

УДК 621.391; 519.21; 519.245

ã Г. Ф. Малыхина, А. В. Меркушева

СОГЛАСОВАННОЕ МНОГОКАНАЛЬНОЕ РАЗДЕЛЕНИЕ СИГНАЛА С ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЕМ И МОДИФИЦИРОВАННОЙ ПОЛИФАЗНОЙ СТРУКТУРОЙ ДЛЯ СПЕКТРАЛЬНОГО ВЫЧИТАНИЯ ФОНА РЕЧЕВОГО СИГНАЛА

Метод спектрального вычитания используется для улучшения качества речевого сигнала (РС), т. к. этот метод способен существенно уменьшить или уничтожить аддитивный шум, присутствующий в сигнале. В отличие от использованного ранее преобразования с равномерной шириной каналов рассматриваются два варианта структуры, в которых деление полного диапазона частот сигнала осуществлено на полосы возрастающей ширины. Такое деление диапазона частот РС лучше соответствует виртуальной системе фильтров слухового аппарата человека, т. е. так называемой перцептуальной модели, отражающей особенности восприятия РС и других звуковых сигналов. Рассмотрены две структуры: одна — на основе дискретного вейвлет-преобразования, вторая — неоднородная система согласованного многоканального разделения на основе полифазной структуры.

ВВЕДЕНИЕ

Обычное спектральное вычитание фона (СВФ)¹⁾ с аналитической точки зрения сводится к введению спектральных весовых коэффициентов, видоизменяющих частотную форму (преобразование Фурье) зашумленного речевого сигнала (РС). Фактически метод реализует вычитание средней величины спектра шума из спектра РС и используется, когда искажение РС происходит за счет аддитивного шума с медленноменяющимися спектральными характеристиками. При этом выполняются прямое и обратное дискретное преобразование Фурье (ДПФ) в форме кратковременного ДПФ с частичным перекрытием "окна" преобразования. Более совершенным является метод СВФ, основанный на структурах согласованного многоканального разделения (СМР) сигнала [1, 2]. Для равномерной (по спектру) ширины каналов этот метод использовался Смитом, Эдинсом (Smith, Eddins) [3] и Аровудом (Arrowood), Смитом [4]. Метод СВФ с неравномерными по частоте каналами, приспособленными к перцептуальной модели²⁾, которая отражает особенности восприятия

звука человеком, использован в [6]. В этой работе СВФ реализовано на основе вейвлет-преобразования и предусмотрена адаптация к изменениям уровня и спектральной структуры фона. СВФ с использованием неравномерных каналов СМР и вейвлет-преобразования выполнено также Лангом, Гуо, Одегардом (Lang, Guo, Odegard) [7]. Однако процедура СВФ проведена только для семи каналов, а для коэффициентов вейвлет-преобразования использована только простая "жесткая" пороговая дискриминация. Применение полужесткого порога для вейвлет-коэффициентов в сочетании с СМР было предложено Сеоком и Бае (Seok, Bae) [8].

В статье проанализированы две структуры согласованного многоканального разделения с неравномерной шириной каналов (ориентированных на перцептуальную модель РС): структура с использованием дискретного вейвлет-преобразования и полифазная структура [6], которая модифицирована за счет применения широкодиапазонных преобразователей вместо элементов задержки (обычно используемых в системах СМР) между входными сигналами отдельных каналов.

В этих структурах частотные диапазоны каналов организованы так, чтобы получить шкалу в барках (см. сноску 2)) и в области низких частот РС иметь возможность осуществлять более эффективное спектральное вычитание фона. Ширина

шкалой частот в барках. Это разбиение в некоторой степени сходно с логарифмической шкалой, но не совпадает с ней.

¹⁾ Иногда используется термин "спектральное выравнивание зашумленного речевого сигнала".

²⁾ Перцептуальная модель отражает неравномерное (по частотному спектру РС) распределение ширины каналов, причем последовательность и ширина каналов соответствуют дискретному анализатору звуковых частот, реализованному так называемой улиткой среднего уха человека [5]. Разбиение частотного диапазона РС в соответствии с перцептуальной моделью называют

полос в низкочастотной области РС (и количество каналов СМР в ней) выбраны таким способом, чтобы компоненты шума в интервале между гармоническими элементами сегментов РС могли быть взвешены по отдельности.

Кроме того, неточность воспроизведения РС на выходе структуры СМР должна быть мала с точки зрения субъективной оценки качества сигнала, переданного через систему СМР. Иначе говоря, РС, прошедший стадии анализа и синтеза в СМР без спектрального взвешивания, должен быть идентичен сигналу, поступившему на вход системы, т. е. по крайней мере не иметь слышимых искажений РС.

В связи с указанными требованиями к системе СМР полное восстановление РС не является безусловно необходимым, также как и прореживание частоты на входных цепях каналов СМР. Это позволяет при построении системы СМР найти рациональный баланс между допустимой ошибкой восстановления РС, частотным разрешением системы, достаточным для спектрального вычитания фона, и вычислительной сложностью алгоритма, который определяет совокупность операций, необходимых для реализации системы.

СПЕКТРАЛЬНОЕ ВЫЧИТАНИЕ ФОНА

При спектральном вычитании фона часто используется модель РС, загрязненного аддитивным гауссовым шумом: $x(k) = s(k) + n(k)$, где $x(k)$ — регистрируемый РС, $s(k)$ — фактический РС, $n(k)$ — шум, k — дискретное время отсчета сигналов. Частотный спектр РС со сниженным уровнем фона $\hat{S}(e^{j\omega})$ получается вычитанием оценки средней величины фона $|\overline{N(e^{j\omega})}|$ из величины спектра $|X(e^{j\omega})|$ регистрируемого РС³⁾:

$$\hat{S}(e^{j\omega}) = \left(|X(e^{j\omega})| - |\overline{N(e^{j\omega})}| \right) e^{j j_x(\omega)}, \quad (1)$$

где $j_x(\omega)$ — фаза регистрируемого РС; $X(e^{j\omega})$, $N(e^{j\omega})$ и $\hat{S}(e^{j\omega})$ — преобразования Фурье (ПФ) регистрируемого РС, шума и сигнала, полученного после спектрального вычитания фона. Средняя величина спектра шума $\overline{N(e^{j\omega})}$ оценивается на интервалах естественных микропауз РС, которые могут детектироваться с помощью нейронной сети [9].

Таким образом, модифицируется только вели-

чина $X(e^{j\omega})$, а фаза сохраняется, какой была до операции спектрального вычитания (СВ). Этот метод СВ основан на том, что ухо человека относительно нечувствительно к возмущению фазы [10], а способ вычитания фазы пока не формализован.

Альтернативный способ оценивания среднего уровня спектра шума (без детектирования пауз) основан на слежении за минимумом сглаженной величины спектра зашумленного РС в некотором временном окне [11, 12]. Длина окна выбирается так, чтобы на нем присутствовали паузы РС. Соотношение (1) можно интерпретировать как спектральное взвешивание зашумленного РС:

$$\hat{S}(e^{j\omega}) = G(e^{j\omega}) X(e^{j\omega}), \quad (2)$$

где

$$G(e^{j\omega}) = \frac{|X(e^{j\omega})| - |\overline{N(e^{j\omega})}|}{|X(e^{j\omega})|}. \quad (3)$$

Поэтому оценка спектра РС, свободного от шума, выполняется путем применения взвешивания, зависящего от спектра регистрируемого сигнала.

Получение в (3) отрицательного значения $G(e^{j\omega})$ трактуется как ошибка оценки уровня фона. Для того чтобы такое $G(e^{j\omega})$ не входило в схему СВ, вместо (3) используется соотношение (4) [13]:

$$G(e^{j\omega}) = \max \left(\frac{|X(e^{j\omega})| - |\overline{N(e^{j\omega})}|}{|X(e^{j\omega})|}, 0 \right) = \max \left(1 - \frac{|\overline{N(e^{j\omega})}|}{|X(e^{j\omega})|}, 0 \right). \quad (4)$$

Для выполнения этого метода входной РС разделяется на мелкие сегменты по 10–50 мс, и применяется спектральное вычитание в пределах каждого сегмента. Чтобы не было артефактов СВ на границах сегментов, сегментация РС делается с наибольшим перекрытием и так же осуществляется восстановление сигнала после процедуры СВ.

Методы СВ в описанной выше форме дают низкоуровневый остаточный шум РС, и для его подавления принимают дополнительные меры:

— частотный спектр входного сигнала может быть немного сглажен для снижения дисперсии компоненты шума;

— используется небольшая переоценка среднего спектра шума умножением его на число, несколько большее единицы, и этот коэффициент может видоизменяться вместе с уровнем шума;

— применяется порог для минимума оцененного спектра сигнала, так что оставшиеся спектральные пики маскируются даже небольшим уровнем широкополосного шума.

³⁾ Частотный спектр (частотная форма сигнала) представляет преобразование Фурье и обозначается прописной буквой, соответствующей обозначению сигнала.

Общим для этих подходов к совершенствованию метода СВ является то, что при уменьшении уровня остаточного шума несколько возрастает искажение РС на выходе или становится меньшим снижение дополнительно возникающих шумов на стадии операций СВ.

Более сложная схема метода СВ предложена Капеланом, Страусом и Вари (Kapelan, Straus, Vary) [14]. Определяются апостериорное отношение сигнал/шум

$$R_{\text{апостер.}b}(e^{ij}) = \frac{|X_b(e^{ij})|^2}{|N_b(e^{ij})|^2} - 1 \quad (5)$$

и априорное отношение сигнал/шум

$$R_{\text{априор.}b}(e^{ij}) = (1-q) \cdot \max(R_{\text{апостер.}b}(e^{ij}), 0) + q \cdot \frac{|G_{b-1}(e^{ij}) \cdot X_b(e^{ij})|^2}{|N_b(e^{ij})|^2}. \quad (6)$$

G_{b-1} означает спектральные веса предыдущего сегмента РС, а b — индекс (номер) сегмента. Тогда спектральные веса текущего сегмента РС выражаются соотношением (7):

$$G_b(e^{ij}) = \frac{\sqrt{p}}{2} \times \sqrt{\left(\frac{1}{1+R_{\text{апостер.}b}(e^{ij})} \right) \left(\frac{R_{\text{априор.}b}(e^{ij})}{1+R_{\text{априор.}b}(e^{ij})} \right)} \times M \left[\left(1+R_{\text{апостер.}b}(e^{ij}) \right) \left(\frac{R_{\text{априор.}b}(e^{ij})}{1+R_{\text{априор.}b}(e^{ij})} \right) \right]. \quad (7)$$

Функция $M[u]$ определяется выражением

$$M[u] = \exp\left(-\frac{u}{2}\right) \cdot \left[\left(1+u\right) \cdot I_0\left(\frac{u}{2}\right) + u \cdot I_1\left(\frac{u}{2}\right) \right], \quad (8)$$

где функции I_0 и I_1 — модифицированные функции Бесселя первого и второго порядка.

НЕЛИНЕЙНЫЙ СПЕКТРАЛЬНЫЙ АНАЛИЗ

Вейвлет-преобразование и структура СМР на его основе

Для спектрального анализа нестационарного процесса наиболее часто используется кратковременное преобразование Фурье (КПФ)

$$F_x^w(t, w) = \int_{-\infty}^{+\infty} x(t) \cdot w(t-t)e^{-iwt} dt. \quad (9)$$

Выбор (постоянной) длины окна этого преобразо-

вания определяется приемлемым компромиссом между величиной разрешения по времени и по частоте, а так называемая "ячейка разрешения" ⁴⁾ имеет одинаковый размер на всей время-частотной плоскости. Разрешение по частоте у КПФ оказывается одинаковым на всем частотном диапазоне спектра сигнала, и это основной недостаток КПФ.

Лучшими (сравнительно с КПФ) характеристиками разрешения обладает вейвлет-преобразование (ВП). Непрерывное ВП сигнала $x(t)$ определяется с помощью единой функции ψ со смещением ее по времени и масштабированием ⁵⁾:

$$W_x^y(b, a) = |a|^{-1/2} \int_{-\infty}^{+\infty} x(t) \psi\left(\frac{t-b}{a}\right) dt. \quad (10)$$

Функция ψ служит прототипом вейвлет-базиса, который получается смещением времени и масштабированием с помощью параметров b и a . При больших значениях a базисные функции $y_{b,a}(t) = |a|^{-1/2} \psi((t-b)/a)$ становятся низкочастотными, при малых a $y_{b,a}(t)$ — высокочастотными, и в разложении (10) они представляют соответствующую область спектра сигнала $x(t)$. Разрешение во время-частотной плоскости не постоянное: в области высоких частот ВП имеет высокое разрешение по времени и небольшое по частоте; в области низких частот — высокое разрешение по частоте и небольшое разрешение по времени.

В частотной области ВП может интерпретироваться как СМР с полосовым пропусканием Δw_i , которое возрастает одновременно с возрастанием центральной частоты w_0 этой полосы. При этом величина относительной полосы $Q = \Delta w_i / w_0$ не зависит от параметра a . В связи с этим использование ВП называют анализом с постоянной величиной относительной разрешающей способности (с постоянным качеством Q). Для речевого сигнала такая форма анализа сходна со структурой частотного анализа, производимого слуховым аппаратом человека.

Избыточность ВП по непрерывным параметрам a и b снимается методом их дискретизации: $a_m = 2^m$, $b_{mn} = a_m nT$, где m, n — целые, T — интервал дискретизации. При этом (10) приобретает форму (11), которая наиболее часто применяется в ВП:

⁴⁾ Ячейку разрешения определяет произведение среднеквадратичных рассеяний по времени и по частоте ($s_t \cdot s_w$). Метод оценки ячейки разрешения время-частотных (включая КПФ) и время-масштабных (вейвлет) преобразований рассмотрен в [15, 16].

⁵⁾ Функция $\psi(t)$ — прототип вейвлет-базиса обладает определенными свойствами и удовлетворяет некоторым условиям [17, 18].

$$W_x^y(b_{mn}, a_m) = |a|^{-m/2} \int_{-\infty}^{+\infty} x(t) y(2^{-m}t - nT) dt. \quad (11)$$

Для повышения разрешающей способности по частоте ВП возможно использовать $1 < a_0 < 2$, однако при этом теряется бинарное свойство масштабирования базисных функций. Другой более эффективный способ повышения частотной разрешающей способности в M раз состоит в использовании

"расщепленной" функции ψ (прототипа базиса) в виде

$$y^j(t) = 2^{-j/2M} y(2^{-j/2M}t), \quad j = 0, \dots, M-1. \quad (12)$$

Поскольку динамический спектральный анализ РС и других нестационарных сигналов производится с использованием их дискретной формы, для этих условий (11) принимает вид (13):

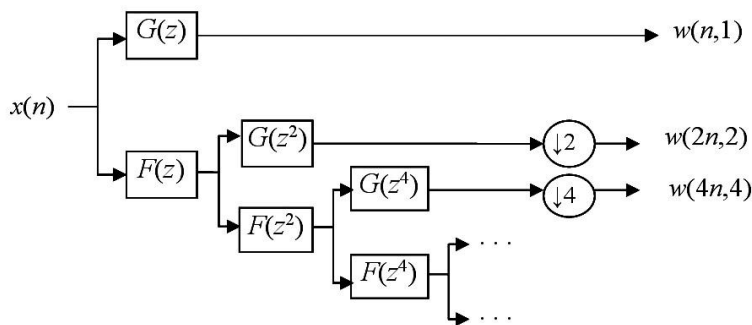


Рис. 1. Структура преобразований при реализации ВП по методу Малата—Шенсы

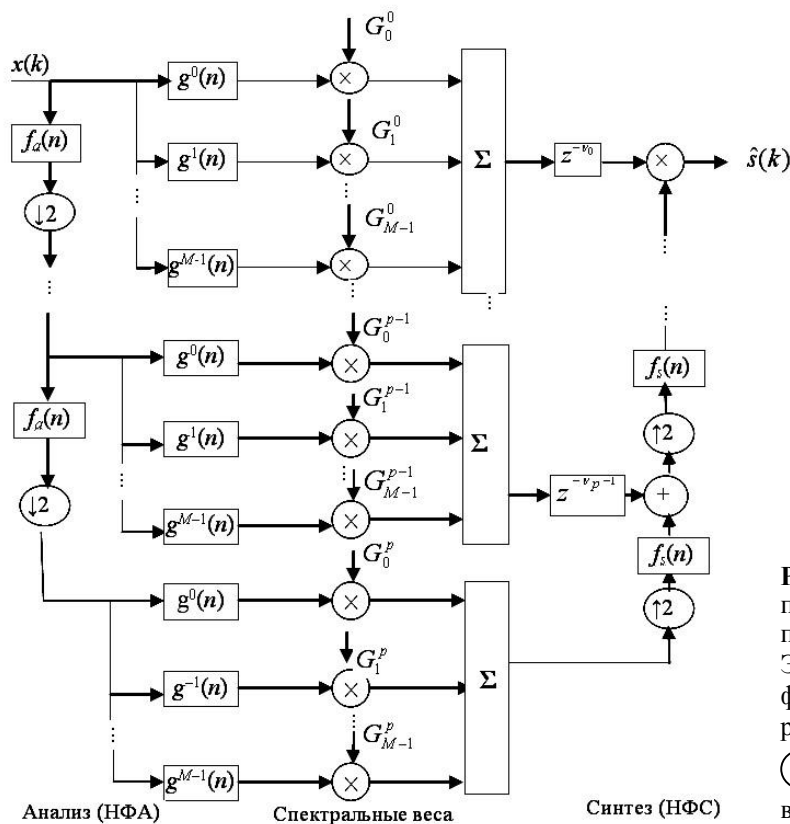


Рис. 2. Структура СМР для вейвлет-преобразования с высоким разрешением по частоте.

Элементы структуры СМР: НФА — набор фильтров анализа и НФС — набор фильтров синтеза описаны в [2] и [4].

$\downarrow 2$ и $\uparrow 2$ — символы операций прореживания частоты и ее восстановления интерполированием

$$W_x^y(2^m n, 2^m) = 2^{-m/2} \sum_k x(k) \cdot y(2^{-m} k - n), \quad (13)$$

где интервал дискретизации T принят за единицу⁶⁾.

Структура преобразований для реализации ВП дискретного сигнала по (13) (т. е. при обычном разрешении по частоте) является своеобразной СМР. Один из видов структуры, отражающий метод Малата—Шенсы (Mallat, Shensa) [19], представлен на рис. 1, где $G(z)$ — z -преобразование прототипа вейвлет-функции; $F(z)$ — передаточная функция фильтра интерполяции.

Для получения более высокого разрешения по частоте методом расщепления прототипа базиса в соотношении (13) вместо функции $y(k)$ следует использовать $y^j(k) = 2^{-j/2M} y(2^{-j/2M} k)$, $j = 0, \dots, M-1$ из соотношения (12). Поэтому для дискретного сигнала ВП с увеличенным в M раз разрешением по частоте должно использоваться выражение (14):

$$W_x^{y^j}(2^m n, 2^m) = 2^{-m/2} \times \sum_k x(k) \cdot 2^{-j/2M} y(2^{-j/2M} (2^{-m} k - n)). \quad (14)$$

При анализе РС с повышенным разрешением по частоте свойство полного восстановления сигнала на выходе структуры СМР позволяет реализовать 2^m -кратное прореживание частоты в m -октаве более эффективно (сравнительно с непосредственным вычислением по соотношению (13)). Для этого может служить структура СМР на основе ВП. Эта вейвлет-структура СМР для $(p+1)$ -октавного спектрального разложения сигнала, дополненная M -кратным расщеплением функции ψ (прототипа базиса), представлена на рис. 2. Для анализируемого сигнала она обеспечивает M -кратное повышение разрешения по частоте.

Функция низкочастотного фильтра $f_a(n)$ состоит в снижении эффекта Гиббса — межканального просачивания частот (элайзинга). Прореживание частоты РС на 2^l в октаве l позволяет использовать в каждой октаве одинаковые фильтры. В цепях набора фильтров синтеза (НФС) восстанавливается частота в канале каждой октавы на выходе СМР, и сигнал суммируется с выходом следующей октавы НФС. При этом производится компенсация запаздывания сигнала в цепях НФА путем введения блоков задержки $z^{-n_{p-1}}, z^{-n_{p-2}}, \dots, z^{-n_0}$. Выполнение синтеза на выходе СМР не обеспечивает 100 %-го полного восстановления сигнала, но качество по-

лучаемого РС вполне достаточно с точки зрения перцептуальной модели. Вместе с тем такая форма спектрального анализа с повышенным частотным разрешением имеет преимущество существенно меньшего объема вычислений (сравнительно с использованием общих аналитических соотношений).

Модифицированная структура СМР с полифазным преобразованием

Получение согласованного многоканального разделения сигнала с неравномерной шириной каналов (требуемой перцептуальной моделью для РС) может быть реализовано на основе полифазной структуры, в которой блоки задержки заменены широкополосными преобразователями (так называемыми "всепропускающими" фильтрами).

Аналитическое описание полифазной структуры СМР приводит к выражению для z -преобразования передаточной функции j -го канала в виде соотношения [1, 2, 4]

$$H_j(z) = \sum_{r=0}^{M-1} \sum_{p=0}^{L_p/M-1} \left\{ (pM + r) \cdot [z^{-1}]^{pM+r} \times \exp \left[-2p i \cdot \left(\frac{r}{M} \cdot j \right) \right] \right\}, \quad j = 0, 1, \dots, M-1, \quad (15)$$

где использованы широкополосные характеристики для преобразования РС в каждом канале на основе единого прототипа фильтра $h(k)$, $k = 0, 1, \dots, L_p - 1$.

Система СМР на основе модифицированной полифазной структуры с заменой блоков задержки на широкополосные преобразователи показана на рис. 3.

В представленной системе СМР на основе полифазной структуры выполнена замена элементов задержки z^{-1} на широкополосные преобразователи степени 1, согласно (16):

$$z^{-1} \rightarrow H(z) = \frac{a z + 1}{z + a}, \quad -1 < a < 1. \quad (16)$$

Это не изменяет величины передаточной функции (ПФ), но вызывает преобразование частоты ω в частотной форме ПФ [14]. Вид преобразования получается из (16) путем подстановки $z \rightarrow e^{i\omega}$. Это дает взаимосвязь первоначальной частоты ω_1 с нелинейно масштабированной частотой ω_2 после преобразования:

$$\omega_2 = 2 \arctg \left[\frac{1+a}{1-a} \cdot \tg \left(\frac{\omega_1}{2} \right) \right]. \quad (17)$$

При значениях $-1 < a < 0$ ширина частотных полос фильтров (— широкополосных преобразователей)

⁶⁾ Шаг временного смещения базисной функции при этом меняется по величине в каждой следующей октаве ВП (т. е. при изменении величины параметра m).

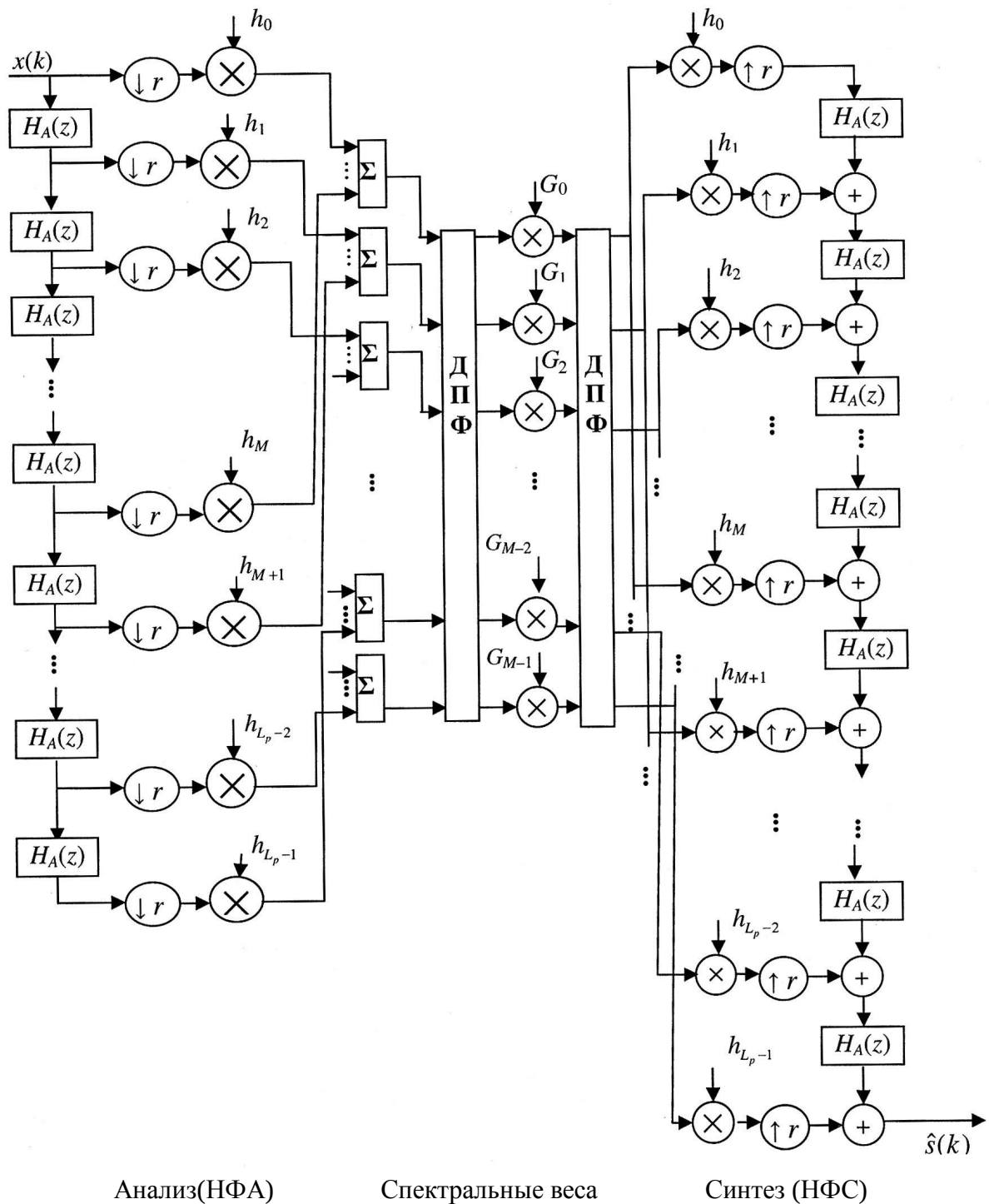


Рис. 3. Система СМР с полифазной структурой и высоким частотным разрешением

возрастает вместе с ростом их центральных частот. При $\alpha = 0$ получаются одинаковые ширины, а при значениях α из интервала $0 < \alpha < 1$ ширина

полос уменьшается по мере увеличения их центральных частот.

Аппроксимация спектра РС при масштабировании частоты в барках

Согласно перцептуальной модели, основанной на психоакустических экспериментах [5], частотный диапазон РС разделен на интервалы ("критические полосы") с примерно одинаковой относительной шириной — шириной, отнесенной к средней частоте интервала. Критические полосы ведут к концепции нелинейного преобразования шкалы частот. Результатом такого преобразования является частотная шкала η с единицей 1 барк. Масштаб шкалы выбран таким образом, что каждая критическая полоса имеет одинаковую ширину $\Delta h = 1$ барк. Преобразование диапазона частот РС в эту шкалу может быть аппроксимировано соотношением:

$$h [\text{барк}] = 13 \cdot \arctg(0.76 \cdot f [\text{кГц}]) + 3.5 \cdot \arctg \left[\left(\frac{f [\text{кГц}]}{7.5} \right)^2 \right]. \quad (18)$$

С помощью подходящего выбора величины a в соотношении (17) получено приближенное описание каналов СМР, соответствующее виртуальным каналам спектрального преобразования в слуховом аппарате человека. Как отмечено выше, характеристика этих каналов достаточно надежно установлена, отражена в перцептуальной модели [5, 6] и выражена частотной шкалой РС в барках. Численное моделирование и анализ системы СМР (рис. 3) показывает, что при значении $a = -0.49$ положение и ширина каналов в этой модифицированной СМР на основе полифазной структуры практически идеально соответствует характеристикам анализа человеком звукового спектра РС в диапазоне до 11 кГц.

Отметим, что простая СМР на основе ВП соответствует логарифмическому масштабу преобразования частоты и поэтому дает значительные отклонения (особенно в области низких частот РС) от перцептуальной модели.

Восстановление РС в системе СМР

Построение рассмотренных систем СМР для речевого сигнала ориентировано на прагматический критерий — отсутствие заметных на слух искажений на выходе СМР сравнительно с сигналом на входе при условии неиспользования в СМР промежуточного преобразования спектра (в СМР без спектральных весов). Требование полного восстановления РС значительно смягчается в задаче улучшения качества сигнала методом спектрального вычитания фона. Это связано с присущими методу небольшими искажениями сигнала и с появлением незначительного остаточного шума. Обе

эти компоненты достаточно хорошо маскируют незначительное искажение РС за счет несовершенного восстановления его в системе СМР. Кроме того, эти небольшие искажения появляются в области высоких частот, а основная часть информационной части РС сосредоточена в пределах 3–4 кГц. Тесты по прослушиванию РС на выходе СМР подтверждают эти положения.

В связи с этим оказывается возможным строить СМР с прореживанием частоты без использования предельного значения коэффициента снижения эффективной частоты дискретизации (ЭЧД [2]) и тем самым достигать дополнительного снижения элайзинга между каналами. Соотношение коэффициента r снижения ЭЧД при прореживании и величины M повышения частотного разрешения является оптимальным для $r = M/4$ — это соотношение обеспечивает высокое качество РС на выходе системы СМР.

ЗАКЛЮЧЕНИЕ

Рассмотрены методы улучшения качества речевого сигнала (РС) за счет спектрального вычитания сопутствующего шума. В связи с этой задачей изучены системы согласованного многоканального разделения (СМР), включающие наборы фильтров анализа (НФА) и синтеза (НФС).

§ Проанализированы модифицированные структуры систем СМР, построенные по схеме повышенного частотного разрешения. Метод повышенного частотного разрешения описан для дискретного вейвлет-преобразования.

§ Приведена структура СМР на основе вейвлет-преобразования, имеющая M -кратно повышенное разрешение по частоте и неравномерную ширину каналов.

§ Приведена система СМР с модифицированной полифазной структурой с заменой блоков задержки фильтрами с широкополосным преобразованием РС, с неравномерной шириной каналов и повышенным частотным разрешением.

§ Для этой системы СМР показан нелинейный характер изменения частоты при использовании широкополосного преобразователя 1-го порядка и зависимость его от величины параметра a .

§ Показан способ эмпирического описания положения и ширины каналов слухового аппарата человека по "перцептуальной модели", основанной на экспериментальных данных и реализованной масштабированием частотной шкалы РС в барках.

§ Приведено соотношение параметра a и коэффициента M — кратности повышения частотного разрешения (в модифицированной структуре СМР), которое оптимальным образом соответствует виртуальным каналам частотного анализа РС в слуховом аппарате человека. Для этой структуры

СМР речевого сигнала дана величина параметра α , при которой достигается лучшее преобразование РС по критерию качества восприятия звука.

СПИСОК ЛИТЕРАТУРЫ

1. *Vaidyanathan P.P.* Orthonormal and Biorthogonal Filter Banks as Convolutional, and Deconvolutional Coding Gain // *IEEE Transactions on Signal Processing*. 1993. V. 41, N 6. P. 2110–2130.
2. Меркушева А.В., Малыхина Г.Ф. Согласованное многоканальное разделение сигнала: фильтрация и мультиплексирование // *Здесь*. С. 98–109.
3. *Smith M.J.T., Eddins S.I.* Analysis/Synthesis Techniques for Sub-Band Image Coding // *IEEE Transactions on Acoustic, Speech and Signal Processing*. 1990. V. 38. P. 1446–1456.
4. *Arrowood J.L., Smith M.J.T.* Exact Reconstruction Analysis/Synthesis Filter Banks with Time-Varying Filters // *Proceedings of IEEE Intern. Conference on Acoustic, Speech and Signal Processing (Minneapolis, USA), April 1993*. V. III. P. 233–236.
5. *Zwicker E.* Psychoacoustik. Berlin: Springer-Verlag, 1982.
6. Меркушева А.В. Нейросетевой метод цифровой обработки нестационарного сигнала (речи) в области вейвлет-отображения // *Цифровая обработка сигналов*. 2006. № 1. С. 18–30.
7. *Lang M., Guo H., Odegard J.E., Burrus C.S., Wells R.O.* Noise Reduction for Speech Signal Using an Undecimated Discrete Wavelet-Transformation // *IEEE Signal Processing Letters*. 1996. V. 3, N 1. P. 10–12.
8. *Seok J.W., Bae K.S.* Speech Enhancement with Reduction of Noise Components in Domain // *International Congress on Acoustic, Speech and Signal Processing, April 1997, Munich, Germany*. V. II. P. 12123–1226.
9. Меркушева А.В. Применение нейронной сети для текущего анализа нестационарного сигнала (речи), представленного его вейвлет-отображением. Основные принципы // *Научное приборостроение*. 2003. Т. 13, № 1. С. 64–71.
10. *Vary P.* Noise Suppression by Spectral Magnitude Estimation — Mechanism and Theoretical Limits // *Signal Processing*. July 1985. V. 8, N 4. P. 387–400.
11. *Dobinger G.* Computationally Efficient Speech Enhancement by Spectral Minima Tracking in Subbands // *Proceedings of EUROSPEECH*. 1995. P. 1513–1516.
12. *Martin R.* Spectral Subtraction Based on Minimum Statistics // *Proceedings of EUROSPEECH*. Edinburg, 1994. P. 1182–1185.
13. *Boll S.F.* Suppression of Acoustic Noise Using Spectral Subtraction // *IEEE Transactions on Acoustic, Speech and Signal Processing*. 1979. V. 27, N 4. P. 113–120.
14. *Kappelan M., Straus B., Vary P.* Flexible Non-uniform Filter Bank Using Allpass Transformatation of Multiple Order // *Proceedings of EUROSPEECH*. 1996. P. 1745–1748.
15. *Malychina G.F., Mercusheva A.V.* Metrological Aspects of Non-Stationary Signal Transformation for Spectrum Dynamic Analysis // *10th IMEKO TC7 International Symposium on Advances of Measurement Science (June 30–July 2, 2004), Saint-Petersburg, Russia*. V. 1. P. 212–216.
16. Малыхина Г.Ф., А.В. Меркушева Метрологические характеристики время-масштабного преобразования для анализа динамического спектра нестационарного сигнала в ИИС // *Измерительные информационные технологии*. Сб. статей факультета технической кибернетики Санкт-Петербургского политехнического университета. Изд. СПбГПУ, 2005. С. 50–62.
17. *Daubechies I.* Ten Lectures on Wavelets // *Conference Board Sciences, Series on Applied Mathematics N 61*. Society of Industrial and Applied Mathematics (SIAM), Philadelphia, 1992. 258 p.
18. *Jawerth B., Sweldens W.* Wavelet-Based Multiresolution Analysis // *SIAM Review*. 1994. V. 36, N 3. P. 337–345.
19. *Shensa M.J.* The Discrete Wavelet Transform: Wedding the a Trous and Mallat algorithms // *IEEE Transactions on Signal Processing*, 1992. V. 40, N 10. P. 2464–2482.

Санкт-Петербург

Материал поступил в редакцию 20.11.2007.

CONSISTENT MULTICHANNEL SIGNAL SEPARATION WITH WAVELET-TRANSFORMATION AND MODIFIED POLYPHASE STRUCTURE FOR SPECTRAL SUBTRACTION OF SPEECH SIGNAL

G. F. Malychina, A. V. Merkusheva

Saint-Petersburg

The spectral subtraction method is used to improve the speech signal (SS) quality since this method can essentially reduce or even eliminate SS additive noises. On the contrary to the earlier used transformation with a uniform channel width, two variants of structure are considered, in which the total signal frequency band is subdivided into bands with increasing widths. Such a method of the SS total band subdivision is more consistent with the virtual filtering system of the human auditory system, namely, the so-called perceptual model that manifests peculiar features of perception of SS and other acoustic signals. The paper considers two structures: one based on the discrete wavelet-transformation, the other being a nonuniform system of consistent multichannel separation based on the polyphase structure.