

МЕТОД ПРЕОБРАЗОВАНИЯ СТЕРЕО ЗВУКА В МНОГОКАНАЛЬНЫЙ

METHOD OF CONVERTING STEREO SOUND TO MULTI-CHANNEL

*M. Gusev
M. Pahomov
V. Rozhnov
N. Treschalov
I. Tolokonnikov*

Summary. Discusses the limitations of the upmixing methods, used in many systems playback multi-channel sound. Provides the method of formation of spatial sound based on the localization of virtual sound sources in the original stereo signal, orienting and positioning them in the sound scene generated by any number of speakers (system 5.1 and higher). The results of expert evaluation of sound generated according to the proposed method and sound systems based on the algorithms Pro Logic and Pro Logic II, confirm the effectiveness of the method developed by the authors.

Keywords: spatial audio, virtual sound source, the expansion of sound scene, modeling the reverberations.

Технологии объёмного звука обязаны своим появлением, прежде всего, развитию звукового кино. Разработчики новых систем кинематографа старались повысить зрелищность не только за счёт увеличения экрана, но и за счёт многоканального звуковоспроизведения, когда направление звука совпадает с изображением его источника. Многоканальные аудиосистемы не только обеспечивают более реалистичные звуки, но также предлагают больше внешних эффектов, чем стандартные стереофонические аудиосистемы.

Число пользователей систем объёмного звучания быстро растёт. К сожалению, в формате 5.1 существует не так много музыкальных записей — большинство записей объёмного звучания создаются кинокомпаниями или производителями видеоигр. Обычно доступен аудиоконтент с количеством каналов меньшим, чем может быть воспроизведено целевой системой.

Сегодня почти любой усилитель системы домашнего кинотеатра имеет возможность воспроизведения стерео для всех динамиков. При этом во многих системах используются простые алгоритмы восходящего микширования, основанные на смешивании каналов исходного стереозвучания в разных пропорциях, иногда дополняемые внесением задержек и применением простейших фильтров.

*Гусев Михаил Николаевич
Пахомов Михаил Олегович
Рожнов Виктор Сергеевич
Трещалов Никита Олегович
Толоконников Иван Сергеевич*

*Филиал корпорации LG Electronics Inc. (Республика Корея) в Санкт-Петербурге, Россия
mikhail.gusev@lge.com*

Аннотация. Рассматриваются ограничения методов восходящего микширования, используемых во многих системах воспроизведения многоканального звука. Предлагается метод формирования пространственного звука, основанный на выделении виртуальных источников звука в исходном стереосигнале, определении направлений на них и позиционировании их в звуковой сцене, формируемой произвольным количеством динамиков (системы 5.1 и выше). Приводятся результаты сравнительной экспертной оценки звука, сформированного согласно предложенному методу и звуковыми системами на основе алгоритмов Pro Logic и Pro Logic II, подтверждающие эффективность разработанного авторами метода.

Ключевые слова: пространственный звук, виртуальный источник звука, расширение звуковой сцены, формирование ревербераций.

Методы восходящего микширования

Рассмотрим типичные подходы к созданию дополнительных каналов, основанные на использовании корреляции между стереоканалами: метод пассивного объёмного декодирования (Passive Surround Decoding Method) [7], метод наименьших квадратов (LMS-based Upmixing Method) [1], метод на основе анализа главных компонент [1] и метод адаптивного панорамирования [2].

На рисунке 1 показана процедура восходящего микширования, где каналы обозначены как FL (передний левый), FR (передний правый), C (центральный), LFE (канал низкочастотных эффектов), RL (задний левый) и RR (правый задний). Как проиллюстрировано на рисунке, FL и каналы FR для 5.1-канального аудиоформата получают непосредственно из исходных стереоканалов, тогда как остальные каналы генерируются из них путем различных преобразований. Поэтому в следующих подразделах описывается, как получить центральный канал и каналы объёмного звучания с использованием каждого метода восходящего микширования.

Метод пассивного пространственного декодирования

Метод пассивного пространственного декодирования (Passive Surround Decoding Method, PSD) является

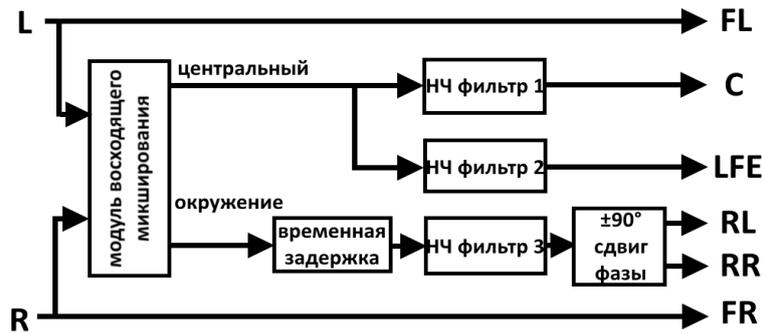


Рис. 1. Процедура восходящего микширования стереозвука в звук 5.1

ранней версией декодера Dolby Surround [7]. В этом методе центральный канал получается путем сложения отсчетов исходных левого и правого каналов, а звук тыловых динамиков — путем вычитания отсчетов правого канала из отсчетов левого канала. Для поддержания энергии сигнала на постоянном уровне, уровни центрального и задних каналов снижаются на 3 дБ, что реализуется путем деления на $\sqrt{2}$:

$$Center(n) = \frac{x_l(n) + x_r(n)}{\sqrt{2}}, \quad (1)$$

$$Surround(n) = \frac{x_l(n) - x_r(n)}{\sqrt{2}}, \quad (2)$$

где $x_l(n)$ и $x_r(n)$ — отсчеты с индексом n левого и правого каналов исходного звука соответственно.

Восходящее микширование на основе метода наименьших квадратов

Восходящее микширование по методу наименьших квадратов (LMS-based Upmixing Method, LMS) создает центральный и задние каналы с использованием алгоритма LMS [1, 4]. При таком подходе один из исходных стереоканалов принимается как желаемый сигнал d_n , а другой считается входным сигналом адаптивного фильтра x_n . Сигнал ошибки e_n определяется как разность выходного сигнала y_n и желаемого сигнала d_n . Выходной сигнал y_n определяется как линейная комбинация входных сигналов с использованием уравнения (3):

$$y(n) = \mathbf{w}^T(n)\mathbf{x}(n) = \mathbf{w}(n)\mathbf{x}^T(n), \quad (3)$$

где

$$\mathbf{x}(n) = [x(n)x(n-1) \dots x(n-N+1)]^T,$$

$$\mathbf{w}(n) = [w_0 w_1 \dots w_{N-1}]^T.$$

В уравнении (3) вектор $w(n)$ — вектор коэффициентов адаптивного фильтра, получаемого на основе LMS-алгоритма согласно уравнению (4):

$$\mathbf{w}(n+1) = \mathbf{w}(n) + 2\mu e(n)\mathbf{x}(n), \quad (4)$$

где N — длина фрейма,

μ — константа, равная, например 10^{-4} .

Значения y_n принимаются за значения отсчетов центрального канала, а значения e_n принимаются за значения отсчетов задних каналов генерируемого 5.1-звука.

Метод восходящего микширования на основе анализа главных компонент

Метод восходящего микширования на основе анализа главных компонент (PCA-based Upmixing Method, Principal Component Analysis, PCA) разлагает исходные стереоканалы на коррелированные и некоррелированные части [1]. Центральный и задний каналы получают на основе ковариационной матрицы:

$$A = \begin{bmatrix} cov(x_l, x_l) & cov(x_l, x_r) \\ cov(x_r, x_l) & cov(x_r, x_r) \end{bmatrix}, \quad (5)$$

где $cov(x_p, x_q)$ — ковариация x_p и x_q , где p (или q) соответствует левому или правому каналу исходного стереосигнала.

Ковариационная матрица A дает два собственных вектора, являющихся базисными векторами для новой системы координат [5]. Эти собственные вектора затем используются в качестве весовых векторов, соответствующих левому и правому каналам, при генерации центрального канала и задних каналов: (6) и (7).

$$Center(n) = c_l x_l(n) + c_r x_r(n), \quad (6)$$

Собственный вектор $[c_l \ c_r]$, соответствующий наибольшему собственному значению, используется в качестве весового вектора для центрального канала. Другой собственный вектор $[s_l \ s_r]$ становится весовым вектором для сигнала задних каналов.

Обработка сигнала по методу PCA выполняется по фреймам. Для устранения артефактов, которые могут возникать на стыках фреймов, вводятся перекрытие и сглаживание [6]. При сглаживании используется весовая функция вида (8), где M — величина перекрытия.

Метод адаптивного панорамирования

Метод адаптивного панорамирования (Adaptive Panning Method, ADP), предложенный в [2], генерирует центральный и задние каналы на основе алгоритма LMS. Вектор весов для ADP вычисляется рекурсивно согласно (3), однако вектора определяются иначе: (9).

Коэффициенты $w_l(n)$ и $w_r(n)$ соответствуют левому и правому каналам и вычисляются по формулам (10) и (11), где μ — константа, приведенная выше. Значения отсчетов центрального и заднего каналов вычисляются как (12) и (13).

Во всех четырех описанных алгоритмах также применяется фильтрация и задержка с инверсией фаз.

Тестирование методов восходящего микширования

В работе [7] приведено описание программного комплекса, выполняющего разложение стереозвука в 5.1 пе-

речисленными методами восходящего микширования и результаты сравнительного тестирования представленных методов.

Эксперименты показали, что прослушивание звукового контента, полученного в результате восходящего микширования, является предпочтительным по сравнению с прослушиванием исходного стереофонического контента. Также было установлено, что метод адаптивного панорамирования превосходит другие рассмотренные методы.

Описанные выше методы не расширяют сцену по сравнению со стерео-сценой: они добавляют объем звучания за счет задних каналов. Изменения панорамирования звука не происходит. Звуковая сцена формируется между фронтальными динамиками, а звук добавляемый в задние динамики создают эффект присутствия в помещении (рисунок 2).

Все перечисленные методы не обеспечивают ни формирования естественной звуковой сцены, ни возможности ее произвольного расширения, ни возможности формирования таких звуковых эффектов, как имитация прослушивания в выбранном акустическом окружении.

Поэтому было принято решение о разработке метода преобразования стереозвука в многоканальный звук, лишённого перечисленных недостатков (рисунок 3).

Предполагается, что звуковая сцена расширяется и создается эффект погружения в любое целевое помещение.

$$Surround(n) = s_l x_l(n) + s_r x_r(n) \quad (7)$$

$$w(n) = \begin{cases} \sin\left(\frac{\pi(n+0.5)}{2(N-M)}\right), & 0 \leq n \leq N-M-1 \\ 1, & N-M \leq n \leq M-1 \\ \sin\left(\frac{\pi(N-n-0.5)}{2(N-M)}\right), & M \leq n \leq N-1 \end{cases} \quad (8)$$

$$\mathbf{x}(n) = [x_l(n) \ x_r(n)]^T, \mathbf{w}(n) = [w_l(n) \ w_r(n)]^T \quad (9)$$

$$w_l(n+1) = w_l(n) + \mu y(n) [x_l(n) - w_l(n) y(n)] \quad (10)$$

$$w_r(n+1) = w_r(n) + \mu y(n) [x_r(n) - w_r(n) y(n)] \quad (11)$$

$$Center(n) = w_l x_l(n) + w_r x_r(n) \quad (12)$$

$$Surround(n) = w_r x_l(n) + w_l x_r(n) \quad (13)$$

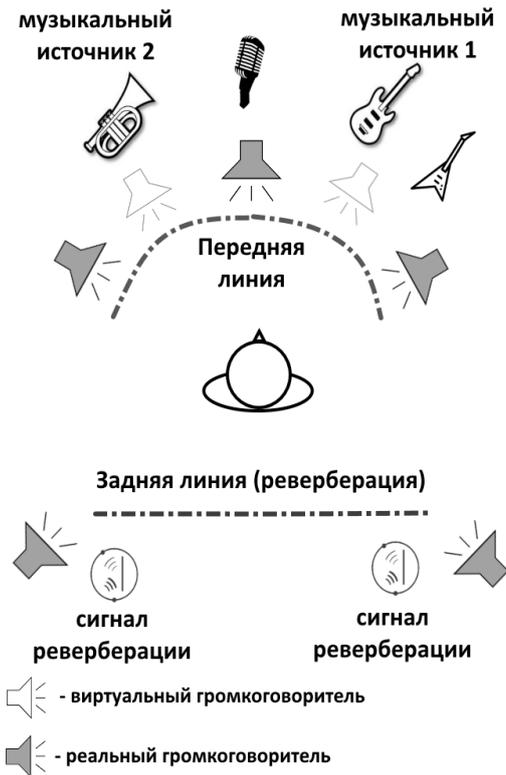


Рис. 2. Формирование звуковой сцены рассмотренными методами



Рис. 3. Желаемая звуковая сцена

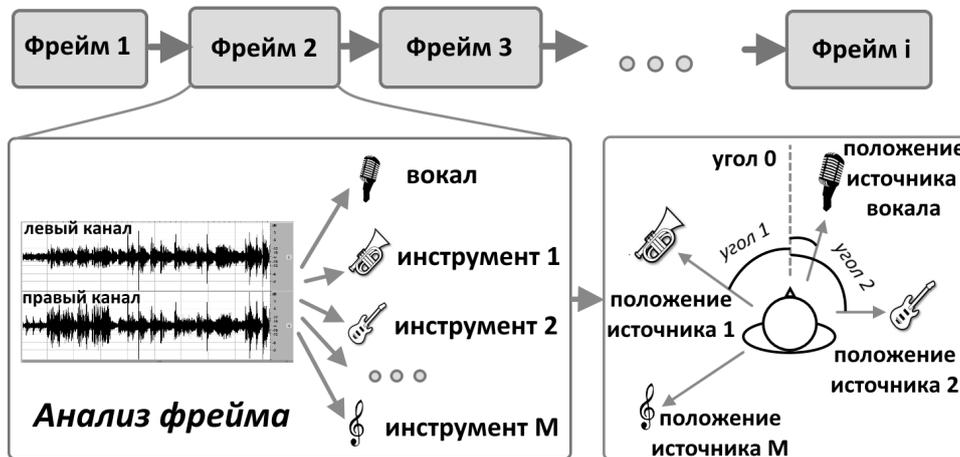


Рис. 4. Схема работы предлагаемого метода

Предлагаемый метод

Локализация звуковых источников возможна благодаря значительным различиям между сигналами, поступающими в левое и правое ухо слушателя. Различия сигналов обуславливаются тремя физическими факторами:

1. Временными, возникающими из-за несовпадения моментов времени прихода одинаковых фаз;
2. Интенсивностными, возникающими из-за различий интенсивности звуковой волны;
3. Спектральными, возникающими из-за разницы в спектральном составе звуков.

Таблица 1. Разбиение звукового сигнала на полосы по учебнику «Аудиотехника»

Номер полосы	Нижняя граница, Гц	Верхняя граница, Гц	Ширина полосы, Гц
0	0	86	86
1	86	172	86
2	172	258	86
3	258	345	86
4	345	517	172
5	517	689	172
6	689	861	172
7	861	1034	172
8–68	1034	22050	345

Таблица 2. Используемое разбиение на полосы

Номер полосы	Область вычисления параметров		Область разделения	
	Нижняя граница, Гц	Верхняя граница, Гц	Нижняя граница, Гц	Верхняя граница, Гц
0	0,0	86,0	0,0	129,0
1	86,0	172,0	43,0	215,0
2	172,0	258,0	129,0	301,5
3	258,0	345,0	215,0	431,0
4	345,0	517,0	301,5	603,0
5	517,0	689,0	431,0	775,0
6	689,0	861,0	603,0	947,5
7	861,0	1034,0	775,0	1206,5
8	1034,0	1379,0	947,5	1551,5
.....				

В рамках предлагаемого метода исследуются только два из трех перечисленных факторов: межканальная разность фаз (Inter-channel phase difference IPD) и межканальная разность интенсивностей (Inter-channel intensity difference, IID).

Предлагаемый метод основан на выделении виртуальных источников звука в исходном стереосигнале, определении направлений на них и позиционировании их в звуковой сцене, формируемой произвольным количеством динамиков (рисунок 4).

Для достижения желаемого результата предлагаемый метод включает следующую последовательность действий:

1. формирование выборок сигнала;
2. Разделение сигнала на полосы;
3. Вычисление параметров сигнала на полосах;
4. Вычисление критерия распределения полос по динамикам;
5. Распределение полос по динамикам;
6. Формирование многоканального сигнала;
7. Формирование эхо-сигнала;
8. Дополнительные обработки.

1) Формирование выборок сигнала

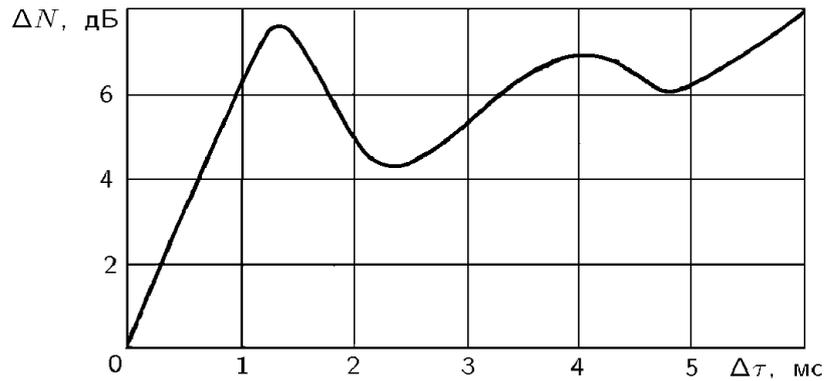
Исходный стерео сигнал разбивается на выборки по N отсчетов и переводится в частотную область:
 X_1 — спектр (Фурье-образ) выборки левого канала,
 X_2 — спектр (Фурье-образ) выборки правого канала.

Длина выборки является параметром обработки, кратным степени двойки. Эксперименты показали, что наилучшее качество звучания достигается при $N = 2048$. Длина выборки в частотной области $M = I + N/2$.

Обработка выполняется со сдвигом, не превышающим длину половины выборки. Величина сдвига является параметром обработки.

2) Разделение сигнала на полосы

Были проведены эксперименты с разными вариантами разбиения на полосы. При разбиении по классическому варианту [8], представленному в таблице 1, звучание получается не самым удачным.

Рис. 5. Связь временного (ΔT) и интенсивностного (ΔN) факторов

Поэтому в качестве рабочего разбиения используется разбиение с перекрытием (таблица 2), при этом при вычислении параметров для разделения используется только часть каждой полосы:

3) Параметры на полосах

Для разделения используются два параметра: межканальная разность интенсивностей и межканальная разность фаз.

Межканальная разность интенсивностей (IID) по формуле (14).

В классическом варианте межканальная разность фаз (IPD) по формуле (15), где X^* означает комплексно сопряженное значение с X (16).

Однако наилучший вариант достигается при вычислении задержки во временной области путем нахождения сдвига максимума корреляции между сигналами левого и правого каналов исходного стереозвука.

4) Вычисление критерия распределения звука по динамикам

В качестве критерия распределения звука по динамикам используется эквивалентное значение межканальной разности уровней сигналов (IIDeq). В зависимости от центральной частоты полосы IIDeq вычисляется так: (17).

Коэффициенты $k1$ и $k2$ позволяют управлять расширением звуковой панорамы путем смещения расчетного положения источников звука.

$$IID[b] = 10 \log_{10} \left(\sum_{k=k_b}^{k_{b+1}-1} (X_1[k].re^2 + X_1[k].im^2) \right) - 10 \log_{10} \left(\sum_{k=k_b}^{k_{b+1}-1} (X_2[k].re^2 + X_2[k].im^2) \right) \quad (14)$$

$$IPD[b] = \sum_{k=k_b}^{k_{b+1}-1} X_1[k] \cdot X_2^*[k] \quad (15)$$

$$X^*.re = X.re, X^*.im = -X.im \quad (16)$$

$$IIDeq = \begin{cases} k_2 \Delta N & , \text{freq} < 500 \\ k_1 IID + k_2 \Delta N & , 500 \leq \text{freq} \leq 5000 \\ k_1 IID & , \text{freq} > 5000 \end{cases} \quad (17)$$

$$\Delta \tau = \frac{500 \cdot |IPD|}{\text{freq} \cdot \pi} \quad (18)$$

Таблица 3. Отрезки и уравнения интерполяции ΔN по ΔT .

Отрезок	Формула
$0 \leq \Delta T < 1.33$	$\Delta N = (7.5f/1.33f) * \Delta T$
$1.33 \leq \Delta T < 2.33$	$\Delta N = 7.5f - (3.0f/1.0f) * (\Delta T - 1.33f)$
$2.33 \leq \Delta T < 4.00$	$\Delta N = 4.5f + (2.5f/1.67f) * (\Delta T - 2.33f)$
$4.00 \leq \Delta T < 4.75$	$\Delta N = 7.0f - (1.0f/0.75f) * (\Delta T - 4.00f)$
$4.75 \leq \Delta T < 6.00$	$\Delta N = 6.0f + (2.0f/1.25f) * (\Delta T - 4.75f)$
$\Delta T \geq 6.00$	$\Delta N = 8.0f + (2.0f/1.25f) * (\Delta T - 6.0f)$

Таблица 4. Группы динамиков и условия их выбора

№ группы	Список динамиков по классике	Расширение	Условие
1	Center, left-front, right-front	Left-side, right-side	$ IIDeq < 2\text{дБ}$
2	Center, left-front	Left-side, right-front	$2 < IIDeq < 7$
3	Center, right-front	Right-side, left-front	$-7 < IIDeq < -2$
4	Left-front	Center, left-side	$7 < IIDeq < 9$
5	Right-front	Center, right-side	$-9 < IIDeq < -7$
6	Left-front, left-side	Center, right-side	$9 < IIDeq < 25$
7	Right-front, right-side	Center, left-side	$-25 < IIDeq < -9$
8	Left-side	Left-front, right-side	$IIDeq > 25$
9	Right-side	Right-front, left-side	$IIDeq < -25$

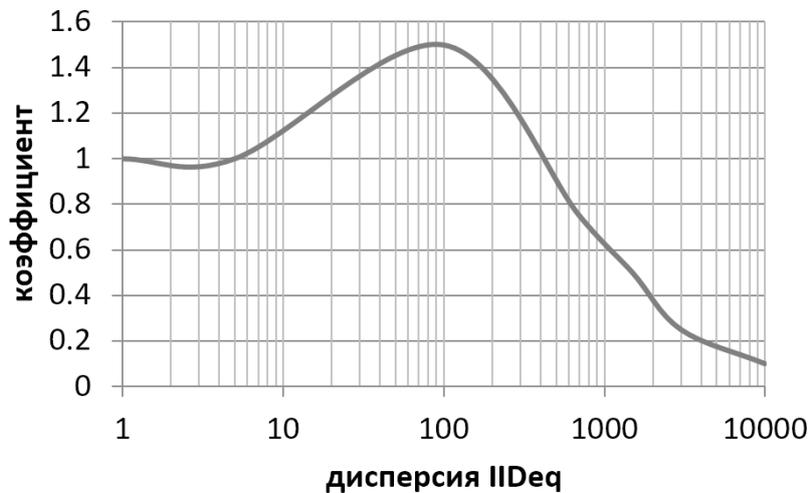


Рис. 6. Зависимость коэффициента коррекции IIDeq от дисперсии IIDeq на полосах

Значение ΔN определяется по графику, представленному на рисунке 5 [8]:

В случае определения эквивалентной разности уровней IIDeq по разности фаз IPD величина ΔT рассчитывается по формуле:

В случае определения эквивалентной разности уровней IIDeq по сдвигу максимума корреляции величина ΔT рассчитывается по формуле:

$$\Delta T = \frac{1000 \cdot |CorPos|}{SampleRate} \tag{19}$$

Представленная выше кривая разбивается на 6 отрезков и пересчет ΔT в разность уровней выполняется путем линейной интерполяции (таблица 3).

Полученное значение ΔN берется со знаком разности фаз (18), либо со знаком сдвига максимума корреляции (19).

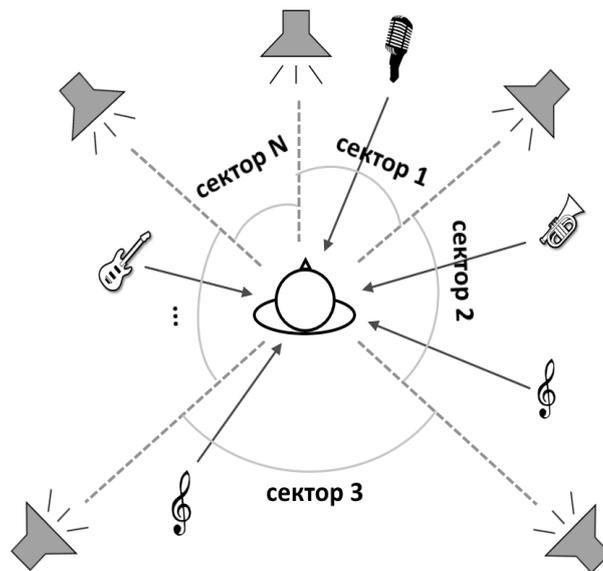


Рис. 7. Разделение пространства на сектора

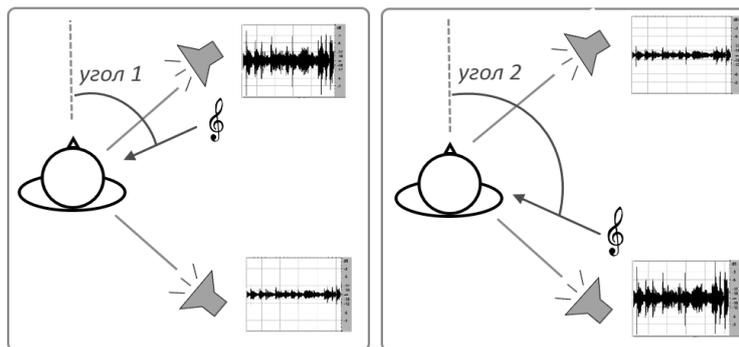


Рис. 8. Позиционирование виртуального источника звука в секторе

Далее, в зависимости от дисперсии II_{Deq} на полосах в текущей выборке, выполняется динамическая коррекция значений II_{Deq} согласно следующей экспериментальной зависимости (рисунок 6):

5) Разделение полос между динамиками

Пространство вокруг слушателя делится на сектора в соответствии с количеством реальных динамиков (рисунок 7). Виртуальные источники звука отображаются в секторах в соответствии со значением II_{Deq} , полученным ранее.

Каждый виртуальный источник звука воспроизводится несколькими реальными динамиками. Основное звучание формируется двумя реальными динамиками, образующими границы сектора (рисунок 8), остальные

динамики формируют звук поддержки, обеспечивающий стабильность источника и гладкость звуковой сцены в целом (при совпадении виртуального источника звука с реальным динамиком основное звучание формируется одним динамиком).

Правильное позиционирование источника звука в секторе обеспечивается выбором соответствующих амплитуд сигналов динамиков, определяющих границы сектора (рисунок 8).

При разложении стереозвуча в многоканальный звук формата 5.1 выделяется 9 групп динамиков. Выбор группы осуществляется по значению II_{Deq} . По результатам экспериментов классические группы динамиков [8] были расширены динамиками, формирующими звук поддержки. Номера групп определяются следующими условиями, также представленными в таблице 4.

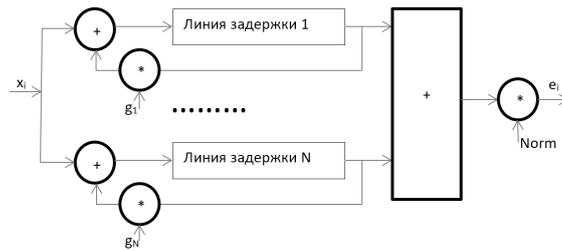


Рис. 9. Схема мультилинейного цифрового ревербератора

Для полос с центральной частотой ниже 2кГц (голосовой диапазон) вводится ограничение по разделению: данные частотные полосы не могут быть направлены полностью назад, чтобы источник вокала не пропадал с привычной для слушателя передней сцены. Поэтому, если они попадают в группы с номером 8 или 9, номер группы уменьшается на 2.

Значение II_{Deq} обладает значительной временной нестабильностью, что может приводить к разрывам формируемой звуковой панорамы. Для коррекции проблемы «непрерывности» звучания было введено сглаживание во времени. Алгоритм сглаживания по 20 выборкам (количество выборок является параметром алгоритма, определенным в результате экспериментов) выбирает группу активных динамиков и коэффициент усиления для них. Если происходит смена активной группы динамиков, то формируется плавное затухание звучания в предыдущей группе активных динамиков и плавное нарастание в новой активной группе. Дополнительно выполняется фильтрация разделения по частотным полосам.

б) Формирование сигналов 5.1

Спектр сигнала разложения формируется для каждого канала многоканального звука. Спектр полосы разложения исходного стереосигнала суммируется со спектрами сигналов выбранных каналов многоканального звука с учетом весов. В случае использования полос с частотным перекрытием при суммировании используется весовая оконная функция:

$$W_i = \frac{1 - \cos((2 \cdot \pi \cdot i) / (\text{windowLength} - 1))}{2}$$

При формировании звука, относящегося к центральному каналу, суммирование ведется без учета фаз звукового сигнала правого канала исходного стереозвука. Специальная процедура суммирования введена из-за того, что при суммировании звука с равными весами, но находящегося в противофазе, приводит к получению нулевого уровня.

Аналогичная процедура суммирования выполняется и при формировании звука LFE-канала. Он формируется как сумма звука левого и правого каналов исходного

стереозвука, взятого с равными весами в полосе частот от 0 до 150Гц.

7) Формирование эха

Разработано несколько схем формирования ревербераций. Наилучшие результаты получены при формировании ревербераций на основе применения свертки с импульсной характеристикой моделируемого зала и мультилинейного цифрового ревербератора.

7.1) Реверберации на основе импульсной характеристики

Эхо формируется из моно-сигнала, получаемого как взвешенная сумма спектров каналов исходного стереосигнала:

$$X_m = w_1 X_1 + w_2 X_2$$

В классическом варианте выполняется декорреляция полученного моно-сигнала путем пропуска его через всепропускающий фильтр с импульсной характеристикой, описываемой уравнением:

$$h_d[n] = \frac{2}{N} \sum_{k=0}^{N/2} \cos\left(\frac{2\pi kn}{N} + \frac{2\pi k(k-1)}{N}\right)$$

Экспертное тестирование показало, что применение всепропускающего фильтра снижает «приятность» синтезируемого сигнала. Поэтому в рабочей версии ПО от применения данного фильтра пришлось отказаться.

Спектр моно-сигнала переводится во временную область, после чего выполняется операция свертки полученного сигнала с импульсной характеристикой выбранного помещения. Для каждого канала может быть задана собственная импульсная характеристика.

Сформированный эхо-сигнал добавляется к основному сигналу соответствующих каналов. Для увеличения согласованности спектров прямого и отраженного сигналов, основной сигнал также может быть обработан фильтром, полученным из выбранной импульсной характеристики.

7.2) Реверберации на основе мультилинейного цифрового ревербератора

Схема мультилинейного цифрового ревербератора представлена на рисунке 9:

Длительности линий задержки возрастают. Коэффициенты усиления определены экспериментально:

$$g_i = -0,9; g_i = g_{i-1} * k, \text{ где } k = 0,99 \dots 0,97.$$

Коэффициент нормировки равен сумме коэффициентов усиления g_i .

Дополнительно входящий сигнал x_i может быть обработан эквалайзером, а результирующий эхо-сигнал e_i задержан на заданное время.

Наиболее естественное эхо формируется при коэффициенте $k = 0,98$ и количестве линий задержки от 10 до 16.

8) Дополнительные обработки

Дополнительная обработка включает в себя:

- ◆ Применение эквалайзера, позволяющего слушателю придать системе звучание, наиболее приятное для себя. Предполагается, что для каждого канала системы может быть подобран собственный эквалайзер. Также эквалайзер может быть использован для исправления АЧХ динамиков;
- ◆ Применение импульсного фильтра, позволяющего компенсировать характеристику помещения, в котором происходит прослушивание, а также моделировать звучание, соответствующее выбранному помещению;
- ◆ Применение лимитера, позволяющего исправлять значения отсчетов сигнала, при выходе их

за допустимый диапазон значений без внесения в сигнал клиппирования.

Аудиторские испытания

Испытания проводились в специально оборудованном помещении с привлечением пяти экспертов по звуку. Для сравнения брались многоканальные записи, полученные на системах Pro Logic, Pro Logic II и записи, полученные в соответствии с предложенным методом. Испытания проводились на записях разных музыкальных жанров (классика, поп-музыка, рок, джаз).

В слепом тесте эксперты единогласно отдали звуку системы Pro Logic третье место. Относительно Pro Logic II и предложенного метода единого мнения получено не было. Трое экспертов отдали предпочтение звуку, сформированному по предложенному методу, один — системе Pro Logic II и один эксперт оценил качество звучания как «одинаково хорошее».

Однако все эксперты отметили большую ширину звуковой сцены, формируемой согласно предложенному методу.

Результаты проведенных испытаний позволяют говорить о возможности применения предложенного метода в системах формирования объемного звучания.

Благодарности

Авторы выражают благодарность заслуженному деятелю науки РФ, доктору технических наук, профессору кафедры радиосвязи и вещания СПбГУТ, Ковалгину Юрию Алексеевичу за ценные советы при планировании и проведении исследования.

ЛИТЕРАТУРА

1. M. R. Bai, G.-Y. Shih, and J.-R. Hong, «Upmixing and downmixing two-channel stereo audio for consumer electronics,» IEEE Trans. on Consumer Electronics, vol. 53, no. 3, pp. 1011–1019, Aug. 2007.
2. R. Irwan, R. M. Aarts, «Two-to-five channel sound processing,» J. Audio Eng. Soc., vol. 50, no. 11, pp. 914–926, Nov. 2002.
3. ITU-R BS.1534–1, Method for the Subjective Assessment of Intermediate Quality Levels of Coding System, Jan. 2003.
4. B. Widrow, S. D. Stearns, Adaptive Signal Processing, Prentice-Hall, 1985.
5. I. T. Jolliffe, Principal Component Analysis, Springer, Heidelberg, 2002.
6. M. Bosi, R. E. Goldberg, Introduction to Digital Audio Coding and Standards, Kluwer Academic Publishers, Dec. 2002.
7. C. J. Chun, Y. G. Kim, J. Y. Yang, H. K. Kim, Real-Time Conversion of Stereo Audio to 5.1 Channel Audio for Providing Realistic Sounds, International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol. 2, No. 4, pp.85–94, Dec. 2009
8. Ковалгин Ю. А., Вологдин Э. И., Аудиотехника. Учебник для вузов. — М.: Горячая линия — Телеком, 2013