АКУСТИЧЕСКИЙ ЖУРНАЛ, 2004, том 50, № 4, с. 562–567

БИОЛОГИЧЕСКАЯ АКУСТИКА

УДК 534.75

СЛУХОВОЙ АНАЛИЗ ПЕРИОДИЧНОСТИ ЗВУКА И ЕГО ОГИБАЮЩЕЙ: МАТЕМАТИЧЕСКАЯ МОДЕЛЬ

© 2004 г. С. М. Ищенко

Акустический институт им. Н.Н. Андреева РАН

117036 Москва, ул. Шверника, 4

E-mail: bvp@akin. ru

Поступила в редакцию 10.02.2004 г.

Представлена математическая модель слухового анализа периодичности звука и его огибающей в виде последовательности математических преобразований, описанных в соответствии со ступенями обработки сигнала. Предлагаемая модель учитывает следующие параметры и свойства слуховой системы:

• грубый спектральный анализ входного акустического сигнала с точностью до ширины критической полосы слуха;

• зависимость ширины критической полосы слуха от частоты;

• спектральный анализ входного сигнала выполняется набором из 3500 плотно расположенных по частоте фильтров;

• значение абсолютных слуховых порогов на данной частоте;

• временной анализ сигналов на выходе каждого фильтра и временной анализ профиля огибающей при помощи функции периодичности;

• импульсную активность нейронов слуховой системы;

• способность слуховой системы запоминать как спектрально-временной образ сигнала, так и отдельные его параметры;

• способность слуховой системы формировать ощущение громкости звука, запоминать и сравнивать значения громкостей звука в различные моменты времени, делать вывод о том, какая из них больше или меньше, или о равенстве их между собой с точностью до некоторого порога;

• зависимость ширины критической модуляционной полосы от частоты модуляции;

• зависимость порогов восприятия амплитудной модуляции от частоты модуляции;

На примере обработки амплитудно-модулированных сигналов с различными отношениями частоты несущего колебания к частоте огибающей показано, что предлагаемая модель удовлетворительно объясняет их высоту и слуховую оценку периода огибающей.

В данной работе обсуждается математическая модель слухового анализа периодичности звука и его огибающей в виде последовательности математических преобразований в соответствии с описанием функций схемы, представленной в работе [1].

В соответствии со схемой модели, акустический сигнал $p(t)$ на входе поступает на ряд из 3500 параллельных, плотно перекрывающихся частотных фильтров $ф_n1$–$ф_nk$, центральные частоты fk которых расположены равномерно на оси частот в логарифмическом масштабе. Диапазон частот, пропускаемых этими фильтрами, равен критической полосе слуха $\Delta F_{кр}$ [2, 3]. Если частоты спектральных компонент сигнала разделены частотным интервалом большим критической полосы слуха, то они воспринимаются слухом раздельно и при этом действует спектральный механизм анализа. Если спектральные компоненты входного сигнала сосредоточены в пределах одной критической полосы слуха, то они воспринимаются слитно, как единый сигнал, точнее как сумма спектральных компонент входного сигнала с соответствующими амплитудами и фазами. Предполагается , что в последнем случае энергия сигнала распределяется по всем плотно перекрывающимся по частоте фильтрам внутри критической полосы слуха в соответствии с формой “слухового фильтра”. В этом случае действует временной механизм анализа звука. Частотный интервал между соседними плотно перекрывающимися фильтрами $\Delta f_{сф}$ в пределах одной критической полосы слуха вычислялся по формуле:

$$\Delta f_{сф} = \Delta F_{кр}(f)/n_ф, \quad (1)$$

где f – частота, Гц, $n_ф$ – число фильтров, укладывающихся в пределах одной критической полосы слуха, $n_ф = 3500/24 \approx 146$, так как число критических (не перекрывающихся) полос слуха, последовательно укладывающихся в звуковом диапазоне частот, равно 24 [2]. Экспериментальная зависимость ширины критической полосы слуха $\Delta F_{кр}(f)$

от частоты приведена в работах [2, 3] и аппроксимирована в данной работе выражениями:

$$\Delta F_{кр}(f) = 40\lg(f) \text{ при } f \le 250, \quad (2)$$

$$\Delta F_{кр}(f) = f/\lg(f + 1.4) \text{ при } 250 < f < 1000 \text{ и при } f > 3000, \quad (3)$$

$$\Delta F_{кр}(f) = f/[\lg(f) + 2.4] \text{ при } 1000 \le f \le 3000. \quad (4)$$

Амплитудно-частотные характеристики (АЧХ) фильтров заимствованы из работ [4, 5], в которых форма фильтров, называемых "слуховыми", представлена в виде:

$$W(g) = (1 + pg)^{1/2}e^{-pg/2}, \quad (5)$$

где p – уровень звукового давления сигнала, $g = |f - f_k|/f_k$ – относительная разность между частотой сигнала f_k и текущей частотой f. При равенстве частоты f_k входного сигнала, на входе k-го фильтра, его центральной частоте, АЧХ указанного фильтра можно представить в виде (6):

$$W_k(g_k) = (1 + u_k g_k)^{1/2}e^{-u_k g_k/2}, \quad (6)$$

где $u_k = \alpha[p_k(f_k) - p_a(f_k)]$ – превышение уровня сигнала $\alpha p_k(f_k)$ на входе k-го фильтра над абсолютным слуховым порогом $\alpha p_a(f_k)$ на той же частоте, $p_a(f_k)$, $g_k = |f - f_k|/f_k$ – относительная разность между частотами f и f_k. α – коэффицент пропорциональности.

Поскольку форма фильтров зависит от параметра u_k, то есть от превышения уровня сигнала на центральной частоте k-го фильтра над абсолютным слуховым порогом, то для вычисления формы фильтров необходимо учитывать зависимость абсолютных слуховых порогов $p_a(f_k)$ от частоты. Усредненная экспериментальная зависимость последних от частоты приведена в работе [2] и приближенно описана следующими пятью выражениями в различных диапазонах частот при $x = \lg(f)$ и $\Delta x = |x - x_0|$:

$$p_a(f), \text{дБ} = 14.0\Delta x - 51x + 140, \text{ где } x_0 = 2.0 \text{ и } 20 < f < 1000; \quad (7)$$

$$p_a(f), \text{дБ} = 3.3\Delta x, \text{ где } x_0 = 3.3 \text{ и } 1000 < f < 3000; \quad (8)$$

$$p_a(f), \text{дБ} = 3.0f\Delta x, \text{ где } x_0 = 3.4 \text{ и } 3000 < f < 12000; \quad (9)$$

$$p_a(f), \text{дБ} = 951.56\Delta x, \text{ где } x_0 = 4.00688 \text{ и } 12000 < f < 15000; \quad (10)$$

$$p_a(f), \text{дБ} = fx, \text{ где } 15000 < f < 20000. \quad (11)$$

Выражения (7)–(11) позволяют приближенно вычислять абсолютные слуховые пороги $p_a(f)$, которые будут отличаться от усредненных экспериментальных не более чем на 1.2 дБ в диапазоне частот от 20 Гц до 12 кГц и не более чем на 4 дБ на частотах выше 12 кГц. Кроме того, нуль децибел в формуле (7) произвольно установлен на частоте 2 кГц.

В соответствии с блок-схемой обобщенной модели слухового механизма анализа периодичности звука [1], ко входному сигналу $p(t)$ прибавляется положительная постоянная составляющая $A_1 = \min|p(t)|$, равная абсолютному значению минимума входного сигнала, что переводит все значения указанных сигналов в положительную область и позволит сохранить временной профиль $\alpha p_k(t)$ спектральной компоненты входного сигнала на входе k-го фильтра. Как показано в работе [6], временной профиль входного сигнала сохраняется в рецепторных потенциалах волосковых клеток органа Корти.

Импульсный характер активности нейронной слуховой сети моделировался следующим образом. При отсутствии сигнала на входе модели, на выходе фильтров ф$_n$1–ф$_n k$ действует только спонтанная импульсная активность, распределение межимпульсных интервалов которой подчиняется закону Пуассона (12):

$$p_{m,n} = (Y^m/m)e^{-Y}, \quad (12)$$

где $p_{m,n}$ – вероятность того, что событие A появится m раз при n независимых испытаниях, $Y = np$.

Плотность спонтанной импульсной активности D_s волокон слухового нерва у человека, приходящуюся на одну критическую полосу слуха, можно оценить по их общему числу (30000), числу критических полос слуха (24), последовательно укладывающихся в звуковом диапазоне частот, и известной средней плотности спонтанного потока импульсов (50 имп./с) одиночных волокон слухового нерва. Таким образом, на одну критическую полосу слуха приходится около 1250 = 30000/24 волокон слухового нерва, а средняя плотность спонтанного потока импульсов волокон слухового нерва человека, приходящаяся на одну критическую полосу слуха оказывается равной $D_s = 1250 \times 50 = 62500$ имп./с. Отметим, что человек с нормальным слухом не ощущает спонтанную импульсную активность нейронов слуховой системы.

Допустим, что $\alpha p_k(\omega_k, t)$ – сигнал на входе k-го фильтра, а $D_{k0}(\omega_k)$ – плотность импульсной активности на выходе того же фильтра на пороге слышимости сигнала. Тогда зависимость плотности импульсной активности $D_k(\omega_{k,t})$ от частоты и времени на выходах фильтров ф$_n$1–ф$_n k$ можно представить в виде выражения:

$$D_k(\omega_k, t) = D_{k0}(\omega_k)[A_1 + \alpha p_k(\omega_k, t)]. \quad (13)$$

Распределение плотности импульсной активности $D_k(\omega_k, t)$ на выходах совокупности фильтров ф$_n$1–ф$_n k$, в данный момент времени , можно рассматривать как мгновенный спектр мощности входного сигнала, а изменение указанного спектра во времени представляет собой спектрально-временной образ входного сигнала, иногда называемый его "полным описанием" [7]. Полное описание сигнала запоминалось и, с выхода каждого фильтра, анализировалось функцией периодичности $P_k(\tau)$ [8]. $P_k(\tau) = S_k(\tau)/S_k(\tau = 0)$, где

$$S_k(\tau) = \\ = (1/T_c)\int_0^{T_c}\chi[|D_k(\omega_k, t) - D_k(\omega_k, t+\tau)| - \varepsilon]dt, \quad (14)$$

где τ – задержка времени, T_c – длительность анализируемого сигнала, t – текущее время, $\varepsilon \ll \ll \max[D_k(\omega_k, t)]$, $\chi(r) = 1$ при $r \le 0$ и $\chi = 0$ при $r > 0$; $r = [|D_k(\omega_k, t) - D_k(\omega_k, t+\tau)| - \varepsilon]$. Период сигнала выделялся по наибольшему максимуму функции периодичности для наименьшей задержки времени τ не равной нулю. Значения выделенных максимумов функции периодичности в различных частотных каналах умножалось на u_k (см. выражение 6). При этом $f = 1/\tau$.

Если уровень спектральной компоненты входного сигнала на центральной частоте f_k данного фильтра был меньше уровня спектральной компоненты сигнала на пороге его слышимости, то на выходе данного фильтра генерировалась плотность импульсной активности, моделирующая спонтанную активность волокон слухового нерва, распределение межимпульсных интервалов которой подчинялось закону Пуассона.

Предполагается, что плотность спонтанной импульсной активности на выходах фильтров постепенно уменьшалась от начальной D_s, при отсутствии сигнала, до нуля, при достижении плотности вынужденного (под действием сигнала) потока импульсов в данном частотном канале определенного значения $D1 = C1D_s$. Допустим, что $C1 \approx 1.5$.

Нетрудно показать, что при умеренных интенсивностях (20–40 дБ) звука его громкость $L(t)$ с точностью до постоянного коэффициента совпадает с его огибающей $M(t)$. Пусть $p(t)$ – амплитудно-модулированное (АМ) колебание вида:

$$p(t) = [1 + d_m\sin(\Omega t)]\sin(\omega t), \quad (15)$$

где d_m – коэффициент модуляции, $M(t) = 1 + + d_m\sin(\Omega t)$ – огибающая входного сигнала $p(t)$, $\sin(\omega t)$ – несущее колебание, Ω – частота модуляции, ω – частота колебаний несущей. Глубина АМ в процентах $d_{mp} = d_m \times 100\%$.

В соответствии с законом Стивенса громкость звука $L(t)$ удовлетворительно описывается следующим выражением:

$$L(t) = Cp(t)^a, \quad (16)$$

где C – постоянный коэффициент, $p(t)$ – звуковое давление, a – показатель степени, зависящий от интенсивности звука. При умеренных интенсивностях звука (20–40 дБ) коэффициент $a = 1$ и выражение (16) можно записать в виде:

$$L(t) = C2[1 + d_m\sin(\Omega t)] = C2M(t), \quad (17)$$

так как громкость синусоидального сигнала постоянной амплитуды постоянна по истечении некоторого времени, после включения импульса звука, то есть $C2 = C\sin(\omega t)$, где $C2$ – константа.

Таким образом, при умеренных интенсивностях звука его громкость оказывается пропорциональной огибающей $M(t)$ входного сигнала $p(t)$. Учитывая это, будем формировать огибающую $M(t)$ входного сигнала, путем вычисления функции $L(t)$, исходя из предположения, что громкость звука определяется суммарной импульсной активностью нейронов слуховой системы, вызываемой данным входным сигналом.

По различным экспериментальным данным громкость импульса звука устанавливается через 25–200 мс после его включения. Чаще всего, используется постоянная времени стабилизации громкости равная 100 мс [9]. Будем вычислять функцию $L(t)$, путем суммирования плотностей импульсной активности с выходов совокупности фильтров ф$_н$1–ф$_н k$ для каждого момента времени за промежуток времени равный $T_{sum} = 5$ мс.

Казалось бы, что временной интервал суммирования плотностей импульсной активности следовало бы положить равным 100 мс. Однако, в этом случае в функции $L(t)$, не будут представлены, в полной мере, частоты огибающих выше 10 Гц, так как указанное суммирование действует как фильтр низких частот.

С учетом того, что плотность импульсной активности $D_k(\omega_k, t)$ на выходе k-го фильтра описывается выражением (13), функцию $L(t)$ можно представить в виде:

$$M(t_i) \sim L(t_i) = \\ = 2\pi/[T_{sum}(\omega_2 - \omega_1)]\sum_{j=0}^{j_m}\sum_{k=1}^{k_m} D_k(\omega_k, t_i + t1_j), \quad (18)$$

где $t_i = i\Delta t$, $i = 0, 1, \ldots, i_m$; $i_m = T_c/\Delta t$; значение Δt определяется необходимой точностью вычисления функции $L(t_i)$, T_c – длительность входного сигнала $f_1 = 20$ Гц, $f_2 = 20$ кГц, $k_m = 3500$, $j = 0, 1, \ldots, j_m$; $j_m = = T_{sum}/\Delta t1$; $t1_j = j\Delta t_1$, $\Delta t_1 \approx 0.01T_{sum}$.

В случае непрерывного изменения частоты и времени, выражение (18) можно представить в виде

$$M(t) \approx L(t) = 2\pi/[T_c T_k(\omega_2 - \omega_1)] \times \\ \times \int_0^{T_{\text{sum}}} \int_{\omega_1}^{\omega_2} D(\omega, t + t_1) d\omega dt_1. \quad (19)$$

Функция $M(t) \approx L(t)$ поступает на ряд плотно перекрывающихся модуляционных фильтров ϕ_{m1} – ϕ_{ml} которые, по традиции, будем называть "слуховыми" фильтрами по огибающей. Наличие модуляционных фильтров [10] и слуховых фильтров, также как и критической полосы слуха и критической модуляционной полосы слуха [11], позволяют предположить, что слуховой механизм анализа периодичности огибающей имеет также спектрально-временной характер и, кроме того, его принципы анализа периодичности огибающей подобны принципам слухового анализа периодичности звука. При этом, если спектральные компоненты огибающей входного сигнала сосредоточены в пределах одной критической модуляционной полосы слуха, то они суммируются с соответствующими амплитудами и фазами, а энергия суммарного колебания распределяется по совокупности модуляционных фильтров, расположенных в пределах одной критической модуляционной полосы слуха в соответствии с формой слуховых фильтров по огибающей.

Если спектральные компоненты огибающей расположены в различных критических модуляционных полосах слуха, то они обрабатываются в каждой полосе отдельно. Форму слуховых фильтров по огибающей представим в виде выражения:

$$W_m(g_m) = (1 + p_m g_m)^{1/2} e^{-p_m g_m/2}, \quad (20)$$

где $g_m = \beta|f_m - f_{ml}|/f_{ml}$ – относительная разность между текущей частотой амплитудной модуляции f_m и частотой f_{ml} спектральной компоненты огибающей в l-ом частотном канале по огибающей, $\beta = \Delta F_{\text{крм}}/\Delta f_{m,\,0.7}$, $\Delta f_{m,\,0.7}$ – ширина модуляционного фильтра $W_m(g_m)$ на уровне 0.7 от его максимального значения на центральной частоте f_{ml}. Коэффициент β введен для выравнивания ширины критической модуляционной полосы $\Delta F_{\text{крм}}(f_m)$, полученной экспериментально, и $\Delta f_{m,\,0.7}$; $p_m = (d_{mp} - p_{ml})$ – разность между глубиной модуляции d_{mp}, вызываемой l-ой спектральной компонентной огибающей с частотой f_{ml} и значением слухового порога $p_a(f_{ml})$ восприятия АМ, на этой же частоте АМ f_{ml}.

Экспериментальная зависимость слуховых порогов восприятия АМ от частоты f_m приведена в работе [11]. Поскольку зависимость $p_{mn}(f_m)$ слуховых порогов восприятия АМ шума от частоты АМ существенно отличается от таковой $p_{mt}(f_m)$ для АМ, тона, то указанные зависимости аппроксимированы следующими пятью выражениями (21)–(25) (в процентах глубины АМ) для АМ шума и для АМ, тона отдельно:

$$p_{mn}(f_m) = 3.00 + (3.00 - f_m)^2 \times 0.2 \text{ при } f_m < 3, \quad (21)$$

$$p_{mn}(f_m) = 3.00 + 0.06 f_m \text{ при } f_m \geq 3, \quad (22)$$

$$p_{mt}(f_m) = 2.5 + (3.0 - f_m)^2 \times 0.2 \text{ при } f_m < 3.0, \quad (23)$$

$$p_{mt}(f_m) = 2.5 + (f_m - 3.0)^2 \times 0.1 \\ \text{при } 3 \text{ Гц} < f_m \leq 11, \quad (24)$$

$$p_{mt}(f_m) = 3.1 + (f_m - 11.0) \times 0.007 \text{ при } f_m > 11. \quad (25)$$

Экспериментальная зависимость ширины критической модуляционной полосы слуха $\Delta F_{\text{крм}}$ от частоты модуляции f_m приведена в работе [12] и представлена тремя выражениями (26)–(28) в различных диапазонах модулирующих частот:

$$\Delta F_{\text{крм}}(f_m) = 0.33 f_m \text{ при } f_m \leq 17, \quad (26)$$

$$\Delta F_{\text{крм}}(f_m) = 0.23 f_m \text{ при } 17 < f_m \leq 30, \quad (27)$$

$$\Delta F_{\text{крм}}(f_m) = 5.00 + 0.025 f_m \text{ при } f_m \leq 30. \quad (28)$$

Число $n_{\text{фм}}$ слуховых фильтров по огибающей, укладывающихся в пределах одной критической модуляционной полосы слуха оценим по отклонению частоты модуляции, которое уверенно обнаруживает слуховая система. Последнее составляет 0.07 на частоте модуляции 1 Гц [8]. Таким образом, $n_{\text{фм}} = 1/0.07 \approx 14$. Частотный интервал между соседними модуляционными фильтрами вычислим по формуле:

$$\Delta f_{\text{мсф}} = \Delta F_{\text{крм}}(f_m)/n_{\text{фм}}. \quad (29)$$

С выхода каждого частотного канала по огибающей сигналы $M_1(t) \approx L_1(t)$ обрабатываются функцией периодичности $P(\tau_1) = S(\tau_1)/S(\tau_1 = 0)$, где

$$S_l(\tau_1) = [1/T_c] \int_0^{T_c} \chi[|L_l(t) - L_l(t + \tau_1)| - \varepsilon_1] dt, \quad (30)$$

где $\tau_1 = j_1 \Delta\tau_1$ – задержка времени, $j_1 = 0, 1, \ldots, j_m$; $j_m = T_l/\Delta\tau_1 \approx 100$, T_l – период спектральной компоненты огибающей в l-ом частотном канале, $\varepsilon_1 \ll \max[L_l(t)]$. Период сигнала $L_l(t)$ определялся по наибольшему максимуму функции периодичности $P_l(\tau_1)$ для наименьшей задержки времени r_1 не равной нулю. Значения выделенных максимумов функции периодичности умножались на значение p_m в данном частотном канале f_{ml} и проецировались на единую частотную ось.

Сигналы с выходов слуховых фильтров и с выходов слуховых модуляционных фильтров по каждому частотному каналу обрабатывались функцией периодичности отдельно и результат анализа с выхода каждого частотного канала имеет свои ось абсцисс (ось частот или ось задержек времени) и ось ординат (от нуля до единицы), поэтому для обобщения результатов анализа периодичности звука и его огибающей все результаты анализа отображались (проецировались) на единую частотную ось в диапазоне частот от 0.01 Гц до 20 кГц. Нижняя граница восприятия частоты а.м., равная 0.01 Гц, определена экспериментально. При наложении одного максимума на другой их значения суммировались для каждого значения частоты отдельно.

Первые 6–7 наибольших выделенных, таким образом, максимумов на общей частотной оси принимались за оценку периодичности входного сигнала. В работе [13] показано, что слуховая система неспособна выделить более 6–7 максимумов сложного периодического звука.

Рассмотрим несколько примеров обработки предлагаемой моделью АМ тона, описываемого выражением (15), $p(t) = [1 + d_m\sin(\Omega t)]\sin(\omega t)$, при различных отношениях частоты ω несущего колебания к частоте Ω модуляции.

Пример 1. Допустим, что на вход модели поступает АМ тон, описываемый выражением (15). При кратном отношении частоты ω несущего колебания к частоте огибающей Ω, то есть $\omega/\Omega = N$, где N – целое число большее или равное 10, спектр такого сигнала будет состоять из трех спектральных компонент с амплитудой центральной компоненты $\sin(\omega t)$ равной 1 и с амплитудами боковых компонент $\sin(\omega - \Omega, t)$ и $\sin(\omega + \Omega, t)$ равных $0.5d_m$. Если частоты спектральных компонент указанного сигнала сосредоточены в пределах одной критической полосы слуха, то входной сигнал будет обрабатываться моделью как сумма трех спектральных компонент входного сигнала. Энергия такого сигнала распределится по совокупности плотно расположенных по частоте слуховых фильтров в пределах одной критической полосы слуха в соответствии с формой слухового фильтра, центральная частота которого равна ω.

Поскольку период такого сигнала равен T_Ω, то на единую частотную ось спроецируется максимум равный 1 на частоте Ω. Огибающая входного сигнала, $m(t) = 1 + d_m\sin(\Omega t)$, поступит в слуховой фильтр по огибающей с центральной частотой Ω, форма которого будет спроецирована на единую частотную ось с максимумом равным d_m, также на частоте Ω. Таким образом, на единой частотной оси будет отражен максимум равный $(1 + d_m)$ на частоте Ω.

В том случае, если Ω превышает ширину критической полосы слуха, то три спектральных компоненты будут обработаны независимо друг от друга, поэтому на частоте ω будет отражена форма слухового фильтра с максимумом равным 1, а на частотах $\omega - \Omega$ и $\omega + \Omega$ будут отражены формы слуховых фильтров с максимумами равными $0.5d_m$.

По огибающей не будет отражено максимумов, так как интенсивность независимых спектральных компонент будет постоянной во времени, так же, как и их сумма, а функция периодичности игнорирует постоянную составляющую анализируемого сигнала.

Пример 2. Допустим, что $N_1 = N + 0.5$ или $N_1 = N - 0.5$. В этом случае период сигнала $p(t)$ будет равен $4\pi/\Omega$. Кроме того, указанный сигнал будет иметь два почти-периода равных $(T_\Omega + T_\omega/2)$ и $(T_\Omega - T_\omega/2)$.

Если три спектральных компоненты указанного сигнала будут расположены в пределах одной критической полосы слуха, то они будут обрабатываться моделью как их сумма и на единую частотную ось будут спроецированы три максимума на частотах $f_1 = 1/2T_\Omega$, $f_2 = 1/(T_\Omega + T_\omega/2)$ и $f_3 = 1/(T_\Omega - T_\omega/2)$ равные 1, $0.5d_m$, и $0.5d_m$, соответственно. По огибающей будет отражен максимум на частоте f_Ω равный $0.5d_m$.

Если спектральные компоненты этого сигнала будут расположены в различных критических полосах слуха, то есть при условии $\Omega > \Delta F_{кр}$, то на единой частотной оси будут отражены три максимума на частотах f_ω, $f_\omega + \Omega$ и $f_\omega - \Omega$ равные 1, $0.5d_m$ и $0.5d_m$, соответственно. По огибающей, по-прежнему, не будет отражено максимумов (см. пример 1 для случая $\Omega > \Delta F_{кр}$).

Пример 3. Допустим, что отношение частоты ω несущего колебания к частоте Ω огибающей в исследуемом сигнале $p(t)$ постепенно увеличивается от $N_1 = N$ до $N_2 = N + 2$.

В случае, когда спектральные компоненты сигнала, с указанными в данном примере параметрами, сосредоточены в пределах одной критической полосы слуха, то есть при $\Omega < \Delta F_{кр}$, то сигнал будет обрабатываться как сумма трех спектральных компонент. При $\omega/\Omega = N$, где N – целое число, на частоте Ω будет отражен максимум равный $(1 + d_m)$ (см. пример 1 при $\Omega < \Delta F_{кр}$). При постепенном увеличении частоты ω от $N_1 = N$ до $N_1 = N + 0.5$ длительность почти-периода будет постепенно увеличиваться от T_Ω до $(T_\Omega + T_\omega/2)$ и при значении $N_1 = N + 0.5$ появится второй почти-период равный $(T_\Omega - T_\omega/2)$ (см. пример 2). При дальнейшем увеличении частоты ω останется только один почти-период равный $(T_\Omega - T_\omega/2)$, который будет постепенно увеличиваться от $(T_\Omega - T_\omega/2)$ до T_Ω при значении $N_1 = N + 1$. При дальнейшем увеличении частоты ω от $\omega/\Omega = N + 1$ до $\omega/\Omega = N + 2$ процесс повторится.

В случае $\Omega > \Delta F_{кр}$ на частотах ω, $(\omega + \Omega)$ и $(\omega - \Omega)$ будут отражены три максимума равные 1, $0.5d_m$ и

$0.5d_m$, соответственно. По огибающей, по-прежнему, не будет отражено максимумов (см. пример 1).

Отметим, что трехмерное, спектрально-временное, представление входного сигнала оказывается эффективным не только при слуховом анализе периодичности звука, но и при выделении основного тона речевого сигнала [14]. Рассмотренные примеры обработки предлагаемой математической моделью АМ сигналов показывают, что предлагаемая модель удовлетворительно объясняет их высоту и слуховую оценку периода их огибающей.

СПИСОК ЛИТЕРАТУРЫ

1. *Ищенко С.М.* Обобщенная модель слухового механизма анализа и выделения периодичности сложных звуков. Блок-схема модели // Акуст. журн. 1998. Т. 44. № 2. С. 226–231.
2. *Фельдкеллер Р., Цвикер Э.* Ухо как приемник информации // Перевод с немецкого. М.: "Связь". 1965. С. 15.
3. *Фрейдин А.А.* Критическая полоса слуха. Измерение КП разными методами // Акуст. журн. 1975. Т. 21. № 5. С. 806–814.
4. *Patterson R.* Auditory filter shape // J. Acoust. Soc. Amer. 1974. V. 55. № 4. P. 802–809.
5. *Patterson R.* Auditory filter shapes derived with noise stimuli // J. Acoust. Soc. Amer. 1976. V. 59. № 3. P. 640–654.
6. *Dallos P.* Response characteristics of manumalian cochlear hair cells // J. Neuroscience. 1985. V. 5. № 6. P. 1591–1608.
7. *Чистович Л.А.* Психоакустика и вопросы теории восприятия речи // В кн. "Распознавание слуховых образов". Новосибирск. Наука. 1970. С. 98–106.
8. *Ищенко С.М.* Функция периодичности, ее основные свойства и сравнение возможностей нормированной автокорреляционной функции и функции периодичности в задаче выделения периода сигнала на фоне шума // Акуст. журн. 1988. Т. 34. № 2. С. 256–260.
9. *Телепнев В.Н.* Громкость // В кн. Слуховая система. Л. "Наука". 1990. С. 14–42.
10. *Ищенко С.М.* Слуховая оценка человеком частоты модуляции амплитудно-модулированного тона // Акуст. журн. 1977. Т. 23. № 1.С. 64–68.9.
11. *Тумаркина Л.Н., Дубровский Н.А.* Некоторые особенности восприятия человеком амплитудно-модулированных сигналов // Биофизика. 1966. Т. 11. Вып. 2. № 4. С. 653–659.
12. *Дубровский Н.А., Тумаркина Л.Н.* Исследование слухового механизма восприятия модуляций. Амплитудная модуляция двумя тонами // Труды Акустического ин-та. 1970. Т. 12. С. 48–63.
13. *Plomp R.* Detectability threshold for combination tones // J. Acoust. Soc. Amer. 1965. V. 37. P. 1110–1123.
14. *Granqvist S., Hammarberg B.* The correlogram: A visual display and periodicity // J. Acoust. Soc. Amer. 2003. V. 114. № 5. P. 2934–2945.

Auditory Analysis of the Periodicity and Envelope of Sound: A Mathematical Model

S. M. Ishchenko

A mathematical model of the auditory analysis of periodicity and envelope of sound is proposed. The model consists of a sequence of mathematical transformations that describe the signal processing stages. The following parameters and properties of the auditory system are taken into account: the crude analysis of the input acoustic signal accurate to the width of the aural critical band; the frequency dependence of the width of the aural critical band; the spectrum of the input signal analysis by a set of 3500 filters closely spaced in frequency; the absolute audibility thresholds at a given frequency; the time-domain analysis of both the output signals of each filter and the envelope profile with the help of the periodicity function; the pulsed activity of auditory neurons; the ability of the auditory system to memorize the frequency-domain and time-domain images of the signal and its individual parameters; the ability of the auditory system to form the perception of the sound volume, to memorize and compare the sound volume at different time moments, and to conclude which of them is higher or lower or whether they are equal accurate to a certain threshold; the dependence of the critical modulation bandwidth on the modulation frequency; and the dependence of the audibility thresholds of amplitude modulation on the modulation frequency. By an example of processing amplitude-modulated signals with various carrier-to-envelope frequency ratios, the model is shown to give a satisfactory explanation of their pitch and auditory estimate of the envelope period.