

Учреждение образования
“БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ”

УКД 681.325.5:621.382.049.77

Башун Ярослав

**АНАЛИЗ И СИНТЕЗ РЕЧЕВОГО ПРОЦЕССОРА
УЛИТКОВОЙ ИМПЛАНТАЦИИ**

Специальность 05.13.05 – Элементы и устройства вычислительной техники
и систем управления

Автореферат

диссертации на соискание ученой степени кандидата технических наук

Минск 2001

Работа выполнена в Белорусском государственном университете информатики и радиоэлектроники и Белостокском техническом университете (Республика Польша)

Научный руководитель:

д.т.н., проф. А. А. Петровский

Официальные оппоненты:

д.т.н., проф. Б. М. Лобанов

к.ф.-м.н., доц. Е. И. Бовбсль

Опонирующая организация:

Научно-инженерное предприятие

информационных технологий НАН РБ

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы диссертации. При восстановлении слуха у людей со значительной степенью глухоты наибольшее распространение нашли имплантанты улитки: процессор обработки речевого сигнала и система электродов электрической стимуляции слуховых нервов. Создание импланта улитки, в частности, его ядра – речевого процессора, является актуальной задачей в области проблемно-ориентированных средств вычислительной техники. Внедрение микропроцессоров в практику проектирования речевых процессоров имплантов улитки – сложная комплексная задача, требующая координированного решения и сбалансированного продвижения по трем направлениям: синтез алгоритмов обработки речевых сигналов в реальном времени; отображение данных вычислительных алгоритмов на архитектуру процессора; построение мобильного программного обеспечения. Успех применения имплантантов улитки – это результат совместных усилий ученых из различных областей науки, включая физиологию, биоинженерию, отоларингологию и цифровую обработку сигналов. Основными направлениями повышения результатов восприятия речи в улитковой имплантации стали развитие способов проектирования электродов и методов обработки входного речевого сигнала. При этом важно, чтобы технология обработки сигнала обеспечивала возможность настройки процессора под конкретного пациента (частотную карту его улитки). Это позволит проектировать речевые процессоры инвариантные к особенностям пациентов. Одним из важных аспектов данной технологии также является разработка алгоритмов повышения качества зашумленной и реверберирующей речи для работы имплантов улитки в неблагоприятных акустических условиях.

Таким образом, разработка методов анализа и синтеза программируемого речевого процессора с препроцессором повышения разборчивости речи для имплантантов улитки человеческого уха, а также его реализация на современной микропроцессорной технике, представляют актуальную научную задачу и способствуют решению важной прикладной проблемы – протезов улитки уха человека.

Связь работы с крупными научными программами, темами. Диссертационная работа выполнена в соответствии с научно-техническими заданиями и планами работ кафедры “Электронных вычислительных средств” Белорусского государственного университета информатики и радиоэлектроники в рамках фундаментальных исследований Министерства образования Республики Беларусь: ГБЦ 97-3077 “Разработка высокопроизводительных модулей для параллельных систем цифровой обработки сигналов на базе DSP процессоров”, ГБЦ 99-3008В “Комбинированная система эхо и шума подавления в системах связи”, ГБЦ 98-3090В “Разработка метода высококачественной компрессии речи для низкой скорости передачи”, а также при финансовой поддержке Белостокского технического

университета в рамках грантов Государственного комитета науки и технологий (Республика Польша): W/П/6/96, W/П/1/98, W/R/99 "Разработка и исследование двумерной дискретной математической модели улитки уха и построение речевого процессора на ее основе".

Цель и задачи исследования. Целью работы является разработка методов анализа и синтеза программируемого речевого процессора с препроцессором повышения разборчивости речи для имплантантов улитки человеческого уха, а также его практическая реализация на современной микропроцессорной технике.

Для достижения данной цели предлагается решение следующих задач:

1. Анализ и модификация частотно- временной цифровой модели улитки уха человека с целью расчета частотной карты улитки.
2. Синтез и анализ банка цифровых кохлеарных фильтров.
3. Разработка метода повышения разборчивости речи как препроцессора для речевого процессора импланта улитки.
4. Разработка программ моделирования частотной карты улитки и инструментальных средств проектирования речевого процессора импланта улитки.
5. Аппаратная реализация речевого процессора импланта улитки на однокристалльных микропроцессорах.

Объект и предмет исследования. Объектом исследования является речевой процессор имплантов улитки человеческого уха на основе однокристалльных микропроцессоров. Предмет исследования - частотно- временная цифровая модель улитки уха человека, ее частотная карта, методы анализа и синтеза кохлеарных фильтров.

Методология и методы проведенного исследования. В ходе выполнения диссертационной работы применялись следующие методы исследования: методы цифровой обработки сигналов, психоакустики, акустики, математического моделирования, теории вероятностей и математической статистики, методы проектирования микропроцессорных систем.

Научная новизна и значимость полученных результатов.

1. На основании анализа существующих моделей имплантантов улитки человеческого уха показано, что наиболее эффективным методом построения процессоров обработки речевого сигнала для данного приложения является аппаратная эмуляция функций улитки.

2. Модифицированная цифровая модель улитки, использующая как асимметричность движения базилярной мембраны, так и представление речевого сигнала в виде конечной суммы синусоидальных сигналов. При этом доказана

возможность моделирования улитки при помощи разностного уравнения второго порядка, а результирующая модель улитки идентична банку параметрически перестраиваемых цифровых полосовых фильтров – кохлеарных фильтров.

3. Частотная карта улитки, составленная на основе математической модели, устанавливающая соотношение между центральной частотой кохлеарного фильтра и положением на базилярной мембране, а также найдена характеристика полосы пропускания данного фильтра как функция позиции фильтра на базилярной мембране.

4. Метод синтеза цифровых параметрически перестраиваемых кохлеарных фильтров высокого порядка, представляемых в виде каскада идентичных рекурсивных полосовых секций второго порядка, центральные частоты и полосы пропускания которых могут независимо друг от друга регулироваться посредством изменения соответствующих коэффициентов фильтра, что позволяет легко настраивать банк фильтров к индивидуальным особенностям пациента без проведения сложных вычислений.

5. Алгоритм повышения качества зашумленной и ревербирующей речи на основе времязависимого модуляционного фильтра в отличие от моделей, построенных на неперестраиваемых фильтрах, обеспечивает адаптацию речевого процессора импланта улитки к соответствующей акустической среде, а малые искажения, вносимые в речевой сигнал, позволяют использовать его в различных других системах мобильной связи.

6. Тестовые процедуры косвенной оценки восприятия пациентом синтезируемого возбуждения, реализованные на восстановлении речевого сигнала из конечной суммы синусоидальных сигналов, частоты которых задает кохлеарная карта, и фаз входного речевого сигнала.

Практическая значимость полученных результатов. Полученные результаты формируют теоретическую и практическую базу для синтеза и анализа речевых процессоров имплантантов улитки, способных работать в неблагоприятной акустической обстановке для пациента. При этом практически реализованы следующие устройства и инструментальные программы.

Одноканальная процедура повышения разборчивости речи как препроцессор речевого процессора импланта улитки инвариантная к акустической среде. Это обеспечивает возможность функционирования имплантанта при наличии шумов и реверберации (соотношение сигнал-шум 0 дБ, время реверберации до 3 сек.). Кроме того, предложенная процедура может быть составной частью мобильных систем автоматического распознавания речи.

Аппаратная реализация речевого процессора имплантантов улитки: банк кохлеарных фильтров с тестовыми процедурами косвенной оценки восприятия

пациентом синтезируемого возбуждения на стековом процессоре DOFIN-1620 (производство ПО "Интеграл") может быть базой для построения миниатюрной модели речевого процессора для ношения за ухом. Речевой процессор с препроцессором повышения разборчивости речи на основе быстрого микропроцессора ЦПОС ADSP-21061x SHARC фирмы Analog Devices с плавающей запятой, а также инструментальной программой в среде Matlab на персональном компьютере, выполняющей моделирование процесса оценки восприятия пациентом синтезируемого возбуждения, могут найти применение в качестве лабораторной исследовательской системы проблемы имплантов, а также мониторинговой системы для проектируемых новых речевых процессоров имплантов улитки человека. Представленные макетные образцы речевых процессоров для имплантов улитки способны работать в комплексе с существующими устройствами стимуляции слухового нерва, реализующими стратегии непрерывного квантования с чередованием (CIS).

Базовые результаты диссертационной работы переданы ОАО Минский приборостроительный завод (г.Минск), внедрены в учебный процесс специальности Т08.02 "Проектирование и технология электронных вычислительных средств" (БГУИР), а также в Институте информатики Белостокского технического университета.

Основные положения диссертации, выносимые на защиту.

1. Модифицированная цифровая модель улитки, использующая как асимметричность движения базилярной мембраны, так и представленные речевого сигнала в виде конечной суммы синусоидальных сигналов.
2. Метод синтеза и анализа цифровых параметрически перестраиваемых кохлеарных фильтров высокого порядка, представляемых в виде каскада идентичных рекурсивных полосовых секций второго порядка, центральные частоты и полосы пропускания которых могут независимо друг от друга регулироваться посредством изменения соответствующих коэффициентов фильтра.
3. Алгоритм повышения качества зашумленной и реверберирующей речи на основе вряязависимого модуляционного фильтра в отличие от моделей, построенных на перестраиваемых фильтрах, обеспечивает адаптацию речевого процессора импланта улитки к соответствующей акустической среде.
4. Речевой процессор с препроцессором повышения разборчивости речи на базе однокристалльных микропроцессоров ЦПОС и инструментальная программа в среде Matlab на персональном компьютере, выполняющей моделирование процесса оценки восприятия пациентом синтезируемого возбуждения.

Личный вклад соискателя. Все новые результаты, изложенные в диссертационной работе, получены автором самостоятельно. В публикациях с

соавторами вклад соискателя определяется рамками излагаемых в диссертации результатов.

Апробация результатов диссертации. Результаты, полученные в ходе выполнения исследований, докладывались и обсуждались на: the International Conference on Structural and Biomedical Acoustics'97 (Zakopane, Poland, 1997); the 4th International Workshop on Systems, Signals and Image Processing (IWSSIP97) (Poznan, Poland, 1997); the 11th International Conference Systems for Automation of Engineering and Research, SEAR'97 (Varna - St. Konstantine, Bulgaria, 1997); the International Conference Modern Methods of Digital Signal Processing for Measure, Inspection, Diagnosis and Control Systems (OC-98) (Minsk, 1998); the 1st and 3rd Russian International Conference Digital Signal Processing and its Applications (Moscow, 1998, 2000); the 2nd European Signal Processing Society Conference "DSP For Multimedia Communications and Services" (FCMCS'99) (Kracow, Poland, 1999); the International Conference Information Technologies for Education, Science and Business, ITESB99 (Minsk, 1999); the 8th International Symposium on Sound Engineering and Mastering, ISSEM'99 (Gdansk, Poland 1999); the international conference SATFERA'99 (Mittweida, Germany, 1999); The PTSK99 conference (Bialystok, Poland, 1999); the 10th European Signal Processing conference (EUSIPCO'2000) (Tampere, Finland, 2000).

Опубликованность результатов. По тематике диссертационной работы опубликовано 19 печатных работ, включая 2 тезиса докладов на научных конференциях и 17 статей в журналах и материалах международных научных конференций и симпозиумов. Суммарный объем публикаций составляет около 130 страниц.

Структура и объем диссертации. Диссертационная работа состоит из введения, пяти глав, заключения и трех приложений. Она содержит 100 страниц основного текста, 30 рисунков на 30 страницах, 7 таблиц на 4 страницах, три приложения на 10 страницах, в списке использованных источников на 7 страницах представлено 129 наименований.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** показана актуальность темы диссертационной работы, формулируются цели и проблемы исследования.

В главе 1 дан детальный анализ принципов функционирования имплантов улитки человеческого уха. При этом функцией речевого процессора в импланте является создание искусственного аналога нормально функционирующей улитки человека. Анализ стратегий обработки речи в имплантах, ориентированных на получение характеристик сигнала, например, как стратегия F0/F1/F2, показал, что возможны ошибки восприятия, особенно в тех случаях, когда речевой сигнал искажается шумом. Поэтому имеет смысл сосредоточиться на стратегиях, выделяющих

информацию о форме сигнала, которые в большей мере соответствуют способу функционирования улитки. Рассмотрены существующие решения в области создания многоканальных имплантов улитки, отдельных их составных частей и показаны новые направления развития обработки речевых сигналов в имплантах: программируемость обработки речи; преобработка сигнала речи с целью подавления неблагоприятных акустических шумов окружающей среды. Предложен новый порядок следования этапов обработки речевого сигнала и соответствующими компоненты системы имплантов улитки, позволяющие адаптироваться к изменениям акустической обстановки и использовать тестовые процедуры косвенной оценки восприятия пациентом синтезируемого возбуждения.

Глава 2 является базовой для последующих частей диссертационной работы. Была модифицирована дискретная частотно-временная модель улитки, использующая как асимметричность движения базилярной мембраны, так и предположение, что речевой сигнал может быть представлен в виде конечной суммы синусоидальных сигналов. Модель базируется на том, что улитка может быть описана как двумерная прямоугольная камера (рис. 1), наполненная вязкой, несжимаемой жидкостью.

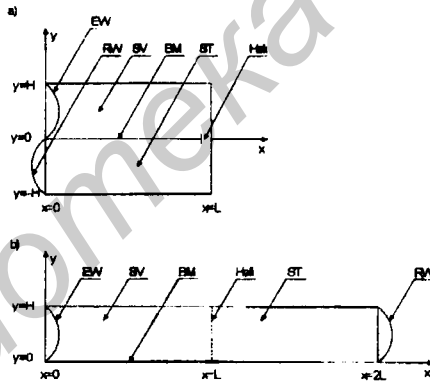


Рис. 1

В эллиптическом окне (EW) при $x=0$, предполагается постоянная скорость. Движение эллиптического окна и круглого окна (RW) рассматриваются как совпадающие (по принципу поршня – движение стремечка вызывает аналогичное смещение круглого окна), что показано на рис. 1a,b при помощи кривых. Эффект геликотермы (Hel) на всей длине базилярной мембраны (BM) не учитывается. Длина базилярной мембраны считается равной L . На основании допущения, сформулированного выше, рис. 1a может быть заменен рис. 1b. Асимметричность движения базилярной мембраны по вертикали позволяет утверждать, что давление жидкости в геликотерме равно нулю, т.е. при $p(x, y, t)|_{x=L} = 0$, где y – это вертикальная пространственная координата по отношению к ВМ. Интегрируя градиент вязкости по

длине базилярной мембраны по граничным условиям получается уравнение, описывающее распределение давления жидкости по ВМ. С другой стороны, известно соотношение между давлением на базилярную мембрану и ее вертикальным смещением. Следовательно, модель улитки в непрерывном пространстве и времени определяется в следующей форме:

$$k(x)z(x,t) + r(x)\frac{\partial z(x,t)}{\partial t} + m(x)\frac{\partial^2 z(x,t)}{\partial t^2} = 2\rho \left[\frac{\partial u_s}{\partial t} + \gamma \left(\frac{\pi}{H} \right)^2 u_s \right] \left[x - L + \frac{8L}{H^2} \sum_{i=1}^{\infty} \frac{\sin\left(\frac{(2i-1)\pi x}{2L}\right)}{\lambda_i^2 (2i-1)} \right],$$

где ρ - плотность жидкости; γ - вязкость жидкости; L - длина базилярной мембраны; H - половина высоты камеры; p - давление жидкости; u_s - скорость перемещения стремечка $u_s = A \sin(\omega t + \varphi) \sin\left(\frac{\pi y}{H}\right)$; x - представление пространственной координаты (от основания к вершине по базилярной мембране ВМ); t - координата времени; $\lambda_i^2 = [(2i-1)\pi/2L]^2 + (\pi/H)^2$; $z(x,t)$ - вертикальное смещение базилярной мембраны; $k(x)$ - жесткость базилярной мембраны; $r(x)$ - затухание базилярной мембраны; $m(x)$ - масса базилярной мембраны.

Проведя дискретизацию пространственных координат модели $x = k\Delta x$, $0 \leq k \leq M$, $\Delta x = L/M$, M количество сегментов ВМ после дискретизации, и выполнив билинейное z -преобразование, получается дискретная модель улитки, которую для краткости назовем SDCM-модель (Second Order Difference Cochlea model):

$$y_k(n) + b_{1k}y_k(n-1) + b_{2k}y_k(n-2) = A_k a_{0k} [u_s(n) - u_s(n-2)],$$

где $y_k(n)$ - перемещение (пучность) ВМ в позиции x_k , а параметры b_{1k} , b_{2k} , A_k и a_{0k} определяются физическими свойствами ВМ в позиции x_k и изменяются вдоль ВМ. Здесь и далее для сокращенной записи времени $n\Delta t$ будет применено обозначение n . Передаточная функция SDCM-модели следующая:

$$\hat{H}_k(z) = A_k \frac{a_{0k}(1-z^{-2})}{1 + b_{1k}z^{-1} + b_{2k}z^{-2}}, \quad 0 \leq k \leq M.$$

SDCM-модель есть математическая модель вибрации базилярной мембраны улитки человеческого уха и представляется в форме банка цифровых полосовых фильтров второго порядка, которые назовем кохлеарными фильтрами. Однако, данный цифровой фильтр будет иметь добротность большую, чем добротность фильтра прототипа, вследствие нелинейности, обусловленной билинейным z преобразованием. С учетом частотной коррекции, передаточная функция k -го

цифрового кохлеарного фильтра второго порядка задается выражением:

$$H_k(z) = a_{ok} \frac{1 - z^{-2}}{1 + (a_{ok} - 1)g_k z^{-1} + (1 - 2a_{ok})z^{-2}},$$

где коэффициенты a_{ok} и g_k определяются только шириной полосы пропускания $\Delta\omega_k$ и центральной частотой ω_{ok} соответственно:

$$a_{ok} = \frac{\Delta\omega_k \Delta t}{2 + \Delta\omega_k \Delta t}, \quad g_k = 2 \cos \omega_{ok} \Delta t.$$

Система разностных уравнений

$$\begin{cases} w_k(n) = x_k(n) - g_k v_k(n-1) + 2v_k(n-2), \\ v_k(n) = a_{ok} w_k(n) + g_k v_k(n-1) - v_k(n-2), \\ y_k(n) = v_k(n) - v_k(n-2), \quad 0 \leq k \leq M \end{cases}$$

представляет собой параметрически перестраиваемую SDCM -модель улитки. Свойство параметрической перестройки SDCM -модели позволяет получить целый ряд очень полезных с практической точки зрения характеристик имплантов улитки: гибкая перестройка частотной карты улитки; простая регулировка кохлеарных фильтров в процессе наладки импланта улитки и его настройки на пациента; универсальность системы имплантов на базе данной модели.

В соответствии с SDCM - моделью определяется частотная карта улитки, т.е. для каждой k -ой позиции на ВМ рассчитываются кохлеарная карта - шкала критических частот (центральные частоты кохлеарных фильтров) и соответствующие им частотные полосы (полосы пропускания кохлеарных фильтров):

$$\omega_{ok} = \frac{1}{\Delta t} \arccos \frac{-b_{1k}}{1 + b_{2k}}, \quad \Delta\omega_k = \frac{2(1 - b_{2k})}{\Delta t(1 + b_{2k})}. \quad \text{На рис. 2 и рис. 3 соответственно}$$

представлены кохлеарная карта (центральные частоты) и полосы пропускания кохлеарных фильтров. На рис. 2 горизонтальная координата определяет центральную частоту кохлеарного фильтра, а вертикальная координата - расстояние от вершины ВМ в мм. Пунктирная линия показывает зависимость частота-положение, полученное по формуле Гринвуда. Из рис. 2 видно, что зависимость между частотой возбуждения и положением, претерпевшим наибольшее смещение, не является однозначно линейной. Другими словами, для кохлеарных фильтров равномерно распределенных по ВМ, вследствие нелинейности кохлеарной карты, они будут расположены неравномерно в частотной области. Обычно кохлеарная карта строится на основании экспериментальных данных. В нашем же случае она получается прямо из математической модели улитки, причем результаты в достаточной степени совпадают с экспериментальными данными. Характеристика полосы пропускания (рис. 3) определяется как соотношение между полосой пропускания по уровню 3 дБ кохлеарного фильтра и его положениями на ВМ, претерпевшими наибольшие

смещения. Штриховая линия показывает критические полосы из Цвикера и Фастла, пунктирная - экспериментальные данные, представленные Грипвудом.

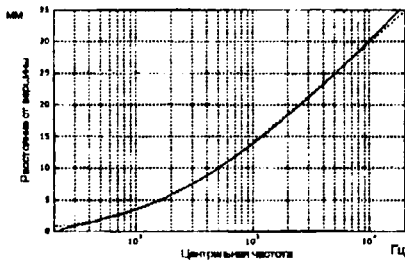


Рис. 2

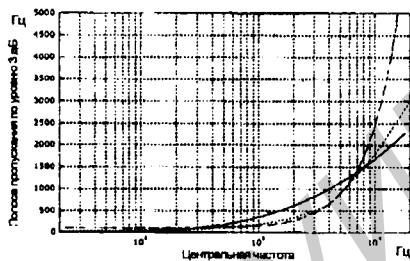


Рис. 3

Новая кохлеарная карта и характеристика полосы пропускания по уровню 3 дБ увеличивают точность реализации функций улитки и позволяют выполнить их количественное обоснование. Благодаря тому, что структура модели проста и обеспечивает получение параметров близких к действительности, она может быть с успехом использована в программируемых речевых процессорах имплантов.

Разработан метод синтеза пифровых параметрически перестраиваемых кохлеарных фильтров высокого порядка, что обусловлено недостаточной избирательностью частотной характеристики фильтра по уровню пдБ. Подобные фильтры строятся в виде каскада идентичных рекурсивных полосовых секций второго порядка. Результатом процесса синтеза фильтра являются коэффициенты секций каскада и количество каскадов.

В главе 3 разработан алгоритм (препроцессор) повышения качества зашумленной и реверберирующей речи на основе времязависимого модуляционного фильтра, который позволяет использовать имплантанты улитки в неблагоприятных акустических условиях. В работах по акустике показана прямая зависимость между плотностью информации (слоговой разборчивостью речи) и энергией в полосе определенных частот модуляции. Причем максимальная ширина полосы частот модуляции равна от 1 до (12 – 16) Гц. Здесь концентрируется наибольший объем информации, который соответствует области речевой информации. На основании изложенного выше, предлагается идея слежения во времени за изменением ширины данной полосы частот модуляции и соответствующей перестройке полосы пропускания модуляционного фильтра, выходной сигнал которого пропорционален оценке коэффициента модуляции чистой речи. Другими словами, модуляционный фильтр в каждый момент времени осуществляет фильтрацию в области модуляционных частот, где концентрируется максимальная энергия сигнала огибающей, т.е. данный фильтр является фильтром со времязависимыми параметрами. Предлагаемая система не нуждается в опорном шумовом сигнале. Это

важно, поскольку реализация дополнительного канала в слуховых аппаратах практически невозможна. Малые искажения, вносимые в речевой сигнал, позволяют использовать данный алгоритм повышения разборчивости речи в различных системах мобильной связи.

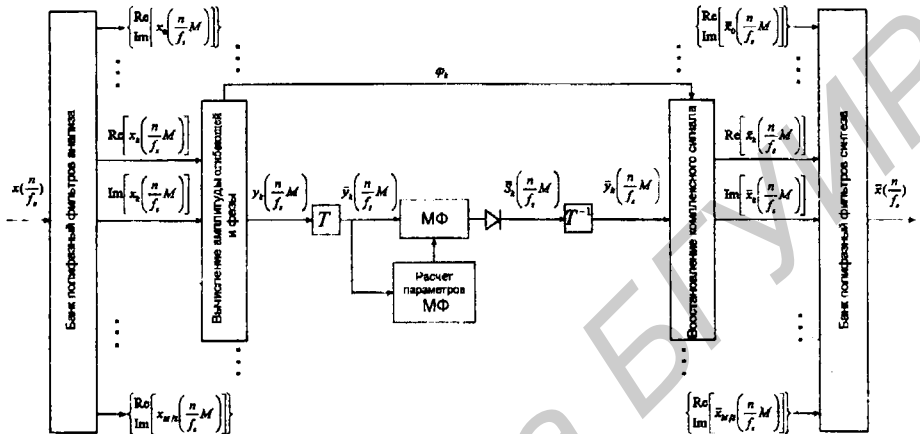


Рис. 4

В качестве модуляционного фильтра используется параметрически перестраиваемый кохлеарный фильтр, у которого в каждом частотном канале k во времени отслеживаются контур изменения коэффициента a_{0k} , являющего функцией полосы пропускания фильтра $\Delta\omega_k$, а центральная частота ω_0 постоянна:

Согласно схеме (см. рис. 4) метода повышения разборчивости речи на основе фильтрации реверберации и шума времязависимым модуляционным фильтром алгоритм повышения разборчивости речи процессора улитковой имплантации состоит из выполнения следующих шагов:

Шаг 1) речевой сигнал $x(n/f_s)$ разделяется на M частотных полос с помощью полифазного банка цифровых фильтров анализа со следующими нормализованными центральными частотами: $(2k+1)\pi/2M$, причем каждый канал характеризуется своей импульсной характеристикой $p_k(n)$, а также частотной характеристикой $P_k(f)$ для $0 \leq k \leq M/2$.

Шаг 2) в каждом канале k ($0 \leq k \leq M/2$) выполнить:

- вычислить огибающую сигнала $x_k(nM/f_s)$:

$$y_k(nM/f_s) = \sqrt{\text{Re}^2[x_k(nM/f_s)] + \text{Im}^2[x_k(nM/f_s)]}, \quad \text{а также фазу сигнала}$$

$$x_k(nM/f_s): \varphi_k = -\arctg \frac{\text{Im}[x_k(nM/f_s)]}{\text{Re}[x_k(nM/f_s)]};$$

- выполнить трансформацию амплитуды огибающей спектра $y_k(nM/f_s)$ путем нелинейного статического сжатия

$$\bar{y}_k(nM/f_s) = \ln(1 + 1000|y_k(nM/f_s)|);$$

- фильтрация огибающей амплитуды спектра $\bar{y}_k(nM/f_s)$ времязависимым модуля-ционным фильтром;
- выполнить трансформацию амплитуды огибающей спектра $\bar{s}_k(nM/f_s)$ в линейный масштаб путем обратного нелинейного преобразования

$$\hat{y}_k(nM/f_s) = \frac{e^{\bar{y}_k(nM/f_s)} - 1}{1000};$$

- выполнить переход из области частоты модуляции в частотно-временную область

$$\text{Re}[\bar{x}_k(nM/f_s)] = \hat{y}_k(nM/f_s) \cos \varphi_k, \quad \text{Im}[\bar{x}_k(nM/f_s)] = \hat{y}_k(nM/f_s) \sin \varphi_k.$$

Процедура слежения за контуром изменения коэффициента $a_{ок}$:

- параллельно с работой модуляционного фильтра определять:

1) оценку среднеквадратического значения S_k компоненты речи $\hat{s}_k(nM/f_s)$;

2) оценку среднеквадратического значения Y_k сигнала $\bar{y}_k(nM/f_s)$ (оценки средне-квадратических значений вычисляются на основе экспоненциального усреднения: $\bar{Q}_i = \bar{Q}_{i-1} + \alpha(Q_i - \bar{Q}_{i-1})$, где $\alpha < 1$, \bar{Q}_i, \bar{Q}_{i-1} - средние оценки величины Q за i и $i-1$ шагов работы алгоритма экспоненциального усреднения соответственно, Q_i - текущая оценка величины Q (постоянная времени данного алгоритма определяет величину задержки в канале модуляционного фильтра);

3) значение действительной весовой функции H_k ; 4) коэффициент модуляционного фильтра $a_{ок}$;

- изменить параметр $a_{ок}$, задающий полосу пропускания модуляционного фильтра.

Шаг 3) восстановить речевой сигнал $\bar{x}_k(nM/f_s)$ посредством банка полифазных фильтров синтеза.

Перфективная реконструкция входного речевого сигнала $x_k(nM/f_s) = \bar{x}_k((n-n_0)M/f_s)$, где n_0 - алгоритмическая задержка алгоритма повышения разборчивости речи, обуславливается перфективной конструкцией банков

анализирующих и синтезирующих фильтров.

Таким образом, основное отличие данного подхода от известной системы RASTA заключается в том, что параметры модуляционного фильтра напрямую зависят от свойств акустической среды (шумы и/или реверберация). Причем здесь отслеживается изменение ее характеристик во времени.

На рис. 5-7 показана работа процедуры повышения разборчивости речи.

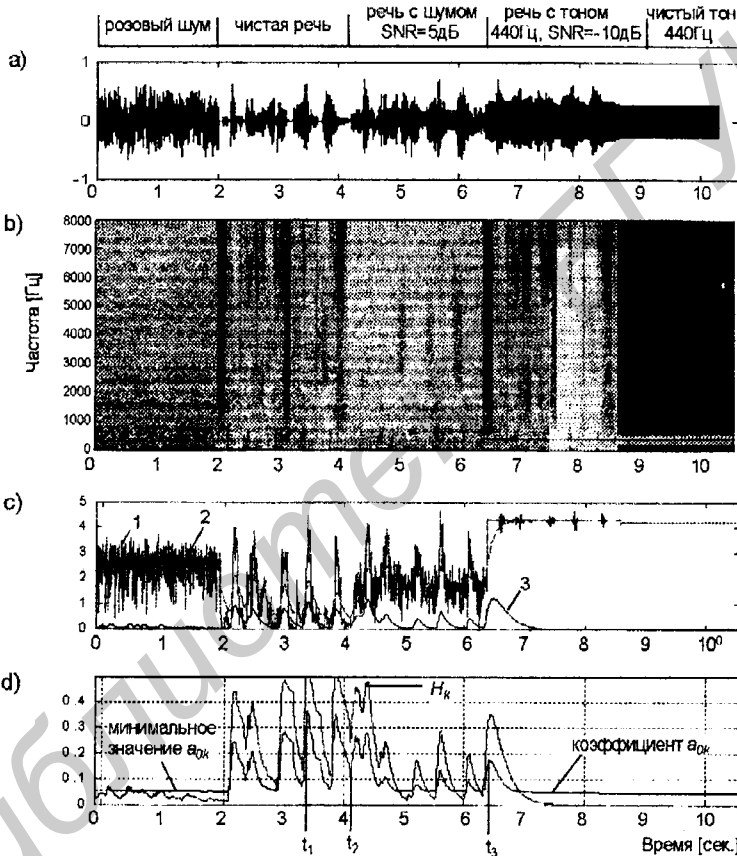


Рис. 5. Работа процедуры слежения за контуром изменения коэффициента a_{0k} модуляционного фильтра: а) входной тестовый сигнал; б) спектрограмма входного тестового сигнала; в) процесс слежения за огибающей $\bar{y}_k(nM/f_s)$ (кривая 1), а также за оценками Y_k (кривая 2) и S_k (кривая 3); д) контур изменения коэффициента a_{0k} .

Для моментов времени t_1 , t_2 и t_3 на рис. 7 приведены амплитудно-частотные характеристики и импульсные характеристики модуляционного фильтра. На рис. 7

также показаны предельные амплитудно-частотные характеристики (для максимальной и минимальной полос пропускания) модуляционного фильтра.

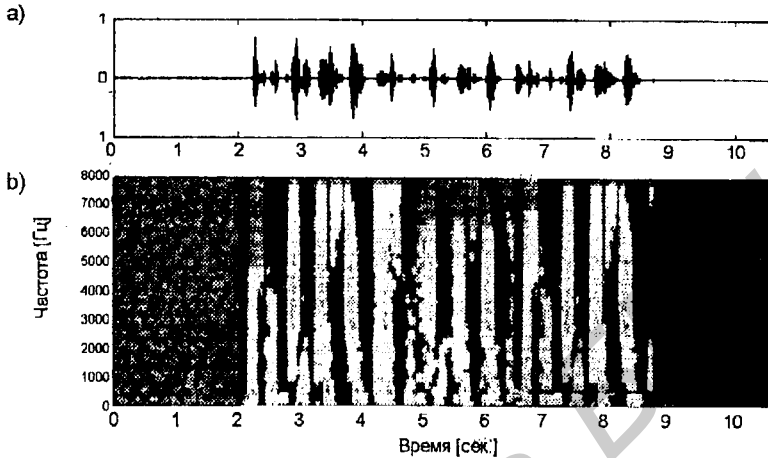


Рис. 6. Реконструированный речевой тестовый сигнал (а) и его спектрограмма (б) на выходе процедуры повышения разборчивости речи

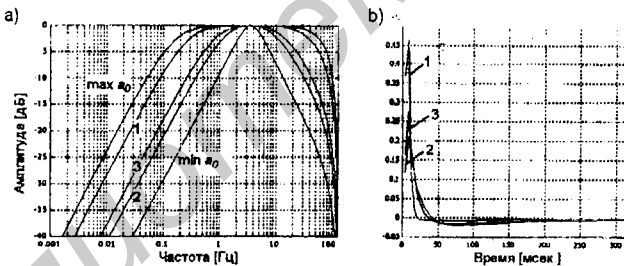


Рис. 7. Амплитудно-частотные характеристики (а) и импульсные характеристики (б) модуляционного фильтра для моментов времени t_1 , t_2 и t_3 .

В главе 4 представлена аппаратная платформа речевого процессора с препроцессором повышения разборчивости речи для имплантов улитки. Проведен анализ вычислительной сложности реализации речевого процессора и системы повышения разборчивости речи. Процессор DOFIN-1620 (ППО "Интеграл") со стекковой организацией был использован для построения миниатюрного речевого процессора для ношения за ухом. Показана реализация речевого процессора с препроцессором повышения разборчивости речи на ЦПОС ADSP-21061x SHARC производства фирмы Analog Devices. Частотная характеристика банка перестраиваемых кохлеарных фильтров приведена на рис. 8. На рис. 9 представлена

блок-схема речевого процессора.

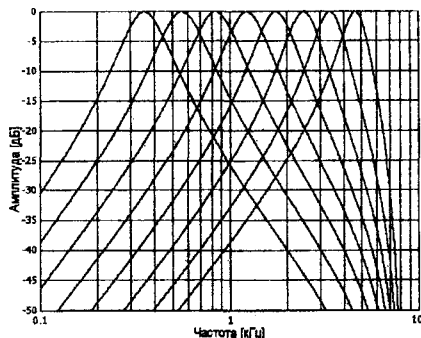


Рис. 8

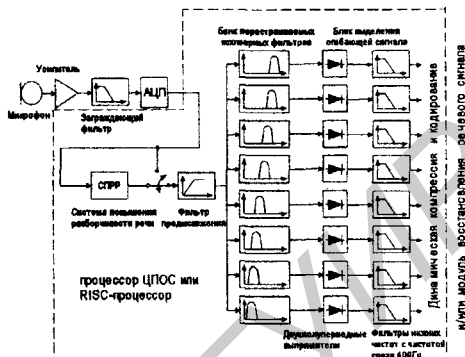


Рис. 9

В ходе анализа возможностей реализации речевого процессора на арифметике с фиксированной запятой детально рассмотрены свойства шума и влияние квантования коэффициентов на центральную частоту, полосу пропускания и фазовую характеристику кохлеарных фильтров, проанализированы колебания предельного цикла. Предложен метод вероятностного округления коэффициентов фильтра и результатов умножения. Проведено его сравнение с обычным округлением и усечением. Исследования установили отсутствие шумов предельного цикла при использовании вероятностного округления. Анализ шумовых характеристик банка кохлеарных фильтров показал, что 16 разрядного процессора достаточно для обеспечения динамического диапазона не более 60 дБ, однако если использовать вероятностное округление, то динамика фильтров может быть увеличена до 80 дБ.

Речевой процессор улитковой имплантации имеет следующие технические характеристики: программно-управляемый процессор банка кохлеарных фильтров; банк кохлеарных фильтров (поддержка коррекции частотной карты улитки; адаптация к пациенту с имплантированными электродами; 8 четвертого порядка кохлеарных фильтров в банке); обеспечивается поддержка CIS стратегии; частотный диапазон обрабатываемых сигналов 100-8000 Гц; частота дискретизации 16000 Гц; питание 5В DC; последовательный интерфейс RS232.

Глава 5 является экспериментальной частью диссертационной работы. Тестирование и симуляция функционирования имплантата улитки были выполнены при помощи программы, написанной в среде MATLAB. Программа позволяет проводить синтез процессоров обработки речевого сигнала для имплантатов улитки и тестирование полученных устройств на основе реконструкции речевого сигнала. На начальных стадиях проектирования речевого процессора имплантов улитки такая методика обеспечивает проведение большей части тестов системы без привлечения пациентов. Тестирование речевого процессора проводилось с использованием

процедуры восстановления речи. Для оценки качества синтезированного процессора были выполнены специальные тесты по распознаванию гласных, согласных звуков и предложений. Были проведены эксперименты для определения оптимального количества каналов. Субъективная оценка показала, что для разработанной модели улитки восьми каналов достаточно.

Препроцессор повышения разборчивости речи был тестирован на зашумленных речевых сигналах с различными значениями отношения сигнал-шум (10дБ, 0дБ, с аддитивными речевыми сигналами других людей, т.е. так называемый “cocktail party effect”) и сигналах с реверберацией (до 3 сек.). При этом для объективной оценки разборчивости речевого сигнала применялась компьютерная программа распознавания речи, не зависящая от голоса диктора. Система показала 30 процентное улучшение результатов распознавания. Таким образом, препроцессор повышения разборчивости речи увеличивает возможности имплантов улитки при их использовании в аддитивных и реверберационных шумах.

Приложения содержат характеристики полифазной структуры банков анализирующих и синтезирующих фильтров препроцессора повышения разборчивости речи; коды программ инструментальной системы, моделирующей функционирование речевого процессора и процесса оценки восприятия пациентом синтезируемого возбуждения (программы разработаны для среды Matlab 5.3); акты внедрения результатов диссертационной работы.

ЗАКЛЮЧЕНИЕ

Основные научные результаты диссертационной работы могут быть сформулированы следующим образом:

1. На основании анализа существующих моделей имплантов улитки человеческого уха показано, что наиболее эффективным методом построения процессоров обработки речевого сигнала является аппаратная эмуляция функций улитки, предложенный же новый порядок следования этапов обработки речевого сигнала и соответствующие компоненты системы имплантов улитки, позволяют адаптироваться к изменениям акустической обстановки и косвенно оценивать восприятие пациентом синтезируемого возбуждения [17].
2. Модифицирована цифровая модель улитки, использующая как асимметричность движения ВМ, так и представление речевого сигнала в виде суммы конечных синусоидальных последовательностей [9, 12]. При этом доказана возможность моделирования улитки при помощи разностного уравнения второго порядка, а результирующая модель улитки идентична банку параметрически

перестраиваемых цифровых полосовых фильтров – кохлеарных фильтров высокого порядка [13, 17], представляемых в виде каскада идентичных рекурсивных полосовых секций второго порядка, центральные частоты и полосы пропускания которых могут независимо друг от друга регулироваться посредством изменения соответствующего коэффициента фильтра [16], что позволяет легко настраивать банк фильтров к индивидуальным особенностям пациента.

3. Разработана частотная карта улитки на основе математической модели, которая устанавливает соотношение между центральной частотой кохлеарного фильтра и положением его на базилярной мембране, а также найдена характеристика полосы пропускания данного фильтра как функция позиции фильтра на базилярной мембране [12, 19].
4. Новый алгоритм повышения качества запумленной и реверберирующей речи на основе времязависимого модуляционного фильтра, который в отличие от моделей, построенных на неперестраиваемых фильтрах, обеспечивает адаптацию речевого процессора импланта улитки к соответствующей акустической среде (соотношение сигнал-шум 0 дБ, время реверберации до 3 сек.), а малые искажения, вносимые в речевой сигнал, позволяют использовать его в различных других системах мобильной связи [14-17].
5. Разработаны тестовые процедуры косвенной оценки восприятия пациентом синтезируемого возбуждения, реализованные на восстановлении речевого сигнала из конечной суммы синусоидальных сигналов, частоты которых задает кохлеарная карта, и фаз входного речевого сигнала [12].
6. Осуществлена аппаратная реализация речевых процессоров имплантантов улитки: банк кохлеарных фильтров с тестовыми процедурами косвенной оценки восприятия пациентом синтезируемого возбуждения на стековом процессоре DOFIN-1620 (производство ПО “Интеграл”) [8], речевой процессор с препроцессором повышения разборчивости речи на основе ЦИОС ADSP-21061x SHARC фирмы Analog Devices с плавающей запятой и инструментальная программа в среде Matlab на персональном компьютере [9, 11, 12], выполняющая моделирование процесса оценки восприятия пациентом синтезируемого возбуждения.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ ПО ТЕМЕ ДИССЕРТАЦИИ

Статьи в журналах и научных сборниках

1. A. Piotrowski, J. Baszun, Pomiar prędkości przepływu krwi z zastosowaniem karty DSP. - Elektronizacja 10/95, Poland, pp. 4-7

2. A. Piotrowski, J. Baszun, V. Serkov, Stacja robocza do syntezy optymalnych FIR filtrów z liniowa faza i ich testowania w czasie rzeczywistym. - Elektronizacja 6/96, Poland, pp. 6-9
3. А.Петровский, В.Серков, А.Иванов, Я.Башун, Психоакустика и обработка речи // Республиканский межведомственный сборник научных трудов, БГУИР, вып. 23, Минск, Беларусь, 1999. – с.110-119
4. A. Petrovsky, A. Ivanov, J. Baszun, 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, Germany, No. 3, 1999, pp. 321-329
5. J. Baszun, A. Petrovsky, Single channel speech enhancement system for voice remote control devices. - Journal of the University of Applied Sciences, Mittweida, Germany, No. 9-E, Digital Broadcast, 2000, pp. 57-65
6. A. Petrovsky, J. Baszun, Design of digital tunable bandpass filters: synthesis and applications. - Kwartalnik Elektroniki i Telekomunikacji, PAN, Poland, vol. 1, no. 1, 2001, pp.5-25
7. A. Petrovsky, J. Baszun, Finite word length effects in tunable recursive bandpass filters. - Kwartalnik Elektroniki i Telekomunikacji, PAN, Poland, vol. 1, no. 1, 2001, pp.26-43

Материалы конференций

8. J. Baszun, A. Petrovsky, Implementacja modelu ślimaka ucha za pomocą banku przestrajanych filtrów cyfrowych na FORTH-procesorze// Structural and Biomedical Acoustics'97 Conference, Proceedings of the Acoustical Methods in Biomedical Engineering, Zakopane, Poland, 21-25 April 1997, pp. 55-60
9. J. Baszun, A. Petrovsky, Evaluation digital signal processing system for the cochlear implants based on the stack machine// Proceedings of the 4th International Workshop on Systems, Signals and Image Processing, IWSSIP'97, Poznan, Poland, 28-29 May 1997, pp. 129-132
10. J. Baszun, A. Petrovsky, The synthesis of 1/b octave digital filters and real-time spectral analyzer of vibroacoustic signals based on it// Proceedings of the 11th International Conference Systems for Automation of Engineering and Research, SAER'97, Varna - St. Konstantine resort, Bulgaria, 20-21 September 1997, pp. 112-116
11. J. Baszun, A. Petrovsky, Sound processor applicable to hearing prostheses// Proceedings of the International Conference Modern Methods of Digital Signal Processing for Measure, Inspection, Diagnosis and Control Systems, OC-98, Minsk, Belarus, 24-27 June 1998, pp. 170-173

12. J. Baszun, A. Petrovsky, Design of speech processor for cochlear implants and method of its testing// Proceedings of the 1st Russian International Conference Digital Signal Processing and its Applications, DSPA '98, Moscow, Russia, 30 June - 3 July 1998, Vol. I, pp.181-189
13. A. Bielawski, J. Baszun, A. Petrovsky, Cochlear spaced filter bank using allpass frequency transformation// Proceedings of the EURASIP Conference DSP For Multimedia Communications and Services, ECMCS'99, Krakow, Poland, 24-26 June, 1999, 4 pages
14. J. Baszun, A. Petrovsky, Speech enhancement system for hearing prosthesis and automatic speech recognition// Proceedings of the International Conference Information Technologies for Education, Science and Business, ITESB'99, Minsk, Belarus, June 24-25, 1999, pp. 50-54
15. J. Baszun, A. Petrovsky, Speech processor applicable to hearing prosthesis in noise and reverberant environment// Proceedings of the 8th International Symposium on Sound Engineering and Mastering, ISSEM'99, Gdansk, Poland, September 09-11, 1999, pp. 37-42
16. J. Baszun, A. Petrovsky, The simulations results of cochlear prosthesis with speech enhancement system: Speech intelligibility// Proceedings of the Szoste Warsztaty Naukowe PTSK'99, Symulacja w badaniach i rozwoju, Bialystok, Poland, 25-27 August 1999, pp. 294-302
17. J. Baszun, A. Petrovsky, Flexible cochlear system based on digital model of cochlea: structure, algorithms and testing// Proceedings of the 10th European Signal Processing Conference (EUSIPCO 2000), Tampere, Finland, 14-17 September 2000, vol. III, pp. 1863-1866

Тезисы докладов научных конференций

18. В.Серков, Я.Башун, А.Петровский, Рабочая станция оптимального синтеза цифровых КИХ фильтров// Труды научной конференции "Современные проблемы в радиотехнике, электронике и телекоммуникации", БГУИР, Минск, Беларусь, 4-5 Мая 1995. – с. 192-193
19. K. Bielawski, J. Baszun, A. Petrovsky, Flexible nonuniform filter banks for cochlear implants and joint system for acoustic echo and noise control communication devices // The Proc. of the 3rd Intern. confer. "Digital signal processing and its applications", DSPA'2000, vol. 1, Moscow, Russia, 2000, pp.190-192

У Baszun

РЭЗЮМЭ

дысертцыі Башуна Яраслава

“Аналіз і сінтэз працэсара мовы слімаковай імплантацыі”

Ключавыя словы: базілярная мембрана, кахлеарная карта, частотная карта слімака, слімаковая імплантацыя, працэсар мовы, лічбовы фільтр, лічбовы працэсар апрацоўкі сігналаў.

Аб’ектам даследвання з’яўляецца працэсар імпланта слімака чалавечага вуха на базе монакрыштальнага мікрапрацэсара. Прадмет даследвання – частотна-часавая дыскрэтная мадэль слімака вуха чалавека, яго частотная карта, метады аналізу і сінтэзу кахлеарных фільтраў.

Мэтай працы з’яўляецца распрацоўка метадаў аналізу і сінтэзу праграмаванага працэсара мовы з прэпрацэсарам павышэння выразнасці мовы для імплантаў слімака вуха чалавека, а таксама яго практычная рэалізацыя на сучаснай мікрапрацэсарнай тэхніцы.

Распрацаваны метады аналізу і сінтэзу праграмаванага працэсара мовы для імплантаў слімака з прэпрацэсарам павышэння выразнасці мовы на падставе двухвымернай дыскрэтнай мадэлі слімака – банка параметрычных рэгуляваных лічбовых фільтраў, складзенных як каскад ідэнтычных рэкурсіўных палосавых секцый другога парадку, цэнтральныя частоты і палосы прапускання якіх могуць рэгулявацца незалежна адзін ад аднаго шляхам замены адпаведных каэфіцыентаў фільтра, якая суадносіць залежнасць паміж цэнтральнымі частотамі, палосамі прапускання кахлеарных фільтраў і іх месцам на базілярнай мембране, а таксама на падставе залежнага ад часу параметрычна рэгуляванага мадуляцыйнага фільтра для паліпшэння якасці мовы з шумама і рэверберацыяй, якая адрозніваецца тым, што забяспечваецца малы ўзровень скажэння, з прычыны высокага дынамічнага дыпазону банка фільтраў, абумоўленага выкарыстаннем імавернаснага квантавання каэфіцыентаў фільтраў, адаптацыяй працэсара мовы імпланта слімака к асаблівасцям пацыента і навакольнаму акустычнаму асяроддзю: шумам і рэверберацыі (суадносіны сігнала да шуму 0дБ, час рэверберацыі да 3 сек.).

Вынікі даследаванняў перададны ААТ “Мінскі прыборабудаўнічы завод” г.Мінск, выкарыстоўваюцца ў навучальным працэсе ў БДУІР, а таксама ў Інстытуце інфарматыкі Беларускага тэхнічнага ўніверсітэта (Рэспубліка Польша).

РЕЗЮМЕ

Диссертации Башуна Ярослава

“Анализ и синтез речевого процессора улитковой имплантации”

Ключевые слова: базилярная мембрана, кохлеарная карта, частотная карта улитки, улитковая имплантация, речевой процессор, цифровой фильтр, цифровой процессор обработки сигналов

Объектом исследования является речевой процессор имплантов улитки человеческого уха на основе однокристальных микропроцессоров. Предметом исследования - частотно-временная цифровая модель улитки уха человека, ее частотная карта, методы анализа и синтеза кохлеарных фильтров.

Целью работы является разработка методов анализа и синтеза программируемого речевого процессора с препроцессором повышения разборчивости речи для имплантов улитки человеческого уха, а также его практическая реализация на современной микропроцессорной технике.

Разработаны методы анализа и синтеза программируемого речевого процессора для имплантов улитки с препроцессором повышения разборчивости речи на основе двумерной цифровой модели улитки - банка параметрически перестраиваемых цифровых кохлеарных фильтров представляемых в виде каскада идентичных рекурсивных полосовых секций второго порядка, центральные частоты и полосы пропускания которых могут независимо друг от друга регулироваться посредством изменения соответствующих коэффициентов фильтра, которая устанавливает соотношение между центральными частотами, полосами пропускания кохлеарных фильтров и их положением на базилярной мембране, а также на базе времязависимого параметрически перестраиваемого модуляционного фильтра для повышения качества зашумленной и ревербирующей речи отличающиеся тем, что обеспечивается малый уровень искажений, вследствие высокого динамического диапазона банка фильтров, обусловленного применением вероятностного квантования коэффициентов фильтров, адаптация речевого процессора импланта улитки к особенностям пациента и окружающей акустической среде: шумам и реверберации (соотношение сигнал-шум 0дБ, время реверберации до 3 сек.).

Результаты исследований переданы ОАО Минский приборостроительный завод (г. Минск), внедрены в учебный процесс БГУИР, а также в Институте информатики Белостокского технического университета (Республика Польша).

SUMMARY

of the dissertation thesis “Analysis and synthesis of cochlear implants speech processor” by Jaroslaw Baszun

Keywords: basilar membrane, cochlear map, frequency map of cochlear, cochlear implant, speech processor, digital signal processing.

Speech processor cochlear implant of human ear based on the microprocessor is the object under investigation. The subject for investigation is the time-frequency digital model of cochlear of human ear, its frequency map, and also methods of cochlear filters synthesis and analysis.

The objective of this work is to design the analysis and synthesis of programmable speech processor with preprocessor of speech intelligibility increase for cochlear implants, and its practical realization based on the modern microprocessor technique.

The analysis and synthesis methods of programmable speech processor with preprocessor of speech intelligibility increase based on the two-dimensional cochlear model (bank of tunable digital bandpass second order sections, where the center frequencies and bandwidths are controlled independently; the cochlear model which relates the center frequency of cochlear filters with basilar membrane and bandwidth characteristics which relates the 3dB bandwidth of cochlear filters with basilar membrane locations) and time-dependence tunable modulation filter for quality increase of speech with noise and reverberation. These methods in contrast to well known methods have low level distortions because the cochlear filters coefficients are stochastic quantized, and speech processor is adapted for individual characteristics of patient and acoustical environment: noise and reverberation (SNR 0dB, reverberation time is no less 3 sec.).

The results of work was used at the public corporation “Minsk instrument-making plant” Minsk, educational process at the BSUIR and at the Institute of Informatics of the Technical university Bialystok (Republic of Poland).

БАШУН Ярослав

**АНАЛИЗ И СИНТЕЗ РЕЧЕВОГО ПРОЦЕССОРА
УЛИТКОВОЙ ИМПЛАНТАЦИИ**

Специальность 05.13.05 – Элементы и устройства вычислительной техники
и систем управления

Автореферат диссертации

на соискание ученой степени кандидата технических наук

Подписано в печать	03. 04. 2001.	Формат 60x84 1/16
Бумага офсетная.	Печать ризографическая.	Усл.печ.л. I, 33.
Уч.-изд.л. I, 2.	Тираж 90 экз.	Зак. 228.

Учреждение образования

“Белорусский государственный университет информатики и радиоэлектроники”

Отпечатано в БГУИР. Лицензия ЛП №156.

220013, г.Минск, ул. П.Бровки, 6