

Министерство образования Республики Беларусь

Учреждение образования
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ

Кафедра сетей и устройств телекоммуникаций

АНАЛИЗ ЧАСТОТНЫХ И ВРЕМЕННЫХ СВОЙСТВ
СЛУХОВОГО АППАРАТА

Методические указания
к лабораторной работе
по дисциплинам «Цифровая обработка речи и изображений» и
«Защита речевых сообщений и объектов связи
от несанкционированного перехвата»
для студентов специальности «Сети телекоммуникаций»
дневной, вечерней и заочной форм обучения

Минск 2003

УДК 621.391.25 (075.8)
ББК 32.811.4 я 73
А64

Составители:
А.А. Борискевич, В.К. Конопелько

А64 Анализ частотных и временных свойств слухового аппарата: Метод. указания к лабораторной работе по дисциплинам «Цифровая обработка речи и изображений» и «Защита речевых сообщений и объектов связи от несанкционированного перехвата» для студентов спец. «Сети телекоммуникаций» дневной, вечерней и заочной форм обучения/Сост. А.А. Борискевич, В.К. Конопелько. – Мн.: БГУИР, 2003. – 19 с.: ил.

В данных методических указаниях рассмотрены основные спектральные и временные свойства психоакустической модели слуховой системы. Приведено подробное описание данной модели в частотной области. Сведения, представленные в настоящих методических указаниях, могут быть использованы для решения задач обработки, сжатия и защиты звуковых сигналов.

УДК 621.391.25 (075.8)
ББК 32.811.4 я 73

1. ЦЕЛЬ РАБОТЫ

Изучение спектральных и временных свойств психоакустической модели слуховой системы для решения задач обработки, сжатия и защиты звуковых сигналов.

2. ТЕОРЕТИЧЕСКАЯ ЧАСТЬ

2.1. Введение

Человек в состоянии принимать огромные потоки информации. Но сознательно он способен обрабатывать лишь около 100 бит/с информации. Поэтому можно говорить о присущей звуковым сигналам существенной избыточности. Важной проблемой при цифровом представлении звуковых сигналов является сокращение имеющейся в них статистической и психофизи-ческой избыточности. Сокращение статистической избыточности основывается на учете свойств самих звуковых сигналов. Она обусловлена наличием корреляционной связи между соседними отсчетами звукового сигнала при его дискретизации. Устранение статистической избыточности не приводит к значительному уменьшению скорости цифрового потока. Наиболее перспективными с этой точки зрения оказались методы сокращения психофизической избыточности звуковых сигналов, учитывающие такие свойства слухового восприятия, как маскировка, предмаскировка и после-маскировка. Если известно, какие составляющие звукового сигнала ухо воспринимает, а какие нет вследствие явления маскировки, то нужно выделить и затем передать по каналу связи лишь те части сигнала, которые ухо способно воспринять, а неслышимые составляющие сигнала можно отбросить. Кроме того, сигналы можно квантовать с возможно меньшим разрешением так, чтобы шумы квантования, изменяясь по величине с изменением уровня самого сигнала, еще оставались бы неслышимыми, т.е. маскировались бы исходным сигналом. Учет психоакустических закономерностей слухового восприятия определяет качество систем кодирования со сжатием цифровых данных. Методами устранения психофизической избыточности можно обеспечить сжатие цифровых аудиоданных в 10-12 раз без существенных потерь в качестве.

2.2. Психоакустические особенности слухового восприятия

Описание характеристик любой слуховой системы осуществляется по частотно-пороговым кривым, которые отражают предельные значения интенсивности (или мощности звукового давления) в зависимости от частоты звуковых колебаний. Частотно-пороговые кривые, полученные для слуховой системы в целом, получили название кривых слышимости. Чувствительность слуха оценивается минимальной интенсивностью звука, при которой человек может отличить звук от постоянно существующего фона собственных шумов. Интенсивность, при которой звук обнаруживается с вероятностью 0,5 (в 50% случаев), называется порогом слышимости или абсолютным порогом для данного сигнала.

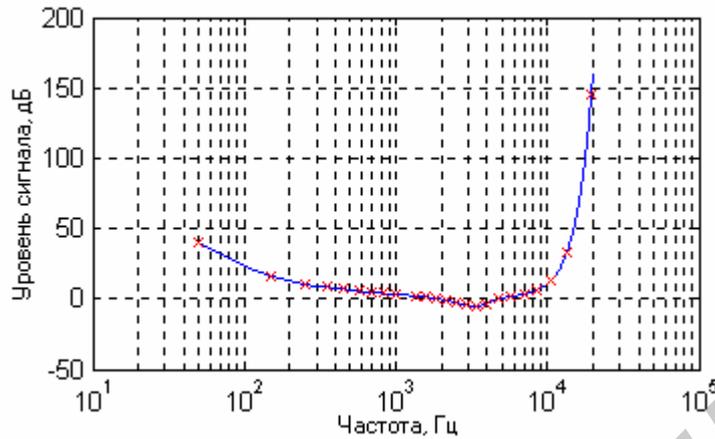


Рис.1. Кривая минимального порога слышимости слухового аппарата человека в тишине

Кривая порога слышимости (рис.1), характеризующая наименьшую интенсивность звука определенной частоты, который может быть услышан человеком в тишине, хорошо аппроксимируется с помощью нелинейной функции

$$T(f) = 3,64(f/1000)^{-0,8} - 6,5e^{-0,6(f/1000-3,3)^2} + 10^{-3}(f/1000)^4, \quad (1)$$

где $T(f)$ – порог слышимости, дБ; f – частота звукового сигнала, Гц. Данная кривая представляет собой геометрическое место точек, которые соответствуют тональным компонентам разных частот, имеющих одинаковую громкость. Анализ усредненного графика порога слышимости показывает, что слуховая система человека чувствительна к частотам в диапазоне 0,01–20 кГц. Наибольшая чувствительность у человека наблюдается на средних частот 2–5 кГц. Порог слышимости сигнала на частоте около 3 кГц составляет приблизительно 0 дБ. Порог слышимости значительно повышается в области низких (менее 200 Гц) и высоких (более 10 кГц) частот и может достигать значения громкости звука 40 дБ и выше. Следовательно, для слабых сигналов амплитудно-частотная характеристика слуха является нелинейной, и, как следствие, слабые низкочастотные и высокочастотные составляющие сигнала могут быть неслышны. Кривые слышимости позволяют оценить работоспособность слуховой системы по ряду признаков: диапазону воспринимаемых частот, полосе частот, в пределах которой чувствительность максимальна, абсолютным значениям звукового давления и их изменениям в зависимости от частоты.

Одно из важных свойств слухового аппарата человека является его ин-тегрирующая способность, т.е. группирование частотных составляющих звука в определенные частотные критические полосы. Слуховая система сравнивает полезный сигнал и мешающий шум по интенсивности в пределах критических полос слуха, оценивая порог слышимости. Критическая полоса слуха является полосой, за пределами которой субъективные ощущения звука сильно изменяются.

На рис.2 и 3 схематически показан метод определения ширины критической полосы слуховой системы.

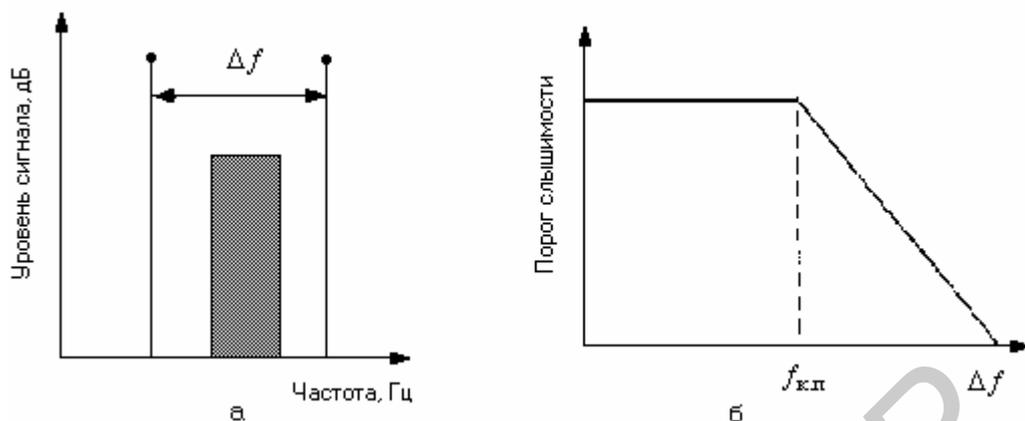


Рис. 2. Определение ширины критической полосы при маскировании узкополосного шумового компонента двумя тональными компонентами

Узкополосный шумовой сигнал маскируется двумя близко расположенными тональными компонентами, расстояние Δf между которыми постепенно увеличивается. На рис.2,б представлена зависимость поведения порога обнаружения шумовой компоненты от изменения расстояния между двумя тональными компонентами Δf . Данный порог остается постоянным до тех пор, пока частотное расстояние между тональными компонентами остается в пределах критической полосы $f_{к.п}$. Вне полосы порог быстро уменьшается.

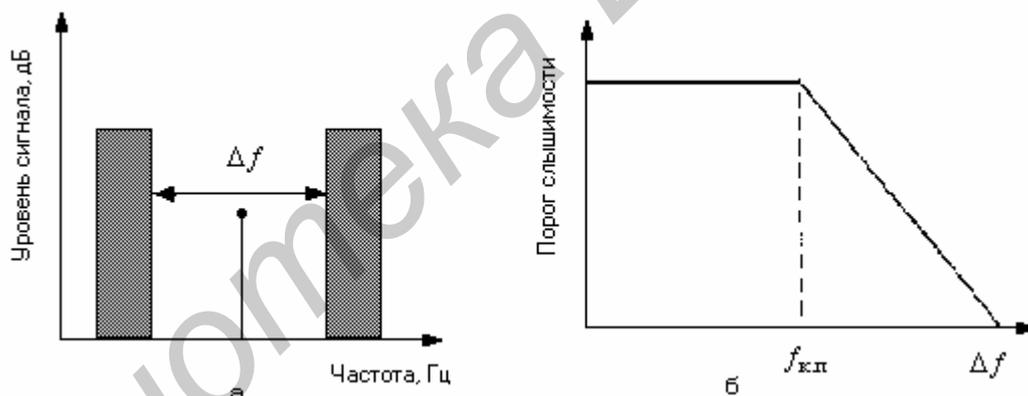


Рис.3. Определение ширины критической полосы при маскировании тонального компонента двумя узкополосными шумовыми компонентами

На рис. 3 приведен случай, когда в качестве маскируемого сигнала выступает тональная компонента, а в качестве маскирующих сигналов – две узкополосные шумовые компоненты (рис.3,а). Поведение порога обнаружения для тональной компоненты (рис.3,б) аналогично рис.2,б.

В таблице представлено разбиение слышимого диапазона частот на критические полосы с соответствующими центральной частотой и шириной полос идеализированного банка фильтров (рис.4) с прямоугольными амплитудно-частотными характеристиками, который позволяет качественно воспроизвести звуковые сигналы.

Разбиение слышимого частотного диапазона на критические полосы в психоакустической модели слухового аппарата

Номер критической полосы	Центральная частота, Гц	Нижняя и верхняя частоты, ширина критической полосы, Гц	
		Нижняя частота	Верхняя частота
1	50	20–100	100

2	150	100–200	100
3	250	200–300	100
4	350	300–400	100
5	450	400–510	110
6	570	510–630	120
7	700	630–770	140
8	840	770–920	150
9	1000	920–1080	160
10	1170	1080–1270	190
11	1370	1270–1480	210
12	1600	1480–1720	240
13	1850	1720–2000	280
14	2150	2000–2320	320
15	2500	2320–2700	380
16	2900	2700–3150	450
17	3400	3150–3700	550
18	4000	3700–4400	700
19	4800	4400–5300	900
20	5800	5300–6400	1100
21	7000	6400–7700	1300
22	8500	7700–9500	1800
23	10500	9500–12000	2500
24	13500	12000–15500	3500
25	19500	15500–22500	7000

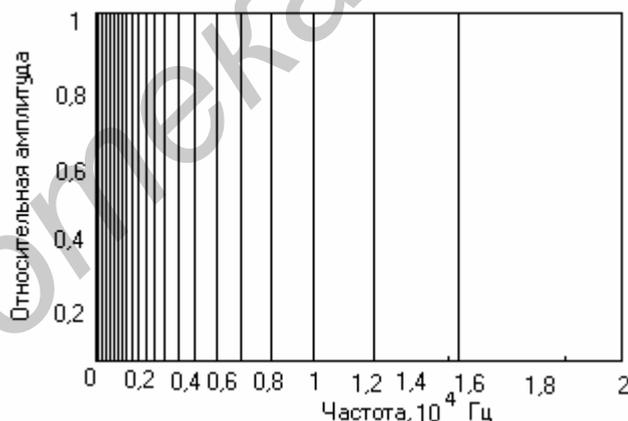


Рис. 4. Идеализированный банк фильтров как упрощенная модель слуховой системы. Из таблицы видно, что критическая полоса не соответствует диапазону с фиксированными нижней и верхней границами. Ширина критической полосы остается приблизительно постоянной (около 100 Гц) вплоть до 500 Гц и увеличивается приблизительно на 20% от центральной частоты при 500 Гц и выше:

$$BW(f) = \begin{cases} 100 \text{ Гц}, & f < 500 \text{ Гц}, \\ 0,2 f \text{ Гц}, & f \geq 500 \text{ Гц}. \end{cases} \quad (2)$$

Из (2) видно, что в частотной области до 500 Гц органы слуха воспринимают интенсивность звука, разделяя ее на участки постоянной абсолютной ширины, а свыше 500 Гц – на участки постоянной относительной ширины. Для среднего слушателя ширина критической полосы слуховых фильтров (рис.5) может быть определена из выражения

$$BW(f) = 25 + 75[1 + 1,4(f / 1000)^2]^{0,69} . \quad (3)$$

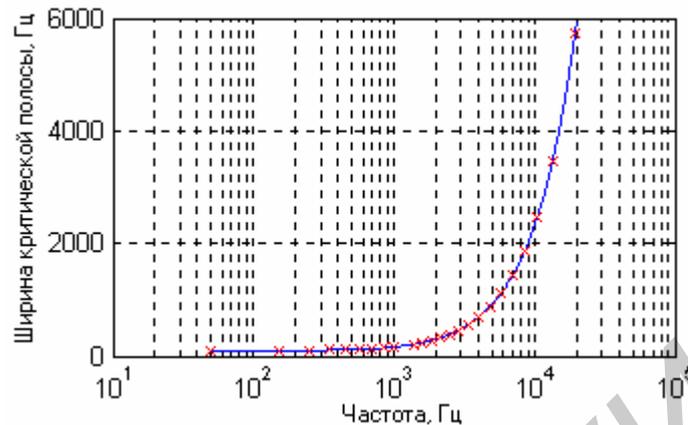


Рис. 5 . Ширина критических полос как функция центральной частоты

Из рис.5. видно, что по мере увеличения центральной частоты критическая полоса расширяется. Поскольку имеет место однозначная зависимость между частотой синусоидального колебания и высотой тона, то по аналогии с частотной группой слуха было введено понятие тональных групп слуха. Ширина тональной группы не зависит от высоты тона и составляет 100 мел (один мел равен ощущаемой высоте звука частотой 1000 Гц при уровне интенсивности 40 дБ), поэтому оказалось удобным ввести единицу высоты тона в одну тональную группу слуха. В честь немецкого физика Г. Баркгаузена она носит название барк (1 барк = 100 мел), равная ширине критической полосы или расстоянию между центральными частотами соседних критических полос. Переход от одной полосы к другой соответствует изменению высоты в 1 барк.

Функция для перехода от частоты в Гц к частоте в барках (рис.6) определяется соотношением

$$z(f) = 13 \operatorname{arctg}(0,00076 f) + 3,5 \operatorname{arctg} \left[\left(\frac{f}{7500} \right) \right]^2, \quad (4)$$

где $z(f)$ представляет собой частоту f в барках.

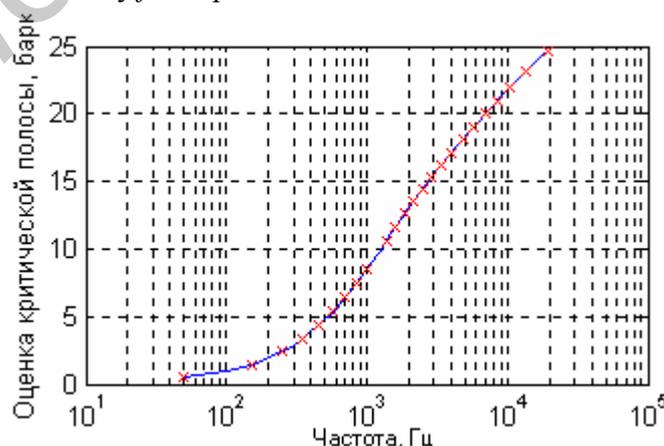


Рис. 6. Расстояние между центральными частотами критических полос

В основе приемов выделения в звуковом сигнале той части информации, которая четко воспринимается ухом, и удаления из сигнала всех остальных неслышимых деталей лежит эффект маскирования. Он связан с процессом взаимодействия сигналов, что приводит к изменению слуховой чувствительности к маскируемому сигналу в присутствии

маскирующего. Эффекты слухового маскирования зависят от спектральных и временных характеристик маскируемого сигнала и сигнала маскирования и могут быть разделены на две основные группы: частотное (одновременное) и временное (неодновременное) маскирования. Маскирование по частоте заключается в следующем: если два сигнала одновременно находятся в ограниченной частотной области, то более слабый сигнал становится неслышимым на фоне более сильного. Маскирование по времени определяет следующий эффект: более слабый сигнал становится неслышимым за 5 – 20 мс до включения сигнала маскирования и становится слышимым через 50 – 200 мс после его включения.

Маскирование непосредственно связано с нелинейностью слуха. Она проявляется в том, что при воздействии на барабанную перепонку достаточно громкого синусоидального звука с частотой f в слуховом аппарате формируются гармоники этого звука с частотами $2f$, $3f$ и т.д. Поскольку в первичном воздействующем тоне этих гармоник нет, они получили название субъективных гармоник.

Анализ кривых маскировки (кривых порога слышимости при наличии маскировки) показывает, что маскировка сигналов, имеющих частоту, лежащую ниже частоты мешающего тона, проявляется значительно слабее чем вышележащих частот. Это дает основание предполагать, что маскирующее действие обусловлено возникновением субъективных гармоник мешающего тона. Несмотря на то, что большую информацию об эффекте маскировки получают при исследованиях с тональными сигналами, при использовании возникает ряд трудностей. Например, если два тона близки по частоте, то между ними возникают биения.

Если в качестве мешающего звука использовать узкополосный белый шум, пороги слышимости определяются более четко, а сами кривые получаются более симметричными. Это связано с отсутствием гармонических биений. Большой практический интерес представляют кривые порога слышимости, полученные при маскировке широкополосными шумами.

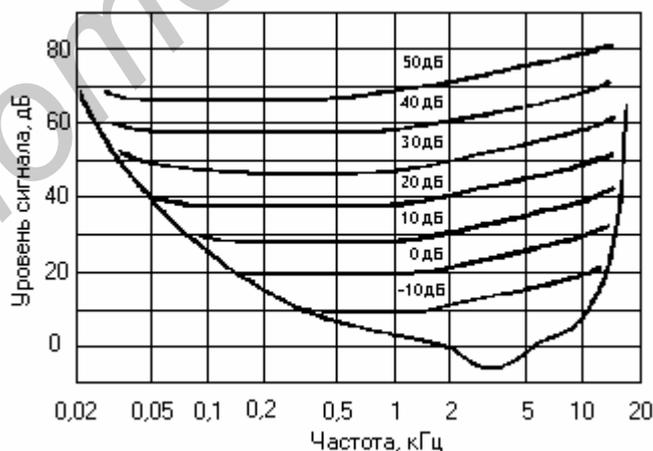


Рис. 7. Кривые порога слышимости при маскировке белым шумом

На рис.7 приведены кривые порогов слышимости синусоидальных звуков, полученные при маскировке белым шумом с различной спектральной плотностью. Из рис.7 видно, что на низких частотах кривые практически не зависят от частоты (примерно до 500 Гц). При дальнейшем увеличении частоты уровень порога слышимости повышается на 3 дБ при каждом удвоении частоты. Это примерно соответствует зависимости расширения ширины критических полос с увеличением центральной частоты (2). Из этого следует, что белый шум неодинаково эффективен для маскировки разных частот. Свойство широкополосных сигналов оказывать максимальное влияние на маскировку сигнала в

пределах критических полос положено в основу современных психоакустических алгоритмов сжатия сигналов.

Эффект маскирования в частотной области связан с тем, что в присутствии больших звуковых амплитуд человеческое ухо нечувствительно к малым амплитудам близких частот (рис.8).

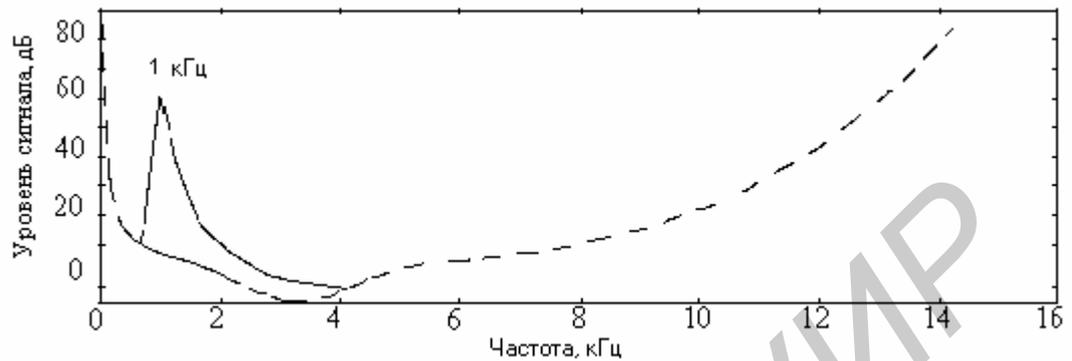


Рис.8. Изменение порога чувствительности слуховой системы при маскировании тональным сигналом определенной частоты

Таким образом, для маскирования важнее не абсолютная громкость звуков, а отношение мощности громкого сигнала к мощности тихого. Кроме того, чем ближе частота к частоте маскирующего сигнала, тем сильнее эффект маскирования (см. рис.8). Степень маскировки – это разность в децибелах между уровнем порога слышимости маскируемого тона в присутствии маскера и его уровнем порога слышимости в тишине. На рис.8 показаны зависимость абсолютного порога чувствительности уха и характер его изменения в присутствии тона с частотой 1 кГц и амплитудой в 60 дБ. Из рис.8 видно, что порог слышимости имеет адаптивный характер: он повышается при появлении каких-либо звуков и понижается в тишине. Для того, чтобы услышать более слабые звуки на фоне более сильных, их частоты должны значительно различаться.

Из рис.9 видно, что с увеличением средней частоты маскирующего сигнала диапазон частот, где проявляется маскировка, становится шире. Кривые маскировки являются несимметричными, они имеют крутой спад в сторону низких частот и пологий спад в сторону высоких частот.

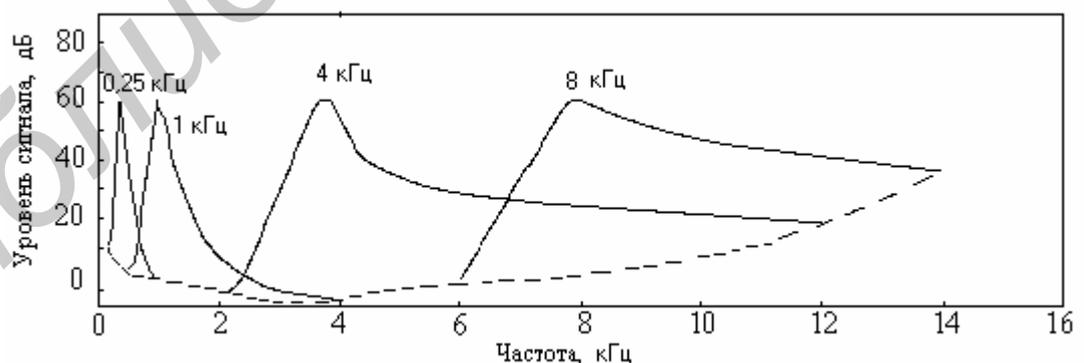


Рис.9. Изменение порога чувствительности слуховой системы при маскировании сигналами со средними частотами полос 0,25, 1, 4 и 8 кГц и уровнем интенсивности 60 дБ

Если при построении кривых маскировки по оси абсцисс откладывать не частоты f , а значения высоты тона z в барках, то кривые маскировки при разных значениях z маскирующего сигнала будут одинаковыми при одном и том же значении уровня маскирующего сигнала (рис.10). Их форма при таком представлении будет зависеть только от маскирующего сигнала, и не будет зависеть от величины z .

Маскирование имеет место не только внутри критической полосы, оно распространяется и на соседние полосы. Маскер, центрированный в пределах критической полосы, оказывает определенное воздействие на порог обнаружения в соседних критических полосах. Функция протяженности маскирования, которая учитывает маскирование как внутри, так и между полосами, задается аналитическим выражением

$$SF(x) = 15,81 + 7,5(x + 0,474) - 17,5\sqrt{1 + (x + 0,474)^2}, \quad (5)$$

где x и $SF(x)$ представлены соответственно в барках и дБ.

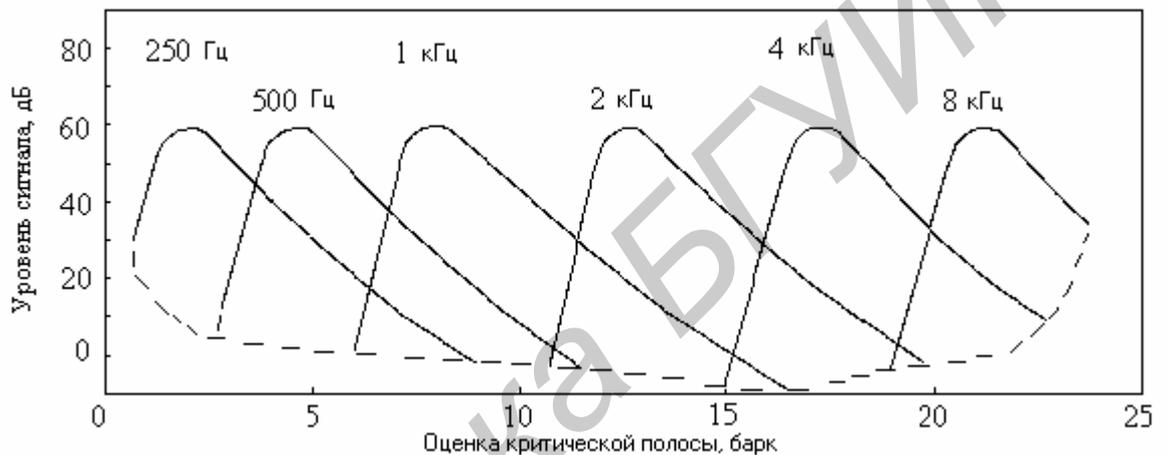


Рис.10. Частотное маскирование с частотной шкалой в барках

Вследствие разделения спектра звукового сигнала на частотные полосы слух реагирует не на общую мощность сигнала, а на мощность, сосредоточенную в отдельных частотных группах. При этом более интенсивные частотные группы при определенных условиях могут маскировать менее интенсивные группы. Применительно к сжатию это означает, что в группах меньшей интенсивности допустим больший уровень шума квантования, т.е. допустимо уменьшение числа разрядов при кодировании соответствующих спектральных составляющих, что в свою очередь снижает скорость передаваемого цифрового потока.

Маскирование во временной области (рис.11) характеризует динамические свойства слуха, показывая изменение во времени относительного порога слышимости (порог слышимости одного сигнала в присутствии другого), когда маскирующий и маскируемый сигналы звучат не одновременно.

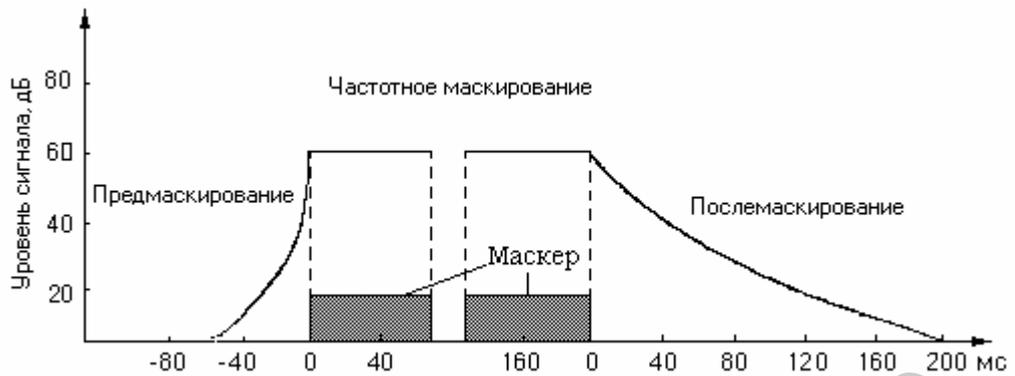


Рис. 11. Зависимость изменения порога слышимости маскируемого сигнала от величины временного интервала между сигналом и маскером

Различают предмаскировку (маскируемый сигнал перед маскером) и послемаскировку (маскируемый сигнал после маскера). Оба вида маскировки наступают, если маскируемый сигнал и маскер одновременно находятся во временном окне шириной 100–200 мс. Считается, что за это время завершается полная обработка сигнала с выделением всех его признаков. В этом случае слуховую систему можно рассматривать как процессор для обработки звуковых сигналов во временном окне шириной примерно 200 мс.

Из рис.11 видно, что длительность типичных промежутков времени, в пределах которых действуют предмаскирование и послемаскирование, составляет соответственно 5 – 50 мс и 50 – 200 мс. Предмаскировка проявляется на значительно более коротком временном интервале, причем ее длительность в очень сильной степени зависит от опыта слушателя. Степень маскирования звукового сигнала зависит от интервала между маскируемым сигналом и маскером, уровня интенсивности и длительности воздействия маскера. Сближение во времени подачи сигнала и маскера увеличивает маскировку. Временное маскирование зависит от частотного взаимоотношения сигнала и маскера точно так же, как и при частотном маскировании. Оно проявится в большей степени, если сигнал и маскер близки по частоте.

2.3. Психоакустическая модель слуховой системы

Алгоритм работы психоакустической модели слуховой системы предназначен для вычисления глобальных маскирующих порогов и отношения сигнал/маска и состоит из следующих этапов.

1. Спектральный анализ и нормализация уровня звукового давления сигнала. Выборки (отсчеты) входного аудиосигнала $s(n)$ нормализуются по формуле

$$x(n) = \frac{s(n)}{N(2^{b-1})}, \quad (6)$$

где N – количество отсчетов или длина окна анализа сигнала; b – количество разрядов или ступеней квантования на выборку; n – индекс выборки входного сигнала.

Нормализованный сигнал $x(n)$, используя оконную функцию Ханна с 1/16 долей перекрытия, разбивается на блоки, содержащие 512 выборок входного сигнала.

Оценка спектральной плотности мощности $P(k)$ в дБ с использованием 512-точечного быстрого преобразования Фурье производится по формуле

$$P(k) = PN + 10 \log_{10} \left| \sum_{n=0}^{N-1} w(n)x(n)e^{-j \frac{2\pi kn}{N}} \right|^2 \quad \text{при } 0 \leq k \leq \frac{N}{2}, \quad (7)$$

где PN – уровень нормализации мощности сигнала, равный 90 дБ; k – индекс спектральной составляющей интервала разрешения (бина);

$w(n) = \frac{1}{2} \left[1 - \cos\left(\frac{2\pi n}{N}\right) \right]$ – оконная функция Ханна.

Поскольку уровни воспроизведения сигнала неизвестны в момент проведения психоакустического анализа сигнала, процедура нормализации (6) и параметр PN в (7) используются для оценки уровня звукового давления независимо от входного сигнала.

2. Идентификация тональных и шумовых маскеров. Локальные максимумы в спектральной плотности мощности $P(k)$, которые превышают соседние компоненты в пределах одного барка по крайней мере на 7 дБ, идентифицируются как тональные. Множество тональных спектральных максимумов S_T определяется следующим выражением:

$$S_T = \left\{ P(k) \left| \begin{array}{l} P(k) > P(k \pm 1), \\ P(k) > P(k \pm \Delta_k) + 7dB \end{array} \right. \right\}, \quad (8)$$

где $P(k)$ – k -й локальный спектральный максимум;

$$\Delta_k \in \begin{cases} 2 & \text{при } 2 < k < 63 \quad (0.17 - 5.5 \text{ кГц}), \\ [2,3] & \text{при } 63 \leq k < 127 \quad (5.5 - 11 \text{ кГц}), \\ [2,6] & \text{при } 127 \leq k \leq 256 \quad (11 - 20 \text{ кГц}). \end{cases}$$

Значения тональных маскеров $P_{TM}(k)$ в дБ вычисляются из спектральных максимумов множества (8) по формуле

$$P_{TM}(k) = 10 \log_{10} \sum_{j=-1}^1 10^{0,1P(k+j)}, \quad (9)$$

где $P(k-1)$ и $P(k+1)$ – значения спектральной плотности мощности частотных составляющих, соседних с $P(k)$.

После идентификации тональных маскеров значение спектральной плотности мощности $P_{NM}(\bar{k})$ в дБ единственного шумового маскера для каждой критической полосы вычисляется из оставшихся спектральных составляющих, находящихся за пределами диапазона тонального маскера $\pm \Delta_k$, по формуле

$$P_{NM}(\bar{k}) = 10 \log_{10} \sum_j 10^{0,1P(j)} \quad (10)$$

для всех $P(j) \notin \{P_{TM}(k, k \pm 1, k \pm \Delta_k)\}$, где $\bar{k} = \left(\prod_{j=1}^u j \right)^{1/(l-u+1)}$ – средняя геометрическая

спектральная составляющая критической полосы; l и u – нижняя и верхняя границы спектральных составляющих критической полосы соответственно.

3. Прореживание и преобразование маскеро́в. На этом шаге количество маскеро́в сокращается на основе двух критериев: абсолютного порога слышимости и частотного прореживания. Тональные и шумовые маскеро́в, расположенные ниже абсолютного порога слышимости, отбрасываются, а остаются лишь маскеро́в, удовлетворяющие условию

$$P_{TM, NM}(k) \geq T_q(k), \quad (11)$$

где $T_q(k)$ – уровень звукового давления абсолютного порога слышимости в тишине для k -й спектральной составляющей.

Процедура скользящего окна шириной 0,5 барка используется для замены любой пары маскеро́в, появляющихся в пределах данного окна, одним маскеро́м с наибольшей мощностью. После выполнения данной процедуры частотные бины (интервалы разрешения) маскеро́в преобразуются согласно схеме субдискретизации

$$P_{TM, NM}(i) = P_{TM, NM}(k), \quad (12)$$

$$P_{TM, NM}(k) = 0, \quad (13)$$

где $i = \begin{cases} k & \text{при } 1 \leq k \leq 48, \\ k + (k \bmod 2) & \text{при } 49 \leq k \leq 96, \\ k + 3 - ((k - 1) \bmod 4) & \text{при } 97 \leq k \leq 232. \end{cases} \quad (14)$

Согласно (14), эффект прореживания бинов маскеро́в в критических полосах 18–22 слуховой системы составляет 2:1, а для критических полос 22–25 он равен 4:1 без потерь маскирующих компонент. Данная процедура уменьшает общее число частотных бинов тональных и шумовых маскеро́в от 256 до 106. Результатом выполнения данного этапа является прореженное множество тональных и шумовых маскеро́в.

Данный этап удаляет спектральные компоненты, которые присутствуют обычно ниже порога слышимости в тишине, и более слабые тональные компоненты, находящиеся в пределах половины ширины критической полосы (0,5 барка) сильного тонального компонента. Данная модель использует прореженные тональные и нетональные (шумовые) компоненты для определения глобального порога маскирования в частотной области.

4. Вычисление индивидуальных порогов маскирования. Каждый индивидуальный порог характеризует маскирующий вклад в i -й частотный бин тонального или шумового маскера, расположенного в j -м бине, преобразованном на этапе 3. Индивидуальные пороги тонального маскера $T_{TM}(i, j)$ вычисляются по формуле

$$T_{TM}(i, j) = P_{TM}(j) - 0,275z(j) + SF(i, j) - 6,025, \quad (15)$$

где $P_{TM}(j)$ – уровень звукового давления тонального маскера в дБ в j -м частотном бине; $z(j)$ – частота j -го бина в барках; $SF(i, j)$ – функция распространения маскирования от j -го бина маскера к i -му бину маскируемого сигнала.

Функция распространения маскирования $SF(i, j)$ в дБ от j -го бина маскера к i -му бину маскируемого сигнала моделируется в виде кусочно-линейной функции уровня маскера $P(j)$ по формуле

$$SF(i, j) = \begin{cases} 17\Delta_z - 0,4P_{TM}(j) + 11 & \text{при } -3 \leq \Delta_z < -1, \\ (0,4P_{TM}(j) + 6)\Delta_z & \text{при } -1 \leq \Delta_z < 0, \\ -17\Delta_z & \text{при } 0 \leq \Delta_z < 1, \\ (0,15P(j)_{TM} - 17)\Delta_z - 0,15P(j)_{TM} & \text{при } 1 \leq \Delta_z < 8, \end{cases} \quad (16)$$

где $\Delta_z = z(i) - z(j)$ – расстояние в барках для разделения маскера и маскируемого сигнала. В данной модели для увеличения вычислительной эффективности протяженность маскирования ограничивается окрестностью шириной до 10 барк.

Индивидуальные пороги шумового маскера $T_{NM}(i, j)$ вычисляются по формуле

$$T_{NM}(i, j) = P_{NM}(j) - 0,175z(j) + SF(i, j) - 2,025, \quad (17)$$

где $P_{NM}(j)$ – уровень звукового давления шумового маскера в дБ в j -м частотном бине; $z(j)$ – частота j -го бина в барках. Функция $SF(i, j)$ вычисляется путем замены $P_{TM}(j)$ на $P_{NM}(j)$ в (16).

5. Вычисление глобальных порогов маскирования. Индивидуальные пороги маскирования объединяются, чтобы оценить глобальный порог маскирования для каждого частотного бина в подмножестве, заданном (14). Данная модель основана на свойстве аддитивности маскирующих эффектов. Глобальный порог маскирования $T_g(i)$ вычисляется следующим образом:

$$T_g(i) = 10 \log_{10} \left(10^{0,1T_q(i)} + \sum_{l=1}^L 10^{0,1T_{TM}(i,l)} + \sum_{m=1}^M 10^{0,1T_{NM}(i,m)} \right), \quad (18)$$

где $T_q(i)$ – абсолютный порог слышимости для i -го частотного бина; $T_{TM}(i, l)$ и $T_{NM}(i, l)$ – индивидуальные пороги маскирования, вычисленные на этапе 4; L и M – количество соответственно тональных и шумовых маскеров, идентифицируемых на этапе 3.

Таким образом, глобальный маскирующий порог для каждого частотного бина представляет собой сигнал-зависимую, аддитивную по мощности модификацию абсолютного порога слышимости в тишине. В этом случае тональная кривая маскирования, нетональная (шумовая) кривая маскирования и кривая абсолютного порога складываются для получения общей кривой маскирования анализируемого входного сигнала.

Маскирующий эффект в каждой критической полосе слуховой системы характеризуется величиной отношения сигнал/маска. В данной модели отношение сигнал/маска задается величиной отношения энергии сигнала в пределах критической полосы к энергии шумов квантования, при которой данный шум еще маскируется полезным сигналом и не замечается слуховым аппаратом человека.

3. ЛАБОРАТОРНОЕ ЗАДАНИЕ

3.1. Изучите теоретическую часть.

3.2. Алгоритм, реализующий психоакустическую модель, выполнен на входном языке инженерного пакета MATLAB в виде набора m-файлов. Головным m-файлом является Test_MPEG.m. Остальные m-файлы, реализующие отдельные функции алгоритма, вызываются в тексте данного головного файла. Работа алгоритма заключается в обработке

сигнала в перекрывающихся окнах длиной 512 отсчетов каждое. Каждое последующее окно смещено относительно предыдущего на 384 отсчета. Запустить MATLAB можно двойным нажатием левой кнопки "мыши" на его пиктограмму. Для запуска головного m-файла необходимо в командной строке среды MATLAB написать его имя Test_MPEG и нажать <Enter>. Запуск этого файла приводит к появлению окошка "Fig. No. 1", в котором появляется амплитудный спектр тестового сигнала (первое окно длиной 512 отсчетов). Тестовый сигнал представляет собой смесь восьми синусоид с различными частотами и амплитудами и гауссовского шума. Далее по нажатию любой клавиши отображаются вычисленные доминирующие тоновые и шумовые компоненты, кривые маскирования для каждой из компонент, глобальный маскирующий порог. Процедура повторяется для следующего окна в 512 отсчетов и сопровождается аналогичным выводом уже в окошко "Fig. No. 2».

3.3. Руководствуясь инструкциями в рабочем окне программы Test_MPEG, реализованной в среде MATLAB, выполните следующие задания:

3.3.1. Рассчитайте и постройте зависимость минимального порога слышимости слухового аппарата человека в тишине от частоты.

3.3.2. Рассчитайте и постройте зависимость ширины критической полосы от ее центральной частоты.

3.3.3. Рассчитайте и постройте зависимость между шкалой частот в Гц и барк-шкалой (шкалой высот тона). Оцените изменение точности идентификации тональных и шумовых маскеров при переходе от частотной шкалы к барк-шкале.

3.3.4. Рассчитайте, постройте и сравните графики нормированной спектральной плотности мощности звукового сигнала, индивидуальных порогов тональных и шумовых маскеров, глобальных порогов маскирования для различных подмножеств спектральных бинов, полученных с использованием временных оконных функций с различной шириной и формой.

3.4. Оформите отчет и сделайте выводы.

4. СОДЕРЖАНИЕ ОТЧЕТА

4.1 Решение задач задания.

4.2. Результаты выполнения работы.

4.3. Анализ результатов и выводы.

5. КОНТРОЛЬНЫЕ ВОПРОСЫ

5.1. Перечислите основные психоакустические особенности слуховой системы, используемые для повышения эффективности системы сжатия.

5.2. Что представляет собой кривая минимального порога слышимости слухового аппарата человека в тишине? Как изменяется кривая порога слышимости при маскировке тональными и широкополосными сигналами?

5.3. Что такое критическая полоса слуха? Как ее ширина связана с частотой сигнала? Какими методами можно измерить ширину критической полосы слуха?

5.4. Что означает единица барк? Почему идентификация тональных и шумовых маскеров, вычисление индивидуальных и глобальных порогов маскирования производятся в барк-шкале?

5.5. В чем состоит явление маскировки? Перечислите особенности маскировки внутри и вне критической полосы слуха.

5.6. В чем проявляется нелинейность слуха? Почему маскировка сигналов, имеющих частоту, лежащую ниже частоты маскирующего тона, проявляется значительно слабее, чем для вышележащих частот?

5.7. При каких условиях наблюдаются эффекты частотной и временной маскировки? От чего зависит степень слухового маскирования?

5.8. Что такое индивидуальный порог маскирования? Для каких типов маскеров он рассчитывается? От чего он зависит? Чем моделируется эффект распространения маскирования?

5.9. Каким образом можно повысить степень сжатия и стеганографической защиты данных с использованием психоакустической модели?

ЛИТЕРАТУРА

1. Слуховая система: Сб. статей / Ред. Я.А. Альтман. – Л.: Наука, 1990. – 620 с.
2. Секунов Н.Ю. Обработка звука на РС. – СПб.: БХВ-Петербург, 2001. – 1248 с.
3. Цвиккер Э., Фельдкеллер Р. Ухо как приемник информации. - М.: Связь, 1971. – 255 с.
4. Загуменков А.П. Компьютерная обработка звука. - М.: ДМК Лайт, 1999. – 382 с.
5. Радзишевский А. Компьютерная обработка звука. – М.: Нолидж, 2000. – 240 с.
6. Смирнов Д.В. Аудиосистема РС. – СПб.: 1999. – 384 с.
7. Гельфанд С.А. Слух: Введение в психологическую и физиологическую акустику. – М.: Медицина, 1984. – 350 с.
8. Дьяконов В.П. Matlab. – СПб.: 2001. – 553 с.