

Построение психоакустической модели в области вейвлет-коэффициентов для перцептуальной обработки звуковых и речевых сигналов

А.А. Петровский,

кандидат технических наук, доцент

В статье предлагается методология построения дерева пакета дискретного вейвлет-преобразования для перцептуальной обработки сигналов речи и звука при минимальном множестве узлов, определённой глубине декомпозиции структуры дерева, заданной частоте дискретизации сигнала, когда ошибка аппроксимации критических частотных полос минимальна в области барков. Также приведена процедура расчёта маскирующих порогов в пространстве вейвлет-коэффициентов.

Abstract

The methodology of a wavelet packet (WP) tree construction for time-frequency processing of speech and audio signals is proposed to obtain a WP tree for the minimal number of nodes, certain WP tree decomposition depth, the required sampling frequency, the minimal approximation error of the critical frequency bands. The procedure of masking thresholds calculation in the wavelet coefficients domain is also described in the given article.

Введение

Пакет дискретного вейвлет-преобразования (ПДВП), или, другими словами, быстрое вейвлет-преобразование [1], является важным инструментом частотно-временной декомпозиции звуковых и речевых сигналов для различных приложений мультимедиа-систем [1, 2]. ПДВП часто используется в субполосной перцептуальной обработке сигналов звука и речи [2-4] в критических частотных полосах согласно психоакустической модели Zwicker [5]:



— расстояние между центральными частотами соседних критических частотных полос определяется формулой

$$z = F(f) = 13 \cdot \arctan(0.00076 \cdot f) + 3.6 \cdot \arctan\left(\left(\frac{f}{7500}\right)^2\right), [\text{Барк}] \quad (1)$$

где f — частота в герцах (единица измерения в данном масштабе — 1 барку);

— ширина критических частотных полос определяется формулой

$$CBW(f) = 25 + 75 \cdot \left(1 + 1.4 \cdot \left(\frac{f}{1000}\right)^2\right)^{0.69} \cdot [\text{Гц}] \quad (2)$$

При этом также вычисляются характеристики психоакустической модели восприятия человеком акустической информации, такие как пороги маскирования, а именно:

— абсолютный порог слышимости АНТ (absolute threshold of hearing), частотная зависимость которого аппроксимируется выражением:

$$ATH(f) = 3.64 \cdot \left(\frac{f}{1000}\right)^{-0.8} - 6.5 \cdot e^{-0.6\left(\frac{f}{1000}-33\right)^2} + 10^{-3} \cdot \left(\frac{f}{1000}\right)^4, [\text{Гц}] \quad (3)$$

где f — частота в герцах;

— частотное маскирование (simultaneous masking), проявляющееся при воздействии маскира в течение некоторого времени одновременно на разных частотах сигнала;

— маскировка во временной области (temporal masking): если громкий звук маскирует следующий за ним слабый звук, то явление называется «маскировкой вперёд» (post-masking) и может продолжаться от 5 мс до 300 мс в зависимости от силы и длительности маскира; при «маскировке назад» (pre-masking) громкий звук маскирует звук, воспроизводимый до него, длительность которого составляет примерно 20 мс.

Итак, пусть $\{\varphi_n(t) : z \in Z\}$ определяет множество структур деревьев ПДВП и пусть $\{E \subset \{(l, n) : 1 \leq l \leq L, 0 \leq n \leq 2^l\}\}$ представляет собой узлы дерева ПДВП, тогда отрезок $[0, 1]$ разделяется на диадические интервалы $I_{l,n} = [n2^l, (n+1)2^l]$, которые соответствуют специфическому множеству узлов E . В частности, $\{\varphi_{l,n,k}(t) : (l, n) \in E, k \in Z\}$, где $\varphi_{l,n,k}(t) \triangleq 2^{-\frac{l}{2}} \varphi_{l,n}(2^{-l}t - k)$ является базовой формой в пространстве сигнала $\overline{\text{span}}\{\varphi_{0,0}(t-k) : k \in Z\}$. Узел $(l, n) \in E$ дерева ПДВП ассоциируется с частотной полосой, у которой центральная частота и полоса пропускания приблизительно задаются следующими соотношениями:

$$f_{l,n} = 2^{-l} (GC^{-1}(n) + 0.5) \cdot \frac{f_s}{2}, \quad (4)$$

$$\Delta f_{l,n} = 2^{-l} \cdot \frac{f_s}{2}, \quad (5)$$

где $GC^{-1}(n)$ — обратный код перестановок Грея, f_s — частота дискретизации сигнала.

ПДВП реализуется на выбранной структуре дерева, поиск которой основывается на известном утверждении [6]: любая комбинация целых индексов $(l, n, k) \in Z$, для которых вейвлеты сконцентрированы на двоичных интервалах $[n2^{-l}, (n+1)2^{-l}]$ из диапазона $[0, \infty)$, соответствует ортогональным базисам $\psi_{l,n,k}(t)$, $\varphi_{l,n,k}(t)$ из пространства $L^2(k)$. Утверждение доказывает существование множества структур ПДВП, причём ПДВП ассоциируется с алгоритмом выбора лучшей структуры преобразования $\{E \subset \{(l, n): 1 \leq l \leq L, 0 \leq n \leq 2^l\}\}$ из множества структур путём изменения и минимизации определённой меры качества.

Цель данной статьи — показать методологию построения дерева ПДВП для минимального множества узлов (l, n) , определённой глубины декомпозиции l , заданной частоты дискретизации сигнала f_s , когда ошибка аппроксимации критических частотных полос минимальна в области барков, а также получить процедуру расчёта маскирующих порогов в пространстве вейвлет-коэффициентов.

1. ПДВП, согласованный со шкалой критических частотных полос

Для того чтобы получить аппроксимацию шкалы критических частотных полос с помощью ПДВП, необходимо осуществить декомпозицию дерева ПДВП таким образом, чтобы расстояние между центрами субполос составляло 1 барк. Следует отметить, что ширина критических частотных полос $CBW(f)$ — монотонно увеличивающаяся функция частоты (2). Для формирования низкочастотных полос требуется интенсивная декомпозиция ПДВП в сравнении с характером изменения дерева ПДВП для аппроксимации высокочастотных полос.

Дано дерево ПДВП $(l, n) \in E_m$ и его вейвлет-коэффициенты $X_{l,n,k}$. Интегральная перцептуально взвешенная ошибка аппроксимации шкалы критических частотных полос деревом $(l, n) \in E_m$ ПДВП в области барков может быть определена следующим образом:

$$Q_E = \frac{1}{L} \sum_{\substack{\text{для} \\ \forall (l, n) \in E_m}} [\widehat{CBW}_{Z_w}(z) - \widehat{CBW}_{E_m}(z_{(l, n)})]^2 \cdot \widehat{W}(z). \quad (6)$$

Здесь ширина критических частотных полос $\widehat{CBW}_{Z_w}(z)$ в Гц — функция центральных частот соседних критических частотных полос, заданных в барках, то есть:

$$\widehat{CBW}_{Z_w}(z) = CBW(F^{-1}(z)), [\text{Гц}] \quad (7)$$

$\widehat{CBW}_{Z_w}(z)$ определяет шкалу критических частотных полос в модели Zwicker [5];

$\widehat{CBW}_{E_m}(z_{(l, n)})$ — аппроксимация критических частотных полос деревом ПДВП $(l, n) \in E_m$;

$z_{(l, n)}$ — центр полосы (l, n) в барках дерева ПДВП E_m вычисляется для центральной частоты $f_{(l, n)}$, заданной в Гц, как $z_{(l, n) \in E_m} = F(f_{(l, n)})$, где F — преобразование (1). Перцеп-

туальная взвешивающая функция $\widehat{W}(z)$, учитывающая определённые частотные свойства наружного и среднего уха, задаёт меньшее распределение ошибки аппроксимации шкалы критических частотных полос в области средних частот по сравнению с низкочастотным и высокочастотным диапазонами и определяется в шкале дБ как функция частоты [7]:

$$W_{\text{дБ}}(f) = -0.6 \times 3.64(10^{-3} f)^{-0.8} + 6.5 \times \exp(-0.6 \times (10^{-3} f - 3.3)^2) - 10^{-3}(10^{-3} f)^4, \quad (8)$$



Также $\widehat{W}(z)$ может быть переопределена для барков как

$$\widehat{W}(z) = \widehat{W}(F(f)) = W(F^{-1}(z)) = W(f), \quad (9)$$

где $W(f) = 10^{(W_{ab}(f)/20)}$.

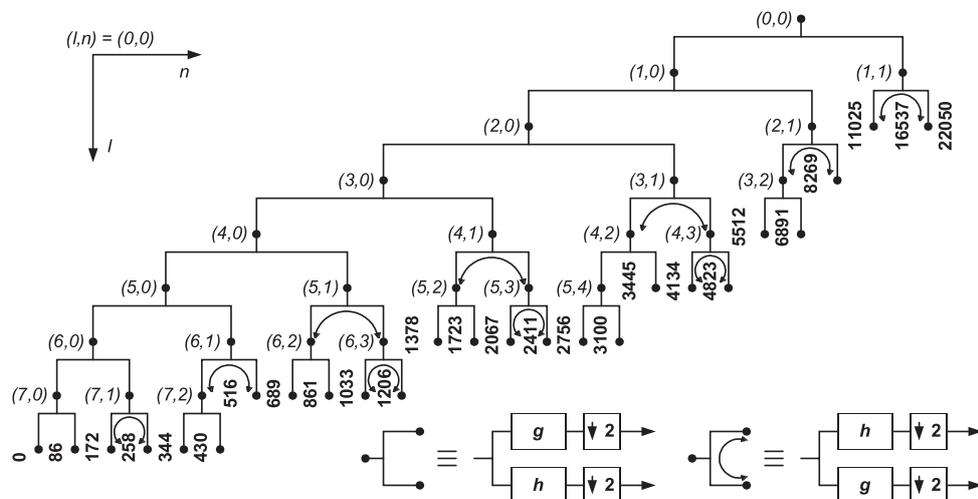
Минимизация ошибки Q_E (6) может позволить автоматизировать процесс построения оптимального дерева ПДВП $(l, n) \in E_{CB}$ для шкалы критических частотных полос.

На **рис. 1** показано дерево ПДВП (Critical Band Wavelet Packet Decomposition (CB – WPD)) [3, 8], полученное эмпирически, которое осуществляет разделение частотного интервала сигнала на полосы согласно критической шкале частот: $CB - WPD: (l, n) \in E_{CB}, l = \overline{0, 8}$, где E_{CB} обозначает множество узлов дерева ПДВП, соответствующего $CB - WPD$. Дерево $CB - WPD$ делит частотный диапазон [0–22,05 кГц] на 25 неравномерных полос $CBW(f)$, то есть на 25 барков. Корневой узел $(l, n) = (0, 0)$ данного дерева соответствует всему частотному диапазону сигнала. Каждый внутренний узел дерева $(l, n) \in E$, названный узлом предка, делится на два потомка: 1-й потомок и 2-й потомок, ассоциируемые соответственно с высокочастотной и низкочастотной фильтрацией, выходные сигналы (вейвлет-коэффициенты) которых децимируются в соотношении два к одному:

$$X_{l,n,k}(t) = \langle x(t), \varphi_{l,n,k}(t) \rangle, (l, n) \in E_{CB}, k \in Z. \quad (10)$$

Для систем перцептуальной обработки широкополосных речевых сигналов предлагается следующее критическое дерево ПДВП $CB - WPD$ [8]: $(l, n) \in E_{CB}, l = \overline{0, 7}$ (см. **рис. 2**), где ширина полосы анализируемого сигнала равна 16 кГц и обработка ведётся в 24 барках.

Разрешающая способность человеческого уха ограничивает длину анализируемого фрейма в пределах 5–10 мс для области верхних частот и 100 мс для нижних частот. Выбор фильтра прототипа преобразования, длины его вейвлет-



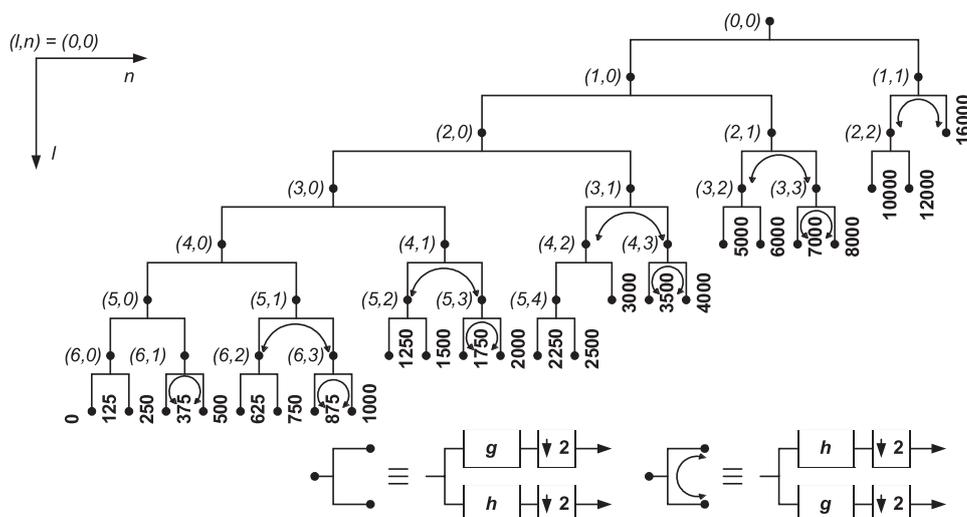


Рис. 2. Структура критического дерева ПДВП СВ – WPD: $(l,n) \in E_{CB}, l=0,7$

функции в общем случае зависит от максимального размера окна обработки (временного разрешения) и нулевых моментов [1]. Использование вейвлет-функций семейства Добеши позволяет обеспечить хорошую частотную избирательность, которая увеличивается с числом масштабных уровней в дереве преобразования.

На рис. 3 и рис. 4 показаны аппроксимации центральной частоты и ширины каждой частотной полосы критической шкалы частот (соответственно) деревом ПДВП, структура которого приведена на рис. 1. Аналогичные результаты для дерева ПДВП (см. рис. 2) иллюстрируются на рис. 5 и рис. 6 соответственно. Здесь непрерывная линия соответствует критической шкале частотных полос согласно модели Zwicker, а кружки — декомпозиции дерева ПДВП. Сопоставление данного результата с другими структурами дерева СВ – WPD [9, 10] даёт определённый выигрыш в величине ошибки Q_E (от 0 до 4 дБ), а неформальные тесты прослушивания показывают лучшее восприятие восстановленных сигналов в перцептуальных системах компрессии и редактирования шумов.

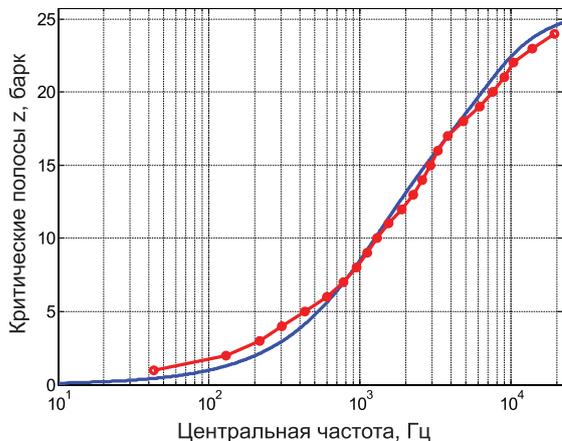


Рис. 3. Аппроксимация центральных частот СВ – WPD: $(l,n) \in E_{CB}, l=0,8$

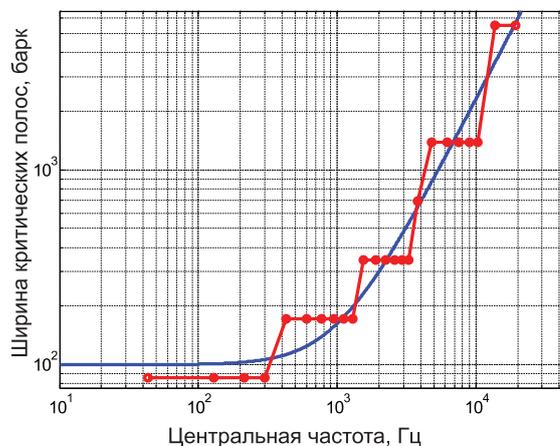


Рис. 4. Аппроксимация ширины критических частотных полос СВ – WPD: $(l,n) \in E_{CB}, l=0,8$

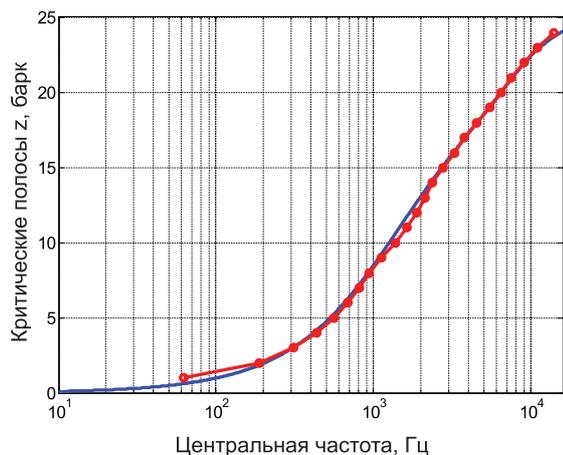


Рис. 5. Аппроксимация центральных частот $CB - WPD: (l, n) \in E_{CB}, l=0,7$

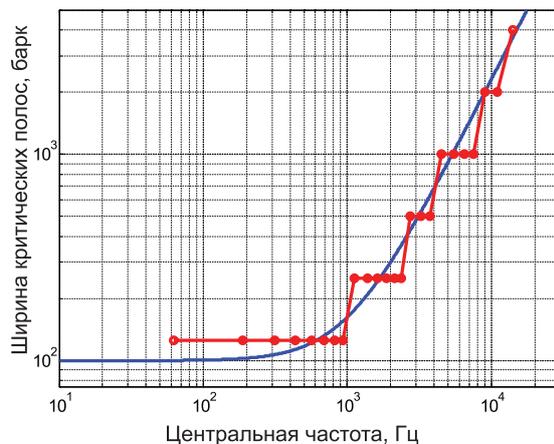


Рис. 6. Аппроксимация ширины критических частотных полос $CB - WPD: (l, n) \in E_{CB}, l=0,7$

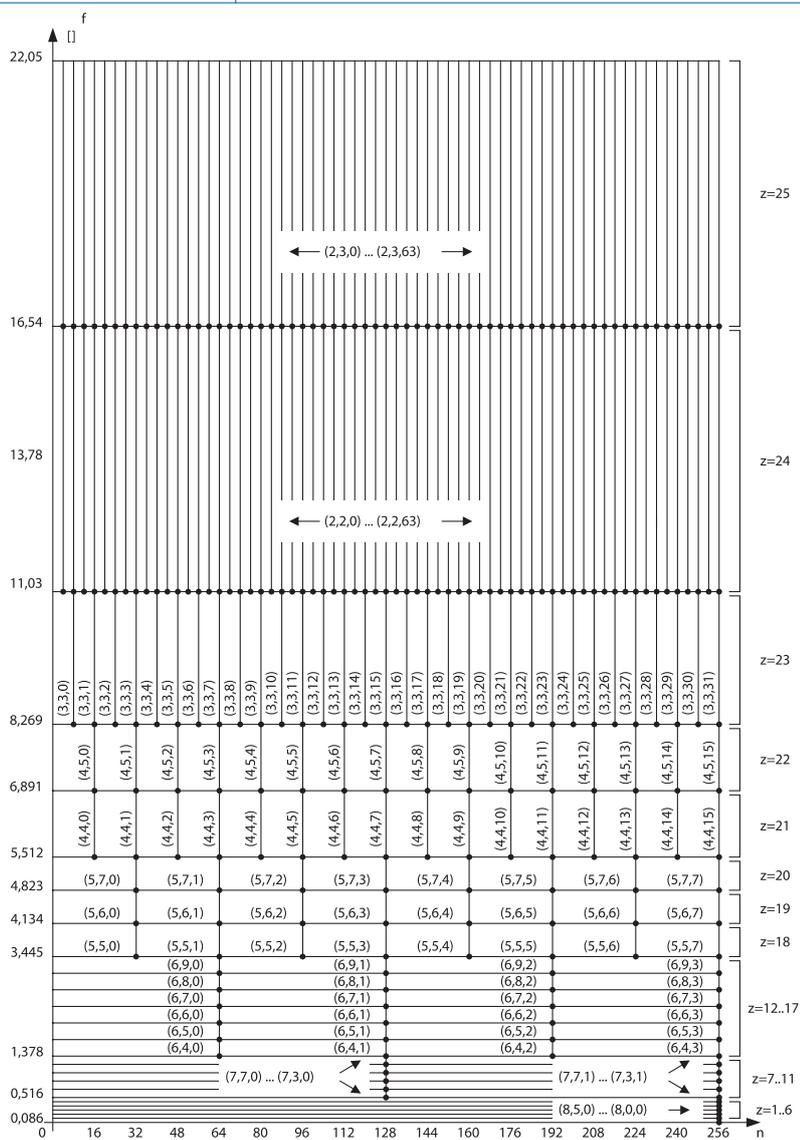


Рис. 7. Частотно-временной план структуры дерева ПДВП $CB - WPD: (l, n) \in E_{CB}, l=0,8, f_s = 44.1 \text{ кГц}$

Частотно-временной план для структуры дерева ПДВП $CB - WPD:(l,n) \in E_{CB}, l = \overline{0,8}, f_s = 44.1$ кГц (рис. 1) [3,8] представлен на **рис. 7**. Ширина каждой клеточки есть длина фрейма и определяется как $F_l = 2^l$ ($F_{min} = 2$ отсчётам и $F_{max} = 256$ отсчётам). Следовательно, длина анализируемого окна равна $W = (P - 1)(F_{l-1}) + 1$ отсчётов. Для первого уровня $l = 1$ преобразования определяющей является область верхних частот и длина окна W равна 40 отсчётам при длине фильтра прототипа $P = 40$. Для уровня $l = 8$ преобразования наибольшая частотная разрешающая способность находится в области нижних частот, а длина окна W равна 9946 отсчётам.

В **табл. 1** и **табл. 2** приведены показатели аппроксимации критических частотных полос структурами деревьев ПДВП $CB - WPD:(l,n) \in E_{CB}, l = \overline{0,8}$ и $CB - WPD:(l,n) \in E_{CB}, l = \overline{0,7}$ соответственно.

Таблица 1

**Аппроксимации критических частотных полос деревом ПДВП
($l,n) \in E_{CB}, l = \overline{0,8}, f_s = 44.1$ кГц**

№ барка	Узел	Кол-во вейвлет-коэфф.		Параметры полосы (Гц)		
		n	K	Нижняя	Центр	Верхняя
1	8	0	1	0.00	43.07	86.13
2	8	1	1	86.13	129.20	172.27
3	8	2	1	172.27	215.33	258.40
4	8	3	1	258.40	301.46	344.53
5	8	4	1	344.53	387.59	430.66
6	8	5	1	430.66	473.73	516.80
7	7	3	2	516.80	602.93	689.06
8	7	4	2	689.06	775.19	861.33
9	7	5	2	861.33	947.46	1033.59
10	7	6	2	1033.59	1119.72	1205.86
11	7	7	2	1205.86	1291.99	1378.13
12	6	4	4	1378.13	1550.39	1722.66
13	6	5	4	1722.66	1895.42	2068.19
14	6	6	4	2068.19	2239.95	2411.72
15	6	7	4	2411.72	2583.98	2756.25
16	6	8	4	2756.25	2928.51	3100.78
17	6	9	4	3100.78	3273.04	3445.31
18	5	5	8	3445.31	3789.84	4134.37
19	5	6	8	4134.37	4478.89	4823.41
20	5	7	8	4823.41	5167.95	5512.50
21	4	4	16	5512.50	6201.56	6890.63
22	4	5	16	6890.63	7579.69	8268.75
23	3	3	32	8268.75	9646.87	11025.00
24	2	2	64	11025.00	13781.16	16537.50
25	2	3	64	16537.50	19293.75	22050.00



Таблица 2

Аппроксимации критических частотных полос деревом ПДВП
 $(l, n) \in E_{CB}, l = \overline{0, 7}, f_s = 32 \text{ кГц}$

№ барка	Узел	Кол-во вейвлет-коэфф.		Параметры полосы (Гц)		
		n	K	Нижняя	Центр	Верхняя
1	7	0	1	0.00	62.5	125.0
2	7	1	1	125.0	187.5	250.0
3	7	2	1	250.0	312.5	375.0
4	7	3	1	375.0	437.5	500.0
5	7	4	1	500.0	562.5	625.0
6	7	5	1	625.0	687.5	750.0
7	7	6	1	750.0	812.5	875.0
8	7	7	1	875.0	937.5	1000.0
9	6	4	2	1000.0	1125.0	1250.0
10	6	5	2	1250.0	1375.0	1500.0
11	6	6	2	1500.0	1675.0	1750.0
12	6	7	2	1750.0	1875.0	2000.0
13	6	8	2	2000.0	2125.0	2250.0
14	6	9	2	2250.0	2375.0	2500.0
15	5	5	4	2500.0	2750.0	3000.0
16	5	6	4	3000.0	3250.0	3500.0
17	5	7	4	3500.0	3750.0	4000.0
18	4	4	8	4000.0	4500.0	5000.0
19	4	5	8	5000.0	5500.0	6000.0
20	4	6	8	6000.0	6500.0	7000.0
21	4	7	8	7000.0	7500.0	8000.0
22	3	4	16	8000.0	9000.0	10000.0
23	3	5	16	10000.0	11000.0	12000.0
24	2	3	32	12000.0	14000.0	16000.0

2. Процедура расчёта порогов маскирования в вейвлет-области

Пусть даны $CB - WPD: (l, n) \in E_{CB}$, частотно-временной план дерева ПДВП E_{CB} и коэффициенты $X_{l,n,k}$. Процедура расчёта порогов маскирования выглядит следующим образом [8]:

- Вычислить спектральную энергию барка:

$$A_{CB}(z) = \sum_{k=0}^{K-1} X_{z,k}^2, \quad (11)$$

- Оценить тональность сигнала в каждой критической частотной полосе и значения индексов $a_{mn}(z)$ и $a_{nmn}(z)$ уменьшения спектральной энергии барка соответственно для тоновых и шумовых маскерров:

— индекс $a_{\text{min}}(z)$, который оценивает отношение маскирования тоном шума, задается так:

$$a_{\text{min}}(z) = -0.275 \cdot z - 15.025 [\text{дБ}]; \quad (12)$$

— индекс маскирования шумом шума оценивается как константа

$$a_{\text{minn}}(z) = -25 [\text{дБ}], \quad (13)$$

так как ПДВП уже внесло некоторое спектральное перекрытие;

— среднее значение тональности маскеров в каждой критической полосе определяется маскирующим индексом

$$a_{CB}(z) = \eta \cdot a_{\text{min}}(z) + (1 - \eta) \cdot a_{\text{minn}}(z) [\text{дБ}], \quad z = \overline{1.25}, \quad (14)$$

где η — тональный коэффициент:

$$\eta = \min\left(\frac{SFM(z)_{\text{дБ}}}{SFM_{\text{дБ max}}}, 1\right), \quad (15)$$

где $SFM_{\text{дБ}}$ — мера спектральной пологости [11], рассчитываемая как

$$SFM(z) = \frac{\left(\prod_{k=0}^{K-1} X_{z,k}^2\right)^{\frac{1}{K}}}{\sum_{k=0}^{K-1} X_{z,k}^2}, \quad (16)$$

$SFM_{\text{дБ max}}$ — максимальное значение меры пологости спектра. Для заданного фильтра прототипа $SFM_{\text{дБ max}} = -25$ дБ.

● Найти спектральную энергию барка z с учётом тональности сигнала:

$$D_{CB}(z) = 10 \cdot \log\left(A_{CB}(z) \cdot 10^{\frac{a_{CB}(z)}{10}}\right) [\text{дБ}], \quad (17)$$

● Вычислить разброс энергии барка $C_{CB}(z)$ как свёртку $D_{CB}(z)$ с функцией разброса $B(z)$ в каждой критической частотной полосе z :

$$C_{CB}(z) = 10 \cdot \log\left(\frac{1}{K} \sum_{k=1}^{25} 10^{\frac{D_{CB}(k)}{10}} \cdot 10^{\frac{B(z-k)}{10}}\right) [\text{дБ}], \quad z = \overline{1.25}, \quad (18)$$

где функция разброса $B(z)$ вычисляется как

$$B(z) = a + \frac{v+u}{2} \cdot (z+c) - \frac{v-u}{2} \cdot \sqrt{d+(z+c)^2}, \quad (19)$$

а параметры функции a, v, u, d, c приведены в первой строке [табл. 3](#).

Таблица 3

Параметры функции разброса

Функция разброса	ν	u	d	c	a
Барк-шкала	30 дБ/барк	-25 дБ/барк	0.3	0.05	15
Временная шкала	$0.0825 \text{ дБ}/F_{min}^*$	$-0.0412 \text{ дБ}/F_{min}^*$	0.3	0.157	$0.032/F_{min}^*$

- Вычислить пороги маскирования во временной области. Аналогично, частотному маскированию, во временном маскировании уже присутствуют некоторые элементы перекрытия в АЧХ, обусловленные деревом ПДВП. Предполагается, что временное маскирование аддитивно к сигналу и определяется через коэффициенты ПДВП в каждой критической частотной полосе (см. рис. 8). Максимальное временное разрешение для ПДВП имеет место в критических частотных полосах верхних частот, которые имеют минимальную протяжённость по времени $F_{min}^* = 2$ отсчёта или 0.0454 мс. Временная функция разброса $B(k)$ в вейвлет-области задаётся так:

$$B(k) = a + \frac{\nu + u}{2} \cdot (k + c) - \frac{\nu - u}{2} \cdot \sqrt{d + (k + c)^2} \text{ [дБ]}. \quad (20)$$

Её параметры вдоль оси времени определяются как $\nu = 20 \text{ дБ}/\text{мс} = 0.0825 \text{ дБ}/F_{min}^*$ и $u = 20 \text{ дБ}/\text{мс} = -0.0412 \text{ дБ}/F_{min}^*$ (см. табл. 3, строка 2. F_{min}^* — минимальная длина анализируемого фрейма). На рис. 8 схематически показано проявление временного маскирования.

- Вычислить энергию вейвлет-коэффициентов в каждой критической частотной полосе z :

$$E_z(k) = X_{z,k}^2, k = \overline{1, K-1}, z = \overline{1.25}. \quad (21)$$

- Определить временную функцию разброса энергии в каждой критической частотной полосе z как свёртку $E_z(k)$ и функции разброса $B(k)$:

$$F_z(m) = \frac{1}{K} \sum_{k=0}^{K-1} E_z(k) \cdot 10^{\frac{B(K-k)}{10}}, m = \overline{1, K-1}. \quad (22)$$

- Найти временной фактор маскирования в полосе z как результат сравнения величин:

$$F_z(m) \geq E_z(m), k = \overline{1, K-1}. \quad (23)$$

Если данное соотношение выполняется, то в соответствующей критической частотной полосе имеет место временное маскирование, в противном случае его нет.

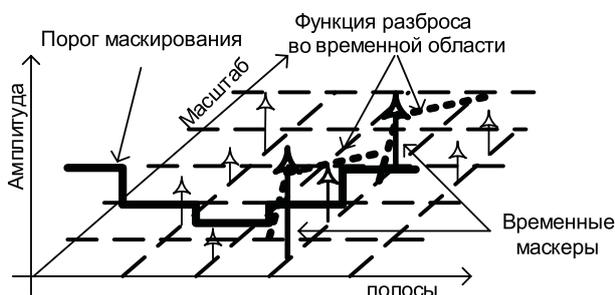


Рис. 8. Схема временного маскирования в соответствующей критической частотной полосе

- Оценить частотно-временной маскирующий порог $M_{CB}(z)$ в каждой критической частотной полосе:

$$M_{CB}(z) = C_{CB}(z) \cdot \max\left(\frac{F_z(k)}{E_z(k)}, 1\right) [\text{дБ}], k = \overline{1, K-1}. \quad (24)$$

- Вычислить значение маскирующего порога $T_{CB}(z)$ в соответствующей критической частотной полосе из сравнения временно-частотного маскирующего порога $M_{CB}(z)$ с минимальным значением абсолютного порога слышимости $ATH(z)$ (3):

$$T_{CB}(z) = \max(ATH(z), M_{CB}(z)) [\text{дБ}]. \quad (25)$$

Заключение

Предложенная процедура расчёта параметров психоакустической модели в области вейвлет-коэффициентов нашла применение при реализации систем кодирования звуковых и широкополосных речевых сигналов [3], а также используется в моделировании переходных компонент сигнала на основе согласованной подгонки с адаптивным словарём, формируемым структурой дерева ПДВП в параметрических кодерах [12]. Ведение обработки сигнала и реализация психоакустической модели в одной области исключает необходимость перерасчёта параметров модели, что позволяет предотвратить дополнительную погрешность.

Литература

1. M. Vetterli, J. Kovacevic. Wavelets and Subband coding. — Englewood Cliffs, NJ: Prentice-Hall, 1995. 488 p.
2. A. Spanias, T. Painter, V. Atti. Audio signal processing and coding. — Hoboken, NJ: John Wiley & Sons, Inc., 2007. 464 p.
3. A.I. Petrovsky, D. Krahe, A.A. Petrovsky. Real-time wavelet packet-based low bit rate audio coding on a dynamic reconfiguration system. // AES, Convention paper 5778, 114th Convention, 2003 March 22-25, Amsterdam, The Netherlands. 22 p.
4. Петровский А. Субполосное перцептуальное кодирование звуковых и речевых сигналов. Мн: Бестпринт, 2008 г. 220 с.
5. E. Zwicker, H. Fastl. Psychoacoustics: Facts and Models. — Berlin, Germany: Springer-Verlag, 1990. 380p.
6. M.V. Wickerhauser. Adaptive Wavelet Analysis from Theory to Software. — A.K. Peters Ltd., Massachusetts, 1994. 486 p.
7. ITU-R Rec. BS.1387, Method for Objective Measurements of Perceived Audio Quality, 998.
8. A.I. Petrovsky, M. Parfieniuk, A. Borowicz, A. Petrovsky. Auditory modeling via frequency warped transforms. // AES, Convention paper 7446, 124th Convention, May 2008, Amsterdam, The Netherlands. 15 p.
9. D. Sinha, A.H. Tewfik. Low bit rate transparent audio compression using adapted wavelets. // IEEE Trans. Signal Processing, vol.41, no.12, pp. 3463–3479, Dec. 1993.
10. B. Carnero, A. Drygajlo. Perceptual speech coding and enhancement using frame-synchronized fast wavelet packet transform algorithms. // IEEE Trans. Signal Processing, vol.47, no.6, pp. 1622–1635, June 1999.
11. Ковалгин Ю.А., Вологдин Э.И. Цифровое кодирование звуковых сигналов. Санкт-Петербург: Корона-Принт, 2004 г. 240 с.
12. A.I. Petrovsky, E. Azarov, A. Petrovsky. Harmonic representation and auditory model-based parametric matching and its application in speech/audio analysis. // AES, Convention paper 7705, 126th Convention, May 2009, Munich, Germany. 13 p.