

На правах рукописи



Скоробогатова Наталья Евгеньевна

**МОДЕЛИ ПРЕДСТАВЛЕНИЯ И АЛГОРИТМЫ
РАСПОЗНАВАНИЯ РУССКИХ ДАКТИЛЕМ**

Специальность: 05.13.01 – Системный анализ, управление
и обработка информации (технические система)

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Рязань 2012

Работа выполнена в ФГБОУ ВПО
«Рязанский государственный радиотехнический университет»

Научный руководитель: **Пылькин Александр Николаевич**
Заслуженный работник высшей школы РФ,
доктор технических наук, профессор
ФГБОУ ВПО «РГРТУ», заведующий
кафедрой «Вычислительная и прикладная
математика» (г. Рязань)

Официальные оппоненты: **Ковшов Евгений Евгеньевич**
доктор технических наук, профессор
ФГБОУ ВПО МГТУ «СТАНКИН»,
заведующий кафедрой «Управления и
информатика в технических системах»
(г. Москва)

Сосулин Юрий Андреевич
кандидат технических наук, доцент
кафедры «Автоматизация технологических
процессов и производств»
ФГБОУ ВПО «РГРТУ» (г. Рязань)

Ведущая организация: ФГБОУ ВПО «Липецкий государственный
педагогический университет» (г. Липецк)

Защита состоится «28» ноября 2012г. В 12 часов на заседании диссертационного совета Д212.211.01 в ФГБОУ ВПО «Рязанский государственный радиотехнический университет» по адресу: 390005, г. Рязань, ул. Гагарина, д 59/1.

С диссертацией можно ознакомиться в библиотеке ФГБОУ ВПО «Рязанский государственный радиотехнический университет».

Автореферат разослан « 22 » октября 2012 г.

Ученый секретарь
диссертационного совета
канд. тех. наук



В.Н. Пржегорлинский

I. ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность работы. В настоящее время активное развитие компьютерных технологий открывает перспективы помощи людям с ограниченными возможностями. Особенно это касается решения проблем общения слабослышащих и глухонемых людей с обществом.

Количество людей, использующих в качестве средства коммуникации жестовую и дактильную речь, достаточно велико. На учете во Всероссийском обществе глухих состоят более 230 тысяч человек. Такой речью пользуются постоянно до трех процентов населения, если учесть тех, кто вынужден с глухими общаться в семье и других местах. Ввод в компьютер и распознавание дактильных знаков автоматически решают важную социальную задачу перевода дактильной речи глухих и слабослышащих людей в текст, понятный всем обычным людям.

Одним из путей помощи таким людям является создание программы распознавания и озвучивания русских дактильных знаков в замедленном шаговом режиме, т.к. в настоящее время перевод русской дактильной речи едва ли возможен в реальном режиме времени.

В настоящее время в большинстве случаев ввод смысловой видеоинформации в компьютер в разных странах реализуется с использованием знаков существующих, сильно различающихся, национальных дактилем.

При этом следует отметить, что сами дактильные знаки легче реализуются и распознаются в одноручном американском, немецком или испанском языке. Сложнее формируются дактилемы русского языка, которые демонстрируются одной рукой, но с разворотом знаков в трех направлениях видеокадра и поэтому сопровождаются увеличенными траекториями движений элементов (пальцев). Это существенно затрудняет автоматизацию процесса распознавания и ввода в компьютер русской дактильной речи.

Сложность построения системы перевода дактильных знаков заключается в особенностях демонстрации (дактилирования) самой речи.

Во-первых, дактильная речь в соответствии с правилами дактилирования, должна быть непрерывной, плавной, с непрерывными переходами от одного демонстрируемого знака к другому, что требует от системы отслеживания всех переходных промежуточных образов и выбора единственного информационного знака.

Во-вторых, дактильная речь может демонстрироваться на различном расстоянии от видеокамеры и людьми с различными размерами руки, поэтому распознающая система должна быть инвариантна к масштабу изображения дактилемы.

В-третьих, положение руки при демонстрации знака не фиксировано в поле кадра и может смещаться и наклоняться в зависимости от привычки воспроизводящего дактильную речь, поэтому распознающая система должна быть инвариантна к смещению и повороту распознаваемого знака.

Существующие методы распознавания дактильной речи за рубежом базируются на упрощающем перекодировании дактилем и приближенном распознавании статистическими методами либо с использованием системы многомерного отслеживания динамических образов с помощью специальных перчаток, что требует специальных стационарных систем. В этом отношении перевод хотя бы только дактильных знаков речи выглядит доступнее. Однако пока неизвестно его практическое решение для русского языка, и поэтому остается проблема автоматизированного перевода русской дактильной речи в обычную.

В диссертации приводится, в рамках общего подхода к построению таких систем, разработка моделей представления дактилем и нескольких алгоритмов шагового распознавания русской дактильной речи, отличающихся методами предварительной обработки и кодирования дактильных знаков, способами селекции дактилем и методами их распознавания. В частности в диссертации разработаны:

- модель представления дактилем в виде вектор-контура, состоящего из элементарных векторов в комплекснозначном пространстве;
- алгоритм распознавания русских дактилем с шаговой демонстрацией на основе контурного анализа (действующий с результатами исследования);
- алгоритм распознавания динамических дактилем с использованием контурного представления и селекцией знаков по изменению направлений траекторий элементов (с моделированием блока селекции дактилем);
- модель представления дактилем в виде веера векторов комплексно-значного пространства с вершиной в центре тяжести контура дактилемы;
- алгоритм с шаговым представлением дактилем и анализом движений элементов на основе веера векторов (действующая с результатами исследования);
- алгоритм распознавания динамических дактилем с использованием веера векторов и селекцией знаков по скорости промежуточных траекторий (с моделированием селектирующей части).

Кроме того, в экспериментальных целях были рассмотрены с целью сравнительного анализа два дополнительных алгоритма:

- алгоритм растрового представления дактилем для распознавания с использованием вейвлет-разложения (результаты эксперимента с растровыми дактилемами);

- алгоритм распознавания растровых полноцветных изображений дактилем и звукового воспроизведения слов дактильной речи (результаты эксперимента с возможностью озвучивания дактилем).

В частности, шаговое представление русских дактилем вводит определенную условность их идентификации и накладывает существенные ограничения на действия демонстрирующего.

Цель диссертационной работы состоит в разработке моделей представления и алгоритмов распознавания в шаговом режиме русских дактильных знаков с помощью компьютера и перевода их в текстовый или звуковой формат.

Для достижения поставленной цели решались следующие задачи:

- анализ существующих методов распознавания иностранных национальных дактилем (американских, английских, немецких, испанских, японских и др.);
- разработка математической модели контурного представления русских дактилем;
- разработка алгоритма распознавания статических шаговых дактилем с контурным представлением;
- разработка математической модели описания веера дактилем с помощью векторов;
- разработка алгоритма распознавания статических дактилем, представляемых веером векторов;
- разработка математической модели селекции динамического знака на основе изменения направления движения траекторий;
- разработка алгоритма селекции динамического знака на основе анализа интервала демонстрации между знаками.

Методы исследования: решение перечисленных задач основывается на системном анализе, теории распознавания образов, теории контурного анализа изображений, теории цифровой обработки изображений.

Научная новизна диссертационной работы определяется тем, что впервые разработаны способы описания контуров русских дактильных знаков глухих и слабослышащих людей и ввода информации в компьютер для распознавания дактильной речи. Конкретно получены следующие научные результаты:

1. Предложена математическая модель представления и дактильных знаков в виде вектор-контуров, обеспечивающая инвариантность к линейным преобразованиям смещения, поворота и масштаба дактилем.
2. Разработаны алгоритмы распознавания статических и динамических шаговых дактильных знаков с контурным представлением.

3. Предложена математическая модель представления дактильных знаков в виде веера векторов, сокращающая объем вычислительных операций и обеспечивающая при распознавании инвариантность к смещению и масштабу.
4. Разработаны алгоритмы распознавания статических и динамических шаговых дактильных знаков в виде веера векторов.
5. Разработан алгоритм выделения (селекции) динамических дактильных знаков из непрерывной последовательности дактилирования по изменению направления движения переходных траекторий элементов.
6. Разработан алгоритм выделения динамических дактилем по скорости движения элементов знака.

Достоверность и обоснованность полученных результатов диссертационной работы подтверждаются компьютерным моделированием и приводимыми результатами компьютерных экспериментальных исследований алгоритмов распознавания шаговых дактилем с контурным и векторным представлением и результатами моделирования алгоритмов селекции динамических дактилем.

Разработанные алгоритмы использованы в Московском психолого-социальном университете (филиал в г.Рязани) в группе глухих слушателей (акт об использовании от 16 февраля 2012 г.), внедрены в учебный процесс на кафедре «Вычислительной и прикладной математики» Рязанского государственного радиотехнического университета и используются студентами специальностей 230105 «Программное обеспечение вычислительной техники и автоматизированных систем» и 080801 «Прикладная информатика (в экономике)» по дисциплинам «Системы искусственного интеллекта», «Интеллектуальные информационные системы» и «Методы и алгоритмы принятия решений» (акт внедрения от 02.02.2012 г.), получена государственная регистрация программного продукта «Программа распознавания русских дактилем» (свидетельство о гос. регистрации программы для ЭВМ № 2012614984 от 4 июня 2012 г.).

На защиту выносятся следующие научные результаты:

1. Математическая модель представления и распознавания дактильных знаков в виде вектор-контур в комплекснозначном пространстве.
2. Алгоритм распознавания русских шаговых дактилем на основе контурного представления в комплекснозначном пространстве.
3. Алгоритм распознавания русских динамических дактилем с контурным представлением и селекцией по изменению направлений траекторий.
4. Математическая модель представления дактильных знаков в виде веера векторов.

5. Алгоритм распознавания шаговых дактильных знаков с представлением веером векторов.
6. Алгоритм распознавания динамических дактилем с представлением веером векторов и селекцией по скорости перемещения элементов.

Практическая значимость работы заключается в следующем:

- разработаны на основе предложенных моделей дактилем два действующих алгоритма распознавания знаков русской дактильной речи, позволяющих также обеспечить в шаговом режиме визуальный дистанционный ввод текстовой информации в компьютер;

- разработанные алгоритмы перевода знаков русской дактильной речи могут быть установлены на компьютерах в общественных приемных, больницах, магазинах и других местах, где глухим требуется объясняться на дактиле;

- разработанные алгоритмы позволяют пользоваться ими во время дистанционных видеоконференций глухих и слабослышащих людей;

- разработанные алгоритмы распознавания русских дактилем используют ограниченные вычислительные ресурсы и поэтому могут быть применены в виде дополнительной функции перевода речи глухих и слабослышащих людей для сотовых устройств, обладающих видеозаписью и звуковым воспроизведением речи;

- разработанные алгоритмы распознавания динамических дактильных знаков позволяют переводить замедленную или записанную на видеокамеру дактильную речь с выделением дактилем из потока непрерывного дактилирования;

Соответствие паспорту специальности. Согласно паспорту специальности 05.13.01 «Системный анализ, управление и обработка информации (технические системы)» проблематика, рассмотренная в диссертации, соответствует следующим областям исследований:

- формализация и постановка задач системного анализа, оптимизации, управления, принятия решений и обработки информации;

- разработка проблемно-ориентированных систем управления, принятия решений и оптимизации технических, экономических, биологических, медицинских и социальных объектов;

- методы и алгоритмы интеллектуальной поддержки при принятии управленческих решений в технических, экономических, биологических, медицинских и социальных системах.

Результаты внедрения. Разработанные алгоритмы использованы в Московском психолого-социальном университете (филиал в г.Рязани) в группе глухих слушателей (акт об использовании от 16 февраля 2012 г.), внедрены в учебный процесс на кафедре «Вычислительной и прикладной математики» Рязанского государственного радиотехнического университета и используются студентами специальностей

230105 «Программное обеспечение вычислительной техники и автоматизированных систем» и 080801 «Прикладная информатика (в экономике)» по дисциплинам «Системы искусственного интеллекта», «Интеллектуальные информационные системы» и «Методы и алгоритмы принятия решений» (акт внедрения от 02.02.2012 г.), осуществлена государственная регистрация программного продукта «Программа распознавания русских дактилем» (свидетельство о гос. регистрации программы для ЭВМ № 2012614984 от 4 июня 2012 г.).

Апробация работы. Результаты диссертационной работы обсуждались на следующих конференциях:

- 16-я международная научно-техническая конференция «Проблемы передачи и обработки информации в сетях и системах телекоммуникаций», Рязань: РГРТУ, 2010 г.

- международная научно-техническая конференция «Информационное общество: идеи, технологии, системы» ИНФО-10, Таганрог: ТТИ ЮФУ, 2010 г.

- 25-я международная научно-техническая конференция «Математические методы и информационные технологии в экономике, социологии и образовании», Пенза: ПГТА, 2010 г.

- международная научно-техническая конференция «Инновационные подходы к применению информационных технологий в профессиональной деятельности», Белгород: БРИПК, 2010 г.

- 66-я международная научно-техническая конференция Российского НТО радиотехники, электроники и связи им. Попова, Научная сессия, Москва, 2011 г.

- 16-я всероссийская НТК «Новые информационные технологии в научных исследованиях», НИТ-2011, Рязань: РГРТУ, 2011 г.

- 3-я всероссийская межвуз. конференция «Наука и образование в промышленной, социальной и экономической сферах регионов России», Муром: РГА, 2011 г.

- 2nd International Scientific Conference «European Science and Technology», Wiesbaden, Germany, 2012г.

Публикации. По результатам исследований опубликовано 17 печатных работ, в том числе 2 научные статьи в изданиях, рекомендованных ВАК для публикации трудов на соискание ученых степеней, 8 в сборниках трудов научных конференций, 7 в межвузовских сборниках научных трудов. Получено свидетельство о государственной регистрации.

Структура и объем работы. Диссертация состоит из введения, 4 глав, заключения и приложения. Работа изложена на 176 странице машинописного текста, содержит 65 рисунка, 5 таблиц, список литературных источников из 86 наименований.

II. СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность работы; изложены цель и задачи исследования, научная новизна, практическая значимость и апробация; сформулированы основные положения, выносимые на защиту.

Первая глава посвящена обоснованию темы диссертации. Проведен анализ способов демонстрации знаков дактильной речи с углубленным рассмотрением особенностей дактилирования, влияющих на принципы построения систем автоматизированного распознавания дактильных знаков.

Рассмотрены вопросы актуальности разработки систем распознавания дактилем.

Дактилемы (буквы из дактильного алфавита) или дактильные знаки формируются с помощью различного положения пальцев и функционально аналогичны буквенному алфавиту.

В диссертации показано, что существующие в мире системы дактильной речи для глухих и слабослышащих людей сильно отличаются по значению дактилем и способам их демонстрации. Одни и те же буквы латинского алфавита, используемые в разных языках, имеют абсолютно разные дактилемы и часто даже демонстрируются разными способами. Поэтому системы перевода дактильных знаков в разных странах не только не могут быть универсальными, но и должны различаться по методам и структуре построения распознающих систем. Более того, даже системы, разрабатываемые для одинаковых целей (ввод информации в компьютер или организация видеоконференций) требуют приспособления к особенностям национальных дактилем и, в общем случае, неприменимы в других странах.

В диссертации описана специфика представления русских дактилем и доказана невозможность использование созданных зарубежных систем распознавания для русской дактильной речи, которая резко отличается демонстрацией из трех разных направлений и увеличенными траекториями перехода между знаками.

При разработке алгоритмов автоматизации распознавания дактилем возникают существенные трудности, вызываемые особенностями дактилирования. При реальной непрерывной демонстрации дактильной речи возникает дополнительная и, пожалуй, основная трудность распознавания – выделение знака из непрерывной последовательности чередования букв и переходов между ними.

Следовательно, в качестве исходной информации для определения момента селекции знака сами распознанные дактилемы не могут быть использованы.

В работе приводится обзор существующих иностранных систем распознавания дактильных знаков и классификация методов распознавания.

В диссертации отмечается, что выделение рук - это только первая часть решения задачи распознавания дактилемы. Основная проблема возникает, если учесть, что в процессе демонстрации жестовой речи осуществляется непрерывное движение рук и при этом сами пальцы могут представлять определенные дактильные образы.

Следуя схожему методу распознавания речи, многие предыдущие попытки машинного распознавания концентрировались на отдельных жестах или на побуквенном вводе. Большая часть попыток была основана на применении специальных перчаток или установленных трех камер и использовала нейронные сети для распознавания.

Таким образом, цель разработки шаговых систем распознавания русских дактилем состоит в том, чтобы появилась возможность использовать сравнительно простое программное обеспечение, пригодное для устройств мобильной сотовой связи. При этом система с использованием контурного анализа обеспечивает инвариантность к масштабу, повороту и сдвигу дактилемы. Система с веером векторов использует значительно меньший объем информации при обработке знаков и, следовательно, имеет большее быстродействие и при этом сохраняет инвариантность к масштабу.

Во второй главе рассматриваются вопросы теории представления контуров дактилем в комплекснозначном пространстве с помощью элементарных векторов, позволяющих формировать вектор-контур и дополнять с ними аналитические преобразования.

При работе с конкретной дактилемой на основе разработанной модели необходимо ее контур представить в виде цифрового сигнала, описывающего границы и, следовательно, форму соответствующего оцифрованного бинарного изображения знака.

Для обработки контура аналитическим путем или на ЭВМ необходимо произвести его кодирование, т.е. поставить в соответствие каждому контурному элементу, каким является элементарный вектор, определенное комплексное число. Последовательность таких чисел и будет называться кодом контура. Фундаментальным требованием предложенной модели для контуров дактильных знаков являются их замкнутость и непрерывность линий.

На рис. 1 приводится бинаризованное изображение дактилемы и ее выделенный контур, представляемый в закодированном виде элементарными векторами, которые записаны в нижней части окна распознающей системы.

Таким образом, в комплекснозначном пространстве формируется для каждого дактильного знака свой вектор-контур, обозначаемый в

ответствующих исходному и преобразованному контурам, поэтому аргумент вектора скалярного произведения равен углу поворота контура;

- модуль нормированного скалярного произведения вектор-контуров дактильных знаков в комплекснозначном пространстве, в отличие от действительного, является характеристикой близости двух контуров, причем, и это главное, инвариантной к линейным преобразованиям сдвига, поворота и масштабирования.

При смещении начальной точки отсчета на d элементов, что происходит в рассматриваемых системах в процессе поиска максимума взаимно-корреляционной функции, у вектор-контура Γ это приводит к сдвигу номеров элементарных векторов на ту же величину $\gamma^{(d)}(n) = \gamma(n + d)$, а у суммарного кода новое значение $\beta'(n)$ определяется как $\beta'(n) = \beta(n) + \beta_0$, где β_0 - комплексное число, определяющее положение начальной точки контура относительно новой системы координат.

При многократном обходе контура, что в системах распознавания происходит в случае неизвестной начальной точки контура, элементы его кода обладают периодичностью с периодом k , равным числу элементарных векторов в вектор-контуре знака, т.е. $\gamma(n) = \gamma(n \pm \alpha k)$, $\beta(n) = \beta(n \pm \alpha k)$, $\alpha = 0, 1, \dots, k$. (2)

Это исключительно важное свойство контурного представления дактильных знаков, т.к. позволяет обойтись без жесткой синхронизации по времени процесса сравнения элементарных векторов распознающих и эталонных дактилем.

Важнейшим свойством комплекснозначного представления дактилем в модели является возможность их преобразования аналитическим путем. В этом отношении можно отследить связь между цепными кодами исходного и преобразованного контуров. Так, при умножении контура на комплексный множитель $\mu = |\mu| \exp\{j\Delta\varphi\}$ происходит растяжение этого контура в $|\mu|$ раз и поворот на угол $\Delta\varphi$

$$\Gamma_\mu = \mu \Gamma = \{|\mu|\} \exp\{j\Delta\varphi\} \gamma(n), \quad n = 0, 1, \dots, (k-1). \quad (3)$$

Это позволяет производить в системах преобразования с образцами дактилем и отслеживать их новые параметры.

При разработке систем распознавания дактилем особая роль принадлежит алгоритмам сравнения анализируемого и эталонного контуров, т.к. от них зависят такие параметры системы, как чувствительность и независимость от искажений. Один из способов определения степени

близости контуров в системах состоит в вычислении в комплекснозначном пространстве их скалярного произведения.

Скалярное произведение может быть определено как:

$$\Psi = (\Gamma, \mathbf{V}) = \sum_{n=0}^{k-1} |\gamma(n)| |v(n)| \exp \{j\Delta\phi(n)\}. \quad (4)$$

В разработанных алгоритмах чаще используется нормированное скалярное произведение Ψ_H , определяемое для комплексного пространства как

$$\Psi_H = \frac{\Psi}{\|\Gamma\| \|\mathbf{V}\|} \quad (5)$$

и являющееся, в общем случае, комплексной величиной.

Алгоритм вычисления взаимно корреляционной функции контуров, так же как и нормированное скалярное произведение, используется в качестве метода сравнения дактильных знаков распознающих систем. При этом взаимно корреляционная функция $\tau(m)$, как полный набор значений скалярных произведений вектор-контура \mathbf{V} на вектор-контур $\Gamma^{(m)}$ со смещенным положением начальной компоненты $\gamma(m)$, $m = 0, 1, \dots, (k-1)$ является достаточно чувствительной характеристикой близости двух контуров.

Таким образом, корреляционная функция $\mathbf{T} = \{\tau(m)\}_{0, k-1}$ периодических комплекснозначных контуров вводится как упорядоченный набор скалярных произведений контура \mathbf{V} одного из сигналов на контур $\Gamma^{(m)}$ другого сигнала со смещением на m элементов. В отличие от вещественных сигналов при нахождении каждого из отсчетов один из контуров берется комплексно-сопряженным.

Для построения распознающих систем важно, что автокорреляционная функция $\tau(m)$ контура Γ служит характеристикой близости этого контура со своей копией $\Gamma^{(m)}$ имеющей в качестве начальной компоненты элементарный вектор $\gamma(m)$. Все особые свойства скалярного произведения относятся также к корреляционным функциям и поэтому могут в процессе распознавания определять степень близости дактилемы, совпадающей с одной из эталонных, с помощью вычисления автокорреляционной функции, а также не совпадающей - путем вычисления взаимно корреляционной функции. Практически в рассматриваемых ниже системах предполагается осуществлять вычисление взаимно корреляционной функции, а автокорреляционная функция вычис-

ляется в единственном случае - в момент совпадения распознаваемого и эталонного знака.

В **третьей главе** приводится разработка двух распознающих алгоритмов, использующих контурное представление дактильных знаков в комплекснозначном пространстве.

В диссертации приводятся четыре конкретные структурные схемы алгоритмов по каждому из рассматриваемых методов распознавания дактилем.

Первый из алгоритмов, приводимый на рис. 2, предназначен для распознавания **статических дактилем** при их шаговой демонстрации.

Во всех алгоритмах анализируемый дактильный знак **V** с выхода теле-, видео- или веб-камеры поступает на входной блок системы, реализующей предварительную обработку по своим алгоритмам, зависящим от конкретного метода.

Частота кадров видеокамеры достаточно высока для обработки всей последовательности изображений, поэтому в следующем блоке обработки осуществляется деление кадров до частоты, обеспечивающей подачу порядка десяти кадров на период между реальными демонстрациями дактилем. Эти промежуточные кадры необходимы для отслеживания траектории перемещения элементов знака.

После предварительной обработки дактилема кодируется в зависимости от выбранного метода: контурами в виде элементарных векторов или веером векторов.

И далее, закодированный в комплекснозначном пространстве по своим алгоритмам дактильный знак **V** поступает на блок сравнения, который сравнивает его со знаками, хранящимися в библиотеке эталонных дактилем в комплексно сопряженном виде $\Gamma^*_{(A,B,...Я)}$ и при совпадении с одним из них, выдает значение соответствующей буквы русского алфавита.

По времени операция распознавания не совпадает с этапом формирования эталонных дактильных знаков, которая осуществляется заранее и в ходе статического представления, возможно с помощью фотокамеры, и с обязательной предварительной обработкой знака по тем же алгоритмам, что и распознаваемый. На рис.2 эта операция проходит по тем же блокам, что и основная операция распознавания, разделяясь только по времени на предварительное формирование эталонных знаков.

Безусловно, у каждой из рассматриваемых систем свои алгоритмы преобразования дактильных знаков и форматы их представления, но есть и общие операции, которые реализуются во всех системах, а именно:

- преобразования цветных растровых полутоновых изображений дактилем, получаемых с выхода видео- или веб-камер, в формат монохромных бинаризованных изображений;
- выделение контуров бинаризованных изображений;
- определение направления демонстрации знака;
- удаление неинформативной части дактильного знака и проверка замкнутости контуров;
- центрирование выделенных знаков дактилем.

Во всех алгоритмах осуществляется операция удаления неинформативной части руки, но перед этим, что особенно важно для русских дактилем, определяется направление демонстрируемой дактилемы. Это учитывается при сопоставлении с отдельным методом хранения эталонных дактилем из трех различных направлений.

При построении алгоритмов распознавания дактильных знаков с контурным представлением важная роль принадлежит выбору способа сравнения анализируемой $\mathbf{V} = \{\nu(n)\}_{0,k-1}$ и эталонной дактилем $\mathbf{\Gamma}^* = \{\gamma(n)\}_{0,mk-1}$. Механизм работы контурного согласованного фильтра заключается в том, что при подаче на его вход настроенной на этот фильтр дактилемы образуется контур $\mathbf{T} = \{\tau(m)\}_{0,k-1}$, определяемый для его составляющих соотношением

$$\tau(m) = \sum_{n=0}^{k-1} \gamma(n) \gamma^*(n - m + k - 1), \quad (6)$$

а это выражение представляет собой смещенную на k отсчетов автокорреляционную функцию фильтруемого контура Γ .

Поэтому в момент $m = k - 1$ получается максимальный отсчет

$$\tau(k - 1) = \sum_{n=0}^{r-1} \gamma(n) \gamma^*(n) = \sum_{n=0}^{k-1} |\gamma(n)|^2 = \|\Gamma\|^2. \quad (7)$$

В связи с настройкой общего алгоритма на шаговое распознавание, дактилемы демонстрируются таким образом, чтобы переход от одного знака к другому занимал минимальное время, а демонстрация самого знака длилась возможно дольше, в зависимости от настройки и, возможно, до индикации завершения распознавания предыдущей буквы.

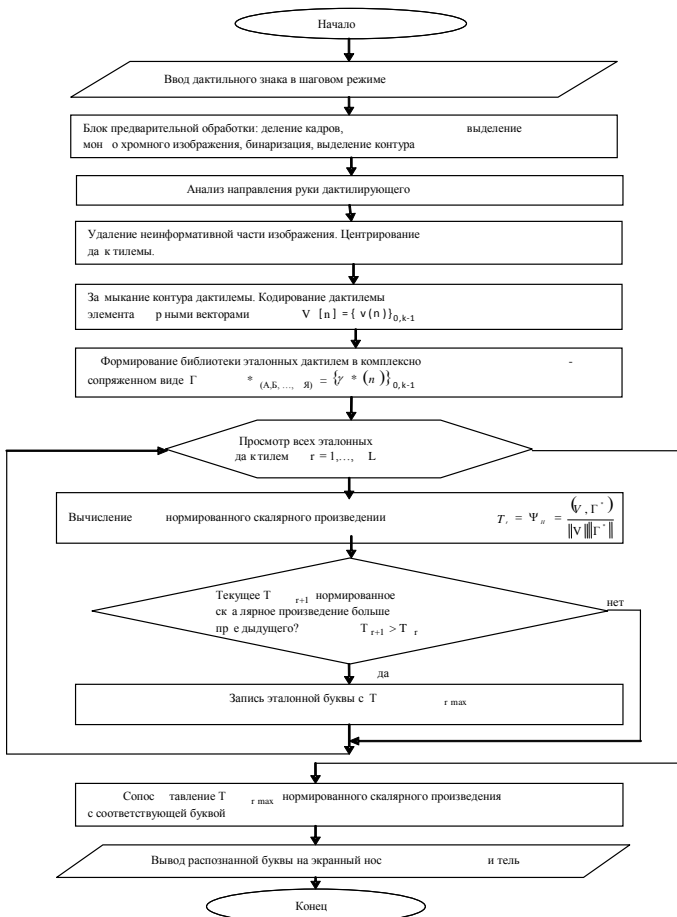


Рис. 2 Алгоритм распознавания шаговых дактилем

Таким образом, в шаговом распознающем алгоритме дактильный знак хотя и может медленно меняться, но т.к. он зафиксирован моментально, то может считаться статичным $V(A)$ и независимым от времени на период его распознавания до подхода к следующей дактилеме.

После разделения библиотеки эталонных дактилем на три секции, в зависимости от направления дактилирующей руки, отпадает необходимость в неинформативной части руки, не участвующей непосредственно в формировании дактильного знака. Поэтому на изображении происходит выделение кисти руки, непосредственно демонстрирующей дактилему. Выделение кисти реализуется по уровню минимальной ши-

рины кисти, в процессе демонстрации искусственной эталонной дактилемы.

Таким образом, на выходе программного блока кодирования контуров формируются контуры распознаваемого $\mathbf{V}_{(A)} = \{v(n)\}_{0,k-1}$ и эталонного $\Gamma_{(A,B,\dots,Y)}$ знаков в комплексно-сопряженном виде. Перед библиотекой эталонов программный блок реализует преобразование элементарных векторов в комплексно-сопряженные $\Gamma_{(A,B,\dots,Y)}^* = \{\gamma^*(n)\}_{0,k-1}$.

Алгоритм осуществляет сравнение анализируемого контура со всеми эталонными контурами дактилем и по максимуму определяет значения их сумм:

$$\Gamma_{(A)}(m) = \sum_{n=0}^{k-1} v(n)\gamma^*(n-m+k-1). \quad (8)$$

На выходе программного блока сравнения алгоритм выдает значение распознаваемой буквы. На рис. 3 идентифицировано буква «В».

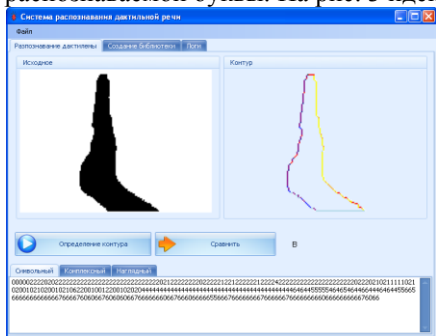


Рис. 3 Результат распознавания буквы «В»

В диссертации приводятся полные экспериментальные исследования направленные на подтверждение работоспособности алгоритмов, положенных в основу разработки системы в целом и отдельных ее частей, принципиально влияющих на параметры распознавания.

В частности, в работе предоставляется результаты исследования, подтверждающие для алгоритма шагового распознавания инвариантность контурного представления дактилем к повороту, сдвигу отсчета вдоль контура и масштабу.

Параметры системы определяются следующими характеристиками:

- Разрешение видеокамеры не менее 320x240;
- Операционная система- Windows 7, XP;
- Язык программирования – С#

- Минимальные требования к техническим параметрам компьютера - CPU не менее 1400 MHz; RAM: 256 MB; HD 10 GB;
- Расстояние до распознаваемого знака – от 0,5 до 1,5 метра;
- Действующая программ – 180МБ.

Среднее время распознавания одной дактилемы алгоритмом с контурным представлением составляет 3,9 сек. Среднее значение вероятности правильного распознавания дактилемы из всего алфавита составляет 85,3% при рассеянном бестеневом освещении и однородном фоне. Яркий контрастный фон с фигурами соизмеримыми с размерами дактилем увеличивает вероятность ошибочного распознавания до 37,4%

В диссертации разработан также алгоритм распознавания **динамических дактилем**, представленных контурами в комплекснозначном пространстве.

Этот алгоритм отличается от рассмотренного выше только включением в структуру программного блока селекции дактилем. На выходе видеокамеры непрерывная последовательность демонстрируемых дактилем $W\{V(t)\}$ в процессе предварительной обработки превращается в поток бинарных дискретизированных и разреженных по времени изображений знаков.

Алгоритм функционирования блока селекции включает в себя: блок предварительного кодирования веером векторов бинаризованного изображения, алгоритм временного анализа траекторий элементов дактилем и блок селекции дактильных знаков. Вместе они образуют алгоритм селекции. Общая структура алгоритма содержит общие с предыдущим алгоритмом распознавания статических шаговых дактилем устройства ввода и блока предварительной обработки со всеми реализуемыми функциями: выделения монохромного канала, бинаризации, анализа направления демонстрации и центрирования знаков.

Предварительное кодирование дактилем веером векторов для селекции знаков осуществляется по упрощенному, но более быстрому алгоритму. Каждый из векторов оценивается размером расстояния от центра тяжести (полюса) до края выделенного бинарного знака.

Размеры векторов веера необходимы для осуществления анализа динамики изменения направлений элементов контуров. Отслеживание динамики изменений элементов контуров происходит путем сравнения трех последовательных отсчетов $l_n(\mu - 1), l_n(\mu), l_n(\mu + 1)$ каждого из векторов. После суммирования по всем векторам команды передаются системе селекции дактилем, которая выбирает момент демонстрации очередного знака $V_{(A)} = \{v(n)\}_{0,k-1}$ для его записи и передачи дальше системе выделения и точного кодирования в комплекснозначном пространстве.

Эталонные дактильные знаки вводятся в систему с фото- или видеокамеры заранее, проходят те же преобразования, что и распознаваемые, но в статическом режиме. Кроме того, после кодирования контуров эталонных дактилем они подвергаются дополнительному преобразованию в комплексно сопряженные, что необходимо для выполнения операции сравнения дактилем.

Команда, соответствующая распознанной букве передается для вывода на экран или на систему звукового воспроизведения.

В рассматриваемом алгоритме анализа динамических дактилем большинство функций распознавания повторяются аналогично приведенному выше. Основное отличие заключается в добавленном алгоритме селекции дактильных знаков. С учетом того, что выбранный метод и алгоритм требуют проверки, была проверена принципиальная способность формирования селектирующей команды. При этом команда должна быть выделена из непрерывной последовательности динамических переходов между дактилемами, значения которых неизвестны и останутся неизвестными вплоть до завершения цикла распознавания.

Работоспособность системы селекции проверялась путем моделирования непрерывного процесса перехода от одного дактильного знака к другому.

На рис.4 показана последовательность кадров с промежутками между дактильными знаками, равными 11 кадрам видеосъемки.

Переход от одного кадра к другому проходил в режиме реальной замедленной дактилируемой речи, и на графике внизу отслеживалось в по-кадровом режиме вычисление суммы векторов, которые в соответствии с алгоритмом, приводимым в соотношениях, изменяли направление движения.

Моделирование показало работоспособность предложенного алгоритма селекции дактильных знаков. При этом алгоритм выделения информационного знака имеет достаточно общий характер и может использоваться в других системах.

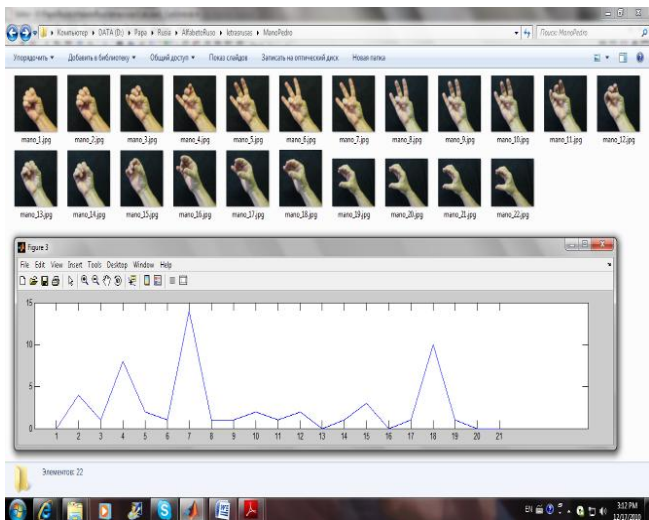


Рис. 4 – Результаты моделирования селекции знака

В четвертой главе приводится разработка алгоритма, в котором дактилемы представлены в виде модели с веером векторов.

Третий разработанный алгоритм предназначенный для распознавания шаговых дактилем, приводится на рис.5.

В диссертации установлена целесообразность совмещения полюса векторов P с центром тяжести контура O дактилемы, формируемого самими векторами.

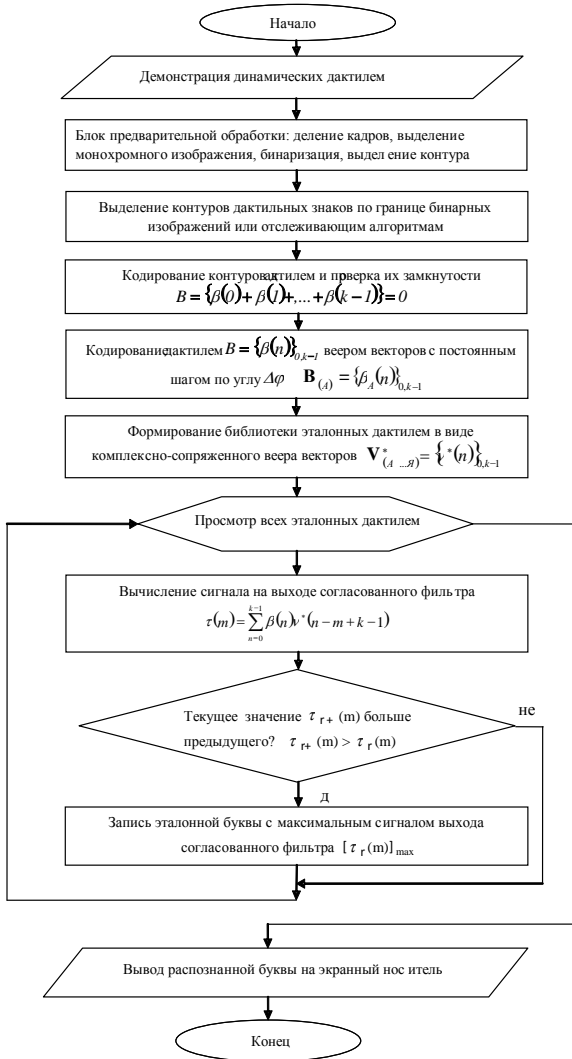


Рис 5 Алгоритм шагового распознавания дактилем на основе вектора

При этом основная информация о знаке заключена не в его контуре $\Gamma = \{\gamma(n)\}_{0,k-1}$, а в длинах его векторов $|\beta(n)|$, образующих вектор $\mathbf{B} = \{\beta(n)\}_{0,k-1}$, как показано на рисунке 6.

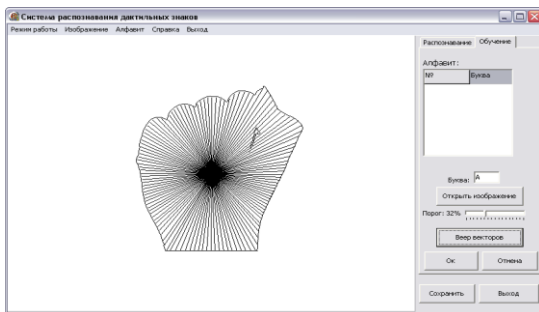


Рис. 6 – Векторное представление информативной части дактилемы

Таким образом, в системах с векторным кодированием контуров алгоритм распознавания основывается на согласованной фильтрации веера векторов.

Модуль выходного отсчета согласованного фильтра представляет меру схожести двух вееров, в том числе и несогласованных.

Если принять за эталонный веер буквы «А» $\mathbf{V}_A = \{\beta_A(n)\}$, задающий импульсную характеристику буквы «А», а анализируемый веер \mathbf{V}_Y будет демонстрировать букву «Я», то согласованный фильтр даст на выходе меру их схожести

$$\tau_{AY}(k-1) = \sum_{n=0}^{k-1} v_Y(n) \beta_A^*(n). \quad (9)$$

В нормированном виде величина этой меры будет равна

$$[\tau_{AY}(k-1)] = \frac{|\tau_{AY}(k-1)|}{\|\mathbf{V}_A\| \|\mathbf{V}_Y\|}. \quad (10)$$

Результат согласованной фильтрации, приведенный на рис. 7, может сам рассматриваться как некоторый веер, и при совпадении сравниваемых векторов его суммарный радиус-вектор будет состоять из коллинеарных векторов, имеющих общее направление, что и определяет эффект согласованной фильтрации.

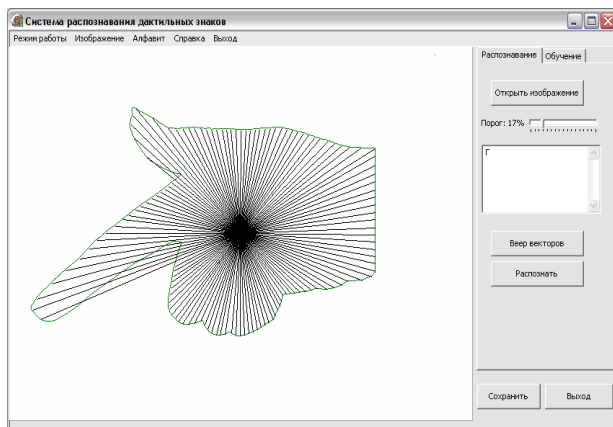


Рис.7 – Результат распознавания буквы «Г»

Из полученных результатов исследований можно сделать вывод, что программа распознает тактильные знаки при повороте кисти руки в ограниченных пределах. Изменение масштаба руки также не влияет на эффективность распознавания демонстрируемых знаков.

В диссертации приводится алгоритм распознавания динамических тактилем, представляемых веером векторов. Число векторов в веере $k=128$ и $k=256$. Среднее время распознавания одной тактилемы составляет, соответственно, 2,6 сек. и 3 сек. Средняя вероятность правильного распознавания алфавита тактилем при оптимальном бестеневом освещении, составляет 80,4% (при $k=256$). Фон с рисунками соизмеримы с размерами тактилем, увеличивает вероятность ошибочного распознавания до 34,1%.

В диссертации разработан четвертый алгоритм – распознавания динамических тактилем, представленных веером векторов.

Основное отличие алгоритма распознавания динамических тактильных знаков от рассмотренного выше состоит в том, что он содержит в своей структуре блок селекции знака и как показали результаты моделирования, может работать с последовательностью движущихся и непрерывно переходящих друг в друга тактилем. Этот алгоритм разрабатывается с целью распознавать в замедленном темпе реально тактилируемую речь.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ И ВЫВОДЫ

По результатам диссертационной работы можно сделать следующие выводы:

1. Подтверждена принципиальная возможность автоматизированного распознавания русских дактильных знаков, несмотря на повышенную сложность их демонстрации и схожесть формы некоторых дактилем.
2. Подтверждена применимость разработанных моделей представления дактильных знаков для решения задачи распознавания русских дактилем.
3. Разработанные алгоритмы шагового распознавания с контурным и веерным представлением дактилем показали возможность их использования в реальных условиях для перевода замедленного (шагового) дактилирования при общении глухих и слабослышащих людей с окружающими. Системы также могут быть использованы для обучения в специализированных школах и для самостоятельного обучения дактильной азбуке глухих и слабослышащих детей на дому и взрослых, вынужденных общаться с ними.
3. Структура, алгоритмы и программное обеспечение позволяют, при соответствующей доработке, быть реализованными в качестве дополнительной функции мобильных устройств связи, имеющих возможность видеосъемки и звукового воспроизведения.
4. Сравнительный анализ качества распознавания показал повышенную точность выполнения этой операции в алгоритмах с контурным представлением при увеличенном времени обработки, превышающем перевод дактилирования в режиме реального времени. Алгоритмы с веерным представлением дактилем при понижении точности распознавания обладают большим быстродействием.
5. Подтверждена принципиальная возможность предложенного способа селекции динамических знаков из непрерывного дактилирования с помощью анализа суммарной смены направления траекторий элементов.
6. Подтверждена возможность предложенного селективного динамических дактильных знаков за счет анализа суммарной скорости движения элементов дактилем при замедленном или предварительно записанном на видеокамеру дактилировании.
7. Проверена и экспериментально подтверждена нецелесообразность использования в алгоритмах распознавания растрового цветного представления дактилем из-за резкого увеличения времени обработки и исключительно высокой чувствительности к режиму освещения.

8. Подтверждена экспериментально возможность озвучивания системами распознавания переведенного текста дактильной речи.

IV. ПУБЛИКАЦИИ ПО ОСНОВНЫМ РЕЗУЛЬТАТАМ ДИССЕРТАЦИИ

Работы, опубликованные в научных журналах, входящих в перечень ведущих рецензируемых журналов и изданий ВАК РФ:

1. Дондик Е.М., Пылькин А.Н., Скоробогатова Н.Е. Математическое представление селекции распознаваемых знаков дактильной речи // Научно-технические ведомости СПб ГПУ выпуск 1(140), 2012. - с.55-60.
2. Дондик Е.М., Пылькин А.Н., Скоробогатова Н.Е. Информационная система распознавания образов переменной структуры // Вестник РГРТУ, №37, 2011.- с.17-22.

Работы, опубликованные в сборниках научных трудов, материалах международных и всероссийских конференций:

3. Скоробогатова Н.Е., Дондик Е.М. Система распознавания и воспроизведения знаков дактильной речи// Инновационные подходы к применению информационных технологий в профессиональной деятельности: сб. науч. тр. междунар. научно-техн. конф.- Белгород: БРИПК, 2010.- с. 287-291.
4. Дондик Е.М., Пылькин А.Н., Скоробогатова Н.Е. Система распознавания и воспроизведения знаков дактильной речи // Труды Российского НТО радиотехники, электроники и связи им. Попова, Научная сессия. – Москва, 2011.- с. 399-401.
5. Скоробогатова Н.Е., Дондик Е.М. Контурный анализ в системе распознавания дактильной речи // Математические методы и информационные технологии в экономике, социологии и образовании: тез. докл. 25-й междунар. научно-техн. конф.- Пенза: ПГТА, 2010.- с.124-127.
6. Скоробогатова Н.Е., Дондик Е.М. Распознавание знаков дактильной речи на основе контурной согласованной фильтрации // Информационное общество: идеи, технологии, системы: тез. докл. междунар. научно-техн. конф., часть 4.- Таганрог: ТТИ ЮФУ, 2010.- с.24-29.
7. Грицай Е.В., Дондик Е.М., Скоробогатова Н.Е. Распознавание знаков дактильной речи // Наука и образование в промышленной, социальной и экономической сферах регионов России: тез. докл. 3-й всерос. межвуз. конф.- Муром, 2011.- с.227-228.

8. Дондик Е.М., Скоробогатова Н.Е. Система распознавания дактильной речи // Информационные технологии в образовании: межвуз. сб. науч. тр.- Рязань: РГРТУ, 2010.- с.127-133.
9. Дондик Е.М., Пылькин А.Н., Родригес Л.М., Скоробогатова Н.Е. Система выделения знаков дактильной речи // математическое и программное обеспечение вычислительных систем: межвуз. сб. науч. тр.- Рязань: РГРТУ, 2011.- с.153-158.
10. Дондик Е.М., Грицай Е.В., Скоробогатова Н.Е. Система распознавания знаков дактильной речи на основе контурного анализа // Математическое и программное обеспечение вычислительных систем: межвуз. сб. науч. тр.- Рязань: РГРТУ, 2011.- с.153-158.
11. Дондик Е.М., Хватова А.С., Скоробогатова Н.Е. Система распознавания знаков дактильной речи на основе вектора векторов // Информационные технологии: межвуз. сб. науч. тр.- Рязань: РГРТУ, 2011.- с. 77-80.
12. Дондик Е.М., Пылькин А.Н., Скоробогатова Н.Е. Система распознавания дактильной речи// Проблемы передачи и обработки информации в сетях и системах телекоммуникаций: тез. докл. 16-й межд. научно-техн. конф. – Рязань, 2010.- с.63-65.
13. Скоробогатова Н.Е. Информационные системы распознавания русских дактилем // Новые информационные технологии в научных исследованиях: тез. док. 16-й всерос. научно-техн. конф. - Рязань: РГРТУ, 2011.- с.24-26.
14. Скоробогатова Н.Е. Структура системы распознавания динамических дактилем // Программные информационные системы: межвуз. сб. науч. тр.- Рязань: РГРТУ, 2011.- с.11-15.
15. Агеев К.А., Дондик Е.М., Некрасова Е.В., Скоробогатова Н.Е. Экспериментальные исследования систем распознавания шаговых дактилем // Математическое и программное обеспечения вычислительных систем: межвуз. сб. науч. тр.- Рязань: РГРТУ, 2011. - с.60-64.
16. Дондик Е.М., Литвинова Т.В., Скоробогатова Н.Е. Анализ динамики перехода между знаками дактильной речи // математическое и программное обеспечения вычислительных систем: межвуз. сб. науч. тр.- Рязань: РГРТУ, 2011. - с.17-21.
17. Skorobogatova N.E., Dondik E.M. Algorithm Structure of Recognition Dynamic Dactyl // European Science and Technology: 2nd International Scientific Conference, Bildungszentrum Rdk E.V. Wiesbaden, Germany, 2012. – Vol. II - p.356-360

Скоробогатова Наталия Евгеньевна

**МОДЕЛИ ПРЕДСТАВЛЕНИЯ И АЛГОРИТМЫ
РАСПОЗНАВАНИЯ РУССКИХ ДАКТИЛЕМ**

Автореферат

диссертации на соискание ученой степени
кандидата технических наук

Подписано в печать 22.10.12 Формат бумаги 60x84 1/16.

Бумага офисная. Печать трафаретная. Усл.печ.л. 1,0

Тираж 100 экз. Заказ

Рязанский государственный радиотехнический университет.

390005, г. Рязань, ул. Гагарина, 59/1.

Редакционно-издательский центр РГРТУ