

Информатика и её применения

Том 4 Выпуск 4 Год 2010

СОДЕРЖАНИЕ

Система управления электронной библиотекой LibMeta А. А. Захаров, В. А. Серебряков	2
Интеграция гетерогенной информации о цветных пикселях и их цветовосприятии О. П. Архипов, З. П. Зыкова	14
Математическая модель и метод восстановления позы человека по стереопаре силуэтных изображений А. К. Цискаридзе	26
Моделирование и классификация многоканальных дистанционных изображений с использованием копул В. А. Крылов	33
Связь между временными и структурно-топологическими характеристиками диаграмм ритма сердца здоровых людей А. А. Кузнецов	38
Semiformal verification for pipelined digital designs based on Algorithmic State Machines S. Varanov, S. Frenkel, and V. Zakharov	48
Алгоритм координации для гибридной интеллектуальной системы решения сложной задачи оперативно-производственного планирования А. В. Колесников, С. А. Солдатов	60
Стационарные характеристики двухканальной системы обслуживания с переупорядочиванием заявок и распределениями фазового типа С. И. Матюшенко	67
Аппроксимация распределения оценки риска пороговой обработки вейвлет-коэффициентов нормальным распределением при использовании выборочной дисперсии О. В. Шестаков	72
Оценки скорости сходимости распределений случайных сумм с безгранично делимыми индексами к нормальному закону С. В. Гавриленко	80
К вопросу о размещении коллективных средств отображения в ситуационном зале с заданными параметрами К. Г. Чупраков	88
Abstracts	96
Об авторах	100
About Authors	101
Авторский указатель за 2010 г.	102
2010 Author Index	105

СИСТЕМА УПРАВЛЕНИЯ ЭЛЕКТРОННОЙ БИБЛИОТЕКОЙ LIBMETA

А. А. Захаров¹, В. А. Серебряков²

Аннотация: В статье рассмотрены вопросы создания современных электронных библиотек (ЭБ). Перечислены основные требования к ЭБ, в частности требования по интеграции с внешними системами и стандартизации, а также предложены некоторые пути их удовлетворения. Рассмотрены основные мировые стандарты в области ЭБ. Представлена к рассмотрению система управления ЭБ (СУЭБ) LibMeta — универсальное средство для создания ЭБ.

Ключевые слова: электронные библиотеки; Единое научное информационное пространство (ЕНИП)

1 Введение

В последние годы объемы информации в сети Интернет в связи с ее бурным развитием лавинообразно увеличиваются [1]. Несмотря на все большее проникновение технологий Semantic Web [2, 3], ощущается серьезная нехватка средств поиска и каталогизации информации, которые позволяли бы искать ее именно по семантике и связям, а не только по ключевым словам и полным текстам, как это делают универсальные поисковые системы. Одним из способов решения данной проблемы видится появление и все большее распространение различного рода ЭБ [4, 5].

Зачастую понятие ЭБ [6] смешивают с уже давно существующими электронными каталогами — информационными системами, предназначенными для использования в классических библиотеках, содержащими только метаинформацию и служащими лишь средством поиска. Однако ЭБ — это совершенно другой класс систем: они хранят не только метаданные, но также и цифровые версии информационных ресурсов. Еще одним важным признаком ЭБ является то, что в качестве информационного ресурса может пониматься фактически все что угодно, например музейные предметы или архивные материалы. В связи с этим можно говорить о сближении задач ЭБ, электронных архивов и цифровых музеев — все эти информационные системы фактически являются ЭБ с различной специализацией [7].

Несмотря на перспективы, в ходе развития ЭБ предстоит преодолеть немало препятствий. Самым большим из них, пожалуй, является недостаток в хороших стандартах в сфере ЭБ. Конечно же, существует много стандартов, так или иначе касающихся разных аспектов ЭБ, однако фактически

нет архитектурных стандартов, которые описывали бы не только концептуальные части ЭБ, но и выдвигали бы требования к хранимой информации. Ситуация со стандартами на форматы данных немногим лучше: практически каждая создаваемая СУЭБ предлагает свой стандарт хранения и обмена данными, несовместимый с другими. Кроме того, в настоящее время во многих системах ЭБ стандарты не только не применяются вовсе, но и, что еще хуже, используются в ограниченном или измененном виде. Немаловажной проблемой видится и то, что уже существующие и вновь создаваемые ЭБ слабо интегрированы друг с другом: поиск и получение данных из нескольких ЭБ возможен только при непосредственном их посещении и повторении поискового запроса [5].

В Вычислительном центре РАН на протяжении нескольких лет ведется исследование вопросов, связанных с ЭБ, и как результат этих исследований в данной статье представлена СУЭБ LibMeta. В первой части статьи более подробно рассмотрены основные проблемы и требования к ЭБ, далее упоминаются наиболее полезные стандарты в сфере ЭБ и в третьей части описана СУЭБ LibMeta.

2 Требования к электронной библиотеке

Современный мир предъявляет высокие требования к информационным системам: с точки зрения пользователей они должны быть удобны в использовании, просты в изучении; а с технической точки зрения должны быть тесно связаны с другими информационными системами и предоставлять стандартизованные службы. Все эти требования в полной мере относятся и к ЭБ.

¹Вычислительный центр им. Дородницына Российской академии наук, andreya@sufiler.ru

²Вычислительный центр им. Дородницына Российской академии наук, serebr@ccas.ru

2.1 Интеграция в другие информационные системы

Фактически любая современная информационная система, и в частности ЭБ, должна быть интегрирована с другими информационными системами. Это дает сразу несколько серьезных преимуществ:

- отсутствие дублирования данных: исходные данные хранятся только в одной информационной системе, в других системах они используются по ссылкам того или иного вида либо реплицируются и автоматически обновляются при обновлении оригинала;
- централизация сервисных служб, например служба аутентификации и авторизации пользователей: пользователи системы вводят свою идентификационную информацию и проходят аутентификацию лишь единожды для всей группы связанных информационных систем;
- интеграция информационных ресурсов: ресурсы, даже хранящиеся в разных системах, представляются связанными друг с другом единой системой навигации.

Для того чтобы снизить сложность разработки таких интегрированных информационных систем, имеет смысл строить их на одной общей архитектурной и технологической основе. Одним из таких архитектурных решений является «Единое научное информационное пространство (ЕНИП) РАН». ЕНИП — это проект, реализуемый в последние несколько лет в рамках программы Президиума РАН «Информатизация». Целью этой программы является создание и объединение информационных систем подразделений и научных институтов РАН для удовлетворения потребностей научных сотрудников как в части поиска качественной информации, так и в выставлении собственной информации в сети Интернет.

Одними из наиболее важных составляющих ЕНИП являются так называемые схемы метаданных, которые представляют собой формализованные описания метаданных, циркулирующих в ЕНИП [8]. Выдвинуты следующие требования к метаданным, описываемым схемами ЕНИП, которые должны:

- включать в себя основные типы информации, требующейся для поддержки работы научного сотрудника;
- быть открытыми, т.е. обеспечивать доступ к соответствующей информации по этим описаниям;
- быть расширяемыми, т.е. обеспечивать возможность детализации описаний;

- обеспечивать возможности интеграции информации;
- обеспечивать возможности уникальной идентификации информации;
- обеспечивать возможности размещения и поиска информации в распределенной среде;
- обеспечивать возможности интероперабельности с внешней средой.

Кроме того, рекомендовано, чтобы схемы метаданных были ориентированы на семантический Веб (Semantic Web) [2, 3].

Схемы метаданных в ЕНИП являются не только обменными схемами разного уровня детализации, но и служат основой для построения конкретных информационных систем, входящих в пространство.

2.2 Распределенность

Одним из наиболее бурно развивающихся направлений информационной индустрии последних лет стала разработка распределенных информационных систем. Причинами их быстрого роста стали достижения одновременно в нескольких областях, среди которых нужно особо отметить:

- значительный рост пропускной способности каналов связи: скорость обмена по ним приближается к скоростям внутренних шин компьютеров;
- рост производительности компьютеров как по скорости, так и по объемам памяти: и оперативной, и внешней;
- широкое проникновение компьютеров и компьютерных технологий в повседневную деятельность как большинства организаций и учреждений, так и граждан;
- развитие сети Интернет, обеспечивающей простой и надежный доступ к огромному числу информационных ресурсов;
- развитие самих информационных технологий; так, можно сказать, что программирование находится на четвертой фазе своего развития: 1) «классическое» программирование (сначала в кодах, затем на ассемблере, и далее на языках высокого уровня) для больших ЭВМ; 2) «классическое» программирование для персональных ЭВМ; 3) программирование с использованием визуальных и CASE средств; 4) «сетевое» программирование.

Тенденция к увеличению доли распределенных систем не обошла и Российскую академию наук. Она имеет разветвленную структуру, которая объединяет большое число научно-исследовательских

учреждений и коллективов, расположенных на всей территории России и вовлеченных во все многообразие видов научной деятельности. Эти учреждения обладают уникальными научными информационными ресурсами. Среди них: опубликованные результаты научных исследований и экспериментов, библиографические и фактографические базы данных, сведения об ученых, их научной деятельности, публикациях, проектах и т. п. Эти ресурсы представляют значительный интерес для сотрудников научных и административных учреждений, членов мирового научного сообщества, для представителей промышленности и предпринимателей, которые заинтересованы во внедрении результатов научных исследований.

В связи с перечисленными факторами при разработке современной ЭБ повышенное внимание следует уделять созданию средств взаимодействия ЭБ друг с другом, а также средств по организации распределенных ЭБ, где данные, объем которых достаточно велик, чтобы их хранить в одной ЭБ, распределены между отдельными хранилищами и существуют специальные средства поиска и каталогизации, работающие «над» всеми данными такой ЭБ.

2.3 Следование стандартам

В настоящее время существует огромное число стандартов на данные и метаданные. В связи с этим возникает вопрос, в каком формате данные должны храниться в ЭБ и должны ли они в том же самом формате предоставляться конечным пользователям? Зачастую оказывается, что формат, удобный для хранения данных (пригодный для полнотекстового поиска, обеспечивающий целостность данных, легкость доступа к ним, содержащий минимальный набор метаданных), не слишком удобен для пользователей, поскольку требует от них установки специального программного обеспечения, либо не приспособлен к передаче по сети, так как не обеспечивает необходимой степени сжатия. Таким образом, возникает необходимость поддержки как минимум двух форматов представления данных для каждого вида ресурсов (тексты, аудио, видео, статические изображения). Также не следует забывать, что с течением времени существующие стандарты устаревают и возникают новые, что приводит к необходимости перехода на новые форматы.

В связи со всем вышесказанным при разработке ЭБ, ориентированной на долговременное хранение, следует учесть необходимость поддержки различных форматов представления данных, а также преобразования между форматами, как в целях

предоставления пользователям ЭБ доступа к данным, так и в целях перехода на новый формат хранения данных в случае необходимости.

2.4 Подготовка ресурсов

В большинстве случаев, когда ЭБ не предназначена для размещения так называемых «изначально цифровых» (born-digital) материалов, подготовка ресурсов к публикации является достаточно сложным и трудоемким процессом, вовлекающим многих участников. За простотой пользовательского интерфейса скрывается целая подсистема подготовки ресурсов, включающая в себя следующие службы:

- оцифровки, создаваемые при библиотеках, музеях и других поставщиках данных там, где возможен непосредственный доступ к оцифровываемым материалам;
- контроля качества оцифровки — единая служба контроля качества оцифровываемых данных, в которой задействованы специалисты по оцифровываемым предметам. Несмотря на то что первичный контроль качества выполняется при самой оцифровке, иногда происходит утеря какой-либо части оцифровываемых данных либо оказывается, что с точки зрения специалиста по оцифровываемым данным оцифровка выполнена не в полном объеме;
- подготовки метаданных. Иногда она может быть объединена со службой оцифровки, однако следует учитывать, что ввод метаданных должен осуществлять не специалист по оцифровке, а специалист в предметной области;
- окончательной подготовки ресурсов и контроля качества, которая выполняет объединение данных и метаданных, следит за обновлением данных и метаданных и осуществляет общий контроль качества подготавливаемых ресурсов.

Перечисленные службы должны обладать своим инструментарием, обеспечивающим выполнение их задач и предоставляющим друг другу необходимые для работы данные. В результате работы подсистемы подготовки ресурсов получают готовые к публикации в ЭБ ресурсы, которые по каналам связи автоматически представляются к публикации.

3 Стандартизация в электронных библиотеках

Проблемой стандартов для ЭБ занимаются многие сообщества и организации [9–13], но,



Рис. 1 Разделение понятий в DELOS

несмотря на значительные усилия и имеющиеся достижения, единства мнений добиться пока не удастся. Необходимость стандартизации в этой области ощущается, пожалуй, сильнее, чем в других видах информационно-поисковых систем, так как многие проекты ЭБ стараются объединить ресурсы нескольких существующих библиотечных и/или архивных систем.

Стандарты, имеющие отношение к ЭБ, условно можно разделить на три большие группы:

- архитектурные стандарты, описывающие принципы и крупные компоненты, применяемые при построении ЭБ;
- стандарты метаданных и их представления, описывающие наборы метаданных и, возможно, способы их представления при передаче и хранении;
- стандарты информационного обмена, описывающие протоколы передачи метаданных и данных между системами в различных целях.

Рассмотрим наиболее интересные и полезные, по мнению авторов, стандарты в сфере ЭБ.

3.1 Архитектура электронных библиотек

Архитектурные стандарты описывают, как должна быть устроена ЭБ. Такие описания могут касаться как интерфейсов взаимодействия с другими системами и пользователями, так и состава и назначения компонентов системы. Стандарты этой группы также пытаются дать четкое определение самому термину ЭБ и всем связанным с ним терминам.

OAIS

Open Archival Information System [14] — это разрабатываемый и поддерживаемый организацией Consultative Committee for Space Data Systems [9] международный стандарт (ISO 14721:2003) на образцовую модель информационной системы

открытого архива. В нем описывается подход к долговременному хранению данных, в частности освещается проблема устаревания форматов данных и физических носителей информации. Также описываются схемы взаимодействия участников архивной системы при различных сценариях работы. Кроме технической стороны работы открытого архива рассматривается также и работа административно-управленческих отделов при ЭБ, занимающихся планированием хранения информации, предоставлением доступа и другими задачами.

DELOS

Группа DELOS [10] (существует под различными названиями с 1996 г.) является одной из наиболее известных и старых организаций, занимающихся стандартизацией в области ЭБ. Основным направлением ее деятельности является исследование и создание стандартов в области архитектуры ЭБ, их интеграции и отчуждения данных.

Самый интересный результат работы DELOS — образцовая модель ЭБ Digital Library Reference Model (DLRM) [15]. В этой модели большое внимание уделяется отделению цифровых объектов (содержимого ЭБ) от ЭБ и СУЭБ (рис. 1). Также вводятся различные классы пользователей ЭБ, решающих свои задачи в рамках ЭБ и которым ЭБ (и СУЭБ) должна предоставлять соответствующую функциональность, приводится весьма обширная концептуальная модель данной области с тщательными определениями важнейших представлений об архитектуре, ресурсах и функциональности ЭБ, в частности в UML-представлении.

3.2 Представление метаданных

Стандарты метаданных играют в ЭБ очень важную роль: они описывают, в каком формате данные представляются для хранения и передачи из одной информационной системы в другую. Следует

заметить, что не все стандарты метаданных в дополнение к самому описанию сущностей, их связей и атрибутов представляют описание контейнера для метаданных, однако большая часть стандартов подразумевает использование XML как одного из видов такого контейнера.

Dublin Core

Стандарт Dublin Core [11] состоит из двух частей: минимальной Dublin Core Metadata Element Set, имеющей статус международного стандарта (ISO 15836:2009), и полной — DCMI Metadata Terms. Обе версии стандарта не привязаны ни к какой конкретной предметной области и могут описывать ресурсы любых видов. Минимальная версия содержит только 15 атрибутов, применяемых к любым сущностям, в которых можно задать название объекта, его описание, автора, административные метаданные и ссылку на источник. Полная версия содержит все атрибуты минимальной, а также наборы дополнительных сущностей, словарей и дополнительное множество атрибутов. Чаще всего используется именно минимальный набор, поскольку именно он позволяет максимально абстрагироваться от предметной области, но при этом, конечно же, проигрывая в деталях.

Также существует рекомендация DCMI Abstract Model, содержащая сведения по расширению DCMI Metadata Terms для нужд различных предметных областей.

Кроме описания схемы метаданных стандарты Dublin Core содержат описания контейнеров, в частности описан формат DC-Text для хранения в чисто текстовом формате, формат DC-XML для XML-представления и DC-RDF для представления в RDF-подмножестве XML.

CIDOC-CRM

С 1994 г. в составе International Council of Museums существует комитет Committee on Documentation of the International Council of Museums (CIDOC) [12], который занимается стандартизацией в области музейных метаданных. Основной целью данного комитета является создание образцовой концептуальной модели (CRM) для описания сущностей и связей, используемых в документировании культурного наследия. В 1999 г. была выпущена первая версия CIDOC-CRM [16], а в 2006 г. модель получила статус международного стандарта (ISO 21127:2006).

Модель CIDOC-CRM чрезвычайно подробно описывает предметную область, и поэтому на практике в полном объеме практически никогда не применяется. Тем не менее отображение метаданных

ЭБ на эту модель представляется полезным, так как многие музейные системы заявляют свою совместимость с CIDOC-CRM.

FRBR и FRBRoo

Изначально разрабатывавшийся The International Federation of Library Associations and Institutions стандарт Functional Requirements for Bibliographic Records (FRBR) [17] имеет примерно такую же роль в области описания публикаций, какую имеет CIDOC-CRM в описаниях предметов культурного наследия. С 2000 г. при содействии комитета CIDOC разрабатывается объектно-ориентированный стандарт FRBRoo [18], заимствующий часть сущностей и идей из CIDOC-CRM.

PRISM

Стандарт Publishing Requirements for Industry Standard Metadata (PRISM) [13] разработан издательскими организациями, входящими в IDEAlliance (International Digital Enterprise Alliance), для обмена метаданными о публикациях. PRISM основан на DCMI, но в большей степени ориентирован на библиографические ресурсы. В стандарте предлагается среда обмена и хранения данных и метаданных и ряд словарей значений этих элементов. В ЕНИП используется схема контролируемых словарей PRISM и основной набор элементов PRISM.

MARC

Стандарт MARC (MAchine-Readable Cataloging) по праву может считаться одним из самых старых стандартов в информационных технологиях. Он был разработан в начале 1960-х гг. в библиотеке конгресса США. Основное предназначение стандарта — хранение библиографических записей для электронных каталогов. В настоящее время оригинальный MARC (USMARC) уже не используется, на смену ему пришли модификации MARC 21 [19] (США и Канада) и национальные модификации (например, RUSMARC [20] в России). Также существует отображение MARC на XML, называемое MARCXML [21]. Хотя MARC основан на устаревших технологиях, он достаточно широко используется и многие современные информационные системы заявляют свою с ним совместимость.

3.3 Информационный обмен

Стандарты информационного обмена предназначены в первую очередь для открытых и распределенных систем. Руководствуясь такими стандартами, можно получать доступ к данным и метаданным ЭБ, ничего не зная об их устройстве,

и, в свою очередь, предоставлять другим системам доступ к собственным данным и метаданным, не раскрывая деталей реализации и способов хранения.

OAI-PMH

Стандарт Open Archives Initiative — Protocol for Metadata Harvesting [22] является признанным лидером среди стандартов распределенного поиска и репликации метаданных. Основу такой популярности обеспечило то, что он достаточно прост в реализации и может инкапсулировать метаданные в любом XML-формате. Каждая OAI-PMH совместимая система должна поддерживать метаданные в формате Dublin Core.

OAI-ORE

Object Reuse and Exchange [23] — Semantic Web-ориентированный стандарт описания и обмена агрегированными ресурсами. Стандарт описывает способы объединения частей объекта (данных, метаданных) в один агрегированный ресурс, доступный по единому URL, а также способы обработки и предоставления частей объекта по такому URL.

4 Система управления электронными библиотеками LibMeta

С 2007 г. в ВЦ РАН ведутся работы по созданию СУЭБ в рамках ЕНИП [24] под названием LibMeta, которая позволила бы библиотекам, архивам и музеям РАН иметь унифицированное решение, позволяющее публиковать полные тексты научных работ и разнообразные мультимедийные материалы, быть интегрированной в существующие информационные системы РАН, а также соответствовать стандартам в области ЭБ.

Портал ЭБ «Научное наследие России» [25] является первой установкой СУЭБ LibMeta, а также площадкой для обкатки технологических и архитектурных решений.

4.1 Архитектура СУЭБ LibMeta

Поскольку предполагается использование СУЭБ LibMeta в научных институтах различной направленности, невозможно предоставить одну единственную схему метаданных, подходящую абсолютно под все задачи. Данная проблема может быть решена двумя способами: внесением избыточности в схему метаданных или предоставлением администратору системы возможности доопределять схему.

Оба подхода не лишены своих недостатков. Так избыточная схема приводит к тому, что в каждой конкретной установке ЭБ используется только часть схемы, как правило достаточно небольшая. Это приводит к нерациональному использованию ресурсов и уменьшению быстродействия. Кроме того, внесение такой избыточности требует глубоких исследований в предметной области, а также наличия общепринятых стандартов на подобные метаданные.

Основным недостатком второго подхода является то, что по метаданным, определенным администратором, поиск может осуществляться только с большими затратами ресурсов. Кроме того, сужаются возможности по связыванию ресурсов друг с другом, так как администратор может определять только содержательные атрибуты.

В СУЭБ LibMeta применяются оба подхода: для ресурсов типа «публикация» и «персона» применяется избыточная схема, разработанная на основе библиотечных стандартов, а для «музейных предметов» используется второй подход. Такое решение основано на том, что для музейных предметов практически невозможно создать единую (избыточную) схему, поскольку каждый музей обладает своей спецификой. Естественно, что для музейных предметов существует и фиксированная часть схемы, позволяющая связать их с другими ресурсами и включающая основные содержательные метаданные.

При разработке СУЭБ LibMeta были проанализированы мировые стандарты в области ЭБ, а также обобщен опыт их использования в других библиотечных системах. На основе этого анализа были выдвинуты предложения по использованию стандартов.

В качестве архитектурных стандартов были использованы как DELOS [15], так и OAIS [14]. Из DELOS взяты общие концепции и варианты использования ЭБ. Из OAIS почерпнуты сведения о процессах, происходящих внутри ЭБ, а также о ее взаимодействии с внешними системами.

Из стандартов метаданных были использованы Dublin Core (unqualified) [11], CIDOC-CRM [16], RUSMARC [20] и др. На эти стандарты существует отображение схемы метаданных СУЭБ LibMeta и возможен информационный обмен по протоколам, поддерживающим инкапсуляцию данных в этих форматах. Перечисленные стандарты являются наиболее распространенными в мире ЭБ, и их использование позволяет достичь максимальной совместимости с существующими системами.

В качестве основного стандарта информационного обмена выбран OAI-PMH [22], поскольку он

поддерживается большей частью библиотечных и архивных систем.

Основой для обмена данными и семантической интероперабельности в ЕНИП служат технологии Semantic Web. Соответственно, оправданным представляется применение в ЕНИП существующих предложений по стандартизации наборов элементов метаданных для Semantic Web [2, 3, 8]. В технологиях Semantic Web широко используется язык RDF, а также его специализация для описания онтологий OWL. Логично было выбрать именно OWL как язык описания метаданных в СУЭБ LibMeta, а RDF — как язык обмена метаданными между системами.

В качестве инфраструктурного решения для реализации СУЭБ LibMeta была выбрана программная платформа, разработанная ВЦ РАН, — система «Научный институт» [26].

«Научный институт» представляет собой типовой программный комплекс автоматизации информационной деятельности научного института в составе Российской академии наук, поддержки научной деятельности его сотрудников, взаимодействия с другими информационными системами в составе ЕНИП. «Научный институт» изначально разрабатывается как модульная, расширяемая система, позволяющая гибко подбирать наиболее подходящий набор функциональных возможностей для каждой конкретной организации.

Основу системы составляет ядро («платформа») — программное решение, предназначенное для создания распределенных информационных систем, веб-порталов, интеграции данных, и набор функциональных модулей, предоставляющих специальные функции. Для СУЭБ LibMeta «Научный институт» является слоем абстракции от подсистемы хранения и контейнером приложений. Кроме того, в СУЭБ LibMeta используется библиографический профиль метаданных ЕНИП, предоставляющий большую часть схемы метаданных в части определения таких ресурсов, как публикации, персоны, проекты и организации. В состав системы «Научный институт» входит также большое число подключаемых модулей, которые могут быть легко включены в установки СУЭБ LibMeta, к примеру для организации форума.

4.2 Электронная библиотека «Научное наследие России»

Электронная библиотека «Научное наследие России» (рис. 2) разрабатывается в рамках одноименной программы Президиума РАН с целью обеспечения сохранности и предоставления публичного доступа к научным трудам известных российских и зарубежных ученых и исследователей, работавших на территории России. Некоторые из подсистем ЭБ (системы хранения и представления электронных изданий конечным пользователям)

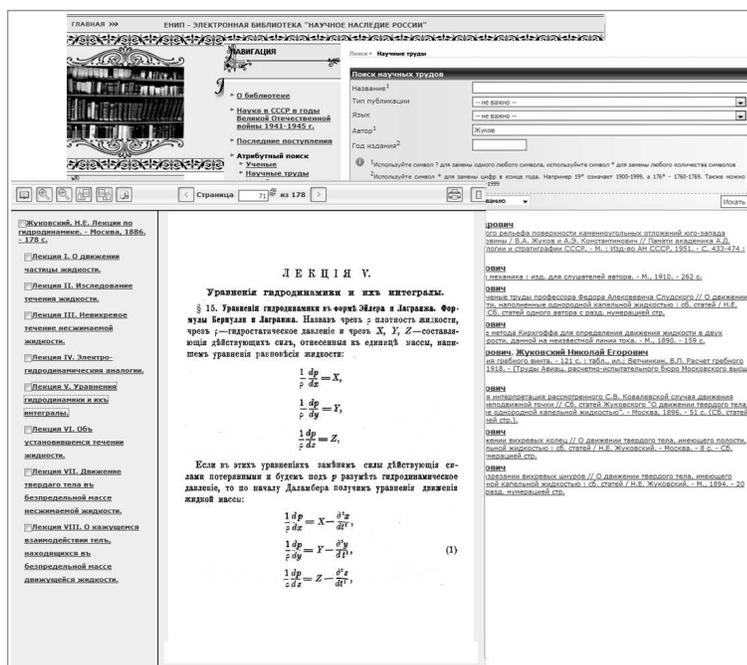


Рис. 2 Система «Научное наследие России»

создаются в рамках программы Президиума РАН «Информатизация». Общая координация и управление проектом осуществляется Межведомственным суперкомпьютерным центром (МСЦ) РАН. Задачами подготовки электронных изданий и сопровождающей информации для размещения в хранилище данных электронной библиотеки занимаются ведущие библиотеки РАН, среди которых БАН, БЕН (Центральная библиотека и ее отделение), ИНИОН.

Другой важной задачей этого проекта является интеграция существующих библиотечных ресурсов в ЕНИП РАН и обеспечение возможности централизованного доступа к ресурсам существующих хранилищ электронных изданий и метаданных об ученых и их научных трудах. Данная задача решается путем определения единой инфраструктуры распределенной системы, унификации форматов данных и протоколов взаимодействия компонентов системы, разработки единых регламентов подготовки и сопровождения электронных изданий.

Профиль метаданных СУЭБ LibMeta

Одним из самых существенных недостатков многих схем метаданных ЭБ является то, что они рассматривают метаданные только в контексте описываемых ими данных и только как набор полей для поиска информации и индексирования ресурсов. В случае, когда ЭБ содержит один вид ресурсов, например только книги, такой подход, возможно, оправдан, однако современные ЭБ содержат разно-

образные ресурсы, и в связи с этим такой подход неприемлем. Метаданные разных ресурсов должны содержать ссылки друг на друга, при этом оставаясь достаточно независимыми. Одним из наиболее удобных подходов к описанию такого рода метаданных является использование OWL-онтологий. Основной частью профиля метаданных в ЕНИП и «Научном институте» являются как раз такие онтологии [8]. Общая схема профилей метаданных, применяемых в СУЭБ LibMeta, а также основных сущностей в этих профилях приведена на рис. 3.

Метаданные публикации

Базовый уровень **публикации** включает следующие свойства (здесь и далее курсивом выделены свойства, являющиеся ссылками на сущности LibMeta, по которым возможна навигация):

- название;
- альтернативный заголовок;
- аннотация;
- ключевые слова;
- *источник* — описание источника информации о данном ресурсе, например: наименование организации, ФИО и пр.;
- авторские права;
- веб-адрес;
- полный текст;
- язык;
- дата издания;

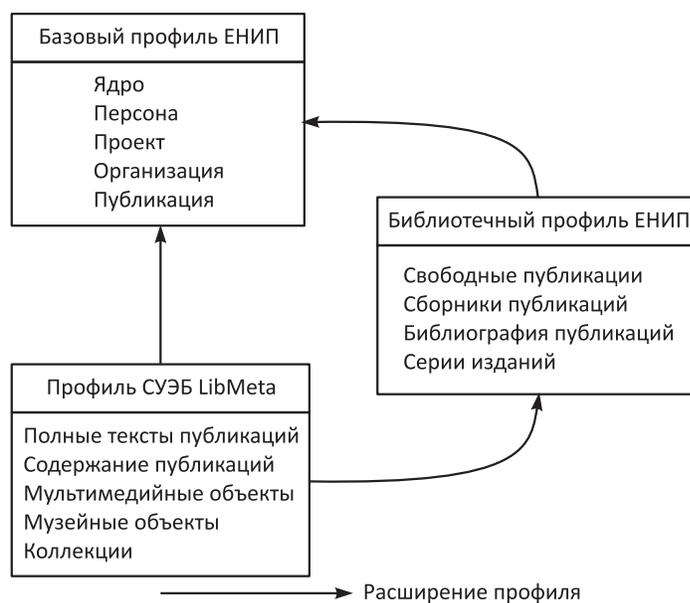


Рис. 3 Профили метаданных в СУЭБ LibMeta

- идентификатор — указание идентификатора ресурса с помощью рекомендуемых стандартных систем идентификации;
- *авторы*;
- *издатель*;
- *редактор*;
- *входит в состав* — данный ресурс является физически или логически частью указанного ресурса;
- *включает* — данный ресурс физически или логически включает указанный ресурс;
- количество страниц;
- реферат — реферат(ы) по данной публикации;
- библиографическое описание — библиографическое описание публикации по ГОСТу целиком, строкой. Может быть указано помимо отдельных элементов описания, указываемых полями «название», «номер тома/выпуска» и пр.;
- *полный код УДК*;
- *рубрика ББК*;
- *основной код УДК*;
- примечания.

При использовании публикаций в научно-исследовательском процессе существует необходимость быстрого ознакомления с содержимым публикации, и аннотации часто оказывается недостаточно. В связи с этим в инструментарии СУЭБ LibMeta разработаны средства полуавтоматического выделения оглавления с обеспечением ссылок на соответствующие разделы документа, а также средства работы с библиографическими ссылками.

Приведем описание фрагмента профиля ЭБ, отражающего решение этих задач. Расширенная схема описания библиографической информации включает:

- список литературы (текстом) — список библиографических ссылок в текстовом виде, если не может быть разобран;
- оглавление — оглавление данной публикации в виде отдельного файла, либо текстового или XML-фрагмента;
- список литературы — список библиографических ссылок, указанных в тексте данной публикации, в виде списка структур «Библиографическая ссылка»:
 - порядковый номер;
 - идентификатор ссылки — краткий идентификатор библиографической ссылки, например «DC» или «12»;

- текст ссылки — исходный текст библиографической ссылки, желательно отформатированный как библиографическое описание по ГОСТу;

- *цитируемая публикация*;

- сведения об издании — сведения, относящиеся к изданию: в какой редакции, данные об оригинале для переводной литературы, место и год издания;
- *составитель*;
- *коллективный автор публикаций*;
- *переводчик публикаций*;
- *редколлегия*.

Метаданные автора

В приведенных описаниях элементов профиля ЭБ можно заметить использование элементов основного профиля ЕНИП: «персона», «организационная единица», «файл данных» и других. Приведем состав наиболее часто используемого класса — «персоны»:

- домашняя страница;
- дата рождения;
- адрес — полный почтовый адрес;
- имя:
 - фамилия;
 - имя;
 - отчество;
 - значение — фамилия, имя отчество полностью, с дополнительными элементами (титулами и пр.);
- пол;
- ученая степень:
 - дата присуждения;
 - ученая степень;
 - специальность ВАК;
- ученое звание:
 - дата присуждения;
 - ученое звание;
 - *присудившая организация*;
- дата смерти;
- место рождения — место рождения данной личности (указывается в произвольной форме). Ввиду сложности поддержки исторической информации об административно-территориальном делении, классификатор регионов не используется для указания места рождения;

- место смерти — место смерти данной исторической личности (указывается в произвольной форме, как и место рождения);
- электронная почта;
- телефон;
- факс;
- www-страница;
- FTP-адрес.

Метаданные предмета

Сближение задач ЭБ, архивов и музеев в представлении научного наследия выдвигает требование стандартизации метаданных физических музейных предметов и их мультимедийных (фото, видео, аудио) представлений. В связи с этим в СУЭБ LibMeta разработан дополнительный прикладной профиль, в котором для сущности «**музейный предмет**» определены следующие свойства и связи:

- название;
- альтернативный заголовок;
- аннотация;
- ключевые слова;
- источник;
- *держатель* (место хранения);
- *состав* — набор ссылок на другие предметы, из которых состоит описываемый;
- *автор описания*;
- состояние (сохранность) — состояние предмета по шкале определяемой его держателем;
- количество предметов;
- номер — идентифицирующие номера предмета:
 - система нумерации;
 - номер;
- *автор сбора* — персона, впервые собравшая предмет;
- дата сбора;
- дата поступления — дата поступления предмета к хранителю;
- география — место находки данного предмета (указывается в произвольной форме); ввиду сложности поддержки исторической информации об административно-территориальном делении, классификатор регионов не используется;
- размеры;
- возраст — предполагаемая дата создания предмета;
- эра — описание эры с указанием в свободной форме ее начала и конца, например «наша эра» ведет отсчет от Рождества Христова;
- время от начала — время от начала эры до наступления события. Формат зависит от эры, например для нашей эры форматом будет дата по григорианскому календарю;

– способ поступления.

В отличие от публикаций, описания музейных объектов могут сильно различаться в разных музеях, и здесь невозможно обеспечить всеобъемлющий набор необходимых свойств. В связи с этим для данных объектов реализуется возможность определения дополнительных свойств в виде связей с двумя вспомогательными объектами: «**дополнительные свойства**» и «**значения дополнительных свойств**». При этом в интерфейсе администратора системы предоставляется возможность определять дополнительные свойства предмета, а в интерфейсах ввода и вывода данных создаются представления соответствующих полей. Введенные значения дополнительных полей выдаются в полных сведениях о предмете, но поиск по ним не производится. В целях унификации подобные наборы дополнительных свойств могут быть приданы не только музейным предметам, но и любым другим видам ресурсов.

Мультимедийные представления

Для обеспечения хранения цифровых представлений ресурсов и абстрагирования от методов хранения данных в СУЭБ LibMeta разработан дополнительный прикладной профиль «Расширенной поддержки хранения данных», в котором вводится ряд новых сущностей.

Сущность «**медиа-объект**» предназначена для описания медиа-объекта как единого целого, состоящего из частей данных с различной функциональной нагрузкой. «**Медиа-объект**» включает в себя единственное свойство:

- части.

Сущность «**часть медиа-объекта**» позволяет в пределах одного целого медиа-объекта, например публикации, иметь несколько частей с различной функциональной нагрузкой, таких как содержание, образы фотографий, текстов или страниц в виде изображений, тексты в чисто текстовом (распознаваемом) формате, отформатированный текст публикации и т. п. Свойствами «**части медиа-объекта**» являются:

- функциональный тип — показывает, какую функциональную нагрузку несет часть, например: «содержание», «страница книги»;
- потоки данных в формате, соответствующем типу данных;
- порядок в медиа-объекте;
- название части.

Сущность «единица хранения» представляет единый и неделимый поток двоичных данных. Позволяет абстрагироваться от конкретных методов хранения данных и собирать медиа-объекты, состоящие из частей, расположенных в разных местах и хранимых различными способами. Содержит следующие свойства:

- тип данных — формат представления данных, хранимых в данной части, например «Документ Microsoft Word» или «Изображение в формате JPEG»; связывается с MIME-типами посредством словаря ЕНИП *IMT* [8];
- тег — может содержать данные о кэшировании, преобразованном формате и т. д.

Введено три вида единиц хранения: ссылки на внешние источники, файлы в файловой системе и ВЛОВ-записи в базе данных.

Принцип использования представленного выше «медиа-объекта» в СУЭБ LibMeta несколько отличается от общепринятого в ЭБ. Для обеспечения цифровых представлений не только публикаций, но и музейных объектов, а также мультимедийных изображений коллекций, фотографий персон, коллективов, зданий организаций и т. п. в сущность «ресурс», являющуюся суперклассом для всех основных объектов онтологии, вводится свойство «медиа-представления». Таким образом, одно или несколько мультимедийных представлений может сопровождать любой ресурс информационной веб-системы.

Коллекции ресурсов

В базовых метаданных ЕНИП предусмотрена поддержка коллекций, однако требования ЭБ, в особенности с поддержкой хранения музейных предметов, не позволяют их полноценно использовать. В связи с этим базовый профиль ЕНИП дополнен коллекциями со следующими атрибутами:

- название;
- тип коллекции;
- ключевые слова;
- описание;

- администратор;
- количество элементов в коллекции;
- место хранения;
- примечание;
- элементы коллекции.

Коллекции такого рода позволяют хранить классические коллекции (архивные, музейные) и иметь любые вложенные наборы объектов (выставочные, выездные, по хранению и пр.).

4.3 Интеграция СУЭБ LibMeta с другими электронными библиотеками

Интеграция СУЭБ LibMeta с другими информационными системами осуществляется несколькими путями: во-первых, это интеграция с ЕНИП, во-вторых, это интеграция с системами подготовки публикаций и, в-третьих, интеграция с универсальными агрегаторами.

Интеграция с системами подготовки публикаций

Для обеспечения интеграции с системами подготовки публикаций в СУЭБ LibMeta существует компонент загрузки и обновления данных и метаданных из внешних источников. Данный компонент ставит в соответствие каждому загружаемому в библиотеку извне ресурсу «состояние размещения», в зависимости от которого с ресурсом производятся действия по размещению и его периодическому обновлению. Когда ресурс уже размещен, также возможно его обновление по запросу от обновляющей (размещающей) стороны.

Описанный компонент настроен для работы с системой подготовки публикаций, применяемой в ЭБ «Научное наследие России», однако ее архитектура рассчитана на простое добавление модулей для любых других систем подготовки публикаций. В плане развития СУЭБ LibMeta стоит разработка подсистемы подготовки публикаций начального уровня, которая, естественно, будет работать с СУЭБ LibMeta.

Интеграция с универсальными агрегаторами

Для интеграции с универсальными агрегаторами в СУЭБ LibMeta полностью реализованы стандарты OAI-PMH [22] и Dublin Core [11]. Кроме того, для интеграции с музейными системами существует отображение метаданных системы на концептуальную модель CIDOC-CRM [16]. Для использования

других протоколов и форматов обмена данными и метаданными, в особенности основанных на XML, не представляется сложной реализация модулей обмена данными.

5 Заключение

В настоящее время идет становление ЭБ как в России, так и в мире. До сих пор не существует всеобъемлющего стандарта на ЭБ, либо реализации ЭБ, удовлетворяющей всем существующим мировым стандартам. В данной статье рассмотрены основные требования, выдвигаемые к современным ЭБ. Также приведен обзор наиболее распространенных стандартов, на которые следует ориентироваться при разработке ЭБ, в частности в среде ЕНИП.

Создаваемая в ВЦ РАН СУЭБ LibMeta наиболее полно удовлетворяет приведенным требованиям и позволяет научным институтам РАН, имеющим собственные библиотечные, архивные или музейные фонды, создавать свои ЭБ, интегрированные с ЕНИП, и легко выставлять данные ресурсы в Интернет как для научного, так и для широкого круга пользователей. На текущий момент работы по СУЭБ LibMeta практически завершены.

Литература

1. *Gantz J., Chute C., Manfrediz A., et al.* Доклад IDC при финансовой поддержке компанией EMC: Обновленный прогноз роста мирового объема информации до 2011 г.
2. *Berners-Lee T., Hendler J., Lassila O.* The semantic Web // *Scientific Am.*, 2001. No. 5. P. 34–43.
3. *Berners-Lee T., Shadbolt N., Hall W.* The semantic Web revisited // *IEEE Intelligent Systems*, 2006. No. 6.
4. *Зацман И. М.* Концептуальный поиск и качество информации. — М.: Наука, 2003.
5. *Галева И. С.* Интернет как инструмент библиографического поиска. — М.: Профессия, 2007.
6. *Kahn R., Cerf V.* The digital library project. Vol. I: The world of knowbots (DRAFT): An open architecture for a digital library system and a plan for its development. — Reston, VA: Corporation for National Research Initiatives, 1988.
7. *Земсков А. И., Шрайберг Я. Л.* Электронная информация и электронные ресурсы: публикации и документы, фонды и библиотеки. — М.: ФАИР, 2007.
8. *Бездушный А. Н., Бездушный А. А., Серебряков В. А., Филиппов В. И.* Интеграция метаданных Единого научного информационного пространства РАН. — М.: ВЦ РАН, 2006.
9. Consultative committee for space data systems. <http://public.ccsds.org/default.aspx>.
10. DELOS an Association for Digital Libraries. <http://www.delos.info>.
11. Dublin core. <http://dublincore.org>.
12. CIDOC CRM Home page. <http://cidoc.ics.forth.gr/index.html>.
13. Publishing requirements for industry standard metadata. <http://prismstandard.org>.
14. CCSDS Secretariat, Program Integration Division (Code M-3). Reference model for an open archival information system (OAIS): Recommendation for space data system standards Blue book. — National Aeronautics and Space Administration, 2002.
15. *Candela L., Castelli D., Ferro N., et al.* The DELOS Digital Library Reference Model — foundations for digital libraries. Version 0.98. — GEIE ERCIM, 2008.
16. *Crofts N., Doerr M., Gill T., Stead S., Stiff M.* Definition of the CIDOC Conceptual Reference Model, 2010.
17. Functional requirements for bibliographic records. <http://www.ifla.org/VII/s13/frbr/frbr.htm>.
18. Functional requirements for bibliographic records object-oriented definition and mapping to FRBR_ER. http://cidoc.ics.forth.gr/docs/frbr_oo/frbr_docs/FRBR_oo_V0.9.pdf.
19. Network Development and MARC Standards Office MARC 21 concise formats. — Cataloging Distribution Service, Library of Congress, 2006.
20. Система форматов RUSMARC. <http://www.rba.ru/rusmarc>.
21. MARC 21 XML schema. <http://www.loc.gov/standards/marcxml>.
22. Open archives initiative protocol for metadata harvesting. <http://www.openarchives.org/pmh>.
23. Open archives initiative object reuse and exchange. <http://www.openarchives.org/ore>.
24. *Захаров А. А., Филиппов В. И.* Поддержка цифровых библиотек и музейных объектов в среде ЕНИП // Электронные библиотеки: перспективные методы и технологии, электронные коллекции: Тр. XI Всероссийской научной конф. RCDL'2009. — Петрозаводск: КарНЦ РАН, 2009. — 487 с.
25. Портал ЭБ «Научное наследие России». <http://e-heritage.ru>.
26. *Бездушный А. Н., Бездушный А. А., Нестеренко А. К. и др.* Информационная Web-система «Научный институт на платформе ЕНИП». — М.: ВЦ РАН, 2007.

ИНТЕГРАЦИЯ ГЕТЕРОГЕННОЙ ИНФОРМАЦИИ О ЦВЕТНЫХ ПИКСЕЛЯХ И ИХ ЦВЕТОВОСПРИЯТИИ

О. П. Архипов¹, З. П. Зыкова¹

Аннотация: Рассмотрена задача интеграции гетерогенной информации о цветных пикселях, их трансформации в пользовательской компьютерной системе и о стандартном и индивидуальном цветовосприятии пользователя в единую инфокоммуникационную среду. Решение задачи необходимо для создания специальных легко адаптируемых к компонентам системы программно-технических инструментов, обеспечивающих произвольным пользователям адекватное восприятие цветной информации, выводимой на периферийные устройства ПЭВМ. При решении задачи используются новые методы представления и анализа гетерогенной (количественной и качественной) информации о связанных цветовых пространствах, многокритериального выбора в части предсказания различия и неразличения цветной информации и классификации пикселей при структурировании цветовых пространств.

Ключевые слова: цветовое пространство; цветовое восприятие; пространство цветовосприятия; аномалия цветового зрения; частичная цветовая слепота; RGB-характеризация

1 Введение

Современные периферийные устройства делают доступной для любого пользователя ПЭВМ цветную информацию, зрительное восприятие которой является одним из основных средств коммуникации человека в окружающем мире.

Известно, что видимое цветовое пространство шире цветового пространства монитора (C_M), которое, в свою очередь, в общем, шире цветового пространства принтера (C_P). Поскольку цветовое пространство наблюдателя является непрерывным, а цветовые пространства периферийных устройств C_M и C_P дискретны, будет рассматриваться его дискретный аналог (совокупность пикселей), обозначаемый в дальнейшем C_H . Это позволит рассматривать цветовосприятие наблюдателя как цветопередачу — отображение цветовых пространств C_M и C_P в C_H . При этих условиях для описания цветовосприятия можно использовать функции цветопередачи, аргументами которых являются пиксели C_M или C_P , а значениями — пиксели C_H .

Цветовосприятие, или цветовое пространство наблюдателя, имеет индивидуальные особенности и отличается у разных людей. То же справедливо и по отношению к цветовым пространствам периферийных устройств и к функциям цветопередачи между связанными цветовыми пространствами.

Общеизвестно, что представление цветного изображения, состоящего из пикселей RGB-куба — исходного цветового пространства (C_H)

$$C_H = \{(R, G, B)\}, \quad 0 \leq R, G, B \leq 255,$$

на двух различных мониторах (или принтерах), имеет заметные различия по цветовому решению. Например, отдельные фрагменты представления изображения могут быть более (или менее) красными, синими и т. д.

Технологии вывода на периферийные устройства, основанные на оцифровке и согласовании цветовых пространств, ориентируются на восприятие цвета «средним стандартным колориметрическим наблюдателем» [1]. При цветовоспроизведении сохраняется корректное соотношение между цветами внутри каждого отдельного представления, а потому по структуре эти представления воспринимаются одинаково большинством наблюдателей (пользователей ПЭВМ). Это обеспечивают им адекватное восприятие цветной информации, включенной в оборот в инфокоммуникационных средах и выводимой на периферийные устройства компьютерных систем.

Если цветовосприятие наблюдателя близко к стандарту, то будем называть его стандартным наблюдателем. Наблюдателей, имеющих различные аномалии цветового зрения, вследствие чего их цветовосприятие отлично от стандарта, будем называть аномальными наблюдателями. В рамках данной работы учитываются такие аномалии цветового зрения, как частичная цветовая слепота.

Наблюдатели с частичной цветовой слепотой не различают некоторые цвета, различаемые стандартными наблюдателями, и поэтому воспринимают цветную информацию в искаженном виде.

¹Орловский филиал Института проблем информатики РАН, ofran@orel.ru

Классификация цветовосприятия таких наблюдателей основана на том, какие именно цвета они не различают [1, 2]:

- дейтеранопия — это цветовосприятие, при котором не различаются зеленый и красный цвета с нормальной функцией спектральной световой эффективности;
- протанопия — это цветовосприятие, при котором не различаются зеленый и красный цвета с ненормально низкой функцией спектральной световой эффективности на длинноволновом участке спектра;
- тританопия — это цветовосприятие, при котором не различаются желтые и синие цвета.

Медицинская диагностика упомянутых классических аномалий проводится с помощью специальных отпечатков, выполненных на полиграфическом оборудовании, — пороговых таблиц для исследования цветового зрения [3, 4]. В этой среде информация об аномалиях представлена в неформальном виде, поэтому не может быть использована при организации вывода на периферийные устройства ПЭВМ.

Из [5, 6] доступно цифровое описание некоторых типичных наблюдателей с аномалиями цветового зрения. Однако они противоречат друг другу и не позволяют учесть индивидуальные особенности цветовосприятия произвольного наблюдателя в конкретных условиях цветовоспроизведения.

Чтобы учесть интересы произвольного наблюдателя хотя бы в жизненно важных областях применения (например, навигация, ориентирование и т. д.), необходимо представить описание его цветовосприятия в виде некоторой функции F_0

$$y = F_0(x), \quad x \in C_n, \quad y \in C_n,$$

обладающей свойствами, соответствующими цветовосприятию наблюдателя.

Например, в случае стандартного наблюдателя каждому аргументу F_0 соответствует только один пиксель его цветового пространства (C'_n), т. е. если

$$\begin{aligned} x' \neq x'', \quad x', x'' \in C_n, \\ y' = F_0(x'), \quad y'' = F_0(x''), \quad y', y'' \in C'_n, \end{aligned}$$

то $y' \neq y''$.

В случае аномального наблюдателя существует, по крайней мере, два различных аргумента F_0

$$x' \neq x'', \quad x', x'' \in C_n,$$

которым соответствует один пиксель цветового пространства аномального наблюдателя (C''_n)

$$y' = y'', \quad y' = F_0(x'), \quad y'' = F_0(x''), \quad y', y'' \in C''_n.$$

Совокупности пикселей C_n , которым в цветовом пространстве наблюдателя соответствует только один цвет, называют зонами толерантности. Для удобства последующих рассуждений предположим, что зоны толерантности и их компоненты упорядочены в некоторые последовательности

$$\begin{aligned} \{G_{0,i}\}, \quad i = 1, 2, \dots, I_0, \quad G_{0,i} \subset C_n, \\ G_{0,i} = \{x_{0,i,j}\}, \quad i = 1, 2, \dots, I_0, \\ 1 \leq j \leq J_{0,i}, \quad x_{0,i,j} \in C_n. \end{aligned}$$

Для стандартного наблюдателя при идеальном цветовоспроизведении существуют только вырожденные (состоящие из одного компонента) зоны толерантности. Для аномального наблюдателя существуют не только вырожденные, но и невырожденные (состоящие хотя бы из двух компонентов) зоны толерантности.

В соответствии с определением одна линия уровня функции F_0 (геометрического места точек пространства аргументов, для которых значения функции одинаковы) совпадает с множеством пикселей одной зоны толерантности C_n . Заметим, что поскольку разным линиям уровня соответствуют разные значения функции, они не имеют общих точек, т. е. не могут пересекаться.

Таким образом, введенная функция F_0 описывает цветовосприятие, поскольку обладает его свойствами, но она не пригодна для его цифрового описания, поскольку не имеет цифровых значений. Для цифрового описания цветовосприятия предлагается использовать его RGB-характеризацию с помощью RGB-функций, аргументами которых являются RGB-пиксели исходного пространства C_n , а значениями — пиксели из RGB-цветовых пространств, характеризующих пространства периферийных устройств.

Для характеристики цветового пространства монитора будет использоваться исходное пространство C_n . Функция RGB-характеристики цветовосприятия представления пикселей на мониторе имеет вид:

$$y = F_1(x), \quad x, y \in C_n,$$

и обладает линиями уровня, совпадающими с зонами толерантности наблюдателя при восприятии им представлений на мониторе,

$$\begin{aligned} \{G_{1,i}\}, \quad i = 1, 2, \dots, I_1, \quad G_{1,i} \subset C_n, \\ G_{1,i} = \{x_{1,i,j}\}, \quad i = 1, 2, \dots, I_1, \\ 1 \leq j \leq J_{1,i}, \quad x_{1,i,j} \in C_n. \end{aligned}$$

Следовательно, если представления двух разных пикселей

$$x' \neq x'', \quad x', x'' \in C_{\text{и}},$$

на мониторе воспринимаются наблюдателем как одинаковые, то значения функции F_1 от них совпадают:

$$y' = y'', \quad y' = F_1(x'), \quad y'' = F_1(x''), \quad y', y'' \in C_{\text{и}}.$$

Если представления двух разных пикселей

$$x' \neq x'', \quad x', x'' \in C_{\text{и}},$$

на мониторе воспринимаются наблюдателем как различные, то значения функции F_1 от них различны:

$$y' \neq y'', \quad y' = F_1(x'), \quad y'' = F_1(x''), \quad y', y'' \in C_{\text{и}}.$$

Для характеристики цветового пространства принтера будет использоваться цветовое пространство сканера (C_c) [7–9]. Функция RGB-характеристики цветовосприятия отпечатков пикселей:

$$y = F_2(x), \quad x \in C_{\text{и}}, \quad y \in C_c,$$

должна обладать линиями уровня, совпадающими с зонами толерантности наблюдателя при восприятии им отпечатков,

$$\begin{aligned} &\{G_{2,i}\}, \quad i = 1, 2, \dots, I_2, \quad G_{2,i} \subset C_{\text{и}}, \\ &G_{2,i} = \{x_{2,i,j}\}, \quad i = 1, 2, \dots, I_2, \\ &1 \leq j \leq J_{2,i}, \quad x_{2,i,j} \in C_{\text{и}}. \end{aligned}$$

Следовательно, если отпечатки двух разных пикселей

$$x' \neq x'', \quad x', x'' \in C_{\text{и}},$$

воспринимаются наблюдателем как одинаковые, то значения функции F_2 от них совпадают:

$$y' = y'', \quad y' = F_2(x'), \quad y'' = F_2(x''), \quad y', y'' \in C_c.$$

Если отпечатки двух разных пикселей

$$x' \neq x'', \quad x', x'' \in C_{\text{и}},$$

воспринимаются наблюдателем как различные, то значения функции F_2 от них различны:

$$y' \neq y'', \quad y' = F_2(x'), \quad y'' = F_2(x''), \quad y', y'' \in C_c.$$

Множество значений таких функций RGB-характеристики цветовосприятия произвольного наблюдателя является подмножеством соответствующего RGB-пространства и может быть использовано в качестве цифрового описания $C_{\text{и}}$, а сама

функция — для цифрового описания цветопередачи из $C_{\text{и}}$ в $C_{\text{н}}$.

Факт принадлежности произвольных пикселей одной или разным зонам толерантности тестируемого наблюдателя может быть установлен при тестировании. Общепринятым подходом к тестированию цветовосприятия произвольного наблюдателя является подход, при котором наблюдателю последовательно предъявляются для визуализации и описания представления тестовых изображений на периферийном устройстве (мониторе, принтере).

Если наблюдатель адекватно воспринимает структуру тестов, то делается вывод о различении им цветных пикселей, использованных при создании изображения. В этом случае можно говорить о том, что им соответствуют различные пиксели в цветовом пространстве тестируемого наблюдателя.

Если структура теста искажается при восприятии его наблюдателем, то, следовательно, он не различает соответствующие цветные пиксели и их можно считать представителями одной зоны толерантности.

Из-за технологических особенностей вывода цветных изображений, проявляющихся на практике в нестабильности цветовоспроизведения, представления пикселей с разными координатами на периферийных устройствах могут иметь идентичные цвета. В этом случае они являются представителями «паразитных» зон толерантности. Далее предполагается, что применяются такие периферийные устройства, для которых этот эффект проявляется только при выводе пикселей с достаточно близкими RGB-координатами.

В то же время из-за нестабильного цветовоспроизведения на периферийном устройстве (обычно принтере) пиксель с одними и теми же RGB-координатами может иметь объективно различные по цвету представления. Далее предполагается, что рассматриваются периферийные устройства со стабильным цветовоспроизведением, при котором имеющиеся различия визуально незаметны для произвольных наблюдателей.

Существование отличных от «паразитных» невырожденных зон толерантности обнаруживает наличие аномалий цветового зрения, а их реквизиты характеризуют основные свойства цветового пространства аномального наблюдателя и его отличие от цветового пространства стандартного наблюдателя.

Границы зон толерантности, возникающих из-за аномалий цветного зрения, размываются из-за паразитных зон толерантности и не могут быть определены точно. Чем выше стабильность вывода на периферийные устройства, тем меньше на прак-

тике может быть погрешность определения указанных границ.

Предположим, что протестировано восприятие представлений на периферийном устройстве каждой пары пикселей $C_{\text{п}}$. Для цифрового описания результатов тестирования достаточно использовать функцию от двух переменных (пикселей $C_{\text{п}}$), принимающую только два значения. Одно из значений используется, если при тестировании установлено, что цвета соответствующей пары пикселей наблюдатель различает, а второе — не различает.

Такую функцию после тестирования можно задать и хранить на диске в виде подходящей таблицы. При автоматической обработке из табличных данных можно, как это сделано, например, в [7], извлечь описание зон толерантности. Как будет показано далее, на этой основе можно построить функцию RGB-характеризации цветовосприятия протестированного пользователя и использовать ее для создания наглядных графических иллюстраций и управления цветопередачей в компьютерной системе в интересах протестированного наблюдателя.

Однако понятно, что на практике такая процедура тестирования не может быть реализована. В разумные временные сроки невозможно протестировать восприятие представлений на мониторе каждой пары из более чем шестнадцати с половиной миллионов пикселей $C_{\text{п}}$, а на принтере препятствием являются не только чрезмерные временные затраты, но и соответствующий объем расходных материалов (бумаги, красителей).

Выход состоит в проведении тестирования восприятия представления на мониторе ограниченного числа пикселей и приближенного вычисления функции RGB-характеризации на этой основе. Если затем построить функцию RGB-характеризации цветопередачи $C_{\text{и}}$ в $C_{\text{п}}$, то функция RGB-характеризации восприятия представления на мониторе может быть использована для аппроксимации функции RGB-характеризации восприятия отпечатков.

2 Постановка задачи

Пусть имеются множества пикселей

$$\{t\} \subset \{t'\} \subset \{t''\} \subset C_{\text{п}},$$

а после тестирования восприятия наблюдателем представлений на мониторе пиксели из $\{t\}$ классифицированы по группам представителей различных зон толерантности

$$\{g'_{1,i}\}, \quad i = 1, 2, \dots, I'_1, \quad g'_{1,i} \subset \{t\},$$

$$\begin{aligned} g'_{1,i} &= \{x_{1,i,j}\}, \quad i = 1, 2, \dots, I'_1, \\ &1 \leq j \leq j'_{1,i}, \quad x_{1,i,j} \in \{t\}, \\ \{t\} &= \bigcup_{i=1}^{I'_1} g'_{1,i} = \bigcup_{i=1}^{I'_1} \bigcup_{j=1}^{j'_{1,i}} x_{1,i,j}. \end{aligned}$$

Пусть далее каждая i -я группа дополнена пикселями из множества $\{t'\}$, не принадлежащими $\{t\}$, если тестируемый наблюдатель не различает их по цвету с ее первым представителем $x_{1,i,1}$. Полученные в результате группы представителей зон толерантности обозначим следующим образом:

$$\left. \begin{aligned} &\{G'_{1,i}\}, \quad i = 1, 2, \dots, I'_1, \quad G'_{1,i} \subset \{t'\}, \\ &G'_{1,i} = \{x_{1,i,j}\}, \quad i = 1, 2, \dots, I'_1, \\ &1 \leq j \leq J'_{1,i}, \quad x_{1,i,j} \in \{t'\}, \\ &\{t'\} \supset \bigcup_{i=1}^{I'_1} G'_{1,i} = \bigcup_{i=1}^{I'_1} \bigcup_{j=1}^{J'_{1,i}} x_{1,i,j}. \end{aligned} \right\} \quad (1)$$

Требуется определить:

- значения функции F_1 от аргументов $\{x_{1,i,j}\}$ из (1) и построить на этой основе аппроксимацию Ψ_1 функции F_1 на $\{t''\}$

$$\Psi_1(t'' = y'' \approx F_1(t''), \quad \{y''\} \subset C_{\text{п}},$$

линии уровня которой на множестве $\{t'\}$, определяемые константами $F_1(t)$,

$$\{G''_{1,i}\}, \quad i = 1, 2, \dots, I''_1, \quad G''_{1,i} \subset C_{\text{п}},$$

удовлетворяют следующим соотношениям:

$$G'_{1,i} \subset G''_{1,i}, \quad i = 1, 2, \dots, I'_1 = I''_1; \quad (2)$$

- значения аппроксимации Ψ_3 функции F_3 , являющейся функцией RGB-характеризации цветопередачи $C_{\text{и}}$ в $C_{\text{п}}$, от аргументов $\{t''\}$ как значения цветовой характеристики [7–9] отпечатков пикселей

$$\Psi_3(t'' = s'' \approx F_3(t''), \quad \{s''\} \subset C_{\text{с}};$$

- аппроксимацию Ψ_2 функции F_2 как суперпозицию функций Ψ_1 и Ψ_3 и вычислить ее значения от аргументов $\{t''\}$ как значения функции Ψ_1 от аргументов $\{s''\}$:

$$\Psi_2(t'') = \Psi_1(s'') = \Psi_1(\Psi_3(t'')).$$

3 Аппроксимация функции RGB-характеристики восприятия представления пикселей на мониторе

Пусть при тестировании обнаружены группы представителей некоторых зон толерантности (1). Тогда, выбирая один из пикселей каждой группы в качестве реперного пикселя, можно определять им значение функции RGB-характеристики F_1 для всех пикселей данной группы.

При выборе реперных пикселей в соответствии с [1] следует руководствоваться правилом: реперным является пиксель, который наиболее близок к серой шкале RGB-куба (прямой, проходящей через вершины куба $(0, 0, 0)$ и $(255, 255, 255)$).

Обозначим μ функцию, реализующую это правило:

$$F_1(x_{1,i,j}) = \mu(G'_{1,i}) = x_{1,i,j''}, \quad i = 1, 2, \dots, I'_1, \quad 1 \leq j, j'' \leq J'_{1,i}, \quad (3)$$

где значение индекса j'' удовлетворяет соотношению:

$$\rho(x_{1,i,j''}, x'_{1,i,j''}) = \min_{1 \leq j' \leq J'_{1,i}} \rho(x_{1,i,j}, x'_{1,i,j'}),$$

а $x'_{1,i,j}$ — это ближайшая для $x_{1,i,j}$ точка серой шкалы.

Заметим, что если пиксель имеет RGB-координаты (r, g, b) , то ближайшая к нему точка серой шкалы лежит на пересечении серой шкалы и перпендикуляра, опущенного из точки (r, g, b) на серую шкалу, и имеет координаты

$$\left(\frac{r+g+b}{3}, \frac{r+g+b}{3}, \frac{r+g+b}{3} \right).$$

Если вершины RGB-куба входят в совокупность пикселей

$$\{x_{1,i,j}\}, \quad i = 1, 2, \dots, I'_1, \quad 1 \leq j \leq J'_{1,i},$$

то их можно использовать в качестве узлов интерполяции для построения аппроксимации Ψ_1 функции F_1 на $\{t''\}$, поскольку в соответствии с (3) в них будут определены значения функции F_1 .

Действительно, в этом случае для определения значения Ψ_1 в произвольной точке $x_0 = (r_0, g_0, b_0)$ из $\{t''\}$, лежащей на некотором ребре куба, достаточно:

- на этом же ребре найти такие точки $x_1 = (r_1, g_1, b_1)$ и $x_2 = (r_2, g_2, b_2)$, в которых значения Ψ_1 уже известны, а точка x_0 является внутренней точкой отрезка $[x_1, x_2]$;

- значения $\Psi_1(x_0)$ определить, например, при линейной интерполяции по формуле

$$\Psi_1(x_0) = \frac{\Psi_1(x_1)/\rho(x_0, x_1) + \Psi_1(x_2)/\rho(x_0, x_2)}{1/\rho(x_0, x_1) + 1/\rho(x_0, x_2)}.$$

Для определения значения Ψ_1 в произвольной точке $x_0 = (r_0, g_0, b_0)$ из $\{t''\}$, лежащей внутри некоторой стороны куба, достаточно:

- на этой же стороне куба найти такие точки $x_1 = (r_1, g_1, b_1)$, $x_2 = (r_2, g_2, b_2)$, $x_3 = (r_3, g_3, b_3)$, $x_4 = (r_4, g_4, b_4)$, которые лежат на пересечении соответствующих ребер куба и опущенных на них перпендикуляров из точки x_0 ;
- вычислить значения $\Psi_1(x_1)$, $\Psi_1(x_2)$, $\Psi_1(x_3)$, $\Psi_1(x_4)$ по ранее описанному алгоритму;
- определить значения $\Psi_1(x_0)$, например, при интерполяции по формуле:

$$\Psi_1(x_0) = \left(\frac{\Psi_1(x_1)}{\rho(x_0, x_1)} + \frac{\Psi_1(x_2)}{\rho(x_0, x_2)} + \frac{\Psi_1(x_3)}{\rho(x_0, x_3)} + \frac{\Psi_1(x_4)}{\rho(x_0, x_4)} \right) / \left(\frac{1}{\rho(x_0, x_1)} + \frac{1}{\rho(x_0, x_2)} + \frac{1}{\rho(x_0, x_3)} + \frac{1}{\rho(x_0, x_4)} \right).$$

Наконец, для определения значения Ψ_1 в произвольной внутренней точке куба $x_0 = (r_0, g_0, b_0)$ из $\{t''\}$ достаточно:

- найти точки, которые лежат на пересечении сторон куба и опущенных на них перпендикуляров из точки x_0 ;
- вычислить в них значения Ψ_1 по ранее описанному алгоритму;
- определить значения $\Psi_1(x_0)$, например, при интерполяции по формуле, аналогичной тем, которые были использованы ранее.

Очевидно, что вычисление приближенных значений функции RGB-характеристики является трудоемкой задачей. Желательно один раз вычислить их и хранить в электронном виде (например, в табличной форме) для последующих многократных использований. Однако в RGB-кубе более шестнадцати с половиной миллионов цветов, что потребует использования таблиц соответствующего размера.

Практика показывает, что удобен компромиссный вариант, когда часть значений функции цветопередачи на некотором подмножестве $\{t''\}$ RGB-куба хранится, а часть — по мере необходимости вычисляется.

Для линий уровня построенной функции условие (2), очевидно, выполняется, причем на практике нельзя добиться точного соответствия между линиями уровня у аппроксимации и линиями уровня у искомой функции. Это обусловлено возможностью появления «паразитных» зон толерантности из-за погрешностей приближенных вычислений, возникающих при интерполяции.

4 Аппроксимация функции RGB-характеристики цветопередачи $C_{и}$ в $C_{п}$

Процедура получения значений цветовой характеристики отпечатков пикселей

$$s'' = \Psi_3(t''), \{s''\} \subset C_c$$

разработана авторами в более ранних работах (см., например, [7–9]), поэтому приведем лишь краткое ее описание:

- на основе пикселей $\{t''\}$ создается специальное изображение — машиночитаемая зона;
- машиночитаемая зона печатается и сканируется по определенным правилам;
- из общего скана извлекается полезная часть, содержащая сканы меток позиционирования и сканы растровых точек исходных пикселей;
- при автоматической обработке полезной части скана определяется местоположение меток позиционирования, а затем на этой основе и местоположение скана каждой растровой точки, и координаты составляющих его пикселей;
- при автоматической обработке координаты пикселей, составляющих скан каждой растровой точки, усредняются, а полученная величина принимается за значение цветовой характеристики отпечатка соответствующего пикселя.

5 Аппроксимация функции RGB-характеристики восприятия отпечатков пикселей

Пусть от аргументов $\{t''\} \subset C_{и}$ известны значения функций

$$\begin{aligned} y'' &= \Psi_1(t''), \{y''\} \subset C_{и}, \\ s'' &= \Psi_3(t''), \{s''\} \subset C_c. \end{aligned}$$

Обозначим $\{z''\} \subset C_{и}$ последовательность RGB-пикселей, имеющих те же RGB-координаты, что и RGB-пиксели $\{s''\} \subset C_c$. Тогда имеем

$$\begin{aligned} v'' &= \Psi_1(z''), \{v''\} \subset C_{и}, \{z''\} \subset C_{и}, \\ s'' &= \Psi_3(t''), \{s''\} \subset C_c. \end{aligned}$$

Обозначим $\{w''\} \subset C_c$ последовательность RGB-пикселей, имеющих те же RGB-координаты, что и RGB-пиксели $\{v''\} \subset C_{и}$. Определим значения функции Ψ_2 от аргументов $\{t''\}$:

$$w'' = \Psi_2(t''), \{t''\} \subset C_{и}, \{w''\} \subset C_c.$$

Если найденные значения хранить на диске, то приближенные значения функции RGB-характеристики восприятия отпечатков пикселей от других значений аргументов, не принадлежащих множеству $\{t''\}$, можно вычислять по мере необходимости путем интерполяции аналогично тому, как это описано ранее.

6 Интегрированные данные о цветовосприятии

В качестве $\{t''\}$ рассмотрим, например, пиксели из множества M

$$\begin{aligned} M &= \{(r_i, g_i, b_i)\} = \{(j \cdot 17, k \cdot 17, n \cdot 17)\}, \\ & \quad j, k, n = 0, \dots, 15, \\ i &= j + k \cdot 16 + n \cdot 16 \cdot 16 = 0, 1, \dots, 4095. \end{aligned}$$

Предположим, что значения функций Ψ_1 и Ψ_2 на M определены и хранятся в электронном виде. Очевидно, что значения функций Ψ_1 и Ψ_2 во всех других пикселях RGB-куба, которые, как можно заметить, принадлежат одному из меньших RGB-кубиков размером $17 \times 17 \times 17$, могут быть вычислены при интерполяции по вершинам кубика, в которых значения функции известны, поскольку эти пиксели принадлежат множеству M .

Если из пикселей множества M и пикселей, являющихся значениями функций Ψ_1 и Ψ_2 построить изображения (обозначим их Img , Img' и Img'' соответственно), то будет получена наглядная иллюстрация, показывающая различие и совпадение цветовосприятия стандартного и протестированного наблюдателя.

Сопоставление трех изображений (Img , Img' и Img'') позволяет получить полное представление о цветовосприятии произвольным наблюдателем вывода на периферийные устройства его компьютерной системы. К тому же электронное

описание цветовосприятия произвольного наблюдателя в виде изображений в BMP-файлах более удобно (по сравнению с табличными данными в текстовых файлах) для интеграции гетерогенной информации о цветных пикселях и цветовосприятии и упрощает проведение визуального и программного анализа индивидуального цветовосприятия.

Пусть выводимая на периферийное устройство цветная информация представляет собой произвольное RGB-изображение. Тогда можно модифицировать его, заменяя пиксели RGB-изображения значениями в них функции Ψ_1 или Ψ_2 , которые либо хранятся на диске, либо вычисляются, если исходный пиксель не принадлежит множеству $\{t''\}$. Сравнивая между собой представления исходного изображения и его модификации, можно более детально оценить степень искажения информации при восприятии произвольным наблюдателем. Если из-за какой-либо аномалии цветового зрения искажения недопустимо велики, необходимо изменить подход к ее графическому представлению.

Если модифицировать RGB-изображение в соответствии с [10], заменяя пиксели RGB-изображения наиболее близкими к ним пикселями из области значений аппроксимирующей функции,

то в результате будет получено RGB-изображение, восприятие которого наблюдателем максимально близко к восприятию исходного изображения стандартным наблюдателем.

7 Пример построения функции Ψ_1

Известно ПО [5, 6], с помощью которого можно получить цифровое описание цветовых пространств таких аномальных наблюдателей, как дейтеранопы, протанопы и тританопы в форме изображений Img и Img' .

Использование файла с Img в качестве входных данных одной из программ [5, 6] позволяет получить на выходе изображение Img' , соответствующее одной из аномалий цветового зрения. Пиксели Img' составляют цифровое описание цветового пространства аномального наблюдателя, а соотношение между ними характеризуют индивидуальные особенности его цветовосприятия.

Обозначим:

φ_1 — функцию RGB-характеризации дейтеранопы [5] и $M'_1: M'_1 = \varphi_1(M)$;

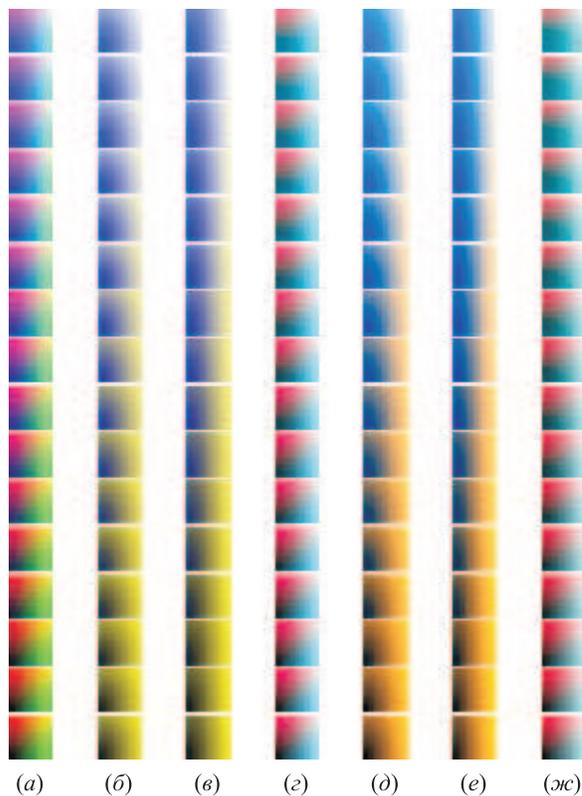


Рис. 1 Вид изображений: (а) M ; (б) M'_1 ; (в) M'_2 ; (г) M'_3 ; (д) M'_4 ; (е) M'_5 ; (ж) M'_6

φ_2 — функцию RGB-характеризации протанопа [5] и $M'_2: M'_2 = \varphi_2(M)$;

φ_3 — функцию RGB-характеризации тританопа [5] и $M'_3: M'_3 = \varphi_3(M)$;

φ_4 — функцию RGB-характеризации дейтеранопа [6] и $M'_4: M'_4 = \varphi_4(M)$;

φ_5 — функцию RGB-характеризации протанопа [6] и $M'_5: M'_5 = \varphi_5(M)$;

φ_6 — функцию RGB-характеризации тританопа [6] и $M'_6: M'_6 = \varphi_6(M)$.

Изображение множеств M и $M'_i, i = 1, 2, \dots, 6$, приведено на рис. 1. Заметим, что эти иллюстрации наглядно демонстрируют отсутствие единого подхода к оцифровке цветовых пространств аномальных наблюдателей в общем случае. Даже для указанных классических аномалий цветного зрения две указанные программы дают различные результаты при одних и тех же входных данных.

Для реализации алгоритма построения функции Ψ_1 было создано специальное ПО, функционирующее на базе одного компьютера типа PC IBM с оболочкой Windows XP. Использовалась периферия: цветной лазерный принтер HP Color LaserJet 4700n и цветной сканер FUJITSU fi-60F.

Практическое тестирование было заменено виртуальным. Рассмотрено шесть примеров в соответствии с [5, 6]. При этом неразличимость произвольных пикселей x' и x'' для i -го виртуального

наблюдателя определялась в соответствии с [11] на основе исследования расстояния между пикселями $\varphi_i(x')$ и $\varphi_i(x''), i = 1, 2, \dots, 6$. Чтобы отличать значения функции Ψ_1 для разных виртуальных наблюдателей, будем использовать второй индекс: значение $\Psi_{1,i}$ функции Ψ_1 соответствует i -му виртуальному наблюдателю.

При расчетах использовались множества:

$$\begin{aligned} \{t\} &= \{(r_i, g_i, b_i)\} = \{(j \cdot 85, k \cdot 85, n \cdot 85)\}, \\ & j, k, n = 0, 1, 2, 3, \\ & i = j + k \cdot 4 + n \cdot 4 \cdot 4 = 0, 1, \dots, 63; \\ \{t'\} &= \{t''\} = \{(r_i, g_i, b_i)\} = \{(j \cdot 5, k \cdot 5, n \cdot 5)\}, \\ & j, k, n = 0, \dots, 51, \\ & i = j + k \cdot 52 + n \cdot 52 \cdot 52 = 0, 1, \dots, 140\,607. \end{aligned}$$

Иллюстрации к полученным результатам построены на основе изображения *Image*, составленного из пикселей $\{t''\}$, принадлежащих сторонам RGB-куба в соответствии с рис. 2. Результаты расчетов отображены на рис. 3–9.

8 Заключение

Рассмотрена задача интеграции гетерогенной информации о цветных пикселях, их трансформации в пользовательской компьютерной системе и о стандартном и индивидуальном цветовосприятии пользователя в единую инфокоммуникационную среду.

Задача решена на основе RGB-характеризации цветовосприятия произвольным пользователем ПЭВМ вывода на периферийные устройства, что позволяет оцифровать его цветовое пространство и цветовосприятия, представить оцифрованные результаты в виде, удобном для организации управления цветопередачей в пользовательской компьютерной системе в интересах наблюдателя, имеющего такую аномалию цветного зрения, как частичная цветовая слепота.

Сравнение применяемого метода с известными подходами к характеристике цветовосприятия, например [5, 6], показывает, что предлагаемая модель позволяет устранить противоречие в характеристике цветовосприятия, поскольку обнаруживает различие в индивидуальных особенностях цветовосприятия, лежащих в их основе. Расхождение между моделями [5, 6] обусловлено различными гипотезами о цветовосприятии дейтеранопов, протанопов и тританопов, в частности к различию в определении зон толерантности одних и тех же типичных наблюдателей с аномалией цветного зрения. Это еще раз подчеркивает необходимость тестирования произвольного наблюдателя для точного определения и дальнейшего корректного уче-



Рис. 2 Вид изображений: (а) Image; (б) Image $\cap \{t\}$

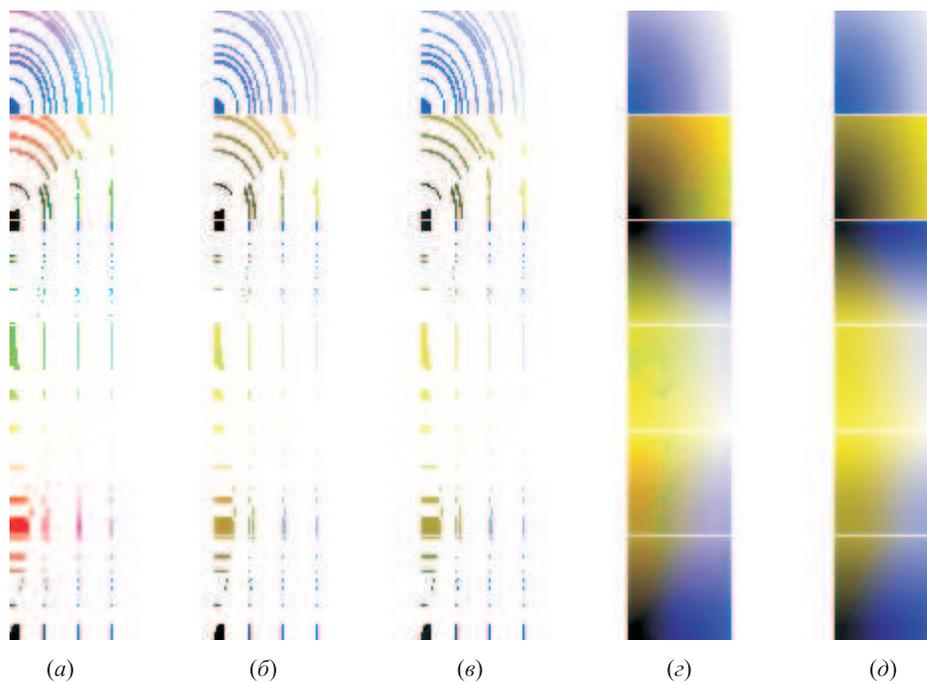


Рис. 3 Вид изображений: (а) линии уровня φ_1 и $\Psi_{1,1}$ на Image; (б) значения $\Psi_{1,1}$ на линиях уровня; (в) значения φ_1 на линиях уровня; (г) $\Psi_{1,1}$ (Image); (д) φ_1 (Image)

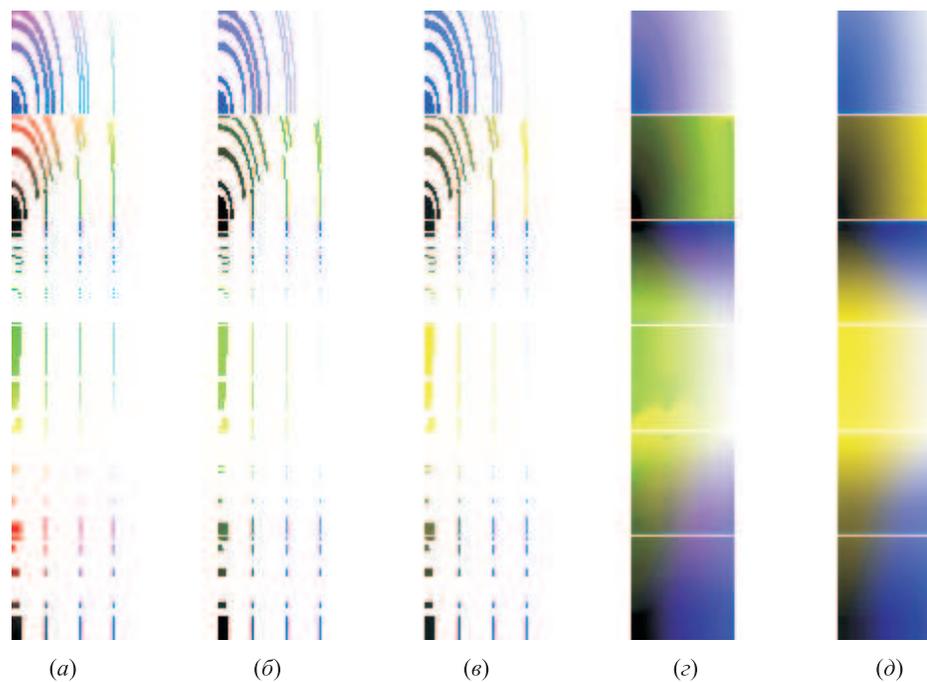


Рис. 4 Вид изображений: (а) линии уровня φ_2 и $\Psi_{1,2}$ на Image; (б) значения $\Psi_{1,2}$ на линиях уровня; (в) значения φ_2 на линиях уровня; (г) $\Psi_{1,2}$ (Image); (д) φ_2 (Image)

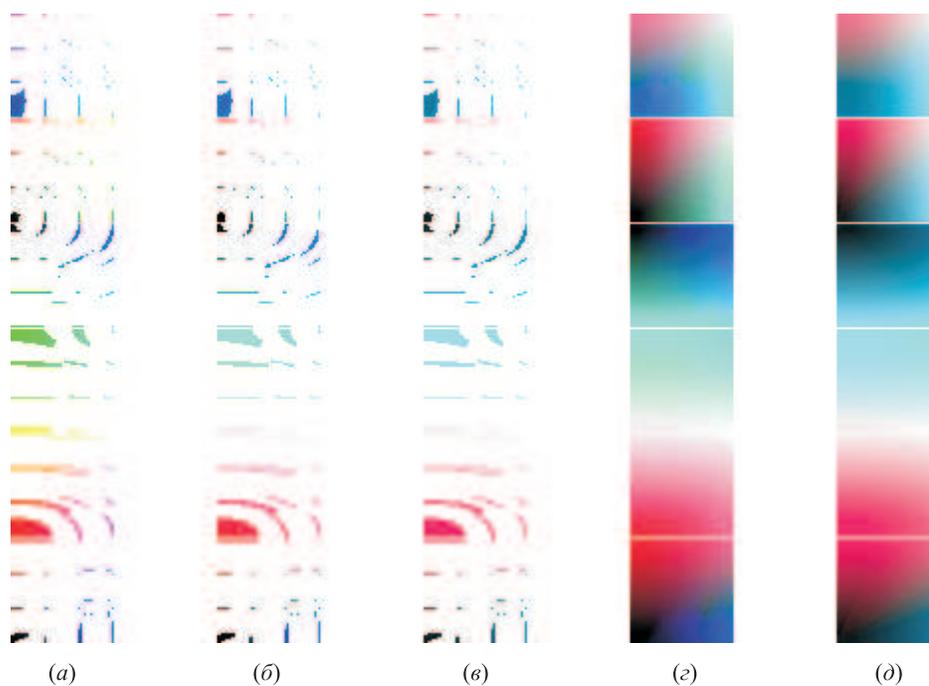


Рис. 5 Вид изображений: (а) линии уровня φ_3 и $\Psi_{1,3}$ на Image; (б) значения $\Psi_{1,3}$ на линиях уровня; (в) значения φ_3 на линиях уровня; (г) $\Psi_{1,3}(\text{Image})$; (д) $\varphi_3(\text{Image})$

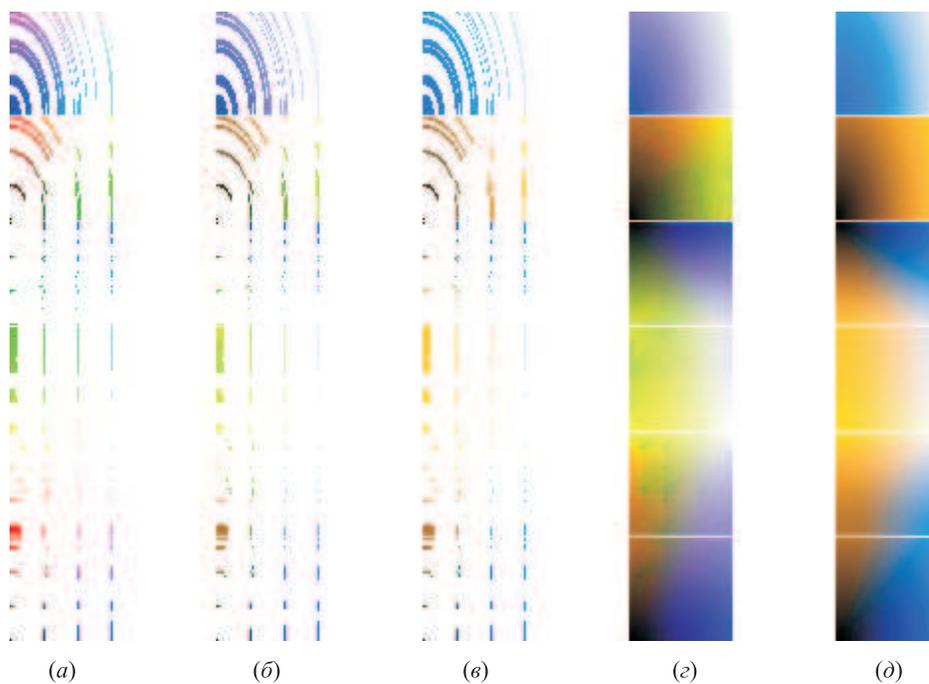


Рис. 6 Вид изображений: (а) линии уровня φ_4 и $\Psi_{1,4}$ на Image; (б) значения $\Psi_{1,4}$ на линиях уровня; (в) значения φ_4 на линиях уровня; (г) $\Psi_{1,4}(\text{Image})$; (д) $\varphi_4(\text{Image})$

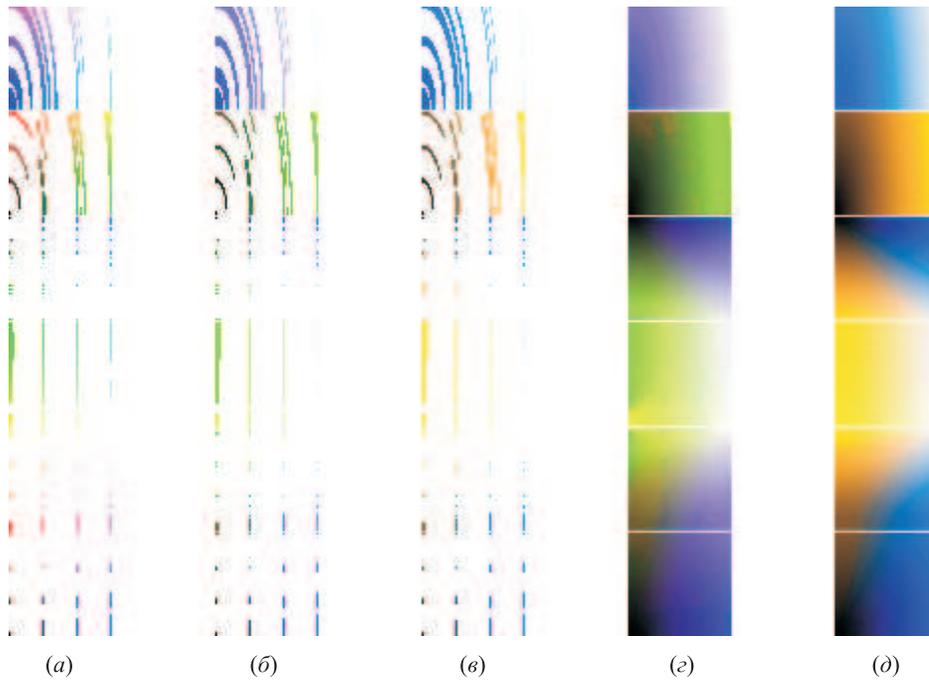


Рис. 7 Вид изображений: (а) линии уровня φ_5 и $\Psi_{1,5}$ на Image; (б) значения $\Psi_{1,5}$ на линиях уровня; (в) значения φ_5 на линиях уровня; (г) $\Psi_{1,5}$ (Image); (д) φ_5 (Image)

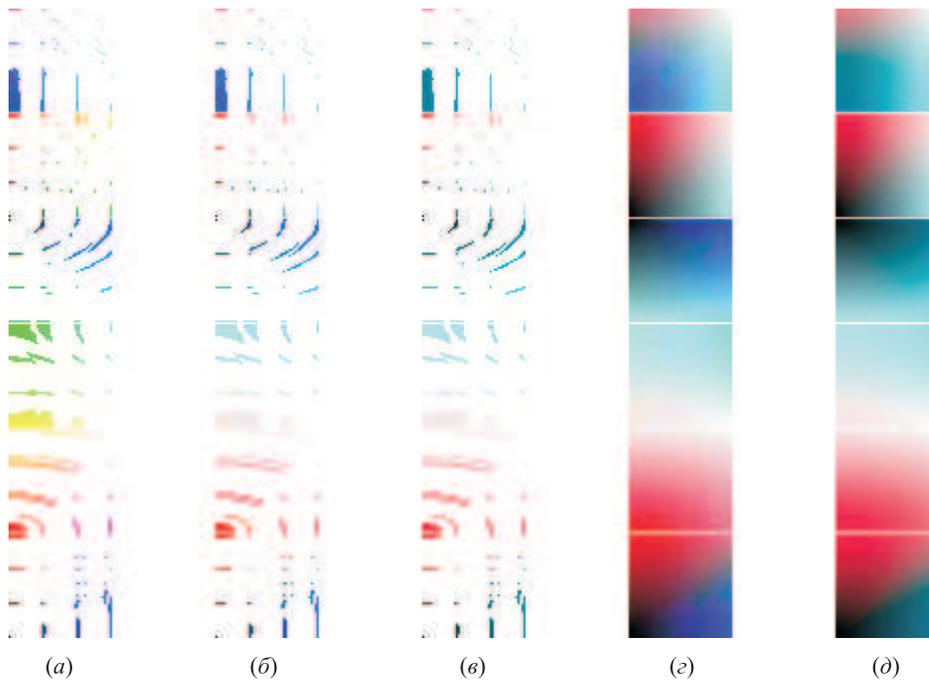


Рис. 8 Вид изображений: (а) линии уровня φ_6 и $\Psi_{1,6}$ на Image; (б) значения $\Psi_{1,6}$ на линиях уровня; (в) значения φ_6 на линиях уровня; (г) $\Psi_{1,6}$ (Image); (д) φ_6 (Image)

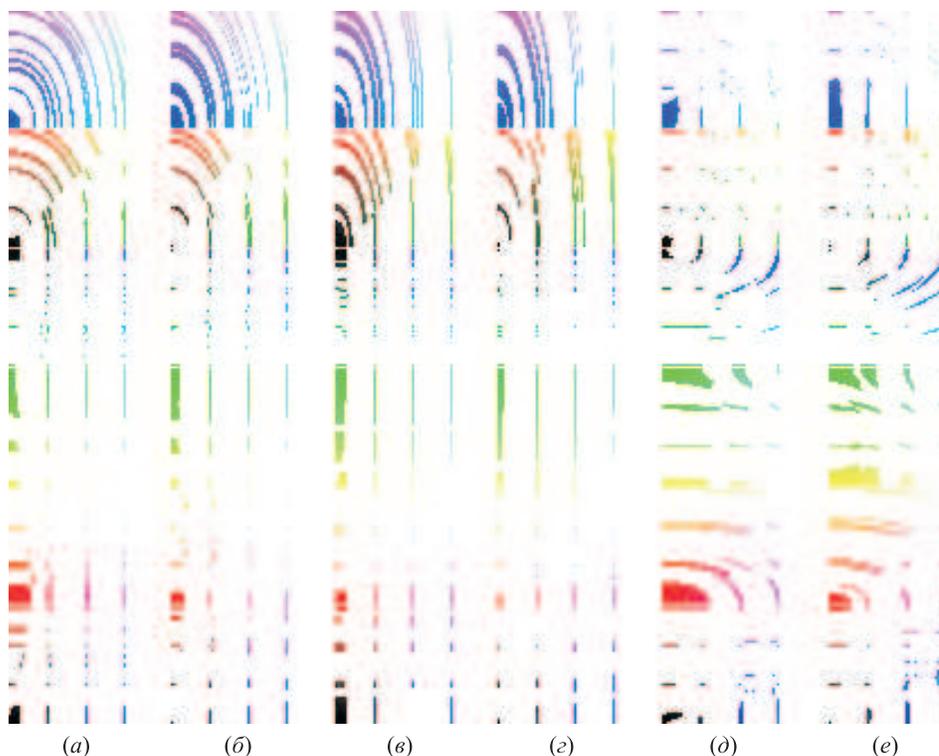


Рис. 9 Вид изображений: (а) линии уровня φ_1 ; (б) линии уровня φ_4 ; (в) линии уровня φ_2 ; (г) линии уровня φ_5 ; (д) линии уровня φ_3 ; (е) линии уровня φ_6

та индивидуальных особенностей его цветовосприятия.

Результаты работы имеют важное практическое значение, поскольку на основе интеграции цифрового описания цветовосприятия произвольного наблюдателя с гетерогенной информацией:

- о цветовых пространствах периферийных устройств;
- о цветопередаче (отображении цветовых пространств друг на друга) в пользовательской компьютерной системе;
- о стандартном цветовосприятии

они применимы для создания специальных легко адаптируемых к компонентам системы программно-технических инструментов, обеспечивающих адекватное восприятие произвольными пользователями вывода на периферийные устройства ПЭВМ.

Литература

1. Джадд Д., Вышецки Г. Цвет в науке и технике. — М.: Мир, 1978.
2. Шашлов Б.А. Цвет и цветовоспроизведение. — М.: Мир, 1986.
3. Юстова Е. Н., Алексеева К. А., Волоков В. В. и др. Пороговые таблицы для исследования цветового зрения: Методическое руководство. — М., 2000.
4. Пороговые таблицы для исследования цветового зрения: Паспорт. Комплект № 4181. — М: Вида, 2008.
5. Jenny B, Vaughn Kelso N. Color Oracle: Design for the color impaired. <http://colororacle.cartography.ch>.
6. Vischeck. <http://www.vischeck.com>.
7. Архипов О. П., Зыкова З. П. Допечатное тестирование индивидуального зрительного восприятия // Вестник компьютерных и информационных технологий, 2008. № 12. С. 2–8.
8. Архипов О. П., Бородина Л. Н., Зыков Р. В. и др. Технология оцифровки цветовосприятия отпечатков. — М.: ИПИ РАН, 2009.
9. Архипов О. П., Бородина Л. Н., Зыков Р. В. и др. Инструментальная оценка цветов отпечатков. — М.: ИПИ РАН, 2009.
10. Соколов И. А., Архипов О. П., Захаров В. Н., Зыкова З. П., Архипов П. О. Способ компьютерного распознавания и визуального воспроизведения цветных изображений. Пат. 2005130683, приоритет 04.10.05. Бюл. № 8 от 20.03.07.
11. Городецкий В. И., Самойлов В. В. Стеганография на основе цифровых изображений // Информационные технологии и вычислительные системы, 2001. № 2/3. С. 51–64.

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ И МЕТОД ВОССТАНОВЛЕНИЯ ПОЗЫ ЧЕЛОВЕКА ПО СТЕРЕОПАРЕ СИЛУЭТНЫХ ИЗОБРАЖЕНИЙ*

А. К. Цискаридзе¹

Аннотация: Работа посвящена проблеме восстановления позы локально симметричного объекта по стереопаре силуэтов в случае отсутствия окклюзии. Рассматриваются две модели фигуры. В первом случае фигура описывается как объединение пространственных жирных линий и предлагается метод ее восстановления. Во втором случае фигура описывается в виде шарнирной модели и предлагается метод ее подгонки по стереопаре силуэтов. Методы основаны на построении непрерывных скелетов силуэтов.

Ключевые слова: стереореконструкция; скелет; цилиндрические объекты; шарнирная модель; серединные оси

1 Введение

Задача восстановления формы пространственного объекта по нескольким двумерным изображениям хорошо известна и имеет множество приложений. В частности, эта проблема возникает при распознавании позы и жестов человека в системах видеонаблюдения. Особенность рассматриваемой нами постановки этой задачи состоит в том, что двумерные изображения являются бинарными и представляют собой лишь силуэты пространственного объекта. Такая задача, в частности, возникает в системах видеонаблюдения, работающих в условиях плохой освещенности. В этом случае камеры плохо передают текстурные особенности изображений и позволяют с достоверностью выявить лишь силуэты представленных на изображении объектов. Для распознавания позы и жестов требуется по этим силуэтам восстановить пространственную форму столь сложного и изменчивого объекта, как фигура человека.

Невозможность анализа изображений на уровне текстур не позволяет применить для такой постановки задачи хорошо известные методы, основанные на автоматическом выявлении общих точек, присутствующих на обоих изображениях стереопары. Очевидно, что если на изображении представлен лишь силуэт объекта, то достоверно на нем можно выявить лишь точки границы этого объекта. Но на двух картинках в стереопаре изображений границы силуэтов полностью различаются, т. е. все точки границы одного силуэта отличаются от всех граничных точек другого силуэта. Поэтому выявление общих точек невозможно.

В общей постановке задачи можно выделить два варианта, различающихся наличием так называемых окклюзий. Силуэтное изображение объекта называется изображением без окклюзии, если в каждую его точку проектируется не более двух точек поверхности объекта. В соответствии с этим первый, более простой, вариант задачи представляют собой работу с изображениями без окклюзий. В этом случае в силуэтах видны голова и все конечности. Второй вариант задачи — когда окклюзии имеют место. В данной работе ограничимся рассмотрением изображений без окклюзий. Не будем также затрагивать вопрос выделения силуэтов на исходных картинах, считая что они уже получены с хорошей (1–2 пикселя) точностью.

Подходы к решению этой задачи в полном объеме для сложных объектов в настоящее время только лишь формируются [1–3]. Можно сослаться на работу [4], в которой описан метод восстановления поверхности сложного пространственного объекта (фигура лошади) по серии силуэтных изображений, полученных в разных ракурсах. Для этого метода требуются изображения хорошего качества, а также большие ресурсы процессорного времени.

Вместе с тем существует определенный класс объектов, чьи структурные особенности позволяют решить задачу восстановления пространственной формы и при отсутствии видимых общих точек. Это объекты, состоящие из круговых цилиндров. Особенности такого пространственного объекта, как фигура человека, позволяют рассматривать его приближенно как объединение нескольких «цилиндрических» элементов, имеющих локальную осевую симметрию. Такие объекты некоторые

* Работа выполнена при поддержке РФФИ, гранты № 08-01-00670, № 08-07-00270.

¹Московский физико-технический институт, AchikoTsi@gmail.com

авторы называют обобщенными цилиндрами [5]. Встречается также термин «трубчатые объекты» (tubular) [6]. Если говорить более строго, под цилиндрическим элементом понимается пространственное тело, образованное семейством шаров, центры которых расположены на некоторой осевой пространственной кривой. В работах [7, 8] такие объекты называются пространственными жирными кривыми. Представляют интерес объекты, которые могут быть представлены в виде объединения небольшого числа пространственных жирных кривых. Человеческая фигура в некотором приближении также может быть составлена из жирных кривых по аналогии с тем, как дети лепят человечков из пластилина.

Предлагаемый в данной работе подход к решению основывается на идее построения скелетов для стереопары силуэтных изображений. Скелет представляет собой совокупность срединных осей силуэта, определяемых как геометрическое место точек — центров вписанных в силуэт окружностей. Использование скелетов открывает несколько возможных путей для решения задачи. Рассмотрим два из них.

Первый путь состоит в построении пространственной циркулярной модели человеческой фигуры. Он основывается на идее конструирования стереопар «невидимых» общих точек обоих изображений. Срединные оси силуэтов предлагается рассматривать как плоские проекции пространственных осевых линий жирных кривых, составляющих фигуру человека. Такой подход позволяет свести задачу восстановления этих пространственных жирных линий к вычислению пространственных кривых по стереопарам их проекций. Результатом решения задачи является циркулярная модель, представляющая собой объединение нескольких пространственных жирных линий.

Второй путь предполагает упрощенное описание фигуры человека в виде «шарнирной» модели заданной структуры. Шарнирная модель также состоит из пространственных жирных линий. Но эти жирные линии являются линейными сегментами постоянной заданной ширины. Поза человека ищется путем подбора некоторого преобразования шарнирной модели, при котором ее плоские проекции будут в наибольшей степени совпадать со скелетами стереопары силуэтов.

2 Циркулярная и шарнирная модели фигуры человека

Рассмотрим множество точек T в евклидовом пространстве R^3 , имеющее вид связного графа,

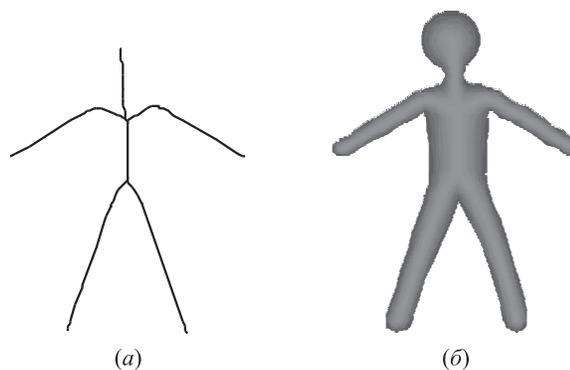


Рис. 1 Циркулярная модель фигуры человека: осевой граф (а) и огибающая поверхность (б)

описывающего человеческую фигуру (рис. 1). Граф имеет пять терминальных вершин и три вершины третьей степени, а его ребра являются непрерывными линиями. При этом ребра не имеют точек пересечения, не совпадающих с их концами.

С каждой точкой $t \in T$ графа T связан некоторый шар c_t с центром в этой точке. Это семейство шаров $C = \{c_t, t \in T\}$ называется *циркулярным графом*, для краткости — *циркуляром* [9]. Граф T называется *осевым графом* циркулярного графа. Объединение $S = \bigcup_{t \in T} c_t$ всех шаров семейства C как точечных множеств является циркулярной моделью фигуры человека. Границей модели является огибающая поверхность семейства шаров C .

Восстановление циркулярной модели сводится к нахождению составляющих циркулярный граф жирных линий, т. е. к описанию их осевых линий и функции ширины, задающей для каждой точки осевой линии радиус шара с центром в этой точке. При этом построение жирных линий должно осуществляться таким образом, чтобы проекции циркулярной модели хорошо согласовывались со стереопарой исходных изображений.

Шарнирная модель описывает фигуру человека как объединение 10 шарнирно закрепленных твердых тел. Каждый элемент этой конструкции представляет собой локально симметрический объект, образованный объединением шаров одинакового радиуса с центрами на прямолинейном отрезке (рис. 2).

С каждым элементом свяжем систему координат, ось X которой направлена вдоль его оси симметрии, а начало координат находится в точке крепления его с родительским телом. Для элемента известна его длина и ширина, а также точка крепления в системе координат родительского тела. По иерархии крепления элементы шарнирной модели

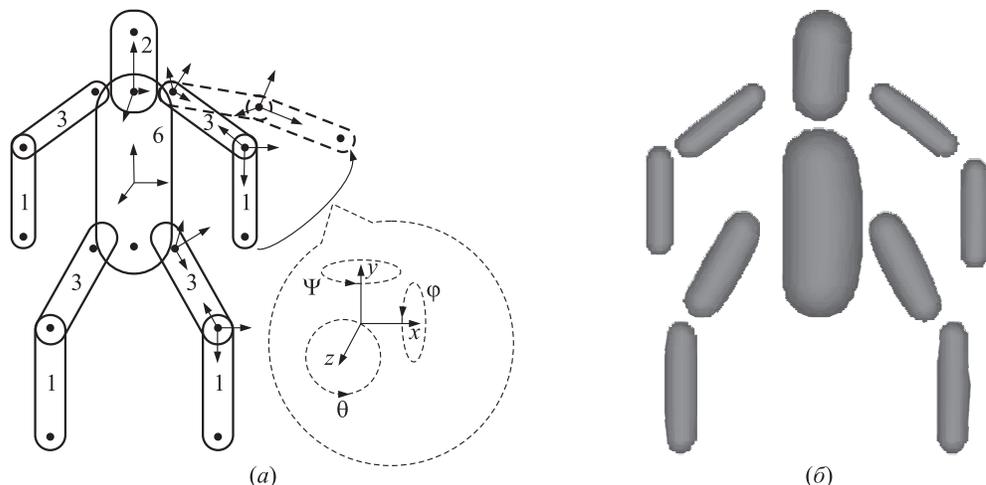


Рис. 2 Шарнирная модель фигуры человека: структура модели (а) и пространственные элементы модели (б)

образуют дерево с корневым элементом, соответствующим туловищу человека.

Вращение каждой части шарнирной модели задается относительно родительской системы координат. Его можно параметризовать с помощью суперпозиции трех вращений $R_\theta \circ R_\psi \circ R_\varphi$ относительно осей X, Y, Z своей системы координат (см. рис. 2). Так как структура скелета человека допускает не всякие вращения, для каждого тела введем ограничения на углы в виде параллелепипеда: $\varphi_{\min} \leq \varphi \leq \varphi_{\max}$, $\psi_{\min} \leq \psi \leq \psi_{\max}$, $\theta_{\min} \leq \theta \leq \theta_{\max}$.

Под позой объекта будем понимать вектор значений динамических параметров модели. Каждой позе соответствует точка в пространстве из 24 динамических параметров фигуры, а все множество поз описывается параллелепипедом Θ в 24-мерном пространстве $\Theta \subset R^{24}$. Таким образом, зада-

ча состоит в том, чтобы по стереопаре бинарных изображений найти вектор динамических параметров шарнирной модели, аппроксимирующей форму пространственной фигуры.

3 Восстановление позы в виде циркулярной модели

3.1 Структура метода

Предлагаемый подход к восстановлению формы человеческой фигуры в виде совокупности жирных линий иллюстрируется на рис. 3–5.

Сначала для силуэтов, полученных на основе сегментации исходных изображений (см. рис. 3), строятся их скелеты в виде серединных осей [10], с которыми связано множество вписанных в силуэты окружностей с центрами на серединных осях (см.

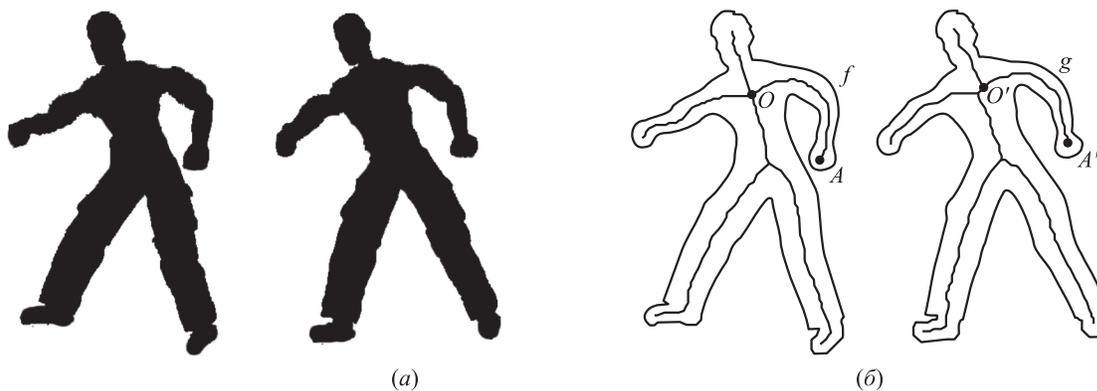


Рис. 3 Стереопара исходных силуэтов (а) и стереопара скелетов (б)

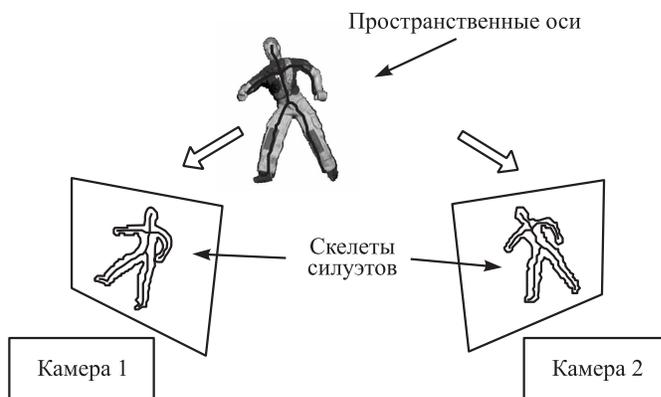


Рис. 4 Предположение, что проекции осей объекта совпадают со скелетом силуэтов

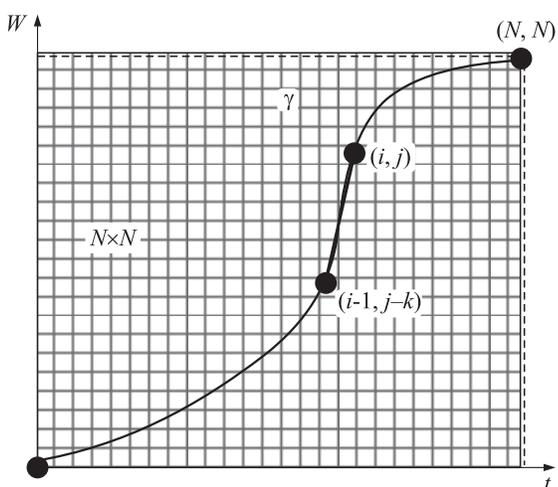


Рис. 5 Кривая, сопоставляющая стереопары точек

рис. 3). После этого из стереопары скелетов конструируется пространственный скелет объекта, а по семействам вписанных кругов вычисляются радиусы сфер с центрами на пространственном скелете. Таким образом, поза человека может быть представлена пространственным скелетом, либо огибающей поверхностью для семейства построенных сфер.

В случае, когда нет окклюзии в силуэтах, в первом приближении можно считать, что направление взгляда на объект не сильно отклоняется от оптической оси камеры. Тогда можно предполагать, что на плоском изображении образ пространственной оси локально симметричного объекта совпадает со скелетом силуэта (см. рис. 4). Строго говоря, центральная проекция сферы на плоскость является эллипсом, но приближенно эти эллипсы будем считать окружностями. По двум плоским проекциям пространственной оси с помощью эпиполярной

геометрии [11] можно восстановить пространственную ось.

3.2 Идентификация реперных точек на скелетах

Рассмотрим стереопару силуэтов и их скелетов (см. рис. 3). Объект «фигура человека» приближенно представляется осесимметричными элементами. Тогда на плоских изображениях образ пространственной оси осесимметричного элемента совпадает с соответствующими ветвями скелетов (см. рис. 3, кривые OA и $O'A'$). Отсюда можно сделать предположение, что множество стереопарных точек ветви OA скелета одного силуэта совпадает с ветвью $O'A'$ скелета другого силуэта, что позволяет построить кривую в пространстве. Задача состоит в том, чтобы наилучшим образом установить соответствие между точками этих скелетных ветвей. Если задать кривую OA как непрерывное отображение $r_1: [0, 1] \rightarrow R^2$, а кривую $O'A'$ как $r_2: [0, 1] \rightarrow R^2$, задача сведется к нахождению непрерывного монотонного отображения $w: [0, 1] \rightarrow [0, 1]$, которое сопоставляет стереопары точек $r_1(t) \leftrightarrow r_2(w(t))$ и при этом минимизирует расхождение, заданное в виде функционала

$$\min \int_0^1 \rho(r_1(t), r_2(w(t))) \sqrt{1 + w'(t)^2} dt.$$

Здесь $\rho(X, Y)$ — функция штрафа — отражает, насколько хорошо сопоставляются точки X и Y . Выбор этой функции осуществляется на основе следующего соображения. Для каждой точки изображения существует луч в пространстве, который в нее проецируется. При идеально правильном сопоставлении точек X и Y проходящие через них лучи должны пересекаться. Поэтому в качестве ρ

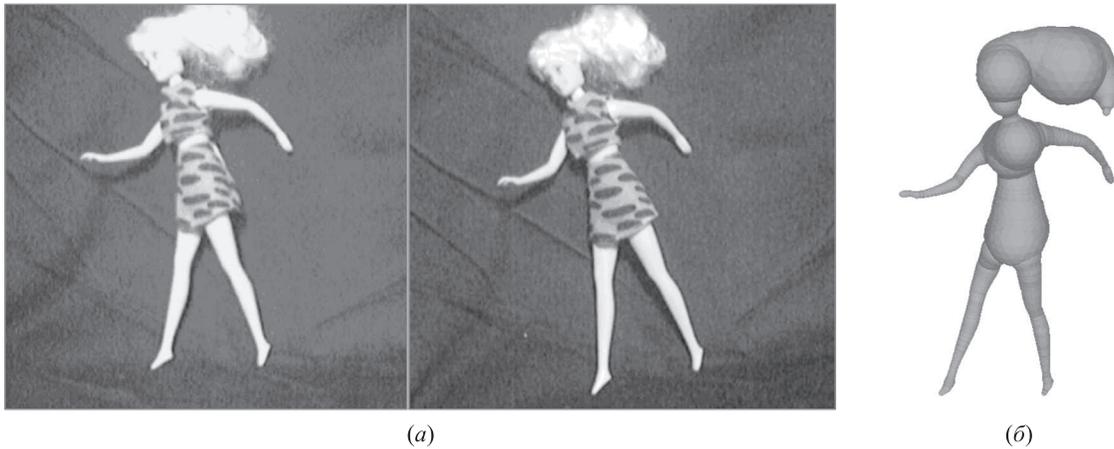


Рис. 6 Стереопара изображений и полученный пространственный объект

можно взять расстояние между скрещивающимися лучами, которые проецируются в точки X и Y .

Полученная задача решается методом динамического программирования. Дискретизируя задачу сеткой $N \times N$, для нахождения точного решения получаем сложность $O(N^2d)$ (см. рис. 5). Здесь d — ограничение, задающее коридор для кривой. Построив пространственные оси, используя ширину скелетов, можно вычислить размеры шаров с центрами на этих осях.

Пример визуализации модели фигуры, полученной по стереопаре изображений, представлен на рис. 6. Как видно из этого примера, визуализация

является не вполне реалистичной, поскольку описание тела человека цилиндрами представляется весьма грубым. Однако для вычислений, связанных с распознаванием позы, такая точность представляется вполне приемлемой.

3.3 Вычислительные эксперименты

Эксперименты с восстановлением пространственной модели фигуры человека проводились с куклами размером 30 см. Это объясняется лишь упрощением съемки в лабораторных условиях. На рис. 6 показана стереопара исходных изо-

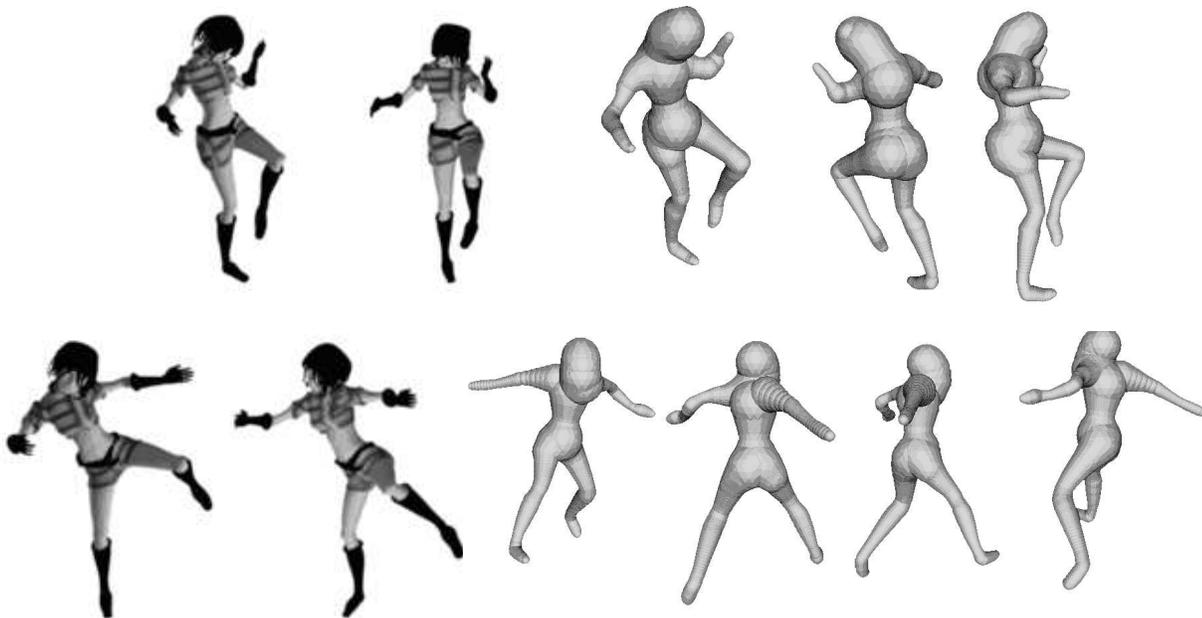


Рис. 7 Стереопары и полученные пространственные объекты для «Kungfu girl»

бражений и полученный пространственный объект. Эксперименты также проводились с синтетическими данными «Kungfu girl», предоставленными группой Graphics-Optics-Vision из института Max-Planck [12]. Эти данные представляют собой синтезированные пространственные модели виртуальных сцен и их проекции в виде двумерных изображений размером 320×240 . Результаты реконструкции показаны на рис. 7.

Эксперименты показывают, что модель локально симметричного объекта хорошо описывает фигуру человека в целом. Скорость работы на компьютере Intel Pentium IV, Core 2 Duo, 2800 МГц составила более 5 кадров/с. Это дало возможность использовать предложенный метод в системах компьютерного зрения в реальном масштабе времени их работы.

4 Восстановление позы в виде шарнирной модели

4.1 Предлагаемый подход к решению задачи

Предположим, что определена шарнирная модель объекта вместе с длинами составляющих ее элементов. Задача подгонки шарнирной модели заключается в следующем:

Дано: Шарнирная модель; стереопара силуэтов s_1, s_2 ; проекционные матрицы камер.

Найти: Вектор динамических параметров шарнирной модели $\theta^* \in \Theta$, который имеет стереопару силуэтов, максимально совпадающую с наблюдаемой стереопарой.

Другими словами, $\theta^* = \arg \min_{\theta \in \Theta} \mu(\vec{\theta}, S_1, S_2)$, где

$\mu(\vec{\theta}, S_1, S_2)$ — величина сходства позы с силуэтами. Такая постановка задачи возможна также для случая с окклюзиями. При этом основная сложность заключается в том, что любая целевая функция для таких сложных объектов, как фигура человека, является многоэкстремальной [13]. В случае, когда нет окклюзии, удается выписать хорошую критериальную функцию. При этом на основе скелетов силуэтов делается предварительная сегментация силуэта.

4.2 Подгонка шарнирной модели под стереопару наблюдаемых силуэтов для задачи без окклюзии

В данном подходе предполагается, что известна шарнирная модель объекта вместе с длинами

отдельных ее элементов. Опишем функцию сходства шарнирной модели с наблюдаемыми силуэтами и метод ее минимизации. Для каждого силуэта $S_i: i \in \{1, 2\}$ построим его базовый скелет. Анализ геометрии скелетов позволяет выделить на них оси для каждой из 6 частей тела человека. На рис. 8 показана такая сегментация для одного из скелетов. Через $A_i, \gamma_i, i = 1, \dots, 5$, обозначены концевые вершины и сегменты соответственно. Для любой позы $\vec{\theta} \in \Theta$ шарнирной модели построим проекции ее осей на плоскость камеры и выделим на них соответствующие ветви. Обозначим их через $B_i, \lambda_i, i = 1, \dots, 5$ (см. рис. 8). Введем расхождение ρ между двумя кривыми γ и λ на плоскости как

$$\rho(\gamma, \lambda) = \int_{\gamma} d(\gamma(t), \lambda) d\gamma + \int_{\lambda} d(\lambda(t), \gamma) d\lambda.$$

Здесь $d(a, \lambda) = \min_{b \in \lambda} |a - b|^2$ задает расстояние от точки a до кривой λ . Тогда функцию сходства $\mu(\theta, S_1, S_2)$ шарнирной модели с наблюдаемыми силуэтами введем как $\mu(\theta, S_1, S_2) = \mu_1(\theta, S_1) + \mu_2(\theta, S_2)$, где $\mu_1(\theta, S_1) = \sum_{i=1}^5 (\rho(\gamma_i, \lambda_i) + \alpha |A_i - B_i|^2)$ задает сходство по первой камере, а μ_2 определяется аналогичным образом для второй камеры. Второй член $\alpha |A_i - B_i|^2$ учитывает расхождение в концевых точках.

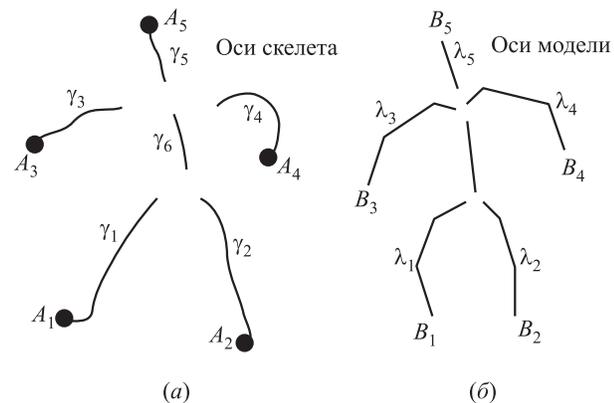


Рис. 8 Сегментация скелета (а) и сегментация проекции осей шарнирной модели (б)

Полученная функция сходства обладает свойством унимодальности. Для простоты реализации использовалась квазиньютоновская схема LBFGS с численным градиентом. Градиент вычислялся разностной схемой по двум точкам. На рис. 9 показан результат подгонки, спроецированный в разных ракурсах. Эксперименты показывают, что шарнирная модель хорошо описывает позу человека.



Рис. 9 Полученная шарнирная модель в разных ракурсах

5 Заключение

В работе рассмотрена задача восстановления позы локально симметричного объекта на примере фигуры человека в случае отсутствия окклюзии в силуэтах. Рассматриваются две модели фигуры. В первом случае фигура описывается как объединение пространственных жирных линий и предлагается метод ее восстановления. Во втором случае фигура описывается в виде шарнирной модели и предлагается метод ее подгонки по стереопаре силуэтов. В обоих случаях эксперименты показали положительный результат. В будущем планируется разработать методы подгонки шарнирной модели в случае окклюзии. Скелетное представление силуэта позволяет анализировать окклюзию, что может быть использовано для построения начального приближения в задаче подгонки.

Автор выражает благодарность своему научному руководителю профессору кафедры интеллектуальных систем Московского физико-технического института Местецкому Леониду Моисеевичу за помощь в постановке задачи и внимание к работе.

Литература

1. Agarwal A., Triggs B. Recovering 3D human pose from monocular images // Pattern Analysis and Machine Intelligence, 2006.
2. Tong M., Liu Yu., Huang T. S. 3D human model and joint parameter estimation from monocular image // Pattern Recognition Letters, 2007. Vol. 28. P. 797–805.
3. Balan A. O., Sigal L., Black M. J., Davis J. E., Haussecker H. W. Detailed human shape and pose from images // IEEE Conference on Computer Vision and Pattern Recognition, 2007.
4. Yezzi A. J., Soatto S. Structure from motion for scenes without features // 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'03) Proceedings. Vol. 1. P. 525–532.
5. Senior A. Real-time articulated human body tracking using silhouette information // IEEE Workshop on Visual Surveillance/PETS Proceedings. France, 2003.
6. Cumani A. Guiducci A. Recovering the 3D structure of tubular objects from stereo silhouettes // Pattern Recognition, 1997. Vol. 30. No. 7. P. 1051–1059.
7. Местецкий Л. М., Шетинин Д. В. Объемные примитивы Безье // Графикон-2001: Тр. 11-й международной конф. — Нижний Новгород, 2001. С. 161–165.
8. Mestetskiy L. Shape comparison of flexible objects // Conference (International) on Computer Vision Theory and Applications, 2007.
9. Местецкий Л. М. Непрерывная морфология бинарных изображений: фигуры, скелеты, циркуляры. — М.: Физматлит, 2009.
10. Местецкий Л. М. Непрерывный скелет бинарного растрового изображения // Графикон-98: Тр. международной конф. — М.: МГУ, 1998. С. 71–78.
11. Форсайт Д., Понс Ж. Компьютерное зрение. — М.: Вильямс, 2004.
12. «Kung-Fu Girl» — a synthetic test sequence for multi-view reconstruction and rendering research. <http://www.mpi-inf.mpg.de/departments/irg3/kungfu>.
13. Sminchisescu C., Triggs B. Kinematic jump processes for monocular 3D human tracking // 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'03) Proceedings. Vol. 1. P. 69–76.

МОДЕЛИРОВАНИЕ И КЛАССИФИКАЦИЯ МНОГОКАНАЛЬНЫХ ДИСТАНЦИОННЫХ ИЗОБРАЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ КОПУЛ*

В. А. Крылов¹

Аннотация: Предложен метод моделирования многоканальных дистанционно полученных изображений (со спутника, самолета) с использованием копул. Суть предлагаемого подхода состоит в применении традиционных статистических моделей для моделирования вероятностных распределений отдельных каналов и построении совместного распределения для многоканального изображения при помощи копул. Рассмотрено также применение разработанного подхода в модели марковского случайного поля (МСП) для байесовской классификации. Эксперименты с реальными изображениями, полученными радаром с синтезированной апертурой, демонстрируя результаты высокой точности, указывают на преимущества предложенного подхода по сравнению с существующими методами.

Ключевые слова: многоканальное изображение; копула; марковское случайное поле; байесовская классификация

1 Введение

Статистический анализ изображений является на сегодняшний день активно развивающейся областью. Одним из важных и активно исследуемых источников изображений служат спутники и радары, предоставляющие данные, широко используемые в задачах картографии, риск-менеджмента (пожары, наводнения), эпидемиологии и др. Современное оборудование позволяет получать многоканальные изображения, число каналов в которых зачастую доходит до сотен (мультиспектральные изображения). Для обработки таких данных требуется адаптация существующих методов обработки отдельных каналов для работы с многоканальными изображениями. В данной работе рассматривается возможный статистический подход к решению подобной задачи многомерного моделирования.

На практике для решения задач требуется иметь аккуратную статистическую модель рассматриваемых данных, в случае многоканальных изображений — многомерную модель. Большинство существующих в литературе подходов строится в предположении, что вектор многоканальных данных имеет какое-то специфическое распределение, например подчиняется многомерному нормальному распределению [1] для мультиспектрального вектора или многомерному распределению Вишарта (Wishart) [2] для комплексного мультиполяризо-

ванного микроволнового изображения. Недостаток подобных подходов состоит в большой потере точности при повышении размерности из-за присутствия шума, погрешностей регистрации данных и аппроксимации. В рамках статистического подхода адекватным способом решения данной проблемы является использование более гибких моделей для построения многомерных распределений. Одной из таких моделей являются копулы [3].

В работе также рассматривается одна из классических проблем обработки дистанционно полученных изображений — задача классификации. Многомерные вероятностные распределения применяются совместно с моделью марковского случайного поля для байесовской классификации. Эксперименты с реальными изображениями, полученными радаром с синтезированной апертурой, демонстрируют преимущества модели с использованием копул.

Статья организована следующим образом. В разд. 2 приводится изложение основ теории копул. В разд. 3 строится модель многомерных распределений с использованием копул. В разд. 4 излагается построение модели марковского случайного поля в задаче классификации. В разд. 5 приводятся эксперименты по применению разработанной модели к изображениям, полученным радаром с синтезированной апертурой. Раздел 6 содержит заключение.

* Работа выполнена при поддержке ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009–2013 гг.

¹ Московский государственный университет им. М. В. Ломоносова, факультет вычислительной математики и кибернетики, vkrylov@cs.msu.ru

2 Копулы

В этом разделе приводится краткий обзор теории двумерных копул. Все сформулированные ниже результаты могут быть обобщены на многомерный случай [3].

Двумерная копула является вероятностным распределением на $[0, 1]^2$ таким, что маргинальные распределения распределены равномерно на $[0, 1]$.

Определение. Двумерной копулой называется отображение $C : [0, 1]^2 \rightarrow [0, 1]$, такое что:

- 1) оба маргинальных распределения являются равномерно распределенными с.в. на $[0, 1]$;
- 2) $\forall u, v \in [0, 1]: C(u, 0) = C(0, v) = 0$ и $C(u, 1) = u, C(1, v) = v$;
- 3) $\forall u_1 \leq u_2, v_1 \leq v_2 \in [0, 1]: C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) + C(u_1, v_1) \geq 0$.

Теоретическое обоснование правомерности использования копул в прикладных задачах обеспечивает теорема Склара [3]:

Теорема. Пусть X, Y — произвольные случайные величины (с.в.) с совместным распределением $H(x, y)$ и маргинальными функциями распределения F и G . Тогда существует копула C такая, что

$$H(x, y) = C(F(x), G(y)) \quad (1)$$

$\forall x, y \in \mathbb{R}$. Если F и G непрерывны, то такая копула C единственна.

Если X, Y имеют плотности $f(x)$ и $g(y)$, то плотность совместного распределения (1) можно представить в виде

$$h(x, y) = f(x)g(y) \frac{\partial^2 C}{\partial x \partial y}(F(x), G(y)), \quad (2)$$

где частная производная $\frac{\partial^2 C}{\partial x \partial y}(x, y)$ задает плотность копулы $C(x, y)$.

На практике широкое применение получил класс архимедовых копул [3].

Определение. Архимедовой копулой называется копула C вида

$$C(u_1, u_2) = \phi^{-1}(\phi(u_1) + \phi(u_2)),$$

где функция $\phi(u)$, называемая функцией-генератором, удовлетворяет следующим требованиям: (1) $\phi(u)$ непрерывна на $[0, 1]$; (2) $\phi(u)$ монотонно убывает, $\phi(1) = 0$; (3) $\phi(u)$ выпуклая.

В рассматриваемых задачах моделирования и классификации дистанционных изображений предлагается использовать следующие копулы: пять архимедовых копул (Clayton, Ali-Mikhail-Наq,

Gumbel, Frank, A12) [3] и одну неархимедову копулу, содержащую абсолютно непрерывную и сингулярную компоненту (Marchal–Olkin) [3]. Такой набор копул обеспечивает достаточно широкий выбор моделируемых структур зависимости [4], тем не менее для некоторых рассматриваемых данных этот набор копул может потребовать пополнения. Эксперименты с более обширным набором копул проводились в [5]. Информация о рассматриваемых копулах собрана в табл. 1.

Простейшим инструментом для оценки копул является коэффициент ранговой корреляции Кендалла τ двух независимых реализаций (Z_1, Z_2) и (\hat{Z}_1, \hat{Z}_2) с общим законом распределения $H(x, y)$: $\tau = \mathbb{P}\{(Z_1 - \hat{Z}_1)(Z_2 - \hat{Z}_2) > 0\} - \mathbb{P}\{(Z_1 - \hat{Z}_1)(Z_2 - \hat{Z}_2) < 0\}$. На практике при наличии реализаций $z_{1,l}$ и $z_{2,l}, l = 1, \dots, N$, эмпирической оценкой коэффициента Кендалла является [3]:

$$\hat{\tau} = \frac{\sum_{l=1}^{N-1} \sum_{k=l+1}^N z_{1,lk} z_{2,lk}}{C_2^N},$$

где

$$z_{n,lk} = \begin{cases} 1, & \text{если } z_{n,l} \leq z_{n,k} \\ -1 & \text{— иначе} \end{cases}, \quad n = 1, 2.$$

Интегрируя в определении τ по (\hat{Z}_1, \hat{Z}_2) , имеем

$$\left. \begin{aligned} \tau &= 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) - 1; \\ \tau &= 1 + 4 \int_0^1 \frac{\phi_A(t)}{\phi_A'(t)} dt, \end{aligned} \right\} \quad (3)$$

где слева получен общий вид зависимости между τ и C , а справа — зависимость в частном случае архимедовых копул с функцией-генератором $\phi_A(t)$. Все рассматриваемые в работе копулы однопараметрические, оценка их параметра θ может быть получена из (3) (см. табл. 1).

3 Моделирование совместных распределений

Рассматривается задача построения M условных распределений изображения с D каналами. В центре внимания в данной работе находится моделирование совместных распределений, поэтому предполагается что маргинальные распределения $p_{dm}(y_d | \omega_m), d = 1, \dots, D, m = 1, \dots, M$, где ω_m — событие, состоящее в попадании наблюдения в

Таблица 1 Рассматриваемые копулы: Clayton, Ali-Mikhail-Haq (AMH), Gumbel, Frank, A12 и Marchal–Olkin (Marchal)

Копула	$C(u, v)$	$\theta(\tau)$	τ -интервал
Clayton	$(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$	$\theta = \frac{2\tau}{1-\tau}$	$\tau \in (0, 1]$
AMH	$\frac{uv}{1 - \theta(1-u)(1-v)}$	$\tau = \frac{3\theta - 2}{3\theta} - \frac{2(1-\theta)^2}{3\theta^2} \ln(1-\theta)$	$\tau \in \left[-0, 1817, \frac{1}{3}\right]$
Gumbel	$\exp\left(-\left[(-\ln(u))^\theta + (-\ln(v))^\theta\right]^{1/\theta}\right)$	$\theta = \frac{1}{1-\tau}$	$\tau \in [0, 1]$
Frank	$-\frac{1}{\theta} \log\left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}\right)$	$\tau = 1 - \frac{4}{\theta^2} \int_0^\theta \frac{t}{e^{-t} - 1} dt$	$\tau \in [-1, 1] \setminus \{0\}$
A12	$\left(1 + [(u^{-1} - 1)^\theta + (v^{-1} - 1)^\theta]^{1/\theta}\right)^{-1}$	$\theta = \frac{2}{3 - 3\tau}$	$\tau \in \left[\frac{1}{3}, 1\right]$
Marchal	$\min(u^{1-\theta}v, uv^{1-\theta})$	$\theta = \frac{2\tau}{\tau + 1}$	$\tau \in [0, 1]$

класс с номером m , уже получены каким-либо методом (включая оценку параметров). Таким образом, обобщая (2), имеем совместную плотность вида

$$p_m(\mathbf{y}|\omega_m) = p_{1m}(y_1|\omega_m) \dots \dots p_{Dm}(y_D|\omega_m) \frac{\partial^D C_m^*}{\partial y_1 \dots \partial y_D}(F_{1m}(y_1), \dots, \dots, F_{Dm}(y_D)). \quad (4)$$

Следующим вопросом является выбор конкретной копулы для каждого класса. В литературе предложен целый ряд методов для решения этой проблемы. Так, для этого применяются подходы, основанные на применении информационных критериев (в частности, Акаике) [3]. В [6] разработан метод для визуального анализа (K-plots) адекватности копулы реальным данным. В [4] применяется функция правдоподобия. В данной работе предлагается использовать критерий согласия χ^2 Пирсона (КСП) для нахождения наиболее подходящей копулы.

Итак, для выбора копулы C_m^* из (4) среди копул в табл. 1 сначала отбрасываются копулы, эмпирическое $\hat{\tau}_m$ которых лежит вне соответствующего им τ -интервала. Затем для оставшихся копул оценивается значение параметра θ и выбирается копула, лучше всего согласующаяся с наблюдениями на основе КСП. Нулевая гипотеза КСП состоит в согласии наблюдаемых частот с частотами, предсказанными теоретической моделью. Статистика КСП имеет следующий вид:

$$X^2 = \sum_{i=1}^N \frac{(O_i - E_i)^2}{E_i},$$

где O_i и E_i — наблюдаемые и предсказанные частоты; n — число классов. P -значение КСП определя-

ется из сравнения наблюдаемого значения X^2 с χ^2 -распределением: $X^2 \sim \chi_{n-r-1}^2$, где r — количество ограничений числа степеней свободы (количество параметров модели).

4 Классификация многоканальных изображений

На базе построенных условных распределений (4) самую примитивную классификацию можно получить, используя метод максимального правдоподобия, выбирая из M классов для каждого пикселя тот, который имеет максимальную вероятность. Однако такого рода классификация слишком неоднородная, и для решения этой проблемы к правдоподобию добавляется регуляризующее слагаемое, обеспечивающее большую однородность решения. Одним из таких способов регуляризации является модель МСП [7]. Концепция МСП получена обобщением марковского свойства на двумерный случай решетки. Классическая теория МСП и скрытых МСП приводится в [7]. Здесь изложение ограничится кратким введением в скрытые МСП.

Задачу классификации относят к классу задач с неполными данными $x = (y, z)$, где y — наблюдаемые данные (исходное изображение, поле Y); z — данные, подлежащие восстановлению (метки классов, поле Z). В решении задачи классификации используется модель скрытого МСП. В этой модели неизвестные метки z_i предполагаются МСП, т.е. зависящими от меток *только* соседних пикселей. Теорема Хаммерсли–Клиффорда (Hammersley–Clifford) [7] предоставляет удобное

представление для совместного распределения с. в., входящих в МСП, в виде распределения Гиббса

$$P_G(z) = W^{-1} \exp(-H(z))$$

с энергией $H(z) = \sum_{c \in C} V_c(z_c)$, (5)

где W — нормирующая константа, $V_c(z_c)$ — потенциалы; C — система множеств (клик) на решетке S [7]. В данной работе рассматриваются двухместные анизотропные потенциалы (т. е. попарно с каждым из 8 соседних пикселей):

$$H(z|\beta) = \sum_c V(z_c|\beta) = \sum_{c=\{s,s'\} \in C} [-\beta \delta_{z_s=z_{s'}}], \quad (6)$$

где $\delta_{z_s=z_{s'}} = 1$, если $z_s = z_{s'}$, и 0 — иначе. В (6) β играет роль веса, т. е. чем больше значение β , тем выше вклад регуляризирующего слагаемого $H(z|\beta)$ в суммарную энергию поля (Y, Z) , что хорошо видно во втором уравнении системы (7), приведенной ниже.

В свою очередь, наблюдения y_i предполагаются условно независимыми $P(y_i|y_{\Omega \setminus \{i\}}, z_i) = P(y_i|z_i)$. Таким образом, при наличии условных вероятностей $p_m(\mathbf{y}|\omega_m)$ распределение скрытого МСП (Y, Z) на S имеет вид:

$$\left. \begin{aligned} P_G(z) &= W^{-1} \exp(-U(\omega_m|\mathbf{y}, \beta)); \\ U(\omega_m|\mathbf{y}, \beta) &= \\ &= \sum_{i \in S} \left[-\ln p_m(\mathbf{y}|\omega_m) - \beta \sum_{s:\{i,s\} \in C} \delta_{z_i=z_s} \right], \end{aligned} \right\} \quad (7)$$

где первое уравнение определяет энергию $U(\cdot)$ аналогично (5), а второе уравнение задает $U(\cdot)$ как сумму энергетических вкладов полей Y и Z .

Для оценки параметра в (7) используется метод имитации отжига [8]. Для нахождения точки минимума функции $U(\omega_m|\mathbf{y}, \beta)$ (и, соответственно, решения с максимальным правдоподобием P_G) необходимо решение задачи глобальной оптимизации, например тем же методом имитации отжига, однако в связи с высокой вычислительной сложностью для экспериментов использовалась локальная оптимизация методом ICM [9].

5 Эксперименты

Для проведения экспериментов используется амплитудное изображение, полученное радаром с синтезированной апертурой системы TerraSAR-X (© Infoterra), с двумя каналами поляризации и разрешением 6 м на пиксель (рис. 1, а). Классификация проводится по трем классам: вода, влажная и сухая почва. В качестве способа оценки маргинальных распределений классов применяется алгоритм с использованием метода конечных смесей [10] на обучающем изображении с заранее предоставленной классификацией. Стоит отдельно отметить, что копулы представляются логичным инструментом для обобщения [10] на многоканальный случай, не требующим изменения самого метода. Для сравнения также приводятся результаты классификации при моделировании совместного распределения моделью двумерного (2D) Накагами–Гамма распределения [2].

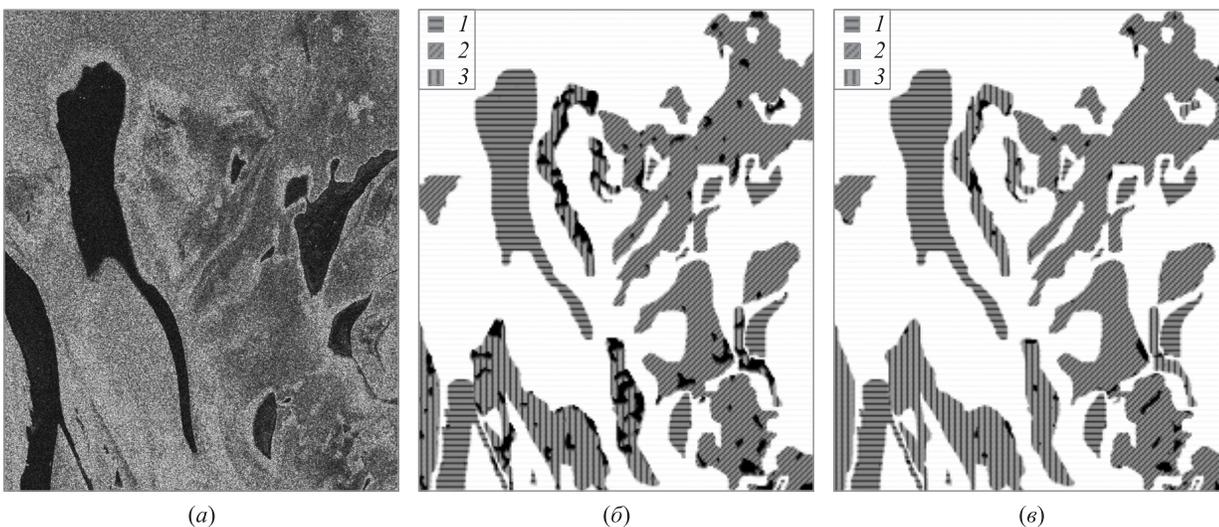


Рис. 1 Исходное изображение (а), классификация с 2D моделью Накагами–Гамма (б) и классификация с моделированием копулами (в): вне карты классификации — белое, правильно классифицированная вода (1), влажная почва (2), сухая почва (3), ошибочная классификация — черное

Доля правильной классификации на рис. 1 составляет 87,1% для 2D Накагами–Гамма модели (рис. 1, б) и 94,8% для предлагаемой модели с использованием копул (рис. 1, в). При помощи КСП были выбраны следующие копулы: Gumbel — для воды и Frank — для сухой и влажной почвы. Такое повышение качества классификации возможно благодаря более точному моделированию совместных распределений, предоставляемому копулами. Дополнительные эксперименты и обсуждение можно найти в [5].

6 Заключение

В работе предлагается подход к моделированию распределений многоканальных изображений с использованием копул. По сравнению с классическими методами оценки многомерных распределений конкретного вида, использование копул предоставляет большую гибкость, позволяя получать более аккуратные модели. Предложенная модель применяется для классификации дистанционно полученных изображений. Эксперименты с изображениями, полученными радаром с синтезированной апертурой, подтверждают высокую описательную точность предложенного подхода.

Использованное в работе изображение системы TerraSAR-X распространяется на свободной основе (©Infoterra, www.infoterra.de).

Автор выражает благодарность своему научному руководителю доц. В. Ф. Матвееву, а также проф.

Г. Мозеру, проф. С. Серпико и проф. Д. Зерубии за помощь в подготовке этой работы.

Литература

1. *Landgrebe D. A.* Signal theory methods in multispectral remote sensing. — N.-Y.: Wiley, 2003.
2. *Lee J.-S., Hoppel K. W., Mango S. A., Miller A. R.* Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery // *IEEE Trans. Geosci. Remote Sens.*, 1994. Vol. 32. No. 10. P. 1017–1028.
3. *Nelsen R. B.* An introduction to copulas. — N.-Y.: Springer, 2007.
4. *Huard D., Évin G., Favre A.-C.* Bayesian copula selection // *Comput. Stat. Data Anal.*, 2006. Vol. 51. No. 2. P. 809–822.
5. *Krylov V., Zerubia J.* High resolution SAR image classification. INRIA, Research Report 7108, 2009.
6. *Genest C., Boies J.-C.* Detecting dependence with Kendall plots // *The American Statistician*, 2003. Vol. 57. No. 4. P. 275–284.
7. *Besag J.* Spatial interaction and the statistical analysis of lattice systems // *J. Roy. Statistical Soc. B*, 1974. Vol. 36. No. 2. P. 192–236.
8. *Yu Y., Cheng Q.* MRF parameter estimation by an accelerated method // *Pattern Recogn. Lett.*, 2003. Vol. 24. No. 9–10. P. 1251–1259.
9. *Besag J.* On the statistical analysis of dirty pictures // *J. Roy. Statistical Soc. B*, 1986. Vol. 48. P. 259–302.
10. *Krylov V., Moser G., Serpico S., Zerubia J.* Dictionary-based probability density function estimation for high-resolution SAR data // *SPIE Proceedings*. — San Jose, USA, 2009. Vol. 7246. P. 72460S-1–72460S-12.

СВЯЗЬ МЕЖДУ ВРЕМЕННЫМИ И СТРУКТУРНО-ТОПОЛОГИЧЕСКИМИ ХАРАКТЕРИСТИКАМИ ДИАГРАММ РИТМА СЕРДЦА ЗДОРОВЫХ ЛЮДЕЙ

А. А. Кузнецов¹

Аннотация: По данным 628 регистраций электрокардиограмм (ЭКГ) у 177 здоровых и больных людей проведен сравнительный анализ параметров реальной и виртуальной диаграмм ритма сердца для оценки влияния системы регуляции на ритм сердца. Между параметрами диаграмм и информационной энтропией в условиях дискретной сезонной адаптации определены функциональные связи. Предложены «формулы функционального состояния организма», связывающие параметры макроструктуры диаграммы ритма сердца с параметрами ее ярусной микроструктуры. Обнаружено, что режим ритма сердца здорового человека вне зависимости от пола имеет цикл календарного года, в течение которого трижды дискретно меняется.

Ключевые слова: диаграмма ритма сердца; функциональное состояние организма; ярусная структура; информационная энтропия; количество информации

1 Введение

При исследовании процессов, характеризующихся большим набором параметров, возникает вопрос о «цене» каждого из них. Поначалу все параметры равноценны, поэтому обычно проводят многофакторный параметрический анализ, одним из возможных инструментов которого служит дискриминантный анализ. По величине вероятности влияния на процесс анализируемых параметров, удовлетворяющих предложенной гипотезе, все они выстраиваются в вариационный ряд по статистической значимости, теряя при этом свою равноценность. С одной стороны, это является хорошей подсказкой в выборе параметров для исследования процесса. С другой стороны, статистическая независимость выделяемых параметров вызывает большие сомнения [1–3]. При исследовании системных процессов в стороне остаются и редко реализуемые параметры, ответственные за функции управления, регуляции и контроля, и нелинейные связи между этими функциями, количественно характеризующиеся указанными параметрами. Вероятно, дискриминантный анализ (ДА) может иметь успех на начальной стадии исследования реализаций результирующих сигналов «искусственных системных процессов».

Известны попытки применения технологии ДА для исследования естественных системных процессов, к которым относится ритм сердца человека, например, по набору параметров вариабельности

ритма сердца (ВРС) [4] или по морфологическим параметрам ЭКГ [5]. Результаты таких исследований являются техническими и не дают никаких предпосылок к пониманию универсальных и индивидуальных физиологических процессов и механизмов управления ими при организации результирующего системного процесса ритма сердца. Вероятно, исследование в указанном направлении следует начинать с определения «критериев нормы» и поиска нормальных физиологических закономерностей ритма сердца как системного процесса, характера и причин искажения или нарушения этих закономерностей в онтогенезе.

В данной работе используется физическая модель исследования, включающая сравнительный качественный и количественный анализ поведения группы параметров ВРС реальной и виртуальной [6] диаграммы ритма сердца (ДРС). Целью его применения является анализ статистических зависимостей общепринятых параметров при поиске общих закономерностей в динамике ритма.

2 Экспериментальная часть

Регистрация ЭКГ проводилась монитором Холтера комплекса амбулаторной регистрации электрокардиосигнала «*Anna Flash 2000*» [7] с использованием накожных электродов для электрокардиографии. При регистрации биопотенциалов применялись двухполюсные отведения по Небу:

¹Владимирский государственный университет, artemi-k@mail.ru

первый электрод располагался во втором межреберном положении у правого края грудины (соответствует V_5^2), второй электрод располагался в области верхушки сердца. Такое расположение электродов позволяет записать переднее грудное отведение (*A-anterior*), соответствующее стандартному отведению II с максимальной амплитудой зубцов на ЭКГ. Данные каждой ЭКГ в лицензированной программе *EScreen* [8] конвертировались в ритмограммы в форме последовательности значений $R-R$ интервалов и далее посредством встроенной процедуры *Heart rate variability* в программе *EScreen* определялись выборочные значения всех параметров ВСР для каждой ритмограммы.

Проведено 628 регистраций ЭКГ у 177 здоровых людей и больных — пациентов реанимационных и кардиореанимационных отделений. Серийные и групповые регистрации здоровых молодых людей выделены отдельно и представляют основной экспериментальный материал данной работы.

Серийные двадцатиминутные посуточные регистрации ЭКГ проводились тремя сезонными сериями в течение 5–7 недель каждая в одинаковых условиях покоя в одно время суток (двумя мониторами) для двух молодых людей 21 года: юноши К. и девушки Ш. Все серийные регистрации ЭКГ в количестве $N_{\text{рег}} = 176$ проводились в домашних условиях при температуре 20–22 °С в положении лежа на спине с периодом адаптации 5–10 мин. Серия из 45 регистраций ЭКГ юноши Р. (21 год) проводилась отдельно несколько раз в сутки в течение первых двух недель февраля в разных условиях покоя и движения.

Групповые двадцатиминутные регистрации ЭКГ проводились в течение 9 недель (февраль–март 2008 г.) для группы из 32 молодых людей 19–24 лет: 20 юношей и 12 девушек. Групповые регистрации проводились в лаборатории университета один раз в неделю в интервале времени 14:00–19:00 в положении покоя: сидя, без адаптации к условиям регистрации.

3 Методика обработки и анализа данных

Индивидуальные особенности ритма сердца найти несложно даже при коротких записях ЭКГ [9]. Сложнее найти общие закономерности ритма сердца одного человека в разные последовательные интервалы времени. Еще более сложно найти общие динамические закономерности ритма сердца разных людей, особенно если записи ЭКГ имеют разную длительность. Сравнение временных

величин, характеризующих общую вариабельность ритма и вычисленных на основе записей различной длительности, является некорректным. Методы оценки общей вариабельности сердечного ритма и ее компонентов с коротким и длинным периодом не могут заменить друг друга [1, 3]. Поэтому при анализе ВСР возникают непреодолимые трудности при сопоставлении данных записей ЭКГ разной длины с нормой [1] для фиксированной короткой или длинной записи. Более того, определение самой нормы функционального состояния организма посредством количественных показателей ВСР становится неоднозначным.

Для решения этой проблемы предлагается перейти от анализа вариабельности ритма сердца по совокупности группы соответствующих показателей, представляющих по отдельности тот или иной информативный признак вариабельности, к функционально-параметрическому анализу их связей. При этом предлагается использовать «во благо» другой проблемный момент метода оценки ВСР: статистическую зависимость и дублирование информации разными параметрами ВСР [1–3]. Тесная корреляционная связь между параметрами ВСР нивелирует их индивидуальные зависимости от длины записи на фазовой плоскости. При этом параметр ритма сердца, выбранный общим аргументом, не должен быть ограничен теми или иными характеристиками ритмограммы. В качестве такого параметра была выбрана информационная энтропия ярусной диаграммы ритма сердца [3, 6, 10].

Существуют разные подходы к понятию «энтропия», связанные с разными объектами и задачами исследования. К наиболее известным относятся [11–14] подходы: (1) Клаузиуса, определяющий энтропию функцией состояния газовой системы при исследовании тепловых потоков; (2) Бриллюэна, определяющий энтропию мерой «деградации» энергии; (3) Пригожина, определяющий энтропию мерой «связанной энергии»; (4) Больцмана, определяющий энтропию мерой интенсивности молекулярного хаоса; (5) Шеннона, связывающий энтропию с количеством информации в информационном сообщении и определяющий ее как меру степени неопределенности состояния физической системы.

Проблема оценки количества информации, содержащегося в сообщении, была решена в 1949 г. [11]. В качестве единицы (бит) информации $I = -\log_2 p$ принимают количество информации в достоверном сообщении о событии, априорная вероятность p которого равна 1/2.

Известно [10, 15], что количество информации I_X , приобретаемое физической системой X при полном выяснении ее состояния, равно энтро-

пии $H(X)$ системы $I_X = H(X)$. Если непрерывную систему свести к дискретной, установив предел точности измерения (шаг дискретизации Δx), это будет равносильно замене плавной кривой на графике функции плотности вероятности $f(x)$ на ступенчатую — в форме гистограммы. При такой замене вероятности попадания p_i в соответствующие разряды определены в форме $p_i = f(x_i)\Delta x$. В таком случае

$$I_X = - \sum_{i=1} p_i \log_2 p_i. \quad (1)$$

При формировании диаграммы ритма сердца количество информации набирается дискретно: от систолы к систоле, поэтому принципиально невозможно использовать понятие скорости набора информации, являющееся основным параметром для технических устройств связи [15]. В этом случае масштабной единицей становится переменный интервал времени события (R – R интервал). Ритмограмма представляет собой номерной ряд последовательности n таких событий. Постоянная частота считывания монитором значений биопотенциалов при формировании ЭКГ задает постоянным шаг дискретизации Δx значений R – R интервала на ДРС. Поэтому значения R – R интервала на ДРС формируют «ярусы микросостояний». Это позволяет трактовать ДРС как реализацию макросостояния системы ритма и применить к ней структурно-топологический анализ неупорядоченности распределения значений R – R интервала по микросостояниям с использованием информационной энтропии [6, 10].

В качестве меры фрактальной размерности странных аттракторов в фазовом пространстве применяют информационную размерность D_I . Мерой непредсказуемости в системе служит информационная энтропия [10, 16–19]:

$$I(\varepsilon) = - \sum_{i=1} p_i \log_2 p_i. \quad (2)$$

Переноса определение этой меры на ярусную ДРС, получим вертикальный размер ε ячейки покрытия, равный шагу дискретизации Δx . В таком представлении категории количества информации (1) и информационной энтропии (2) становятся тождественными. Чтобы не допускать путаницы и для конкретного объекта исследования (ярусной структуры ДРС) в обозначении информационной энтропии будем использовать символ I^* .

Степень неопределенности состояния системы ритма может определяться и вероятностями (p_i) ее возможных состояний, и их количеством [15], поэтому возникает возможность перехода от вероятностных категорий к макропараметрам ярусной

ДРС. После несложных алгебраических преобразований формулы (1) и (2), примененные к ярусной структуре ДРС, можно представить в виде [6]:

$$I^* = \frac{A}{n} [\ln \Gamma + B], \quad (3)$$

где $A = 1/\ln 2$ — полиномиальный коэффициент, $\Gamma = N! / \prod_{i=1}^n N_i!$, N — число дискретных значений R – R интервала в анализируемой выборке ритмограммы, N_i — число дискретных значений R – R интервала на i -м ярусе ДРС.

При принятой точности расчета (до двух значащих цифр) величиной B можно пренебречь уже при $n > 100$, так как величина абсолютной погрешности $\Delta I_X = AB_{\max}/n$ с ростом n асимптотически стремится к нулю [6]. С учетом этого формула (3) принимает окончательный расчетный вид:

$$I^* = \frac{A}{n} \ln \Gamma. \quad (4)$$

В формуле (4) полиномиальный коэффициент Γ приобретает смысл термодинамической вероятности и определяет число микросостояний (комбинаций), посредством которых реализуется макросостояние системы X . Величина I^* , определенная с точностью до величины AB_{\max}/n , определяет среднее количество информации, недостающее до полного описания одного отсчета.

С одной стороны, информационная энтропия I^* обладает основными свойствами физической энтропии — при фиксированных внешних условиях растет с ростом n , принимая максимальное совместимое с внешними условиями значение. С другой стороны, она определена отношением количества информации I_Σ , недостающего для полного описания ДРС к объему выборки n и по смыслу является средним количеством информации, недостающим для описания одного микроперехода на ДРС [6]. По сравнению с другими параметрами variability сердечного ритма (ВСР), информационная энтропия I^* не теряет адекватного физического смысла для многомодального распределения и имеет постоянную, четко выраженную «правую границу условного здоровья» [6]. Очевидно, что разные заболевания могут дать один и тот же результат по величине I^* для ДРС. Это может означать, что разные стадии разных заболеваний подобны по результатирующему параметру I^* , т. е. по неупорядоченности ритма сердца обследуемых людей. В таком случае параметр I^* может служить количественной оценкой общего функционального состояния человека (ФСО) [6].

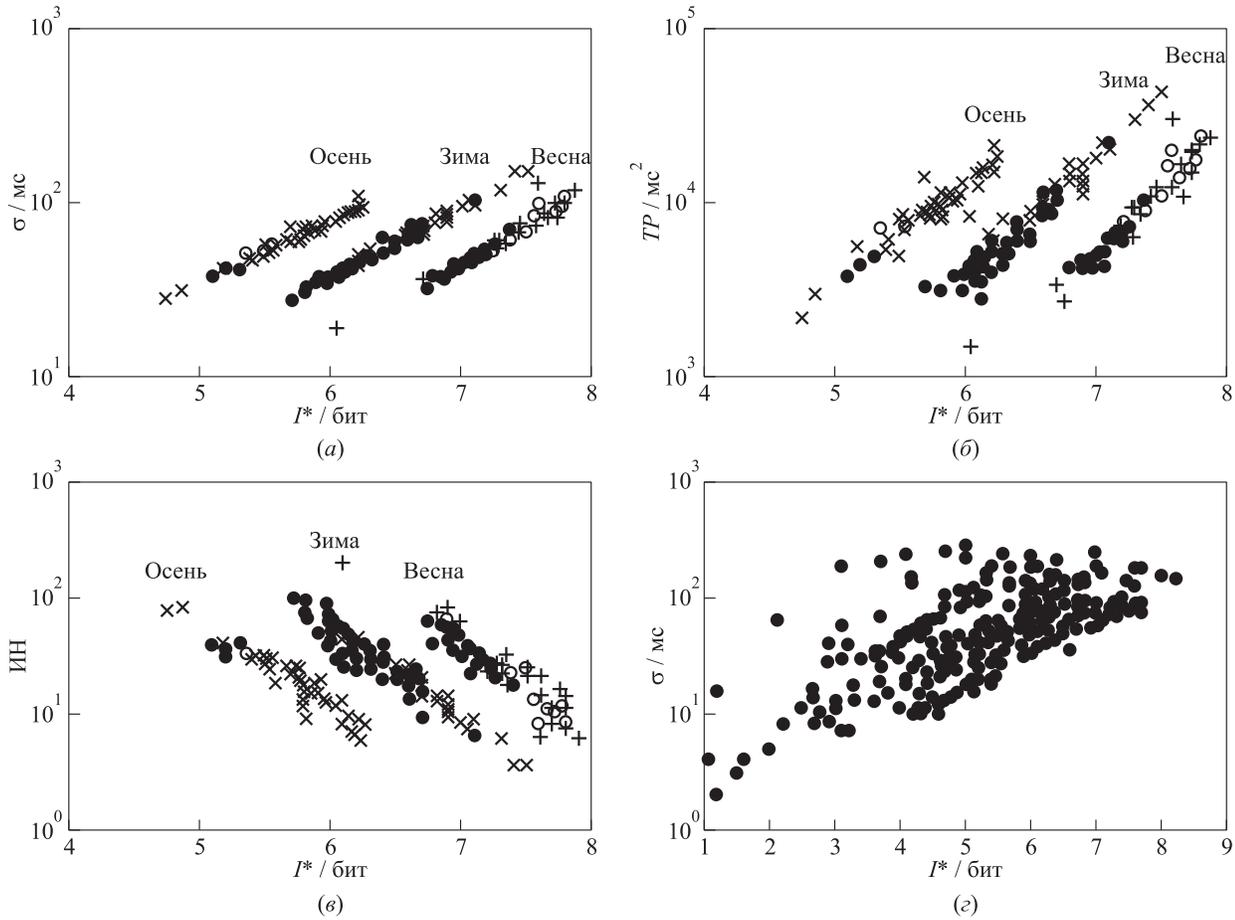


Рис. 1 Графики функциональных связей параметров ВСП и информационной энтропии для пяти серий ($N = 253$) сезонных регистраций ЭКГ здоровых молодых людей (а)–(в) и для 375 регистраций ЭКГ здоровых людей и пациентов отделений реанимации за 9 лет (г)

На рис. 1, а–г в полулогарифмическом масштабе приведены точечные графики зависимости основных расчетных выборочных параметров ВСП от соответствующих значений информационной энтропии по всем N ритмограммам. В качестве основных параметров ВСП представлены: из временной области анализа ДРС — стандартное отклонение (σ , мс), из частной области анализа — полная спектральная мощность (*Total Power*, или TP , мс²), из набора производных показателей Баевского — индекс напряжения (ИН), или стресс-индекс (SI), характеризующий степень централизации управления ритмом [20].

Графики на рис. 1, а–в представлены условно-сезонными линиями одинакового наклона. Это указывает на наличие прочных функциональных связей между параметрами ВСП и информационной энтропией в условиях дискретной сезонной адаптации. В зависимости от уровня ФСО в интервалах времени условно-сезонного исследования точки на

графиках соответствующих параметров перемещаются вдоль линий функциональных кривых «как по монорельсу». При смене сезона заполняются новые «функциональные уровни», соответствующие обретению ритмом качественно новых стационарных режимов (рис. 1, а–г). Данные юношей (\times , $+$) и девушек (\bullet , \circ) для каждого условно-сезонного интервала времени принадлежат соответствующему «функциональному уровню» (см. рис. 1, а–в) без разделения по полу. На основании этого результата при одинаковом возрасте здоровых молодых людей, обследуемых и в группе, и серийно, оказалось возможным предположить, что ритм сердца меняет режим дискретно при изменении сезонных внешних условий. Механизм адаптации является стабилизирующим каждое новое качество ритма сердца.

При изменении длины записи функциональные кривые лишь меняют свою длину изменением координат правой или левой границы. Ве-

личина $I^* = 6$ бит характеризует режимы ритма здорового молодого человека на всех ступенях сезонной адаптации при двадцатиминутной записи ЭКГ. Обращает внимание, что определенная ранее нелинейная динамика выборочных коэффициента асимметрии и эксцесса распределений на ДРС при серийных регистрациях [6] не сказывается на форме графиков $\sigma(I^*, N)$ в интервале времени любого условно-сезонного исследования.

При серийных и групповых исследованиях обнаружен нелинейный характер связи стандартного отклонения σ и информационной энтропии I^* со средне-выборочным значением $R-R$ интервала ($\langle X \rangle$) и моды (Mo) ритмограмм. Внешне форма функций $\langle X \rangle(I^*, N)$ и $\langle X \rangle(\sigma, N)$ на соответствующих графиках напоминает «полет мухи под люстрой». При внутривыборочных исследованиях ритмограмм, проведенных с использованием «метода скользящих средних» [6], оказалось, что на малые флуктуации среднего уровня ритма сердца ни σ , ни I^* практически не откликаются. Если средний уровень ритма на некотором интервале времени ($\Delta n < 100$) постоянен, то и указанные параметры не меняются. Однако если постоянство $\langle X \rangle(n)$ во времени затягивается, то величины обоих параметров начинают медленно монотонно падать. На любое относительно резкое и длительное изменение $\langle X \rangle$ функция $\sigma(n)$ откликается импульсно таким образом, что σ_{\max} всегда приходится на точку перегиба графика $\langle X \rangle(n)$. В таком случае $\sigma(n)$ определяется скоростью изменения функции $\langle X \rangle(n)$. При монотонном росте $\langle X \rangle$ значения I^* и σ обретают тенденцию к росту за счет набора нерабочих («пустых») ярусов. При этом процентный состав рабочих ярусов падает, а неупорядоченность ДРС слабо растет. Таким образом, функция $I^*(n)$ «следит» за средним уровнем $\langle X \rangle(n)$ комплексно: за величиной, за длительностью его характерных динамических участков и за скоростью изменения.

4 Формулы функционального состояния организма

Подавляющее количество моделей и приемов исследования временных рядов относится к стационарным в широком смысле рядам, т. е. к рядам, для которых первые четыре момента не зависят от времени [21, 22]. В общем случае даже в интервале одной двадцатиминутной регистрации ЭКГ не удастся исключить тренд в дисперсии, асимметрии и эксцессе разностным дифференцированием рядов ДРС [6, 21]. Таким образом, реальные процессы ритма сердца не являются стационарными. Для

того чтобы в первом приближении для задач краткосрочного прогноза считать их таковыми, необходимо создать определенные экспериментальные условия, а именно в состав анализа включать выборки только молодых здоровых людей, находящихся в состоянии стабильного психоэмоционального покоя. В рамках такого приближения к стационарным рядам можно применить теорему Вольда о разложении [23], согласно которой всякий стационарный процесс может быть единственным образом представлен в виде суммы двух не коррелирующих между собой процессов: детерминированного (сингулярного) и случайного (регулярного белого шума).

Если принять в качестве гипотезы утверждение, что ритм сердца является нестационарным процессом в той мере, в которой в него включены составляющие внешнего влияния механизмов управления и регуляции [2, 6], то стационарным ритмическим процессом может характеризоваться состояние «неуправляемого сердца». Обычно этот физиологический термин применяется для описания работы изолированного от организма сердца с перфузией. В применении к описанию работы сердца, изолированного только от детерминистской информации внешнего влияния, жесткий ритм, обеспечиваемый функцией автоматии, дополняется на заданном уровне регулярным белым шумом, определяющим лишь присутствие и функциональную готовность разных механизмов управления и регуляции. Приближением к такому режиму ритма может быть ритм сердца молодого здорового человека, находящегося в условиях адаптации к условиям регистрации и в состоянии устойчивого психоэмоционального покоя. При этом распределение $R-R$ интервалов на ДРС приближается к нормальному распределению [6, 24–28]. Крайней идеализацией такого режима ритма может служить «виртуальный ритм» с реализацией в форме цифрового ряда, полученного генерацией случайных чисел по нормальному закону и по заданным значениям стандартного отклонения и шага дискретизации [6, 28]. Для построения виртуальной диаграммы ритма сердца используется нормальный генератор случайных чисел в программе Excel. Оказалось, что величина информационной энтропии I^*_Γ виртуальной ДРС (ВДРС) при $n \rightarrow \infty$ монотонно приближается к значению энтропии

$$H(X) = \log_2 \frac{(2\pi e)^{1/2} \sigma}{\Delta x}$$

в форме математического ожидания для непрерывного множества случайных чисел, распределенных по нормальному закону [6, 15] (рис. 2).

Очевидно, что величины энтропии $H(X)$ и информационной энтропии I^*_Γ сближаются при раз-

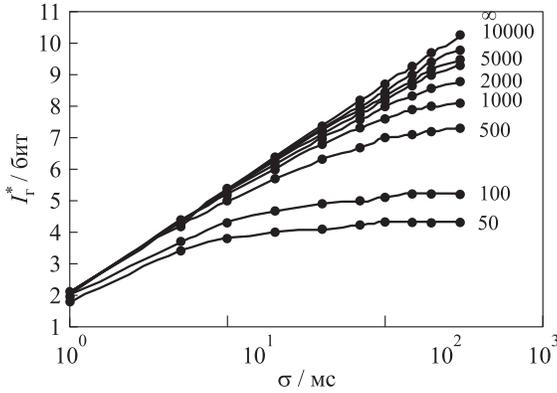


Рис. 2 Графики зависимости $I_r^*(\sigma, n)$ в сравнении с функцией $H(X)$

мерах n цифровых рядов, стремящихся к бесконечности (см. рис. 2). Если величина выборки конечна, то результаты расчета по этим двум величинам расходятся. Они расходятся тем больше, чем меньше величина n . Например, при $n = 1000$ отсчетов расхождение достигает 20%. Такое расхождение объясняется тем, что связующим звеном между двумя формами записи энтропии является формула Стирлинга для случая $n \rightarrow \infty$ [6, 29]. При уменьшении величины n расчетная погрешность этой формулы нарастает.

Уравнения трендовых линий функциональных кривых (см. рис. 1, а, з) могут быть представлены в общем виде $3\sigma = 2^{I^*+i}$ при коэффициенте достоверности аппроксимации $R^2 > 0,9$ (см. рис. 1, а) для ряда значений $i = 0, 1, 2, 3$. Эта формула, с одной стороны, иллюстрирует «правило 3σ » для ДРС в интерпретации, отличной от общепринятой [15, 23]: величина 3σ представляет полное число комбинаций (число кодонов) при переходе от «алфавита» с двумя буквами к алфавиту с k буквами ($k < 3\sigma$) и переменным размером кодона ($I^* + i$). С другой стороны, она может быть представлена в виде: $I^* = \log_2 3\sigma - i$. Такую форму записи автор назвал «формулой ФСО», так как по величине i определяются и уровень ФСО здорового человека, и степень тяжести заболевания больного (см. далее).

При сопоставлении теоретической и экспериментальной форм записи энтропии $H(X)$ и I^* величина σ оказывается одинаковой, поэтому в качестве переменных параметров могут служить Δx и i . Если в формуле математического ожидания энтропии цифрового ряда случайных величин, распределенных по нормальному закону

$$H(X) \approx \log_2 \left(\frac{4,13\sigma}{\Delta x} \right), \quad (5)$$

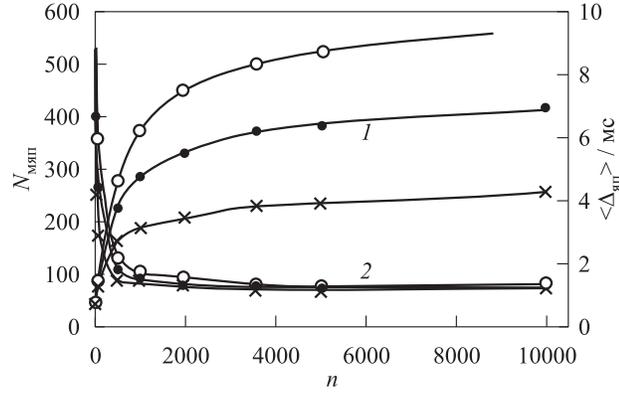


Рис. 3 Графики $N_{\text{МЯП}}(n)$ (1) и $\langle \Delta_{\text{ЯП}} \rangle (n)$ (2) виртуальных диаграмм ритма сердца при заданных $\sigma = 40$ (x), 70 (•) и 100 мс (o)

принять шаг дискретизации исходной непрерывной функции $\Delta x = 1$ мс, то

$$H(X) \approx \log_2(4,13\sigma),$$

где единица измерения $[\sigma] = 1$ мс. При $\Delta x = 2$ мс получим

$$H(X) \approx \log_2(4,13\sigma) - 1;$$

при $\Delta x = 4$ мс —

$$H(X) \approx \log_2(4,13\sigma) - 2$$

и т. д. Точка на графике $H(\sigma, \Delta x)$ сдвигается влево на единицу при изменении величины Δx кратно 2. Если $\Delta x = 0,5$ мс, точка на указанном графике (см. рис. 2) сместится на единицу вправо. При фиксированном значении σ график $H(\Delta x)$ будет линейным в полулогарифмическом масштабе.

Величина шага дискретизации $\Delta x = 1$, определенная приборной частотой, не равна выборочному среднему расстоянию между ярусами, т. е. средней величине межъярусного промежутка $\langle \Delta_{\text{ЯП}} \rangle$. Результаты анализа, проведенные для виртуальной ДРС, показали, что с ростом объема выборки n число межъярусных промежутков $N_{\text{МЯП}}$ на диаграмме монотонно растет, а величина $\langle \Delta_{\text{ЯП}} \rangle \rightarrow \Delta x$ при $n \rightarrow \infty$ (рис. 3).

Если для виртуальных цифровых рядов формуле (5) поставить в адекватное соответствие форму записи

$$I_r^* \approx \log_2 \frac{4,13\sigma}{\langle \Delta_{\text{ЯП}} \rangle}, \quad (6)$$

то при $n \rightarrow \infty$ они совпадут. Можно оценить, что $4,13\sigma / \langle \Delta_{\text{ЯП}} \rangle \approx 3\sigma / \Delta x$ (см. рис. 3), и формула (6)

описывает предельный режим ритма, информационная энтропия ДРС которого определена формулой

$$I_{\max}^* = \log_2 \frac{3\sigma}{\Delta x}$$

(см. выше). Ясно, что всегда выполняется тройное неравенство: $H(X) > I_r^* \geq I_{\max}^* \geq I^*$. Верхняя «математическая граница», с одной стороны, является идеализацией, а с другой — эталоном для ритма сердца как маркер «правой границы нормы условного здоровья».

Следуя той же логике, для реальных ДРС

$$I^* \approx \log_2 \frac{4,13\sigma}{\langle \Delta_{яп} \rangle},$$

или с учетом экспериментальных данных (см. рис. 1, а)

$$I^* \approx \log_2 \frac{3\sigma}{2^i \Delta x},$$

где $\Delta x = 1$ мс. Эти формулы связывают параметры макроструктуры ДРС (I^* и σ) с параметрами ее ярусной микроструктуры ($\langle \Delta_{яп} \rangle$ и i). При $i = 0$ данные условно соответствуют весенним, при $i = 1$ — зимним, при $i = 2$ — осенним, при $i = 3$ и выше — болезни. Для людей одного возраста получается возможность организации шкалы ФСО по функции $\langle \Delta_{яп} \rangle(N)$ или по величине i как по группе, так и по серии N опытов. Предлагается величину i определять показателем ФСО: $i = 0$ — норма, $i = 1$ — обратимое угнетение в рамках сезонной адаптации, $i = 2$ — обратимое донозологическое состояние в рамках сезонной адаптации, $i = 3$ — необратимое состояние вне рамок сезонной адаптации (патогенез).

Уравнением, связывающим среднюю величину микроперехода на ДРС и информационную энтропию ДРС, может служить $\langle \Delta_{яп} \rangle = -2I^* + (13 \div 17)$ (рис. 4). Для здорового молодого человека гради-

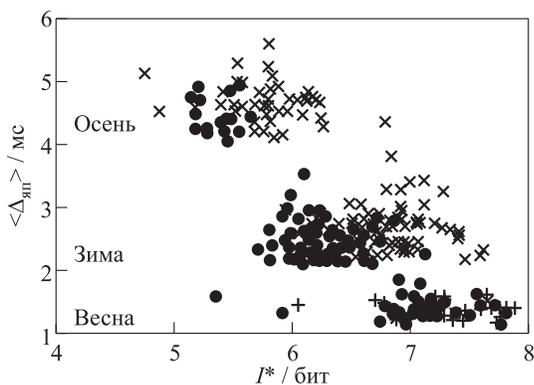


Рис. 4 Графики $\langle \Delta_{яп} \rangle(I^*, N)$ по данным регистраций ЭКГ здоровых людей

ент средней величины микроперехода направлен в сторону убывания I^* и определен величиной — 2 мс/бит. При $\langle \Delta_{яп} \rangle \rightarrow (\Delta x = 1)$ показатель ФСО $i = 0$ и величине I^* разрешено варьировать в пределах 6–8 бит; при $\langle \Delta_{яп} \rangle \rightarrow 2$ ($i = 1$) — $I^* = 5,5$ –7,5 бит; при $\langle \Delta_{яп} \rangle \rightarrow 4$ ($i = 2$) $I^* = 4,5$ –6,5 бит. Переход «весна–осень» соответствует дискретному четырехкратному увеличению величины $\langle \Delta_{яп} \rangle$ (см. рис. 4).

5 Относительная информационная энтропия

Оба энтропийных параметра: I^* для ДРС и I_r^* для виртуальной ДРС — зависят от n [3, 6]. Это доставляет неудобства при сравнении цифровых рядов разной длины. Следовательно, необходим параметр, который бы сохранял информацию о неупорядоченности ярусной структуры ДРС и не зависел от n .

При сравнении ДРС с ВДРС информационная энтропия I_r^* имеет ту же функциональную погрешность по n , что и I^* для реальной ДРС [6]. Поэтому их отношение $i_r = I/I_r^*$ становится свободным от функциональной погрешности, связанной с конечностью числа измерений. Оба энтропийных параметра I^* и I_r^* зависят и от n , и от σ , но в первом из них содержится информация о функции $\sigma(n)$, т. е. о детерминистской составляющей сигнала. Поэтому их отношение I^*/I_r^* , оценивая это влияние, является мерой неупорядоченности ярусной структуры ДРС по отношению к максимально возможной неупорядоченности, ограниченной задаваемыми параметрами, сводя влияние объема выборки к пренебрежимо малому.

Таким образом, относительную информационную энтропию (I^*/I_r^*) можно использовать как индикатор уровня ФСО по фактору регуляции ритма в смысле отклонения от нормального закона распределения. При этом можно исходить из утверждения, что отклонение гомеостатической функции от стационарного уровня вызвано внешним управляющим влиянием и в любом проявлении приводит к появлению и интенсификации работы механизмов регуляции. Механизмы адаптации тормозят процессы отклонения, а механизмы регуляции возвращают его в норму. Механизмы контроля удерживают его около нормы в разрешенных пределах флуктуаций. Например, для состояния «весна» механизмы контроля удерживают отношение I^*/I_r^* около индивидуального значения, близкого к единице (или 100%) и должны быть определены процессами непрерывного действия. По величине это-

Таблица 1 Показатели ФСО по всем регистрациям

Время Серия или группа	$N_{\text{рег}}$	ЧСС, уд./мин	Рабочие ярусы, %	$\langle \Delta_{\text{яп}} \rangle$, мс	I^*/I_{Γ}^* , %	Формула ФСО (см. рис. 1, $a-z$)
Декабрь—январь 2008 г. К. (21 г.) Ш. (21 г.)	34 48	74 ± 2 75 ± 2	$34,4 \pm 2,7$ $40,4 \pm 1,6$	$2,5 \pm 0,1$ $3,2 \pm 0,3$	$89,8 \pm 0,7$ $87,2 \pm 1,8$	$I^* = \log_2 3\sigma - 1$ $I^* = \log_2 3\sigma - 1$
Январь—февраль 2008 г. Р. (21 г.)	45	63 ± 2	$36,7 \pm 1,4$	$2,8 \pm 0,1$	$92,1 \pm 0,6$	$I^* = \log_2 3\sigma - 1$
Февраль—март 2008 г. Гр. 32 (19–24 г.) Юноши Девушки	32 20 12	81 ± 5 82 ± 7 79 ± 5	$72,6 \pm 2,6$ $72,3 \pm 3,3$ $73,2 \pm 4,2$	$1,4 \pm 0,1$ $1,4 \pm 0,1$	$97,9 \pm 2$ $99,3 \pm 0,3$ $95,5 \pm 4,5$	$I^* = \log_2 3\sigma$ $I^* = \log_2 3\sigma$ $I^* = \log_2 3\sigma$
Октябрь—ноябрь 2008 г. К. (22 г.) Ш. (22 г.)	34 33	68 ± 3 81 ± 2	$21,9 \pm 1,3$ $41,7 \pm 1,4$	$4,6 \pm 0,2$ $2,4 \pm 0,1$	$79,0 \pm 1,1$ $88,5 \pm 0,4$	$I^* = \log_2 3\sigma - 2$ $I^* = \log_2 3\sigma - 1$
Апрель—май 2009 г. Ш. (22 г.)	27	79 ± 2	$73,2 \pm 2,6$	$1,4 \pm 0,1$	$99,4 \pm 0,2$	$I^* = \log_2 3\sigma$
1999–2009 гг. Группа 330 (17–75 лет)	375	—	2–75	(1,4) (2,8) (4,2) (5,6)	20–99	$I^* = \log_2 3\sigma$ $I^* = \log_2 3\sigma - 1$ $I^* = \log_2 3\sigma - 2$ $I^* = \log_2 3\sigma - 3$...

() — предполагаемые расчетные значения сертификации больных по ФСО.

го отношения и его динамике человек принимает решение: переходить к нозологическим процедурам или в этом нет необходимости. Иными словами, величина I^*/I_{Γ}^* и ее динамика позволяют сопоставить субъективные оценки состояния человека с относительным уровнем хаотичности ритма сердца.

В табл. 1 представлены данные по сериям регистраций ЭКГ обследуемых К., Ш., Р. и группы здоровых молодых людей (32 человека), а также данные массовых нерегулярных во времени регистраций по группе из 330 человек. В таблице приведены: число регистраций $N_{\text{рег}}$ и расчетные средние значения частоты сердечных сокращений (ЧСС), относительного количества рабочих ярусов, величины межъярусного промежутка и относительной информационной энтропии. Все расчеты проведены с уровнем значимости $\alpha = 0,05$. В последнем столбце таблицы приведены формулы ФСО, представляющие зависимости $I^*(\sigma, N)$ отдельно по ДРС выделенных в строках серий и групп регистраций.

Данные по группе из 330 человек разного возраста (140 здоровых людей, 235 пациентов отделений реанимации за период 1999–2009 гг.) выделены в отдельную нижнюю строку таблицы. Соответствующие графики представлены ранее (см. рис. 1, z). В сравнении с графиком на рис. 1, a и данными табл. 1 для них характерными являются следующие отличия:

- (1) слабо выражены данные, определенные по формуле ФСО «весна»;
- (2) отчетливо проявляются данные, соответствующие формуле ФСО $I^* = \log_2 3\sigma - 3$;
- (3) заметно рассеяние данных, связанное в основном с большим расхождением возраста обследуемых.

Визуальный анализ графиков (см. рис. 1, $a-v$) и расчетных данных табл. 1 приводит к следующим общим результатам:

1. В течение одного года режим ритма меняется скачкообразно трижды.
2. С позиции $(I^*/I_{\Gamma}^*)_{\text{min}}$ условный год начинается в начале октября. Первый триместр, «осень» (октябрь—декабрь), и второй триместр, «зима» (декабрь—февраль), характеризуются относительным постоянством индивидуальных величин I^*/I_{Γ}^* , но заканчиваются их скачкообразным ростом, характеризующим изменение качества режима ритма и структуры ДРС смещением в сторону превалирования хаотической составляющей. Третий триместр, «весна» (март—сентябрь), характеризуется слабым рассеянием величины I^*/I_{Γ}^* около единицы. В ритме сердца превалирует хаотическая составляющая со слабыми признаками проявления механизмов управления и регуляции.

3. В начале октября режим ритма сердца столь значительно и скачкообразно меняет качество, что переходный процесс претендует на категорию «катастрофы» [30].

Предварительные исследования показали, что вне зависимости от того, болеет человек или нет, с возрастом I^* падает и все дальше отходит от эталонного значения. Отношение I^*/I_{Γ}^* с возрастом уменьшается, но не монотонно, а ступенчато, как бы задерживаясь на определенных функциональных уровнях, которые могут служить индикаторами биологического возраста человека. По колебанию величины этого отношения около индивидуального уровня соответствующего возраста можно определить уровень ФСО, поэтому данное отношение предлагается в качестве индикатора донозологической диагностики вне зависимости от возраста человека.

6 Выводы

1. Между параметрами ВСП и информационной энтропией ДРС в условиях дискретной сезонной адаптации существуют прочные функциональные связи. При смене сезона заполняются новые «функциональные уровни», соответствующие обретению ритмом качественно новых стационарных режимов.
2. При изменении длины записи ЭКГ графики функциональных кривых сохраняются, меняя длину изменением координат правой или левой границы.
3. Сравнительный анализ реальной ДРС и соответствующей ей виртуальной ДРС позволяет оценивать влияние системы регуляции на ритм сердца в форме отклонения распределения значений $R-R$ интервалов на ДРС от нормального закона.
4. Всегда выполняется тройное неравенство: $H(X) > I_{\Gamma}^* \geq I_{\max}^* \geq I^*$. Верхняя «математическая граница», с одной стороны, является идеализацией, а с другой — эталоном для ритма сердца как маркер «правой границы нормы условного здоровья».
5. Формулы ФСО адекватно связывают параметры макроструктуры ДРС (I^* и σ) с параметрами ее ярусной микроструктуры ($\langle \Delta_{\text{яп}} \rangle$, Δx и i).
6. В формуле ФСО $I^* \approx \log_2[(3\sigma)/(2^i \Delta x)]$ величину i предлагается определять показателем ФСО: $i = 0$ — норма, $i = 1$ — обратимое угнетение в рамках сезонной адаптации, $i = 2$ — обратимое донозологическое состояние в рамках сезонной адаптации, $i = 3$ — необратимое состояние вне рамок сезонной адаптации (патогенез).
7. Уравнением, связывающим среднюю величину микроперехода и информационную энтропию ДРС, может служить $\langle \Delta_{\text{яп}} \rangle = -2I^* + (13 \div 17)$.
8. Относительную информационную энтропию (I^*/I_{Γ}^*) можно использовать как индикатор уровня функционального состояния организма по фактору регуляции ритма. Режим ритма сердца здорового молодого человека вне зависимости от пола в течение календарного года трижды дискретно меняет свое качество: от $(I^*/I_{\Gamma}^*)_{\min} \approx 0,8$ в интервале октябрь—ноябрь к $(I^*/I_{\Gamma}^*) \approx 0,9$ в интервале декабрь—февраль и до $I^*/I_{\Gamma}^* \leq 1$ в интервале март—сентябрь. Замыкает годовой цикл изменений режима ритма наиболее резкое изменение качества при возврате к значению $(I^*/I_{\Gamma}^*)_{\min} \approx 0,8$ в начале октября.
9. Обнаружено сильное и направленное влияние возраста человека на качество режима ритма сердца. Отношение I^*/I_{Γ}^* с возрастом уменьшается, но не монотонно, а ступенчато, как бы задерживаясь на определенных функциональных уровнях, которые могут служить индикаторами биологического возраста человека.

Литература

1. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Task Force of The European Society of Cardiology and The North American Society of Pacing and Electrophysiology // *European Heart J.*, 1996. Vol. 17. P. 354–381.
2. *Амиров Н. Б., Чухнин Е. В.* Применение метода изучения вариабельности сердечного ритма при различных состояниях (Обзор литературы) // *Диагностика и лечение нарушений регуляции сердечно-сосудистой системы.* — М.: ГКГ МВД России, 2008. С. 63–75.
3. *Кузнецов А. А.* Методы анализа и обработки электрокардиографических сигналов: Новые подходы к выделению информации. — Владимир: ВлГУ, 2008. — 140 с.
4. *Малиновский Л. Г.* Классификация объектов средствами дискриминантного анализа. — М.: Наука, 1979. — 260 с.
5. *Зозуля Е. П.* Геометрический анализ нелинейных хаотических колебаний кардиоритма как новый метод для автоматического обнаружения фибрилляции предсердий // *Физика и радиоэлектроника в медицине и экологии.* Кн. 1. — Владимир—Суздаль: ВлГУ, 2008. С. 172–175.
6. *Кузнецов А. А.* Энтропия ритма сердца. — Владимир: ВлГУ, 2009. — 172 с.

7. Прилуцкий Д. А., Кузнецов А. А., Плеханов А. А., Чепенко В. В. Накопитель ЭКГ «AnnA Flash 2000» // Методы и средства измерений физических величин. — Н. Новгород: НГТУ, 2006. С. 31.
8. Medical Computer Systems, Zelenograd, Moscow. <http://www.mks.ru>.
9. Кушаковский М. С., Журавлева Н. Б. Аритмии и блокады сердца (атлас электрокардиограмм). — Л.: Медицина, 1981. — 340 с.
10. Мун Ф. Хаотические колебания: Вводный курс для научных сотрудников и инженеров / Пер. с англ. Ю. А. Данилова и А. М. Шукурова. — М.: Мир, 1990. — 312 с.
11. Shannon C. E., Weaver W. The mathematical theory of communication. — Urbana, IL: The University of Illinois Press, 1949.
12. Матвеев А. Н. Молекулярная физика: Учеб. пособие для физ. спец. вузов. — М.: Высшая школа, 1987. — 360 с.
13. Биофизика: Учебник / Под ред. акад. П. Г. Костюка. — Киев: Выща школа, 1988. — 504 с.
14. Блюменфельд Л. А. Информация, термодинамика и конструкция биологических систем // СОЖ, 1996. № 7. С. 88–92.
15. Венциель Е. С. Теория вероятностей: Учебник для вузов. — М.: Высшая школа, 1999. — 576 с.
16. Shaw R. Strange attractors, chaotic behavior and information flow // Z. Naturforsch., 1981. Vol. A36. P. 80–112.
17. Farmer J. D., Ott E., Yorke J. A. The dimension of chaotic attractors // Physica, 1983. Vol. 7D. P. 153–170.
18. Grassberger P., Procaccia I. Characterization of strange attractors // Phys. Rev. Lett., 1983. Vol. 50. P. 346–349.
19. Кузнецов А. А. Энтропия, количество информации и информационная размерность *RR*-интервалограммы // Биомедицинские технологии и радиоэлектроника, 2008. № 6. С. 15–19.
20. Баевский Р. М., Берсенева А. П. Введение в донозологическую диагностику. — М.: Слово, 2008. — 176 с.
21. Кобзарь А. И. Прикладная математическая статистика. — М.: Физматлит, 2006. — 816 с.
22. Орлов Ю. Н., Осминин К. П. Методика определения оптимального объема выборки для прогнозирования нестационарного временного ряда // Информационные технологии и вычислительные системы, 2008. № 3. С. 3–13.
23. Королюк В. С., Портенко Н. И., Скороход А. В., Турбин А. Ф. Справочник по теории вероятностей и математической статистике. — М.: Наука, 1985. — 640 с.
24. Babloyantz A., Destexhe A. Is the normal heart a periodic oscillator // Biol. Cybern., 1988. Vol. 58. P. 203.
25. Pool R. Is it healthy to be chaotic? // Science, 1989. Vol. 243. P. 604.
26. Эйдукайтис А., Варонецкас Г., Жемайтите Д. Применение теории хаоса для анализа сердечного ритма в различных стадиях сна у здоровых лиц // Физиология человека, 2004. Т. 30. № 5. С. 56–61.
27. Кузнецов А. А. Проверка возможности применения функциональных уравнений для оценки состава и распределения ритма сердца // Биомедицинские технологии и радиоэлектроника. — М.: Радиотехника, 2008. № 3. С. 17–20.
28. Кузнецов А. А. Структурно-топологические особенности диаграмм ритма сердца // Инфокоммуникационные технологии, 2009. Т. 7. № 3. С. 80–85.
29. Математическая энциклопедия / Гл. ред. И. М. Виноградов. Т. 5. — М.: Сов. энциклопедия, 1984. — 1248 с.
30. Арнольд В. И. Теория катастроф. — М.: Наука, 1990. — 128 с.

SEMIFORMAL VERIFICATION FOR PIPELINED DIGITAL DESIGNS BASED ON ALGORITHMIC STATE MACHINES*

S. Baranov¹, S. Frenkel², and V. Zakharov³

Abstract: The method and tools to use Algorithmic State Machine (ASM) formalism for high-level specification of complex digital designs and their Model-Checking-based verification are described. This methodology is based on the possibility of hierarchical description of the target digital designs at algorithmic level of abstraction, and ability to generate finite state machines (FSM) models of the systems from the ASM flowcharts. The software tool was developed for automatic generation of SMV (Symbolic Modes Verifier) codes from the ASMs and corresponding FSMs. A way of this approach application to design verification is demonstrated for a pipelined microprocessor.

Keywords: formal verification; Model Checking; finite state machines

1 Introduction

Standards for safety and security increasingly demand the use of formal development methods and automated verification. To apply the formal methods in a systematic way, the nature of the system has to be taken into account. For example, one of the key optimizations used in hardware systems is pipelining, the idea being that the functionality of a system is partitioned into several stages, which can work in parallel resulting in an increase in the throughput of the pipelined implementation of the system.

Existing techniques for verifying of pipelined machines, models that describe the pipelined behavior of hardware systems, effort and resources, paired with the cost of software tools supporting this task lead to an activity that is mostly ad-hoc, labor intensive, and accessible only to a few verification specialists (with PhD degree, usually) within a design team.

In the past few decades, much research has been dedicated to improving the quality and the effectiveness of verification algorithms, Model Checking, first of all [1]. However, much less effort has been devoted to support team activity in the design in order to reduce the efforts mentioned above.

In this paper, which is a continuation of the authors' work [2] a possibility to use ASM [3] as a semiformal specification of pipelined systems is shown, which can provide essential economy of human resources in the design verification process. This semiformality is based on Model Checking with support by well-known SMV (Symbolic Model Verifier) Cadence tool.

2 Verification Techniques

Key point for designers involved in the activity of design verification is to choose appropriate verification algorithms to verify a design relatively properties, which would guarantee the lack of well-known possible hazards in systems. Finally, these problems are affecting the state space explosion problem.

Among the techniques and methodologies available for functional verification, simulation-based verification is prevalent in the industry because of its linear and predictable complexity [4]. The common methodology in this context is *random simulation*, where stimuli are provided by a constraint-based random generator. Such generators can automatically produce random legal input for the design at a very high rate, based on a set of rules (or constraints) derived from the specification document. In random simulation, bugs are detected by means of assertion statements, or checkers, embedded in the design. However, because of the randomized nature of this methodology, simulations cannot provide guarantees on the worst case behavior of a given system as they try to generalize the behavior of the system from a few observed test cases.

Besides, real-life Systems-on-Chip (SoCs) are usually highly nondeterministic; therefore, they cannot be efficiently verified by simulations.

To overcome the limitations of simulation, the industry is starting to use formal and semiformal methods [1, 4, 5], often in conjunction with simulation-based techniques. Verification approaches based on formal methods essentially provide a mathematical proof of

*The research was partially supported by RFBR (project No. 09-07-12032) and RAS Presidium program "Fundamental Problems of the System Programming (project 1.1)."

¹Holon Institute of Technology, Holon, Israel, samary@012.net.il

²Institute of Informatics Problems RAS, fsergei@mail.ru

³Institute of Informatics Problems RAS, VZakharov@ipiran.ru

correctness for the system and are therefore very exhaustive. For example, Intel's used of formal verification on a large-scale was during the Pentium 4 design cycle and consisted of about 60 person years [5]. Formal methods were used to verify that the design satisfied various properties describing the expected behavior of the microprocessor and, to date, no bugs have been discovered in the parts of the design that were formally verified.

Further, only formal and semiformal approaches will be shortly outlined as a principal topic of this paper.

2.1 Formal verification summary

Let us consider the most popular formal verification algorithm via Model Checking [1]. In general, a *model checking problem* is a problem of checking whether a given model, describing a design verified, satisfies a given property:

$$M \models \phi$$

which means that a property ϕ is satisfied for the model M . The properties are expressed in terms of "atomic propositions" (i. e., variables in the design), Boolean connectives (AND, OR, NOT, IMPLY, etc.), and so-called "Temporal operators."

The model M is a *finite state transition system*, and its properties which must be true for an *implementation* of the system at the lower level of the system design (e. g., by a *structural* FSM, that is, with binary coding of its inputs, states, outputs [1]) are expressed in a *propositional temporal logic* like *Linear temporal logic* (LTL) and *Computational Tree Logic* (CTL) [1]. Linear temporal logic is a temporal logic which can describe logical formulae about the future of paths such as that a condition will eventually be true, that a condition will be true until another fact becomes true. Computational tree logic is a branching-time logic, meaning that its model of time is a tree-like structure in which the future is not determined. That is, there are different paths in the future, any one of which might be an actual path.

Temporal Operators:

- Xp "Formula p holds next time;"
- Fp "Formula p holds sometime in the future;"
- Gp "Formula p holds globally in the future;"
- $p Uq$ "Formula p holds until q holds;"
- $p R q$ "Formula q is released when Formula p becomes false."

Temporal connectives according to their character can occur as single symbols such as X , F , G , U (linear time temporal logic LTL) or as a pair of symbols where each symbol of X , F , G , U must be preceded by path quantifiers an A or an E (branching time temporal logic CTL), similar to well-known quantifiers of predicate

calculus, that is \forall ("All quantor") and \exists ("Existense quantor"). Quantifier $A\langle PF \rangle$ means that for all paths leading from the current state, the path formula PF holds. For example, AXg means that state formula g is true in the next state of all paths leading from the current state. $E\langle PF \rangle$ means that for some of the paths leading from the current state, the path formula PF holds. For example, EFg means that there exists a possible timeline where the formula ("statement") g will eventually become true.

Various temporal operators allow us to relate properties of the current state of a given digital system model with the properties of succeeding states of the model.

Both CTL and LTL are interpreted at the states of a *transition system*, which constitute different paths s_0, s_1, \dots

By this mean, a design is modeled as a state transition graph, called a *Kripke structure* [1], which is a four-tuple model $K = (S, S_0, R, L)$. S is a finite set of states. S_0 is a set of initial states, where $S_0 \in S$, $R : S \times S$ is a transition relation between states, where for every state $s \in S$, there is a state $s' \in S$ such that for the state transition $(s, s') \in R$, $L : S \subseteq 2^{AP}$ is the labeling function to mark each state with a set of atomic propositions (AP) that hold in that state. It is basically a graph whose nodes represent the reachable states of the system and whose edges represent state transitions. A labeling function is used to assign a set of properties to each state.

If the transition system is described as a FSM, as a Mealy automaton M , in particular, $M = \langle S, I, O, \delta, \lambda \rangle$, where S and I are the state space and the input space; O is the output space, δ is the state transition function, and λ is the output function, then the Kripke structure, corresponding to the FSM M can be described as following [6]:

$$\left. \begin{aligned} &(\langle s, i \rangle, \langle s', i' \rangle) \in R \text{ iff } \delta(s, i) = s'; \\ &L(\langle s, i \rangle) = S \cup i \cup \lambda(s, i) \end{aligned} \right\} \quad (1)$$

for any $s, s' \in S$ and any $i, i' \in I$.

In terms of the structure, the properties of a target system can be expressed as propositional temporal logic that describes sequences of transitions on the computation paths of expected design behavior.

For example, the modulo-3 counter in Fig. 1 with input $I = \{\text{rst}\}$ (corresponding to *reset* signal with active logical value 1), state space $S = \{S_0, S_1, S_2\}$ is translated to the Kripke structure shown in Fig. 2. Here, each transition $\delta(s, i) \rightarrow s'$ in the FSM corresponds to a node $\langle s_k, i \rangle$, $k = 0, 1, 2$; $i \in I$ in the Kripke structure one-by-one. $\{S_{00}, S_{10}, S_{20}, S_{01}, S_{11}, S_{21}\}$ is a list of transitions of the Kripke structure. For example, S_{10} in the Kripke structure represents the transition from S_1 to S_2 in the FSM. The structural automaton that is an implementation of this abstract one (see above) can be

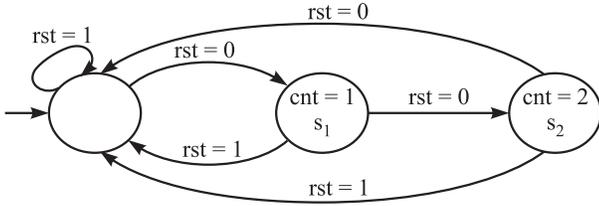


Figure 1 Automaton Mealy of modulo-3 counter

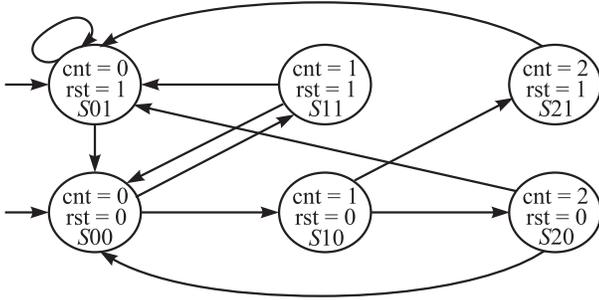


Figure 2 Modulo-3 counter Kripke structure

performed by coding of its states by vector of logical values of the counter bits $\{\text{bit}[0], \text{bit}[1]\}$, initial state $S_0 = S0$ (with the vector $\{\text{bit}[0] = \text{bit}[1] = 0\}$, state transition function δ is represented on the transition diagram in Fig. 1. The variable cnt in Fig. 1 is the number of clocks (*step of the FSM transition*).

Correspondingly, there are the following atomic propositions in this case:

- $\{p1 : (\text{cnt} = 0 \& \text{rst} = 1) \rightarrow S01\}$;
- $\{p2 : (\text{cnt} = 1 \& \text{rst} = 1) \rightarrow S11\}$;
- $\{p3 : (\text{cnt} = 2 \& \text{rst} = 1) \rightarrow S21\}$;
- etc. (total up to $p6$).

In fact, the Kripke structure nodes constitute global states of the state space of the design and contain all the propositional atoms, which are true in that particular state. The edges of the graph are oriented global state transitions.

Having this structure, desirable properties of the FSM implementation can be described. For example, if CTL is used

$$\text{AG} ((\text{bit}[0] = 0) \& (\text{bit}[1] = 0) \& (\text{rst} = 0)) \supset \supset \text{AX} ((\text{bit}[0] = 1) \& (\text{bit}[1] = 0))$$

which means that designed correctly counter with the states encoded as above must transit from the state $\{\text{bit}[0] = 0, \text{bit}[1] = 0\}$ to $\{\text{bit}[0] = 1, \text{bit}[1] = 0\}$, if there is no reset signal.

So, the Model-Checking-based formal verification is performed as following: Given a formal model

$M = (S, S0, R, L)$ of a design and a propositional temporal logic p of a property, the model checking problem is to find the set of all states in S that satisfy p , $\{s \in S | M, s \models p\}$. The design satisfies the property if all initial states are in the set. If the property does not hold for the design, error trace from the error state to an initial state is given as a counterexample that helps designers debug the error.

To achieve 100 percent confidence of correctness of the design, the specification should include all the properties that the design should satisfy.

In general, hardware verification involves comparing two descriptions of a design for “consistency.” This requires (i) the two descriptions; (ii) formal models for each (since these may or may not be the same); (iii) a formal notion of the “consistency” relation between them; and (iv) some way of checking or proving the consistency relation. Such a proof does not need necessarily directly resemble a “traditional” logic proof (e. g., it may simply be a mechanical enumeration of all possibilities).

2.2 Semiformal verification

Pure formal verification techniques, such as symbolic simulation, bounded model checking (BMC) or reachability analysis [1, 5], would be ideal to generate verification solutions. Unfortunately, they do not scale well and can only be applied to very small designs.

This problem leads to the concept of semiformal verification [5]. The idea is to combine the strengths of simulation, namely, ease of use and the ability to handle large designs with the thoroughness of formal verification along with various combinations of formal and simulation-based approaches to verification cost reduction.

In fact, there are two possibilities to perform such combination:

- (1) a “mechanical” combination of the verification techniques: part of design is verified by simulation, while another by a formal method; and
- (2) by using both formal and informal specifications.

Thus, in very general sense, Model-Checking-based semiformal verification can deal either with formal test examples generation with an informal results analysis, or with formal generation counterexamples using some results of informal simulation [7], or in using of a description with well-formalized syntax as an input of simulation. An example of the first approach is BMC [5], which is a semiformal one as it includes a choice of a trajectory length k .

Note that the lack of precise semantics for modeling notations may lead to incorrect interpretations and

misunderstands for models. So, the interpretations of model's designer may differ from model's reader. The lack of formal semantics results in the lack of analysis tools that check the consistency of the resulting models and their satisfaction to their properties. Below, it will be shown that using of ASM flowcharts as description is a good point of the semiformal verification as it allows, on one hand, formal description and, on the other hand, a combination of formal model checking and informal simulation.

3 Pipelined Microprocessor Model-Checking-Based Verification

There are a lot of papers on verifying pipelined processors (systems) with Model-Checking-based formal methods [8, 9]. Using Model Checking, one achieves an essential amount of automation but one suffers from the state explosion problem when the number of properties checked grows.

In general, the properties verified have to meet to the lack of the following hazards:

- **Data Hazard** — occurs when data needed by instruction is not yet available because it needs to be calculated by another instruction;
- **Structural Hazard** — occurs when there is a conflict over a common resource between several pipeline stages, e.g., conflict over the basic registers and operand memory address buses; and
- **Control Hazard** — hazard caused by flow control. When the processor needs to execute conditional branch, it cannot continue execution till a message whether it needs to take the branch and to where it should be jumped (or make a guess and purge the pipeline if the guess is wrong). Various methods were developed to deal with control hazards with varying success.

It is obvious that plethora of such possible conditions leads to enormous the Model-Checking exploration space.

One of the ways which could help by reducing the state space is to use so-called abstractions, refinement technique, and compositions [1, 5].

In general, an abstraction is used to build from a model M an abstract model M' such that

$$M' \models \varphi \text{ implies } M \models \varphi$$

where φ is an CTL /LTL specification formula.

Intuitively speaking, existent abstraction allows partitioning the states of a Kripke structure into some states

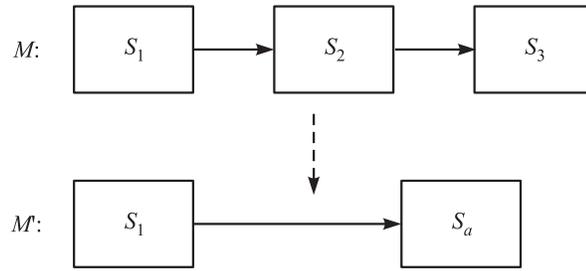


Figure 3 An abstraction $M \rightarrow M'$

agglomerations (or clusters), and treating the clusters as new abstract states. That is, the model M' contains states set S' which is a surjection $f: S \rightarrow S'$, and S' in this case is the set of abstract states [1].

For example, in Fig. 3, $S_a = S_2 \cup S_3$.

From the practical point of view, the main difficulty is that there may be variables in the model with large or unbounded ranges (such as memory addresses) and arrays with large or unbounded number of elements (such as memory arrays). These difficulties can be overcome by using abstract interpretation to reduce each data type to a small number of abstract values [9]. For example, suppose that one has a property with a parameter i ranging over memory addresses. The type A of memory addresses is reduced to a set containing two values: parameter value i , and symbol A/i representing all values other than i . In the abstract interpretation, accessing an array at location i will produce the value of that location, whereas accessing the array at A/i will produce a symbol representing an unknown value [1].

Obviously that use of these abstractions deals with a lot of manual work in order to bridge the gap in abstraction level between the implementation and the specification. The “specification” is a description of the design at more abstract level, say, as the instruction set architecture (ISA) [9], ASM, etc. As the SMV Model checking is considered as the basis of formal verification, let us outline the possibilities of reducing SMV models by the abstractions.

4 Specification Based on Algorithmic State Machine

Let us consider the ASM description as a specification of a digital design.

4.1 Algorithmic State Machines

Algorithmic state machine is a directed connected graph (Fig. 4a) containing an initial vertex (Begin), a final vertex (End), a finite set of operators (microinstructions) and conditional vertices (rhombuses) [3]. Operators

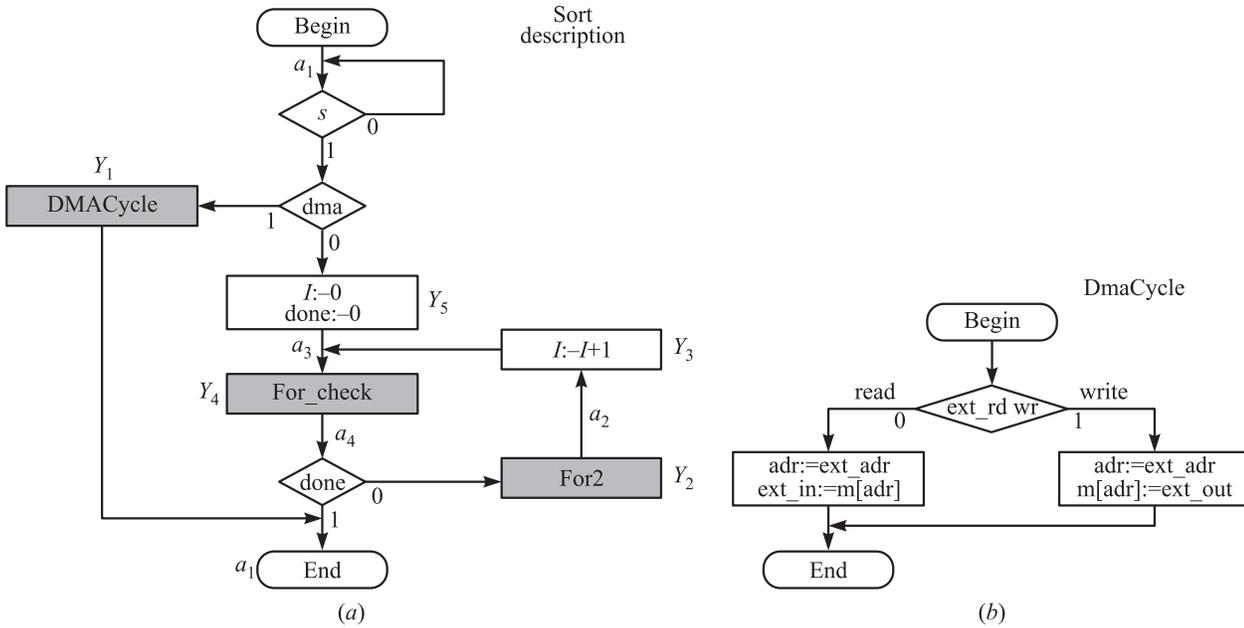


Figure 4 Example of an ASM: (a) bubble sort processor; and (b) DMACycle

and conditional vertices have only one input; the initial vertex has no input. The initial vertex and operators have only one output; a conditional vertex has two outputs marked by “1” and “0.” The final vertex has no outputs. Each operator is, in fact, a microinstruction, and it includes some body consisting of so-called microoperations. These microinstructions correspond to the instruction set architecture ISA mentioned above.

Execution of each operator vertex takes exactly one clock cycle, in accordance with true values of the Boolean variables in the rhombuses. For example, in Fig. 4a, the variables s , dma , and $done$ are the Boolean ones, which can be marked, say, as x_1 , x_2 , and x_3 .

The gray vertexes are the nested ASMs corresponding to pieces of the target design algorithm (see below).

The ASM-based design methodology is implemented in ABELITE software tool [10], which, in particular, generates from ASM a Mealy automaton which is used in a Control Unit design (synthesis), describing the sequence of the microinstruction executions and represented in the form

$$\{a_m, a_s, f_{ms}(x_{i1}, \dots, x_{ik}), Y\} \quad (2)$$

where a_m and a_s are the current and the next states of the automaton, and the transitions $a_m \rightarrow a_s$ corresponding to execution of microoperations in the vertexes (see Table 1 and some explanations after as well) labeled as components Y_j of a vector Y (index j ranged as number of the vertexes ($j = 1, \dots, 5$ for ASM in Fig. 4a)), which are considered as the automaton output variables constituted by set of the microoperations (labeled by ABELITE as y_k ,

Table 1 Finite-state machine built from ASM flowchart

a_m	a_s	$X(a_m, a_s)$	Y
a_1	a_1	$x_1 x_2$	y_1
a_1	a_3	$x_1 \sim x_2$	$y_5 y_6$
a_1	a_1	$\sim x_1$	—
a_2	a_3	1	y_3
a_3	a_4	1	y_4
a_4	a_1	x_3	—
a_4	a_2	$\sim x_3$	y_2

k are the indexes of the microoperations in the vertex) executed in the given vertex. These transitions are performed on the automaton inputs (column $X(a_m, a_s)$ in Table 1) in accordance with conjunction of all Boolean variables x_{i1}, \dots, x_{ik} , in the conditional vertexes (rhombuses), where i marks a path between some rectangles (operational vertex). These variables in the column $X(a_m, a_s)$ are ordered (with one index) by ABELITE as in the field “Logical Conditions” of the following structure:

$$\left. \begin{array}{l} \text{Microinstructions : } \langle Y_j = \{y_1, y_2, \dots, y_k\} \rangle \\ \text{Microoperations : } \langle y_j : \text{Op}_j \rangle \\ \text{Logical Conditions : } \langle x_i : R_v^i \rangle \end{array} \right\} \quad (3)$$

where Op_j are the signatures of microoperations mentioned above (see below the detailed description of Fig. 4), whereas in the set of strings “Logical Conditions” symbols x_i are Boolean variables (where indexes are ranged in accordance with number of rhombus

labels in a given ASM flowchart, 1–3 in Fig. 4a) referring the rhombuses, which correspond to some control variables R_v^i used in the target systems.

Column $Y(a_m, a_s)$ contains the $Y_j = \{y_1, y_2, \dots, \dots, y_{k_j}\}$ mentioned above.

For example, the ASM in Fig. 4a represents an algorithm of the bubble sort processor. Figure 4b is a detailization of one of its gray vertexes, the DMACycle (Direct Memory Access) block. Direct Memory Access is the special mode that does not execute an operation but connects its memory with the outside storage device to read or write information; `ext_rdw` is the signal from outside. When `ext_rdw = 1`, external system writes information to memory from its output `ext_out`. When `ext_rdw = 0`, external system reads information from memory to its input `ext_in`.

Recall that the transition table of the automata Mealy (FSM) obtained from the ASM of Fig. 4a (without any gray blocks translation) is represented in Table 1. The table represents the FSM behavior in well-known *cubic* form, that is, each vector of the third column may contain free components (which do not affect the output values). For example, one can say that the second row of Table 1 presents transitions from a_1 with any vector which is covered by cube $10x$. Here, the states marking (generated by ABELITE) can be easily seen from the Microinstruction/Microoperations structure. For example, transition to the state a_3 corresponds to execution of microinstruction (vertex) $i := 0, done = 0$.

Note that state a_1 is assigned to vertexes “Begin” and “End” (this assignment deals with some logical synthesis issues, see [3]). One can look at the condition s as a signal from the button Start-Stop; so, the instruction will be executed only when $s = 1$. There is a possibility to transit to state a_1 both from state a_1 (after execution of microoperations $i := 0, done := 0$ in the vertex Y_5) and from the state a_2 (after execution of microoperation $i := i + 1$ in the vertex Y_3).

Structures (3) formed by ABELITE for this example are as follows:

Microinstructions:

$Y1 = y1$
 $Y2 = y2$
 $Y3 = y3$
 $Y4 = y4$
 $Y5 = y5 y6$

Microoperations:

$y1$: DMACycle
 $y2$: For2
 $y3$: $i := i + 1$
 $y4$: For_Check
 $y5$: $i := 0$
 $y6$: $done := 0$

Logical Conditions:

$x1$: s
 $x2$: dma
 $x3$: $done$

As it can be seen from this example, ABELITE supports the *hierarchical* design style, that is, each of gray vertexes can be included in the automaton either as a microoperation Y_j or it can be translated with its own ASM (like DMA vertex in Fig. 4b).

4.2 Decomposition-Based Design

Decomposition-based design (Fig. 5) means a possibility to get unified model of target system from the models of each individual blocks (both on system high-level and logical level). At the system level, the models are some abstract automata and set of microoperations (y_i in the example above).

Since the principal stage of Model checking is the analysis of the Kripke structure (see Section 2), one could use the ASM flowcharts as an input of formal specification if they would have included all information needed for the Kripke structure building. Comparison of formula (1) and structure (2) shows the possibility of this mapping. Example in Table 1 shows all possible transitions from each pair of (input, state) to all reachable states.

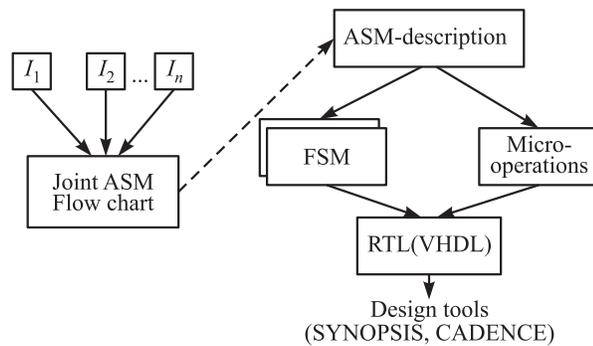


Figure 5 General scheme of hierarchical decomposition design with ABELITE where I_1, \dots, I_n are some ASMs

In other words, the Kripke structure, whose state is defined by a collection of state variables, which may be of Boolean or scalar type, can be described by Transitions Table of FSM, generated from the ASM flowchart (like Table 1), as well as by structures “Microoperation,” “Microinstruction,” and “Logical Condition” of Section 4 (structure (3)).

Also, ASM enables a convenient way to abstraction description. A possibility to use the ASM-based formalized specification is due to some formal rules, used for ASM flowchart construction [10]:

- (1) state boxes should contain only register statements, control signals in parentheses;
- (2) all operations within a state box should be concurrently executable in one clock cycle;
- (3) if the operations in two consecutive state boxes can be executed in the same clock cycle, then these two state boxes can be combined into one state box; and
- (4) for each register-transfer statement, there must be a path between the source and destination registers.

Therefore, it is possible to obtain automatically a Kripke structure from the ASM description, and the ASM description can be used in the Model-checking-based formal verification of the designs. An example of the ASM transformation in SMV code will be done in Section 6.

Below, a methodology of ASM-based semiformal verification, based on the possibility to use ASM as an input of Model Checking, is considered.

5 Model Checking Based on Algorithmic State Machine

Let us consider the verification methodology in detail. The design verification scheme is shown in Fig. 6.

The verification process starts from the informal description (specification) of the target design. The description is manually transformed into ASM flowcharts (see Section 4). The ASM formalism is supported by ABELITE tool [10], which enables rapidly prototyping hardware systems, performing design space exploration, and simplifying hardware verification. It should be noted that main purpose of ABELITE tool is high-level hardware synthesis using ASM flowchart as an input [10] but this topic is out of the scope of this paper. The ASM flowcharts can be automatically transformed into SMV code, corresponding to microarchitecture (MA) level (see Section 3) using the program ASM2 SMV devel-

oped in the Institute of Informatics Problems of the Russian Academy of Sciences (the team of Alexander Kurts and Dmitry Liburkin) [11]. ABELITE output is the automata tables (in the format of structure (2), subsection 4.1 — see example in Table 1). Some desirable temporal properties on the design are described formally by the CTL language in SMV format. Such properties are supplemented by specification consistency conditions, like absence of hazards and deadlocks between parallel operations. Then, temporal properties are verified by the SMV model checking tool [12].

If the ASM description satisfies all the specified properties, a verification specialist prepares some cycle-accurate specifications of the design (e.g., as in [13]) using semiautomated refinement from ASM flowcharts (see Section 3). Finally, the specifications are used for simulation-based verification of the design with some simulation tool.

Before receiving the SMV code, the ASM flowchart can be transformed into a cubic representation of an FSM [4] (formula (1) and structure (2)), where variables from conditional vertexes of ASM correspond to FSM input variables, and results of operators correspond to its outputs.

It should be emphasized that during compilation, all generalized operational vertexes (see Section 4) will be included in the overall ASM. Thus, a microinstruction can be either a set of microoperations or a generalized operator.

At early stages of the design process, when some operations (or parts of operations) are not clearly defined, we may informally outline some desirable actions in the vertexes using textual comments. This allows us to capture a timing structure of the target operations without specifying their functionality. On the base of such partial flowcharts, one can automatically construct various test benches for verification via simulation.

Usage of ASM for graphical representation of target design is more convenient in comparison with informal description in natural language. It allows verification engineers to present visually a timing structure of operations and their functionality (totally or partially). It should be emphasized that the ASM flowchart is not only a picture describing the target design; it has formal semantics and can be used for formal verification.

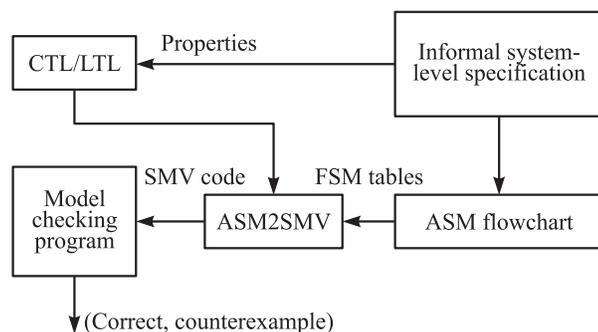
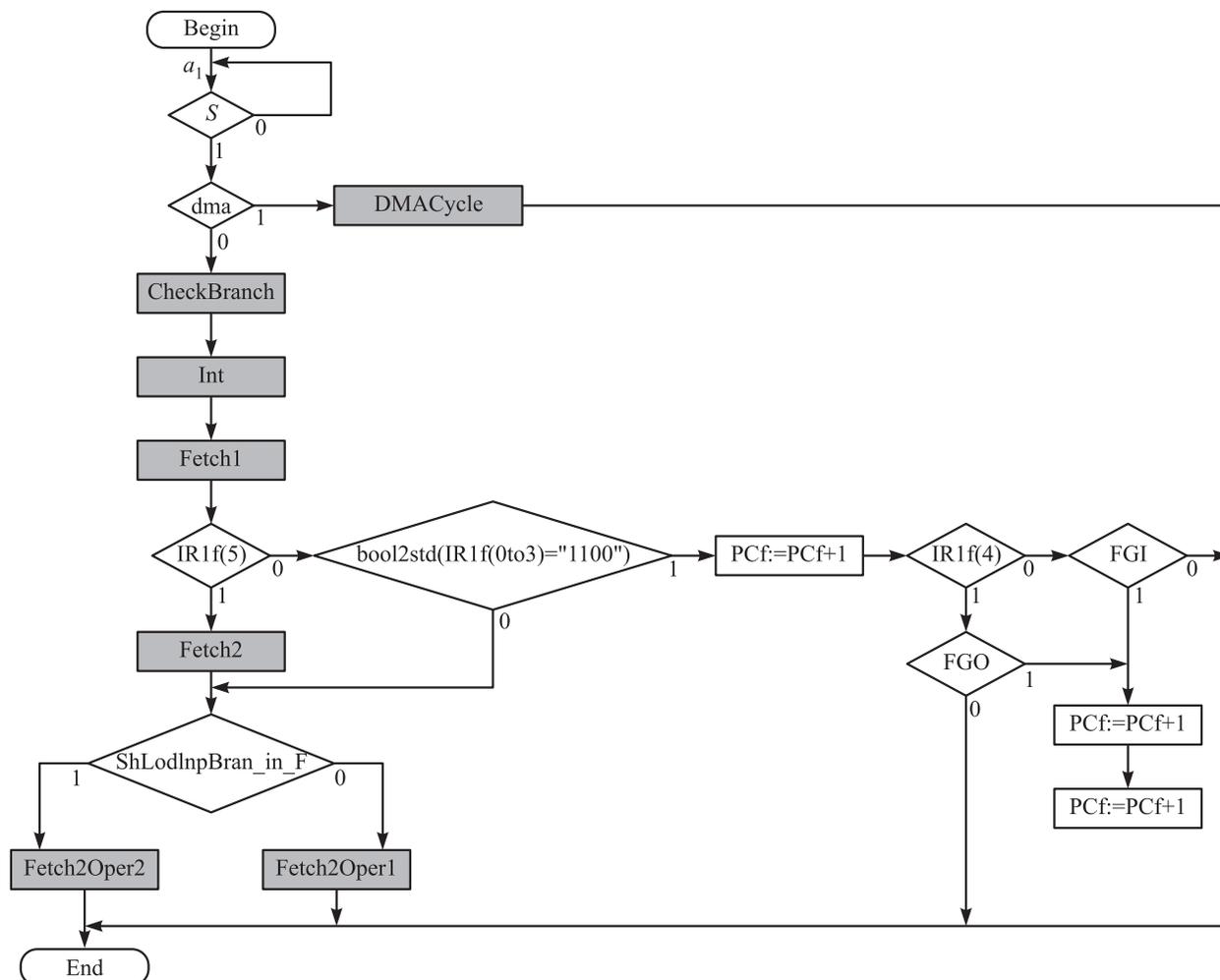


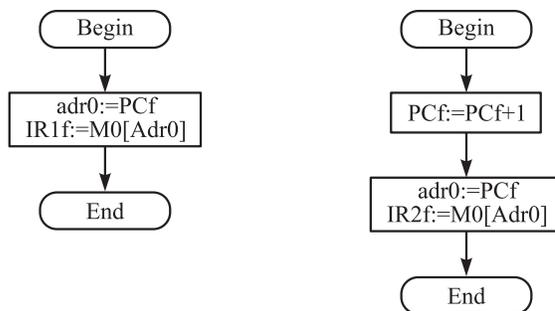
Figure 6 Semiformal design verification based on ASM

6 Pipelined Microprocessor Algorithmic-State-Machine-Based Specification and Verification

Let us consider the pipelined processor [14]. In the processor, every pipeline stage has its data path and



(a)



(b)

Figure 7 ASMM of Fetch stage (a), and (Fetch 1/ Fetch 2) blocks (b)

control unit. Instruction execution begins by fetching the instruction from program memory which address is in the program counter (PC). The instruction is then decoded to determine the operation type, the operand registers, the branch target, etc. The program counter is updated by incrementing its current value.

Let us consider the fetch stage in the pipeline (Fig. 7), which is responsible for fetching instructions from the

instruction memory. Note that the gray vertices are the nested ASMs (not expanded here).

The instruction is read into two Instruction Registers, IR1 and IR2, used for short and long instructions, correspondingly. The address of the instruction is stored in the PC, which is promoted every time when instruction is read. Program counter, IR1 and IR2 are then passed to the next stage.

The fetch stage also deals with branch execution, interrupts, and DMA which allows outside devices to read or write data directly into each of two memory blocks. The CheckBranch vertex checks whether the fetched instruction is a branch one, and in dependence of the result, produces corresponding flag. “Int” block fetches an instruction from memory, say, a memory block M0 (Fig. 7b), and if it is a short instruction (length of 16 bits), then this is indicated by zero in bit 5 of the IR1f register (suffix *f* denotes that the register copy is modeled on the Fetch stage), and the following action will correspond to the right branch of the flowchart. But if this bit is one, Fetch 2 block is activated, and the remaining bits of the instruction are written to the 16-bit register IR2f.

The Logical condition ShLodInpBran.in_F checks if the fetched instruction is one of the types “Shift,” “Load,” “Input,” or “Branch.” They require for their processing algorithm, described in the nested vertex Fetch2Oper2, another instructions are processed in the nested vertex Fetch2Oper1. Note that thanks to hierarchical properties of the ASM/ABELITE, we can abstract away all data from gray blocks when we are aimed to study properties on the high level description.

The condition variable bool2std(IR1f(0to3)) checks condition of skip of next instruction.

In distinct of more specialized systems like SystemC, ABELITE does not support directly the modeling of several simultaneous process, but possibility of reading/writing description in numerous register blocks, using different models of their interactions, as well as hierarchical description of target design by ASM allows description and verification of the pipelining stages execution. In particular, in the framework of Model-Checking verification, the idea is to show that a machine modeled at the MA level, which corresponds to microoperations set and connections of ABELITE ASM model, implements (refines) a machine modeled at the ISA level. That is, a mapping function from MA states to ISA states. The refinement map, *r*, shows how to view an MA state as an ISA state, e. g., the refinement map has to hide some MA components that do not appear in the ISA. That is, the verification problem can be considered as a checking of equivalence of ISA and MA.

In terms of formal verification, SMV model obtained from ASM with generalized blocks (hierarchical ASM) is an *implementation* of the system designed (see subsection 2.1) in comparison with the high-level ASM. States of the Pipelined machine FSM model, generated from ASM by the FSM2SMV program mentioned above contain both all visible by a programmer components and the pipeline registers, which should be represented in ASM specification.

In fact, ASM captures the structure, behavior, and mapping (between the structure and behavior) of the architecture pipelines. The structure is defined by its

components (units, storages, ports, connections) and the connectivity (pipeline and data paths) between these components. We can write CTL/LTL-based properties for each category of the hazard mentioned above.

Designers can be interested to check various properties, which show how instructions are transferred from fetch pipeline stage to Fetch2Oper1 and Fetch2Oper2 stages. Taking into account that the authors verify their design via Model-Checking methodology, that is, compare instruction level description of the target processor and its MA level FSM behavior (represented in terms of microoperations generated by ABELITE), the program ASM2SMV with fetch flowchart as input (see Fig. 4) is used and one gets the SMV model of FSM, implemented all microoperations as a result.

It is interesting, for example, to check the fulfillment of the property, described by LTL as

$$\text{assert } G((\text{ShLodInpBran.in_F} = 0) \rightarrow F(\text{state} = n))$$

where “state” is a state of FSM, generated from ASM, with a number $n = 1, 2, \dots, 7$. The specific variable “state” is assigned from the ASM, by the program, which translated ASM to SMV. In this example, state 1 corresponds to the DMA feature, that is, one checks that before the DMA is performed, fetch stage waits for all other stages to stop so there are no conflicts between the stages.

The resulted SMV code of the Fetcher is:

```

module main() {
  S : boolean;
  ShLodInpBran_in_F : boolean;
  bool2std : array 1..100 of boolean;
  to3 : boolean;
  PCf : 0..100;
  DMA : boolean;
  IR1f :array 1..100 of boolean;
  FGO : boolean;
  FGI : boolean;
  state : 1..7;

  init(state):=1;
  case {
    state=1 :
      if ((S) & !(DMA)) {
        ; /* CheckBranch */
        next(state):=7;
      }
      else if (!(S))
        next(state):=1;
      else {
        ; /* DMACycle */
        next(state):=1;
      }
  }
}

```

```

    }
state=2 :
  if (!(ShLodInpBran_in_F)) {
    ; /* Fetch2Oper1 */
    next(state):=1;
  }
  else {
    ; /* Fetch2Oper2 */
    next(state):=1;
  }
state=3 :
  if (!(IR1f[5]) &
      (bool2std[IR1f[0to3]=1100])) {
    next(state):=5;
    next(PCf):=PCf+1;
  }
  else if (!(IR1f[5]) &
           !(bool2std[IR1f[0to3]=1100]) &
           (ShLodInpBran_in_F)) {
    ; /* Fetch2Oper2 */
    next(state):=1;
  }
  else if (IR1f[5]) {
    ; /* Fetch2 */
    next(state):=2;
  }
  }
  else {
    ; /* Fetch2Oper1 */
    next(state):=1;
  }
state=4 : {
  ; /* Fetch1 */
  next(state):=3;
}
state=5 :
  if ((IR1f[4]) & !(FGO))
    next(state):=1;
  else if (!(IR1f[4]) &
           (FGI)) {
    next(state):=6;
    next(PCf):=PCf+1;
  }
  }
  else if (!(IR1f[4]) & !(FGI))
    next(state):=1;
  else {
    next(state):=6;
    next(PCf):=PCf+1;
  }
  }
state=6 : {
  next(state):=1;
  next(PCf):=PCf+1;
}
}

```



Figure 8 Window of SMV program execution

```

state=7 : {
    ; /* Int */
    next(state):=4;
}
}
assert G ((ShLodInpBran_in_F=0) ->
F (state = 1));
}

```

Let us describe briefly the states constituting the Kripke structure built from the ASM flowchart in Fig. 7. Totally 7 following states were generated: state 1 is an initial state of the fetcher, and it corresponds both to the vertexes “Begin” and “End,” execution of the vertexes Fetch2Oper1 or Fetch2Oper2 corresponds to transition from the state 2 to state 1 (end of this fetch stage), execution of the vertex Fetch1 means the transition either from state 3 to state 5 with the PC promotion (in dependence of the flags FGO/FG1 values), or execution of the Fetch2Oper1/Fetch2Oper2 functions with ending of the fetch stage, etc.

Execution of the SMV code by CADENCE SMV program shows (see the SMV code of the Fetcher above) that given property is carried out with necessity (true is indicated in the “Result” position in Fig. 8).

Here, structure Kripke is represented by SMV operators next() which correspond to transitions represented by Table 1.

Note that as it was mentioned above, the main difficulty is that there may be variables in the model with large or unbounded ranges (such as memory addresses) and arrays with a large or unbounded number of elements (such as memory arrays). Therefore, the program ASM2SMV allows definition of any type of variables during synthesis of SMV code.

7 Concluding Remarks

This paper introduces a new semiformal methodology for functional verification of pipelined hardware designs. This methodology has two specification levels: (i) ASM (upper level) and (ii) microarchitectural specifications (lower level). Algorithmic state machine is used for the graphic specification of target design. Using the ASM descriptions, one can automatically check temporal properties of the design using model checking techniques. In fact, ASM captures the structure, behavior, and mapping (between the structure and behavior) of the pipelines MA. The structure is defined by its components (units, storages, ports, connections) and the connectivity (pipeline and data-transfer paths) between these components. Usage of ASM for graphical representation of target design is more convenient in comparison with informal description in natural language. It allows verification engineers to visually present

a timing structure of operations and their functionality (totally or partially). It should be emphasized that the ASM is not only a picture describing the target design; it has formal semantics and can be used for formal verification.

References

1. Clarke E. M., Grumberg O., Peled D. Model checking. — MIT Press, 2000.
2. Baranov S., Frenkel S., Sinelnikov V., Zakharov V. Concurrent design and verification of digital hardware // Informatics and Its Applications, 2008. Vol. 2. Issue 3. P. 7–25.
3. Baranov S. Logic and system design of digital systems. — TTU Press, 2008.
4. Lee R., Tsien B. Pre-silicon verification of the Alpha 21364 microprocessor error handling system // Design Automation Conference, DAC’01, Proceedings. 2001. P. 822–827.
5. Bhadra J., Abadir M. S., Wang L.-C., Ray S. A survey of hybrid techniques for functional verification // IEEE Design & Test of Computers, March–April 2007. P. 112–122.
6. Xu X., Kimura S., Horika K., Tsuchiya T. Coverage estimation using transition perturbation for symbolic model checking in hardware verification // IEICETRANS. Fundamentals, 2006. Vol. E89-A. No. 12. P. 3451–3457.
7. Kirxshsteiger Chr., Grinshel J., Trummer Chr., Steger Chr., Weil R. Automatic test generation from semi-formal specification for function // IEEE International System Conference SysCon2008. Montreal, Canada. April 7–10, 2008.
8. McMillan K. L. Verification of an implementation of Tomasulo’s algorithm by compositional model checking // CAV98 Proceedings. 1998. P. 110–121.
9. McMillan K. L. A methodology for hardware verification using compositional model checking // Sci. Comp. Prog., 2000. Vol. 37. No. 1–3. P. 279–309.
10. Baranov S. ASMs in high level synthesis of EDA tool ABELITE // DESDes’09 Int. IFAC Workshop Proceedings. Valencia, Spain, 2009. P. 195–200.
11. Frenkel S., Kurts A., Liburkin D., Anders B., Fandjushina N. Certificate of the State Registration No. 2010612828 on the software program “Translator of the Mealy automaton table representations into programs on SMV language for the automation of the process of digital devices verification basing on Model Checking”. 26 April, 2010.
12. McMillan K. Getting started with SMV. — Cadence, 2001.
13. Frenkel S., Kamkin A. Verification methodology based on algorithmic state machine and cycle-accurate contract specifications // 7th IEEE East–West Design & Test Symposium EWDTs Proceedings. 2009. P. 39–42.
14. Kutzman I., Raitzin A. Design of fast pipelined processor with complex addressing mode. M.Sc. Thesis. Bar Ilan University, 2007.

ПОЛУФОРМАЛЬНАЯ ВЕРИФИКАЦИЯ ЦИФРОВОГО УСТРОЙСТВА С КОНВЕЙЕРОМ, ОСНОВАННАЯ НА ИСПОЛЬЗОВАНИИ АЛГОРИТМИЧЕСКИХ МАШИН СОСТОЯНИЯ

С. Баранов¹, С. Л. Френкель², В. Н. Захаров³

¹Холонский технологический институт, Израиль, samary@012.net.il

²Институт проблем информатики Российской академии наук, fsergei@mail.ru

³Институт проблем информатики Российской академии наук, VZakharov@ipiran.ru

Аннотация: Статья описывает методы и средства использования Алгоритмических машин состояний (ASM) для высокоуровневой спецификации проектов сложных цифровых систем и их верификации методом Проверки моделей (Model Checking). Данный подход использует иерархическое описание проектов на алгоритмическом уровне абстракции и возможность построения конечно-автоматных моделей по ASM-диаграммам алгоритмов. Для автоматической генерации программ Проверки моделей на языке SMV (Symbolic Modes Verifier) по ASM-диаграмме проектируемого устройства была разработана специальная программа. Способ применения данного подхода и программы демонстрируется на примере верификации проекта конвейерного микропроцессора.

Ключевые слова: формальная верификация; метод Проверки моделей; конечные автоматы

АЛГОРИТМ КООРДИНАЦИИ ДЛЯ ГИБРИДНОЙ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ РЕШЕНИЯ СЛОЖНОЙ ЗАДАЧИ ОПЕРАТИВНО-ПРОИЗВОДСТВЕННОГО ПЛАНИРОВАНИЯ

А. В. Колесников¹, С. А. Солдатов²

Аннотация: Рассмотрена задача оперативно-производственного планирования на машиностроительном предприятии с заказным, мелкосерийным характером производства и описан подход к решению подобных задач на основе методологии функциональных гибридных интеллектуальных систем (ГиИС) с координацией.

Ключевые слова: машиностроительное производство; координация; задача оперативного планирования; гибридная интеллектуальная система

1 Введение

Задачи оперативно-производственного планирования рассматриваются в работах многих отечественных и зарубежных ученых: С. Н. Петракова, Г. Н. Кальянова, Е. В. Фрейдиной, В. П. Заболотского, Ю. Е. Звягинцева, Н. С. Сачко, Д. Тейлора и др. Но, несмотря на имеющееся разнообразие научных методов и инженерного инструментария [1–7], планирование производства по-прежнему остается плохо изученным объектом.

Как показал анализ, выработанное в исследовании операций, искусственного интеллекта, теории принятия решения и системном анализе представление о задаче для индустриального общества устарело. В известных методах и моделях отсутствует либо имеет ограниченную область применения важный для решения задачи оперативно-производственного планирования механизм взаимодействия (координации) подзадач в ходе решения сложной задачи [8].

Опыт применения функциональных интеллектуальных гибридных систем [8] в планировании также показал нерелевантность применяемых представлений о задаче уровню сложности явлений и процессов в системах поддержки принятия решений (СППР).

Настоящая работа развивает предложенное в [8] двухуровневое представление сложных задач и предлагает новый подход к построению функциональных ГиИС, основанный на моделировании координации в ходе коллективного обсуждения решаемых проблем.

2 Эволюция понятия «координация»

Исследование понятия «координация» для количественных, хорошо формализуемых задач исследования операций прослеживается в работах Дж. Фаркаша и Л. В. Кантаровича [1], Дж. Данцига и П. Вулфа [9], Р. Беллмана [10], посвященных математическому программированию. При этом декомпозиция сложной задачи на составные части сводится к математическому приему поиска для исходной матрицы ее «блочной» структуры, вычислениям на частях-блоках и численным методам интеграции (объединения) частных решений в общее. В работах М. Месаровича, Д. Мако и И. Такахары предложен иной подход к координации в сложных формализованных задачах [11]. Сложная задача рассматривается как последовательность решаемых подзадач без структурно выделенного координирующего элемента. В этом случае только завершение решения всех подзадач проясняет, получено ли решение общей задачи. Применение Ф. И. Перегудовым, Ф. Л. Тарасенко, Р. Акоффом, Ф. Эмери [13, 12] системного подхода к анализу решения сложных задач дало более совершенные теоретико-множественные представления. Сложные задачи обрели состав, структуру и эмерджентность. При этом возникли проблемы с количественной мерой сложности и исследованием причинно-следственных связей качественных (количественных) параметров задачи с ее интегративными свойствами.

Фактическое положение вещей на примере коллективного решения задачи оперативного планирования в СППР показывает обязательное

¹Калининградский филиал Института проблем информатики РАН, avkolesnikov@yandex.ru

²ООО «Лайтон», Москва, ssa@west-automatica.com

наличие координирующего элемента — лица, принимающего решения (ЛПР), диспетчера, в частности выполняющего специфические, малоизученные функции, связанные с самоорганизацией в ходе коллективного обсуждения.

В этой связи в рамках системного подхода предлагается новая модель сложной задачи: метод моделирования решения сложных задач с координацией подзадач, а также архитектура информационно-вычислительной системы, построенной по методологии функциональных гибридных интеллектуальных систем [8].

3 Самоорганизация в системах поддержки принятия решений

Системы поддержки принятия решений (система «ЛПР–эксперты») — это выработанный жизнью способ решения сложных практических задач [8]. Применительно к планированию на машиностроительных предприятиях они названы планерками. Главная особенность таких решений — самоорганизация в процессе коллективного обсуждения. Здесь многое зависит не только от экспертов и решения частных задач (задач-элементов), но и от ЛПР, его знаний и опыта работы со сложными задачами-системами и управления ходом коллективного обсуждения.

Обозначим задачу-систему π^u , а задачу-элемент — π^h (рис. 1, а). Тогда $\Pi^h = \{\pi_1^h, \dots, \pi_N^h\}$ — множество задач-элементов, входящих в π^u ; $\dot{\Pi}^u = \{\hat{\pi}_1^u, \dots, \hat{\pi}_M^u\}$ — множество декомпозиций задачи π^u ; $R^{wq}|w, q = 1, \dots, N; w \neq q$ — отношения

между задачами-элементами; N — здесь и далее мощность множества. Тогда модель задачи-системы представим в следующем виде:

$$\pi^u = \langle \Pi^h, \dot{\Pi}^u, R^{wq} \rangle. \quad (1)$$

При решении задачи-системы задачи-элементы преимущественно отделены от внешней среды или ее состояние зафиксировано, т. е. выполняется требование о том, что связи внутри системы намного сильнее, чем с внешней средой.

Модель (1) имеет недостатки, и основной из них — нерелевантное отображение связей между элементами. Учитывать только связи R^{wq} недостаточно, а простое суммирование решений задач-элементов не дает решения задачи-системы. Исследования СППР показали, что в большинстве случаев эксперты не могут дать профессиональных решений в условиях, заданных им ЛПР изначально. А для изменения первоначальных условий в модели (1) необходим существенный элемент — образ ЛПР, который выполнял бы функцию координатора, как «перераспределителя» ресурсов и переформулировал бы в зависимости от ситуации цели экспертов. Это позволило бы отобразить ситуацию, когда реальная СППР «приспосабливается» к различного рода обстоятельствам во внешней среде. Согласно синергетической парадигме это должно происходить путем самоорганизации СППР.

В связи с вышесказанным в предлагаемом подходе задача рассматривается не только как отображение последовательности решений подзадач, но и как система с новым элементом — координатором π^k . Его функция — мониторинг и управление процессом решения подзадач π_1^h, \dots, π_N^h экспертами в ходе коллективного обсуждения. Координатор

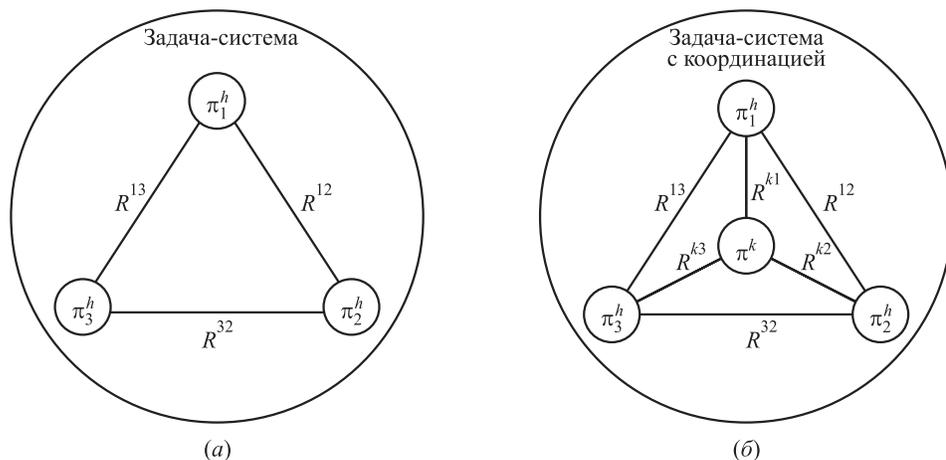


Рис. 1 Пример традиционного представления задачи в системном подходе (а) и с учетом координации (б): π_1^h, \dots, π_3^h — задачи-элементы (подзадачи); π^k — задача-координатор; $R^{wq}|w, q = 1, \dots, 3; w \neq q$ — отношения подзадач; R^{kq} — отношения координатора с подзадачами

связан отношениями $R^{kq}|q = 1, \dots, N$ с каждой задачей π^h в системе π^u , посредством которых собирает информацию о состоянии процесса решения экспертами задач-элементов и в конкретные моменты времени (планерки) выдает координирующие воздействия для изменения входного набора данных (ресурсов, целей). Тогда модель сложной задачи с координацией представим в следующем виде:

$$\pi^{uk} = \langle \Pi^h, \dot{\Pi}^u, \pi^k, R^{wq}, R^{kq} \rangle, \quad (2)$$

где π^k — координатор; $R^{kq}|q = 1, \dots, N$ — отношения между координатором и задачами-элементами.

Таким образом, можно дать следующее определение сложной практической задачи — это задача, включающая взаимодействующие элементы-подзадачи, между которыми происходит обмен данными (значениями переменных, синхроимпульсами и т.п.), управляемый специальным элементом — координатором. Его присутствие отображает самоорганизацию СППР при решении задач-систем.

Сравнение (1) и (2) показывает, что (2) носит более общий характер и легко сводится к (1). По сути, элемент-координатор может быть представлен «координирующей задачей» (k -задачей), которая должна быть «добавлена» в декомпозицию $\dot{\pi}^u \in \dot{\Pi}^u$ сложной задачи π^u , чтобы релевантно отображать в модели особенности коллективного решения задач планирования. Отметим, что задача-координатор не дополняет имеющиеся подзадачи, а существует как подзадача сложной задачи, решение которой традиционно возлагалось на ЛПР.

Можно также отметить, что с увеличением количества задач-элементов, актуальность координации их решения возрастает, так как комбинаторно растет число отношений между задачами-элементами.

4 Алгоритм координации в системах поддержки принятия решений

Предлагаемый алгоритм отображает самоорганизацию в СППР по принципу «как есть» (лат. *ad hoc*) и относится к алгоритмам, основанным на знаниях, где акцент смещается с использования формализованной математической схемы на извлечение профессиональных знаний и рассуждениям с их применением. Для разработки алгоритма такие знания были извлечены на примере СППР машиностроительного предприятия с мелкосерийным характером производства. Все знания представлены продукционными системами. База знаний ЛПР (начальника производственного центра)

о k -задаче содержит 24 правила. Базы знаний пяти экспертов (начальников отделов) содержат от 10 до 40 правил. В качестве оболочки ЭС выбрана программа Visual Rule Studio [14] и метод рассуждений в прямом направлении [8].

Приведенный ниже алгоритм имитирует последовательность заседаний СППР в относительном, модельном времени. При этом линии рассуждений экспертов координируются ЛПР. Полученные экспертами после каждой итерации (планерки) решения частных подзадач передаются ЛПР как исходные данные для задачи координации (k -задачи). Лицо, принимающее решения, использует в процессе решения задачи координации данные, полученные после декомпозиции исходной сложной задачи. Решив задачу координации, ЛПР дает рекомендации каждому из экспертов. Эти рекомендации в совокупности с первоначально известными после декомпозиции данными, служат исходными данными для решаемых экспертами подзадач на следующей итерации (планерке). В случаях, когда не требуется вносить изменения в ход решения подзадачи экспертом, ЛПР выдает «пустую команду».

Дано: СППР, состоящая из N экспертов и ЛПР. Число планерок в плановом периоде — k . Установлено однозначное соответствие

$$\psi_1 : \text{Out}|\pi^h_1 \cup \dots \cup \text{Out}|\pi^h_N \rightarrow \text{In}|\pi^k$$

между выходными параметрами $\text{Out}|\pi^h_i = \{\text{Out}_1, \dots, \text{Out}_m\}|\pi^h_i$ подзадач π^h_i , $i = 1, \dots, N$, и входными параметрами $\text{In}|\pi^k = \{\text{In}_1, \dots, \text{In}_n\}|\pi^k$ задачи координации π^k . Установлено однозначное соответствие

$$\psi_2 : \text{Out}|\pi^k \rightarrow \text{In}|\pi^h_1 \cup \dots \cup \text{In}|\pi^h_N$$

между выходными параметрами $\text{Out}|\pi^k = \{\text{Out}_1, \dots, \text{Out}_n\}|\pi^k$ задачи координации π^k и входными параметрами $\text{In}|\pi^h_i = \{\text{In}_1, \dots, \text{In}_n\}|\pi^h_i$ подзадач π^h_i , $i = 1, \dots, N$. Соответствия ψ_1 и ψ_2 сюръективны и инъективны.

Найти: результаты решения (выходные параметры) подзадач π^h_i , $i = 1, \dots, N$, экспертами с учетом координации их решения ЛПР.

Обозначения: БЗ_i , $i = 1, \dots, N$, — базы знаний экспертов; $\text{БЗ}_{\text{лпр}}$ — база знаний ЛПР; БФ_i , $i = 1, \dots, N$, — базы фактов экспертов; $\text{БФ}_{\text{лпр}}$ — база фактов ЛПР; j — счетчик планерок; i — счетчик экспертов; $\text{Run}(\text{БФ}, \text{БЗ})$ — процедура имитации рассуждений ЛПР или экспертов, где БФ и БЗ — базы фактов и знаний соответственно; $\text{Search}(\text{БЗ})$ — процедура-просмотр правил из БЗ

и сопоставление фактов с образцами с целью определения множества правил, которые могут быть активированы; Conf — процедура разрешения конфликтов в «плане решения» (Agenda) экспертной системы, когда возникает необходимость выбора между несколькими правилами из БЗ для продолжения рассуждений (метод разрешения конфликтов — «первый в Agenda»); Execute — процедура применения найденных правила для продолжения рассуждений, т. е. выполнения действий указанных в правой части выбранных правил.

Алгоритм

- 1) $j = 0$;
- 2) $i = 1$;
- 3) Заполнить $БФ_i$ с учетом $(In_1, \dots, In_n)^{\pi_i^h}$ и счетчика планерок j ;
- 4) Run($БФ_i$, $БЗ_i$):
 - 1.1) Search($БЗ_i$);
 - 1.2) Conf;
 - 1.3) Execute;
 - 1.4) Если для продолжения рассуждений не выбрано ни одно из правил, то закончить имитацию рассуждений и перейти к п. 5. Иначе — к п. 4.1;
- 5) Присвоить $(Out_1, \dots, Out_m)^{\pi_i^h}$ значения из фактов — результатов рассуждений i -го эксперта;
- 6) Присвоить, используя ψ_1 , соответствующим входным параметрам $(In_1, \dots, In_n)^{\pi^k}$ задачи координации π^k значения выходных параметров $(Out_1, \dots, Out_m)^{\pi_i^h}$ задачи π_i^h ;
- 7) Если $i < N + 1$, то $i = i + 1$ и перейти к п. 3. Иначе перейти к п. 8;
- 8) Заполнить $БФ_{лпп}$ с учетом значений входных параметров $(In_1, \dots, In_n)^{\pi^k}$ задачи координации π^k и счетчика планерок j ;
- 9) Run($БФ_{лпп}$, $БЗ_{лпп}$):
 - 9.1) Search($БЗ_{лпп}$);
 - 9.2) Conf;
 - 9.3) Execute;
 - 9.4) Если для продолжения рассуждений не выбрано ни одно из правил, то закончить имитацию рассуждений и перейти к п. 10. Иначе — к п. 9.1;
- 10) Присвоить $(Out_1, \dots, Out_n)^{\pi^k}$ задачи координации π^k значения из фактов-результатов рассуждений ЛПР;

- 11) Присвоить, используя ψ_2 , соответствующим входным параметрам $(In_1, \dots, In_n)^{\pi^k}$ подзадач π_1, \dots, π_N значения выходных параметров задачи координации π^k ;
- 12) Если $j < k$, то $j = j + 1$ и перейти к п. 2. Иначе перейти к п. 13;
- 13) Вывести значения выходных параметров подзадач, полученные на j -й планерке:
 - 13.1) $i = 1$;
 - 13.2) Печать $(Out_1, \dots, Out_m)^{\pi_i^h}$;
 - 13.3) Если $i < N + 1$, то $i = i + 1$ и перейти к п. 13.2. Иначе перейти к п. 14;
- 14) Конец.

Разработанный алгоритм дискретен и представляет координацию как последовательность простых шагов. Ему присуща определенность — каждое правило алгоритма детерминированное, четкое и не оставляет места для многозначности. Алгоритм приводит к решению сложной задачи за конечное число шагов (результативность), так как внешний цикл (имитация планерок) выполняется не более $k - 1$ раз. Вложенные циклы (имитация рассуждений каждого эксперта) и цикл вывода фактов-результатов выполняются не более N раз. Отсутствие циклов в графе И/ИЛИ рассуждений экспертов и ЛПР гарантирует отсутствие заикливания. Алгоритм координации разработан в общем виде (массовость), однако он не лишен недостатка «хрупкости» любой экспертной системы.

Порядок опроса экспертов не важен, так как пока последний эксперт не сообщит результат решения своей подзадачи, ЛПР не может приступить к решению задачи координации. В случае, если в результате имитации рассуждений экспертов или ЛПР не могут быть выведены факты, необходимые для спецификации выходных параметров решаемых подзадач, то используются значения по умолчанию, извлеченные у экспертов и ЛПР при разработке конкретной информационно-вычислительной системы. Существенное отличие базы фактов ЛПР от баз фактов экспертов в том, что в ней происходит интеграция результатов решения подзадач экспертами, основанная на причинно-следственных связях, частично-формализованных продукционными правилами и представленными базой профессиональных знаний ЛПР, о качестве которых можно судить исключительно по результатам проведенной апробации. Для таких экспериментов была разработан программный продукт «Гибридная система планирования».

5 Модель гибридной интеллектуальной системы с учетом координации

В качестве модели ГиИС как абстрактного автомата для решения сложной задачи оперативно-производственного планирования принята функциональная крупнозернистая ГиИС $\alpha^u(t)$ [8]. Ее расширение выполнено исходя из следующих посылок. В процессе координации контролируются промежуточные состояния процесса решения подзадач. Под этими состояниями понимаются состояния функциональных элементов α^h , имитирующих решение подзадач π^h , а также состояния технологических элементов α^τ . На основании анализа этих состояний в ходе координации изменяются свойства «вход» \hat{x}_1^2 одного или нескольких функциональных и технологических элементов α^h и α^τ . Для учета этого факта введем в модель крупнозернистой функциональной ГиИС $\alpha^u(t)$ следующую тройку: $\hat{x}_3^2(t)R^{22}\hat{x}_1^2(t+1)$. Иными словами, на основании состояния ГиИС $\hat{x}_3^2(t)$ в момент времени t меняются исходные данные $\hat{x}_1^2(t+1)$ для ГиИС в момент времени $t+1$, т. е. для следующей итерации. Множество R^{22} устанавливает отношения между состоянием $\hat{x}_3^2(t)$ гибрида α^u на данный момент модельного времени t и состоянием входов одного или нескольких функциональных и технологических элементов α^h и α^τ на следующем шаге. Чтобы произвести необходимое изменение входов \hat{x}_1^2 одного или нескольких функциональных и технологических элементов α^h и α^τ в $\alpha^u(t)$ введем тройку $\hat{x}_3^2(t)R^{23}X^3$, где $X^3 = \{x_1^3, \dots, x_6^3\}$ — множество понятий, обозначающих координирующие действия (интегральная координация, четкая координация, интервальная координация, лингвистическая координация, координация по времени, «пустое действие»), которое тождественно множеству координирующих действий, введенных в [15]. В алгоритме координации эти действия (рекомендации экспертам) описаны в базе знаний ЛПР. Множество R^{23} — это отношения между состоянием \hat{x}_3^2 гибрида α^u в момент модельного времени t и необходимыми координирующими действиями X^3 . Ниже приведена модифицированная схема ролевых концептуальных моделей α^{uk} для спецификации крупнозернистой функциональной ГиИС с координацией

$$\alpha^{uk}(t) = \alpha^u(t) \circ \hat{x}_3^2(t)R^{22}\hat{x}_1^2(t+1) \circ \hat{x}_3^2(t)R^{23}X^3, \quad (3)$$

где \circ — знак конкатенации; 1, 2, 3 в качестве правого верхнего индекса X или x — признак ресурса, свойства, действия соответственно; нижний правый индекс для x — порядковый номер класса по-

нятий; верхний правый индекс для R обозначает между какими понятиями категориального ядра [8] установлены отношения (22 — «свойство—свойство», 23 — «свойство—действие»).

Отношения R^{22} и R^{23} не задаются заранее, а фиксируются в ходе функционирования ГиИС и поиска результата решения задачи-координатора (k -задачи). Поскольку в соответствии с [8] технологические элементы управляют порядком работы функциональных элементов и обменом информации между ними, целесообразно возложить решение задачи координации на технологический элемент.

Рассмотрим пример ГиИС, состоящей из трех функциональных элементов $\alpha^h|_1^1, \alpha^h|_1^7, \alpha^h|_1^6$ и одного технологического элемента $\alpha^\tau|_1^7$, где $w = 1, \dots, 7$ для $\alpha^h|_j^w$ и $\alpha^\tau|_j^w$ обозначает базовые классы методов функциональных ГиИС [8], $j = 1$ — порядковый номер задачи в классе. На вход ГиИС подаются исходные данные, разделенные между функциональными элементами в соответствии с декомпозицией сложной задачи. На выходе имеем результаты работы функциональных элементов, агрегированные в общее решение задачи.

На рис. 2, а изображена структурная схема ГиИС, построенная для решения сложной задачи в соответствии с $\alpha^u(t)$. Здесь моделируется только логически увязанная последовательность решения подзадач π^h из декомпозиции $\hat{\pi}^u$ сложной задачи π^u . Это соответствует модели сложной задачи на рис. 1, а. В этом случае по каналу обратной связи ЛПР получает от компьютерной СППР решение сложной задачи. Если общее единое решение дает с точки зрения ЛПР ошибочный результат, то он на основании своих оценок по каналу прямой связи вносит изменения во множества входных данных и условий задачи π^u . Далее ЛПР инициирует новый синтез ГиИС и повторное решение.

На рис. 2, б изображена принципиально иная структурная схема ГиИС. Ее отличие от вышеприведенной в том, что технологический элемент $\alpha^\tau|_1^7$ определяет не только порядок работы функциональных элементов и обмен информации между ними, но и в соответствии с (3) по состоянию всех функциональных элементов итерационно корректирует для каждого из них входной набор данных и условий. Таким образом, часть функций ЛПР передается технологическому элементу, что отражено на рис. 2, б изменением размеров структурных блоков ЛПР и технологического элемента, а также толщины линий прямой и обратной связи в контуре управления.

В представленной на рис. 2, б структурной схеме ЛПР по каналу обратной связи получает от компьютерной СППР результат решения сложной задачи.

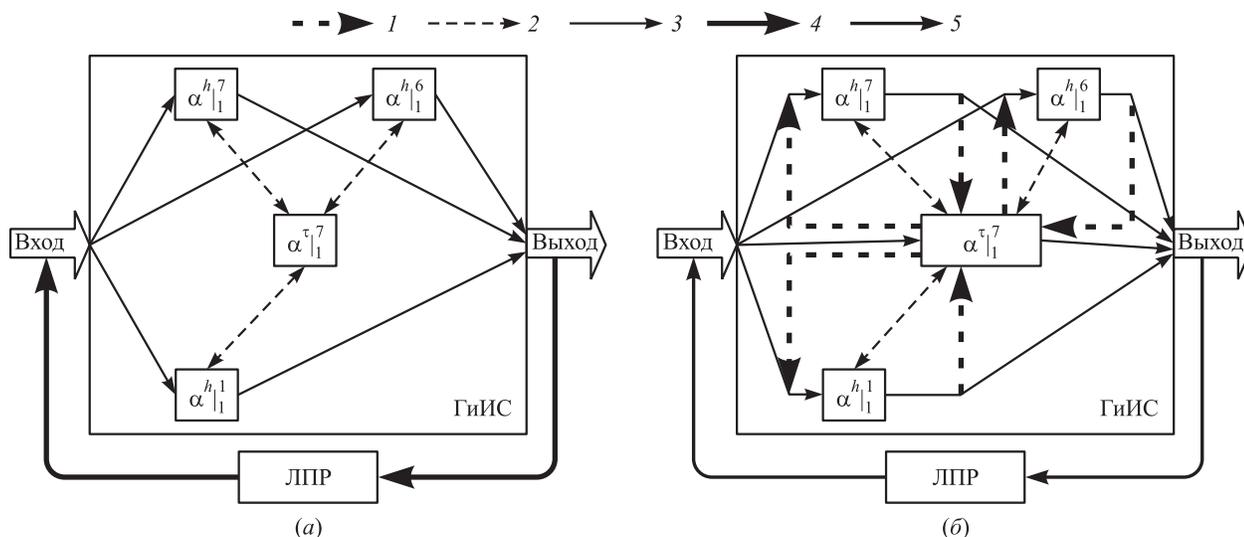


Рис. 2 Структурная схема ГиИС без координации (а) и с координацией (б): 1 — прямая и обратная координирующие связи между технологическим элементом и функциональными элементами; 2 — отношения порядка работы и обмена информацией между элементами; 3 — редукция сложной задачи (на входе) и интеграция результатов работы функциональных элементов; 4 — прямая и обратная связи между ГиИС и ЛПР при отсутствии координации внутри ГиИС; 5 — прямая и обратная связи между ГиИС и ЛПР при наличии координации внутри ГиИС

Если решение не устраивает ЛПР, например оно ведет к увеличению стоимости производимых изделий и т. д., то он вмешивается в обсуждение и меняет условия координации, т. е. модель задачи-координатора. Далее ЛПР инициирует новый синтез ГиИС и повторное решение сложной задачи. В итоге это позволяет моделировать процесс самоорганизации, о котором было сказано выше.

В модели (3) развито предположение, что включение в компьютерную модель СППР модели ЛПР приводит к возникновению синергетических эффектов — самоорганизации. При этом появляется возможность увязать результаты работы отдельных функциональных элементов СППР еще в процессе синтеза решения сложной задачи, а не после, как в известных моделях. Тем самым достигается большая релевантность компьютерной СППР реальному процессу коллективного обсуждения проблем.

6 Заключение

Гибридные интеллектуальные системы с координацией элементов — новый шаг в синергетическом искусственном интеллекте, позволяющий полнее раскрыть и исследовать многообразие отношений в системе «ЛПР—эксперты». Это также результат и в моделировании категории «время» в гибридных системах на примере производственных

планерок динамических процессов оперативно-производственного планирования.

По сравнению с известными алгоритмами координации [1, 8–10, 16, 17], предлагаемый алгоритм имеет следующие достоинства: итеративность позволяет имитировать обмен информацией в процессе решения сложной задачи, а применение профессиональных баз знаний делает модель релевантной сложной практической задаче.

Апробация ГиИС, решающих k -задачу, на реальных данных показала положительный эффект от увеличения релевантности моделирования сложной задачи оперативно-производственного планирования, что улучшило качественные и количественные показатели машиностроительного предприятия с мелкосерийным заказным характером производства. В частности, ожидаемый экономический эффект от внедрения ГиИС на одном из предприятий в процентном отношении к ожидаемой прибыли предприятия за год составил 11%.

Литература

1. Канторович Л. В. Математические методы организации и планирования производства. — Л.: ЛГУ, 1959.
2. Татевосов К. Г. Основы оперативно-производственного планирования на машиностроительном предприятии. — Л.: Машиностроение, 1985.

3. *Заболотский В. П., Оводенко А. А., Степанов А. Г.* Математические модели в управлении: Учеб. пособие. — СПб.: СПбГУАП, 2001.
4. *Туровец О. Г., Родионов В. Б., Бухалков М. И.* Организация производства и управление предприятием. — М.: ИНФРА-М, 2005.
5. *Кальянов Г. Н.* Моделирование, анализ, реорганизация и автоматизация бизнес-процессов. — М.: Финансы и статистика, 2007.
6. *Сачко Н. С.* Организация и оперативное управление машиностроительным производством. — Минск: Новое Знание, 2008.
7. *Тейлор Д., Рэйден Н.* Почти интеллектуальные системы. Как получить конкурентные преимущества путем автоматизации принятия скрытых решений. — СПб.: Символ-Плюс, 2009.
8. *Колесников А. В., Кириков И. А.* Методология и технология решения сложных задач методами функциональных гибридных интеллектуальных систем. — М.: ИПИ РАН, 2007.
9. *Данигс Дж.* Линейное программирование, его обобщения и применения. — М.: Прогресс, 1966.
10. *Беллман Р., Дрейфус С.* Прикладные задачи динамического программирования. — М.: Наука, 1965.
11. *Месарович М., Мако Д., Такахара И.* Теория иерархических многоуровневых систем. — М.: Мир, 1973.
12. *Акофф Р., Эмери Ф.* О целеустремленных системах. — М.: Советское радио, 1974.
13. *Перегудов Ф. И., Тарасенко Ф. Л.* Введение в системный анализ. — М.: Высшая школа, 1989.
14. Сайт компании Rule Machines Corporation. <http://www.rulemachines.com>.
15. *Колесников А. В., Солдатов С. А.* Теоретические основы решения сложной задачи оперативно-производственного планирования с учетом координации // Вестник Российского государственного ун-та им. И. Канта. Вып. 10. Сер. Физико-математические науки. — Калининград: РГУ им. И. Канта, 2009. С. 82–98.
16. *Beat F., Schmid K., Stanoevska S., Lei Yu.* Supporting distributed corporate planning through new coordination technologies, 1998. <http://www.alexandria.unisg.ch/Publikationen/9453>.
17. *Geun-Sik Jo, Kang-Hee L., Hwi-Yoon L., Sang-Ho H.* Ramp activity expert system for scheduling and coordination at an airport // Innovative Application of Artificial Intelligence '99, AAAI, July, 1999. P. 807–812. <http://www.aaai.org/Papers/IAAI/1999/IAAI99-114.pdf>.

СТАЦИОНАРНЫЕ ХАРАКТЕРИСТИКИ ДВУХКАНАЛЬНОЙ СИСТЕМЫ ОБСЛУЖИВАНИЯ С ПЕРЕУПОРЯДОЧИВАНИЕМ ЗАЯВОК И РАСПРЕДЕЛЕНИЯМИ ФАЗОВОГО ТИПА

С. И. Матюшенко¹

Аннотация: Рассматривается двухканальная система обслуживания ограниченной емкости с распределениями фазового типа и переупорядочиванием заявок. Получено выражение для преобразования Лапласа–Стилтьеса функции распределения задержки переупорядочивания в рассматриваемой системе в стационарном режиме работы. Разработан алгоритм для расчета факториальных моментов числа заявок в буфере переупорядочивания.

Ключевые слова: система массового обслуживания; распределение фазового типа; переупорядочивание заявок

1 Постановка задачи

Рассмотрим двухканальную систему массового обслуживания (СМО) с общим накопителем ограниченной емкости r , $r < \infty$, на которую поступает рекуррентный поток заявок с функцией распределения (ФР) фазового типа $A(x)$,

$$A(x) = 1 - \alpha^T e^{\Lambda x} \mathbf{1}, \quad x \geq 0, \quad \alpha^T \mathbf{1} = 1,$$

с неприводимым РН-представлением (α, Λ) порядка l [1].

Будем предполагать, что времена обслуживания на приборе j независимы между собой и имеют общую ФР фазового типа

$$B_j(x) = 1 - \beta_j^T e^{M_j x} \mathbf{1}, \quad x \geq 0, \quad \beta_j^T \mathbf{1} = 1$$

с неприводимым РН-представлением (β_j, M_j) порядка m_j , $j = 1, 2$.

Далее, без ограничения общности примем, что интенсивность обслуживания на приборе 1 выше, чем на приборе 2. Заявка, поступающая в свободную СМО, направляется на первый прибор. При наличии очереди действует дисциплина FCFS.

Предположим, что всем заявкам при поступлении в систему присваивается порядковый номер. На выходе из СМО будем требовать сохранения порядка между заявками, установленного при входе в нее. Заявки, прошедшие обслуживание и нарушившие установленный порядок, будут накапливаться на выходе системы в буфере переупорядочивания (БП).

В соответствии с обозначениями Кендалла рассматриваемую систему будем кодировать как

$PH/PH/2/r/res$, где res — сокращение от английского *resequence* — переупорядочивание. Такая система уже рассматривалась автором в [2]. В той работе функционирование системы описывалось однородным марковским процессом (ОМП) над пространством состояний без учета содержимого БП. В итоге был разработан рекуррентный матричный алгоритм для расчета стационарных вероятностей состояний указанного процесса. Основная задача данной работы состоит в том, чтобы, опираясь на результаты [2], получить показатели, характеризующие стационарное состояние БП. Перейдем к решению этой задачи.

2 Построение математической модели

Построим марковский процесс, описывающий функционирование рассматриваемой системы. Для этого введем понятие упорядоченности. Будем считать, что система находится в упорядоченном состоянии (упорядочена), если на приборах 1 и 2 обслуживаются заявки с номерами N_1 и N_2 , $N_1 < N_2$, в противном случае, т.е. при $N_1 > N_2$, система не упорядочена. Система также упорядочена (не упорядочена) при наличии в ней одной заявки на приборе 1 (приборе 2).

Теперь рассматриваемую СМО с учетом введенного понятия упорядоченности, а также с учетом вероятностной интерпретации РН-распределения ([1], с. 104) можно описать ОМП $Y(t)$, $t \geq 0$, над пространством состояний

¹Российский университет дружбы народов, кафедра теории вероятностей и математической статистики, matushenko@list.ru

$$\mathcal{Y} = \bigcup_{k=0}^{r+2} \mathcal{Y}_k, \quad \mathcal{Y}_k = \mathcal{Y}_{k1} \cup \mathcal{Y}_{k2},$$

$$\mathcal{Y}_{ki} = \bigcup_{n=0}^{\infty} \mathcal{Y}_{kin}, \quad k = \overline{1, r+2}, \quad i = 1, 2,$$

где

$$\mathcal{Y}_0 = \{(s, 0), s = \overline{1, l}\};$$

$$\mathcal{Y}_{1in} = \{(s, j_i, i, n), s = \overline{1, l}, j_i = \overline{1, m_i}\},$$

$$i = 1, 2, \quad n \geq 0;$$

$$\mathcal{Y}_{kin} = \{(k, s, j_1, j_2, i, n), s = \overline{1, l}, j_1 = \overline{1, m_1},$$

$$j_2 = \overline{1, m_2}\}, \quad k = \overline{2, r+2}, \quad i = 1, 2, \quad n \geq 0.$$

Здесь для некоторого момента времени t : $Y(t) = (s, 0)$, если в момент времени t система пуста, а процесс генерации заявки проходит фазу s ; $Y(t) = (s, j_i, i, n)$, если в системе имеется одна заявка, обслуживаемая на первом приборе при $i = 1$ либо на втором приборе при $i = 2$, процесс обслуживания находится на фазе j_i , а в БП содержится n заявок; $Y(t) = (k, s, j_1, j_2, i, n)$, если в очереди и на приборах имеется k заявок, процессы обслуживания заявок на приборах находятся на фазах j_1 и j_2 соответственно, причем система упорядочена, если $i = 1$, либо не упорядочена, если $i = 2$, а индексы s и n имеют прежний смысл.

В предположении, что интенсивности потока и обслуживания конечны, процесс $Y(t)$ эргодичен и, следовательно, существуют вероятности

$$p_y = \lim_{t \rightarrow \infty} P\{Y(t) = y\}, \quad y \in \mathcal{Y},$$

совпадающие со стационарными.

Введем векторы

$$\mathbf{p}_{1in}^T = (p_{11in}, \dots, p_{1m_i in}, \dots, p_{2m_i in}, \dots,$$

$$\dots, p_{l1in}, \dots, p_{lm_i in}), \quad i = 1, 2, \quad n \geq 0;$$

$$\mathbf{p}_{kin}^T = (p_{ki1in}, \dots, p_{ki1m_2in}, \dots, p_{ki1m_1m_2in}, \dots,$$

$$\dots, p_{klm_1m_2in}), \quad k = \overline{2, r+2}, \quad i = 1, 2, \quad n \geq 0.$$

Будем использовать обозначение $\mathbf{p}_{ki\cdot} = \sum_{n \geq 0} \mathbf{p}_{kin}$ и введем матрицы:

$$\mathbf{D}_{1i} = \mathbf{\Lambda}_1 \oplus \mathbf{M}_i, \quad i = 1, 2;$$

$$\mathbf{D}_{ki} = \mathbf{\Lambda}_k \oplus \mathbf{M}_i, \quad k = \overline{2, r+1}, \quad i = 1, 2;$$

$$\mathbf{D}_{r+2, i} = (\mathbf{\Lambda}_{r+2} + \lambda_{r+2} \alpha^T) \oplus \mathbf{M}_i, \quad i = 1, 2;$$

$$\mathbf{E}_{1i} = \begin{cases} \lambda_1 \alpha^T \otimes \mathbf{I} \otimes \beta_2^T, & i = 1, \\ \lambda_1 \alpha^T \otimes \beta_1^T \otimes \mathbf{I}, & i = 2; \end{cases}$$

$$\mathbf{E}_{ki} = \lambda_k \alpha^T \otimes \mathbf{I} \otimes \mathbf{I}, \quad k = \overline{2, r+2}, \quad i = 1, 2;$$

$$\mathbf{G}_{2i} = \begin{cases} \mathbf{I} \otimes \mathbf{I} \otimes \mu_2, & i = 1, \\ \mathbf{I} \otimes \mu_1 \otimes \mathbf{I}, & i = 2; \end{cases}$$

$$\mathbf{G}_{ki} = \begin{cases} \mathbf{I} \otimes \mathbf{I} \otimes \mu_2 \beta_2^T, & i = 1, \\ \mathbf{I} \otimes \mu_1 \beta_1^T \otimes \mathbf{I}, & i = 2. \end{cases}$$

Здесь и далее $U \otimes V$ — кронекерово произведение, а $U \oplus V$ — кронекерова сумма матриц U и V .

Стационарное распределение вероятностей $\{p_y, y \in \mathcal{Y}\}$ удовлетворяет следующей системе уравнений равновесия (СУР):

$$\mathbf{0}^T = \mathbf{p}_0^T \mathbf{\Lambda}_0 + \mathbf{p}_{11, \cdot}^T (\mathbf{I} \otimes \mu_1) + \mathbf{p}_{12, \cdot}^T (\mathbf{I} \otimes \mu_2); \quad (1)$$

$$\mathbf{0}^T = u(1-n)u(2-i)\mathbf{p}_0^T (\lambda_0 \alpha^T \otimes \beta_1^T) + \mathbf{p}_{1in}^T \mathbf{D}_{1i} +$$

$$+ [u(1-n)\mathbf{p}_{2,3-i, \cdot}^T + u(n)\mathbf{p}_{2,i,n-1}^T] \mathbf{G}_{2i},$$

$$i = 1, 2, \quad n \geq 0; \quad (2)$$

$$\mathbf{0}^T = u(3-k)\mathbf{p}_{1in}^T \mathbf{E}_{1i} + u(k-2)\mathbf{p}_{k-1, in}^T \mathbf{E}_{k-1, i} +$$

$$+ \mathbf{p}_{kin}^T \mathbf{D}_{ki} + [u(1-n)\mathbf{p}_{k+1,3-i, \cdot}^T +$$

$$+ u(n)\mathbf{p}_{k+1, i, n-1}^T] \mathbf{G}_{k+1, i},$$

$$k = \overline{2, r+1}, \quad i = 1, 2, \quad n \geq 0; \quad (3)$$

$$\mathbf{0}^T = \mathbf{p}_{r+1, in}^T \mathbf{E}_{r+1, i} + \mathbf{p}_{r+2, in}^T \mathbf{D}_{r+2, i},$$

$$i = 1, 2, \quad n \geq 0, \quad (4)$$

с условием нормировки

$$\mathbf{p}_0^T \mathbf{1} + \sum_{k=1}^{r+2} \sum_{i=1}^2 \sum_{n=0}^{\infty} \mathbf{p}_{kin}^T \mathbf{1} = 1. \quad (5)$$

Здесь и далее $u(x) = 1$ при $x > 0$ и $u(x) = 0$ при $x \leq 0$.

Система уравнений (1)–(5) понадобится в дальнейшем для определения стационарных характеристик рассматриваемой СМО.

3 Факториальные моменты числа заявок в буфере переупорядочивания

Для определения факториальных моментов числа заявок в БП воспользуемся аппаратом производящих функций (ПФ). Положим

$$\left. \begin{aligned} F_{sj_i i}(z) &= \sum_{n \geq 0} p_{sj_i in} z^n, \\ F_{ksj_1 j_2 i}(z) &= \sum_{n \geq 0} p_{ksj_1 j_2 in} z^n, \quad s = \overline{1, l}, \\ j_i &= \overline{1, m_i}, \quad i = 1, 2, \quad z \in \mathbb{C}, \quad |z| \leq 1, \end{aligned} \right\} \quad (6)$$

и введем векторы $F_{1i}(z)$ и $F_{ki}(z)$, аналогичные по структуре векторам p_{1i} и p_{ki} .

Умножая уравнения (2)–(4) для каждого фиксированного $n \geq 0$ справа на z^n и суммируя полученные равенства при каждом фиксированном $k = \overline{1, r+2}$, $i = 1, 2$, по всем возможным значениям n , с учетом (6) приходим к следующей системе уравнений:

$$\mathbf{0}^T = u(1-n)u(2-i)p_0^T (\lambda_0 \alpha^T \otimes \beta_1^T) + F_{1i}^T(z) D_{1i} + [p_{2,3-i}^T + z F_{2i}^T(z)] G_{2i}, \quad i = 1, 2; \quad (7)$$

$$\mathbf{0}^T = u(3-k)F_{1i}^T(z)E_{1i} + u(k-2)F_{k-1,i}^T(z)E_{k-1,i} + F_{ki}^T(z)D_{ki} + [p_{k+1,3-i}^T + z F_{k+1,i}^T(z)] G_{k+1,i}, \quad k = \overline{2, r+1}, \quad i = 1, 2; \quad (8)$$

$$\mathbf{0}^T = F_{r+1,i}^T(z)E_{r+1,i} + F_{r+2,i}^T(z)D_{r+2,i}, \quad i = 1, 2. \quad (9)$$

Далее введем обозначения:

$$v_{kiv} = F_{ki}^{(\nu)}(1), \quad k = \overline{1, r+2}, \quad i = 1, 2, \quad \nu \geq 0, \quad (10)$$

$$v_\nu = \sum_{k=1}^{r+2} \sum_{i=1}^2 v_{kiv} \mathbf{1}, \quad \nu \geq 1. \quad (11)$$

Заметим, что v_ν — факториальный момент порядка ν , $\nu = 1, 2, \dots$, числа заявок в БП, а $v_{ki0} = \sum_{n \geq 0} p_{kin}$, $k = \overline{1, r+2}$, $i = 1, 2$. Причем система уравнений для определения v_{ki0} получается из (7)–(9) после подстановки в них $z = 1$. Кроме того, эта система с учетом (1) и (5) полностью совпадает с СУР из работы [2] и в силу единственности ее решения получаем

$$v_{ki0} = p_{ki}, \quad k = \overline{1, r+2}, \quad i = 1, 2.$$

Теперь получим систему уравнений для определения v_{kiv} , $k = \overline{1, r+2}$, $i = 1, 2$, $\nu \geq 1$. Для этого продифференцируем (7)–(9) ν раз по z и положим $z = 1$. В результате приходим к следующим уравнениям:

$$\mathbf{0}^T = v_{1i\nu}^T D_{1i} + [v_{2i\nu}^T + \nu v_{2i,\nu-1}^T] G_{2i}, \quad i = 1, 2; \quad (12)$$

$$\mathbf{0}^T = u(3-k)v_{1i\nu}^T E_{1i} + u(k-2)v_{k-1,i\nu}^T E_{k-1,i} + v_{kiv}^T D_{ki} + [v_{k+1,i\nu}^T + \nu v_{k+1,i,\nu-1}^T] G_{k+1,i}, \quad k = \overline{2, r+1}, \quad i = 1, 2; \quad (13)$$

$$\mathbf{0}^T = v_{r+1,i\nu}^T E_{r+1,i} + v_{r+2,i\nu}^T D_{r+2,i}, \quad i = 1, 2, \quad \nu = 1, 2, \dots \quad (14)$$

Решение системы (12)–(14) получим с помощью блочного LU-разложения матрицы коэффициентов. Для этого введем векторы

$$\begin{aligned} v_{i\nu}^T &= (v_{1i\nu}^T, \dots, v_{r+2,i\nu}^T), \quad i = 1, 2, \quad \nu \geq 1; \\ d_{kiv}^T &= \begin{cases} -\nu v_{k+1,i,\nu-1}^T G_{k+1,i}, & k = \overline{1, r+1}; \\ \mathbf{0}^T, & k = r+2, \quad i = 1, 2, \quad \nu \geq 1; \end{cases} \\ d_{i\nu}^T &= (d_{1i\nu}^T, \dots, d_{r+2,i\nu}^T), \quad i = 1, 2, \quad \nu \geq 1, \end{aligned}$$

и матрицы

$$D_i = \begin{pmatrix} D_{1i} & E_{1i} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ G_{2i} & D_{2i} & E_{2i} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & G_{3i} & D_{3i} & E_{3i} & \dots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & G_{r+1,i} & D_{r+1,i} & E_{r+1,i} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & G_{r+2,i} & D_{r+2,i} \end{pmatrix}, \quad i = 1, 2.$$

С учетом введенных обозначений систему уравнений (12)–(14) можно записать в следующем виде:

$$v_{i\nu} D_i = d_{i\nu}^T, \quad i = 1, 2, \quad \nu \geq 1. \quad (15)$$

Матрица коэффициентов D_i для любого фиксированного $i = 1, 2$ неразложима и обладает свойствами инфинитезимальной матрицы с диагональным преобладанием. Следовательно, для решения системы (15) можно воспользоваться выводами, полученными в [3]. Непосредственно из [3] вытекает результат, который сформулируем в виде теоремы.

Теорема 1. Для каждого фиксированного $i = 1, 2$ и $\nu = 1, 2, \dots$ решение системы (15) представимо в виде

$$\begin{aligned} v_{r+2,i\nu}^T &= y_{r+2,i\nu}^T; \\ v_{kiv}^T &= y_{kiv}^T + v_{k+1,i\nu}^T H_{k+1,i}, \quad k = \overline{r+1, 1}, \end{aligned}$$

где

$$\begin{aligned} y_{1i\nu}^T &= d_{1i\nu}^T S_{1i}^{-1}; \\ y_{kiv}^T &= [d_{kiv}^T - y_{k-1,i\nu}^T E_{k-1,i\nu}] S_{ki}^{-1}, \quad k = \overline{2, r+2}; \\ H_{ki} &= -G_{ki} S_{k-1,i}^{-1}, \quad k = \overline{2, r+2}, \end{aligned}$$

а невырожденные матрицы S_{ki} , $k = \overline{1, r+2}$, задаются соотношениями:

$$S_{k,i} = D_{ki} + H_{ki} E_{k-1,i} u(k-1).$$

Итак, получен алгоритм, позволяющий для каждого $\nu = 1, 2, \dots$ вычислять факториальный момент v_ν числа заявок в БП через моменты низшего порядка. При этом, начальный шаг алгоритма состоит в определении векторов p_{ki} , для чего следует обратиться к результатам [2].

4 Преобразования Лапласа–Стилтьеса стационарной функции распределения задержки переупорядочивания

Обозначим через δ задержку переупорядочивания заявки в стационарном режиме работы СМО, а через $f(s)$ — преобразование Лапласа–Стилтьеса (ПЛС) ФР $F_\delta(t)$ случайной величины (с.в.) δ . Далее, обозначим через $f_j(s)$ ПЛС условной стационарной ФР задержки переупорядочивания при условии, что задержана заявка, обслуженная прибором j , $j = 1, 2$.

Очевидно, что заявка, обслуженная прибором j , $j = 1, 2$, не будет задержана для переупорядочивания, если после ее ухода система останется пустой либо если в момент $\tau = 0$ окончания обслуживания этой заявки на приборе $3 - j$ будет обслуживаться заявка, пришедшая в систему позже данной заявки. Если же в момент $\tau = 0$ на приборе $3 - j$ будет обслуживаться заявка, пришедшая в систему раньше данной, то данная заявка будет задержана до момента окончания обслуживания на приборе $3 - j$. Обозначим через $\pi_{D,j}^-(s, j)$ и $\pi_{D,j}^-(k, s, j_{3-j}, i)$ стационарные вероятности макросостояний (s, j) , (k, s, j_{3-j}, i) в момент $\tau = 0$ выхода заявки из прибора j , $j = 1, 2$, $k = \overline{2, r+2}$, $s = \overline{1, l}$, $j_{3-j} = \overline{1, m_{3-j}}$, $i = 1, 2$, и введем векторы:

$$\begin{aligned} \pi_{D,j}^{-T}(1, j) &= \left(\pi_{D,j}^-(1, j), \dots, \pi_{D,j}^-(l, j) \right); \\ \pi_{D,j}^{-T}(k, i) &= \left(\pi_{D,j}^-(k, 1, 1, i), \dots \right. \\ &\quad \left. \dots, \pi_{D,j}^-(k, 1, 2, i), \dots, \pi_{D,j}^-(k, l, m_{3-j}, i) \right), \\ &\quad j = 1, 2, \quad i = 1, 2, \quad k = \overline{2, r+2}. \end{aligned}$$

Тогда на основании вышеизложенного и с учетом [1, с. 104] получаем

$$\begin{aligned} f_1(s) &= \sum_{k=1}^{r+2} \pi_{D,1}^{-T}(k, 1) \mathbf{1} + \\ &+ \sum_{k=2}^{r+2} \pi_{D,1}^{-T}(k, 2) (\mathbf{1} \otimes \mathbf{I}) (s\mathbf{I} - M_2)^{-1} \boldsymbol{\mu}_2; \quad (16) \end{aligned}$$

$$\begin{aligned} f_2(s) &= \sum_{k=1}^{r+2} \pi_{D,2}^{-T}(k, 2) \mathbf{1} + \\ &+ \sum_{k=2}^{r+2} \pi_{D,1}^{-T}(k, 1) (\mathbf{1} \otimes \mathbf{I}) (s\mathbf{I} - M_1)^{-1} \boldsymbol{\mu}_1. \quad (17) \end{aligned}$$

Для определения стационарных вероятностей $\pi_{D,j}^-(k, j_{3-j}, i)$, $\pi_{D,j}^-(s, j)$, $j = 1, 2$, $k = \overline{2, r+2}$, $s = \overline{1, l}$, $j_{3-j} = \overline{1, m_{3-j}}$, $i = 1, 2$, воспользуемся результатами [4], согласно которым

$$\pi_{D,j}^{-T}(1, j) = \frac{1}{\lambda_D(j)} \mathbf{p}_{1j}^T (\mathbf{I} \otimes \boldsymbol{\mu}_j), \quad j = 1, 2; \quad (18)$$

$$\begin{aligned} \pi_{D,1}^{-T}(k, i) &= \frac{1}{\lambda_D(1)} \mathbf{p}_{ki}^T (\mathbf{I} \otimes \boldsymbol{\mu}_1 \otimes \mathbf{I}), \\ &\quad k = \overline{2, r+2}, \quad i = 1, 2; \quad (19) \end{aligned}$$

$$\begin{aligned} \pi_{D,2}^{-T}(k, i) &= \frac{1}{\lambda_D(2)} \mathbf{p}_{ki}^T (\mathbf{I} \otimes \mathbf{I} \otimes \boldsymbol{\mu}_2), \\ &\quad k = \overline{2, r+2}, \quad i = 1, 2, \quad (20) \end{aligned}$$

где $\lambda_D(j)$ — интенсивность выхода заявок, обслуженных прибором j , определяемая выражением

$$\begin{aligned} \lambda_D(j) &= \mathbf{p}_{1j}^T (\mathbf{1} \otimes \boldsymbol{\mu}_j) + \\ &+ \sum_{k=2}^{r+2} \mathbf{p}_{k.}^T [u(2-j)(\mathbf{1} \otimes \boldsymbol{\mu}_1 \otimes \mathbf{1}) + \\ &\quad + u(j-1)(\mathbf{1} \otimes \mathbf{1} \otimes \boldsymbol{\mu}_2)], \quad j = 1, 2. \end{aligned}$$

Далее заметим, что вероятность выхода заявок из прибора j равна $\lambda_D(j)/\lambda_D$, $j = 1, 2$, где интенсивность выходящего из системы потока

$$\lambda_D = \lambda_D(1) + \lambda_D(2).$$

Таким образом, подытоживая рассуждения, получаем, что справедлива

Теорема 2. Преобразование Лапласа–Стилтьеса ФР задержки переупорядочивания в СМО РН/РН/2/г/рес в стационарном режиме ее работы определяется выражением

$$f(s) = \frac{1}{\lambda_D} [\lambda_D(1)f_1(s) + \lambda_D(2)f_2(s)],$$

где $f_j(s)$, $j = 1, 2$, задаются формулами (16) и (17).

Обозначим через δ_ν начальный момент порядка ν , $\nu = 1, 2, \dots$, задержки переупорядочивания заявки в исследуемой СМО. Тогда из теоремы 2 с учетом (18)–(20) и [1, с. 104] получаем очевидное

Следствие. Начальный момент порядка ν задержки переупорядочивания в СМО РН/РН/2/г/рес определяется выражением

$$\begin{aligned} \delta_\nu &= \frac{(-1)^\nu \nu!}{\lambda_D} \sum_{k=2}^{r+2} [\mathbf{p}_{k2}^T (\mathbf{1} \otimes \boldsymbol{\mu}_1 \otimes M_2^{-\nu} \mathbf{1}) + \\ &\quad + \mathbf{p}_{k1}^T (\mathbf{1} \otimes M_1^{-\nu} \mathbf{1} \otimes \boldsymbol{\mu}_2)]. \quad (21) \end{aligned}$$

Теперь остановимся на связи средней задержки переупорядочивания со средним числом заявок в БП.

Теорема 3. Средняя величина задержки переупорядочивания δ_1 и среднее число заявок v_1 в БП СМО $PH/PH/2/r/res$ связаны соотношением

$$\delta_1 = \frac{1}{\lambda_d} v_1. \quad (22)$$

Доказательство. Умножим уравнения системы (12)–(14) при фиксированном $i = 1, 2$ и $\nu = 1$ справа на матрицу Z_i :

$$Z_i = \begin{cases} \mathbf{1} \otimes I \otimes \mathbf{1}, & i = 1, \\ \mathbf{1} \otimes \mathbf{1} \otimes I, & i = 2. \end{cases}$$

В результате умножения получим следующую систему:

$$\mathbf{0}^T = \mathbf{v}_{1i1}^T (-L_{1i} + U_i) + (\mathbf{v}_{2i1}^T + \mathbf{p}_{2i}^T) N_i, \quad i = 1, 2; \quad (23)$$

$$\begin{aligned} \mathbf{0}^T = & u(3 - k) \mathbf{v}_{1i1}^T L_{1i} + u(k - 2) \mathbf{v}_{k-1,i1}^T L_{k-1,i} + \\ & + \mathbf{v}_k^T (-L_{ki} + U_i - N_i) + (\mathbf{v}_{k+1,i1}^T + \mathbf{p}_{k+1,i}^T) N_i, \\ & i = 1, 2, \quad k = \overline{2, r+1}; \quad (24) \end{aligned}$$

$$\mathbf{0}^T = \mathbf{v}_{r+1}^T L_{r+1,i} + \mathbf{v}_{r+2}^T (U_i - N_i), \quad i = 1, 2, \quad (25)$$

где

$$\begin{aligned} L_{ki} &= \begin{cases} \lambda_k^T \otimes I \otimes \mathbf{1}, & i = 1, \\ \lambda_k^T \otimes \mathbf{1} \otimes I, & i = 2, \quad k = \overline{1, r+1}; \end{cases} \\ U_i &= \begin{cases} \mathbf{1} \otimes M_1 \otimes \mathbf{1}, & i = 1, \\ \mathbf{1} \otimes \mathbf{1} \otimes M_2, & i = 2; \end{cases} \\ N_i &= \begin{cases} \mathbf{1} \otimes I \otimes \mu_2, & i = 1, \\ \mathbf{1} \otimes \mu_1 \otimes I, & i = 2. \end{cases} \end{aligned}$$

Последовательно суммируя уравнения системы (23)–(25), после несложных алгебраических

преобразований приходим к следующим соотношениям:

$$-\sum_{k=1}^{r+2} \mathbf{v}_{ki1}^T U_i = \sum_{k=2}^{r+2} \mathbf{p}_{ki}^T N_i, \quad i = 1, 2. \quad (26)$$

Умножим (26) при фиксированном $i = 1, 2$ справа на матрицу

$$T_i = \begin{cases} I \otimes M_1^{-1} \mathbf{1} \otimes I, & i = 1 \\ I \otimes I \otimes M_2^{-1} \mathbf{1}, & i = 2, \end{cases}$$

и просуммируем по i полученные выражения. В результате, учитывая (10) и (11), приходим к следующему равенству:

$$\begin{aligned} v_1 = & -\sum_{k=2}^{r+2} [\mathbf{p}_{k1}^T (\mathbf{1} \otimes M_1^{-1} \mathbf{1} \otimes \mu_2) + \\ & + \mathbf{p}_{k2}^T (\mathbf{1} \otimes \mu_1 \otimes M_2^{-1} \mathbf{1})]. \quad (27) \end{aligned}$$

Из сравнения (27) и (21) при $\nu = 1$ очевидным образом вытекает (22). Таким образом, теорема доказана.

Заметим, что (22) является аналогом известной формулы Литтла и имеет вполне очевидную физическую интерпретацию.

Литература

1. Бочаров П. П., Печинкин А. В. Теория массового обслуживания. — М.: РУДН, 1995.
2. Матюшенко С. И. Анализ двухканальной системы обслуживания ограниченной емкости с буфером переупорядочивания и с распределениями фазового типа // Вестник РУДН: Прикладная математика. Информатика. Физика, 2010. № 4. С. 84–88.
3. Наумов В. А. Численные методы анализа марковских систем. — М.: УДН, 1985.
4. Наумов В. А. О предельных вероятностях полумарковского процесса // Современные задачи в точных науках. — М.: УДН, 1975. С. 35–39.

АППРОКСИМАЦИЯ РАСПРЕДЕЛЕНИЯ ОЦЕНКИ РИСКА ПОРОГОВОЙ ОБРАБОТКИ ВЕЙВЛЕТ-КОЭФФИЦИЕНТОВ НОРМАЛЬНЫМ РАСПРЕДЕЛЕНИЕМ ПРИ ИСПОЛЬЗОВАНИИ ВЫБОРОЧНОЙ ДИСПЕРСИИ*

О. В. Шестаков¹

Аннотация: Исследуются асимптотические свойства оценки риска при пороговой обработке коэффициентов вейвлет-разложения функции сигнала. Получены некоторые оценки скорости сходимости распределения оценки риска к нормальному закону.

Ключевые слова: вейвлеты; пороговая обработка; оценка риска; нормальное распределение; оценка скорости сходимости

1 Введение

Популярность методов вейвлет-обработки сигналов и изображений значительно возросла за последние десятилетия. Объясняется это тем, что вейвлет-разложение представляет собой удобный математический аппарат, способный решать те задачи, в которых применение традиционного Фурье-анализа оказывается неэффективным. В настоящее время вейвлеты применяются в самых разнообразных областях, включая геофизику, оптику, физику плазмы, вычислительную томографию, компьютерную графику и т. д.

Основные задачи, для решения которых используются вейвлеты, — это сжатие сигналов/изображений и удаление шума. При этом чаще всего используется пороговая обработка вейвлет-коэффициентов, которая обнуляет коэффициенты, не превышающие заданного порога. Наличие шума и процедуры пороговой обработки неизбежно приводят к погрешностям в оцениваемом сигнале/изображении. Свойства оценки таких погрешностей (риска) исследовались во многих работах (см., например, [1–7]). При определенных условиях оценка риска является асимптотически нормальной (см. [6]).

В данной работе исследуется вопрос о скорости сходимости распределения оценки риска к нормальному закону в одномерном случае (т. е. при обработке одномерных сигналов).

2 Свойства коэффициентов вейвлет-разложения

При использовании вейвлет-разложения функция $f \in L^2(\mathbf{R})$, описывающая сигнал, представляется в виде ряда из сдвигов и растяжений некоторой вейвлет-функции ψ :

$$f = \sum_{j,k \in \mathbf{Z}} \langle f, \psi_{j,k} \rangle \psi_{j,k}, \quad (1)$$

где $\psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k)$, а семейство $\{\psi_{j,k}\}_{j,k \in \mathbf{Z}}$ образует ортонормированный базис в $L^2(\mathbf{R})$. Индекс j в (1) называется масштабом, а индекс k — сдвигом. Функция ψ должна удовлетворять определенным требованиям [8], однако ее можно выбрать таким образом, чтобы она обладала некоторыми полезными свойствами, например была дифференцируемой нужное число раз и имела заданное число M нулевых моментов (см. [8]):

$$\int_{-\infty}^{\infty} x^k \psi(x) dx = 0, \quad k = 0, \dots, M-1.$$

В дальнейшем будут рассматриваться функции сигнала $f \in L^2(\mathbf{R})$ на конечном отрезке $[a, b]$, равномерно регулярные по Липшицу с некоторым параметром $\gamma > 0$, т. е. такие функции, для которых существует константа $L > 0$ и полином P_γ степени

* Работа выполнена при финансовой поддержке РФФИ (грант 08-01-00567).

¹ Московский государственный университет им. М. В. Ломоносова, кафедра математической статистики факультета вычислительной математики и кибернетики, oshestakov@cs.msu.su

$n = \lfloor \gamma \rfloor$ такой, что для любого $y \in [a, b]$ и любого $x \in \mathbf{R}$

$$|f(x) - P_y(x)| \leq L |x - y|^\gamma.$$

Для таких функций f известно [9], что если вейвлет-функция M раз непрерывно дифференцируема ($M \geq \gamma$), имеет M нулевых моментов и быстро убывает на бесконечности вместе со своими производными, т. е. для всех $0 \leq k \leq M$ и любого $m \in N$ найдется константа C_m , что при всех $x \in \mathbf{R}$

$$|\psi^{(k)}(x)| \leq \frac{C_m}{1 + |x|^m},$$

то найдется такая константа $A > 0$, что

$$\langle f, \psi_{j,k} \rangle \leq \frac{A}{2^{j(\gamma+1/2)}}. \quad (2)$$

На практике функции сигнала всегда заданы в дискретных отсчетах на конечном отрезке. Не ограничивая общности, будем считать, что это отрезок $[0, 1]$ и функция f задана в точках i/N ($i = 1, \dots, N$, где $N = 2^J$ для некоторого J): $f_i = f(i/N)$. Дискретное вейвлет-преобразование представляет собой умножение вектора значений функции f (обозначим его через \bar{f}) на ортогональную матрицу W , определяемую вейвлет-функцией ψ : $\bar{f}^W = W\bar{f}$ (см. [9]). При этом если перейти к двойному индексу (j, k) , как в непрерывном случае, то дискретные вейвлет-коэффициенты будут связаны с непрерывными следующим образом: $f_{j,k}^W \approx \sqrt{N} \langle f, \psi_{j,k} \rangle$ (см., например, [2] или [10]). Это приближение тем точнее, чем больше N . Не будем обсуждать методы борьбы с краевыми эффектами, связанными с использованием вейвлет-разложения на конечном отрезке. Познакомиться с этими методами можно, например, в [11]. В дальнейшем для удобства будем нумеровать дискретные вейвлет-коэффициенты так же, как отсчеты функции f , одним индексом i вместо двойного индекса (j, k) .

В реальных наблюдениях всегда присутствует шум. Будем рассматривать следующую модель:

$$Y_i = f_i + z_i, \quad i = 1, \dots, N,$$

где z_i — независимые случайные величины, имеющие нормальное распределение с нулевым средним и дисперсией σ^2 . Тогда в силу ортогональности матрицы W для дискретных вейвлет-коэффициентов примем следующую модель:

$$Y_i^W = f_i^W + z_i^W, \quad i = 1, \dots, N,$$

где z_i^W также независимы и нормально распределены с нулевым средним и дисперсией σ^2 , а f_i^W равны соответствующим непрерывным вейвлет-коэффициентам, умноженным на \sqrt{N} .

3 Пороговая обработка и оценка риска

Смысл пороговой обработки вейвлет-коэффициентов заключается в удалении достаточно маленьких коэффициентов, которые считаются шумом. Будем использовать так называемую мягкую пороговую обработку с порогом T . К каждому вейвлет-коэффициенту применяется функция $\rho_T(x) = \text{sgn}(x)(|x| - T)_+$, т. е. при такой пороговой обработке коэффициенты, которые по модулю меньше порога T , обнуляются, а абсолютные величины остальных коэффициентов уменьшаются на величину порога. Погрешность (или риск) мягкой пороговой обработки определяется следующим образом:

$$R_N(f) = \sum_{i=1}^N \mathbf{E} (f_i^W - \rho_T(Y_i^W))^2. \quad (3)$$

В выражении (3) присутствуют неизвестные величины f_i^W , поэтому вычислить значение $R_N(f)$ нельзя. Однако его можно оценить. В каждом слагаемом если $|Y_i^W| > T$, то вклад этого слагаемого в риск составляет $\sigma^2 + T^2$, а если $|Y_i^W| \leq T$, то вклад составляет $(f_i^W)^2$. Поскольку $\mathbf{E}(Y_i^W)^2 = \sigma^2 + (f_i^W)^2$, величину $(f_i^W)^2$ можно оценить разностью $(Y_i^W)^2 - \sigma^2$.

Таким образом, в качестве оценки риска можно использовать следующую величину:

$$\tilde{R}_N(f) = \sum_{i=1}^N F[(Y_i^W)^2], \quad (4)$$

где

$$F[x] = (x - \sigma^2)\mathbf{1}_{|x| \leq T} + (\sigma^2 + T^2)\mathbf{1}_{|x| > T}.$$

Для так определенной оценки риска справедливо следующее утверждение (см. [9]).

Теорема 1. $\mathbf{E}\tilde{R}_N(f) = R_N(f)$, т. е. $\tilde{R}_N(f)$ является несмещенной оценкой для $R_N(f)$.

В работах [1] и [3] было предложено использовать порог $T = \sigma\sqrt{2 \ln N}$. Было показано, что при таком пороге риск близок к минимальному (см. [1]). Этот порог получил название «универсальный». В дальнейшем будем ориентироваться именно на такой вид порога.

Зачастую дисперсия σ^2 не известна и ее также необходимо оценивать, при этом выражения (4) принимают вид

$$\hat{R}_N(f) = \sum_{i=1}^N \hat{F}[(Y_i^W)^2], \quad (5)$$

где

$$\widehat{F}[x] = (x - \hat{\sigma}^2) \mathbf{1}_{|x| \leq \hat{T}} + (\hat{\sigma}^2 + \hat{T}^2) \mathbf{1}_{|x| > \hat{T}},$$

а

$$\hat{T} = \hat{\sigma} \sqrt{2 \ln N}.$$

Обычно дисперсия σ^2 оценивается по выборке сигнала, однако ее можно оценить и по независимой выборке. Для этого следует произвести измерение пустого сигнала, тогда наблюдения будут представлять собой чистый шум, по которому и оценивается σ^2 . В следующих разделах будут рассмотрены оба случая.

4 Оценка скорости сходимости распределения оценки риска к нормальному закону

В работах [6, 7] исследуется поведение оценки риска $\widehat{R}_N(f)$ при использовании различных оценок дисперсии $\hat{\sigma}^2$. В частности, показано, что при достаточно общих требованиях к $\hat{\sigma}^2$ оценка риска является асимптотически нормальной. В этом разделе будут получены некоторые оценки скорости сходимости распределения $\widehat{R}_N(f)$ к нормальному закону при использовании в качестве $\hat{\sigma}^2$ выборочной дисперсии, которая строится по независимой выборке (Y'_1, \dots, Y'_N) :

$$\hat{\sigma}^2 = \frac{1}{N-1} \sum_{i=1}^N (Y'_i - \bar{Y}')^2, \text{ где } \bar{Y}' = \frac{1}{N} \sum_{i=1}^N Y'_i.$$

Далее для удобства будем обозначать Y_i^W через X_i , а f_i^W — через a_i .

Теорема 2. Пусть $f \in L^2(\mathbf{R})$ задана на отрезке $[0, 1]$ и является равномерно регулярной по Липшицу с параметром $\gamma = \frac{1}{2} + \alpha$ ($\alpha > 0$) и пусть выборочная дисперсия $\hat{\sigma}^2$ не зависит от наблюдений X_i , тогда существует такая константа C_0 (зависящая от α , A и σ), что

$$\sup_{x \in \mathbf{R}} \left| P \left(\frac{\widehat{R}_N(f) - R_N(f)}{\hat{\sigma}^2 \sqrt{2N}} < x \right) - \Phi_2(x) \right| \leq \frac{C_0 (\ln N)^{1/2}}{N^{1/4-1/(4(\alpha+1))}}, \quad (6)$$

где $\Phi_2(x)$ — функция распределения нормального закона с нулевым средним и дисперсией, равной 2.

Доказательство. Пользуясь теоремой 1, запишем разность $\widehat{R}_N(f) - R_N(f)$ в виде

$$\widehat{R}_N(f) - R_N(f) = S_N + V_N,$$

где

$$S_N = \sum_{i=1}^N \left(X_i^2 \mathbf{1}_{|X_i| \leq \hat{T}} - \mathbf{E} X_i^2 \mathbf{1}_{|X_i| \leq \hat{T}} \right) + 2 \sum_{i=1}^N \left(\hat{\sigma}^2 \mathbf{1}_{|X_i| > \hat{T}} - \mathbf{E} \sigma^2 \mathbf{1}_{|X_i| > \hat{T}} \right) + \sum_{i=1}^N \left(\hat{T}^2 \mathbf{1}_{|X_i| > \hat{T}} - \mathbf{E} T^2 \mathbf{1}_{|X_i| > \hat{T}} \right),$$

а

$$V_N = N (\sigma^2 - \hat{\sigma}^2).$$

Рассмотрим S_N . Разобьем это слагаемое на три суммы U_N , W_N и Z_N :

$$U_N = \sum_{i \in I_1} (X_i^2 - \mathbf{E} X_i^2);$$

$$W_N = - \sum_{i \in I_1} \left(X_i^2 \mathbf{1}_{|X_i| > \hat{T}} - \mathbf{E} X_i^2 \mathbf{1}_{|X_i| > \hat{T}} \right) + 2 \sum_{i \in I_1} \left(\hat{\sigma}^2 \mathbf{1}_{|X_i| > \hat{T}} - \mathbf{E} \sigma^2 \mathbf{1}_{|X_i| > \hat{T}} \right) + \sum_{i \in I_1} \left(\hat{T}^2 \mathbf{1}_{|X_i| > \hat{T}} - \mathbf{E} T^2 \mathbf{1}_{|X_i| > \hat{T}} \right);$$

$$Z_N = \sum_{i \in I_2} \left(X_i^2 \mathbf{1}_{|X_i| \leq \hat{T}} - \mathbf{E} X_i^2 \mathbf{1}_{|X_i| \leq \hat{T}} \right) + 2 \sum_{i \in I_2} \left(\hat{\sigma}^2 \mathbf{1}_{|X_i| > \hat{T}} - \mathbf{E} \sigma^2 \mathbf{1}_{|X_i| > \hat{T}} \right) + \sum_{i \in I_2} \left(\hat{T}^2 \mathbf{1}_{|X_i| > \hat{T}} - \mathbf{E} T^2 \mathbf{1}_{|X_i| > \hat{T}} \right),$$

где I_1 — множество тех i , для которых в силу (2) выполнено $|a_i| \leq A/N^\alpha$, а I_2 — множество остальных i . Оценим сумму $W_N + Z_N$.

При произвольном $\varepsilon > 0$, используя неравенство Чебышева, имеем

$$P(|W_N + Z_N| > \varepsilon) \leq \frac{\mathbf{E}|W_N| + \mathbf{E}|Z_N|}{\varepsilon}.$$

Рассмотрим $\mathbf{E}|Z_N|$.

$$\mathbf{E}|Z_N| \leq \sum_{i \in I_2} \mathbf{E} \left| X_i^2 \mathbf{1}_{|X_i| \leq \hat{T}} - \mathbf{E} X_i^2 \mathbf{1}_{|X_i| \leq \hat{T}} \right| + 2 \sum_{i \in I_2} \mathbf{E} \left| \hat{\sigma}^2 \mathbf{1}_{|X_i| > \hat{T}} - \mathbf{E} \sigma^2 \mathbf{1}_{|X_i| > \hat{T}} \right| + \sum_{i \in I_2} \mathbf{E} \left| \hat{T}^2 \mathbf{1}_{|X_i| > \hat{T}} - \mathbf{E} T^2 \mathbf{1}_{|X_i| > \hat{T}} \right|.$$

Поскольку f регулярна по Липшицу с $\gamma = 1/2 + \alpha$, число слагаемых в каждой из этих трех сумм не превосходит $B_1 N^{1/(2(\alpha+1))}$, где B_1 — некоторая константа, зависящая от α . Слагаемые в первой и третьей суммах не превосходят $\sigma^2 \ln N$, а слагаемые во второй сумме не превосходят σ^2 . Следовательно, $E|Z_N|$ не превосходит $B_2 N^{1/(2(\alpha+1))} \ln N$ для некоторой константы B_2 .

Оценим теперь $E|W_N|$.

$$\begin{aligned} E|W_N| &\leq \sum_{i \in I_1} E \left| X_i^2 \mathbb{1}_{|X_i| > \hat{T}} - EX_i^2 \mathbb{1}_{|X_i| > T} \right| + \\ &+ 2 \sum_{i \in I_1} E \left| \hat{\sigma}^2 \mathbb{1}_{|X_i| > \hat{T}} - E\sigma^2 \mathbb{1}_{|X_i| > T} \right| + \\ &+ \sum_{i \in I_1} E \left| \hat{T}^2 \mathbb{1}_{|X_i| > \hat{T}} - ET^2 \mathbb{1}_{|X_i| > T} \right|. \quad (7) \end{aligned}$$

Оценим первую сумму. Имеем

$$\begin{aligned} E \left| X_i^2 \mathbb{1}_{|X_i| > \hat{T}} - EX_i^2 \mathbb{1}_{|X_i| > T} \right| &\leq \\ &\leq E \left| X_i^2 \mathbb{1}_{|X_i| > \hat{T}} - X_i^2 \mathbb{1}_{|X_i| > T} \right| + \\ &+ E \left| X_i^2 \mathbb{1}_{|X_i| > T} - EX_i^2 \mathbb{1}_{|X_i| > T} \right|. \quad (8) \end{aligned}$$

Рассмотрим первое слагаемое. В силу независимости X_i и \hat{T} при достаточно больших N (таких, что $T - a_i > 0$)

$$\begin{aligned} E \left| X_i^2 \mathbb{1}_{|X_i| > \hat{T}} - X_i^2 \mathbb{1}_{|X_i| > T} \right| &= \\ &= EX_i^2 \left| \mathbb{1}_{\hat{T} > |X_i| > T} + \mathbb{1}_{T \geq |X_i| > \hat{T}} \right| \leq \\ &\leq E\hat{T}^2 \mathbb{1}_{\hat{T} > |X_i| > T} + ET^2 \mathbb{1}_{T \geq |X_i| > \hat{T}} \leq \\ &\leq \frac{1}{\sqrt{2\pi\sigma^2}} \left(E\hat{T}^2 e^{-(T-a_i)^2/(2\sigma^2)} |\hat{T} - T| + \right. \\ &+ ET^2 |\hat{T} - T| \mathbb{1}_{\hat{T} \leq a_i} + \\ &+ \left. ET^2 e^{-(\hat{T}-a_i)^2/(2\sigma^2)} |\hat{T} - T| \mathbb{1}_{\hat{T} > a_i} \right). \end{aligned}$$

Обозначим $\beta_{i,N} = a_i/(\sigma\sqrt{2\ln N})$. Тогда начиная с некоторого N

$$\begin{aligned} ET^2 e^{-(\hat{T}-a_i)^2/(2\sigma^2)} |\hat{T} - T| \mathbb{1}_{\hat{T} > a_i} &= \\ &= 2\sqrt{2}\sigma^2 (\ln N)^{3/2} E|\hat{\sigma} - \sigma| e^{-\ln N(\hat{\sigma}/\sigma - \beta_{i,N})^2} \mathbb{1}_{\hat{T} > a_i} = \\ &= 2\sqrt{2}\sigma^2 (\ln N)^{3/2} E|\hat{\sigma} - \sigma| e^{-\ln N(\hat{\sigma}/\sigma - \beta_{i,N})^2} \times \\ &\times \left(\mathbb{1}_{\hat{\sigma}/\sigma \geq 1/\sqrt{2} + \beta_{i,N}} + \mathbb{1}_{\beta_{i,N} < \hat{\sigma}/\sigma < 1/\sqrt{2} + \beta_{i,N}} \right) \leq \\ &\leq 2\sqrt{2}\sigma^2 (\ln N)^{3/2} \left(e^{-\ln N/2} E|\hat{\sigma} - \sigma| + \right. \\ &+ \left. E|\hat{\sigma} - \sigma| \mathbb{1}_{\beta_{i,N} < \hat{\sigma}/\sigma < 1/\sqrt{2} + \beta_{i,N}} \right) \leq \end{aligned}$$

$$\begin{aligned} &\leq 2\sqrt{2}\sigma (\ln N)^{3/2} \left(\frac{\sqrt{D\hat{\sigma}^2}}{N^{1/2}} + \right. \\ &+ \left. \sqrt{D\hat{\sigma}^2 \mathbf{P} \left(\beta_{i,N} < \frac{\hat{\sigma}}{\sigma} < \frac{1}{\sqrt{2}} + \beta_{i,N} \right)} \right). \end{aligned}$$

При переходе к последнему неравенству использовалось то, что

$$E|\hat{\sigma} - \sigma|^2 = E \frac{(\hat{\sigma}^2 - \sigma^2)^2}{(\hat{\sigma} + \sigma)^2} \leq E \frac{(\hat{\sigma}^2 - \sigma^2)^2}{\sigma^2} = \frac{D\hat{\sigma}^2}{\sigma^2}.$$

Оценим вероятность, входящую в подкоренное выражение.

$$\begin{aligned} \mathbf{P} \left(\beta_{i,N} < \frac{\hat{\sigma}}{\sigma} < \frac{1}{\sqrt{2}} + \beta_{i,N} \right) &\leq \\ &\leq \mathbf{P} \left(\frac{\hat{\sigma}}{\sigma} < \frac{1}{\sqrt{2}} + \beta_{i,N} \right) = \\ &= \mathbf{P} \left(\frac{\hat{\sigma}^2}{\sigma^2} < \left(\frac{1}{\sqrt{2}} + \beta_{i,N} \right)^2 \right) = \\ &= \mathbf{P} \left(\sigma^2 - \hat{\sigma}^2 > \sigma^2 \left(1 - \left(\frac{1}{\sqrt{2}} + \beta_{i,N} \right)^2 \right) \right) \leq \\ &\leq \mathbf{P} \left(|\sigma^2 - \hat{\sigma}^2| > \sigma^2 \left(1 - \left(\frac{1}{\sqrt{2}} + \beta_{i,N} \right)^2 \right) \right) \leq \\ &\leq \frac{D\hat{\sigma}^2}{\sigma^4 \left(1 - \left(\frac{1}{\sqrt{2}} + \beta_{i,N} \right)^2 \right)^2}. \end{aligned}$$

Последнее неравенство справедливо, поскольку $1 - (1/\sqrt{2} + \beta_{i,N})^2 > 0$ начиная с некоторого N . В результате, учитывая порядок дисперсии оценки $\hat{\sigma}^2$ [12], имеем, что для некоторой константы C_1 справедливо

$$ET^2 e^{-(\hat{T}-a_i)^2/(2\sigma^2)} |\hat{T} - T| \mathbb{1}_{\hat{T} > a_i} \leq \frac{C_1 (\ln N)^{3/2}}{N}. \quad (9)$$

Далее, начиная с некоторого N

$$\begin{aligned} ET^2 |\hat{T} - T| \mathbb{1}_{\hat{T} \leq a_i} &= \\ &= 2\sqrt{2}\sigma^3 (\ln N)^{3/2} E \left| \frac{\hat{\sigma}}{\sigma} - 1 \right| \mathbb{1}_{\hat{\sigma}/\sigma \leq \beta_{i,N}} \leq \\ &\leq 2\sqrt{2}\sigma^3 (\ln N)^{3/2} E \mathbb{1}_{\hat{\sigma}/\sigma \leq \beta_{i,N}} = \\ &= 2\sqrt{2}\sigma^3 (\ln N)^{3/2} \mathbf{P}(\sigma^2 - \hat{\sigma}^2 \geq \sigma^2(1 - \beta_{i,N}^2)) \leq \\ &\leq 2\sqrt{2} (\ln N)^{3/2} \frac{D\hat{\sigma}^2}{\sigma(1 - \beta_{i,N}^2)^2} \leq \frac{C_2 (\ln N)^{3/2}}{N} \quad (10) \end{aligned}$$

для некоторой константы C_2 .

Наконец, для некоторых констант C_3 и C_4

$$\begin{aligned} E\hat{T}^2 e^{-(T-a_i)^2/(2\sigma^2)} |\hat{T} - T| &\leq \\ &\leq \frac{C_3(\ln N)^{3/2}}{N} E\hat{\sigma}^2 |\hat{\sigma} - \sigma| \leq \\ &\leq \frac{C_3(\ln N)^{3/2}}{N\sigma} \sqrt{D\hat{\sigma}^2 E\hat{\sigma}^4} \leq \frac{C_4(\ln N)^{3/2}}{N^{3/2}}. \end{aligned} \quad (11)$$

Рассмотрим теперь второе слагаемое в (8). Если в разности оба члена положительны, то модуль разности не превосходит максимума из модулей каждого из них, поэтому

$$\begin{aligned} E|X_i^2 \mathbb{1}_{|X_i|>T} - EX_i^2 \mathbb{1}_{|X_i|>T}| &\leq E|X_i^2 \mathbb{1}_{|X_i|>T}| \leq \\ &\leq \frac{2}{\sqrt{2\pi\sigma^2}} \int_T^\infty u^2 e^{-(u-|a_i|)^2/(2\sigma^2)} du = \\ &= \frac{2}{\sqrt{2\pi\sigma^2}} \int_{T-|a_i|}^\infty (v+|a_i|)^2 e^{-v^2/(2\sigma^2)} dv = \\ &= \frac{2}{\sqrt{2\pi\sigma^2}} \int_{T-|a_i|}^\infty (v^2 + 2v|a_i| + a_i^2) e^{-v^2/(2\sigma^2)} dv; \\ &\int_{T-|a_i|}^\infty v^2 e^{-v^2/(2\sigma^2)} dv = -\sigma^2 v e^{-v^2/(2\sigma^2)} \Big|_{T-|a_i|}^\infty + \\ &+ \sigma^2 \int_{T-|a_i|}^\infty e^{-v^2/(2\sigma^2)} dv \leq \frac{C_5(\ln N)^{1/2}}{N} + \frac{C_6}{N(\ln N)^{1/2}} \end{aligned}$$

для некоторых констант C_5 и C_6 . Далее, учитывая, что $|a_i| < A/N^\alpha$ для $i \in I_1$, имеем

$$\begin{aligned} |a_i| \int_{T-|a_i|}^\infty v e^{-v^2/(2\sigma^2)} dv &= \\ &= \sigma^2 |a_i| e^{-(T-|a_i|)^2/(2\sigma^2)} \leq \frac{C_7}{N^{1+\alpha}}; \\ a_i^2 \int_{T-|a_i|}^\infty e^{-v^2/(2\sigma^2)} dv &\leq \frac{C_8}{N^{1+2\alpha}(\ln N)^{1/2}}. \end{aligned}$$

Следовательно, существует такая константа C_9 , что

$$E|X_i^2 \mathbb{1}_{|X_i|>T} - EX_i^2 \mathbb{1}_{|X_i|>T}| \leq \frac{C_9(\ln N)^{1/2}}{N}. \quad (12)$$

Объединяя (8)–(12), получаем, что существует такая константа C^* , что первая сумма в (7) не превосходит $C^*(\ln N)^{3/2}$. Вторая и третья сумма в (7)

оцениваются аналогично. Таким образом, существует такая константа C^{**} , что

$$E|W_N| \leq C^{**}(\ln N)^{3/2}. \quad (13)$$

Далее, для произвольного $\varepsilon > 0$ справедливо

$$\begin{aligned} \sup_{x \in \mathbf{R}} \left| \mathbf{P} \left(\frac{\hat{R}_N(f) - R_N(f)}{\sigma^2 \sqrt{2N}} < x \right) - \Phi_2(x) \right| &= \\ = \sup_{x \in \mathbf{R}} \left| \mathbf{P} \left(\frac{V_N + U_N + W_N + Z_N}{\sigma^2 \sqrt{2N}} < x \right) - \Phi_2(x) \right| &\leq \\ \leq \sup_{x \in \mathbf{R}} \left| \mathbf{P} \left(\frac{V_N + U_N}{\sigma^2 \sqrt{2N}} < x \right) - \Phi_2(x) \right| &+ \\ + \frac{\varepsilon}{2\sqrt{\pi}} + \mathbf{P} \left(|W_N + Z_N| > \varepsilon \sigma^2 \sqrt{2N} \right). \end{aligned}$$

Выберем $\varepsilon = (\ln N)^{1/2} N^{1/(4(\alpha+1)) - 1/4}$, тогда, учитывая оценку для $E|Z_N|$ и (13), получаем, что для некоторой константы B_3 справедливо

$$\begin{aligned} \sup_{x \in \mathbf{R}} \left| \mathbf{P} \left(\frac{\hat{R}_N(f) - R_N(f)}{\sigma^2 \sqrt{2N}} < x \right) - \Phi_2(x) \right| &\leq \\ \leq \sup_{x \in \mathbf{R}} \left| \mathbf{P} \left(\frac{V_N + U_N}{\sigma^2 \sqrt{2N}} < x \right) - \Phi_2(x) \right| &+ \\ + \frac{B_3(\ln N)^{1/2}}{N^{1/4 - 1/(4(\alpha+1))}}. \end{aligned} \quad (14)$$

Так как V_N и U_N независимы, имеем [13]

$$\begin{aligned} \sup_{x \in \mathbf{R}} \left| \mathbf{P} \left(\frac{V_N + U_N}{\sigma^2 \sqrt{2N}} < x \right) - \Phi_2(x) \right| &\leq \\ \leq \sup_{x \in \mathbf{R}} \left| \mathbf{P} \left(\frac{V_N}{\sigma^2 \sqrt{2N}} < x \right) - \Phi(x) \right| &+ \\ + \sup_{x \in \mathbf{R}} \left| \mathbf{P} \left(\frac{U_N}{\sigma^2 \sqrt{2N}} < x \right) - \Phi(x) \right|, \end{aligned} \quad (15)$$

где $\Phi(x)$ — функция распределения стандартного нормального закона. Величина $\hat{\sigma}^2(N-1)/\sigma^2$ имеет распределение χ^2 с $N-1$ степенью свободы (см. [12]). Учитывая результаты работы [14], можно показать, что

$$\sup_{x \in \mathbf{R}} \left| \mathbf{P} \left(\frac{V_N}{\sigma^2 \sqrt{2N}} < x \right) - \Phi(x) \right| \leq \frac{B_4}{\sqrt{N}} \quad (16)$$

для некоторой константы B_4 .

Для второго слагаемого, учитывая, что $DX_i^2 = 2\sigma^4 + 4\sigma^2 a_i^2$, имеем

$$\sup_{x \in \mathbf{R}} \left| \mathbf{P} \left(\frac{U_N}{\sigma^2 \sqrt{2N}} < x \right) - \Phi(x) \right| = \sup_{x \in \mathbf{R}} |\mathbf{P}(H_N \delta_N < x) - \Phi(x)|,$$

где

$$H_N = \frac{U_N}{\sigma^2 \sqrt{2N_1 + (4/\sigma^2) \sum_{i \in I_1} a_i^2}},$$

а

$$\delta_N = \frac{\sqrt{N_1 + (2/\sigma^2) \sum_{i \in I_1} a_i^2}}{\sqrt{N}}.$$

Здесь N_1 — число индексов в I_1 ($N \geq N_1 \geq N - B_1 N^{1/(2(\alpha+1))}$). Далее,

$$\begin{aligned} \sup_{x \in \mathbf{R}} |\mathbf{P}(H_N \delta_N < x) - \Phi(x)| &\leq \\ &\leq \sup_{x \in \mathbf{R}} |\mathbf{P}(H_N < x) - \Phi(x)| + \\ &+ \sup_{x \in \mathbf{R}} \left| \Phi \left(\frac{x}{\delta_N} \right) - \Phi(x) \right| \leq \\ &\leq \sup_{x \in \mathbf{R}} |\mathbf{P}(H_N < x) - \Phi(x)| + \\ &+ \frac{1}{\sqrt{2\pi e}} \max \left(\frac{1}{\delta_N} - 1, \delta_N - 1 \right). \end{aligned} \quad (17)$$

Используя тот факт, что сумма $\sum_{i \in I_1} a_i^2$ при $N \rightarrow \infty$ фактически представляет из себя остаток ряда из квадратов вейвлет-коэффициентов функции f , умноженного на N , а f принадлежит $L^2(\mathbf{R})$ и регулярна по Липшицу с $\gamma = 1/2 + \alpha$, можно показать, что существует такая константа B_5 , что

$$\max \left(\frac{1}{\delta_N} - 1, \delta_N - 1 \right) \leq \frac{B_5}{N^{1/2-1/(2(\alpha+1))}}. \quad (18)$$

Далее, поскольку $DX_i^2 = 2\sigma^4 + 4\sigma^2 a_i^2 \geq 2\sigma^4$ и величины $E|X_i^2 - EX_i^2|^3$ ограничены некоторой константой B_6 для $i \in I_1$, то в силу неравенства Берри—Эссеена для сумм разнораспределенных независимых случайных величин [15], получаем

$$\sup_{x \in \mathbf{R}} |\mathbf{P}(H_N < x) - \Phi(x)| \leq \frac{B_7}{\sqrt{N}} \quad (19)$$

для некоторой константы B_7 . Объединяя (14)–(19), получаем (6). Теорема доказана.

В (6) разность $\hat{R}_N(f) - R_N(f)$ нормируется величиной, зависящей от σ^2 . Однако, поскольку в (5) в $\hat{R}_N(f)$ вместо σ^2 подставляется $\hat{\sigma}^2$, естественнее

подставить $\hat{\sigma}^2$ и в эту нормировку. При этом из доказанной теоремы можно получить следующее следствие.

Следствие. Если при выполнении условий теоремы 2 вместо σ^2 в (6) подставить $\hat{\sigma}^2$, то для константы C_0 из теоремы 2 и некоторой константы B_0 справедливо

$$\begin{aligned} \sup_{x \in \mathbf{R}} \left| \mathbf{P} \left(\frac{\hat{R}_N(f) - R_N(f)}{\hat{\sigma}^2 \sqrt{2N}} < x \right) - \Phi_2(x) \right| &\leq \\ &\leq \frac{C_0 (\ln N)^{1/2}}{N^{1/4-1/(4(\alpha+1))}} + \frac{B_0}{N^{1/4}}. \end{aligned} \quad (20)$$

Доказательство.

$$\begin{aligned} \sup_{x \in \mathbf{R}} \left| \mathbf{P} \left(\frac{\hat{R}_N(f) - R_N(f)}{\hat{\sigma}^2 \sqrt{2N}} < x \right) - \Phi_2(x) \right| &= \\ = \sup_{x \in \mathbf{R}} \left| \mathbf{P} \left(\frac{\hat{R}_N(f) - R_N(f)}{\sigma^2 \sqrt{2N}} \frac{\sigma^2}{\hat{\sigma}^2} < x \right) - \Phi_2(x) \right|. \end{aligned}$$

Для произвольного $0 < \varepsilon < 1$

$$\begin{aligned} \sup_{x \in \mathbf{R}} \left| \mathbf{P} \left(\frac{\hat{R}_N(f) - R_N(f)}{\sigma^2 \sqrt{2N}} \frac{\sigma^2}{\hat{\sigma}^2} < x \right) - \Phi_2(x) \right| &\leq \\ &\leq \sup_{x \in \mathbf{R}} \left| \mathbf{P} \left(\frac{\hat{R}_N(f) - R_N(f)}{\sigma^2 \sqrt{2N}} < x \right) - \Phi_2(x) \right| + \\ &+ \mathbf{P} \left(\left| \frac{\sigma^2}{\hat{\sigma}^2} - 1 \right| > \varepsilon \right) + \frac{\varepsilon}{(1-\varepsilon)\sqrt{2\pi e}}. \end{aligned} \quad (21)$$

Далее,

$$\begin{aligned} \mathbf{P} \left(\left| \frac{\sigma^2}{\hat{\sigma}^2} - 1 \right| > \varepsilon \right) &\leq \frac{E|\hat{\sigma}^2 - \sigma^2|(1+\varepsilon)}{\varepsilon \sigma^2} \leq \\ &\leq \frac{\sqrt{D\hat{\sigma}^2}(1+\varepsilon)}{\varepsilon \sigma^2}. \end{aligned}$$

Выберем $\varepsilon = N^{-1/4}$. Тогда найдется такая константа B_0 , что

$$\mathbf{P} \left(\left| \frac{\sigma^2}{\hat{\sigma}^2} - 1 \right| > \varepsilon \right) + \frac{\varepsilon}{(1-\varepsilon)\sqrt{2\pi e}} \leq \frac{B_0}{N^{1/4}}. \quad (22)$$

Объединяя (21), (22) и (6), получаем (20).

Замечание 1. Хотя в утверждениях этого раздела требуется равномерная регулярность по Липшицу, когда дисперсия оценивается по независимой выборке, это требование можно ослабить, позволив функции быть разрывной в конечном числе точек, если потребовать, чтобы вейвлет-функция имела компактный носитель. При этом порядок оценок в теореме 2 и ее следствии не изменится.

5 Оценивание дисперсии по выборке сигнала

Если дисперсия оценивается по выборке сигнала и функция f удовлетворяет требуемым условиям регулярности, то в силу (2) обычно ее оценивают по половине всех вейвлет-коэффициентов для $j = J - 1$ (напомним, что $N = 2^J$), так как эти коэффициенты фактически содержат только шум. Для доказательства утверждений этого пункта будем использовать две оценки дисперсии, каждая из которых построена по половине вейвлет-коэффициентов из указанного множества, т.е. по четверти всех вейвлет-коэффициентов. Если перейти к одному индексу i , то эти оценки будут выглядеть следующим образом:

$$\hat{\sigma}_1^2 = \frac{1}{N/4 - 1} \sum_{i=N/2+1}^{3N/4} (X_i - \bar{X}_1)^2, \quad (23)$$

где

$$\bar{X}_1 = \frac{4}{N} \sum_{i=N/2+1}^{3N/4} X_i;$$

$$\hat{\sigma}_2^2 = \frac{1}{N/4 - 1} \sum_{i=3N/4+1}^N (X_i - \bar{X}_2)^2, \quad (24)$$

где

$$\bar{X}_2 = \frac{4}{N} \sum_{i=3N/4+1}^N X_i.$$

При пороговой обработке для построения порога \hat{T} будем использовать $\hat{\sigma}_1^2$ для тех наблюдений X_i , которые не зависят от $\hat{\sigma}_1^2$, и $\hat{\sigma}_2^2$ — для тех наблюдений X_i , которые не зависят от $\hat{\sigma}_2^2$. Для наблюдений, не зависящих ни от $\hat{\sigma}_1^2$, ни от $\hat{\sigma}_2^2$, будем использовать одну из этих оценок так, чтобы каждая из них использовалась одинаковое число раз. Таким образом, многие рассуждения, изложенные в теореме 2, останутся справедливыми.

Можно показать, что если f регулярна по Липшицу с параметром $\gamma = 1/2 + \alpha$, то

$$\left. \begin{aligned} E\hat{\sigma}_k^2 &= \sigma^2 + O\left(\frac{1}{N^{1+2\alpha}}\right); \\ D\hat{\sigma}_k^2 &= O\left(\frac{1}{N}\right), \quad k = 1, 2. \end{aligned} \right\} \quad (25)$$

Справедлив аналог теоремы 2.

Теорема 3. Пусть $f \in L^2(\mathbf{R})$ задана на отрезке $[0, 1]$ и является равномерно регулярной по Липшицу с параметром $\gamma = 1/2 + \alpha$ ($\alpha > 0$) и пусть оценка σ^2 задается соотношениями (23) и (24), тогда

существует такая константа \tilde{C}_0 (зависящая от α , A и σ), что

$$\sup_{x \in \mathbf{R}} \left| P \left(\frac{\hat{R}_N(f) - R_N(f)}{\sigma^2 \sqrt{2N}} < x \right) - \Phi(x) \right| \leq \frac{\tilde{C}_0 (\ln N)^{1/2}}{N^{1/4-1/(4(\alpha+1))}}. \quad (26)$$

Замечание 2. В отличие от теоремы 2, предельный закон в (26) имеет дисперсию, равную 1.

Доказательство. Так же, как в теореме 2, запишем разность $\hat{R}_N(f) - R_N(f)$ в виде $\hat{R}_N(f) - R_N(f) = V_N + U_N + W_N + Z_N$. Рассмотрим сумму $V_N + U_N$:

$$V_N + U_N = \sum_{i \in I_1} (X_i^2 - EX_i^2) + N \left(\sigma^2 - \frac{1}{2} (\hat{\sigma}_1^2 + \hat{\sigma}_2^2) \right).$$

Заметим, что индексы всех слагаемых в суммах (23) и (24) содержатся в I_1 . Обозначим множество этих индексов через I'_1 . Таким образом, имеем

$$V_N + U_N = \sum_{i \in I_1 \setminus I'_1} (X_i^2 - EX_i^2) - \sum_{i \in I'_1} (X_i^2 - EX_i^2) - 2 \sum_{i \in I'_1} a_i^2 - \frac{8}{N-4} \sum_{i \in I'_1} X_i^2 + \frac{N^2}{2(N-4)} (\bar{X}_1^2 + \bar{X}_2^2).$$

Пусть

$$U'_N = \sum_{i \in I_1 \setminus I'_1} (X_i^2 - EX_i^2) - \sum_{i \in I'_1} (X_i^2 - EX_i^2);$$

$$\Delta_N = -2 \sum_{i \in I'_1} a_i^2 - \frac{8}{N-4} \sum_{i \in I'_1} X_i^2 + \frac{N^2}{2(N-4)} (\bar{X}_1^2 + \bar{X}_2^2).$$

Так же, как в теореме 2, убеждаемся, что для некоторой константы \tilde{C}_1

$$\sup_{x \in \mathbf{R}} \left| P \left(\frac{U'_N}{\sigma^2 \sqrt{2N}} < x \right) - \Phi(x) \right| \leq \frac{\tilde{C}_1}{\sqrt{N}} + \frac{\tilde{C}_2}{N^{1/2-1/(2(\alpha+1))}}. \quad (27)$$

Далее, учитывая соотношения (25), можно оценить сумму $E|\Delta_N| + E|W_N| + E|Z_N|$ аналогично тому, как это было сделано в теореме 2 для $E|W_N| + E|Z_N|$.

Используя неравенство, аналогичное (14), и учитывая (27), получаем (26). Теорема доказана.

Из теоремы 3 можно сделать такое же следствие, как и из теоремы 2.

Следствие. Если при выполнении условий теоремы 3 вместо σ^2 в (26) подставить $\hat{\sigma}^2 = (\hat{\sigma}_1^2 + \hat{\sigma}_2^2)/2$, то для константы \tilde{C}_0 из теоремы 3 и некоторой константы \tilde{B}_0 справедливо

$$\sup_{x \in \mathbf{R}} \left| \mathbf{P} \left(\frac{\hat{R}_N(f) - R_N(f)}{\hat{\sigma}^2 \sqrt{2N}} < x \right) - \Phi(x) \right| \leq \frac{\tilde{C}_0 (\ln N)^{1/2}}{N^{1/4-1/(4(\alpha+1))}} + \frac{\tilde{B}_0}{N^{1/4}}. \quad (28)$$

Доказательство неравенства (28) аналогично доказательству следствия из теоремы 2.

Замечание 3. Из доказательств приведенных утверждений следует, что оценки на самом деле имеют более сложную структуру. Можно заметить, что если уточнить структуру оценок, то константа при главном члене не будет зависеть от σ^2 .

Замечание 4. Метод разложения суммы на две компоненты, одна из которых ведет себя как сумма независимых случайных величин, а другая стремится к нулю по вероятности, как правило, приводит к оценкам, не являющимся оптимальными по порядку. Следовательно, можно ожидать, что порядок оценок в приведенных утверждениях может быть улучшен.

Литература

1. Donoho D., Johnstone I. M. Ideal spatial adaptation via wavelet shrinkage // *Biometrika*, 1994. Vol. 81. No. 3. P. 425–455.
2. Donoho D., Johnstone I. M. Adapting to unknown smoothness via wavelet shrinkage // *J. Amer. Stat. Assoc.*, 1995. Vol. 90. P. 1200–1224.
3. Donoho D. L., Johnstone I. M., Kerkycharian G., Picard D. Wavelet shrinkage: Asymptopia? // *J. R. Statist. Soc. Ser. B.*, 1995. Vol. 57. No. 2. P. 301–369.
4. Marron J. S., Adak S., Johnstone I. M., Neumann M. H., Patil P. Exact risk analysis of wavelet regression // *J. Comput. Graph. Stat.*, 1998. Vol. 7. P. 278–309.
5. Antoniadis A., Fan J. Regularization of wavelet approximations // *J. Amer. Statist. Assoc.*, 2001. Vol. 96. No. 455. P. 939–967.
6. Маркин А. В. Предельное распределение оценки риска при пороговой обработке вейвлет-коэффициентов // *Информатика и её применения*, 2009. Т. 3. № 4. С. 57–63.
7. Маркин А. В., Шестаков О. В. О состоятельности оценки риска при пороговой обработке вейвлет-коэффициентов // *Вестн. Моск. ун-та. Сер. 15. Вычисл. матем. и киберн.*, 2010. № 1. С. 26–34.
8. Добеши И. Десять лекций по вейвлетам. — Ижевск: НИЦ Регулярная и хаотическая динамика, 2001.
9. Mallat S. A wavelet tour of signal processing. — Academic Press, 1999.
10. Abramovich F., Silverman B. W. Wavelet decomposition approaches to statistical inverse problems // *Biometrika*, 1998. Vol. 85. No. 1. P. 115–129.
11. Boggess A., Narkowich F. A first course in wavelets with Fourier analysis. — Prentice Hall, 2001.
12. Крамер Г. Математические методы статистики. — М: Мир, 1975.
13. Senatov V. V. Normal approximation: New results, methods, and problems. — VSP, 1998.
14. Ульянов В. В., Кристоф Г., Фуджикоши Я. О приближениях преобразований хи-квадрат распределений в статистических приложениях // *Сиб. матем. журн.*, 2006. Т. 47. № 6. С. 1401–1413.
15. Петров В. В. Предельные теоремы для сумм независимых случайных величин. — М: Наука, 1987.

ОЦЕНКИ СКОРОСТИ СХОДИМОСТИ РАСПРЕДЕЛЕНИЙ СЛУЧАЙНЫХ СУММ С БЕЗГРАНИЧНО ДЕЛИМЫМИ ИНДЕКСАМИ К НОРМАЛЬНОМУ ЗАКОНУ*

С. В. Гавриленко¹

Аннотация: Построены новые оценки скорости сходимости распределений случайных сумм с безгранично делимыми индексами, справедливые при существенно более широких условиях, нежели известные. Показано, что новые оценки также могут быть заметно более точными, нежели известные. В качестве примера применения этих результатов построены оценки точности нормальной аппроксимации для распределений случайных сумм с индексами, имеющими отрицательное биномиальное распределение.

Ключевые слова: случайная сумма; целочисленное безгранично делимое распределение; обобщенное пуассоновское распределение; отрицательное биномиальное распределение; нормальная аппроксимация

1 Оценки скорости сходимости распределений случайных сумм с безгранично делимыми индексами

Формально задача оценивания точности нормальной аппроксимации для распределений случайных сумм с индексами, имеющими целочисленные безгранично делимые (т.е. обобщенные пуассоновские) распределения, имеет менее общий характер по сравнению с рассмотренной, например, в работах [1–3] ситуацией, когда индекс M может иметь произвольное распределение. Тем не менее данный частный случай представляет существенный интерес с точки зрения практического применения соответствующих моделей. Так, целочисленный случайный процесс с независимыми приращениями представляет собой совокупность случайных величин M_t , имеющих целочисленное безгранично делимое (и, значит, обобщенное пуассоновское) распределение. Класс целочисленных безгранично делимых распределений включает в себя много широко применяемых представителей, в частности отрицательное биномиальное распределение и др. Задачи исследования случайных сумм с индексами, являющимися значениями упомянутых процессов, возникают при анализе систем массового обслуживания, в частности вычислительных и телекоммуникационных систем, процессов рис-

ка в страховой и финансовой математике и многих других областях.

Прежде всего введем некоторые обозначения, упрощающие запись результатов данной статьи и их доказательств.

Предположим, что все рассматриваемые в данной статье случайные величины заданы на одном и том же вероятностном пространстве (Ω, \mathcal{A}, P) . Совпадение распределений случайных величин X и Y будем обозначать $X \stackrel{d}{=} Y$. Функцию распределения и характеристическую функцию любой случайной величины X обозначим $F_X(x)$ ($x \in \mathbb{R}$) и $f_X(t)$ ($t \in \mathbb{R}$) соответственно, а производящую функцию любой неотрицательной целочисленной случайной величины M обозначим $\psi_M(z)$ ($|z| \leq 1$).

Пусть M — некоторая неотрицательная целочисленная случайная величина, X — произвольная случайная величина. Случайную величину, характеристическая функция которой равна $\psi_M(f_X(t))$, будем обозначать символом $\{M, X\}$. Несложно убедиться (см., например, [4]), что если случайная величина Z может быть представлена в виде $Z \stackrel{d}{=} \sum_{j=1}^M X_j$, где X_1, X_2, \dots — одинаково распределенные случайные величины, причем случайные M, X_1, X_2, \dots независимы, то $Z \stackrel{d}{=} \{M, X\}$, где $X \stackrel{d}{=} X_j$ (для определенности будем считать, что $\sum_{j=1}^0 X_j = 0$). Случайную величину $\{M, X\}$ будем

* Работа выполнена при поддержке Министерства образования и науки (государственный контракт 16.740.11.0133 от 02.09.2010).

¹Московский государственный университет им. М. В. Ломоносова, факультет вычислительной математики и кибернетики, gavrilenko.cmc@gmail.com

называть случайной суммой. При этом случайная величина M будет называться индексом, а случайная величина X — случайным слагаемым.

Везде далее будем считать, что случайное слагаемое X удовлетворяет условиям

$$EX = 0, \quad DX = 1. \quad (1)$$

Предположим, что распределение случайной величины M является безгранично делимым в классе распределений неотрицательных целочисленных случайных величин, т. е. для любого натурального числа n существует такая неотрицательная целочисленная случайная величина M'_n , что

$$M \stackrel{d}{=} \{n, M'_n\}.$$

Как известно (см., например, [5], гл. XII, § 3), распределение является безгранично делимым в классе распределений неотрицательных целочисленных случайных величин тогда и только тогда, когда оно является обобщенным пуассоновским, т. е. соответствующая ему характеристическая функция имеет вид

$$f_M(t) = \exp[\lambda(f_Y(t) - 1)]$$

для некоторого $\lambda > 0$, где $f_Y(t)$ — характеристическая функция некоторой неотрицательной целочисленной случайной величины Y . Другими словами,

$$M \stackrel{d}{=} \{N_\lambda, Y\}, \quad (2)$$

где N_λ — пуассоновская случайная величина с параметром λ . Сразу же обратим внимание на то, что для производящей функции случайной величины M справедливо соотношение (см. [5])

$$\psi_M(s) = \exp\{\lambda(\psi_Y(s) - 1)\}, \quad (3)$$

где $\psi_M(s)$ — это производящая функция случайной величины M , а $\psi_Y(s)$ — производящая функция случайной величины Y , той самой, что определена в (2).

Будем считать, что

$$E|X|^3 = \beta_3 < \infty \quad (4)$$

и, кроме того, существуют первые три момента случайной величины M .

Отношением Ляпунова (или ляпуновской дробью) будем называть величину

$$L(X) = \frac{E|X|^3}{(EX^2)^{3/2}}.$$

В данной статье рассмотрим оценку точности аппроксимации распределения случайной величины $S_M \stackrel{d}{=} \{M, X\}$ стандартным нормальным законом в ситуации, когда случайная величина M

имеет целочисленное безгранично делимое распределение. Очевидно, что в такой ситуации

$$S_M \stackrel{d}{=} \{\{N_\lambda, Y\}, X\}. \quad (5)$$

Центрированную и нормированную случайную сумму S_M обозначим

$$\tilde{S}_M = \frac{S_M - ES_M}{\sqrt{DS_M}}.$$

Хорошо известно, что в сделанных предположениях

$$ES_M = EX \cdot EM, \quad DS_M = EM \cdot DX + DM \cdot (EX)^2.$$

Учитывая (1), замечаем, что в рассматриваемой ситуации

$$ES_M = 0, \quad DS_M = EM, \quad (6)$$

а значит,

$$\tilde{S}_M = \frac{S_M}{\sqrt{EM}}.$$

Обозначим

$$\Delta = \sup_x |F_{\tilde{S}_M}(x) - \Phi(x)|,$$

где $\Phi(x)$ — стандартная нормальная функция распределения.

Оценки точности нормальной аппроксимации для распределений случайных сумм с произвольными индексами изучались во многих работах, см., например, [1–4] и списки литературы в этих работах. Оценки из работы [2] для случая $EX = 0$ представляются близкими к окончательным. Однако при дополнительной информации о распределении индекса эти оценки можно уточнить. В частности, в ситуации, когда случайная величина M имеет распределение Пуассона, справедлива оценка, которую можно сформулировать в виде следующей теоремы.

Теорема 1. *Если M имеет распределение Пуассона с параметром λ , а случайное слагаемое X имеет конечный третий абсолютный момент, то справедливо неравенство*

$$\Delta \leq 0,3041 \frac{L(X)}{\sqrt{\lambda}}. \quad (7)$$

Данная оценка с наилучшей на текущий момент константой 0,3041 получена в работе [6], где также описана предыстория вопроса. В данной статье оценка (7) будет обобщена на случай, когда индекс M имеет целочисленное безгранично делимое

распределение. Аналогичная ситуация рассматривалась и в работе [7]. Здесь будут построены оценки, альтернативные приведенным в статье [7] и справедливые при более широких условиях.

В статье [7] доказана теорема, которая с учетом оценки абсолютной константы, полученной в работе [6], имеет следующий вид.

Теорема 1.2. Пусть неотрицательная целочисленная случайная величина M является безгранично делимой и, кроме того, обладает тремя конечными первыми моментами. Пусть случайное слагаемое X также имеет первые три конечных момента. Тогда справедлива оценка

$$\begin{aligned} \Delta &= \sup_x |P(S_M < \sqrt{DS_M}x + ES_M) - \Phi(x)| \leq \\ &\leq \frac{0,3041}{\sqrt{\lambda}} (E[Y(Y-1)(Y-2)](E|X|)^3 + \\ &\quad + 3E[Y(Y-1)]E|X|EX^2 + \\ &\quad + EYE|X|^3) / [EY^2(EX)^2 + EYDX]^{3/2}, \quad (8) \end{aligned}$$

где λ и Y — элементы представления (2) случайной величины M .

Обратим внимание, что в теореме требуется только конечность первых трех моментов случайного слагаемого, при этом совсем не обязательно выполнение требований нулевого математического ожидания и единичной дисперсии. В случае, когда условия (1) соблюдены, имеет место

Следствие 1.1. Пусть неотрицательная целочисленная случайная величина M является безгранично делимой и, кроме того, обладает тремя конечными первыми моментами. Пусть также выполнены условия (1) и (4). Тогда

$$\begin{aligned} \Delta &= \sup_x |P(S_M < \sqrt{EM}x) - \Phi(x)| \leq \\ &\leq \frac{0,3041}{\sqrt{\lambda}} (E[Y(Y-1)(Y-2)](E|X|)^3 + \\ &\quad + 3E[Y(Y-1)]E|X| + EYE|X|^3) / (EY)^{3/2}, \quad (9) \end{aligned}$$

где λ и Y — элементы представления (2) случайной величины M .

В конце данного раздела сравним полученную оценку с оценкой (9), обсудим ее преимущества и недостатки.

Перейдем теперь непосредственно к получению оценки для случайных сумм с индексом, имеющим целочисленное безгранично делимое (т.е. обобщенное пуассоновское) распределение. Для этого понадобится несколько вспомогательных результатов.

Лемма 1.1. Пусть M — целочисленная неотрицательная случайная величина, имеющая обобщенное пуассоновское распределение. Тогда для случайной величины $S_M = \{M, X\}$ справедливы представления

$$S_M = \{M, X\} \stackrel{d}{=} \{\{N_\lambda, Y\}, X\} \stackrel{d}{=} \{N_\lambda, U\},$$

где

$$U \stackrel{d}{=} \{Y, X\}. \quad (10)$$

Доказательство. (См., например, [4].) Представление $S_M \stackrel{d}{=} \{\{N_\lambda, Y\}, X\}$ (см. (5)) вытекает из определения случайной величины M с целочисленным безгранично делимым распределением и упоминавшегося выше результата из [8] о представимости распределения каждой такой случайной величины в виде обобщенного пуассоновского. Представление $S_M \stackrel{d}{=} \{N_\lambda, \{Y, X\}\}$ легко получить из вида характеристических функций соответствующих случайных сумм (см., например, [4], с. 83):

$$f_{S_M}(t) = \psi_M(f_X(t)).$$

Напомним, что здесь $f_X(t)$ — это характеристическая функция случайного слагаемого X . С учетом представления (3) можно записать

$$\psi_M(f_X(t)) = \exp\{\lambda(\psi_Y(f_X(t)) - 1)\}.$$

Таким образом, $f_{S_M}(t)$ является характеристической функцией пуассоновской случайной суммы $\{N_\lambda, U\}$, где N_λ имеет распределение Пуассона с параметром λ , а случайная величина U имеет характеристическую функцию $\psi_Y(f_X(t))$, т.е. имеет место представление (10), где Y — случайная величина, фигурирующая в представлении (2) случайной величины M .

Лемма 1.1 позволяет представить случайную сумму с безгранично делимым индексом в виде пуассоновской случайной суммы

$$S_M \stackrel{d}{=} \{N_\lambda, U\},$$

что, в свою очередь, позволяет использовать теорему 1.1 для оценивания скорости сходимости распределения случайной величины S_M к нормальному закону. При этом согласно теореме 1.1

$$\Delta \leq \frac{0,3041}{\sqrt{\lambda}} \frac{E|U|^3}{(EU^2)^{3/2}}.$$

Оценим $E|U|^3$ и $E|U|^2$. С учетом (10) по формуле полной вероятности имеем

$$E|U|^3 = E \left| \sum_{k=0}^Y X_k \right|^3 = \sum_{n=1}^{\infty} P(Y = n) E \left| \sum_{k=1}^n X_k \right|^3. \quad (11)$$

Далее потребуется еще одна лемма.

Лемма 1.2. Пусть X_1, X_2, \dots, X_n — независимые случайные величины с математическим ожиданием, равным нулю, и конечными моментами порядка $u \geq 2$. Тогда

$$\mathbb{E} \left| \sum_{i=1}^n X_i \right|^u \leq C_u n^{u/2-1} \sum_{k=1}^n \mathbb{E}|X_k|^u,$$

где

$$C_u = \frac{u}{2} (u-1) \max(1, 2^{u-3}) \left(1 + \frac{2}{u} K_{2m}^{(u-2)/2m} \right),$$

целое число m удовлетворяет условию

$$2m \leq u < 2m + 2, \quad K_{2m} = \sum_{j=1}^m \frac{j^{2m-1}}{(j-1)!}.$$

Доказательство. Доказательство можно найти в [9] (также см. [10]).

Продолжим (11). Применим лемму 1.2 к оцениванию математического ожидания в правой части (11). Заметим, что $m = 1$ для $u = 3$, поэтому $K_{2m} = 1$ и, следовательно, $C_3 = 5$. Поэтому получаем

$$\begin{aligned} \sum_{n=1}^{\infty} P(Y = n) \mathbb{E} \left| \sum_{k=1}^n X_k \right|^3 &\leq \\ &\leq C_3 \mathbb{E}|X|^3 \sum_{n=1}^{\infty} P(Y = n) n^{3/2} = 5\beta_3 \mathbb{E}Y^{3/2}. \end{aligned} \quad (12)$$

Далее,

$$\begin{aligned} (\mathbb{E}Y^2)^{3/2} &= [DU + (\mathbb{E}Y)^2]^{3/2} = [\mathbb{E}Y \cdot DX + \\ &+ (\mathbb{E}X)^2 \cdot DY + (\mathbb{E}Y \cdot \mathbb{E}X)^2]^{3/2} = (\mathbb{E}Y)^{3/2}. \end{aligned} \quad (13)$$

Таким образом, из (12) и (13) получаем, что

$$\Delta \leq 5 \cdot 0,3041 \frac{\beta_3}{\sqrt{\lambda}} \frac{\mathbb{E}Y^{3/2}}{(\mathbb{E}Y)^{3/2}}.$$

Теперь все готово, чтобы сформулировать основной результат данного раздела.

Теорема 1.3. Пусть целочисленная неотрицательная случайная величина M является безгранично делимой, причем случайная величина Y в представлении (2) удовлетворяет условию $\mathbb{E}Y^{3/2} < \infty$. Предположим, что выполнены также условия (1) и (4). Тогда справедлива оценка

$$\begin{aligned} \Delta &= \sup_x |P(S_M < \sqrt{\mathbb{E}M} x) - \Phi(x)| \leq \\ &\leq 1,5205 \frac{\beta_3}{\sqrt{\lambda}} \frac{\mathbb{E}Y^{3/2}}{(\mathbb{E}Y)^{3/2}}. \end{aligned} \quad (14)$$

Сравним полученную оценку с результатом работы [7]. Во-первых, сразу заметим, что теорема 1.3 имеет гораздо более широкую область применимости, чем следствие 1.1, поскольку в ней требуется существование только момента $\mathbb{E}Y^{3/2}$, тогда как для справедливости следствия 1.1 требуется существование конечного третьего момента случайной величины Y .

Во-вторых, рассмотрим вопрос о том, какая из оценок точнее в ситуации, когда $\mathbb{E}Y^3 < \infty$. Для того чтобы сравнить оценки (14) и (9) в этой ситуации, очевидно, достаточно сравнить величины

$$A \equiv 5\mathbb{E}|X|^3 \mathbb{E}Y^{3/2}$$

и

$$\begin{aligned} B &\equiv \mathbb{E}[Y(Y-1)(Y-2)](\mathbb{E}|X|)^3 + \\ &+ 3\mathbb{E}[Y(Y-1)]\mathbb{E}|X| + \mathbb{E}Y\mathbb{E}|X|^3. \end{aligned}$$

Пусть X имеет двухточечное распределение

$$X = \begin{cases} 1 & \text{с вероятностью } \frac{1}{2}, \\ -1 & \text{с вероятностью } \frac{1}{2}. \end{cases}$$

Тогда, как несложно видеть,

$$\mathbb{E}X = 0, \quad \mathbb{E}|X|^k = 1$$

для любого $k = 1, 2, \dots$. Таким образом, указанное распределение X удовлетворяет требованиям (1) и (4). Вычислим A и B для данного распределения. Имеем

$$\begin{aligned} A &= 5\mathbb{E}Y^{3/2}, \\ B &= \mathbb{E}[Y(Y-1)(Y-2)] + 3\mathbb{E}[Y(Y-1)] + \mathbb{E}Y = \\ &= \mathbb{E}Y^3 - 3\mathbb{E}Y^2 + 2\mathbb{E}Y + 3\mathbb{E}Y^2 - 3\mathbb{E}Y + \mathbb{E}Y = \mathbb{E}Y^3. \end{aligned}$$

Теперь в качестве распределения неотрицательной целочисленной случайной величины Y также возьмем двухточечное распределение

$$Y = \begin{cases} n & \text{с вероятностью } \frac{1}{2}, \\ 0 & \text{с вероятностью } \frac{1}{2}, \end{cases}$$

где n — некоторое натуральное число. Легко видеть, что

$$\mathbb{E}Y^k = \frac{n^k}{2}$$

для любого $k = 1, 2, \dots$. Значит,

$$\frac{B}{A} = \frac{EY^3}{5EY^{3/2}} = \frac{n^3}{5n^{3/2}} = \frac{n^{3/2}}{5}.$$

Отсюда следует, что если $n \geq 3 > 5^{2/3}$, то $B > A$, т. е. оценка (14) лучше оценки (9). Если же $n \leq 2 < 5^{2/3}$ (т. е. $n = 1$ или $n = 2$), то оценка (9) точнее, чем (14).

Из приведенных примеров можно сделать вывод, что, вообще говоря, оценки (14) и (9) не сравнимы. Но в ряде случаев оценка (14) является лучшей. Более того, область применимости новой оценки намного шире. В отличие от оценки из работы [7], ей можно пользоваться в случаях, когда $EY^{3/2} < \infty$, но $EY^3 = \infty$.

2 Оценки точности нормальной аппроксимации для распределений отрицательных биномиальных случайных сумм

В данном разделе рассмотрим случайные суммы с конкретным целочисленным безгранично делимым индексом — случайные суммы с индексом, имеющим отрицательное биномиальное распределение (отрицательные биномиальные случайные суммы). Напомним, что случайная величина M имеет отрицательное биномиальное распределение с параметрами $r > 0$ и $p \in (0, 1)$, если

$$P(M = n) = \frac{\Gamma(r+n)}{n!\Gamma(r)} p^r (1-p)^n, \quad n = 0, 1, \dots \quad (15)$$

Отрицательные биномиальные случайные суммы находят широкое применение в качестве математических моделей в теории надежности, теории страхового риска, финансовой математике, являясь наиболее естественной моделью, например, для суммарного количества затраченных на проект средств до получения первых r положительных денежных потоков (прибылей) от него или для дохода от кредитного портфеля до первых r дефолтов по входящим в портфель обязательствам.

Так известно, отрицательное биномиальное распределение может быть представлено в виде обобщенного пуассоновского, т. е. является целочисленным безгранично делимым. Этот факт позволяет использовать результат, полученный в предыдущем разделе, с целью построения оценки точности нормальной аппроксимации для распределений отрицательных биномиальных случайных сумм.

Итак, пусть M — случайная величина, имеющая отрицательное биномиальное распределение с параметрами $r > 0$ и $p \in (0, 1)$. Как и прежде, везде далее будем требовать выполнения условий (1)

и (4). Далее потребуются некоторые вспомогательные результаты.

Лемма 2.1. Пусть M — случайная величина, имеющая отрицательное биномиальное распределение (15). Тогда M является обобщенной пуассоновской случайной величиной.

Доказательство. Этот хорошо известный результат приведен во многих источниках (см., например, [5], гл. XII, § 2). Для удобства последующего изложения воспроизведем его доказательство еще раз. Производящая функция случайной величины M имеет вид

$$\psi_M(s) = \left(\frac{p}{1 - (1-p)s} \right)^r.$$

Запишем ее немного иначе:

$$\begin{aligned} \psi_M(s) &= \exp \left\{ r \log \frac{p}{1 - (1-p)s} \right\} = \\ &= \exp \left\{ r \left(\log \frac{1}{1 - (1-p)s} + \log p \right) \right\} = \\ &= \exp \left\{ r \log \frac{1}{p} \left[\frac{1}{\log(1/p)} \cdot \log \frac{1}{1 - (1-p)s} - 1 \right] \right\}. \end{aligned}$$

Обозначим

$$\lambda = r \log \frac{1}{p}, \quad \psi(s) = \frac{r}{\lambda} \log \frac{1}{1 - (1-p)s}. \quad (16)$$

Таким образом,

$$\psi_M(s) = \exp \{ \lambda(\psi(s) - 1) \}.$$

Убедимся, что $\psi(s)$ является производящей функцией. Действительно, при разложении в ряд Тейлора получаем

$$\begin{aligned} \psi(s) &= \frac{r}{\lambda} \log \left(\frac{1}{1 - (1-p)s} \right) = \\ &= \frac{1}{\log(1/p)} \sum_{k=1}^{\infty} \frac{[(1-p)s]^k}{k}. \end{aligned}$$

Заметим, что, во-первых, $\psi(s)$ не зависит от r , а во-вторых,

$$\frac{1}{\log(1/p)} \sum_{k=1}^{\infty} \frac{(1-p)^k}{k} = 1,$$

откуда следует, что набор

$$\frac{1}{\log(1/p)} \frac{(1-p)^k}{k}, \quad k = 1, 2, \dots$$

задает дискретное вероятностное распределение (такое распределение принято называть логарифмическим или распределением логарифмического ряда [11, 12]). Поэтому

$$\psi(s) = \psi_Y(s) = \mathbb{E}s^Y, \quad (17)$$

где

$$P(Y = k) = \frac{1}{\log(1/p)} \frac{(1-p)^k}{k}, \quad k = 1, 2, \dots,$$

и

$$M \stackrel{d}{=} \{N_\lambda, Y\},$$

что и требовалось доказать. \square

Лемма 2.1 показывает, что отрицательная биномиальная случайная величина M является безгранично делимой, что дает право воспользоваться теоремой 1.3 для оценивания скорости сходимости отрицательной биномиальной случайной суммы S_M . Согласно теореме 1.3 имеем

$$\begin{aligned} \Delta &= \sup_x |P(S_M < \sqrt{\mathbb{E}Mx}) - \Phi(x)| \leq \\ &\leq 1,5205 \frac{\beta_3}{\sqrt{\lambda}} \frac{\mathbb{E}Y^{3/2}}{(\mathbb{E}Y)^{3/2}}, \end{aligned}$$

что, если вернуться к обозначениям (16), примет вид

$$\Delta \leq 1,5205 K^*(p) \frac{\beta_3}{\sqrt{r}}, \quad (18)$$

где

$$K^*(p) = \frac{\mathbb{E}Y^{3/2}}{(\mathbb{E}Y)^{3/2}} \frac{1}{\sqrt{\log(1/p)}}. \quad (19)$$

Как известно,

$$\mathbb{E}M = \frac{r(1-p)}{p},$$

что с учетом (6) дает

$$\mathbb{E}S_M = 0, \quad \mathbb{D}S_M = \frac{r(1-p)}{p}.$$

Таким образом,

$$\Delta = \Delta_r \equiv \sup_x \left| P \left(S_M < x \sqrt{\frac{r(1-p)}{p}} \right) - \Phi(x) \right|.$$

Зная распределение случайной величины Y , описанное в (17), приступим к оценке $K^*(p)$ в (18). Имеем

$$\begin{aligned} \mathbb{E}Y &= \frac{1}{\log(1/p)} \sum_{k=1}^{\infty} k \frac{(1-p)^k}{k} = \\ &= \frac{1}{\log(1/p)} \sum_{k=1}^{\infty} (1-p)^k = \frac{1-p}{p \log(1/p)}, \end{aligned}$$

$$\begin{aligned} \mathbb{E}Y^{3/2} &= \frac{1}{\log(1/p)} \sum_{k=1}^{\infty} k^{3/2} \frac{(1-p)^k}{k} = \\ &= \frac{1}{\log(1/p)} \sum_{k=1}^{\infty} \sqrt{k} (1-p)^k. \end{aligned}$$

Рассмотрим функцию $g(x) = x^{1/2}(1-p)^x$. Она возрастает на отрезке $[0, (1/2)(\log(1/(1-p)))^{-1}]$ и убывает при $x > (1/2)(\log(1/(1-p)))^{-1}$. Зная поведение $g(x)$, можно оценить $\mathbb{E}Y^{3/2}$ следующим образом. Имеем

$$\begin{aligned} \mathbb{E}Y^{3/2} &\leq \frac{1}{\log(1/p)} \left[\int_0^{\infty} x^{1/2} (1-p)^x dx + \right. \\ &\left. + (1-p)^{(1/2)(\log(1/(1-p)))^{-1}} \frac{1}{\sqrt{2}} \left(\log \frac{1}{1-p} \right)^{-1/2} \right]. \end{aligned}$$

Далее,

$$\begin{aligned} &\int_0^{\infty} x^{1/2} (1-p)^x dx = \\ &= \left(\log \frac{1}{1-p} \right)^{-1/2} \left(\log \frac{1}{1-p} \right)^{-1} \int_0^{\infty} t^{1/2} e^{-t} dt = \\ &= \left(\log \frac{1}{1-p} \right)^{-3/2} \Gamma(3/2) = \frac{\sqrt{\pi}}{2} \left(\log \frac{1}{1-p} \right)^{-3/2}, \end{aligned}$$

где $\Gamma(a)$, $a > 0$, — эйлерова гамма-функция,

$$\Gamma(a) = \int_0^{\infty} x^{a-1} e^{-x} dx.$$

Таким образом,

$$\begin{aligned} \mathbb{E}Y^{3/2} &\leq \frac{1}{\log(1/p)} \left(\log \frac{1}{1-p} \right)^{-3/2} \times \\ &\times \left[\frac{\sqrt{\pi}}{2} + \frac{1}{\sqrt{2}} (1-p)^{(1/2)(\log(1/(1-p)))^{-1}} \log \frac{1}{1-p} \right]. \end{aligned}$$

Так как

$$\log \frac{1}{1-p} \leq p \left(1 + \frac{p}{2(1-p)} \right),$$

то

$$\begin{aligned} \mathbb{E}Y^{3/2} &\leq \frac{1}{2 \log(1/p)} \left(\log \frac{1}{1-p} \right)^{-3/2} \times \\ &\times \left[\sqrt{\pi} + \frac{p}{\sqrt{2}} \left(\frac{2-p}{1-p} \right) \right]. \quad (20) \end{aligned}$$

Приведем еще одну оценку для $EY^{3/2}$. Напомним, что

$$EY^{3/2} = \frac{1}{\log(1/p)} \sum_{k=1}^{\infty} \sqrt{k}(1-p)^k.$$

Оценим сумму в правой части последнего соотношения с помощью неравенства Коши–Буняковского:

$$\begin{aligned} \sum_{k=1}^{\infty} \sqrt{k}(1-p)^k &= \sum_{k=1}^{\infty} \sqrt{k(1-p)^k} \sqrt{(1-p)^k} \leq \\ &\leq \left[\sum_{k=1}^{\infty} k(1-p)^k \sum_{k=1}^{\infty} (1-p)^k \right]^{1/2}. \end{aligned}$$

Осталось заметить, что

$$\sum_{k=1}^{\infty} (1-p)^k = \frac{1-p}{p}$$

и

$$\sum_{k=1}^{\infty} k(1-p)^k = \frac{1-p}{p^2}.$$

В итоге получаем, что

$$EY^{3/2} \leq \frac{1-p}{p^{3/2} \log(1/p)}. \quad (21)$$

Итак, получены две разные оценки для $EY^{3/2}$. Подставив (20) в выражение для $K^*(p)$, получим

$$\begin{aligned} K^*(p) &\leq \frac{1}{(\log(1/p))^{1/2}} \frac{p^{3/2}}{(1-p)^{3/2}} \left(\log \frac{1}{p}\right)^{3/2} \times \\ &\times \frac{1}{2 \log(1/p)} \left(\log \frac{1}{1-p}\right)^{-3/2} \times \\ &\times \left[\sqrt{\pi} + \frac{p}{\sqrt{2}} \left(\frac{2-p}{1-p}\right) \right]. \quad (22) \end{aligned}$$

Так как

$$\log \frac{1}{1-p} > p \left(1 + \frac{p}{2}\right),$$

то в продолжение (22) получаем

$$\begin{aligned} K^*(p) &\leq \frac{p^{3/2}}{2(1-p)^{3/2}} \left(\log \frac{1}{1-p}\right)^{-3/2} \times \\ &\times \left[\sqrt{\pi} + \frac{p}{\sqrt{2}} \left(\frac{2-p}{1-p}\right) \right] \leq \\ &\leq \frac{\sqrt{2}}{(1-p)^{3/2}(2+p)^{3/2}} \left[\sqrt{\pi} + \frac{p}{\sqrt{2}} \left(\frac{2-p}{1-p}\right) \right]. \quad (23) \end{aligned}$$

Подставив (21) в то же самое выражение, имеем

$$\begin{aligned} K^*(p) &\leq \frac{1}{(\log(1/p))^{1/2}} \frac{p^{3/2}}{(1-p)^{3/2}} \left(\log \frac{1}{p}\right)^{3/2} \times \\ &\times \frac{1}{\log(1/p)} \frac{1-p}{p^{3/2}} = \frac{1}{\sqrt{1-p}}. \quad (24) \end{aligned}$$

Очевидно, что при достаточно малых значениях p оценка (23) лучше, так как $\sqrt{\pi}/2 < 1$. Однако при p , заметно отличающихся от нуля, лучшей будет оценка (24). Несложно проверить, что уже при $p > 0,15$ оценка (24) выигрывает. Таким образом, можно записать, что

$$\begin{aligned} K^*(p) &\leq \frac{1}{\sqrt{1-p}} \times \\ &\times \min \left\{ 1; \frac{\sqrt{2}}{(1-p)(2+p)^{3/2}} \left[\sqrt{\pi} + \frac{p}{\sqrt{2}} \left(\frac{2-p}{1-p}\right) \right] \right\}. \end{aligned}$$

Теперь, наконец, можно сформулировать теорему, являющуюся основным результатом данного раздела.

Теорема 2.1. Пусть выполнены условия (1) и (4). Тогда для распределения отрицательной биномиальной случайной суммы S_M с параметрами $r > 0$ и $p \in (0, 1)$ справедливо неравенство

$$\begin{aligned} \Delta_r = \sup_x \left| \mathbb{P} \left(S_M < x \sqrt{\frac{r(1-p)}{p}} \right) - \right. \\ \left. - \Phi(x) \right| \leq K(p) \frac{\beta_3}{\sqrt{r}}, \end{aligned}$$

где

$$\begin{aligned} K(p) &= \frac{1,5205}{\sqrt{1-p}} \times \\ &\times \min \left\{ 1; \frac{\sqrt{2}}{(1-p)(2+p)^{3/2}} \left[\sqrt{\pi} + \frac{p}{\sqrt{2}} \left(\frac{2-p}{1-p}\right) \right] \right\}. \end{aligned}$$

Подводя итог, отметим, что при фиксированных значениях параметра p распределение отрицательных биномиальных случайных сумм, находящих столь широкое практическое применение, можно приблизить стандартным нормальным законом. При этом чем больше значение параметра r , тем точнее приближение. Более конкретно, точность данной аппроксимации имеет порядок $O(r^{-1/2})$.

Литература

1. *Englund G.* A remainder term estimate in a random-sum central limit theorem // Теория вероятностей и ее применения, 1983. Т. 28. Вып. 1. С. 143–149. Письмо в редакцию // Теория вероятностей и ее применения, 1984. Т. 29. Вып. 1. С. 200–201.
2. *Королев В. Ю.* О точности нормальной аппроксимации для распределений сумм случайного числа независимых случайных величин // Теория вероятностей и ее применения, 1988. Т. 33. Вып. 3. С. 577–581.
3. *Круглов В. М., Королев В. Ю.* Предельные теоремы для случайных сумм. — М.: МГУ, 1990.
4. *Королев В. Ю., Бенинг В. Е., Шоргин С. Я.* Математические основы теории риска. — М.: Физматлит, 2007.
5. *Феллер В.* Введение в теорию вероятностей и ее приложения. Т. 1. — М.: Мир, 1984.
6. *Королев В. Ю., Шевцова И. Г.* Уточнение неравенства Берри–Эссеена с приложениями к пуассоновским и смешанным пуассоновским случайным суммам // Обзорное прикладное и промышленное математическое, 2010. Т. 17. Вып. 1. С. 25–56.
7. *Шоргин С. Я.* О точности нормальной аппроксимации распределений случайных сумм с безгранично делимыми индексами // Теория вероятностей и ее приложения, 1996. Т. 41. Вып. 4. С. 920–926.
8. *Феллер В.* Введение в теорию вероятностей и ее приложения. Т. 2. — М.: Мир, 1984.
9. *Dharmadhikari S. W., Jogdeo K.* Bounds on moments of certain random variables // Ann. Math. Statist., 1969. Vol. 40. No. 4. P. 1506–1508.
10. *Петров В. В.* Суммы независимых случайных величин. — М.: Наука, 1972.
11. *Fisher R. A., Corbet A. S., Williams C. B.* The relation between the number of species and the number of individuals // J. Animal Ecology, 1943. Vol. 12. P. 42–58.
12. *Кендалл М. Стьюарт А.* Теория распределений. — М.: Наука, 1966.

К ВОПРОСУ О РАЗМЕЩЕНИИ КОЛЛЕКТИВНЫХ СРЕДСТВ ОТОБРАЖЕНИЯ В СИТУАЦИОННОМ ЗАЛЕ С ЗАДАННЫМИ ПАРАМЕТРАМИ

К. Г. Чупраков¹

Аннотация: Установлены некоторые зависимости между основными параметрами ситуационного зала: его размерами, информативностью отображаемого контента, числом одновременных наблюдателей и размерами экрана. Основой для таких зависимостей стали рекомендации, закрепленные в ГОСТах, и простые геометрические соображения.

Ключевые слова: системы отображения информации; ситуационный зал; область наилучшего наблюдения; аналитические зависимости

1 Введение

Использование ситуационных центров доказало их практическую значимость для решения задач управления крупными и сложными объектами, в числе которых государственные учреждения [1], а также крупные корпорации и предприятия [2]. Однако создание ситуационных центров выявило ряд проблем, возникающих и в момент их разработки, и в процессе эксплуатации [1, 3–7].

Ситуационный зал является одним из приложений ситуационного центра для решения задач ситуационного управления с помощью двух основных методик: обсуждение возможных решений экспертной группой и доклад (презентация) некоторого материала с помощью средств отображения информации [8, 9]. В ситуационном зале могут анализироваться и разрабатываться различные варианты решения стратегического характера. Поэтому создание ситуационного зала — это не просто оснащение комнаты презентационным оборудованием с обеспечением некоторого респектабельного облика помещения, но и создание максимального уровня комфорта, где присутствие даже самой незначительной детали должно быть обосновано на уровне технического задания.

Для создания ситуационного зала в первую очередь необходимо оценить целый ряд параметров: размеры помещения, максимальное число человек, участвующих в переговорах, необходимую производительность средств отображений информации с точки зрения статической информативности экрана, а также геометрические размеры экранов этих средств отображения.

В статье предложен подход к определению взаимосвязей между основными параметрами ситуационного зала и оценке относительного расположения рабочих мест сотрудников и коллективного экрана.

2 Общий подход. Термины и определения

Создание и оборудование ситуационного зала среди прочих требований должно опираться на существующие стандарты по эргономике, действующие на территории РФ. В табл. 1 перечислены основные термины и понятия, которые будут использованы в рамках статьи [10–12].

Опираясь на эти понятия и термины, можно сформулировать и проанализировать требования и зависимости, описанные в государственных стандартах.

В качестве одного из основных будет использовано понятие **области наилучшего наблюдения**, под которой понимается область в трехмерном или двумерном пространстве, удовлетворяющая эргономическим требованиям ГОСТов.

2.1 Угол обзора человека

Согласно [13, 14]:

- очень часто используемые средства отображения информации, требующие точного и быстрого считывания показаний, следует располагать так, чтобы в вертикальном сечении они

¹Институт проблем информатики Российской академии наук, chkos@rambler.ru

Таблица 1 Основные термины и понятия

Термин	Обозначение	Определение
Активная часть экрана		Часть экрана, ограниченная пикселями
Стягиваемый угол (угловой размер)	ψ	Размер визуального объекта при данном конкретном расстоянии наблюдения: $\psi = 2 \arctg (h/(2D))$, где h — высота объекта; D — расстояние наблюдения
Высота знака	h	Линейная высота знака, м
Ширина знака	w	Линейная ширина знака, м
Формат знака		Число пикселей по горизонтали и вертикали в матрице, используемой для построения символа
Проектное расстояние наблюдения	D	Расстояние или диапазон расстояний между экраном и глазами наблюдателя, при котором изображение соответствует требованиям разборчивости и удобочитаемости
Рабочая площадь помещения		Множество положений в помещении, в которых сохраняется разборчивость, удобочитаемость материала, отображенного экраном
Угол обзора человека	γ	Угол, стягивающий точки пространства, которые человек способен наблюдать без напряжения для глазных мышц и поворотов головы

были видны под углом $\pm 15^\circ$ от нормальной линии взгляда и в горизонтальном сечении — под углом $\pm 15^\circ$ от плоскости симметрии человеческого тела (рис. 1);

- часто используемые средства отображения информации, требующие менее точного и быстрого считывания показаний, — $\pm 30^\circ$ по вертикали и горизонтали;
- редко используемые средства отображения информации — $\pm 60^\circ$ по вертикали и горизонтали.

2.2 Угловой размер знака

Стягиваемый угол определяется согласно [12, 15]. Рекомендуемые показатели для обеспечения разборчивости (для латинского алфавита): минимальный $\psi_{\min} = 16'$; предпочтительный $\psi_{\text{предп}} = 20' - 22'$.

Из определения угла ψ следует, что

$$\frac{h}{D} = 2 \operatorname{tg} \frac{\psi}{2} \approx \psi. \quad (1)$$

Формула (1) верна, так как угол ψ , измеряемый здесь и далее в радианах, очень мал. Данное соотношение позволяет оценить минимальное и рекомендуемое отношение высоты знака и проектного расстояния: минимальное $(h/D)_{\min} = \psi_{\min} = 4,7 \cdot 10^{-3}$; рекомендуемое $(h/D)_{\text{рек}} = \psi_{\text{рек}} = 5,8 - 6,4 \cdot 10^{-3}$.

2.3 Формат букв

Отношение линейных параметров ширины знака к его высоте должно соответствовать следующим значениям (для латинского алфавита) [12]: допустимый диапазон — от 0,5:1 до 1:1, предпочтительный диапазон — от 0,6:1 до 0,9:1.

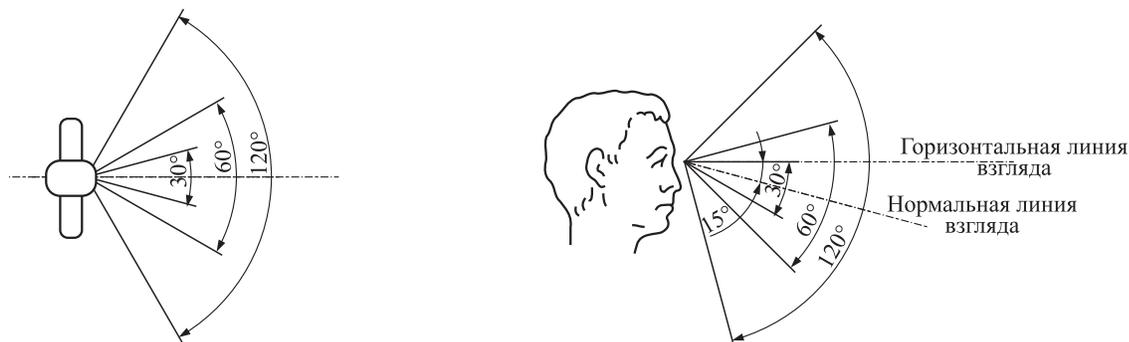


Рис. 1 Рекомендуемый угол обзора человека в зависимости от интенсивности наблюдения

Таблица 2 Основные параметры, используемые в зависимостях

Термины	Обозначение	Определение
Информативность	I	Максимальное число знаков в одном кадре контента
Число рабочих мест	N	Число рабочих мест для одного коллективного экрана
Диаметр помещения	P	Максимальное расстояние между двумя точками помещения, в том числе обусловленное его геометрией, м
Ширина экрана	W	Геометрическая ширина активной части экрана, м
Высота экрана	H	Геометрическая высота активной части экрана, м

3 Формирование зависимостей между основными параметрами ситуационного зала

На основании рекомендаций, обозначенных в п. 2, можно приступить к формированию взаимосвязей между основными параметрами системы «помещение—экран—наблюдатели». Этими основными параметрами являются: информативность, число рабочих мест, диаметр помещения, ширина и высота экрана (табл. 2).

Стоит отметить разницу между проектным расстоянием наблюдения D и диаметром помещения P . Первый параметр характеризуется свойствами системы «экран—наблюдатели», а параметр P — исключительно свойствами помещения.

3.1 Область наилучшего наблюдения для коллективного экрана

Пусть φ — угол направления обзора элемента экрана относительно нормали к поверхности экрана из центра элемента. Тогда размер элемента при наблюдении под углом φ уменьшится и может быть приближенно вычислен как

$$w(\varphi) = w \cos \varphi, \quad (2)$$

где w — линейный размер элемента, а $w(\varphi)$ — его линейный размер при наблюдении под углом φ . Поэтому из сохранения углового размера символа следует соотношение:

$$D(\varphi) = D \cos \varphi.$$

Следовательно, область наилучшего наблюдения элемента экрана (в плоском горизонтальном сечении) — круг, касающийся экрана в точке, соответствующей этому элементу и диаметром, равным D — проектному расстоянию (рис. 2).

Рассматривая экран как множество активных точек [16], для каждой из которых строится область

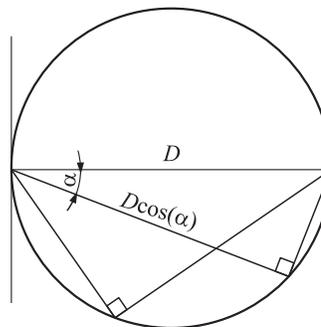


Рис. 2 Область наилучшего наблюдения для точечного элемента экрана

наилучшего наблюдения, можно построить область наилучшего наблюдения для всего экрана — пересечение множества построенных кругов (рис. 3). Далее рассматривается случай, когда экран плоский и рабочие места располагаются в горизонтальной плоскости.

В случае, когда экран плоский, областью наилучшего наблюдения будет пересечение двух наиболее далеких друг от друга кругов.

Другим существенным ограничением, которое необходимо наложить на область наилучшего наблюдения, являются пропорции наблюдаемых символов, которые, согласно данным п. 2.3, не могут меняться более чем в 2 раза, так как максимальное и минимальное допустимые значения пропорций отличаются друг от друга ровно в 2 раза (максимальная ширина равна одной высоте знака, а минимальная — ее половине). Значит, в силу формулы (2) максимальный угол отклонения от нормали не должен превышать 60° .

Пусть α — максимальный угол наблюдения для любого малого элемента поверхности экрана из некоторой точки пространства, расположенной со стороны активной поверхности экрана. Ясно, что этот максимум в горизонтальной плоскости будет достигаться возле правого или левого края экрана. Ввиду вышесказанного угол α не должен превосходить 60° .

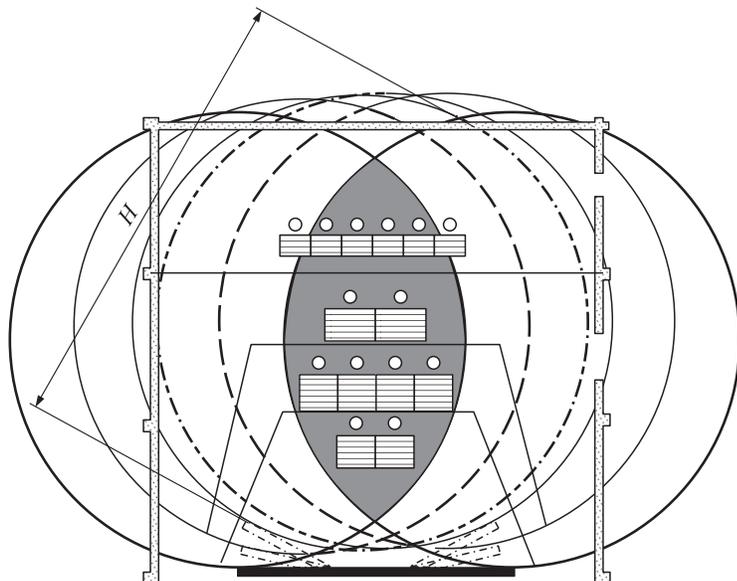


Рис. 3 Область наилучшего наблюдения экрана

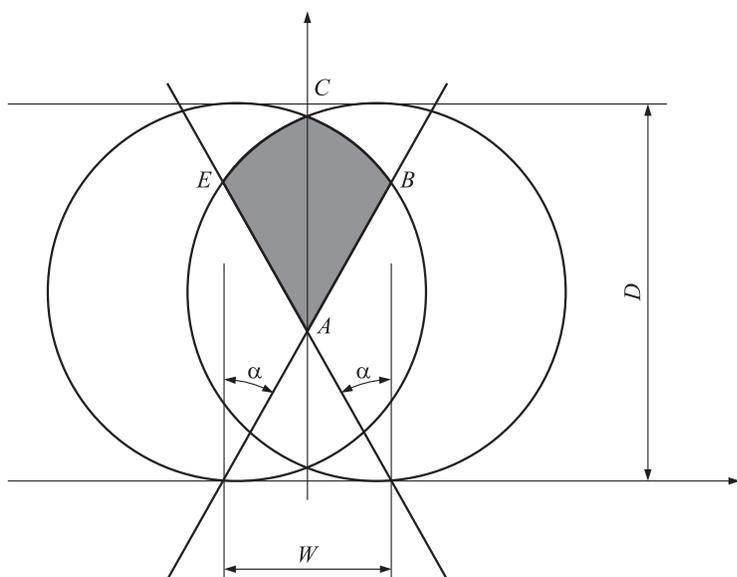


Рис. 4 Область наилучшего наблюдения экрана

Получается, что область наилучшего наблюдения заключена внутри угла пересечения двух лучей, проведенных из крайних точек экрана под углом α к нормали. С учетом всех ограничений область наилучшего наблюдения будет являться фигурой пересечения двух уже построенных областей (рис. 4).

Точки A , B , C и E имеют следующие координаты:

$$A \left(0; \frac{W \operatorname{ctg} \alpha}{2} \right);$$

$$B \left(\frac{D \sin 2\alpha - W}{2}; D \cos^2 \alpha \right);$$

$$C \left(0; \frac{D + \sqrt{D^2 - W^2}}{2} \right);$$

$$E \left(\frac{-D \sin 2\alpha + W}{2}; D \cos^2 \alpha \right).$$

Площадь области $ABCE$ можно оценить снизу с помощью треугольников ABC и ACE . Площадь каждого из них может быть посчитана по формуле

$$S_{\Delta} = \frac{1}{2} AC \cdot BH =$$

$$= \frac{1}{8} (D \sin 2\alpha - W) \left(D - W \operatorname{ctg} \alpha + \sqrt{D^2 - W^2} \right),$$

где BH — высота треугольника ABC , опущенная из вершины B на сторону AC . С учетом того, что таких треугольника два, получаем оценку площади области наилучшего наблюдения:

$$S_{\text{о.н.н.}} = \frac{D^2}{4} (\sin 2\alpha - C) \times$$

$$\times \left(1 - C \operatorname{ctg} \alpha + \sqrt{1 - C^2} \right) =$$

$$= \frac{W^2}{4C^2} (\sin 2\alpha - C) \left(1 - C \operatorname{ctg} \alpha + \sqrt{1 - C^2} \right), \quad (3)$$

где $C = W/D$ — константа системы, которая, как будет видно далее, зависит от информативности контента I , отношения k высоты экрана к ширине, отношения p ширины знака к его высоте и угла ψ , стягиваемого одним символом.

3.2 Число рабочих мест в области наилучшего наблюдения

Пусть D и W зафиксированы. Тогда определена область наилучшего наблюдения и можно оценить максимальное число человек N , которые смогут находиться в области наилучшего наблюдения, используя один коллективный экран. Главным условием комфортной работы условимся считать отсутствие помех со стороны других пользователей на расстоянии вытянутой руки. Таким образом, каждому пользователю сопоставим круг радиусом, равным половине маховой сажени (маховая сажень $\approx 1,8$ м). Наиболее плотной расстановкой этих кругов на плоскости будет расположение, когда каждый из кругов касается других шести. При этом центры кругов образуют сетку из равносторонних треугольников со стороной, равной одной маховой сажени, или 1,8 м (рис. 5).

Для оценки числа точек равномерной треугольной сетки, которые могут попасть внутрь области наилучшего наблюдения, воспользуемся формулой Пика, связывающей число узлов квадратной сетки с шагом 1, попавших внутрь и на границу многоугольника, с площадью этого многоугольника [17]:

$$S = B + \frac{\Gamma}{2} - 1,$$

где S — целочисленная площадь многоугольника, B — число узлов квадратной сетки с шагом 1, попавших внутрь многоугольника; Γ — число узлов этой сетки, которые попали на границу.

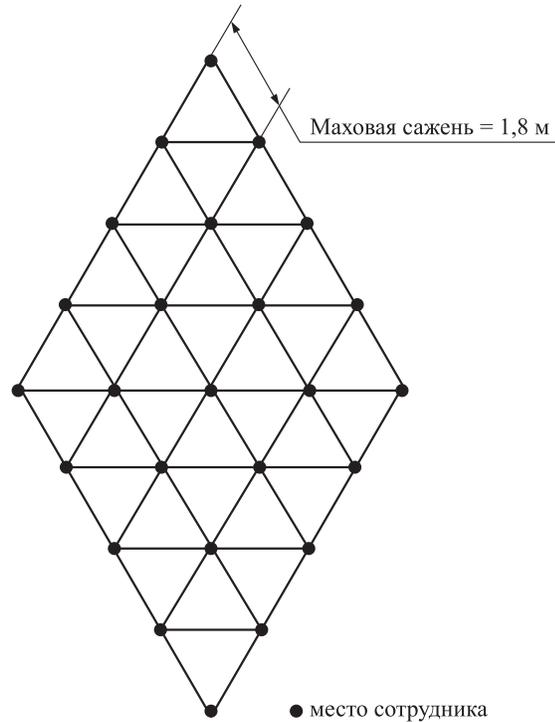


Рис. 5 Сетка из центров кругов

Треугольную сетку с шагом 1,8 можно рассматривать как квадратную сетку с шагом 1, которая претерпела два преобразования:

- сжатие по направлению вектора, параллельного диагонали квадрата сетки с коэффициентом $\sqrt{2}$;
- растяжение по двум любым взаимно перпендикулярным осям с коэффициентом 1,8.

Поэтому для применения формулы Пика необходимо рассматривать не исходное значение площади, а ее преобразование, обратное преобразованиям сетки. То есть в случае равномерной треугольной сетки с шагом 1 будет выполнено следующее соотношение:

$$\frac{\sqrt{2}}{1,8^2} S = B + \frac{\Gamma}{2} - 1. \quad (4)$$

Число рабочих мест в области наилучшего наблюдения при этом составляет

$$N = B + \Gamma = \frac{\sqrt{2}}{1,8^2} S + 1 + \frac{\Gamma}{2}. \quad (5)$$

Таким образом, для получения оценки числа рабочих мест в области наилучшего наблюдения необходимо оценить максимальное значение Γ . Важно понимать, что формула Пика описывает площадь

многоугольника с вершинами в узлах сетки. Ясно, что в случае области наилучшего наблюдения такое обеспечить не всегда возможно. Поэтому формула (5) представляет собой верхнюю оценку числа рабочих мест.

Из оценки периметра области наилучшего наблюдения получаем, что

$$\Gamma \leq \frac{1}{1,8} \left(2|AB| + 2\widehat{CB} \right) = \frac{D}{1,8} \left(\frac{\sin 2\alpha - C}{\sin \alpha} + 2\alpha - \arcsin C \right).$$

Таким образом, число рабочих мест в области наилучшего наблюдения может быть оценено сверху следующей величиной:

$$N = B + \Gamma = 0,44S + \frac{D}{3,6} \left(\frac{\sin 2\alpha - C}{\sin \alpha} + 2\alpha - \arcsin C \right) + 1. \quad (6)$$

3.3 Оценка максимальной площади области наилучшего наблюдения и максимального числа рабочих мест в этой области

Пусть I определено. Тогда исходя из пространственных ограничений, характеризуемых диаметром P , можно оценить максимальную площадь области наилучшего наблюдения при некоторых значениях проектного расстояния D и ширины экрана W .

Ввиду того, что соотношение сторон экрана у большинства производителей на настоящий момент составляет от 16:9 до 16:12, можно исключить из рассмотрения параметр высоты экрана. Дальнейшие оценки получаются в результате подсчета с двух сторон площади активной части экрана

$$S_{\text{display}} = WH = Iwh.$$

Пусть

$$H = kW; \quad w = ph,$$

где $k \in [9/16; 12/16]$, $p \in [0,6; 0,9]$ (см. п. 2.3). Тогда

$$kW^2 = Iph^2 \quad \text{или} \quad W = h\sqrt{\frac{Ip}{k}}. \quad (7)$$

Из формулы (1) следует, что

$$h = D\psi. \quad (8)$$

Из формул (7) и (8) получается, что при фиксированных значениях информативности I , отношения k высоты экрана к ширине, отношения p

ширины знака к его высоте и угла ψ , стягиваемого одним символом, существует константа C , удовлетворяющая следующим соотношениям:

$$C = \frac{W}{D} = \psi\sqrt{\frac{Ip}{k}}. \quad (9)$$

Из формулы (9) следует, что

$$C_{\min} = \psi_{\min}\sqrt{\frac{p_{\min}}{k_{\max}}}\sqrt{I} = \frac{\sqrt{I}}{193}. \quad (10)$$

Из-за ограниченности помещения проектное расстояние D не может превышать диаметра помещения P , поэтому ввиду оценок (3) и (10)

$$\begin{aligned} S_{\max} &= \frac{D^2}{4} (\sin 2\alpha - C) \left(1 - C \operatorname{ctg} \alpha + \sqrt{1 - C^2} \right) \leq \\ &\leq \frac{P^2}{2} \left(\sin 2\alpha - \frac{\sqrt{I}}{193} \right) \times \\ &\times \left(1 - \frac{\operatorname{ctg} \alpha \cdot \sqrt{I}}{193} + \sqrt{1 - \frac{I}{193^2}} \right). \quad (11) \end{aligned}$$

Отметим, что информативность не должна превосходить $(193 \sin 2\alpha)^2$ знаков. В противном случае область наилучшего наблюдения окажется пустым множеством.

На основании формулы (11) можно оценить число рабочих мест, которые можно разместить в области наилучшего наблюдения с учетом возможностей помещения, ограниченных его диаметром P . Согласно оценкам (6) и (11)

$$N_{\max} = 0,44S_{\max} + \frac{P}{3,6} \left(\frac{\sin 2\alpha - \sqrt{I}/193}{\sin \alpha} + 2\alpha - \arcsin \frac{\sqrt{I}}{193} \right) + 1. \quad (12)$$

3.4 Оценка минимальной ширины активной поверхности экрана

Чем меньше размер экрана, тем меньше его стоимость при прочих равных условиях. Поэтому размеры экрана, в том числе и его ширина, должны быть по возможности минимизированы. Пусть известна информативность контента I , угол наблюдения α и диаметр помещения P . Рассмотрим два принципиально разных случая:

1. Число наблюдателей неизвестно, необходимо оценить размеры экрана (его ширину), позволяющие эффективно использовать пространство помещения.
2. Число наблюдателей N известно, но оно меньше, чем полученное по формуле (12) для известных параметров P , I и α . Необходимо оценить минимальные размеры экрана (его ширину), позволяющие обеспечить расположение всех наблюдателей в области наилучшего наблюдения с учетом отсутствия взаимных помех.

Случай 1. Из соотношения (9) следует, что

$$W = PC \geq PC_{\min} = \frac{P\sqrt{I}}{193}.$$

Случай 2. Оценим минимальную площадь, на которой могут разместиться N наблюдателей. Для этого воспользуемся тем соображением, что минимальная площадь равна количеству внутренних рабочих мест (тех, которые не попали на границу), умноженному на площадь двух треугольников, ограниченных сеткой. Такое соображение следует из того, что каждой внутренней точке можно сопоставить 6 треугольников сетки, а каждому треугольнику сетки — 3 точки. Это значит, что каждой точке сопоставляется по 2 треугольника.

Из формул (4) и (5) следует, что

$$B = \frac{2\sqrt{2}}{1,8^2} S + 2 - N.$$

С другой стороны, как было замечено,

$$S = B \cdot 2S_{\Delta}.$$

Поэтому

$$S_{\min} = (N - 2) \frac{2 \cdot 1,8^2 \cdot 1,4}{4\sqrt{2} \cdot 1,4 - 1,8^2} = 1,94(N - 2).$$

Из формулы (3) получаем:

$$\begin{aligned} W_{\min} &= 2C_{\min} \left(S_{\min} / \left((\sin 2\alpha - C_{\min}) \times \right. \right. \\ &\quad \left. \left. \times (1 - C_{\min} \operatorname{ctg} \alpha + \sqrt{1 - C_{\min}^2}) \right) \right)^{1/2} = \\ &= 2 \left(1,94I(N - 2) / \left((193 \sin 2\alpha - \sqrt{I}) \times \right. \right. \\ &\quad \left. \left. \times (193 - \sqrt{I} \operatorname{ctg} \alpha + \sqrt{193^2 - I}) \right) \right)^{1/2}. \end{aligned}$$

4 Заключение

В статье сформулированы основные определения и термины, используемые при проектировании средств отображения информации, а также описаны рекомендации к ним, действующие в рамках государственных стандартов.

На основании этих требований и рекомендаций, а также на основании простых геометрических соображений получены оценки основных параметров для проектирования средства отображения и рабочих мест. В качестве основной определяющей величины выступает информативность контента, которая формируется на основании задач ситуационного зала и ситуационных моделей отображения.

В качестве основного использован параметр отношения ширины экрана к проектному расстоянию наблюдения. Часто этот параметр для удобства считается фиксированным, но в общем случае показано, что это отношение прямо пропорционально корню квадратному от величины информативности контента.

Если, помимо информативности контента, известен диаметр помещения, в котором располагается ситуационный зал, то можно определить максимальную площадь области наилучшего наблюдения и оценить число рабочих мест, которые могут быть расположены в области наилучшего наблюдения с помощью формулы Пика и аффинных преобразований.

В качестве обратной задачи, в которой известно количество человек, одновременно работающих с экраном, получена зависимость для определения минимально необходимой ширина экрана.

Литература

1. Ильин Н. И. Основные направления развития ситуационных центров органов государственной власти // ВКСС Соппест! (Ведомственные корпоративные сети и системы), 2007. № 6(45). С. 2–9.
2. Лисица К. В. Опыт создания и применения Автоматизированной системы стратегического управления в ОАО «Российские железные дороги» // Ситуационные центры: модели, технологии, опыт практической реализации: Мат-лы науч.-практич. конф. — М.: РАГС, 2007.
3. Филитович А. Ю. Ситуационная система — что это такое? // PCWeek/RE, 2003. No. 26.
4. Зацаринный А. А., Ионенков Ю. С., Кондрашев В. А. Об одном подходе к выбору системотехнических решений построения информационно-телекоммуникационных систем // Системы и средства информатики. Вып. 16. — М.: Наука, 2006. С. 65–72.

5. Зацаринный А. А., Сучков А. В., Босов А. В. Ситуационные центры в современных информационно-телекоммуникационных системах специального назначения // ВКСС Connect! (Ведомственные корпоративные сети и системы), 2007. № 5(44). С. 64–76.
6. Зацаринный А. А., Ионенков Ю. С. Некоторые аспекты выбора технологии построения информационно-телекоммуникационных сетей // Системы и средства информатики. Вып. 17. — М.: Наука, 2007. С. 5–16.
7. Зацаринный А. А. Тенденции развития ситуационных центров как компонентов информационно-телекоммуникационных систем в условиях глобальной информатизации общества // Докл. XXXV междунар. конф. «Информационные технологии в науке, образовании, телекоммуникации и бизнесе (IT + S&E'08)». Ялта–Гурзуф, Украина. 2008.
8. Ситуационные центры (СЦ) и их история. http://ta.interrussoft.com/s_centre.html.
9. Зацаринный А. А. Организационные принципы системного подхода к разработке, проектированию и внедрению современных информационно-телекоммуникационных сетей // ВКСС Connect! (Ведомственные корпоративные сети и системы), 2007. № 1(40). С. 60–67.
10. ГОСТ 26387-84 Система «Человек—машина». Термины и определения. — М.: Стандартинформ, 2006.
11. ГОСТ 27833-88 Средства отображения информации. Термины и определения. — М.: Стандартинформ, 2005.
12. ГОСТ Р 52324-2005 (ИСО 13406-2:2001) Эргономические требования к работе с визуальными дисплеями, основанными на плоских панелях. — М.: Стандартинформ, 2005.
13. ГОСТ 12.2.032-78 Система стандартов безопасности труда. Рабочее место при выполнении работ сидя. Общие эргономические требования. — М.: Изд-во стандартов, 2001.
14. ГОСТ 12.2.033-78 Система стандартов безопасности труда. Рабочее место при выполнении работ стоя. Общие эргономические требования. — М.: Изд-во стандартов, 2001.
15. ГОСТ Р ИСО 9241-3 Эргономические требования при выполнении офисных работ с использованием видеодисплейных терминалов. — М.: Изд-во стандартов, 2003.
16. ГОСТ 21958-76 Зал и кабины операторов. Взаимное расположение рабочих мест. — М.: Изд-во стандартов, 1976.
17. Прасолов В. В. Задачи по планиметрии. — М.: МЦНМО, 2001.

LibMeta — DIGITAL LIBRARY MANAGEMENT SYSTEM

A. A. Zakharov¹ and V. A. Serebryakov²

¹Dorodnicyn Computing Center of the Russian Academy of Sciences, andreya@sufler.ru

²Dorodnicyn Computing Center of the Russian Academy of Sciences, serebr@ccas.ru

The problems of modern digital libraries (DLs) creation are addressed. General requirements on integration with external data sources and standartization are proposed among with some means to satisfy them. Leading world standards for DLs are considered. As a result of practical implementation of studies, DL management system LibMeta is presented as an universal tool for DL creation.

Keywords: digital libraries; distributed information system

INTEGRATION OF HETEROGENEOUS INFORMATION ABOUT COLOR PIXELS AND THEIR COLOR PERCEPTION

O. P. Arkhipov and Z. P. Zykova

IPI RAN, Orel Branch, ofran@orel.ru

The problem of integration of heterogeneous information about color pixels, their transformation in user computer system, standard and individual color perception of user in uniform information and communication environment has been considered. It is necessary to solve this problem in order to create special easy-adaptable to system components program-technical tools, providing the adequate perception by arbitrary users of output on peripheral devices of PC. When solving the problem, the new representation methods and analysis of heterogeneous (quantitative and qualitative) information on bounded color spaces, multicriterion choice for prediction of difference in color information and pixels classification for structuring of color spaces are used.

Keywords: color space; chromatic sensation; color perception space; anomaly of color vision; partial color blindness; RGB-characterization

MATHEMATICAL MODEL AND HUMAN POSE RECONSTRUCTION METHOD BASED ON STEREOMATE IMAGE SILHOUETTES

A. Tsiskaridze

Moscow Institute of Physics and Technology, AchikoTsi@gmail.com

A problem of locally symmetric object reconstruction based on a stereomate of silhouettes when occlusion is absent is discussed. Two models of an object are considered. In the first model, an object is described as a union of fat spatial curves and a method of restoration of such objects is presented. In the second case, an object is described as a joint model and a fitting method based on a stereomate of silhouettes is proposed. The methods are based on a construction of continuous skeletons. The methods have been compared for human body pose reconstruction.

Keywords: stereoreconstruction; skeleton; cylindrical objects; joint model; medial axes

MODELING AND CLASSIFICATION OF MULTICHANNEL REMOTELY SENSED IMAGES VIA COPULAS

V. A. Krylov

Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics,
M. V. Lomonosov Moscow State University, vkrylov@cs.msu.ru

A novel approach to modeling of multichannel remotely sensed images is proposed. This approach suggests to use the classical statistical probability distribution estimation methods for single channels and then the construction of the joint probability distribution of a multichannel image via copulas. An integration of the developed copula-based approach with a Markov random field model is proposed for supervised Bayesian image classification. Experiments with real remotely sensed images captured by a synthetic aperture radar demonstrate high accuracy classification results proving the efficiency of the developed approach as compared to state-of-the-art methods.

Keywords: multichannel image; copula; Markov random field; Bayesian classification

COMMUNICATION BETWEEN TIME AND STRUCTURAL-TOPOLOGICAL CHARACTERISTICS OF HEALTHY PEOPLE HEART RHYTHM DIAGRAMS

A. A. Kuznetsov

Vladimir State University, artemi-k@mail.ru

According to 628 electrocardiogram registration at 177 healthy and ill people, the comparative analysis of parameters of real and virtual heart rhythm diagrams for a regulation system influence estimation on a heart rhythm is carried out. Between diagrams parameters and information entropy in conditions of discrete seasonal adaptation, functional communications are determined. "Organism functional condition formulas" connecting a heart rhythm diagram macrostructure parameters with its storey microstructures parameters are offered. It is revealed that the mode of a healthy person heart rhythm without sex dependence has a calendar year cycle, during which it three times discretely varies.

Keywords: heart rhythm diagram; organism functional condition; storey structure; information entropy; quantity of the information

SEMIFORMAL VERIFICATION FOR PIPELINED DIGITAL DESIGNS BASED ON ALGORITHMIC STATE MACHINES

S. Baranov¹, S. Frenkel², and V. Zakharov³

¹Holon Institute of Technology, Holon, Israel, samary@012.net.il

²IPI RAN, fsergei@mail.ru

³IPI RAN, VZakharov@ipiran.ru

The method and tools to use Algorithmic State Machine (ASM) formalism for high-level specification of complex digital designs and their Model-Checking-based verification are described. This methodology is based on the possibility of hierarchical description of the target digital designs at algorithmic level of abstraction, and ability to generate finite state machines (FSM) models of the systems from the ASM flowcharts. The software tool was developed for automatic generation of SMV (Symbolic Modes Verifier) codes from the ASMs and corresponding FSMs. A way of this approach application to design verification is demonstrated for a pipelined microprocessor.

Keywords: formal verification; Model Checking; finite state machines

COORDINATION ALGORITHM FOR HYBRID INTELLECTUAL SYSTEM SOLUTIONS OF A COMPLEX PROBLEM OF OPERATIONAL INDUSTRIAL PLANNING

A. V. Kolesnikov¹ and S. A. Soldatov²

¹Department of Computer Modeling and Information Systems, “Immanuel Kant Russian State University;”
Kaliningrad Branch of the IPI RAN, avkolesnikov@yandex.ru

²Limited Liability Company “Lighton,” Moscow, ssa@west-automatica.com

The problem of operational industrial planning at the machine-building enterprise with custom-made, small-scale character of manufacture is considered and the approach to the solution of similar problems on the basis of methodology of functional hybrid intellectual systems with coordination is described. The description of practical realization of hybrid intellectual system and the short analysis of the received results are given.

Keywords: machine-building enterprise; problem of operational planning; coordination; hybrid intellectual systems

STATIONARY CHARACTERISTICS OF THE TWO-CHANNEL QUEUEING SYSTEM WITH REORDERING CUSTOMERS AND DISTRIBUTIONS OF PHASE TYPE

S. I. Matyushenko

Department of Probability Theory and Mathematical Statistics,
Peoples' Friendship University of Russia, matushenko@list.ru

The two-channel finite-capacity queueing system with the distributions of phase type and reordering customers is considered. The Laplace–Stieltjes transform of the distribution function that characterizes the delay of reordering is obtained. The algorithm is developed to calculate the factorial moments of the number of customers being at the buffer of reordering.

Keywords: queueing system; distributions of phase type; reordering customers

NORMAL APPROXIMATION FOR DISTRIBUTION OF RISK ESTIMATE FOR WAVELET COEFFICIENTS THRESHOLDING WHEN USING SAMPLE VARIANCE

O. V. Shestakov

Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics,
M. V. Lomonosov Moscow State University, oshestakov@cs.msu.su

The asymptotic properties of risk estimate for thresholding wavelet coefficients of signal function are analyzed. Some estimates for rate of convergence to the normal law are obtained.

Keywords: wavelets; thresholding; risk estimate; normal distribution; rate of convergence

ESTIMATES FOR CONVERGENCE RATE OF DISTRIBUTIONS OF RANDOM SUMS WITH INFINITELY DIVISIBLE INDICES TO THE NORMAL DISTRIBUTION

S. V. Gavrilenko

Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics,
M. V. Lomonosov Moscow State University, gavrilenko.cmc@gmail.com

New estimates for convergence rate of distributions of random sums with infinitely divisible indices were obtained. These estimates remain true under weaker conditions than the known ones. As an example of this result, the article contains estimates of the accuracy of the normal approximation for the distribution functions of random sums with indices that have negative binomial distribution.

Keywords: random sum; integer-valued infinitely divisible distribution; generalized Poisson distribution; negative binomial distribution; normal approximation

ON COLLECTIVE DISPLAY FACILITIES PLACED IN A SITUATIONAL HALL WITH PRESCRIBED PARAMETERS

K. G. Chuprakov

IPI RAN, chkos@rambler.ru

There are some dependences between main parameters of situational (or congress) hall. It is possible to link the diameter of hall (the largest distance between any two points of the hall), the informativity of demonstrated on screen content, the quantity of people working with a screen and also screen dimensions. The base for describing these dependences are the recommendations fixed in the Russian state standards and simple geometric considerations.

Keywords: visualization systems; situational hall; good observation area; algebraic dependences

Об авторах

Архипов Олег Петрович (р. 1948) — кандидат технических наук, директор Орловского филиала ИПИ РАН

Баранов Самарий Иосифович (р. 1938) — доктор технических наук, профессор Холонского технологического института, Израиль

Гавриленко Семен Васильевич (р. 1985) — аспирант кафедры математической статистики факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова

Захаров Андрей Андреевич (р. 1983) — аспирант Вычислительного центра им. А. А. Дородницына РАН

Захаров Виктор Николаевич (р. 1948) — кандидат технических наук, доцент, ученый секретарь ИПИ РАН

Зыкова Зоя Петровна (р. 1953) — кандидат физико-математических наук, заведующая лабораторией технологий и методов ввода-вывода информации Орловского филиала ИПИ РАН

Колесников Александр Васильевич (р. 1948) — доктор технических наук, профессор кафедры компьютерного моделирования и информационных систем ФГОУ ВПО «Российский государственный университет им. Иммануила Канта»; старший научный сотрудник Калининградского филиала ИПИ РАН

Крылов Владимир Андреевич (р. 1985) — аспирант кафедры математической статистики факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова

Кузнецов Артемий Артемьевич (р. 1954) — кандидат физико-математических наук, доцент кафедры общей и прикладной физики факультета прикладной математики и физики Владимирского государственного университета

Матюшенко Сергей Иванович (р. 1963) — кандидат физико-математических наук, доцент кафедры теории вероятностей и математической статистики Российского университета дружбы народов (СНРS)

Серебряков Владимир Алексеевич (р. 1946) — доктор физико-математических наук, заведующий отделом Вычислительного центра им. А. А. Дородницына РАН

Солдатов Сергей Александрович (р. 1982) — инженер, ООО «Лайтон» Москва

Френкель Сергей Лазаревич (р. 1951) — кандидат технических наук, доцент, старший научный сотрудник ИПИ РАН; доцент факультета информационных технологий Московского института радиотехники, электроники и автоматики (МИРЭА)

Цискаридзе Арчил Константинович (р. 1983) — аспирант кафедры «Интеллектуальные системы» факультета управления и прикладной математики Московского физико-технического института

Чупраков Константин Григорьевич (р. 1985) — аспирант, ИПИ РАН

Шестаков Олег Владимирович (р. 1976) — кандидат физико-математических наук, старший преподаватель кафедры математической статистики факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова

About Authors

Arkhipov Oleg P. (b. 1948) — Candidate of Science (PhD) in technology, Director of Oryol Branch of Institute of Informatics Problems, Russian Academy of Sciences

Baranov Samary I. (b. 1938) — Doctor of Sciences in technology, Professor of Holon Institute of Technology, Israel

Chuprakov Konstantin G. (b. 1985) — PhD student, Institute of Informatics Problems, Russian Academy of Sciences

Frenkel Sergey L. (b. 1951) — Candidate of Sciences (PhD) in technology, senior scientist, Institute of Informatics Problems, Russian Academy of Sciences; associate professor, Faculty of Computer Science, Moscow Institute of Radio, Electronics, and Automation

Gavrilenko Semyon V. (b. 1985) — PhD student, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University

Kolesnikov Alexander V. (b. 1948) — Doctor of Science in technology, professor, Department of Computer Modeling and Information Systems, “Immanuel Kant Russian State University;” senior scientific associate, Kaliningrad Branch of the Institute of Informatics Problems, Russian Academy of Sciences

Krylov Vladimir A. (b. 1985) — PhD student, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University

Kuznetsov Artemiy A. (b. 1954) — Candidate of Science (PhD) in physics and mathematics, associate professor, Faculty of Applied Mathematics and Physics, Department of General and Applied Physics, Vladimir State University

Matyushenko Sergey I. (b. 1963) — Candidate of Science (PhD) in physics and mathematics, associate professor, Department of Probability Theory and Mathematical Statistics, Peoples’ Friendship University of Russia

Serebryakov Vladimir A. (b. 1946) — Doctor of Science in physics and mathematics, Head of Department, Dorodnicyn Computing Center, Russian Academy of Sciences

Shestakov Oleg V. (b. 1976) — Candidate of Science (PhD) in physics and mathematics, senior lecturer, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University

Soldatov Sergey A. (b. 1982) — engineer, Limited Liability Company “Lighton,” Moscow

Tsiskaridze Archil K. (b. 1983) — PhD student, Department of Intelligent Systems, Faculty of Control Management and Applied Mathematics, Moscow Institute of Physics and Technology

Zakharov Andrey A. (b. 1983) — PhD student, Dorodnicyn Computing Center, Russian Academy of Sciences

Zakharov Victor N. (b. 1948) — Candidate of Sciences (PhD) in technology, associate professor, Scientific secretary, Institute of Informatics Problems, Russian Academy of Sciences

Zykova Zoya P. (b. 1953) — Candidate of Science (PhD) in physics and mathematics, Head of Laboratory of technology and methods of information input-output, Oryol Branch of Institute of Informatics Problems, Russian Academy of Sciences

АВТОРСКИЙ УКАЗАТЕЛЬ ЗА 2010 г.

		Выпуск	Стр.
Арутюнян А. Р. Моделирование влияния деформаций отпечатков пальцев на точность дактилоскопической идентификации	1		51
Архипов О. П., Зыкова З. П. Интеграция гетерогенной информации о цветных пикселях и их цветовосприятии	4		15
Баранов С. И., Френкель С. Л., Захаров В. Н. Полуформальная верификация цифрового устройства с конвейером, основанная на использовании алгоритмических машин состояния	4		49
Бекетова И. В. см. Каратеев С. Л.			
Белоусов В. В. см. Сеницын И. Н.			
Бенинг В. Е., Королев Р. А. О предельном поведении мощностей критериев в случае распределения Лапласа	2		63
Бенинг В. Е., Сипина А. В. Асимптотическое разложение для мощности критерия, основанного на выборочной медиане, в случае распределения Лапласа	1		18
Бондаренко А. В. см. Каратеев С. Л.			
Бородина А. В., Морозов Е. В. Об оценивании асимптотики вероятности большого отклонения стационарной регенеративной очереди с одним прибором	3		29
Бунтман Н. В., Минель Ж.-Л., Ле Пезан Д., Зацман И. М. Типология и компьютерное моделирование трудностей перевода	3		77
Визильтер Ю. В. см. Каратеев С. Л.			
Гавриленко С. В. Оценки скорости сходимости распределений случайных сумм с безгранично делимыми индексами к нормальному закону	4		81
Григорьева М. Е., Шевцова И. Г. Уточнение неравенства Каца—Берри—Эссеена	2		75
Грушо А. А., Грушо Н. А., Тимонина Е. Е. Поиск конфликтов в политиках безопасности: модель случайных графов	3		38
Грушо Н. А. см. Грушо А. А.			
Гудков В. Ю. Математические модели изображения отпечатка пальца на основе описания линий	1		58
Гуртов А. В. см. Лукьяненко А. С.			
Желтов С. Ю. см. Каратеев С. Л.			
Захаров А. А., Серебряков В. А. Система управления электронной библиотекой LibMeta	4		2
Захаров В. Н. см. Баранов С. И.			
Захарова Т. В. см. Матвеева С. С.			
Зацаринный А. А., Чупраков К. Г. Некоторые аспекты выбора технологии для построения систем отображения информации ситуационного центра	3		59
Зацман И. М. см. Бунтман Н. В.			
Зейфман А. И., Коротышева А. В., Сатин Я. А., Шоргин С. Я. Об устойчивости нестационарных систем обслуживания с катастрофами	3		9
Зыкова З. П. см. Архипов О. П.			
Илюшин Г. Я., Соколов И. А. Организация управляемого доступа пользователей к разнородным ведомственным информационным ресурсам	1		24
Кавагучи Ю., Ульянов В. В., Фуджикоши Я. Приближения для статистик, описывающих геометрические свойства данных большой размерности, с оценками ошибок	1		12
Каратеев С. Л., Бекетова И. В., Ососков М. В., Князь В. А., Визильтер Ю. В., Бондаренко А. В., Желтов С. Ю. Автоматизированный контроль качества цифровых изображений для персональных документов	1		65

	Выпуск	Стр.
Козеренко Е. Б. Лингвистические фильтры в статистических моделях машинного перевода	2	83
Козеренко Е. Б., Кузнецов И. П. Когнитивно-лингвистические представления в системах обработки текстов	3	69
Князь В. А. см. Каратеев С. Л.		
Колесников А. В., Солдатов С. А. Алгоритм координации для гибридной интеллектуальной системы решения сложной задачи оперативно-производственного планирования	4	61
Коновалов М. Г. О планировании потоков в системах вычислительных ресурсов	2	3
Конущин А. С. см. Конущин В. С.		
Конущин В. С., Кривовязь Г. Р., Конущин А. С. Алгоритм распознавания людей в видео-последовательности по одежде	1	74
Корепанов Э. Р. см. Синицын И. Н.		
Королев В. Ю. см. Соколов И. А.		
Королев Р. А. см. Бенинг В. Е.		
Коротышева А. В. см. Зейфман А. И.		
Кривенко М. П. Непараметрическое оценивание элементов байесовского классификатора	2	13
Кривовязь Г. Р. см. Конущин В. С.		
Крылов А. С. см. Павельева Е. А.		
Крылов В. А. Моделирование и классификация многоканальных дистанционных изображений с использованием копул	4	34
Крючин О. В. Разработка параллельных эвристических алгоритмов подбора весовых коэффициентов искусственной нейронной сети	2	53
Кудрявцев А. А., Шоргин С. Я. Байесовские модели массового обслуживания и надежности: характеристики среднего числа заявок в системе $M M 1 \infty$	3	16
Кузнецов А. А. Связь между временными и структурно-топологическими характеристиками диаграмм ритма сердца здоровых людей	4	39
Кузнецов И. П. см. Козеренко Е. Б.		
Ле Пезан Д. см. Бунтман Н. В.		
Лукьяненко А. С., Морозов Е. В., Гуртов А. В. Анализ сетевого протокола с общей функцией расширения окна передачи сообщения при конфликтах	2	46
Лямин О. О. О предельном поведении мощностей критериев в случае обобщенного распределения Лапласа	3	47
Маркин А. В., Шестаков О. В. Асимптотики оценки риска при пороговой обработке вейвлет-вейвлет коэффициентов в задаче томографии	2	36
Матвеева С. С., Захарова Т. В. Сети массового обслуживания с наименьшей длиной очереди	3	22
Матюшенко С. И. Стационарные характеристики двухканальной системы обслуживания с переупорядочиванием заявок и распределениями фазового типа	4	68
Минель Ж.-Л. см. Бунтман Н. В.		
Морозов Е. В. см. Бородина А. В.		
Морозов Е. В. см. Лукьяненко А. С.		
Ососков М. В. см. Каратеев С. Л.		
Павельева Е. А., Крылов А. С. Поиск и анализ ключевых точек радужной оболочки глаза методом преобразования Эрмита	1	79
Печинкин А. В. см. Френкель С. Л.,		
Протасов В. И. Составление субъективного портрета с использованием эволюционного морфинга и квалиметрии метода	1	83
Рудаков К. В., Торшин И. Ю. Вопросы разрешимости задачи распознавания вторичной структуры белка	2	25
Сатин Я. А. см. Зейфман А. И.		
Сейфуль-Мулюков Р. Б. Нефть как носитель информации о своем происхождении, структуре и эволюции	1	41

Семендяев Н. Н. см. Сеницын И. Н.		
Серебряков В. А. см. Захаров А. А.		
Сеницын В. И. см. Сеницын И. Н.		
Сеницын И. Н., Сеницын В. И., Корепанов Э. Р., Белоусов В. В., Семендяев Н. Н. Оперативное построение информационных моделей движения полюса Земли методами линейных и линеаризованных фильтров	1	2
Сипина А. В. см. Бенинг В. Е.		
Соколов И. А. О работах заслуженного деятеля науки Российской Федерации И. Н. Си- ницына в области информационных технологий и автоматизации (к 70-летию со дня рождения)	3	84
Соколов И. А. см. Илюшин Г. Я.		
Соколов И. А., Королев В. Ю. Предисловие	2	2
Солдатов С. А. см. Колесников А. В.		
Степанов С. Ю. Использование координатного метода фрагментации коммутаторной нейронной сети для сокращения трафика	2	57
Тимонина Е. Е. см. Грушо А. А.		
Торшин И. Ю. см. Рудаков К. В.		
Ульянов В. В. см. Кавагучи Ю.		
Фазекаш И. см. Чупрунов А. Н.		
системах после сбоев, вызываемых переходными помехами	3	2
Френкель С. Л. см. Баранов С. И.		
Френкель С. Л., Печинкин А. В. Оценка времени самовосстановления в цифровых Фуджикоши Я. см. Кавагучи Ю.		
Цискаридзе А. К. Математическая модель и метод восстановления позы человека по стереопаре силуэтных изображений	4	27
Чупраков К. Г. К вопросу о размещении коллективных средств отображения в ситуа- ционном зале с заданными параметрами	4	89
Чупраков К. Г. см. Зацаринный А. А.		
Чупрунов А. Н., Фазекаш И. Законы повторного логарифма для числа безошибочных блоков при помехоустойчивом кодировании	3	42
Шевцова И. Г. см. Григорьева М. Е.		
Шестаков О. В. Аппроксимация распределения оценки риска пороговой обработки вейвлет-коэффициентов нормальным распределением при использовании выбо- рочной дисперсии	4	73
Шестаков О. В. см. Маркин А. В.		
Шоргин С. Я. см. Зейфман А. И.		
Шоргин С. Я. см. Кудрявцев А. А.		

2010 AUTHOR INDEX

	Issue	Page
Arkhipov O. P. and Zyкова Z. P. Integration of Heterogeneous Information About Color Pixels and Their Color Perception	4	15
Arutyunyan A. R. Modeling of Elastic Deformations Impact on Fingerprint Recognition Performance	1	51
Baranov S. I., Frenkel S. L., and Zakharov V. N. Semiformal Verification for Pipelined Digital Designs Based on Algorithmic State Machines	4	49
Beketova I. V. see Karateev S. L.		
Belousov V. V. see Sinitsyn I. N.		
Bening V. E. and Korolev R. A. On Asymptotic Behavior of the Powers of the Tests for the Case of Laplace Distribution	2	63
Bening V. E. and Sipina A. V. Asymptotic Expansion for the Power of Test Based on Sample Median in the Case of Laplace Distribution	1	18
Bondarenko A. V. see Karateev S. L.		
Borodina A. V. and Morozov E. V. On Estimation of the Large Deviation Asymptotic of a Single Server Regenerative Stationary Queue	3	29
Buntman N. V., Minel J.-L., Le Pesant D., and Zatsman I. M. Typology and Computer Modeling of Translation Difficulties	3	77
Chuprakov K. G. On Collective Display Facilities Placed in a Situational Hall with Prescribed Parameters	4	89
Chuprakov K. G. see Zatsarinny A. A.		
Chuprunov A. N. and Fazekas I. Laws of Iterated Logarithm for Numbers of Nonerror Blocks Under Error Corrected Coding	3	42
Fazekas I. see Chuprunov A. N.		
Frenkel S. L. and Pechinkin A. V. Estimation of Self-Healing Time for Digital Systems Under Transient Faults	3	2
Frenkel S. L. see Baranov S. I.		
Fujikoshi Y. see Kawaguchi Y.		
Gavrilenko S. V. Estimates for Convergence Rate of Distributions of Random Sums with Infinitely Divisible Indices to the Normal Distribution	4	81
Grigorieva M. E. and Shevtsova I. G. An Improvement of the Katz–Berry–Esseen Inequality. . .	2	75
Grusho A. A., Grusho N. A., and Timonina E. E. Search of the Conflicts in Security Policies: A Model of Random Graphs	3	38
Grusho N. A. see Grusho A. A.		
Gudkov V. Yu. Mathematical Models of Fingerprint Image on the Basis of Lines Description . . .	1	58
Gurtov A. V. see Lukyanenko A. S.		
Ilyushin G. Y. and Sokolov I. A. Organization of Users' Manageable Access to Heterogeneous Departmental Informational Resources	1	24
Karateev S. L., Beketova I. V., Ososkov M. V., Knyaz V. A., Vizilter Yu. V., Bondarenko A. V., and Zheltov S. Yu. Automated Quality Testing of Digital Imagery for Personal Documents	1	65
Kawaguchi Y., Ulyanov V. V., and Fujikoshi Y. Asymptotic Distributions of Basic Statistics in Geometric Representation for High-Dimensional Data and Their Error Bounds	1	12
Knyaz V. A. see Karateev S. L.		
Kolesnikov A. V. and Soldatov S. A. Coordination Algorithm for Hybrid Intellectual System Solutions of a Complex Problem of Operational Industrial Planning	4	61
Konovalov M. G. On Task Flow Planning in Computational Resource Systems	2	3

	Issue	Page
Konushin A. S. see Konushin V. S.		
Konushin V. S., Krivovyaz G. R., and Konushin A. S. An Algorithm for Clothes-Based Human Recognition in Video	1	74
Korepanov E. R. see Sinitsyn I. N.		
Korolev R. A. see Bening V. E.		
Korotysheva A. V. see Zeifman A. I.		
Kozerenko E. B. Linguistic Filters in Statistical Machine Translation Models	2	83
Kozerenko E. B. and Kuznetsov I. P. Cognitive Linguistic Presentations in the Text Processing Systems	3	69
Krivenko M. Nonparametric Estimation of Bayesian Classifier Elements	2	13
Krivovyaz G. R. see Konushin V. S.		
Krylov A. S. see Pavelyeva E. A.		
Krylov V. A. Modeling and Classification of Multichannel Remotely Sensed Images via Copulas	4	34
Kryuchin O. V. Development of Parallel Heuristic Algorithms of Weights Coefficients Selection for Artificial Neural Network	2	53
Kudriavtsev A. A. and Shorgin S. Ya. Bayesian Queueing and Reliability Models: Characteristics of Mean Number of Claims in the System $M M 1 \infty$	3	16
Kuznetsov A. A. Communication Between Time and Structural-Topological Characteristics of Healthy People Heart Rhythm Diagrams	4	39
Kuznetsov I. P. see Kozerenko E. B.		
Le Pesant D. see Buntman N.		
Lukyanenko A. S., Morozov E. V., and Gurtov A. V. Analysis of a Link Protocol with a General Contention Window Backoff Function	2	46
Lyamin O. O. On an Asymptotic Behavior of Power of Tests in Case of Generalized Laplace Distribution	3	47
Markin A. V. and Shestakov O. V. Asymptotic Properties of Risk Estimate of Wavelet-Vaguelette Coefficients Thresholding in Tomographic Reconstruction Problem	2	36
Matveeva S. S. and Zakharova T. V. Queueing Systems with Minimized Average Waiting Line ...	3	22
Matyushenko S. I. Stationary Characteristics of the Two-Channel Queueing System with Reordering Customers and Distributions of Phase Type	4	68
Minel J.-L. see Buntman N. V.		
Morozov E. V. see Borodina A. V.		
Morozov E. V. see Lukyanenko A. S.		
Ososkov M. V. see Karateev S. L.		
Pavelyeva E. A. and Krylov A. S. Hermite Transform Based Iris Key Points Selection and Analysis	1	79
Pechinkin A. V. see Frenkel S. L.		
Protasov V. Drawing Up of an Identikit Using Evolutionary Morphing and Qualimetry Method	1	83
Rudakov K. V. and Torshin I. Yu. Solvability Problems in the Protein Secondary Structure Recognition	2	25
Satin Ya. A. see Zeifman A. I.		
Seiful-Mulukov R. B. Petroleum as a Carrier of Information on Its Origin, Structure, and Evolution	1	41
Semendyaev N. N. see Sinitsyn I. N.		
Serebryakov V. A. see Zakharov A. A.		
Shestakov O. V. Normal Approximation for Distribution of Risk Estimate for Wavelet Coefficients Thresholding when Using Sample Variance	4	73
Shestakov O. V. see Markin A. V.		
Shevtsova I. G. see Grigorieva M. E.		
Shorgin S. Ya. see Kudriavtsev A. A.		
Shorgin S. Ya. see Zeifman A. I.		
Sinitsyn I. N., Sinitsyn V. I., Korepanov E. R., Belousov V. V., and Semendyaev N. N. On-Line Information Model Building of the Earth Pole Motion by Linear and Linearized Filters	1	2
Sinitsyn V. I. see Sinitsyn I. N.		

	Issue	Page
Sipina A. V. see Bening V. E.		
Sokolov I. A. see Ilyushin G. Y.		
Soldatov S. A. see Kolesnikov A. V.		
Stepanov S. Yu. Application of the Coordinate Method of Commutated Neural Network Fragmentation for Traffic Reduction	2	57
Timonina E. E. see Grusho A. A.		
Torshin I. Yu. see Rudakov K. V.		
Tsiskaridze A. Mathematical Model and Human Pose Reconstruction Method Based on Stereo- mate Image Silhouettes.....	4	27
Ulyanov V. V. see Kawaguchi Y.		
Vizilter Yu. V. see Karateev S. L.		
Zakharov A. A. and Serebryakov V. A. Libmeta — Digital Library Management System	4	2
Zakharov V. N. see Baranov S. I.		
Zakharova T. V. see Matveeva S. S.		
Zatsarinny A. A. and Chuprakov K. G. Some Issues of Technology Selection to Build Visualization System in Situational Center.....	3	59
Zatsman I. M. see Buntman N. V.		
Zeifman A. I., Korotysheva A. V., Satin Ya. A., and Shorgin S. Ya. On Stability for Nonstationary Queueing Systems with Catastrophes	3	9
Zhelto S. Yu. see Karateev S. L.		
Zykova Z. P. see Arkhipov O. P.		

Технический редактор Л. Кокушкина

Титульный редактор Т. Торжкова

Художественный редактор М. Седакова

Сдано в набор 04.10.10. Подписано в печать 03.12.10. Формат 60 × 84/8

Бумага офсетная. Печать цифровая. Усл.-печ. л. 13,5. Уч.-изд. л. 11,1. Тираж 200 экз.

Заказ №

Издательство «ТОРУС ПРЕСС», Москва 119991, ул. Косыгина, д. 4

torus@torus-press.ru; <http://www.torus-press.ru>

Отпечатано в ППП «Типография «Наука» с готовых файлов

Москва 121099, Шубинский пер., д. 6

Правила подготовки рукописей статей для публикации в журнале «Информатика и её применения»

Журнал «Информатика и её применения» публикует теоретические, обзорные и дискуссионные статьи, посвященные научным исследованиям и разработкам в области информатики и ее приложений. Журнал издается на русском языке. Тематика журнала охватывает следующие направления:

- теоретические основы информатики;
- математические методы исследования сложных систем и процессов;
- информационные системы и сети;
- информационные технологии;
- архитектура и программное обеспечение вычислительных комплексов и сетей.

1. В журнале печатаются результаты, ранее не опубликованные и не предназначенные к одновременной публикации в других изданиях. Публикация не должна нарушать закон об авторских правах. Направляя свою рукопись в редакцию, авторы автоматически передают учредителям и редколлегии неисключительные права на издание данной статьи на русском языке и на ее распространение в России и за рубежом. При этом за авторами сохраняются все права как собственников данной рукописи. В связи с этим авторами должно быть представлено в редакцию письмо в следующей форме: Соглашение о передаче права на публикацию:

«Мы, нижеподписавшиеся, авторы рукописи « _____ », передаем учредителям и редколлегии журнала «Информатика и её применения» неисключительное право опубликовать данную рукопись статьи на русском языке как в печатной, так и в электронной версиях журнала. Мы подтверждаем, что данная публикация не нарушает авторского права других лиц или организаций. Подписи авторов: (ф. и. о., дата, адрес)».

Редколлегия вправе запросить у авторов экспертное заключение о возможности опубликования представленной статьи в открытой печати.

2. Статья подписывается всеми авторами. На отдельном листе представляются данные автора (или всех авторов): фамилия, полное имя и отчество, телефон, факс, e-mail, почтовый адрес. Если работа выполнена несколькими авторами, указывается фамилия одного из них, ответственного за переписку с редакцией.

3. Редакция журнала осуществляет самостоятельную экспертизу присланных статей. Возвращение рукописи на доработку не означает, что статья уже принята к печати. Доработанный вариант с ответом на замечания рецензента необходимо прислать в редакцию.

4. Решение редакционной коллегии о принятии статьи к печати или ее отклонении сообщается авторам. Редколлегия не обязуется направлять рецензию авторам отклоненной статьи.

5. Корректурa статей высылается авторам для просмотра. Редакция просит авторов присылать свои замечания в кратчайшие сроки.

6. При подготовке рукописи в MS Word рекомендуется использовать следующие настройки. Параметры страницы: формат — А4; ориентация — книжная; поля (см): внутри — 2,5, снаружи — 1,5, сверху — 2, снизу — 2, от края до нижнего колонтитула — 1,3. Основной текст: стиль — «Обычный»: шрифт Times New Roman, размер 14 пунктов, абзацный отступ — 0,5 см, 1,5 интервала, выравнивание — по ширине. Рекомендуемый объем рукописи — не свыше 25 страниц указанного формата. Ознакомиться с шаблонами, содержащими примеры оформления, можно по адресу в Интернете: <http://www.ipiran.ru/journal/template.doc>.

7. К рукописи, предоставляемой в 2-х экземплярах, обязательно прилагается электронная версия статьи (как правило, в форматах MS WORD (.doc) или LaTeX (.tex), а также — дополнительно — в формате .pdf) на дискете, лазерном диске или по электронной почте. Сокращения слов, кроме стандартных, не применяются. Все страницы рукописи должны быть пронумерованы.

8. Статья должна содержать следующую информацию на русском и английском языках: название, Ф.И.О. авторов, места работы авторов и их электронные адреса, аннотация (не более 100 слов), ключевые слова. Ссылки на литературу в тексте статьи нумеруются (в квадратных скобках) и располагаются в порядке их первого упоминания. Все фамилии авторов, заглавия статей, названия книг, конференций и т. п. даются на языке оригинала, если этот язык использует кириллический или латинский алфавит.

9. Присланные в редакцию материалы авторам не возвращаются.

10. При отправке файлов по электронной почте просим придерживаться следующих правил:

- указывать в поле subject (тема) название журнала и фамилию автора;
- использовать attach (присоединение);
- в случае больших объемов информации возможно использование общеизвестных архиваторов (ZIP, RAR);
- в состав электронной версии статьи должны входить: файл, содержащий текст статьи, и файл(ы), содержащий(е) иллюстрации.

11. Журнал «Информатика и её применения» является некоммерческим изданием, и гонорар авторам не выплачивается.

Адрес редакции: Москва 119333, ул. Вавилова, д. 44, корп. 2, ИПИ РАН

Тел.: +7 (499) 135-86-92 Факс: +7 (495) 930-45-05 E-mail: rust@ipiran.ru