

Часть 15.2

Слуховое восприятие пространственных систем, часть 2

Ирина Алдошина

Развитие систем стереовоспроизведения и современных систем пространственного звуковоспроизведения (Home Theatre и др.) основано на создании слуховых иллюзий – большом обмане слуховой системы. Две или большее число акустических систем (например, 5.1) воспроизводят звук из разных точек помещения, а слушатель слышит звуки, исходящие из несуществующих (мнимых) источников, воспринимая определенную звуковую панораму. Скорее всего, этот обман оказывается возможным потому, что все системы звуковоспроизведения являются искусственными, сравнительно недавно созданными устройствами, поэтому слух в процессе длительного биологического развития оказался неприспособленным для распознавания этого обмана (возможно, существует и более глубокое объяснение этого явления).

Однако бурное развитие пространственных систем звуковоспроизведения и многочисленные психоакустические эксперименты показали, что обман происходит не полностью: слух действительно локализует мнимый источник там, где никакого реального источника не существует, – но вот слышит ли он тембр от этого источника таким же, как от реального источника, находящегося на этом же месте, – это очень большой вопрос.

В рамках большого исследовательского международного проекта Medusa по исследованию восприятия пространственных систем звуковоспроизведения были поставлены работы в исследовательских центрах различных стран (Англии, Дании, США и др.) по анализу тембров мнимых источников в разных условиях звуковоспроизведения. Исследования эти особенно актуальны в настоящее время не только в связи с развитием пространственных систем, но и созданием виртуального акустического пространства (например, систем аурализации, "Звукорежиссер", 7/2000). Результаты, полученные на первых этапах этих исследований, могут быть полезны звукорежиссерам, осваивающим новые технологии пространственной звукозаписи.

Как уже было показано в статьях, посвященных тембру ("Звукорежиссер", 2, 3/2001), тембр определяется в соответствии с международными стандартами как "атрибут слухового ощущения, с помощью которого слушатель может судить, что два звука, одинаковые по высоте и громкости, отличаются друг от друга". Оценка тембра зависит от многих факторов: амплитудного спектра и его изменения во времени, фазового спектра, процессов установления и спада звука и др.

На начальном этапе исследований тембров виртуальных источников и их отличий от реальных источников были проведены работы по сравнительной оценке только амплитудных спектров. Было высказано предположение, что для оценки тембров используется формирование спектральной плотности, которое лежит в основе определения громкости, а именно – распределение спектральных уровней внутри третьоктавных полос. Распределение третьоктавных полос очень близко к частотно-зависимой ширине критических полос слуха (рисунок 1). Под "критической полосой" ("Звукорежиссер", 9/2000) понимается ширина полосы пропускания слуховых фильтров, с помощью которых происходит спектральный анализ звуковых сигналов на базилярной мембране во внутреннем ухе.

Ширина этих полос зависит от частоты и интенсивности сигнала. Наличие критических полос представляет собой очень важное свойство слуховой системы. Внутри каждой критической полосы происходит интеграция энергии независимо от вида сигнала: отрезок шума или спектральные компоненты тональных сигналов, находящиеся внутри критической полосы, если они имеют одинаковый уровень интенсивности, интегрируются и создают ощущение одинакового уровня удельной громкости. Поэтому общее ощущение громкости связано с распределением спектрального уровня сигнала по критическим полосам и определяется возникающими при этом ощущениями удельной громкости.

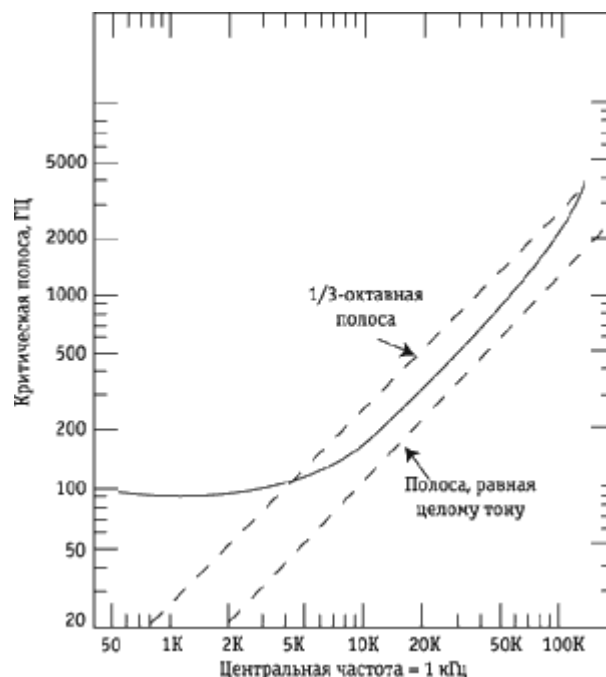


Рис.1 Зависимость ширины третьоктавных полос и критических полос слуха от частоты

В работах Цвиккера было показано, что бинауральные эффекты при восприятии тембров зависят от сумм спектральных удельных громкостей, поступающих из каждого ушного канала в центральные отделы слуховой системы.

Компьютерная модель периферической слуховой системы, построенная для оценки бинауральной громкости, была использована и для оценки различия тембров реальных и виртуальных источников. Принцип

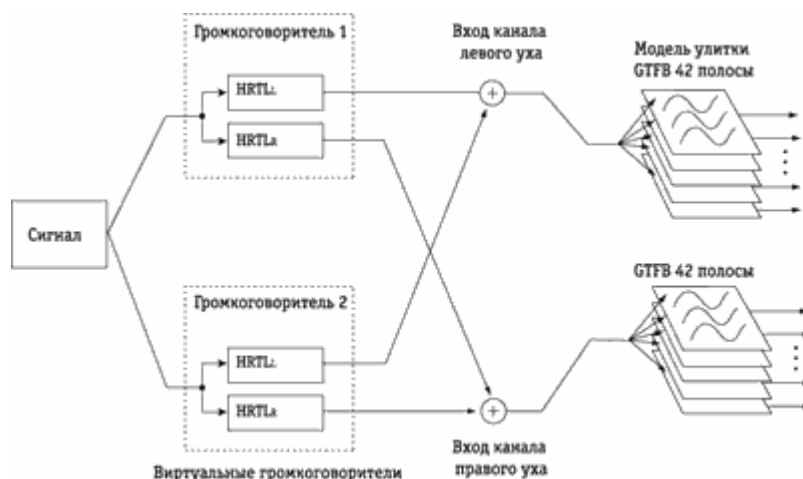


Рис.2 Принцип работы компьютерной слуховой модели

работы компьютерной модели слуховой системы, показанный на рисунке 2, использует построение передаточных функций левого и правого уха (HRTF) с учетом дифракции на голове и ушной раковине, и фильтрации сигнала во внешнем и среднем ухе. Кроме того, в модели происходит спектральная обработка сигнала с помощью системы полосовых фильтров в 42 полосах, характеристики которых аналогичны слуховым фильтрам на базилярной мембране. Затем подсчитывается энергия сигнала на выходе каждого фильтра, и из него определяется удельная спектральная громкость. Полученное таким образом спектральное распределение удельной громкости и используется в дальнейших экспериментах для оценки

тембров мнимых источников.

Для оценки параметров виртуального источника (его локализации и тембра) были использованы различные конфигурации расположения громкоговорителей в помещении: два громкоговорителя, расположенные на равном расстоянии от слушателя (обычная стереопара, рисунок 3); три громкоговорителя, расположенные в виде треугольника; громкоговорители, расположенные в различной пространственной позиции, в том числе в вертикальной плоскости. Кроме того, оценки выполнялись при поворотах головы слушателя (рисунок 3а, 3б). На громкоговорители подавался когерентный (одинаковый) сигнал, панорамирование (создание и размещение виртуального источника) происходило за счет изменения амплитуд сигнала в каждом канале громкоговорителя путем изменения коэффициента усиления.

Как было указано в первой части этой статьи, локализация мнимого источника при расположении двух громкоговорителей в горизонтальной плоскости происходит за счет разницы сигналов во времени (ITD) и разницы по интенсивности (IID) между сигналами, попадающими на разные уши. Как показано на рисунке 3, определение направления на источник звука в горизонтальной плоскости производится с помощью расчета угла θ_{cc} . Для расположения двух громкоговорителей в горизонтальной плоскости угол на мнимый источник при интенсивностном (амплитудном) панорамировании (только за счет изменения амплитуд сигналов от громкоговорителей) определяется по формуле:

$$\frac{\lg \theta_{cc}}{\lg \theta_0} = \frac{g_1 - g_2}{g_1 + g_2}$$

где θ_{cc} и θ_0 – углы на мнимый источник и на реальный громкоговоритель, g_1 и g_2 – коэффициенты усиления сигналов в громкоговорителях. Этот широко известный в акустике закон дает достаточную точность при определении направления на мнимый источник для стереопары громкоговорителей.

Для произвольно расположенных в трехмерном пространстве громкоговорителей направление на мнимый источник определяется с помощью вектора пространственного панорамирования P_T , который выражается в матричной форме: $P_T = g L_{nmk}$, где L_{nmk} – вектор направлений на громкоговорители (рисунок 4), P_T – вектор направлений на мнимые источники, g – вектор коэффициентов усилений. Этот метод

определения положения мнимых источников (VBAP) был предложен сравнительно недавно, в 1997 году финским ученым В. Пулки. С его использованием были созданы цифровые процессоры, позволяющие управлять в реальном времени положением мнимых источников в трехмерном пространстве (рисунок 5). В частности, было реализовано цифровое устройство на восемь входов и восемь выходов, то есть восемь громкоговорителей, расположенных в пространстве, которые были успешно использованы в системах мультимедиа и даже в больших залах на концертах компьютерной музыки. Возможно, сейчас имеются устройства и для большего числа каналов.

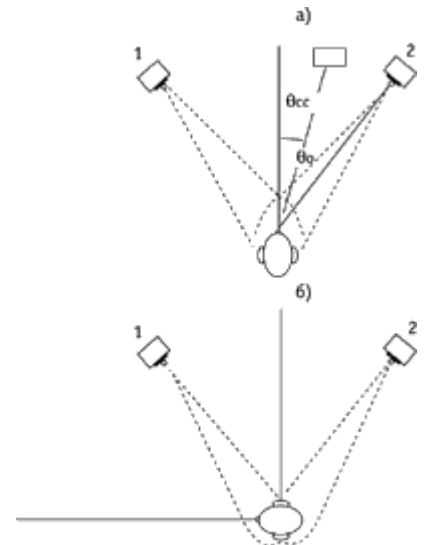


Рис.3 Система расположения громкоговорителей

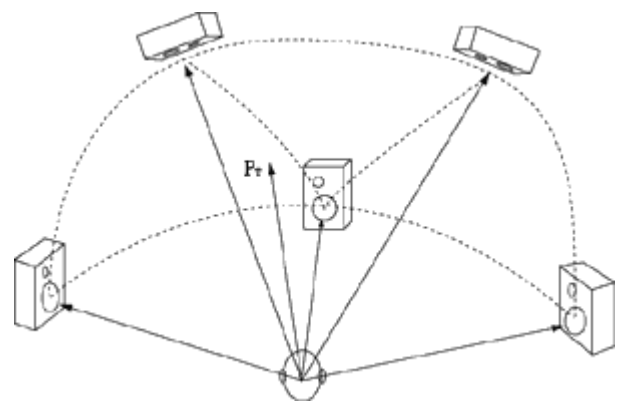


Рис.4 Трехмерная система расположения громкоговорителей

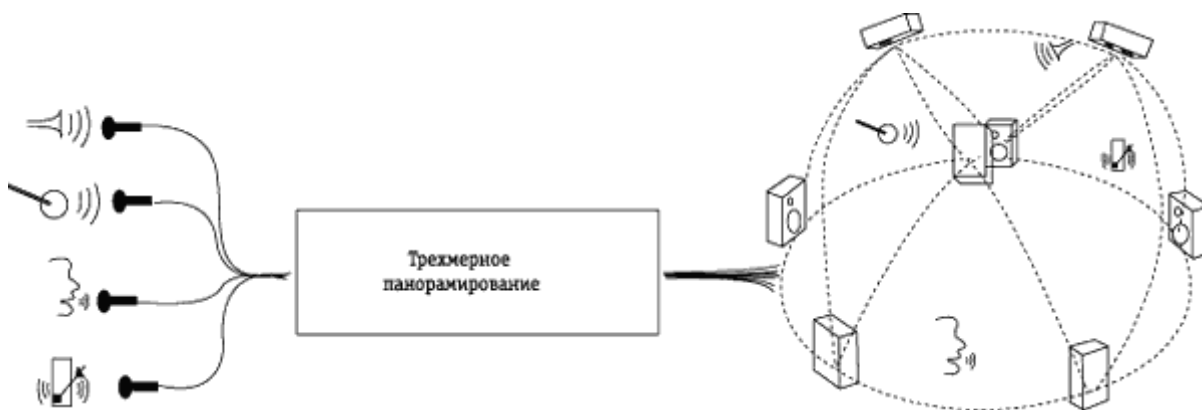


Рис.5 Панорамирование пространственного источника

Кроме того, как было показано в одной из предыдущих статей ("Звукорежиссер", 10/1999), при локализации источников, находящихся в боковой плоскости, существует так называемый "конус неопределенности", где локализация происходит с большим трудом, не подчиняется вышеуказанным законам, носит очень индивидуальный характер, и обычно требует дополнительных поворотов головы.

Для выполнения сравнительной оценки тембров виртуальных и реальных источников на первом этапе определялось положение виртуального источника для заданной конфигурации громкоговорителей расчетным путем по вышеуказанным формулам, а также на компьютерных моделях слуховой системы, с помощью которых производился расчет спектра звукового сигнала в ушном канале в каждой критической полосе (аналогично определению средней громкости). Затем в эту же позицию помещался реальный источник, и для него также проводился аналогичный расчет и дополнительные измерения спектра реального звука в слуховом канале. После этого из спектра мнимого источника вычитался спектр реального источника, и по полученной разности оценивалась разница спектров и, соответственно, разница в воспринимаемых тембрах. Конечно, одной информации о спектрах недостаточно для полного суждения о разнице в тембрах, однако для первичной оценки это может служить основанием.

Результаты, полученные для двух громкоговорителей в стереопаре в заглушенной камере, представлены на рисунке 6. Здесь показана разность между суммарными сигналами, поступающими в каждое ухо с учетом передаточных функций головы, определяемых дифракцией на голове и ушной раковине. Они были измерены у двадцати различных слушателей от мнимого и реального источника, помещенного на место локализованного мнимого источника. По оси абсцисс отложены центральные частоты для каждой критической полосы, по оси ординат – уровень громкости в фонах (значение громкости в фонах равно значению уровня сигнала в дБ на частоте 1 кГц).

Как видно из рисунка 6, эта разница очень значительна в области 2 кГц (имеется широкий провал более 10 фон), имеются также существенные различия и в области высоких частот: на частоте 8 кГц – более 4 дБ. Эта разница появляется за счет так называемого эффекта "гребенчатой фильтрации", которая возникает при сложении когерентных (одинаковых) сигналов от двух (или нескольких) громкоговорителей на каждом слуховом канале. Этот эффект происходит при определении

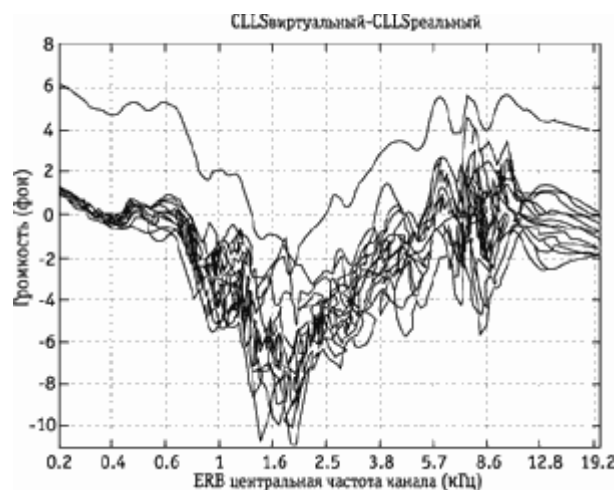


Рис.6 Разность спектральных уровней громкости реального и мнимого источника

положения мнимого источника, и отсутствует, когда звук идет от одного громкоговорителя, находящегося в этом же месте – при таком сложении появляются глубокие провалы в спектре суммарных сигналов.

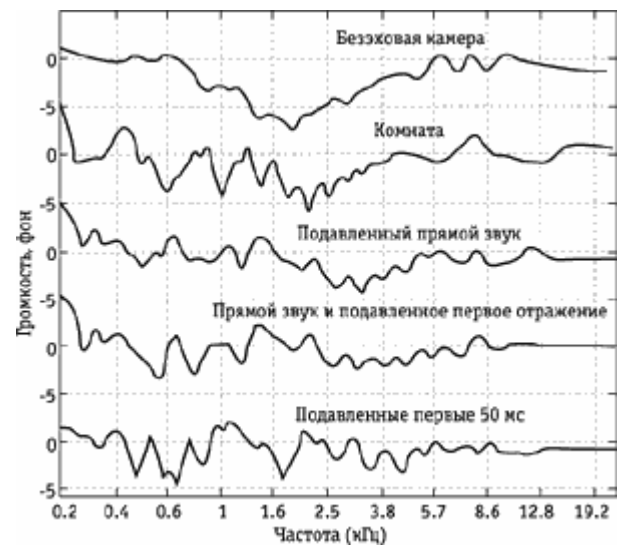


Рис.7 Разность спектральных уровней громкости реального и мнимого источника в заглушенной камере, реальном помещении, при подавлении первых отражений

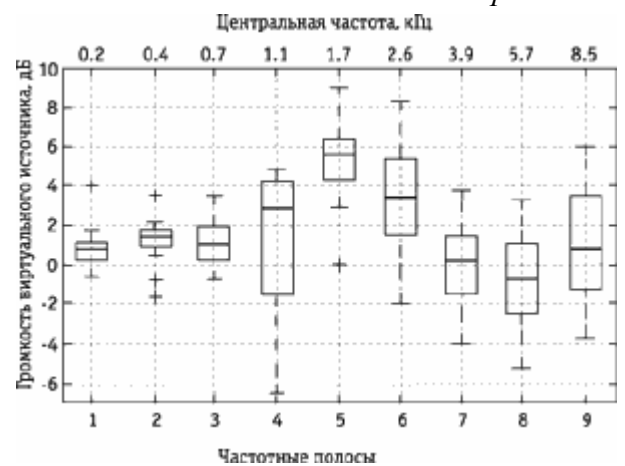
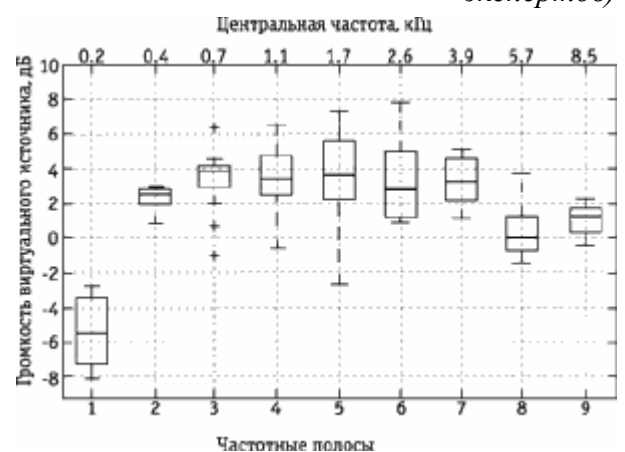


Рис.8 Разность спектральных уровней громкости по результатам субъективных экспертиз в заглушенной камере (8 экспертов)



Аналогичные эксперименты и расчеты были повторены для реальных условий в реверберирующем помещении. Время реверберации было выбрано $0,3 \pm 0,05$ с, а

громкоговорители размещались в горизонтальной плоскости под углом 30° на расстоянии 2 м. При расчетах разности в спектрах реального и мнимого источников на модели слуховой системы учитывались бинауральные импульсные характеристики помещения, результаты показаны на рисунке 7. Видно, что наличие в комнате отраженных сигналов уменьшает эффект гребенчатой фильтрации, а также уменьшает разницу между спектрами реальных и мнимых источников – однако она все равно достаточно велика в области средних частот. Если подавить прямой сигнал и первые отражения на промежутке 50 мс, т.е. оставить только диффузный отрезок реверберации, то видно, что никакой разницы (и, соответственно, окрашивания звука, изменения тембра) практически нет, поскольку в диффузном поле эффект гребенчатой фильтрации отсутствует.

Все результаты расчетов и экспериментов были проверены с помощью субъективных экспертиз. На первом этапе прослушивания производились для стереопары, установленной в заглушенной камере под углом $\pm 30^\circ$ на расстоянии трех метров. На место мнимого источника устанавливался сначала реальный источник с определенным уровнем звукового давления, затем слушателей просили установить усиление таким образом, чтобы субъективно выровнять громкости в каждой полосе частот. Испытания производились на разных сигналах: синусоида, узкополосный шум, импульсные сигналы и др.

Результаты показаны на рисунке 8, и они совпадают с результатами, полученными на моделях слуховой системы: в области 2 кГц имеет место подъем усиления примерно на 6 дБ (по-видимому, для компенсации провала за счет эффекта гребенчатой фильтрации); в области низких частот разницы практически нет, поэтому дополнительное усиление для мнимого источника практически оказывается равным нулю (ось ординат на рисунке 8).

Таким образом, слушатели по-разному воспринимают спектральную громкость реального и мнимого источника и, соответственно, по-разному оценивают его тембр.

Испытания были повторены для обычной комнаты прослушивания. Использовались два монитора фирмы Genelec на расстоянии двух метров от слушателей, которые располагались в центре. Результаты тестов для узкополосных шумовых сигналов показаны на рисунке 9. Из полученных результатов видно, что в области 2 кГц мнимый источник имеет громкость ниже, и требуется его усиление – как и в заглушенной камере, но разница меньше. Однако в области низких частот мнимый источник кажется выше по уровню, чем реальный источник в этой же точке (на оси), и поэтому его уровень усиления понижается. Таким образом, помещение снижает отклонения за счет эффекта гребенчатой фильтрации, но вносит свои дополнительные искажения.

Полученные результаты отражают субъективно воспринимаемое различие в спектрах мнимых и реальных источников, а это служит основой для различий в восприятии тембров.

Среди специалистов по звуку существует мнение, что если сравнивать мнимый источник не с реальным источником, находящимся на его месте (например, в центре), а с одним из боковых громкоговорителей из стереопары, то различие в тембре проявляется не так сильно. Действительно, проведенные эксперименты показали, что различие с субъективно воспринимаемой спектральной громкостью будет меньше, если сравнивать мнимый источник в центре с одним из боковых громкоговорителей. Получается интересное психоакустическое явление: слух локализует мнимый источник в одном месте (например, в центре), но формирует представление о тембре, ориентируясь на реальный источник, находящийся сбоку (то есть на громкоговоритель, воспроизводящий этот звук), а не на тембр реального громкоговорителя, который должен был быть в центре.

Испытания были продолжены для панорамирования мнимого источника под разными углами к оси: 0° , 10° , 20° и 30° . Чем ближе мнимый источник подвигается к одному из громкоговорителей (которые расположены под углами $\pm 30^\circ$), тем меньше проявляются

Рис.9 Разность спектральных уровней по результатам субъективных экспертиз в реальном помещении

отличия в спектрах и, соответственно, окрашивание тембра. Кроме того, было исследовано влияние поворотов головы слушателя, находящегося в центре под углами 0° , 30° , 60° к направлению на мнимый источник. Установлено, что область наибольшего окрашивания при поворотах головы смещается в район 4...8 кГц, что соответствует области максимального проявления эффектов дифракции на ушной раковине.

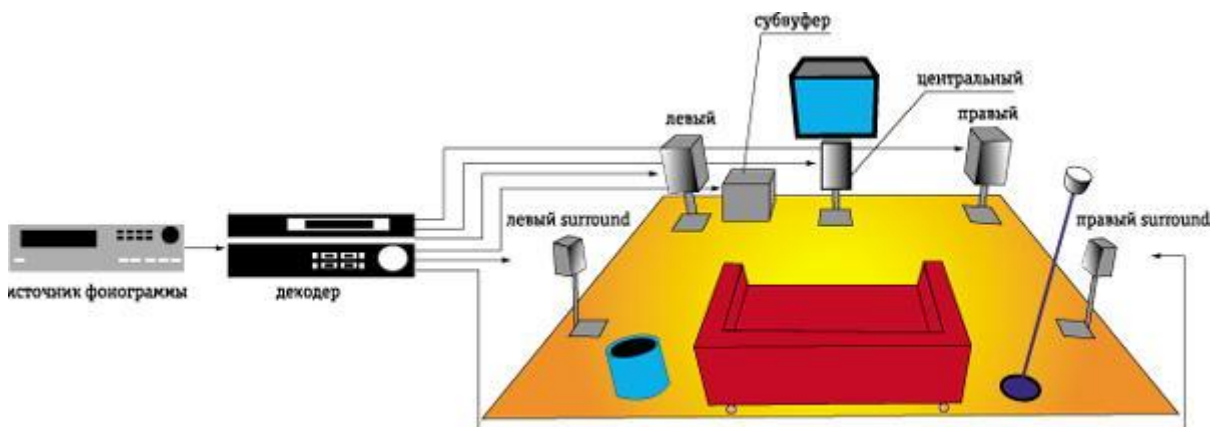


Рис.10 Расположение громкоговорителей по системе Home Theatre

Можно предположить, что когда для создания мнимого источника (распределенного виртуального пространственного образа) используется несколько громкоговорителей, размещенных различным образом в трехмерном пространстве, например, в системах Home Theatre (рисунок 10), то окрашивание, т.е. отличие в воспринимаемом тембре мнимого источника, будет значительно больше. Во-первых, потому, что большее число когерентных сигналов прибывает на каждый ушной канал от нескольких громкоговорителей, поэтому эффекты гребенчатой фильтрации проявляются гораздо сильнее. Во-вторых, когда несколько громкоговорителей разнесены дальше в пространстве, их отличия друг от друга по тембру будут сильнее, и это также увеличит окрашивание мнимого источника. Эти предположения были проверены на моделях слуховой системы, и результаты тестов прослушивания будут опубликованы в ближайшее время.

Громкоговорители были расположены по системе треугольник (рисунок 11) под углами 0° , $+30^\circ$, -30° в горизонтальной плоскости, и $0^\circ +15^\circ$, -15° в вертикальной плоскости. Мнимый источник перемещался в точки, указанные на рисунке 11 кружком: $(0, 0)$, $(0, +15)$, $(0, -15)$. Результаты измерений показаны на рисунке 12. На оси абсцисс отложен уровень громкости в фонах, по оси ординат – центральная частота, как на предыдущем рисунке. На рисунке 11 видно, что провалы за счет гребенчатой фильтрации стали значительно шире, от 1,5 до 4 кГц, при изменении угла подъема они смещаются по частотной шкале. Естественно, что это увеличивает разницу при субъективном прослушивании в уровнях громкости, и требует большей компенсации, что и приводит к большему окрашиванию тембра.

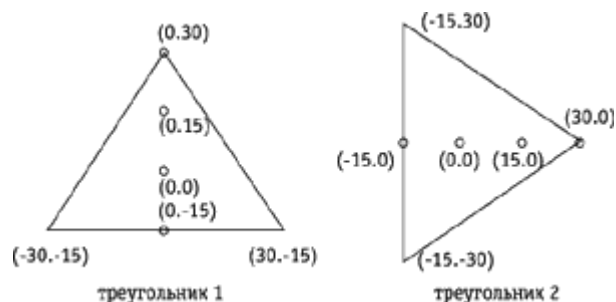


Рис.11 Расположение громкоговорителей по системе "треугольник"

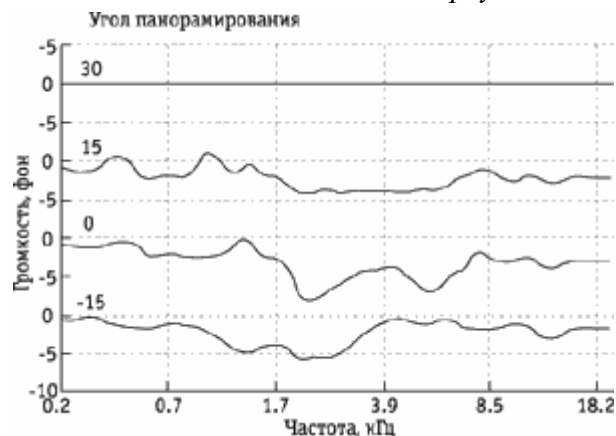


Рис.12 Расположение громкоговорителей по системе "треугольник"

В целом эксперименты по оценке тембров виртуальных источников при размещении различных конфигураций систем излучателей в заглушенной камере и в реальном помещении, выполненные как на моделях слуховой системы, так и с помощью субъективных тестов, показали, что результаты, полученные на моделях и с помощью слушательских тестов, достаточно хорошо коррелируют друг с другом, и могут служить основой для предсказания изменения тембров мнимых источников при различных установках громкоговорителей.

Проведенные эксперименты позволили сделать ряд важных для практики звукозаписи выводов:

- окрашивание тембра виртуального источника при различном положении громкоговорителей происходит за счет двух эффектов: гребенчатой фильтрации и влияния реверберационных процессов в помещении. Звуковые волны от нескольких громкоговорителей прибывают к ушным каналам в разное время, что служит причиной интерференционных эффектов типа "гребенчатого фильтра" при восприятии мнимого источника. При стереорасположении громкоговорителей этот эффект создает глубокие провалы в области 1...4 кГц в спектре сигналов, поступающих на ушные каналы. Эти провалы приводят к существенным отличиям при восприятии тембра и громкости виртуального и реального источников, особенно при прослушивании в заглушенной камере или сильно заглушенном помещении;
- при увеличении заглушения помещения, то есть при сильном уменьшении времени реверберации, разница в восприятии тембров увеличивается;
- влияние этого эффекта особенно сильно, когда громкоговорители находятся во фронтальной плоскости перед слушателем; когда же они сдвигаются в сторону или слушатель поворачивает голову, то провалы сдвигаются в сторону высоких частот, и становятся менее заметными на слух;
- когда пространственные системы размещаются в реальном помещении, эффект гребенчатой фильтрации проявляется меньше – в диффузном поле он вообще не проявляется, но на восприятие тембра виртуального источника начинают влиять процессы ранних отражений и поздней реверберации в помещении;
- измерения, выполненные для двух громкоговорителей, для системы из трех громкоговорителей ("треугольник"), а также с использованием разных видов сигналов (шумовые, импульсные и др.), показали, что общие закономерности восприятия тембров и громкости виртуального источника сохраняются.

Таким образом, воспринимаемый пространственный виртуальный источник будет по-разному восприниматься в разных помещениях, и степень его различия от реального источника может существенно меняться в зависимости от степени заглушенности помещения. В настоящее время идет активный процесс перехода на системы пространственного многоканального воспроизведения, и соответственно происходит процесс переоснащения контрольных комнат студийными мониторами, расположенными по системе 5.1 (или другими вариантами систем Surround Sound). Поэтому регулирование панорамы, баланса и других параметров происходит при слуховой оценке именно виртуального пространственного звукового образа. В связи с этим изменения его тембрального окрашивания при изменении положения слушателя и условий в окружающем помещении, а также процесс увеличения отличий его тембра от тембра реального источника, имеют существенное значение для качества звукозаписи. Все это требует продолжения серьезных исследований в данной области, и разработки рекомендаций для практической работы.