

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ
ФЕДЕРАЦИИ**

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ

**САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ**

НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК

Выпуск 42

**ФУНДАМЕНТАЛЬНЫЕ
И ПРИКЛАДНЫЕ ИССЛЕДОВАНИЯ
ИНФОРМАЦИОННЫХ СИСТЕМ
И ТЕХНОЛОГИЙ**



**САНКТ-ПЕТЕРБУРГ
2007**

Выпуск содержит материалы XXXVI научной и учебно-методической конференции СПбГУ ИТМО «Достижения ученых, аспирантов и студентов университета в науке и образовании». Конференция была проведена 30 января - 2 февраля 2007 года Санкт-Петербургским государственным университетом информационных технологий, механики и оптики в сотрудничестве с

- ФГУП ВНЦ «ГОИ им. С.И. Вавилова»
- ФГУП НПК «ГОИ им. С.И. Вавилова»
- ФГУП НИТИОМ «ГОИ им. С.И. Вавилова»
- Институтом аналитического приборостроения РАН (ИАНП РАН)
- Институтом проблем машиноведения РАН (ИПМаш РАН)
- Комитетом по науке и высшей школе Санкт-Петербурга (КНВШ)
- ФГУП СПб ОКБ «Электроавтоматика»
- ВНИИМ им. Д.И. Менделеева
- ОАО «ЛОМО»
- ОАО «Техприбор»
- ЦНИИ «Электроприбор»
- ЗАО «БИ ПИТРОН»

В выпуске представлены работы, поддержанные финансированием в рамках:

- Федеральной целевой научно-технической программы «Исследования и разработки по приоритетным направлениям развития науки и техники на 2002–2006 гг» (Федеральное агентство по науке и инновациям);
 - аналитической ведомственной целевой программы «Развитие научного потенциала высшей школы (2006–08 гг.) (Федеральное агентство по образованию);
 - Федеральной целевой программы развития образования на 2006–2010 гг. (Федеральное агентство по образованию),
 - Российского фонда фундаментальных исследований,
- а также инициативные разработки.

Программный комитет конференции:

Васильев В.Н. (СПбГУ ИТМО) – председатель

Аронов А.М. (ОАО «ЛОМО»)	Маслов Ю.В. (ОАО «Техприбор»)
Викторов А.Д. (КНВШ)	Максимов А.С. (КНВШ)
Гатчин Ю.А. (СПбГУ ИТМО)	Медунецкий В.М. (СПбГУ ИТМО)
Дукельский К.В. (ФГУП ВНЦ «ГОИ им. С.И. Вавилова»)	Немолочнов О.Ф. (СПбГУ ИТМО)
Жигулин Г.П. – начальник ИКВО	Никифоров В.О. (СПбГУ ИТМО)
Иванов А.Ю. (СПбГУ ИТМО)	Парамонов П.П. (ФГУП СПб ОКБ «Электроавтоматика»)
Иванов А.В. (СПбГУ ИТМО)	Парфенов В.Г. (СПбГУ ИТМО)
Карасев В.Б. (ФГУП ВНЦ «ГОИ им. С.И. Вавилова»)	Пешехонов В.Г. (ЦНИИ «Электроприбор»)
Козлов С.А. (СПбГУ ИТМО)	Семенов А.Н. (СПбГУ ИТМО)
Колесников Ю.Л. (СПбГУ ИТМО)	Смирнов С.Б. (СПбГУ ИТМО)
Курочкин В.Е. (ИАНП РАН)	Стафеев С.К. (СПбГУ ИТМО)
Латыев С.М. (СПбГУ ИТМО)	Ткалич В.Л. (СПбГУ ИТМО)
Лукьянов Г.Н. (СПбГУ ИТМО)	Ханов Н.И. (ВНИИМ им. Д.И. Менделеева)
	Шехонин А.А. (СПбГУ ИТМО)
	Яблочников Е.И. (СПбГУ ИТМО)

Организационный комитет конференции:

Никифоров В.О. – председатель, Студеникин Л.М. – зам. председателя,
Казар Л.Н. – ученый секретарь, Горкина Н.М. Гусарова Н.Ф. Савельева Л.П.

ISSN 1819-222X

© Санкт-Петербургский государственный университет
информационных технологий, механики и оптики, 2007

ПЛЕНАРНОЕ ЗАСЕДАНИЕ

СОВРЕМЕННОЕ СОСТОЯНИЕ И ПЕРСПЕКТИВЫ РАЗВИТИЯ ТОМОГРАФИИ

М.Я. Марусина, А.О. Казначеева

В докладе представлены различные виды томографических технологий, их аналитические возможности, достоинства и недостатки. Подробно рассмотрены основные виды томографии: рентгеновская компьютерная (РКТ), магнитно-резонансная (МРТ), позитронно-эмиссионная (ПЭТ) томография, ПЭТ/КТ сканеры. Проведена классификация томографических комплексов по назначению. Проанализированы тенденции развития томографических комплексов. Приведены результаты комплексного мониторинга рынка томографических технологий в Санкт-Петербурге.

Введение

Мировые тенденции в развитии медицинского приборостроения в последние годы претерпели значительные изменения. Это связано с появлением качественно новых требований к уровню медицинской диагностики, которые продиктованы необходимостью раннего обнаружения различных заболеваний. В настоящее время состояние медицинской техники позволяет визуализировать структурные и функциональные изменения одного и того же объекта с помощью устройств, имеющих различные физические принципы действия.

На современном этапе одним из наиболее информативных методов является томография (послойное исследование структуры различных объектов), дающая намного больше информации о каждом элементарном объеме объекта, чем другие известные медицинские диагностические методы. В настоящий момент развивается несколько видов томографии: рентгеновская компьютерная (РКТ), магнитно-резонансная (МРТ), позитронно-эмиссионная (ПЭТ), электронно-лучевая, ультразвуковая, оптическая когерентная томография и др. Но суть всех способов одина: по суммарной информации (в РКТ это интенсивность сигналов на детекторах, в МРТ это интенсивность эхосигналов), полученной от некоторого сечения (слоя) объекта, нужно определить локальную информацию, а именно плотность вещества в каждой точке сечения. В разных видах томографии и суммарная информация, и математическое описание (в РКТ это интегральные уравнения Радона или Фредгольма, в МРТ это двухмерное преобразование Фурье) качественно различны, хотя существуют попытки описать различные способы получения изображений единым, так называемым основным уравнением томографии [1, 14, 15].

Качество томографических изображений зависит от целого ряда факторов: физических процессов, на основе которых проводится сбор данных; аппаратного и программного обеспечения; параметров сканирования; свойств исследуемого вещества (плотности, процессов релаксации, диффузии); различных влияющих факторов, например, температуры окружающей среды, перепады которой приводят к повышению уровня шума на РКТ-изображениях и снижают их качество и т.д. [1,2].

Одно из направлений развития кафедры измерительных технологий и компьютерной томографии – научно-исследовательская работа в области разработки и совершенствования измерительных томографических комплексов.

Основные виды томографии Компьютерная томография

Из всех существующих томографических методов особые успехи в настоящее время достигнуты в рентгеновской компьютерной томографии. Предпосылкой ее появления послужили недостатки обычной рентгенографии, породившие идею получения не одного, а ряда снимков, выполненных под разными углами, и определения путем математической обработки плотностей исследуемого вещества в ряде сечений.

Выходные данные КТ-сканера даются в КТ-числах или единицах Хаунсфилда. Компьютерная обработка изображения позволяет различать более ста степеней изменения плотности исследуемых тканей: от нуля – для воды, до ста и более – для костей. Это дает возможность дифференцировать различия нормальных и патологических участков тканей в пределах 0,5–1%, т.е. в 20–30 раз больше, чем на обычных рентгенограммах.

Основным достоинством РКТ является небольшая продолжительность исследования (всего 1–5 мин) при достаточно высоком пространственном разрешении изображения (до сотых долей мм). К достоинствам можно также отнести: возможность построения качественных 3D реконструкций объекта; низкие эксплуатационные затраты РКТ; оперативность; возможность исследования промышленных объектов

Основные недостатки РКТ: наличие рентгеновского излучения; получение срезов только в поперечной плоскости; присутствие на томограммах артефактов от металла; влияние температуры окружающей среды; необходимость калибровки аппарата из-за дрейфа КТ-чисел.

Современные многосрезовые РКТ-сканеры имеют до 64 рядов детекторов и обеспечивают высокое изотропное разрешение изображений. Например, томограф Somatom Sensation 64-slice (“Siemens”) позволяет проводить исследования с изотропным разрешением 0,24 мм, при этом время одного оборота трубки составляет 0,33 с, а скорость движения стола – 87 мм/с. Подобные системы наиболее часто используются в кардиологии, пульмонологии и при исследованиях сосудистой системы [3, 4].

Существенно повысить информативность томографических данных можно, применив различные методы трехмерной реконструкции, дающие возможность рассмотреть участки исследуемого объекта под произвольным углом (рис. 1).

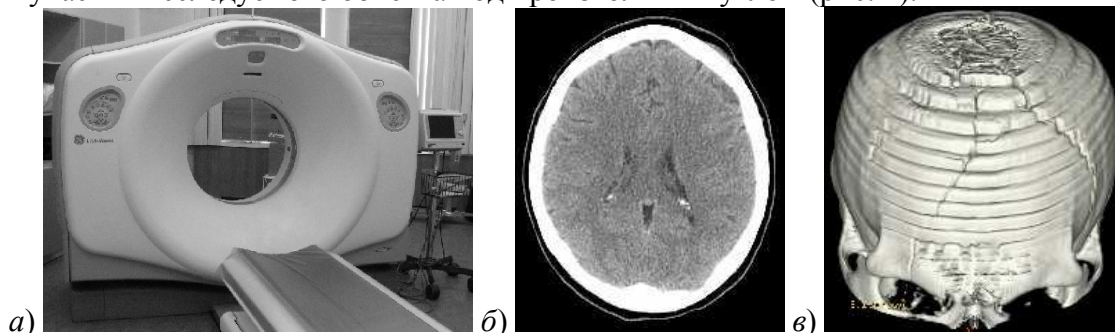


Рис. 1. Компьютерный томограф (а), КТ-изображение (б) и трехмерная реконструкция (в)

Для получения изображений внутренних полостей (например, сосудов) в перспективе, а иногда и для отображения областей, не доступных обычной эндоскопии (например, цистерны мозга), используется виртуальная эндоскопия. В этом режиме интересные полости отображаются с помощью объемного представления в перспективе, что дает ощущение «полета» через отображаемую область.

Использование различных методов реконструкции позволяет существенно повысить информативность полученных данных, в том числе за счет наглядности пространственного расположения исследуемых тканей.

Магнитно-резонансная томография

Магнитно-резонансная томография (МРТ) – это метод отображения внутренней структуры объектов, основанный на явлении ядерного магнитного резонанса (ЯМР) и используемый преимущественно для медицинских визуализаций.

Основные достоинства МРТ: более высокая разрешающая способность, контрастность и большие возможности по дифференциации тканей по сравнению с РКТ; возможность получения изображений в произвольной плоскости; возможность осуществления функциональных исследований; относительная безвредность (отсутствие лучевого воздействия на пациента); получение изображений сосудов; определение спектрального состава веществ.

Основные недостатки МРТ: достаточно высокая продолжительность исследования (15–30 мин.); присутствие на томограммах значительных артефактов от металла и от движения объекта; появление ложных изображений; наличие ряда противопоказаний для проведения исследований; невозможность проведения МР-исследований промышленных объектов.

По типу источника основного магнитного поля МР-томографы разделяются на постоянные, резистивные, сверхпроводящие и гибридные системы. В зависимости от напряженности основного магнитного поля МР-томографы классифицируются на: сверхнизкие (менее 0,1 Тл); низкочастотные (0,1–0,4 Тл); среднечастотные (0,5 Тл); высокочастотные (1–2 Тл); сверхвысокочастотные (свыше 2 Тл). По виду конструкции МР-томографы бывают открытые и закрытые.

Низкочастотные МРТ обычно имеют резистивные или постоянные магниты. Их преимущество заключается в меньшем количестве противопоказаний для пациентов и персонала. Иногда низкочастотные системы имеют специализированную область применения, например, только для исследования конечностей, или используются как открытые томографы. Недостатками таких систем являются низкое соотношение сигнал/шум и большое время сканирования, необходимое для получения изображения хорошего качества.

Оптимальная величина напряженности поля для клинического отображения лежит в диапазоне от 0,5 до 2,0 Тл. Свыше верхнего предела 2,0 Тл, магнитные поля потенциально опасны для пациентов и могут использоваться только в исследовательских лабораториях. Вопрос об оптимальной напряженности поля – предмет постоянной дискуссии специалистов.

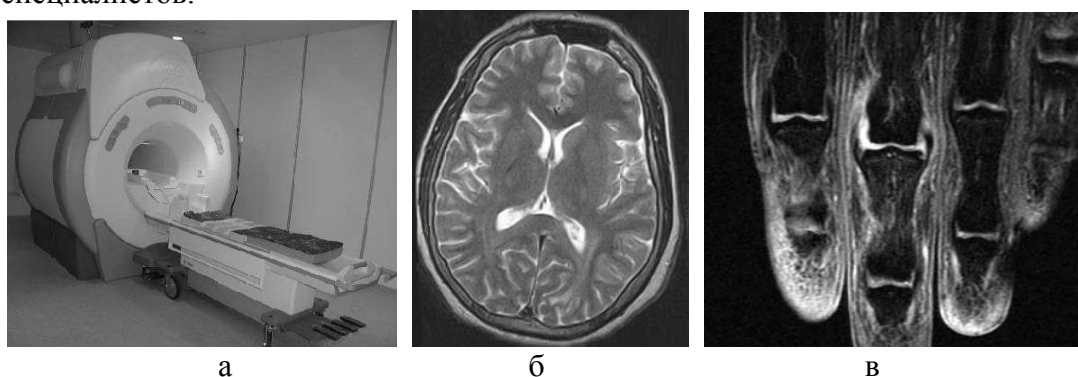


Рис. 2. Магнитно-резонансный томограф (а), МР-изображение головного мозга (б) и кисти (в)

Более 90% парка МР-томографов составляют модели со сверхпроводящими магнитами. В середине 80-х гг. фирмы-производители делали упор на выпуск моделей с полем 1,5 Тл и выше, но уже через несколько лет стало ясно, что в большинстве областей применения они не имеют существенных преимуществ перед моделями со средней индукцией поля. Поэтому сейчас основные производители МР-томографов уделяют

особое внимание выпуску моделей со средним и низким полем, отличающихся компактностью и экономичностью при удовлетворительном качестве изображений и меньшей стоимости. Высокопольные системы используются преимущественно в научно-исследовательских центрах [6].

Совершенствование МР-томографов и создание высокопольных систем сделало возможным проведение в клинических условиях магнитно-резонансной спектроскопии (МРС), позволяющей с высокой точностью измерять спектры различных химических элементов. Благодаря этому МРС дает возможность проследить за региональными метаболическими изменениями.

Позитронно-эмиссионная томография

В клинической практике ПЭТ используется с начала 1990-х гг. и в отличие от классических методов, таких, как РКТ и МРТ, позволяющих получать только изображения анатомических структур и выявлять изменения в них, ПЭТ допускает проведение количественного анализа биохимических или физиологических функций. Эта информация часто позволяет выявить функциональные изменения, вызванные заболеванием, задолго до появления каких-либо морфологических изменений. В частности, ПЭТ – единственный метод, позволяющий диагностировать опухоль на "нулевой" стадии.

Методика ПЭТ-отображения является комбинацией двух изобретений, удостоенных Нобелевской премии – радиоактивного индикатора и принципов томографии.

Получение изображения базируется на внешнем обнаружении совпадающих испускаемых объектом γ -лучей в течение 10 нс между двумя датчиками, расположенными на противоположных сторонах сканера. Реконструкция изображений производится путем обработки данных, полученных детекторами, расположенными в виде нескольких колец, что позволяет одновременно получать изображения до пятнадцати срезов. Отличительной особенностью ПЭТ-изображений является их представление в цветной шкале. Цвет или степень яркости каждого пикселя пропорциональны концентрации изотопа в соответствующей точке объекта, что создает более наглядную картину [5].

Дальнейшее усовершенствование ПЭТ-сканеров состоит в повышении пространственного разрешения, чувствительности детекторов, увеличении числа одновременно получаемых срезов и разработке новых алгоритмов реконструкции изображений.

Хотя в последние 20 лет ПЭТ использовалась преимущественно для научных исследований, ее роль как метода диагностики находится на стадии становления. В настоящее время основными областями клинического применения ПЭТ являются онкология, кардиология и неврология. Также ПЭТ используется при изучении механизмов мозговой деятельности, лежащей в основе человеческого сознания.

Ценность ПЭТ возрастает при совместном проведении различных томографических исследований, позволяющих получить привязку к анатомическим структурам (например, КТ, МРТ).

ПЭТ/КТ-сканеры

Сравнительно недавно появились новые диагностические устройства – ПЭТ/КТ сканеры. Они позволяют в одном исследовании получать функциональную (ПЭТ) и структурную (КТ) информацию, что достигается совмещением ПЭТ- и КТ-изображений. Кроме того, данные КТ могут использоваться для коррекции аттенуации ПЭТ-изображений, что сокращает время исследования на 30–40%. Эти особенности делают ПЭТ/КТ более предпочтительным методом, чем обычное ПЭТ-исследование. Недостатком метода является возникновение артефактов и количественных ошибок на

ПЭТ-изображениях при коррекции аттенюации на основе данных КТ. Например, использование контрастирующих веществ и наличие металлических имплантатов может привести к переоценке активности препарата в исследуемой области. Дыхание пациента приводит к несоответствию данных КТ- и ПЭТ-исследований (разные поля сбора данных) и артефактам [5].

Конструктивно ПЭТ/КТ-сканер представляет собой позитронно-эмиссионный и компьютерный томографы, размещенные в одном гентри (обычно КТ-блок расположен перед ПЭТ-блоком). Современные ПЭТ/КТ сканеры имеют апертуру гентри 70 см и ширину 100 см. Оба сканера могут использоваться как вместе, так и независимо друг от друга. КТ-сканер может быть многосрезовым с пошаговым или спиральным режимами сбора данных и различными скоростями вращения трубки, ПЭТ-сканер может работать в двух- или трехмерном режимах.

Первым этапом сбора данных является получение базового изображения, служащего для задания границ области ПЭТ/КТ-исследования. КТ-изображения получают при напряжении 100-140 кВ и различных силах тока (в зависимости от цели исследования). Затем пациент автоматически перемещается в ПЭТ-сканер и позиционируется в то же анатомическое положение. Сбор данных на этом этапе занимает 3-5 минут, затем ПЭТ-изображения реконструируются с учетом коррекции на аттенюацию по КТ-изображениям. Полученные КТ-изображения имеют матрицу 512×512, а ПЭТ-изображения – матрицу 128×128.

Одновременное получение функциональной и анатомической информации позволяет повысить качество исследования за счет улучшения качества изображения и повышения точности определения пространственной локализации образований до 98%. Анализ спроса на медицинскую технику показывает увеличение продаж ПЭТ/КТ-сканеров на 65% за последние 3 года, что вызвано, прежде всего, преимуществами данного метода.

Все крупные производители медицинского диагностического оборудования разработали и выпускают позитронно-эмиссионные томографы, комбинированные с компьютерными томографами. Другие виды томографов (лазерные, ультрозвуковые) пока не столь широко представлены на рынке.

Если в первые годы появления томографических установок реализуемые алгоритмы давали возможность проведения узкого спектра диагностических процедур, то развитие аппаратного и программного обеспечения позволило реализовать более сложные и информативные исследования, подчас не реализуемые другими методами неинвазивной диагностики.

Диаграмма, построенная по данным фирм-производителей томографической аппаратуры, иллюстрирует экспоненциальный рост количества томографических аппаратов в мире. За последние 5 лет количество сканеров в мире увеличилось примерно в 2 раза (рис. 3).

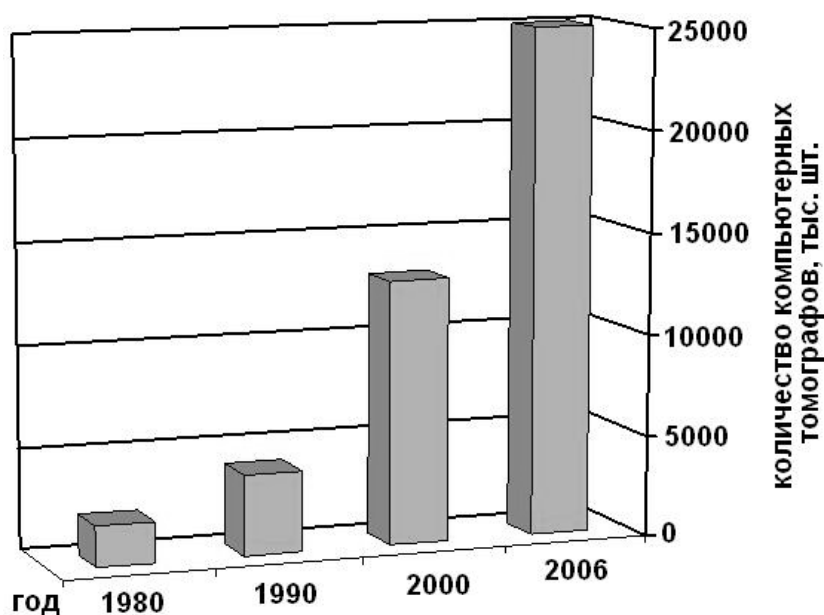


Рис. 3. Примерная диаграмма развития компьютерной томографии

Классификация медицинских томографов по назначению

Безусловно, основной областью применения томографии является медицинская диагностика. При этом в зависимости от назначения томографического комплекса можно выделить:

- томографы общего назначения;
- специализированная область диагностики (ортопедия, педиатрия, кардиология);
- контроль хирургических операций;
- динамические исследования;
- мобильные томографы.

Большинство современных разрабатываемых медицинских МРТ используют для формирования основного поля сверхпроводящие магниты. Подобные системы позволяют реализовать ряд преимуществ: достичь высокой однородности основного поля, высокого соотношения сигнал/шум, реализовывать сложные методы получения изображения (спектроскопия, диффузия, тензор диффузии, функциональные исследования). Как правило, такие системы имеют закрытую конфигурацию. Последние научные разработки в данной области направлены на создание высокопольных (свыше 0,5 Тл) открытых сверхпроводящих МР-систем.

Другим направлением развития является создание узкоспециализированных (относительно дешевых) МР-систем, позволяющих проводить исследования отдельных анатомических областей, например, в области ортопедических исследований. Такие томографы позволяют проводить исследования с высоким соотношением сигнал/шум и пространственным разрешением, при этом требования к помещению и эксплуатационные расходы малы. Существуют как сверхпроводящие, так и постоянные магнитные системы для исследований суставов конечностей. Вес таких систем составляет от 600 кг.

Одним из наиболее известных производителей в данной отрасли является итальянская фирма ESAOTE (рис. 4, а), выпускающая линейку МР-томографов для исследований суставов (C-scan, E-scan, G-scan, Artroscan-M).

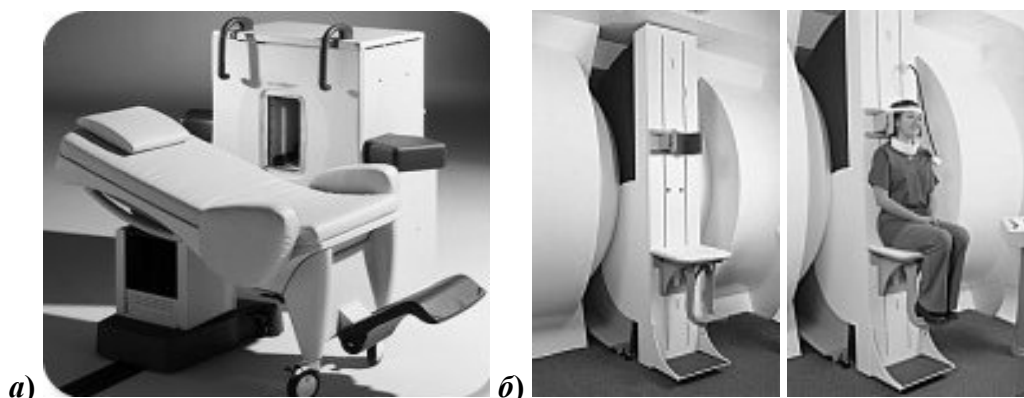


Рис. 4. МР-томографы Artroskan-M (ESAOTE-GE) (а) и UpRight (Fonar, США) (б)

Для ряда медицинских исследований бывает необходимо получить томограммы одного и того же среза в различных положениях пациента. Такие динамические исследования (исследования с функциональными пробами), часто требуются при диагностике заболеваний суставов и позвоночника. Лидером в разработке и производстве данной категории МРТ-сканеров является американская компания FONAR, выпускающая модель UpRight (рис. 4, б). Эта система представляет собой открытый резистивный магнит, позволяющий проводить томографию при различных положениях пациента, и часто является незаменимой при различных заболеваниях, сопровождающихся болевыми ощущениями.

В отдельную группу можно выделить томографы, используемые для контроля проведения хирургических операций. Как правило, это КТ или МРТ сканеры, передвигаемые по специальным направляющим между помещениями операционного блока, и позволяющие проводить диагностику в операционной. На рис. 5 представлен МР-томограф iMotion (IMRIS, США), созданный на базе сверхпроводящего магнита фирмы Siemens и один из вариантов схемы его перемещения между операционными.

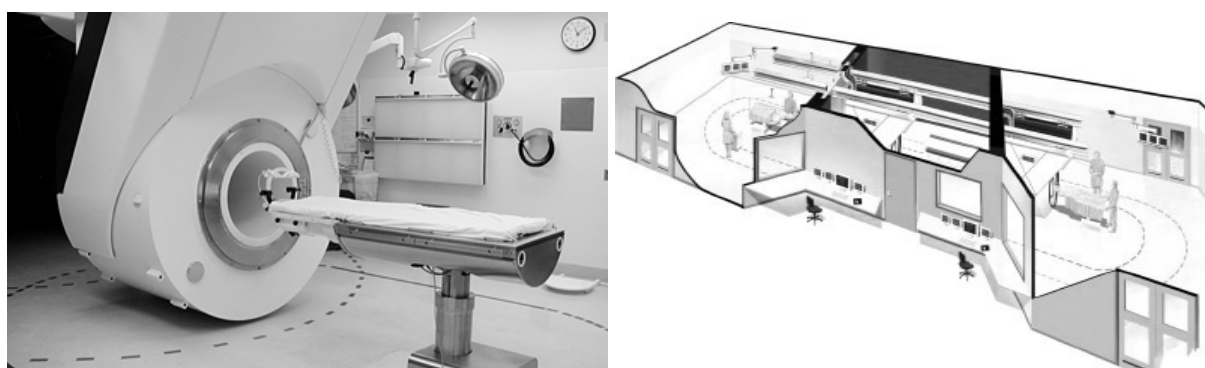


Рис. 5. МР-томограф iMotion (IMRIS, США) и схема помещений оперблока на две операционные

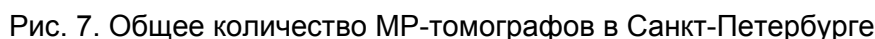
Томограф закрепляется на направляющих под потолком операционной. В случае необходимости проведения томографии, сканер перемещается по этим направляющим. При этом положение пациента на операционном столе не меняется, а сам операционный стол со встроенными элементами РЧ-катушки оказывается внутри туннеля магнита.

Сложность разработки таких систем связана с проблемами точности позиционирования томографа и высокими требованиями, предъявляемыми к данному параметру.

Необходимость проведения исследований в небольших населенных пунктах привела к разработке мобильных томографических комплексов, размещаемых в трейлерах. Подобные комплексы не уступают стационарным установкам по информативности и

Одной из областей применения рентгеновской компьютерной томографии является неразрушающий контроль промышленных объектов, где она является одним из методов контроля структуры таких изделий техники как турбинные лопатки, лопасти вертолетов, ракетные клапаны, аккумуляторы, электродвигатели и т.д.

Проведенный комплексный мониторинг томографических технологий в мире и в Санкт-Петербурге показывает рост общего количества томографических установок в различных областях диагностики (рис. 7).



В медицинских учреждениях Санкт-Петербурга наиболее широко применяются томографы фирм GE и Siemens (рис. 8). Томографы отечественных фирм, производя-

щих МРТ, таких как ТОРОСС (Москва), НИИЭФА (СПб), ВНИИКТ (Москва), занимают менее 5% от общего числа томографов в Санкт-Петербурге.

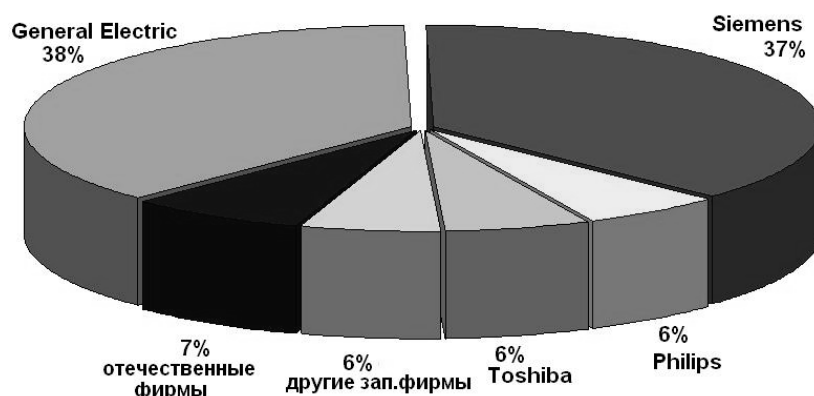


Рис. 8. Производители томографов на рынке Санкт-Петербурга

Современные требования, предъявляемые к уровню медицинской диагностики, таковы, что любое крупное медицинское учреждение, больница или диагностический центр должны быть оснащены томографическими установками. С этой точки зрения была проанализирована оснащенность томографами медицинских учреждений (клиник, больниц и поликлиник) Санкт-Петербурга. Лишь 10% из них имеют хотя бы 1 установку.

В Санкт-Петербурге за последний год наблюдается увеличение количества медицинских учреждений, которые приобрели томографическую аппаратуру. Если учесть, что томографами оснащено только 10% учреждений СПб, то можно спрогнозировать значительное увеличение темпов этого роста. Только в 2007 году в Санкт-Петербурге планируется установка не менее 20 различных томографов.

Научно-исследовательская работа кафедры ИТиКТ в области неинвазивных технологий

Кафедра измерительных технологий и компьютерной томографии (ИТиКТ) на настоящий момент имеет связи более чем с десятью медицинскими учреждениями Санкт-Петербурга, обладающими разноплановой томографической техникой, с которой работают студенты и выпускники кафедры. Потребности рынка в кадрах по эксплуатации и ремонту аппаратуры очевидны и растут с каждым годом.

На кафедре ведется научно-исследовательская работа по ряду направлений в области неинвазивных технологий:

- разработка и создание ЯМР-томографа [7,8,13];
- расчет и проектирование электромагнитных систем МР-томографов [7];
- совмещение мультимодальных данных [10, 11];
- теоретико-групповые методы коррекции пространственных искажений при совмещении изображений [12,16];
- исследования с высоким пространственным разрешением [2,10,11];
- функциональные исследования [1];
- стереотаксическая нейрохирургия [1].

Работы в области совмещения мультимодальных данных связаны с разработкой программного обеспечения, позволяющего объединить достоинства функциональных и анатомических исследований, т.е. совместить различные томограммы – функциональную картину без дифференциации конкретных анатомических структур (ПЭТ изображения) и анатомическое строение костей (КТ-изображения) и мягких тканей (МР-изображения).

Исследования с высоким пространственным разрешением позволяют более четко (по сравнению с анализом среза макропрепарата) визуализировать структурные изменения (отеки, границы опухолей, сосудистые изменения). Пространственное разрешение в таких исследованиях составляет 0,3 мм, в повседневной медицинской практике – 1 мм.

МРТ позволяет проводить исследования диффузных процессов, перфузии жидкости, спектральный анализ. Получив диффузно-взвешенные изображения в разные моменты времени, можно рассчитать коэффициент диффузии и выявить области, функциональная активность которых снижена.

Заключение. Перспективы развития томографии

Перспективы развития томографии связаны с совершенствованием аппаратуры и алгоритмов обработки и анализа данных. В магнитно-резонансной томографии создание высоких магнитных полей позволит повысить чувствительность метода, а использование мощных градиентов приведет к сокращению времени TE и снижению потерь сигнала, вызванных расфазированием спинов [6–8].

В области материаловедения МРТ можно использовать для исследования новых материалов – стекол, керамики, полимеров, синтетических мембран и сверхпроводников.

В пищевой химии МРТ может использоваться для проверки старения вина и его подлинности, для идентификации масел. Она может внести вклад в неразрушающее исследование механизмов, ответственных за разложение пищи. МР-методы высокого разрешения применяются при анализе сложных смесей различных экстрактов натуральных продуктов

В области клинических исследований развитие МРТ связано с созданием алгоритмов быстрого сбора данных и совершенствованием методов их обработки. МР-микроскопия может применяться при локализации и характеристике метаболитов в биологических жидкостях, в диагностике заболеваний и как неразрушающая процедура для идентификации тканей [9]. Также МР-микроскопию можно использовать для обнаружения микродефектов в пластмассовых трубах или для определения локализации алмазов в породе.

Для химиков, физиков и биологов МР-микроскопия возможно самый мощный инструмент изучения веществ на молекулярном уровне, так как регистрируемые сигналы изменяются в соответствии с ядерной плотностью и свойствами ядерного окружения. Поэтому можно локализовать в 3D объеме магнитные ядра, позволяющие получать изображения и наблюдать объекты с разрешением, достигающим 10^{-6} м.

Перспективы использования компьютерной томографии связаны с совершенствованием аппаратной части КТ-сканера, разработкой новых конструкций, например, КТ-сканеров с двумя рентгеновскими трубками, а также развитие программного обеспечения, позволяющего проводить исследования на функциональном уровне.

В ПЭТ большие перспективы связаны с созданием конструкции сканера, представляющего собой гибрид магнитно-резонансного и позитронно-эмиссионного томографа, что позволит вывести неинвазивную диагностику на качественно новый уровень.

Литература

1. Марусина М.Я. Казначеева А.О. Современные виды томографии: Учебное пособие. – СПб.: СПбГУ ИТМО, 2006. – 152 с.
2. Марусина М.Я., Казначеева А.О. Шумоподавление в томографии с помощью вейвлет-фильтров //Изв. Вузов. Приборостроение. – Т. 49. – №10. – С. 51 – 57.

3. Chang H., Fitzpatrick J.M. A technique for accurate magnetic resonance imaging in the presence of field inhomogeneities // IEEE Transactions on Medical Imaging. – 1995. – P. 152 – 191.
4. Dawson P., Lees W.R. Multi-slice technology in computed tomography // Clinical Radiology. – 2001. – Vol. 56 – P. 302 – 309.
5. Fokas A.S., Iserles A., Marinakis V. Reconstruction algorithms for positron emission tomography and single photon emission computed tomography and their numerical implementation // J. Royal Soc. Interface. – 2006. – №3. – P. 45–54.
6. Spencer S.S., Theodore W.H., Berkovic S.F. Clinical application: MRI, SPECT and PET // Magnetic resonance Imaging. – 1995. – Vol. 13. – №8. – P. 1119 – 1124.
7. Галайдин П.А., Иванов В.А., Марусина М.Я. Расчет и проектирование электромагнитных систем магниторезонансных томографов: Учебное пособие. – СПб: СПбГУ ИТМО, 2004. – 87 с.
8. Иванов В.А., Марусина М.Я., Рущенко Н.Г., Сизиков В.С. Реконструкция МР-изображений с учетом неоднородностей // Научное приборостроение. – 2003. – Т. 13. – № 2. – С. 17 – 21.
9. Индейкин Е.Н. Телемедицина – настоящее и будущее // Главный врач – 1997. – № 3. – С. 11 – 15.
10. Казначеева А.О. Устранение искажений МР-изображений. Сборник статей "Современные технологии" /Под. ред. С.А. Козлова. – СПб: СПбГУ ИТМО, 2003. – С. 140 – 145.
11. Марусина М.Я., Казначеева А.О. Информационная совместимость данных в томографии: Материалы IX международной конференции «Региональная информатика – 2004». – Санкт-Петербург, 2004 – С. 333 – 334.
12. Марусина М.Я., Скалецкая Н.Д., Казначеева А.О. Коррекция пространственных искажений в томографии // Научное приборостроение. – 2005. – Т. 15. – № 3. – С. 77 – 82.
13. Сизиков В.С. Математические методы обработки результатов измерений: Учебник для вузов. – СПб: Политехника, 2001. – 240 с.
14. Тихонов А.Н., Арсенин В.Я., Тимонов А.А. Математические задачи компьютерной томографии. – М.: Наука, 1987. – 160 с.
15. Троицкий И.Н. Статистическая теория томографии. – М.: Радио и связь, 1989. – 240 с.
16. Марусина М.Я. Инвариантный анализ и синтез в моделях с симметриями. – СПб.: СПбГУ ИТМО, 2004. – 144 с.

**ФУНДАМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ
В РАМКАХ ТЕМАТИЧЕСКОГО ПЛАНА
НАУЧНО-ИССЛЕДОВАТЕЛЬСКИХ РАБОТ
УНИВЕРСИТЕТА, ФИНАНСИРУЕМЫХ
ФЕДЕРАЛЬНЫМ АГЕНТСТВОМ
ПО ОБРАЗОВАНИЮ,
И НАУЧНО-ИССЛЕДОВАТЕЛЬСКИХ РАБОТ
ПО КОНТРАКТАМ, ФИНАНСИРУЕМЫМ
ФЕДЕРАЛЬНЫМ АГЕНТСТВОМ ПО НАУКЕ
В 2006 ГОДУ**

**ОПТИМИЗАЦИЯ ИНФОРМАЦИОННЫХ ПОДСИСТЕМ
СЛЕДЯЩИХ ЭЛЕКТРОПРИВОДОВ ПРЕЦИЗИОННЫХ
КОМПЛЕКСОВ СЛЕЖЕНИЯ**

И.Н. Жданов, В.М. Домрачев, К.М. Денисов

В докладе представлены различные виды томографических технологий, их аналитические возможности, достоинства и недостатки. Подробно рассмотрены основные виды томографии: рентгеновская компьютерная (РКТ), магнитно-резонансная (МРТ), позитронно-эмиссионная (ПЭТ) томография, ПЭТ/КТ сканеры. Проведена классификация томографических комплексов по назначению. Проанализированы тенденции развития томографических комплексов. Приведены результаты комплексного мониторинга рынка томографических технологий в Санкт-Петербурге.

При проектировании современных следящих электроприводов систем наведения квантово-оптических систем, к которым предъявляются повышенные требования по статической и динамической точности, целесообразно отказаться от согласующих редукторов, малонадежных исполнительных двигателей постоянного тока и перейти к безредукторным следящим электроприводам на основе бесколлекторных двигателей постоянного тока или вентильных двигателей.

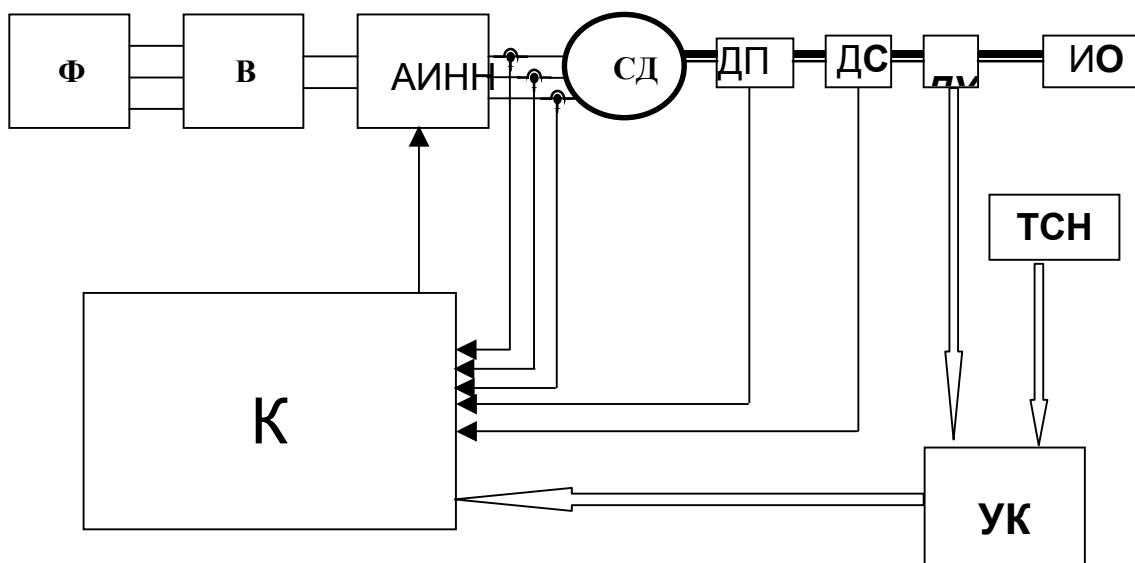


Рис.1. Функциональная схема безредукторного электропривода

Функциональная схема безредукторного электропривода каждой из осей опорно-поворотного устройства (ОПУ) комплекса представлена на рис.1. Структурно электропривод содержит энергетическую и информационную подсистемы. Энергетическая подсистема определяет предельно достижимые динамические характеристики электропривода (максимальные скорости и ускорения следящей оси) и, в значительной степени, массогабаритные показатели систем управления. Информационная подсистема обеспечивает необходимые точности в заданных режимах работы привода.

Энергетическая подсистема содержит силовой источник питания, автономный инвертор напряжения (АИН), синхронный электромеханический преобразователь (СД) и исполнительную ось как нагрузку электропривода (ИО). С валом СД жестко связаны роторы первичного датчика углового положения ротора СД (ДПР), первичного датчика скорости (ДС) и первичного датчика углового положения исполнительной оси (ДУ).

Микропроцессорная система управления (МПСУ) обеспечивает аналого-цифровое преобразование сигналов первичных датчиков координат, реализацию алгоритмов цифровых регуляторов, формирование шести каналов широтно-модулированных сигналов для управления ключами АИН, связь с компьютером или пультом управления и реализацию функций защиты и диагностики системы. МПСУ привода построена на основе управляющего компьютера (УК) и контроллера (К).

Система может работать в трех основных режимах: в режиме программного наведения, в режиме автосопровождения и в комбинированном режиме. Смена режимов осуществляется по командам от ЭВМ верхнего уровня.

Первичным вычислительным устройством системы является управляющий компьютер. Необходимость использования в данном случае именно полноценного компьютера, а не встроенного контроллера объясняется требованием реализации на нем в режиме реального времени, помимо задачи управления приводом, целого ряда задач визуализации, регистрации и управления работой системы в целом. Это, во-первых, требует реализации многозадачного режима работы в реальном времени, что подразумевает наличие многозадачной операционной системы. Во-вторых, необходимым является наличие большого объема оперативной и энергонезависимой памяти (накопители на жестких дисках или Flash-карты), а также мощных графических устройств визуализации и манипуляции. Все перечисленные требования в совокупности могут быть выполнены лишь на вычислительных устройствах класса управляющих рабочих станций.

Требование подключения датчика положения к управляющему компьютеру в режиме программного наведения определяет необходимость реализации алгоритма цифрового регулятора положения на этом компьютере. Он может быть реализован на центральном процессоре управляющего компьютера или, в случае повышенной вычислительной загрузки последнего, на встроенном сопроцессоре.

Выходным сигналом управляющего компьютера, с точки зрения системы управления приводом, является выходной сигнал контура цифрового регулирования положения, который, в свою очередь, является входным сигналом задания для контура регулирования скорости.

Задача контроллера (К) заключается в реализации алгоритмов цифровых регуляторов скорости и фазных токов двигателя, алгоритмов обработки сигналов бесконтактного тахогенератора (БТГ) и ДПР, формировании широтно-модулированных сигналов управления силовыми ключами АИН, а также в реализации функций защиты преобразователя и аварийного выключения привода. Для взаимодействия управляющего компьютера и управляющего контроллера в стойке управления приводом они должны быть связаны посредством скоростного информационного канала.

Сигналы обратных связей для контуров регулирования фазных токов формируются с помощью датчиков тока фаз двигателя и поступают в аналоговом виде в контроллер (К). Они должны быть преобразованы в цифровые коды с помощью анало-

го-цифровых преобразователей. Выходные сигналы цифровых регуляторов тока преобразуются в набор широтно-модулированных сигналов для управления силовыми ключами преобразователя. Поэтому в состав управляющего контроллера приводом должен входить многоканальный широтно-импульсный модулятор.

В результате анализа современной микропроцессорной элементной базы была установлена целесообразность и возможность реализации всех перечисленных функций на основе DSP-контроллеров, специализированных для управления полупроводниковыми преобразователями. К их числу относятся микроконтроллеры семейств ADMC300, ADMC400 фирмы Analog Devices. Большинство современных микроконтроллеров, обладающих встроенными широтно-импульсными модуляторами, способны формировать не более 6 каналов ШИМ первого рода. В микроконтроллерах ADMC имеется возможность изменять требуемую длительность импульса ШИМ дважды за период, что является первым приближением к ШИМ2.

В качестве ДУ наиболее широко применяются вращающиеся трансформаторы. Одна из особенностей ВТ – относительная сложность обработки их выходных сигналов для получения информации о положении вала двигателя. Существует несколько способов решения данной задачи [2, 3]. Для обеспечения требуемых точностных и динамических показателей, предъявляемых к организации канала обратной связи по положению, рекомендуется использовать микропроцессорную систему обработки сигналов СКВТ [4]. Структурная схема данной системы приведена на рис. 2. Использование данного метода позволяет уменьшить погрешность определения угла до единиц угловых секунд [5]. Для повышения точности предлагается использовать способ, предложенный в [1]. Данный способ позволяет компенсировать погрешность первичного преобразователя с характером изменения по углу вида $\sin 4\alpha$ и $\sin 8\alpha$. При реализации данного способа с помощью микроконтроллера, выполняющего вторичное преобразование сигналов СКВТ в код, не требуется производить каких-либо аппаратных изменений схемы. В отличие от предложенного в [1] конструктивного решения, программная реализация данного алгоритма позволяет без существенных затрат компенсировать и более высокочастотные составляющие погрешности. Результирующая погрешность не превышает уровня нескольких десятых долей угловой секунды.

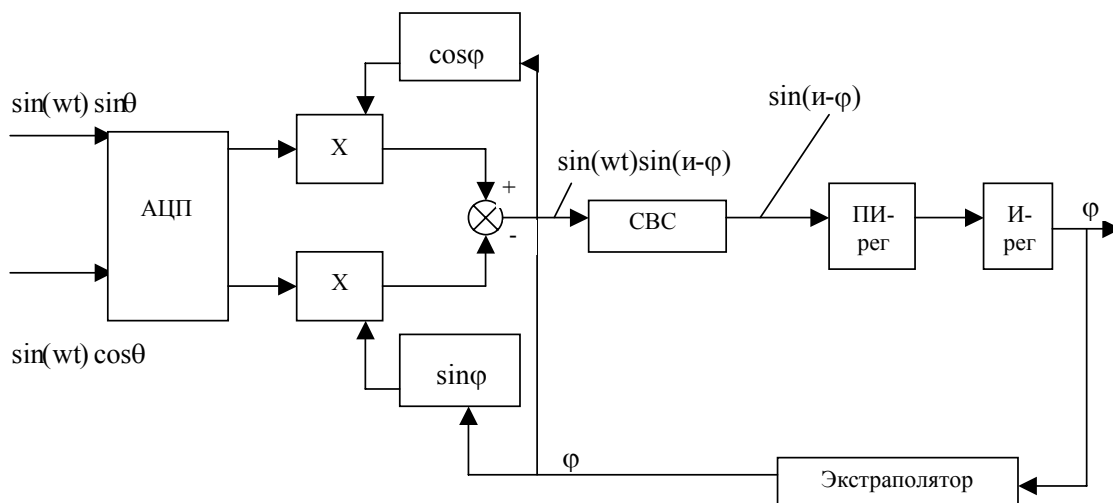


Рис. 2. Структурная схема обработки сигналов СКВТ

Важным фактором является скорость передачи информации об угле положения вала двигателя от вторичного преобразователя контроллеру и управляющему компьютеру. По данному вопросу было реализовано несколько аппаратных решений.

1. Разработана микропроцессорная система, состоящая из платы с контроллером PCI-интерфейса и контроллером, обеспечивающим обмен информацией между вторичным преобразователем, управляющим компьютером и контроллером. Выходной параллельный 22-битный код вторичного преобразователя с помощью мультиплексора и регистра преобразуется в 3 последовательные серии по 8 бит и передается на контроллер связи (КС). КС передает информацию по PCI интерфейсу в УК и получает от него угол задания, который по последовательному (RS-232, RS-485) интерфейсу передается в контроллер К (рис.3,а). Такая схема полностью себя оправдывает при управлении двигателем постоянного тока, однако при использовании СД возникают трудности, поскольку информация о положении также необходима для формирования сигналов управления в контуре тока. Таким образом, необходима передача информации сразу двум приемникам.

2. Информация от вторичного преобразователя считывается напрямую контроллером связи. Дальнейшая передача информации происходит по CAN-интерфейсу (рис.3,б). В качестве контроллера связи использован микроконтроллер ADuC7026 и микросхема CAN-интерфейса. Использование данного интерфейса позволяет передавать информацию на большие расстояния со сравнительно высокой скоростью и оставляет возможность для подключения большего числа приемников/передатчиков. Контроллер управления приводом также должен иметь CAN интерфейс.

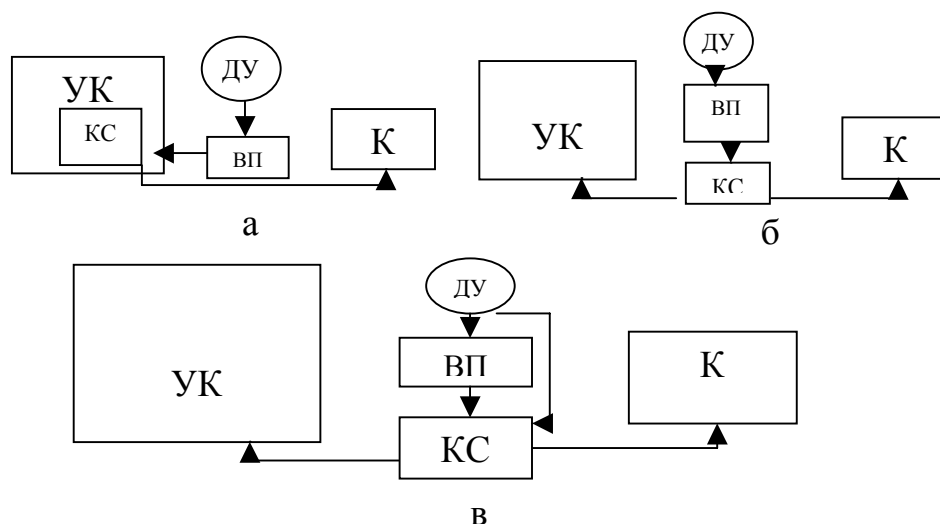


Рис. 3. Функциональные схемы канала обратной связи по углу положения

3. Информация с ДПР поступает только на контроллер управления приводом К для формирования контура тока, а с ДУ через КС – на УК. Такой вариант – наиболее затратный.

4. Информация, необходимая для замыкания контура положения, вырабатывается вторичным преобразователем ВП, а затем через контроллер связи поступает в УК. Необходимую для замыкания контура тока информацию об угле можно получить, реализовав соответствующий алгоритм на КС (рис.3,в). Для замыкания контура тока 22-битный код углового положения является избыточным, 14-разрядного кода более чем достаточно. С другой стороны, ВП выдает код с частотой единицы кГц, что является недостаточным. Реализация менее точной обработки сигналов СКВТ позволит повысить скорость обновления информации до 20 кГц.

Проблемным является вопрос получения информации о скорости исполнительной оси при инфранизких скоростях слежения (до единиц угловых секунд за секунду). Перспективно применение для этих целей бесконтактных шестифазных синхронных тахогенераторов.

На выходе шестифазного тахогенератора имеются 6 сигналов, изменяющихся при вращении вала по синусоидальному закону и сдвинутые друг относительно друга на 60 градусов. Для получения информации о скорости вращения вала двигателя необходимо выполнить преобразование этих 6 сигналов к одному сигналу постоянного тока. При решении такой задачи традиционными способами уровень пульсаций выходного напряжения недопустимо велик, а применение фильтра невозможно вследствие того, что приходится работать при вращении с инфранизкими скоростями.

Для уменьшения пульсаций выходного сигнала предлагается перейти от шести-фазной системе к 12-фазной. Это легко достигается при попарном суммировании сигналов, сдвинутых на 60 градусов (U_1 и U_2 , U_2 и U_3 , U_3 и U_4 , U_4 и U_5 , U_5 и U_6 , U_6 и U_1). Затем осуществляется суммирование модульных значений напряжений полученной 12-фазной системы (рис.4). Причем напряжения, проходящие на сумматор, определяются в зависимости от текущего угла положения вала двигателя. Коды, соответствующие прохождению каждой из фаз, представлены на рис.5.

Расчеты показывают, что сформированный таким образом сигнал $U_{ТГ}$ в виде напряжения постоянного тока при исходных синусоидальных сигналах $U_1 - U_6$ имеет пульсацию, не превышающую 3,4 % от единичного сигнала, частота изменения которой в двенадцать раз превышает частоту изменения фазных сигналов ТГ (пульсации без перехода к 12-фазной системе составляют 16%). Сглаживание такой пульсации до допустимого уровня не обеспечивает качественного управления. Отсюда возникает требование по минимизации пульсаций.

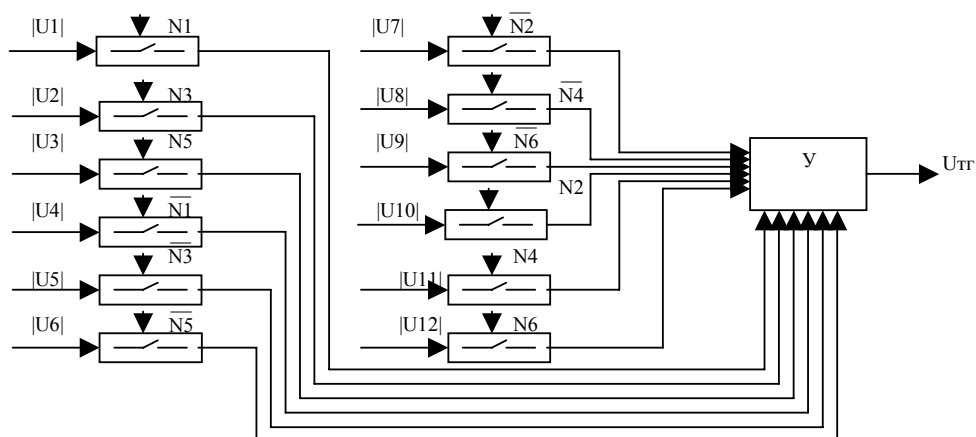


Рис. 4. Структурная схема обработки сигналов тахогенератора

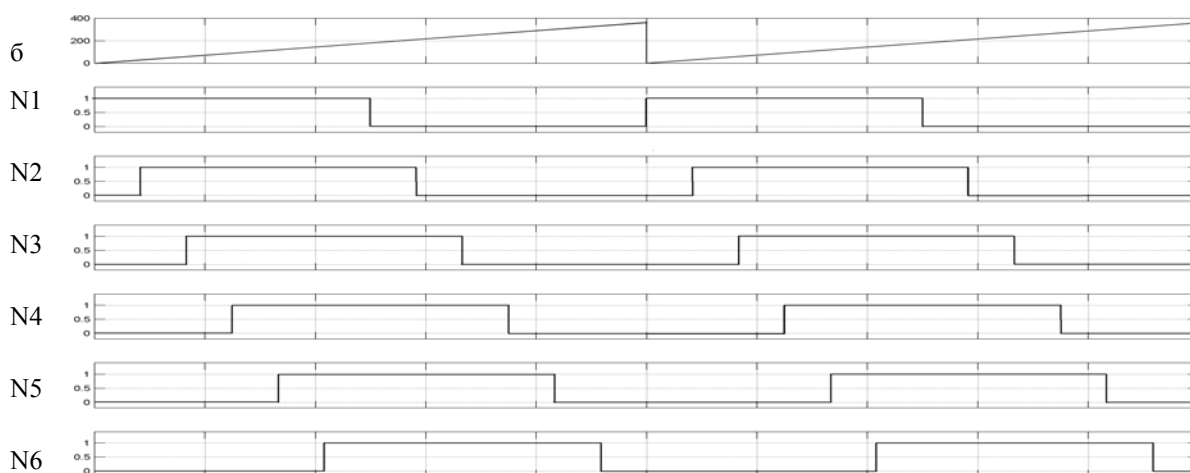


Рис. 5. Сигналы управления коммутацией фаз

Первый вариант коррекции пульсаций начального сигнала $U_{ТГ}$ состоит в следующем. Корректирующее напряжение U_K образуется определенной коммутацией модульных значений 12-фазной системы (рис.6,7). С учетом тождественности (с определенной степенью точности) характера изменения сигналов пульсации и U_K происходит понижение уровня пульсации.

Расчеты показывают, что при таком варианте коррекции суммарный сигнал $U_{ТГ1} = U_{ТГ} + U_K$ имеет пульсацию Δk на уровне 0,9%, т.е. обеспечивается почти четырехкратное снижение пульсации по отношению к начальной пульсации. Необходимо отметить следующий положительный момент - частота пульсации сигнала $U_{ТГ1}$ в два раза превышает частоту пульсаций сигнала $U_{ТГ}$.

При втором варианте предполагается скорректировать сигнал $U_{ТГ}$ следующим образом. Зависимость пульсаций $U_{ТГ}$ в пределах электрического периода ТГ с определенной степенью достоверности соответствует зависимости вида $|\sin 6\alpha|$, которая может быть сформирована соответствующим преобразованием значений кода угла в код поправки. Расчетные значения пульсации напряжения после коррекции не превышают уровня 0,21%.

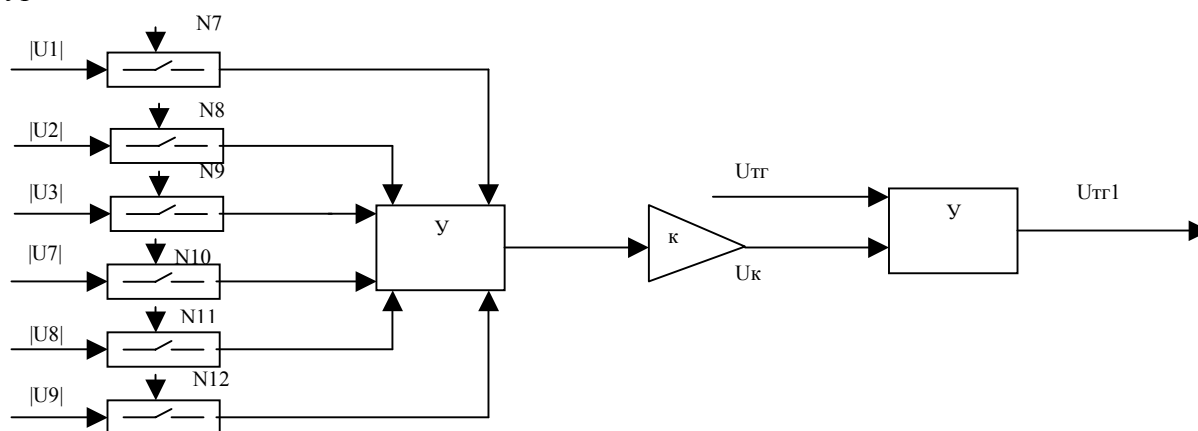


Рис. 6. Структурная схема обработки сигналов ТГ с коррекцией

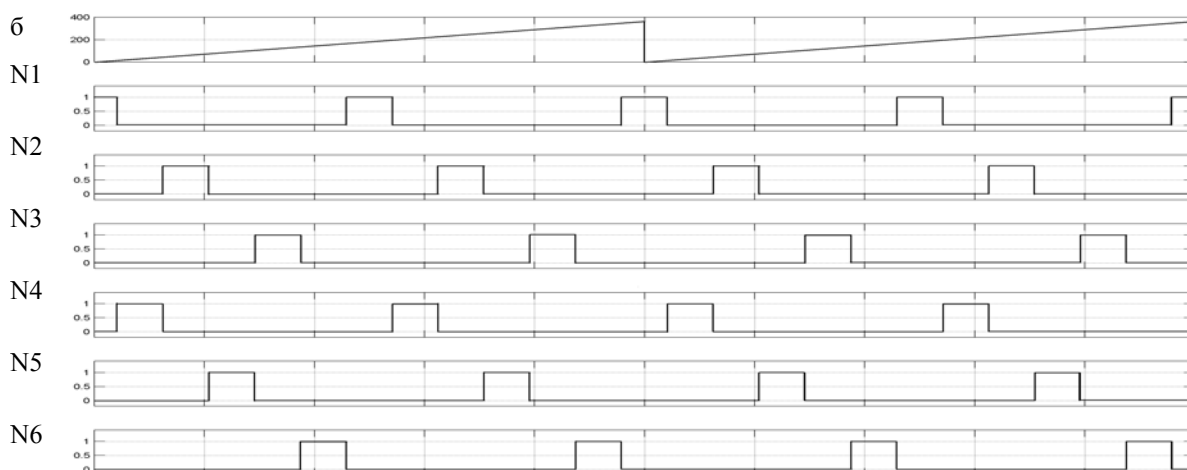


Рис. 7. Сигналы управления коммутацией фаз в схеме с коррекцией

Третий вариант коррекции по своей технической сути соответствует ранее рассмотренному варианту, но в нем уже коррекции подвергается напряжение $U_{ТГ1}$, пульсация которого в пределах электрического периода ТГ характеризуется зависимостью $\cong |\sin 12\alpha|$. Код коррекции формируется по значениям кода угла. Максимальное расчетное значение пульсации тахометрического сигнала соответствует значению порядка

0,08%. Характер пульсаций при первом, втором и третьем вариантах коррекции представлен на рис.8.

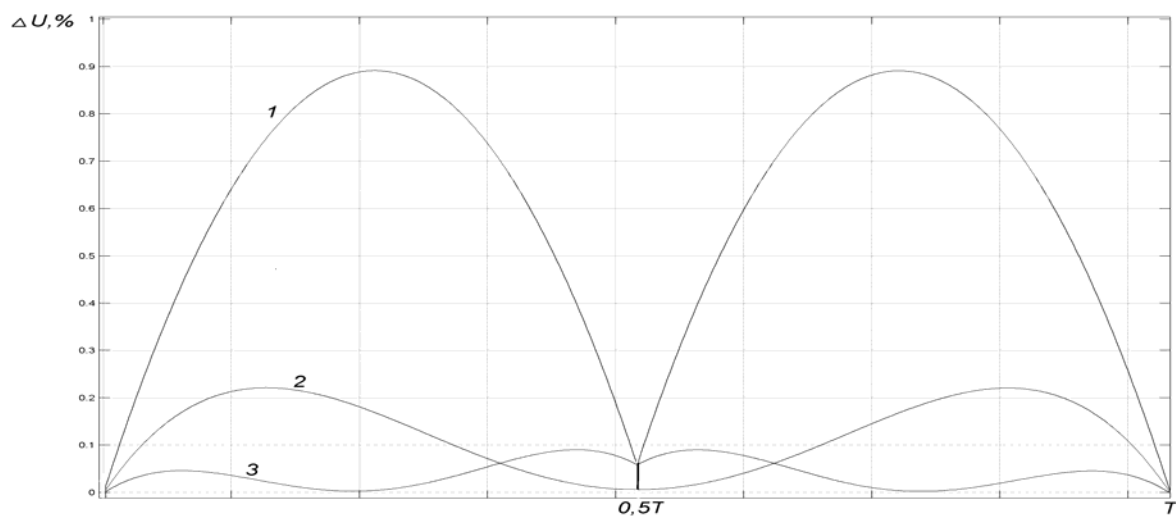


Рис. 8. Характер пульсаций сигнала ТГ после обработки

При работе с шестифазной системой и последующим выпрямлением и суммированием напряжений не удастся обеспечить выполнение требований, предъявляемых к современным прецизионным позиционным приводам; после осуществления предложенной коррекции напряжение $U_{ТГ}$ соответствует требованиям, предъявляемым к выходному напряжению ТГ постоянного тока, а динамические и точностные показатели системы существенно повышаются и удовлетворяют заданным параметрам.

Реализация алгоритмов коррекции также может быть возложена на КС, либо, в зависимости от конкретной задачи, может быть выполнена на интегральной схеме Altera.

Таким образом, в работе показано, что для обеспечения высокой точности слежения и качественной динамики системы необходимо обладать точной информацией об угле положения и скорости вращения вала двигателя. Предложены перспективные алгоритмы обработки информации датчиков угла поворота и скорости исполнительной оси, позволяющие устранить шумы и получить код угла и скорости с требуемой точностью. Реализованы аппаратные решения организации информационной подсистемы следящего электропривода. Использование предложенных методов обеспечивают точность наведения на уровне десятков угловых секунд.

Литература

1. Воронин Н.Н., Домрачев В.М., Сигачев И.П., Тимашов Н.А. Высокоточный цифровой преобразователь угла. // Измерительная техника. – 2004. – №2. – С. 20–24.
2. Дж. Вульвет. Датчики в цифровых системах. – М.: Энергоиздат, 1981.
3. Домрачев В.Г., Матвеевский В.Р., Смирнов Ю.С. Схемотехника цифровых преобразователей перемещений. – М.: Энергоиздат, 1987.
4. Жданов И.Н., Денисов К.М. Микропроцессорная система обработки данных СКВТ в системе электропривода. // XXIX неделя науки СПбГТУ. Часть V. СПб: СПбГТУ, 2001
5. Жданов И.Н. «Моделирование микропроцессорной системы обработки сигналов СКВТ для прецизионного электропривода». // Вестник конференции молодых ученых СПбГУ ИТМО. Сборник научных трудов. Том 2. СПб: СПбГУИТМО, 2004

ОПТИКО–ГЕОМЕТРИЧЕСКИЙ МЕТОД ИССЛЕДОВАНИЯ НАПРЯЖЕННО-ДЕФОРМИРОВАННЫХ СОСТОЯНИЙ ЭЛЕМЕНТОВ МИКРОЭЛЕКТРОНИКИ И МИКРОСЕНСОРНОЙ ТЕХНИКИ

А.Г. Коробейников, В.Л. Ткалич

На базе поляризационно-оптического метода предлагается оптико-геометрический метод определения напряженно-деформированных состояний элементов микроэлектроники и микросенсорной техники. Статья подготовлена по результатам НИР “Разработка неразрушающих бесконтактных оптических методов исследования стереометрии и внутренних структурных дефектов элементной базы микроэлектроники и микросенсорной техники”

Введение

Задачи повышения прочности, износостойкости и надежности микроэлектроники и микросенсорной техники в настоящее время входят в основную группу приоритетов современного научно-технического прогресса. Одним из важнейших показателей надежного функционирования элементов и узлов технологического оборудования является оценка его напряженно-деформированного состояния (НДС). Определение НДС выполняется экспериментальными исследованиями, аналитическими и расчетно-экспериментальными методами [1].

Все методы оценки НДС имеют свои достоинства и недостатки по ряду показателей: чувствительности, точности, глубине определения измерения, времени, требуемого на измерение и т.д. Поэтому выбор конкретного метода диктуется в основном техническим заданием.

В развитии современной техники наблюдается тенденция использования функциональных объектов малых размеров. Примером может служить электронная техника, в которой микроминиатюризация приборов уже сейчас требует применения элементов, размеры которых составляют несколько микрон. Уникальность свойств таких объектов (наноструктур) во многом определяется атомными и электронными процессами, протекающими как в объеме, так и на их границах и имеющими уже квантовый характер. Задача осложняется еще и тем, что наноструктуры являются достаточно неравновесными системами.

Поэтому дальнейшее развитие электроники ставит задачу разработки методов исследования НДС, с помощью которых можно получать информацию и обнаруживать структурные неоднородности на границе раздела фаз подложка – пленка и пленка – окружающая среда. Это можно делать на базе неразрушающих бесконтактных поляризационно-оптических методов интерферометрии, которые позволяют определить коэффициенты концентрации механических напряжений элементной базы микроэлектроники и микросенсорной техники с выявлением причин, вида и мест локализации.

В данной работе предложен оптико-геометрический метод (ОГМ) исследования деформаций и напряжений. Физическим основанием выбора ОГМ служит тот факт, что НДС, например, является причиной оптических искажений (фотоупругих, термических) в градиентных и однородных элементах технической оптики. Эти явления могут быть обусловлены технологическими (остаточными напряжениями вязко-упругого происхождения при охлаждении градиентного изделия после высокотемпературного синтеза) или/и эксплуатационными (силовыми, температурными) факторами. Таким образом, работы, посвященные разработке и анализу ОГМ исследования НДС элементов микроэлектроники и микросенсорной техники, являются актуальными.

Оптико-геометрический метод

Для исследования НДС и выявления взаимосвязей ряда геометрических параметров с характеристиками надежности в работе предложен модифицированный поляризационно-оптический метод. Он характеризуется тем, что не требует больших деформаций образца и, тем более, его разрушения. Кроме того, достаточно лишь небольших количеств образца для исследования упругих свойств (например, модуль Юнга, напряжение слома и т.д.). В особенности этот метод удобен при определении внутренних остаточных напряжений, которые прозрачные тела приобретают в процессе своего создания и существования.

При проведении экспериментальных работ использовалась установка, структурная схема которой изображена на рис. 1. В установке применялся источник когерентного монохроматического излучения – лазер (Л), свет от которого пропусклся через объектив (О), на выходе которого получается параллельный пучок света. После этого свет проходил через поляризатор (П). Затем свет пропусклся через исследуемый образец (ИО), где его поляризация неким образом изменялась. После этого волна, проходя через анализатор (А), скрещенный с поляризатором, попадала в фотодатчик (ФД).

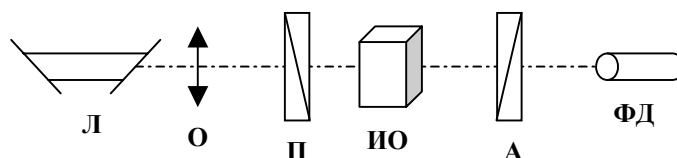


Рис. 1. Схема экспериментальной установки

После выхода из П луч имеет линейную поляризацию, причем ее плоскость повернута под углом α к оси исследуемого образца. В этом случае для интенсивности на выходе получим:

$$I = I_0 \sin^2(2\alpha) \sin^2(\pi\Delta/(2\lambda\alpha)),$$

где I_0 – интенсивность на выходе из поляризатора, I – интенсивность на выходе из анализатора, Δ – оптическая разность хода, λ – длина волны, $2\pi\Delta/\lambda$ – разность хода между лучами.

В случае $\alpha = \pi/4$ получается: $I = I_0 \sin^2(\pi\Delta/\lambda)$.

Для того чтобы связать разность хода с напряжениями σ_i , воспользуемся уравнением Вертгейма [2]:

$$\Delta = lc(\sigma_1 + \sigma_2 - \sigma_3),$$

где $\sigma_1, \sigma_2, \sigma_3$ – главные диагональные напряжения в тензоре механических напряжений, l – длина пути луча в веществе, c – коэффициент фотоупругости.

При проведении исследований напряжения σ_2 и σ_3 были равны нулю, поэтому уравнение Вертгейма запишется так: $\Delta = lc\sigma_1$.

При проведении измерений использовалась струбцина, на которую была нанесена шкала углов и у которой был померен шаг резьбы. Для того чтобы напряжения в объектах были распределены равномерно, использовались стекла в качестве прокладок. Таким образом, деформации могли быть измерены с большой точностью. Далее был изготовлен кубик из смеси эпоксидной смолы (ЭС) марки ЭДП с ацетоном. Измерения зависимости интенсивности от угла поворота ручки струбцины сначала проводились для ЭС, положенной последовательно с эталонным образцом (ЭО), модуль Юнга которой был заранее измерен (рис.2, а), а затем просто с ЭС (рис. 2, б). По экспериментальным данным был определен период колебаний интенсивности на графике зависимости ин-

тенсивности от деформации для обоих случаев (рис. 2, в–г). За один период у ЭС напряжения меняются на одну и ту же величину независимо от того, что лежит последовательно с ней, следовательно, и деформации у ЭС одинаковы. Обозначим изменения высоты в первом и во втором случаях соответственно Δh_1 и Δh_2 , а площади ЭО и ЭС – S_1 и S_2 . Модуль Юнга ЭО обозначим за E_1 , а ЭС – за E_2 :

$$E_2 = E_1(S_1/S_2)((\Delta h_1 - \Delta h_2)h/(d\Delta h_2)), \sigma_0 = E_1(S_1/S_2)((\Delta h_1 - \Delta h_2)/(d)),$$

где σ_0 – напряжение в ЭС, приходящееся на один период, $\Delta h_i = \alpha_i \delta / 2\pi$, δ – шаг резьбы, α_i – угол поворота струбицы.

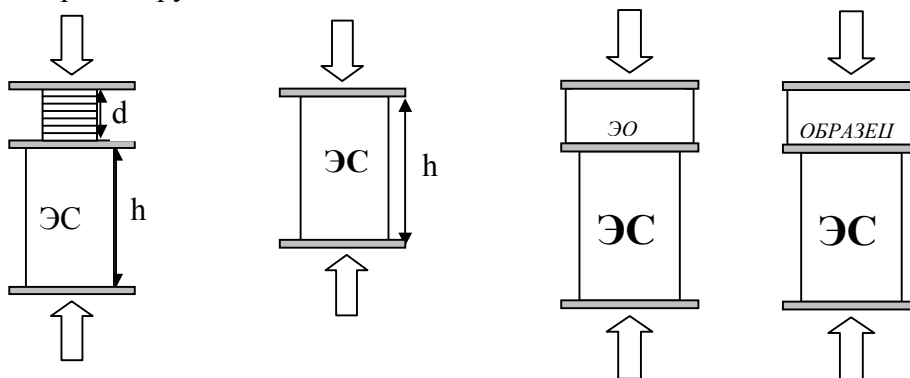


Рис. 2, а–г. Различные варианты организации измерений

Зная σ_0 , можно определить модуль Юнга любого вещества, размещая его, как на рис. 4, т.е. $E_k = \sigma_0(S_2/S_k)(2\pi d_k/(\delta(\alpha_k - \alpha_2)))$.

Заключение

С позиций физической мезомеханики легко объяснить, почему столь пристальное внимание уделяется структуре и свойствам поверхностных слоев нагруженных твердых тел. Деформация, обусловленная локальной потерей сдвиговой устойчивости твердого тела, начинается в зонах концентраторов напряжений соответствующего масштаба и, в первую очередь, в поверхностных слоях материала, так как они имеют наименьшее сдвиговое сопротивление. В результате в поверхностных слоях интенсивно развиваются полосы локализованной деформации, снижающие сопротивление сдвигу твердого тела. Поверхностные концентраторы напряжений генерируют деформационные дефекты (точечные дефекты, дислокации, дисклинации, мезополосы сдвига), которые, зарождаваясь на поверхности, перемещаются затем в глубь образца, обуславливая развитие пластического течения в объеме материала. Модифицируя поверхностный слой путем нанесения покрытий с различной кристаллической структурой, можно создать условия, позволяющие предотвратить процесс зарождения деформационных дефектов и, как следствие, повысить механические свойства материала.

Литература

1. Кузьменко Б.П., Садаков О.С. Применение оптико-геометрических методов к исследованию некоторых задач механики композитов // В кн.: Оптико-геометрические методы исследования деформаций и напряжений. – Челябинск, 1986.
2. Сивухин Д.В. Общий курс физики. Механика. – М.: Наука, 1979

РАЗРАБОТКА АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ КОНТРОЛЯ КАЧЕСТВА МИКРО - И НАНОСТРУКТУР

А.Г. Коробейников, В.Л. Ткалич, Б.М. Папченко

На базе интерферометрического метода разработана автоматизированная система определения параметров шероховатости микро – и наноструктур. Статья подготовлена по результатам НИР “Разработка методов исследования качественных и количественных характеристик наносистемных структур”.

Введение

В настоящее время развитие интегральной микроэлектроники вышло на такой уровень, когда наиболее продвинутые в инновациях фирмы ведут исследования, направленные на разработку технологии, которая обеспечивала бы масштабирование быстродействующих СБИС. Например, в январе 2007 г. компания Intel продемонстрировала действующую модель нового чипа, созданного с использованием транзисторов изготовленных по 45-нанометровой технологии, в которых вместо диоксидов кремния применены металлические затворы. По словам соучредителя компании Гордона Мура (Gordon Moore), новая разработка является крупнейшим изменением в транзисторной технологии за последние 40 лет. Intel изменила основу своих микросхем с тем, чтобы соответствовать современным требованиям индустрии персональных компьютеров. Прежние разработки, такие как Pentium, хотя и имеют высокую скорость вычислений, но требуют сильного охлаждения. Новые же транзисторы используют металл гафний и потребляют энергию на 30% меньше с повышением производительности на 20%. Они построены на основе архитектуры Core, которая в настоящее время используется в процессорах семейства Core 2 Duo. Эта технология была разработана компанией IBM совместно с Toshiba, Sony и AMD. Компания Intel намерена приступить к коммерческому производству новых чипов уже в 2007 г. В продаже они появятся, вероятно, в 2008 году.

Компания Toshiba сообщила о разработке транзистора по 32-нм технологии. Устройство и соответствующие спецификации, по сообщению Dusiness Nikkei Daily, компания представила конференции US International Electron Devices. Транзистор получил название FinFET – field effect transistor. Его реальная длина составляет всего 20 нм. Руководство Toshiba возлагает на эту технологию самые радужные надежды.

Как видно из вышесказанного, по некоторым параметрам миниатюризация СБИС процессоров достигла значений, которые являются предельными с точки зрения физических ограничений процессов, происходящих в структурах СБИС. В то же время наличие многослойной коммутации и большого числа слоёв межслойной изоляции обуславливают сложный рельеф поверхности кристалла, в котором формируется СБИС. Этот рельеф непрерывно изменяется с переходом от одного техпроцесса к другому, и требуется производить контроль высоты и площади рельефа.

Для обеспечения контроля СБИС с такой степенью интеграции разработаны методы и аппаратура, предназначенные для формирования изображений, элементного анализа, построения профилей распределения элементов и др. Как утверждается в [1], методы формирования изображения относятся к наиболее ценным, как минимум, по двум причинам. Во-первых, изображение «выглядит» так, как оно представляется аналитику, и тогда приводится в действие весь ранее накопленный опыт визуальных наблюдений, и изображение предоставляет максимум информации за минимальный промежуток времени. Во-вторых, открывается возможность применения мощных средств обработки изображений, повышающих информативность и уровень количественной интерпретации.

В данной работе рассматривается автоматизированная система определения параметров шероховатости микро- и наноструктур на базе бесконтактного метода неразрушающего контроля поверхности с использованием микроинтерферометра МИИ-4, созданного по схеме академика В.П. Линника, и представляющего собой развитие схемы интерферометра Майкельсона для случая системы сильного увеличения [2].

Разработка измерительно-вычислительного комплекса для анализа шероховатости поверхности микро- и наноструктур

Используемый в настоящее время для бесконтактного определения шероховатости поверхности микроинтерферометр МИИ-4М имеет недостатки как в осветительном, так и в измерительном каналах. Коротко их рассмотрим.

1. Источником света в МИИ-4М является лампа накаливания. При ее работе подавляющее количество потребляемой энергии выделяется не в виде световой энергии видимого диапазона, а в виде тепловой энергии. Вследствие этого лампа накаливания нагревает часть оптических деталей и корпуса прибора, контактирующего с узлом осветителя, создавая нестационарный градиент температуры в ветвях интерферометра. Это приводит к дополнительной разности хода лучей, а, значит, и к появлению случайной ошибки измерений.

2. Лампа накаливания по своей конструкции и технологии изготовления не имеет постоянного положения нити накаливания относительно посадочного места лампы. В процессе работы лампы из-за разогрева спирали меняется ее жесткость, происходит смещение нити накала под действием силы тяжести.

3. Оператор должен выполнить измерения профиля кривой, находящейся в поле зрения. При этом измерения проводятся в двух взаимно перпендикулярных направлениях. Оператор оценивает на глаз расстояния и толщины, составляющие около 0,1 ширины изображения. Это требует от оператора значительного напряжения органов зрения. Для одного измерения требуется снять более 10 отсчетов, что занимает несколько минут напряженной зрительной работы.

4. Поскольку каждый снятый отсчет оператор должен записать, у него постоянно меняется режим аккомодации глаза, меняется характер освещения. Все это приводит к сильному утомлению органов зрения и как следствие – снижению точности и скорости измерений.

Первые два указанных недостатка требуют наличия узла подвижки лампы и проведения операции настройки освещения как перед работой, так и периодически в процессе работы. Качество настройки освещения существенно влияет на качество изображения и, соответственно, на точность измерений.

Третий и четвертый недостаток сказываются на эффективности работы оператора.

Таким образом, указанные недостатки делают актуальной задачу разработки автоматизированной системы определения параметров шероховатости микро- и наноструктур на базе бесконтактного метода неразрушающего контроля поверхности с использованием микроинтерферометра МИИ-4. Цель создания – повышение информативности и точности измерений шероховатости поверхности, ускорение процесса измерений и снижение утомляемости оператора.

Для начала был подвергнут модернизации узел осветителя, в котором применен принципиально новый тип источника света. Затем разработан блок, в который входит цифровая телевизионная камера и канал связи с персональным компьютером (ПК). Кроме этого, было разработано специализированное программное обеспечение (ПО).

В узле осветителя использован отечественный светодиод, так как формирование необходимой оптической характеристики светодиода (форма колпака) практически можно реализовать только при непосредственном взаимодействии с изготовителем изделия.

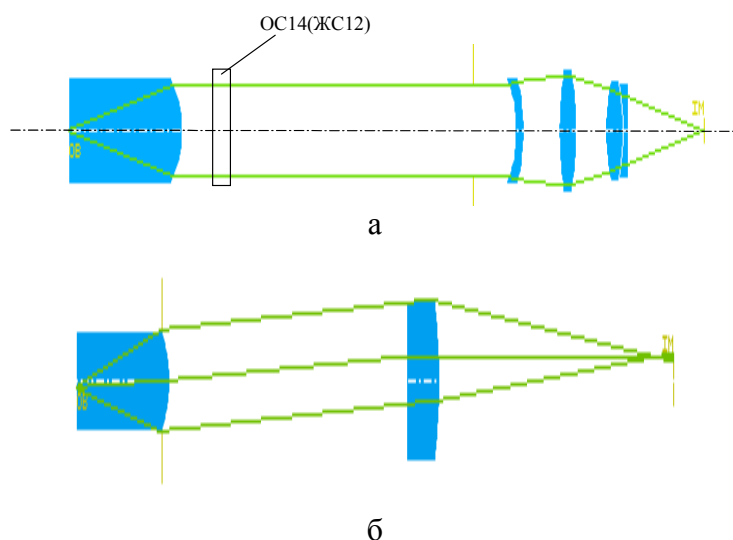


Рис. 1. Светодиод с эллиптическим колпаком

При использовании светодиода УЗ37 с эллиптическим колпаком в осветительном канале МИИ-4М возможны два варианта решения поставленной задачи. В первом варианте (рис. 1, а) используется светодиод со сложным коллектором и конденсорной линзой. Таким образом, мы полностью используем осветительный канал МИИ-4М, который рассчитан для работы с лампой накаливания. Во втором варианте рассмотрен случай использования только конденсорной линзы (рис. 1, б). В этом случае светодиод с эллиптическим колпаком уже представляет из себя осветитель с коллектором, и можно ожидать, что изъятие линз конденсора не приведет к ухудшению характеристик освещения при упрощении конструкции. Анализ проведенных расчетов показал, что схема (б) имеет лучшие характеристики по светораспределению при одновременном упрощении конструкции. Поэтому было принято решение разрабатывать конструкцию по схеме (б).

При расчете осветительного канала были определены зависимости между фокусным расстоянием колпака, показателем преломления материала и радиусом при вершине.

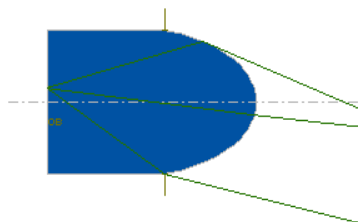


Рис. 2. Вид рассчитанного колпака

В результате расчета получены данные для изготовления эллиптического колпака для светодиода УЗ37 (рис. 2). Оптическая схема (ОС) осветителя по существу является упрощенной ОС осветителя МИИ-4М, из которой исключены две линзы коллектора. Их функцию выполняет эллиптический колпак светодиода. Светодиод с колпаком крепится теплопроводящей пастой к несущей детали, представляющей собой полый стержень. Внутри этого стержня проходят провода питания светодиода. Сам стержень имеет значительную массу и выполнен из материала, имеющего высокую теплопроводность. Это позволяет обеспечить температуру корпуса кристалла не более чем на 2–3° выше температуры окружающей среды.

Для устранения третьего и четвертого недостатков МИИ-4М была установлена цифровая телевизионная камера, передающая информацию через блок измерения на ПК. Разрешающая способность цифровой камеры не должна быть меньше разрешающей способности при измерении с помощью микрометра dM . Ее можно определить из соотношения:

$$dM = D/\Delta,$$

где D – диаметр поля зрения в плоскости изображения, Δ – цена деления микрометра.

При $D = 10$ мм и $\Delta = 0,01$ мм получается $dM = 1000$ мм.

Отсюда следует, что разрешающая способность цифровой камеры должна быть не менее 1000 пикселей на 1 см.

Линейное увеличение дополнительного канала $\beta_{\text{дк1}}$ должно обеспечивать использование разрешающей способности объектива и цифровой камеры, определяемое формулой

$$\beta_{\text{дк1}} = 2\ell A / (n\lambda\beta_0),$$

где ℓ – длина наименьшей стороны приемника, A – числовая апертура объектива, n – число пикселей, λ – длина волны, β_0 – линейное увеличение объектива.

Линейное увеличение дополнительного канала $\beta_{\text{дк1}}$ должно также обеспечивать использование поля зрения оптического прибора. Для описанной окружности это

$$\beta_{\text{дк2}} = D_{\text{цк}} / (D\beta_0),$$

где $D_{\text{цк}}$ – диаметр приемника цифровой камеры, а для вписанной окружности –

$$\beta_{\text{дк3}} = \ell / (D_M \beta_0),$$

где D_M – линейное поле зрения оптического прибора в пространстве предметов.

Благодаря наличию цифровой телевизионной камеры с разрешающей способностью не менее 1000 пикселей по координатам X и Y появляется дополнительная возможность определения координат точек пересечения профиля с линией эквидистантной средней линии на заданном уровне.

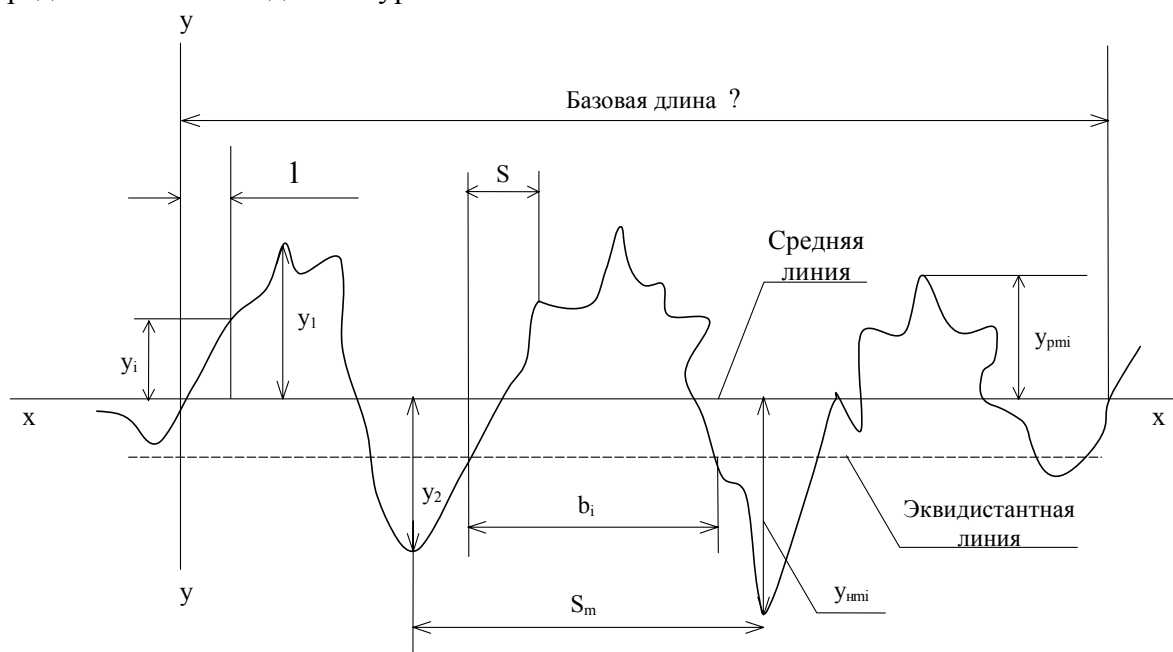


Рис. 3. Профиль поверхности объекта

Блок измерения позволяет фиксировать результаты измерения не только одной наводки на резкое изображение, но двух и более, обеспечивая запоминание результатов измерения профиля до линии, заданной оператором по высоте профиля. Последующие измерения профиля производятся от заданной линии и суммируются с предыдущими результатами измерений. На рис. 3 показан профиль поверхности объекта.

Заключение

Разработанная система существенно ускоряет определение параметров опорной длины профиля, что важно при исследовании износа, трения и других требуемых технических характеристик поверхностей. Ее позволяет за счет автоматизации процесса измерения и обработки результатов измерения сократить время определения всех регламентируемых ГОСТом параметров шероховатости в одном поле зрения не менее, чем в 10÷20 раз. При этом значительно уменьшается утомляемость оператора благодаря ликвидации ручной, визуальной наводки на точки профиля поверхности.

Основываясь на данной работе, можно разрабатывать конкурентоспособное на рынке оборудование, которое позволит проводить качественный анализ микро – и наноструктур элементной базы отечественной электроники

Литература

1. Мейеран Ю.С., Флинн П.А., Каррузерс Дж. Р. Аналитические методы в технологии СБИС // ТИИЭР. – 1987. – Том 75. – № 7. – С. 50–101.
2. Китайгородский А.И. Введение в физику. М.: Наука, 1973. – 685 с.

ТЕКСТОВЫЙ ЯЗЫК АВТОМАТНОГО ПРОГРАММИРОВАНИЯ

В. С. Гуров, М. А. Мазин, А. А. Шалыто

В статье описывается текстовый язык автоматного программирования, созданный с помощью системы метапрограммирования *JetBrains MPS*. Этот язык дополняет визуальное инструментальное средство автоматного программирования *UniMod* текстовым входом. При этом диаграммы состояний строятся по мере набора текста программы автоматически, что позволяет назвать предлагаемый подход к построению программ текстово-визуальным.

В рамках проекта *UniMod* [1] предложены метод и средство для визуального моделирования и реализации объектно-ориентированных программ со сложным поведением на основе автоматного подхода. При этом системы предлагалось строить с помощью двух типов *UML*-диаграмм: диаграмм классов и диаграмм состояний [2]. На диаграмме классов, которая представляется в виде схемы связей и взаимодействия автоматов, изображаются источники событий, автоматы и объекты управления, которые реализуют функции входных и выходных воздействий. Поведение автоматов описывается с помощью диаграмм состояний. Код для источников событий и объектов управления пишется на языке *Java* [3]. Указанные диаграммы и написанный вручную код могут интерпретироваться или компилироваться. Поэтому в рамках проекта *UniMod* указанные диаграммы являются не картинками, а визуальным языком программирования. Таким образом, этот проект поддерживает концепцию "Исполняемый *UML*" (*Executable UML*) [4, 5].

С использованием инструментального средства *UniMod* выполнен ряд проектов, который доступен по адресу <http://is.ifmo.ru/unimod-projects/>. Эти проекты показали эффективность применения автоматного программирования и средства *UniMod* при реализации систем со сложным поведением, но также выявили и ряд недостатков:

- ввод диаграмм состояний с помощью графического редактора трудоемок;
- многие программисты предпочитают работать с текстовым представлением программы, несмотря на то, что диаграммы позволяют представлять информацию более компактно и обозримо;
- невозможно в одном *Java*-классе совместить автомат и объект управления, что не позволяет прозрачно использовать автоматное программирование совместно с объектно-ориентированным, так как в настоящее время код, генерируемый из автоматной модели, не является в полной мере объектно-ориентированным.

Для устранения перечисленных недостатков авторы предложили новый подход к разработке автоматных программ и применению автоматов в объектно-ориентированных системах.

В рамках этого подхода используется система метапрограммирования *JetBrains Meta Programming System (MPS)* [6–8], которая позволяет создавать проблемно-ориентированные языки (*Domain Specific Language — DSL*) [8, 9]. Для задания языка в системе *MPS* требуется разработать:

- структуру абстрактного синтаксического дерева (АСД) [10] для разрабатываемого языка. Узлам АСД могут соответствовать такие понятия как «объявление класса», «вызов метода», «операция сложения» и т.п.;
- модель текстового редактора для каждого типа узла АСД. Задание редактора для узла АСД равноценно заданию конкретного синтаксиса для этого узла. При этом, если для традиционных текстовых языков программирования создание удобного редактора — отдельная сложная задача, то для языков, созданных с помощью средства *MPS*, редакторы являются частью языка. Эти редакторы поддерживают автоматическое завершение ввода текста и проверку корректности программы;
- модель ограничений на экземпляры АСД;

- модель системы типов [11] для языка;
- модель трансформации программы на задаваемом языке в исполняемый код.

Система *MPS* позволяет как создавать новые языки, так и расширять языки уже созданные с помощью этой системы.

В отличие от традиционных языков, языки, созданные с помощью системы *MPS*, не являются текстовыми в традиционном понимании, так как при программировании на них пользователь пишет не текст программы, а вводит ее в виде АСД с помощью специальных редакторов. Структура и внешний вид этих редакторов таковы, что работа с моделью программы для пользователя выглядит, как традиционная работа с текстом программы.

Отказ от традиционного текстового ввода программ значительно упрощает создание новых языков [12] — исчезает необходимость в разработке лексических и синтаксических анализаторов, и, как следствие, перестают действовать ограничения на класс грамматик языков. Недостатком такого подхода является зависимость языков от системы *MPS* — невозможно разрабатывать программы без этой системы. Однако подобное ограничение присуще и традиционным, чисто текстовым языкам, которые зависят от компиляторов. Впрочем, после трансляции программы, написанной на языке, созданном в системе *MPS*, исполняемый код перестает зависеть от этой системы.

С помощью системы *MPS* авторами созданы два варианта текстового языка для автоматного программирования.

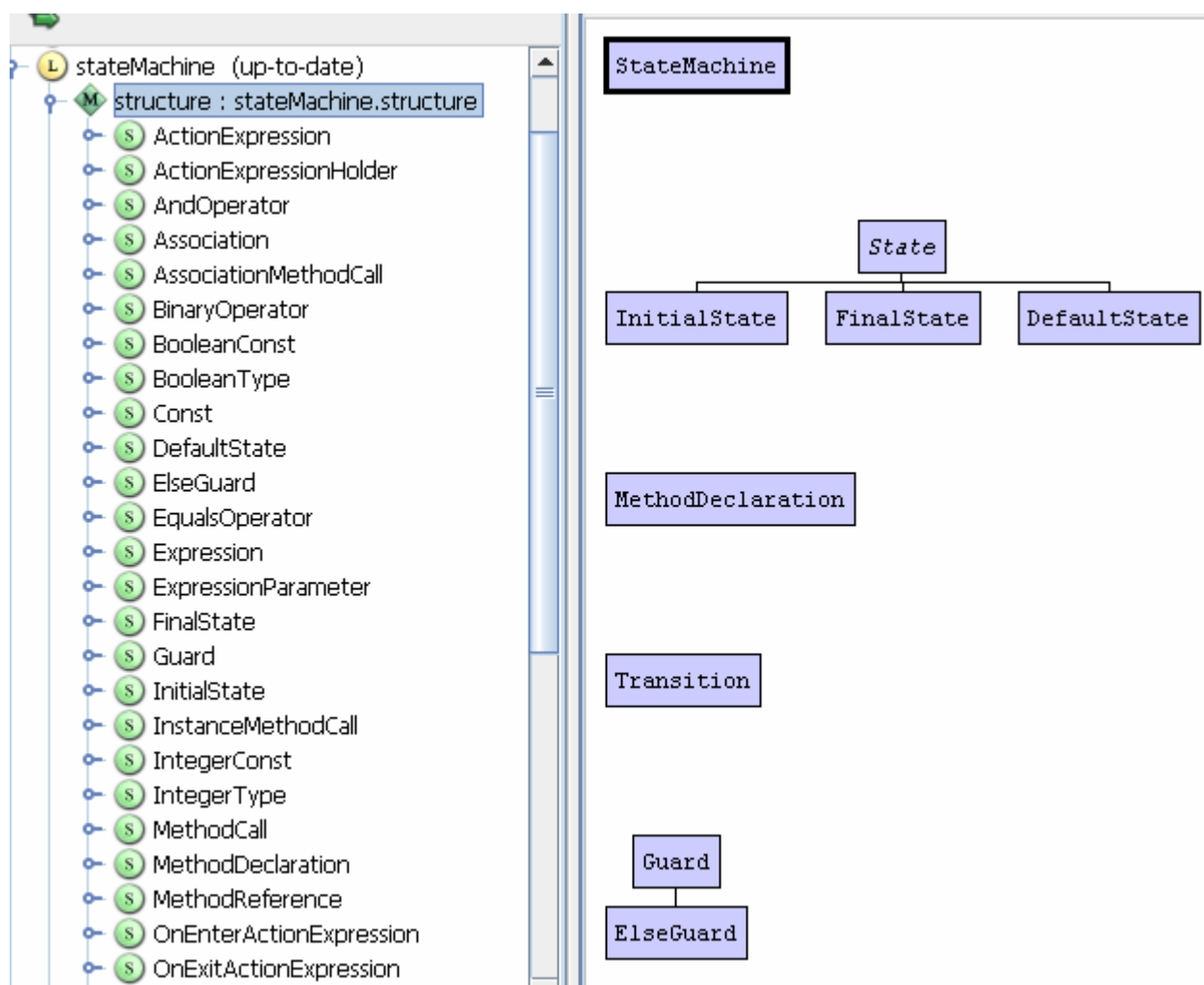


Рис. 1. Структура абстрактного синтаксического дерева текстового языка автоматного программирования первого типа в среде *MPS*

Первый язык выполнен в виде самостоятельного языка, а второй – в виде расширения языка *Java*. Эти языки позволяют описывать состояния и логику переходов автоматов, а также события, обрабатываемые автоматами. При этом, также как и в инструментальном средстве *UniMod*, функции входных и выходных переменных реализуются на другом языке программирования.

На рис. 1 показана структура абстрактного синтаксического дерева текстового языка автоматного программирования первого типа.

Текстовый язык автоматного программирования второго типа расширяет язык *Java* и позволяет в одном *Java*-классе совмещать автоматные и не автоматные аспекты. После написания программы на этом языке, она сначала транслируется в *Java*-код, но уже без автоматного расширения, а затем компилируется стандартным *Java*-компилятором. Преимуществом этого языка является простота его использования в объектно-ориентированных приложениях, написанных на языке *Java*. При применении этого языка проверка корректности программы осуществляется на стадии ее написания, а не в процессе компиляции.

В то же время для приложений, написанных на других языках, этот язык не подходит. Для таких приложений предпочтительно использовать независимую от языка *Java*, первую версию автоматного языка. Эта версия менее зависима от платформы и полностью соответствует понятию запускаемых спецификаций.

Использование специальных языков для автоматного программирования упрощает разработку автоматов, избавляя программиста от необходимости описывать их средствами языков общего назначения.

```

statemachine a_b_c_3 {
  << associations >>
  void next ( string );
  void end ( );
  void el ( );
  initial state s0 {
    transiteto p
  }
  state p {
    << onenter >>
    << onexit >>
    on next ( s ) else transiteto error ;
    on end ( ) else transiteto error ;
    initial state p1 {
      transiteto a
    }
    state a {
      << onenter >>
      << onexit >>
      on next ( " a " ) transiteto a ;
      on next ( " b " ) transiteto b ;
      << inner states >>
    }
    state b {
      << onenter >>
      << onexit >>
      on next ( " b " ) transiteto b ;
      on next ( " c " ) transiteto c ;
      << inner states >>
    }
    state c {
      << onenter >>
      << onexit >>
      on next ( " c " ) transiteto c ;
    }
  }
}

```

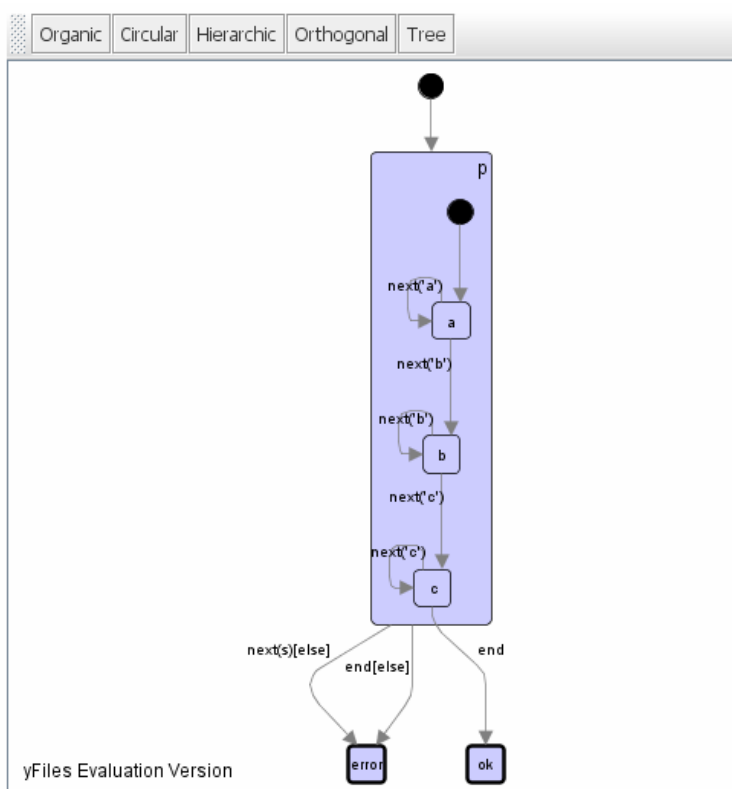


Рис. 2. Пример программы на текстовом языке автоматного программирования и автоматически построенной по ней диаграммы состояний

Так как диаграммы состояний являются более наглядными по сравнению с текстовым представлением автоматов, то при использовании рассматриваемых языков может быть обеспечена возможность автоматического построения диаграммы состояний по мере набора текста. Однако эти диаграммы доступны только в режиме просмотра, а не редактирования.

На рис. 2 слева показан пример программы, на разработанном текстовом языке автоматного программирования первого типа, который распознает цепочки символов вида $a*b*c^*$. Автоматически построенная диаграмма состояний изображена на рис. 2 справа.

С применением текстового языка автоматного программирования второго типа можно подробно ознакомиться в работе [13].

Исходя из изложенного, можно утверждать, что в настоящей работе предложен подход к построению программ, который может быть назван текстово-визуальным.

Литература

1. Гуров В.С., Мазин М.А., Нарвский А.С., Шалыто А.А. UML. SWITCH-Технология. Eclipse // Информационно-управляющие системы. – 2005. – № 6 (13). – С. 12–17. <http://is.ifmo.ru/works/uml-switch-eclipse/>.
2. Гуров В.С., Мазин М.А., Шалыто А.А. Операционная семантика UML-диаграмм состояний в программном пакете UniMod // Труды XII Всероссийской научно-методической конференции "Телематика-2005". СПб: СПбГУ ИТМО. – Т. 1. – С. 74–76. http://tm.ifmo.ru/tm2005/db/doc/get_thes.php?id=224
3. Eckel B. Thinking in Java. – NJ: Prentice Hall, 2006.
4. Executable UML. http://en.wikipedia.org/wiki/Executable_UML
5. Гуров В.С., Нарвский А.С., Шалыто А.А. Исполняемый UML из России // PC Week/RE". – 2005. – № 26. – С. 18, 19. http://is.ifmo.ru/works/_umlrus.pdf
6. Дмитриев С. Языково-ориентированное программирование: следующая парадигма // RSDN Magazine. – 2005. – № 5. <http://www.rsdn.ru/article/philosophy/LOP.xml>
7. Fowler M. A Language Workbench in Action – MPS. <http://martinfowler.com/articles/mpsAgree.html>
8. Фаулер М. Языковой инструментарий: новая жизнь языков предметной области (Domain Specific Languages) // Клиент-сервер. – 2005. – №3. <http://www.optim.su/cs/2005/3/fowler/fowler.asp>
9. Ward M. Language Oriented Programming // Software — Concepts and Tools. – 1994. – № 15.
10. Ахо А., Сети Р., Ульман Дж. Компиляторы. Принципы, технологии, инструменты. – М.: Вильямс, 2003.
11. Luo Z. Computation and Reasoning: A Type Theory for Computer Science. – Oxford University Press, 1994.
12. Simonyi C. The Death of Computer Languages, the Birth of Intentional Programming. /The Future of Software. Univ. of Newcastle upon Tyne. England. Dept. of Computing Science, 1995.
13. Красильников Н.Н., Шалыто А.А. Мультиагентная система дорожного движения. Реализации на языке Java и текстовом языке автоматного программирования. СПбГУ ИТМО, 2007. <http://is.ifmo.ru>, раздел "Автоматный язык".

ВВЕДЕНИЕ В ВЕРИФИКАЦИЮ АВТОМАТНЫХ ПРОГРАММ НА ОСНОВЕ МЕТОДА *MODEL CHECKING*

С. Э. Вельдер, А. А. Шалыто

В статье применительно к автоматным программам (их поведение описывается одним конечным автоматом) излагается техника верификации, которая базируется на *темпоральных логиках* и называется *Model Checking*. Для автоматных программ удастся автоматизировать процесс построения модели программы, подлежащей верификации.

In this article the verification of simple automata-based programs (which behavior can be described with a single finite automaton) is considered. The applied verification technique is based on temporal logics and is known as Model Checking. For automata-based programs it is possible to automate the process of building program model subject to verification.

Введение

Model Checking – это автоматизированный подход, позволяющий для заданной модели поведения системы с конечным (возможно, очень большим) числом состояний и логического свойства (требования) проверить, выполняется ли это свойство в рассматриваемых состояниях данной модели.

Алгоритмы для *Model Checking* обычно базируются на полном переборе пространства состояний модели. При этом для каждого состояния проверяется, удовлетворяет ли оно сформулированным требованиям. Алгоритмы гарантированно завершаются, так как модель программы конечна.

Принципиальная схема *Model Checking* приведена на рис. 1.



Рис. 1. Model Checking

В проблематике верификации [1] сформировалось два направления: аксиоматическое и алгоритмическое. При использовании первого из них разрабатывается набор аксиом, с помощью которого может быть описана как сама система, так и ее свойства [2]. Основу второго направления составляет метод *Model Checking*.

Перечислим достоинства этого метода.

1. Эффективность. Программы для верификации моделей способны работать с большими пространствами состояний благодаря концепции упорядоченных двоичных разрешающих деревьев [3], которая также упоминается в данной работе.
2. Возможность генерации контрпримеров.

Перечислим ограничения рассматриваемого метода.

1. Поддержка только моделей с конечным числом состояний. Поэтому для большинства классов систем с бесконечным числом состояний необходимо выполнять формальную верификацию системы – математическое доказательство свойств самой программы, а не ее модели.
2. Ограниченность верификации. С использованием метода *Model Checking* проверяется модель системы вместо реальной системы. Таким образом, любое применение метода *Model Checking* настолько же качественно, как и сама модель системы. Кроме того, с помощью этого метода проверяются не любые свойства модели, а только темпоральные.
3. Для многопроцессорных систем размер пространства состояний в худшем случае пропорционален произведению размеров пространств состояний их индивидуальных компонент. Этот эффект называется проблемой показательного (экспоненциального) взрыва состояний (*state-space explosion problem*).

В рамках данной работы рассматривается автоматное программирование (программирование с явным выделением управляющих состояний) [4, 5], и поэтому ограничения 1 и 3 в нашем случае несущественны.

В заключение раздела отметим, что настоящая работа выполнялась «параллельно» с работой [6], которая появилась после знакомства ее авторов с автоматным подходом.

Технология верификации одного класса автоматных программ

Технология автоматного программирования использует такие модели [7], как автомат *Мили* [8], автомат *Мура* [9] и смешанный автомат, которые легко интерпретируются с помощью моделей *Крипке*.

Первым шагом в процессе верификации автоматной программы является преобразование графа переходов исходного автомата в модель *Крипке*, для которой удобно формулировать проверяемые свойства программ. В данной работе отдается предпочтение автомату *Мили*, который при необходимости всегда может быть преобразован в автомат *Мура*.

Исследования в данной области (моделирование автомата и конвертация его в структуру *Крипке*) проводились в работах [10–13]. При этом конвертация была сопряжена со следующими проблемами:

- трудности с выполнением композиции автоматов;
- неоднозначность интерпретации формулы языка *Computational Tree Logic (CTL)* [14] в исходном автомате.

При решении первой проблемы, как правило, возникала вторая. Для ее решения применялась модификация языка *CTL*.

Методы моделирования, рассматриваемые в настоящей работе (в частности, «редуцированная» схема), проводят изменение семантики языка *CTL* для того, чтобы воспрепятствовать экспоненциальному росту числа состояний. При этом пути, построенные в качестве сценариев для *CTL*-формул, однозначно преобразуются из модели в автомат *Мили*. Это удобно, особенно если моделирование производится совместно с исполнением автомата, его визуализацией и отладкой [15, 16].

Кроме того, при использовании автоматного программирования число управляющих состояний относительно невелико. Это позволяет не применять в данной работе

специальные техники для сжатия автоматов с большим числом состояний (упорядоченные двоичные разрешающие диаграммы), а использовать достаточно простые и наглядные алгоритмы, которые работают быстро при небольшом числе состояний, один из которых будет дополнен алгоритмом генерации сценариев.

Перечислим основные положения настоящей работы.

1. В автоматных программах [4, 5] поведение специфицируется с помощью конечных автоматов. В настоящей работе в основном применяются спецификации, состоящие из одного автомата *Мили*. В общем случае модель может состоять из нескольких взаимодействующих автоматов. Для верификации таких систем применяется композиция исходных автоматов или моделей *Крипке*.
2. Использование подхода *Model Checking* для таких программ связано с преобразованием автомата *Мили* в структуру *Крипке* (модель *Крипке*), так как она, в отличие от автомата, приспособлена для верификации.
3. Использование структуры *Крипке* предполагает применение темпоральной логики для записи требований, которые должны быть проверены. В настоящей работе при написании программ верификации требования описываются на языке *CTL*.
4. Собственно верификация модели (рис. 1) выполняется по структуре *Крипке*, построенной по автомату *Мили*, и требованиям, записанным в виде формулы на языке темпоральной логики *CTL*.

Верификация осуществляется с использованием двух алгоритмов. Первый из них предназначен для определения набора состояний в структуре *Крипке*, в которых выполняется заданное формулой требование, а второй – по заданному исходному состоянию и подформуле заданного требования с помощью построенного набора состояний формирует сценарий, который подтверждает или опровергает эту подформулу.

5. Сценарий для структуры *Крипке* преобразуется в сценарий для автомата *Мили*.
6. Все этапы изложенной технологии верификации рассматриваемого класса автоматных программ иллюстрируется на примере программы для «Универсального инфракрасного пульта для бытовой техники» [17].

Таким образом, схематично процесс, исследуемый в данной работе, можно представить так, как изображено на рис. 2.

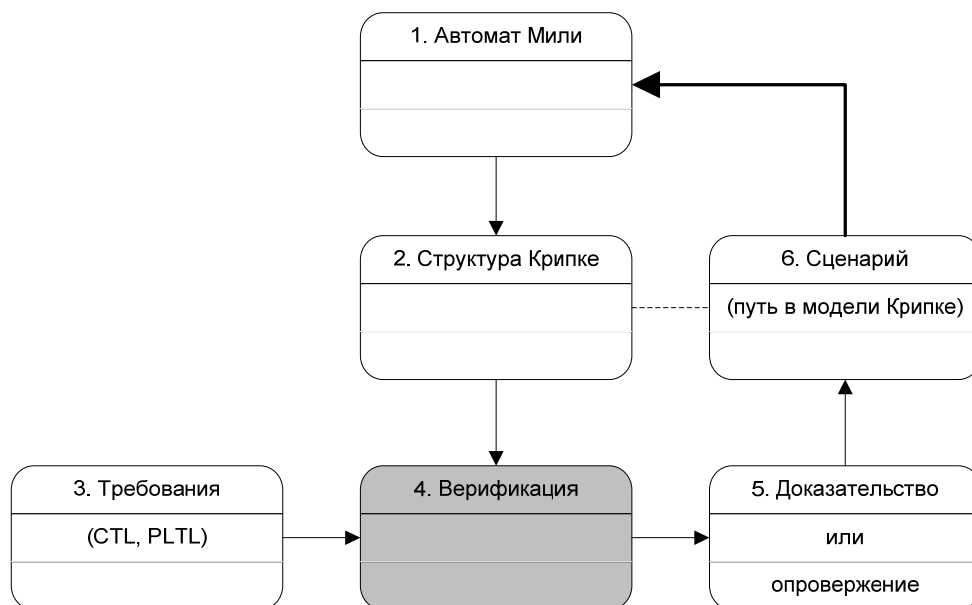


Рис. 2. Этапы верификации автоматных программ

Преобразование автомата *Мили* в структуру *Крипке* и разработка требований

В этом разделе будут построены несколько различных схем для генерации множества атомарных предложений автоматной программы, преобразования автомата *Мили* в модель *Крипке* и записи требований к программе.

Выделим три основных схемы такого преобразования:

1. Установка состояний на событиях и выходных воздействиях (переменных).
2. Создание полного графа переходов.
3. Редукция полного графа переходов с внесением тесных отрицаний внутрь атомарной формулы.

Каждая из этих схем будет описана в данном разделе.

Учтем для любой схемы, что если конечная формула ее спецификации представима в виде конъюнкции нескольких подформул, то эту конъюнкцию целесообразно разбивать на операнды и рассматривать их по отдельности, так как при этом удобнее (и правильнее) исследовать, адекватны ли формальные требования к модели соображениям разработчика о них.

При этом отметим, что автоматы, в которых состояния могут содержать внутри себя другие автоматы, можно исследовать тремя способами:

1. Для внешних и внутренних автоматов можно выполнять моделирование, спецификацию и верификацию независимо (конечно, этот способ влечет утрату определенных характеристик автомата при моделировании).
2. Также можно «раскрыть» состояние S автомата A , внутри которого (состояния) находится другой автомат B , добавив для каждого перехода из состояния S в состояние T по одному эквивалентному переходу из каждого состояния автомата B в состояние T . Все переходы, которые ведут в состояние S , следует перенаправить в стартовое состояние автомата B . В результате, внутренний автомат B превращается в часть автомата A , и для него можно выполнять верификацию вместе с автоматом A .
3. Систему взаимодействующих автоматов можно привести к одному автомату с помощью композиции (произведения) [10]. Также можно выполнять сначала моделирование каждого автомата, а после него – композицию моделей *Крипке* (рис. 3).

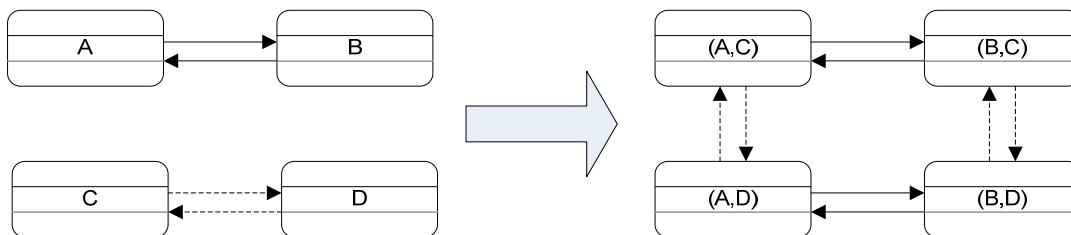


Рис. 3. Композиция структур *Крипке*

Выбор способа определяется соображениями эффективности и удобства. В примерах, исследуемых далее, считается, что данный вопрос уже решен, и рассматриваются системы, описываемые одним автоматом без вложенных состояний.

Во всех трех схемах, которые будут построены, состояния исходного автомата изоморфно перейдут в состояния модели.

Для каждого перехода между состояниями S и T исходного автомата создадим не менее одного состояния в модели *Крипке* (назовем его *состоянием-событием*), атомарным предложением которого будет событие E , инициировавшее переход. При наличии выходных воздействий на переходе также создадим по одному состоянию на каждое воздействие Z , атомарным предложением которого (состояния) будет Z (такие состояния будем называть *состояниями-выходными воздействиями*). Добавим в модель переходы: между состоянием S и состоянием-событием E ; между состоянием-событием E и первым

состоянием-выходным воздействием; далее последовательно (в порядке выполнения) между соседними состояниями для выходных воздействий, и, наконец, между последним таким состоянием-выходным воздействием и состоянием T (далее будет приведен пример такой конвертации).

Если выходное воздействие Z размещалось в состоянии T и выполнялось при входе в него, то при конвертации добавляется еще одно состояние, соответствующее воздействию Z . В это состояние должен вести каждый переход, который первоначально вел в состояние T . Кроме этого, добавляется переход из состояния, соответствующего Z , в состояние T . Само же это воздействие после генерации состояний уничтожается.

Для всех полученных состояний модели *Кринке* естественным образом устанавливаются атомарные предложения. Добавим также три «управляющих» атомарных предложения: *InState*, *InEvent*, *InAction* – для состояний модели, построенных, соответственно, из *состояний*, *событий*, *выходных воздействий* исходного автомата. Это сделано для того, чтобы при записи формулы в темпоральной логике можно было различать тип исследуемого состояния.

Таким образом, множество атомарных предложений во всех трех схемах содержит объединение множеств состояний, событий, выходных воздействий и трех описанных выше атомарных предложений.

Далее будут рассмотрены индивидуальные особенности каждой из трех схем.

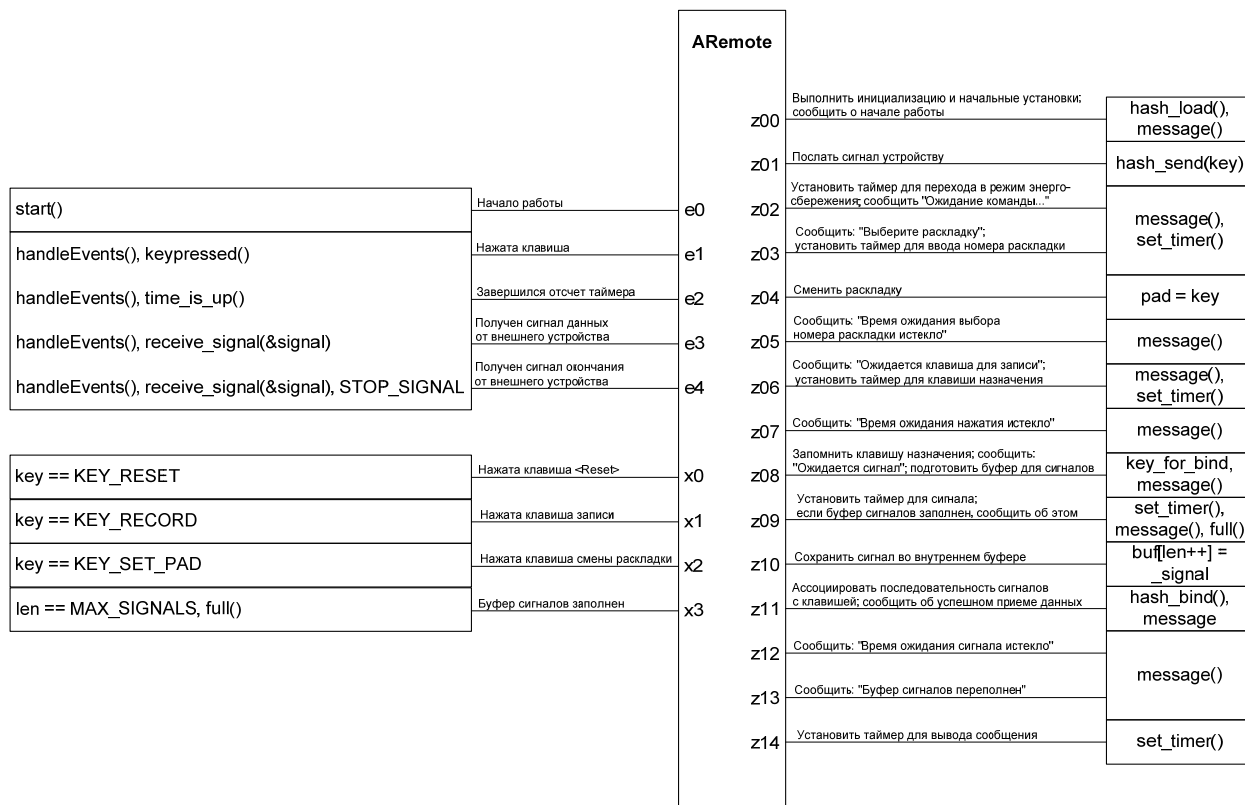


Рис. 4. Схема связей автомата *ARemote*

Схема «Состояния на событиях и выходных воздействиях» (CCBB). Данная схема, как и две другие, наследует общую идеологию моделирования, описанную выше. От других схем ее отличает то, что кроме указанного общего принципа в ней больше ничего не содержится. Таким образом, применяя схему *CCBB* для автоматной программы, можно полностью абстрагироваться от понятия входных переменных, оставляя только состояния (без них не обойдется ни одна базовая модель), события и выходные воздействия. Это самый простой подход.

Рассмотрим **пример**. Пусть исходное автоматное приложение эмулирует (в довольно упрощенной форме) универсальный инфракрасный пульт для бытовой техники [17]. Эмулятор представлен с помощью одного автомата *ARemote*. Схема связей этого автомата приведена на рис. 4.

Граф переходов этого автомата приведен на рис. 5.

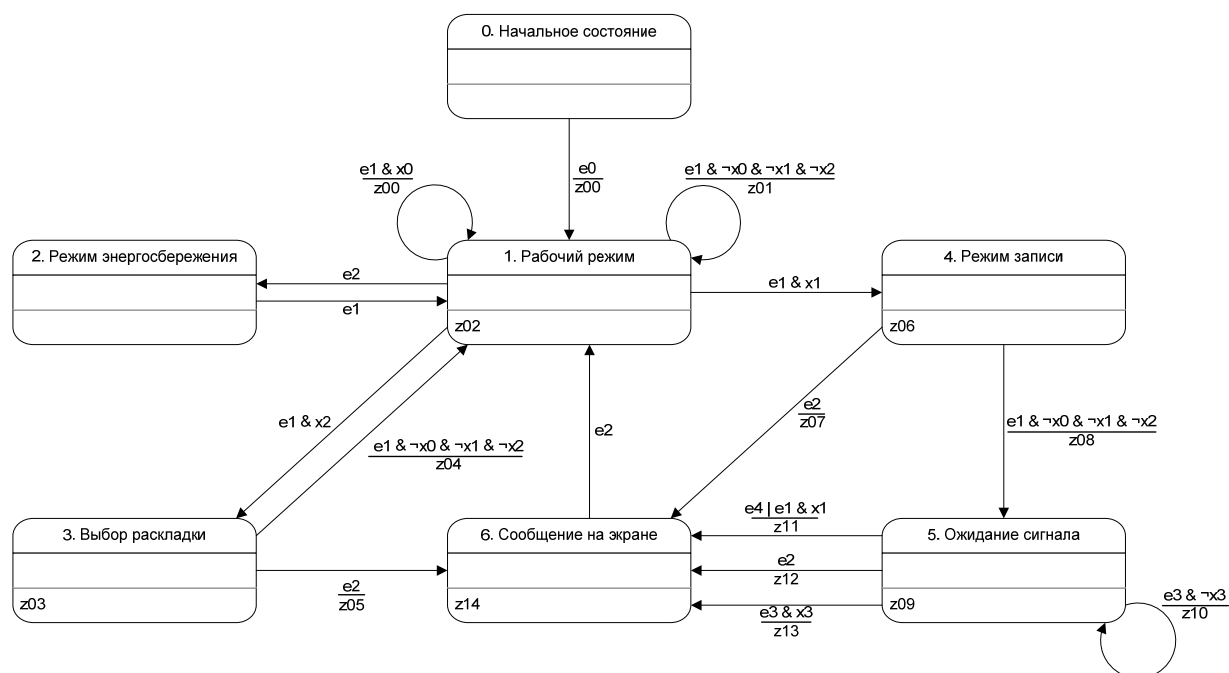


Рис. 5. Граф переходов автомата *ARemote*

В рассматриваемом примере модель *Крипке* для автомата, построенная по схеме *ССВВ*, будет изоморфна графу на рис. 6.

В модели *Крипке*, изображенной на рис. 6, состояния-события и состояния-выходные воздействия указаны явно. При интерактивном моделировании совместно с исполнением и визуализацией [14, 15] их целесообразно обозначать, как и в исходном автомате, в виде меток на дугах. Таким образом, модель построена.

Схема «Полный автомат». Во второй схеме не будем абстрагироваться от входных переменных, а представим автомат моделью *Крипке* «со всей полнотой» относительно входных воздействий. В исходном автомате переходы могут быть заданы не полностью – могут существовать не указанные петли. Это означает, что для некоторого состояния (некоторых состояний) дизъюнкция формул, составленных из входных переменных, которые помечают переходы из него по одному и тому же событию *E*, не является тавтологией.

Снабдим это состояние (состояния) петлевыми переходами по событию *E*, соответствующими дополнению к рассматриваемой дизъюнкции. Это, конечно же, не изменит семантику автомата, а лишь полностью опишет его поведение. В конечном счете, в автомате из каждого состояния по каждому событию должно исходить 2^n переходов, где n – общее число входных переменных автомата. При этом каждому переходу соответствует набор значений всех переменных. После получения полного автомата преобразуем его в модель *Крипке* по общей схеме с одной модификацией: для каждого состояния-события добавим во множество его атомарных предложений набор входных переменных, истинных на том переходе, на котором находится рассматриваемое состояние-событие. Таким образом, во множество атомарных предложений по отношению к обобщенной схеме добавились еще и входные переменные. Достоинство такой схемы (несмотря на ее расточительность, освобождение от которой будет описано ниже) в том, что она и только она позволяет модели *Крипке* **полностью** отражать поведение исходного автомата.

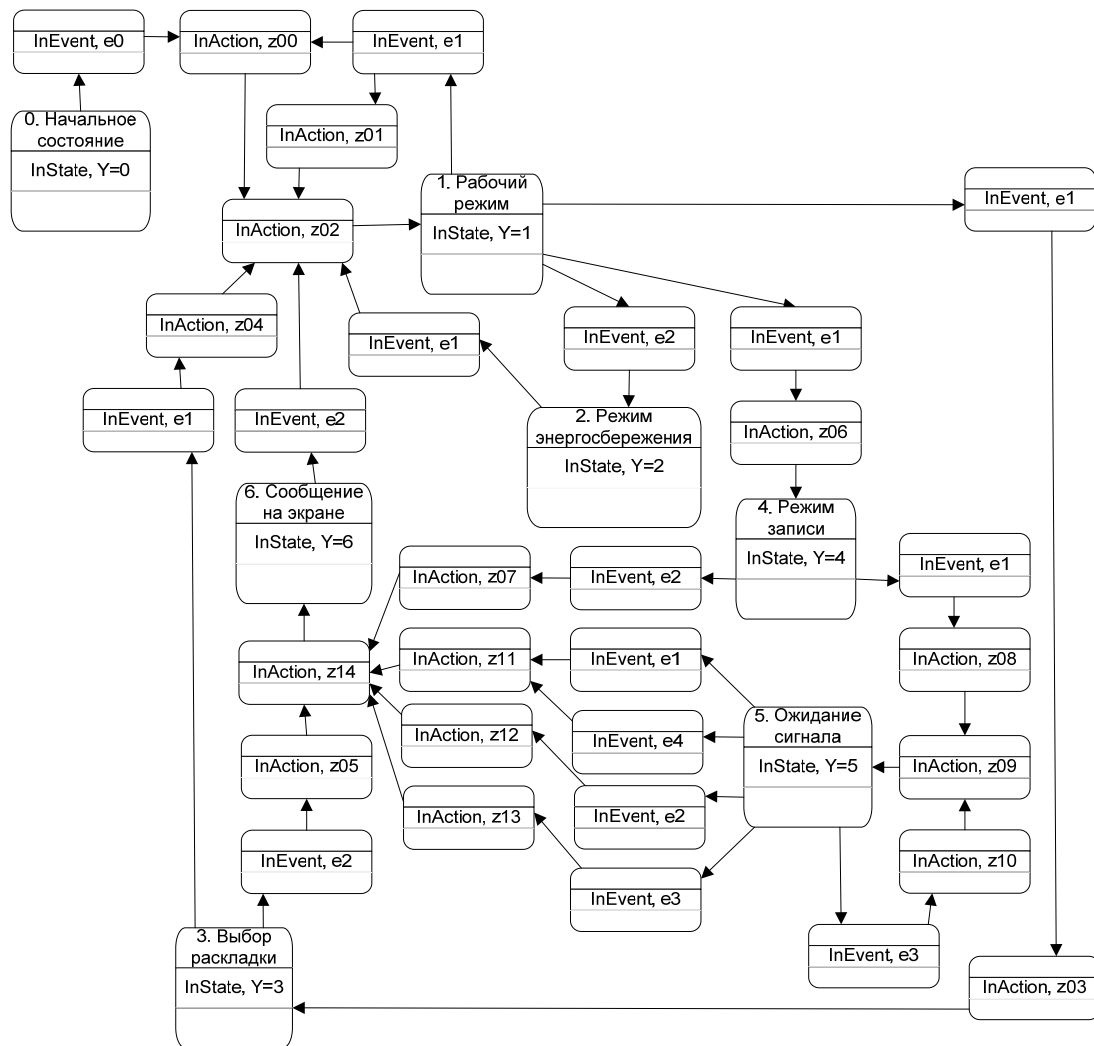


Рис. 6. Модель *Крипке*, построенная по схеме *CCBB*

Приведем теперь **пример** *CTL*-формулы, справедливость которой можно устанавливать верификацией: $\neg E[\neg(Y=6)U(Y=1)]$. Смысл этой формулы состоит в следующем: в состояние 1 нельзя попасть, минуя состояние 6 (нельзя попасть в рабочий режим, минуя сообщение на экране). Эта формула справедлива для состояний 4, 5, 6 *исходного автомата* и только для них (для модели *Крипке* таких состояний больше).

«**Редуцированная**» схема. Основным недостатком предыдущей схемы было большое число генерируемых состояний для модели *Крипке*, а достоинством – ее полнота.

В «редуцированной» схеме семантика моделей будет изменена таким образом, чтобы число состояний в них можно было уменьшить, не потеряв при этом их выразительные возможности. Это можно сделать так, что размер модели изменится асимптотически *линейно* по отношению к размеру графа переходов исходного автомата и к числу переменных (*билинейно*), в отличие от предыдущей схемы, где размер модели увеличивался *экспоненциально* от числа входных переменных.

Множество атомарных предложений по отношению к предыдущей схеме также будет видоизменено.

Рассмотрим исходный автомат без дизъюнкций на переходах. Если такие переходы существуют, создадим эквивалентные переходы для каждого дизъюнкта, а сами переходы с дизъюнкциями удалим. В качестве **примера** можно разбить переход

“(e4 | e1&x1) / z11” графа *ARemote*, изображенного на рис. 5, на два перехода: “e4 / z11” и “e1&x1 / z11”.

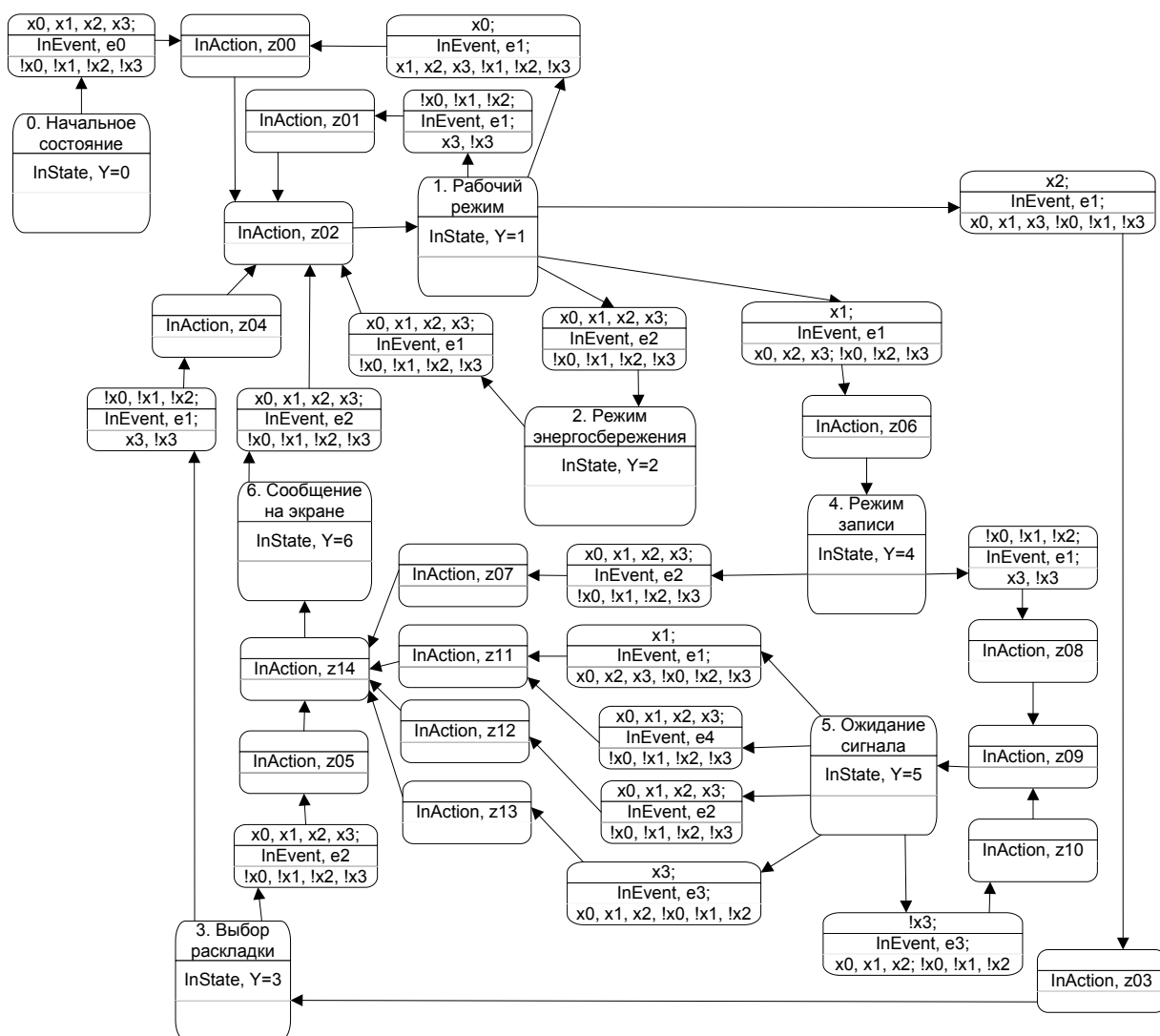


Рис. 7. Редуцированная модель *Крипке* для автомата *ARemote*

Добавим в автомат состояния, соответствующие событиям, входным и выходным переменным так, как это было сделано в первой схеме (*CCBB*), но с одним отличием: во множества атомарных предложений на состояниях-событиях добавим входные переменные **в том виде, в котором они присутствуют на переходах** (вместе с отрицаниями, если они есть). Таким образом, в состав множества всех атомарных формул модели входят следующие элементы, и только они: состояния; события; выходные воздействия; все литералы, составленные из входных переменных (сами переменные и их отрицания). Кроме того, в атомарные предложения каждого полученного состояния-события добавим **все литералы, составленные из несущественных входных переменных для данного перехода** (несущественными будем называть те переменные исходного автомата, которые не обозначены на рассматриваемом переходе). Таким образом, будем добавлять на одно и то же состояние-событие и несущественные переменные, и их отрицания. С точки зрения синтаксиса и семантики темпоральной логики это допустимо: процесс обработки модели *Крипке* не предполагает совместность множества атомарных предложений состояния, так как интерпретирует эти предложения просто как строки. Причина такого обращения с несущественными переменными ясна: требуется обеспечить, чтобы любая

ссылка на несущественную в данном состоянии-событии переменную, упомянутая в *CTL*-формуле, давала истинный результат.

Не обязательно хранить все литералы, составленные из несущественных переменных в состоянии в явном виде. Важно лишь то, что во время обработки модели существенные и несущественные входные переменные интерпретируются раздельно: первые – в том виде, в каком они записаны на переходах исходного автомата, а вторые – «в двух экземплярах» (в прямом и инверсном виде).

Результат конвертации графа переходов автомата *ARemote*, выполненного с применением данной схемы, изображен на рис. 7.

Размер модели на рис. 7 совпадает с размером модели, созданной по схеме «Состояния на событиях и выходных воздействиях». Вообще говоря, первая схема может рассматриваться по аналогии с третьей или второй, в которой полностью исключены входные воздействия. Аналогично, третью схему можно рассматривать как видоизменение второй, при котором отождествляются наборы значений несущественных переменных.

Теперь рассмотрим построение и интерпретацию *CTL*-формул для «редуцированных» моделей.

CTL-семантика в данной схеме будет немного отличаться от общепринятой: все отрицания, стоящие непосредственно перед атомарными предложениями в *CTL*-формуле (их также называют *тесными отрицаниями*), **следует внести внутрь атомарных предложений**. При этом только результирующая формула в рассматриваемой схеме подлежит верификации методами, предназначенными для *CTL*-логики.

Рассмотрим пример для автомата *ARemote*. Пусть требуется проверить свойство: «существует способ провести инициализацию устройства, не нажимая кнопку *Reset*». В терминах языка *CTL* с исходной семантикой данное свойство может быть записано следующим образом: $E[(InEvent \rightarrow \neg x0) \cup z00]$.

Эта формула не выполняется в состоянии $Y=0$ (рис. 7). На это, правда, и не стоило рассчитывать. Преобразуем формулу согласно третьей схеме: $E[(InEvent \rightarrow !x0) \cup z00]$. Вместо отрицания в языке *CTL* в формулу было внесено другое атомарное предложение, являющееся отрицанием исходного. Преобразованная формула уже верна для состояния $Y=0$.

Подведем итог. Для уменьшения числа состояний и из соображений практичности была предложена схема моделирования автомата и изменена семантика языка *CTL*. Однако такое изменение семантики неудобно для верификации. В результате был предложен способ преобразования исходной формулы, соответствующей новой семантике, в новую формулу, для которой применима общепринятая семантика языка *CTL*.

Выполненные примеры показывают, что такой подход не существенно снижает выразительность модели по сравнению с предыдущим (схема «Полный автомат»). Опыт показывает, что для многих формул такая схема может быть использована.

Другие абстракции. Основным недостатком всех описанных выше схем моделирования автоматов является то, что при составлении требований к модели, разработчику не всегда удобно различать, где состояния, которые перенесены из исходного графа, где состояния-события, а где состояния-выходные воздействия. Для различения состояний используются атомарные предложения *InState*, *InEvent* и *InAction*, но их применение может быть связано с дополнительными проверками. Для этого, а также для уменьшения числа состояний модели в принципе, можно при построении модели абстрагироваться от каких-либо других ее характеристик, помимо тех, которые были рассмотрены в описанных выше схемах.

Например, можно абстрагироваться не только от входных переменных, но и от событий, а также от выходных воздействий. Можно вообще преобразовать автомат в модель *Крипке* в один этап, что может быть сделано, например, с помощью исклю-

чения событий и выходных переменных на переходах. Для автомата *ARemote* результатом такого преобразования является модель на рис. 8.

Выбор альтернативного метода можно осуществлять, руководствуясь представлениями о производительности и результативности. Основное внимание необходимо уделять атомарности переходов. Если они слишком большие (по числу действий), то разработчик может пропустить ошибку, если же слишком маленькие – то размер модели может немotивированно увеличиться за счет появления несущественных свойств.

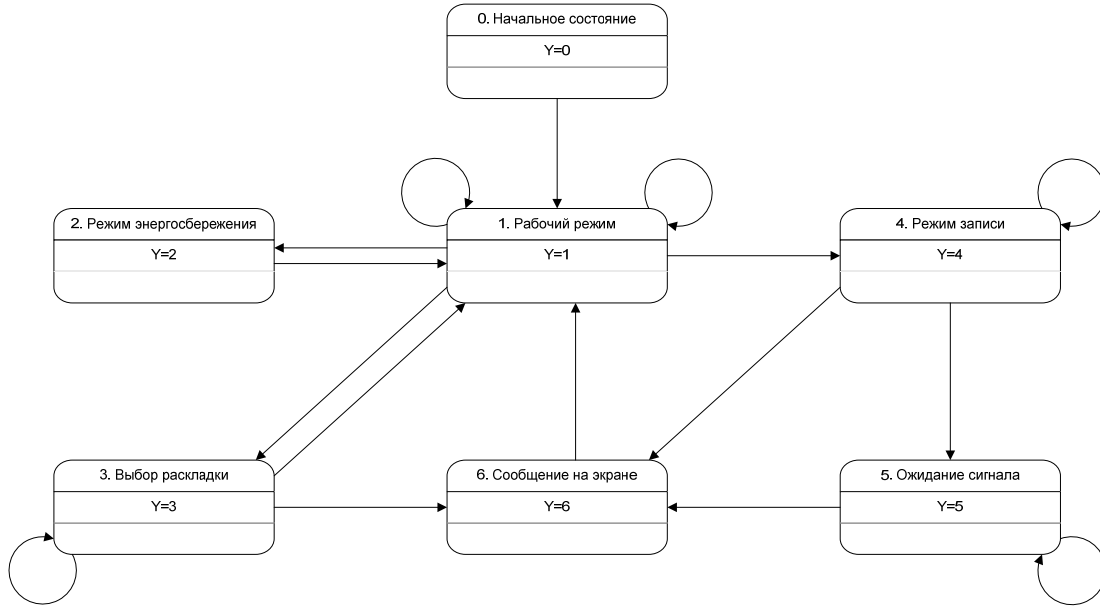


Рис. 8. Сокращенная модель *ARemote* (без событий и выходных переменных)

CTL-верификация автоматных программ

В данном разделе будет описана идея алгоритма *CES* (*Clarke, Emerson, Sistla*) [18], который основан на переформулировке синтаксиса языка *CTL*. Этот алгоритм дополнен таким образом, что позволяет строить подтверждающие сценарии для проверяемых формул. Применять этот алгоритм будем для изображенных явно моделей *Krinke*.

Под локальной задачей верификации обычно понимается вопрос: выяснить для данной модели и состояния в ней, выполняется ли в этом состоянии заданная формула.

При построении алгоритма формулируется глобальная задача верификации: для данной модели и проверяемой формулы построить множество всех состояний модели, в которых верна эта формула.

Когда число состояний невелико, как в случае автоматных программ, это множество можно строить в явном виде.

Запишем одну из форм определения синтаксиса и семантики языка *CTL* (в ней темпоральная часть будет целиком выражена через операции **EX**, **EU**, **EG**):

$$\begin{aligned}
 \mathbf{A} (f \mathbf{U} g) &= \neg (\mathbf{E}[(\neg g) \mathbf{U} \neg (f \vee g)] \vee \mathbf{EG} \neg g) \\
 \phi &::= p \in AP \mid \neg \phi \mid \phi \vee \psi \mid \mathbf{EX} \phi \mid \mathbf{E}[\phi \mathbf{U} \psi] \mid \mathbf{EG} \phi.
 \end{aligned}$$

(здесь AP – множество атомарных предложений)

CTL-моделью для множества состояний S называется тройка:

$$M = (S; R \subseteq S \times S; Label \subseteq S \times AP).$$

Здесь R – тотальное отношение на множестве S (отношение переходов между состояниями), а $Label$ – отношение, определяющее атомарные предложения, соответствующие каждому состоянию.

Множество выполняющих состояний алгоритм строит для каждой подформулы входной формулы (для каждого состояния создается список выполненных в нем подформул). Идея алгоритма отражена в псевдокоде (рис. 9).

```

function  $Sat(\phi : Formula) : \text{set of } State;$ 
begin
  if       $\phi = 1$        $\rightarrow$  return       $S$ 
         $\phi = 0$        $\rightarrow$  return       $\emptyset$ 
         $\phi \in AP$      $\rightarrow$  return       $\{s \mid Label(s, \phi)\}$ 
         $\phi = \neg \phi_1$   $\rightarrow$  return       $S \setminus Sat(\phi_1)$ 
         $\phi = \phi_1 \vee \phi_2$   $\rightarrow$  return  $Sat(\phi_1) \cup Sat(\phi_2)$ 
         $\phi = EX \phi_1$    $\rightarrow$  return  $\{s \in S \mid \exists (s, s') \in R \mid s' \in Sat(\phi_1)\}$ 
         $\phi = E[\phi_1 U \phi_2]$   $\rightarrow$  return  $Sat_{EU}(\phi_1, \phi_2)$ 
         $\phi = EG \phi_1$    $\rightarrow$  return  $Sat_{EG}(\phi_1)$ 
  end if
end

```

Рис. 9. Индукция по построению формулы в алгоритме CES

Как следует из рассмотрения текста этой программы, множество состояний, выполняющих формулу ϕ , строится индукцией по построению ϕ .

Будем считать для удобства, что из исходного графа *Кринке* построен симметричный ему граф – граф, в котором все переходы заменены на противоположные. В алгоритме CES нетривиальными являются последние два шага, которые могут быть реализованы с помощью построения деревьев «обратных путей» и определением компонент сильной связной связности у графа модели.

Теперь осталось только дополнить этот алгоритм методами *предоставления подтверждений* истинности формул в моделях. Иными словами, требуется построить способ генерации сценариев.

Алгоритм генерации сценария для CTL-формулы f в модели *Кринке*. Итак, требуется *показать*, что в данном состоянии s модели M выполняется (или не выполняется) формула f .

1. Если f – атомарное предложение, то просто предъявим описание состояния s в модели M – множество его атомарных предложений. В нем, в частности, содержится информация о выполнимости формулы f в данном состоянии s .
2. Доказательство $\neg f$ сводится к опровержению формулы f , и наоборот.
3. Для доказательства формулы $f \vee g$ докажем одну из формул f или g , а для опровержения – опровергаем обе формулы f и g .
4. Для доказательства $EX f$ предъявим вершину в модели *Кринке*, в которую из вершины s имеется переход и которая выполняет f . Такая вершина обязательно существует, иначе на этапе верификации не обнаружилось бы, что формула $EX f$ верна. Опровержение $EX f$ (доказательство $AX \neg f$) подтверждается весьма просто, так как **любой** переход, который ведет из вершины s , будет вести только в вершину, выполняющую $\neg f$. Таким образом, любой переход из этой вершины можно предъявить пользователю в качестве опровержения.
5. Доказательство формул $E[f U g]$ и $EG f$ выполняется рекуррентным способом с использованием описанных пунктов 1–4. Выполним рекуррентное разложение для этих формул:

$$\begin{aligned} E[fU g] &= g \vee f \wedge EX E[fU g] \\ EG f &= f \wedge EX EG f \end{aligned}$$

Тогда для доказательства формулы $E[fU g]$ достаточно построить путь в графе, применяя шаг за шагом пункты 1–4 к рекуррентному разложению этой формулы, до тех пор, пока не попадем в вершину, выполняющую g .

Для доказательства формулы $EG f$ сделаем то же самое, пока не попадем в вершину, в которой уже были. Путь в этом случае, начиная с некоторого состояния, становится периодическим (рис. 10).

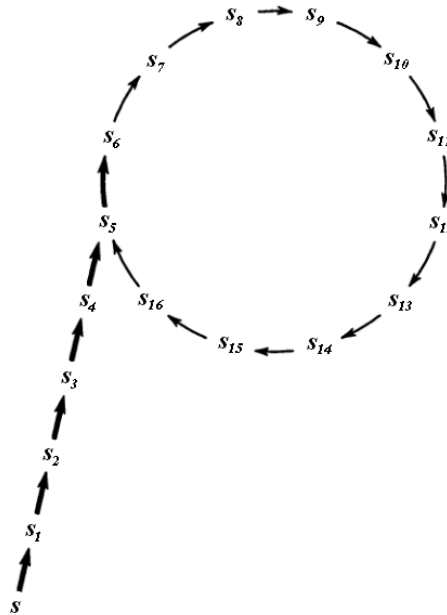


Рис. 10. Бесконечный p -путь

Опровержение формул $E[fU g]$ и $EG f$ выполняется аналогично опровержению формулы $EX f$. Любой (бесконечный) путь, который начинается в текущей вершине, можно предъявить пользователю для рассмотрения, так как путь не выполняет введенную формулу. Иначе говоря, вместо доказательства CTL^* -формул $A\neg[fU g]$ и $A\neg G f$ следует доказывать формулы $E\neg[fU g]$ и $E\neg G f$. Проще всего предъявлять пути, замыкающиеся начиная с некоторого состояния в цикл, так как такие пути однозначно задаются конечным числом вершин.

На этом изложение алгоритма завершено.

Анализ построенного алгоритма формирования сценариев, а также семантики языка CTL позволяет сформулировать следующее утверждение.

Утверждение. Если в модели Крипке существует бесконечный путь, выполняющий заданную CTL -формулу или являющийся контрпримером к ней, то существует и путь «в p -форме» (аналогично, выполняющий или опровергающий ее), представимый в виде объединения «предциклической» и «циклической» частей (рис. 10).

Доказательство этого утверждения является конструктивным и целиком опирается на применение описанного алгоритма. Достаточно только заметить, что алгоритм всегда завершается, выдавая сценарий, который необходимо обладает свойством периодичности.

Преобразование сценария для модели Крипке в сценарий для автомата Мили

Переходим к последней фазе процесса верификации в автоматном программировании (рис. 2). Ниже будет описано, как представлять путь (последовательность вершин) модели *Крипке* в виде пути исходного автомата *Мили*. При этом будем предполагать, что модель *Крипке* была сгенерирована по «редуцированной» схеме. Остальные схемы, для которых было дано описание, позволяют применять к себе аналогичный интуитивно ясный способ перехода от модели к автомату.

При использовании нестандартных методов моделирования интерпретировать результаты разработчику приходится самому – ему придется проводить анализ путей прямо на модели *Крипке*, которую он сам (вручную) и построил.

После того, как отработала программа-верификатор, необходимо определить выполнимость формул спецификации на определенных участках автомата. Среди этих участков могут быть состояния, события, выходные воздействия. Сценарий для любой подформулы спецификации представляет собой бесконечный путь в модели *Крипке*, иллюстрирующий справедливость или ошибочность данной подформулы. Требуется, чтобы сценарий, предъявленный программой, был представлен в исходном автомате. Изображать этот путь следует конечным – вспомним утверждение из предыдущего раздела.

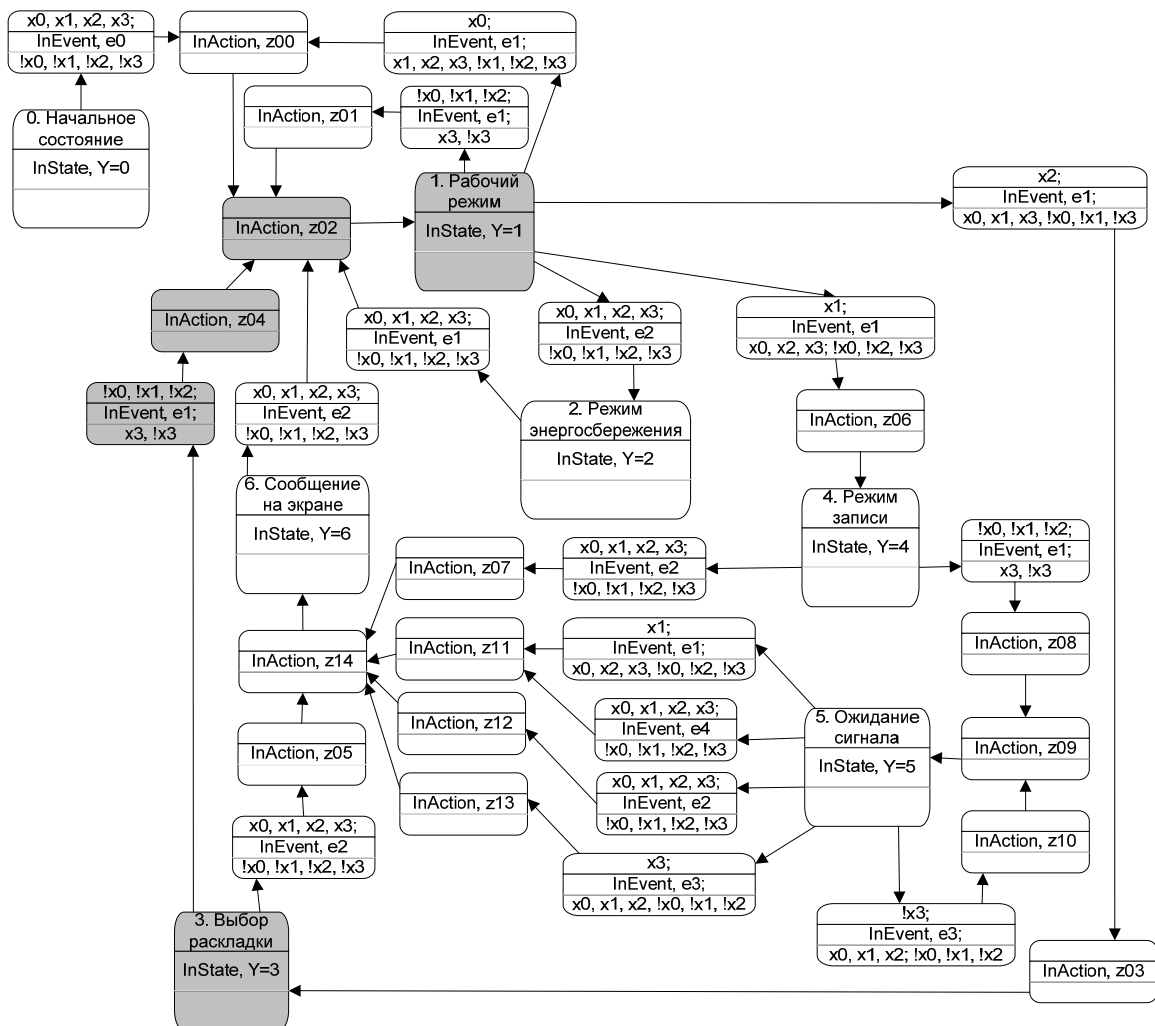


Рис. 11. Контрпример. Путь в модели *Крипке*

Что касается «переноса» пути из модели *Крипке* в автомат, то данная операция (скажем, для редуцированной схемы) выполняется однозначно. Действительно, состояния модели, содержащие атомарное предложение *InState*, однозначно преобразуются в

соответствующие им состояния автомата. Путь же между любыми двумя соседними состояниями проходит ровно через одно состояние-событие, из атомарных предложений которого можно узнать, какое событие ведет по данному пути из исходного состояния, а также значения существенных и список несущественных входных переменных в момент, когда произошло это событие. Эта информация однозначно определяет направление, вдоль которого строится путь в исходном автомате *Мили*. Если же путь (или его участок) начинается не в состоянии *InState*, то обратная трассировка пути позволяет узнать состояние *InState*, предшествующее текущему, и всю необходимую информацию относительно того, как попасть в текущее состояние.

Рассмотрим **пример** для автомата *ARemote*. Пусть для состояния 3 выполняется верификация формулы $\neg E[\neg(Y=6)U(Y=1)]$ (в состоянии 1 нельзя попасть, минуя состояние 6). Эта формула в состоянии 3 не выполняется. Верификатор сгенерировал (кратчайший и единственный в данном случае) контрпример, который на рис. 11 выделен серым цветом. Это конечный путь, любое продолжение которого удовлетворяет формуле $\neg E[\neg(Y=6)U(Y=1)]$.

Этот же путь, но представленный в исходном автомате, можно увидеть на рис. 12.

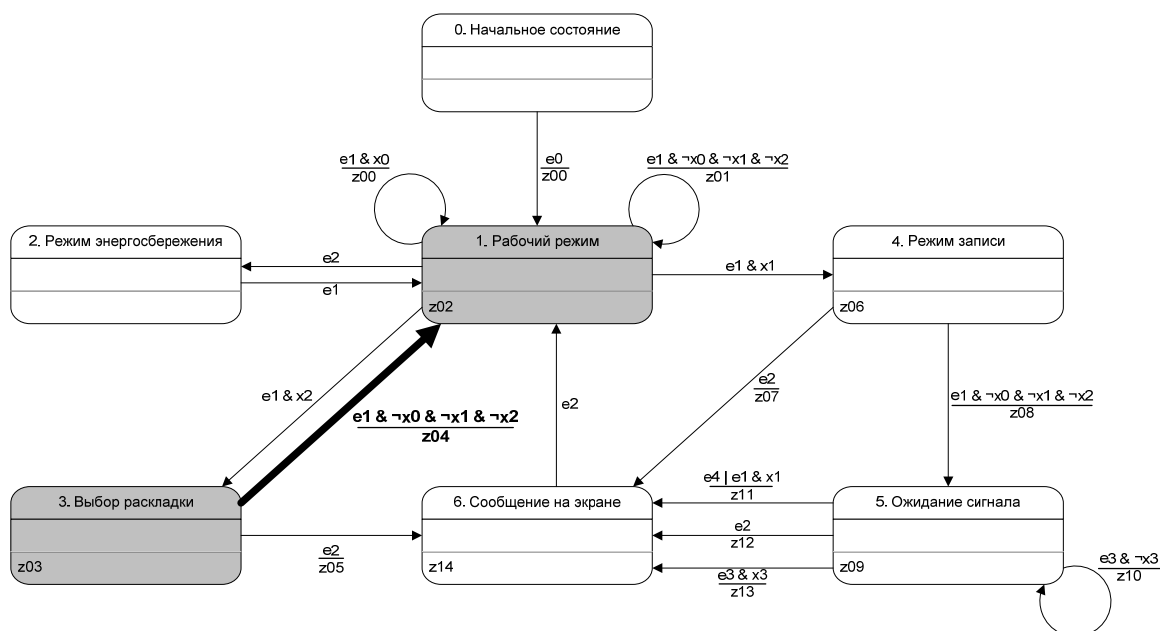


Рис. 12. Контрпример. Путь в автомате *Мили*

В случае, когда при моделировании выполнялась композиция автоматов/моделей *Крипке*, независимая нумерация их состояний позволит для каждого перехода в пути, представленном в окончательной модели, однозначно решить вопрос о том, в какой именно индивидуальной компоненте системы взаимодействующих автоматов произошел переход. Это, опять же, дает возможность отобразить путь на модели в путь на исходном автомате.

Заключение

В данной работе были предложены методы для моделирования автомата *Мили* структурами *Крипке*. Был разработан алгоритм для построения сценариев и их интерпретации в исходном автомате. В связи с созданием этого алгоритма было сформулировано утверждение, позволяющее привести все сценарии к общему виду.

Составление сценариев (в том числе, контрпримеров) с помощью верифицирующих инструментов позволяет проводить исследования в области *автоматической или интерактивной коррекции* модели или автомата с целью удовлетворить предъявляе-

мым условиям. Например, если программа-верификатор предъявила путь, опровергающий некоторое желательное свойство для системы, она может предложить разработчику исказить/ликвидировать этот путь, например, за счет удаления какого-либо перехода. При этом, разумеется, не гарантируется, что в модели при этом не возникнет других противоречий со спецификацией, хотя, не исключается возможность и более интеллектуальной коррекции.

Исходя из изложенного, можно кратко сформулировать основные достоинства автоматных программ в части их верификации [19]:

1. Класс автоматных программ является наиболее удобным для верификации методом *Model checking*, так как в этом случае модель программы может быть автоматически построена по спецификации ее поведения, задаваемой в общем случае системой взаимодействующих конечных автоматов, в то время как для программ других классов модель приходится строить вручную.
2. Структура автоматных программ, в которых функции входных и выходных воздействий почти полностью отделены от логики программ, делает практичным верификацию этих функций на основе формальных доказательств с использованием пред- и постусловий [20, 21].

Литература

1. Джексон Д. Программы проверяют программы // В мире науки. – 2006. – №10. – С.52–57.
2. Вудкок Дж. Первые шаги к решению проблемы верификации программ //Открытые системы. – 2006. – № 8. – С. 36–57.
3. Katoen J.-P. Concepts, Algorithms, and Tools for Model Checking. Lehrstuhl für Informatik VII, Friedrich-Alexander Universität Erlangen-Nürnberg. Lecture Notes of the Course (Mechanised Validation of Parallel Systems) (course number 10359). Semester 1998/1999.
4. Шалыто А. А. SWITCH-технология. Алгоритмизация и программирование задач логического управления. – СПб.: Наука, 1998. <http://is.ifmo.ru/books/switch/1>
5. Шалыто А. А. Логическое управление. Методы аппаратной и программной реализации. – СПб.: Наука, 2000. http://is.ifmo.ru/books/log_upr/1
6. Кузьмин Е. В., Соколов В. А. Верификация автоматных программ с использованием LTL. / Моделирование и анализ информационных систем. – Ярославль: ЯрГУ. 2007. – № 1. – С. 3–14. <http://is.ifmo.ru>, раздел «Верификация».
7. Finite state machine. http://en.wikipedia.org/wiki/Finite_state_machine
8. Mealy machine. http://en.wikipedia.org/wiki/Mealy_machine
9. Moore machine. http://en.wikipedia.org/wiki/Moore_machine
10. Sebastiani R. Introduction to Formal Methods, 2005–2006. http://dit.unitn.it/~rseba/DIDATTICA/fm2005/02_transition_systems.pdf
11. Margaria T. Model Structures. Service Engineering – SS 06. <https://www.cs.uni-potsdam.de/sse/teaching/ss06/sveg/ps/2-ServEng-Model-Structures.pdf>
12. Roux C., Encrenaz E. CTL May Be Ambiguous when Model Checking Moore Machines. UPMC LIP6 ASIM. CHARME, 2003. <http://sed.free.fr/cr/charme2003-presentation.pdf>
13. Hull R. Web Services Composition: A Story of Models, Automata and Logics. Bell Labs, Lucent Technologies, 2004. <http://edbtss04.dia.uniroma3.it/Hull.pdf>
14. Миронов А. М. Математическая теория программных систем. <http://intsys.msu.ru/study/mironov/mthprogsys.pdf>
15. Сайт проекта UniMod. <http://unimod.sf.net>

16. Сайт eVelopers Corporation. <http://www.evelopers.com>
17. Вельдер С.Э., Бедный Ю.Д. Универсальный инфракрасный пульт для бытовой техники. – СПбГУ ИТМО. Курсовая работа. 2005. <http://is.ifmo.ru/projects/irrc/>
18. Clarke E.M., Emerson E.A., Sistla A.P. Automatic Verification of Finite-State Concurrent Using Temporal Logic Specifications // ACM Transactions on Programming Languages and Systems (TOPLAS). – 1986. – V.8.– № 2, – P. 244–263.
19. Switch-technology. <http://en.wikipedia.org/wiki/Switch-technology>
20. Дейкстра Э. Заметки по структурному программированию / Дал У., Дейкстра Э., Хоор К. Структурное программирование. – М.: Мир, 1975.
21. Мейер Б. Объектно-ориентированное конструирование программных систем. – М.: Русская редакция. 2005.

АЛГОРИТМ АДАПТАЦИИ С АСТАТИЗМОМ ПЕРВОГО ПОРЯДКА ДЛЯ СТАБИЛИЗАЦИИ ЛИНЕЙНОГО ВОЗМУЩЕННОГО НЕСТАЦИОНАРНОГО ОБЪЕКТА

А.А. Бобцов, А.Г. Наговицина, В.О. Никифоров

Данная статья посвящена проблеме управления по выходу линейными нестационарными системами с неизвестными быстроизменяющимися параметрами. Предполагая, что объект управления подвержен влиянию внешнего неизвестного постоянного возмущения, представлен подход, обеспечивающий решение задачи стабилизации выходной переменной. Теоретические результаты проиллюстрированы на практическом примере.

1. Введение

Данная работа посвящена актуальной проблеме анализа и синтеза алгоритмов адаптивного управления неопределенными нестационарными линейными системами по выходу (т.е. без измерения производных выходной переменной или вектора переменных состояния). Среди результатов, посвященных управлению нестационарными системами, следует выделить работы известных отечественных и зарубежных ученых таких как: Андреев [1], Барабанов [2], Иоанну [3 – 5], Первозванский [6], Юркевич [7] и др. Несмотря на то, что проблема управления нестационарными системами является не новой и ей посвящено множество публикаций, следует отметить, что ряд актуальных задач все еще не имеет удовлетворительных решений. Следует отметить, что на сегодняшний день получен ряд интересных результатов для линейных систем, затрагивающих проблему управления в условиях медленного изменения параметров [6], периодического изменения параметров [1], а также для случая специальных структур матриц описания нестационарных моделей [8]. Однако особое место занимают задачи управления линейными нестационарными системами с неизвестными коэффициентами.

Среди методов управления в условиях неопределенности, как правило, преобладают алгоритмы, обеспечивающие заданное поведение системы для класса математических моделей определенной структуры. К таким подходам, в частности, можно отнести схемы адаптивного и робастного управления, позволяющие решать задачи стабилизации и слежения для нестационарных объектов, в которых неопределенность согласована с управляющим входом [9 – 11]. Также можно выделить работу, посвященную стабилизации нестационарных моделей более общей структуры [2]. Однако класс неопределенностей рассматриваемых в [2], ограничен случаем, когда матрица неопределенности по норме меньше единицы. Также существует ряд публикаций, посвященных разработке адаптивных регуляторов, результаты которых основаны на предположении, что параметры объекта медленно изменяются и действуют на систему как внешнее возмущение [3, 4, 12, 13]. Позднее, анализ моделей изменения параметров и наличие некоторой априорной информации об изменении параметров привели к разработке новых алгоритмов адаптации, позволяющих управлять системами с быстрыми изменениями параметров (см., например работы [14, 15, 4]). В статье [5] указанные проблемы были решены, с использованием итеративной процедуры синтеза закона управления. Однако процедура синтеза закона управления, предложенная в работе [5], отличается достаточной сложностью, а сам регулятор обладает высокой размерностью.

В данной статье будет рассмотрен подход, позволяющий решать задачу управления по выходу линейными нестационарными системами с неизвестными ограниченными параметрами. Также следует отметить, что данная схема управления позволяет синтезировать адаптивный регулятор фиксированной размерности, которая в свою очередь, не зависит от числа неизвестных параметров как в работе [5].

2. Постановка задачи

Рассмотрим линейную нестационарную систему

$$\dot{x} = Fz + L(u + w) + \theta(t)y, \quad (1)$$

$$y = Rz, \quad (2)$$

где $z(t) \in R^n$ – вектор переменных состояния модели (1), (2); F , L и R – неизвестные постоянные матрицы, соответственно $n \times n$, $n \times 1$ и $1 \times n$ размерности; $\theta(t) \in R^n$ – вектор неизвестных переменных параметров; $y(t) \in R$ – выходная переменная системы (1), (2); $w \in R$ – неизвестное постоянное возмущение.

Будем полагать, что измеряется только выходная переменная системы (1), (2), но не ее производные, вектор состояния $z(t)$ и возмущение w не измеряются, а параметры вектора $\theta(t) \in R^n$ гладкие и ограниченные функции. Также будем допускать, что передаточная функция $H(p) = R(pI - F)^{-1}L = \frac{b(p)}{a(p)}$ минимально фазовая, т.е. полином

$b(p)$ – гурвицев, $p = \frac{d}{dt}$ – оператор дифференцирования.

Целью управления является решение задачи синтеза регулятора, обеспечивающего для любых начальных условий стремление выходной переменной $y(t)$ к нулю при $t \rightarrow \infty$.

3. Модельные преобразования и предварительные результаты

Запишем систему (1), (2) следующим образом

$$\dot{x} = Fz + L(u + w) + \sum_{i=1}^n D_i \theta_i(t)y, \quad (3)$$

$$y = Rz, \quad (4)$$

где векторы $D_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 7 \\ 0 \end{bmatrix}$, $D_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 7 \\ 0 \end{bmatrix}$, ..., $D_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 7 \\ 1 \end{bmatrix}$ имеют размерность $(n \times 1)$; $\theta_1, \theta_2, \dots$,

θ_n – компоненты вектора неизвестных переменных параметров $\theta(t) = \begin{bmatrix} \theta_1 \\ \theta_2 \\ 7 \\ \theta_{n-1} \\ \theta_n \end{bmatrix}$.

Запишем модель вход-состояние-выход (3), (4) в виде вход-выход

$$\begin{aligned} y &= \frac{b(p)}{a(p)}(u + w) + \frac{c_1(p)}{a(p)}\theta_1(t)y + \frac{c_2(p)}{a(p)}\theta_2(t)y + \dots \\ &+ \frac{c_n(p)}{a(p)}\theta_n(t)y = \frac{b(p)}{a(p)}(u + w) + \sum_{i=1}^n \frac{c_i(p)}{a(p)}\theta_i(t)y, \end{aligned} \quad (5)$$

где $p = \frac{d}{dt}$ – оператор дифференцирования; передаточная функция

$$\frac{c_i(p)}{a(p)} = R(pI - F)^{-1} D_i; \text{ размерность полинома } c_i(p) \text{ для всех } i = \overline{1, n} \text{ строго меньше}$$

размерности полинома $a(p)$; параметры $\theta_1, \theta_2, \dots, \theta_n$ являются гладкими и ограниченными функциями.

Прежде, чем приступить к синтезу управления сформулируем вспомогательный результат, опубликованный в работе [16]. Рассмотрим линейную стационарную систему

$$\dot{x}' = A'x' + B'u', \quad (6)$$

$$y' = C'x', \quad (7)$$

где $x' \in R^n$, $y' \in R$, $u' \in R$, а матрицы A' , B' и C' имеют соответствующие размерности. Передаточная функция системы (6), (7) от u' к y' определяется выражением

$$\chi(p) = C'(pI - A')^{-1} B'.$$

Пусть система (8), (9) замкнута линейной обратной связью по выходу

$$u' = -ky'. \quad (8)$$

где число $k > 0$.

Поставим вопрос о существовании положительно определенной матрицы $M = M^T$ и числа k , удовлетворяющих соотношениям

$$M(A' + kB'C') + (A' + kB'C')^T M \leq -G, \quad (9)$$

$$MB' = (C')^T \quad (10)$$

для некоторой положительно определенной матрицы $G = G^T$.

$$\text{Лемма. Пусть } \chi(p) = \frac{b'(p)}{a'(p)}, \text{ где } b'(p) = b'_{n-1}p^{n-1} + \dots + b'_0 \text{ и } a'(p) = a'_n p^n + \dots + a'_0$$

– соответственно, числитель и знаменатель передаточной функции $\chi(p)$. Пусть полином $b'(p)$ гурвицев и $b'_{n-1} > 0$, тогда существует такое число $k_0 > 0$, для которого соотношения (9), (10) разрешимы для любого $k \geq k_0$.

4. Синтез алгоритма управления

Выберем управление следующим образом

$$u = \frac{-\phi(p)}{p}(k + \lambda)y', \quad (11)$$

где число $k > 0$; положительный параметр k предназначен для компенсации неопределенностей $\sum_{i=1}^n D_i \theta_i(t)y$ и w ; полином $\phi(p)$ выбирается таким образом, чтобы полином $\beta(p) = \phi(p)b(p)$ был гурвицев и имел n порядок; функция $y'(t)$ является оценкой сигнала $y(t)$, которая формируется алгоритмом вида

$$\begin{cases} \dot{\xi}_1 = \sigma \xi_2, \\ \dot{\xi}_2 = \sigma \xi_3, \\ \dots \\ \dot{\xi}_p = \sigma(-k_1 \xi_1 - k_2 \xi_2 - \dots - k_p \xi_p + k_1 y), \\ y' = \xi_1, \end{cases} \quad (12)$$

$$y' = \xi_1, \quad (13)$$

где число $\sigma > k + \lambda$, а коэффициенты k_i рассчитываются из соображений асимптотической устойчивости модели (12) при нулевом входе y .

Подставляя (11) в уравнение (5), получаем

$$\begin{aligned} y &= \frac{b(p)}{a(p)} \left[\frac{-\phi(p)}{p} (k + \lambda) y + w \right] + \sum_{i=1}^n \frac{c_i(p)}{a(p)} \theta_i(t) y = \\ &= \frac{b(p)}{a(p)} \left[\frac{-\phi(p)}{p} (k + \lambda) y + \frac{\phi(p)}{p} (k + \lambda) \varepsilon + w \right] + \sum_{i=1}^n \frac{c_i(p)}{a(p)} \theta_i(t) y, \end{aligned} \quad (14)$$

где функция отклонений $\varepsilon(t)$ равна

$$\varepsilon = y - y^{\setminus}. \quad (15)$$

Преобразуем уравнение (14) следующим образом

$$\begin{aligned} a(p)y + k \frac{\phi(p)}{p} b(p)y &= b(p) \frac{\phi(p)}{p} [(k + \lambda)\varepsilon - \lambda y] + \\ &+ b(p)w + \sum_{i=1}^n c_i(p) \theta_i(t) y, \end{aligned}$$

Примем следующие обозначения

$$\gamma(p) = a(p) + k \frac{\phi(p)}{p} b(p), \quad \beta(p) = \frac{\phi(p)}{p} b(p), \quad f = \frac{p}{\phi(p)} w,$$

тогда для (14) получаем

$$y = \frac{\beta(p)}{\gamma(p)} [-\lambda y + (k + \lambda)\varepsilon + f] + \sum_{i=1}^n \frac{c_i(p)}{\gamma(p)} \theta_i(t) y. \quad (16)$$

Перепишем модель вход-выход (16) в виде модели вход-состояние-выход

$$\dot{x} = Ax + b(-\lambda y + (k + \lambda)\varepsilon + f) + \sum_{i=1}^n q_i \theta_i(t) y, \quad (17)$$

$$y = c^T x, \quad (18)$$

где $x \in R^n$ – неизмеряемый вектор переменных состояния модели (17); A , b , q_i и c – соответствующие матрицы перехода от модели вход-выход к модели вход-состояние-выход.

Так как $\beta(p)$ – гурвицев полином степени $n-1$, то в силу представленной выше леммы, существует число k_0 такое, что можно указать число $k \geq k_0$ и симметрическую положительно определенную матрицу P , удовлетворяющую двум следующим матричным уравнениям:

$$A^T P + PA = -Q_1, \quad Pb = c, \quad (19)$$

где $Q_1 = Q_1^T$ положительно определенная матрица.

Заметим, что значения матрицы Q_1 зависят от параметра k и не зависят от κ .

Перепишем модель (12), (13) в векторно-матричной форме

$$\dot{\xi} = \sigma(\Gamma \xi + dk_1 y), \quad (20)$$

$$y^{\setminus} = h^T \xi, \quad (21)$$

$$\text{где } \Gamma = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 7 & 7 & 7 & 9 & 7 \\ -k_1 & -k_2 & -k_3 & \dots & -k_p \end{bmatrix}, \quad d = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 7 \\ 1 \end{bmatrix} \text{ и } h = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 7 \\ 0 \end{bmatrix}.$$

Введем в рассмотрение новую переменную

$$\eta = hy - \xi, \quad (22)$$

тогда в силу структуры матрицы h невязка ε примет вид

$$\varepsilon = y - \hat{y} = h^T hy - h^T \xi = h^T (hy - \xi) = h^T \eta \quad /$$

Для производной от η получим

$$\dot{\eta} = h\dot{y} - \sigma(\Gamma(hy - \eta) + dk_1 y) = h\dot{y} + \sigma\Gamma\eta - \sigma(dk_1 + \Gamma h)y. \quad (23)$$

Так как $dk_1 = -\Gamma h$ (проверяется подстановкой), то

$$\dot{\eta} = h\dot{y} + \sigma\Gamma\eta, \quad (24)$$

$$\varepsilon = h^T \eta, \quad (25)$$

где матрица Γ , в силу расчета коэффициентов k_i модели (12), имеет собственные числа с отрицательной вещественной частью и удовлетворяет уравнению Ляпунова:

$$\Gamma^T N + N\Gamma = -Q_2, \quad (26)$$

где $N = N^T$ и $Q_2 = Q_2^T$ положительно определенные матрицы.

Теорема. Существуют числа $\sigma > k + \lambda$ и $\lambda > 0$ такие, что все траектории системы (17), (18), (24), (25) ограничены и цель управления выполнена.

Доказательство теоремы аналогично, приведенному в [16].

Теперь рассмотрим проблему выбора коэффициентов k, λ, σ регулятора (11) – (13). Возможным вариантом настройки коэффициентов k, λ, σ является их увеличение до тех пор, пока не будет выполнено целевое условие следующего вида:

$$|y(t)| < \nu \quad \text{при } t \geq t_1, \quad (27)$$

где число ν задается разработчиком системы.

Для реализации этой идеи воспользуемся алгоритмом настройки вида

$$\tilde{k}(t) = \int_{t_0}^t \mu(\tau) d\tau, \quad (28)$$

где $\tilde{k} = k + \lambda$, а функция $\mu(t)$ рассчитывается следующим образом

$$\mu(t) = \begin{cases} \mu_0 & \text{при } |y(t)| \geq \nu, \\ 0 & \text{при } |y(t)| < \nu, \end{cases}$$

где число $\mu_0 > 0$.

Выберем σ следующим образом

$$\sigma = \sigma_0 \tilde{k}^2, \quad (29)$$

где число $\sigma_0 > 0$.

Для иллюстрации предложенной схемы адаптивного управления (11) – (13), (28), (29) рассмотрим следующий числовой пример.

Пример. Рассмотрим пример управления нестационарным объектом следующего вида:

$$\begin{cases} \dot{x}_1 = z_2 + \theta_1(t)z_1, \\ \dot{x}_2 = u + w + \theta_2(t)z_1, \end{cases} \quad (30)$$

$$y = z_1, \quad (31)$$

Выберем закон управления в силу уравнений (11) – (13)

$$u = -\frac{\phi(p)}{p}(k + \lambda)y, \quad (32)$$

$$\begin{cases} \dot{\xi}_1 = \sigma \xi_2, \\ \dot{\xi}_2 = \sigma(-k_1 \xi_1 - k_2 \xi_2 + k_1 y), \end{cases} \quad (33)$$

где полином $\phi(p) = p^2 + p + 1$ и коэффициенты $k_1 = 1, k_2 = 10$.

Приняв $\nu = 0,2$ промоделируем систему при $\mu_0 = 4$ и $\sigma_0 = 0,6$, $y(0) = 1$, $\dot{y}(0) = 0$. Результаты компьютерного моделирования для неизвестных нестационарных параметров $\theta_1(t) = 2 + \sin 0,1t + \sin 10t$, $\theta_2(t) = 2 \cos t$ и возмущения $w(t) = 5$, представлены на рис. 1 а, б.

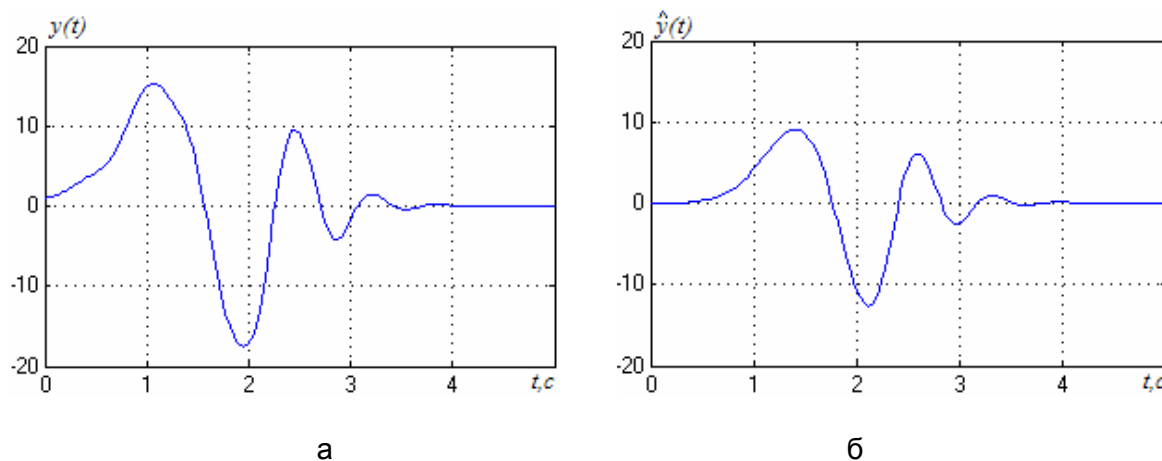


Рис. 1. Результаты компьютерного моделирования

Графики компьютерного моделирования демонстрируют достижение заданной цели управления.

5. Заключение

Рассмотрена задача синтеза закона управления по выходу нестационарной системы вида (1), (2). Был синтезирован алгоритм адаптации вида (11) – (13), обеспечивающий выполнение цели управления. Основные отличия представленного подхода от наиболее популярных аналогов состоят в следующем:

- в отличие от работы [1], относительно вектора переменных параметров $\theta(t)$ системы (1), (2) не выдвигается допущение о периодичности его изменения;
- в сравнение с работами [8 – 11], рассматривается общий вид нестационарной системы, а не частный случай структур матриц описания нестационарных моделей, как в [8] или случай, когда неопределенность согласована с управляющим входом, как в [9 – 11];
- в отличие от работ [1, 2, 6, 8 – 11], синтез алгоритма адаптивного управления произведен по выходу, а не по состоянию;
- усиливая результат, представленный в статье [5], было выдвинуто предположение о том, что объект управления подвержен влиянию внешнего неизвестного постоянного возмущения;
- предложенная схема управления является проще в реализации и позволяет синтезировать адаптивный регулятор фиксированной размерности $\rho + 2$, которая в свою очередь, зависит только от относительной степени передаточной функции $H(p) = R(pI - F)^{-1}L$, но не от числа неизвестных параметров как в работе [5].

Литература

1. Андреев Ю.Н. Управление конечномерными линейными системами. – М.: Наука, 1976. 424 с.
2. Барабанов Н.Е. О стабилизации линейных нестационарных систем с неопределенностью в коэффициентах // АиТ. – 1990. – №10.
3. Tsakalis K. S. and Ioannou P.A. Adaptive control of linear time-varying plants // Automatica. – 1987. – V. 23. – № 4. – P. 459–468,
4. Tsakalis K. S. and Ioannou P.A. Linear time varying systems: control and adaptation. – Upper Saddle River, NJ: Prentice-Hall, 1993.
5. Zhang Y., Fidan B., Ioannou P.A. Backstepping control of linear time-varying systems with known and unknown parameters // IEEE Trans. Automat. Contr. – 2003. – V. 48. – № 11. – P. 1908–1925.
6. Первозванский А.А. Курс теории автоматического управления. – М.: Наука, 1986. 615 с.
7. Юркевич В.Д. Синтез нелинейных нестационарных систем управления с разнотемповыми процессами. – СПб: Наука 2000.
8. Мирошник И.В., Никифоров В.О., Фрадков А.Л. Нелинейное и адаптивное управление сложными динамическими системами. – СПб: Наука, 2000.
9. Цыкунов А.М. Робастное управление нестационарными объектами // АиТ. – 1996. – №2.
10. Бобцов А.А., Лямин А.В., Сергеев К.А. Синтез закона адаптивного управления для стабилизации не точно заданных нестационарных объектов // Известия вузов. – Приборостроение. – 2001. – №4.
11. Никифоров В.О. Робастная следящая система // Известия вузов. – Приборостроение. – 1998. – №7.
12. Kreisselmeier G. Adaptive control of a class of slowly time varying plants // Syst. Control Lett. – 1986. – V. 8 – № 2 – P. 97–103.
13. Middleton R. H., Goodwin G. C. Adaptive control of time-varying linear systems // IEEE Trans. Automat. Contr. – 1988. – V. 33/ – P. 150–155.
14. Tsakalis K. S., Ioannou P. A. Adaptive control of linear time-varying plants: A new model reference controller structure // IEEE Trans. Automat. Contr. – 1989. – V. 34. – P. 1038–1047.
15. Tsakalis K. S., Ioannou P. A. A new indirect adaptive control scheme for time-varying plants // IEEE Trans. Automat. Contr. – 1990. – V. 35. – P. 697–705,
16. Бобцов А.А., Николаев Н.А. Синтез управления нелинейными системами с функциональными и параметрическими неопределенностями на основе теоремы Фрадкова // АиТ. – 2005. – №1.

ФОРМИРОВАНИЕ НАНОСВИТКОВ В ВЯЗКОЙ ЖИДКОСТИ

И.В. Блинова, И.Ю. Попов, А.И. Свитенков, С.А. Чивилихин

Процесс получения нанотрубок гидротермальным методом, основанный на скручивании предварительно напряженного слоя в вязкой жидкости, является предметом многочисленных экспериментальных исследований последних лет. В настоящей работе построена математическая модель этого процесса.

Введение

Механика формирования нанотрубок путем скручивания напряженных нанослоев рассматривалась в ряде работ (см., например, [1–3]). Вместе с тем динамика процесса скручивания нанотрубок при их образовании во флюидных средах теоретически не анализировалась. Особенностью образования нанотрубок в гидротермальных или в других флюидных средах является влияние вязкой среды на процесс скручивания. Сворачивание двойного слоя происходит вследствие внутренних напряжений, возникающих из-за неполного структурного соответствия составляющих его слоев. Двойной слой толщины δ' при увеличении толщины межслоевого пространства δ'' , вследствие интеркалирования в него воды или других компонентов окружающей среды, скручивается в круговой цилиндр – нанотрубку с начальным радиусом R_0 (рис. 1) порядка нескольких нанометров.

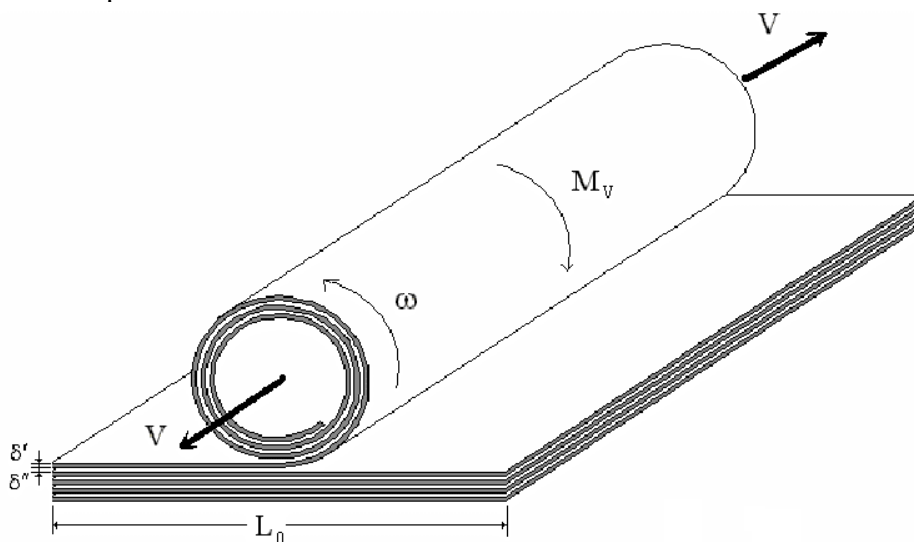


Рис.1. Схема скручивания нанотрубки

Угловая скорость вращения нанотрубки ω определяется балансом двух моментов – момента внутренних сил $M_E \sim \frac{E(\delta')^3}{R_0}$ и момента сил вязкого трения $M_V \sim 2\pi R_0^2 \mu \omega$.

Число Рейнольдса вязкого течения вокруг скручивающейся нанотрубки может быть оценено как

$$Re = \frac{\rho \omega R_0^2}{\mu}.$$

Здесь E – модуль Юнга нанотрубки, μ и ρ – динамическая вязкость и плотность жидкости. Для характерных значений параметров нанотрубки $R_0 \sim 5 \div 10^{-9} \text{ м}$, $\delta' \sim 2 \cdot 10^{-10} \text{ м}$, $E \sim 10^{11} \text{ Па}$, $\mu \sim 10^{-2} \frac{\text{кг}}{\text{м} \cdot \text{сек}}$, $\rho \sim 300 \frac{\text{кг}}{\text{м}^3}$ получаем $Re \sim 10^{-4}$. Столь ма-

лое значения числа Рейнольдса позволяет использовать для описания течения квазистационарное приближение Стокса [4].

Построение модели скручивания нанотрубки

Решение бигармонического уравнения для функции тока Ψ с условиями прилипания на плоской границе слоя и на поверхности вращающейся трубки имеет вид:

$$\Psi = 2R^2\omega \left[\frac{4Ry^3}{(x^2 + y^2)^2} - \frac{3y^2}{x^2 + y^2} \right]. \quad (1)$$

Используя (1) и корректируя поле скоростей в окрестности точки качания нанотрубки с плоскостью с использованием условия частичного прилипания [5], получаем

$$M_V = 2\pi\alpha R^2\mu\omega, \text{ где } \alpha \approx \frac{12R_0}{\pi\delta'}.$$

Первый виток нанотрубки имеет радиус R_0 . В ходе скручивания внутренний радиус наносвитка R_1 принимает значения, меньшие исходного радиуса R_0 , а ее внешний радиус R_2 оказывается большим, чем R_0 . Уменьшение внутреннего радиуса нанотрубки после образования более чем одного витка, определяется уменьшением при этом механической энергии нанотрубки вследствие проскальзывания скрученных слоев друг относительно друга и приводит к выдавливанию жидкости изнутри нанотрубки.

При этом суммарная потенциальная энергия нанотрубки и еще не скрученной части двойного слоя имеет вид

$$U = \frac{E\delta'^3 L_0}{24R_0^2} L + \frac{E\delta'^3 L_0}{24} \int_0^{L_0-L} \left(\frac{1}{R(\zeta)} - \frac{1}{R_0} \right)^2 d\zeta, \quad (2)$$

где L , L_0 – длина нескрученной части нанослоя и ее начальное значение, ζ – координата вдоль витков нанотрубки, R_0 – равновесный радиус витка нанотрубки. Выражение (2) позволяет рассчитать момент упругих сил M_E , приложенный к нанотрубке, и нормальное напряжение на ее внутренней поверхности. Тогда равенство моментов $M_E = M_V + M_A$, где M_A – момент сил адгезии нанотрубки к нанослою, дает уравнение динамики внешнего радиуса R_2 вида

$$\frac{dr_2}{d\tau} = \frac{1}{r_2} \left[\left(1 + \lambda r_1^4 \right) \frac{r_2 - 0.5}{r_2^2} - \lambda r_1^2 (r_1 - 0.5) \right] - \frac{m_A}{r_2^2}, \quad m_A = \frac{12R_0}{E\delta'^3} M_A, \quad (3)$$

где $r_1 = \frac{R_1}{R_0}$, $r_2 = \frac{R_2}{R_0}$, $\tau = \frac{E\delta'^3\delta}{48\pi^2\mu\alpha R_0^4} t$, $\lambda = \frac{2\pi^2\alpha R_0^4}{L_0^2\delta^2}$. При скручивании нанотрубки ее

внутренний радиус R_1 убывает, что приводит к выдавливанию вязкой жидкости из внутренней полости трубки. Приравняв механическое напряжение на внутренней поверхности нанотрубки и среднее напряжение в вязкой жидкости внутри нанотрубки, получаем уравнение динамики внутреннего радиуса

$$\frac{dr_1}{d\tau} = -\lambda r_1^3 \left(\frac{r_1 - 0.5}{r_1^2} - \frac{r_2 - 0.5}{r_2^2} \right). \quad (4)$$

Исследование процесса скручивания нанотрубки на основе построенной модели

На рис. 2 представлены результаты расчетов динамики внешнего и внутреннего радиусов нанотрубки с начальными условиями $R_1(0) = R_0 - 0.5\delta$, $R_2(0) = R_0 + 0.5\delta$ при различных значениях момента сил адгезии. При этом предполагается, что длина и

ширина исходного нанослоя одинаковы и имеют порядок 2 мкм. В качестве предельного, рассмотрен вариант скручивания нанотрубки без взаимного проскальзывания ее витков (в этом случае уравнения (3) и (4) существенно упрощаются). Отметим, что характерное время скручивания прямо пропорционально вязкости жидкости.

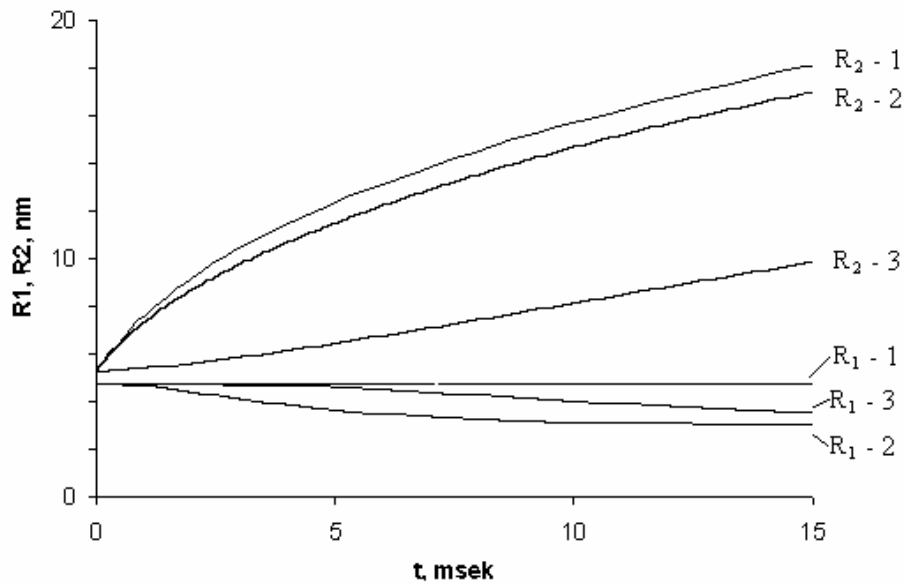


Рис.2. Зависимость внутреннего (R_1) и внешнего (R_2) радиусов нанотрубки от времени скручивания: 1 – без взаимного проскальзывания витков, безразмерный момент сил адгезии $m_A = 0$; 2 – с учетом проскальзывания витков, $m_A = 0$; 3 – с учетом проскальзывания витков, $m_A = 0.5$

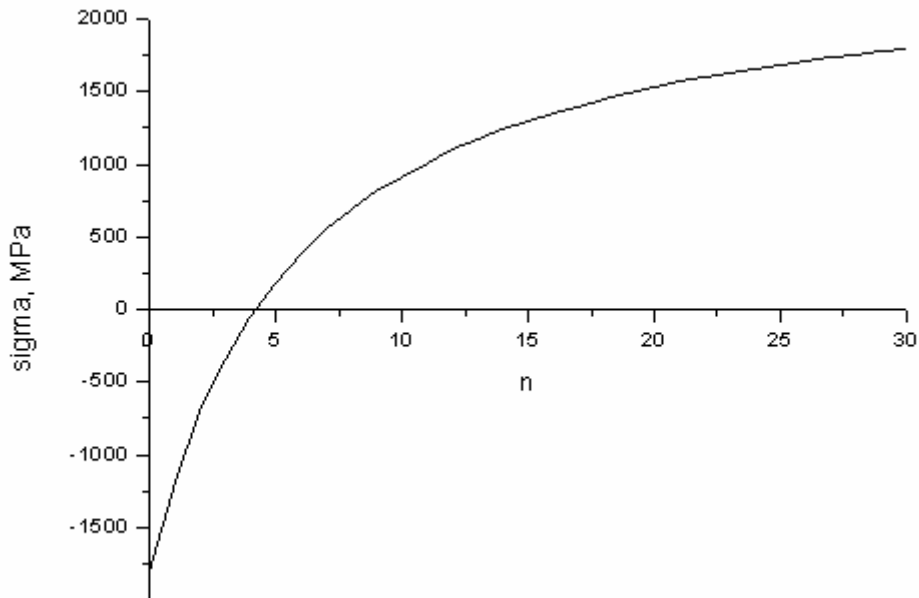


Рис.3. Зависимость напряжения от номера слоя в нанотрубке

После того, как весь исходный нанослой смотался в нанотрубку, крутящий момент и момент сил адгезии исчезают. Однако упругая энергия нанотрубки не является минимальной. Поэтому нанотрубка продолжает подскручиваться за счет сил упругости вплоть до асимптотического установления равновесной конфигурации. На рис. 3 изо-

бражена зависимость напряжения от номера слоя в равновесной нанотрубке, сформированном из слоя длиной $L_0 = 2 \text{ мкм}$.

Выводы

В работе построена математическая модель процесса получения нанотрубок, основанном на скручивании предварительно напряженного слоя в вязкой жидкости. Показано, что вязкость существенно сказывается на времени скручивания слоя. Характерное время скручивания слоя со структурой серпентина шириной несколько микрометров, в рассматриваемом диапазоне значений вязкости – $\mu \sim 10^{-2} - 10^{-1} \frac{\text{кг}}{\text{м} \cdot \text{сек}}$, составляет от десятых долей до единиц миллисекунд. Момент сил адгезии также может значительно повлиять на скорость скручивания нанотрубки. Нанотрубка, сформированная в результате скручивания, оказывается напряженной. Знак напряжения меняется при переходе от внешних витков нанотрубки к ее внутренним виткам.

Литература

1. Ивановский А.Л. Неуглеродные нанотрубки: синтез и моделирование // Успехи химии. – 2002. – Т. 71. – № 3. – С. 203–224.
2. Маслов А.В., Корыткова Э.М., Гусаров В.В. Гидротермальный синтез неорганических нанотрубок // В кн.: Химия поверхности и синтез низкоразмерных систем. – СПб.: РИОСПбГТИ(ТУ). 2002. – С.54–59.
3. Корыткова Э.Н., Маслов А.В., Пивоварова Л.Н., Дроздова И.А., Гусаров В.В. Формирование нанотрубок состава $\text{Mg}_3\text{Si}_2\text{O}_5(\text{OH})_4$ в гидротермальных условиях // Физика и химия стекла. – 2004. – Т.30. – №1. – С.72–78.
4. Корыткова Э.Н., Маслов А.В., Пивоварова Л.Н., Полеготченкова Ю.В. Павинич В.Ф., Гусаров В.В. Образование нанотрубчатых гидросиликатов системы $\text{Mg}_3\text{Si}_2\text{O}_5(\text{OH})_4 - \text{Ni}_3\text{Si}_2(\text{OH})_4$ при повышенных температурах и давлениях // Неорган. матер. – 2005. – Т.41. – №7. – С.849–855.
5. Joly L., Ybert C., Bocquet L. Probing the Nanohydrodynamics at Liquid-Solid Interfaces Using Thermal Motion // Phys. Rev. Letters. – 2006. – №96. – 046101-1 - 046101-1.

**РАЗРАБОТКА АЛГОРИТМИЧЕСКОГО И ПРОГРАММНОГО
ОБЕСПЕЧЕНИЯ ДЛЯ МОДЕЛИРОВАНИЯ АПЕРТУРЫ
СБОМ – ЗОНДОВ НА ОСНОВЕ ЭКСПЕРИМЕНТАЛЬНОГО
ИССЛЕДОВАНИЯ ДАЛЬНЕПОЛЬНОЙ ОСВЕЩЕННОСТИ**

Ле Зуй Туан

Разработано программное обеспечение для регистрации излучения зонда в дальнем поле с расширенным диапазоном регистрируемой освещенности. Распределения освещенности в дальней зоне сопоставлены с качеством зондов. Реализован метод компьютерного моделирования структуры ближнего поля на основе математической обработки полученной экспериментальной информации.

Введение

Одной из актуальных задач быстро развивающейся в настоящее время сканирующей ближнепольной оптической микроскопии (СБОМ) [1] является контроль или аттестация ближнепольных зондов, т.е. определение их размеров и формы. Эта задача непосредственно связана с определением разрешающей способности СБОМ, поэтому крайне важно найти наиболее оптимальные и простые пути ее решения. Непосредственное измерение параметров зонда возможно лишь средствами растровой электронной микроскопии (РЭМ), что по ряду причин не является наилучшим способом аттестации. Альтернативой электронному микроскопу является оптический анализ параметров вторичного источника света по характеристикам его излучения в дальней зоне, реализация которого может быть значительно более простой и дешевой, чем применение РЭМ. Кроме того, решение задачи аттестации зондов на основе этого подхода влечет за собой и другие важные преимущества, как, например, возможность экспериментального определения передаточной функции СБОМ.

Определение параметров апертуры зондов по характеристикам излучения в дальней зоне возможно, если разработать установку для обеспечения уверенной регистрации распределения интенсивности света и математический аппарат для восстановления структуры вторичного источника света [2, 3].

**Установка и программное обеспечение для регистрации распределения
излучения зонда в дальнем поле**

На рис. 1 показана схема измерительной установки. Лазерное излучение ($\lambda = 680\text{ нм}$) фокусируется с помощью линзы (микрообъектива) и направляется в зонд, который находится на расстоянии h от центра ПЗС матрицы, расположенной перпендикулярно его оси. Зонд и ПЗС матрица при работе расположены в закрытой коробке с целью ликвидации шумовых сигналов. Результат измерения выводится на экран компьютера.

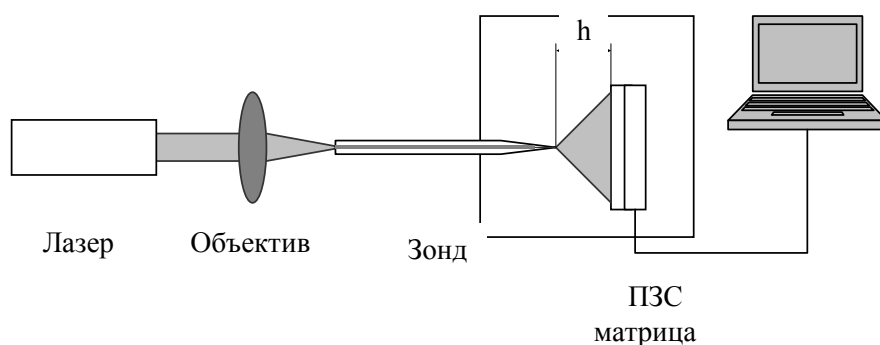


Рис. 1. Схема экспериментальной установки для регистрации излучения от зонда в дальнем поле

Для расширения диапазона регистрируемых освещенностей и устранения нелинейной световой характеристики ПЗС матрицы было разработано алгоритмическое и программное обеспечение, которое основано на регистрации серии фотоснимков пятна рассеяния, полученных последовательно на одном и тоже фотоприемнике с переменным временем экспозиции для получения совокупности фотометрических сечений, соответствующих различным уровням распределения освещенности в дальнем поле. Из каждого фотометрического сечения формируется контур, который называется изофотой.

Из системы изофот можно получить карту распределения освещенности в пятне рассеяния (рис. 2). Такой метод назван методом изофотометрии с изменяющимся временем накопления [6]. Диапазон регистрации освещенности по данному методу не зависит от протяженности рабочего участка световой характеристики фотоприемника, а зависит от диапазона изменения времени накопления, который составляет $10^3 - 10^5$. В экспериментальной установке была использована матричная видеокамера VAC-135 с диапазоном относительного времени накопления свыше трех порядков ($1 \div 2047$).

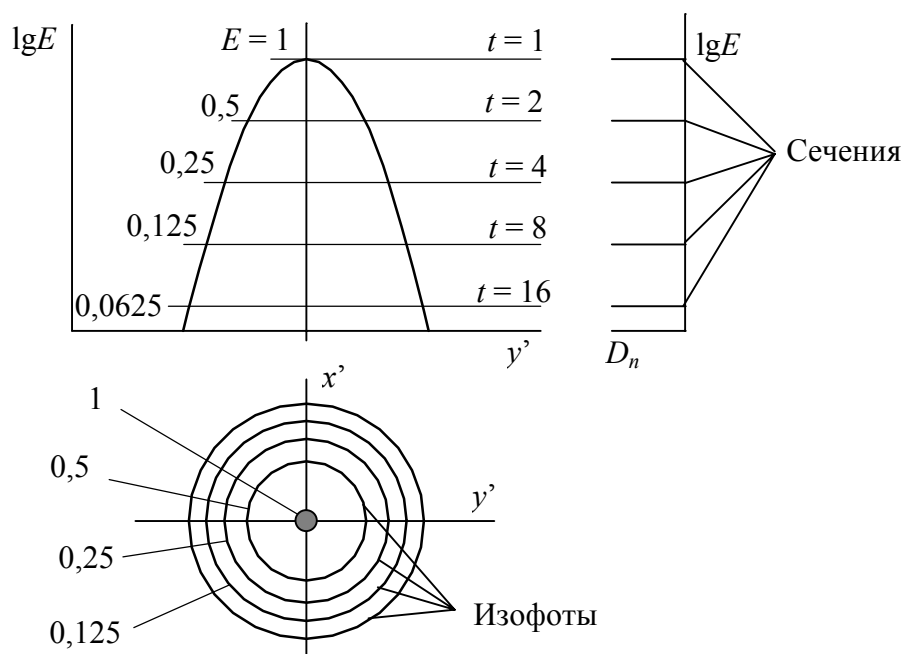


Рис. 2. Схема метода изофотометрии с переменным временем накопления и вид изофотограммы

На рис. 3 показано алгоритмическое решение компьютерной версии изофотометрии с переменным накоплением. В исходном состоянии параметр “exposure” камеры установлен на минимальное значение t_1 , а чувствительность камеры отрегулирована так, что световое поле в дальней зоне формирует изображение, дающее изофоту в виде точки (первая изофота). Эта изофота соответствует максимуму освещенности $E_{отн1}$, которое принимаем равным единице.

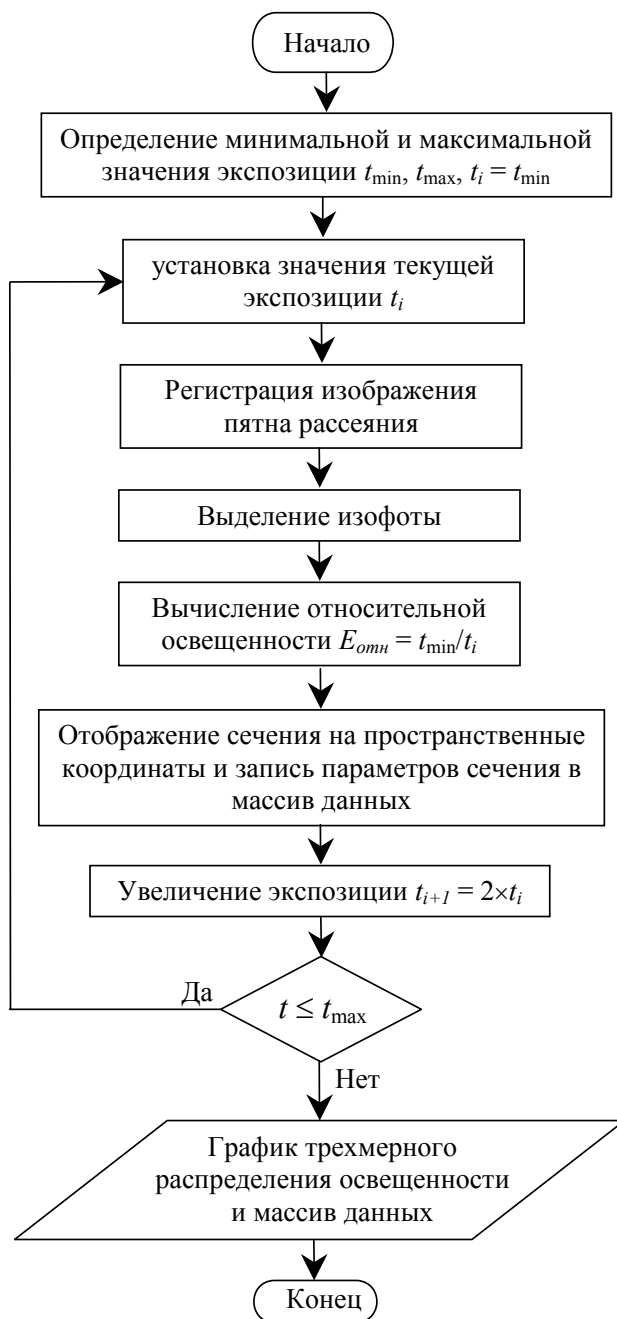


Рис. 3. Алгоритмическое решение компьютерной версии изофотометрии с переменным накоплением

В процессе измерения, после установки увеличенного времени экспозиции $t_2 = k \cdot t_1$, программа выполняет регистрацию последующей фотографии и вводит ее в модуль выделения изофоты. Эта изофота является контуром фотометрического сечения, соответствующего новому уровню относительной освещенности $E_{отн2} = t_1 / t_2$.

Таким образом, $E_{отн\ n} = 1/t_{отн\ n} = t_1/t_n$, поэтому помещаем изофоту в систему координат на уровень $E_{отн\ n}$ и записываем его в память. Когда t превышает максимальное значение t_{max} для ПЗС-камеры, работа программы заканчивается, и на мониторе получается графическое отображение всей суммы изофот распределения освещенностей.

В результате запоминания ряда фотометрических сечений и построения в пространственных координатах их исходного взаимного расположения формируется трехмерное распределение освещенности в дальнем поле.

Результат измерения

При регистрации распределения интенсивности излучения зонда в дальнем поле были получены разные картины распределения с особенностями, соответствующими различным зондам. Результаты измерения представлены на рис. 4.

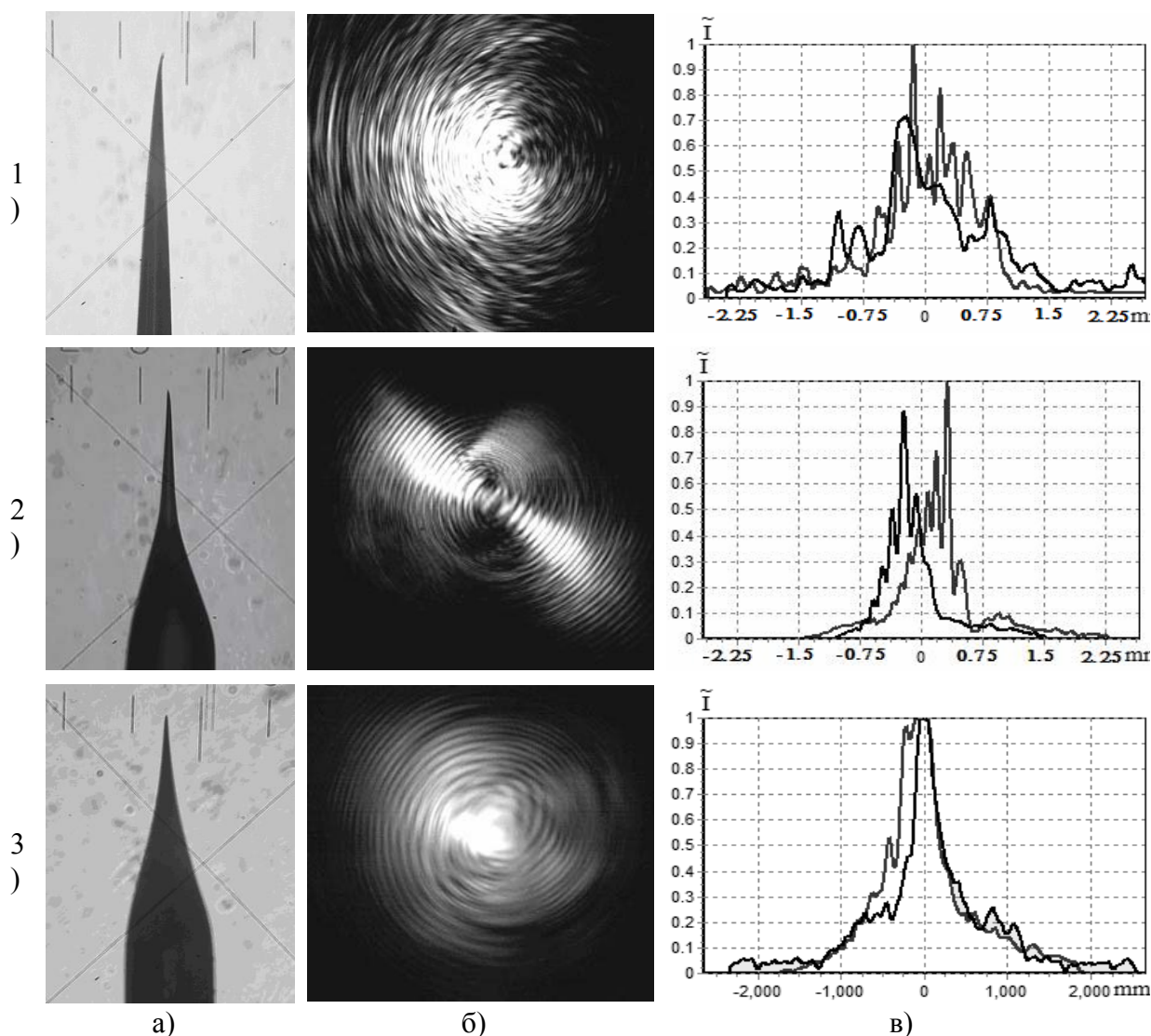


Рис. 4. Зонды и соответственно полученные картины пятна, регистрируемые на расстоянии 4 мм от зонда: а) внешний вид зондов; б) картины распределения освещенности в дальнем поле; в) кривые распределения освещенности по двум сечениям

Из этих картин мы можем сделать некоторые выводы.

- Изгиб острия (зонд 1 на рис. 4) оказывает влияние на распределение интенсивности таким образом, что половина картины имеет низкую интенсивность.
- Эллиптическая апертура приводит на картине распределения интенсивности двух лепестков с повышенной интенсивностью, которые направлены перпендикулярно большому радиусу эллипса (зонд 2 на рис. 4).
- Лучшее распределение дает зонд 3, который имеет круглую и симметричную апертуру, что проявляется в круглой и симметричной картине распределения.

Оптический метод определения формы и размера апертуры зонда по распределению излучения в дальнем поле

Как уже сказано, качество зонда имеет большое значение в разрешении и качестве изображения, полученного в ближнепольном оптическом микроскопе. Поэтому знание параметров его апертуры составляет важную задачу. При определении размера апертуры зонд нельзя различить в обычный оптический микроскоп, а другие методы микроскопии высокого разрешения (электронная сканирующая микроскопия или атомно-силовая микроскопия) не могут дать однозначной информации об апертуре вследствие неоптической природы процессов наблюдения. Обычно зонд тестируют прямо на ближнепольном микроскопе, сканируя тест-объект. Этот метод требует сложной и дорогой техники. Более простым и оперативным методом представляется метод определения размера и качества апертуры по распределению излучения в дальнем поле [3, 4].

Излучение, проходящее через субволновую апертуру, имеет две составляющие. Первая составляющая, соответствующая высокой частоте в пространственном спектре источника, является эванесцентной волной, которая быстро затухает. Вторая составляющая, соответствующая низкой частоте в пространственном спектре источника, является излучающей волной, которая может распространяться [5]. Регистрируется излучающая волна, и по ее распределению в дальнем поле вполне возможно определить распределение этой волны в апертуре путем решения обратной задачи, от чего можно определить ее форму и размер.

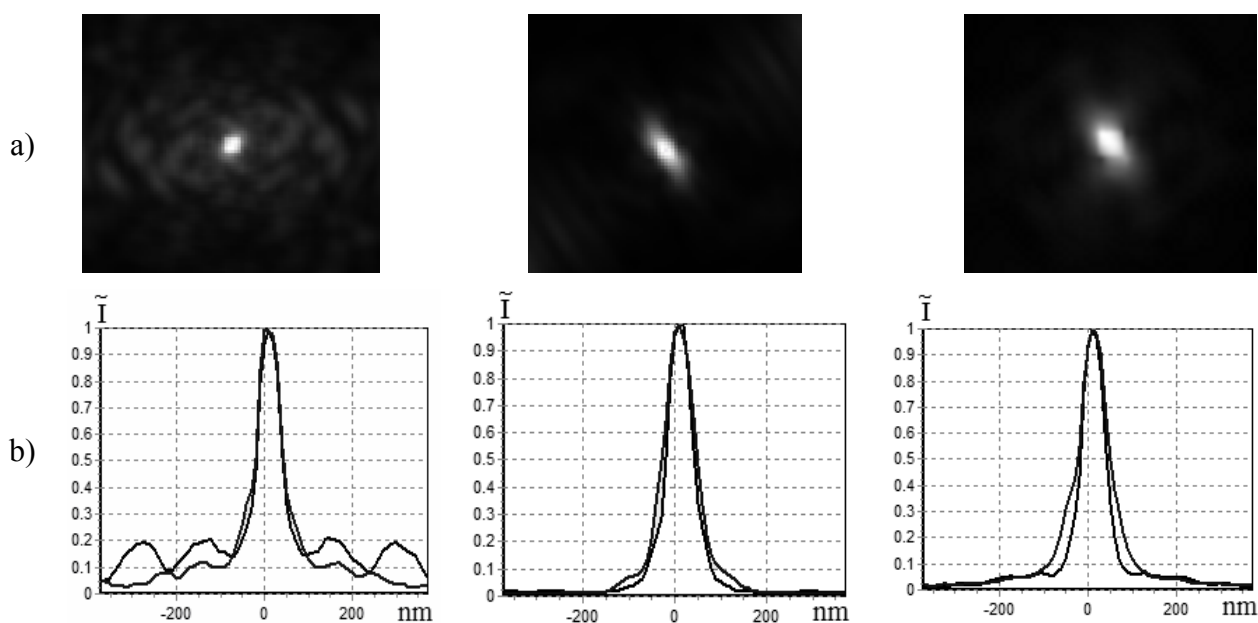


Рис. 5. Результаты восстановления светового поля в апертурах зондов, показанных на рис. 4
а) световые поля в апертурах зондов; б) соответствующие профили распределения полей в центре апертур

В работах [3, 8] был разработан метод сверхразрешения для восстановления светового поля в плоскости апертуры, которая служит вторичным источником для излучения в дальней зоне. Однако алгоритм восстановления в этих работах требует процесса итерационного расширения спектра с помощью полиномов Цернике для восстановления спектра за предел регистрируемой (видимой) области. Это не позволяет избежать субъективных погрешностей при выборке методов итерации.

Как уже говорилось, ПЗС-матрица, управляемая с помощью программного обеспечения, написанного на основе метода изофотометрии, имеет расширенные возможности и может регистрировать большой диапазон пространственно-частотного спектра. Таким образом, нам не требуется итерационное расширение спектра. Регистрируемые распределения интенсивности в дальней зоне могут прямо использоваться для восстановления светового поля в апертуре. Распределения интенсивности представлены на рис. 4.

По алгоритму восстановления параметров субдлинноволновых вторичных источников света, представленному в [7], было разработано программное обеспечение для математического моделирования. Результат восстановления показан на рис. 5. Из профилей распределения поля в апертуре можно определить размер апертур. Например, на уровне 0.1 относительной интенсивности размеры апертур составляют соответственно 180, 150 и 210 нм.

Выводы

1. Экспериментально показано, что изучение распределения интенсивности излучения зонда в дальнем поле может быть полезным для характеристики БОЗ. С использованием этого распределения определен размер апертуры БОЗ.

2. Разработано программное обеспечение для реализации метода компьютерной изофотометрии путем управления режимами ПЗС-камеры для регистрации серии кадров с переменным временем экспозиции.

3. Разработан пакет программ для обработки данных изофотометрии и определения параметров поля в зоне апертуры зонда, что позволяет судить о размере и качестве апертуры зонда.

Литература

1. Heinzmann, D. W. Pohl. Scanning Near-field optical microscopy // *Applied. Physics. A.* 1994 – Vol.59. – P. 89–101.
2. Christian Obermuller, Khaled Karrai. Far field characterization of diffracting circular apertures // *Appl. Phys. Lett.* – December 1995. – 67 (23). – p. 3408-3410.
3. Вейко В.П., Вознесенский Н.Б., Гусев А.Е., Иванова Т.В., Родионов С.А. Возможность определения параметров вторичных источников света, меньших длины волны, по характеристикам дальнего поля // *Оптический журнал.* – 1998. – Т. 65. – № 10. – С. 49-53.
4. Вознесенский Н. Б. . Математическое моделирование процессов распространения и дифракции света в наноструктурах // *Сборник материалов всероссийского совещания “Зондовая микроскопия-2000”.* Нижний Новгород. 28 февраля–2 марта 2000 г. С.142–146.
5. D. J. Shin, A. Chavez-Pirson, S. H. Kim, S. T. Jung, and Y. H. Lee. Diffraction by a sub-wavelength-sized aperture in a metal plane. // *J. Opt. Soc. Am. A.* – July 2001. – Vol. 18, – No 7.
6. Кирилловский В.К. Оптические измерения Часть 4. Оценка качества оптического изображения и измерение его характеристик. – СПб: СПб ГУ ИТМО., 2005.

7. Иванова Т.А. Разработка математических методов моделирования и восстановления характеристик субдлинноволновых вторичных источников света // Диссертация на соискание ученой степени кандидата технических наук. СПб ГУ ИТМО. 1999.
8. Veiko V.P., Voznessenski N.B., Domnenko V.M., Ivanova T.V., Rodionov S.A., Goussev A.E. A new approach to analysis of subwavelength sized secondary light sources. // Proceedings of SPIE – Vol. 3467. – 1998.

ИСПОЛЬЗОВАНИЕ АВТОМАТОВ С ФЛАГАМИ ДЛЯ РЕШЕНИЯ ЗАДАЧИ О «ФЛИБАХ»

П.Г. Лобанов, А.А. Шалыто

Предложен метод реализации флибов, моделирующих поведение простейших живых существ в простейшей среде, с помощью автоматов с флагами. Выполнены вычислительные эксперименты, демонстрирующие эффективность предложенного метода по сравнению с известным.

Введение

Существует класс задач (итерированная дилемма узников [1], задача про «умного» муравья [2], задачи синхронизации [3] и классификации плотности [4, 5] для клеточных автоматов), в качестве решения которых может выступать автомат. Для этих задач существуют генетические алгоритмы [6], которые позволяют автоматически строить автоматы, являющиеся решением данных задач.

В работе [7] также была рассмотрена задача, для решения которой требуется построить автомат. Это задача о флибах. В работе [8] был предложен подход, позволивший получить более точное решение задачи о флибах.

Целью данной работы является разработка нового метода, позволяющего улучшить результаты, достигнутые в работах [7, 8].

1. Задача о флибах

Флиб моделирует поведение простейшего живого существа. Это существо способно предсказывать имеющие периодичность изменения простейшей окружающей среды. В качестве модели живого существа используется конечный автомат. Генетические алгоритмы позволяют построить автомат, который предсказывает изменения среды с достаточно высокой точностью. Таким образом, при решении данной задачи требуется построить «устройство» (предсказатель), которое предсказывает изменения среды с наибольшей вероятностью.

2. Постановка задачи

Одна из важнейших способностей живых существ – предсказание изменений окружающей среды. В качестве простейшей модели живого существа можно использовать конечный автомат. В работе [7] такие конечноавтоматные модели названы *флибами* (сокращение от *finite living blobs* – конечные живые капельки).

На вход флиба подается переменная, которая принимает одно из двух значений – ноль или единица. Эта переменная соответствует состоянию окружающей среды в текущий момент времени. Изменения окружающей среды задаются повторяющейся последовательностью битов (битовой маской). Флиб изменяет свое состояние и формирует значение выходной переменной. Это значение соответствует возможному состоянию среды в следующий момент времени.

Флиб должен предсказать какое состояние окружающей среды наступит в следующий момент времени. Благодаря периодичности входного сигнала можно построить автомат, который будет правильно предсказывать состояние среды заданное определенной маской.

Точность, с которой флиб предсказывает изменения среды, используется в качестве оценочной функции в генетическом алгоритме.

В работе [7] для решения указанной задачи используется один из генетических алгоритмов. Предложенный в работе [8] подход, позволяет улучшить этот алгоритм за счет изменения способа формирования поколения (используется турнирный отбор [9] и

принцип элитизма [10]) и применения операции восстановления связей между состояниями.

Данная работа базируется на генетическом алгоритме [8], для улучшения точности решений которого предлагается использовать в качестве флиба вместо обычного автомата автомат с флагами [11]. В связи с отсутствием изменений в генетическом алгоритме, предложенном в [8], его описание и реализация в данной работе не приводятся.

3. Известный способ реализации флиба

Поведение флиба описывается с помощью таблицы переходов и выходов. В табл. 1 в качестве примера приведен флиб с тремя состояниями А, В и С, который всегда имеет одну входную и одну выходную переменную.

Состояния	Входная переменная	
	0	1
А	1, В	0, А
В	0, С	0, А
С	1, А	0, В

Таблица 1. Таблица переходов и выходов для флиба с тремя состояниями

На рис. 1 изображен граф переходов для флиба, таблица переходов и выходов которого приведена в табл. 1.

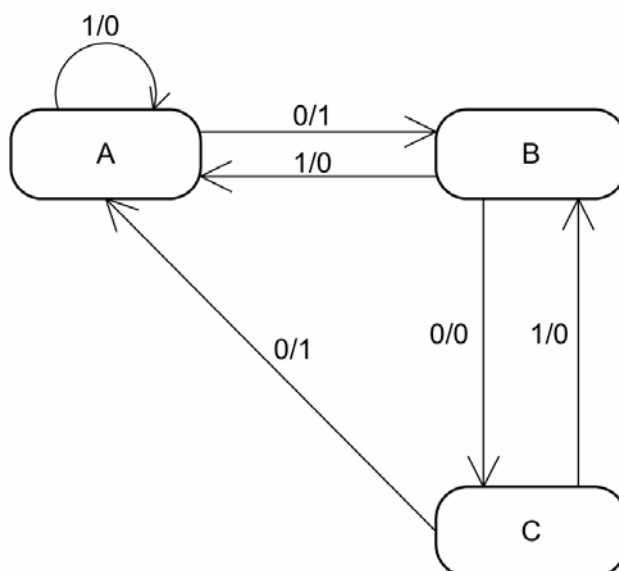


Рис 1. Граф переходов для флиба с тремя состояниями

В работе [8] применялся способ реализации флиба с помощью трех классов: Flib, State и Branch, которые приведены на сайте <http://is.ifmo.ru> в разделе "Статьи".

В качестве главного класса при реализации флиба применяется класс `Flib`. Классы `State` и `Branch` реализуют его состояния и переходы соответственно. В каждом из этих классов имеется метод `Clone`, предназначенный для создания копий объектов.

Массив `_states` в классе `Flib` содержит состояния флиба. Поле `_curStateIndex` используется для хранения номера текущего состояния флиба в массиве `_states`. Число правильно предсказанных входных символов размещается в поле `_guessCount`. Метод `Step` переводит флиб в новое состояние и, при необходимости, изменяет количество правильно предсказанных символов. Метод `Nulling` возвращает флиб в начальное состояние, и обнуляется число правильно предсказанных символов.

В массиве `_branches` класса `State` содержатся дуги переходов из данного состояния. Номер элемента в массиве соответствует входной переменной.

4. Использование автоматов с флагами для реализации флиба

В работе [11] предложен класс автоматов, названный “автоматы с флагами”. В таких автоматах используется вектор многозначных флаговых переменных F , который содержит информацию о “глубокой предыстории” – предыдущих состояниях автомата. Автоматы данного класса взаимодействуют с внутренними ячейками многозначной памяти.

Флаговые переменные используются в качестве внутренних входных и выходных переменных автомата. В таких автоматах переход в следующее состояние зависит не только от текущего состояния и соответствующих значений входных переменных, но и от значений флаговых переменных, определяемых состояниями, в которых автомат находился ранее. Как показано в работе [11], использование флаговых переменных в общем случае может позволить уменьшить число состояний в автомате. Это, правда, затрудняет понимание алгоритма работы автомата. Однако, этот недостаток не является существенным для случая автоматической генерации автоматов.

В настоящей работе предлагается модификация флаговых автоматов, в которой флаговые переменные используются для **хранения предыдущих состояний окружающей среды**. Во флаговой переменной с номером один будем хранить состояние среды на предыдущий момент времени, а во флаговой переменной с номером n , будем хранить состояние среды на n шагов назад.

Пусть s и $F = (f_1, \dots, f_n)$ соответственно состояние среды и вектор флаговых переменных в текущий момент времени. Тогда вектор флаговых переменных для следующего момента времени будет иметь вид $F' = (f_1' = s, f_2' = f_1, \dots, f_n' = f_{n-1})$. При программной реализации флаговые переменные хранятся в массиве. Переход от F к F' выполняется с помощью удаления последнего элемента массива и вставки в начало массива текущего состояния среды s .

Состояния	Флаговая и входная переменные			
	00	01	10	11
A	1, A	1, B	0, B	0, A
B	1, B	1, C	0, C	0, A
C	0, B	0, C	1, A	1, B

Таблица 2. Флиб с тремя состояниями

Введение каждой новой флаговой переменной увеличивает число дуг из каждого состояния вдвое. При n флаговых переменных число дуг из каждого состояния флиба будет равно $(1 + n)^2$. Номер дуги, которая должна использоваться в текущий момент времени для перехода, можно вычислить по формуле $s + 2^1 f_1 + 2^2 f_2 + \dots + 2^n f_n$. Номер этой дуги может быть записан также как двоичное число $f_n \dots f_2 f_1 s$.

Далее флиб, который моделируется с помощью автомата с флагами, будем называть флибом с флагами.

В табл. 2 приведен пример таблицы переходов и выходов для флиба с тремя состояниями и одной флаговой переменной.

На рис. 2 изображен граф переходов для флиба, таблица переходов которого приведена в табл. 2.

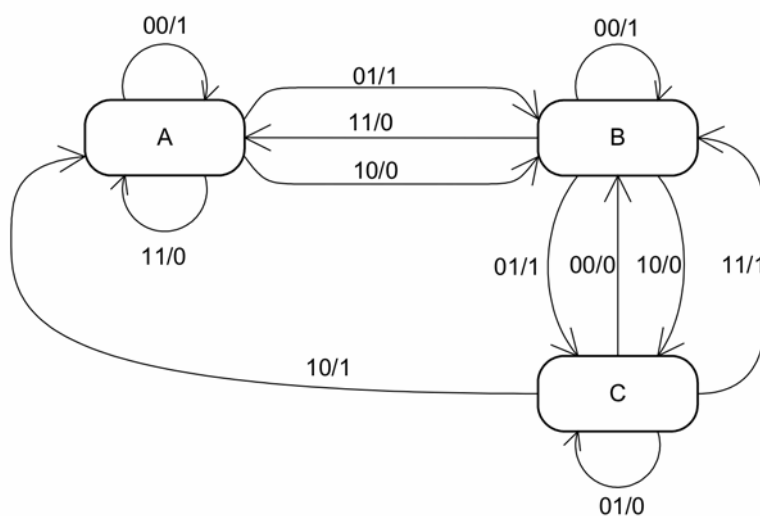


Рис 2. Граф переходов для флиба с тремя состояниями и одной флаговой переменной

Для реализации флиба с флагами, число которых в общем случае больше одного, используется класс `FlibWithMem`, который приведен на сайте <http://is.ifmo.ru> в разделе "Статьи". Данный класс унаследован от класса `Flib` и перегружает методы `Step`, `Nulling` и `Clone`. Массив `_flags` используется для хранения значений флаговых переменных.

Для вычисления номера дуги перехода из текущего состояния используется метод `GetBranchIndex`, принимающий на вход текущее состояние среды.

5. Программа для проведения экспериментов

В программу, написанную в работе [8], была добавлена возможность проведения экспериментов над флибами с флагами. Обновленное окно пользовательского интерфейса программы приведено на рис. 3.

Программа позволяет выполнять алгоритмы, предложенные как в настоящей работе, так и в работах [7, 8]. Выбор алгоритма выполняется в окне с помощью выпадающего списка *Способ формирования нового поколения*.

В выбираемом алгоритме может быть изменен элитизм в диапазоне от нуля до размера поколения. Программа позволяет выбрать размер поколения и число поколений. Может быть выбрана вероятность применения оператора мутации (в диапазоне от нуля до единицы) и тип оператора скрещивания (одноточечный и двухточечный). В программе имеется возможность задавать битовую маску, определяющую среду. Для

каждого флиба может быть задано число его состояний, и определено будет ли использоваться алгоритм восстановления связей между состояниями.

Помимо этого в новой версии программы с помощью выпадающего списка *Тип флиба* можно выбрать тип автомата, использующегося для моделирования флиба (простой автомат или автомат с флагами). Также можно указать число флагов, использующихся во флибе.

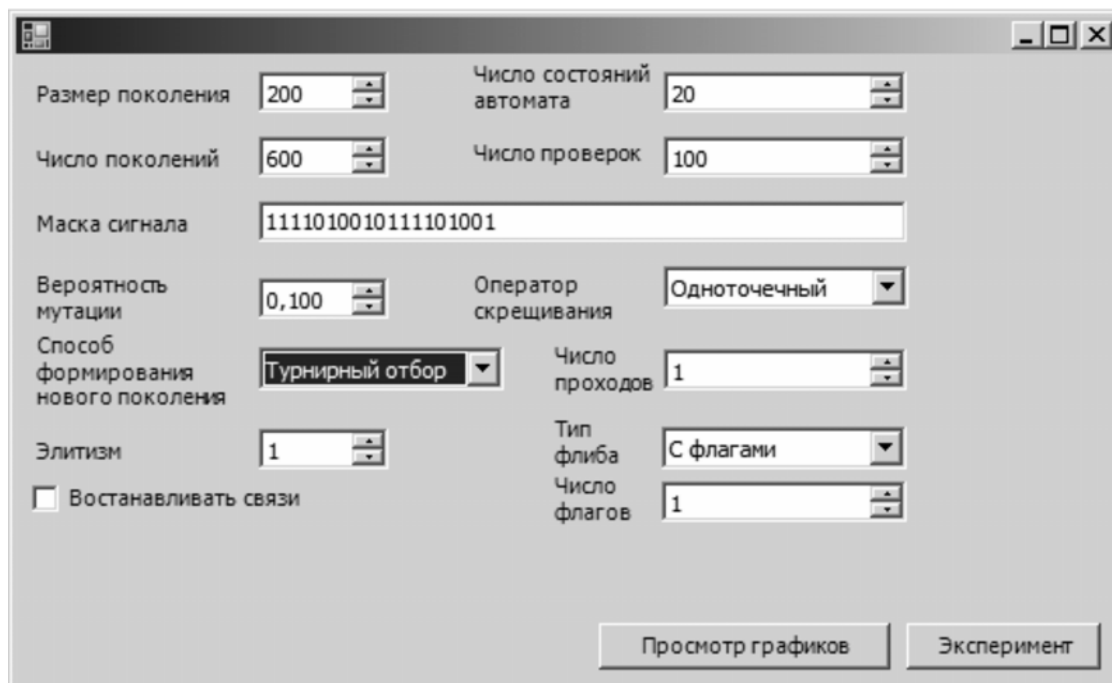


Рис. 3. Окно пользовательского интерфейса

6. Эксперименты

Все эксперименты, о которых идет речь ниже, проводились при размере поколения равном 500 и числе поколений равном 1000. Вероятность применения оператора мутации была выбрана 0,03. Число воздействий среды на флиб выбрано равным 100. Благодаря этому число правильно предсказанных символов по величине равно точности предсказания символов в процентах. В таблицах с результатами экспериментов приводятся минимальные, максимальные и средние точности предсказания автоматически построенных флибов. В качестве битовой маски использовалась строка 1111010010111101001.

Число состояний	Худший результат	Усредненный результат	Лучший результат
10	89	93,76	95
20	90	94,3	95
40	89	94,58	100
80	89	96	100

Таблица 3. Результаты экспериментов над флибами с разным числом состояний

На графиках по оси абсцисс отложены номера поколений, по оси ординат – число правильно предсказанных символов лучшим предсказателем в каждом поколении. При

построении графиков применялись **усредненные данные**, полученные при проведении 50 экспериментов с одинаковыми начальными параметрами.

Чем больше состояний у автомата, тем более сложным может быть его поведение. Исходя из этого, можно попытаться улучшить точность предсказания с помощью увеличения числа состояний флиба. В табл. 3 приведены результаты экспериментов для флибов, заданных автоматом без флагов и имеющих число состояний от 10 до 80.

Как следует из рассмотрения этой таблицы, увеличение числа состояний действительно несколько улучшает точность найденного предсказателя.

Графики для экспериментов над флибами без флагов с разным числом состояний приведены на рис. 4. Точками изображен график для результатов эксперимента над флибами с 10 состояниями. Штрихпунктирная линия используется для флибов с 20 состояниями, а пунктирная – для флибов с 40 состояниями. График для результатов эксперимента над флибами, имеющими 80 состояний, изображен сплошной линией.

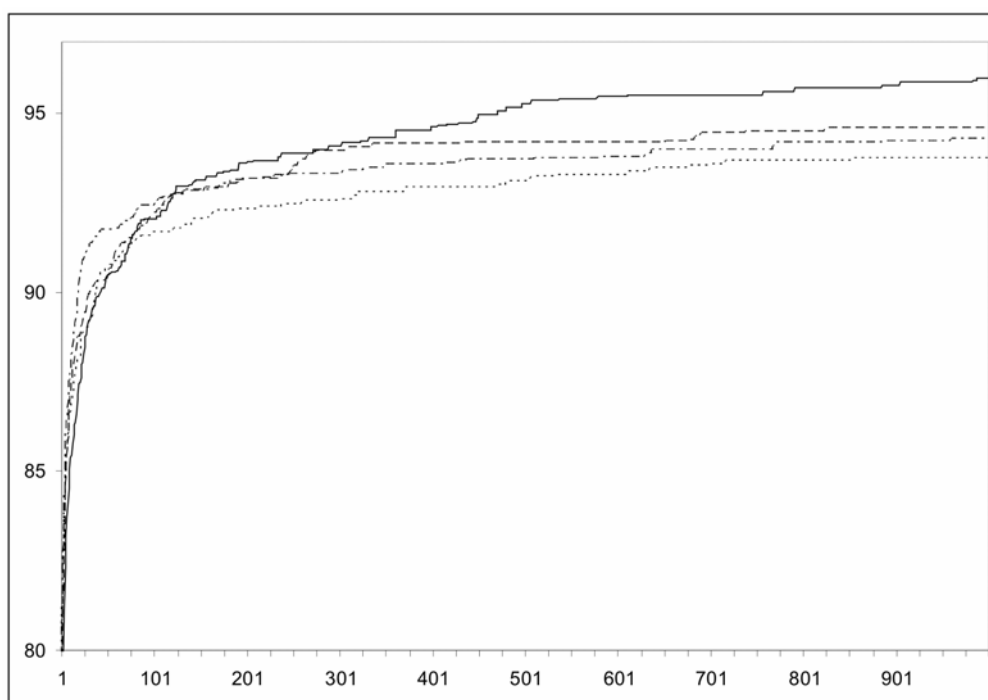


Рис. 4. Графики усредненных результатов предсказаний для эксперимента над флибами с разным числом состояний

Основным недостатком увеличения числа состояний является рост объема вычислений и, как следствие, увеличение времени, необходимого для построения предсказателя

Число состояний	Число флагов	Худший результат	Усредненный результат	Лучший результат
80	0	89	96	100
40	1	90	96,82	100
20	2	95	98,54	100
10	3	95	99,74	100

Таблица 4. Результаты экспериментов над флибами с разным числом состояний и флагов

В табл. 4 приведены результаты экспериментов над флибами с флагами (вторая – четвертая строки таблицы) для разного числа состояний. Следует отметить, что число

флагов и состояний в этих экспериментах подобрано таким образом, что общее число дуг переходов во всех флибах одинаково.

Графики для экспериментов над флибами с разным числом флагов и состояний приведены на рис. 5. Точками изображен график для результатов эксперимента над флибами с 80 состояниями без флагов. Штрихпунктирная линия используется для флибов с 40 состояниями и одним флагом, а пунктирная линия используется для флибов с 20 состояниями и двумя флагами. График результатов эксперимента над флибами с 10 состояниями и тремя флагами изображен сплошной линией.

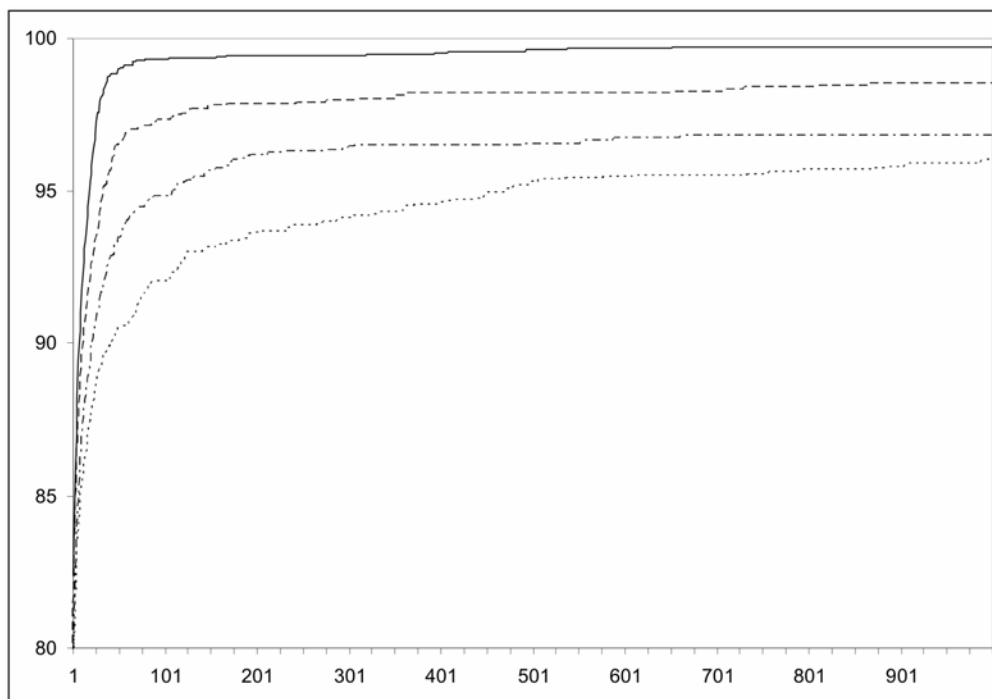


Рис. 5. Графики усредненных результатов предсказаний для экспериментов над флибами с разным числом состояний и флагов

Из рассмотрения графиков на рис.5 и данных в табл. 4 следует, что увеличение числа флагов приводит к улучшению качества предсказания.

Таким образом, применение автоматов с флагами для моделирования флибов позволяет строить более точные предсказатели.

Заключение

В работе показано, что добавление флагов к автоматам, моделирующим флибы, является более эффективным методом для улучшения точности работы алгоритма, по сравнению с простым увеличением числа состояний флиба.

Литература

1. Mitchell M. An Introduction to Genetic Algorithms. – MIT Press. Cambridge. MA, 1996.
2. Langdon W., Poli R. Better Trained Ants for Genetic Programming, 1998.
3. Wolfram S. A New Kind of Science. – Wolfram Media, 2002.
4. Mitchell M., Crutchfield J., Hraber P. Evolving cellular automata to perform computations. // Physica D. – 1993. – 75. – P.361–391.

5. Бедный Ю.Д. Применение генетических алгоритмов для решения одной задачи на клеточных автоматах. Задача классификации плотности для клеточных автоматов. СПбГУ ИТМО. Бакалаврская работа. 2006. <http://is.ifmo.ru>, раздел «Работы».
6. Whitley D. A Genetic Algorithm Tutorial. Statistics and Computing. 1994. 4, pp. 65-85. <http://www.cs.colostate.edu/~genitor/Pubs.html>.
7. Воронин О., Дьюдни А. Дарвинизм в программировании // Мой компьютер. – 2004. – № 35. <http://www.mycomp.kiev.ua/text/7458>
8. Лобанов П.Г., Шалыто А.А. Использование генетических алгоритмов для автоматического построения конечных автоматов в задаче о «флибах» / Материалы 1-й Российской мультikonференции по проблемам управления. Сборник докладов 4-й Всероссийской научной конференции "Управление и информационные технологии" (УИТ-2006). – СПбГЭТУ "ЛЭТИ", 2006. – С. 144–149. http://is.ifmo.ru/works/_flib.pdf
9. Miller B., Goldberg M. Genetic algorithms, tournament selection, and the effects of noise // Complex Systems. – 1995. – 3. – P. 193–212.
10. De Jong K. An analysis of the behavior of a class of genetic adaptive systems. PhD thesis. Univ. Michigan. Ann Arbor, 1975.
11. Шалыто А.А. Switch-технология. Алгоритмизация и программирование задач логического управления. – СПб.: Наука. 1998, 628 с. <http://is.ifmo.ru/books/switch/1>

НАШИ АВТОРЫ

Блинова Ирина Владимировна – ассистент кафедры высшей математики

Бобцов Алексей Алексеевич – доктор технических наук, доцент кафедры систем управления и информатики

Вельдер Сергей Эдуардович – студент кафедры компьютерных технологий

Гуров Вадим Сергеевич – ассистент кафедры компьютерных технологий

Денисов Константин Михайлович – старший преподаватель

Домрачев Владимир Михайлович – кандидат технических наук, доцент ФГУП ЦНИИАГ

Жданов Иван Николаевич – научный сотрудник кафедры электротехники

Коробейников Анатолий Григорьевич – доктор технических наук, профессор кафедры: проектирования компьютерных систем

Ле Зуй Туан – аспирант кафедры прикладной и компьютерной оптики

Лобанов Павел Геннадьевич – аспирант кафедры компьютерных технологий

Мазин Максим Александрович – аспирант кафедры компьютерных технологий

Марусина Мария Яковлевна – доктор технических наук, заведующая кафедрой измерительных технологий и компьютерной томографии

Наговицина Анна Георгиевна – аспирант кафедры систем управления и информатики

Никифоров Владимир Олегович – доктор технических наук, профессор, проректор

Папченко Борис Петрович – старший научный сотрудник кафедры: проектирования компьютерных систем

Попов Игорь Юрьевич – доктор физико-математических наук, профессор, заведующий кафедрой высшей математики

Свитенков Андрей Игоревич – студент 3 курса

Ткалич Вера Леонидовна – доктор технических наук, профессор кафедры: проектирования компьютерных систем

Царев Федор Николаевич – студент кафедры компьютерных технологий

Чивилихин Сергей Анатольевич – кандидат физико-математических наук, доцент

Шалыто Анатолий Абрамович – доктор технических наук, профессор, заведующий кафедрой технологий программирования

СОДЕРЖАНИЕ

ПЛЕНАРНОЕ ЗАСЕДАНИЕ.....	3
Марусина М.Я., Казначеева А.О. Современное состояние и перспективы развития томографии	3
Секция 1. ФУНДАМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ В РАМКАХ ТЕМАТИЧЕСКОГО ПЛАНА НАУЧНО- ИССЛЕДОВАТЕЛЬСКИХ РАБОТ УНИВЕРСИТЕТА, ФИНАНСИРУЕМЫХ ФЕДЕРАЛЬНЫМ АГЕНТСТВОМ ПО ОБРАЗОВАНИЮ, И НАУЧНО-ИССЛЕДОВАТЕЛЬСКИХ РАБОТ ПО КОНТРАКТАМ, ФИНАНСИРУЕМЫМ ФЕДЕРАЛЬНЫМ АГЕНТСТВОМ ПО НАУКЕ В 2006 ГОДУ	14
Жданов И.Н., Домрачев В.М., Денисов К.М. Оптимизация информационных подсистем следящих электроприводов прецизионных комплексов слежения.....	14
Коробейников А.Г., Ткалич В.Л. Оптико-геометрический метод исследования напряженно-деформированных состояний элементов микроэлектроники и микросенсорной техники	21
Коробейников А.Г., Ткалич В.Л., Папченко Б.П. Разработка автоматизированной системы контроля качества микро - и наноструктур	24
Гуров В.С., Мазин М.А., Шалыто А.А. Тестовый язык автоматного программирования.	29
Вельдер С.Э., Шалыто А.А. Введение в верификацию автоматных программ на основе метода MODEL CHECKING.....	33
Бобцов А.А., Наговицина А. Г., Никифоров В.О. Алгоритм адаптации с астатизмом первого порядка для стабилизации линейного возмущенного нестационарного объекта	49
Блинова И.В., Попов И.Ю., Свитенков А И., Чивилихин С.А. Формирование наносвитков в вязкой жидкости	56
Секция 2. ФУНДАМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ В ОБЛАСТИ ЕСТЕСТВЕННЫХ, ТЕХНИЧЕСКИХ И ГУМАНИТАРНЫХ НАУК. НАУЧНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ РАЗВИТИЯ ИНФРАСТРУКТУРЫ ВУЗОВСКОЙ НАУКИ.....	60
Ле Зуй Туан (асп.). Разработка алгоритмического и программного обеспечения для моделирования апертуры БРОМ – зондов на основе экспериментального исследования дальнепольной освещенности.....	60
Лобанов П.Г. (асп.), Шалыто А.А. Использование автоматов с флагами для решения задачи о «флибах»	67
НАШИ АВТОРЫ	75

Научно-технический вестник СПбГУ ИТМО. Выпуск 42.
Фундаментальные и прикладные исследования информационных систем и технологий / Главный редактор д.т.н., проф. В.Н. Васильев. – СПб: СПбГУ ИТМО, 2006. 77 с.

НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК СПбГУ ИТМО
Выпуск 42

**ФУНДАМЕНТАЛЬНЫЕ
И ПРИКЛАДНЫЕ ИССЛЕДОВАНИЯ
ИНФОРМАЦИОННЫХ СИСТЕМ
И ТЕХНОЛОГИЙ**

Главный редактор
доктор технических наук, профессор
В.Н. Васильев
Дизайн обложки В.А. Петров, А.А. Колокольников
Редакционно-издательский отдел СПбГУ ИТМО
Зав. РИО Н.Ф. Гусарова
Лицензия ИД № 00408 от 05.11.99.
Подписано в печать 15.11.07.
Заказ 1100. Тираж 100 экз.