

УДК 534.773

И.В. ПРАСОЛ, А.С. НЕЧИПОРЕНКО

*Харьковский национальный университет радиоэлектроники, Украина***ОСОБЕННОСТИ ОБРАБОТКИ РЕЧЕВЫХ СИГНАЛОВ
В ЦИФРОВЫХ СЛУХОВЫХ АППАРАТАХ**

Статья посвящена проблеме обработки речевых сигналов в цифровых слуховых аппаратах. Рассмотрены особенности восприятия речи больными нейросенсорной тугоухостью, а также существующие методы цифровой обработки сигналов, используемые в современных цифровых слуховых аппаратах для компенсации слуховых потерь. Предложен адаптивный метод цифровой обработки речевых сигналов, основанный на фильтрации сигнала, с выделением формантных областей и последующей транспозиции спектра сигнала в область остаточного слухового восприятия пациента, который позволяет достоверно повысить разборчивость речи у больных нейросенсорной тугоухостью.

Ключевые слова: цифровой слуховой аппарат, сигнальный процессор, методы цифровой обработки сигналов, разборчивость речи, нейросенсорная тугоухость, транспозиция спектра.

Введение

В настоящее время всё большее число людей с различными нарушениями слуха нуждается в слухопротезировании. Эффективность слухопротезирования зависит от многих факторов, таких как точная диагностика нарушений слуха, квалификация специалиста, занимающегося слухопротезированием, а также от технических возможностей слуховых аппаратов. Благодаря применению методов цифровой обработки сигналов появилась возможность более точной компенсации слуховых потерь. Например, согласование динамического диапазона сигналов с областью остаточного восприятия, что достигается с помощью многополосной компрессии, применение методов шумоподавления, а также алгоритмов обработки информативных признаков речевых сигналов, предназначенных для повышения разборчивости речи, алгоритмов формирования пространственных диаграмм направленности микрофонов [1]. Предпринимались также попытки переноса компонентов сигнала из зоны полной глухоты конкретного пациента в зону слышимости, однако применяемые методы переноса спектра приводили к ухудшению помехоустойчивости слуховых аппаратов и поэтому не получили широкого распространения [2, 3]. В работе [4] описан алгоритм частотной транспозиции спектра речевого сигнала в область остаточной слышимости, применение которого, однако, не позволяет достичь значительного улучшения разборчивости речи. Также следует отметить, что для пациентов с определёнными особенностями патологии данной транспозиции сигнала недостаточно. Число больных, которые оценивают результаты слухопротезирования как неудовлетворительные, в настоя-

щее время достаточно велико. Существует ряд патологий, таких как нейросенсорная тугоухость, когда имеют место специфические поражения звукопринимающего отдела слухового анализатора, для компенсации которых необходимы специальные алгоритмы обработки сигнала, осуществляющие преобразование сигналов в представление, обеспечивающее максимальную разборчивость речи остаточным слухом больного [5]. В данной работе предложен адаптивный метод обработки речевых сигналов, который позволяет повысить разборчивость речи для больных, у которых, вследствие поражения наружных волосковых клеток, нарушено либо полностью отсутствует слуховое восприятие различных диапазонов частот. Метод состоит из трёх основных этапов предварительной обработки сигнала – это реализация шумоподавления, фильтрация речевого сигнала с помощью адаптивно перестраиваемых, в зависимости от расположения формант, полосовых фильтров, а также транспозиция полученного спектра сигнала в область остаточного слухового восприятия пациента. Варианты реализации данного метода варьируются от математического моделирования в программной среде типа Matlab [6] до реализации предложенной модели в режиме реального времени с применением специализированных сигнальных процессоров.

1. Шумоподавление

Для решения практической задачи шумоочистки целесообразно использовать методы динамического шумоподавления, а также методы, основанные на использовании характеристик шума [7]. В данной работе осуществляется двухступенчатое шумопо-

давление. На первом этапе происходит динамическое шумоподавление, (рис. 1), суть которого заключается в том, что выбирается определенный порог сигнала – порог шума, а все сигналы ниже данного уровня считаются шумом и отсекаются.

Следует отметить, что обрабатывается весь спектр сигнала, а не только речевой диапазон. Выявленные участки относительной тишины сохраняются и используются в дальнейшем шумоподавлении. Входной сигнал должен иметь нормированный уровень, для чего в микрофонном усилителе используется автоматическая регулировка усиления, что также позволяет наиболее эффективно использовать АЦП, задействовав всю его разрядность.

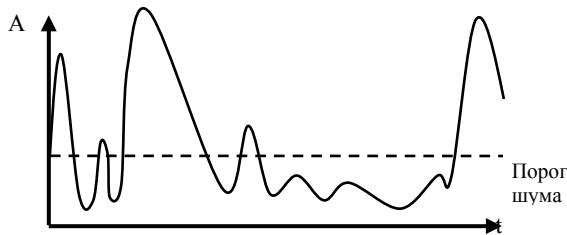


Рис. 1. Динамическое шумоподавление

Применение данного метода обосновано тем, что одновременно с подавлением шумов происходит выделение так называемых образцов шума, которые используются на следующем этапе для спектрального вычитания. На втором этапе используются методы, которые основаны на различных модификациях алгоритма вычитания амплитудных спектров. Такой подход оптимален в случае широкополосных непрерывных и импульсно-непрерывных помех, пересекающихся с областью спектра речи. Шумы данного типа не могут быть устранены другими методами (например, адаптивной фильтрацией), поскольку такие помехи являются рассредоточенными по спектру и пересекаются с областью спектра речи.

Пусть $X_p(k)$ – спектр зашумленного сигнала на p -м фрейме, $Y(k)$ – спектр шума, $S_p(k)$ – спектр восстановленного сигнала на p -м фрейме. На практике шум $Y(k)$ вычисляется на шумовых фреймах сигнала $X(k)$. Это связано с тем, что обычно известен только зашумленный сигнал. Таким образом, воспользуемся формулой фильтрации по Боллу [8].

$$S_p(k) = \begin{cases} H_p(k)X_p(k), & |X_p(k)| - \alpha |Y(k)| > 0 \\ \beta X_p(k), & \text{иначе} \end{cases}$$

$$H_p(k) = \frac{|X_p(k)| - \alpha |Y(k)|}{|X_p(k)|} \quad (1)$$

Данное выражение описывает спектральное вычитание амплитуды шума по модулю. На начальном этапе осуществляется разложение сигнала с по-

мощью быстрого преобразования Фурье (БПФ), затем производится оценка спектра шума, далее "вычитание" амплитудного спектра шума из амплитудного спектра сигнала. Восстановление речевого сигнала во временной области осуществляется с помощью обратного преобразования Фурье. Под оценкой спектра шума подразумевается выборка сигналов без полезной информации, содержащей только шум. Данные выборки получены в результате динамического шумоподавления.

При выборе метода шумоочистки авторы руководствовались такими критериями как разборчивость обработанного сигнала и аппаратные возможности процессора. Следует отметить, что в случае, когда шум или помеха имеют стационарный (или квазистационарный) характер и их спектр имеет гармоническую структуру, удаётся достичь значительного повышения как качества, так и разборчивости речи. Однако, если шум имеет быстроизменяющуюся спектральную характеристику, то необходимо применять пост-обработку сигнала в спектрально-временной области, для чего используются различные алгоритмы [8]. Реализация вышеописанного двухступенчатого шумоподавления, которое представляет собой комбинацию двух методов, позволяет повысить эффективность очистки полезного сигнала от шумов. Для оценки эффективности очистки сигнала использовались меры качества звучания сигнала и объективные меры типа усреднённых евклидовых расстояний между отфильтрованным и незашумленным сигналами. Следует также отметить, что важным достоинством двухступенчатого шумоподавления является отсутствие необходимости больших вычислительных ресурсов.

2. Фильтрация речевого сигнала

Для спектрального анализа сигнала используется дискретное преобразование Фурье (ДПФ), с помощью которого строится ключевая импульсная характеристика фильтра. В данном случае используется алгоритм быстрого преобразования Фурье (БПФ) как наименее ресурсоемкий, что необходимо для вычислений в реальном масштабе времени. Современные сигнальные процессоры производят вычисление БПФ за минимальное число машинных циклов [9]. Процессоры содержат дополнительные внутренние функции, такие как, например, наличие двух генераторов адреса с возможностью бит-реверсивной адресации и другие аппаратные возможности, что также позволяет оптимизировать вычислительный процесс. Алгоритмы БПФ по постоянным основаниям обеспечивают рекурсивность, регулярность вычислений, а также возможность гибкого изменения формата преобразования. Алго-

ритм БПФ по постоянному основанию 2 базируется на следующем разложении основной формулы вычислений ДПФ:

$$X(k) = \sum_{n=0}^{N-1} x(nT)W_N^{kn}, k = 0, \dots, N-1, \quad (2)$$

где $X(k)$ ($k = 0, \dots, N-1$) – последовательность из N частотных отсчётов; $x(nT)$, ($n = 0, \dots, N-1$) – последовательность из N временных отсчётов с периодом T ; $W_N = e^{-i2\pi/N}$, $i = \sqrt{-1}$ – дискретное преобразование Фурье (ДПФ).

Исходная последовательность $x(nT)$ разбивается на две последовательности $x_1(nT) = x(nT)$ и $x_2(nT) = x((n + \frac{N}{2})T)$, где $n = 0, \dots, N/2 - 1$ и отдельно вычисляются чётные и нечётные отсчёты ДПФ, согласно формулам:

$$X(2k) = \sum_{n=0}^{N/2-1} (x_1(nT) + x_2(nT))W_{N/2}^{nk}, \quad (3)$$

$$X(2k+1) = \sum_{n=0}^{N/2-1} (x_1(nT) - x_2(nT))W_{N/2}^{nk}. \quad (4)$$

Таким образом, в рассматриваемом алгоритме выходная последовательность $X(k)$ располагается в двоично-инверсионном порядке, и именно такая последовательность используется для вычислений в сигнальных процессорах.

Для фильтрации сигнала используются 3 полюсовых фильтра с конечной импульсной характеристикой (КИХ), квантованная импульсная характеристика определяет коэффициенты фильтра, порядок фильтра выбирается в зависимости от необходимой крутизны амплитудно-частотной характеристики (АЧХ) фильтра. КИХ-фильтр должен работать в соответствии с уравнением, задающим свёртку, согласно формуле:

$$Y(n) = h(k) * x(n) = \sum_{k=0}^{N-1} h(k)x(n-k), \quad (5)$$

где $h(k)$ – массив коэффициентов фильтра;

$x(n-k)$ – входной массив данных фильтра;

N – число звеньев фильтра.

Проектирование фильтров осуществляется с помощью метода Паркса-Макклиллана [10], который использует алгоритм обнѳмена Ремеза, проектирование КИХ-фильтров с фиксированной неравномерностью полосы пропускания, полосы задержки. Граничные частоты полюсовых фильтров выбираются таким образом, чтобы выделить частотные области, в которых с наибольшей вероятностью сосредоточены форманты большинства голосовых типов, согласно табл. 1, а также произвести фильтрацию сонорных звуков, тем самым обеспечивая их

выделение из шумов. Частота дискретизации в данном случае выбрана 8000 Гц, неравномерность полосы пропускания 0,01 дБ. При условии соблюдения точности в формулировании технических требований, которые зависят от особенностей программируемого процессора, число звеньев фильтра рассчитывается автоматически.

Таблица 1

Области локализации формант

№ форманты	Мужской голос (интервал в Гц)	Женский голос (интервал в Гц)
1	200-800	250-1000
2	600-280	700-3300
3	300-3400	1500-4000

Полученные квантованные коэффициенты фильтра являются исходными данными для ассемблерного кода либо любого другого кода на языке низкого уровня для ряда различных процессоров. Следует отметить, что в слуховом аппарате необходимо, чтобы выделение формант происходило в режиме реального времени, т.е. с задержкой не более 50 мс. При оценке необходимой вычислительной мощности процессора необходимо учитывать время вычисления базовой операции БПФ («бабочка»), а также количество обрабатываемых числовых значений в процессе вычислений. Для предотвращения переполнения разрядной сетки данные следует масштабировать, заранее оставляя достаточное количество дополнительных разрядов для увеличения значений обрабатываемых данных. Альтернативный подход заключается в том, что данные могут масштабироваться после каждого каскада вычисления БПФ. Использование сигнального процессора с плавающей точкой устраняет потребность в масштабировании данных и поэтому приводит к более простой реализации алгоритма БПФ, но следствием этого упрощения является увеличение времени обработки, которое требуется для сложных арифметических вычислений с плавающей точкой. Кроме того, 32-разрядный процессор с плавающей точкой, очевидно, будет иметь меньший уровень шумов округления, чем 16-разрядный процессор с фиксированной точкой. В табл. 2 приведены данные по реализации БПФ для популярных сигнальных процессоров Analog Devices.

Таблица 2

БПФ на популярных DSP

Процессор	Время выполнения БПФ
ADSP-2189M, 16 разрядов, фиксированная точка	453 мкс (1024 точки)
ADSP-21160 SHARC™, 32 разряда, плавающая точка	90 мкс (1024 точки)
ADSP-TS001 TigerSHARC™ 150 MHz 16 разрядов, режим с фиксированной точкой	7,3 мкс (256 точек БПФ) 69 мкс (1024 точки)

В частности, процессор ADSP-TS001 TigerSHARC™ обеспечивает оба режима: и с плавающей, и с фиксированной точкой, обеспечивая, таким образом, исключительную гибкость программирования.

3. Транспозиция спектра сигнала

Основная причина нарушения восприятия речи у больных нейросенсорной тугоухостью – это невозможность восприятия информационно значимых компонентов звука, расположенных в зоне с нарушенной чувствительностью слуха либо в зоне полной глухоты. С целью компенсации данных нарушений необходимо осуществить реконструкцию спектра звуков таким образом, чтобы информативные признаки речевого сигнала попадали в частотную область с сохранившейся чувствительностью слуха для конкретного пациента. Данную реконструкцию предлагается производить путём транспозиции спектра сигнала, в основе которой лежит принцип амплитудной модуляции. Линейная транспозиция спектра описывается формулой:

$$f_n(t) = f_o(t) \pm f_{sp}, \quad (6)$$

где $f_n(t)$ – частотный спектр сигнала после транспозиции;

$f_o(t)$ – частотный спектр сигнала до транспозиции;

$f_{sp}(t)$ – частота сдвига.

При проектировании цифровых слуховых аппаратов основными требованиями при переносе спектра являются высокая помехоустойчивость, а также равномерный последовательный перенос спектра. Для этого предлагается использовать метод амплитудной модуляции, при котором в соответствии с модулирующим сигналом изменяется амплитуда несущего колебания. Амплитудную модуляцию можно описать формулой:

$$S_{AM}(t) = A(t) \cos(\omega_0 t + \varphi_0), \quad (7)$$

где $S_{AM}(t)$ – амплитудно-модулированный сигнал;

$A(t)$ – амплитудная функция;

ω_0 – несущая частота;

φ_0 – начальная фаза.

Спектр сигнала после транспозиции представлен на рис. 2.

Как видно из рис. 2, спектр сигнала после амплитудной модуляции состоит из спектра частоты модуляции и верхней и нижней боковых полос модулированного сигнала [11]. Поскольку одним из основных требований к алгоритмам обработки речевых сигналов в слуховых аппаратах является повышение разборчивости речи на фоне шумов, были проанализированы три варианта модуляции: использование обеих результирующих полос, нижней полосы, а также верхней полосы частот. Экспери-

ментально установлено, что использование верхней полосы спектра (справа на рис. 2) обеспечивает максимальную разборчивость речи.

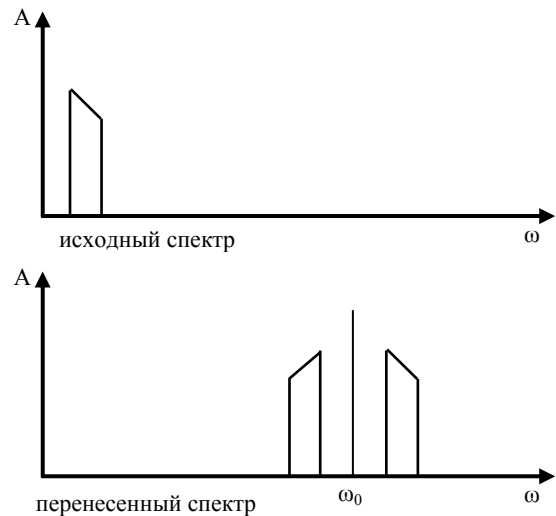


Рис. 2. Спектр модулированного сигнала

Следует отметить также, что при этом удаляется модулирующая частота. Естественно, что качество звучания реконструированных таким образом звуков исказится, однако после специальной тренировки (абилитации), пациент обучается распознавать звуки с достаточно высокой вероятностью (0,9). Таким образом, амплитудная модуляция является простой и нересурсоемкой операцией для современных сигнальных процессоров, в то же время обеспечивая перенос спектра в нужные диапазоны частот. Следует отметить, что применение среды Matlab позволяет быстро и наглядно оценить эффективность метода цифровой обработки сигнала, а также оптимизировать алгоритм обработки перед выбором аппаратных средств его реализации.

4. Анализ полученных результатов

Для оценки эффективности предложенного метода обработки речевых сигналов был сконструирован рабочий макет, который включает в себя микрофон, микрофонный усилитель, блок АЦП – процессор – ЦАП, представленный микроконтроллером Atmel Mega 88, усилитель НЧ и динамик. С помощью данного макета производится транспозиция спектра из области полной либо частичной глухоты в область слышимости для конкретного пациента. Микроконтроллер с помощью специальной программы осуществляет перенос спектра сигнала в необходимый диапазон частот. Результаты исследований представлены на рис. 3 – 6.

В результате адаптивной фильтрации выделяются области частот с наиболее информативными признаками речи, после чего данные участки спек-

тра переносятся в диапазон слышимых частот, в соответствии с аудиограммой пациента. На рис. 3 и рис. 5 представлены аудиограммы двух пациентов с диагнозом нейросенсорная тугоухость, с различными особенностями патологии, на рис. 4 и рис. 6 представлены кривые разборчивости речи больных, полученные после прослушивания исходного сигнала, а также сигнала, полученного в результате транспозиции спектра.

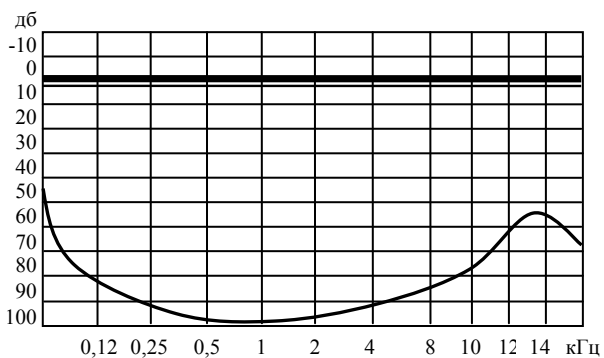


Рис. 3. Аудиограмма больного кохлеарным невритом

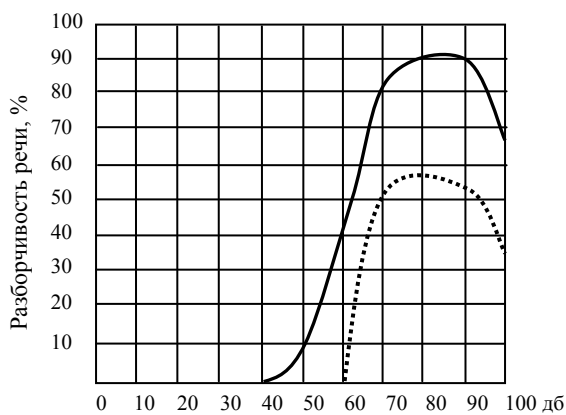


Рис. 4. Кривая разборчивости речи больного кохлеарным невритом:

- до транспозиции (пунктирная линия);
- после транспозиции (сплошная линия)

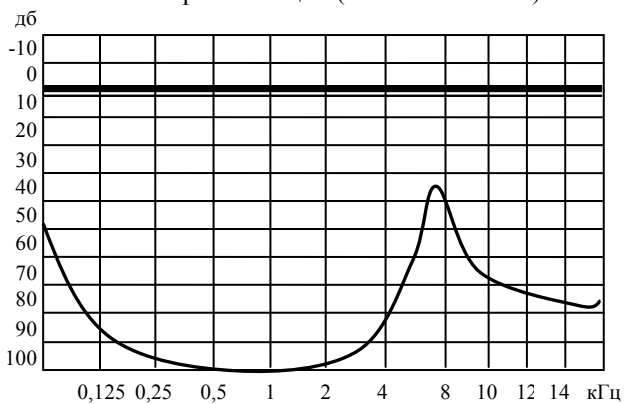


Рис. 5. Аудиограмма больного хронической нейросенсорной тугоухостью

В эксперименте участвовало 20 пациентов с диагнозом нейросенсорная тугоухость, с различными

степенями тугоухости. Разборчивость речи оценивалась с помощью речевой аудиметрии. В итоге, после предварительной тренировки, пациенты начинали распознавать речь с достаточно высокой вероятностью. Следовательно, можно сделать вывод о том, что в результате такой транспозиции спектра сигнала в области слышимости конкретного пациента происходит достоверное повышение разборчивости речи.

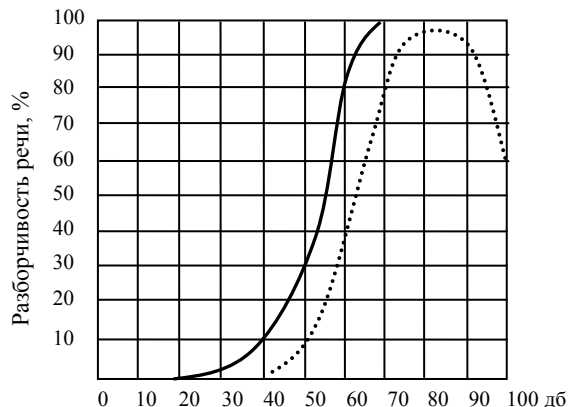


Рис. 6. Кривая разборчивости речи больного хронической нейросенсорной тугоухостью:

- до транспозиции (пунктирная линия);
- после транспозиции (сплошная линия)

Заключение

В данной работе предложен метод цифровой обработки речевых сигналов в слуховых аппаратах, позволяющий осуществлять реконструкцию речевых сигналов таким образом, чтобы достичь повышения разборчивости речи для пациентов с нейросенсорной тугоухостью, у которых имеет место частичное нарушение либо полное отсутствие восприятия речевых элементов. Метод реализован в программной среде Matlab 7.5, а также проанализирована возможность его реализации в современных сигнальных процессорах. Использование макета цифрового слухового аппарата в аудиологической практике позволит повысить эффективность процесса слухопротезирования. Таким образом, можно сделать вывод, что предложенная последовательность этапов обработки сигналов позволяет достичь повышения разборчивости речи, более оптимально использовать технические ресурсы современных сигнальных процессоров, а также осуществлять необходимую обработку сигналов в режиме реального времени. Следует отметить, что развитие методов цифровой обработки сигналов в скором времени позволит применять ещё более сложные алгоритмы реконструкции и адаптивной фильтрации сигналов, характеристики которых будут максимально отображать механизмы функционирования периферического отдела слухового анализатора.

Литература

1. Killion Mead C.A. High fidelity hearing aid / A.C. Killion Mead // *Hearing Instruments*. – 1990. – Vol. 41, № 8. – P. 25-27.
2. Лисовский В.А. Слуховые приборы и аппараты / В.А. Лисовский, В.А. Елисеев. – М.: Радио и связь, 1991. – 190 с.
3. Sokol B.J. *Electr and Wirel world* / B.J. Sokol, M. Velmans // *Hearing Instruments*. – 1998. – № 94. – P. 174-175.
4. *Critical factors in ensuring efficacy of frequency transposition I: individualizing the start frequency* / F. Kuk, D. Keenan, H. Peeters, P. Korhonen, O. Nau, H. Andersen // *Hearing Review*. – 2007. – № 14(3). – P. 60-67.
5. Бабкина Л.Н. Пути использования цифровой обработки сигналов для расширения возможностей цифровых слуховых аппаратов / Л.Н. Бабкина, А.П. Молчанов // *Вестник оториноларингологии*. – 1998. – №4. – С. 18-21.
6. Чен К. Матлаб в математических исследованиях: пер. с англ. / К. Чен, П. Джублин, А. Ирвинг. – М.: Мир, 2001. – 346 с.
7. Hansen J.H.L. *Analysis and compensation of speech under stress and noise for environmental robustness in speech recognition* / J.H.L. Hansen // *Speech Communication*. – 1996. – Vol. 20, № 2. – P. 151-173.
8. Чучупал В.Я. Алгоритмы и программы для предобработки зашумленных речевых сигналов / В.Я. Чучупал, А.В. Чичагов, К.А. Маковкин // *Тезисы 8-й Всероссийской конф. "Математические методы распознавания образов"*. – М., 1997. – С. 222.
9. *Signal Processing in High-End Hearing Aids: State of the Art, Challenges, and Future Trends* / V. Hamacher, J. Chalupper, J. Eggers, E. Fischer, U. Kornagel, H. Puder, U. Rass // *Eurasip Journal on Applied Signal processing*. – 2005. – №18. – P. 2911-2914.
10. Smith Steven W. *The Scientist and Engineer's Guide to Digital Signal Processing, Second Edition* / Steven W. Smith. – San Diego: California Technical Publishing, 1999. – 643 p.
11. *Frequency lowering processing for listeners with significant hearing loss*. In / C. Aguilera-Munoz, B. Peggy, C. Rutledge, A. Gago // *Proceedings of the Sixth IEEE International Conference on Electronics, Circuits and Systems, (Cat. No.99EX357)*. – Part 2(2). – 1999. – P. 741-744.

Поступила в редакцию 28.11.2008

Рецензент: д-р техн. наук, проф., проф. каф. ПОЭВМ Г.Г. Четвериков, Харьковский национальный университет радиоелектроники, Харьков.

ОСОБЛИВОСТІ ОБРОБКИ МОВНИХ СИГНАЛІВ У ЦИФРОВИХ СЛУХОВИХ АПАРАТАХ

І.В. Прасол, А.С. Нечипоренко

У статті розглянуто проблему обробки мовних сигналів у цифрових слухових апаратах. Розглянуто особливості сприйняття мови хворими на нейросенсорну тугоухість, а також існуючі методи цифрової обробки сигналів, що використовуються у сучасних цифрових слухових апаратах з метою компенсації слухових втрат. Запропоновано адаптивний метод цифрової обробки мовних сигналів, який базується на фільтрації сигналу з метою відокремлення формантних ділянок та подальшій транспозиції спектру сигналу у діапазон залишкового слуху пацієнта.

Ключові слова: цифровий слуховий апарат, сигнальний процесор, методи цифрової обробки сигналів, розбілівність мови, нейросенсорна тугоухість, транспозиція спектру.

FEATURES OF SPEECH PROCESSING IN DIGITAL HEARING AIDS

I.V. Prasol, A.S. Nechiporenko

The article is devoted the problem of speech processing in digital hearings aids. The features of perception of speech for patients with sensorineural hearing loss are considered. Existent methods of digital speech processing, in-use in modern digital hearing aids for indemnification of auditory losses are considered too. A new adaptive method to improve phrase intelligibility of people sick sensorineural hearing loss is suggested. It is based on the filtering and frequency transposition the most significant areas of speech spectrum which affects on speech perception patients are detected.

Key words: digital hearing aid, signaling processor, methods of the digital signal processing, phrase intelligibility, sensorineural hearing loss, spectrum transposition.

Прасол Игорь Викторович – канд. техн. наук, проф. кафедры биомедицинских электронных устройств и систем, Харьковский национальный университет радиоелектроники, Харьков, Украина, e-mail: nkc@kture.kharkov.ua.

Нечипоренко Алина Сергеевна – аспирант кафедры биомедицинских электронных устройств и систем, Харьковский национальный университет радиоелектроники, Харьков, Украина, e-mail: alinanechiporenko@gmail.com.