электронный журнал
МОЛОДЕЖНЫЙ НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК
Издатель ФГБОУ ВПО "МГТУ им. Н.Э. Баумана". Эл No. ФС77-51038.

УДК 534.756.2

Эволюция развития модели механизма слухового восприятия в частотной области (критические полосы слуха)

***Сальников А.М.**, студент*
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана,
кафедра «Защита информации»

Научный руководитель: Козлачков С.Б., к.т.н, доцент
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана,
кафедра «Защита информации»
runc@bmstu.ru

Одной из групп признаков, по которым производят распознавание речи, является с овокупность амплитудно-частотных признаков. К ним, в частности, относятся: частота ос новного тона (ЧОТ), формантные частоты и т.д. Все эти признаки появились на основани и моделей восприятия звука в частотной области, связанных с объяснением механизмов р аботы периферийной слуховой системы.

Одной из первых была выдвинута теория Г. Гельмгольца (нем. Hermann Ludwig Fer dinand von Helmholtz) в 1863-ем году[1], заключавшаяся в том, что базилярная мембрана с остоит из волокон разной длины и натянутости, которые резонируют на различных частот ах. Этим же ученым была предложена модель представления периферийной слуховой сис темы как линейки фильтров, а в дальнейшем такое же предположение выдвинул Х. Флет чер (англ. Harvey Fletcher**)** в 1940-ом году [2] после выполнения ряда экспериментов, суть которых будет изложена ниже, и ввел новый термин – критическая полоса слуха. Под кри тической полосой слуха понимается минимальная полоса частот, которая возбуждает одн у и ту же часть базилярной мембраны.

С момента начала изучения механизма слухового восприятия и по настоящее время объектом исследований является модель критических полос слуха. Рассмотрение этапов э волюции данной модели в исторической перспективе важно для понимания процессов, пр оисходящих в периферийной слуховой системе.

Х. Флетчер измерял пороги обнаружения тональных сигналов как функцию от шир ины полосы маскирующего сигнала. В качестве маскирующего сигнала использовался узк ополосный белый шум. При увеличении полосы маскирующего сигнала порог обнаружен

ия сначала повышается, но затем принимает постоянное значение. Графическое представл ение полученной зависимости представлено на рис. 1. [3]

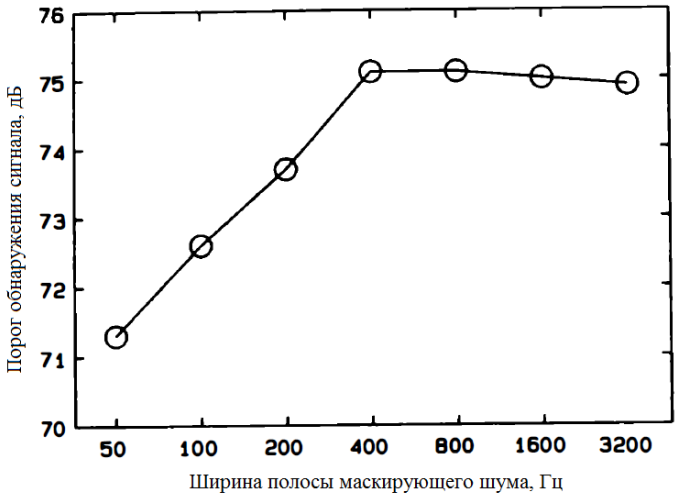


Рис. 1. Характер зависимости, полученный Х. Флетчером в 1940-ом году, частота тонального сигн ала – 2000 Гц

Этот эксперимент положил начало большому числу других по измерению ширины критической полосы слуха как функции от частоты. Так, Э. Цвикер (нем. Eberhard Zwicker) в 1961-ом [4] году представил результаты своих экспериментов, которые стали общепри нятыми.

Однако исследования в этой области после публикации Э. Цвикера продолжаются до сих пор. Необходимость в уточнении полученных данных обуславливается следующей причиной: ограничениями технологий того времени, которые не позволяли получать дост оверные и точные результаты в области низких частот. Одним из существенных ограниче ний являлась невозможность создания узкополосных шумовых сигналов, необходимых дл я проведения экспериментов.

Потребность в создании таких сигналов обусловлена проблемой, возникавшей при проведении экспериментов, а именно явления внечастотного прослушивания (подобное яв ление в области обработки сигналов называется просачиванием). Оно заключается в том, что периферийная слуховая система перестраивает фильтр на частоту, близкую к частоте с игнала, тем самым незначительно ослабляя воздействие тонального сигнала и значительно – шумового, тем самым значительно искажая измеряемое соотношение сигнал/шум (SNR).

Для уменьшения влияния этого явления при проведении экспериментов, Р. Паттерс он (англ. Roy D. Patterson) в 1976-ом году предложил метод, который называется в англояз ычной литературе «Notched-Noise Method» [5]. Данный метод (см. рис. 2) заключается в сл

едующем: тональный сигнал располагается на частотной сетке в «выемке» в шумовом сиг нале, причем ширина этой выемки переменна. Тем самым, появляется возможность измер ения ширины КПС как функции частоты с предотвращением внечастотного прослушиван ия.

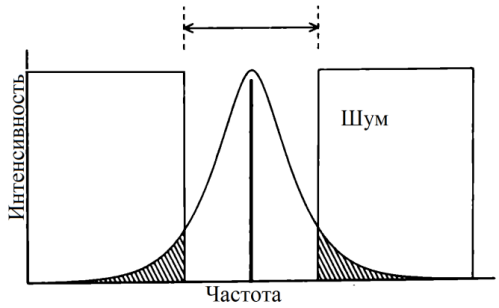


Рис. 2. Notched-Noise Method (Из-за наличия «выемки» в полосе шума данный метод называется зу бчатым – notched, англ.)

Перед тем, как представить результаты, полученные с помощью этого метода, необ ходимо отметить, что ученые Б. Глазберг и Б. Мур (англ. Brian R. Glasberg, Brian C J Moor e), проводившие эти эксперименты, иначе понимали физический смысл термина «КПС». К ак правило ширина полосы фильтра, в нашем случае слухового фильтра, определяется как разность верхней и нижней граничных частот участка АЧХ, на котором мощность равна 0, 5 от максимальной, что соответствует уровню -3 дБ.

Однако существует и другой оценочный метод, называемый эквивалентной прямоугольно й полосой пропускания, который в англоязычной литературе обозначается аббревиатурой ERB (Equivalent Rectangular Bandwidth). Метод заключается в следующем: для любого фи льтра существует такой фильтр с идеальной прямоугольной АЧХ, который пропускает бел ый шум той же мощности, что и рассматриваемый фильтр. Под ERB понимается ширина п олосы частот такого идеального фильтра.

Б. Глазберг и Б. Мур в 1990-ом году [6] провели свои исследования, по результатам которых были уточнены виды зависимостей КПС (см. рис. 3) и выведена формула (1), являющаяся функцией ширины критической полосы от частоты, которая называется ERB. Данной формулой (1) пользуются по настоящее время, поскольку она хорошо согласуется с характерами зависимостей, полученными в результате последующих экспериментов.

$$\mathrm{ERB} = 24.7(4.37\,\mathrm{F} + 1), \qquad (1)$$

где F – частота (кГц).

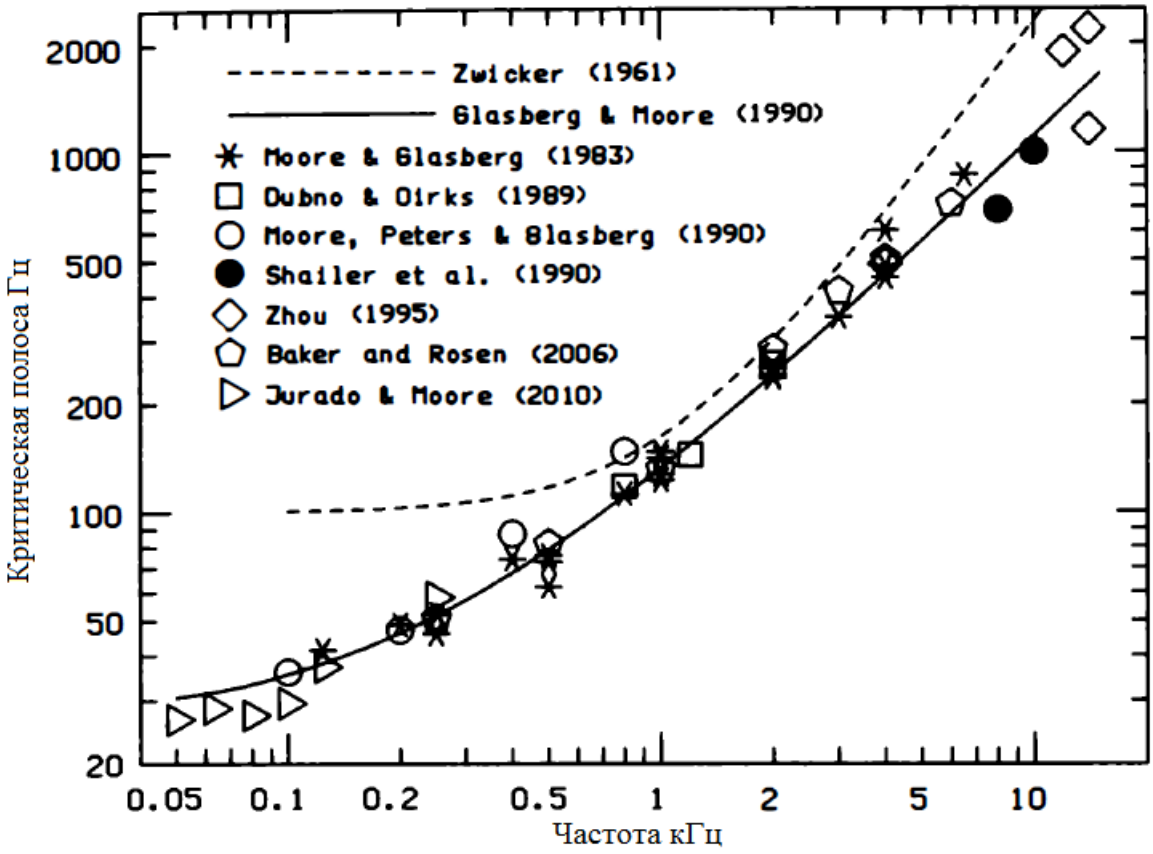


Рис. 3. Вид зависимостей ширины КПС от частоты, полученный в результате проведения эксперим ентов Э. Цвикером в 1961-ом году, Б. Глазбергом и Б. Муром в 1990-ом году и др.

В 2000-ом Б. Глазберг и Б. Мур [7] провели новый эксперимент с целью уточнения вида АЧХ слухового фильтра в зависимости от уровня сигнала, попадающего в полосу фи льтра. Вид полученной зависимости представлен на рис.4.

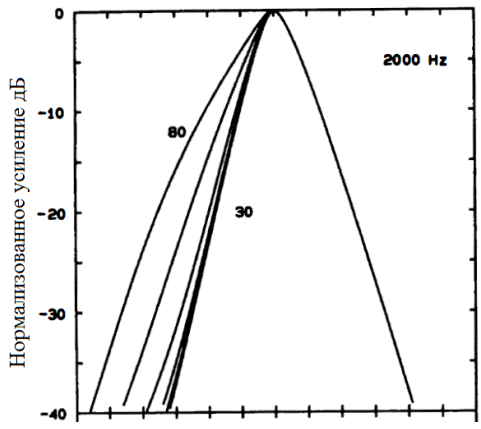


Рис. 4. Вид зависимости АЧХ слухового фильтра от уровня сигнала

Анализ зависимости указывает на то, что при уменьшении интенсивности сигнала АЧХ слухового фильтра сужается. Это согласуется с ранее выдвинутой ими моделью, нос ящей название «Энергетическая модель спектра» (Power Spectrum Model). Она сводится к тому, что АЧХ фильтра «формируется» вокруг возмущения, причем параметры этого фил ьтра изменяются для улучшения восприятия таким образом, чтобы получить наилучшее с оотношение SNR.

Ранее модель периферийной слуховой системой представляли как банк фильтров с фиксированными параметрами, в т. ч. центральными частотами, однако до сих пор не пол учено каких-либо экспериментальных подтверждений физического существования таких фильтров с их стационарным положением на базилярной мембране. Напротив, современн ыми исследованиями подтверждается, что в процессе приема сигналов слуховая система а даптируется под каждый для наилучшего восприятия.

Поскольку адаптивная фильтрация позволят повысить SNR, она снижает влияние п омех при анализе сигнала в смеси с шумом. Таким образом адаптивная фильтрация парам етров КПС реализует один из механизмов повышения помехоустойчивости механизма слу хового восприятия.

Вывод

Приведенные выше исследования и модели показали, что при анализе речевого сиг нала следует использовать не строго определенные частотные полосы, а определять их ис ходя из каждого конкретного случая по локальным максимумам спектра сигнала.

Дальнейшие исследования нацелены на уточнение зависимости ширины КПС от ч астоты, а также на определение функции, описывающей вид АЧХ слухового фильтра в зав исимости от уровня сигнала [8]. Также необходимо определить временные характеристик и механизма адаптивной фильтрации, в частности время «формирования» АЧХ и изменен ия параметров слуховых фильтров согласно «Энергетической модели спектра».

Перспективным направлением исследования является программная реализация мех анизма адаптивной фильтрации, базирующаяся на вышеизложенных принципах, а также п роверка ее применимости в процессе распознавания речи.

Список литературы

1. Helmholtz H. L. F. Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik. Braunschweig: Braunschweig, 1913. 696 p.
2. Fletcher H. Auditory patterns // Rev. Mod. Phys. 1940. № 12. P. 47–65.
3. Moore B.C.J. An Introduction to the Psychology of Hearing. Bingley: Emerald Group

Publishing Limited, 2012. 457 p.

4. Zwicker E. Subdivision of the audible frequency range into critical bands // The Journal of the Acoustical Society of America. 1961. № 33. P. 248.
5. Patterson R.D. Auditory filter shapes derived with noise stimuli // The Journal of the Acoustical Society of America. 1976. № 59. P. 640–654.
6. Glasberg B.R., Moore B.C.J. Derivation of auditory filter shapes from notched-noise data // Hearing research. 1990. № 47. P. 103–138.
7. Glasberg B.R., Moore B.C.J. Frequency selectivity as a function of level and frequency measured with uniformly exciting notched noise // The Journal of the Acoustical Society of America. 2000. № 108. P 2318–2328.
8. Unoki M., Irino T., Glasberg B.R., Moore B.C.J., Patterson R.D. Comparison of the roex and gammachirp filters as representations of the auditory filter // The Journal of the Acoustical Society of America. 2006. № 120. P. 1474–1492.