

Неравнополосный банк фильтров с фазовым преобразованием и объединением субполос для обработки речевых сигналов



Парфенюк М.,

доктор, доцент

Вашкевич М.И.,

аспирант

Петровский А.А.,

доктор технических наук, профессор



В работе рассматриваются вопросы проектирования неравнополосного банка фильтров для обработки речевых сигналов. Особенностью построения данного банка фильтров является использование фазового преобразования, а также объединение субполос, что позволяет повысить точность аппроксимации критических частотных полос слуховой системы человека. Показана полифазная структура рассматриваемого банка фильтров, позволяющая заметно снизить его вычислительную сложность. Также уделено внимание таким аспектам проектирования банка фильтров, как расчёт фильтра-прототипа, выбор коэффициентов децимации субполосных сигналов и выравнивание групповой задержки в банке фильтров. В качестве основной области применения неравнополосного банка фильтров рассматриваются слуховые аппараты.

In this paper the design of nonuniform cosine-modulated filter bank for speech signal processing is considered. Essential features of proposed filter bank are employment of allpass transform and subband merging, which allows improve accuracy of critical bands approximation. The polyphase structure with reduced computational complexity of filter bank under consideration is outlined. Also special attention is given to design of filter-prototype, subsampling ration selection and equalization of group delay in filter bank analysis/synthesis system. Hearing aids are considered as a main field of application of proposed nonuniform filter bank.

Ключевые слова: банк фильтров, фазовое преобразование, субполосная обработка речевых сигналов.



Введение

В большом количестве приложений обработки речевых сигналов необходимо выполнять декомпозицию сигнала на частотные составляющие. В качестве примера такого приложения рассмотрим слуховой аппарат (СА), схема которого изображена на **рис. 1**. Акустический сигнал поступает в устройство с нескольких микрофонов, из которых формируется моносигнал. В дальнейшем в СА ведётся субполосная обработка полученного моносигнала. Важнейшие шаги субполосной обработки — редактирование шума и усиление отдельных субполос для корректировки частотной характеристики слуха пациента. Технически сложна задача компенсации эхо-сигнала. Почти все упомянутые компоненты СА могут быть различным образом настроены в зависимости от ситуации. Для подобной настройки необходима система классификации, в которой анализируется акустический сигнал. Анализ заключается в извлечении признаков, по которым классифицируется, в какой акустической обстановке находится пользователь. Это влечёт за собой изменение параметров, которые регулируют блок редактирования шумов и модуль усиления субполосных сигналов.

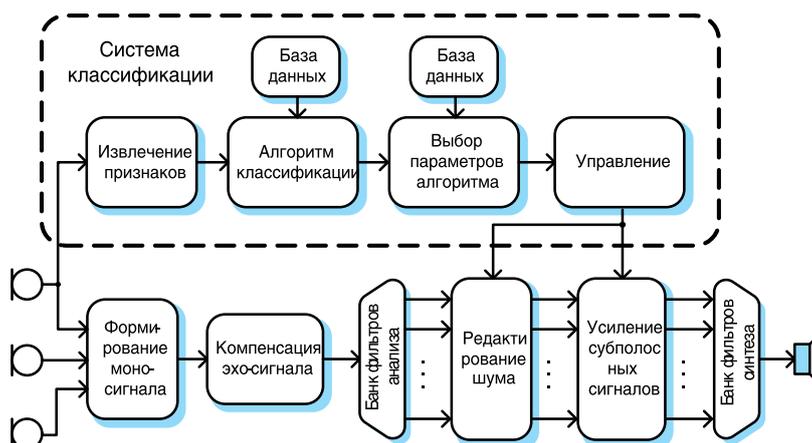


Рис. 1. Стадии обработки сигнала в слуховом аппарате

Качество работы слухового аппарата во многом определяется способом декомпозиции сигнала на субполосные составляющие. Поэтому в последнее время в литературе появляется множество публикаций, которые посвящены методам проектирования банков фильтров для слуховых аппаратов [3–5]. В данной работе будет рассмотрен банк фильтров для слуховых аппаратов, использующий фазовое преобразование и объединение субполос [6].

В отличие от многих авторов [7–8], которые фокусируются на исследовании неравнополосных ДПФ-модулированных банков фильтров, мы будем рассматривать банки фильтров основанные на косинусной модуляции [9]. Их преимущество в том, что для действительных входных сигналов канальные сигналы также действительны, что значительно облегчает их последующую обработку и техническую реализацию банка фильтров.

1. Косинусно-модулированный банк фильтров с фазовым преобразованием

1.1. Механизм неравнополосного частотного разбиения

Неравнополосный косинусно-модулированный банк фильтров на основе фазового преобразования образуется из своего равнополосного аналога [9]. При этом для последнего импульсные характеристики фильтров анализа и синтеза имеют следующий вид:

$$\begin{aligned} h_k(n) &= 2p(n)\cos\left[(2k+1)\frac{\pi}{2M}\left(n - \frac{N-1}{2}\right) + (-1)^k \frac{\pi}{4}\right], \\ f_k(n) &= 2p(n)\cos\left[(2k+1)\frac{\pi}{2M}\left(n - \frac{N-1}{2}\right) - (-1)^k \frac{\pi}{4}\right], \end{aligned} \quad (1.1)$$

$$0 \leq k < M,$$

где $p(n)$ — это коэффициенты фильтра-прототипа порядка N с частотой среза $\omega_c = \frac{\pi}{2M}$, а M — количество каналов банка фильтров. На порядок фильтра-прототипа накладывается ограничение $N = 2mM$, где m — произвольное положительное число. Далее для получения неравнополосного банка фильтров применяется фазовое преобразование (англ. *allpass transform*), которое заключается в замене в системе всех элементов задержки z^{-1} на фазовые звенья $A(z)$. Получаемые при этом передаточные функции фильтров анализа и синтеза M -канального косинусно-модулированного банка фильтров, основанного на деформации оси частот, описываются следующими выражениями [9–10]:

$$H_k(z) = a_k b_k U_k(A^{-1}(z)) + \bar{a}_k \bar{b}_k V_k(A^{-1}(z)), \quad (1.2)$$

$$F_k(z) = \bar{a}_k b_k U_k(A^{-1}(z)) + a_k \bar{b}_k V_k(A^{-1}(z)),$$

где верхняя черта означает комплексное сопряжение,

$$a_k = e^{j(-1)^k \frac{\pi}{4}}, \quad (1.3)$$

$$b_k = W_{2M}^{-(k+\frac{1}{2})(\frac{N-1}{2})}, \quad (1.4)$$

$$U_k(z) = P(zW_{2M}^{\frac{k+1}{2}}), \quad (1.5)$$

$$V_k = P\left(zW_{2M}^{-(k+\frac{1}{2})}\right), \quad (1.6)$$

где W_{2M} — корень степени $2M$ из единицы, а $P(z)$ — z -преобразование фильтра-прототипа. Для расчёта фильтра-прототипа могут использоваться методы, разработанные для равнополосных банков фильтров, см. [11–12] и литературу к этим статьям.

Простая замена $z^{-1} \rightarrow A(z)$ имеет примечательный эффект: частотная характеристика равнополосного банка фильтров деформируется в соответствии с ФЧХ $\varphi(\omega)$ фазового



звена $A(z)$, как показано на рис. 2. Простейшая деформация оси частот основывается на применении передаточной функции фазового звена первого порядка

$$A(z) = \frac{\alpha + z^{-1}}{1 + \alpha z^{-1}}, \quad (1.7)$$

фазовая характеристика которого зависит только от действительного коэффициента $|\alpha| \leq 1$:

$$\varphi(\omega) = -\omega + 2 \arctg \left(\frac{\alpha \sin \omega}{\alpha \cos \omega - 1} \right). \quad (1.8)$$

Чем меньше абсолютное значение коэффициента, тем слабее деформация оси частот. Для $\alpha = 0$ фактически не происходит деформации оси частот, поскольку в этом случае $A(z) = z^{-1}$. За исключением данного вырожденного случая, для реализации такого фильтра требуется одно умножение и два элемента задержки либо два умножения и один элемент задержки.

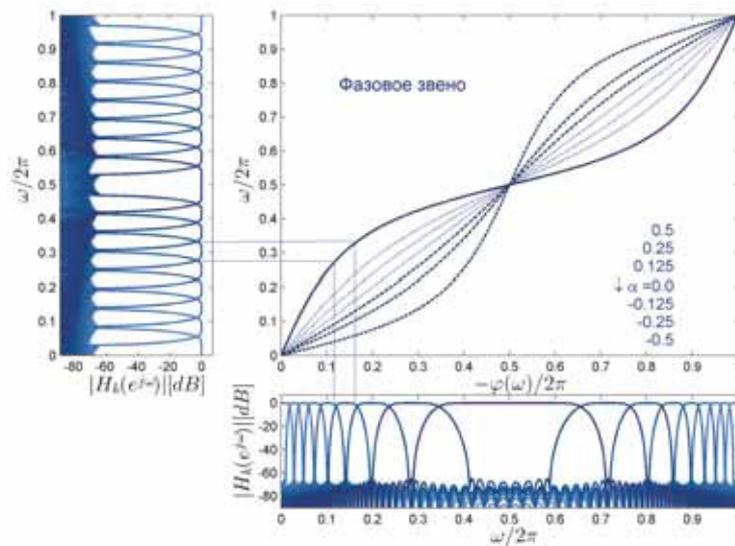


Рис. 2. Фазовое преобразование косинусно-модулированного банка фильтров

1.2. Полифазная структура банка фильтров

В работе [13] показано, что косинусно-модулированный банк фильтров можно эффективно реализовать, используя полифазное представление фильтра-прототипа:

$$P(z) = \sum_{n=0}^{N-1} p(n)z^{-n} = \sum_{q=0}^{2M-1} \sum_{r=0}^{m-1} p(q+2rM)z^{-(q+2rM)} = \sum_{q=0}^{2M-1} z^{-q} G(z^{2M}), \quad (1.9)$$

где $G(z)$ это полифазные компоненты I типа.

Введём обозначение

$$c_{k,l} = 2 \cos \left(\frac{(2k+1)\pi}{2M} \left(l - \frac{N-1}{2} \right) + (-1)^k \frac{\pi}{4} \right), \quad (1.10)$$

из которого, используя периодичность функции косинус, получаем тождество:

$$c_{k(l+2rM)} = (-1)^r c_{k,l}. \quad (1.11)$$

Объединяя полифазное представление фильтра-прототипа (1.9) с выражениями (1.10) — (1.11), получаем формулу для эффективной реализации косинусно-модулированного банка фильтров анализа:

$$H_k(z) = \sum_{n=0}^{N-1} h_k(n) z^{-n} = \sum_{n=0}^{N-1} p(n) c_{k,n} = \sum_{q=0}^{2M-1} \sum_{r=0}^{m-1} p(q+2rM) c_{k,(q+2rM)} z^{-(q+2rM)}. \quad (1.12)$$

После чего, применяя упрощение (1.10) к (1.12), получим:

$$H_k(z) = \sum_{n=0}^{N-1} z^{-q} c_{k,q} \sum_{r=0}^m (-1)^r p(q+2rM) z^{-2rM} = \sum_{q=0}^{2M-1} c_{k,q} G_q(-z^{2M}). \quad (1.13)$$

Данное выражение можно переписать в матричной форме:

$$\begin{bmatrix} H_0(z) \\ H_1(z) \\ \vdots \\ H_{M-1}(z) \end{bmatrix} = C_A \times \begin{bmatrix} G_0(-z^{2M}) \\ z^{-1} G_1(-z^{2M}) \\ \vdots \\ z^{-(2M-1)} G_{2M-1}(-z^{2M}) \end{bmatrix}, \quad (1.14)$$

здесь $C_A = [c_{k,l}]_{0 \leq k < M, 0 \leq l < 2M}$ — это матрица косинусной модуляции, которая, как было показано в [13], сводится к вычислению M -точечного дискретного косинусного преобразования типа IV (ДКП-IV).

Аналогично, представляя фильтр-прототип в виде суммы полифазных компонент

II типа $H(z) = \sum_{n=0}^{2M-1} z^{-(2M-1-q)} R_q(z^{2M})$, выводятся формулы для эффективной реализации банка фильтров синтеза:

$$\begin{bmatrix} F_0(z) \\ F_1(z) \\ \vdots \\ F_{M-1}(z) \end{bmatrix}^T = \begin{bmatrix} z^{-(2M-1)} R_0(-z^{2M}) \\ z^{-(2M-2)} R_1(-z^{2M}) \\ \vdots \\ R_{2M-1}(-z^{2M}) \end{bmatrix}^T \times C_S, \quad (1.15)$$

где $C_S = C_A^T$.

На основании выражений (1.14) и (1.15) строится полифазная структура банка анализа и синтеза соответственно. Далее, чтобы получить неравнополосный косинусно-модулированный банк фильтров, необходимо в полифазной структуре заменить все элементы задержки на фазовые звенья, т.е. $z^{-1} \rightarrow A(z)$. Получаемая после этого структура косинусно-модулированного банка

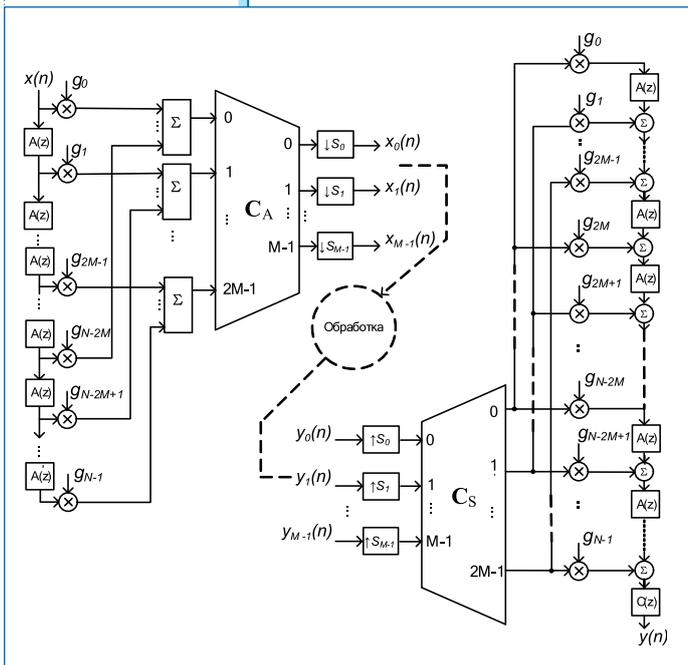


Рис. 3. Структура косинусно-модулированного банка фильтров с деформированной осью частот

фильтров с деформированной осью частот показана на рис. 3. На этом рисунке через $g_i, i = 0 \dots 1 - N$ обозначены коэффициенты полифазных компонент фильтра-прототипа, которые находятся, исходя из формулы (1.13).

Отличие данной полифазной структуры от той, что применяется для равнополосных банков фильтров, в том, что все элементы задержки заменены фазовыми звеньями, частота дискретизации канальных сигналов изменяется за пределами банка фильтров, а после реконструкции сигнала следует фильтр $C(z)$, выравнивающий групповую задержку. Эти особенности усложняют техническую реализацию банка фильтров и являются негативными последствиями, которые сопровождают полезный эффект деформации оси частот.

1.3. Выбор коэффициентов децимации

В системах на основе деформации оси частот предельная децимация без внесения искажений невозможна. Тем не менее полная реконструкция хорошо аппроксимируется путём выбора избирательного фильтра-прототипа для ослабления компонент наложения спектров. Кроме того, для предотвращения появления алайзинговых эффектов требуется наложить ограничения на коэффициенты децимации канальных сигналов в соответствии с их шириной полосы и расположением на частотной оси.

В работе [14] были выведены следующие неравенства:

$$\left[\frac{n_k}{2f_{Uk}} f_s \right] \geq S \geq \left[\frac{n_k - 1}{2f_{Lk}} f_s \right], \quad 1 \leq n_k \leq \left[\frac{f_{Uk}}{f_{Uk} - f_{Lk}} \right], \quad (1.16)$$

в которых полагается, что сигнал k-го канала является действительным, занимает частотный диапазон $[-f_{Uk}, -f_{Lk}] \cup [f_{Uk}, f_{Lk}]$ и имеет частоту дискретизации f_s . Выражения определяют диапазон значений коэффициента децимации S_k , для которых не возникает значительного, необратимого наложения спектров. Разумеется, наибольший интерес представляет максимальное допустимое значение S_k . Нижняя и верхняя частоты f_{Lk} и f_{Uk} соответственно, должны включать (хотя бы частично) переходную полосу фильтров анализа, образуемую после деформации частотной оси.

1.4. Выравнивание групповой задержки

Наиболее важная проблема использования банков фильтров в слуховых аппаратах, работающих в реальном времени, — то, что субполосная система

анализ/синтез вносит задержку. Для того, чтобы пользователя не беспокоил эхо-сигнал его собственной речи или не было несоответствия между движением губ говорящего и словами, которые слышит пользователь, временной сдвиг между входным сигналом и обработанным должен удерживаться в определённых рамках.

Упомянутые временные границы зависят от конкретного человека, и поэтому затруднительно дать их определённое значение. Имеющиеся работы различаются в этом отношении [3, 15]. Общее заключение представляется таким, что задержка свыше 10 мс объективно вызывает эхо, в то время как задержка выше 20–30 мс делает систему непригодной для использования в двухстороннем общении, потому что теряется согласованность восприятия с движением губ.

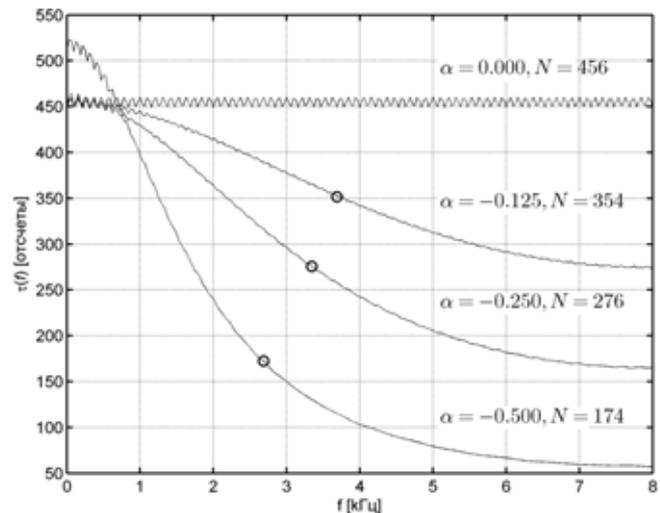


Рис. 4. График групповой задержки для различных банков фильтров

Растяжение частотной оси приводит к тому, что групповая задержка становится нелинейной функцией, поскольку банк фильтров на основе фазового преобразования всегда является БИХ системой. Групповая задержка зависит от силы деформации частотной оси и является монотонной функцией, с направлением, определяемым знаком коэффициента фазового преобразования. Для $\alpha < 0$ нижние частоты задерживаются на число отсчётов, большее порядка фильтра, в то время как высокие частоты задерживаются меньше. Для $\alpha > 0$, ситуация изменяется в противоположную сторону. Заметим также, что отклик системы анализ/синтез с деформированной осью частот эквивалентен отклику цепочки фазовых звеньев из N секций [16].

На рис. 4 показаны групповые задержки банков фильтров с различными фильтрами прототипами (порядка N), также с различающимися коэффициентами фазового преобразования (α). Кружками на графиках обозначены точки, в которых групповая задержка такая же, как у соответствующих равнополосных банков фильтров (с таким же фильтром-прототипом). Легко заметить, что отличие между этими значениями существенное, поэтому остро стоит вопрос выравнивания групповой задержки в системах на основе деформации оси частот.

Как известно, групповая задержка линейной системы $\tau(\omega)$ равна производной от ФЧХ системы с обратным знаком:

$$\tau(\omega) = -\frac{d\varphi(\omega)}{\omega}, \quad (1.17)$$

из этого следует, что изменение групповой задержки в зависимости от частоты (рис. 4) — это следствие нелинейности ФЧХ системы анализ/синтез неравнополосного косинусно-модулированного банка фильтров. Нелинейная ФЧХ системы недопустима, поскольку банк фильтров предполагается использовать для обработки звуковых сигналов. Ниже будет рассмотрен вопрос построения фильтра для коррекции ФЧХ системы анализ/синтез.

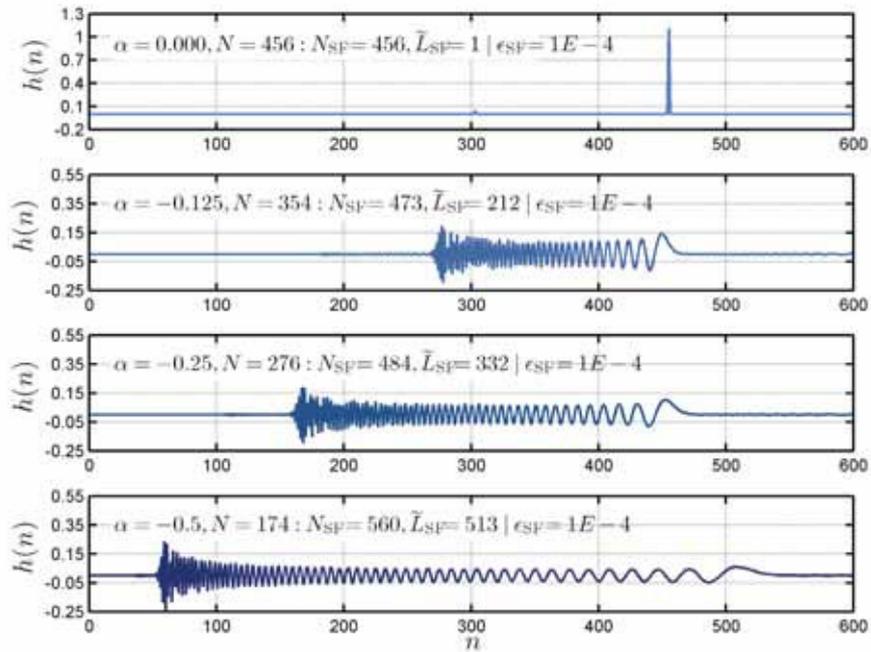


Рис. 5. Импульсные характеристики систем анализ/синтез с различными параметрами N и α

В [6] показано, что КИХ-фильтр $C(z)$ для выравнивания групповой задержки можно получить путём реверсии импульсной характеристики системы анализ/синтез банка фильтров, для которой требуется корректировка. Даже в случае бесконечного отклика в нём есть определяющий значимый фрагмент, что проиллюстрировано на рис. 5. Следовательно, можно принять, что существуют только отсчёты с индексами между 0 и N_{SF} , которые определяют передаточную функцию полиномиального типа, почти эквивалентную исходной:

$$A^N(z) \approx \sum_{n=0}^{N_{SF}} a(n)z^{-n}. \quad (1.18)$$

Реверсируя порядок следования отсчётов, получаем КИХ-фильтр, который аппроксимирует ряд соединённых обратных передаточных функций и задержек:

$$C(z) \approx \sum_{n=0}^{N_{SF}} a(N_{SF} - n)z^{-n} \approx A^N(z^{-1})z^{-N_{SF}}. \quad (1.19)$$

Так как $A^N(z^{-1}) = A^{-N}(z)$, то

$$A^N(z)C(z) \approx A^N(z)A^{-N}(z)z^{-N_{SF}} = z^{-N_{SF}}. \quad (1.20)$$

Очевидно, что отклонение произведения, стоящего слева от чистой задержки, зависит от остатка, который остаётся после усечений отклика. Мы экспериментально установили, что достаточно определить последний значащий отсчёт, после которого абсолютные значения становятся меньше $\epsilon_{SF} = 10^{-4}$. Для широких диапазонов α и N колебания групповой задержки после корректировки с использованием этого порогового значения не превышает нескольких

отсчётов для низких частот и быстро уменьшается с увеличением частоты до дробных частей отсчёта.

1.5. Расчёт фильтра-прототипа

Разработано много различных вариантов расчёта фильтра-прототипа для косинусно-модулированного банка фильтров (см. [11, 12, 18] и литературу к этим статьям), большинство из которых включают различные оптимизационные процедуры с нелинейными ограничениями. В этом разделе мы рассмотрим наиболее простой метод расчёта фильтра-прототипа, описанный в работе [11]. Достоинство данного метода в том, что ограничиваясь поиском фильтра-прототипа среди фильтров, получаемых методом оконного взвешивания, задача проектирования сводится к оптимизации всего одного параметра.

На рис. 6 схематично изображена АЧХ фильтра-прототипа для косинусно-модулированного банка фильтров. Допустим, что фильтр-прототип $P(e^{j\omega})$ имеет линейную ФЧХ. Тогда условие почти полного восстановления в банке фильтров может быть выражено в терминах $P(e^{j\omega})$ следующим образом:

$$P(e^{j\omega}) \approx 0, \quad |\omega| > \frac{\pi}{2M}, \quad (1.21)$$

$$T(e^{j\omega}) \approx 1, \quad \text{где } T(e^{j\omega}) = \sum_{k=0}^{2M-1} \left| P \left(e^{j(\omega - \frac{k\pi}{M})} \right) \right|^2, \quad (1.22)$$

Точность первого приближения (1.21) определяет меру ошибки, возникающую в результате наложения спектров (алайзинга), в то время как точность второго приближения (1.22) задаёт меру ошибки искажения (от англ. «distortion error»).

Запишем фильтр прототип, получаемый методом оконного взвешивания

$$p(n) = h_i(n)w(n), \quad (1.23)$$

где

$$h(n) = \frac{\sin(\omega_c(n - 0,5N))}{\pi(n - 0,5N)}, \quad n = 0, 1, \dots, N - 1, \quad (1.24)$$

это импульсная характеристика ФНЧ с частотой среза ω_c , а $w(n)$ — оконная функция (в [11] рассматривался случай, когда $w(n)$ — окно Кайзера). Введём фильтр $G(e^{j\omega}) = |P(e^{j\omega})|^2$, тогда исходя из условия (1.22), можно сделать вывод, что $G(e^{j\omega})$ аппроксимирует фильтр Найквиста ($2M$). Это значит, что должно выполняться условие $g(2Mn) \approx \left(\frac{1}{2M}\right)\delta(n)$,

где $\delta(n)$ — это дельта-функция. Теперь введём простую целевую функцию

$$\phi = \max_{n, n \neq 0} |g(2Mn)|. \quad (1.25)$$

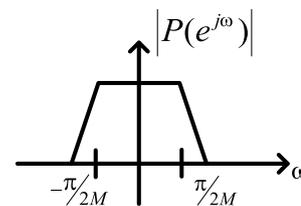


Рис. 6. Типичная АЧХ фильтра-прототипа



Единственный регулируемый параметр, который влияет на значение функции ϕ , это частота среза ω_c . Оптимизация заключается в нахождении значения ω_c , которое бы минимизировало функцию (1.25).

2. Объединение субполос

2.1. Обоснование

Использование фазовых звеньев даже первого порядка, для которых деформация оси частот зависит только от одного коэффициента α , позволяет получить частотное разбиение, соответствующее психоакустическим частотным шкалам. Эта проблема была изучена в [18] и выведено выражение в аналитической форме, которое даёт коэффициент фазового преобразования для заданной частоты дискретизации.

Рассматриваемый метод с объединением субполос точнее метода [19], основанного на использовании равнополосного банка фильтров, но фактическое его применение к системам с деформированной осью частот не позволяет получить частотное разделение, которое соответствует границам критических частотных полос. Другими словами, точное отображение из линейной оси частот в шкалу барков не обязательно «растягивает» банк фильтров так, что результирующие субполосы соответствуют критическим полосам, особенно, если число последних сравнимо с числом каналов банка фильтров. Хороших результатов можно достичь, если имеется большое число узких субполос, чьи выходы можно интерпретировать как отсчёты в частотной области, и затем просто группировать, чтобы оценить содержимое критической частотной полосы [20].

Критические полосы можно аппроксимировать, более точно используя деформацию оси частот на основе фазовых фильтров высших порядков, которые, однако, труднее рассчитывать, применять и реализовывать в аппаратуре [21].

Другая проблема, связанная с деформацией оси частот, это то, что в предыдущих статьях коэффициент фазового звена определялся очень точно, в виде десятичной дроби (например, $\alpha = 0,5756$ для частоты дискретизации 16 Гц), при этом предполагалась реализация фазового звена на основе умножителя. Однако ещё никто не рассматривал возможности упростить цепочку фазовых звеньев так, чтобы реализовать её без умножителей. В данной статье предлагается решить проблему аппроксимации критических полос, как по точности, так и по вычислительной сложности путём совмещения деформации оси частот с использованием фазовых фильтров первого порядка с объединением субполос. Идея заключается в проектировании банка фильтров с более либо менее деформированной осью частот, но с числом каналов большим, чем требуется субполос в частотном разбиении, после чего каналы объединяются для аппроксимации критических полос. До сих пор этот подход применялся без особого успеха лишь для равнополосных систем [22–24]. Для систем с деформированной осью частот он представляется более пригодным, как это будет показано в дальнейшем.

2.2. Точность частотного разбиения

В нашем подходе при аппроксимации критических частотных полос главная задача состоит в определении коэффициента фазового преобразования α и необходимого числа каналов так, чтобы число субполос соответствующего банка фильтров после объединения точно соответствовало заданному числу критических полос. Очевиден даже более общий вопрос — как разбить каналы на группы для объединения в субполосы в банке фильтров с деформацией оси частот.

Для решения этой задачи можно объединить субполосы так, чтобы нижняя частота первой из них и высшая частота последней были как можно ближе к краям данной критической полосы. При этом не возникает трудностей для отдельных субполос, попадающих внутрь аппроксимируемой критической частотной полосы, но возникают затруднения, когда одна субполоса принадлежит двум критическим полосам одновременно. Судя по всему, эта задача не имеет удовлетворительного решения, набор расчётных параметров для которых данные ситуации возникают, можно просто опускать.

Среди оставшихся пар коэффициентов фазового преобразования и числа каналов можно найти оптимальные комбинации, однако для этого необходимо иметь критерий точности. Мы рассмотрим такие метрики, как максимум и среднеквадратичная ошибка, основываясь на отклонении верхней частоты от заданной. Данное отклонение при этом нормализуют к ширине критической частотной полосы [6]:

$$\epsilon_{MAX} = \max_{k=0 \dots \tilde{M}-2} \frac{|f_{U_k} - \hat{f}_{U_k}|}{|\hat{f}_{U_k} - \hat{f}_{L_k}|}, \quad (2.1)$$

$$\epsilon_{RMS} = \sqrt{\frac{1}{\tilde{M}-2} \sum_{k=0}^{\tilde{M}-1} \left(\frac{|f_{U_k} - \hat{f}_{U_k}|}{|\hat{f}_{U_k} - \hat{f}_{L_k}|} \right)^2}. \quad (2.2)$$

В данных выражениях границы k -ой субполосы, полученной объединением, обозначены как f_{L_k} и f_{U_k} , соответственно, при этом \hat{f}_{L_k} и \hat{f}_{U_k} обозначают соответствующие границы критических частотных полос. Во всех выражениях ниже $\tilde{M} \leq M$ обозначает число каналов после объединения.

На основании (2.1) и (2.2), можно узнать, как они зависят от α и \tilde{M} . Перебор широкого диапазона значений параметров: $-0,9 \leq \alpha \leq 0$ и $20 \leq \tilde{M} \leq 200$, даёт поверхность ошибки. Эксперимент повторялся для четырёх наиболее часто используемых частот дискретизации: 8, 16, 32 и 44.1 кГц, которые соответствуют телефонному диапазону, широкополосной речи, FM и CD качеству соответственно. Полученные поверхности для ошибок максимума показаны на **рис. 7** с использованием графиков поверхностей.

Чистое объединение (при $\alpha = 0$) позволяет аппроксимировать любое частотное разбиение с наперёд заданной точностью. Однако при совместном использовании и объединении, и деформации оси частот необходимо меньшее число каналов для получения той же точности.

Для фиксированного числа каналов с увеличением коэффициента «растяжения» достигается улучшение точности, но лишь до определённой степени: слишком большая

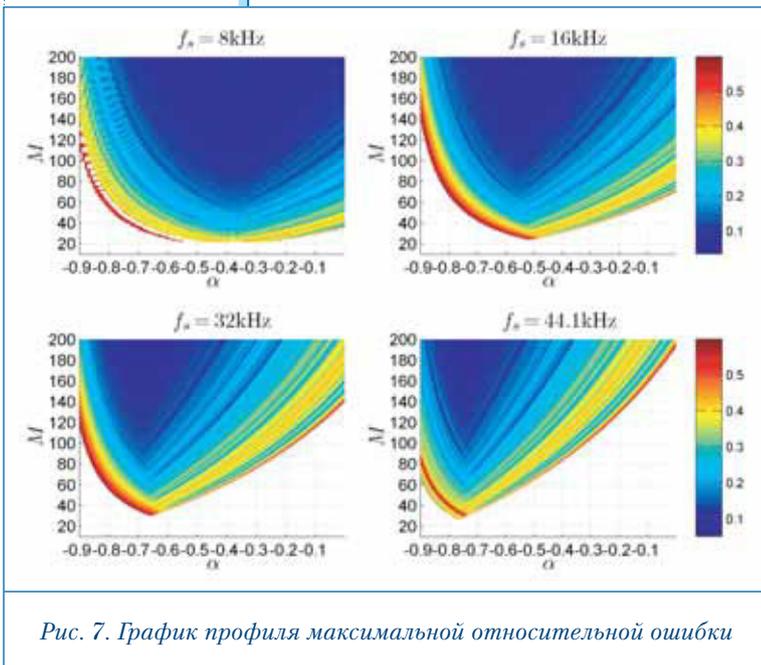


Рис. 7. График профиля максимальной относительной ошибки

деформация оси частот ухудшает аппроксимацию. Если число каналов невелико, то деформация оси частот необходима, чтобы сделать возможным аппроксимацию.

Для фиксированного коэффициента деформации оси частот α точность всегда может быть улучшена путём добавления каналов. Чем выше частота дискретизации, тем больше каналов необходимо для достижения той же точности, или чтобы сделать возможной аппроксимацию. С другой точки зрения, чем больше разница между самой узкой и самой широкой полосами, которая вызвана расширением частотного диапазона, тем больше деформация оси частот превосходит объединение.

Выше были охарактеризованы только общие тенденции. Ошибка аппроксимации колеблется, создавая много изолированных максимумов и минимумов, формирующих овраги. Из графиков становится понятно, что одна и та же точность может быть получена с использованием различных комбинаций коэффициента «растяжения» и числа каналов (α и M), которые соответствуют точкам, нерегулярно распределённым по плоскости параметров. Эти сочетания параметров можно использовать для получения оптимальных банков фильтров с точки зрения точности аппроксимации критических частотных полос и вычислительной сложности.

3. Неравнополосный косинусно-модулированный банк фильтров с частотной деформацией и объединением субполос на основе фазовых звеньев без умножителей

В этом разделе рассматриваются неравнополосные банки фильтров, для которых цепочка фазовых звеньев может быть реализована без использования аппаратного умножителя, т.е. при $\alpha = \left\{0, -\frac{1}{2}, -\frac{1}{4}, -\frac{1}{8}\right\}$. Таким образом,

применяя теорию из предыдущих разделов, были рассчитаны четыре банка фильтров для аппроксимации критических полос при частоте дискретизации входного сигнала 16 кГц. В табл. 1 приведены границы критических частотных полос и информация об объединении каналов.

Частотные характеристики банков фильтров до и после объединения каналов показаны на рис. 8–11, штриховыми и сплошными линиями соответственно. Дополнительно изображены графики частотной характеристики $|T(f)|$ и групповой задержки $\tau(f)$. Расчётные параметры, порядок фильтров прототипов и значения ошибок аппроксимации показаны на графиках.

Фильтры-прототипы для банков фильтров были рассчитаны по методу [11], а чтобы ограничить алгоритмическую задержку, порядок фильтра N выбирался равным. Приведённые банки фильтров характеризуются различными компромиссными решениями между степенью деформации оси частот и числом каналов исходного банка фильтров.

Таблица 1

Критические частотные полосы и объединения субполос, используемые для их аппроксимации

Критическая частотная полоса		Объединяемые субполосы/ Коэффициенты передискретизации S_k для наборов параметров			
№	Частотный диапазон, кГц	$\alpha = -0,5$ $M = 29$	$\alpha = -0,25$ $M = 46$	$\alpha = -0,125$ $M = 59$	$\alpha = 0$ $M = 76$
1	0–100	0 — 0/57	0 — 0/51	0 — 0/50	0 — 0/50
2	100–200	1 — 1/34	1 — 1/31	1 — 1/30	1 — 1/30
3	200–300	2 — 2/24	2 — 2/21	2 — 2/21	2 — 2/21
4	300–400	3 — 3/37	3 — 3/33	3 — 3/33	3 — 3/33
5	400–510	4 — 4/30	4 — 4/27	4 — 4/27	4 — 4/27
6	510–630	5 — 6/22	5 — 5/35	5 — 5/34	5 — 5/35
7	630–770	7 — 7/28	6 — 6/30	6 — 6/30	6 — 6/30
8	770–920	8 — 8/25	7 — 8/15	7 — 8/15	7 — 8/24
9	920–1080	9 — 10/20	9 — 9/28	9 — 9/28	9 — 9/36
10	1080–1270	11 — 11/30	10 — 11/17	10 — 11/17	10 — 11/24
11	1270–1480	12 — 12/27	12 — 13/20	12 — 13/20	12 — 13/20
12	1480–1720	13 — 14/13	14 — 15/17	14 — 15/17	14 — 15/23
13	1720–2000	15 — 15/15	16 — 17/15	16 — 17/15	16 — 18/15
14	2000–2320	16 — 17/9	18 — 19/13	18 — 20/13	19 — 21/13
15	2320–2700	18 — 18/11	20 — 22/11	21 — 23/11	22 — 25/11
16	2700–3150	19 — 20/7	23 — 24/15	24 — 27/7	26 — 29/12
17	3150–3700	21 — 21/8	25 — 28/6	28 — 31/8	30 — 34/8
18	3700–4400	22 — 23/5	29 — 31/7	32 — 36/7	35 — 41/7
19	4400–5300	24 — 24/4	32 — 35/4	37 — 42/4	42 — 49/6
20	5300–6400	25 — 26/2	36 — 39/2	43 — 49/2	50 — 60/2
21	6400–7700	27 — 27/4	40 — 44/4	50 — 56/4	61 — 72/4
22	7700–8000	28 — 28/6	45 — 45/18	57 — 58/18	73 — 75/21

Все банки фильтров, кроме последнего, характеризуются схожей точностью аппроксимации, групповой задержкой и уровнем искажений. Главное отличие между системами — то, как переходная полоса зависит от числа субполос. Чем меньше степень деформации, тем менее она возрастает с увеличением частоты. В двух системах с $\alpha = -\frac{1}{8}$ и $\alpha = 0$ как узкие, так и широкие субполосы характеризуются одинаковой шириной переходной полосы, что нежелательно во многих приложениях.

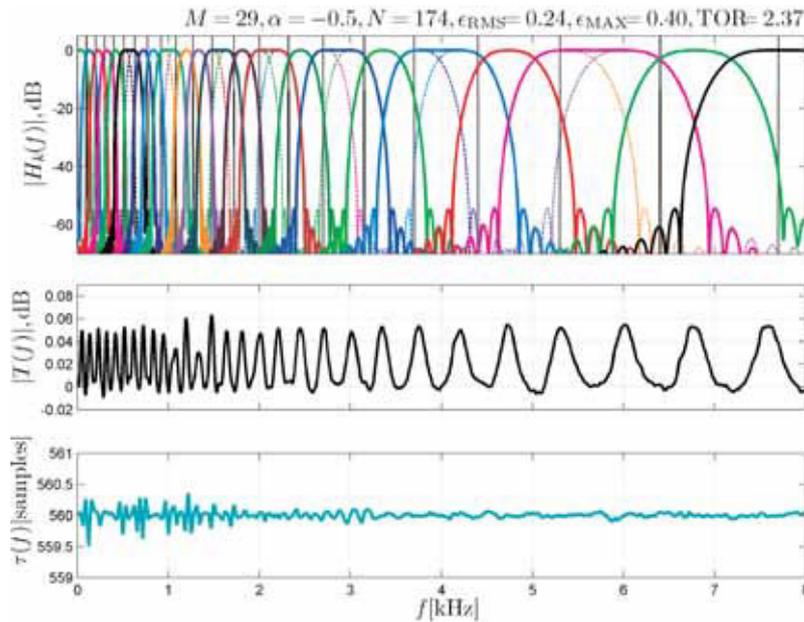


Рис. 8. Частотные характеристики банка фильтров, основанных на деформации оси частот и объединении субполос ($M = 29, N = 174$)

Деформация оси частот позволяет сделать перекрытие между субполосами пропорциональным их ширине полос, так, что длина фильтра-прототипа может быть уменьшена без влияния на уровень ослабления в полосе заграждения. Этот фактор также влияет на *общее отношение передискретизации*

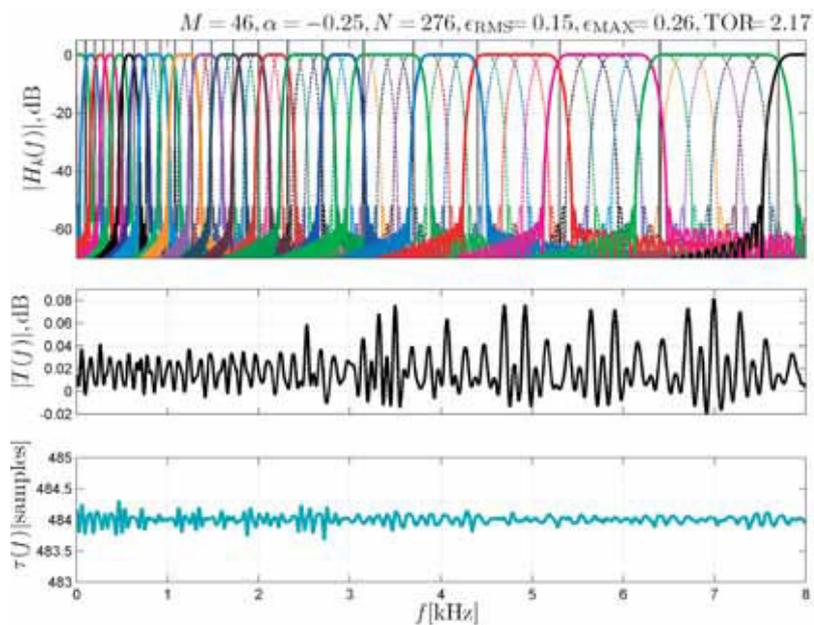


Рис. 9. Частотные характеристики банков фильтров, основанных на деформации оси частот и объединении субполос ($M = 46, N = 276$)

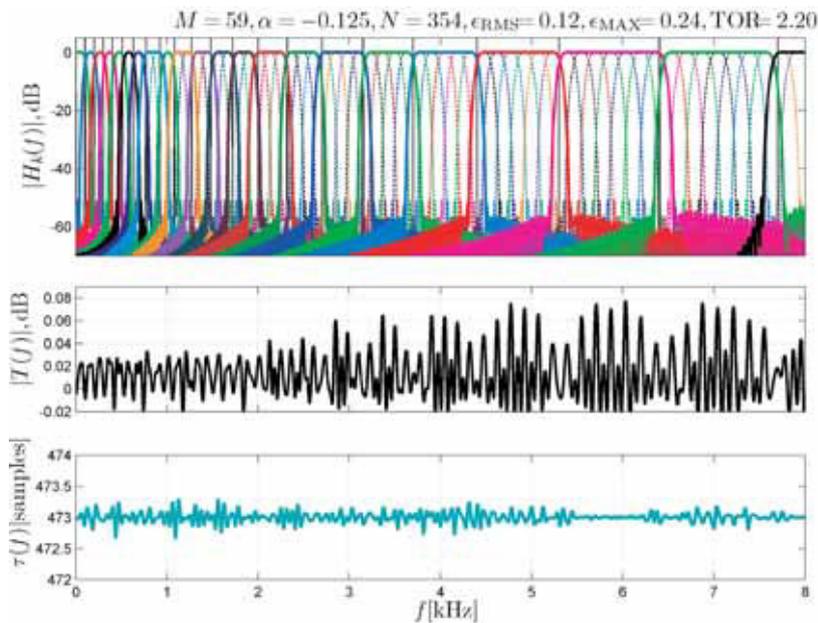


Рис. 10. Частотные характеристики банков фильтров, основанных на деформации оси частот и объединении субполос ($M = 59, N = 354$)

(ООП, от англ. *TOR* — *total oversampling ratio*), которое вычисляется как сумма величин обратных коэффициентам децимации:

$$TOR = \sum_{k=0}^{M-1} \frac{1}{S_k}. \quad (3.1)$$

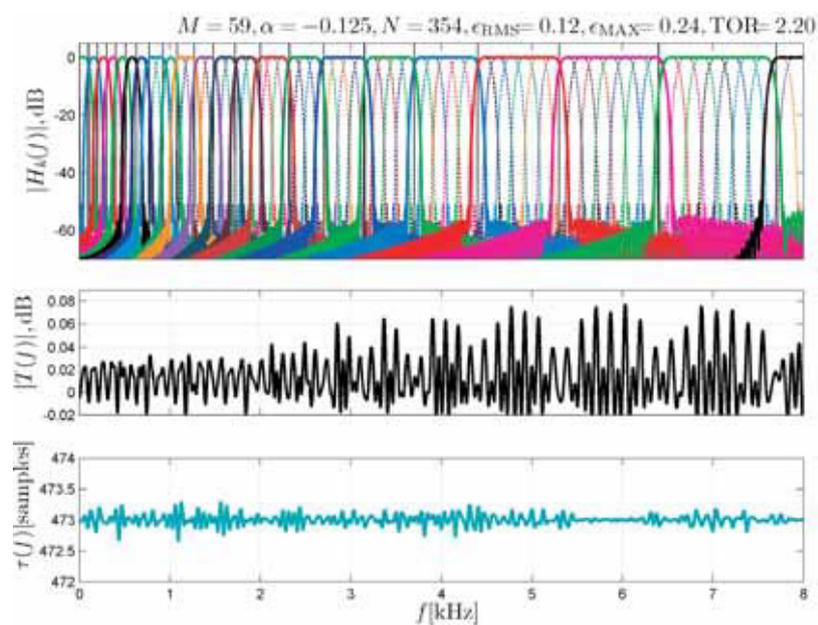


Рис. 11. Частотные характеристики банков фильтров, основанных на деформации оси частот и объединении субполос ($M = 76, N = 456$)



Расширение переходных полос увеличивает риск наложения спектров, так что переходные полосы фильтров должны быть добавлены к краям субполос, перед тем как применять выражение (1,16), что, к сожалению, приводит к более низким коэффициентам децимации.

Эти примеры расчёта банков фильтров показывают, что достигаются следующие характеристики банка при минимальном уровне вносимых искажений: ослабление в полосе задерживания 50 дБ, неравномерность частотной характеристики банка фильтров 0,1 дБ, задержка 29–35 мс. Эти результаты хороши для построения систем кодирования речевых и звуковых сигналов. Что касается слуховых аппаратов, то необходим компромисс между вычислительной сложностью банка фильтров (для минимизации потребляемой мощности рекомендуется частота задающего генератора 1 МГц [3]) и уровнем вносимых искажений.

Работа частично финансировалась по гранту Министерства образования Республики Беларусь ГБЦ 09-3105 и гранту W/WI/06/10 Белостокского политехнического института.

Литература

1. Эфрусси М.М. Слуховые аппараты и аудиометры / М.М. Эфрусси. М.: Энергия, 1975.
2. Marpatin, R. Advances in digital speech transmission / R. Martin, U. Heute, Ch. Antweiler. — Chichester. : John Wiley & Sons Ltd, 2008.
3. Bäuml R.W., W. Sörgel «Uniform polyphase filter banks for use in hearing aids: design and constraints», in the proc. 16th European signal processing conference (EUSIPCO-2008), Lausanne, Switzerland, 25–29 august 2008.
4. Kurbiel T., Göckler H.G., Alfsmann D. «Oversampling complex-modulated digital filter bank pairs suitable for extensive subband-signal amplification» in the proc. of 17th European signal processing conference (EUSIPCO-2009), Glasgow, Scotland, 24–28 august 2009. P. 2658–2662.
5. Alfsmann D., Gockler H.G., Kurbiel T. «Filter banks for hearing aids applying subband amplification: a comparison of different specification and design approaches», in proc. of 17th European signal processing conference (EUSIPCO-2009), Scotland, Glasgow, August 24–28, 2009. P. 2663–2667.
6. Parfieniuk M., Petrovsky A.A., Wan W. «Frequency warping and subband merging for approximating the critical bands with cosine modulated filter banks,» in the proc. Internationals conference on audio, language and image processing (ICALIP'08), China, Shanghai, July 78, 2008. vol. 2, P. 1159–1166.
7. Galijasevic E., Kliewer J. «Non-uniform near-perfect-reconstruction oversampled DFT filter banks based on all-pass transform» The Proc. of IEEE Digital Signal Processing Workshop'2000, Hunt, TX, USA, 2000 P. 1–6.
8. Löllmann H. W., Vary P. «Improved design of oversampled allpass transformed DFT filter banks with near-perfect reconstruction,» in the proc. of 15th European signal processing conference (EUSIPCO-2007), Poland, Pozan, September 3–7, 2007. P. 50–54.
9. Parfieniuk M., Petrovsky A.A. «Tunable non-uniform filter bank mixing cosine modulation with perceptual frequency warping by allpass transformation» Automatic Control and Computer Sciences. 2004. vol. 38, no. 4. P. 44–52.
10. Piotrowski A., Parfieniuk M. «Digital filter banks: analysis, synthesis and implementation for multimedia systems»/ Wydawnictwo Politechniki Bialostockiej, Bialystok, 2006. P. 389.
11. Lin Y.P., P.P. Vaidynathan «A Kaiser Window Approach for the Design of prototype Filters of Cosine Modulated Filterbanks» IEEE Signal Processing Letters. 1998. vol. 5, no. 6. P. 132–134.
12. Bergen, S.W., Antoniou A. «An efficient closed-form design method for cosine-modulated filter banks using window function,» Signal Processing. 2007. vol. 87, no. 5. P. 811–823.

13. Koipillai D., P.P. Vaidynathan «Cosine-Modulated FIR Filter Banks Satisfying Perfect Reconstruction,» IEEE Trans. on Signal Processing. 1992. vol. 40, no. 40. pp. 770–783.
14. Parfieniuk M., Petrovsky A.A. «Simple rule of selection of subsampling ratios for warped filter banks» Известия Белорусской Инженерной Академии. 2003. № 1. С. 130–134.
15. Kates M., Arehart K.H. «Multichannel dynamic range compression using digital frequency warping» EURASIP Journal on applied processing. 2005. vol.2005. P. 3003–3014.
16. Parfieniuk, M. Petrovsky A.A. «Reduced Complexity Synthesis Part of Non-Uniform Near-Perfect-Reconstruction DFT Filter Bank Based on All-Pass Transformation,» in the proc. of the European conf. on Circuit theory and Design (ECCTD'03) Cracow, Poland, 2003. vol. III. P. 5–8.
17. Furtado M.B., Diniz P.S., Netto S.L. «Numerically efficient optimal design of cosine-modulated filter banks with peak-constrained least-squares behavior» IEEE trans. on Circuits and Systems 2006. vol. 52, no. 3, P. 597–608.
18. Smith III, J.O Abel J.S. «Bark and ERB bilinear transform» IEEE Trans. On Speech and Audio Processing, 1999. vol. 7, No. 40. P. 697–708.
19. Karmakar A., Kumar A., Patney R.K. «Design of optimal wavelet packet trees based on auditory perception criterion» IEEE signal processing letters, 2007. vol. 14, no. 4, P. 240–243.
20. Parfieniuk M., Petrovsky A.A. «Warped DFT as the basis for psychoacoustic model» Proc. of the International Conference of Acoustic, Speech and Signal Processing (ICASSP'04) Moreal, Canada, 2004. vol. IV. P. 185–188.
21. Kappellan M., Strauss B., Vary P. «Flexible nonuniform filterbanks using allpass transformation of multiple order,» in the proc. of 8th European signal processing conference (EUSIPCO-1996), Trieste, Italy, 7–13 september 1996. P. 1745–1748.
22. Lee J.J., Lee B.G. «A design of nonuniform cosine-modulated filter banks» IEEE Trans. Circuits and Syst. II., 1995. vol. 42, no. 8. P. 2105–2109.
23. Li J., Nguyen T. Q., Tantaratana S. «A simple design method for near-perfect reconstruction nonuniform filter banks,» IEEE Trans. Signal Processing, 1997. vol. 45, no.8. P. 2105–2109.
24. Niamut O.A., Heusdens R. «Subband merging in cosine-modulated filter banks,» IEEE Signal processing letters, 2003. vol. 10, no.4. P. 111–114.

Парфенюк Марек —

доктор, доцент. Работает в Белостокском политехническом институте, г.Белосток, Польша. Основное образование: Белостокский политехнический институт, факультет информатики. Специальность – информатик. Область интересов: цифровая обработка сигналов (многоскоростная обработка, анализ/синтез банков фильтров), проектирование систем мультимедиа. Член общества IEEE.

Вашкевич Максим Иосифович —

аспирант, окончил Белорусский государственный университет информатики и радиоэлектроники, факультет компьютерного проектирования, специальность — «электронные вычислительные средства». Область научных интересов: многоскоростная обработка сигналов, проектирование банков фильтров, аппаратная реализация алгоритмов цифровой обработки сигналов на ПЛИС.

Петровский Александр Александрович —

доктор технических наук, профессор. Работает в Учреждении образования «Белорусский государственный университет информатики и радиоэлектроники», кафедра «Электронные вычислительные средства». Закончил Учреждение образования «Белорусский государственный университет информатики и радиоэлектроники» по специальности «Электронные вычислительные машины». Главные научные интересы лежат в области цифровой обработки сигналов речи и звука для целей компрессии, распознавания, редактирования шума, а также в области проектирования проблемно-ориентированных средств вычислительной техники реального времени для систем мультимедиа. Член НТО РЭС им. А.С.Попова, IEEE, EURASIP, AES.