И.А., Одегов С.В. Оценка информационной безопасности в облачных вычислениях на основе байесовского подхода // Научно-технический вестник информационных технологий, механики и оптики. 2012. № 4 (80). С. 121–126.
- 3. Лебедев И.С., Борисов Ю.Б. Анализ текстовых сообщений в системах информационной безопасности. // Информационно-управляющие системы. – 2011. – № 2 (51). – С. 37–43.
- Артамонов В.А. Модели безопасности информационных технологий критичных информационноизмерительных систем [Электронный ресурс]. – Режим доступа: http://itzashita.ru/publications/modelibezopasnosti-informacionnyx-texnologij-kritichnyx-informacionno-izmeritelnyx-sistem-chast-1.html, свободный. Яз. рус. (дата обращения 26.02.2013).
- Manning C.D., Raghavan P., Schutze H. Introduction to Information Retrieval. Cambridge University Press, Cambridge, England. – 2009. – 504 p.
- 6. Медведовский И.Д., Семьянов П.В., Леонов Д.Г., Лукацкий А.В. Атака из Internet. М.: Солон-Р, 2002. С. 140–144.

Сухопаров Михаил Евгеньевич	_	Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, sukho- parovm@gmail.com
Соловьев Игорь Николаевич	-	Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, solovyev.i.n@yandex.ru
Лебедев Илья Сергеевич	_	Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, доцент, lebedev@cit.ifmo.ru
Комаров Игорь Иванович	-	Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат физмат. наук, доцент, Komarov@cit.ifmo.ru

10

НОВЫЕ ТЕХНОЛОГИИ В ОБРАЗОВАНИИ

УДК 371.26 АВТОМАТИЗАЦИЯ СОСТАВЛЕНИЯ ЗАДАНИЙ ПО МАТЕМАТИКЕ МЕТОДОМ ВАРИАЦИИ ПАРАМЕТРОВ Е.В. Костюченко

В работе освещены методические аспекты автоматического выбора параметров при составлении типовых заданий по математике. Как следствие, сформулированы необходимые с точки зрения математика-методиста требования к обеспечивающей автоматизацию технической системе. Рассмотренный подход может использоваться в других дисциплинах, допускающих аналогичную формализацию контрольно-измерительных материалов. Ключевые слова: генерация заданий, метод вариации параметров, система генерации вариантов.

Введение

Современное образование все быстрее движется по пути оптимизации затрат квалифицированного труда преподавателей. Это с неизбежностью влияет на контролирующий процесс, где меньше времени отводится традиционным формам контроля при увеличении доли самостоятельно выполняемых работ и применении электронных систем тестирования. Одновременно совершенствуется коммуникативное поле студентов и учащихся школ, в силу чего многократное использование одних и тех же контрольноизмерительных материалов искажает результаты контроля. Составление заданий вручную – трудоемкий и высокочувствительный к ошибкам составителя процесс, результаты которого требуют дополнительных затрат на переформатирование при использовании в разных условиях. Применение средств автоматизации позволяет разрабатывать не задание в отдельности, а его шаблон с переменными параметрами. Задача машины – перебрать значения параметров и дать на выходе комплект однотипных, но отличающихся друг от друга заданий. Преимуществами подхода становятся высокая производительность и хорошая защищенность от рутинных ошибок. Недостаток – информационный барьер, состоящий в необходимости предварительного освоения среды разработки (системы генерации вариантов).

Задача построения системы генерации вариантов (СГВ) не нова. Первые из известных автору решений, опубликованных в нашей стране, получены приблизительно двадцать лет назад. Большая часть разработок делалась применительно к школьной программе, хотя прогресс последних лет не обошел математические дисциплины в вузах. Краткий обзор современных достижений дан в [1]. Различные подходы к проблеме отражены в [2–5]. Тем не менее, львиная доля существующих СГВ не выходит из сферы своего локального применения, не отражена в публикациях, и судить о размерах этого айсберга можно лишь по косвенным упоминаниям. Из-за недостаточного обмена опытом и множества частных решений более или менее универсальный подход к построению СГВ на сегодняшний день не сформирован. Настоящая работа призвана сделать шаг в этом направлении.

Метод вариации параметров

Разнообразие генерируемых из одного шаблона вариантов задания достигается путем автоматической вариации указанных в нем целочисленных параметров. Это не значит, что итоговые задания могут различаться только числовыми данными. Меняться в зависимости от значений параметров могут символьные строки и даже формулировки задания в целом. Ярко выраженная однотипность получается при изменении в условиях только числовых данных, но не является следствием метода вариации как такового. Один хорошо методически проработанный шаблон может порождать различные по структуре задания, а чередование нескольких шаблонов позволяет получать варианты, уже практически неотличимые от составленных вручную.

Параметризация заданий по математике является сложной методической задачей со специфическими особенностями. Прежде всего, необходимо в некотором принятом за основу задании определить допускающие изменения составляющие. Ими могут быть как числовые данные, так и структурные элементы, вплоть до перечня близких по смыслу формулировок. Переменным составляющим следует сопоставить набор числовых параметров так, чтобы каждый из наборов их значений однозначно соответствовал одной из допустимых модификаций задания. Тем самым предварительно очерчивается область допустимых значений параметров.

Рассмотрим пример из курса математического анализа, в котором требуется найти предел функ- $1 - \sqrt{1-x}$

ции: $\lim_{x\to 0} \frac{1-\sqrt{1-x}}{x}$. Студентам могут быть известны несколько способов решения, но ответом является

число, что удобно как для ручной, так и для автоматической проверки (контрольные задания в большинстве разделов математики допускают постановку, приводящую к числовому ответу). Заметим, что смысл задания сохраняется при любом целом значении параметра a, натуральном c и $b=c^2-a$, если их подста-

вить в выражения $\lim_{x \to a} \frac{\sqrt{x+b}-c}{x-a}$, $\lim_{x \to a} \frac{x-a}{\sqrt{x+b}-c}$ и $\lim_{x \to a} \frac{c-\sqrt{x+b}}{x-a}$. Разумеется, не следует брать громозд-

кие числа. Здесь вполне достаточно менять a от -9 до 9 и c от 1 до 10. Можно исключить a=0 и ограничить b по абсолютной величине (или уменьшить диапазон изменения c), затем добавить формулу для ответа, и шаблон готов к внесению в СГВ.

Шаблоны с двумя параметрами очень просты и используются редко. Рассмотренный пример можно продолжить: варьировать количество слагаемых и корней, квадратные корни заменить переменными дробными степенями и т.д. При этом необходимо:

- контролировать сочетания значений, приводящие к вырожденным или некорректным ситуациям;
- сохранять единый уровень трудоемкости типовых заданий;
- следить за тем, чтобы разные наборы параметров не давали практически одинаковые результаты.

Последние два пункта хорошо иллюстрируются обычными трудностями с едва ли не самым распространенным объектом генерации в математике – системами линейных алгебраических уравнений. Если предполагается решение методом Гаусса, то наличие «хороших» коэффициентов и ответов (т.е. целых и небольших по абсолютной величине) вовсе не гарантирует отсутствия трехзначных чисел в промежуточных выкладках даже для системы из трех уравнений. При этом в некоторых вариантах система может решаться буквально на первом шаге. Далее, разные наборы параметров без специального контроля могут давать системы, отличающиеся друг от друга перестановкой строк, или обозначением неизвестных, и т.п. Множество всевозможных нюансов учитывается путем тщательного отбора допустимых значений параметров при создании шаблона и посредством дополнительных тестовых условий при его отладке. Базовые возможности СГВ, позволяющей методисту выполнять эту работу, рассмотрены в следующем разделе.

Требования к СГВ

СГВ – это совокупность программных средств, обеспечивающих обработку введенного составителем шаблона с получением на выходе различных вариантов задания. В состав СГВ включаются также средства ввода и хранения шаблонов, комплектования контрольно-измерительных материалов и выдачи их в требуемом формате. Выбор вспомогательных технологий зависит от контекста использования системы и часто так же неоднозначен, как выбор между визуальной средой редактирования, наподобие MS Word, и TeX-нотацией.

По принципу работы СГВ условно делятся на два класса. Первый класс – простейшие – генерируют шаблонные задания независимо, без учета меры их различия. Значения параметров в таких СГВ обычно подбираются с помощью датчика псевдослучайных чисел. СГВ второго класса позволяют создавать комплекты с максимально различающимися заданиями. Оценка различий выполняется соответственно принятой в системе метрике и специальным указаниям методиста. СГВ второго класса не генерируют «на лету», а создают базу всевозможных заданий для каждого шаблона и формируют комплекты путем извлечения из нее выборки на основе оценок разнообразия. Однократность обработки дает пре-имущество в возможности составления ресурсоемких шаблонов.

Перспективной является разработка СГВ второго класса, позволяющих методисту вне зависимости от условностей инструментальной среды выполнять следующие действия:

- указывать диапазон изменения целочисленных параметров;
- использовать переменные (целочисленные, с плавающей точкой и символьные строки);
- работать с одномерными и двумерными массивами;
- вводить функции, зависящие от параметров и переменных, и использовать возвращаемые ими значения наравне с параметрами;
- задавать условия допустимости набора значений параметров;
- задавать специальную числовую характеристику (вес) параметров и переменных, чьи значения критичны для определения меры различия наборов значений;
- разбивать параметры на группы, проверять допустимость их значений по условиям для каждой группы в отдельности, а затем проверять условия на параметры в совокупности (этот пункт важен при большом количестве параметров, сплошной перебор по диапазону которых нецелесообразен, а подчас и невыполним);
- оформлять текст заданий с формулами, таблицами и графиками, включая в него ссылки на подлежащие генерации данные;
- давать несколько формулировок задания на базе одного набора параметров с указанием режима обращения к ним: либо чередовать в разных вариантах, либо обращаться адресно, как, например, к ответам и проверочным материалам;
- получать перечень сгенерированных заданий и сопутствующую техническую информацию (максимально возможное количество вариантов, частоты принимаемых параметрами значений и пр.).

Как видно из требований, в основе СГВ лежит адаптированный язык программирования, который по возможностям математических расчетов сравним с языком программирования общего назначения. Необходимость программирования резко ограничивает круг методистов, потенциально заинтересованных в возможностях автоматизации. В полной мере этот барьер непреодолим, но сделать его минимальным – основная задача разработчиков СГВ.

Заключение

Данное в работе описание характеристик технической системы, обеспечивающей генерацию типовых вариантов, может быть использовано при разработке таких систем применительно к конкретным условиям. Одним из результатов автора в этой области является интегрированная вычислительная среда «Математический практикум» (свидетельство о регистрации № 2005610503 от 24.02.2005 г.). С ее помощью были созданы тесты по математике для системы дистанционного обучения НИУ ИТМО. Опыт подтвердил эффективность метода вариации параметров и показал возможность его использования при составлении высококачественных контрольно-измерительных материалов, в которых различие вариантов не исчерпывается только числовыми значениями. Вопрос создания доступных и понятных широкому кругу пользователей системы генерации вариантов на сегодняшний день остается открытым.

Литература

- Гинко В.И., Гинко М.С. Автоматизация обучения и контроля знаний с использованием скриптовых технологий для генерации задач (в АСУ образовательного назначения) // Научный поиск. – 2012. – № 3 (5). – С. 48–51.
- 2. Посов И.А., Рукшин С.Е. Генерируемые задачи в системе для организации удаленной работы с задачами // Научно-технический вестник СПБГУ ИТМО. – 2010. – № 5 (69). – С. 130.
- 3. Братчиков И.Л. Генерация тестовых заданий в экспертно-обучающих системах // Вестник РУДН. 2012. № 2 [Электронный ресурс]. Режим доступа: http://imp.rudn.ru/vestnik.aspx?id=2012_2/, свободный. Яз. рус. (дата обращения 24.03.2013).
- 4. Кручинин В.В., Магазинников Л.И., Морозова Ю.В. Модели и алгоритмы компьютерных самостоятельных работ на основе генерации тестовых заданий // Известия ТПУ. – 2006. – № 8 – [Электронный ресурс]. – Режим доступа: http://cyberleninka.ru/article/n/modeli-i-algoritmy-kompyuternyhsamostoyatelnyh-rabot-na-osnove-generatsii-testovyh-zadaniy /, свободный. Яз. рус. (дата обращения 24.03.2013).
- 5. Кулик А.С., Чухрай А.Г., Вагин Е.С., Педан С. И. Формализация генерации заданий для комплекса интерактивных web-тестов по математике // Радіоелектронні і компь'ютерні системи. 2010. № 7 (48). С. 86–89.

Костюченко Евгений Викторович

Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат физ.-мат. наук, доцент, ev28@yandex.ru

КРАТКИЕ СООБЩЕНИЯ

УДК 535.21; 681.7.03 ФОТОДЕСТРУКЦИЯ НАНОЧАСТИЦ СЕРЕБРА В ФОТО-ТЕРМО-РЕФРАКТИВНЫХ СТЕКЛАХ

Д.А. Игнатьев, А.И. Игнатьев, Н.В. Никоноров

Исследованы процессы разрушения серебряных наночастиц в фото-термо-рефрактивных стеклах при облучении импульсным лазером (532 нм) наносекундной длительности в зависимости от дозы облучения. Показано, что под действием лазерного излучения происходит фотофрагментация наночастиц серебра с образованием серебряных кластеров, а также фотоионизация нейтральных атомов серебра с переходом их в ионное состояние. Ключевые слова: фотодеструкция, фото-термо-рефрактивные стекла, серебряные наночастицы.

Фото-термо-рефрактивные (ФТР) стекла представляют собой перспективный класс фоточувствительных материалов, разработанных для записи высокоэффективных объемных фазовых голограмм. ФТР стекла – это сложные гетерофазные структуры, в которых под действием света и последующей термической обработки выделяются наночастицы серебра, на которых вырастают нанокристаллы NaF-AgBr [Л]. На основе этих стекол создаются фазовые голограммы, используемые в качестве узкополосных селекторов, комбайнеров световых пучков, фильтров, повышающих спектральную яркость лазерных диодов. Однако полоса плазмонного резонанса серебряных наночастиц с максимумом 450–480 нм существенно ограничивает применение голограмм в видимом диапазоне спектра. По этой причине актуальной задачей при записи объемных фазовых голограмм в ФТР стекле является уменьшение поглощения серебряных наночастиц за счет их фотолитического разрушения (фотодеструкции).

Образцы в виде плоскопараллельных пластин ФТР стекол облучались ультрафиолетовым излучением ртутной лампы высокого давления и термически обрабатывались в муфельной печи. Измерение спектров поглощения осуществлялось на спектрофотометре Lambda 650. Фотодеструкция серебряных наночастиц производилась излучением импульсного YAG:Nd³⁺ лазера (Solar LQ-129) с длиной волны 532 нм. Энергия импульса составляла 68 мДж, длительность импульса 13 нс, частота следования импульсов 10 Гц.

Впервые обнаружено, что:

- при облучении ФТР стекол, содержащих наночастицы серебра, происходит существенное изменение спектров поглощения. Амплитуда полосы плазмонного резонанса серебряных частиц уменьшается в зависимости от дозы излучения, вплоть до полного ее исчезновения (рисунок);
- 2. фотолитическому разрушению подвержены наночастицы серебра, находящиеся как в окружении стеклообразующей матрицы, так и внутри оболочки из кристаллической фазы NaF-AgBr.

На основании анализа спектров поглощения сделан вывод, что под действием импульсного лазерного излучения происходит фотофрагментация наночастиц серебра с образованием серебряных кластеров, а также фотоионизация нейтральных атомов серебра с переходом их в ионное состояние. Эффект фотодеструкции может быть использован для создания высокоэффективных фазовых голограмм, работающих в видимом диапазоне спектра.



Рисунок. Спектры поглощения ФТР стекла при разных дозах излучения лазера

[Л]. Никоноров Н.В., Панышева Е.И., Туниманова И.В., Харченко М.В. Особенности окрашивания мультихромных стекол под действием лазерного излучения // Физика и химия стекла. – 1993. – Т. 19. – № 3. – С. 442–448.

Игнатьев Дмитрий Александрович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, инженер, ignatiev_d_a@mail.ru

Игнатьев Александр Иванович — Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, зав. лабораторией, ignatiev@oi.ifmo.ru

Никоноров Николай Валентинович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор физ.-мат. наук, профессор, зав. кафедрой, Nikonorov@oi.ifmo.ru

УДК 622.7, 528.854, 004.93 ОПТИКО-ЭЛЕКТРОННЫЙ КОМПЛЕКС ДЛЯ ЭКСПРЕСС-АНАЛИЗА ОБОГАТИМОСТИ РУД ТВЕРДЫХ ПОЛЕЗНЫХ ИСКОПАЕМЫХ ОПТИЧЕСКИМ МЕТОДОМ А.А. Алёхин

Приведены результаты работ по созданию экспериментального образца оптико-электронного комплекса, предназначенного для оценки степени обогатимости руд твердых полезных ископаемых оптическим методом. Ключевые слова: обогатимость, твердые полезные ископаемые, оптический метод обогащения, цвет, обработка изображений.

Оптический (в зарубежной практике – color sorting) метод сепарации является наиболее универсальным среди радиометрических и широко используется при переработке минерального сырья, твердых бытовых и промышленных отходов, сельскохозяйственных культур. Обогащение минерального сырья данным методом основывается на использовании современных видеоинформационных технологий и анализа в реальном времени таких оптических характеристик минералов, как цвет, блеск, прозрачность, отражательная способность.

Несмотря на применение в существующих оптических сепараторах методов машинного обучения, они, главным образом, эффективны для решения простых задач: в случае ярко выраженного цветового контраста объектов контроля, при разделении по нескольким цветовым оттенкам или же по однородности одного из цветов. Однако при необходимости различения в автоматическом режиме тонких цветовых оттенков, контроле минеральных структур со сложной поверхностной и (или) внутренней структурой (образцы с рельефной поверхностью, дающей вторичные тени на изображении; частично и неоднородно прозрачные образцы), потенциально перспективный метод часто пасует, не позволяя достигать должного качества сортировки.

В большинстве случаев причиной неэффективности использования оптического метода является пренебрежение особенностями получения и обработки цветного изображения (например, используемые методы цветовой интерполяции) в блоке регистрации сепаратора, а также свойствами используемых цветовых моделей (RGB, L*a*b* и т.п.). Особенности получения и обработки цветного изображения обусловливают максимальный для данного конкретного сепаратора диапазон различаемых цветовых оттенков анализируемого объекта, а используемая цветовая модель – минимальную границу их различения.

Кроме того, несмотря на активное использование оптического метода и обогатительного оборудования, его реализующего, для сепарации твердых полезных ископаемых разных типов до сих пор не существует как методов оценки обогатимости (за исключением прямой пробы на сепараторе), так и критериев выбора в пользу того или иного сортировочного комплекса для решения конкретной задачи обогащения.

Коллективом кафедры оптико-электронных приборов и систем НУИ ИТМО разработаны принципы организации процесса экспресс-анализа обогатимости руд оптическим методом, а также предложена конструкция экспериментального образца соответствующего измерительного комплекса, в настоящее время не имеющего аналогов на рынке оборудования для горнодобывающей промышленности.

Разработанный экспериментальный образец содержит два канала регистрации с разрешением 800×600 пикселей и позволяет проводить двусторонний анализ минеральных образцов крупностью от 5 до 150 мм. При этом сушка поверхности образцов не требуется, что очень удобно с точки зрения организации технологического процесса оценки обогатимости. Другой особенностью предлагаемого решения является то, что анализ производится с использованием сразу трех цветовых моделей – RGB, YUV и HLS. При этом выбирается модель, наиболее подходящая для сепарации конкретного вида минерального сырья. Кроме того, помимо цвета, анализируется возможность использования при сепарации и других селективных признаков, характеризующих поверхностную структуру минералов, например, наличие «блесток» (точечных объектов) и «прожилок» (линейных объектов).

Результатом анализа рудной пробы являются выбранный селективный признак (или набор признаков), а также параметры настройки сепаратора, оптимальные для обогащения данного типа руды – наиболее эффективная цветовая модель и границы разделения.

При разработке экспериментального образца отдельное внимание было уделено принципам освещения рабочей зоны [1] и методикам настройки цветопередачи [2, 3] системы регистрации аналитического комплекса, имеющим важное значение для обеспечения эффективности анализа и выбора оптимальных условий сепарации. Для проверки выдвинутых теоретических положений, разработанных методик анализа и настройки с помощью разработанного экспериментального образца проведены экспериментальные исследования проб золотосодержащих руд на базе научно-производственного предприятия «ГеоТестСервис» (г. Москва).

Работа проводилась при финансовой поддержке Министерства образования и науки Российской Федерации в рамках ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009– 2013 годы.

- Chertov A., Gorbunova E., Korotaev V., Serikova M., Peretyagin V. Simulation of the multicomponent radiation source with the required irradiance and color distribution on the flat illuminated surface // Proc. SPIE. - 2012. - V. 8429. - 84290D.
- 2. Горбунова Е.В., Коротаев В.В., Тимофеев А.Н., Чертов А.Н. Коррекция цветопередачи камеры с искажающей оптикой посредством анализа цвета изображения фона. Свидетельство о гос. регистрации программы для ЭВМ № 2012613938 от 27.04.2012.
- 3. Алёхин А.А., Горбунова Е.В., Коротаев В.В., Чертов А.Н. Основные принципы настройки цветовых оптико-электронных систем технического зрения промышленного назначения // Изв. вузов. Приборостроение. – 2012. – Т. 55. – № 4. – С. 33–36.

Алехин Артем Андреевич – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, temka-pk@mail.ru

УДК 681.2

ОРГАНИЗАЦИЯ БИБЛИОТЕК В ALTIUM DESIGNER О.В. Кузнецова, Е.Б. Романова

Описана структура библиотек в Altium Designer. Предложены подходы к именованию составляющих библиотеки электронных компонентов – условно-графических обозначений и посадочных мест. Описаны достоинства применения рекомендуемых подходов.

Ключевые слова: печатная плата, библиотека, УГО, посадочное место, Altium Designer.

Наиболее популярной системой автоматизированного проектирования (САПР) печатных плат (ПП) в России является P-CAD, но последняя версия этой САПР была выпущена в 2006 году, и дальнейшее продвижение этого программного продукта прекращено. По этой причине разработчикам ПП приходится внедрять другие САПР – зачастую Altium Designer, так как P-CAD и Altium Designer – это продукты одного разработчика (компании Altium Limited), и переход из P-CAD в Altium Designer проще, чем в другие САПР.

Сквозное проектирование ПП в САПР предполагает использование библиотек условнографических обозначений (УГО) и библиотек посадочных мест под корпус компонента. Библиотека УГО используется для разработки принципиальных электрических схем, а библиотека посадочных мест – для разработки топологии печатной платы. При сквозном проектировании ПП сначала разрабатывается принципиальная электрическая схема, а затем список электрических связей, список электронных компонентов (ЭК) переносится в проект топологии ПП. Для реализации сквозного подхода каждому УГО должно быть присвоено посадочное место. В Р-САД библиотечный компонент включает информационную таблицу, УГО, посадочное место и таблицу выводов. В Altium Designer библиотечный компонент представляет собой УГО с подключенной к нему ссылкой на посадочное место. В связи с таким существенным различием структур библиотек перед проектировщиками ПП встает задача поиска новых подходов к организации библиотек. В первую очередь рассматриваются подходы, применяемые в готовых библиотеках Altium Designer, но УГО в них имеют названия, трудно ассоциируемые с реальными компонентами (например, резисторы названы Res1, Res2, Res3... (рис. 1)). Трудно, а порой и невозможно понять по такому наименованию, какой тип резистора туда заложен. Ориентироваться отечественному разработчику в такой библиотеке очень трудно. Помимо этого, УГО в готовых библиотеках не соответствуют ЕСКД, а посадочные места не подходят для российских заводов-изготовителей ПП, ориентированных на отечественный ГОСТ. По этой причине на каждом российском предприятии разрабатывается своя библиотека ЭК.

Целью проведенной авторами работы стало описание подходов к организации собственных библиотек в Altium Designer, что позволило бы централизованно пользоваться едиными библиотеками в рамках одного предприятия всем участникам проекта – разработчикам библиотек, схемотехникам и конструкторам. Новизна заключается в том, что УГО и посадочные места предлагается именовать более информативно, чем в готовых библиотеках Altium Designer: УГО – с учетом типа ЭК, а посадочные места – на основе анализа классификации корпусов.

При проектировании ПП важно информативно именовать составляющие библиотеки ЭК, чтобы каждый пользователь библиотеки легко ориентировался по наименованиям [Л]. На практике наименования составляющих библиотеки ЭК настолько разнообразны и неоднозначны, что иногда проще создать

свою библиотеку, чтобы свободно в ней ориентироваться, чем выискивать нужные компоненты в готовой библиотеке (зачастую на одном предприятии в разных отделах используются разные библиотеки).

Так как в Altium Designer УГО является основой библиотечного компонента, к которому можно подключать посадочное место в виде ссылки, то авторами рекомендуется в библиотеке в наименовании УГО отражать тип ЭК и добавлять к этому УГО ссылку на одно посадочное место, соответствующее указанному типу. Например, диод Шотки типа КД270А надо назвать в библиотеке УГО (schlib) КД270А и подключать к нему посадочное место с именем TO-220AC, при этом имя посадочного места присваивается в соответствии с типом корпуса (см. рис. 2).

Altium Designer (13.0) - C:\Users\Public\Docume	ents\Altium\AD13\Library\M	Aiscellaneous Devices\Miscellaneous De	evices.SchLib - Miscellaneous Devices.Lib	Pkg
DXP Eile Edit View Project Place I	ools <u>R</u> eports <u>W</u> indow	Help	C:\Users\Public\Documents	s\Altiur • 🔘 • 🔘 - 🐢
	B B Ø □ + >	χ^{2} $ $ $ $ $ $ $ $ $ $ $ $ $ $ Mode \star $ $ $+$ $=$ $ $ $+$	- ⇒ a • ⊠ • ⊞ • ⊡	
SCH Library 🔻 🖡 🗙	😤 диоды.SchLib 🧟 Mis	scellaneous Devices.SchLib		Fav
Components / Description		1K		Cipbo
Place Add Delete Edit Pins Name Type AXIA RESL Res	· ·			ubraries
Add Delete Edit	Editor	• Окно		Mask Level Clear
Model / Type Description	Model / Type AXIAL-0.4 Footprint Res Signal Int RESISTOR Simulation	Loc Description Axial Device, Thru-Hole; 2 Leads; Resistor <u>B</u> emove <u>E</u> dit	o () —	-0
X:0 Y:0 Grid:10			System Design Compiler Help	S⊆H Instruments >>

Рис. 1. Библиотечный компонент Res2 в schlib в Altium Designer

Одно посадочное место может быть подключено к различным УГО. В связи с этим посадочные места тоже важно именовать максимально информативно, чтобы другие разработчики могли подключить готовое посадочное место к своему УГО. Посадочные места в этом случае именуют по типу корпуса, а при именовании важно разобраться в классификации однотипных корпусов.

Корпуса типа SO (SOIC) часто именуют SO8, SO16 и т.д. (исходя из количества ножек в корпусе), хотя такие корпуса отличаются еще и шириной. Существуют три распространенные ширины – 150, 208 и 300 тысячных дюйма (3,81; 5,28 и 7,6 мм). Обычно они обозначаются SOxx-150, SOxx-208 и SOxx-300. В библиотеке их рекомендуется называть SO8-150, SO8-208 и т.п.

Рекомендуется при разработке каждого посадочного места изучить классификацию однотипных (подобных) корпусов и присвоить посадочному месту наиболее информативное наименование (в данном случае будет лучше, если наименование будет длиннее, чем неверная интерпретация наименования и впоследствии, – ошибка в плате).

🤏 Altium Designer (13.0) - D:\Eva\ИТМО\Публикации\2013 Вестник\Jib\диоды.SchLib - Диоды.LibPkg. Not signed in.	-	
JZP Eile Edit View Project Place Tools Reports Window Help	D:\Eva\NTMO\D	убликации\2013 🔹 🔘 - 💮 - 🕐
D 22 21 2 02 2 03 02 2 03 03 1 + x x 1 0 0 Mode - + - + + 1	8•⊠•⊪•	<u>a</u>
SCH Library 🔻 🕈 😤 диоды. SchLib		Fav
		orites
Components / Description	I	Cipboan
Place Add Delete Edit Pins Name Type To-220AC		
		Mask Level Clear
Add Delete Edit Model / Type Location Description TO-220AC Footprint TO, Thru-Hole, Vertical, Heat Add Delete Edit		
Files Projects Navigator SCH Library Sc Add Footprint <u>Remove</u> <u>Edit</u>	A -	
X:-7.5mm Y:2.5mm Grid:2.5mm System	Design Compiler	Help SCH Instruments >>

Рис. 2. Библиотечный компонент КД270А в schlib в Altium Designer

Рекомендуемые подходы к именованию составляющих библиотек электронных компонентов позволяют централизованно использовать единые библиотеки (резисторов, конденсаторов, диодов, микросхем и т.д.) в рамках одного предприятия. Применение рекомендуемых авторами подходов позволит автоматизированно создавать перечни элементов к схемам и спецификации на платы. Предложенные в рамках работы подходы включены в курс обучения «Altium Designer – PCB Профессиональный» в компании «НИП-Информатика» с августа 2012 года и успешно реализуются на предприятиях, сотрудники которых окончили данный курс.

[Л]. Романова Е.Б. Создание библиотеки электронных компонентов в соответствии с ГОСТ. Молодые ученые – промышленности Северо-Западного региона // Материалы конференций политехнического симпозиума. Декабрь 2006 года. – СПб: Изд-во Политехн. ун-та, 2006. – С. 62–63.

Кузнецова Ольга Валерьевна – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, olunchik 1989@mail.ru

Романова Ева Борисовна – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат технических наук, доцент, eva_rom@mail.ru

УДК 004.42

ПРИМЕНЕНИЕ ГИПЕРСПИСКОВ В РАЗРАБОТКЕ ВЕБ-ПОРТАЛА И.А. Бессмертный, Е.П. Пономарева

Описывается опыт применения механизма гиперсписков инструментальной платформы MS Sharepoint для построения корпоративного веб-портала. Особенностью предлагаемого подхода является хранение в гиперсписках не только контента веб-портала, но и конфигурационных данных, что существенно упрощает сопровождение веб-портала. Ключевые слова: гиперсписки, веб-портал, MS SharePoint.

В течение жизненного цикла любого веб-портала существенная часть затрат требуется для его сопровождения как в части информационного наполнения, так и конфигурирования. В этой связи целесообразно использовать инструментальные средства, максимально стандартизующие указанные манипуляции. Одним из таких инструментальных средств является MS SharePoint – механизм гиперсписков для хранения данных. Применение гиперсписков позволяет нескольким пользователям совместно редактировать данные, отслеживать версии и идентифицировать авторов изменений, а также создавать различные представления для просмотра подмножеств всех данных. В работе описывается опыт создания корпоративного портала с использованием гиперссылок для хранения не только данных информационного наполнения, но и параметров конфигурации.

Гиперсписок – это хранилище данных, которое построено в виде списка элементов с определенным набором полей, причем поле может ссылаться на элемент того же или любого другого списка. Перечислим преимуществом использования гиперсписков MS SharePoint для хранения данных веб-портала:

- гиперсписки позволяют пользователю манипулировать данными без вмешательства разработчиков;
- MS SharePoint предоставляет удобные средства настройки различных представлений гиперсписков (например, отображение только элементов, автором которых является текущий пользователь);
- в механизме гиперсписков имеется интерфейс разделения прав доступа;
- имеются серверная и клиентская модели MS SharePoint, которые упрощают доступ к данным.

На рисунке приведен фрагмент разрабатываемого на платформе MS SharePoint веб-портала, а именно гиперсписок «Базы данных» (рисунок, б), который хранит все необходимые данные для подключения к внешним базам данных, а также пример кода, в котором происходит извлечение данных из списка для генерации строки подключения к базе данных (рисунок, а).

Извлечение данных из гиперсписка	Гиперсписок «Базы данных»
<pre>private string GetConnectionString(SPWeb web, string listUrl,</pre>	
<pre>SPList list = web.GetListExt(listUrl);</pre>	Базы данных
SPQuery query = new SPQuery(); query.Query = @" <where></where>	Name User Id Password Provider Data Source
<eq> <pre> <fieldref name="Name"></fieldref> <value type="Text">" +</value></pre></eq>	AD admin ****** Microsoft.Jet.OLEDB.4.0 c:\Data\AD.mdb
<pre>db_name + @ </pre>	LD admin ****** Microsoft.Jet.OLEDB.4.0 c:\Data\LDb.mdb
<pre>string connString = String.Empty; if (items != null && items.Count > 0) {</pre>	Ф Добавить элемент
<pre>connString = String.Format(@"Provider={0};Data Source={1}; User Id={2};Password={3};", items[0]["Provider"], items[0]["DataSource"], items[0]["UserId"], items[0]["Password"]);</pre>	
ر return connString;	
ĵ	
а	б

Рисунок. Фрагмент веб-портала. Извлечение данных из гиперсписка (а); гиперсписок «Базы данных» (б)

Элемент гиперсписка «Базы данных» хранит всю необходимую информацию для подключения к конкретной базе данных. Администратор при помощи удобного пользовательского интерфейса может легко отредактировать данные, добавить подключение к новой базе данных и (или) удалить старое подключение. Подключение к базам данных было необходимо, так как перед авторами стояла задача синхронизации пользователей из нескольких баз данных с профилями пользователей MS SharePoint. Синхронизацию выполняет Timer Job (класс, наследуемый от SPJobDefinition), который запускается раз в неделю. Администратор веб-портала может исправить график работы Timer Job в случае необходимости. Каждый пользователь может добавлять или изменять строки данных в гиперсписке. Если это свойство может быть полезным в базе данных веб-портала, то для хранения конфигурационных параметров оно совершенно недопустимо. Самое простое решение данной проблемы состоит в том, чтобы сделать гиперсписок с конфигурационными данными скрытым. Однако, зная URL, можно получить к этому списку доступ. По этой причине здесь следует использовать специальные классы объектной модели MS SharePoint, такие как SPPermission, SPRoleDefinition и т.д. Необходимо создать группу для администраторов веб-портала и этой группе дать доступ к гиперсписку; для остальных же пользователей, не входящих в эту группу, закрыть доступ на добавление / редактирование / удаление / просмотр элементов гиперсписка. После этого можно быть уверенным, что необходимый уровень безопасности будет достигнут.

Опыт использования указанных инструментальных средств показывает, что использование гиперсписков в среде MS SharePoint позволяет существенно, не менее чем в два раза сократить издержки на сопровождение крупных веб-порталов.

Бессмертный Игорь Александрович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат технических наук, доцент, Igor_bessmertny@hotmail.com Пономарева Екатерина Павловна – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, студент, katerin.ponomareva@gmail.com

REVIEW ARTICLE

MODERN STATE AND DEVELOPMENT PERSPECTIVES OF FOREIGN INFRARED IMAGERS V.Tarasov, Y. Yakushenkov

Short information about the foreign infrared imagers (II) development programs and their element base (foremost, about the optical systems and focal plane arrays – FPA) is stated. Main attention is devoted to the third generation of II working in two or more spectral bands with FPA. Some examples of military, civil and «dual assignment» of II are cited. Perspectives of II development during the nearest years are considered. **Keywords**: infrared imager, optical system, focal plane array.

Victor Tarasov – Moscow State University of Geodesy and Cartography, Professor, D.Sc.; Central R&D Institute «Cyclone», General director, cyclone@asvt.ru

Yuri Yakushenkov – Moscow State University of Geodesy and Cartography, Department head, D.Sc., Professor, yakush@miigaik.ru

OPTICAL AND OPTICAL ELECTRONIC SYSTEMS. OPTICAL TECHNOLOGIES CORRECTION METHOD OF FIELD CURVATURE IN WIDE-ANGLE OBJECTIVE LENS V. Bezrukov, G. Karpova

The article deals with the method of field curvature correction in wide-angle objectives and technique of its implementation. The correctional elements were used with definite values of longitudinal magnification along the chief ray to solve the problem of correcting components synthesis. Design parameters of these elements were related mathematically with its real aberration, in particular, the aberration of astigmatism.

Keywords: wide-angle objectives, field curvature, astigmatism, correcting lens.

Vyacheslav Bezrukov – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Associate professor, PhD, Associate professor, bezrukov@mail.ifmo.ru

Galina Karpova – Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Associate professor, PhD, Associate professor, Karpova3101@mail.ru

METHOD OF ANGULAR SELECTIVE REGULATION FOR DIRECTIONAL LIGHT TRANSMISSION

R. Zakirullin

The article deals with a method of angular selective regulation for directional light transmission, when a light source moves with respect to the glazed object. An optical filter with grating layers on the two surfaces of the glass substrate provides angular regulation by the relative position of gratings. Both gratings consist of absorptive, scattering or reflective strips, alternating with strips of raw surface of the filter glass. A graphic-analytical method for angular characteristics calculating of light transmission for the plane shape filter is described. At a pre-known trajectory of the light source the light transmission characteristics can be pre-adapted to the change in the incidence angles. The graphic-analytical calculation results are confirmed experimentally. The regulation method potential in architectural glazing, in optical and lighting devices is considered.

Keywords: directional light transmission, optical filter, grating with alternating strips, graphic-analytical calculation, angular selective regulation.

Rustam Zakirullin – Orenburg State University, Associate professor, PhD, Associate professor, rustam.zakirullin@gmail.com

CONTINUOUS SPECTRA RESTORATION BY AN ADAPTIVE WAY OF COMPUTING EXPERIMENTS WITH REGULARIZATION V. Sizikov, A. Krivykh

The article deals with inverse spectroscopy problem - restoration of continuous spectra by mathematical treatment of measured spectra, distorted by the spectrometer instrument function and noises. The problem is reduced to solution of the first-kind Fredholm integral equation. This problem is illegal; therefore, the Tikhonov regularization method is used for stable numerical solving of the equation. Furthermore, the adaptive way of computing experiments is used, according to which, one treats the model spectrum Q with prescribed true spectrum z and simulated measured spectrum u with regard to additional (a priori) information about the true spectrum P. This makes it possible to choose the regularization parameter α . The proposed technique can be used to enhance the spectrometer resolving power. Numerical illustrations are given.

Keywords: continuous spectrum, inverse spectroscopy problem, integral equation, Tikhonov regularization method, way of computing experiments, spectrometer resolving power enhancement.