

**На правах рукописи**



**ШАМИН Евгений Анатольевич**

**КОЛЛЕКТИВНЫЙ АЛГОРИТМ НА БАЗЕ  
НЕЧЕТКОЙ ЛОГИКИ ДЛЯ ОБРАБОТКИ  
ЭЛЕКТРОКАРДИОЛОГИЧЕСКОЙ ИНФОРМАЦИИ  
В СЛОЖНЫХ СИГНАЛЬНО-ПОМЕХОВЫХ СИТУАЦИЯХ**

**Специальность 05.13.17 – теоретические основы информатики**

**Автореферат  
диссертации на соискание ученой степени  
кандидата технических наук**

**Пенза – 2012**

Работа выполнена в ФГБОУ ВПО «Пензенская государственная технологическая академия» на кафедре «Информационные технологии и менеджмент в медицинских и биотехнических системах».

**Научный руководитель –**

**Истомина Татьяна Викторовна,**  
доктор технических наук, профессор

**Официальные оппоненты:**

**Михеев Михаил Юрьевич,**  
доктор технических наук, профессор,  
ФГБОУ ВПО «Пензенская государственная  
технологическая академия», заведующий  
кафедрой «Информационные технологии и  
системы»,

**Солодимова Галина Анатольевна,**  
кандидат технических наук,  
ОАО «Научно-исследовательский институт  
электронно-механических приборов»,  
начальник научно-технического центра  
испытаний и метрологического  
обеспечения

**Ведущая организация –**

ФГБОУ ВПО «Астраханский  
государственный университет»,  
г. Астрахань.

Защита диссертации состоится 27 декабря 2012 г., в 13 часов, на заседании диссертационного совета Д 212.337.01 при Пензенской государственной технологической академии по адресу: 440039, г. Пенза, пр. Байдукова / ул. Гагарина, д. 1а / 11, корпус 1, конференц-зал.

С диссертацией можно ознакомиться в библиотеке ФГБОУ ВПО «Пензенская государственная технологическая академия».

Автореферат разослан 26 ноября 2012 г.

Ученый секретарь  
диссертационного совета



**Чулков Валерий Александрович**

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность работы.** В современном мире огромное значение имеет постоянное совершенствование информационного обеспечения различных сфер деятельности человека, в том числе, обеспечения обработки медицинской информации для принятия решений по результатам диагностических исследований. Создание эффективных медицинских информационных систем требует разработки новых информационных моделей процессов воспроизведения биосигналов в условиях помех, алгоритмической формализации процессов диагностики и принятия решений, а также совершенствования информационно-методического обеспечения для решения оценочных задач на стадии верификации полученных результатов.

Одним из наиболее информативных биосигналов является электрокардиосигнал (ЭКС). Многие проблемы, возникающие при автоматической обработке ЭКС, к настоящему времени уже решены. Однако решение задачи получения информации о состоянии сердечно-сосудистой системы в режиме свободной двигательной активности до сих пор сопряжено с преодолением значительных затруднений в процессе выделения и обработки информативной составляющей ввиду влияния интенсивных помех различного типа, энергетический спектр которых перекрывается со спектром ЭКС.

Для выявления путей эффективного ослабления влияния помех в системах обработки электрокардиографической (ЭКГ) информации необходимо проведение исследований электрической активности сердца в рамках системного подхода к описанию сложных прикладных объектов, к которым относятся биотехнические системы. Это предполагает разработку методов и средств обработки ЭКГ-информации на основе системного подхода с учетом взаимных связей элементов сложноорганизованной медицинской системы и биообъекта. Чтобы успешно проводить обработку ЭКГ-информации в этих условиях, необходимо установить причины возникновения помех, изучить структуру, спектральные и статистические характеристики ЭКС, его параметры в сложных сигнально-помеховых ситуациях (ССПС). Среди известных методов обработки ЭКГ-информации системный подход наиболее полно реализуется с помощью методов коллективного распознавания, основанных на согласованной агрегации решений отдельных алгоритмов. Общим недостатком известных технических решений подобного рода является их низкая эффективность в условиях неопределенности, обусловленных естественным разбросом параметров ЭКС в норме и при патологиях, а также влиянием помех.

Точное диагностическое заключение базируется на достоверной информации о функционировании сердечно-сосудистой системы пациента. В настоящее время не существует достаточно эффективных алгоритмов выделения информативных элементов (ИЭ) ЭКС в сложных сигнально-помеховых ситуациях, а также недостаточно проработаны вопросы сравнительного анализа эффективности различных алгоритмов обработки ЭКГ-информации в условиях интенсивных помех. Кроме того, существующие системы для кардиодиагностики не обеспечивают адаптацию к естественным изменениям сигнала в норме и при патологиях, а также к флуктуациям помех. Все это не позволяет повысить достоверность результатов диагностики и, как следствие, улучшить качество медицинских услуг в целом.

В настоящее время в России и за рубежом достаточно активно развивается направление по разработке алгоритмов обработки ЭКГ-информации. Большой вклад в создание теоретических основ медицинских информационных систем внесли российские ученые: Н.А. Корневский, Л.А. Манило, А.П. Немирко, С.В. Селищев, Л.Т. Сушкова, Л.И. Титомир, С.И. Щукин и др. Многие исследователи отмечают эффективность применения для обработки биомедицинской информации относительно нового математического подхода – теории нечеткой логики. Основным отличием такого подхода является возможность учета естественной вариабельности параметров биообъектов в реальном масштабе времени. Однако вопросы применения нечеткой математики для обработки ЭКГ-информации в

настоящее время проработаны недостаточно, поэтому в данном направлении необходимы дальнейшие исследования.

На основании сказанного можно заключить, что повышение эффективности обработки ЭКГ-информации коллективным алгоритмом на базе нечеткой логики в сложных сигнально-помеховых ситуациях является актуальной задачей научного исследования.

**Объектом исследования** является электрокардиографическая информация о состоянии биообъекта.

**Предмет исследования** – модели информативных элементов электрокардиосигнала, информационные модели помех и сигнально-помеховых ситуаций, а также алгоритм обработки, повышающий эффективность работы систем кардиодиагностики.

**Целью** диссертационного исследования является разработка и верификация коллективного алгоритма на базе нечеткой логики для обработки электрокардиологической информации, обеспечивающего повышение точности диагностики состояния биообъекта в сложных сигнально-помеховых ситуациях.

**Задачи исследования.**

1. Разработка системы информационно-лингвистических моделей информативных элементов электрокардиосигнала, помех и сложных сигнальных и сигнально-помеховых ситуаций на основе анализа методов обработки электрокардиографической информации в условиях интенсивных помех, возникающих в процессе ее получения.

2. Разработка основанного на нечеткой логике коллективного алгоритма обработки электрокардиологической информации, позволяющего адаптироваться к естественным изменениям электрокардиосигнала в норме и при патологиях, а также к флуктуациям помех различного вида.

3. Исследование и оптимизация параметров разработанного коллективного алгоритма обработки электрокардиографической информации путем моделирования его работы в пакете *MatLab*.

4. Разработка информационно-методического и программного обеспечения проведения верификации разработанного алгоритма в условиях интенсивных помех, анализ результатов исследования.

**Практическая направленность исследования.** Исследование включает анализ проблем информационного обеспечения своевременной диагностики состояния биообъекта, разработку коллективного алгоритма для анализа электрокардиологической информации, разработку системы информационно-лингвистических моделей сигнала, помех и сложных сигнально-помеховых ситуаций, разработку методики верификации результатов исследований.

Диссертационное исследование направлено на совершенствование программного обеспечения для повышения эффективности работы медицинских информационных систем для кардиологии, обеспечивающих их конкурентоспособность на современном рынке медицинской техники.

**Методы исследования.** Для решения поставленных задач использовались методы цифровой обработки сигналов, теория сигналов, теоретический аппарат нечеткой логики, коллективные методы распознавания, элементы математической статистики, а также методы имитационного и математического моделирования в среде *MatLab*.

**Научная новизна полученных результатов заключается в следующем.**

1. Разработана иерархическая система информационно-лингвистических моделей электрокардиосигнала, различных помех и сложных сигнально-помеховых ситуаций, позволяющих формализовать характерные признаки электрокардиографической информации и создать информационно-методическое обеспечение верификации алгоритмов и систем для кардиодиагностики.

2. Впервые синтезирован коллективный алгоритм на базе нечеткой логики для обработки электрокардиологической информации, позволяющий повысить диагностическую эффективность за счет адаптации к естественным изменениям электрокардиосигнала в норме и при патологиях, а также к флуктуациям помех различного вида.

3. По результатам имитационного моделирования функционирования предложенного коллективного алгоритма на основе нечеткой логики оптимизированы его параметры в следующем сочетании: структура Сугено; система входных признаков ЭКС – мощность, коэффициент подобия, производная и продолжительность кардиоимпульса; схема дефаззификации по взвешенному среднему.

4. Разработана удовлетворяющая международным нормативам многоуровневая методика оценки эффективности методов обработки электрокардиографической информации в условиях нормы и в сложных сигнально-помеховых ситуациях, отвечающая требованиям универсальности, полноты и иерархичности, отличающаяся расширенными функциональными возможностями и позволяющая проводить верификацию результатов. Методика обеспечивает сравнительный анализ эффективности различных вариантов алгоритмов и систем для кардиодиагностики.

#### **Положения, выносимые на защиту.**

1. Система информационно-лингвистических моделей информативных элементов электрокардиосигнала, сопутствующих помех и сложных сигнально-помеховых ситуаций, позволяющая формализовать логику врача и создать универсальное иерархическое описание электрокардиосигнала.

2. Коллективный алгоритм на базе нечеткой логики для обработки электрокардиографической информации на основе структуры Сугено, отличающийся сочетанием оригинальной системы входных признаков электрокардиосигнала и схемы дефаззификации по взвешенному среднему.

3. Многоуровневая методика проведения верификации результатов исследований алгоритмов и систем для обработки электрокардиографической информации, позволяющая осуществлять их поэтапное тестирование в условиях сложных сигнальных и сигнально-помеховых ситуаций.

**Обоснованность и достоверность научных результатов** обеспечивается аргументированным применением математического аппарата, корректностью основных допущений и ограничений, положительными результатами испытаний и внедрениями, подтвержденными соответствующими актами.

#### **Практическая значимость** работы состоит в следующем.

1. Применение созданных адекватных информационных лингвистических моделей информативных элементов электрокардиосигнала, помех и сложных сигнально-помеховых ситуаций для разработки и тестирования медицинских систем обработки электрокардиографической информации повышает точность медицинской диагностики пациентов.

2. Использование новых модулей программного обеспечения повышает эффективность работы медицинских информационных систем для кардиологических исследований в режиме свободной двигательной активности.

**Апробация работы.** Основные положения и результаты диссертационной работы обсуждены: на I и III научных конференциях студентов, аспирантов и молодых ученых “Актуальные проблемы науки и образования” (г. Пенза, 2008, 2010 гг.); на II и III всероссийских научно-технических конференциях “Информационные и управленческие технологии в медицине” (г. Пенза, 2008–2009 г.); на международной конференции “Биомедсистемы–2008” (г. Рязань, 2008 г.); на XI международной научно-технической конференции “Медико-экологические информационные технологии” (г. Курск, 2008 г.); на V, VI и VII межрегиональных научно-практических конференциях студентов и аспирантов “Инновационные технологии в экономике, информатике и медицине” (г. Пенза, 2008–2010 г.), на международном симпозиуме “Надежность и качество” (Пенза, 2010 г.).

**Публикации и личный вклад автора.** По теме диссертации опубликовано 13 печатных работ, из них 3 публикации в журналах, рекомендованных ВАК. Все опубликованные результаты, составляющие содержание диссертации, получены автором самостоятельно.

Научному руководителю принадлежит постановка задач и участие в обобщении результатов исследований. Автор лично получал аналитические соотношения, проводил экспериментальные исследования, интерпретировал и обобщал полученные результаты, формулировал выводы.

**Структура и объем работы.** Диссертационная работа состоит из введения, четырех глав, выводов, списка использованной литературы из 153 наименований. Диссертация изложена на 178 листах машинописного текста и содержит 59 рисунков и 6 таблиц.

## ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

**Во введении** обоснована актуальность темы, сформулированы цели и задачи исследования, изложены основные положения, выносимые на защиту, показана научная новизна и практическая ценность диссертационной работы.

**В первой главе** на основе анализа проблем информационного обеспечения современной кардиодиагностики выявлено, что на результат принятия диагностического решения большое влияние оказывает неинформативная составляющая в виде помех различного вида и существенный индивидуальный разброс параметров в норме и при патологиях. С целью определения наиболее эффективных методов обработки ЭКГ-информации проведен их аналитический обзор, на основе которого выявлены преимущества и недостатки существующих методов и определены пути их совершенствования.

Разнообразие применяемых алгоритмов обработки ЭКГ-информации свидетельствует о том, что задача надежного обнаружения и качественного распознавания информативных элементов при обработке сигналов сердца в настоящее время остается актуальной задачей для научных исследований. Каждый из рассмотренных подходов имеет определенные достоинства в частных условиях, но практически не дает возможности эффективно решать поставленные задачи в сложных сигнально-помеховых ситуациях, поскольку не учитывает индивидуального варьирования параметров в процессе мониторинга.

Таким образом, необходимы разработка и исследование алгоритмов обработки ЭКГ-информации, основанных на применении коллективов решающих правил, позволяющих объединить сильные стороны различных подходов. Кроме того, необходима модернизация коллективного подхода с целью придания ему способности адаптации решающих правил к индивидуальным изменениям информативных параметров ЭКС, что можно реализовать на основе применения нечеткой логики.

Еще одной проблемой является верификация результатов, так как многие из представленных алгоритмов не были протестированы на стандартных ЭКГ-базах или вообще на каких-либо базах данных. Это делает невозможной сравнительную оценку результатов и обуславливает необходимость выработки универсальной методики верификации результатов работы авторских алгоритмов и алгоритмов-аналогов с применением международных аннотированных баз ЭКГ-данных и установления системы критериев оценки этих результатов.

Для этого необходимо создать модель сложных сигнально-помеховых ситуаций, которая позволит описать ЭКС в следующих случаях: зашумленный сигнал; электродные артефакты; дрейф изолинии; сетевые наводки; патологические сигналы; низкоамплитудные кардиоимпульсы (КИ); непредсказуемые скачки амплитуды КИ.

Общим недостатком всех рассмотренных методов является то, что они, будучи основанными на применении жесткой логики, недостаточно эффективно работают в ССПС, поскольку не учитывают индивидуального разброса параметров биосигналов как в норме, так и при патологиях. Поэтому необходима разработка нового алгоритма, обеспечивающего разрешение указанной проблемы.

**Во второй главе** проведен анализ математических основ обработки ЭКГ-информации как инструмента совершенствования алгоритмов. По результатам анализа сделан вывод о том, что для решения поставленных задач наиболее перспективным оказывается применение коллективных методов, основанных на согласованной агрегации решений. Выявлены некоторые недостатки таких методов, связанные с отсутствием учета индивидуального

разброса параметров, характерного для биосигналов как в норме, так и при патологиях вследствие применения жесткой логики.

Анализ разнообразия коллективных классификаторов и используемых ими признаков является одним из путей решения проблемы выбора системы информативных признаков. Именно эта проблема остро стоит при обработке ЭКГ-информации, особенно при трудных для анализа сочетаниях нормальных и патологических ИЭ ЭКС и в сложных сигнально-помеховых ситуациях при сочетании внешних и внутренних шумов разной природы. Однако наличие пересечений между классами образов ЭКС, вызывающих погрешности при анализе, связано с зашумлением граничных областей классов. В ситуациях, когда рассматриваемый образец попадает в области пересечений классов образов и не удастся дать однозначный ответ о его принадлежности к определенному классу, важно получить ответ о степени родства образца с каждым из предполагаемых классов образов, а иногда и определить, что это точно не образец определенного класса образов. Таким образом, можно предположить, что модификация коллективного алгоритма обработки ЭКГ-информации в вариант с нечетким выводом позволит повысить эффективность диагностики сердечно-сосудистых патологий. В связи с этим предлагается задачу обработки ЭКГ-информации рассматривать как нечеткую модификацию задачи коллективного распознавания образов.

Представим схему системы коллективного принятия решения в виде системы нечеткого вывода (СНВ) (рисунок 1), тогда решатели примут вид фаззификаторов  $F_1, F_2, \dots, F_n$  (функция принадлежности – пороговая функция), а решатель уровня  $F$ , или метарешатель, реализующий алгоритм принятия коллективного решения, можно представить в виде оставшейся части СНВ, содержащей базу нечетких правил и дефаззификатор.

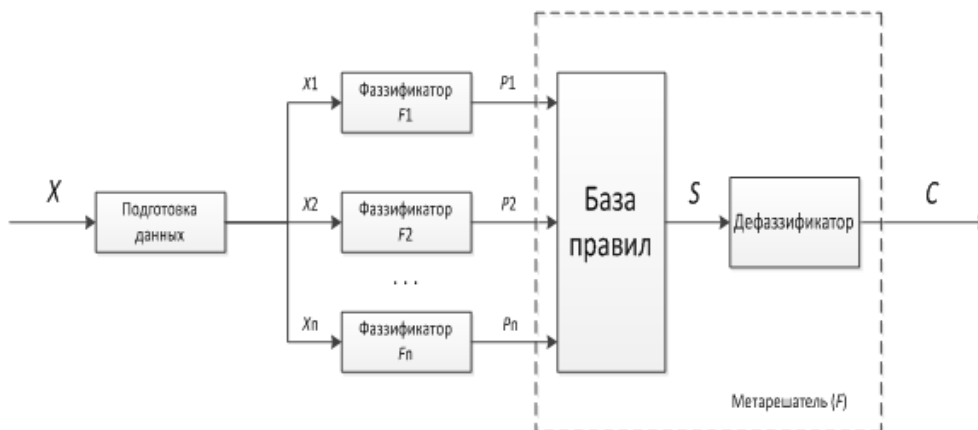


Рисунок 1 – Схема коллективного принятия решения

Здесь  $X$  – ЭКС,  $X_k$  – выбранные признаки,  $P_k$  – нечеткие решатели 1-го уровня, показывающие результаты сравнения признаков с их функциями принадлежности в виде степени принадлежности к определенным термам,  $S$  – результат согласованной агрегации функций принадлежности входных признаков,  $C$  – результат коллективного решения на основе нечеткой логики.

Тогда решение задачи распознавания будет состоять из двух этапов – отнесение текущего импульса к одному из выходных термов согласно базе нечетких правил и дефаззификации:

$$\begin{aligned}
 \Phi_1: & \text{ ЕСЛИ } x_1 \text{ это } A_{11} \dots \text{ И } \dots x_n \text{ это } A_{1n}, \text{ ТО } S \text{ это } B_1 \\
 \dots & \\
 \Phi_i: & \text{ ЕСЛИ } x_1 \text{ это } A_{i1} \dots \text{ И } \dots x_n \text{ это } A_{in}, \text{ ТО } S \text{ это } B_i \\
 \dots & \\
 \Phi_m: & \text{ ЕСЛИ } x_1 \text{ это } A_{m1} \dots \text{ И } \dots x_n \text{ это } A_{mn}, \text{ ТО } S \text{ это } B_m,
 \end{aligned} \tag{1}$$

где  $P_k, k=1..n$  – входные переменные;  $S$  – выходная переменная;  $A_{ik}$  – заданные нечеткие множества с функциями принадлежности.

Результатом нечеткого вывода является четкое значение переменной  $C$  на основе заданных четких значений  $x_k, k=1..n$ :

$$C = \text{ДФ} (S_1, S_2, \dots, S_L), \quad (2)$$

где ДФ – алгоритм дефаззификации.

Дефаззификация нечеткого множества  $\tilde{y} = \int_{\min}^{\max} \frac{\mu_x(x)dx}{x}$  по методу центра тяжести осуществляется по формуле:

$$y = \frac{\int_{\min}^{\max} x \cdot \mu(x) dx}{\int_{\min}^{\max} \mu_x(x) dx}, \quad (3)$$

где  $y$  – результат дефаззификации,  $x$  – переменная, соответствующая выходной лингвистической переменной,  $\mu(x)$  – функция принадлежности выходной переменной после этапа аккумуляции.

Физическим аналогом процедуры является нахождение центра тяжести плоской фигуры между осями координат и функцией принадлежности.

Уравнения вида (1–3) представляют собой коллективное решающее правило на базе нечеткой логики, позволяющее осуществлять принцип согласованной агрегации решений и обеспечивать повышение эффективности обработки ЭКГ-информации в условиях неопределенности. На основе решающего правила создан коллективный алгоритм на базе нечеткой логики (КАНЛ) для обработки электрокардиологической информации, представленный на рисунке 2.

Для оценки эффективности работы предложенного алгоритма необходимо создание иерархической системы моделей ЭКГ-информации, отражающей конфигурацию, временную последовательность и длительность реальных КИ. В результате применения системного подхода к решению проблемы повышения эффективности обработки ЭКГ-информации в ССПС создана система информационных моделей ситуаций на входе средства для кардиодиагностики, отвечающая следующим требованиям: адекватное описание ИЭ ЭКС в норме и при патологиях; обеспечение компромисса между простотой и количеством ИЭ ЭКС; иерархичность; обеспечение поэтапного тестирования алгоритмов и систем с целью выявления узких мест и недостатков функционирования для ускорения поиска неисправностей.

Основой для синтеза и тестирования алгоритмов распознавания ЭКС является выбор необходимого и достаточного количества классов форм КИ с близкими свойствами. Для этого на базе данных о наиболее часто встречающихся формах КИ сформулировано их лингвистическое описание:

- односегментный КИ положительной полярности;
- односегментный КИ отрицательной полярности;
- двухсегментный двухполярный КИ с первым положительным сегментом;
- двухсегментный двухполярный КИ с первым отрицательным сегментом;
- двухсегментный положительный КИ;
- двухсегментный отрицательный КИ;
- трехсегментный КИ с положительным средним сегментом;
- трехсегментный КИ с отрицательным средним сегментом;
- четырехсегментный КИ с положительным средним сегментом;
- четырехсегментный КИ с отрицательным средним сегментом.



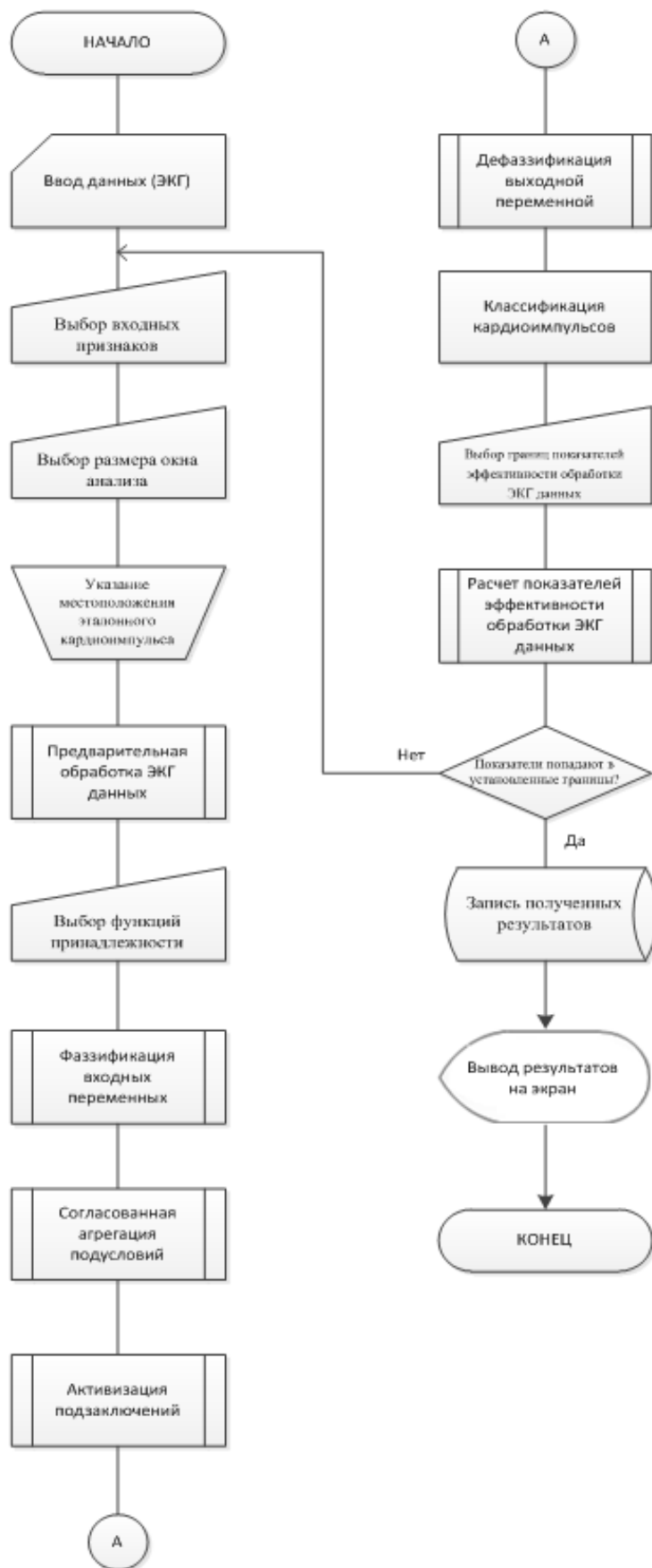


Рисунок 2 – Коллективный алгоритм обработки ЭКГ-информации

Графическая интерпретация лингвистических моделей КИ приведена на рисунке 3. Выбранная система признаков достаточно полно характеризует существующие структуры КИ ЭКС, однако она слишком громоздка, подобно известным системам, составленным по описательному принципу, и не позволяет производить логические (математические) операции. Поэтому предложено преобразовать данную систему лингвистических моделей в информационно-лингвистическое описание КИ на основе применения многоуровневой иерархической системы моделей, базирующейся на определенных выше информативных элементах ЭКС 1-го, 2-го, 3-го, 4-го и 5-го уровней.

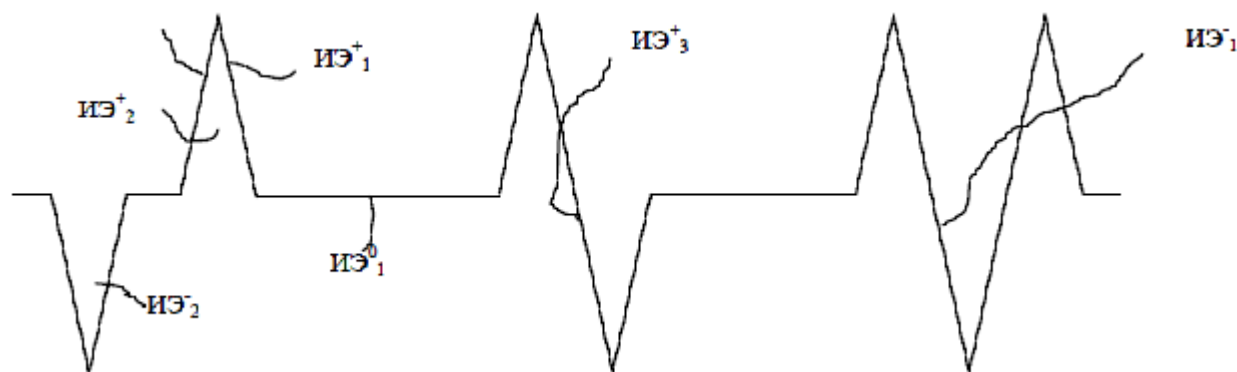


Рисунок 3 – Примеры ИЭ ЭКС разного уровня

Так как в соответствии с предложенным системным подходом все КИ образованы из ИЭ ЭКС 2-го уровня (положительных и отрицательных сегментов), которые в свою очередь состоят из ИЭ ЭКС 1-го уровня (фронтов, спадов и участков изолинии), то можно записать следующие выражения:

$$P(X_i) = ИЭ_1^+ \cap ИЭ_1^-;$$

$$D(X_i) = (ИЭ_1^+ \cap ИЭ_1^-) \cap (ИЭ_1^- \cap ИЭ_1^+);$$

$$G(X_i) = (ИЭ_1^+ \cap ИЭ_1^-) \cap (ИЭ_1^+ \cap ИЭ_1^-);$$

$$V(X_i) = (ИЭ_1^+ \cap ИЭ_1^-) \cap (ИЭ_1^- \cap ИЭ_1^+) \cap (ИЭ_1^+ \cap ИЭ_1^-);$$

$$W(X_i) = (ИЭ_1^+ \cap ИЭ_1^-) \cap (ИЭ_1^- \cap ИЭ_1^+) \cap (ИЭ_1^+ \cap ИЭ_1^-) \cap (ИЭ_1^- \cap ИЭ_1^+);$$

$$P^*(X_i) = ИЭ_1^- \cap ИЭ_1^+;$$

$$D^*(X_i) = (ИЭ_1^- \cap ИЭ_1^+) \cap (ИЭ_1^+ \cap ИЭ_1^-);$$

$$G^*(X_i) = (ИЭ_1^- \cap ИЭ_1^+) \cap (ИЭ_1^- \cap ИЭ_1^+);$$

$$V^*(X_i) = (ИЭ_1^- \cap ИЭ_1^+) \cap (ИЭ_1^+ \cap ИЭ_1^-) \cap (ИЭ_1^- \cap ИЭ_1^+);$$

$$W^*(X_i) = (ИЭ_1^- \cap ИЭ_1^+) \cap (ИЭ_1^+ \cap ИЭ_1^-) \cap (ИЭ_1^- \cap ИЭ_1^+) \cap (ИЭ_1^+ \cap ИЭ_1^-).$$

Эти выражения будут справедливы лишь при условии, что ИЭ КИ находятся в конкретном временном интервале определения ЭКС.

Тогда участок ЭКС любой необходимой для исследований длины можно записать с помощью системы иерархических моделей ЭКС, соединенных между собой на основе конкатенации ИЭ 1-го и 0-го порядков, например:

$$ЭКС = \langle ИЭ_1^0, ИЭ_1^+, ИЭ_1^-, ИЭ_1^+, ИЭ_1^-, ИЭ_1^0, ИЭ_1^+, ИЭ_1^-, ИЭ_1^0, \dots \rangle.$$

Недостатком такого описания является отсутствие учета временного масштаба. Избежать этого поможет введение принадлежности интервалам:

$$ЭКС = \langle ИЭ_1^0 \in [t_i, t_{i+j}], ИЭ_1^+ \in [t_{i+j}, t_{j+k}], ИЭ_1^- \in [t_{j+k}, t_{k+l}], ИЭ_1^+ \in [t_{k+l}, t_{m+n}], ИЭ_1^- \in [t_{m+n}, t_{o+p}], ИЭ_1^0 \in [t_{p+q}, t_{q+r}], ИЭ_1^+ \in [t_{q+r}, t_{s+t}], ИЭ_1^- \in [t_{s+t}, t_{u+v}], ИЭ_1^0 \in [t_{u+v}, t_{w+x}], \dots \rangle.$$

При этом порядковый номер ИЭ ЭКС определяется его местом в кортеже, индексы переменной  $t$  характеризуют целочисленное количество дискретных отсчетов, приходящееся на каждый из временных интервалов, которым принадлежат информативные элементы ЭКС 1-го уровня градации. При построении системы моделей ЭКС в сложных ситуациях “норма/патология” (ССНП) без учета помех за основу целесообразно взять следующее описание:

- НСП – высокоамплитудный КИ;
- ЛСП – низкоамплитудный КИ;
- НТ – высокоамплитудный Т-зубец;
- НТ – любой Т-зубец, кроме высокоамплитудного;
- СП – узкий КИ;
- ДСП – широкий КИ;
- ВЛ – участок изолинии любой длительности;
- НРР – нормальный интервал между КИ;
- СРР – укороченный интервал между КИ;
- МРР – удлиненный интервал между КИ;
- ВЕС – одиночная желудочковая экстрасистола.

Данные примитивы представляют средний уровень иерархии, так как сами могут быть описаны на основе ИЭ ЭКС нижних уровней, например, как было показано выше.

По результатам анализа атласов ЭКГ и данных по тестированию алгоритмов обработки ЭКГ-информации автором предложен следующий проблемный перечень (лингвистическое описание ССНП):

- ССНП<sub>1</sub> = “низкоамплитудный КИ на фоне высокоамплитудных”,
- ССНП<sub>2</sub> = “высокоамплитудный КИ на фоне низкоамплитудных”,
- ССНП<sub>3</sub> = “высокоамплитудный Т зубец на фоне низкоамплитудных КИ”,
- ССНП<sub>4</sub> = “экстрасистолический КИ, подобный артефакту движения”,
- ССНП<sub>5</sub> = “узкий КИ на фоне широких”,
- ССНП<sub>6</sub> = “широкий КИ на фоне узких”,
- ССНП<sub>7</sub> = “широкий КИ на фоне учащенного ритма”,
- ССНП<sub>8</sub> = “вставочные (интерполированные) экстрасистолы на фоне учащенного ритма”,
- ССНП<sub>9</sub> = “экстрасистолы на фоне сердечных блокад”.

Эти ситуации целесообразно представить с помощью предложенного выше формализованного описания в виде информационно-лингвистических моделей, например:

ССНП<sub>1</sub> = <BLP, НСП, НТ, ВЛ, Р, НСП, НТ, ВЛ, Р, ЛСП, НТ, ВЛ, Р, НСП, НТ, ВЛ> – низкоамплитудный КИ на фоне высокоамплитудных.

Для учета временной координаты необходимо определять каждый элемент кортежа на конкретном интервале времени, например, в виде:

$$\text{ССНП}_1 = \langle \text{BLP} \in [t_i, t_{i+j}], \text{НСП} \in [t_{i+j}, t_{j+k}], \text{НТ} \in [t_{j+k}, t_{k+l}], \text{ВЛ} \in [t_{i+m}, t_{m+n}], \text{Р} \in [t_{n+o}, t_{o+p}], \\ \text{НСП} \in [t_{p+q}, t_{q+r}], \text{НТ} \in [t_{r+s}, t_{s+t}], \text{ВЛ} \in [t_{r+u}, t_{u+v}], \text{Р} \in [t_{v+w}, t_{w+x}], \text{ЛСП} \in [t_{x+y}, t_{y+z}], \dots \rangle.$$

При этом амплитудные значения описываются моделями ИЭ низкого уровня, а участки, характеризующиеся цикличностью, могут повторяться многократно в зависимости от выбора длительности времени тестирования.

Наиболее сложным уровнем формализованного описания ЭКС является моделирование помех различного типа, возникающих в процессе получения и обработки ЭКГ-информации. Обычно применяют стандартные модели синусоидального сигнала при описании сетевой помехи и дрейфа изолинии и аддитивно накладывают их на сигнал. У такого подхода есть ряд недостатков, так как результирующие сигналы имеют существенные отличия от реальных. Автором предлагается формировать модели ССПС, собирая тестовую последовательность из реальных участков сигнала и помех различного вида, при этом

основываться на заранее определенных наиболее сложных для работы алгоритмов и систем ситуациях.

Лингвистическое описание сложных сигнально-помеховых ситуаций дается следующим образом.

Внешние помехи:

MA – артефакт движения одиночный;

DI – дрейф изолинии;

SP – сетевая помеха;

DA – дыхательная аритмия.

Внутренние помехи:

EKG – помеха от электрической активности желудка;

EEG – помеха от электрической активности мозга;

EMG – помеха от электрической активности мышц;

ED – кривая дыхания;

STI – стимулирующий импульсы.

Система информационно-лингвистических моделей ССПС:

ССПС<sub>1</sub> = “низкоамплитудный КИ на фоне артефакта движения”,

ССПС<sub>2</sub> = “низкоамплитудный КИ на фоне дрейфа изолинии”,

ССПС<sub>3</sub> = “низкоамплитудный КИ на фоне дыхательной аритмии”,

ССПС<sub>4</sub> = “экстрасистола на фоне помехи от активности мышц”,

ССПС<sub>5</sub> = “узкий КИ на фоне стимулирующих импульсов”,

ССПС<sub>6</sub> = “широкий КИ на фоне дрейфа изолинии”,

ССПС<sub>7</sub> = “широкий КИ на фоне помехи от активности желудка”,

ССПС<sub>8</sub> = “низкоамплитудный узкий КИ на фоне помехи от электрической активности мозга”.

Пример информационно-лингвистической модели ССПС:

ССПС<sub>1</sub> = <BLP, LCP, NT, BL, P, HCP, NT, BL, P, LCP, NT, BL, P, MA, NT, BL> – низкоамплитудный КИ на одиночного артефакта движения.

Для учета временной координаты в системе моделей ССПС также необходимо определять каждый элемент кортежа на конкретном интервале времени, например в виде:

$$\text{ССПС}_1 = \langle \text{BLP} \in [t_i, t_{i+j}], \text{LCP} \in [t_{i+j}, t_{j+k}], \text{NT} \in [t_{j+k}, t_{k+l}], \text{BL} \in [t_{i+m}, t_{m+n}], \text{P} \in [t_{n+o}, t_{o+p}], \\ \text{HCP} \in [t_{p+q}, t_{q+r}], \text{NT} \in [t_{r+s}, t_{s+t}], \text{BL} \in [t_{t+u}, t_{u+v}], \text{P} \in [t_{v+w}, t_{w+x}], \text{MA} \in [t_{x+y}, t_{y+z}], \dots \rangle.$$

Модели участков помех, включаемые в ССПС, формируются с помощью включения в кортежи массивов дискретных отсчетов, заданных на конкретных интервалах времени и максимально соответствующих реальным сигнально-помеховым ситуациям. Эти неинформативные участки затем объединяются с информативными с помощью операции конкатенации. Для формирования тестовых моделей большой длительности любого уровня можно получать с помощью многократного повторения базовых моделей.

В процессе разработки универсальной многоуровневой системы входных моделей ЭКС были получены иерархические формализованные описания трех уровней: система моделей формы КИ; система моделей сложных ситуаций, связанных с различными вариантами нормы и патологии, учитывающая последовательность следования КИ во времени и нарушения их периодичности; и система моделей, описывающая сложные ситуации, связанные с возникновением помех различного вида, как внутренних, так и внешних. Каждый новый уровень системы информационно-лингвистических моделей включает формализованное описание, созданное для более низкого уровня, таким образом, в ней учитываются принципы полноты и иерархичности. Предложенная система моделей позволяет

производить поэтапное тестирование алгоритмов обработки ЭКГ-информации и локализацию при поиске неисправностей, а также обеспечивает всестороннюю проверку эффективности их работы в процессе верификации кардиомониторных систем.

В третьей главе проведена оптимизация разработанного КАНЛ ЭКГ-информации с помощью имитационного моделирования в пакете *MatLab*. На первом этапе для нечеткой обработки (НО) использовался алгоритм Мамдани. Анализ проводился по трем признакам. На рисунке 4 представлена структурная схема СНО с тремя анализируемыми параметрами – “Мощность”, “Производная” и “Схожесть”.

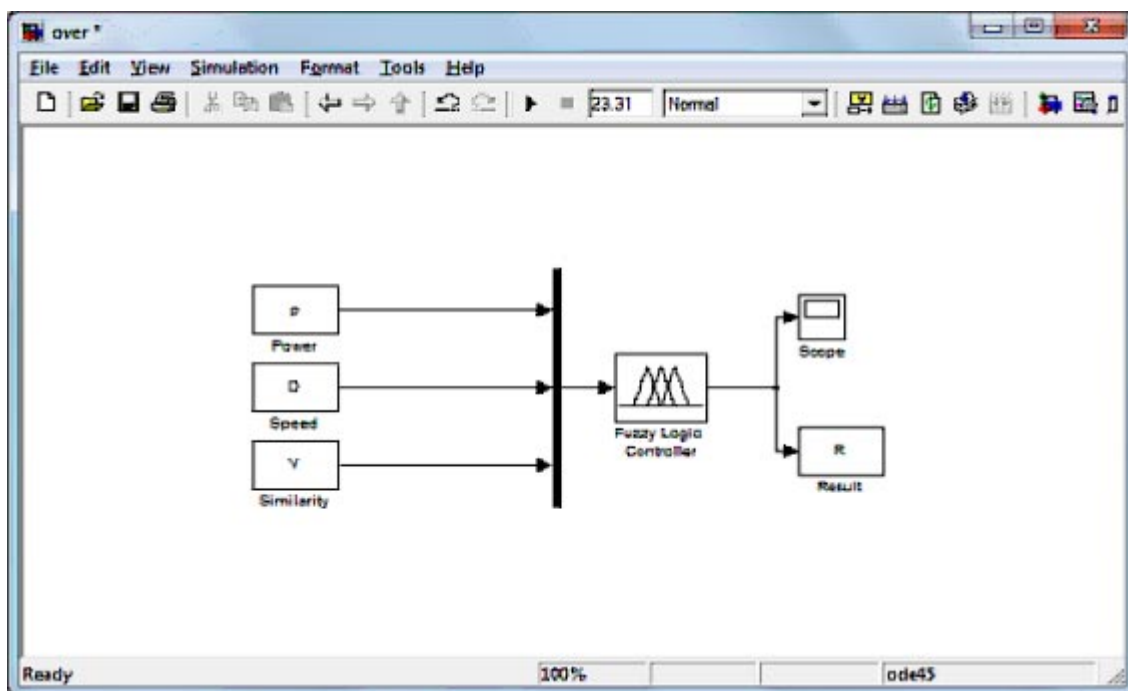


Рисунок 4 – Структурная схема НО с тремя анализируемыми параметрами

На рисунке 5 представлена схема нечеткого контроллера на основе алгоритма Мамдани в среде *Fuzzy Logic Toolbox MatLab*.

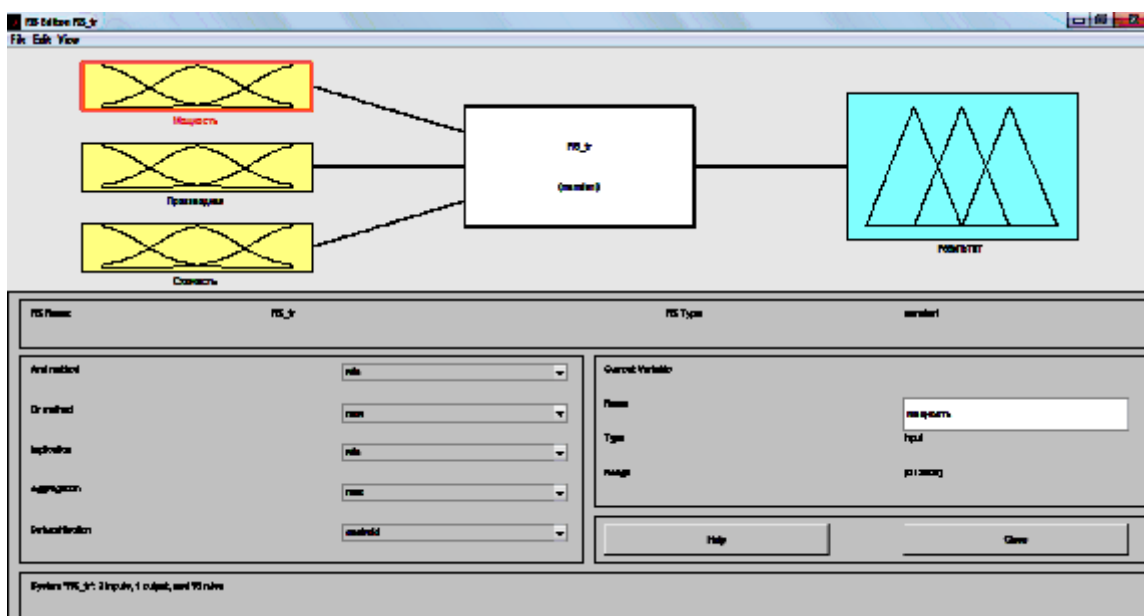


Рисунок 5 – Схема нечеткого контроллера на основе алгоритма Мамдани

Затем из входного ЭКС синтезируются выбранные признаки. На рисунке 6 представлены графики ЭКС и трех выбранных признаков.

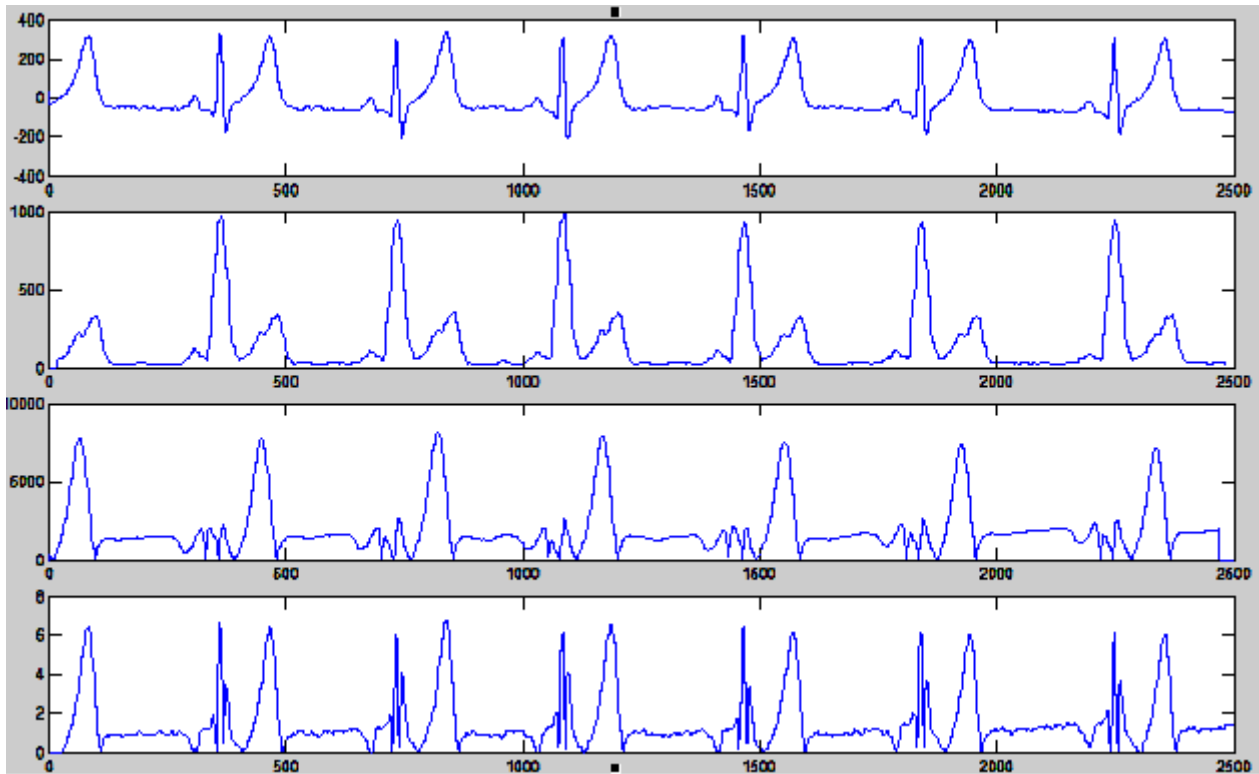


Рисунок 6 – Результаты этапа предварительной обработки ЭКС.  
Сверху вниз: 1) исходный сигнал, 2) “Производная”, 3) “Мощность”, 4) “Схожесть”

Каждый признак (“лингвистическая переменная”) описывается несколькими термами со своими функциями принадлежности. На рисунке 7 показаны функции принадлежности лингвистической переменной “Мощность”.

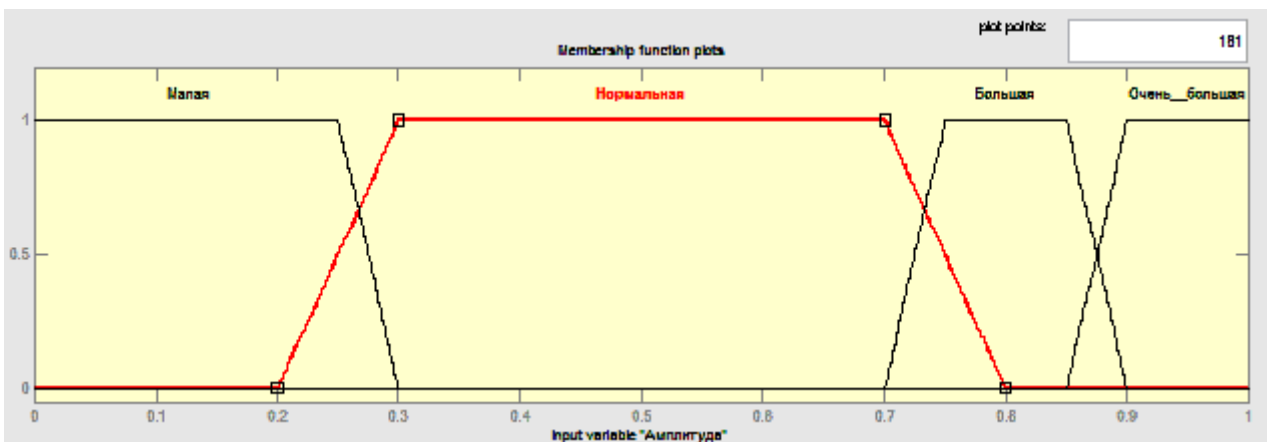


Рисунок 7 – Функции принадлежности признака “Мощность”

В результате анализа входных данных схемы НО получаются значения выходной лингвистической переменной “Результат”. Анализируя ее, можно классифицировать тот или иной отсчет как относящийся к классу “Да” или “Возможно”. Результат классификации представлен на рисунке 8.

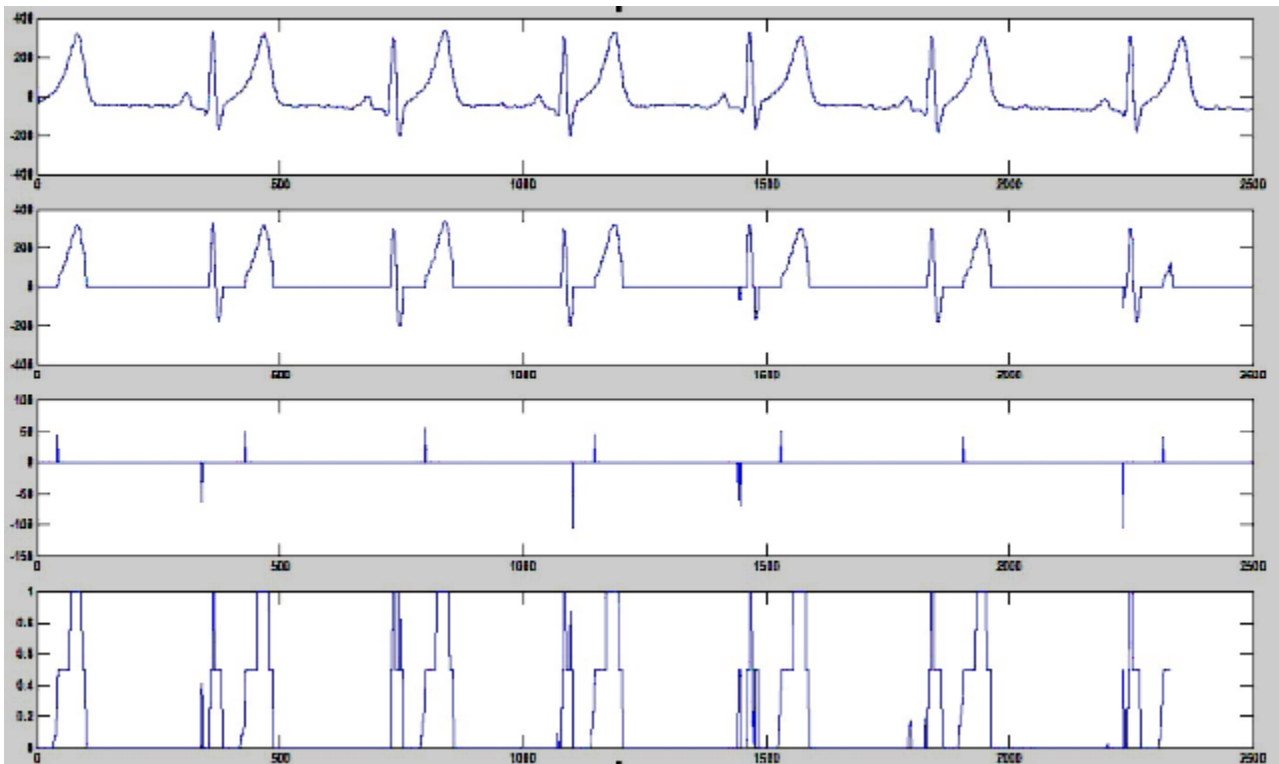


Рисунок 8 – Результаты этапа классификации КИ. Сверху вниз:  
 1) исходный сигнал, 2) область, классифицированная как “Да”, 3) область,  
 классифицированная как “Возможно”, 4) результат на выходе

В связи с особенностью алгоритма Мамдани, и, как следствие, с высокой сложностью схемы НО на этапе анализа результатов, было принято решение применить в предложенном алгоритме НО на основе алгоритма Сугено. Результат моделирования на основе алгоритма Сугено показал, что данный алгоритм наиболее удачно реализует поставленную задачу, и позволяет избежать постоянной подстройки порога на этапе анализа данных на выходе.

С целью повышения качества обнаружения в ССНП и ССПС было принято решение дополнительно использовать четвертый признак – “Длительность импульса”. Структура системы представлена на рисунке 9.

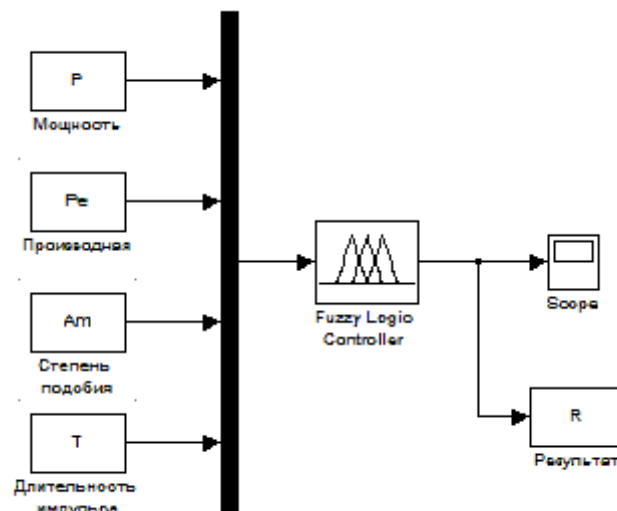


Рисунок 9 – Структурная схема алгоритма НО (схема Сугено, 4 признака)

В качестве анализируемого использовался реальный оцифрованный сигнал с КИ (Т-зубцами) патологической формы, наличие которых существенно осложняет анализ. Входной сигнал представлен на рисунке 10.

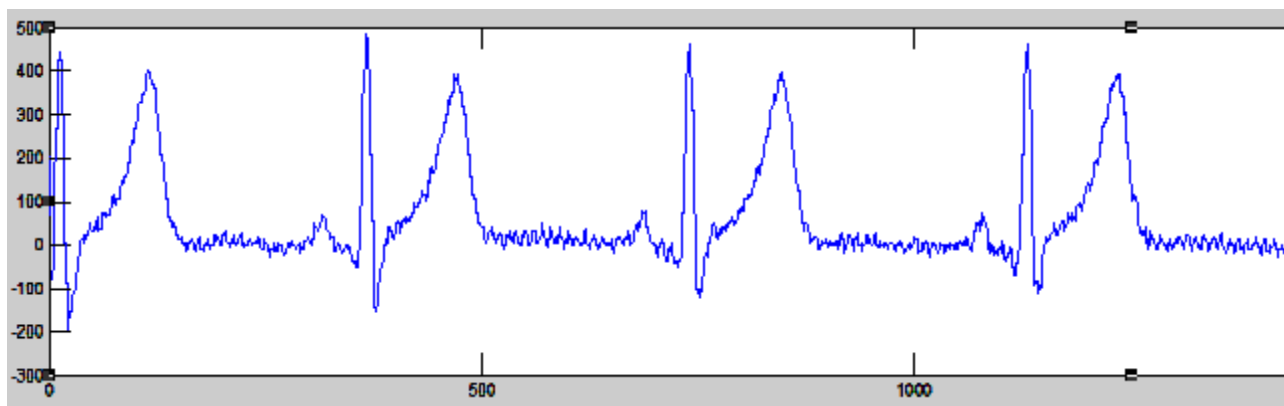


Рисунок 10 – Входной сигнал с аномальным Т-зубцом

За счет четвертого признака “Длительность импульса” стало возможным отделить R-зубцы от Т-зубцов (рисунок 11).

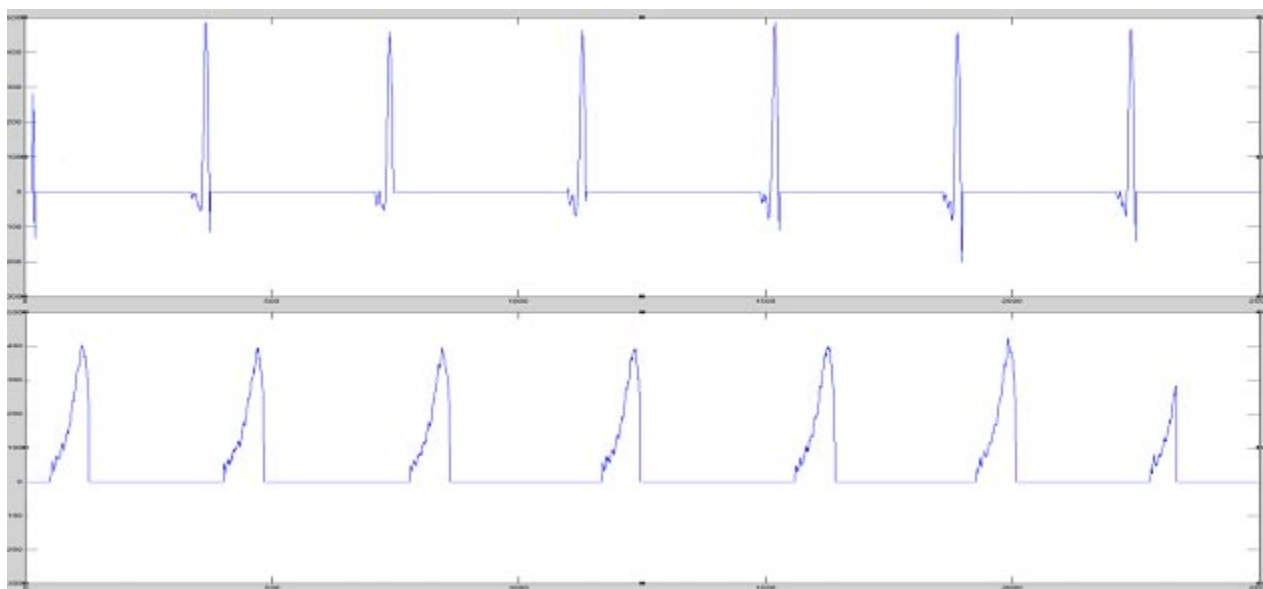


Рисунок 11 – Результаты этапа классификации КИ. Сверху вниз:

1) по классу “Да”, 2) по классу “Возможно”

Таким образом, исследования по подбору и оптимизации параметров предложенного КАНЛ ЭКГ-информации показали целесообразность применения следующих параметров: алгоритм Сугено; система входных признаков: мощность, коэффициент подобия, производная и продолжительность КИ; метод дефаззификации: взвешенное среднее.

**В четвертой главе** разработана исследовательская методика для сравнительной оценки эффективности алгоритмов обработки ЭКГ-информации. Методика включает несколько этапов и имеет многоуровневую структуру, включающую предложенную систему информационно-лингвистических моделей ЭКС в сочетании с реальными сигналами с платы ввода-вывода и из международных аннотированных баз данных, она позволяет производить тестирование и верификацию результатов работы алгоритмов и систем для кардиодиагностики.

Для верификации алгоритмов и программного обеспечения для кардиодиагностики доступны несколько баз данных стандартных ЭКС. Эти базы содержат аннотацию и обеспечивают воспроизводимые и сопоставимые результаты. Кроме того, эти базы данных



содержат самые разнообразные ЭКС, в том числе сигналы, которые редко наблюдаемы, но представляют высокую клиническую значимость. С целью выбора сертифицированной базы для тестирования проведен поиск и изучение аннотированных международных баз данных ЭКС, доступных в ресурсах Internet, по результатам изучения принято решение в процессе верификации разработанного алгоритма использовать базу MIT-BIH, которая входит в состав средств, рекомендованных для тестирования электрокардиографических систем по стандарту ANSI.

Разработана многоуровневая база тестовых сигналов, дающая возможность комбинировать реальные ЭКС и помехи различного вида с репрезентативным набором всех классов формы КИ и патологических аритмий.

Для проведения сравнительной оценки эффективности функционирования алгоритмов обработки ЭКГ-информации разработано программное приложение в пакете *MatLab*, позволяющее моделировать известные и разработанные алгоритмы, реализовать предложенную систему моделей и многоуровневую методику, а также произвести расчет выбранных показателей эффективности алгоритмов и проанализировать результаты. Пример экранной формы разработанного программного приложения представлен на рисунке 12, отдельные модули программ приведены в приложении к диссертации.

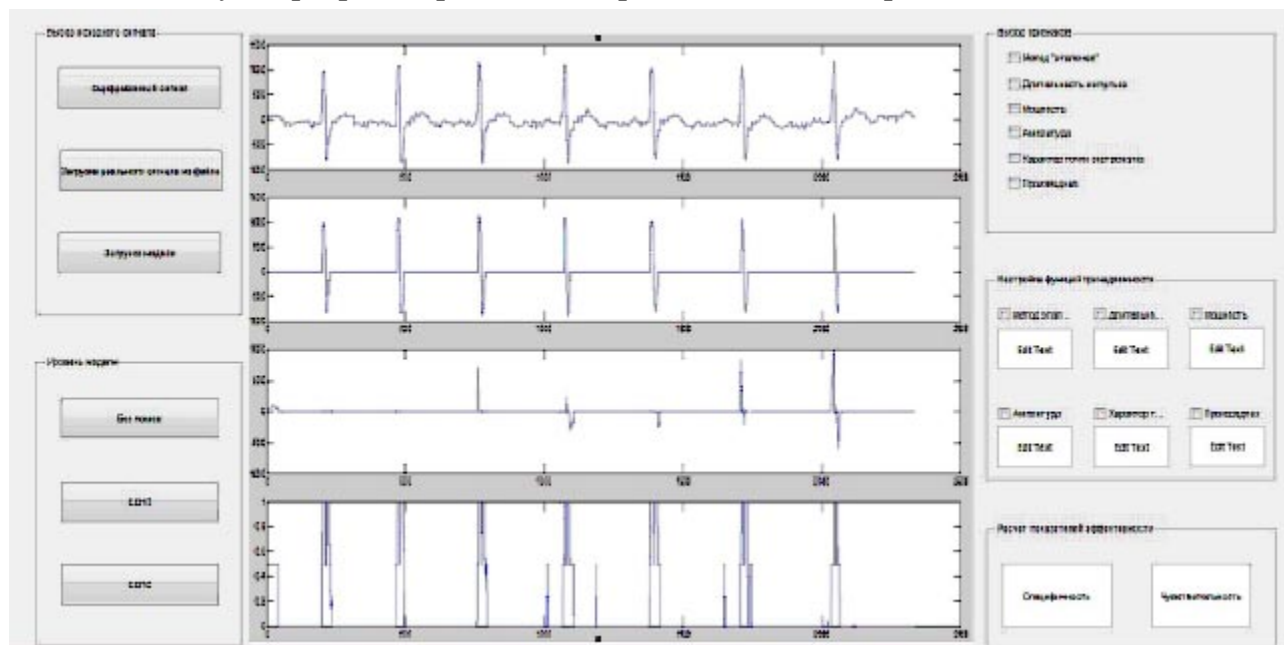


Рисунок 12 – Пользовательский интерфейс программного приложения

Результаты практической оценки эффективности лучшего из отдельных алгоритмов Пана-Томпкинса (ПТ), коллективного алгоритма (КА) и разработанного КАНЛ с использованием показателя диагностической эффективности (ДЭ) в качестве критерия эффективности приведены в таблице, представляющей фрагмент результатов исследований для 35 записей, содержащих не менее 2 тысяч КИ ЭКС каждая, с включением помех различного типа и интенсивности.

Расчет показателей эффективности выполнен по общепринятой методике: TP (*True Positives*) – верно классифицированные положительные примеры (истинно положительные случаи); TN (*True Negatives*) – верно классифицированные отрицательные примеры (истинно отрицательные случаи); FN (*False Negatives*) – положительные примеры, классифицированные как отрицательные (ошибка I рода или ложный пропуск); FP (*False Positives*) – отрицательные примеры, классифицированные как положительные (ошибка II рода или ложное обнаружение). Для обобщения результатов рассчитано отношение истинных результатов обнаружения к общему числу КИ ЭКС или ДЭ:

$$\text{ДЭ} = ((\text{TP} + \text{TN}) / \text{N}) \times 100\%,$$

где  $\text{N} = \text{TP} + \text{TN} + \text{FP} + \text{FN}$ .

Таблица 1 – Результаты сравнения эффективности алгоритмов

№ записи ЭКС	Кол-во КИ	Алгоритм ПТ			Алгоритм КА			Алгоритм КАНЛ		
		TP	FP	FN	TP	FP	FN	TP	FP	FN
200	2605	1823	154	782	2187	107	418	1201	98	404
201	1858	1304	98	554	1564	68	294	1564	66	294
202	2038	1425	112	613	1710	78	328	1752	76	286
203	2882	2015	160	867	2418	112	464	2482	99	400
205	2491	1742	56	1245	2090	39	401	2401	43	90
207	2215	1549	86	666	1938	60	277	2002	60	213
208	2823	1830	95	993	2196	66	627	2555	40	288
209	2860	2003	76	857	2403	53	457	2666	55	194
210	2502	1751	59	751	2101	41	401	2221	39	281
...	...	...	...	...	...	...	...	...	..	...
214	2134	1493	56	641	1927	39	207	2011	30	123
215	3172	2220	123	952	2664	86	508	2673	77	499
ДЭ		0,72			0,84			0,88		

Как следует из таблицы 1, значение ДЭ алгоритма КАНЛ существенно превышает этот показатель для алгоритмов ПТ и КА, демонстрируя положительный эффект применения согласованной агрегации на основе нечеткой логики. Результаты обработки многоуровневой представительной выборки ЭКС также показывают, что алгоритм КАНЛ дает наилучшие результаты, особенно в условиях интенсивных помех, обеспечивая увеличение показателя ДЭ в среднем на 5 % по сравнению с отдельными алгоритмами и алгоритмом КА. Применение алгоритма позволяет повысить надежность работы средств обнаружения КИ ЭКС в составе медицинских информационных систем для кардиодиагностики.

## ОСНОВНЫЕ РЕЗУЛЬТАТЫ И ВЫВОДЫ

1. Проанализированы методы помехоустойчивой обработки электрокардиологической информации и показаны перспективы применения коллективов решающих правил, основанных на согласованной агрегации решений частных обнаружителей и позволяющих объединить сильные стороны различных алгоритмов в сложных сигнально-помеховых ситуациях.

2. На базе применения теории нечеткой логики разработаны решающее правило и алгоритм, основанный на модернизации коллективного подхода, обеспечивающий повышение эффективности обработки электрокардиографической информации в сложных сигнально-помеховых ситуациях, обеспечивающие адаптацию к индивидуальным изменениям информативных параметров электрокардиосигнала.

3. На основе аналитического обзора, эмпирического материала и применения системного подхода разработана многоуровневая система информационно-лингвистических моделей, описывающая электрокардиосигнал в норме и в сложных сигнальных и сигнально-помеховых ситуациях, которая позволяет производить поэтапное тестирование и верификацию алгоритмов и систем для кардиодиагностики. Определены системы критических сигнальных и сигнально-помеховых ситуаций для областей компетенции отдельных алгоритмов обработки ЭКГ-информации, в которых существенно

снижается эффективность принятия решения и повышается вероятность ошибок обнаружения и распознавания кардиосигналов.

4. Результаты исследований по подбору и оптимизации параметров предложенного коллективного алгоритма нечеткой обработки электрокардиологической информации показали целесообразность применения следующих параметров СНВ: алгоритм нечеткого вывода Сугено; оригинальная система входных признаков: мощность; коэффициент подобия; производная и продолжительность импульса; метод дефаззификации; взвешенное среднее.

5. Разработана методика верификации результатов функционирования алгоритмов обработки ЭКГ-информации, позволяющая проводить поэтапную сравнительную оценку эффективности на основе верификации результатов работы авторских алгоритмов и алгоритмов-аналогов с применением международных аннотированных баз ЭКГ-данных. Определены показатели эффективности этих результатов в различных условиях, в том числе, в сложных сигнальных и сигнально-помеховых ситуациях. Алгоритм КАНЛ в условиях интенсивных помех обеспечивает увеличение диагностической эффективности в среднем на 5% по сравнению с алгоритмами-аналогами и позволяет повысить надежность работы медицинских информационных систем для кардиодиагностики.

## **ОСНОВНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ**

### **Публикации в изданиях, рекомендованных ВАК:**

1. Шамин, Е.А. Перспективные направления ЭКГ-анализа / Е.А. Шамин, Е.В. Истомина, Б.А. Истомин, А.А. Лавреев // Известия ЮФУ. Технические науки. Тематический выпуск “Медицинские информационные системы”. – Таганрог, 2009. – № 9. – С. 89–93.

2. Шамин, Е.А. Преимущества применения нечеткого алгоритма при обнаружении информационных участков электрокардиосигнала / Е.А. Шамин, Т.В. Истомина // Известия ЮФУ. Технические науки. Тематический выпуск “Медицинские информационные системы”. – Таганрог, 2010. – № 8. – С. 63–68.

3. Шамин Е.А. Нечеткая модификация коллективных методов обработки биомедицинской информации / Е.А. Шамин, Т.В. Истомина // XXI век: итоги прошлого и проблемы настоящего плюс: Научно-методический журнал. Серия: технические науки. – 2012. – № 05(09). – С. 21–29.

### **Публикации в других изданиях:**

4. Шамин, Е.А. Применение информационных технологий в медицине. / Е.А. Шамин, Т.В. Истомина // Информационные и управленческие технологии в медицине: Сборник статей II Всероссийской научно-технической конференции. – Пенза: Приволжский Дом знаний, 2008. – С. 23–27.

5. Шамин, Е.А. Перспективы применения нечеткой логики для обработки биосигналов / Е.А. Шамин, Т.В. Истомина // Сборник материалов научной конференции студентов, аспирантов и молодых ученых “Актуальные проблемы науки и образования”. – Пенза: Изд-во ПГТА, 2008. – С. 202–205.

6. Шамин, Е.А. Применение аппарата нечеткой логики для анализа электрокардиосигнала / Е.А. Шамин // Инновационные технологии в экономике, информатике и медицине: Сборник статей V межрегиональной научно-практической конференции. – Пенза: ПГТА, 2008. – С. 183–185.

7. Шамин Е.А. Интернет-технологии в медицинском образовании / Е.А. Шамин, Т.В. Истомина // Медико-экологические информационные технологии: Сборник статей XI международной научно-технической конференции, посвященной 15-летию кафедры “Биомедицинская инженерия”. – Курск: КГТУ, 2008. – С. 17–20.

8. Шамин Е.А. Применение информационных технологий с использованием нечеткой логики в поиске информативных участков при исследовании ЭКС / Е.А. Шамин, Т.В. Истомина / Биомедсистемы–2008: Материалы международной конференции. – Рязань, 2008. – С. 26–27.

9. Шамин, Е.А. Нечеткий алгоритм анализа электрокардиосигнала / Е.А. Шамин, Т.В. Истомина, Л.Ю. Кривоногов // Информационные и управленческие технологии в медицине и экологии: Сборник статей III Всероссийской научно-технической конференции. – Пенза: ПГТА, 2009. – С 29–33.

10. Шамин Е.А. Помехоустойчивый нечеткий алгоритм обнаружения кардиоимпульсов / Е.А. Шамин // Инновационные технологии в экономике, информатике и медицине: Сборник статей VI межрегиональной научно-практической конференции. – Пенза: ПГТА, 2009. – С. 71–73.

11. Шамин, Е.А. Методология проектирования распределенных медицинских систем на основе концептуальных спецификаций / Е.А. Шамин, Т.В. Истомина, В.В. Горюнова, Ю.В. Молодцова, С.А. Аленин // Надежность и качество: Труды международного симпозиума. – Пенза, 2010. – Т. 2. – С. 11–17.

12. Шамин, Е.А. Нечеткий алгоритм обнаружения вершины кардиоимпульса / Е.А. Шамин, Т.В. Истомина // Актуальные проблемы науки и образования: Сборник материалов III научной конференции студентов, аспирантов и молодых ученых. – Пенза: Изд-во ПГТА, 2010. – С. 59–63.

13. Шамин, Е.А. Проблема выбора параметров для анализа электрокардиосигнала / Е.А. Шамин, Т.В. Истомина // Инновационные технологии в экономике, информатике и медицине: Сборник статей VII Межрегиональной научно-практической конференции студентов и аспирантов. – Пенза: ПГТА, 2010. – С. 42–44.

## **ШАМИН ЕВГЕНИЙ АНАТОЛЬЕВИЧ**

### **КОЛЛЕКТИВНЫЙ АЛГОРИТМ НА БАЗЕ НЕЧЕТКОЙ ЛОГИКИ ДЛЯ ОБРАБОТКИ ЭЛЕКТРОКАРДИОЛОГИЧЕСКОЙ ИНФОРМАЦИИ В СЛОЖНЫХ СИГНАЛЬНО-ПОМЕХОВЫХ СИТУАЦИЯХ**

**Специальность 05.13.17 – теоретические основы информатики**

**Автореферат  
диссертации на соискание ученой степени  
кандидата технических наук**

Компьютерная верстка Д.Б. Фатеева, Е.В. Рязановой

Сдано в производство 23.11.12. Формат 60x84 <sup>1/16</sup>  
Бумага типогр. №1. Печать трафаретная. Шрифт Times New Roman Cyr.  
Усл. печ. л. 1,16. Уч.-изд. л. 1,17. Заказ № 2235. Тираж 100.

---

Пензенская государственная технологическая академия.  
440605, Россия, г. Пенза, пр. Байдукова/ ул. Гагарина, 1<sup>а</sup>/11.