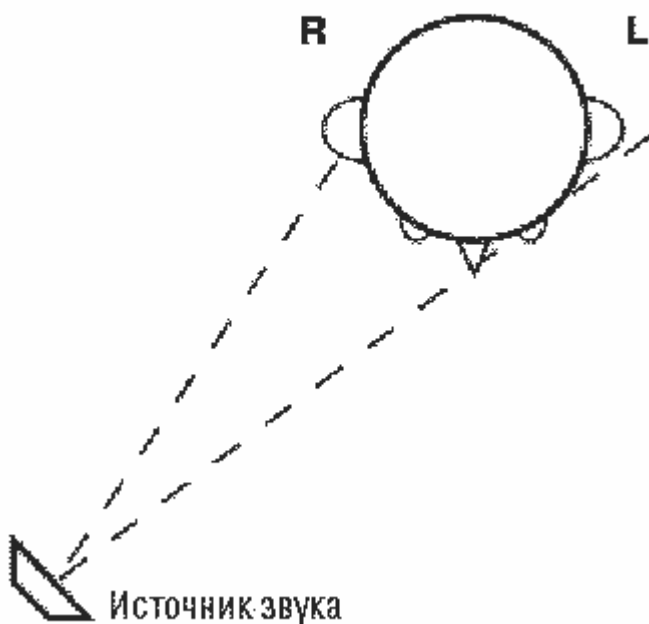


Рис 1а Расстояние от звукового источника до правого и левого ушей



: архив : архив журнала  
"Звукорежиссер" : 1999 : №10

## Основы психоакустики часть 4 Бинауральный слух и пространственная локализация

*Ирина Алдошина*

Наличие двух приемников слуха обеспечивает человеку возможность воспринимать пространственный звуковой мир и оценивать перемещение звуковых сигналов в пространстве. Информация, которая поступает на оба слуховых канала, обрабатывается в периферической части слуховой системы (подвергается спектрально-временному анализу) и затем передается в высшие отделы головного мозга, где путем сравнения этой информации из двух разных каналов

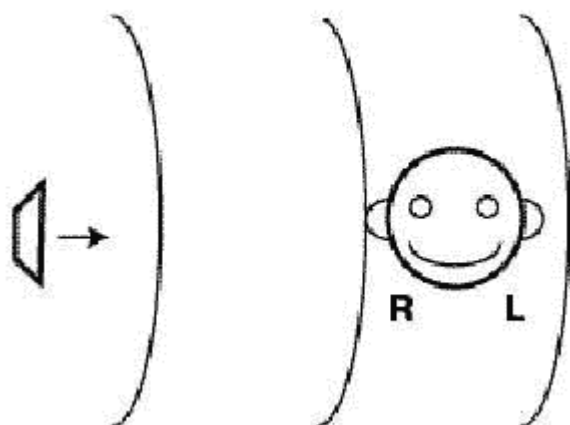
формируется единый пространственный слуховой образ.

Восприятие через два приемника информации, иначе называемое бинауральным слухом, дает человеку огромные преимущества, основные из которых следующие:

- локализация сигналов как от одиночных, так и от множественных источников, что позволяет формировать пространственную перспективу и оценивать пространственное звуковое поле (например, в помещении).
- разделение сигналов, приходящих от различных звуковых источников из различных точек пространства.
- выделение сигналов выбранного звукового источника на фоне других звуковых сигналов, например выделение прямого звука на фоне реверберирующих сигналов в помещении, выделение речи на фоне шумов и т.д.

Анализ бинауральных слуховых эффектов представляет особый научный интерес, в частности для изучения функционирования и спецификации полушарий головного мозга, а также громадный практический интерес в связи с развитием и промышленным внедрением бинауральных технологий для создания систем пространственной звукозаписи и звуковоспроизведения (стереофонические системы, пространственные системы типа Dolby Digital и др.), для синтеза трехмерных виртуальных звуковых полей (технология 3D-Sound, техника аурализации, создание адаптивных процессоров и др.), для развития новых методов метрологии и оценки звуковой аппаратуры.

Рис 1б Огибание головы на низких частотах



локализацию, эффект предшествования, бинауральное суммирование громкости, бинауральную демаскировку, бинауральные биения и слияние звуков при определении высоты, эффекты "правого" и "левого" уха при восприятии речи и музыки и др.

Начнем рассмотрение этих свойств с пространственной локализации.

### Бинауральная пространственная локализация

Прослушивая звучание симфонического оркестра в концертном зале (или пение хора в большом соборе), слушатель отчетливо воспринимает и разделяет расположение инструментов в горизонтальной плоскости на сцене, их расположение по глубине, а также ощущает пространственность окружающего звукового образа. Эта способность и называется пространственной бинауральной локализацией. Причем механизмы локализации в горизонтальной, вертикальной плоскости и по глубине несколько различаются.

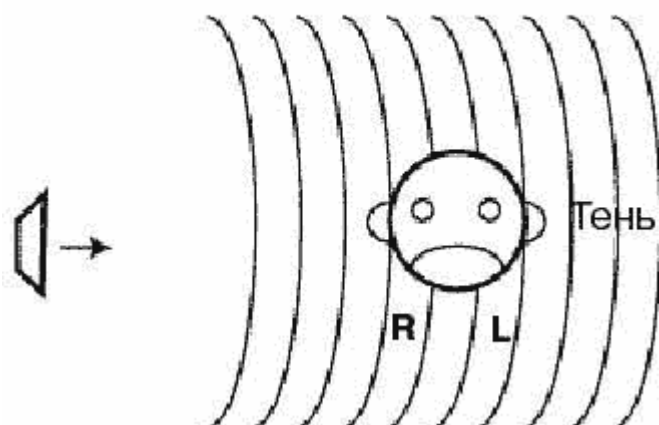
### Горизонтальная (азимутальная) локализация

На рисунке 1а представлены различительные признаки направленности при прослушивании источника звука (например, громкоговорителя при его различных положениях относительно головы слушателя). Звук, исходящий из громкоговорителя, расположенного справа от слушателя, должен пройти большее расстояние к левому уху, чем к правому. Как показано на рисунке 1б, низкие звуковые частоты имеют длину волны больше, чем диаметр головы, поэтому они огибают голову, поступая в ухо, расположенное дальше (дифракция). Однако звуки высокой частоты (Рис. 1в) имеют длину волны меньше, чем диаметр головы, поэтому они "блокируются" на пути к левому уху. Эта "акустическая" тень головы уменьшает интенсивность звука, поступающего в ухо, расположенное дальше от источника звука.

Обеспечение пространственной панорамы, делимости и выделения сигналов на фоне других сигналов и шумов является важнейшей задачей звукорежиссера при записи и обработке звука, а поскольку это требует использования бинауральных свойств слуха, то анализ этих свойств и является целью данной статьи.

К числу основных свойств бинаурального слуха можно отнести: пространственную

Рис 1в Акустическая тень на высоких частотах

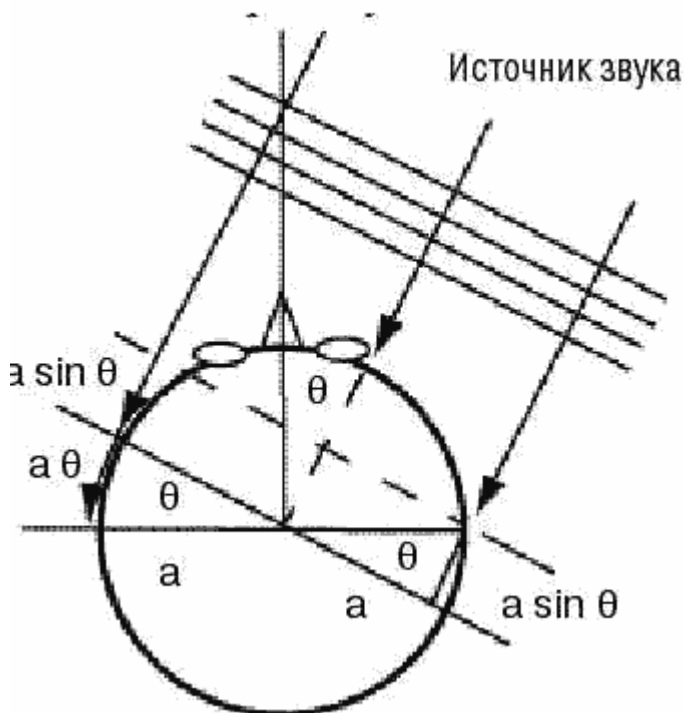


Частота, Гц	Длительность (-)периода, мс	град	Частота, Гц	Длительность (-)половины периода, мс	град
400	1,250	90	2000	0,250	24

800	0,625	90	2400	0,208	19
1200	0,417	42	3200	0,151	14
1600	0,313	30	4000	0,125	11

Пространственная разнесенность двух слуховых приемников (ушных раковин), и экранирующее влияние головы и торса за счет дифракционных эффектов приводит к значительным различиям между сигналами, поступающими в правое и левое ухо, что позволяет произвести локализацию звукового источника в пространстве, обусловленную тремя физическими факторами:

Рис. 2 Разность хода лучей до левого и правого ушей



а) временным (Interaural Time Difference - ITD) - возникающим из-за несовпадения по времени моментов прихода одинаковых фаз звука к левому и правому уху;

б) интенсивностным (Interaural Intensity Difference - IID) - возникающим из-за неодинаковой величины интенсивностей звуковой волны вследствие дифракции ее вокруг головы и образования "акустической тени" со стороны, обратной источнику звука, как показано на рисунке 1а;

в) спектральным - возникающим из-за разницы в спектральном составе звуков, воспринимаемых левым и правым ухом, вследствие неодинакового экранирующего влияния головы и ушных раковин на низкочастотные и высокочастотные составляющие сложного звука.

а) временная разность - ITD

Разность времени прихода одинаковых фаз звука к ушам (ITD) можно легко рассчитать, зная разность хода  $dx$  звуковой волны до левого и правого уха  $ITD=dx/C$ , где  $C$ -скорость распространения звуковой волны.

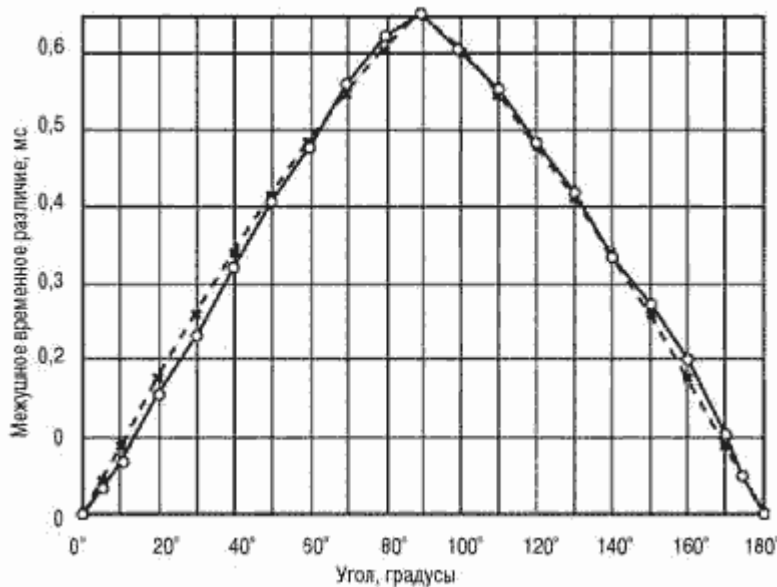
Смысл величины ITD можно понять из рисунка 2. Исследования зависимости между направлением локализации источника звука в горизонтальной плоскости, определяемым углом  $\theta$  и временем задержки ITD, приводят к следующему простому соотношению:

$$ITD=a/C (\theta+\sin \theta) \text{ при } -90^\circ < \theta < +90^\circ, (1)$$

где  $\theta$  - азимутальный угол, отсчитываемый в горизонтальной плоскости от плоскости симметрии головы (Рис. 2);  $a$  - радиус головы. Разность времени прихода одинаковых фаз звука к ушам (ITD) равна  $10^{-6}$  секунд при расположении звукового источника точно посередине и равна  $a/c(\theta/2+1)$  для расположения источника точно напротив одного уха, что составляет  $\sim 0,7$  мс (средний радиус головы  $\sim 9$  см, кратчайшее расстояние вокруг головы от одного уха до другого  $\sim 26$  см).

Различия по времени прихода звуковых волн для разных углов расположения источника для

Рис 3 Различия междушного времени при разных углах расположения источника звука



частоты 1500 Гц показаны на рисунке 3. Как видно из рисунка, при перемещении источника звука вокруг головы максимальная разница во времени возникает при  $\alpha = 90^\circ$ . На низких частотах эта временная разница увеличивается.

Для синусоидальных колебаний при частоте 800 Гц максимальное время запаздывания ИТД становится равным половине периода колебания  $T/2$ , а при более высоких частотах - превышает половину периода ( $ITD > T/2$ ). В этом случае возникает неясность в фазовых соотношениях колебаний, действующих на правое и левое

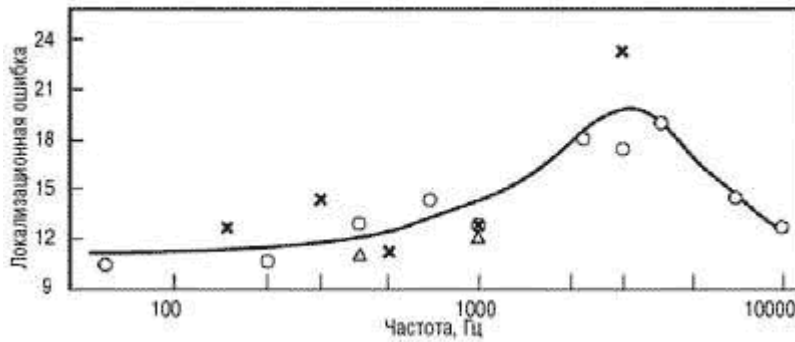
уши: с одинаковым основанием можно считать, что одна волна отстает по фазе от другой на время  $dT$  пачунд или опережает ее на это же время. Следовательно, предельное значение времени запаздывания, правильно воспринимаемое слухом, не должно превышать половину периода.

В соответствии с этим наибольшее значение азимутального угла  $\alpha_{max}$ , определяемое временным бинауральным эффектом, с повышением частоты уменьшается. Это иллюстрируется данными табл. 1, в которой приведены расчетные значения  $\alpha_{max}$ , вычисленные для разных частот по формуле (1) путем подстановки  $ITD = T/2$ . Например, при частоте 3200 Гц время запаздывания  $ITD = T/2$  создает ощущение углового перемещения всего лишь на  $14^\circ$ . Однако это обстоятельство не столь существенно, так как в этой области частот при изменении направления прихода звуковых волн уже достаточно сильно сказывается дифракция звука вокруг головы, то есть вступает в силу интенсивностный фактор.

#### б) интенсивностная разность - ИД

Как видно из рисунка 1а, по мере повышения частоты за счет дифракции образуется "акустическая тень" и интенсивность звуков, достигающих противоположного по отношению к источнику уха становится меньше. Наибольшая разность уровней звуковых давлений, действующих на левое и правое ухо, возникает при боковом положении источника ( $90^\circ$ ). Для этого случая на рисунке 4 приведен полученный экспериментально график частотной зависимости разности уровней звуковых давлений  $dN$  у левого и правого уха. Из графика видно, что по мере повышения частоты эта разность существенно возрастает, достигая на 5000 Гц величины  $\sim 20$  дБ.

Рис 5 Локализационная ошибка как функция частоты

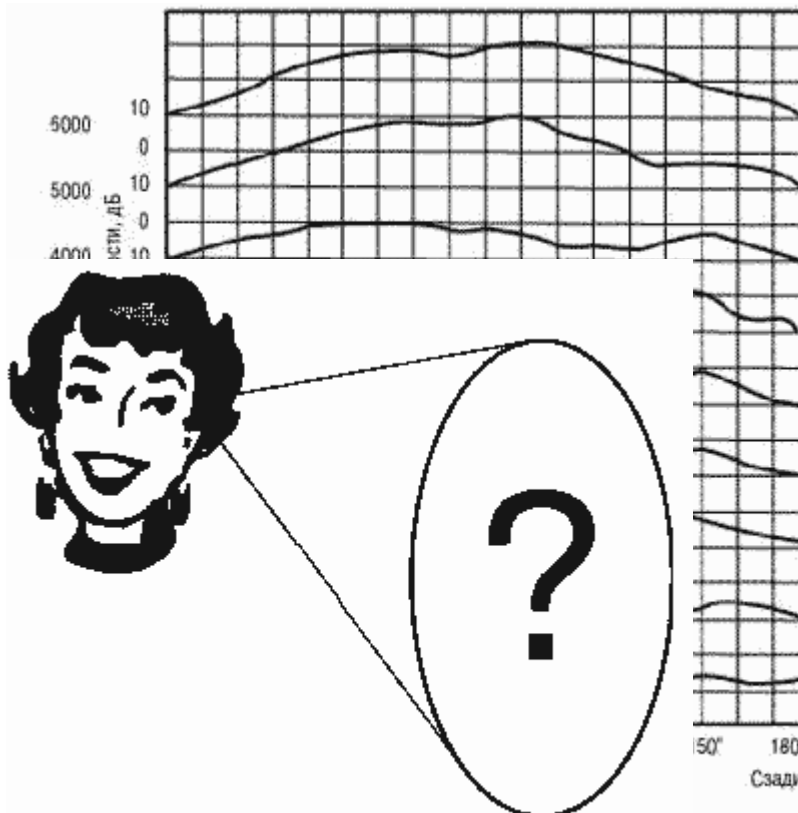


Последнее обстоятельство, однако, не означает, что при повышении частоты звука обостряется локализация. Напротив, чистые тоны очень высоких частот (свыше 8000 Гц) почти не поддаются локализации. Так же слабо выражена способность человека определять направление на источник синусоидальных звуков низкой частоты (ниже 300 Гц она становится значительно хуже, а ниже 150 Го отсутствует вообще), поэтому в современных системах "домашний театр" расположение низкочастотных блоков (subwoofer) может выбираться произвольно.

Исследования ошибок при локализации положения синусоидального источника показали (Рис. 5), что наибольшие ошибки человек совершает в области 2000-4000 Гц, где, по-видимому, происходит смена механизмов локализации от временного к интенсивностному.

Анализ способности к угловому различию двух источников, находящихся в горизонтальной плоскости, также подтвердил, что в области частот 1500-2000 Гц резко возрастает наименьшая различимая величина угла между источниками.

Интересно отметить, что минимальное различие в азимуте (угле) воспринимается, когда источники находятся перед испытуемым. В этом случае он достигает 2i. Наибольшее различие возникает, когда источники находятся справа или слева: возникает так называемый "конус неопределенности" с каждой



стороны уха (Рис. 6), внутри которого изменение положения источника звука не вызывает ощущение изменения его положения. Это объясняется тем, что при расположении источника сбоку получается большая разница и в интенсивности, и во времени, поэтому сдвиги источника дают малое относительное изменение общей разности. И поэтому для локализации очень важно движение головы - это изменяет положение конуса и сводит на нет его влияние.

с) спектральные различия

Наибольшая острота локализации достигается при восприятии сложных звуков и звуковых импульсов, когда, кроме рассмотренных ранее причин,

Рис 6 Конус неопределенности

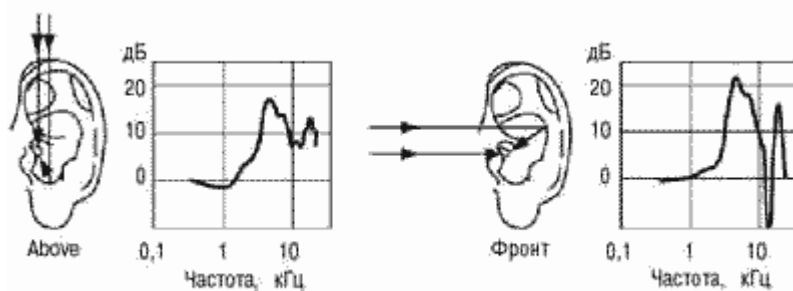


Рис 7 Форма АЧХ при разных углах расположения звукового источника теневого действия головы.

сказывается еще и спектральный фактор. Например, если звук, приходящий под углом  $\theta = 90^\circ$ , содержит как низкочастотные, так и высокочастотные составляющие, то в спектре звука, действующего на дальнее ухо, высокочастотных составляющих будет меньше, так как на этих частотах скажется

Кроме того, сами ушные раковины производят сложную фильтрацию звука, зависящую от его частоты, что будет рассмотрено дальше. Существенное значение для локализации имеет также энергия переходных процессов, причем наибольшее значение имеет наличие в звуке низкочастотных составляющих переходного процесса. Поэтому при прослушивании музыкальных и речевых сигналов изменение спектрального состава сигнала, а, следовательно, и его тембра, в зависимости от его расположения, помогает в локализации.

В целом анализ способности к локализации в горизонтальной плоскости показал, что наименьший ощутимый угол отклонения источника при восприятии звуковых импульсов составляет около  $3^\circ$ . Эту величину следует считать угловой, или бинауральной разрешающей способностью слуха. Однако слух замечает угловое смещение на  $3^\circ$ , но при определении направления совершает ошибку в среднем на  $12^\circ$ . Поэтому точность локализации имеет величину  $12^\circ$  для источников, находящихся в передней полуплоскости, а для источников, расположенных позади слушателя, эта точность еще меньше.

#### Вертикальная (высотная) локализация

Способность определять направление прихода звука в вертикальной плоскости у человека развита значительно слабее, чем в горизонтальной. Она составляет  $10-15^\circ$  (по сравнению с  $3^\circ$  в

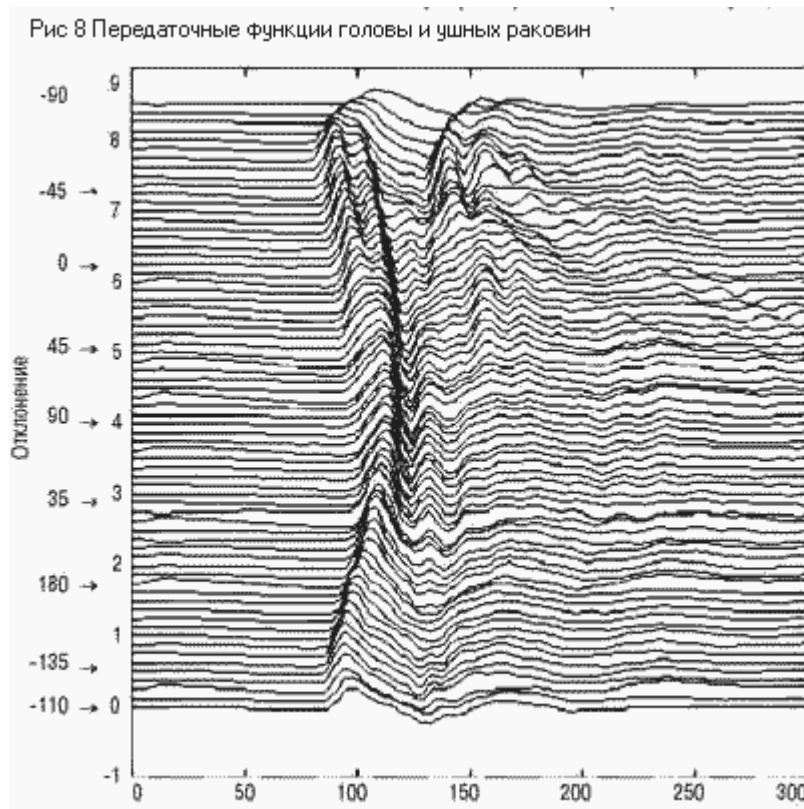


Рис 8 Передаточные функции головы и ушных раковин

горизонтальной). Эту способность связывают обычно с ориентацией и формой ушных раковин: если в ушной канал поставить микрофоны и записать звук от источника, находящего в разных точках медианной плоскости (также и в горизонтальной плоскости), то АЧХ (Рис. 7) будет разной при приходе звука спереди - сверху и сзади на АЧХ отчетливо видны пики за счет отражения от ушной раковины в области 4 - 8 кГц, хотя есть пики и ниже 2 кГц за счет отражения от грудной клетки и спины слушателя.

Ушная раковина имеет сложную геометрию, она действует как акустическая антенна: на низких частотах она усиливает общую энергию сигнала, на средних и

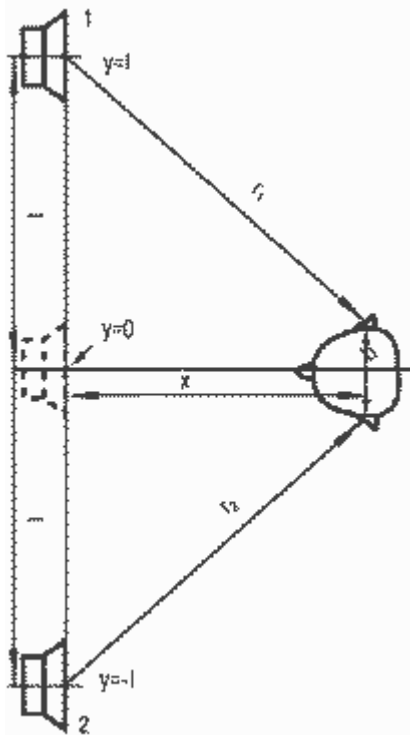


Рис 9 Расположение мнимого источника

высоких частотах начинают сказываться резонансы ее внутренних полостей, усиливая некоторые частоты. Кроме того, происходит интерференция прямого звука со звуком, отраженным от отдельных участков ушной раковины, то есть ушная раковина действует как фильтр, внося максимальные искажения в области 6-16 кГц, причем форма этих искажений зависит от того, спереди или сзади находится источник звука, и под каким углом подъема он расположен в медианной плоскости. Вид АЧХ сигнала, записанного на микрофоны, находящиеся в ушных раковинах при разных положениях источника, показан на рисунке 8 (они называются бинауральными передаточными функциями головы - HRTF).

Эта зависимость АЧХ звукового давления, поступающего на барабанную перепонку левого и правого уха, от положения источника, используется для сравнения спектральных компонент сигнала, приходящего спереди и сзади или и сверху, и их локализации. Поэтому широкополосные сигналы лучше локализируются, чем узкополосный шум.

Если звуковые сигналы подавать через наушники, то ушные раковины оказываются прижатыми к голове. Поскольку такая ситуация для мозгового процессора является неестественной, человек теряет способность производить локализацию в пространстве, помещая при этом источник звука как бы внутрь головы. Это свойство называется латерализацией и служит причиной значительной утомляемости людей, долгое время работающих в наушниках. В настоящее время созданы цифровые процессоры, которые производят предварительную фильтрацию сигналов в наушниках, аналогичную тому, как это делает ушная раковина. Это дает возможность "выносить" звуковой образ из головы, облегчая работу звукорежиссеров, операторов и др.

#### Глубинная локализация (оценка расстояния до источника)

Чувствительность слуха к расстоянию до источника имеет жизненно важное значение - гудок автомобиля, находящегося сзади близко или далеко, должен вызывать разную реакцию. Однако именно это свойство слуховой системы изучено явно недостаточно. Среди основных факторов, определяющих оценку глубины можно выделить следующие:

- уменьшение уровня звукового давления с расстоянием - на низких частотах, где длина волны большая (\* \*5-15 м), любой источник можно считать точечным, и звуковые волны вокруг него - сферическими. В сферической волне площадь поверхности увеличивается пропорционально квадрату расстояния, и соответственно давление падает обратно пропорционально расстоянию, то есть на 6 дБ при каждом удвоении расстояния.

Многочисленные эксперименты по смещению источника и оценке кажущегося расстояния до слухового образа (выполненные в заглушенной камере и на открытом пространстве) показали, что, при удалении источника-громкоговорителя на расстояние от 1 до 10 м, слуховой образ у экспертов (в заглушенной камере при отсутствии визуального контроля) также смещался в этом же направлении, но имело место отставание слухового образа от реального источника - чем дальше, тем больше.

Ощущение удвоения расстояния до звукового объекта возникало только при уменьшении уровня звукового давления на 20 дБ (а не на 6 дБ, как при объективном измерении). При этом

точность локализации была не очень велика: ошибка для широкополосного сигнала (щелчки, часы и т.д.) составляла от 3,5 до 30 см при изменении расстояния от 1 до 8 м. Если при увеличении расстояния повышать напряжение на громкоговорителе так, чтобы уровень звукового давления у слухового канала эксперта не менялся, то способность определять расстояние до источника (глубинная локализация) исчезает.

Таким образом, при отсутствии визуального контроля в условиях свободного поля, когда отраженные сигналы поглощаются (например, в заглушенной камере или в свободном пространстве), уровень звукового давления в месте расположения эксперта является решающим признаком, по которому и оценивается расстояние до источника.

При больших расстояниях (больше 15 м) начинает сказываться затухание, зависящее от расстояния, проходимого звуковой волной. При этом высокочастотные составляющие затухают быстрее, и спектральный состав сигнала при удалении источника меняется (тембр становится "темнее"). Кроме того, на распространение звука оказывает влияние влажность воздуха и направление ветра на открытом пространстве.

Следует отметить, что возможности слуха по определению глубины расположения источника ограничены, имеется "акустический горизонт".

На близком расстоянии (менее 3 м), на глубинную локализацию начинает оказывать влияние также дифракция на ушной раковине и голове, то есть сказываются разности уровней интенсивностей (выше 1500 Гц) и временные задержки (ниже 1500 Гц), как и в предыдущих случаях.

Приблизительно локализацию по глубине при расстояниях меньше 3 м можно оценить по формуле:

$L=2C \cdot dT (\ln I_0 / dI)$ , где  $dT$  - временная разность сигналов,  $dI$  - интенсивностная.

При этом на близких расстояниях меняется спектральный состав при смещении звукового источника за счет дифракционных эффектов, то есть меняется тембр ("тускнеет" при приближении к источнику).

Таким образом, при изменении расстояния до источника меняется одновременно громкость и тембр, что и служит различительными признаками.

Общая точность глубинной локализации не очень велика, при смещении широкополосного звукового источника от 50 до 150 см ошибки составляют 15-30%.

Существенную роль для глубинной локализации играет личный опыт, если слушателю знаком сигнал, а если он имеет возможность сделать визуальную оценку, то точность глубинной локализации многократно увеличивается.

Точность глубинной локализации звукового источника значительно повышается в закрытом реверберирующем помещении. Роль реверберации в оценке удаленности источника, например, распределения музыкантов по глубине оркестра, исключительно велика. При перемещении звукового источника по глубине меняется отношение энергии прямого звука к энергии отраженного (реверберационного) звука, что помогает точнее определить расстояние до источника. Важнейшее значение имеет также разность по времени между прямым звуком и приходом первых отражений и соотношение их по уровням.



$L = \frac{\alpha S}{50(1-\alpha)} \sqrt{\frac{E_{\text{рев}}}{E_{\text{пр}}}}$  Приблизенно, глубинную локализацию в помещении можно оценить следующим образом:

где  $\alpha$  - коэффициент поглощения,  $S$  - площадь поверхности,  $E_{\text{рев}}/E_{\text{пр}}$  - отношение плотностей отраженной и прямой энергии.

Субъективное ощущение "акустики зала" определяется целым рядом параметров, некоторые из них прямо связаны с пространственной локализацией:

Пространственное впечатление (камерность, интимность, близость) - определяет для слушателя кажущийся размер пространства. Разные стили музыки требуют разных его значений. Композитор (звукорежиссер, исполнитель и др.) должен иметь в виду этот параметр, иначе будет несоответствие стиля музыки размеру помещения (например, звучание органа в маленькой комнате), что очень четко ощущается слушателями.

Пространственное впечатление определяется разницей во времени между прямым звуком и первыми отражениями. В залах с "интимной" акустикой эта разница составляет для слушателей в центре зала 15-30 мс. Если эти отражения имеют похожие спектр и огибающую, и их громкость не выше прямого звука, то в пределах этого времени они не воспринимаются как отдельные отражения, а помогают в улучшении локализации прямого звука, в том числе глубинной. Малая разница во времени прихода первых отражений характерна для музыкальных комнат XXVIII столетия, средняя - для концертных залов 19 века, большая - для соборов.

Амбиентность - ощущение слушателя, что музыка от источника (например, оркестра) идет от всего фронта сцены, и звук окружает его со всех сторон.

Тренированный слушатель различает две составляющие в восприятии амбиентности: кажущееся расширение площади источника звука, и окружение (обволакивание), когда слушатель чувствует себя погруженным в звук, окруженным им со всех сторон.

По мнению многих экспертов, кажущееся расширение площади источника является одним из главных индикаторов акустического качества концертных залов и помещений прослушивания. Оно связано с уровнем боковых отражений - чем выше этот уровень, тем больше кажущееся расширение источника.

Кроме того, высокую связь с этим параметром показали результаты измерения на искусственной голове коэффициента внутрислуховой кросс-корреляции сигнала, усредненного в интервале 0-80 мс и измеренного в третьоктавных полосах с центральными частотами 500, 1000, 2000 Гц. Значения этого коэффициента (в соответствии с измерениями Беранека) для девятнадцати лучших залов мира составляют от 0,35 до 0,6. Кажущаяся ширина звукового источника связана также с уровнем звукового давления на низких частотах, в основном в области частот 125 и 200 Гц.

Обволакивание (окружение) - связано с ощущением позднего реверберирующего звука, поступающего со всех сторон (после 80 мс). Оно определяется конструкцией зала: наличием нерегулярностей стен, балконов и пр., то есть всеми конструктивными элементами, которые обеспечивают приход звука с разных сторон. Ощущения от звучания музыки у слушателя, к которому отраженные звуки приходят со всех сторон: от потолка, стен, пола и т.д., будут существенно отличаться от ощущений слушателя, сидящего под балконом, к которому звук приходит только с фронта. Оно связано с коэффициентом внутрислуховой кросс-корреляции, усредненного за период времени от 80 мс до 1 с.

Таким образом, наш слуховой аппарат, используя разные механизмы обработки звуковых сигналов, позволяет определить и локализовать положение звукового источника в трехмерном пространстве. Именно эта способность используется при создании современных систем компьютерного моделирования трехмерных звуковых пространств (системы аурализации).

Это же свойство слуха используется и в современных системах пространственного звуковоспроизведения. Создавая искусственные условия, к которым наша слуховая система не была приспособлена в процессе естественной эволюции, например, помещая два одинаковых громкоговорителя на одинаковом расстоянии от левого и правого ушей, подавая на них одинаковые сигналы, (Рис. 9), мы заставляем наш слуховой аппарат помещать слышимый (мнимый) источник звука посередине между реальными звуковыми источниками.

Пространство таких мнимых источников, создаваемых различными пространственными системами воспроизведения (стереофоническими, Surround и др.), и создает стереоэффект - по существу, это "большой обман" нашего слухового аппарата. Вопрос о том, как формируется и как управляется этот пространственный образ мнимых (виртуальных) источников, может служить предметом рассмотрения отдельной статьи.