

УДК 004.4'277

## ПОВЫШЕНИЕ КАЧЕСТВА СЛУХОВОГО ВОСПРИЯТИЯ В ШУМОВОЙ ОБСТАНОВКЕ НА ОСНОВЕ ЧАСТОТНОЙ ДЕКОМПОЗИЦИИ И АДАПТИВНОЙ КОМПРЕССИИ ДИНАМИЧЕСКОГО ДИАПАЗОНА СИГНАЛА

*Азаров И.С., к.т.н., Белорусский государственный университет информатики и радиоэлектроники, e-mail: azarov@bsuir.by;*

*Вашкевич М.И., аспирант кафедры Электронных вычислительных средств Белорусского государственного университета информатики и радиоэлектроники, e-mail: vashkevich@bsuir.by;*

*Лихачев Д.С., Белорусский государственный университет информатики и радиоэлектроники;*

*Петровский А.А., д.т.н., профессор, зав. кафедрой Электронных вычислительных средств Белорусского государственного университета информатики и радиоэлектроники, e-mail: palex@bsuir.by.*

**Ключевые слова:** качество слухового восприятия, звуковой сигнал, шумовая обстановка, нелинейное усиление, адаптивная компрессия.

### Введение

При использовании мобильных звуковоспроизводящих устройств в шумовой обстановке пользователь воспринимает смесь чистого сигнала (от устройства) и внешнего акустического шума. Во многих практических ситуациях интенсивность шума столь высока, что приводит к существенному ухудшению слухового восприятия за счет эффекта маскирования полезного сигнала шумом. Данный эффект возникает как во время разговора с удаленным абонентом, так и при прослушивании различных звуковых записей (музыка, аудиокниги и проч.).

На рис. 1 показаны сигналы чистой речи и шума метрополитена для которых отношение сигнал/шум (ОСШ) составляет -5 дБ. В результате наложения сигналов спектральные компоненты речи полностью или частично маскируются шумом, что существенно ухудшает разбор-

*Исследуется возможность автоматизированного повышения качества восприятия звуковых сигналов в шумовой обстановке. С целью повышения качества восприятия предлагается выполнять нелинейное усиление частотно-временных компонент воспроизводимого сигнала, зависящее от спектральных характеристик окружающего шума. Предлагаемый метод обработки является универсальным, поскольку одинаково хорошо применим для обработки различных звуковых сигналов, в том числе речевых и музыкальных. Выполнено исследование практической применимости предложенного метода для речевых сигналов с использованием объективных оценок.*

чивость и общее качество восприятия.

В описываемой ситуации шум невозможно устранить при помощи классических способов шумоподавления (таких, например, как спектральное взвешивание), поскольку шум добавляется к полезному сигналу после его воспроизведения устройством. Таким образом, единственной практической возможностью улучшить восприятие полезного сигнала является его предварительная обработка, с учетом характеристик окружающего акустического шума. На практике пользователь воспроизводящего устройства вынужден самостоятельно регулиро-

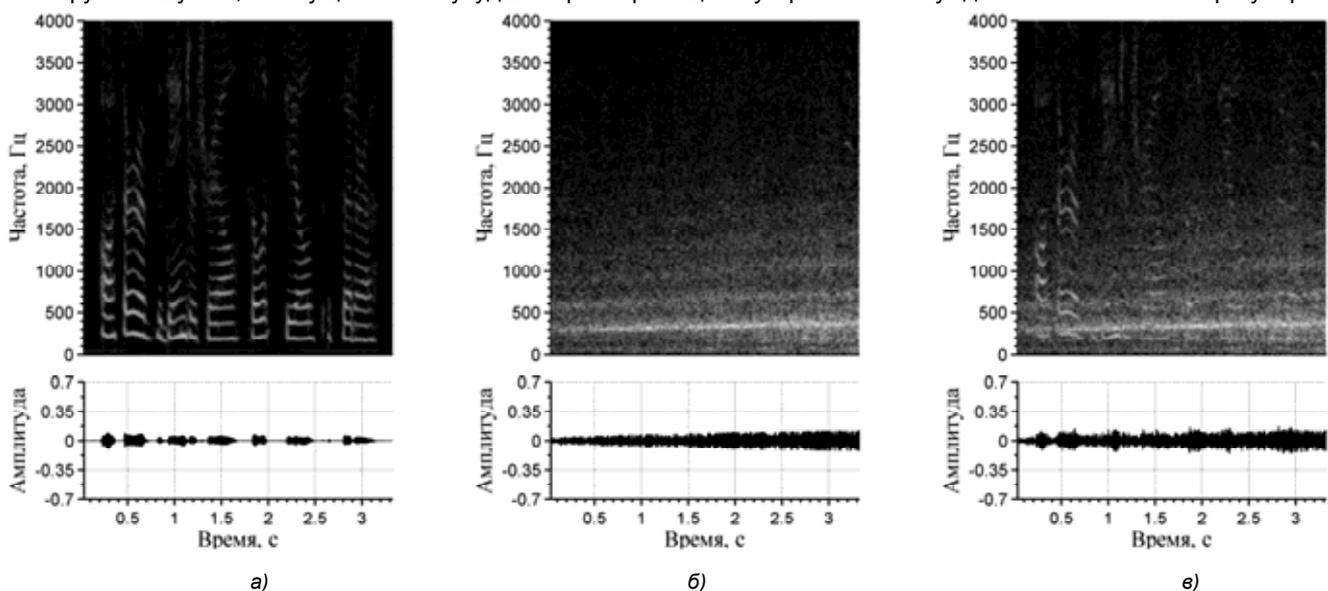


Рис. 1. Частотно-временное представление сигнала:  
а) чистая речь; б) шум метрополитена; в) сумма речи и шума

вать громкость воспроизведения в зависимости от шумовой обстановки, однако при этом возникают следующие неудобства: 1) при наличии нестационарного источника шума требуется частая ручная регулировка; 2) регулировка громкости одинаково усиливает все частотные компоненты вне зависимости от того маскируются они шумом или нет, и в результате некоторые важные компоненты полезного сигнала остаются неслышными, в то время как другие получают чрезмерное усиление, вызывающее дискомфорт и быстрое утомление при прослушивании.

Современные мобильные телекоммуникационные устройства, часто используемые для воспроизведения звука, имеют внешний микрофон, который можно использовать для оценки параметров шума. Это делает возможным создание мобильной системы повышения качества воспроизведения звуковых сигналов с автоматическим учетом шума окружающей среды (рис. 2).



Рис. 2. Система для прослушивания сигнала в шумовой обстановке на основе мобильного телекоммуникационного устройства

В настоящее время предложено несколько подходов к решению указанной задачи. Например, в [1] предлагается способ выборочного усиления частотно-временных составляющих полезного сигнала. Целью обработки является обеспечение заданного минимального ОСШ между усредненным спектром полезного сигнала и измеренным спектром шума. Недостатком метода является то, что в результате обработки между всеми тихими звуками теряется относительная громкость – они становятся одинаково громкими на фоне шума. Также теряется относительная громкость и между спектральными составляющими, поскольку они усиливаются независимо друг о друга, что приводит к нежелательному эффекту «выбеления»<sup>1</sup> полезного сигнала. Похожий способ, предложенный в [2], состоит в избирательном усилении частотно-временных составляющих полезного сигнала для увеличения индекса разборчивости речи (ИРП) ( $SII$  –

*speech intelligibility index*). Здесь показано, что обработка сигнала с учетом ИРП позволяет получить более качественный результат, чем обработка с учетом ОСШ [1]. Однако ИРП применим только в случае, когда полезный сигнал является речевым, в то время как на практике, сигналы часто содержат неречевые фрагменты (музыка, звуки природы, аудио-эффекты и пр.). Общим недостатком обоих методов является то, что для анализа и обработки сигнала используется кратковременное дискретное преобразование Фурье (ДПФ), которое обеспечивает равномерное частотное разрешение, т.е. не учитывается нелинейное восприятие частот слуховой системой человека.

Недавно был предложен способ повышения разборчивости речи в шумах, сохраняющий энергию исходного сигнала [3] и состоящий из двух последовательных этапов: адаптивного изменения формы спектральной огибающей и расширения/компрессии динамического диапазона. Адаптивное изменение формы спектральной огибающей выполняется путем подчеркивания формантной структуры речи и уменьшения наклона спектральной огибающей корректирующим фильтром. Степень обработки зависит от уровня вокализованности речевого фрагмента – чем ближе он к гласному звуку, тем сильнее вносимый эффект. Способ предназначен для обработки исключительно речевых сигналов и предполагает наличие в сигнале специфических для речи характеристик, таких как формантные частоты и вокализованные фрагменты, а также подразумевает возможность произвести классификацию диктора (мужчина либо женщина). Поскольку сигнал сохраняет свою энергию, невозможно достичь удовлетворительной разборчивости при низких ОСШ. Недостатком способа является также и то, что даже при наличии идеальных условий (в полной тишине) в обрабатываемый сигнал вносятся искажения. Способ плохо применим для нестационарных шумов, поскольку в нем не учитываются изменяющиеся спектральные и динамические характеристики шума.

Целью данной работы является исследование возможности построения схемы обработки сигнала, обеспечивающей повышение разборчивости и натуральности звучания речи и музыки при прослушивании в шумной акустической обстановке за счет снижения эффекта маскирования полезного звукового сигнала нестационарными акустическими шумами путем использования частотно-зависимого адаптивного усиления. Важным условием является использование минимального необходимого усиления, достаточного для устранения эффекта маскирования, но при этом не создающего дискомфорта для восприятия и не приводящего к быстрой утомляемости слуха.

Предлагаемая схема обработки использует декомпозицию сигнала неравнополосным банком фильтров, позволяющим моделировать неравномерное частотное разрешение слуховой системы человека. Усиление каждой субполосы сигнала выполняется в зависимости от интенсивности шума в этой полосе при помощи адаптивного компрессора динамического диапазона с изменяющимися параметрами.

<sup>1</sup> Спектр сигнала становится плоским, как для белого шума, отсюда и название эффекта

**Повышение качества слухового восприятия звуковых сигналов в шумовой обстановке**  
**Схема обработки сигнала**

Ухудшение восприятия звукового сигнала в шумовой обстановке вызвано тем, что шум маскирует [4] часть составляющих полезного сигнала. Эффект маскирования возможно уменьшить за счет компрессии (сужения) динамического диапазона (ДД) полезного сигнала [4, 5]. На рис. 3 показаны плотности распределения вероятности уровня частотных компонент речевого сигнала (рис. 3, а), шума метрополитена (рис.3, б) и речевого сигнала после компрессии (рис.3, в). Заметно, что частотные составляющие речи лежат в широком динамическом диапазоне, причем средний уровень компонент шумового сигнала для большого числа частотных полос выше среднего уровня компонент речевого сигнала. В результате компрессии громкие звуки получают малое усиление, а тихие – большое за счет чего эффект маскирования тихих компонент речевого сигнала становится менее выраженным.

Поскольку интенсивность шума зависит от частоты, то целесообразно использовать отдельные компрессоры для разных частотных диапазонов сигнала. Предлагаемая схема обработки заключается в выполнении следующих шагов: 1) декомпозиция полезного сигнала и шума на субполосы; 2) вычисление кратковременной энергии в каждой субполосе; 3) компрессия каждой субполосы полезного сигнала с учетом кратковременной энергии шума; 4) реконструкция выходного широкополосного сигнала из обработанных субполосных сигналов.

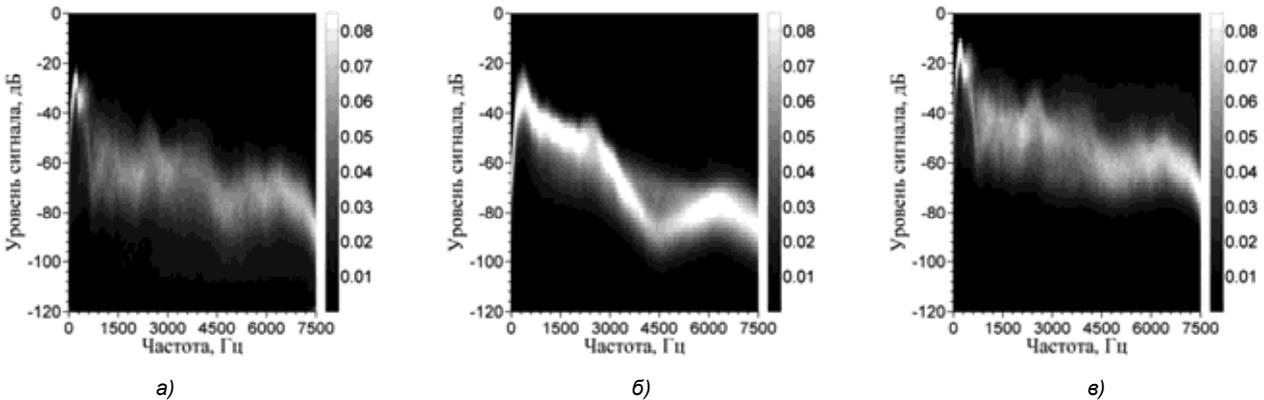


Рис. 3. Плотность распределения вероятности уровня частотных составляющих: а) речевого сигнала; б) шума метрополитена; в) речевого сигнала после компрессии



Рис. 4. Схема обработки сигнала для повышения качества его воспроизведения в шумовой обстановке

Декомпозиция и реконструкция полезного сигнала и шума выполняется при помощи неравнополосного косинусно-модулированного банка фильтров [6]. Для улучшения восприятия ширина полос выбирается в соответствии с разрешающей способностью слуховой системы человека. На рис. 4 показана предлагаемая схема обработки сигнала.

Схема имеет два входа, на которые поступают полезный и шумовой сигналы. На выходе формируется обработанный полезный сигнал. Оба входных сигнала раскладываются на субполосы при помощи банка фильтров анализа. Отсчеты субполосных сигналов набираются с прореживанием для того, чтобы снизить вычислительные затраты при практической реализации. Прореживание не приводит к потере информации, поскольку каждый субполосный сигнал имеет узкий частотный диапазон. Для восстановления полезного сигнала используется банк фильтров синтеза.

**Адаптивный компрессор динамического диапазона**

Уровень сигнала на выходе компрессора зависит от уровня входного сигнала и амплитудной характеристики компрессора. Амплитудная характеристика определяет вносимое усиление в зависимости от уровня входного сигнала (часто амплитудную характеристику отображают в виде графика, который показывает зависимость выходного уровня от входного).

На рис. 5, а показана амплитудная характеристика КДД, в соответствии с которой выполняется компрессия входного сигнала, если его уровень не превосходит по-

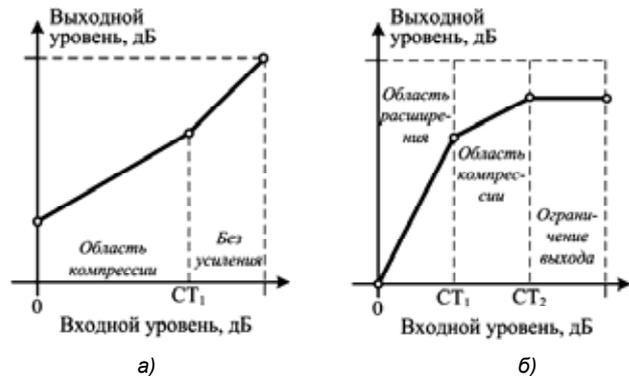
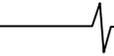


Рис. 5. Типовые амплитудные характеристики КДД



рога срабатывания  $CT_1$ . В тоже время сигналы, чей уровень выше  $CT_1$ , пропускаются компрессором без изменения. Более сложная амплитудная характеристика, сочетающая в себе три стратегии обработки, показана на рис. 5, б. Низкоуровневые сигналы (меньшие порога  $CT_1$ ) попадают в область расширения, в которой усиление сигнала увеличивается по мере роста входного уровня. Применение операции расширения динамического диапазона делает более различимыми сигналы, обладающие малой энергией. Сигналы с уровнем, попадающим в диапазон от  $CT_1$  до  $CT_2$ , подвергаются компрессии, т.е. на этом участке коэффициент усиления падает с возрастанием уровня сигнала. Третий участок функции на рис. 5, б соответствует работе компрессора в режиме ограничителя, при этом выполняется ослабление, а не усиление сигнала.

Обычно амплитудная характеристика является фиксированной и не изменяется в процессе работы компрессора, что не подходит для решения поставленной задачи, поскольку в данном случае степень компрессии должна зависеть от уровня регистрируемого шума – при агрессивном шуме она должна быть высокой, а в тишине должна быть нулевой. Необходимо использовать перестраиваемый (адаптивный) компрессор, с изменяющейся амплитудной характеристикой.

Математически адаптивный компрессор описывается действием нелинейного оператора

$$y[n] = \psi \{x[n], w[n]\},$$

где  $x[n]$  – полезный сигнал,  $w[n]$  – шумовой сигнал, а  $y[n]$  – выходной сигнал компрессора. Для обработки компрессором входной и шумовой сигналы разбиваются на фреймы размером  $R$  отсчетов:

$$x^{(m)}[r] = x[mR + r],$$

$$w^{(m)}[r] = w[mR + r],$$

где  $m$  – номер фрейма, а  $r = 1, 2, \dots, R$  – индекс отсчета внутри фрейма. Выходной сигнал компрессора рассчитывается следующим образом

$$y^{(m)}[r] = x^{(m)}[r]g^{(m)}[r]. \quad (1)$$

Коэффициенты усиления удовлетворяют условию  $g^{(m)}[r] \geq 1, \forall r$  и обеспечивают ослабление эффекта маскирования полезного сигнала шумовым. Для того, чтобы избежать появления слышимых артефактов при состыковке коэффициентов  $g^{(m)}[r]$  необходимо, чтобы последовательность коэффициентов была без разрывов на границах соседних фреймов, что достигается путем интерполяции.

Адаптивный КДД, предлагаемый в данной работе, имеет изменяющуюся амплитудную характеристику (см. рис. 6), которая зависит от уровня шума  $W_{nos}$  (дБ) и определяется двумя параметрами:

1)  $\Delta G$  (дБ) определяет динамический диапазон сигналов (в зависимости от уровня шума), к которым применяется усиление и компрессия;

2)  $\Delta R$  (дБ) совместно с  $\Delta G$  определяет коэффициент компрессии (угол наклона амплитудной характеристики).

Особенностью работы адаптивного КДД является то, что в случае отсутствия шума (в тихой обстановке), он не вносит искажений в полезный сигнал (характеристика имеет вид прямой с углом наклона  $45^\circ$ ). Для слежения за мощностью шумового сигнала используется схема экспоненциального усреднения:

$$W = 10 \log_{10} \left( \frac{1}{R} \sum_{r=1}^R (w^{(m)}[r])^2 \right),$$

$$W_{nos}^{(m)} = \alpha W_{nos}^{(m-1)} + (1 - \alpha)W, \quad 0 < \alpha < 1,$$

где  $W_{nos}^{(m)}$  – мощность текущего фрейма,  $W_{nos}^{(m-1)}$  – мощность предыдущего фрейма,  $\alpha$  – параметр усреднения.



Рис.6. Амплитудная характеристика адаптивного КДД

Для текущего  $m$ -го фрейма полезного сигнала в адаптивном компрессоре рассчитывается мощность

$$X^{(m)} = 10 \log_{10} \left( \frac{1}{R} \sum_{r=1}^R (x^{(m)}[r])^2 \right),$$

которая вместе с текущей мощностью шума  $W_{nos}^{(m)}$  используется для определения выходного уровня сигнала  $Y^{(m)}$  (см. рис. 6). Коэффициент усиления для  $m$ -го фрейма рассчитывается следующим образом:

$$g_{avg}^{(m)} = 10^{(Y^{(m)} - X^{(m)})/20}.$$

Корректирующие коэффициенты  $g^{(m)}[r]$ , применяемые к входному сигналу (1) компрессора, рассчитываются путем интерполяции коэффициентов усиления текущего, последующего и предыдущего фреймов (т.е.  $g_{avg}^{(m-1)}$ ,  $g_{avg}^{(m)}$  и  $g_{avg}^{(m+1)}$ ). При этом каждый коэффициент  $g_{avg}$  ассоциируется с центром соответствующего фрейма. Процедура интерполяции показана на рис. 7.

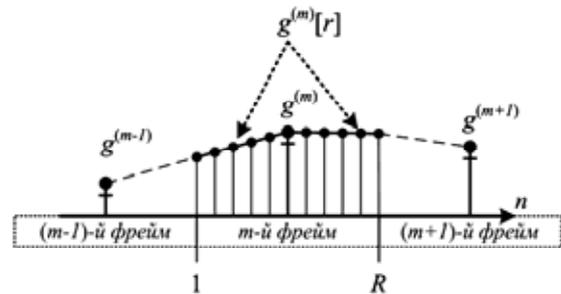


Рис. 7. Интерполяция коэффициентов усиления

Общая схема обработки сигнала в блоке адаптивного компрессора представлена на рис. 8.

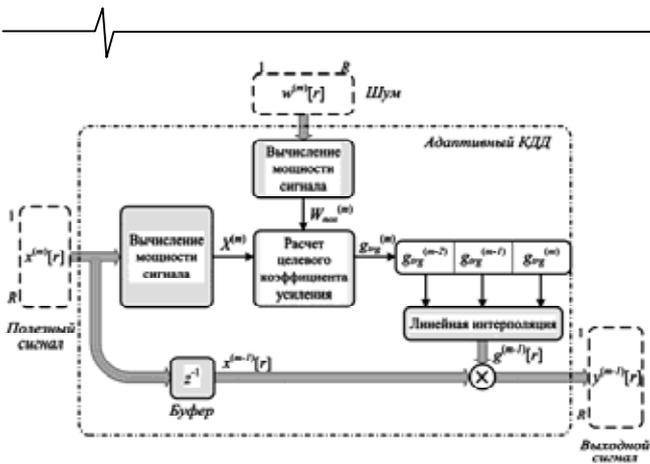


Рис. 8. Блок адаптивного КДД

Достоинством предлагаемого адаптивного КДД является то, что он позволяет ослабить эффект маскирования тихих звуков сигнала, не искажая громких звуков, которые не маскируются акустическим шумом.

**Неравнополосный банк фильтров**

Для реализации субполосного метода повышения качества звукового сигнала (рис. 4) используется неравнополосный косинусно-модулированный банк фильтров (НКМБФ), который аппроксимирует психоакустическую частотную шкалу барков [7]. В [8] показано, что повышение разборчивости речевого сигнала достигается путем разбиения входного сигнала на небольшое число субполос с последующей независимой компрессией каждой из них. Установлено [8], что максимизация разборчивости происходит уже при разбиении сигнала на 5 каналов, а дальнейшее увеличение их числа разборчивости практически не вызывает. Кроме того, небольшое количество полос позволяет избежать эффекта «выбеливания» и сохранить натуральное звучание полезного. Исходя из этих фактов и стремясь к уменьшению вычислительных затрат в практической реализации схемы обработки сигнала (рис. 4), предлагается использовать 5-канальный НКМБФ.

В банке фильтров используется принцип модуляции [9], позволяющий получить из одного фильтра-прототипа  $h[n]$  импульсные характеристики полосовых фильтров  $h_k[n]$ , равномерно покрывающих весь частотный диапазон. Импульсные характеристики фильтров анализа  $h_k[n]$  и синтеза  $f_k[n]$  имеют следующий вид

$$h_k[n] = 2h[n] \cos\left( (2k+1) \frac{\pi}{2M} \left( n - \frac{N-1}{2} \right) + (-1)^k \frac{\pi}{4} \right),$$

$$f_k[n] = 2h[n] \cos\left( (2k+1) \frac{\pi}{2M} \left( n - \frac{N-1}{2} \right) - (-1)^k \frac{\pi}{4} \right),$$

где  $k = 0, 1, \dots, M-1$  – номер каналов банка фильтров;  $M$  – число каналов банка фильтров, а  $N$  – длина фильтра-прототипа, которая должна удовлетворять уравнению  $N = 2sM$ , где  $s$  – произвольное натуральное число.

Преимуществом такого способа формирования импульсных характеристик является то, что для фильтрации сигнала существует эффективная схема (полифазная реализация). Схема основывается на следующих положениях:

- 1) вычисление выходных сигналов выполняется одновременно для всех каналов;
- 2) модулирующая функция является симметричной

$$c(k, n) = \cos\left( (2k+1) \frac{\pi}{2M} \left( n - \frac{N-1}{2} \right) \pm (-1)^k \frac{\pi}{4} \right),$$

для которой из свойств периодичности косинуса имеем

$$c(k, l + 2pM) = (-1)^p c(k, l). \tag{2}$$

Для наглядности описания реализации необходимо выполнить запись в терминах  $z$ -преобразования:

$$\begin{bmatrix} H_0(z) \\ H_1(z) \\ \vdots \\ H_{M-1}(z) \end{bmatrix} = \begin{bmatrix} c(0,0) & c(0,1) & \dots & c(0,N-1) \\ c(1,0) & c(1,1) & \dots & c(1,N-1) \\ \vdots & \vdots & \ddots & \vdots \\ c(M-1,0) & c(M-1,1) & \dots & c(M-1,N-1) \end{bmatrix} \times \tag{3}$$

$$\times \text{diag}(1, z^{-1}, \dots, z^{-N}) \cdot \begin{bmatrix} h[0] \\ h[1] \\ \vdots \\ h[N-1] \end{bmatrix},$$

где  $H_{0,M-1}(z)$  –  $z$ -преобразования импульсных характеристик полосовых фильтров.

В процессе фильтрации входной сигнал умножается на коэффициенты фильтра-прототипа и на матрицу косинусной модуляции  $M_C$ . Используя (2), можно показать, что данная матрица обладает регулярной структурой

$$M_C = \underbrace{\begin{bmatrix} C & -C & C & -C & \dots \end{bmatrix}}_{m \text{ повторов}},$$

где  $C = [c(k, n)]_{0 \leq k < M, 0 \leq n < 2M}$ , и может быть представлена в виде произведения

$$M_C = C \cdot \underbrace{\begin{bmatrix} I & -I & I & -I & \dots \end{bmatrix}}_{m \text{ повторов}} = C \cdot A, \tag{4}$$

где  $I$  – единичная матрица размера  $2M \times 2M$ . Факторизация (4), по сути, представляет запись быстрого алгоритма умножения на матрицу  $M_C$ . При этом количество требуемых операций умножения уменьшается в  $s$  раз, поскольку умножение на матрицу  $A$  требуют только операций сложения. Учитывая (4), выражение (3) можно переписать в следующем виде

$$\begin{bmatrix} H_0(z) \\ H_1(z) \\ \vdots \\ H_{M-1}(z) \end{bmatrix} = C \cdot \begin{bmatrix} h[0] - z^{-2M} h[2M] + \dots \\ z^{-1} h[1] - z^{-(2M+1)} h[2M+1] + \dots \\ \vdots \\ z^{-(2M-1)} h[2M-1] - z^{-(4M+1)} h[4M-1] + \dots \end{bmatrix} =$$

$$= \mathbf{C} \cdot \begin{bmatrix} \sum_{p=0}^{m-1} d[2Mp]z^{-2Mp} \\ \sum_{p=0}^{m-1} d[2Mp+1]z^{-(2Mp+1)} \\ \vdots \\ \sum_{p=0}^{m-1} d[2M(p+1)-1]z^{-(2M(p+1)-1)} \end{bmatrix} = \mathbf{C} \cdot \mathbf{P}, \quad (5)$$

где  $d(q+2pM) = (-1)^p h[q+2pM]$ ,  $0 \leq q < 2M$ ,  $0 \leq p < m$ . Элементами матрицы  $\mathbf{P}$  являются полифазные компоненты фильтра-прототипа. Дополнительное уменьшение вычислительной сложности косинусно-модулированного банка фильтров (КМБФ) заключается в поиске эффективного способа умножения на матрицу  $\mathbf{C}$ , которое может быть выполнено с использованием дискретного косинусного преобразования [7, 10].

Выражение (5) служит основой для построения полифазной реализации косинусно-модулированного банка фильтров, которая показана на рис. 9.

Выше описан способ получения эффективной реализации только анализирующей части банка фильтров. Используя аналогичные построения можно получить поли-

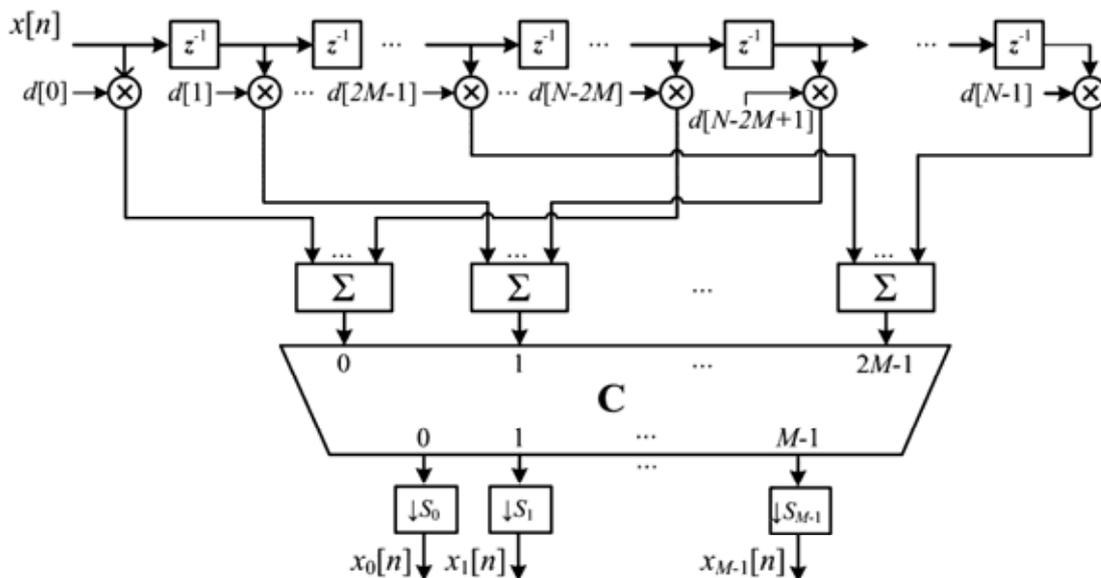


Рис. 9. Полифазная реализация косинусно-модулированного банка фильтров анализа

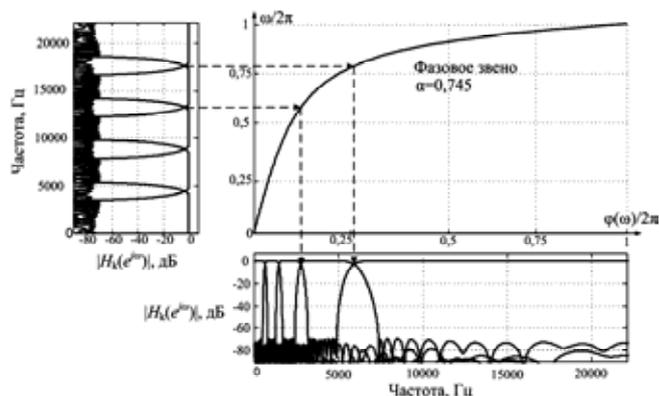


Рис. 10. Формирование неравнополосного банка фильтров (частота дискретизации 44,1 кГц)

фазную реализацию и для банка фильтров синтеза [7].

Для получения неравнополосной версии банка фильтров необходимо все элементы задержки структуры на рис. 9 заменить на фазовые звенья [7]:

$$A(z) = \frac{\alpha + z^{-1}}{1 + \alpha z^{-1}}, \quad |\alpha| \leq 1, \quad A(e^{j\omega}) = e^{j\varphi(\omega)},$$

где

$$\varphi(\omega) = -\omega + 2 \arctg(\alpha \sin \omega / (\alpha \cos \omega - 1))$$

Замена  $z^{-1} \rightarrow A(z)$  вызывает нелинейное отображение частотной оси  $\omega \mapsto \varphi(\omega)$ , благодаря которому банк фильтров становится неравнополосным. Процесс синтеза 5-канального банка фильтров для практической реализации предлагаемого метода повышения качества звукового сигнала показан на рис. 10.

Коэффициент деформации частотной оси  $\alpha = 0,745$  (при частоте дискретизации 44,1 кГц) позволяет получить банк фильтров, у которого частотные полосы имеют равную ширину в шкале барков. Для банка фильтров также определены коэффициенты децимации/интерполяции  $S_k = \{21, 7, 3, 1, 1\}$ . Подробности процедуры расчета НКМБФ можно найти в работах [6, 7].

### Экспериментальные исследования

На рис. 11 показан результат обработки сигнала предлагаемой схемой (рис. 4), причем использован тот же сигнал, который был показан ранее на рис. 1. Сопоставляя спектрограммы речи до (рис. 1, а) и после (рис. 11, а) обработки со спектральными характеристиками шума (рис. 11, б), можно заметить, что усиление частотных компонент речи тем больше, чем больше интенсивность шума в соответствующей полосе. Также видно, что усиление полезного сигнала плавно возрастает с усилением интенсивности шума, что сохраняет для пользователя комфортное соотношение сигнала и шума при изменении уровня зашумленности внешней среды. Из рис. 11, в видно, что в отличие от (рис. 1, в), все компоненты речи легко различимы при прослушивании.

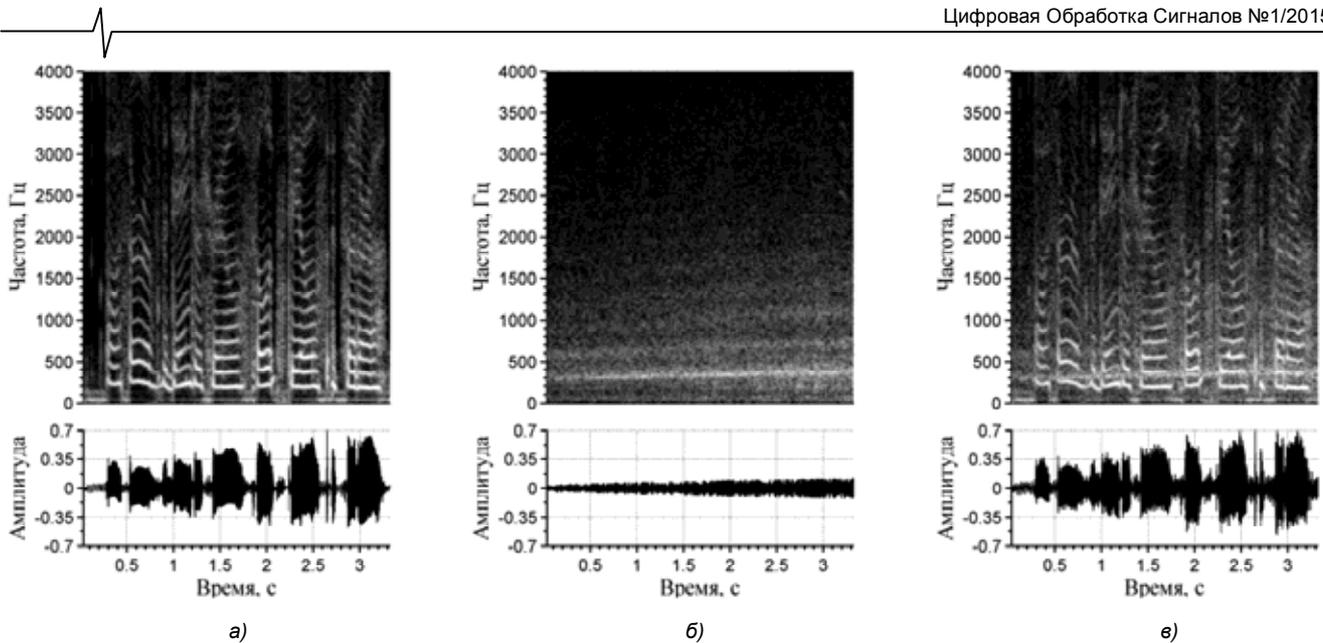


Рис. 11. Частотно-временное представление сигнала:

а) обработанная чистая речь; б) шум метрополитена; в) сумма обработанного сигнала и шума

Объективная оценка предлагаемого метода повышения качества восприятия звукового сигнала выполнена при помощи индекса разборчивости речи (ИРР) [11]. ИРР представляет собой объективную величину, которая коррелирует с разборчивостью речи в различных неблагоприятных условиях прослушивания. ИРР основан на вычислении энергии сигнала в критических частотных полосах. При расчете ИРР анализируется эквивалентный уровень спектра речевого сигнала и эквивалентный уровень спектра шумового сигнала в каждой из критических полос, вносящих вклад в разборчивость. Эквивалентный уровень рассчитывается, как усредненная по времени мощность сигнала в критической полосе, поделенная на ширину полосы. Поскольку в рассматриваемой задаче отдельно доступны и полезный сигнал и шумовой, то ИРР можно вычислить достаточно легко. Величина ИРР принимает значение от 0 до 1. Считается, что ИРР выше 0,75 является хорошим показателем, в то время как значение меньше 0,45 является неудовлетворительным.

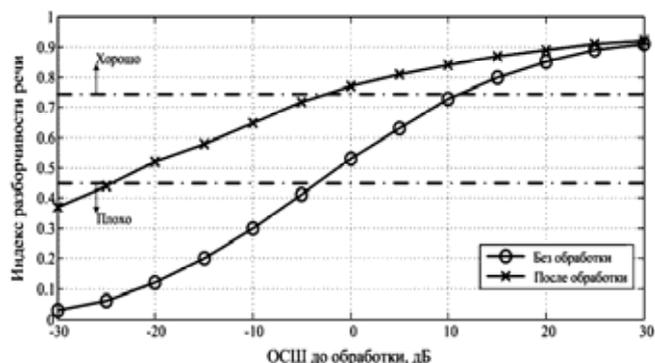


Рис. 12. Сравнение среднего значения ИРР для речевых сигналов

до и после обработки предложенным методом

Речевой материал для выполнения экспериментов взят из свободно распространяемой речевой базы CMU ARCTIC [12]. В качестве шумового сигнала использована запись акустической обстановки в метрополитене, запи-

санная при помощи смартфона. Работа предложенного способа проверялась для различных ОСШ, во всех случаях измерялась разборчивость речевого сигнала на фоне шума до обработки и после обработки. Полученные результаты экспериментов показаны на рис. 12.

В агрессивной шумовой обстановке (ОСШ < -20 дБ) метод обеспечивает удовлетворительную разборчивость, а при снижении интенсивности шума (ОСШ > 0 дБ) – хорошую. Необработанный сигнал имеет заметно меньшую разборчивость для ОСШ < 0 дБ и становится неразборчивым уже при ОСШ < -10 дБ.

Предлагаемая схема обработки сигнала реализована в виде приложения для iPhone<sup>2</sup>.

### Заключение

В работе предложена схема субполосной обработки звуковых сигналов для улучшения их восприятия в шумной акустической обстановке. Обработка выполняется при помощи неравнополосного банка фильтров, аппроксимирующего разрешающую способность слуховой системы, за счет чего достигается натуральное звучание выходного сигнала. Каждая субполоса полезного сигнала обрабатывается адаптивным компрессором динамического диапазона, который регулируется в зависимости от текущей энергии шума в этой субполосе. Эффективность применения предложенной схемы обработки сигнала подтверждается субъективными наблюдениями и объективными оценками на основе индекса разборчивости речи. Данная схема обработки может применяться в мобильных телекоммуникационных устройствах имеющих внешний микрофон для регистрации шума окружающей среды.

<sup>2</sup> Сайт приложения доступен по ссылке: <http://urbandenoiser.com/>

## Литература

1. Sauert B., Vary P. Near and listening enhancement: Speech intelligibility improvement in noisy environments // Processing of International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Toulouse, France, 14–19 May, 2006. – Toulouse, 2006. – P. 493–496.
2. Sauert B., Vary P. Near and listening enhancement optimized with respect to speech intelligibility index // Processing of 17th European Signal Processing Conference (EUSIPCO 2009), Glasgow, Scotland, 24–28 August, 2009. – Glasgow, 2009. – P. 1844–1848.
3. Zorila, T.-C. Speech-in-noise intelligibility improvement based on spectral shaping and dynamic range compression / T.-C. Zorila, V. Kandida, Y. Stylianou // In Proc. Interspeech, 2012. – Portland, Oregon, 2012. – P. 635–638
4. Фонлантен А. Слуховые аппараты / А. Фонлантен, А. Хорст.– Ростов н/Д.: Феникс, 2009. – 304 с.
5. Blesser B.A. Audio dynamic range compression for minimum perceived distortion / B.A. Blesser // IEEE Trans. Audio and Electroacoustics. – 1969. – vol. 17. – no. 1. – P. 22-32.
6. Вашкевич, М.И. Проектирование передискретизированного неравнополосного косинусно-модулированного банка фильтров / М.И. Вашкевич, А.А. Петровский // Информатика. – 2011. – № 2 (30). – С. 21–39.
7. Вашкевич М.И. Косинусно-модулированные банки фильтров с фазовым преобразованием: реализация и применение в слуховых аппаратах / М.И. Вашкевич, И.С. Азаров, А.А. Петровский – М.: Горячая линия-Телеком, 2014. – 210 с.
8. Trine T.D., Tasell D.V. Digital hearing aid design: facts vs. fantasy // The Hearing Journal. – 2002. – Vol. 55, №2. – P. 36–42.
9. Koilpillai D. Cosine-modulated FIR filter banks satisfying perfect reconstruction / D. Koilpillai, P.P. Vaidyanathan // IEEE transaction on signal processing. – 1992. – Vol. 40, № 4. – P. 770–783.
10. Вашкевич, М.И. Применение полиномиальных алгебр и теории Галуа для синтеза быстрых алгоритмов дискретных косинусных преобразований / М.И. Вашкевич, А.А. Петровский // Цифровая обработка сигналов. – 2011. – № 3. – С. 2–10.
11. American National Standard. Methods for the Calculation of the Speech Intelligibility Index. ANSI S3.5-1997, 1997.
12. Kominek J., Black A. The CMU ARCTIC speech databases for speech synthesis research / Language Technologies Institute, Carnegie Mellon University, Pittsburgh, PA, Tech. Rep. CMU-LTI-03-177, 2003.

### LISTENING ENHANCEMENT IN NOISY ENVIRONMENT BASED ON SPECTRAL DECOMPOSITION AND ADAPTIVE DYNAMIC RANGE COMPRESSION OF THE SIGNAL

*Azarov E.S., Vashkevich M.I., Lihachev D.S., Petrovsky A.A.*

The paper investigates possibility of automatic listening enhancement in noisy environment. In order to improve subjective quality time-frequency components of the signal are nonlinearly amplified according to spectral components of surrounding noise. The method, proposed in the paper, is applicable to different signal including speech and audio. Practical value of the method is evaluated for speech signals using objective measures.

### Уважаемые коллеги!

*Приглашаем Вас принять участие в формировании тематических выпусков журнала «Цифровая обработка сигналов» и размещению рекламы продукции (услуг) Вашей организации на его страницах. В случае положительного решения просим представить в редакцию журнала Ваши предложения по плановому размещению информационных материалов и макет рекламы продукции (услуг) с указанием желаемого её месторасположения: обложка (2-я, 3-я или 4-я стр.), цветная внутренняя полоса (объем полосы).*

Журнал «Цифровая обработка сигналов» издается с 1999 года. Выходит ежеквартально, тиражом – 700 экз. Распространяется по подписке через агентство «Роспечать» в России (индекс 82185), СНГ и странах Балтии (индекс 20630), а также на Конференции: «Цифровая обработка сигналов и ее применение – DSPA'».

Научно-технический журнал «Цифровая обработка сигналов» включен в Перечень изданий, рекомендуемый ВАК РФ для публикации результатов научных исследований соискателями ученой степени доктора и кандидата технических наук в области радиотехники, связи, вычислительной техники, электроники, приборостроения, информационных технологий, информационно-измерительных и управляющих систем. По предварительным итогам за 2014 год по рейтингу Science Index базы РИНЦ (3,394) журнал «Цифровая обработка сигналов» занимает 344-ю позицию из почти 3000 представленных изданий.

#### Планируемые сроки издания отдельных номеров журнала:

№ 1 – апрель 2015 г. Тематический выпуск: «ЦОС в информационно-измерительных системах»

№ 2 – июнь 2015 г. Тематический выпуск: «Теория и методы цифровой обработки сигналов» (по материалам международной научно-технической конференции «Цифровая обработка сигналов и ее применение – DSPA' 2015».

№ 3 – сентябрь 2015 г. Тематический выпуск: «Цифровая обработка изображений».

№ 4 – декабрь 2015 г. Тематический выпуск: «ЦОС в радиотехнике и системах телекоммуникаций».

#### Ориентировочная стоимость рекламных услуг:

4-я (внешняя) страница цветной обложки – 25 тысяч рублей.

2-я и 3-я (внутренние) страницы цветной обложки – 15 тысяч рублей.

1/2 цветной внутренней полосы – 8 тысяч рублей.

Ждем Ваших предложений.

С наилучшими пожеланиями, зам. главного редактора  
д.т.н., профессор Витязев Владимир Викторович, телефон 8-903-834-81-81.

Предложения прошу направлять по адресу: E-mail: vityazev.v.v@rsreu.ru или info@dspa.ru