

(19)



**Евразийское  
патентное  
ведомство**

(11) **041656**

(13) **B1**

(12) **ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ**

(45) Дата публикации и выдачи патента  
**2022.11.18**

(51) Int. Cl. **H04S 7/00** (2006.01)

(21) Номер заявки  
**201992556**

(22) Дата подачи заявки  
**2016.08.23**

---

(54) **АУДИОДЕКОДЕР И СПОСОБ ДЕКОДИРОВАНИЯ**

---

(31) **62/209,742; 15189008.4**

(32) **2015.08.25; 2015.10.08**

(33) **US; EP**

(43) **2021.03.31**

(62) **201890557; 2016.08.23**

(71)(73) Заявитель и патентовладелец:  
**ДОЛБИ ЛЭБОРЕТЕРИЗ  
ЛАЙСЕНСИНГ КОРПОРЕЙШН  
(US); ДОЛБИ ИНТЕРНЭШНЛ АБ  
(NL)**

(56) HERRE J. и др. MPEG Surround-The ISO/MPEG Standard for Efficient and Compatible Multichannel Audio Coding, JAES, AES, NEW YORK, USA, vol. 56, no. 11, 2008-11-01, pages 932-955, XP040508729, 1 ноября 2008, раздел 3.1, 4.4, 4.7, фиг. 16

US-A-5757931

US-A1-2008319765

(72) Изобретатель:  
**Бребарт Дирк Ерун, Купер Дэвид  
Мэттью (AU), Самуэльссон Лейф  
Йонас (SE)**

(74) Представитель:  
**Медведев В.Н. (RU)**

---

(57) Способ представления второго представления аудиоканалов или объектов как потока данных, причем способ содержит следующие этапы: (а) обеспечение множества базовых сигналов, базовые сигналы представляют первое представление аудиоканалов или объектов; (b) обеспечение множества параметров преобразования, параметры преобразования предназначены для преобразования первого представления во второе представление; параметры преобразования также заданы по меньшей мере для двух частотных полос и включают в себя множество параметров матрицы свертки с несколькими отводами по меньшей мере для одной из частотных полос.

---

**B1**

**041656**

**041656**

**B1**

Настоящая заявка испрашивает приоритет согласно предварительной заявке на патент США № 62/209,742, поданной 25 августа 2015 года, и заявке на европейский патент № 15189008.4, поданной 8 октября 2015 года, содержание каждой из которых полностью включено в настоящую заявку посредством ссылки.

#### **Область техники**

Изобретение относится к области обработки сигналов и, в частности, раскрывает систему для эффективной передачи аудиосигналов, имеющих компоненты для ориентации в пространстве.

#### **Уровень техники**

Любое обсуждение уровня техники по тексту описания никоим образом не должно рассматриваться как признание того, что такая область техники широко известна или является частью общедоступных сведений в предметной области.

Создание контента, кодирование, распространение и воспроизведение аудиоинформации традиционно выполняется в формате на основе каналов, то есть, одна конкретная целевая система воспроизведения предусматривается для контента в экосистеме контента. Примерами таких целевых форматов аудиосистем воспроизведения являются моно, стерео, 5.1, 7.1 и т.п.

Если контент должен быть воспроизведен на другой системе воспроизведения, отличающейся от намеченной, может быть применен процесс понижающего микширования или повышающего микширования. Например, контент в формате 5.1 может быть воспроизведен на системе воспроизведения стерео с использованием заданных уравнений понижающего микширования. Другим примером является воспроизведение контента, закодированного в формате стерео, на конфигурации динамиков 7.1, которое может содержать процесс так называемого повышающего микширования, которым может управлять или не управлять информация, присутствующая в стереосигнале. Системой, способной к повышающему микшированию, является система Dolby Pro Logic от Dolby Laboratories Inc (Roger Dressier, "Dolby Pro Logic Surround Decoder, Principles of Operation", www.Dolby.com).

Когда контент в формате стерео или в многоканальном формате должен быть воспроизведен в наушниках, часто желательно смоделировать многоканальную конфигурацию динамиков посредством импульсных характеристик слухового аппарата (HRIR) или бинауральных импульсных характеристик комнаты (BRIR), которые моделируют путь прохождения звука от каждого динамика до барабанных перепонок в (смоделированной) окружающей среде без эха и с эхом, соответственно. В частности, аудиосигналы могут быть подвергнуты свертке с помощью характеристик HRIR или BRIR, чтобы восстановить разности уровней между ушами (ILD), разности времени между ушами (ITD) и спектральные признаки, которые позволяют слушателю определять местоположение каждого индивидуального канала. Моделирование акустической окружающей среды (реверберация) также помогает достигнуть некоторого восприятия расстояния.

#### **Локализация источника звука и моделирование виртуального динамика**

Когда контент в формате стерео, в многоканальном формате или в формате на основе объектов должен быть воспроизведен в наушниках, часто желательно смоделировать многоканальную конфигурацию динамиков или множество отдельных виртуальных акустических объектов посредством свертки с помощью импульсных характеристик слухового аппарата (HRIR) или бинауральных импульсных характеристик комнаты (BRIR), которые моделируют путь прохождения звука от каждого динамика до барабанных перепонок в (смоделированной) окружающей среде без эха и с эхом, соответственно.

В частности, аудиосигналы подвергаются свертке с помощью характеристик HRIR или BRIR, чтобы восстановить разности уровней между ушами (ILD), разности времени между ушами (ITD) и спектральные признаки, которые позволяют слушателю определять местоположение каждого индивидуального канала или объекта. Моделирование акустической окружающей среды (ранние отражения и поздняя реверберация) помогает достигнуть некоторого восприятия расстояния.

Обратимся к фиг. 1, на которой проиллюстрирован схематический обзор 10 потока обработки для рендеризации двух объектов или канальных сигналов  $x_1$ , 13, 11, считываемых из хранилища 12 контента для обработки, например, посредством четырех характеристик 14 HRIR. Выходные данные характеристик HRIR затем суммируются 15, 16 для каждого канального сигнала, чтобы произвести выходные данные динамиков наушников для воспроизведения слушателю через наушники 18. Основной принцип характеристик HRIR, например, разъяснен в литературе Wightman et al. (1989).

Подход со сверткой с помощью характеристик HRIR/BRIR сопровождается несколькими недостатками, одним из которых является значительное количество обработки, которая требуется для воспроизведения с помощью наушников. Свертка с помощью характеристики HRIR или BRIR должна быть применена для каждого входного объекта или канала отдельно, и поэтому сложность, как правило, растет линейно в зависимости от количества каналов или объектов. Поскольку наушники, как правило, используются вместе с работающими от аккумулятора портативными устройствами, высокая вычислительная сложность не желательна, поскольку она в значительной степени сокращает время работы от батареи. Кроме того, с введением аудиоконтента на основе объектов, который может содержать более чем 100 активных объектов одновременно, сложность свертки с помощью характеристики HRIR может быть в значительной степени выше, чем для традиционного контента на основе каналов.

### Методика параметрического кодирования

Вычислительная сложность не является единственной проблемой для доставки контента на основе каналов или объектов в экосистеме, включающей в себя создание, распределение и воспроизведение контента. Во многих ситуациях на практике, и особенно для мобильных приложений, сильно ограничена скорость передачи данных, доступная для доставки контента. Потребители, телевизионные компании и поставщики контента доставляли аудиоконтент в формате стерео (с двумя каналами) с использованием перцепционных аудиокодеков с потерями с типичными битрейтами между 48 и 192 кбит/с. Эти традиционные основанные на каналах аудиокодеки, такие как уровень MPEG-1 layer 3 (Brandenberg et al., 1994), MPEG AAC (Bosi et al., 1997) и Dolby Digital (Andersen et al., 2004) имеют битрейт, который масштабируется приблизительно линейно в зависимости от количества каналов. В результате доставка десятков или даже сотен объектов приводят к битрейтам, которые затруднительны или даже недоступны для целей доставки потребителю.

Чтобы сделать возможной доставку сложного контента на основе объектов при битрейтах, которые сопоставимы с битрейтом, требуемым для доставки контента в формате стерео, с использованием традиционных перцепционных аудиокодеков, за прошлое десятилетие подверглись научным исследованиям так называемые параметрические методы. Эти параметрические методы дают возможность воссоздания большого количества каналов или объектов из относительно низкого количества базовых сигналов. Эти базовые сигналы могут быть перенесены от отправителя к получателю с использованием традиционных аудиокодеков с дополнительной (параметрической) информацией, чтобы сделать возможным воссоздание исходных объектов или каналов. Примерами таких методик являются Parametric Stereo (Schuijers et al., 2004), MPEG Surround (Herre et al., 2008) и MPEG Spatial Audio Object Coding (Herre et al., 2012).

Важным аспектом таких методик, как Parametric Stereo и MPEG Surround, является то, что эти методы нацелены на параметрическое воссоздание единственного, предварительно заданного представления (например, динамики стерео в Parametric Stereo и динамики 5.1 в MPEG Surround). В случае MPEG Surround виртуализатор наушников может быть интегрирован в декодер, который генерирует конфигурацию виртуальных динамиков 5.1 для наушников, в которых виртуальные динамики 5.1 соответствуют конфигурации динамиков 5.1 для воспроизведения с помощью динамиков. Следовательно, эти представления весьма зависимы в том, что представление для наушников представляет ту же самую (виртуальную) схему размещения динамиков, что и представление для динамиков. MPEG Spatial Audio Object Coding (пространственное кодирование аудиообъектов MPEG), с другой стороны, нацелено на воссоздание объектов, которые требуют последующей рендеризации.

Обратимся теперь к фиг. 2, на которой будет обзорно описана параметрическая система 20, поддерживающая каналы и объекты. Система разделена на кодер 21 и декодер 22. Кодер 21 принимает каналы и объекты 23 в качестве входных данных и генерирует звукозапись 24 с понижающим микшированием с ограниченным количеством базовых сигналов. Кроме того, вычисляется последовательность параметров 25 воссоздания объектов/каналов. Кодер 26 сигналов кодирует базовые сигналы из блока 24 понижающего микширования и включает в итоговый битовый поток вычисленные параметры 25, а также метаданные 27 объектов, указывающие, каким образом объекты должны быть рендеризированы.

Декодер 22 сначала декодирует 29 базовые сигналы, затем выполняется воссоздание 30 каналов и/или объектов с помощью переданных параметров 31 воссоздания. Полученные в результате сигналы могут быть воспроизведены непосредственно (если это каналы) или могут быть рендеризированы 32 (если это объекты). В последнем случае каждый воссозданный сигнал объекта рендеризуется в соответствии со своими привязанными метаданными 33 объекта. Одним примером таких метаданных является вектор позиции (например, координаты  $x$ ,  $y$  и  $z$  объекта в трехмерной системе координат).

#### Матрицирование декодера

Воссоздание 30 объектов и/или каналов может быть достигнуто посредством переменных по времени и частоте матричных операций. Если декодированные базовые сигналы 35 обозначены как  $z_s[n]$ , где  $s$  - индекс базового сигнала, и  $n$  - индекс отсчета, первый этап, как правило, содержит преобразование базовых сигналов посредством преобразования или набора фильтров.

Могут быть использованы самые различные преобразования и наборы фильтров, такие как дискретное преобразование Фурье (DFT), модифицированное дискретное косинусное преобразование (MDCT) или набор квадратурных зеркальных фильтров (QMF). Выходные данные такого преобразования или набора фильтров обозначены как  $Z_s[k,b]$ , где  $b$  - индекс подполосы или спектральный индекс и  $k$  - индекс кадра, слота или времени или отсчета подполосы.

В большинстве случаев индексы подполос или спектральные индексы отображаются на меньшее множество параметрических полос  $p$ , которые имеют общие параметры воссоздания объектов/каналов. Это может быть обозначено как  $b \in B(p)$ . Другими словами,  $B(p)$  представляет множество последовательных подполос  $b$ , которые принадлежат параметрической полосе с индексом  $p$ . С другой стороны,  $p(b)$  относится к параметрической полосе с индексом  $p$ , на которую была отображена подполоса  $b$ . Подполосы или воссозданные каналы или объекты  $\hat{Y}_j$  в области преобразования затем получаются посредством матрицирования сигналов  $Z_i$  с помощью матриц  $M[p(b)]$ :

$$\begin{bmatrix} \hat{Y}_1[k,b] \\ \vdots \\ \hat{Y}_j[k,b] \end{bmatrix} = M[p(b)] \begin{bmatrix} Z_1[k,b] \\ \vdots \\ Z_s[k,b] \end{bmatrix}$$

Сигналы  $y_j[n]$  воссозданных каналов и/или объектов во временной области затем получаются посредством обратного преобразования или набора фильтров для синтеза.

Приведенный выше процесс, как правило, применяется к некоторому ограниченному диапазону отсчетов, слотов или кадров к подполосы. Другими словами, матрицы  $M[p(b)]$ , как правило, обновляются/модифицируются со временем. Для простоты обозначения эти обновления здесь не обозначены. Однако считается, что обработка множества отсчетов  $k$ , связанного с матрицей  $M[p(b)]$ , может являться переменным по времени процессом.

В некоторых случаях, в которых количество воссозданных сигналов  $J$  значительно больше, чем количество базовых сигналов  $S$ , часто полезно использовать выходные данные  $D_m[k,b]$  факультативного декоррелятора, воздействующие на один или более базовых сигналов, которые могут быть включены в воссозданные выходные сигналы:

$$\begin{bmatrix} \hat{Y}_1[k,b] \\ \vdots \\ \hat{Y}_j[k,b] \end{bmatrix} = M[p(b)] \begin{bmatrix} Z_1[k,b] \\ \vdots \\ Z_s[k,b] \\ D_1[k,b] \\ \vdots \\ D_M[k,b] \end{bmatrix}$$

Фиг. 3 более подробно схематично иллюстрирует одну форму блока 30 воссоздания каналов или объектов, показанного на фиг. 2. Входные сигналы 35 сначала обрабатываются наборами 41 фильтров для анализа, затем выполняется факультативная декорреляция (D1, D2) 44 и матрицирование 42, и применяется набор 43 фильтров для синтеза. Матричной манипуляцией  $M[p(b)]$  управляют параметры 31 воссоздания.

#### Предсказание с минимальной среднеквадратической ошибкой (MMSE) для воссоздания объекта/канала

Хотя существуют разные стратегии и способы для воссоздания объектов или каналов из множества базовых сигналов  $Z_s[p(b)]$ , один конкретный способ часто упоминается как прогноз с минимальной среднеквадратической ошибкой (MMSE), который использует корреляционные и ковариационные матрицы для вывода матричных коэффициентов  $M$ , которые минимизируют норму  $L_2$  между желаемым и воссозданным сигналами. Для этого способа предполагается, что базовые сигналы  $z_s[n]$  генерируются в блоке 24 понижающего микширования кодера как линейная комбинация сигналов  $x_i[n]$  входных объектов или каналов:

$$z_s[n] = \sum_i g_{i,s} x_i[n]$$

Для входного контента на основе каналов коэффициенты  $g_{i,s}$  усиления амплитудного панорамирования обычно являются постоянными, в то время как для контента на основе объектов, в которой отмеченная позиция объекта обеспечивается посредством переменных по времени метаданных объектов, коэффициенты  $g_{i,s}$  усиления, следовательно, могут быть переменными по времени. Это уравнение также может быть сформулировано в области преобразования или подполос, и в этом случае множество коэффициентов  $g_{i,s}[k]$  усиления используется для каждого элемента разрешения по частоте/полосы  $k$ , и, таким образом, коэффициенты  $g_{i,s}[k]$  усиления могут быть сделаны переменными по частоте:

$$z_s[k,b] = \sum_i g_{i,s}[k] x_i[k,b]$$

Матрица 42 декодера, игнорируя на данный момент декорреляторы, производит:

$$\begin{bmatrix} \hat{Y}_1[k,b] \\ \vdots \\ \hat{Y}_j[k,b] \end{bmatrix}^T = \begin{bmatrix} Z_1[k,b] \\ \vdots \\ Z_s[k,b] \end{bmatrix}^T M[p(b)]$$

или в матричной формулировке, опуская индекс  $b$  подполосы  $b$  и индекс  $p$  параметрической полосы для ясности:

$$Y = ZM$$

$$Z = XG$$

Критерием вычисления матричных коэффициентов  $M$  кодером является минимизация среднеквадратической ошибки  $E$ , которая представляет квадратическую ошибку между выходными данными  $\hat{Y}_j$  декодера и первоначальными входными объектами/каналами  $X_j$ :

$$E = \sum_{j,k,b} (\hat{Y}_j[k,b] - X_j[k,b])^2$$

Тогда матричные коэффициенты, которые минимизируют  $E$ , заданы в матричном представлении как:

$$M = (Z^* Z + \epsilon I)^{-1} Z^* X$$

где  $\epsilon$  - константа регуляризации, и  $(*)$  - оператор комплексно-сопряженного транспонирования. Эта операция может быть выполнена для каждой параметрической полосы  $p$  независимо, получая матрицу  $M[p(b)]$ .

### Предсказание с минимальной среднеквадратической ошибкой (MMSE) для преобразования представления

Помимо воссоздания объектов и/или каналов параметрические методики могут использоваться для преобразования одного представления в другое представление. Примером такого преобразования представления является преобразование звукозаписи в формате стерео, предназначенной для воспроизведения с помощью динамиков, в бинауральное представление для наушников, или наоборот.

Фиг. 4 иллюстрирует поток управления для способа 50 для одного такого преобразования представления. Аудиоданные объектов или каналов сначала обрабатываются в кодере 52 посредством набора 54 гибридных квадратурных зеркальных фильтров для анализа. Матрица  $G$  рендеризации динамиков вычисляется и применяется 55 к сигналам  $X_i$  объектов, сохраненным в запоминающем носителе 51, на основе метаданных объектов с использованием методики амплитудного панорамирования, чтобы получить в результате представление  $Z_s$  для динамиков в формате стерео. Это представление для динамиков может быть закодировано с помощью аудиокодера 57.

Кроме того, бинауральная матрица  $H$  рендеризации генерируется и применяется 58 с использованием базы 59 данных передаточных функций слухового аппарата (HRTF). Эта матрица  $H$  используется для вычисления бинауральных сигналов  $Y_j$ , которые позволяют воссоздать бинауральную звукозапись с использованием стереофонической звукозаписи для динамиков в качестве входных данных. Матричные коэффициенты  $M$  кодируются аудиокодером 57.

Переданная информация передается от кодера 52 к декодеру 53, где она распаковывается 61 для включения компонентов  $M$  и  $Z_s$ . Если в качестве системы воспроизведения используются динамики, представление для динамиков воспроизводится с использованием информации  $Z_s$  каналов, и, следовательно, матричные коэффициенты  $M$  отбрасываются. С другой стороны, для воспроизведения с помощью наушников представление для динамиков сначала преобразовывается 62 в бинауральное представление посредством применения переменной по времени и частоте матрицы  $M$  перед применением гибридных фильтров QMF для синтеза и воспроизведением 60.

Если желаемые бинауральные выходные данные из элемента 62 матрицирования записаны в матричном представлении как:

$$Y = XH$$

тогда матричные коэффициенты  $M$  могут быть получены в кодере 52 как:

$$M = (G^* X^* XG + \epsilon I)^{-1} G^* X^* XH$$

В этом применении коэффициенты матрицы  $H$  кодера, применяемой в блоке, обозначенном номером 58, обычно являются комплекснозначными, например, имеющими элемент модификации задержки или фазы, чтобы позволить восстановить разности времени между ушами, которые играют очень большую роль при восприятии для локализации источника звука в наушниках. Другими словами, бинауральная матрица  $H$  рендеризации является комплексной, и, таким образом, матрица  $M$  преобразования является комплексной. Для прозрачного для восприятия восстановления признаков локализации источника звука было показано, что желательна частотное разрешение, которое имитирует частотное разрешение слуховой системы человека (Breebaart 2010).

В упомянутых выше разделах для определения матричных коэффициентов  $M$  используется критерий минимальной среднеквадратической ошибки. Без потери общности аналогичным образом могут использоваться другие известные критерии или способы вычисления матричных коэффициентов, чтобы заменить или дополнить принцип минимальной среднеквадратической ошибки. Например, матричные коэффициенты  $M$  могут быть вычислены с использованием остаточных членов более высокого порядка или посредством минимизации нормы  $L1$  (например, критерий наименьших абсолютных отклонений). Кроме того, могут использоваться различные способы, в том числе методики неотрицательной факторизации или оптимизации, непараметрическая оценка, оценка максимального правдоподобия и т.п. Кроме того, матричные коэффициенты могут быть вычислены с использованием итерационных процессов или процессов градиентного спуска, методов интерполяции, эвристических методов, динамического программирования, машинного обучения, нечеткой оптимизации, имитации отжига или решений в замкнутой форме, и может использоваться методика анализа посредством синтеза. Наконец, что не менее важно, на оценку матричных коэффициентов могут быть наложены ограничения различным образом, например, посредством ограничения диапазона значений, условий регуляризации, наложения требований сохранения энергии и т.п.

### Требования преобразования и набора фильтров

В зависимости от применения и от того, должны ли быть воссозданы объекты или каналы, могут быть наложены определенные требования на частотное разрешение преобразования или набора фильтров для блока 41 набора фильтров на фиг. 3. В наиболее практическом применении частотное разрешение соответствует предполагаемому разрешению слуховой системы человека, чтобы дать наилучшее воспринимаемое качество звука для данного битрейта (определяемого несколькими параметрами) и сложности. Известно, что слуховая система человека может рассматриваться как набор фильтров с нелинейным частотным разрешением. Эти фильтры упоминаются как критические полосы (Zwicker, 1961), и они являются приблизительно логарифмическими по природе. На низких частотах критические полосы составляют

менее 100 Гц по ширине, в то время как на высоких частотах критические полосы могут составлять более 1 кГц по ширине.

Это нелинейное поведение может создать проблемы, когда речь идет о конструкции набора фильтров. Преобразования и наборы фильтров могут быть очень эффективно реализованы с использованием симметрии в их структуре обработки при условии, что частотное разрешение является постоянным по частоте.

Это подразумевает, что длина преобразования или количество подполос будут определены критической шириной полосы на низких частотах, и отображение элементов разрешения DFT на так называемые параметрические полосы может использоваться для имитации нелинейного частотного разрешения. Такой процесс отображения, например, разъяснен в литературе Breebaart et al., (2005) и Breebaart et al. (2010). Один недостаток этого подхода состоит в том, что требуется очень длинное преобразование, чтобы удовлетворить ограничение низкочастотной критической ширины полосы, в то время как преобразование является относительно длинным (или неэффективным) на высоких частотах. Альтернативное решение для расширения частотного разрешения на низких частотах состоит в том, чтобы использовать структуру гибридного набора фильтров. В такой структуре используется каскад из двух наборов фильтров, в котором второй набор фильтров расширяет разрешение первого, но только в нескольких самых нижних подполосах (Schuijers и др., 2004).

Фиг. 5 иллюстрирует одну форму структуры 41 гибридного набора фильтров, подобного изложенному в литературе Schuijers et al. Входной сигнал  $z[n]$  сначала обрабатывается набором 71 комплекснозначных квадратурных зеркальных фильтров (CQMF) для анализа. Затем сигналы подвергаются дискретизации с понижением с коэффициентом  $Q$ , например, в блоке 72, и в результате получаются сигналы  $Z[k,b]$  подполос, где  $k$  - индекс отсчета подполосы и  $b$  - индекс частоты подполосы. Кроме того, по меньшей мере один из полученных в результате сигналов подполосы обрабатывается вторым набором 74 фильтров (Найквиста), в то время как остальные сигналы подполос подвергаются задержке 75 для компенсации задержки, вносимой набором фильтров Найквиста. В этом конкретном примере каскад наборов фильтров дает в результате 8 подполос ( $b=1, \dots, 8$ ), которые отображаются на 6 параметрических полос  $p=(1, \dots, 6)$  с нелинейным частотным разрешением. Полосы 76 объединяются вместе, чтобы сформировать одну параметрическую полосу ( $p=6$ ).

Преимуществом этого подхода является более низкая сложность по сравнению с использованием единственного набора фильтров с намного большим количеством (более узких) подполос. Однако недостаток состоит в том, что задержка всей системы значительно увеличивается, и, следовательно, использование памяти также значительно больше, что вызывает увеличение потребления энергии.

#### Ограничения предшествующего уровня техники

Возвращаясь к фиг. 4 предлагается на предшествующем уровне техники использовать концепцию матрицирования 62, возможно дополненную использованием декорреляторов, чтобы воссоздавать каналы, объекты или сигналы представления  $\hat{Y}_j$  из множества базовых сигналов  $Z_s$ . Это приводит к следующей матричной формулировке для описания предшествующего уровня техники в обобщенном виде:

$$\begin{bmatrix} \hat{Y}_1[k,b] \\ \vdots \\ \hat{Y}_j[k,b] \end{bmatrix}_T = \begin{bmatrix} Z_1[k,b] \\ \vdots \\ Z_s[k,b] \\ D_1[k,b] \\ \vdots \\ D_M[k,b] \end{bmatrix}_T M[p(b)]$$

Матричные коэффициенты  $M$  либо передаются непосредственно от кодера к декодеру, либо выводятся из параметров локализации источника звука, например, как описано в литературе Breebaart et al. 2005 для параметрического кодирования стерео или Herge et al., (2008) для многоканального декодирования. Кроме того, этот подход также может использоваться для восстановления разности фаз между каналами посредством использования комплекснозначных матричных коэффициентов (см., например, Breebaart et al., 2010 и Breebaart, 2005).

Как проиллюстрировано на фиг. 6, на практике использование комплекснозначных матричных коэффициентов подразумевает, что желаемая задержка 80 представлена кусочным постоянным приближением 81 фазы. В предположении, что желаемая характеристика фазы представляет собой чистую задержку 80 с линейно уменьшающейся фазой с ростом частоты (пунктирная линия), операция комплекснозначного матрицирования предшествующего уровня техники приводит в результате к кусочному постоянному приближению 81 (сплошная линия). Приближение может быть улучшено посредством увеличения разрешения матрицы  $M$ . Однако это имеет два важных недостатка. Требуется увеличение разрешения набора фильтров, что приводит к увеличению использования памяти, более высокой вычислительной сложности, более длительному времени ожидания и, таким образом, большему потреблению энергии. Также требуется отправка большего количества параметров, что приводит к более высокому битрейту.

Все эти недостатки являются преимущественно проблематичными для мобильных устройств и устройств с батарейным питанием. Было бы выгодно иметь более оптимальное доступное решение.

### Сущность изобретения

Задача изобретения в его предпочтительной форме состоит в том, чтобы обеспечить улучшенную форму кодирования и декодирования аудиосигналов для воспроизведения в разных представлениях.

В соответствии с первым аспектом настоящего изобретения предложен способ представления второго представления аудиоканалов или объектов как потока данных, причем способ содержит следующие этапы: (а) обеспечение множества базовых сигналов, базовые сигналы представляют первое представление аудиоканалов или объектов; (б) обеспечение множества параметров преобразования, параметры преобразования предназначены для преобразования первого представления во второе представление; параметры преобразования также заданы по меньшей мере для двух частотных полос и включают в себя множество параметров матрицы свертки с несколькими отводами по меньшей мере для одной из частотных полос.

Множество коэффициентов фильтра может представлять фильтр с конечной импульсной характеристикой (КИХ; FIR). Множество базовых сигналов предпочтительно разделяется на последовательность временных сегментов, и множество параметров преобразования может быть обеспечено для каждого временного сегмента. Коэффициенты фильтра могут включать в себя по меньшей мере один коэффициент, который может являться комплекснозначным. Первое или второе представление могут быть предназначены для воспроизведения с помощью наушников.

В некоторых вариантах осуществления параметры преобразования, соответствующие более высоким частотам, не модифицируют фазу сигнала, в то время как для более низких частот параметры преобразования модифицируют фазу сигнала. Множество коэффициентов фильтра может быть предпочтительно выполнено с возможностью обработки матрицы свертки с несколькими отводами. Множество коэффициентов фильтра может быть предпочтительно использовано для обработки низкочастотной полосы.

Множество базовых сигналов и множество параметров преобразования предпочтительно объединяются для формирования потока данных. Параметры преобразования могут включать в себя матричные коэффициенты высокочастотных аудиоданных для матричной манипуляции высокочастотной части множества базовых сигналов. В некоторых вариантах осуществления для среднечастотной части высокочастотной части множества базовых сигналов матричная манипуляция предпочтительно может включать в себя комплекснозначные параметры преобразования.

В соответствии с дополнительным аспектом настоящего изобретения предложен декодер для декодирования закодированного аудиосигнала, закодированный аудиосигнал включает в себя первое представление, включающее в себя множество базовых аудиосигналов, предназначенных для воспроизведения аудиоданных в первом формате представления аудиоданных; и множество параметров преобразования для преобразования базовых аудиосигналов в первом формате представления во второй формат представления, причем параметры преобразования включают в себя, по меньшей мере, высокочастотные параметры преобразования аудиоданных и низкочастотные параметры преобразования аудиоданных, низкочастотные параметры преобразования включают в себя параметры матрицы свертки с несколькими отводами, при этом декодер включает в себя первый блок разделения для разделения множества базовых аудиосигналов и множества параметров преобразования, блок матричного умножения для применения параметров матрицы свертки с несколькими отводами к низкочастотным компонентам базовых аудиосигналов; для применения свертки к низкочастотным компонентам, производя свернутые низкочастотные компоненты; и блок скалярного умножения для применения высокочастотных параметров преобразования аудиоданных к высокочастотным компонентам базовых аудиосигналов, чтобы произвести скалярные высокочастотные компоненты; выходной набор фильтров для объединения свернутых низкочастотных компонентов и скалярных высокочастотных компонентов, чтобы произвести выходной сигнал во временной области во втором формате представления.

Блок матричного умножения может модифицировать фазу низкочастотных компонентов базовых аудиосигналов. В некоторых вариантах осуществления параметры преобразования матрицы свертки с несколькими отводами предпочтительно являются комплекснозначными. Высокочастотные параметры преобразования аудиоданных также предпочтительно являются комплекснозначными. Множество параметров преобразования дополнительно может содержать действительные значения параметров преобразования аудиоданных более высокой частоты. В некоторых вариантах осуществления декодер может дополнительно включать в себя фильтры для разделения базовых аудиосигналов на низкочастотные компоненты и высокочастотные компоненты.

В соответствии с дополнительным аспектом настоящего изобретения предложен способ декодирования закодированного аудиосигнала, причем закодированный аудиосигнал включает в себя первое представление, включающее в себя множество базовых аудиосигналов, предназначенных для воспроизведения аудиоданных в первом формате представления аудиоданных; и множество параметров преобразования для преобразования базовых аудиосигналов в первом формате представления во второй формат представления, причем параметры преобразования включают в себя, по меньшей мере, высокочастотные параметры преобразования аудиоданных и низкочастотные параметры преобразования аудиоданных, низкочастотные параметры преобразования включают в себя параметры матрицы свертки с несколькими

отводами, при этом способ включает в себя следующие этапы: свертка низкочастотных компонентов базовых аудиосигналов с помощью низкочастотных параметров преобразования, чтобы произвести свернутые низкочастотные компоненты; умножение высокочастотных компонентов базовых аудиосигналов на высокочастотные параметры преобразования, чтобы произвести умноженные высокочастотные компоненты; объединение свернутых низкочастотных компонентов и умноженных высокочастотных компонентов, чтобы произвести частотные компоненты выходного аудиосигнала для воспроизведения во втором формате представления.

В некоторых вариантах осуществления закодированный сигнал может содержать несколько временных сегментов, способ дополнительно предпочтительно может включать в себя следующие этапы: интерполяция параметров преобразования нескольких временных сегментов закодированного сигнала, чтобы произвести интерполированные параметры преобразования, включающие в себя интерполированные низкочастотные параметры преобразования аудиоданных; и свертка нескольких временных сегментов низкочастотных компонентов базовых аудиосигналов с помощью интерполированных низкочастотных параметров преобразования аудиоданных, чтобы произвести несколько временных сегментов свернутых низкочастотных компонентов.

Множество параметров преобразования закодированного аудиосигнала может являться предпочтительно переменным по времени, и способ дополнительно предпочтительно может включать в себя следующие этапы: свертка низкочастотных компонентов с помощью низкочастотных параметров преобразования для нескольких временных сегментов, чтобы произвести несколько множеств промежуточных свернутых низкочастотных компонентов; интерполяция нескольких множеств промежуточных свернутых низкочастотных компонентов, чтобы произвести свернутые низкочастотные компоненты.

Интерполяция может использовать метод наложения и суммирования нескольких множеств промежуточных свернутых низкочастотных компонентов.

#### **Краткое описание чертежей**

Теперь будут описаны варианты осуществления изобретения лишь в качестве примера со ссылкой на прилагаемые чертежи.

Фиг. 1 иллюстрирует схематический обзор процесса свертки HRIR для двух объектов источников, каждый канал или объект обрабатывается парой характеристик HRIR/BRIR;

фиг. 2 схематично иллюстрирует обобщенную параметрическую систему кодирования, поддерживающую каналы и объекты;

фиг. 3 схематично более подробно иллюстрирует одну форму блока 30 воссоздания каналов или объектов, показанного на фиг. 2;

фиг. 4 иллюстрирует поток данных способа преобразования представления для динамиков в формате стерео в бинауральное представление наушников;

фиг. 5 схематично иллюстрирует структуру гибридного набора фильтров для анализа в соответствии с предшествующим уровнем техники;

фиг. 6 иллюстрирует сравнение желаемой (пунктирная линия) и фактической (сплошная линия) характеристик фазы, полученных на предшествующем уровне техники;

фиг. 7 схематично демонстрирует иллюстративную систему набора фильтров и отображения параметров кодера в соответствии с вариантом осуществления изобретения;

фиг. 8 схематично иллюстрирует набор фильтров и отображение параметров декодера в соответствии с вариантом осуществления; и

фиг. 9 иллюстрирует кодер для преобразования представления в формате стерео в бинауральное представление.

Фиг. 10 схематично иллюстрирует декодер для преобразования представления в формате стерео в бинауральное представление.

#### **Список литературы**

- Wightman, F. L., and Kistler, D. J. (1989). "Headphone simulation of free-field listening. I. Stimulus synthesis," J. Acoust. Soc. Am. 85, 858-867.
- Schuijers, Erik, et al. (2004). "Low complexity parametric stereo coding." Audio Engineering Society Convention

116. Audio Engineering Society.

Herre, J., Kjörling, K., Breebaart, J., Faller, C., Disch, S., Purnhagen, H.,... & Chong, K. S. (2008). MPEG surround—the ISO/MPEG standard for efficient and compatible multichannel audio coding. *Journal of the Audio Engineering Society*, 56(11), 932-955.

Herre, J., Purnhagen, H., Koppens, J., Hellmuth, O., Engdegård, J., Hilpert, J., & Oh, H. O. (2012). MPEG Spatial Audio Object Coding—the ISO/MPEG standard for efficient coding of interactive audio scenes. *Journal of the Audio Engineering Society*, 60(9), 655-673.

Brandenburg, K., & Stoll, G. (1994). ISO/MPEG-1 audio: A generic standard for coding of high-quality digital audio. *Journal of the Audio Engineering Society*, 42(10), 780-792.

Bosi, M., Brandenburg, K., Quackenbush, S., Fielder, L., Akagiri, K., Fuchs, H., & Dietz, M. (1997). ISO/IEC MPEG-2 advanced audio coding. *Journal of the Audio engineering society*, 45(10), 789-814.

Andersen, R. L., Crockett, B. G., Davidson, G. A., Davis, M. F., Fielder, L. D., Turner, S. C.,... & Williams, P. A. (2004, October). Introduction to Dolby digital plus, an enhancement to the Dolby digital coding system. In *Audio Engineering Society Convention 117*. Audio Engineering Society.

Zwicker, E. (1961). Subdivision of the audible frequency range into critical bands (Frequenzgruppen). *The Journal of the Acoustical Society of America*, (33 (2)), 248.

Breebaart, J., van de Par, S., Kohlrausch, A., & Schuijers, E. (2005). Parametric coding of stereo audio. *EURASIP Journal on Applied Signal Processing*, 2005, 1305-1322.

Breebaart, J., Nater, F., & Kohlrausch, A. (2010). Spectral and spatial parameter resolution requirements for parametric, filter-bank-based HRTF processing. *Journal of the Audio Engineering Society*, 58(3), 126-140.

Breebaart, J., van de Par, S., Kohlrausch, A., & Schuijers, E. (2005). Parametric coding of stereo audio. *EURASIP Journal on Applied Signal Processing*, 2005, 1305-1322.

#### **Подробное описание**

Этот предпочтительный вариант осуществления обеспечивает способ воссоздания объектов, каналов или "представлений" из множества базовых сигналов, который может быть применен в наборах фильтров с низким частотным разрешением. Одним примером является преобразование представления в формате стерео в бинауральное представление, предназначенное для воспроизведения с помощью наушников, которое может быть применено без (гибридного) набора фильтров Найквиста. Сокращенное частотное разрешение декодера компенсируется матрицей свертки с несколькими отводами. Эта матрица свертки требует лишь небольшое количество отводов (например, два) и в практических случаях требуется только на низких частотах. Этот способ (1) сокращает вычислительную сложность декодера, (2) сокращает использование памяти декодера и (3) сокращает битрейт параметров.

В предпочтительном варианте осуществления обеспечены система и способ для преодоления нежелательной вычислительной сложности и требований к памяти на стороне декодера. Это реализовано посредством обеспечения высокого частотного разрешения в кодере при использовании ограниченного (более низкого) частотного разрешения в декодере (например, использование частотного разрешения, которое значительно хуже, чем использованное в соответствующем кодере) и использования матрицы (свертки) с несколькими отводами для компенсации сокращенного частотного разрешения декодера.

Как правило, поскольку матрица с высоким частотным разрешением требуется только на низких частотах, матрица (свертки) с несколькими отводами может использоваться на низких частотах, в то время как традиционная (без состояний) матрица может использоваться для остальных (более высоких)

частот. Другими словами, на низких частотах матрица представляет множество КИХ-фильтров (FIR-фильтров), действующих на каждую комбинацию ввода и вывода, в то время как на высоких частотах используется матрица без состояний.

#### Набор фильтров и отображение параметров кодера

Фиг. 7 демонстрирует иллюстративную систему 90 набора фильтров и отображения параметров кодера в соответствии с вариантом осуществления. В этом иллюстративном варианте 90 осуществления 8 подполос ( $b=1, \dots, 8$ ), например, обозначенных номером 91, первоначально генерируются посредством гибридного (каскадного) набора 92 фильтров и набора 93 фильтров Найквиста.

Затем первые четыре подполосы отображаются 94 на одну и ту же параметрическую полосу ( $p=1$ ) для вычисления матрицы свертки  $M[k,p=1]$ , например, матрица теперь имеет дополнительный индекс  $k$ . Остальные подполосы ( $b=5, \dots, 8$ ) отображаются на параметрические полосы ( $p=2,3$ ) с использованием матрицы  $M[p(b)]$  95, 96 без состояний.

#### Набор фильтров и отображение параметров декодера

Фиг. 8 демонстрирует соответствующую иллюстративную систему 100 набора фильтров и отображения параметров декодера. В отличие от кодера, набор фильтров Найквиста не присутствует, и при этом нет никаких задержек для компенсации задержки набора фильтров Найквиста. Набор фильтров 101 для анализа декодера генерирует только 5 подполос ( $b=1, \dots, 5$ ), например, обозначенных номером 102, которые подвергнуты дискретизации с понижением с коэффициентом  $Q$ . Первая подполоса обрабатывается матрицей 103 свертки  $M[k,p=1]$ , в то время как остальные полосы обрабатываются матрицами 104, 105 без состояния в соответствии с предшествующим уровнем техники.

Хотя приведенный выше пример применяет набор фильтров Найквиста в кодере 90 и соответствующую матрицу свертки только для первой подполосы SQMF в декодере 100, тот же самый процесс может быть применен ко множеству подполос, не обязательно ограниченных только самой нижней подполосой (подполосами).

#### Вариант осуществления кодера

Один вариант осуществления, который является особенно полезным, состоит в преобразовании представления для динамиков в бинауральное представление. Фиг. 9 иллюстрирует кодер 110, использующий предложенный способ для преобразования представления. Множество входных каналов или объектов  $x_i[n]$  сначала преобразовывается с использованием набора 111 фильтров. Набор 111 фильтров представляет собой гибридный набор комплексных квадратурных зеркальных фильтров (HCQMF), но в равной степени могут использоваться другие структуры наборов фильтров. Полученные в результате представления подполос  $X_i[k,p]$  обрабатываются дважды 112, 113.

Во-первых (113), чтобы сгенерировать множество базовых сигналов  $Z_s[k,b]$  113, предназначенных для выходных данных кодера. Эти выходные данные, например, могут быть сгенерированы с использованием методики амплитудного панорамирования, чтобы полученные в результате сигналы были предназначены для воспроизведения с помощью динамиков.

Во-вторых (112), чтобы генерировать множество желаемых преобразованных сигналов  $Y_j[k,b]$  112. Эти выходные данные, например, могут быть сгенерированы с использованием обработки HRIR, чтобы полученные в результате сигналы были предназначены для воспроизведения с помощью наушников. Такая обработка HRIR может использоваться в области набора фильтров, но в равной степени может быть выполнена во временной области посредством свертки HRIR. Характеристики HRIR получают из базы 114 данных. Матрица  $M[k,p]$  свертки затем получается посредством подачи базовых сигналов  $Z_s[k,b]$  через линию 116 задержки с отводами. Каждый из отводов линий задержки служит в качестве дополнительных входов в модуль 115 предсказания MMSE. Этот модуль предсказания MMSE вычисляет матрицу  $M[k,p]$  свертки, которая минимизирует ошибку между желаемыми преобразованными сигналами  $Y_j[k,b]$  и выходными данными декодера 100 на фиг. 8 посредством применения матриц свертки. После этого матричные коэффициенты  $M[k,p]$  заданы как:

$$M = (Z^* Z + \epsilon I)^{-1} Z^* Y$$

В этой формулировке матрица  $Z$  содержит все входы линий задержки с отводами.

Взяв первоначально случай для воссоздания одного сигнала  $\hat{Y}[k]$  для данной подполосы  $b$ , когда имеется  $A$  входов от линий задержки с отводами, имеем:

$$Z = \begin{bmatrix} Z_1[0,b] & \dots & Z_1[-(A-1),b] & Z_s[0,b] & \dots & Z_s[-(A-1),b] \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ Z_1[K-1,b] & \dots & Z_1[K-1-(A-1),b] & Z_s[K-1,b] & \dots & Z_s[K-1-(A-1),b] \end{bmatrix}$$

$$Y = \begin{bmatrix} Y_1[0,b] \\ \vdots \\ Y_1[K-1,b] \end{bmatrix}$$

$$M = \begin{bmatrix} m_1[0,b] & \dots & m_s[0,b] \\ \vdots & \ddots & \vdots \\ m_1[A-1,b] & \dots & m_s[A-1,b] \end{bmatrix} = (Z^* Z + \epsilon I)^{-1} Z^* Y$$

Полученные в результате коэффициенты  $M[k,p]$  матрицы свертки квантуются, кодируются и передаются вместе с базовыми сигналами  $z_s[n]$ . Тогда декодер может использовать процесс свертки для вос-

создания  $\hat{Y}[k,b]$  из входных сигналов  $Z_s[k,b]$ :

$$\hat{Y}[k,b] = \sum_s Z_s[k,b] * m_s[., b]$$

или в другой записи с использованием выражения свертки:

$$\hat{Y}[k,b] = \sum_s \sum_{a=0}^{A-1} Z_s[k-a,b] m_s[a,b]$$

Подход свертки может быть комбинирован с линейным (без состояний) матричным процессом.

Дополнительное различие может быть сделано между комплекснозначным и действительнозначным матрицированием без состояний. На низких частотах (как правило, ниже 1 кГц) процесс свертки ( $A > 1$ ) является предпочтительным, чтобы обеспечить возможность точного воссоздания свойств между каналами в соответствии с перцепционной шкалой частот. На средних частотах, приблизительно до 2 или 3 кГц, слуховая система человека чувствительна к разностям фаз между каналами, но не требует очень высокого частотного разрешения для воссоздания такой фазы. Это подразумевает, что комплекснозначная матрица с одним отводом (без состояний) является достаточной. Для более высоких частот слуховая система человека фактически не чувствительна к фазе с тонкой структурой форм волны, и применение действительнозначной матрицы без состояний является достаточным. С ростом частот количество выходов набора фильтров, отображаемых на параметрическую полосу, как правило, увеличивается, чтобы отразить нелинейное частотное разрешение слуховой системы человека.

В другом варианте осуществления первое и второе представления в кодере меняются местами, например, первое представление предназначено для воспроизведения с помощью наушников, и второе представление предназначено для воспроизведения с помощью динамиков. В этом варианте осуществления представление для динамиков (второе представление) генерируется посредством применения зависимых от времени параметров преобразования по меньшей мере в двух частотных полосах к первому представлению, причем параметры преобразования дополнительно задаются как включающие в себя множество коэффициентов фильтра по меньшей мере для одной из частотных полос.

В некоторых вариантах осуществления первое представление может быть разделено во времени на последовательность сегментов с отдельным множеством параметров преобразования для каждого сегмента. При дальнейшем уточнении, когда параметры преобразования сегментов недоступны, параметры могут быть интерполированы на основе предыдущих коэффициентов.

#### Вариант осуществления декодера.

Фиг. 10 иллюстрирует вариант осуществления декодера 120. Входной битовый поток 121 разделен на битовый поток 131 базовых сигналов и данные 124 параметров преобразования. Далее декодер 123 базовых сигналов декодирует базовые сигналы  $z[n]$ , которые затем обрабатываются набором 125 фильтров для анализа. Полученные в результате сигналы  $Z[k,b]$  в частотной области с подполосой  $b=1, \dots, 5$  обрабатываются блоками 126, 129 и 130 матричного умножения. В частности, блок 126 матричного умножения применяет комплекснозначную матрицу  $M[k, p=1]$  свертки к сигналу  $Z[k, b=1]$  в частотной области. Кроме того, блок 129 матричного умножения применяет комплекснозначные коэффициенты  $M[p=2]$  матрицы с одним отводом к сигналу  $Z[k, b=2]$ . Наконец, блок 130 матричного умножения применяет действительнозначные матричные коэффициенты  $M[p=3]$  к сигналам  $Z[k, b=3, \dots, 5]$  в частотной области. Выходные сигналы блока матричного умножения преобразовываются в выходные данные 128 во временной области посредством набора 127 фильтров для синтеза. Ссылки на  $z[n]$ ,  $Z[k]$  и т.д. относятся ко множеству базовых сигналов, а не к какому-либо конкретному базовому сигналу. Таким образом,  $z[n]$ ,  $Z[k]$  и т.д. могут быть интерпретированы как  $z_s[n]$ ,  $Z_s[k]$  и т.д., где  $0 \leq s < N$ , и  $N$  - количество базовых сигналов.

Другими словами, блок 126 матричного умножения определяет выходные отсчеты подполосы  $b=1$  выходного сигнала  $\hat{Y}_j[k]$  из взвешенных комбинаций текущих отсчетов подполосы  $b=1$  базовых сигналов  $Z[k]$  и предыдущих отсчетов подполосы  $b=1$  базовых сигналов  $Z[k]$  (например,  $Z[k-a]$ , где  $0 < a < A$ , и  $A$  больше 1). Весовые коэффициенты, используемые для определения выходных отсчетов подполосы  $b=1$  выходного сигнала  $\hat{Y}_j[k]$ , соответствуют комплекснозначной матрице  $M[k, p=1]$  свертки для сигнала.

Кроме того, блок 129 матричного умножения определяет выходные отсчеты подполосы  $b=2$  выходного сигнала  $\hat{Y}_j[k]$  из взвешенных комбинаций текущих отсчетов подполосы  $b=2$  базовых сигналов  $Z[k]$ . Весовые коэффициенты, используемые для определения выходных отсчетов подполосы  $b=2$  выходного сигнала  $\hat{Y}_j[k]$ , соответствуют комплекснозначным коэффициентам  $M[p=2]$  матрицы с одним отводом.

Наконец, блок 130 матричного умножения определяет выходные отсчеты подполос  $b=3, \dots, 5$  выходного сигнала  $\hat{Y}_j[k]$  из взвешенных комбинаций текущих отсчетов подполос  $b=3, \dots, 5$  базовых сигналов  $Z[k]$ . Весовые коэффициенты, используемые для определения выходных отсчетов подполос  $b=3, \dots, 5$  выходного сигнала  $\hat{Y}_j[k]$ , соответствуют действительнозначным матричным коэффициентам  $M[p=3]$ .

В некоторых случаях декодер 123 базовых сигналов может воздействовать на сигналы с таким же частотным разрешением, которое обеспечено набором 125 фильтров для анализа. В таких случаях декодер 125 базовых сигналов может быть выполнен с возможностью выдавать сигналы  $Z[k]$  в частотной области, а не сигналы  $z[n]$  во временной области, в этом случае набор 125 фильтров для анализа может быть опущен. Кроме того, в некоторых случаях может быть предпочтительно применить комплексно-

значные коэффициенты матрицы с одним отводом вместо действительных матричных коэффициентов к сигналам  $Z[k, b=3, \dots, 5]$ .

На практике матричные коэффициенты  $M$  могут обновляться со временем; например, посредством привязки индивидуальных кадров базовых сигналов к матричным коэффициентам  $M$ . В качестве альтернативы или дополнительно матричные коэффициенты  $M$  дополняются метками времени, которые указывают, в какое время или интервал базовых сигналов  $z[n]$  должны быть применены матрицы. Чтобы сократить битрейт передачи, связанный с обновлениями матриц, количество обновлений оптимально ограничено, что приводит к редкому распределению обновлений матриц. Такие нечастые обновления матриц требуют специальной обработки, чтобы гарантировать плавные переходы от одного экземпляра матрицы к следующему. Матрицы  $M$  могут обеспечиваться с привязанными заданными сегментами (кадрами) времени и/или частотными регионами базовых сигналов  $Z$ . Декодер может использовать множество методов интерполяции, чтобы гарантировать плавный переход от последующих экземпляров матрицы  $M$  по времени. Один пример такого метода интерполяции состоит в вычислении накладываемых оконных кадров сигналов  $Z$  и вычислении соответствующего множества выходных сигналов  $Y$  для каждого такого кадра с использованием матричных коэффициентов  $M$ , привязанных к этому конкретному кадру. Последующие кадры затем могут быть агрегированы с использованием методики наложения с суммированием, обеспечивающей плавный переход со взаимным наложением. В качестве альтернативы декодер может принять метки времени, привязанные к матрице  $M$ , которые описывают желаемые матричные коэффициенты в заданные моменты времени. Для отсчетов аудиоданных между метками времени матричные коэффициенты матрицы  $M$  могут быть интерполированы с использованием линейной, кубической, ограниченной по полосе интерполяции или других средств для интерполяции, чтобы гарантировать плавные переходы. Помимо интерполяции по времени подобные методики могут использоваться для интерполяции матричных коэффициентов по частоте.

Таким образом, настоящий документ описывает способ (и соответствующий кодер 90) для представления второго представления аудиоканалов или объектов  $X_i$  как потока данных, который должен быть передан или обеспечен соответствующему декодеру 100. Способ содержит этап обеспечения базовых сигналов  $Z_s$ , упомянутые базовые сигналы представляют первое представление аудиоканалов или объектов  $X_i$ . Как описано в общих чертах выше, базовые сигналы  $Z_s$  могут быть определены на основе аудиоканалов или объектов  $X_i$  с использованием сначала рендеризации параметров  $G$  (а именно с использованием первой матрицы усиления, например, для амплитудного панорамирования). Первое представление может быть предназначено для воспроизведения с помощью динамиков или для воспроизведения с помощью наушников. С другой стороны, второе представление может быть предназначено для воспроизведения с помощью наушников или для воспроизведения с помощью динамиков. Таким образом, может быть выполнено преобразование из воспроизведения с помощью динамиков в воспроизведение с помощью наушников (или наоборот).

Способ дополнительно содержит обеспечение параметров  $M$  преобразования (а именно одной или более матриц преобразования), упомянутые параметры  $M$  преобразования предназначены для преобразования базовых сигналов  $Z_s$  упомянутого первого представления в выходные сигналы  $\hat{Y}_j$  упомянутого второго представления. Параметры преобразования могут быть определены, как описано в общих чертах в настоящем документе. В частности, желаемые выходные сигналы  $Y_j$  для второго представления могут быть определены из аудиоканалов или объектов  $X_i$  с использованием вторых параметров  $H$  рендеризации (как описано в общих чертах в настоящем документе). Параметры  $M$  преобразования могут быть определены посредством минимизации отклонения выходных сигналов  $\hat{Y}_j$  от желаемых выходных сигналов  $Y_j$  (например, с использованием критерия минимальной среднеквадратической ошибки).

Также, в частности, параметры  $M$  преобразования могут быть определены в области подполос (т.е., для разных частотных полос). С этой целью базовые сигналы  $Z[k, b]$  в области подполос могут быть определены для частотных полос  $B$  с использованием набора 92, 93 фильтров кодера. Количество  $B$  частотных полос больше одного, например,  $B$  равно или больше 4, 6, 8, 10. В примерах, описанных в настоящем документе,  $B=8$  или  $B=5$ . Как описано в общих чертах выше, набор 92, 93 фильтров кодера может содержать гибридный набор фильтров, который обеспечивает, что низкочастотные полосы из  $B$  частотных полос имеют более высокое частотное разрешение, чем высокочастотные полосы из  $B$  частотных полос. Кроме того, могут быть определены желаемые выходные сигналы  $Y[k, b]$  в области подполос для  $B$  частотных полос. Параметры  $M$  преобразования для одной или более частотных полос могут быть определены посредством минимизации отклонения выходных сигналов  $\hat{Y}_j$  от желаемых выходных сигналов  $Y_j$  в одной или более частотных полос (например, с использованием критерия минимальной среднеквадратической ошибки).

Каждый из параметров  $M$  преобразования, таким образом, может быть задан для по меньшей мере двух частотных полос (а именно для  $B$  частотных полос). Кроме того, параметры преобразования могут включать в себя множество параметров матрицы свертки с несколькими отводами по меньшей мере для одной из частотных полос.

Таким образом, описан способ (и соответствующий декодер) для определения выходных сигналов второго представления аудио каналов/объектов из базовых сигналов первого представления аудио кана-

лов/объектов. Первое представление может использоваться для воспроизведения с помощью динамиков, и второе представление может использоваться для воспроизведения с помощью наушников (или наоборот). Выходные сигналы определяются с использованием параметров преобразования для разных частотных полос, причем параметры преобразования по меньшей мере для одной из частотных полос содержат параметры матрицы свертки с несколькими отводами. В результате использования параметров матрицы свертки с несколькими отводами по меньшей мере для одной из частотных полос вычислительная сложность декодера 100 может быть сокращена, а именно посредством сокращения частотного разрешения набора фильтров, используемого декодером.

Например, определение выходного сигнала для первой частотной полосы с использованием параметров матрицы свертки с несколькими отводами может содержать определение текущего отсчета первой частотной полосы выходного сигнала как взвешенной комбинации текущего и одного или более предыдущих отсчетов первой частотной полосы базовых сигналов, причем весовые коэффициенты, используемые для определения взвешенной комбинации, соответствуют параметрам матрицы свертки с несколькими отводами для первой частотной полосы. Один или более параметров матрицы свертки с несколькими отводами для первой частотной полосы обычно являются комплекснозначными.

Кроме того, определение выходного сигнала для второй частотной полосы может содержать определение текущего отсчета второй частотной полосы выходного сигнала как взвешенной комбинации текущих отсчетов второй частотной полосы базовых сигналов (а не на основе предыдущих отсчетов второй частотной полосы базовых сигналов), причем весовые коэффициенты, используемые для определения взвешенной комбинации, соответствуют параметрам преобразования для второй частотной полосы. Параметры преобразования для второй частотной полосы могут быть комплекснозначными или в качестве альтернативы могут быть действительностнозначными.

В частности, одно и то же множество параметров матрицы свертки с несколькими отводами может быть определено по меньшей мере для двух смежных частотных полос из  $N$  частотных полос. Как проиллюстрировано на фиг. 7, одно множество параметров матрицы свертки с несколькими отводами может быть определено для частотных полос, обеспеченных набором фильтров Найквиста (т.е. для частотных полос, имеющих относительно высокое частотное разрешение). Посредством этого использование набора фильтров Найквиста в декодере 100 может быть опущено, тем самым сокращая вычислительную сложность декодера 100 (при поддержании качества выходных сигналов для второго представления).

Кроме того, один и тот же действительностнозначный параметр преобразования может быть определен по меньшей мере для двух смежных высокочастотных полос (как проиллюстрировано в контексте фиг. 7). Посредством этого вычислительная сложность декодера 100 может быть дополнительно сокращена (при поддержании качества выходных сигналов для второго представления). Интерпретация

Упоминание по тексту этого описания "одного варианта осуществления", "некоторых вариантов осуществления" или "варианта осуществления" означает, что конкретный признак, структура или характеристика, описанные в связи с вариантом осуществления, включены по меньшей мере в один вариант осуществления настоящего изобретения. Таким образом, появления фраз "в одном варианте осуществления", "в некоторых вариантах осуществления" или "в варианте осуществления" в различных местах по тексту этого описания не обязательно все относятся к одному и тому же варианту осуществления, но могут. Кроме того, конкретные признаки, структуры или характеристики могут быть объединены любым подходящим образом, как будет очевидно для специалиста в области техники из этого раскрытия, в одном или более вариантах осуществления.

Если не указано иначе, использование в настоящем документе порядковых числительных "первый", "второй", "третий" и т.д. для описания общего объекта указывает лишь на то, что упоминаются разные экземпляры подобных объектов, и не подразумевает, что описанные таким образом объекты должны находиться в данной последовательности ни во времени, ни в пространстве, ни по ранжированию, и ни каким-либо другим образом.

В приведенной ниже формуле изобретения и в описании в настоящем документе любой из терминов "содержащий", "состоящий из" или "который содержит" является неограничивающим термином, который означает включение по меньшей мере следующих элементов/признаков, но не исключая других. Таким образом, термин "содержащий", когда он используется в формуле изобретения, не должен интерпретироваться как ограничивающий для перечисленных после него средств, элементов или этапов. Например, контекст выражения "устройство, содержащее А и В" не должен быть ограничен устройствами, состоящими только из элементов А и В. Любой из терминов "включающий в себя" или "который включает в себя", используемых в настоящем документе, также является неограничивающим термином, который также означает включение по меньшей мере элементов/признаков, которые следуют за термином, но не исключая других. Таким образом, "включающий в себя" является синонимом и означает "содержащий".

Используемый в настоящем документе термин "иллюстративный" используется в смысле обеспечения примеров, в противоположность указанию на качество. Таким образом, "иллюстративный вариант осуществления" является вариантом осуществления, обеспеченным в качестве примера, в противоположность тому, чтобы обязательно являться вариантом осуществления иллюстративного качества.

Следует понимать, что в приведенном выше описании иллюстративных вариантов осуществления изобретения различные признаки изобретения иногда группируются в единственном варианте осуществления, фигуре или их описании с целью оптимизации раскрытия и помощи в понимании одного или более различных аспектов изобретения. Однако этот метод раскрытия не должен интерпретироваться как отражение намерения, что заявленное изобретение требует большего количества признаков, чем явно изложено в каждом пункте формулы изобретения. Вместо этого, как отражает последующая формула изобретения, аспекты изобретения лежат менее чем во всех признаках единственного предшествующего раскрытого варианта осуществления. Таким образом, формула изобретения после подробного описания тем самым явно включена в это подробное описание, и каждый пункт формулы изобретения представляет собой отдельный вариант осуществления этого изобретения.

Кроме того, хотя некоторые варианты осуществления, описанные в настоящем документе, включают в себя некоторые, но не другие признаки, включенные в другие варианты осуществления, подразумевается, что комбинации признаков разных вариантов осуществления находятся в рамках объема изобретения и формируют другие варианты осуществления, как будет понятно специалистам в области техники. Например, в следующей формуле изобретения любой из заявленных вариантов осуществления может использоваться в любой комбинации.

Кроме того, некоторые варианты осуществления описаны в настоящем документе как способ или комбинация элементов способа, которые могут быть реализованы процессором компьютерной системы или другим средством выполнения функции. Таким образом, процессор с необходимыми инструкциями для выполнения такого способа или элемента способа формирует средство для выполнения способа или элемента способа. Кроме того, описанный в настоящем документе элемент варианта осуществления устройства является примером средства для выполнения функции, выполняемой элементом с целью выполнения изобретения.

В обеспеченном в настоящем документе описании сформулированы многочисленные конкретные подробности. Однако подразумевается, что варианты осуществления изобретения могут быть осуществлены без этих конкретных подробностей. В других случаях известные способы, структуры и методики подробно не показаны, чтобы не препятствовать пониманию этого описания.

Аналогичным образом, следует отметить, что термин "присоединенный", когда он используется в формуле изобретения, не должен интерпретироваться как ограничиваемый только прямыми соединениями. Может использоваться термин "присоединенный" и "соединенный" вместе с их производными. Следует понимать, что эти термины не подразумеваются как синонимы друг для друга. Таким образом, контекст выражения "устройство А, присоединенное к устройству В" не должен быть ограничен устройствами или системами, в которых выход устройства А непосредственно соединен с входом устройства В. Это означает, что существует путь между выходом А и входом В, который может представлять собой путь, включающий в себя другие устройства или средства. "Присоединенный" может означать, что два или более элементов либо находятся в прямом физическом или электрическом контакте, либо эти два или более элементов не находятся в прямом контакте друг с другом, но все же сотрудничают или взаимодействуют друг с другом.

Таким образом, хотя были описаны предполагаемые предпочтительные варианты осуществления изобретения, специалисты в области техники поймут, что в них могут быть внесены другие и дополнительные модификации без отступления от сущности изобретения, и имеется в виду, что все такие изменения и модификации находятся в пределах объема изобретения. Например, любые приведенные выше формулы являются лишь репрезентативными для процедур, которые могут использоваться. Функциональность может быть добавлена или удалена из блок-схем, и операции могут меняться местами среди функциональных блоков. Этапы могут быть добавлены или удалены из способов, описанных в рамках объема настоящего изобретения.

Различные аспекты настоящего изобретения могут быть очевидны на основе следующих перечисленных иллюстративных вариантов осуществления (ЕЕЕ):

ЕЕЕ 1. Способ представления второго представления аудиоканалов или объектов как потока данных, при этом способ содержит следующие этапы:

(a) обеспечение множества базовых сигналов, упомянутые базовые сигналы представляют первое представление аудиоканалов или объектов;

(b) обеспечение множества параметров преобразования, причем упомянутые параметры преобразования предназначены для преобразования упомянутого первого представления в упомянутое второе представление; упомянутые параметры преобразования также заданы по меньшей мере для двух частотных полос и включают в себя множество параметров матрицы свертки с несколькими отводами по меньшей мере для одной из частотных полос.

ЕЕЕ 2. Способ согласно ЕЕЕ 1, в котором упомянутое множество коэффициентов фильтра представляет фильтр с конечной импульсной характеристикой (КИХ; FIR).

ЕЕЕ 3. Способ согласно любому предыдущему ЕЕЕ, в котором упомянутое множество базовых сигналов разделяется на последовательность временных сегментов, и множество параметров преобразования обеспечено для каждого временного сегмента.

EEE 4. Способ согласно любому предыдущему EEE, в котором упомянутые коэффициенты фильтра включают в себя по меньшей мере один коэффициент, который является комплекснозначным.

EEE 5. Способ согласно любому предыдущему EEE, в котором первое или второе представление предназначено для воспроизведения с помощью наушников.

EEE 6. Способ согласно любому предыдущему EEE, в котором параметры преобразования, соответствующие более высоким частотам, не модифицируют фазу сигнала, в то время как для более низких частот параметры преобразования модифицируют фазу сигнала.

EEE 7. Способ согласно любому предыдущему EEE, в котором упомянутое множество коэффициентов фильтра выполнено с возможностью обработки матрицы свертки с несколькими отводами.

EEE 8. Способ согласно EEE 7, в котором упомянутое множество коэффициентов фильтра используется для обработки низкочастотной полосы,

EEE 9. Способ согласно любому предыдущему EEE, в котором упомянутое множество базовых сигналов и упомянутое множество параметров преобразования объединяются для формирования упомянутого потока данных.

EEE 10. Способ согласно любому предыдущему EEE, в котором упомянутые параметры преобразования включают в себя матричные коэффициенты высокочастотных аудиоданных для матричной манипуляции высокочастотной части упомянутого множества базовых сигналов.

EEE 11. Способ согласно EEE 10, в котором для среднечастотной части высокочастотной части упомянутого множества базовых сигналов матричная манипуляция включает в себя комплекснозначные параметры преобразования.

EEE 12. Декодер для декодирования закодированного аудиосигнала, где закодированный аудиосигнал включает в себя

первое представление, включающее в себя множество базовых аудиосигналов, предназначенных для воспроизведения аудиоданных в первом формате представления аудиоданных; и

множество параметров преобразования для преобразования упомянутых базовых аудиосигналов в упомянутом первом формате представления во второй формат представления, причем упомянутые параметры преобразования включают в себя, по меньшей мере, высокочастотные параметры преобразования аудиоданных и низкочастотные параметры преобразования аудиоданных, упомянутые низкочастотные параметры преобразования включают в себя параметры матрицы свертки с несколькими отводами, декодер включает в себя первый блок разделения для разделения множества базовых аудиосигналов и множества параметров преобразования, блок матричного умножения для применения упомянутых параметров матрицы свертки с несколькими отводами к низкочастотным компонентам базовых аудиосигналов, чтобы применить свертку к низкочастотным компонентам, производя свернутые низкочастотные компоненты; и

блок скалярного умножения для применения упомянутых высокочастотных параметров преобразования аудиоданных к высокочастотным компонентам базовых аудиосигналов, чтобы произвести скалярные высокочастотные компоненты;

выходной набор фильтров для объединения упомянутых свернутых низкочастотных компонентов и упомянутых скалярных высокочастотных компонентов, чтобы произвести выходной сигнал во временной области в упомянутом втором формате представления.

EEE 13. Декодер согласно EEE 12, в котором упомянутый блок матричного умножения модифицирует фазу низкочастотных компонентов базовых аудиосигналов.

EEE 14. Декодер согласно EEE 12 или 13, в котором упомянутые параметры преобразования матрицы свертки с несколькими отводами являются комплекснозначными.

EEE 15. Декодер согласно любому из EEE 12-14, в котором упомянутые высокочастотные параметры преобразования аудиоданных являются комплекснозначными.

EEE 16. Декодер согласно EEE 15, в котором упомянутое множество параметров преобразования дополнительно содержит действительные параметры преобразования аудиоданных более высокой частоты.

EEE 17. Декодер согласно любому из EEE 12-16, дополнительно содержащий фильтры для разделения базовых аудиосигналов на упомянутые низкочастотные компоненты и упомянутые высокочастотные компоненты.

EEE 18. Способ декодирования закодированного аудиосигнала, где закодированный аудиосигнал включает в себя

первое представление, включающее в себя множество базовых аудиосигналов, предназначенных для воспроизведения аудиоданных в первом формате представления аудиоданных; и

множество параметров преобразования для преобразования упомянутых базовых аудиосигналов в упомянутом первом формате представления во второй формат представления, причем упомянутые параметры преобразования включают в себя по меньшей мере высокочастотные параметры преобразования аудиоданных и низкочастотные параметры преобразования аудиоданных, упомянутые низкочастотные параметры преобразования включают в себя параметры матрицы свертки с несколькими отводами,

способ включает в себя следующие этапы:

свертка низкочастотных компонентов базовых аудиосигналов с помощью низкочастотных параметров преобразования, чтобы произвести свернутые низкочастотные компоненты;

умножение высокочастотных компонентов базовых аудиосигналов на высокочастотные параметры преобразования, чтобы произвести умноженные высокочастотные компоненты;

объединение упомянутых свернутых низкочастотных компонентов и упомянутых умноженных высокочастотных компонентов, чтобы произвести частотные компоненты выходного аудиосигнала для воспроизведения во втором формате представления.

EEE 19. Способ согласно EEE 18, в котором упомянутый закодированный сигнал содержит несколько временных сегментов, причем упомянутый способ дополнительно включает в себя следующие этапы:

интерполяция параметров преобразования нескольких временных сегментов закодированного сигнала, чтобы произвести интерполированные параметры преобразования, включающие в себя интерполированные низкочастотные параметры преобразования аудиоданных; и

свертка нескольких временных сегментов низкочастотных компонентов базовых аудиосигналов с помощью интерполированных низкочастотных параметров преобразования аудиоданных, чтобы произвести несколько временных сегментов упомянутых свернутых низкочастотных компонентов.

EEE 20. Способ согласно EEE 18, в котором множество параметров преобразования упомянутого закодированного аудиосигнала является переменным по времени, и упомянутый способ дополнительно включает в себя следующие этапы:

свертка низкочастотных компонентов с помощью низкочастотных параметров преобразования для нескольких временных сегментов, чтобы произвести несколько множеств промежуточных свернутых низкочастотных компонентов;

интерполяция нескольких множеств промежуточных свернутых низкочастотных компонентов, чтобы произвести упомянутые свернутые низкочастотные компоненты.

EEE 21. Способ согласно либо EEE 19, либо EEE 20, в котором упомянутая интерполяция использует метод наложения и суммирования нескольких множеств промежуточных свернутых низкочастотных компонентов.

EEE 22. Способ согласно любому из EEE 18-21, дополнительно содержащий фильтрацию базовых аудиосигналов в упомянутые низкочастотные компоненты и упомянутые высокочастотные компоненты.

EEE 23. Машиночитаемый неизменяемый запоминающий носитель, включающий в себя программные команды для работы компьютера в соответствии со способом согласно любому из EEE 1-11 и 18-22.

Положение 1. Способ представления второго представления аудиоканалов или объектов в виде потока данных, при этом способ содержит этапы, на которых:

(а) обеспечивают базовые сигналы, каковые базовые сигналы представляют первое представление аудиоканалов или объектов;

(б) обеспечивают параметры преобразования, каковые параметры преобразования предназначены для преобразования базовых сигналов первого представления в выходные сигналы второго представления; каждый из упомянутых параметров преобразования задан для по меньшей мере двух частотных полос и включает в себя множество параметров матрицы свертки с несколькими отводами для по меньшей мере одной из частотных полос; первое представление предназначено для воспроизведения с помощью динамиков, и второе представление предназначено для воспроизведения с помощью наушников, или наоборот.

Положение 2. Способ согласно положению 1, в котором параметры матрицы свертки с несколькими отводами указывают фильтр с конечной импульсной характеристикой (КИХ; FIR).

Положение 3. Способ согласно любому предыдущему положению, в котором базовые сигналы разделены на последовательность временных сегментов, и параметры преобразования обеспечиваются для каждого временного сегмента.

Положение 4. Способ согласно любому предыдущему положению, в котором параметры матрицы свертки с несколькими отводами включают в себя по меньшей мере один коэффициент, который является комплекснозначным.

Положение 5. Способ согласно любому предыдущему положению, в котором упомянутое обеспечение базовых сигналов содержит этап, на котором определяют базовые сигналы из аудиоканалов или объектов с использованием первых параметров рендеризации;

при этом способ содержит этап, на котором определяют желаемые выходные сигналы для второго представления из аудиоканалов или объектов с использованием вторых параметров рендеризации; и

упомянутое обеспечение параметров преобразования содержит этап, на котором определяют параметры преобразования посредством минимизации отклонения выходных сигналов от желаемых выходных сигналов.

Положение 6. Способ согласно положению 5, в котором упомянутое определение параметров преобразования содержит этапы, на которых

определяют базовые сигналы в области подполос для В частотных полос с использованием набора фильтров кодера;

определяют желаемые выходные сигналы в области подполос для В частотных полос с использованием набора фильтров кодера; и

определяют одно и то же множество параметров матрицы свертки с несколькими отводами для по меньшей мере двух смежных частотных полос из В частотных полос.

Положение 7. Способ согласно положению 6, в котором упомянутый набор фильтров кодера содержит гибридный набор фильтров, который обеспечивает, что низкочастотные полосы из В частотных полос имеют более высокое частотное разрешение, чем высокочастотные полосы из В частотных полос; и

упомянутые две по меньшей мере смежных частотных полосы являются низкочастотными полосами.

Положение 8. Способ согласно положению 7, в котором упомянутое определение параметров преобразования содержит этап, на котором определяют один и тот же действительный параметр преобразования для упомянутых по меньшей мере двух смежных высокочастотных полос.

Положение 9. Способ согласно любому предыдущему положению, в котором упомянутые по меньшей мере две частотных полосы содержат полосу более низкой частоты и полосу более высокой частоты,

параметры преобразования, заданные для полосы более высокой частоты, не модифицируют фазу сигнала базовых сигналов, и

параметры преобразования, заданные для полосы более низкой частоты, модифицируют фазу сигнала базового сигнала.

Положение 10. Способ согласно любому предыдущему положению, в котором упомянутые параметры матрицы свертки с несколькими отводами используются для обработки низкочастотной полосы.

Положение 11. Способ согласно любому предыдущему положению, в котором упомянутые базовые сигналы и упомянутые параметры преобразования объединяются для формирования упомянутого потока данных.

Положение 12. Способ согласно любому предыдущему положению, в котором упомянутые параметры преобразования включают в себя матричные коэффициенты высокочастотных аудиоданных для матричной манипуляции высокочастотной частью упомянутых базовых сигналов.

Положение 13. Способ согласно положению 12, в котором для среднечастотной части высокочастотной части упомянутых базовых сигналов матричная манипуляция включает в себя комплекснозначные параметры преобразования.

Положение 14. Декодер для декодирования закодированного аудиосигнала, причем закодированный аудиосигнал включает в себя:

первое представление, включающее в себя базовые аудиосигналы, предназначенные для воспроизведения

закодированного аудиосигнала в первом формате представления аудиоданных; и

параметры преобразования для преобразования базовых аудиосигналов в первом формате представления в выходные сигналы второго формата представления, при этом параметры преобразования содержат высокочастотные параметры преобразования аудиоданных и низкочастотные параметры преобразования аудиоданных, причем низкочастотные параметры преобразования включают в себя параметры матрицы свертки с несколькими отводами, при этом первый формат представления предназначен для воспроизведения с помощью динамиков, и второй формат представления предназначен для воспроизведения с помощью наушников, или наоборот,

при этом декодер включает в себя

первый блок разделения для разделения базовых аудиосигналов и параметров преобразования,

блок матричного умножения для применения параметров матрицы свертки с несколькими отводами к низкочастотным компонентам базовых аудиосигналов, чтобы применить свертку к низкочастотным компонентам, производя свернутые низкочастотные компоненты;

блок скалярного умножения для применения высокочастотных параметров преобразования аудиоданных к высокочастотным компонентам базовых аудиосигналов, чтобы получить скалярные высокочастотные компоненты; и

выходной набор фильтров для объединения свернутых низкочастотных компонентов и скалярных высокочастотных компонентов, чтобы получить выходной сигнал во временной области второго формата представления.

Положение 15. Декодер согласно положению 14, в котором блок матричного умножения модифицирует фазу низкочастотных компонентов базовых аудиосигналов.

Положение 16. Декодер согласно положению 14 или 15, при этом параметры преобразования матрицы свертки с несколькими отводами являются комплекснозначными.

Положение 17. Декодер согласно любому из положений 14-16, при этом высокочастотные параметры преобразования аудиоданных являются комплекснозначными.

Положение 18. Декодер согласно положению 17, при этом параметры преобразования дополнительно содержат действительные высокочастотные параметры преобразования аудиоданных.

Положение 19. Декодер согласно любому из положений 14-18, дополнительно содержащий фильтры для разделения базовых аудиосигналов на низкочастотные компоненты и высокочастотные компоненты.

Положение 20. Способ декодирования закодированного аудиосигнала, причем закодированный аудиосигнал включает в себя

первое представление, включающее в себя базовые аудиосигналы, предназначенные для воспроизведения

закодированного аудиосигнала в первом формате представления аудиоданных; и

параметры преобразования для преобразования базовых аудиосигналов в первом формате представления в выходные сигналы второго формата представления, при этом параметры преобразования содержат высокочастотные параметры преобразования аудиоданных и низкочастотные параметры преобразования аудиоданных, причем низкочастотные параметры преобразования включают в себя параметры матрицы свертки с несколькими отводами, при этом первый формат представления предназначен для воспроизведения с помощью динамиков, и второй формат представления предназначен для воспроизведения с помощью наушников, или наоборот,

при этом способ включает в себя этапы, на которых

выполняют свертку низкочастотных компонентов базовых аудиосигналов с помощью низкочастотных параметров преобразования, чтобы получить свернутые низкочастотные компоненты;

умножают высокочастотные компоненты базовых аудиосигналов на высокочастотные параметры преобразования, чтобы получить умноженные высокочастотные компоненты;

объединяют свернутые низкочастотные компоненты и умноженные высокочастотные компоненты, чтобы получить частотные компоненты выходного аудиосигнала для второго формата представления.

Положение 21. Способ согласно положению 20, в котором закодированный аудиосигнал содержит несколько временных сегментов, при этом способ дополнительно включает в себя этапы, на которых

выполняют интерполяцию параметров преобразования нескольких временных сегментов закодированного аудиосигнала, чтобы получить интерполированные параметры преобразования, включающие в себя интерполированные низкочастотные параметры преобразования аудиоданных; и

выполняют свертку нескольких временных сегментов низкочастотных компонентов базовых аудиосигналов с помощью интерполированных низкочастотных параметров преобразования аудиоданных, чтобы получить несколько временных сегментов свернутых низкочастотных компонентов.

Положение 22. Способ согласно положению 20, в котором параметры преобразования закодированного аудиосигнала являются переменными по времени, и упомянутая свертка низкочастотных компонентов базовых аудиосигналов включает в себя этапы, на которых

выполняют свертку низкочастотных компонентов базовых аудиосигналов с помощью низкочастотных параметров преобразования для нескольких временных сегментов, чтобы получить несколько множеств промежуточных свернутых низкочастотных компонентов; и

выполняют интерполяцию этих нескольких множеств промежуточных свернутых низкочастотных компонентов, чтобы получить свернутые низкочастотные компоненты.

Положение 23. Способ согласно положению 20 или 22, в котором упомянутая интерполяция использует метод наложения и суммирования нескольких множеств промежуточных свернутых низкочастотных компонентов.

Положение 24. Способ согласно любому из положений 20-23, дополнительно содержащий этап, на котором выполняют фильтрацию базовых аудиосигналов в упомянутые низкочастотные компоненты и упомянутые высокочастотные компоненты.

Положение 25. Машиночитаемый неизменяемый носитель, включающий в себя программные команды для работы компьютера в соответствии со способом по любому из положений 1-13 и 20-24.

## ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Компьютерно-реализуемый способ формирования потока данных для обеспечения представления аудиоканалов или аудиообъектов, содержащий этапы, на которых

получают базовые сигналы, каковые базовые сигналы представляют представление аудиоканалов или аудиообъектов;

определяют параметры преобразования, каковые параметры преобразования предназначены для преобразования базовых сигналов упомянутого представления в выходные сигналы; при этом параметры преобразования включают в себя по меньшей мере одни из высокочастотных параметров преобразования, заданных для полосы более высокой частоты, и низкочастотных параметров преобразования, заданных для полосы более низкой частоты, причем низкочастотные параметры преобразования включают в себя параметры матрицы свертки с несколькими отводами для свертки низкочастотных компонентов базовых сигналов с помощью низкочастотных параметров преобразования, чтобы получить свернутые низкочастотные компоненты, и высокочастотные параметры преобразования включают в себя параметры матрицы без состояний для умножения высокочастотных компонентов базовых сигналов на высокочас-

- тотные параметры преобразования, чтобы получить умноженные высокочастотные компоненты; и объединяют базовые сигналы и параметры преобразования, чтобы сформировать поток данных.
2. Способ по п.1, в котором параметры матрицы свертки с несколькими отводами указывают фильтр с конечной импульсной характеристикой (КИХ; FIR).
3. Способ по п.1, в котором базовые сигналы разделены на последовательность временных сегментов и по меньшей мере часть параметров преобразования обеспечивается для каждого временного сегмента.
4. Способ по п.1, в котором параметры матрицы свертки с несколькими отводами включают в себя по меньшей мере один коэффициент, который является комплекснозначным.
5. Способ по п.1, в котором упомянутое получение базовых сигналов содержит этап, на котором определяют базовые сигналы из упомянутых аудиоканалов или объектов с использованием первых параметров рендеризации.
6. Способ по п.5, содержащий этап, на котором определяют желаемые выходные сигналы из упомянутых аудиоканалов или объектов с использованием вторых параметров рендеризации.
7. Способ по п.6, в котором при упомянутом определении параметров преобразования параметры преобразования определяют посредством минимизации отклонения выходных сигналов от желаемых выходных сигналов.
8. Долговременный машиночитаемый носитель, на котором сохранены команды, которые при их исполнении устройством предписывают устройству выполнять операции, содержащие получение базовых сигналов, каковые базовые сигналы представляют представление аудиоканалов или аудиообъектов; определение параметров преобразования, каковые параметры преобразования предназначены для преобразования базовых сигналов упомянутого представления в выходные сигналы; при этом параметры преобразования включают в себя по меньшей мере одни из высокочастотных параметров преобразования, заданных для полосы более высокой частоты, и низкочастотных параметров преобразования, заданных для полосы более низкой частоты, причем низкочастотные параметры преобразования включают в себя параметры матрицы свертки с несколькими отводами для свертки низкочастотных компонентов базовых сигналов с помощью низкочастотных параметров преобразования, чтобы получить свернутые низкочастотные компоненты, и высокочастотные параметры преобразования включают в себя параметры матрицы без состояний для умножения высокочастотных компонентов базовых сигналов на высокочастотные параметры преобразования, чтобы получить умноженные высокочастотные компоненты; и объединение базовых сигналов и параметров преобразования, чтобы сформировать поток данных.
9. Долговременный машиночитаемый носитель по п.8, при этом параметры матрицы свертки с несколькими отводами указывают фильтр с конечной импульсной характеристикой (КИХ; FIR).
10. Долговременный машиночитаемый носитель по п.8, при этом базовые сигналы разделены на последовательность временных сегментов и по меньшей мере часть параметров преобразования обеспечивается для каждого временного сегмента.
11. Долговременный машиночитаемый носитель по п.8, при этом параметры матрицы свертки с несколькими отводами включают в себя по меньшей мере один коэффициент, который является комплекснозначным.
12. Долговременный машиночитаемый носитель по п.8, при этом упомянутое получение базовых сигналов содержит определение базовых сигналов из упомянутых аудиоканалов или объектов с использованием первых параметров рендеризации.
13. Долговременный машиночитаемый носитель по п.12, при этом операции содержат определение желаемых выходных сигналов из упомянутых аудиоканалов или объектов с использованием вторых параметров рендеризации.
14. Долговременный машиночитаемый носитель по п.13, при этом при упомянутом определении параметров преобразования параметры преобразования определяются посредством минимизации отклонения выходных сигналов от желаемых выходных сигналов.
15. Вычислительная система, выполненная с возможностью формирования потока данных для обеспечения представления аудиоканалов или аудиообъектов, содержащая процессор; и долговременный машиночитаемый носитель, на котором сохранены команды, которые при их исполнении процессором предписывают процессору выполнять операции, содержащие получение базовых сигналов, каковые базовые сигналы представляют представление аудиоканалов или аудиообъектов; определение параметров преобразования, каковые параметры преобразования предназначены для преобразования базовых сигналов упомянутого представления в выходные сигналы; при этом параметры преобразования включают в себя по меньшей мере одни из высокочастотных параметров преобразования, заданных для полосы более высокой частоты, и низкочастотных параметров преобразования, заданных для полосы более низкой частоты, причем низкочастотные параметры преобразования включают в себя параметры матрицы свертки с несколькими отводами для свертки низкочастотных компонентов

базовых сигналов с помощью низкочастотных параметров преобразования, чтобы получить свернутые низкочастотные компоненты, и высокочастотные параметры преобразования включают в себя параметры матрицы без состояний для умножения высокочастотных компонентов базовых сигналов на высокочастотные параметры преобразования, чтобы получить умноженные высокочастотные компоненты; и

объединение базовых сигналов и параметров преобразования, чтобы сформировать поток данных.

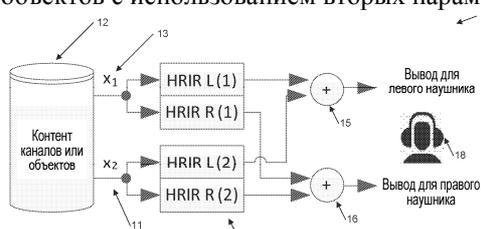
16. Система по п.15, при этом параметры матрицы свертки с несколькими отводами указывают фильтр с конечной импульсной характеристикой (КИХ; FIR).

17. Система по п.15, при этом базовые сигналы разделены на последовательность временных сегментов и по меньшей мере часть параметров преобразования обеспечивается для каждого временного сегмента.

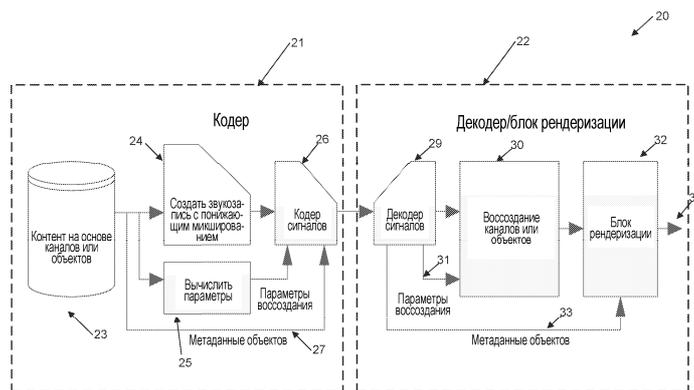
18. Система по п.15, при этом параметры матрицы свертки с несколькими отводами включают в себя по меньшей мере один коэффициент, который является комплекснозначным.

19. Система по п.15, в которой упомянутое получение базовых сигналов содержит определение базовых сигналов из упомянутых аудиоканалов или объектов с использованием первых параметров рендеризации.

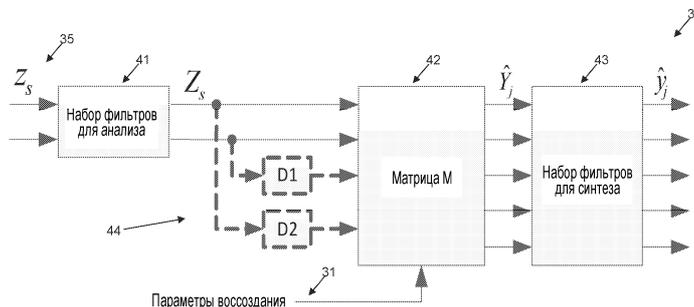
20. Система по п.19, в которой операции содержат определение желаемых выходных сигналов из упомянутых аудиоканалов или объектов с использованием вторых параметров рендеризации.



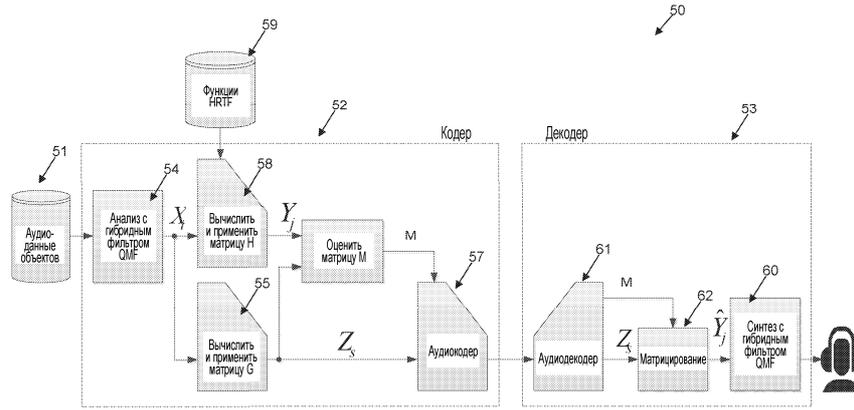
Фиг. 1



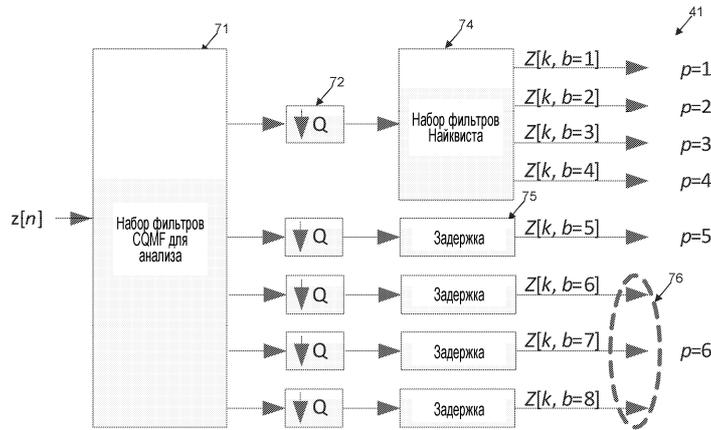
Фиг. 2



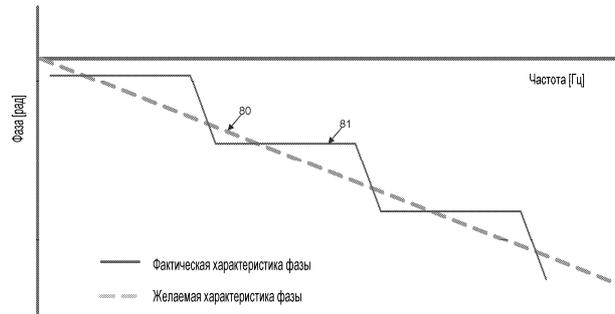
Фиг. 3



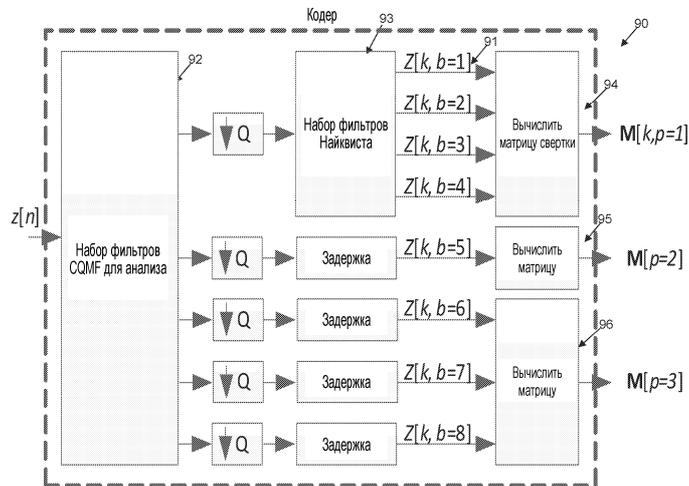
Фиг. 4



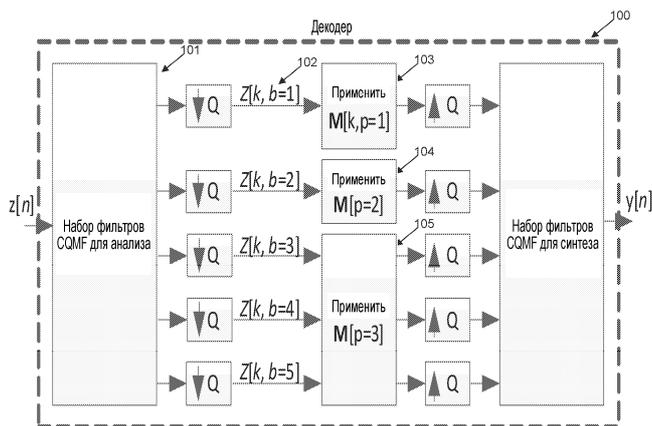
Фиг. 5



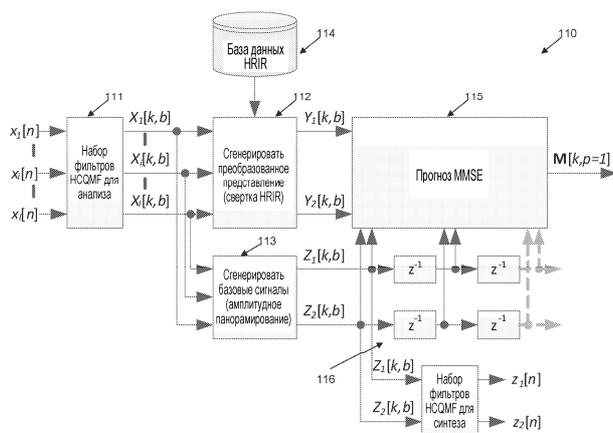
Фиг. 6



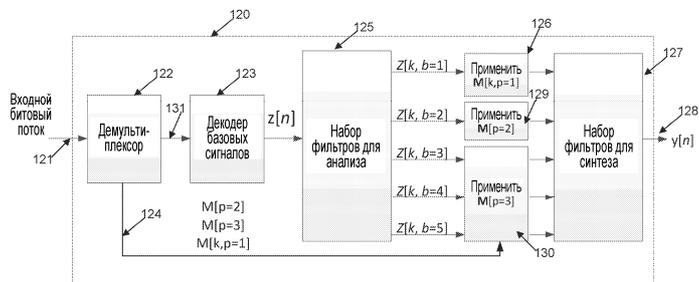
Фиг. 7



Фиг. 8



Фиг. 9



Фиг. 10