

---

УДК 534.843.742

С.М. Порошин, И.С. Беликов

Национальный технический университет "ХПИ", Харьков

## РАЗРАБОТКА АКТИВНОЙ СИСТЕМЫ КОРРЕКЦИИ МЕСТОПОЛОЖЕНИЯ КАЖУЩЕГОСЯ ИСТОЧНИКА ЗВУКА В ПРОСТРАНСТВЕ

*В статье предложена структурная схема системы автоматизированного управления перемещения кажущегося источника звука в пространстве. Рассматривается возможность расширения стереофонии согласно перемещению человека в зоне прослушивания.*

*Ключевые слова:* акустическая система, стереофония, кажущийся источник звука, OpenNI, Microsoft Kinect.

### Введение

**Постановка проблемы.** В современных мультимедийных акустических системах (АС) на сегодняшний день не задействованы следующие возможности:

- распознавания и слежение за перемещением человека в пространстве;
- влияние временной и интенсивностной стереофонии на кажущийся источник звука;
- интерактивное взаимодействие человека и АС.

**Целью данного исследования является** разработка метода автоматического управления параметрами АС с целью коррекции местоположения кажущегося источника звука в пространстве.

**Анализ достижений по теме исследований.** Современные системы распознавания образа человека в пространстве, такие как OpenNI и OpenCV позволяют в условиях реального времени получать данные о местоположении человека в пространстве и следить за его перемещением. При отклонении

человека от оптимальной оси направленности громкоговорителя (ГГ), изменяется воображаемое расположение источников звука в помещении, где происходит озвучивание. Данные о местоположении головы человека в пространстве, поступающие в виде системы трёхмерных координат с камеры Microsoft Kinect, используются для коррекции параметров обработки звука в акустическом тракте [1].

Внесение корректировок в амплитудно- и фазо-частотную характеристику акустической системы в реальном времени даёт возможность человеку постоянно находиться в оптимальной зоне прослушивания при перемещении в комнате, где он прослушивает мультимедийный контент.

Позиционирование КИЗ при озвучивании пространства рассчитывается исходя из свойств человеческого слуха. При расположении слушателя напротив центра стереобазы (точка О на рис. 1) АС, воспроизводимые сигналы с правого громкоговорителя (Гр1) и левого громкоговорителя (Гр2) не имеют различий по времени ( $\Delta t=0$ ) и по уровню ( $\Delta L=0$ ).

При таком условии звучание обоих громкоговорителей сливается в единый звуковой образ, который соответствует изначально задуманному звукорежиссером звучанию. В процессе слияния образуется КИЗ [2].

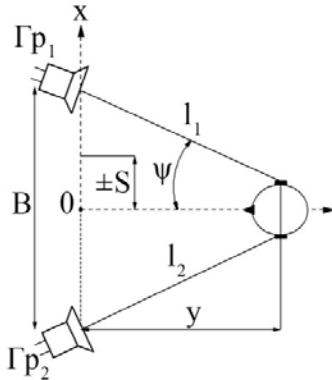


Рис. 1. Расположение слушателя относительно базы АС: В – база АС; О – центр стереобазы и место расположения камеры Kinect; ±S – диапазон смещения КИЗ от центрального положения; у – расстояние от центра базы АС до головы слушателя; l<sub>1</sub> – расстояние от головы слушателя до правого громкоговорителя Gr<sub>1</sub>; l<sub>2</sub> – расстояние от головы слушателя до левого громкоговорителя Gr<sub>2</sub>

Положение КИЗ на линии базы громкоговорителей зависит от временных (Δτ) и интенсивностных (ΔL) различий между сигналами левого и правого каналов АС, достигающих ушей слушателя (рис. 1). Применение интенсивностной и временной стереофонии расширяет возможности воспроизведения звука акустической системы. Внесение задержки во времени в один из каналов акустической системы, также как и уменьшение уровня звукового сигнала, ведёт к смещению местонахождения кажущегося источника звука в противоположную сторону от слушателя.

Величина уровня сигнала ΔL для левого (L) и правого (R) каналов определяется по формуле (1).

$$\Delta L_{L,R} = 10 \lg \frac{l_2(x,y)D_2(\psi)}{l_1(x,y)D_1(\psi)}, \quad (1)$$

где l<sub>1</sub> – расстояние до первого громкоговорителя в координатах x, y; l<sub>2</sub> – расстояние, соответственно, до второго громкоговорителя; D<sub>1</sub>, D<sub>2</sub> – амплитудная характеристика громкоговорителя относительно угла отклонения ψ от оси АС.

Расстояние от головы слушателя до Gr<sub>1</sub> и Gr<sub>2</sub> определяется следующим образом:

$$l_{1,2}(x,y) = \sqrt{\left(\frac{B}{2} \pm S(x)\right)^2 + y^2}, \quad (2)$$

где x, y – координаты слушателя в пространстве, В – размер установленной базы АС, S(x) – смещение

местоположения относительно центра стереобазы влево-вправо, у – расстояние отдаления слушателя от центра стереобазы.

Величина задержек Δτ, вносимых в левый и правый канал соответственно, определяется взаимным местоположением громкоговорителей и слушателя в пространстве x, y (формула (2)).

$$\Delta\tau_{L,R} = \frac{l_1(x,y) - l_2(x,y)}{c}, \quad (3)$$

где c – скорость звука, l<sub>1</sub>, l<sub>2</sub> – расстояние до соответственно первого и второго громкоговорителя в координатах x, y.

Уменьшение амплитуды сигнала АС<sub>1</sub> будет воспринято слухом как перемещение КИЗ в сторону АС<sub>2</sub> и наоборот. Таким образом, варьируя громкость звучания левой и правой АС, можно вызывать и поддерживать иллюзию перемещения виртуального источника звука. Это явление называют интегральной локализацией (или локализацией суммы).

### Основная часть

Разработана структурная схема автоматизированной системы корректировки местоположения КИЗ в пространстве между каналами АС и головой слушателя (рис. 2).

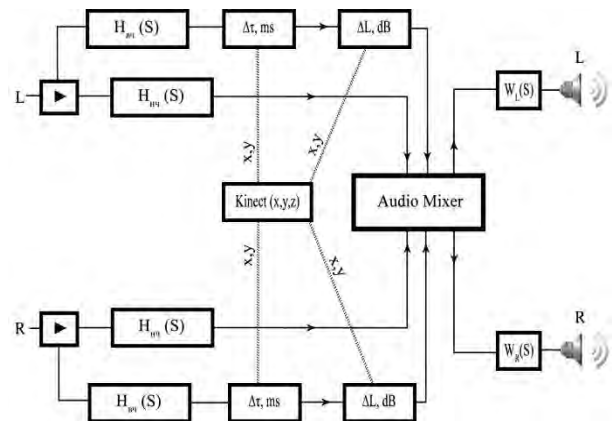


Рис. 2. Структурная схема автоматизированной коррекции перемещения КИЗ в пространстве

Это необходимо для усиления эмоционального восприятия человека, при прослушивании аудио материала, а также при непосредственном интерактивном участии слушателя в создании акустической картины мультимедийного контента.

При маршрутизации стереосигнал разделяется на две составляющих по частотам. Бинауральный эффект восприятия звуков человеческом не позволяет точно определять источник КИЗ в пространстве при воспроизведении низких частот (НЧ). Это обусловлено большой длиной волны звуковых колебаний, а также с особенностью строения человеческого слуха.

На частотах от 300 Гц и выше становится заметным сдвиг фаз звуковых волн, попадающих в

правое и левое ухо. В данной работе предлагается вводить временные задержки  $\Delta\tau_{x,y}$  и изменение интенсивности уровня сигнала  $\Delta L_{xy}$  в область частот от 300 Гц и выше.

На рис. 2 представлена маршрутизация акустического стереосигнала в системе по корректированию местоположения КИЗ относительно перемещения слушателя.

Наибольшая точность локализации достигается при восприятии сложных звуков и звуковых импульсов. При этом важным фактором является спектральный состав звуков. Наименьший ощутимый угол восприятия отклонения источника звуковых импульсов равен  $3^\circ$ . Эту величину следует считать разрешающей способностью слуха для фронтального направления. Точность локализации источников звука, расположенных слева и справа, значительно меньше и составляет примерно  $12^\circ$ . Для тылового направления эта величина около  $6^\circ$  [4].

В процессе обработки из акустического сигнала (рис. 2) отфильтровывается две части фильтрами нижних (НЧ) и верхних частот:

- область нижних частот (от 20 Гц до 300 Гц);
- область рабочих частот (от 300 Гц и выше).

Индивидуально для левого и правого каналов в области рабочих частот (от 300 Гц и выше) выполняется последующая корректировка временных и интенсивностных характеристик. В область нижних частот не вносятся каких-либо изменений, так как человеческий слух обладает низкой степенью локализации источника звука с низкой частотой колебаний. Отфильтрованная НЧ часть поступает без изменений на финальный аудио микшер (Audio Mixer на рис. 2). Отфильтрованная область частот от 300 Гц и выше далее в статье будет называться областью рабочей частоты [2, 3].

Фильтр нижних частот ( $H_{НЧ}(S)$ ) на рис. 2) представляет собой устройство, пропускающее сигналы низких частот и подавляющее сигналы высоких частот [4]. Передаточная функция фильтра низких частот имеет вид:

$$H_{НЧ}(S) = \frac{K C \omega_c^2}{s^2 + E \omega_c s + C \omega_c^2}, \quad (4)$$

где  $\omega_c$  – частота среза, В и С представляют собой нормированные коэффициенты, поскольку для  $\omega_c = 1$  эта передаточная функция приводится к виду (5) при  $n=2$ .

$$H_{НЧ}(S) = \frac{K b_c}{s^n + b_{n-1} s^{n-1} + \dots + b_1 s + b_n}. \quad (5)$$

Фильтр верхних частот ( $H_{ВЧ}(S)$ ) на рис. 2) представляет собой блок, пропускающий сигналы высоких частот и подавляющий сигналы низких частот. Передаточную функцию фильтра верхних частот с частотой среза  $\omega_c$  можно получить из передаточной

функции нормированного фильтра нижних частот (имеющего  $\omega_c$ , равную 1 рад/с) с помощью замены переменной  $s$  на  $\omega_c / s$ . Следовательно, функция фильтров верхних частот Баттерворта и Чебышева будет содержать следующие сомножители второго порядка:

$$H_{ВЧ}(S) = \frac{K_s^2}{s^2 + \left(\frac{B \omega_c}{C}\right) s + \frac{\omega_c^2}{C}}, \quad (6)$$

где  $\omega_c$  – частота среза фильтра, а В и С представляют собой нормированные звена фильтра частот второго порядка [11].

В область рабочей частоты сигнала вносятся временные задержки, исходя из данных об удалении головы слушателя от центрального местоположения стереобазы. На рис. 2 этот процесс указан в блоке  $\Delta t$ . Данные о местоположении слушателя в пространстве обозначаются как  $x, y$ . Согласно рассчитанным данным, в аудио канал громкоговорителя, к которому человек расположен ближе, вносится временная задержка в мс. Величина  $\Delta t$  рассчитывается по формулам (2) и (3). Сигнал из противоположного громкоговорителя воспроизводится без задержки во времени. Это даёт возможность сместить КИЗ обратно в центр акустической базы, не нарушая целостности акустического восприятия материала.

Интенсивностная стереофония (блок  $\Delta L$  на рис. 2) позволяет скорректировать отклонение КИЗ путём усиления уровня сигнала канала, который подвергается маскированию, относительно местоположения головы слушателя. При приближении слушателя к одному из громкоговорителей, КИЗ смещается в сторону громкоговорителя, возле которого расположился слушатель. Согласно формулам (1) и (3), применение интенсивностной стереофонии позволяет скомпенсировать отклонение КИЗ путём ослабления уровня сигнала, находящегося ближе к слушателю.

Получение выходного сигнала выполняется путём микширования ранее отфильтрованной низкочастотной составляющей и диапазона частот, в которые вводились корректировки (Audio Mixer на рис. 2).

Аудио сигнал на выходе из АС позволяет слушателю ощущать КИЗ, в соответствии с задумкой звукорежиссера на этапе создания звуковой композиции, независимо от своего местоположения относительно ГГ в помещении.

## Выводы

Авторами ведется разработка математической модели и экспериментальных расчётов для дальнейшего исследования передаточных функций акустических систем и корректирующих фильтров.

## Список литературы

1. Catuhe D. *Programming with the Kinect for Windows [Текст]* / D. Catuhe. – Microsoft Press, Redmond, Washington, 2012. – 224 с.

2. Кононович Л.М. *Стерефоническое восприятие звука [Текст]* / Л.М. Кононович, Ю.А. Ковалгин. – М.: Радио и связь, 1981. – 184 с.

3. Алдошина И.А. *Электроакустика и звуковое вещание: Учебное пособие для вузов [Текст]* / И.А. Алдошина, Э.И. Вологдин, А.П. Ефимов и др.; под ред. Ю. А. Ко-

валгина. – М.: Горячая линия-Телеком, Радио и связь, 2007. – 872 с.

4. Алябьев, С.И. *Радиовещание и электроакустика [Текст]* / С.И. Алябьев, А.В. Выходец, Р. Гермер и др. – М.: Радио и связь, 1998. – 783 с.

Поступила в редколлегию 4.06.2014

**Рецензент:** д-р техн. наук, проф. Ю.В. Стасев, Харьковский университет Воздушных Сил им. И. Кожедуба, Харьков.

## РОЗРОБКА АКУСТИЧНОЇ СИСТЕМИ КОРЕКЦІЇ МІСЦЕПОЛОЖЕННЯ УЯВНОГО ДЖЕРЕЛА ЗВУКУ У ПРОСТОРІ

С.М. Порошин, І.С. Біліков

У статті розглядається розробка методу використання активної стереофонії при інтерактивному прослуховуванні аудіо контенту.

**Ключові слова:** акустична система, стереофонія, уявне джерело звуку, OpenNI, Microsoft Kinect.

## DEVELOPMENT OF ACOUSTIC CORRECTION SYSTEM OF LOCATION APPARENT SOUND SOURCE IN SPACE

S.M. Poroshin, I.S. Bielikov

This article discusses how to enhance the interactivity of human involvement in formatting apparent sound source in space in real time.

**Keywords:** acoustic system, stereophony, apparent sound source, OpenNI, Microsoft Kinect.