

На правах рукописи



БУЛАНОВА ЮЛИЯ АНАТОЛЬЕВНА

**СИСТЕМА ДИАГНОСТИКИ ЗАБОЛЕВАНИЙ ПО
МАММОГРАФИЧЕСКИМ СНИМКАМ**

**Специальность: 05.11.17 – Приборы, системы и изделия медицинского
назначения**

АВТОРЕФЕРАТ

диссертации на соискание ученой степени

кандидата технических наук

Рязань - 2014

Работа выполнена на кафедре «Информационные системы» Муромского института (филиала) ФГБОУ ВПО «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых».

Научный руководитель: **Садыков Султан Сидыкович**
доктор технических наук, профессор

Официальные оппоненты:

Фролов Сергей Владимирович, доктор технических наук, профессор, заведующий кафедрой «Биомедицинская техника», ФГБОУ ВПО Тамбовский государственный технический университет, г. Тамбов

Мельник Ольга Владимировна, кандидат технических наук, доцент кафедры БМПЭ, ФГБОУ ВПО Рязанский государственный радиотехнический университет, г. Рязань

Ведущая организация: Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Юго-Западный государственный университет», г. Курск

Защита состоится **5 декабря 2014 г. в 12⁰⁰** на заседании диссертационного совета Д212.211.04 в ФГБОУ ВПО «Рязанский государственный радиотехнический университет» по адресу: 390005, г. Рязань, ул. Гагарина, 59/1.

С диссертацией можно ознакомиться в библиотеке РГРТУ.

Автореферат разослан «___» _____ 2014 г.

Ученый секретарь
диссертационного совета
д-р техн. наук, доцент



Г.В. Овечкин

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность работы. Самой опасной болезнью женского населения является рак молочной железы (РМЖ). По данным Всемирной организации здравоохранения (ВОЗ) каждая 5 женщина в Европе заболевает раком до достижения 75 лет. По данным Министерства здравоохранения России число впервые выявленных злокачественных новообразований ежегодно растет: показатель впервые выявленного РМЖ у женского населения в 2007 г. равен 67,9, в 2008 г. – 68,8, в 2010 г. – 75.

Маммография является самым распространенным неинвазивным рентгенологическим методом выявления заболеваний молочной железы. Большинство медицинских учреждений России оснащены аналоговыми маммографическими комплексами. Маммограммы, полученные с помощью этих комплексов имеют большое количество недостатков: частая нерезкость изображения; повышенная зернистость фона; наличие мелких белых точек на снимках и т.д. Анализ маммограмм выполняет врач-рентгенолог визуально и качество оценки снимка, а, следовательно, и результат диагноза зависит от состояния здоровья врача, его квалификации, времени осмотра снимка, достаточной подсветки изображения и т.д. Цифровые маммографические комплексы устраняют большинство недостатков пленочной маммографии, однако дороги, как в приобретении, так и в обслуживании.

Таким образом, основным недостатком пленочной маммографии является ручная обработка и визуальный анализ маммограмм.

Зарубежными ведущими производителями пленочных и цифровых маммографических аппаратов являются фирмы Agfa Healthcare, Carestream, FujiFilm, GE Healthcare, Hologic, IMS, Philips Healthcare, Planmed, Siemens Medical Solutions, IMS Giotto.

На мировом рынке существует несколько фирм, предлагающих компьютерные системы обработки маммограмм (от англ. Computer Aided Detection). Самые распространенные CAD системы: AccuDetect, Second Look Digital, The MAMMEX MammoCAD, TotalLook MammoAdvantage, syngo MammoCAD, Fujifilm's Digital Mammography System, но ни одна из фирм, производящих коммерческие CAD системы не раскрывает алгоритмы предварительной обработки, сегментации, распознавания маммографических снимков. Стоимость коммерческой CAD системы не отличается от стоимости цифрового маммографического аппарата.

Алгоритмы обработки и анализа маммограмм представлены в работах Moradmand H., Setayeshi S., Karimian A.R., Sirous M., Esmaeil M. Garg R., Mittal B., Garg S. Anguh M., A. Silva и др. В России практически не ведутся исследования по автоматизированной обработке и анализу маммографических снимков.

В связи с вышесказанным актуальной является разработка автоматизированной системы обработки и анализа маммограмм, которая будет недорогим и тиражируемым аналогом существующих зарубежных компьютерных систем диагностики новообразований в молочной железе, позволит провести обработку как скрининговых, так и снимков обследования, сократить время обработки и анализа каждого снимка, получить новые данные о динамике развития опухолей, повысить точность и объективность диагноза за счет использования большего количества данных, сформировать единую базу данных обследований и вести статистические исследования по различным параметрам.

Целью диссертационной работы является создание системы, обеспечивающей обработку и анализ оцифрованных маммографических снимков, позволяющей повысить процент правильной диагностики новообразования молочной железы.

Для достижения цели необходимо решить следующие задачи:

- обзор и анализ алгоритмов, методов, систем обработки и анализа маммографических снимков, основанных на методах цифровой обработки изображений, распознавания образов и информационных технологий;
- разработка и исследование алгоритмов предварительной обработки маммограмм;
- разработка и исследование алгоритмов выделения областей 3 типов новообразований (киста, фиброаденома, рак молочной железы) на маммографических снимках;
- разработка и исследование алгоритмов идентификации новообразований 3 типов (киста, фиброаденома, рак молочной железы) на маммографических снимках;
- создание автоматизированной системы на основе разработанных алгоритмов, программных средств и современных компьютеров и формирование базы данных обработанных маммограмм для научно-статистических исследований;
- разработка технологии проведения маммографического обследования с использованием автоматизированной системы;
- проведение исследований на группе пациентов и экспериментальная оценка возможностей разработанной системы;
- практическое применение созданной системы в маммографическом отделении больницы.

Объект исследования – аналоговые и цифровые маммографические снимки, полученные в результате маммографических обследований и скрининговых исследований.

Методы исследования. В работе использованы методы дискретной математики, математической логики, теории множеств, цифровой обработки изображений, математической статистики и распознавания образов.

Научная новизна работы

– предложена и разработана научная идея использования текстурных признаков Харалика второго порядка полутоновых маммографических изображений для описания новообразований, отличающаяся простотой расчетов матрицы GLCM и формировании на ее основе вектора-признаков и позволившая распознать быстро и достаточно точно все три типа новообразований - киста, фиброаденома, рак молочной железы;

– предложен подход выделения области молочной железы на маммограмме, основанный на случайных марковских полях, учитывающий индивидуальные особенности строения молочной железы каждой пациентки; выделения новообразований типов «Киста», «Фиброаденома», «Рак молочной железы» с использованием новых и достаточно простых алгоритмов предварительной обработки, позволяющие выделять место локализации новообразований даже на фоне фиброзно-кистозной болезни и аденоза, а также идентификации новообразований трех типов, использующий дискриминантный анализ Фишера, отличающийся расчетом дискриминантных функций и оценок независимо от количества экспериментов в группах;

– доказана экспериментально на реальных маммографических снимках перспективность практического применения предложенных научной идеи использования текстурных признаков Харалика второго порядка полутоновых изображений и подходов выделения области молочной железы, областей новообразований трех типов, а также идентификации этих новообразований в медицинских системах маммографического скрининга.

Теоретическая значимость исследования обоснована тем, что:

– применительно к задаче диссертации результативно использованы комплексы базовых методов теорий цифровой обработки изображений, распознавания образов, математической статистики;

– изложены новые идея, понятие, алгоритмы, доказательства возможности распознавания на маммограммах новообразований трех основных типов;

– установлено отсутствие подходов, алгоритмов и систем в цифровой обработке изображений и распознавании образов, позволяющих решить задачу идентификации новообразований трех основных типов;

– изучены применительно к задачам диссертации методы математической статистики, цифровой обработки изображений и распознавания образов;

– проведена модернизация существующих алгоритмов предварительной обработки, а также случайных марковских полей, обеспечивающие получение новых научных и практически результатов по теме диссертации.

Значение полученных соискателем результатов исследования для практики подтверждается тем, что:

– разработаны и внедрены технология обработки и анализа маммограмм и автоматизированная система обработки и анализа маммографических снимков в НУЗ Отделенческой больницы на станции Муром, ОАО РЖД (г. Муром, Владимирская область) и в учебном процессе на кафедре информационных систем Муромского института (филиала) ФГБОУ ВПО «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых»(г. Муром, Владимирская область);

– определены возможности использования полученных теоретических и практических разработок для идентификации других новообразований при изменении коэффициентов алгоритмов и перспективы практического использования при маммографическом скрининге;

– создана автоматизированная система обработки и анализа маммографических снимков, позволяющая выделять и идентифицировать три основных типа новообразований;

– представлены рекомендации по дальнейшему совершенствованию алгоритмов и системы распознавания по выделению других типов новообразований, а также сокращению времени и повышению точности их распознавания.

Оценка достоверности результатов исследования выявила , что:

– результаты экспериментальных работ получены на сертифицированном оборудовании и программном обеспечении; показана воспроизводимость результатов исследований, как на тестовых, так и на реальных маммографических снимках, выполненных в разных проекциях;

– использованная теория построена на известных методах цифровой обработки изображений и распознавания образов, на проверяемых данных, полученных по теме диссертации традиционным экспериментальным путем;

– идея использования текстурных признаков Харалика второго порядка полутоновых маммографических снимков для распознавания новообразований, базируется на анализе практики распознавания изображений в медицине;

– проведено сравнение результатов предварительной обработки алгоритмами автора и известными алгоритмами общего назначения, такими как метод линейного контрастирования, эквализации, нормализации и установлено, что алгоритмы автора дают результаты более высокого качества, что важно для обеспечения нужной точности распознавания новообразований;

– применены современные методики сбора и обработки исходной информации, принятые в цифровой обработке изображений и распознавании образов и использованы презентабельные выборочные данные при выполнении экспериментальных исследований.

Личный вклад соискателя состоит в:

– участии автора на всех этапах процесса исследований по реализации темы диссертации;

– личном участии в теоретических разработках, связанных с созданием и реализацией алгоритмов выделения новообразований нескольких типов;

– непосредственном участии соискателя в получении исходных изображений реальных маммографических снимков, а также научных экспериментальных исследованиях;

– личном участии в апробации результатов исследования по выделению и идентификации новообразований на маммографических снимках;

– обработке и интерпретации экспериментальных данных по распознаванию новообразований трех основных типов и моделированию технологии обработки и анализа маммографических снимков, выполненных лично автором;

– подготовке основных публикаций по предварительной обработке, выделению новообразований типа «киста», «фиброаденома» и «рак молочной железы», а также разработке автоматизированной системы обработки и анализа маммограмм;

– личном участии по внедрению результатов диссертационной работы в маммографическом отделении лечебно-профилактического учреждения и в учебном процессе.

Диссертация охватывает основные вопросы поставленной научной задачи по созданию алгоритмов и системы распознавания новообразований на маммографических снимках и соответствует критерию внутреннего единства, что подтверждается наличием последовательного плана исследований, основной идейной линии по созданию новых быстродействующих и точных алгоритмов распознавания новообразований, взаимосвязи выводов по результатам теоретических и экспериментальных исследований.

Практическая ценность работы. Разработанные алгоритмы и программные модули составляют основу АСОиАМС, применение которой позволяет решать следующие задачи практической медицины:

– постановка диагноза с использованием информации, содержащейся во всех снимках, полученных в результате обследования;

– существенное уменьшение субъективных ошибок врача при выявлении новообразований на скрининговых маммограммах;

– сокращения времени, повышения точности, объективности обследования пациентов.

Реализация результатов работы. Разработанная автоматизированная система используется в маммографическом отделении больницы на станции Муром ОАО «РЖД» и в учебном процессе Муромского института (филиала) ФГБОУ ВПО «Владимирский государственный университет имени А. Г. и Н. Г. Столетовых» в виде цикла практических занятий по дисциплине «Методы и системы цифровой обработки изображений» и «Научно-исследовательской работе студентов (НИРС)», что подтверждается соответствующими актами.

На защиту выносятся следующие основные результаты работы.

1. Алгоритм выделения областей молочной железы, основанный на случайных марковских полях, учитывающий индивидуальные особенности строения молочной железы каждой пациентки

2. Алгоритмы и программные модули выделения новообразований типа «Киста молочной железы», «Фиброаденома», «Рак молочной железы» с использованием новых и достаточно простых алгоритмов предварительной обработки, позволяющие выделять место локализации новообразований даже на фоне фиброзно-кистозной болезни и аденоза;

3. Алгоритмы и программные модули распознавания новообразований 3-х типов (киста, фиброаденома, рак молочной железы) на маммографических снимках, использующие дискриминантный анализ Фишера, отличающийся расчетом дискриминантных функций и оценок независимо от количества экспериментов в группах.

4. Тиражируемая система автоматизированной обработки и анализа маммографических снимков, обеспечивающая как автоматическую, так и диалоговую диагностику заболеваний молочной железы;

5. Результаты экспериментов исследований и практического применения системы, которые экспериментально доказали на реальных маммографических снимках применение предложенных научной идеи использования текстурных признаков и подходов выделения области молочной железы, областей новообразований трех типов, а также идентификации этих новообразований в медицинских системах маммографического скрининга.

Апробация работы. Материалы диссертационной работы докладывались и обсуждались на: Международном симпозиуме «Надежность и качество» (г. Пенза, 2012 г., 2013 г.); Международных конференциях «Распознавание 2012», «Распознавание 2013» (г. Курск, 2012 г., 2013 г.); XX Международной научной конференции студентов, аспирантов и молодых учёных «ЛОМОНОСОВ-2013», (г. Москва, 2013 г.), а также на IV, V, VI Всероссийских межвузовских научных конференциях «Наука и образование в развитии промышленной, социальной и экономической сфер регионов России» (г. Муром, 2011-2013 гг.); XVIII Международной конференции по вычислительной механике и современным прикладным программным системам (г. Алушта, 2013 г.); VII всероссийском национальном конгрессе лучевых диагностов и терапевтов «РАДИОЛОГИЯ–2013» (г. Москва, 2013 г.); 11-ой международной конференции «Распознавание образов и анализ изображений: Новые информационные технологии»

(РОАИ-11-2013) (г. Самара, 2013 г.); «Искусственный интеллект. Интеллектуальные системы» (пос. Кацивели, АР Крым, Украина, 2013 г.)

Публикации. Основные результаты диссертации опубликованы в 26 работах, в том числе в 7 статьях в журналах из перечня ВАК, одной монографии, 9 статьях в других научных журналах и 6 тезисах. Получены 2 патента РФ на полезные модели и 1 свидетельство о регистрации программ для ЭВМ.

Структура и объем работы. Диссертация состоит из введения, четырех глав, заключения, списка литературы из 104 наименований, 99 рисунков, 18 таблиц. Общий объем диссертации 149 страниц, в том числе 130 страниц основного текста, 9 страниц литературы, 10 страниц приложений.

Благодарности. Автор выражает искреннюю благодарность врачу-рентгенологу Захаровой Е.А. за подробные разъяснения медицинской литературы, постоянную помощь в разработке алгоритмов и написании статей, главному врачу НУЗ «Отделенческой больницы на станции Муром, ОАО «РЖД» Сафиулову А.Н. за постоянное внимание и поддержку, а также сотрудникам Муромского института (филиала) ВлГУ за содействие в выполнении работы.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы диссертации, сформулированы цель и задачи исследований, научная новизна, практическая ценность, приведены результаты реализации работы.

В первой главе представлен обзор состояния задачи развития и диагностики заболеваний молочной железы: рассмотрены основные способы диагностики, выделены самые доступные и распространённые, определены роль и место маммографии в диагностике заболеваний молочной железы; представлен обзор существующих аппаратных, программных средств, алгоритмов, методов, систем обработки и анализа маммографических снимков. Материалы данной главы позволили сформулировать следующие основные выводы:

1. Анализ состояния развития злокачественных заболеваний молочной железы в мире, в России и по Владимирской области показал что, рак молочной железы лидирует среди онкологических заболеваний у женщин – это общая тенденция.

2. Для выявления различных заболеваний молочных желез используются пленочная и цифровая маммография. В России большинство лечебно-профилактических учреждений оснащены устаревшей аналоговой техникой, которая не позволяет получить маммограммы с высоким качеством и использовать современные информационные технологии, методы цифровой обработки изображений и распознавания. Существующая пленочная технология получения и последующая ручная обработка маммограмм требует много времени, и высока вероятность появления шума и артефактов на снимках.

3. В развитых странах осуществляется переход от аналоговой к цифровой технике, однако в России это процесс выполняется очень медленно. Цифровые маммографические комплексы позволяют получить цифровую маммограмму высокого качества, снизить долю облучения пациенток, хранить снимки с предыдущих обследований, отправка маммограмм по сети Интернет для коллективной постановки диагноза.

4. Молодым женщинам (до 40 лет) с плотными тканями молочных желез требуется ультразвуковая маммография, позволяющая выявить заболевания молочных желез, которые не способна показать рентгеновская маммография.

5. Компьютерная томография как вид обследований МЖ в России вообще не используется в связи с большой дозой облучения женщин.

6. Магнитно-резонансная томография используется для выявления злокачественных областей на снимках с мастопатией, когда размер опухоли очень мал. Данный вид обследований не несет рентгеновского излучения. Несмотря на достоинства данного метода, не все женщины могут позволить себе обследование МЖ с помощью МРТ, в виду дороговизны либо отсутствия МРТ во многих регионах.

7. Сложность обработки аналоговых маммограмм зависит от следующих факторов: наличие повреждений на экранах кассет, правильность промывки пленки, состояние реактивов в проявочных машинах и прочее. Кроме того, наличие той или иной формы мастопатии, из-за которой плотная молочная железа имеет однородную структуру, осложняет автоматическую диагностику молочной железы. Кроме того, возникает сложность при диагностике рака молочной железы, так как на ранней стадии это заболевание болезненных ощущений не вызывает, и имеет малый размер опухоли.

8. Существующие алгоритмы предварительной обработки молочных желез не учитывают наличие мастопатии на снимках, что плохо сказывается на результате обработки снимков.

9. Непосредственное применение существующих алгоритмов распознавания для анализа маммографического снимка не дает требуемой точности классификации новообразований. Это связано с тем, что на снимках не выделены информационные признаки и области новообразований.

10. В некоторых исследовательских центрах и учебных заведениях России ведутся работы по созданию алгоритмов предварительной обработки, выделения новообразований и их классификации для выявления заболеваний на ранних стадиях. Однако до сих пор не создан алгоритм автоматической обработки рентгеновских снимков, однозначно применяющийся к маммограммам.

11. Отсутствуют алгоритмы выделения областей новообразований на снимках, удовлетворяющие требования обеспечения максимального уровня правильного распознавания.

Вторая глава посвящена разработке алгоритмов обработки и анализа маммографических снимков: выделения области молочной железы на снимке, предварительной обработки маммограмм, содержащих три типа новообразований: киста, фиброаденома и рак молочной железы (РМЖ). Поскольку рак молочной железы практически не выявляется на ранних стадиях, ему уделено особое внимание: рассмотрены фоны, которые мешают выявлению: жировая инволюция, фиброзно-кистозная болезнь и аденоз, кроме того, представлены алгоритмы выделения РМЖ даже при наличии этих осложнений. Особое внимание уделено выделению новообразований на маммограммах и их дальнейшей классификации (алгоритмы текстурной сегментации, распознавания на основе методов сравнения с эталонами).

Алгоритм определения областей на маммограмме на основе случайных марковских полей состоит в следующем.

Пусть распределение яркостей на изображении имеет нормальное распределение. Требуется элементы исходной матрицы $F_{m \times n}$ разделить на 3 множества: 1) M – множество, состоящее из векторов, соответствующих пикселям грудной мышцы; 2) R – множество, состоящее из векторов, соответствующих пикселям фона и 3) G – множество векторов, соответствующих пикселям области, на которой изображена молочная железа.

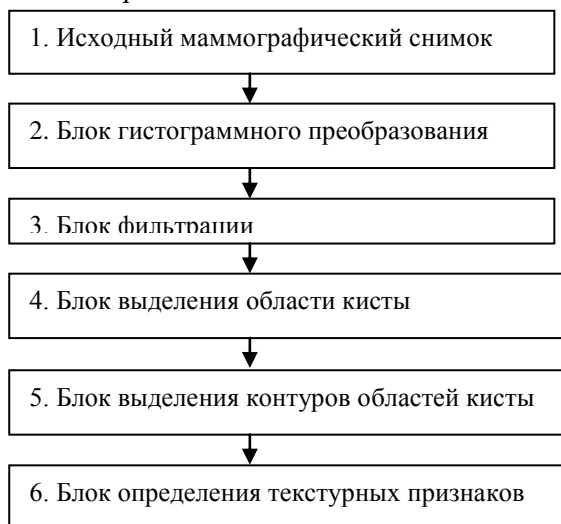
Изображение разбивается на блоки 7×7 пикселей. Формируются компоненты марковского случайного поля из набора блоков $S = \{s_1, s_2, \dots, s_m\}$, где S – множество блоков, m – количество блоков, на которое разбито изображение.

Основными характеристиками поля являются среднеквадратическое отклонение (σ) и математическое ожидание (f_{cp}). Исследования показали, что средняя яркость области молочной железы колеблется от 120 до 150, а СКО – от 12 до 30.

$$\begin{aligned} \text{Если } (0 \leq f_{cp} \leq 100) \ \& \ (0 \leq \sigma \leq 10) & \quad M_l = \text{фон} \\ \text{Если } (120 \leq f_{cp} \leq 150) \ \& \ (30 \leq \sigma \leq 70) & \quad M_l = \text{молочная_железа} \\ \text{Если } (130 \leq f_{cp} \leq 170) \ \& \ (12 \leq \sigma \leq 30) & \quad M_l = \text{грудная_мышца} \end{aligned} \quad (1)$$

Данный алгоритм сегментации, основанный на случайных марковских полях, позволяет безошибочно выделять по 3 области во всех 322 снимках из базы MIAS и 136 снимках реальных пациентов.

Алгоритм выделения области кисты молочной железы представлен на рис. 1.



маммографического снимка, который определяется следующим образом:

$$k = \frac{f_{\max}}{f_{cp}}, \quad (2)$$

где k – коэффициент гистограммных преобразований; f_{\max} , f_{cp} – максимальное и среднее значения яркости исходного снимка, соответственно.

Линейное гистограммное преобразование, учитывающее значение коэффициента гистограммных преобразований имеет следующий вид:

Рисунок 1 - Общая блок-схема алгоритма выделения области кисты

Введем понятие коэффициента гистограммных преобразований

$$g(x, y) = k * t[i] - \frac{t[i] - t[0]}{m * m - t[0]}, \quad (3)$$

где $g(x, y)$ – текущее значение яркости выходного снимка; k – коэффициент гистограммных преобразований исходного снимка;

$$t[i] = \sum_{j=0}^i \frac{h[j]}{m * m},$$

i – значение градации яркости точки выходного снимка ($i=0, 1, 2, \dots, 255$); j – значение градации яркости точки исходного снимка ($j=0, 1, 2, \dots, 255$), $j \leq i$; $h[j]$ – значение элемента гистограммы исходного снимка с яркостью j ; T – область изменения элементов

гистограммы выходного изображения $g(x,y)$; $t[i]$ – значение элемента гистограммы выходного изображения при яркости i ($t[i] \in T$); $t[0] = t[f(x,y)=0]$ – значение элемента гистограммы выходного изображения при яркости точки, равной 0; $m \cdot m$ – размер исходного изображения.

Расчет текстурных признаков изображения производится по матрице Харалика (матрица смежности, или GLCM), вычисленные в 4 направлениях: 0° , 45° , 90° , 135° , что позволяет учитывать различные ориентации текстуры. Далее формируется вектор признаков выделенной области $V[v_1, v_2, \dots, v_{64}]$, а на следующем этапе полученный вектор-признак сравнивается с эталонными векторами $V_э[v_{э,1}, v_{э,2}, \dots, v_{э,64}]$ – признаками из базы знаний, при этом вычисляются минимумы средних сумм квадратов отклонений. Длина вектора признаков не зависит от размера области.

Алгоритм выделения области рака молочной железы (РМЖ) на фоне жировой инволюции

Алгоритм выделения области РМЖ на фоне жировой инволюции представлен на рис. 2.

Алгоритм повышения контраста таких изображений выглядит следующим образом:

$$q(x, y) = K * \frac{v[i] - v[0]}{m \cdot m} + (q_{\max} - q_{\min}) * \frac{I(x, y) - I_{cp}}{I_{cp} - I_{\min}}, \quad (4)$$

где $q(x,y)$ – текущее значение яркости выходного снимка; q_{\max} – максимальное желаемое значение яркости выходного изображения; q_{\min} – минимальное желаемое значение яркости выходного изображения; K – экспериментально подобранный коэффициент ($1 \leq K \leq 2$); $I(x,y)$ – текущее значение яркости исходного снимка; I_{\max} , I_{\min} , I_{cp} – максимальное, минимальное и среднее значения яркости исходной маммограммы, соответственно;

$$v[i] = \sum_{j=0}^i \frac{s[j]}{m \cdot m},$$

i – значение градации яркости точки выходного снимка ($i=0, 1, 2, \dots, 255$); j – значение градации яркости точки исходного снимка ($j=0, 1, 2, \dots, 255$), $j \leq i$; $s[j]$ – значение элемента гистограммы исходного снимка при яркости j ; S – область изменения элементов гистограммы выходного изображения $q(x,y)$; $s[i]$ – значение элемента гистограммы выходного изображения при яркости i ($s[i] \in S$); $s[0] = s[I(x,y)=0]$ – значение элемента гистограммы выходного изображения при яркости точки, равной 0; $m \cdot n$ – размер исходного изображения.

Обычно преобразование гистограммы вызывает некоторое дополнительное зашумление изображения, поэтому изображение $g(x,y)$ подвергается фильтрации с помощью сигма-фильтра маской 3×3 .

Блок выделения контуров на изображении осуществляет локализацию нужной области с четкими границами.

Прежде чем выделить контура на изображении, текстурное изображение снимка подвергается бинаризации.

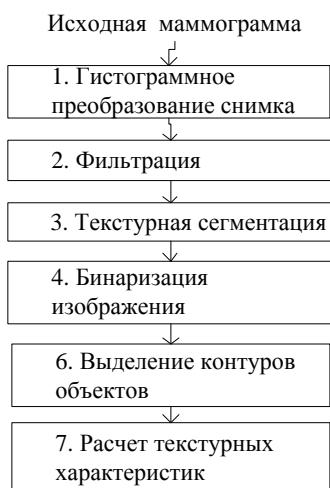


Рисунок 2 – Общая блок-схема алгоритма выделения области РМЖ на маммограмме на фоне жировой инволюции

Алгоритм выделения области рака молочной железы (РМЖ) на фоне аденоза

Алгоритм гистограммных преобразований для изображений молочных желез с преобладанием железистой ткани, можно представить следующим образом:

$$k(x, y) = T * \frac{w[e] - w[0]}{m * n} + (k_{\max} - k_{\min}) * \frac{R(x, y) - R_{cp}}{R_{cp} - R_{\min}}, \quad (6)$$

где $k(x, y)$ – выходное изображение; k_{\max} , k_{\min} – требуемое максимальное и минимальное значения яркости соответственно; T – коэффициент ($T \in [1, 2]$); $R(x, y)$ – текущее значение яркости исходного снимка; R_{\max} , R_{\min} , R_{cp} – максимальное, минимальное и среднее значения исходной маммограммы, соответственно,

$$w[e] = \sum_{p=0}^e h[p]$$

где e – значение яркости точки выходного снимка ($e=0, 1, 2, \dots, 255$); p – значение градации яркости точки исходного снимка ($p=0, 1, 2, \dots, 255$); H – область распределения элементов гистограммы исходного изображения; $h[p]$ – значение элемента гистограммы исходного снимка при яркости p ($h[p] \in H$); W – область изменения элементов гистограммы выходного изображения $k(x, y)$; $w[e]$ – значение элемента гистограммы выходного изображения при яркости e ($w[e] \in W$); $w[0]=w[R(x, y)=0]$ – значение элемента гистограммы обработанного изображения при яркости точки, равной 0; $m \times n$ – размер изображения.

$$b(x, y) = w * e(x, y)^5 - h[0], \quad (5)$$

где $b(x, y)$ – текущее значение яркости выходного снимка; $e(x, y)$ – текущее значение яркости текстурного изображения снимка; $h[0] = h[e(x, y)=0]$ – значение элемента гистограммы текстурного изображения при яркости точки, равной 0; w – экспериментально подобранный коэффициент, от которого зависит отображение различных типов тканей (если $w=0,01$, то на снимке остаются только области грудной мышцы и РМЖ), ($0 < w \leq 10$)).

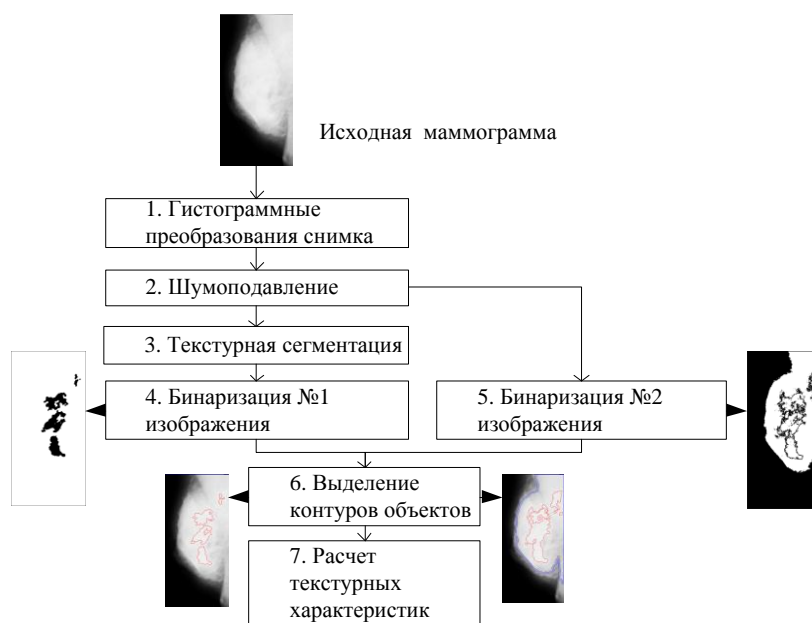


Рисунок 3 – Общая блок-схема алгоритма выделения области РМЖ на маммограмме на фоне аденоза

Экспериментально было установлено, что классификацию новообразований следует выполнять по методу ближайших соседей, который заключается в поиске минимального расстояния от вектора признаков найденной области до эталона, и дискриминантному анализу Фишера: вычисление средних переменных в каждой группе и объединенной дисперсионной матрицы; обращение объединенной дисперсионной матрицы; вычисление коэффициентов дискриминантных функций и оценки функций для каждого наблюдения.

Третья глава посвящена разработке автоматизированной системы обработки и анализа маммографических снимков (АСОиАМС), а также ее отдельных подсистем.

Поставленные в первой главе задачи исследования и полученные во второй главе основные научные результаты позволяют сформулировать набор требований к системе, к маммограммам, к техническому и программному обеспечению.

Требования к техническим и программным средствам АСОиАМС: 1) настройка аналогового маммографа; 2) аналогово-цифровое преобразование исходной маммограммы должно обеспечивать 8 битные полутоновые изображения размером 1024×1024 пикселей. В этом блоке необходимо использовать функцию импорта оцифрованного изображения из различных форматов (bmp, png, tiff и так далее) в формат DICOM-3.0; 3) наличие подсистемы предварительной обработки, которая должна фильтровать артефакты, увеличивать яркость и контраст снимков. Поскольку маммограмма состоит из трех областей: фон, грудная мышца и область молочной железы, обработка всего снимка не целесообразна, поэтому подсистема предварительной обработки должна уметь выделять область молочной железы, учитывая индивидуальные особенности каждой женщины; 4) выделение локального фрагмента (участка), где располагается новообразование; 5) подсистема расчета текстурных характеристик областей новообразований; 6) осуществить распознавание (классификацию) типа новообразования; 7) АСОиАМС должна обеспечивать работу, как в диалоговом, так и в автономном режимах; 8) иметь модуль работы с базой данных о пациентах. Эта база должна содержать основную информацию о пациенте, результаты медицинского обследования, рентгенограммы молочной железы; 9)

формирование медицинских отчетов по результатам обследования каждого пациента и другие статистические данные; 10) система должна быть ориентирована на платформу Microsoft Windows. 11) язык программирования C++.

На рис. 4 представлена общая структурно-функциональная схема разработанной системы, отвечающая выше поставленным требованиям и решаемым задачам.

Основные функции АСОиАМС: 1. регистрация аналоговых маммографических снимков (подсистема регистрации аналогового маммографического снимка); 2. аналогово-цифровое преобразование пленочной маммограммы (блок оцифровки); 3. автоматическое определение области молочной железы на снимке (подсистема обработки изображения); 4. предварительная обработка снимков, содержащих три типа новообразований (подсистема обработки изображения); 5. выделение областей новообразований киста, фиброаденома, рак молочной железы (подсистема выделения области новообразования); 6. расчет текстурных признаков области новообразований для последующего распознавания типа новообразования (подсистема вычисления текстурных признаков областей новообразований); 7. идентификация выделенных областей новообразований (подсистема классификации новообразований); 8. визуальный контроль результатов распознавания с возможностью редактирования и выполнения других этапов обработки на любом из шагов алгоритмов распознавания (подсистема ручного выделения ROI); 9. выполнение стандартных операций с базой данных (БД) пациентов (подсистемы ввода информации и БД); 10. хранение всех изменений, внесенных пользователем в процесс редактирования данных (подсистема БД); 11. формирование и печать отчета (протокола) маммографического исследования (подсистема подготовки отчетов).

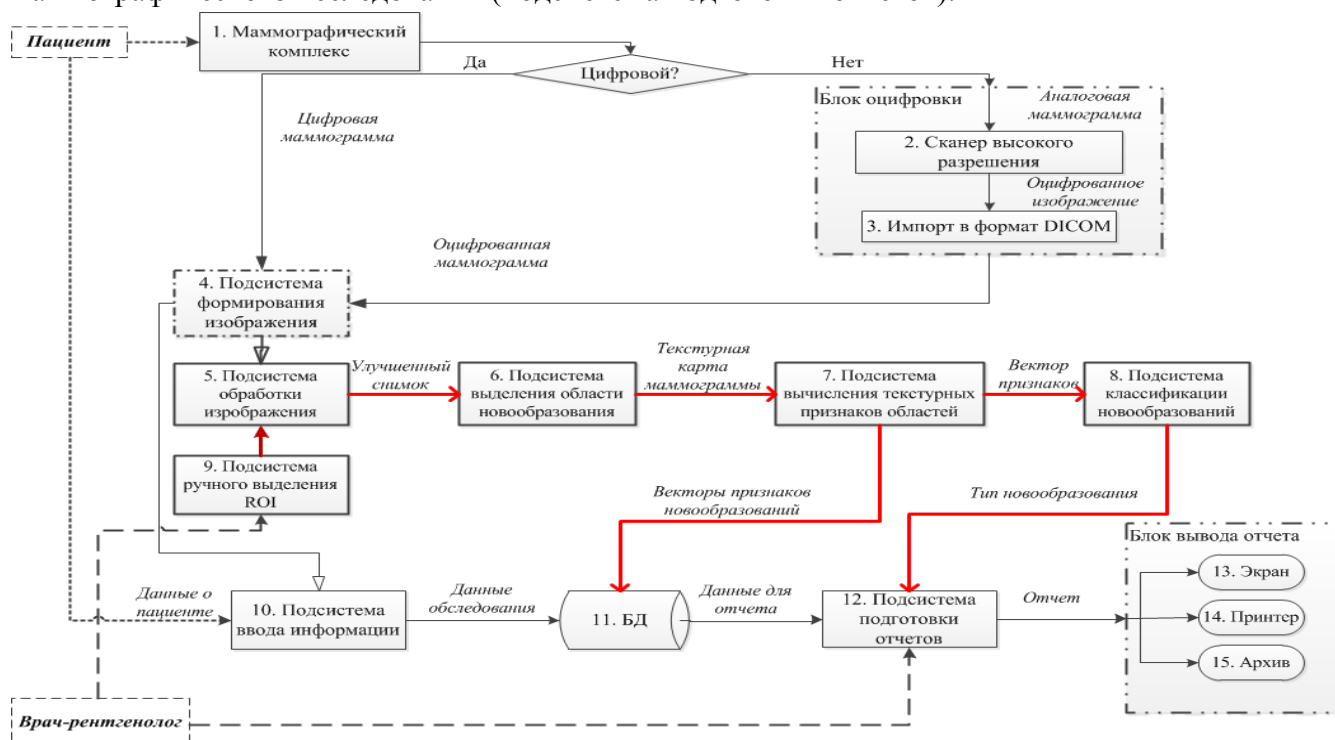


Рисунок 4 – Общая структурно-функциональная схема АСОиАМС

В четвертой главе приведены результаты экспериментальных исследований АСОиАМС на тестовых и реальных снимках.

Исследования проводились на реальных маммографических снимках, взятых из базы маммограмм MIAS (всего обработано 322 снимка).

Были сформированы обучающая и контрольные выборки снимков. Все снимки были обработаны и проанализированы вручную, а затем с помощью АСОиАМС.

Результаты экспериментальных исследований показали, что:

1) Основными эндогенными факторами риска развития РМЖ в г. Муром являются: а) возраст 55 – 65 лет; б) репродуктивная функция женщины: первая беременность в возрасте 26-35 лет, а также первые роды в 25 лет и старше; в) наличие и количество аборт, а также срок прерывания беременности.

2) Сравнительная оценка результатов распознавания на этапе обучения свидетельствует тому, что наиболее точно работает система в сложных случаях, например: а) фиброаденома (аденоз) – 83,3%, у эксперта – 50%; б) фиброаденома (ФКБ) – 83%, у эксперта – 46%; в) РМЖ (аденоз) – 89%, у эксперта - 40%; г) РМЖ (ФКБ) – 85%, у эксперта – 40%.

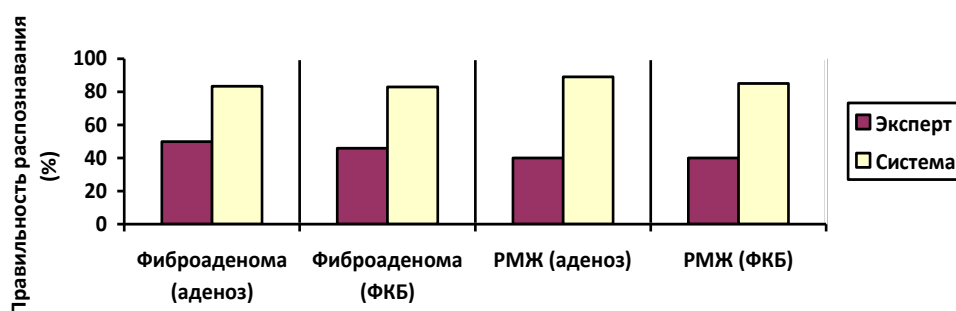


Рисунок 5 – Классификация снимков с «фиброаденомой» и «РМЖ» системой и врачом-экспертом

3) При контрольном распознавании оказалось, что система превышает значение правильного распознавания эксперта по довлеющему большинству видов новообразований. Правильное распознавание наиболее сложных случаев следующее: а) норма (аденоз) у эксперта составляет $\approx 43\%$, а у системы $\approx 86\%$; б) фиброаденома (аденоз) у эксперта составляет 40%, а у системы – 80%; в) РМЖ (аденоз) у эксперта составляет 40%, а у системы – 90.

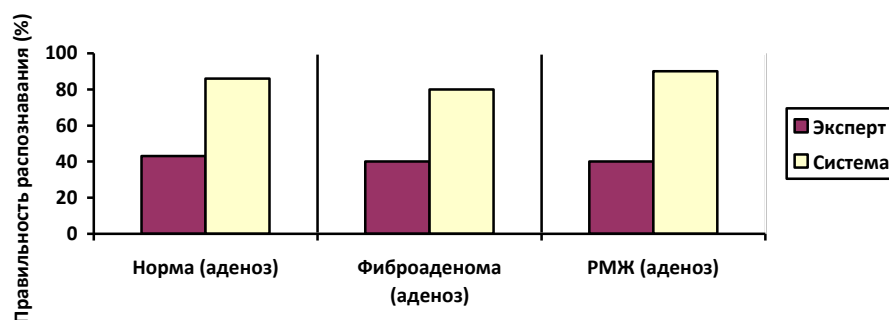


Рисунок 6 – Классификация снимков контрольной выборки с «нормой» системой и врачом-экспертом

Кроме того, система обеспечивает 100% распознавание по 7 видам из 12. Самый низкий уровень распознавания у системы составляет 75%, а у эксперта – 40%.

4) Создана достаточно простая и тиражируемая система автоматизированной обработки и классификации заболеваний молочной железы, позволяющая повысить точность и объективность диагностики новообразований и она является надежным инструментом в руках врача-рентгенолога для ранней диагностики заболеваний молочной железы.

5) Отсутствует возможность сравнения разработанной системы с существующими, поэтому за основу были взяты знания и умения эксперта – врача-рентгенолога.

В заключении сформулированы основные результаты работы.

1. Проведен анализ состояния заболеваемости молочной железы в мире, в России, во Владимирской области и в г. Муром. Показано, что рак молочной железы лидирует среди онкологических заболеваний женщин, и это является общей тенденцией всюду в мире.

2. Установлено, что в России большинство лечебно-профилактических учреждений оснащены устаревшей аналоговой техникой, которая не позволяет получить маммограммы с высоким качеством, в то время как в развитых странах широко используются современные системы, дающие на выходе высококачественные цифровые маммографические снимки, что позволяет для их обработки и анализа использовать методы цифровой обработки изображений и распознавания, способные обеспечить объективность результатов и высокую их точность.

3. Показано, что из современных систем визуализации для диагностики новообразований молочной железы использованы также УЗИ и томографические аппараты, когда возможности цифровых маммографических систем оказываются недостаточными для обнаружения новообразований.

4. Установлено, что в России в настоящее время нет систем автоматизированной обработки и распознавания новообразований молочной железы.

5. На основе экспериментальных исследований из множества методов и алгоритмов выбрано ограниченное количество, обеспечивающие требуемое качество предварительной обработки снимков.

6. Разработаны алгоритмы выделения областей кисты, фиброаденомы и рака на маммограммах, являющиеся основным компонентом разработанной системы автоматизированной обработки и анализа маммографических снимков.

7. Создана система автоматизированной обработки и анализа маммографических снимков, обеспечивающая как автоматическую, так и диалоговую диагностику заболеваний молочной железы. Система построена по модульному типу, легко тиражируется, проста в использовании и имеет невысокую стоимость.

8. Экспериментально на тестовых и реальных маммограммах доказаны широкие возможности разработанной системы автоматизированной обработки и анализа маммограмм. Система диагностирует новообразования значительно точнее, чем врач-рентгенолог, особенно в сложных случаях.

9. Сравнительная оценка результатов распознавания на этапе обучения показала, что наиболее точно работает система в сложных случаях, например:

- фиброаденома (аденоз) – 83,3%, у эксперта – 50%;
- фиброаденома (ФКБ) – 83%, у эксперта – 46%;
- РМЖ (аденоз) – 89%, у эксперта - 40%;
- РМЖ (ФКБ) – 85%, у эксперта – 40%.

10. При распознавании контрольной выборки оказалось, что система превышает значение правильного распознавания эксперта по довлеющему большинству видов новообразований. Отметим, что правильное распознавание наиболее сложных случаев следующее:

- норма (аденоз) у эксперта составляет - 43%, а у системы - 86%;

- фиброаденома (аденоз) у эксперта составляет 40%, а у системы – 80%;
- РМЖ (аденоз) у эксперта составляет 40%, а у системы – 90%.

Кроме того, система обеспечивает 100% распознавание по 7 видам из 12. Самый низкий уровень распознавания у системы составляет 75%, а у эксперта – 40%.

В приложениях приведены: эталоны новообразований типа киста, фиброаденома и рак молочной железы; копии патентов на полезные модели и свидетельство о регистрации программ для ЭВМ; документы, подтверждающие практическое использование результатов диссертационной работы.

Публикации по теме диссертации

Статьи в журналах из перечня ВАК

1 Садыков С.С., Захарова Е.А., Буланова Ю.А. Технология выделения кисты на маммограмме // Вестник РГРТУ. № 1 (выпуск 43), 2013. С.7-11. ISSN 1995-4565

2 Садыков С.С., Буланова Ю.А., Захарова Е.А. Алгоритм выделения области кисты на малоконтрастных маммограммах//Информационные технологии, №8, 2013. С. 53-57. ISSN 1684-6400

3 Садыков С.С., Захарова Е.А., Буланова Ю.А. Использование информационных технологий для выявления области кисты молочной железы на маммограммах //Вестник рентгенологии и радиологии №3, 2013. С. 15-20. ISSN 0042-4676

4 Буланова Ю.А. Использование информационных технологий для локализации области рака молочной железы на маммограммах с преобладанием железистого компонента // Прикаспийский журнал: управление и высокие технологии, №3 (23), 2013. С. 100-111. ISSN 2074-1707

5 Садыков С.С., Захарова Е.А., Буланова Ю.А. Идентификация новообразований молочной железы с использованием скрытых марковских полей //Вестник РГРТУ. №4 (выпуск 46), 2013. С. 44-47. ISSN 1995-4565

6 Садыков С.С., Буланова Ю.А., Захарова Е.А. Компьютерная диагностика новообразований на маммографических снимках // Компьютерная оптика, т.38, №1, 2014. С. 131-138. ISSN 0134-2452

7 Буланова Ю.А. Экспертно-аналитическая система обработки и анализа маммограмм // Прикаспийский журнал: управление и высокие технологии, №1 (25), 2014. С. 92-102. ISSN 2074-1707

Монография

8 Садыков С.С. Автоматизированная обработка и анализ маммографических снимков: монография/ С.С. Садыков, Ю.А. Буланова, Е.А. Захарова; Владим. гос. Ун-т им. А.Г. и Н.Г. Столетовых.- Владимир: Изд-во ВлГУ, 2014. – 208 с. ISBN 978-5-9984-0474-0

Свидетельство о регистрации программы для ЭВМ

9 Свидетельство о государственной регистрации программы для ЭВМ № 2013616037 РФ. Система обработки и анализа маммографических снимков / Буланова Ю.А., Захарова Е.А., Садыков С.С.: зарег. в реестре прог. для ЭВМ 25.06.2013 г.

Патенты на полезные модели

10 Патент на полезную модель № 136612. Устройство диагностики области новообразования типа «Рак молочной железы на маммографических снимках» / Садыков С.С., Буланова Ю.А.; зарег. в Государственном реестре полезных моделей РФ 10.01.2014 г.

11 Патент на полезную модель №137139. Устройство диагностики области новообразования типа «Киста молочной железы» на маммограммах / Садыков С.С., Буланова Ю.А.; зарег. в Государственном реестре полезных моделей РФ 27.01.2014 г.

Статьи в других журналах и научных сборниках

12 Sadykov S.S., Bulanova Y.A. Algorithm of localization of breast cancer in the background of mastopathy // 11th International Conference on Pattern Recognition and Image Analysis: New Information Technologies (PRIA-11-2013), 2013. – vol. 2. – P. 717-721.

13 Садыков С.С., Буланова Ю.А., Захарова Е.А., Каряев В.В. Автоматизация обнаружения новообразований на маммографических снимках. Труды международного симпозиума "Надежность и качество". 2012. Т. 2. С. 376-379.

14 Садыков С.С., Захарова Е.А., Буланова Ю.А. Методика выявления области кисты молочной железы на маммограмме//Алгоритмы, методы и системы обработки данных. 2012. № 22. С. 89-103. ISSN 2220-878X

15 Садыков С.С., Буланова Ю.А., Селюминов И.И. Применение информационных технологий для выявления рака молочной железы на маммограммах // Труды XI Международной научно-технической конференции "Распознавание 2013", Курск, 2013. - С. 423-426.

16 Садыков С.С., Буланова Ю.А., Захарова Е.А.. Диалоговая система анализа маммографических снимков //Алгоритмы, методы и системы обработки данных. №19 2012 г., С. 155-187. ISSN 2220-878X

17 Захарова Е.А., Буланова Ю.А. Обзор медицинской системы КМИС и формирование статистики заболеваемости молочных желез//Алгоритмы, методы и системы обработки данных. 2012. №19. С. 54-61. ISSN 2220-878X

18 Садыков С.С., Буланова Ю.А., Захарова Е.А. Методика выявления рака молочной железы//Алгоритмы, методы и системы обработки данных. 2012. № 19. С. 168-178. ISSN 2220-878X

19 Садыков С.С., Буланова Ю.А., Романов А.Г. Алгоритм текстурной сегментации для выявления областей кисты на маммограммах// Алгоритмы, методы и системы обработки данных. 2013. №19. С. 50-55. ISSN 2220-878X

20 Садыков С.С., Буланова Ю.А., Захарова Е.А., Яшков В.С. Исследование маркерного водораздела для выделения области рака молочной железы //Алгоритмы, методы и системы обработки данных. 2013. №23. С. 56-64. ISSN 2220-878X

21 Садыков С.С., Буланова Ю.А., Яшков В.С. Предварительная обработка маммографических снимков // Труды международного симпозиума "Надежность и качество". 2013. Т. 1, С. 340-343

22 Садыков С.С., Буланова Ю.А., Яшков В.С. Алгоритм локализации области фиброаденом на маммограмме // Алгоритмы, методы и системы обработки данных. 2013. № 3. С. 64-73. ISSN 2220-878X

23 Садыков С.С., Буланова Ю.А., Захарова Е.А. Алгоритм распознавания рака на маммографических снимках // Оптико-электронные приборы и устройства в системах распознавания образов, обработки изображений и символьной информации. Распознавание -2012: сб. матер. X Междунар. науч.-техн. конф. Курск, 2012, С. 306-307

24 Буланова Ю.А. Алгоритм выделения области кисты на маммограмме //Ломоносов-2013: Материалы XX Международной научной конференции студентов, аспирантов и молодых ученых: секция «Вычислительная математика и кибернетика»; 9-12

апреля; Москва, МГУ имени М.В. Ломоносова, факультет ВМК: Сборник тезисов/Сост. Месяц А.И., Шевцова И.Г. – М.: Издательский отдел факультета ВМК МГУ (лицензия ИД 05899 от 24.09.2001)), 2013.С. 95

25 Садыков С.С., Захарова Е.А., Буланова Ю.А. Автоматическая система обработки и анализа маммограмм //Материалы XVIII Международной конференции по вычислительной механике и современным прикладным программным системам (ВМСППС 2013), 22-31 мая 2013 г., Алушта.- М.:Изд-во МАИ, 2013. С. 113

26 Захарова Е.А., Буланова Ю.А. Применение информационных технологий для выявления рака молочной железы на маммограмме //Материалы VII всероссийского национального конгресса лучевых диагностов и терапевтов «РАДИОЛОГИЯ–2013», 29-31 мая, 2013. С. 310-311

Буланова Юлия Анатольевна

СИСТЕМА ДИАГНОСТИКИ ЗАБОЛЕВАНИЙ ПО МАММОГРАФИЧЕСКИМ
СНИМКАМ

Автореферат
диссертации на соискание ученой степени
кандидата технических наук

Подписано в печать ____ . ____ . ____ . Формат бумаги 60×84 1/16.
Бумага офсетная. Печать трафаретная. Усл.печ. л. 1,0.
Тираж 130 экз. Заказ

Рязанский государственный радиотехнический университет.
390005, Рязань, ул. Гагарина, 59/1.

Издательство Муромского института (филиала) федерального
государственного бюджетного образовательного учреждения высшего
профессионального образования «Владимирский государственный
университет имени Александра Григорьевича и Николая Григорьевича
Столетовых»
602264, Владимирская обл., г. Муром, ул. Орловская, 23.