

Учреждение образования  
**“БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ”**

УДК 681.327.12:534.782

**ИВАНОВ**  
Алексей Валерьевич

**ФОРМИРОВАНИЕ ПРОСТРАНСТВА ПРИЗНАКОВ НА ОСНОВЕ  
АНТРОПОМОРФИЧЕСКОЙ ОБРАБОТКИ ИНФОРМАЦИИ В  
РАСПОЗНАВАТЕЛЯХ РЕЧИ В УСЛОВИЯХ ПРОТИВОДЕЙСТВИЯ**

Специальность 05.13.17 – Теоретические основы информатики

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

Минск, 2004

Работа выполнена в учреждении образования “Белорусский государственный университет информатики и радиоэлектроники”.

Научный руководитель - д.т.н., профессор Петровский А.А.  
(Учреждение образования “Белорусский государственный университет информатики и радиоэлектроники”).

Официальные оппоненты:

д.т.н., с.н.с Лобанов Б.М.  
(Объединённый Институт Проблем Информатики  
НАН Республики Беларусь)

к.т.н., доцент Бовбель Е.И.  
(Белорусский государственный университет)

Опонирующая организация - Национальный центр информационных  
ресурсов и технологий НАН Республики  
Беларусь

Защита состоится 22 декабря 2004 г. в 16 часов на заседании совета по защите диссертаций Д02.15.04 при учреждении образования “Белорусский государственный университет информатики и радиоэлектроники” по адресу: 220013, г. Минск, ул. П. Бровки, 6, БГУИР, 1<sup>ый</sup> учебн. корпус, ауд. 232-1, тел. 239-89-89.

С диссертацией можно ознакомиться в библиотеке Учреждения образования “Белорусский государственный университет информатики и радиоэлектроники”.

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность темы диссертации.** Существующие искусственные системы распознавания речи не обладают достаточным уровнем эффективности распознавания в условиях неидеальной акустической обстановки для их применения в качестве интерфейсов окружающих человека технических устройств. Обладая сравнимым с человеческим уровнем эффективности в отсутствии противодействия, искусственные распознаватели существенно проигрывают человеку в способности распознавать речь в присутствии различных типов помех, характерных для реальной акустической обстановки. Поэтому повышение эффективности распознавания искусственными системами в шумах является необходимым для их широкого применения в технике.

Модификация алгоритмов выделения признаков с целью увеличить степень инвариантности получаемых с их помощью признак-векторов является одним из возможных путей решения поставленной проблемы. В рамках этого подхода, доступный инструментарий повышения эффективности в основном ограничивается набором линейных методов и алгоритмов обработки сигналов. Такая ситуация объясняется принципиальной трудностью синтеза нелинейных алгоритмов обработки сигналов с желаемыми свойствами. Принципиальная возможность достижения большей эффективности распознавания при помощи использования нелинейных методов и алгоритмов обработки сигналов является следствием сравнения уровней производительности существующих искусственных систем и человека в реальных акустических условиях. Развитие аудиторного аппарата в ходе биологической эволюции сделало его универсально-эффективным средством оценки окружающей акустической обстановки и выделения полезного сигнала из входного потока акустических колебаний. Сведения, полученные при исследованиях аудиторного аппарата человека доказывают нелинейность его свойств.

Антропоморфический принцип моделирования аудиторной системы человека открывает принципиальную возможность расширить круг методов и алгоритмов обработки сигналов подходящих для повышения эффективности распознавания. Он добавляет нелинейные методы и алгоритмы, построенные по "образу и подобию" важнейших процессов, происходящих внутри аудиторной системы человека и обеспечивающих эффективную обработку поступающей информации биологической системой.

Успешное применение антропоморфических алгоритмов выделения признаков в составе искусственных систем распознавания речи способно серьезно продвинуть эти системы по пути их широкого применения в качестве удобного и производительного естественного интерфейса между человеком и машиной. Более тесное и естественное взаимодействие человека и техники при помощи искусственных систем распознавания речи позволит интенсифицировать промышленное производство и повысить доступность услуг потребления и связи, повысить безопасность труда и снизить

информационную нагрузку человека в его повседневной жизни. Прогресс по всем перечисленным направлениям позволит увеличить экономический потенциал и качество жизни населения Республики Беларусь.

**Связь работы с крупными научными программами, темами.**

Диссертационная работа выполнена в соответствии с научно-техническими заданиями и планами работ кафедры «Электронные вычислительные средства» и НИГ 3.6 «Микропроцессорные системы реального времени» Белорусского Государственного университета информатики и радиоэлектроники в рамках фундаментальных исследований Министерства образования Республики Беларусь: по гранту фонда фундаментальных исследований РБ ГБЦ 01-1089 «Разработать процедуры сжатия речевой информации, обеспечивающие коммерческое качество восстановленной речи», выполняемой в соответствии с Государственной научно-технической программой «Развитие методов и средств системы комплексной защиты информации», в рамках научно-исследовательской работы ГБЦ 04-7041 «Антропоморфические методы и алгоритмы обработки речи и аудиосигналов»; при финансовой поддержке Фонда фундаментальных исследований Республики Беларусь Т99-308 «Новый подход к низкоскоростному кодированию речевых сигналов в условиях зашумления с учётом особенностей слуха человека», а также, в рамках международного сотрудничества с компанией Lernout & Hauspie Speech Products NV (Бельгия) «Automated Speech Recognition Systems in the Adverse Acoustic Environment».

**Цель и задачи исследования.** Целью работы является разработка метода повышения эффективности распознавания речи в условиях неидеальной акустической обстановки в результате применения антропоморфического принципа, пригодного для использования в составе мобильных систем распознавания речи.

Для достижения поставленной цели необходимо решить следующие задачи:

1. Обосновать применение антропоморфического принципа с научной и практической точек зрения в подходе к повышению эффективности распознавания речи в условиях неидеальной акустической обстановки;
2. Разработать подход к синтезу антропоморфических алгоритмов обработки сигналов, имеющих желаемые свойства с точки зрения повышения эффективности распознавания речи искусственными системами;
3. Осуществить выбор наиболее важных процессов, происходящих внутри аудиторного аппарата человека, с точки зрения обработки сигналов биологической системой для их моделирования в составе антропоморфического алгоритма обработки сигналов;
4. Исследовать свойства моделей, описывающих выбранные физиологические процессы, выделить их предположительный алгоритмический смысл и сделать возможной интеграцию сконструированных на этой основе антропоморфических методов и

алгоритмов обработки сигналов в составе антропоморфического алгоритма выделения признаков;

5. Осуществить выбор подходящей схемы организации классификатора с точки зрения задачи распознавания речи в составе мобильных систем;
6. Исследовать возможность повышения вычислительной эффективности реализации антропоморфического алгоритма выделения признаков в виде цифрового устройства;
7. Реализовать программно антропоморфический алгоритм выделения признаков и мобильный распознаватель и исследовать эффективность этих алгоритмов для решения задачи распознавания речи в шумах.

**Объект и предмет исследования.** Объектом исследований являются механизмы выделения признаков искусственных систем распознавания речи. Предметом исследования являются методы и алгоритмы выделения признаков в составе искусственных систем распознавания речи, построенные на основе антропоморфического принципа и обеспечивающие повышение эффективности распознавания речи искусственными системами в условиях шумов.

**Гипотеза.** В диссертации выдвигается гипотеза о том, что уточнённое моделирование нелинейных свойств аудиторного аппарата человека в составе алгоритмов выделения признаков систем распознавания речи может привести к созданию распознавателей, характеризующихся большей эффективностью работы в условиях неидеальной акустической обстановки по сравнению с известными системами.

**Методология и методы проведенного исследования.** Совокупность использованных в диссертации средств исследований базируется на аппарате линейной алгебры, математического анализа, теории дифференциальных уравнений, теоретической механики, численных математических методов, теории вероятностей, теории информации, математической статистики и статистического моделирования. Экспериментальная проверка основных положений и результатов работы проводилась при помощи компьютерного моделирования.

**Научная новизна и значимость полученных результатов.** К числу новых научных результатов относятся следующие:

1. Подход к синтезу алгоритмов обработки сигналов на основе антропоморфического принципа, который, в отличие от известных, позволяет создавать нелинейные алгоритмы обработки сигналов с малым количеством оцениваемых в ходе их синтеза параметров. Оптимальное соответствие таких алгоритмов поставленной задаче основывается на достаточном уровне точности моделирования биологической системы обработки сигналов и предположении, что моделируемая биологическая система оптимально соответствует этой задаче в результате процесса биологической эволюции.
2. Метод синтеза цифровой активной кохлеарной модели, на базе банка перестраиваемых цифровых фильтров, который, в отличие от известных,

минимизирует необходимое количество вычислительных ресурсов и является пригодным для применения в составе алгоритмов выделения признаков. В рамках этого метода предусмотрено рассмотрение критериев устойчивости перестраиваемых цифровых фильтров в составе модели и оценка уровня их аппроксимации в условиях минимизации частоты дискретизации представления сигналов в устройстве.

3. Метод синтеза цифровой реализации резервуарной модели синапса между волосковой клеткой и волоконном аудиторного нерва, который, в отличие от известных методов, основывается на эквивалентном представлении модели в виде комбинации фильтров с постоянными параметрами и механизма регулировки усиления, что позволяет использовать получаемые цифровые реализации модели синапса внутри технических устройств с минимизацией частоты дискретизации представления сигналов.
4. Антропоморфический метод выделения признаков-векторов, который, в отличие от известных методов, является нелинейным способом обработки сигналов. Алгоритмы, созданные в рамках этого метода, точнее соответствуют процессам, происходящим во внутреннем ухе человека по сравнению с известными методами, и, на основании предположения об оптимальности соответствия аудиторной системы человека задаче оценки акустической обстановки, обладают желательными для систем распознавания речи в шумах свойствами, что подтверждается проведёнными экспериментами.
5. Метод распознавания речи на базе гибрида искусственная нейронная сеть – скрытая Марковская модель с антропоморфическим выделением признаков, который, в отличие от существующих, является эффективным при распознавании речи в условиях противодействия характерных для неидеальной акустической обстановки, что признаётся необходимым при решении задачи построения мобильных распознавателей речи.

**Практическая значимость полученных результатов.** Результаты диссертационной работы подготовлены в виде компьютерных моделей для практического применения. Предлагаемые методы могут быть использованы при создании средств распознавания речи повышенной эффективности в шумах, характерных для реальной акустической обстановки, в составе мобильных электронно-вычислительных средств.

Скорейшее промышленное внедрение средств распознавания речи повышенной эффективности в условиях шумов необходимо для обеспечения удобного и производительного взаимодействия человека с окружающей оболочкой технических средств, которая имеет тенденцию к быстрому усложнению и проникновению во все сферы человеческой деятельности и более тесной интеграции с человеком. Распространение технических устройств с голосовым интерфейсом позволит повысить производительность труда человека практически во всех без исключения отраслях промышленности и

транспорта, позволит во многих случаях высвобождать людские ресурсы в торговле и сфере услуг повышая при этом уровень их доступности, повысит эффективность и снизить стоимость средств связи. Более естественный речевой интерфейс позволит снизить общую информационную нагрузку на операторов сложной техники путём частичного замещения визуального интерфейса и, помимо уже отмеченного повышения производительности труда, позволит сделать эксплуатацию этой техники более безопасной.

Результаты диссертационной работы переданы для внедрения в составе средств распознавания речи, систем речевой идентификации абонента и выделения ключевых слов в ООО "ОДИТЕК" (г. Санкт-Петербург, Россия), а также, приняты к использованию в процессе проектирования мобильных мультимедийных устройств и голосовых телекоммуникационных порталов в компании ЗАО "НПП Белсофт"(г. Минск, Республика Беларусь), в процессе разработки мультимедийных устройств, мобильных устройств с голосовым интерфейсом и средств аудио-телекоммуникаций ИЧУТПП "ЦДЛБЕЛ" (г. Минск, Республика Беларусь), в процессе проектирования измерительных систем и систем контроля Минским научно-исследовательским приборостроительным институтом (МНИПИ), а также, используются в учебном процессе на кафедре электронно-вычислительных средств Белорусского государственного университета информатики и радиоэлектроники в дисциплинах "Речевые интерфейсы ЭВС" и "Проектирование проблемно-ориентированных электронных вычислительных средств".

**Основные положения диссертации, выносимые на защиту.** На защиту выносятся следующие научные положения:

- Метод синтеза цифровой активной кохлеарной модели, учитывающей сигнал-зависимую чувствительность слуховой улитки, на базе банка перестраиваемых цифровых фильтров, пригодной для использования в технических устройствах с ограниченным вычислительным ресурсом;
- Метод синтеза цифровой резервуарной модели синапса между волосковой клеткой и волокном аудиторного нерва в виде механизма сигнал-зависимой регуляции усиления;
- Антропоморфический метод и алгоритм выделения признаков-векторов, оптимизированный на ограниченный вычислительный ресурс систем искусственного распознавания речи в условиях шумов окружающей среды;
- Метод и алгоритм распознавания речи на базе гибрида искусственной нейронной сети – скрытая Марковская модель, эффективно использующий ресурс памяти системы, ориентированный на использование в составе мобильных средств распознавания речи.

**Личный вклад соискателя.** Все вошедшие в диссертацию результаты получены при непосредственном личном участии автора. Все предлагаемые методы разработаны в соавторстве с научным руководителем и программно реализованы лично автором. В публикациях с соавторами вклад соискателя

определяется рамками излагаемых в диссертации результатов. Научный руководитель принимал участие в постановке задач и определении возможных путей их решения.

**Апробация результатов диссертации.** Результаты исследований, включённые в диссертацию докладывались и обсуждались на 9-ой Международной Конференции “Речь и Компьютер” SPECOM’2004 (Санкт-Петербург, Россия, 2004); 116<sup>th</sup> Convention of the Audio Engineering Society (Berlin, Germany, 2004); IEEE Workshop Signal Processing’2003 (Poznan, Poland, 2003); 9<sup>th</sup> International Workshop on Systems, Signals and Image Processing IWSSIP’02 (Manchester, UK, 2002); Internationale Konferenz an der Hochschule Mittweida SATERRA’99 (Mittweida, Germany, 1999); International Conference on Neural Networks and Artificial Intelligence ICNNAI’99 (Brest, Belarus, 1999); A Workshop on Text, Speech, Dialog TSD’99 (Plzen, Czech Republic, 1999); 5th International Conference on Pattern Recognition and Information Processing PRIP’99 (Minsk, Belarus, 1999); 6th International Workshop on Systems, Signals and Image Processing IWSSIP’99 (Bratislava, Slovakia, 1999); III-ей Международной Научно-Технической Конференции “СОВРЕМЕННЫЕ СРЕДСТВА СВЯЗИ” (Нарочь, Беларусь, 1999); 5th Conference Komputerowe Wspomaganie Badan Naukowych, V KK KOWBAN’98 (Polanica Zdroj, Poland, 1998)

**Опубликованность результатов.** По материалам диссертационной работы опубликовано 13 печатных работ, включая 1 препринт, 5 статей в научно-технических журналах и научных сборниках и 7 статей в материалах международных конференций. Суммарный объём публикаций составляет примерно 107 печатных страниц.

**Структура и объём диссертации.** Диссертационная работа состоит из введения, общей характеристики работы, пяти глав, заключения, списка использованных источников и приложений. Она содержит 100 страниц основного текста; 52 рисунка и 18 таблиц, расположенных в тексте диссертации; 3 приложения на 28 страницах; в списке использованных источников на 12 страницах представлено 173 наименования.



## СОДЕРЖАНИЕ РАБОТЫ

**Во введении** показана актуальность темы диссертационной работы, формулируются цель и задачи исследования.

**В главе 1** производится анализ текущего состояния речевых технологий их места и роли в развитии устройств обработки и передачи информации. На примере сравнения эффективности человека и автоматических распознавателей в различных условиях и формулировках задачи распознавания показано, что при уменьшении соотношения сигнал-шум до 0 дБ процент ошибок распознавания человеком остаётся практически неизменным (в пределах 1%), а искусственные системы практически полностью перестают работать (количество ошибок составляет ~ 100%). Делается вывод, что повышение эффективности искусственных распознавателей речи в условиях неидеальной акустической обстановки представляет собой актуальную задачу, решение которой необходимо для широкого применения средств распознавания речи в составе технических устройств.

Математическая формулировка задачи распознавания предоставляет возможность разделить структуру распознавателя речи на три основных этапа: выделение признаков, акустическое и лингвистическое моделирование:

$$P(W_i | \{X\}^k) = \sum_{\forall \{C\}^k \Rightarrow W_i} P(W_i, \{C\}^k | \{X\}^k) = \sum_{\forall \{C\}^k \Rightarrow W_i} P(\{C\}^k | \{X\}^k) P(W_i | \{C\}^k, \{X\}^k) = \\ = \sum_{\forall \{C\}^k \Rightarrow W_i} P(\{C\}^k | \{X\}^k) P(W_i | \{C\}^k),$$

где  $\{X\}^k$  обозначает последовательность признак-векторов полученных на этапе выделения признаков;  $\{C\}^k$  - гипотетическую последовательность состояний, получаемую на этапе акустического моделирования;  $W_i$  - гипотезу о произнесённом слове (словарном отрывке), получаемую на этапе лингвистического моделирования. Общая структура распознавателя речи показана на рис. 1.



Рис. 1 Общая структура распознавателя речи

Каждый из структурных элементов предоставляет возможности для поиска способов повышения эффективности распознавателей: модификация алгоритма выделения признаков, адаптация акустической модели к предположительным условиям работы распознавателя, применение средств языкового моделирования для использования грамматического контекста при выдвижении гипотез о произнесённом отрывке.

Модификация алгоритма выделения признаков наиболее универсальна и подходит для повышения эффективности распознавателей во всех без исключения вариантах формулировки задачи распознавания речи. Основным критерием качества механизма выделения признаков является максимизация взаимной информации между признак-вектором и источником речи (диктором) вне зависимости от акустической обстановки в процессе распознавания:

$$J(X, S) = H(S) - H(S|X) = H(X) - H(X|S) = \\ = - \int_{G(X)} p(X) \log_2 p(X) dX + \sum_{\forall C_i \in C} \int_{G(X)} p(X|C_i) P(C_i) \log_2 p(X|C_i) dX,$$

где  $X$  обозначает некоторый признак;  $S$  - источник, а  $C_i$  - его состояние.

Дополнительно, для облегчения процесса производства распознавателей в условиях априорной неопределённости конкретного типа помех при распознавании, выдвигается требование точной или приближительной инвариантности признак-вектора к противодействию акустической обстановки.

Сравнение производительности человека и машин также приводит к предположению, что более точное моделирование механизма обработки информации человеческим аудиторным аппаратом приведёт к созданию искусственных систем, в большей степени соответствующих уровню человека, т.е. обладающих свойством универсальной эффективности распознавания в широком спектре возможных противодействий среды. Современные методики выделения признаков широко используют знания о свойствах аудиторного аппарата человека. Примерами перцептуально обоснованных методик могут служить алгоритмы RASTA и PLP. Но, в основном, эти знания почерпнуты из психоакустического субъективного описания целостной аудиторной системы в экспериментах с простыми звуками.

Психоакустика доказывает, что аудиторный аппарат человека – нелинейная система, параметры которой зависят от времени. Следовательно, не существует возможности точно предсказать её отклик на такой сложный и нестационарный сигнал как речь при использовании только той информации, которая заключена в психоакустическом описании.

Путь решения данной проблемы лежит в использовании объективно зафиксированных физиологических свойств элементов аудиторной системы. Этот подход к моделированию и созданию алгоритмов был назван “антропоморфическим принципом”, т.е. принципом создания искусственных устройств обработки информации, работающих по “образу и подобию” человека, в надежде на то, что при достаточно точном уровне моделирования искусственные системы приобретут свойства, характерные для их прообраза.

**В главе 2** формулируется постановка задачи конструирования антропоморфического алгоритма обработки информации. Дается математическая характеристика нелинейных систем как не удовлетворяющих принципу суперпозиции и на её основе делается вывод о сложности применения концепции “чёрного ящика” к моделированию аудиторного

аппарата. Критерий выбора конкретных физиологических свойств элементов аудиторной системы для моделирования сформулирован на основе предположения о том, что в процессе биологической эволюции аудиторный аппарат человека стал оптимальным средством оценки окружающей акустической обстановки и его важнейшие свойства обусловлены алгоритмической необходимостью.

На основе обзора исследований нарушений слуха человека показано, что недостаточность функции слуховой улитки, выражающаяся в нарушении механизмов взаимного влияния компонент сложного входного сигнала, служит причиной пониженной способности пациентов к распознаванию речи в условиях противодействия. Соответственно, область поиска физиологических механизмов ограничена механизмами внутреннего уха, вносящими вклад в психоакустически наблюдаемое маскирование. Среди них выделены явления супрессии и адаптивности откликов внутренних волосковых клеток.

Аудиторная супрессия заключается в объективном снижении отклика участка базилярной мембраны на стимуляцию тоном с частотой, совпадающей с центральной для данного участка, в присутствии других компонент сложного сигнала. Это явление совместно с распространением возбуждения служат основными механизмами частотного маскирования. Аудиторная супрессия моделируется в рамках представления об активной улитке.

Адаптивность откликов внутренних волосковых клеток является одним из механизмов временного маскирования. Это явление заключается в неравномерности отклика с течением времени: мгновенный отклик на только что представленный стимул значительно превышает установившееся значение отклика на стационарный сигнал. Адаптивность отклика объясняется свойствами химического синапса между волосковой клеткой и волокном аудиторного нерва и количественно описывается при помощи т.н. резервуарной модели химического синапса, которая описывает циркуляцию нейротрансмиттера в химическом синапсе (см. рис. 2 и 3).

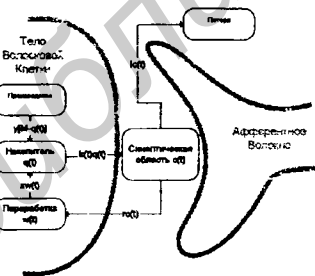


Рис. 2 Течение нейротрансмиттера в резервуарной модели химического синапса

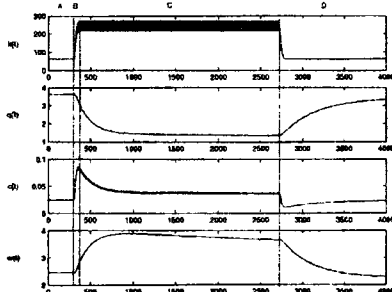


Рис. 3 Отклик  $c(t)$  резервуарной модели на возбуждение  $k(t)$

**В главе 3** производится конструирование антропоморфического алгоритма, основанного на модели активной улитки и резервуарной модели химического синапса. Каждый из эквивалентных кохлеарных фильтров, характеризующих отклик отдельных участков базилярной мембраны, моделируется каскадом цифровых линейных постоянных и перестраиваемых полюсных гамма-тон фильтров. Параметры перестраиваемых фильтров контролируются через цепь обратной связи, которая является феноменологической моделью активного процесса в улитке и включает в себя описание электромотильных откликов внешних волосковых клеток.

Отдельное внимание уделяется вопросам устойчивости и аппроксимации перестраиваемых цифровых фильтров в составе алгоритма при низкой частоте дискретизации входного сигнала. Показано, что реализация перестраиваемых фильтров в виде нормированной решетчатой структуры позволяет гарантировать их асимптотическую устойчивость во всех точках их параметрического пространства, в которых устойчивы фильтры аналогичного типа с постоянными параметрами. Предложен способ оценки аппроксимации обобщённых амплитудно- и фазо-частотных характеристик перестраиваемых фильтров в условиях произвольного закона изменения их параметров. И на примере линейного закона приведены оценки аппроксимации при различной скорости изменения параметров. На основе рассмотрения результатов обработки речевого сигнала моделью активной улитки дана алгоритмическая интерпретация эффекта суппрессии. В результате суппрессии повышается контраст перцептуальной сонограммы (см. рис. 4), усиливаются тоновые компоненты в низкочастотной области и шумовые – в высокочастотной, что хорошо согласуется со спектральной структурой речи.



Рис. 4 Сонограммы речевого отрывка без модели суппрессии (слева) и с моделью суппрессии (справа)

Данный вывод подтверждается результатами психоакустических экспериментов, измеряющих эффект автосуппрессии гласных звуков в процессе восприятия их человеком.

Эквивалентная цифровая схема резервуарной модели синапса получена на основе анализа дифференциальных уравнений модели волосковой клетки Меддиса. Она представляет собой алгоритм сигнало-зависимой регулировки

усиления и, таким образом, является нелинейным алгоритмом обработки сигнала (см. рис. 5).

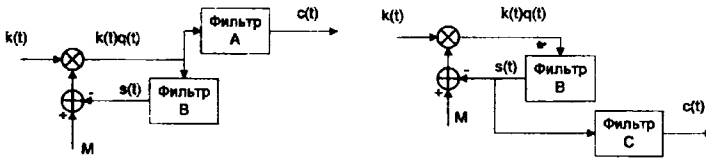


Рис. 5 Эквивалентные схемы резервуарной модели.

В условиях малых отклонений системы от положения равновесия получена линейная аппроксимация эквивалентной цифровой схемы. Качественное сходство линейной аппроксимации и алгоритмов модуляционной фильтрации позволяет сделать вывод об алгоритмическом смысле эквивалентной цифровой схемы. Он заключается в усилении нестационарностей, модуляционные частоты которых лежат в полосе пропускания фильтра линейной аппроксимации. Анализ свойств линейных цифровых фильтров, входящих в состав эквивалентной цифровой схемы а также фильтра линейной аппроксимации позволяет сделать количественную оценку критерия их устойчивости. В работах авторов исходной резервуарной модели давалось лишь качественное указание на необходимость достаточно высокой частоты дискретизации сигналов при использовании цифровой модели волосковой клетки.

**Глава 4** посвящена оптимизированной с точки зрения потребления вычислительных ресурсов реализации разработанного антропоморфического алгоритма выделения признаков в составе распознавателя речи на базе гибрида нейронная сеть – скрытая Марковская модель.

В качестве подходящей структуры реализации выбрана схема банка фильтров с последовательной октавной децимацией (см. рис. 6).

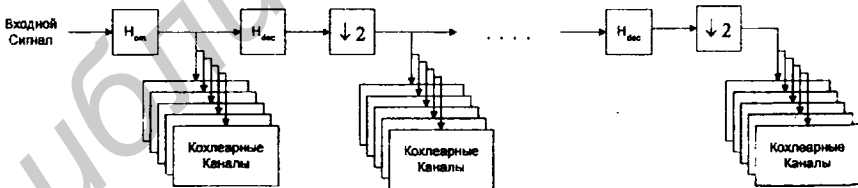


Рис. 6 Структурная схема реализации кохлеарной модели при помощи банка фильтров с последовательной децимацией

В соответствии с принципами построения такой структуры произведено распределение кохлеарных каналов по частотным диапазонам (см. таблицу 1). Предложена схема подсчёта частотных диапазонов, гарантирующая наиболее равномерное распределение вычислений по времени.

Распределение кохлеарных каналов по частотным диапазонам

Номер диапазона	1	2	3	4	5	6	7	8
Диапазон, Гц	0-62,5	0-125	0-250	0-500	0-1000	0-2000	0-4000	0-8000
Начало, № канала	1	7	17	32	51	74	98	125
Конец, № канала	6	16	31	50	73	97	124	152
Количество каналов	6	10	15	19	23	24	27	28

Для дальнейшей экономии вычислительных ресурсов все нелинейные функции в составе модели предложено реализовывать в виде кусочно-линейных аппроксимаций. Ввиду сложности формул перехода от сигнала обратной связи к параметрам перестраиваемых фильтров в составе кохлеарных каналов, их предлагается выбирать из предварительно посчитанной таблицы соответствия. При этом шаг дискретизации параметра качества выбирается в соответствии с допустимым уровнем погрешностей аппроксимации, который определяется при помощи методики описанной в главе 3.

Алгоритм усиления нестационарных компонент был дополнительно модифицирован для его реализации в составе структуры с последовательной октавной децимацией. Необходимость модификации обусловлена требованием стабильности фильтров, входящих в состав эквивалентной цифровой реализации, при любой, сколь угодно низкой частоте дискретизации сигналов.

Произведенные численные оценки показывают, что реализация алгоритма в виде структуры с последовательной октавной децимацией позволяет добиваться более чем 3-х кратного сокращения вычислительных потребностей алгоритма. Прямая реализация требует пересчёта всех 152 кохлеарных каналов при поступлении нового отсчёта входного сигнала. Необходимое количество пересчётов кохлеарных каналов на протяжении 128 отсчётов выражается

$$R_0 = 128 * 152 = 19456.$$

Применение схемы с последовательной децимацией позволяет сократить количество пересчётов кохлеарных каналов до

$$R_D = 128 * 28 + 64 * 27 + 32 * 24 + 16 * 23 + 8 * 19 + 4 * 15 + 2 * 10 + 6 = 6686.$$

Дополнительный выигрыш достигается объединением компонент движущегося среднего ПТФ фильтров кохлеарных каналов в каждом из частотных диапазонов.

Для окончательного формирования признак-векторов результаты работы антропоморфического алгоритма децимируются до частоты 125 Гц и коэффициентов, их первых и вторых производных.

Классификатор в антропоморфическом распознавателе речи основан на основе гибрида многоуровневый перцептрон – скрытая Марковская модель. При этом многоуровневый перцептрон используется для акустического моделирования, а скрытая Марковская модель – для лингвистического. Для их объединения выбран первый тип гибрида МУП/СММ ввиду его простоты.

Распределение вероятности порождения акустического вектора данным классом  $p(X|c_i)$  моделируется с помощью нейронной сети. Выходы натренированной сети оценивают  $P(c_i|X)$  и применяя формулу Байеса, можно получить искомую вероятность  $\frac{p(X|c_i)}{p(X)} = \frac{P(c_i|X)}{P(c_i)}$ .

Значение  $p(X)$  - константа в процессе распознавания и можно получить взвешенное распределение  $p(X|c_i)$ , которое в дальнейшем используется в скрытой Марковской модели. Выбранная схема построения классификатора обладает рядом преимуществ по отношению к классическим классификаторам на основе СММ, т.к. объединяет сильные стороны указанных способов статистического моделирования и позволяет избавиться от ряда модельных предположений, присущих классическому подходу.

**Глава 5** является экспериментальной частью работы. В ней производятся численные оценки величин, обозначенных в главе 1 как критерии качества механизма выделения признаков. Полученные оценки сравниваются для случая антропоморфического алгоритма, случая базового алгоритма выделения признаков и при применении алгоритма модуляционной фильтрации RASTA в составе выделения признаков. Оценка взаимной информации между источником речи и компонентами признак-вектора производится для случая отсутствия противодействия. Она основана на непараметрической оценке распределений вероятности в исходной формуле при помощи гистограмм и замене интегрирования конечной суммой:

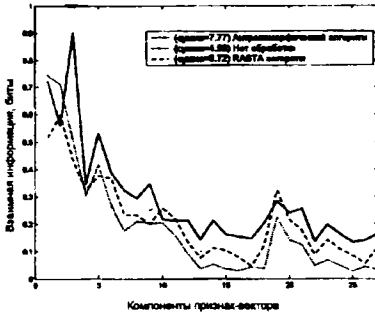
$$J_{\Delta X}(X, S) \approx J_{\Delta X}(X, S) = \sum_{\forall i} \sum_{\forall j} \frac{N(C_i)}{N} \frac{N(\Delta X_j, C_i)}{N(C_i)} \log_2 \frac{N(\Delta X_j, C_i)}{N(C_i)} - \sum_{\forall j} \frac{N(\Delta X_j)}{N} \log_2 \frac{N(\Delta X_j)}{N} =$$

$$= \log_2 N + \frac{1}{N} \left( \sum_{\forall i} \sum_{\forall j} N(C_i, \Delta X_j) \log_2 N(C_i, \Delta X_j) - \sum_{\forall i} N(C_i) \log_2 N(C_i) - \sum_{\forall j} N(\Delta X_j) \log_2 N(\Delta X_j) \right),$$

где  $N$  обозначает полное количество испытаний;  $N(\Delta X_j)$  - количество испытаний, в которых значение величины  $X$  принадлежало интервалу  $\Delta X_j$ ;  $N(C_i)$  - количество испытаний, в которых источник находился в состоянии  $C_i$ ;  $N(C_i, \Delta X_j)$  - количество испытаний, в которых источник находился в состоянии  $C_i$ , а значение величины  $X$  принадлежало интервалу  $\Delta X_j$ .

Критерием обоснованности гистограммных оценок, т.е. критерием выбора количества отсчётов  $m$  гистограммы, покрывающих область определения признака, принят информационный критерий Акаике  $m \sim N^{\frac{1}{k+2}}$ , но также, возможно использование критерия минимума длины описания.

Численная оценка взаимной информации между источником и компонентами признак-вектора приводит к выводу, что антропоморфический алгоритм позволяет получать наиболее информативные признаки среди всех сравниваемых методов их выделения. Эта оценка представлена на рис. 7.



**Рис. 7** Взаимная информация между компонентами признака-вектора и источником речи в отсутствие противодействия



**Рис. 8** Проверка статистической независимости компонентов признака-вектора

При помощи подсчёта статистики Кендалла проверена статистическая зависимость между компонентами признака-вектора, при условии, что входным сигналом служит речь. Результат проверки указывает на весьма вероятное (доверительная вероятность  $\beta = 0.999$ ) наличие этой зависимости (см. рис. 8). Оценка полной взаимной информации между признаком-вектором и источником речи при помощи непараметрического гистограммного метода исключительно затруднена ввиду большой вычислительной сложности этой операции.

В качестве меры инвариантности признака-векторов выбрано евклидово расстояние в пространстве кепстральных коэффициентов без использования нулевого коэффициента. Помимо полного набора свойств, присущих евклидовым расстояниям эта мера обладает рядом полезных свойств, с перцептуальной точки зрения: она инвариантна к средней громкости входного сигнала и отражает интегральную непохожесть двух воздействий таким, как оно выглядит после преобразования к перцептуальной громкости. При помощи большого количества примеров речи производится измерение среднего расстояния, возникающего между признака-векторами при добавлении к исходному речевому сигналу определённого количества помех окружающей среды (см. таблицу 2).

**Таблица 2**

**Средние расстояния между звуками (чистыми и в условиях противодействия)**

$\bar{m}[d_c]$	Противодействие			
	Шум SNR=60 дБ	Шум SNR=30 дБ	Шум SNR=10 дБ	Телефонный канал
Отсутствие обработки	0.41597	0.78894	1.05047	0.49298
RASTA	0.09842	0.17563	0.22338	0.05300
Автопоморфический алгоритм	0.26853	0.44951	0.42615	0.16665



Это измерение подтверждает, что применение антропоморфического алгоритма позволяет существенно повысить инвариантность признак-вектора к противодействию окружающей среды. Приведенная оценка доверительного интервала инструментальной погрешности (доверительная вероятность  $\beta = 0.999$ ) при измерении среднего расстояния указывает на то, что данный эксперимент исключительно надёжно зафиксировал положительный эффект от применения антропоморфического алгоритма в составе механизма выделения признаков. Половина ширины доверительного интервала инструментальной погрешности составляет  $\varepsilon[\bar{m}[\Delta d]] = 3,29 * \sigma[\bar{m}[\Delta d]] \approx 0.004$ .

Окончательная оценка эффективности антропоморфического алгоритма выделения признаков произведена при помощи измерения уровня производительности гибридного распознавателя. Она подтверждает, что применение антропоморфического алгоритма позволяет добиваться производительности, по крайней мере, не худшей, чем в случае использования алгоритма RASTA и значительно лучшей, чем в случае базового алгоритма (см. рис. 9). Процедура оптимизации параметров антропоморфического алгоритма может далее улучшить его свойства для того, чтобы он в большей степени соответствовал задаче распознавания речи искусственными системами.

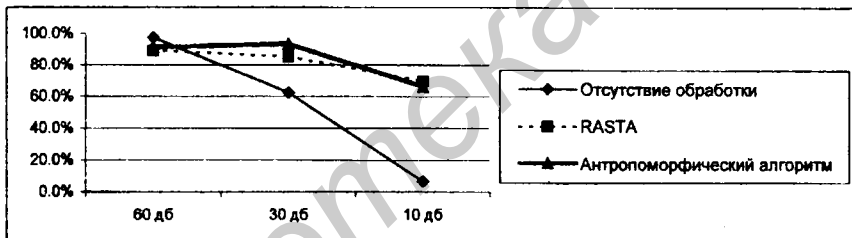


Рис. 9 График производительности гибридного распознавателя в шумах в сочетании с различными алгоритмами переднего края

**Приложения** содержат краткое психоакустическое описание аудиторной системы; сжатую анатомико-физиологическую характеристику внутреннего уха человека; акты внедрения результатов диссертационной работы.

## ЗАКЛЮЧЕНИЕ

Основные результаты диссертационной работы следующие:

1. Предложен подход к синтезу нелинейных алгоритмов обработки сигналов на основе антропоморфического принципа, пригодных для реализации в составе технических устройств с целью улучшения их характеристик при работе в шумах, характерных для реальной акустической обстановки [11, 13], который, в отличие от известных, позволяет создавать нелинейные алгоритмы обработки сигналов с малым количеством оцениваемых в ходе синтеза алгоритма параметров. Оптимальность соответствия этих алгоритмов поставленной задаче основывается на достаточном уровне точности

моделирования биологической системы обработки сигналов и предположении, что моделируемая биологическая система оптимально соответствует этой задаче в результате процесса биологической эволюции.

2. Разработан метод синтеза цифровой активной кохлеарной модели на базе банка перестраиваемых цифровых фильтров [1, 12], который, в отличие от известных, минимизирует необходимое количество вычислительных ресурсов и является пригодным для применения в составе алгоритмов выделения признаков. В рамках этого метода предусмотрено рассмотрение критериев устойчивости перестраиваемых цифровых фильтров в составе модели и оценка уровня их аппроксимации в условиях минимизации частоты дискретизации представления сигналов в устройстве.

3. Разработан метод синтеза цифровой реализации резервуарной модели синапса между волосковой клеткой и волокном аудиторного нерва в виде механизма сигнала-зависимой регулировки усиления, пригодной для использования в технических устройствах с ограниченным вычислительным ресурсом [1], который, в отличие от известных методов, основывается на эквивалентном представлении модели в виде комбинации фильтров с постоянными параметрами и механизма регулировки усиления, что позволяет использовать получаемые цифровые реализации модели синапса внутри технических устройств с минимизацией частоты дискретизации представления сигналов.

4. Разработан антропоморфический алгоритм выделения признаков-векторов, состоящий из алгоритмов повышения контраста перцептуальной сонограммы и усиления нестационарных компонентов речевого сигнала и оптимизированный с точки зрения потребления вычислительных ресурсов, для систем искусственного распознавания речи повышенной эффективности распознавания в условиях шумов [13], который, в отличие от известных методов, является нелинейным способом обработки сигналов. Алгоритмы, созданные в рамках этого метода, точнее соответствуют процессам, происходящим во внутреннем ухе человека по сравнению с известными методами, и, на основании предположения об оптимальности соответствия аудиторной системы человека задаче оценке акустической обстановки, обладают желательными для систем распознавания речи в шумах свойствами.

5. Разработан метод распознавания речи на базе гибрида искусственная нейронная сеть – скрытая Марковская модель с антропоморфическим выделением признаков [2, 5, 7-10], который, в отличие от существующих, является эффективным при распознавании речи в условиях противодействия, характерных для неидеальной акустической обстановки, что признаётся необходимым при решении задачи построения мобильных распознавателей речи.

## СПИСОК ОПУБЛИКОВАННЫХ РАБОТ ПО ТЕМЕ ДИССЕРТАЦИИ

### Препринт

1. Ivanov A.V., Petrovsky A.A. A composite physiological model of the inner ear for audio coding. - Berlin: 116th AES Convention, 2004. – 20 p. - (Preprint / 6082).

### Статьи в журналах и научных сборниках

2. Иванов А.В., Петровский А.А. Тренировка многоуровневых перцептронов в задаче статистической идентификации фонем с использованием базы звуковых данных ТИМТ // Известия Белорусской Инженерной Академии. - 1998. - № 2(6)/1. - С.46-52.
3. Petrovsky A., Ivanov A., Baszun J. An Attempt To Adequately Estimate Intelligibility Of The Speech Perceived Through Cochlear Implant In Noisy Environment Based On The Neural Network Approach // Journal of the University of Applied Sciences Mittweida. - 1999. - N 3. - P.321-330.
4. Ivanov A., Petrovsky A. MLPs and Mixture Models for the Estimation of the Posterior Probabilities of Class Membership // Lecture notes in Artificial Intelligence. - Berlin: Springer-Verlag. - 1999.- P.215-218.
5. Иванов А.В., Петровский А.А. Методы Построения Устройств Распознавания Речи на Базе Гибрида Нейронная Сеть/Скрытая Марковская Модель // Нейрокомпьютеры: разработка и применение. - Москва: 2002. - N 12. - С.26-36.
6. Петровский А.А., Серков В.В., Иванов А.В., Башун Я.Н. Психоакустика и обработка речевых сигналов // Радиотехника и Электроника: Сб. ст. - 1999. - Вып.23. - С.110-119.

### Статьи в материалах международных конференций

7. Ivanov A., Petrovsky A. Software Kit for Simulation and Evaluation of the ASR Front-End algorithms based on MatLab Integrated Shell // Komputerowe Wspomaganie Badan Naukowych: Proc. 5th Conf. / V KK KOWBAN'98, Polanica Zdroj, Poland, 15-17 October 1998.
8. Ivanov A., Petrovsky A. Training Multi-Layer Perceptrons in the problem of Static Phoneme Identification with the use of TIMIT Speech Corpus // Systems, Signals and Image Processing: Proc. 6th International Workshop on, Bratislava, Slovakia, 2-4 June 1999. - P.118-121.

9. Ivanov A., Petrovsky A. Experiments with Neural Networks for Sequence Recognition in Application to Automatic Speech Recognition // Pattern Recognition and Information Processing: Proc. 5th International Conference on / PRIP'99, Minsk, Belarus, 18-20 May 1999. - P.149-154.
10. Ivanov, A. Petrovsky A. Temporal Processing Neural Networks for Speech Recognition // Neural Networks: Proc. International Conference on / ICNN'99, Brest, Belarus, 12-15 October 1999. - P.117-125.
11. Ivanov A.V., Likhachev D.S., Petrovsky A.A. Spiking Neuron Auditory Model for Speech Processing Systems // Systems, Signals and Image Processing: Proc. of 9th International Workshop on, Manchester, UK, 7-8 November 2002.
12. Ivanov A., Petrovsky A. Auditory Models for Robust Feature Extraction: Suppression // Signal Processing: Proc. of IEEE Workshop, Poznan, Poland, 10 October 2003. - P.23-28.
13. Ivanov A.V., Petrovsky A.A. Anthropomorphic feature extraction algorithm for speech recognition in adverse environments // Speech and Computer: Proc. 9<sup>th</sup> International Conference on / SPECOM'2004, St. Petersburg, Russia, 20-22 September 2004. – P.17-21.



Иванов А. В. 1

## РЭЗЮМЭ

дысертацыі Іванова Аляксея Валер'евіча

### **“Фарміраванне прастранства прызнакаў на аснове антрапамарфічнай апрацоўкі інфармацыі ў распазнавальніках маўлення ва ўмовах процівадзеяння”**

**Ключавыя словы:** аўтаматычнае распазнаванне маўлення, вылучэнне прызнакаў, антрапамарфічная апрацоўка сігналаў, псіхаакустыка, фізіялогія слыху, унутранае вуха, аўдыторныя мадэлі, мадэль актыўнага слімака, мадэль унутранай валасковай клеткі Медзіса, аўдыторная супрэсія, адаптацыя водгукаў унутраных валасковых клетак, нелінейныя залежныя ад часу сістэмы, аўтаматычная рэгуліроўка узмацнення, скрытыя Маркаўскія мадэлі, шматузроўневыя перцептроны.

Аб'ектам даследавання з'яўляюцца механізмы вылучэння прызнакаў штучных сістэм распазнавання маўлення. Прадметам даследавання з'яўляюцца метады і алгарытмы вылучэння прызнакаў у саставе штучных сістэм распазнавання маўлення, якія пабудаваны на аснове антрапамарфічнага прынцыпа і якія забяспечваюць павялічэнне эфектыўнасці распазнавання маўлення штучнымі сістэмамі ва ўмовах шуму.

Мэтай работы з'яўляецца распрацоўка метада павялічэння эфектыўнасці распазнавання маўлення ва ўмовах неідэальных акустычных абставін у выніку прымянення антрапамарфічнага прынцыпу, які можа быць выкарыстан у саставе мабільных сістэм распазнавання маўлення.

Распрацаван метада павялічэння эфектыўнасці распазнавальнікаў маўлення ва ўмовах неідэальных акустычных абставін на аснове выкарыстання антрапамарфічнага прынцыпу ў працэсе праектавання алгарытму вылучэння прызнакаў. Антрапамарфічны алгарытм вылучэння прызнакаў уключае алгарытм павялічэння кантрастнасці перцептуальнай санаграммы, які заснаваны на мадэлі актыўнага слімака і звязаная з аб'ектыўна зафіксаванай з'явай аўдыторнай супрэсіі, а таксама, алгарытма узмацнення нестацыянарных кампанент моўнага сігнала, які заснаваны на рэзервуарнай мадэлі хімічнага сінапса, якая адлюстроўвае аб'ектыўна зафіксаваную з'яву адаптацыі ўнутраных валасковых клетак. Праведзеныя даследаванні даказваюць, што антрапамарфічны алгарытм з'яўляецца больш эфектыўным сродкам вылучэння прызнакаў у саставе штучных сістэм распазнавання маўлення ва ўмовах розных тыпаў процівадзеяння навакольных акустычных абставін, забяспечваю атрыманне больш інфарматыўных, чым пры класічных метадах, прызнакаў, якія дадаткова характэрныя павялічанай ступенню інварыянтнасці да процівадзеяння.

Вынікі даследаванняў перададаны ў Мінскі навукова-даследчы прыборабудаўнічы інстытут (МНДП), ЗАТ “НПП Белсофт”, ІЧУГПІ “ЦДЛБЕЛ” (г. Мінск), ТАА “Одітэк” (г. Санкт-Пецярбург, Рассія) і ўкараняны ў навучальны працэс БДУІР.

## РЕЗЮМЕ

диссертации Иванова Алексея Валерьевича

### **“Формирование пространства признаков на основе антропоморфической обработки информации в распознавателях речи в условиях противодействия”**

**Ключевые слова:** автоматическое распознавание речи, выделение признаков, антропоморфическая обработка сигналов, психоакустика, физиология слуха, внутреннее ухо, аудиторные модели, модель активной улитки, модель внутренней волосковой клетки Меддиса, аудиторная суппрессия, адаптация откликов внутренних волосковых клеток, нелинейные зависящие от времени системы, автоматическая регулировка усиления, скрытые Марковские модели, многоуровневые перцептроны.

Объектом исследований являются механизмы выделения признаков искусственных систем распознавания речи. Предметом исследования являются методы и алгоритмы выделения признаков в составе искусственных систем распознавания речи, построенные на основе антропоморфического принципа и обеспечивающие повышение эффективности распознавания речи искусственными системами в условиях шумов.

Целью работы является разработка метода повышения эффективности распознавания речи в условиях неидеальной акустической обстановки в результате применения антропоморфического принципа, пригодного для использования в составе мобильных систем распознавания речи.

Разработан метод повышения эффективности распознавателей речи в условиях неидеальной акустической обстановки на основе применения антропоморфического принципа в процессе проектирования алгоритма выделения признаков. Антропоморфический алгоритм выделения признаков включает в себя алгоритм повышения контрастности перцептуальной сонограммы, основанный на модели активной улитки, связанной с объективно зафиксированным явлением аудиторной суппрессии, а также, алгоритм усиления нестационарных компонент речевого сигнала, основанный на резервуарной модели химического синапса, отражающей объективно зафиксированное явление адаптации внутренних волосковых клеток. Проведенные исследования доказывают, что антропоморфический алгоритм является более эффективным средством выделения признаков в составе искусственных систем распознавания речи в условиях различных типов противодействия окружающей акустической обстановки, обеспечивая получение более информативных, чем в случае классических методов, признаков, дополнительно характеризующихся повышенной степенью инвариантности к противодействию.

Результаты исследований переданы для внедрения в Минский научно-исследовательский приборостроительный институт (МНИПИ), ЗАО “НПП Белсофт”, ИЧУТПП “ЦДЛБЕЛ” (г. Минск), ООО “Одитек” (г. Санкт-Петербург, Россия) и внедрены в учебный процесс БГУИР.

## SUMMARY

of the dissertation thesis by Alexei V. Ivanov

### **“Feature Space Building Based on the Anthropomorphic Information Processing for Speech Recognisers in Adverse Environments”**

**Key words:** automated speech recognition, feature extraction, anthropomorphic signal processing, psychoacoustics, hearing physiology, inner ear, auditory models, active cochlea model, Meddis inner hair cell model, auditory suppression, inner hair cell response adaptation, non-linear time-variant systems, automated gain control, hidden Markov models, multilayer perceptrons.

The object under investigation comprises the mechanisms of feature extraction for automated speech recognition systems. The subject of the research comprises the speech recognition feature extraction algorithms, which are constructed on the basis of the anthropomorphic principle and are able to ensure increased performance of the speech recognisers in the adverse conditions.

The goal of the research is to develop a method to ensure increased performance of automated speech recognition systems in adverse conditions in the result of application of the anthropomorphic principle, which is particularly suitable for the mobile speech recognition systems.

A method to increase speech recogniser performance in the adverse acoustical environment conditions was developed on the basis of the anthropomorphic principle application to the feature extraction algorithm development. The resulting anthropomorphic feature extraction algorithm includes perceptual sonogram contrast enhancement algorithm, which is based on the active cochlea model and associated with it objectively observed effect of auditory suppression, as well as speech non-stationary component enhancement algorithm, which is based on the reservoir model of chemical synapse, reflecting the objectively observed phenomenon of inner hair cell response adaptation. Performed investigation proves that anthropomorphic algorithm is a more effective, than conventional approaches, tool of feature extraction in the adverse acoustical environment, which allows for more informative features that additionally are characterised by the increased degree of invariance to the adverse interference.

The results of work were used at the Minsk research institute of instrument-making (MNIPI), companies “Belsoft”, “CDLBel” (Minsk, Belarus), “Auditech” (St. Petersburg, Russia), as well as educational process at the BSUIR.

**ИВАНОВ**  
Алексей Валерьевич

**ФОРМИРОВАНИЕ ПРОСТРАНСТВА ПРИЗНАКОВ НА ОСНОВЕ  
АНТРОПОМОРФИЧЕСКОЙ ОБРАБОТКИ ИНФОРМАЦИИ В  
РАСПОЗНАВАТЕЛЯХ РЕЧИ В УСЛОВИЯХ ПРОТИВОДЕЙСТВИЯ**

Специальность 05.13.17 – Теоретические основы информатики

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

---

Подписано в печать	16.11.2004.	Формат 60x84 1/16.	Бумага офсетная.
Гарнитура «Таймс».		Печать ризографическая.	Усл. печ. л. 1,63.
Уч.-изд. л. 1,3.		Тираж 80 экз.	Заказ 676.

---

Издатель и полиграфическое исполнение: Учреждение образования  
«Белорусский государственный университет информатики и радиоэлектроники»  
Лицензия на осуществление издательской деятельности №02330/0056964 от 01.04.2004.  
Лицензия на осуществление полиграфической деятельности №02330/0133108 от 30.04.2004.  
220013, Минск, П. Бровки, 6.