

Часть 10.

Аурализация - виртуальный звуковой мир

В статье "Научные результаты 108 конвенции AES" ("Звукорежиссер" №3/2000) мною было обещано сделать три вещи:

- передать CD-ROM с докладами конвенции в редакцию (что было выполнено);
- подробнее рассказать о принципиально новом направлении в создании пространственных звуковых полей компьютерной технологии аурализации (об этом в данной статье);
- познакомить с направлениями исследований в мировом центре компьютерной музыки и акустики IRCAM (а об этом в следующих номерах).

Итак, начнем с самого "горячего" направления в современной звукотехнике - "аурализации".

В одной из статей фирмы "Брюэль и Кьер" была высказана любопытная мысль, что отношения человека со звуком можно разбить на три крупных этапа:

I. От начала эпохи "человека разумного" до начала XX когда звук прошел эволюцию от средства чисто сигнального, необходимого для выживания, через средство коммуникации (речь) к средству эмоционального и эстетического воздействия, то есть к музыке.

К началу XX века музыка достигла невиданных высот, стала мощным средством передачи величайших глубин человеческой мысли и эмоций. Гениальные композиторы (Бах, Бетховен, Моцарт и др.) подняли музыкальное творчество на необыкновенную высоту, разработали особый язык (код), способный передавать не меньшее богатство мыслей и нюансов, чем письменная и устная речь (литература).

Человечество создало широкую палитру музыкальных инструментов, отработало вокальную технику, построило великолепные концертные залы, театры, соборы и др. Однако, эти величайшие достижения искусства были доступны очень ограниченному кругу людей лишь доли процента населения могли слушать хорошую "естественную" музыку в хороших "естественных" залах);

II. от начала XX века до 80-х годов с момента изобретения радио и телевидения музыкальное и вокальное искусство стало доступно миллионам, но, как всегда, при массовом тиражировании качество звука резко упало отставали технические средства. Главной задачей в тот период было передача смысловой (семантической) вербальной информации.

К середине века техника звукозаписи, воспроизведения и звукопередачи значительно выросла, и это позволило поднять проблему передачи эмоциональной и эстетической информации на новый уровень родилось движение Hi-Fi (high-fidelity, высокая верность воспроизведения), идеология которого состояла в том, чтобы акустическая аппаратура могла воспроизводить звук максимально похоже на натуральный "живой" источник. Постановка проблемы достоверной передачи звука в записи дала мощный толчок к развитию акустики и созданию крупной индустрии, производящей звукотехническую аппаратуру.

Современная акустика представляет мощное и развитое направление науки во всех странах

мира, и имеет огромную промышленную базу: сотни научных институтов, тысячи фирм, разрабатывающих и производящих огромное разнообразие звукотехники:

- студийное оборудование микшеры, микрофоны, усилители, мониторы и т.д.;
- передающее оборудование радио- и телепередатчики;
- воспроизводящее оборудование акустические системы, проигрыватели, магнитофоны и др.

Объемы мирового производства например, громкоговорителей, достигают более 500 млн в год; акустические системы только на рынке США представлены 300 фирмами, выпускающими более 3000 моделей) и т.д.

III. от 80 лет до настоящего времени. В начале 80-х известным акустиком проф.Олсоном (Olson) были сформулированы проблемы третьего этапа развития акустики "перенос атмосферы первичного поля в любое вторичное помещение прослушивания"

Сама постановка такой проблемы была бы в принципе невозможна без создания новой научной и технической базы развития цифровых компьютерных технологий обработки звука и соответствующих технических средств: звуковых процессоров, цифровых станций обработки, монтажа, редактирования, архивирования и т.д., цифровых магнитофонов и лазерных проигрывателей и т.д.

Разработка этой проблемы уже привела к развитию пространственных систем звукопередачи (Dolby, бинауральная стереофония и т.д.), появлению адаптивных цифровых процессоров, систем пространственного звуковоспроизведения Dolby Stereo, Dolby surround и др. Однако она потребовала решения целого ряда новых задач в области психоакустики, и привела к появлению новых направлений в цифровой акустике.

Одним из принципиальных этапов в решении этой проблемы было создание технологии аурализации.

Термин "аурализация" (auralization) появился несколько лет тому назад и еще не определен окончательно. Его определение дал Мендель Клейнер (Mendel Kleiner) по аналогии с термином "визуализация" на конгрессе AES в 1989 году. Он звучит так:

"Аурализация процесс превращения звукового поля источника в пространстве в "слышимый звук" путем физического или математического моделирования таким образом, чтобы смоделировать бинауральное слуховое ощущение на заданной позиции моделируемого пространства".

Сейчас этой проблеме уделяется очень большое внимание в специальной литературе: появляется много научных статей и докладов, в том числе и на последнем конгрессе AES; разработаны специальные пакеты программ для реализации этой идеи; появились фирмы, которые специализируются на создании и внедрении таких программных продуктов, например, фирма Одеон, которая предложила свое определение:

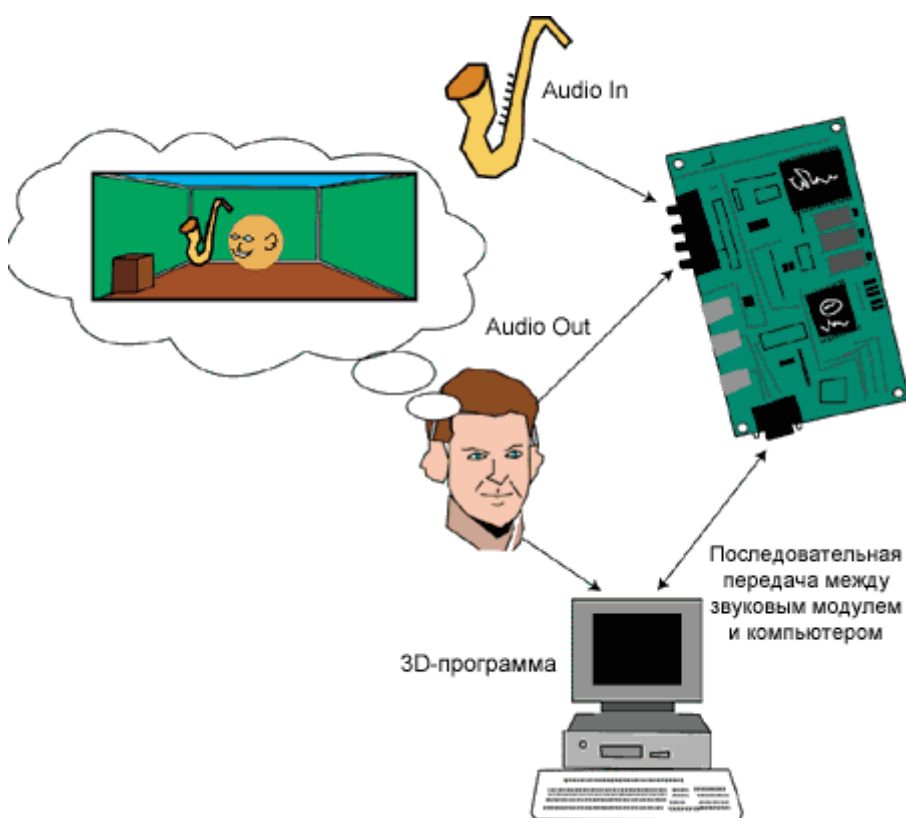
"Аурализация искусство создания цифровых моделей бинауральных записей в несуществующих помещениях".

Иначе говоря, аурализация - это способ воссоздать трехмерное звуковое поле, пытаясь с помощью компьютерных программ повторить способы обработки звука, которые слуховая система применяет к звуковому сигналу в помещении, чтобы создать ощущение пространства.

Необходимо отметить, что точное определение этого процесса пока еще не принято окончательно, а в русской технической литературе его вообще еще нет.

Нужно сказать, что трехмерное визуальное пространство удалось создать раньше, что нашло уже широкое применение в компьютерных играх, видеоклипах, системах обнаружения и др., поэтому моделирование трехмерного звукового пространства стало необходимым этапом, поскольку вместе они могут полностью воссоздать "пространственный виртуальный мир". К чему это приведет в XXI веке сказать трудно...

Попробуем рассмотреть, что надо сделать с музыкальным сигналом, чтобы после его компьютерной обработки слушатель, находящийся при воспроизведении в любом помещении, воспринимал звуковое пространство таким же, как если бы он слушал музыку в реальном концертном зале.



По существу, задача ставится таким образом: как надо "обмануть" мозг, чтобы создать у слушателя ощущение трехмерного звукового пространства концертного зала вот для чего нужна психоакустика. Рис.1.

Рис.1. Общая структура системы аурализации

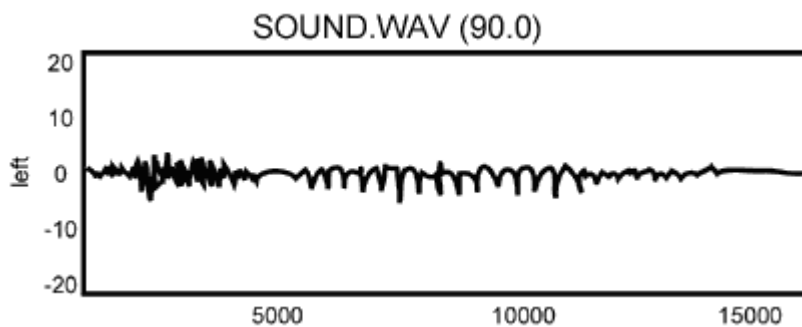
Как уже было рассмотрено в предыдущих статьях по психоакустике (и еще будет рассмотрено в следующих), только два физических параметра сигнала воспринимаются нашей слуховой системой: интенсивность (т.е. энергия

или звуковое давление), и время начало и конец сигнала, и его повторяемость во времени (периодичность или частота).

Человек "слышит" звук, воспринимая изменения величины звукового давления, воздействующие на его барабанную перепонку, во времени. Вся информация, которую мы получаем о звуке, содержится в звуковых волнах, являющихся сжатием-разрежением воздуха. Все остальное, что мы оцениваем в звуке: его громкость, высота, тембр, звуковое пространство, тонкие музыкальные нюансы и др. - это результат обработки его нашим слуховым аппаратом и мозгом.

Рассмотрим, что происходит, когда слушатель воспринимает звук в помещении.

Музыкальный инструмент (голос, оркестр и т.п.) создают определенный акустический сигнал, который представляет собой некоторую звуковую волну, с определенной зависимостью звукового давления от времени $p_1(t)$. Например, осциллограмма звучания мужского голоса при произнесении слова "sound" (записанная в заглушенной камере) показана на рис.2. Затем

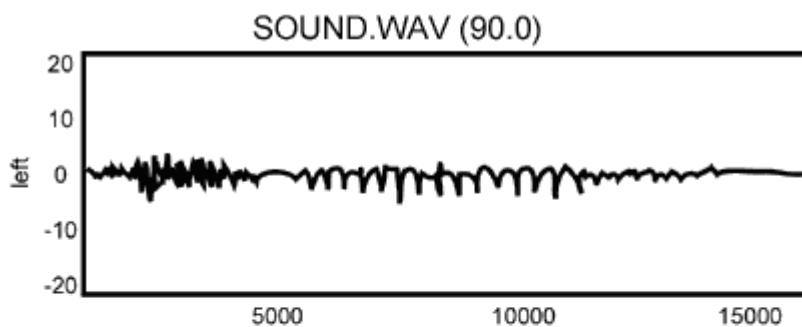


этот сигнал определенным образом изменяется помещением за счет отражений звуковых волн, процессов затухания, дифракции и т.д. Если рассматривать помещение как линейный фильтр, который имеет свои характеристики

$p_{\text{пом}}(t)$, то в каждой точке пространства суммарный сигнал получается как "свертка" сигнала источника и характеристик помещения (термином "свертка" называется результат обработки одного сигнала другим, например, в данном случае

$$p_{\Sigma}(\tau) = p_1 \oplus p_{\text{пом}} = \int_{-\infty}^{\infty} p_1(\tau) p_{\text{пом}}(\tau-t) dt.$$

Рис.2. Спектрограмма звука sound, записанная в заглушенной камере



Наконец, голова и ушные раковины производят свою обработку звукового сигнала (см. "Звукорежиссер", №10/1999).

Пример осциллограммы того же самого звука после обработки его помещением и слуховой системой в левом и правом ушах показан на рис.3.

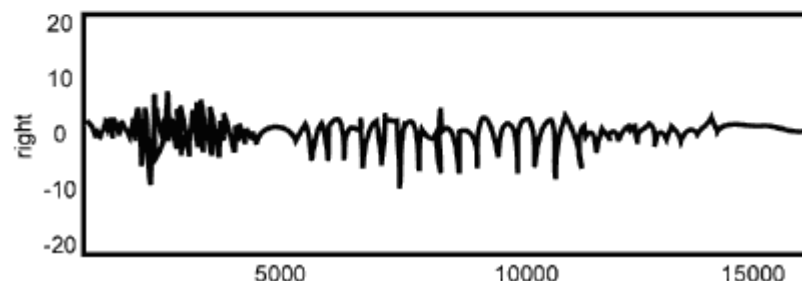
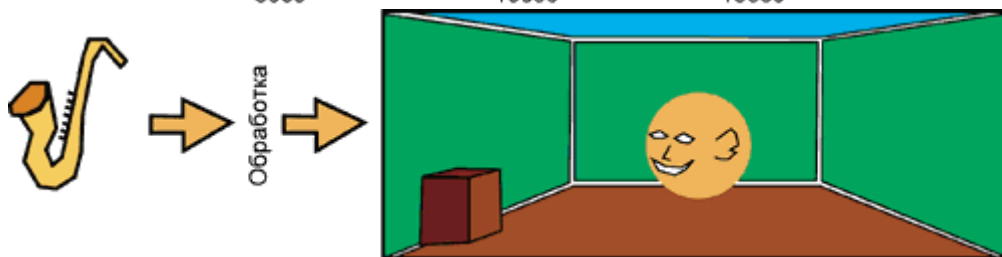


Рис.3. Спектрограмма звука sound, снятая на двух входах в ушной канал в помещении

Таким образом, чтобы вызвать такие же ощущения у слушателя,

надо сделать следующее:

записать оригинальный звуковой источник, изменить программно его



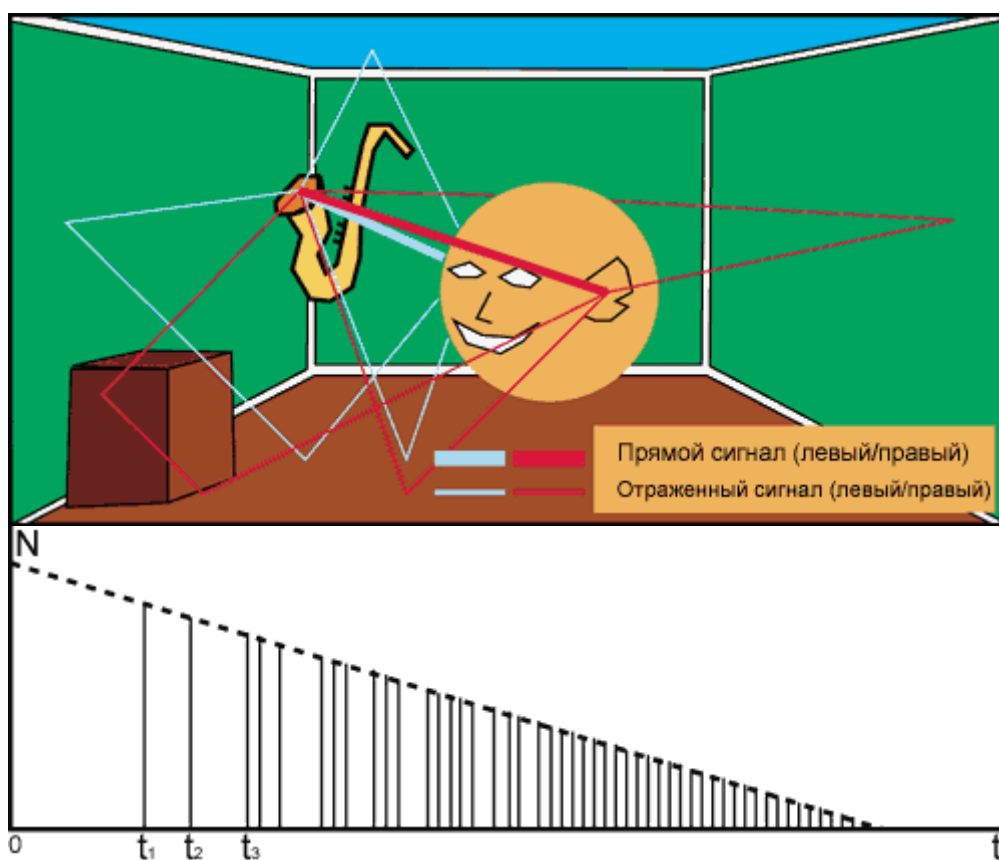
импульсную характеристику (или спектр) так, как это делает заданное помещение (концертный зал, собор или подвал), а затем произвести его фильтрацию аналогично тому, как это происходит в ушных раковинах и голове. После того как это сделано, можно воспроизвести эти сигналы, например, через головные телефоны, и получить ощущение того, что слушатель находится в реальном трехмерном акустическом пространстве вместе с источником звука (певцом, музыкантом, оркестром) см. рис.4 .

Рис.4. Схема обработки сигнала

Процесс формирования звуковых сигналов при аурализации проходит следующие последовательные стадии:

- моноуральная запись источника звука обычно записи делают в заглушенной камере, но допустимо и в полузаглушенных помещениях, если запись делается там, где отражения не являются определяющими. Расстояние для записи выбирается обычно 1 м на оси. Уровень записи должен соответствовать среднему уровню данного источника звука при естественном звучании. Допускается использование синтезированных музыкальных записей, а также различных банков семплов. Это дает возможность прослушать, как в данном помещении, существующем или желаемом, будут звучать различные виды музыки.

- создание компьютерных моделей звукового поля в помещении как известно из архитектурной акустики, при прослушивании любого звукового источника в помещении к слушателю поступает прямой звук и его многочисленные отражения от стен, потолка, пола и др. (рис.5). Для расчета структуры этих отражений в помещениях различной формы и



размеров, с различными поглощающими или отражающими материалами и конструкциями, мебелью, элементами декораций и др., используются пакеты программ, построенные как на приближенных методах геометрической акустики (лучевой метод или метод мнимых источников), так и на более точных методах волновой акустики (МКЭ, МГЭ и др.).

Рис.5. Структура отражений в помещении

Если записать микрофоном звуковой сигнал, например, короткий импульс, воспроизведенный через акустическую систему установленную в помещении, то сигнал в любой точке помещения будет иметь вид, показанный на рис.6, т.е. наряду с прямым сигналом в данную точку приходят ранние дискретные отражения, затем число их увеличивается, и процесс приобретает сплошной характер. Важнейшей характеристикой этого реверберационного процесса является время реверберации T_s , т.е. время, в течение которого сигнал затухает на 60 дБ.

Рис.6. Структура реверберационного процесса в помещении

Время реверберации, структура ранних отражений, характер затухания их на последнем этапе и др. вызывают у слушателя субъективное ощущение размеров пространства, полноты звука, ясности, тембра и др. параметров, по которым отличается акустически хороший зал от плохого. (О связях объективных параметров реверберационного процесса и субъективных ощущениях акустики залов получено за последнее время много новых результатов).

Таким образом для определения импульсных характеристик помещения (под импульсной характеристикой понимается реакция системы на воздействие в виде короткого импульса) можно произвести измерения микрофоном в разных точках пространства, а можно, задав геометрические размеры помещения, поглощающие свойства его стен, потолков и др., рассчитать его для любой точки помещения, что и делается в программах аурализации.

- определение бинауральных импульсных характеристик помещения (BRIR) после того как звуковой сигнал, созданный источником звука, изменен помещением (т.е. к прямому сигналу добавлены его отражения), он обрабатывается двумя слуховыми приемниками, и только после этого он поступает на барабанную перепонку и проходит дальнейшие стадии обработки в периферической слуховой системе и в высших отделах мозга.

Импульсные характеристики, которые получают на входе левого и правого слуховых каналов, называются бинауральными импульсными характеристиками помещения BRIR (binaural room impulse response).

Эти функции BRIR несут в себе всю необходимую информацию: о положении и свойствах источника звука, о свойствах помещения и свойствах приемника звука, то есть обо всех процессах обработки звука, которые происходят в голове, ушных раковинах и др.

Для того чтобы описать эти свойства приемника (т.е. головы и ушных раковин), используются передаточные HRTF (АЧХ и ФЧХ) или импульсные функции слуховой системы - HRIR.

Для определения этих передаточных функций обычно используют библиотеки уже выполненных измерений АЧХ и ФЧХ внутри слухового канала на моделях "искусственной головы" в заглушенной камере при разном расположении источника вокруг головы (Рис.7). Вид передаточных функций существенно меняется (особенно в области частот 5...16 кГц) в зависимости от положения источника в вертикальной и горизонтальной плоскости и по глубине по отношению ко входу в левый и правый слуховые каналы головы. Иначе говоря, ушные раковины, голова и торс действуют как частотно-зависимые дифракционные фильтры. Физические причины этого были рассмотрены в статье о бинауральном слухе (№10/1999).

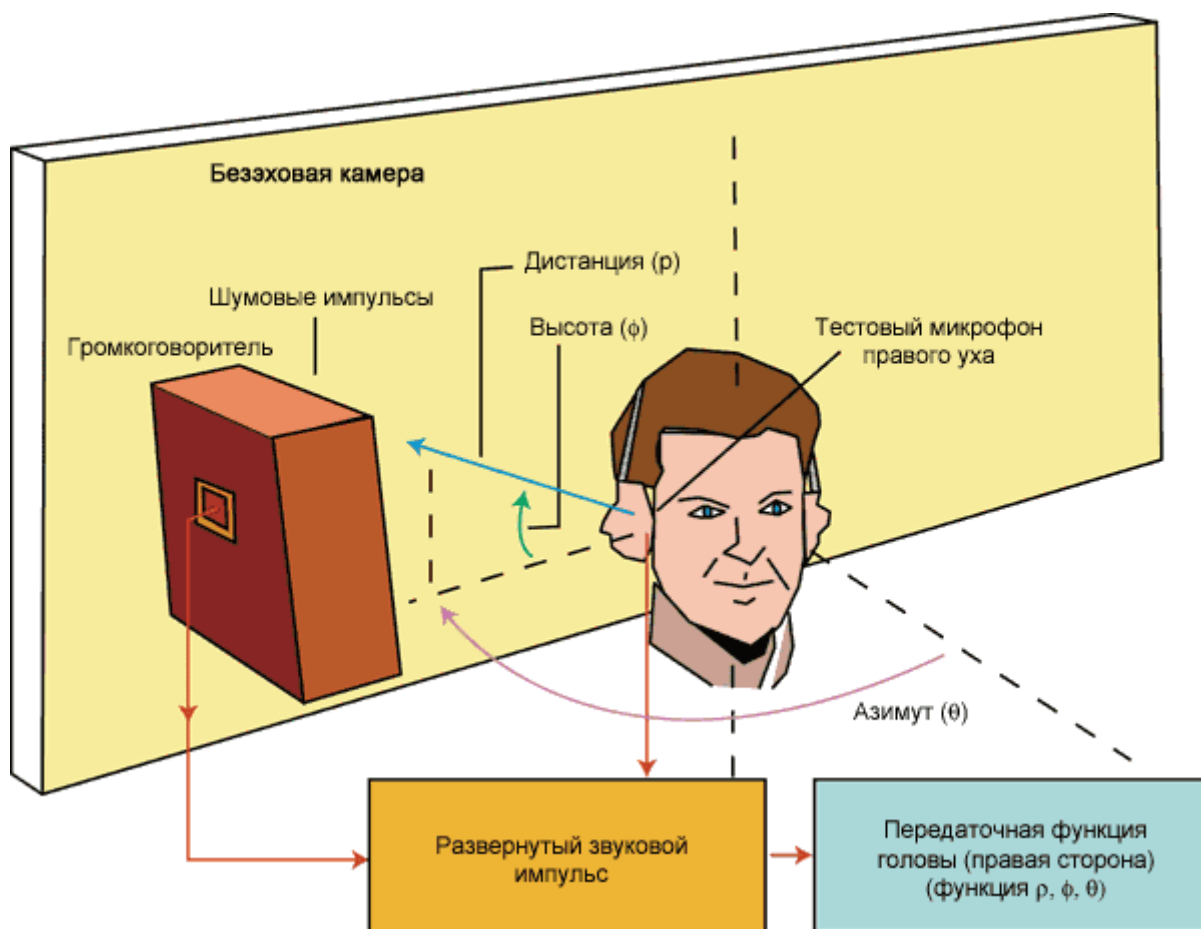


Рис.7. Схема записи передаточных функций слуховой системы при разных положениях источника

В результате для воссоздания пространственного звукового образа необходимо ввести в компьютер следующую информацию:

- выбор источника сигнала: музыка, речь, пение и т.п. Это может быть запись в полузаглушенном или заглушенном помещении или синтезированный сигнал;
- выбор помещения, в которое "помещается" этот источник звука. Могут быть заданы параметры известного помещения или помещения, которое еще предстоит построить;
- положение источника звука в помещении на сцене, на полу, в любой другой точке;
- положение слушателя в каком-либо месте помещения в партере, на балконе и т.д.

Для реализации "виртуального звукового образа" созданы пакеты компьютерных программ. Наиболее известны программы фирм Odeon и CATT, которые последовательно выполняют следующие операции (Рис.8):

- вводят свойства источника звука из библиотеки записанных или синтезированных звуков в моноварианте;
- производят расчет структуры звукового поля в заданном помещении и вычисляют импульсную характеристику в заданных точках расположения правого и левого ушей слушателя;
- используют из заранее составленной библиотеки значения передаточных функций головы, соответствующих данному положению источника и слушателя;
- производят "свертку", т.е. последовательную обработку фильтрацию сигнала источника с помощью импульсных характеристик помещения и импульсных (передаточных) характеристик головы.

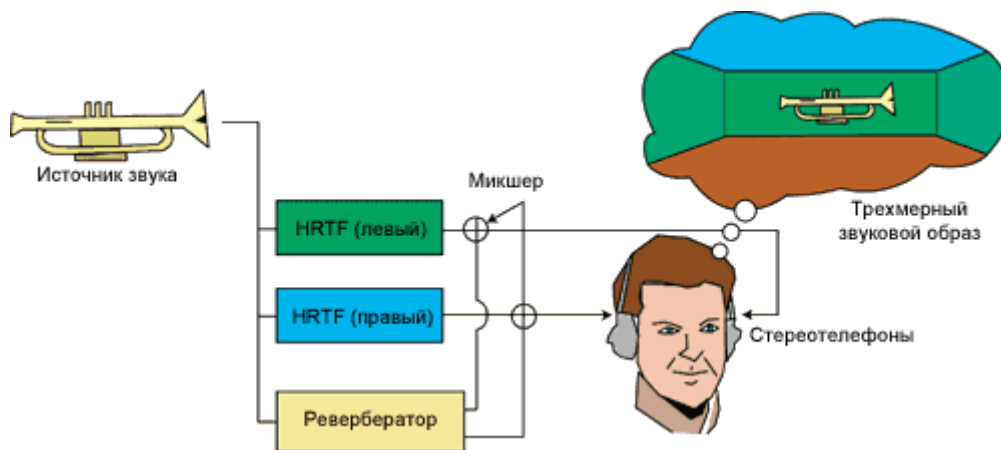


Рис.8. Структура алгоритма

Полученные стереосигналы подаются на головные телефоны, это дает возможность слушателю почувствовать, что он находится на определенном месте

внутри зала, и звук окружает его со всех сторон. При смене положения слушателя или источника производится пересчет передаточных функций.

Как уже было показано в статье, посвященной бинауральному слуху, для нашей слуховой системы существует несколько наиболее важных признаков, по которым она определяет пространственное положение источника. Для локализации в горизонтальной плоскости основное значение имеет разница по времени прибытия сигнала в правое и левое уши, и разница по интенсивности за счет дифракции на голове. Для определения глубины важна разница в уровне звукового давления и разница в спектральном составе, а для локализации в вертикальной плоскости разница в форме АЧХ и ФЧХ за счет дифракции на ушной раковине. Измеренные значения передаточных функций несут в себе всю необходимую для слуховой системы информацию о локализации источника. Разумеется, эти значения передаточных функций сделаны для некоторых усредненных параметров головы и ушных раковин, это вносит определенную погрешность, т.к. каждый человек имеет некоторые индивидуальные особенности. Но, во-первых, исследования показали, что погрешности не слишком велики, а, во-вторых, уже созданы компьютерные модели ушной раковины, в которых можно учесть индивидуальные параметры слушателя. Таким образом, компьютерная модель обработки звука, аналогичная работе бинауральных слуховых приемников, должна включать последовательный ряд следующих моделей:

Моноуральная запись => расчет структуры реверберационного процесса (в т.ч. ранние отражения) => модель локализации в вертикальной плоскости (эхо на ушной раковине) => модель локализации глубины (реверберация, громкость) => азимутальная модель локализации (временная и интенсивностная) => выход на левый и правый канал стереотелефонов.

Техника ауризации стремительно развивается - на последней 108-й конвенции AES в Париже было несколько докладов и демонстраций, посвященных программам ауризации с учетом движения головы (система BRS), которая позволяет при поворотах или подъеме головы пересчитать соответствующие параметры звукового поля и дает возможность услышать в наушниках, как соответственно перемещается источник звука в пространстве. Для этого нужен постоянный мониторинг (система обратной связи), которая отслеживает движение головы и пересчитывает соответствующие бинауральные импульсные характеристики. Существуют разные системы обратной связи, с помощью которых это можно делать: от простейших инфракрасных датчиков, с использованием которых уже давно выпускаются стереотелефоны, до сложных и дорогостоящих систем обратного контроля. Когда эта технология окончательно созреет, качество систем пространственного звуковоспроизведения перейдет на новый уровень.

На 108-й конвенции AES были представлены специальные демонстрационные системы фирмы Studer, когда слушатель мог через мониторы сначала прослушать пространственный звук,

который при этом перемещался от одной акустической системы к другой (в комнате были установлены передние, задние, боковые и центральные системы). Затем с помощью компьютерной системы аурализации производилось прослушивание записей через головные телефоны с системой обратной связи.

При этом можно было услышать полную пространственную картину внешнего окружающего звукового поля, которое перемещалось при повороте головы это действительно впечатляет!

Конечно, хотелось бы использовать в системах аурализации не только головные телефоны, но и иметь возможность прослушивать обработанные записи через акустические системы. Однако для этого надо решить еще несколько дополнительных проблем: во-первых, вторичное помещение накладывает свое реверберационное поле, что в данном случае является помехой, ведь вся необходимая информация о помещении уже закодирована в сигнале. Во-вторых, сигналы от левого канала попадают не только на левое ухо, но и на правое, т.е. возникают перекрестные связи, которые разрушают звуковой образ.

В настоящее время активно развивается техника бинауральной стереофонии, в которой необходимо решение тех же проблем. За последние годы разработаны различные методы проектирования бифонических процессоров, которые реализуют подавление перекрестных связей в реверберирующих помещениях; и адаптивных процессоров, которые могут подавлять отражения во вторичном помещении. Правда, осталась проблема расширения зоны прослушивания, так как пока удастся это сделать для фиксированной позиции слушателя. (О бинауральной стереофонии постараюсь рассказать в дальнейшем). Когда будет достигнут прогресс и этих направлениях, можно ожидать, что появится возможность прослушивания через акустические системы. Следует отметить, что достаточно будет двух акустических систем для левого и правого каналов для воссоздания пространственного звучания это-то и заманчиво!

Новую технологию компьютерного создания пространственных звуковых полей "аурализацию" несомненно, ожидает много применений:

- в архитектурной акустике-для оценки акустических свойств различных существующих залов и моделирования еще не построенных помещений, для оценки влияния различных элементов звукопоглощающих конструкции на качество звучания, в том числе и студий звукозаписи, для проектирования систем звукоусиления и др.;
- в технике звукозаписи открывается много новых возможностей для звукорежиссеров в создании пространственных эффектов, совершенствовании систем пространственной звукопередачи;
- в системах мультимедиа для создания "виртуальных реальностей" как видео-, так и звуковых и не только в компьютерных играх;
- для обучения музыкантов, певцов, артистов для получения возможности предварительного прослушивания различных видов исполнения в залах с разной акустикой;
- для тренировки слепых в распознавании и локализации источников звука;
- в постановке научных экспериментов, в частности, в психоакустике;
- в системах обнаружения и распознавания различных источников сигнала в пространстве (в авиации и др.)

По мере развития этой технологии, несомненно, появятся новые применения. Хотелось бы пожелать, чтобы она нашла себе широкое применение и в отечественной звукотехнике.

