

БИНАУРАЛЬНАЯ СТЕРЕОФОНИЯ НА СОВРЕМЕННОМ ЭТАПЕ РАЗВИТИЯ БИНАУРАЛЬНЫХ ТЕХНОЛОГИЙ

Игнатов П.В., Иванов Ю.М., Алдошина И.А.

НОУ ВПО «Санкт-Петербургский гуманитарный университет профсоюзов», Санкт-Петербург, soundline@fromru.com

Развитие таких пространственных систем записи и передачи звуковых сигналов с сохранением 3-D акустических характеристик, как бинауральные технологии, включающие в себя системы бинауральной стереофонии, трансауральной стереофонии и бинаурального синтеза, дают возможность при воспроизведении только через два излучателя в стереотелефонах или через стереогромкоговорители воссоздать иллюзию виртуального трехмерного звукового пространства с сохранением тембральных и пространственных характеристик музыкальных и речевых сигналов. Это позволяет подойти к вопросам пространственного микширования на новом технологическом уровне, задействовать принципы реализации таких систем в задачах мультимедиа, в интерактивных и аудиовизуальных искусствах. В данной статье рассматривается технологии бинауральной стереофонии, включая принципы ее построения и алгоритмы реализации. Другие направления (трансауральная стереофония и бинауральный синтез) будут рассмотрены в следующих статьях, посвященных бинауральным технологиям.

Ключевые слова: бинауральные технологии, бинауральная стереофония, бинауральная звукозапись, звуковые сигналы, пространственное микширование, интерактивное и аудиовизуальное искусство.

BINAURAL STEREOPHONY AT THE PRESENT STAGE OF DEVELOPMENT OF BINAURAL TECHNOLOGIES

Ignatov P.V., Ivanov Y.M., Aldoshina I.A.

Saint-Petersburg University of Humanities and Social Sciences, Saint-Petersburg, soundline@fromru.com

Development of spatial sound recording systems traced the 3-D acoustic characteristics as binaural technology, including binaural system stereophonic, transaural stereophonic, binaural synthesis enable the playback uses only two emitter in stereo headphones or through stereo speakers to recreate the illusion of three-dimensional virtual sound space preserving timbre and spatial characteristics of music and speech signals. This allows to get approaches of spatial mixing to realize a new technological level, to use the principles of the implementation of such systems in problems of multimedia, interactive and audiovisual arts. This article discusses the binaural stereo and technology, including the principles of its construction and algorithms. Other topics (transaural stereo and binaural synthesis) will be considered in the following articles on binaural technology.

Keywords: binaural technology, binaural to stereo, binaural sound recording, sound signals, spatial mixing, interactive and audio-visual arts.

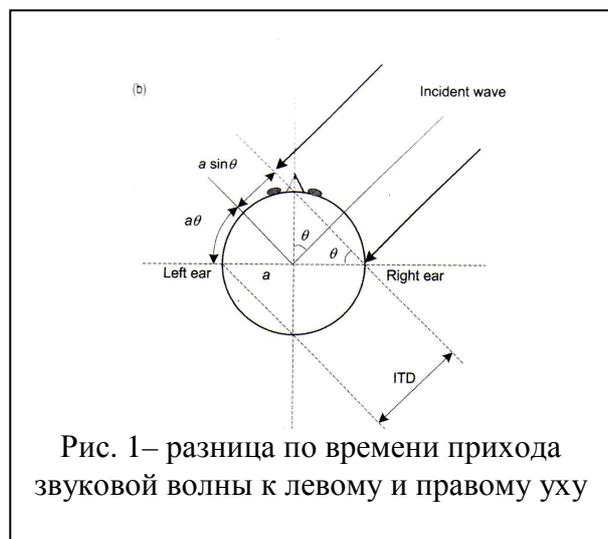
Бинауральные технологии включают программные и аппаратные средства для записи, передачи и воспроизведения пространственного звука [10]. Отличим их от других систем пространственного звучания (Dolby Digital, Ambisonic, Ambiphonic, Wave Field Synthesis и др.) является то, что в них используются звуковые сигналы, записанные (или синтезированные) внутри слухового канала человека. Это позволяет выделить такие параметры сигнала, опираясь на которые слуховая система формирует восприятие трехмерного звукового пространства. Запись и передача звуковых сигналов с сохранением этих параметров дает возможность при воспроизведении только через два излучателя в стереотелефонах или через стереогромкоговорители воссоздать иллюзию виртуального

звукового пространства с сохранением тембральных и пространственных характеристик музыкальных и речевых сигналов.

Бинауральные технологии включают в себя как технику записи сигналов внутри ушных каналов на естественной или «искусственной» голове (см. ниже) и воспроизведение этих сигналов через стереотелефоны и громкоговорители, получившую название «бинауральная стереофония», так и технику бинаурального синтеза этих сигналов из монофонического источника путем процессорной обработки, получившую название *виртуальное слуховое пространство* (дисплей) – virtual auditory display VAD (иногда она называется 3-D Sound) [4].

Развитие бинауральных технологий сдерживалось такими причинами как: невозможностью коллективного прослушивания через стереотелефоны и через громкоговорители, так и отсутствием необходимых цифровых процессорных устройств для обработки звука в реальном масштабе времени. Однако в настоящее время в связи с широким развитием мобильных устройств (телефонов, плееров и др.), передачей высококачественного звука по компьютерным сетям, развитием компьютерных игр, мультимедиа систем виртуальной реальности и др., где предполагается именно индивидуальное прослушивание, а также значительным прогрессом в создании цифровых процессоров, потребность в бинауральных технологиях резко возросла и области их применения существенно расширились. В основе бинауральных технологий лежит использование локализационных возможностей слуховой системы человека. Как известно, в основе локализации звука в горизонтальной, вертикальной плоскости и по глубине лежат следующие явления [2, 3]: разность по времени прихода сигналов на левое и правое ухо (ITD – Interaural time difference), разность по интенсивности сигналов между левым и правым

ухом (IID – interaural intensive difference), изменения спектральных характеристик сигнала в зависимости от направления прихода звука (Spectral cues), изменение уровня сигнала (L) и отношения прямого звука к отраженному в помещении.



Разница по времени прихода (ITD) используется слуховой системой для локализации в горизонтальной плоскости, в основном до частоты 1500Гц (рис. 1) ее значения достигают 0.7мс для расположения источника под углом 90град. Разница по интенсивности (IID), которая возникает за счет дифракции звуковой волны на голове (и торсе) на частотах выше 1500Гц, на частоте 5кГц она достигает значений 17дБ (рис. 2). Эти признаки позволяют локализовать звуковой источник в горизонтальной плоскости с точностью до 3град. Для локализации в вертикальной



Рис. 2 – разница по интенсивности между сигналами на левое и правое ухо на разных

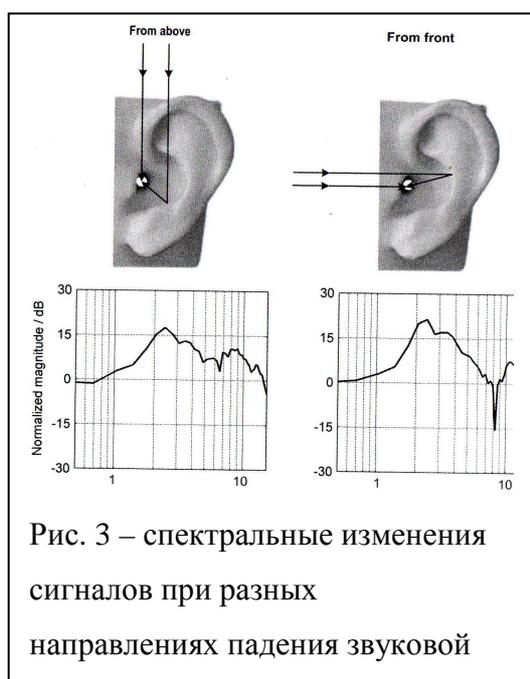


Рис. 3 – спектральные изменения сигналов при разных направлениях падения звуковой

плоскости используются в основном спектральные изменения сигнала за счет дифракции на ушных раковинах, голове и торсе (рис. 3), что позволяет определить расположение источника с точностью до 15 град, и, наконец, для локализации по глубине используются такие признаки, как изменение уровня сигнала, его спектральных характеристик и отношение прямого и отраженного звуков в помещении.

В целом, способность слуховой системы к локализации звуковых источников проявляется в диапазоне от 200Гц до 8кГц.

Все эти признаки содержатся в звуковых сигналах, которые поступают на барабанную перепонку внутри слухового канала правого и левого уха, они затем декодируются мозгом и позволяют ему определить положение звукового источника в трехмерном пространстве. Детальный анализ этих процессов приведен в работах Y.Ando [1].

Все бинауральные технологии основаны на гипотезе, что одно и то же бинауральное звуковое давление в

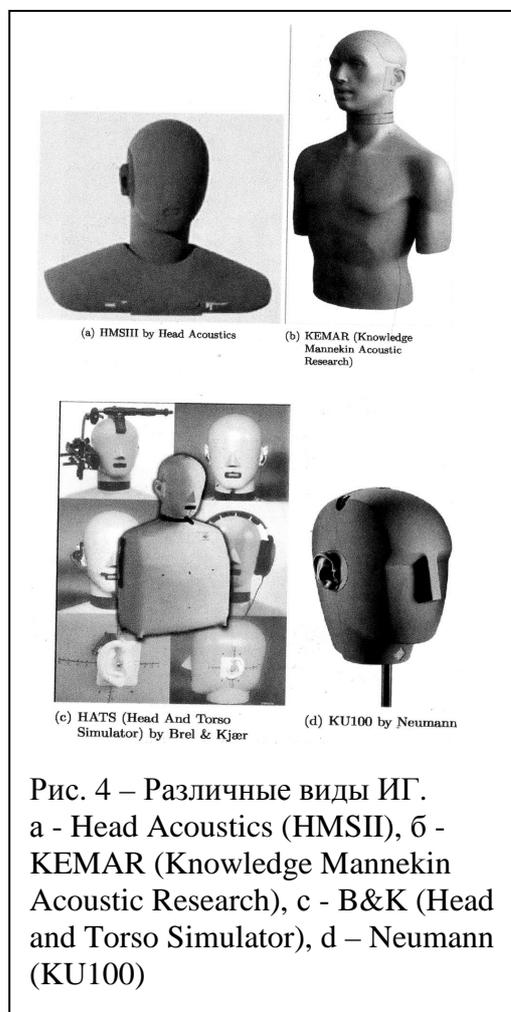


Рис. 4 – Различные виды ИГ. а - Head Acoustics (HMSII), б - KEMAR (Knowledge Mannekin Acoustic Research), в - B&K (Head and Torso Simulator), д – Neumann (KU100)

слуховых каналах создает одно и то же слуховое восприятие звукового пространства, и поэтому их главная задача состоит в том, чтобы записать (или синтезировать) это давление, передать, по возможности, без искажений и воспроизвести в слуховых каналах слушателя.

Для определения звукового давления в левом и правом слуховом канале используется понятие «передаточная функция головы» (Head Related Transfer Function –HRTF), которая определяется как отношение давления, измеренного внутри слухового канала (у барабанной перепонки) на левом P_L и правом P_R ухе к давлению P_0 измеренному в свободном поле в точке, соответствующей центру головы (при ее отсутствии) [4].

$$H_L = \frac{P_L(r, \varphi, \theta, f, a)}{P_0(r, f)} \quad H_R = \frac{P_R(r, \varphi, \theta, f, a)}{P_0(r, f)} \quad (1)$$

Передаточные функции H_L , H_R зависят от угла падения звукового сигнала в горизонтальной – φ и вертикальной плоскости – θ , от расстояния до источника – r , от частоты – f , индивидуальных анатомических особенностей человека – коэффициент – a .

Методика измерения передаточных функций достаточно подробно изложена в работах [8], в частности было показано, что измерения можно производить на входе в ушной канал, а не у барабанной перепонки, пространственные характеристики при этом сохраняются. Измерения могут производиться на «искусственной голове» – ИГ (приборе, имитирующем формы головы и ушных раковин человека со встроенными микрофонами). В настоящее время они выпускаются различными фирмами Neumann, B&K, Head Acoustics и др. Примеры ИГ различных фирм показаны на рис. 4.

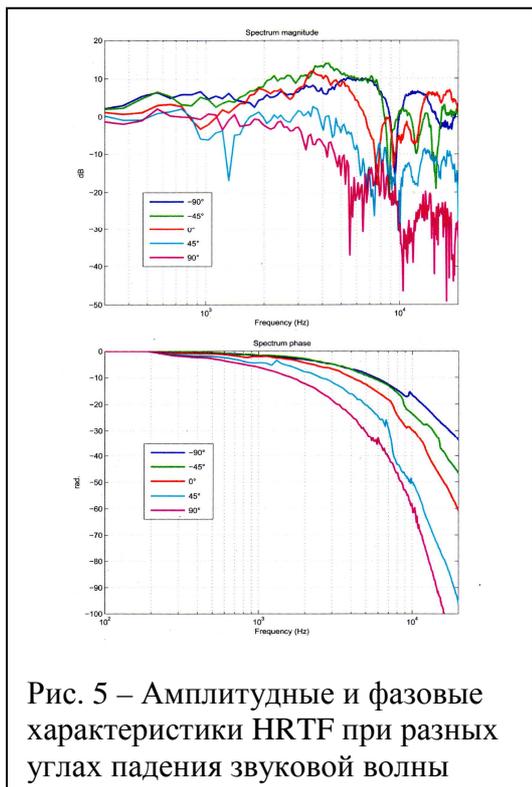


Рис. 5 – Амплитудные и фазовые характеристики HRTF при разных углах падения звуковой волны

Примеры передаточных функций для разных углов падения звуковой волны показаны на рис.5. Если для измерения используется импульсный сигнал, то в слуховом канале измеряются импульсные характеристики – HRIR (Head Related Impulse Responses) (рис. 6), которые могут пересчитываться в HRTF с помощью преобразования Фурье.

По данному выше определению (1) [4, 3] передаточные функции головы HRTF измеряются в свободном поле (т.е. в заглушенной камере), следовательно, в них не содержится информация о реверберационной структуре поля в помещении. Для учета этого влияния измерения передаточных функций должны производиться в реальных

помещения прослушивания, в таком случае они называются BRTF (Binaural Related Transfer Function) и BRIR (Binaural Related Impulse Responses).

Использование передаточных функций лежит в основе таких направлений в бинауральных технологиях как бинауральная (трансауральная) стереофония, бинауральный синтез пространственного звука, создание систем виртуальной реальности, создание аурализации моделей помещения и др.

Остановимся на некоторых из них более подробно:

Бинауральная стереофония

Бинауральная стереосистема – это

пространственная система звукозаписи и звукопередачи, при которой звуковые сигналы записываются внутри слухового канала, передаются по каналам связи и воспроизводятся в той же точке ушного канала, где они были записаны.

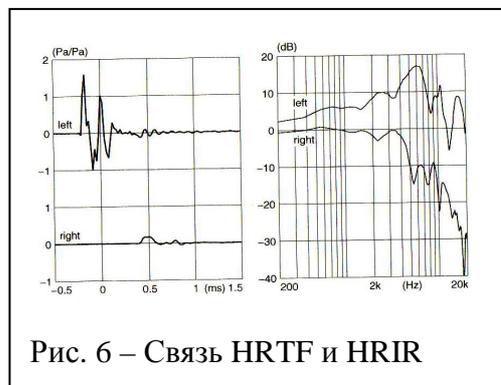
При этом к слушателю поступает полная информация о пространственном расположении источника (источников) звука, что позволяет воссоздать трехмерный звуковой образ.

Практическая реализация этой системы выполняется следующим образом, звуковые сигналы записываются обычно с помощью «искусственной головы», передаются и воспроизводятся через стереотелефоны.

Попытки создания таких систем начались еще в 1881г. (опыты С. Adera), продолжались в лабораториях Белла (Bell Laboratories) в 40г, в 80г появилось достаточно много записей по этой системе, однако она не получала широкого распространения, т.к. предназначалась только для индивидуального прослушивания. История развития бинауральной стереофонии и перспективы ее развития детально проанализированы в работе Paul S. [9].

Оценка качества звучания таких систем показала их несомненные преимущества перед обычными стереосистемами по таким параметрам как: пространственность (включая оценку расположения источников в вертикальной плоскости, перемещения источников и др.); прозрачность (разделимость инструментов); натуральность тембров; правильное восприятие прямого и отраженных звуков в помещении (с учетом эффектов «первой волны») и др.

Первые эксперименты по оценке качества бинауральных систем проводились в 1931-1932 годах в лабораториях А. Белла под руководством Н. Fletcher'a.



В экспериментах принимал участие симфонический оркестр Филадельфии под управлением Леопольда Стоковски. Во время концертов и репетиций была организована трансляция из концертного зала в фойе и специальные помещения для прослушивания. Звукозапись осуществлялась на систему микрофонов, установленных на подвесах, и бинаурально, т.е. на искусственную голову (модель «Oscar»); воспроизведение проводилось через громкоговорители и головные телефоны. Слушатели и эксперты при прослушивании бинауральных записей отмечали высокую реалистичность звукопередачи, точную передачу акустических характеристик концертного зала, что создавало эффект присутствия [7]. Сам Н. Fletcher отмечал так же эффект «окружения звуком» (auditory perspective), что обозначал как «трехмерный звук в головных телефонах» [5].

Особый интерес представляет исследование А. Farina «Binaural for popular music» [6], целью которого было изучение реакции слушателей на различные методы звукозаписи. В рамках исследования были проведены экспериментальные прослушивания звукозаписей в стереофоническом и бинауральном форматах. Музыкальный материал был записан в студии на микрофонную систему Neumann-KU100 и пару микрофонов Schoeps МК-2, установленных по системе АВ. Звукозаписи оценивались по следующим критериям: пространственное впечатление, локализация источников звука, натуральность передачи тембров и воспринимаемая ширина источника (ASW). Бинауральная звукозапись по результатам экспертизы получила более высокие средние оценки: 4.5 за пространственное впечатление, что на 2 балла выше, чем у стереофонической записи (оценки выставлялись по пятибалльной шкале).

Результаты исследований, представленные в работе F. Wanderley, J. Sousa [11] также подтвердили достоинства бинауральной звукозаписи при воспроизведении через стереотелефоны. Материалы для эксперимента были записаны в формате Ambisonics B-format на микрофон Soundfield ST-250. Записанные сигналы (были выполнены записи симфонического оркестра с хором и фортепианного соло в концертном зале) декодировались как для стереофонической, так и бинауральной звукопередачи. Результаты слушательских тестов свидетельствуют о преимуществах бинауральной звукопередачи над стереофонией по таким параметрам как ширина виртуального источника (на 30%), глубина источника (на 20%), передача «объема» концертного зала (на 20%). Результаты приведены в процентах от общего числа оценок. Также при воспроизведении бинаурального материала эксперты отмечали высокую точность в локализации источников звука и различимость инструментальных групп.

В настоящее время в связи с развитием мобильных устройств (как указано выше) интерес к таким системам вырос, и появились технические возможности для решения

основных проблем. К числу наиболее существенных из них можно отнести несоответствие условий записи и воспроизведения звука. Идеальное воссоздание пространственного звукового поля возможно только, если запись сигнала производилась у барабанной перепонки слушателя и на его же голове воспроизводилась, причем сигналы должны быть доставлены в ту же точку слухового канала, где они были записаны, технически это сделать практически невозможно. При практической реализации бинауральной стереосистемы возникают различные искажения сигнала, которые могут приводить к ошибкам типа «фронт-тыл», локализации звука внутри головы, недостоверности воспроизведения тембров и др.

Причины этих ошибок заключаются в следующем:

- микрофоны, тракты передачи и стереотелефоны вносят свои искажения в звуковой сигнал, особенно значительный вклад вносят стереотелефоны (влияет выбор их характеристик, расположение на голове слушателя, дополнительная нагрузка на ушной канал, создаваемая ими и др.);
- при записи на искусственной голове не учитываются индивидуальные особенности ушных раковин, а они имеют значительные отличия и могут вносить искажения;
- отсутствует учет движения головы, который в реальной ситуации помогает избежать ошибок типа «фронт-тыл»;
- отсутствуют (или не соответствуют слуховым) визуальные признаки.

Для решения этих проблем в настоящее время используются следующие методы: как было показано в многочисленных исследованиях [4] пространственная информация сохраняется даже, если запись сигналов производилась не у барабанной перепонки, а в любой точке слухового канала, в частности на входе слухового канала (даже если слуховой канал при этом блокирован), это существенно упрощает процедуру записи на голове (искусственная голова может быть сделана вообще без ушных каналов) и процедуру воспроизведения, т.к. можно воспроизводить сигнал на входе в слуховой канал (на барабанной перепонке его невозможно воспроизвести с помощью стереотелефонов). При этом при записи на заблокированных ушных каналах, индивидуальные отличия формы ушных раковин влияют меньше. Стереотелефоны вносят искажения в слуховой канал, во-первых, потому что их излучатели имеют не плоскую АЧХ, во-вторых, они находятся не в той точке слухового канала, где производилась запись, в-

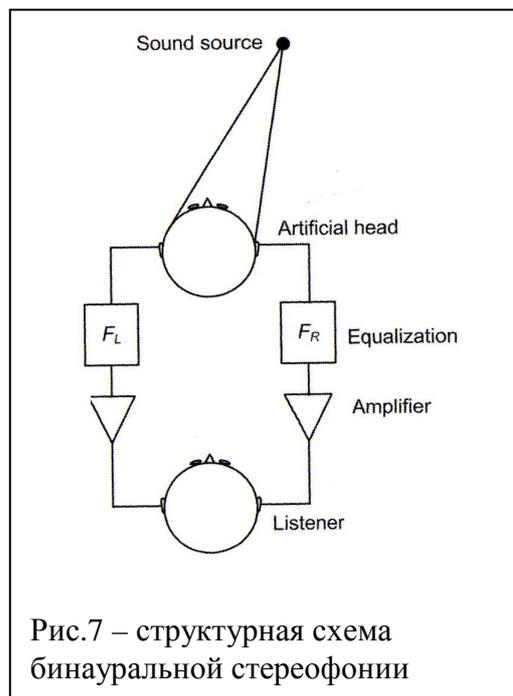
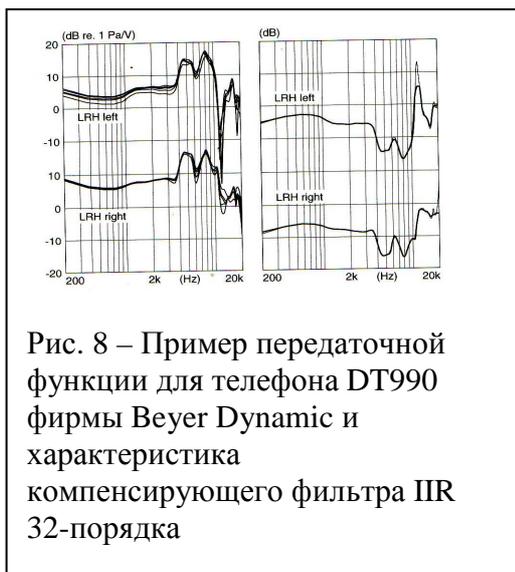


Рис.7 – структурная схема бинауральной стереофонии

третьих, они создают нагрузку на слуховой канал, которая отличается от свободного пространства (т.е. меняется импеданс излучения со стороны слухового канала).

Для ликвидации этих искажений необходимо измерить передаточную функцию стереотелефона H_pTF (Headphone Transfer Function), которая есть отношение звукового давления $P(f)$, измеренного в той же точке слухового канала, где была сделана запись, к напряжению - E_v - подводимому ко входу стереотелефона $H_pTF = P_{сл.кан} / E_v$.

Полученная таким образом передаточная функция должна быть компенсирована, т.е. построена H_pTF^{-1} и умножена на H_pTF (метод инверсной фильтрации [11]). Общая схема



бинауральной записи и воспроизведения с калибровкой стереотелефонов показана на рис. 7.

Пример передаточной функции для телефона DT990 фирмы Beyerdynamic и характеристика компенсирующего фильтра IIR 32-порядка показана на рис. 8. Наилучшие результаты в смысле экстернализации (выноса звукового образа из головы) достигаются, если калибровка передаточных функций стереотелефонов происходит индивидуально для данного слушателя (что на практике реализуется редко). Выбор типа и позиции стереотелефонов также

может существенно уменьшить погрешности воспроизведения: телефоны должны быть открытыми, в том смысле, чтобы нагрузка, создаваемая со стороны телефона на слуховой канал (т.е. его импеданс) не сильно отличается от нагрузки без телефонов. Для оценки этого вводится понятие PDR (pressure division ratio) отношение давление у входа в слуховой канал без телефона к давлению при наличии телефона (не включенного), оно должно быть близко к единице. Примеры таких телефонов есть, среди промышленно выпускаемых, например, DT990 фирмы Beyerdynamic [10, 4]. Позиция телефонов на ушных раковинах при прослушивании может меняться и это влияет на их передаточные функции H_pTF , особенно на высоких частотах. Менее чувствительными к изменению позиции являются охватывающие телефоны (circum aural). Выбор типа «Искусственной головы», которые сейчас производятся разными фирмами показывает [4], что чем ближе ИГ к структуре человеческой головы, тем меньше ошибок в локализации при прослушивании бинауральных записей, поэтому конструкции ИГ все время усложняются. Индивидуальность HRTF и H_pTF оказывает существенное влияние на ошибки, возникающие при локализации источников звука, особенно в вертикальной плоскости, в определении фронт-тыл и в экстернализации источника [10]. При прослушивании бинауральных записей музыки через стереотелефоны,

они не так критичны, но при создании систем виртуальной реальности с помощью бинаурального синтеза они оказывают существенное влияние, методы их учета рассмотрены в работе [10]. Следует отметить, еще один момент – слушатель воспринимает звуковое пространство только с тех позиций, на которые была ориентирована голова (естественная или искусственная) в первичном помещении. Если источник в первичном помещении находился точно по оси, то он только с этого направления будет восприниматься и во вторичном помещении, при этом при повороте головы слушателя с телефонами источник будет поворачиваться вместе со слушателем, и будет постоянно находиться на оси. Если нужно передать восприятие источника (и зала) с другого направления необходимо для нового направления записать и передать другие передаточные функции. В экспериментах, выполненных в институте IRT (Германия), где повороты искусственной головы при записи были синхронизированы с поворотами головы слушателя во вторичном помещении, было отмечено существенное уменьшение ошибок типа «фронт-тыл».

Таким образом, развитие пространственных систем записи и передачи звуковых сигналов с сохранением 3-D акустических характеристик дает возможность при воспроизведении только через два излучателя в стереотелефонах или через стереогромкоговорители воссоздать иллюзию виртуального звукового пространства с сохранением тембральных и пространственных характеристик музыкальных и речевых сигналов. Это позволяет подойти к вопросам пространственного микширования на новом технологическом уровне, задействовать принципы реализации таких систем в задачах мультимедиа, в интерактивных и аудиовизуальных искусствах.

Список литературы

1. Ando Y. Audio-Video interection. Springer, Berlin, 2009.
2. Blauert J. Communication Acoustics .Springer, Berlin, 2005.
3. Blauert J. Spatial Hearing. The Mit Press, Cambridge, 1997.
4. Bosun Xie. Head-Related Transfer Function and Virtual Auditory Display. J.Ross Publishing, USA, 2013.
5. Fletcher H. 1884-1981 a bibliographical memoir by S.Flethcer. Memoir, 1992. – P. 20-21.
6. Fontana S., A. Farina, Y.Grenier. Binaural for popular music: a case of study. Proceedings of the 13th International Conference on Auditory Display, Montreal, Canada. 2007.
7. Hammer K., Snow W. Binaural transmission system at academy of music on Philadelphia. Memorandum MM-3950, Bell laboratories, Nov. 1932.

8. Moller H., Hammershoi D., Jensen C. Transfer characteristics of headphones measured on human ears. JAES 43(4), 203-217, 1995.
9. Paul S. Binaural recording technology: a historical review and possible future development. Acta Acustica unitedwith Acustica 95(5)767-788.
10. Rozenn N. Binaural Technology. Audio Eng.Society, N.Y., 2010.
11. Wanderley F., Sousa J. Subjective comparison between stereo and binaural processing from B-format Ambisonic raw audio material. AES, convention paper 8374, presented at 130th convention, may13-16, 2011, London, UK.

Рецензенты:

Денисов А.В., доктор искусствоведения, профессор кафедры теории и истории культуры Российского государственного педагогического университета им. Герцена, г. Санкт-Петербург;

Соломонова Н.А., доктор искусствоведения, профессор кафедры звукорежиссуры Негосударственного образовательного учреждения высшего профессионального образования «Санкт-Петербургский гуманитарный университет профсоюзов», г. Санкт-Петербург.