

(19)



**Евразийское
патентное
ведомство**

(21) **202291846** (13) **A1**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2022.10.28

(22) Дата подачи заявки
2018.03.19

(51) Int. Cl. **G06F 17/10** (2006.01)
G10L 19/02 (2013.01)
G10L 19/22 (2013.01)
G10L 19/24 (2013.01)
G10L 19/26 (2013.01)

**(54) ОБРАТНО СОВМЕСТИМАЯ КОМПОНОВКА ГАРМОНИЧЕСКОГО ТРАНСПОЗЕРА
ДЛЯ РЕКОНСТРУКЦИИ ВЫСОКИХ ЧАСТОТ ЗВУКОВЫХ СИГНАЛОВ**

(31) 62/475,619

(32) 2017.03.23

(33) US

(62) 202191156; 2018.03.19

(71) Заявитель:
ДОЛБИ ИНТЕРНЭШНЛ АБ (NL)

(72) Изобретатель:
**Виллемоес Ларс, Пурнхаген Хейко,
Экстранд Пер (US)**

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

(57) Раскрыт способ для декодирования кодированного звукового битового потока. Способ включает в себя прием кодированного звукового битового потока и декодирование звуковых данных для формирования декодированного звукового сигнала в полосе низких частот. Способ дополнительно включает в себя извлечение метаданных реконструкции высоких частот и фильтрацию декодированного звукового сигнала в полосе низких частот гребенкой фильтров разложения сигнала для формирования фильтрованного звукового сигнала в полосе низких частот. Способ также включает в себя извлечение флажкового признака, указывающего, должны ли выполняться над звуковыми данными спектральное преобразование или гармоническая транспозиция, и восстановление части высоких частот звукового сигнала с использованием фильтрованного звукового сигнала в полосе низких частот и метаданных реконструкции высоких частот в соответствии с флажковым признаком.

A1

202291846

202291846

A1

**ОБРАТНО СОВМЕСТИМАЯ КОМПОНОВКА ГАРМОНИЧЕСКОГО ТРАНСПОЗЕРА ДЛЯ
РЕКОНСТРУКЦИИ ВЫСОКИХ ЧАСТОТ ЗВУКОВЫХ СИГНАЛОВ**

ОБЛАСТЬ ТЕХНИКИ, К КОТОРОЙ ОТНОСИТСЯ ИЗОБРЕТЕНИЕ

Варианты осуществления имеют отношение к обработке звуковых сигналов, а точнее, к кодированию, декодированию или перекодированию звуковых битовых потоков с управляющими данными, задающими то, должна ли выполняться над звуковыми данными базовая форма реконструкции высоких частот ("HFR") или усовершенствованная форма HFR.

УРОВЕНЬ ТЕХНИКИ ИЗОБРЕТЕНИЯ

Типичный звуковой битовый поток включает в себя как звуковые данные (например, кодированные звуковые данные), указывающие один или более каналов звукового контента, так и метаданные, указывающие по меньшей мере одну характеристику звуковых данных или звукового контента. Одним из широко известных форматов для формирования кодированного звукового битового потока является формат Перспективного звукового кодирования (AAC) MPEG-4, описанный в стандарте ISO/IEC 14496-3:2009 MPEG. В стандарте MPEG-4, AAC обозначает «перспективное звуковое кодирование», а HE-AAC обозначает «высокоэффективное перспективное звуковое кодирование».

Стандарт AAC MPEG-4 определяет несколько звуковых профилей, которые определяют, какие объекты и средства кодирования присутствуют в совместимом кодировщике или декодере. Три из этих звуковых профилей являются (1) профиль AAC, (2) профиль HE-AAC и (3) профиль HE-AAC v2. Профиль AAC включает в себя тип объекта малой сложности AAC (или «AAC-LC»). Объект AAC-LC является аналогом для профиля малой сложности AAC MPEG-2 с некоторыми корректировками и не включает в себя ни тип объекта репликации полос спектра («SBR»), ни тип объекта параметрической стереофонии («PS»). Профиль HE-AAC является надмножеством профиля AAC и дополнительно включает в себя тип объекта SBR. Профиль HE-AAC v2 является надмножеством профиля HE-AAC и дополнительно включает в себя тип объекта PS.

Тип объекта SBR содержит в себе средство репликации полос спектра, которое является важным средством кодирования реконструкции высоких частот («HFR»), которое значительно улучшает эффективность сжатия перцепционных звуковых кодеков. SBR реконструирует высокочастотные составляющие звукового сигнала на стороне приемника (например, в декодере). Таким образом, кодировщику нужно кодировать и передавать только низкочастотные составляющие, предоставляя возможность для гораздо более высокого качества звука на низких скоростях передачи данных. SBR основана на репликации последовательностей гармоник, ранее усеченных, для того чтобы уменьшить скорость передачи данных, из имеющегося в распоряжении сигнала с ограниченной шириной полосы частот и управляющих данных, полученных из кодировщика. Соотношение между тональными и шумоподобными составляющими поддерживается посредством адаптивной обратной фильтрации, а также необязательным добавлением шума и синусоидальных сигналов. В стандарте AAC MPEG-4, средство SBR выполняет спектральную вклейку (также называемое линейным преобразованием или спектральным преобразованием), в котором некоторое количество следующих друг за другом поддиапазонов квадратурных зеркальных фильтров (КЗФ, QMF) копируются (или «вклеиваются») из переданной части полосы низких частот звукового сигнала в часть полосы высоких частот звукового сигнала, которая формируется в декодере.

Спектральная вклейка или линейное преобразование могут не быть идеальными для некоторых типов звукового сигнала, таких как музыкальный контент с относительно низкими частотами разделения. Поэтому, необходимы технологии для улучшения репликации полос спектра.

Краткое описание вариантов осуществления изобретения

Раскрыта первая группа вариантов осуществления, которая относится к способу для декодирования кодированного звукового битового потока. Способ включает в себя прием кодированного звукового битового потока и декодирование звуковых данных для формирования декодированного звукового сигнала в полосе низких частот. Способ дополнительно включает в себя извлечение

метаданных реконструкции высоких частот и фильтрацию декодированного звукового сигнала в полосе низких частот гребенкой фильтров разложения сигнала для формирования фильтрованного звукового сигнала в полосе низких частот. Способ дополнительно включает в себя извлечение флажкового признака, указывающего, должны ли выполняться над звуковыми данными спектральное преобразование или гармоническая транспозиция, и восстановление части высоких частот звукового сигнала с использованием фильтрованного звукового сигнала в полосе низких частот и метаданных реконструкции высоких частот в соответствии с флажковым признаком. В заключение, способ включает в себя комбинирование фильтрованного звукового сигнала в полосе низких частот и восстановленной части полосы высоких частот для формирования широкополосного звукового сигнала.

Вторая группа вариантов осуществления относится к звуковому декодеру для декодирования кодированного звукового битового потока. Декодер включает в себя входной интерфейс для приема кодированного звукового битового потока, где кодированный звуковой битовый поток включает в себя звуковые данные, представляющие собой часть низких частот звукового сигнала, и основной декодер для декодирования звуковых данных, чтобы формировать декодированный звуковой сигнал в полосе низких частот. Декодер также включает в себя демультиплексор для извлечения из кодированного звукового битового потока метаданных реконструкции высоких частот, где метаданные реконструкции высоких частот включают в себя рабочие параметры для процесса реконструкции высоких частот, который линейно преобразует некоторое количество следующих друг за другом поддиапазонов из части полосы низких частот звукового сигнала в часть полосы высоких частот звукового сигнала, и гребенку фильтров разложения сигнала для фильтрации декодированного звукового сигнала в полосе низких частот, чтобы формировать фильтрованный звуковой сигнал в полосе низких частот. Декодер дополнительно включает в себя демультиплексор для извлечения из кодированного звукового битового потока флажкового признака, указывающего, должны ли выполняться над звуковыми данными линейное преобразование или

гармоническая транспозиция, и восстановитель высоких частот для восстановления части высоких частот звукового сигнала с использованием фильтрованного звукового сигнала в полосе низких частот и метаданных реконструкции высоких частот в соответствии с флажковым признаком. В заключение, декодер включает в себя гребенку фильтров синтеза сигнала для комбинирования фильтрованного звукового сигнала в полосе низких частот и восстановленной части полосы высоких частот, чтобы формировать широкополосный звуковой сигнал.

Другие группы вариантов осуществления относятся к кодированию и перекодированию звуковых битовых потоков, содержащих в себе метаданные, идентифицирующие, должна ли выполняться обработка усовершенствованной репликации полос спектра (eSBR).

Краткое описание чертежей

Фиг. 1 - структурная схема варианта осуществления системы, которая может быть выполнена с возможностью выполнять вариант осуществления обладающего признаками изобретения способа.

Фиг. 2 - структурная схема кодировщика, который является вариантом осуществления обладающего признаками изобретения блока обработки звукового сигнала.

Фиг. 3 - структурная схема системы, включающей в себя декодер, который является вариантом осуществления обладающего признаками изобретения блока обработки звукового сигнала и, по выбору, также постпроцессор, присоединенный к нему.

Фиг. 4 - структурная схема декодера, который является вариантом осуществления обладающего признаками изобретения блока обработки звукового сигнала.

Фиг. 5 - структурная схема декодера, который является еще одним вариантом осуществления обладающего признаками изобретения блока обработки звукового сигнала.

Фиг. 6 - структурная схема еще одного варианта осуществления обладающего признаками изобретения блока обработки звукового сигнала.

Фиг. 7 - схема блока битового потока AAC MPEG-4, включающего в себя сегменты, на которые он разделен.

Система обозначений и терминология

На всем протяжении данного описания изобретения, в том числе в формуле изобретения, выражение выполнение операции «над» сигналом или данными (например, фильтрацией, масштабированием, преобразованием или применением коэффициента передачи к сигналу или данным) используется в широком смысле для обозначения выполнения операции непосредственно над сигналом или данными, или над обработанным вариантом сигнала или данных (например, над вариантом сигнала, который подвергся предварительной фильтрации или предварительной обработке перед выполнением операции над ним) .

На всем протяжении этого описания изобретения, в том числе в формуле изобретения, выражение «блок обработки звукового сигнала» или «звуковой процессор» используется в широком смысле для обозначения системы, устройства или установки, выполненных с возможностью обрабатывать звуковых данных. Примеры блоков обработки звукового сигнала включают в себя, но не в качестве ограничения, кодировщики, перекодировщики, декодеры, кодеки, системы предварительной обработки, системы постобработки и системы обработки битового потока (иногда упоминаемые как средства обработки битового потока). Практически вся бытовая электроника, такая как мобильные телефоны, телевизоры, дорожные компьютеры, планшетные компьютеры, содержит в себе блок обработки звукового сигнала или звуковой процессор.

На всем протяжении этого описания изобретения, в том числе, в формуле изобретения, термин «присоединяется» или «присоединен» используется в широком смысле, чтобы означать прямое или опосредованное соединение. Таким образом, если первое устройство присоединяется к второму устройству, соединение может происходить через прямое соединение или через опосредованное соединение с помощью других устройств и соединений. Более того, компоненты, которые объединены в или с другими компонентами, также соединены друг с другом.

Подробное описание вариантов осуществления изобретения

Стандарт AAC MPEG-4 предполагает, что кодированный битовый поток AAC MPEG-4 включает в себя метаданные, указывающие каждый

тип обработки реконструкции высоких частот («HFR»), которая должна выполняться (если какая-нибудь должна выполняться) декодером для декодирования звукового контента битового потока, и/или которые управляют такой обработкой HFR и/или указывают по меньшей мере одну характеристику или параметр по меньшей мере одного средства HFR, подлежащего применению для декодирования звукового контента битового потока. В материалах настоящей заявки, мы используем выражение «метаданные SBR» для обозначения метаданных этого типа, который описан или упоминается в стандарте AAC MPEG-4 для использования с репликацией полос спектра («SBR»). Как принимается во внимание специалистом в данной области техники, SBR является разновидностью HFR.

SBR предпочтительно используется в качестве системы с двойной частотой дискретизации, причем, лежащий в основе кодек работает на половине исходной частоты дискретизации, в то время как SBR функционирует на исходной частоте дискретизации. Кодировщик SBR работает параллельно с лежащим в основе основным кодеком, хотя и на более высокой частоте дискретизации. Хотя SBR главным образом является постобработкой в декодере, важные параметры извлекаются в кодировщике, для того чтобы обеспечивать наиболее точную реконструкцию высоких частот в декодере. Кодировщик оценивает огибающую спектра диапазона SBR, что касается частотно-временного диапазона/разрешения по времени и частоте, подходящих для характеристик текущих сегментов входного сигнала. Огибающая спектра оценивается посредством разложения сигнала с помощью комплексных КЗФ и последующего расчета энергии. Разрешения по времени и частоте огибающих спектра могут выбираться с высокой степенью свободы, для того чтобы обеспечивать наилучшее разрешение по времени и частоте для данного сегмента входного сигнала. Оценка огибающей должна учитывать, что переходный процесс в оригинале, главным образом расположенный в области высоких частот (например, высокий выброс), будет присутствовать в незначительной степени в сформированной SBR полосе высоких частот перед коррекцией огибающей, поскольку полоса высоких частот в декодере основана на полосе низких частот, где переходный процесс выражен в

гораздо меньшей степени по сравнению с полосой высоких частот. Этот аспект накладывает иные требования на разрешение по времени и частоте данных огибающей спектра по сравнению с традиционной оценкой огибающей спектра, как используемая в других алгоритмах звукового кодирования.

Кроме огибающей спектра, извлекаются несколько дополнительных параметров, представляющих собой спектральные характеристики входного сигнала для разных частотно-временных областей. Поскольку кодировщик естественно имеет доступ к исходному сигналу, а также к информации о том, каким образом блок SBR в декодере будет создавать полосу высоких частот, при условии определенного набора параметров управления, можно, чтобы система справлялась с ситуациями, где полоса низких частот составляет мощный гармонический ряд, а полоса высоких частот, которая должна быть воссоздана, главным образом составляет случайные составляющие сигнала, а также ситуации, где мощные тональные составляющие присутствуют в исходной полосе высоких частот без аналогов в полосе низких частот, на которой основана область полосы высоких частот. Более того, кодировщик SBR работает в тесной зависимости от лежащего в основе основного кодека, чтобы оценивать, какой частотный диапазон должен быть охвачен SBR в данный момент времени. Данные SBR рационально кодируются перед передачей посредством использования энтропийного кодирования, а также зависимостей каналов управляющих данных, в случае стереофонических сигналов.

Алгоритмы извлечения параметров управления типично необходимо тщательно настраивать под лежащий в основе кодек на заданной скорости передачи битов и заданной частоте дискретизации. Это происходит вследствие того обстоятельства, что более низкая скорость передачи битов обычно полагается на больший диапазон SBR по сравнению с высокой скоростью передачи битов, и разные частоты дискретизации соответствуют разным разрешениям по времени кадров SBR.

Декодер SBR типично включает в себя несколько разных частей. Он содержит модуль декодирования битового потока, модуль восстановления высоких частот (HFR), модуль дополнительных

высокочастотных составляющих и модуль корректора огибающей. Система основана на гребенке фильтров из комплекснозначных КЗФ. В модуле извлечения битового потока, управляющие данные считываются из битового потока и декодируются. Время/частотная сетка получается для текущего кадра перед считыванием данных огибающей из битового потока. Лежащий в основе основной декодер декодирует звуковой сигнал текущего кадра (хотя и с более низкой частотой дискретизации), для создания отсчетов звукового сигнала во временной области. Результирующий кадр звуковых данных используется для реконструкции высоких частот модулем HFR. Декодированный сигнал в полосе низких частот затем подвергается разложению сигнала с использованием гребенки фильтров из КЗФ. Реконструкция высоких частот и коррекция огибающей впоследствии выполняются над отсчетами поддиапазонов гребенки фильтров из КЗФ. Высокие частоты реконструируются из полосы низких частот гибким образом на основании заданных параметров управления. Более того, реконструированная полоса высоких частот адаптивно фильтруется на основе каналов поддиапазона согласно управляющим данным, чтобы обеспечить надлежащие спектральные характеристики данной частотно/временной области.

Верхним уровнем битового потока AAC MPEG-4 является последовательность блоков данных (элементов «raw_data_block»), каждый из которых является сегментом данных (в материалах настоящей заявки упоминаемым как «block»), который содержит в себе звуковые данные (типично для периода времени в 1024 или 960 отсчетов) и связанную информацию и/или другие данные. В материалах настоящей заявки, мы используем термин «блок» для обозначения сегмента битового потока AAC MPEG-4, содержащего звуковые данные (и соответствующих метаданных и, по выбору, также других связанных данных), которые определяют или являются указывающими один (но не более чем один) элемент «raw_data_block».

Каждый блок битового потока AAC MPEG-4 может включать в себя некоторое количество синтаксических элементов (каждый из которых также материализован в битовом потоке в виде сегмента данных). Семь типов таких синтаксических элементов определено в

стандарте AAC MPEG-4. Каждый синтаксический элемент идентифицируется разным значением элемента «id_syn_ele» данных. Примеры синтаксических элементов включают в себя «single_channel_element()», «channel_pair_element()» и «fill_element()». Элемент одиночного канала является контейнером, включающим в себя звуковые данные одиночного звукового канала (монофонического звукового сигнала). Элемент пары каналов включает в себя звуковые данные двух звуковых каналов (то есть, стереофонический звуковой сигнал).

Элемент заполнения является контейнером информации, включающей в себя идентификатор (например, значение отмеченного выше элемента «id_syn_ele»), сопровождаемый данными, которые упоминаются как «данные заполнения». Элементы заполнения исторически использовались для коррекции мгновенной скорости передачи битов битовых потоков, которые должны передаваться по каналу с постоянной скоростью передачи. Добавлением надлежащего количества данных заполнения в каждый блок, может достигаться постоянная скорость передачи данных.

В соответствии с вариантами осуществления у изобретения, данные заполнения могут включать в себя одну или более полезных нагрузок расширения, которые расширяют тип данных (например, метаданные), допускающих передачу в битовом потоке. Декодер, который принимает битовые потоки будут с данными заполнения, содержащими в себе новый тип данных, по выбору может использоваться устройством, принимающим битовый поток (например, декодером) для расширения функциональных возможностей устройства. Таким образом, как может быть принято во внимание специалистом в данной области техники, элементы заполнения являются специальным типом структуры данных и отличаются от структур данных, типично используемых для передачи звуковых данных (например, полезных нагрузок звукового сигнала, содержащих в себя каналные данные).

В некоторых вариантах осуществления изобретения, идентификатор, используемый для идентификации элемента заполнения, может состоять из трехбитного беззнакового целого числа, передаваемого старшим битом вперед («uimsbf»), имеющего

значение 0×6 . В одном блоке могут появляться несколько экземпляров одного и того же типа синтаксического элемента (например, несколько элементов заполнения).

Еще одним стандартом для кодирования звуковых битовых потоков является стандарт унифицированного кодирования речи и звуковых сигналов (USAC) MPEG (ISO/IEC 23003-3:2012). Стандарт USAC MPEG описывает кодирование и декодирование звукового контента с использованием обработки репликации полос спектра (в том числе, обработки SBR, как описанная в стандарте AAC MPEG-4, и, к тому же, включающую в себя другие усовершенствованные формы обработки репликации полос спектра). Эта обработка применяет средства репликации полос спектра (иногда в материалах настоящей заявки упоминаемые как «средства усовершенствованной SBR» или «средства eSBR») расширенного и усовершенствованного варианта набора средств SBR, описанных в стандарте AAC MPEG-4. Таким образом, eSBR (как определенная в стандарте USAC) является усовершенствованием в отношении SBR (как определенная в стандарте AAC MPEG-4).

В материалах настоящей заявки, мы используем выражение «обработка усовершенствованной SBR» (или «обработка eSBR») для обозначения обработки репликации полос спектра с использованием по меньшей мере одного средства eSBR (например, по меньшей мере одного средства eSBR, которое описано или упомянуто в стандарте USAC MPEG), которое не описано и не упомянуто в стандарте AAC MPEG-4. Примерами таких средств eSBR являются гармоническая транспозиция и дополнительная предварительная обработка вклейки с помощью КЗФ или «предварительное выравнивание».

Гармонический транспозер целочисленного порядка T отображает синусоиду с частотой ω в синусоиду с частотой $T\omega$, тем временем сохраняя длительность сигнала. Три порядка, $T=2, 3, 4$, типично используются последовательно для создания каждой части желательного выходного частотного диапазона с использованием наименьшего возможного порядка транспозиции. Если требуется выходной сигнал выше диапазона транспозиции четвертого порядка, он может формироваться посредством сдвигов по частоте. Когда

возможно, почти критически дискретизированные временные области основной полосы создаются для обработки, чтобы минимизировать вычислительную сложность.

Битовый поток, сформированный в соответствии со стандартом USAC MPEG (в материалах настоящей заявки иногда упоминаемый как «битовый поток USAC»), включает в себя кодированный звуковой контент и типично включает в себя метаданные, указывающие каждый тип обработки репликации полос спектра, которая должна применяться декодером для декодирования звукового контента битового потока USAC, и/или метаданные, которые управляют такой обработкой репликации полос спектра или указывают по меньшей мере одну характеристику или параметр по меньшей мере одного средства SBR и/или средства eSBR, которое должно применяться для декодирования звукового контента битового потока USAC.

В материалах настоящей заявки используем выражение «метаданные усовершенствованной SBR» (или «метаданные eSBR») для обозначения метаданных, указывающих каждый тип обработки репликации полос спектра, которая должна применяться декодером для декодирования звукового контента кодированного звукового битового потока (например, битового потока USAC), и/или которые управляют такой обработкой репликации полос спектра и/или указывает по меньшей мере одну характеристику или параметр по меньшей мере одного средства SBR и/или средства eSBR, подлежащего применению для декодирования такого звукового контента, но которые не описаны и не упомянуты в стандарте AAC MPEG-4. Примером метаданных eSBR являются метаданные (указывающие или предназначенные для управления обработкой репликации полос спектра), которые описаны или упомянуты в стандарте USAC MPEG, но не в стандарте AAC MPEG-4. Таким образом, метаданные eSBR в материалах настоящей заявки обозначают метаданные, которые не являются метаданными SBR, а метаданные SBR в материалах настоящей заявки обозначают метаданные, которые не являются метаданными eSBR.

Битовый поток USAC может включать в себя как метаданные SBR, так и метаданные eSBR. Точнее, битовый поток USAC может включать в себя метаданные eSBR, которые управляют выполнением

обработки eSBR декодером, и метаданные SBR, которые управляют выполнением обработки SBR декодером. В соответствии с типичными вариантами осуществления настоящего изобретения, метаданные eSBR (например, специфичные eSBR конфигурационные данные) (в соответствии с настоящим изобретением) включенные в битовый поток AAC MPEG-4 (например, в контейнере `sbr_extension()` в конце полезной нагрузки SBR).

Выполнение обработки eSBR во время декодирования кодированного битового потока с использованием средства eSBR (содержащего по меньшей мере одно средство eSBR), установленного декодером, восстанавливает полосу высоких частот звукового сигнала на основании репликации последовательностей гармоник, которые были усечены во время кодирования. Такая обработка eSBR типично корректирует огибающую спектра сформированной полосы высоких частот и применяет обратную фильтрацию, и добавляет шумовые и синусоидальные составляющие, для того чтобы воссоздать спектральные характеристики исходного звукового сигнала.

В соответствии с типичными вариантами осуществления изобретения, метаданные eSBR включены (например, небольшое количество управляющих битов, которые являются метаданными eSBR, включено) в один или более сегментов метаданных кодированного звукового битового потока (например, битового потока AAC MPEG-4), который также включает в себя кодированные звуковые данные в других сегментах (сегментах звуковых данных). Типично, по меньшей мере один такой сегмент метаданных каждого блока битового потока представляет собой (или включает в себя) элемент заполнения (включающий в себя идентификатор, указывающий начало элемента заполнения), и метаданные eSBR включены в элемент заполнения после идентификатора. Фиг. 1 – структурная схема примерной цепи обработки звукового сигнала (системы обработки звуковых данных), в которой один или более из элементов системы может быть выполнен в соответствии с вариантом осуществления настоящего изобретения. Система включает в себя следующие элементы, соединенные друг с другом, как показано: кодировщик 1, подсистему 2 доставки, декодер 3 и блок 4 постобработки. В разновидностях показанной системы, не включены в состав один или

более элементов, или включены в состав дополнительные блоки обработки звуковых данных.

В некоторых реализациях, кодировщик 1 (который по выбору включает в себя блок предварительной обработки) выполнен с возможностью принимать отсчеты РСМ (временной области), содержащие звуковой контент, в качестве входного сигнала и выдавать кодированный звуковой битовый поток (имеющий формат, который совместим со стандартом AAC MPEG-4), который является указывающим звуковой контент. Данные битового потока, которые указывают звуковой контент, в материалах настоящей заявки иногда упоминаются как «звуковые данные» или «кодированные звуковые данные». Если кодировщик выполнен в соответствии с типичным вариантом осуществления настоящего изобретения, звуковой битовый поток, выдаваемый из кодировщика, включает в себя метаданные eSBR (и, к тому же, типично другие метаданные), а также звуковые данные.

Один или более кодированных звуковых потоков, выдаваемых из кодировщика 1, могут предъявляться в подсистему 2 доставки кодированного звукового сигнала. Подсистема 2 выполнена с возможностью сохранять и/или доставлять каждый кодированный битовый поток, выдаваемый из кодировщика 1. Кодированный звуковой битовый поток, выведенный из кодировщика 1, может сохраняться подсистемой 2 (например, в виде DVD или диска Bluray) или передаваться подсистемой 2 (которая может реализовывать канал или сеть передачи), или может как сохраняться, так и передаваться подсистемой 2.

Декодер 3 выполнен с возможностью декодировать кодированный звуковой битовый поток AAC MPEG-4 (сформированный кодировщиком 1), который он принимает через подсистему 2. В некоторых вариантах осуществления, декодер 3 выполнен с возможностью извлекать метаданные eSBR из каждого блока битового потока и декодировать битовый поток (в том числе, посредством выполнения обработки eSBR с использованием извлеченных метаданных eSBR) для формирования декодированных звуковых данных (например, потоков декодированных звуковых отсчетов РСМ). В некоторых вариантах осуществления, декодер 3 выполнен с возможностью извлекать

метаданные SBR из битового потока (но игнорировать метаданные eSBR, включенные в битовый поток) и декодировать битовый поток (в том числе, посредством выполнения обработки SBR с использованием извлеченных метаданных SBR) для формирования декодированных звуковых данных (например, потоков декодированных отсчетов звукового сигнала PCM. Типично, декодер 3 включает в себя буфер, который хранит (например, незфемерным образом) сегменты кодированного звукового битового потока, принятого из подсистемы 2.

Блок 4 постобработки по фиг. 1 выполнен с возможностью принимать поток декодированных звуковых данных из декодера 3 (например, отсчеты звукового сигнала PCM) и выполнять постобработку над ним. Блок постобработки также может быть выполнен с возможностью воспроизводить подвергнутый постобработке звуковой контент (или декодированный звуковой сигнал, принятый из декодера 3) для проигрывания одним или более громкоговорителей.

Фиг. 2 – структурная схема кодировщика (100), который является вариантом осуществления обладающего признаками изобретения блока обработки звукового сигнала. Любые из компонентов или элементов кодировщика 100 могут быть реализованы в виде одного или более процессов и/или одной или более схем (например, ASIC (специализированных интегральных схем), FPGA (программируемых пользователем вентильных матриц) или других интегральных схем) в аппаратных средствах, программном обеспечении или комбинации аппаратных средств и программного обеспечения. Кодировщик 100 включает в себя кодировщик 105, каскад 107 заполнителя/формatera, каскад 106 формирования метаданных и буферную память 109, соединенные как показано. К тому же, типично, кодировщик 100 включает в себя другие элементы обработки (не показаны). Кодировщик 100 выполнен с возможностью преобразовывать входной звуковой битовый поток в кодированный выходной битовый поток AAC MPEG-4.

Генератор 106 метаданных присоединен и выполнен с возможностью формировать (и/или пропускать насквозь в каскад 107) метаданные (в том числе, метаданные eSBR и метаданные SBR),

которые должны быть включены каскадом 107 в кодированный битовый поток, подлежащий выводу из кодировщика 100.

Кодировщик 105 присоединен и выполнен с возможностью кодировать (например, выполняя сжатие над ними) входные звуковые данные и предъявлять результирующий кодированный звуковой сигнал в каскад 107 для включения в кодированный битовый поток, который должен выдаваться из каскада 107.

Каскад 107 выполнен с возможностью мультиплексировать кодированный звуковой сигнал из кодировщика 105 и метаданные (в том числе, метаданные eSBR и метаданные SBR) из генератора 106 для формирования кодированного битового потока, который должен выдаваться из каскада 107, предпочтительно так, чтобы кодированный битовый поток имел формат, как заданный одним из вариантов осуществления настоящего изобретения.

Буферная память 109 выполнена с возможностью хранить (например, неэфемерным образом) по меньшей мере один блок кодированного звукового битового потока, выведенного из каскада 107, и последовательность блоков кодированного звукового битового потока затем предъявляется из буферной памяти 109 в качестве выходного сигнала из кодировщика 100 в систему доставки.

Фиг. 3 – структурная схема системы, включающей в себя декодер (200), который является вариантом осуществления обладающего признаками изобретения блока обработки звукового сигнала и, по выбору, также постпроцессор (300), присоединенный к нему. Любые из компонентов или элементов декодера 200 и постпроцессора 300 могут быть реализованы в виде одного или более процессов и/или одной или более схем (например, ASIC (специализированных интегральных схем), FPGA (программируемых пользователем вентильных матриц) или других интегральных схем) в аппаратных средствах, программном обеспечении или комбинации аппаратных средств и программного обеспечения. Декодер 200 содержит буферную память 201, деформатер 205 (синтаксический анализатор) полезной нагрузки битового потока, подсистему 202 декодирования звукового сигнала (иногда упоминаемую как каскад «основного» декодирования или подсистема «основного»

декодирования), каскад 203 обработки eSBR и каскад 204 формирования управляющих битов, соединенные как показано. К тому же, типично, декодер 200 включает в себя другие элементы обработки (не показаны).

Буферная память 201 (буфер) хранит (например, неэфемерным образом) по меньшей мере один блок кодированного звукового битового потока AAC MPEG-4, принятого декодером 200. При эксплуатации декодера 200, последовательность блоков битового потока предъявляется из буфера 201 в деформатер 205.

В разновидности варианта осуществления на фиг. 3, APU, который не является декодером (например, APU 500 по фиг. 6) включает в себя буферную память (например, буферную память, идентичную буферу 201), которая хранит (например, неэфемерным образом) по меньшей мере один блок кодированного звукового битового потока (например, звукового битового потока AAC MPEG-4) того же самого типа, принятого буфером 201 по фиг. 3 или фиг. 4 (то есть, кодированного звукового битового потока, который включает в себя метаданные eSBR).

Вновь со ссылкой на фиг. 3, деформатер 205 присоединен и выполнен с возможностью демультиплексировать каждый блок битового потока для извлечения метаданных SBR (в том числе, данные квантованной огибающей) и метаданных eSBR (и типично также других метаданных) из него, предъявлять по меньшей мере метаданные eSBR и метаданные SBR в каскад 203 обработки eSBR и типично также предъявлять другие извлеченные метаданные в подсистему 202 декодирования (и, по выбору, также в генератор 204 управляющих битов). Деформатер 205 также присоединен и выполнен с возможностью извлекать звуковые данные из каждого блока битового потока, и предъявлять извлеченные звуковые данные в подсистему 202 декодирования (каскад декодирования).

Система по фиг. 3 по выбору также включает в себя постпроцессор 300. Постпроцессор 300 включает в себя буферную память 301 (буфер) и другие элементы обработки (не показаны), в том числе, по меньшей мере один элемент обработки, присоединенный к буферу 301. Буфер 301 хранит (например, неэфемерным образом) по меньшей мере один блок (или кадр)

декодированных звуковых данных, принятых постпроцессором 300 из декодера 200. Элементы обработки постпроцессора 300 присоединены и выполнены с возможностью принимать и адаптивно обрабатывать последовательность блоков (или кадров) декодированного звукового сигнала, выведенного из буфера 301, с использованием метаданных, выведенных из подсистемы 202 декодирования (и/или деформатора 205), и/или управляющих битов, выведенных из каскада 204 декодера 200.

Подсистема 202 декодирования звукового сигнала декодера 200 выполнена с возможностью декодировать звуковые данные, извлеченные синтаксическим анализатором 205 (такое декодирование может упоминаться как операция «основного» декодирования), для формирования декодированных звуковых данных и предъявлять декодированные звуковые данные в каскад 203 обработки eSBR. Декодирование выполняется в частотной области и типично включает в себя обращенное квантование, сопровождаемое спектральной обработкой. Типично, завершающий каскад обработки в подсистеме 202 применяет преобразование из частотной области во временную область к декодированным звуковым данным частотной области, так чтобы выходными данными подсистемы были декодированные звуковые данные временной области. Каскад 203 выполнен с возможностью применять средства SBR и средства eSBR, указанные метаданными SBR и метаданными eSBR (извлеченными синтаксическим анализатором 205), к декодированным звуковым данным (то есть, выполнять обработку SBR и eSBR над выходными данными подсистемы 202 декодирования с использованием метаданных SBR и eSBR) для формирования полностью декодированных звуковых данных, которые выводятся (например, в постпроцессор 300) из декодера 200. Типично, декодер 200 включает в себя память (доступную из подсистемы 202 и каскада 203), которая хранит деформированные звуковые данные и метаданные, выданные из деформатора 205, а каскад 203 выполнен с возможностью осуществлять доступ к звуковым данным и метаданным (в том числе, метаданным SBR и метаданным eSBR) по мере надобности во время обработки SBR и eSBR. Обработка SBR и обработка eSBR в каскаде 203 может считаться постобработкой на выходе из подсистемы 202 основного

декодирования. По выбору, декодер 200 также включает в себя подсистему завершающего повышающего микширования (которая может применять средства параметрической стереофонии («PS»), определенные в стандарте AAC MPEG-4, с использованием метаданных PS, извлеченных деформатером 205, и/или управляющих битов, сформированных в подсистеме 204), которая присоединена и выполнена с возможностью выполнять повышающее микширование над выходным сигналом каскада 203, чтобы формировать полностью декодированный подвергнутый повышающему микшированию звуковой сигнал, который выводится из декодера 200. В качестве альтернативы, постпроцессор 300 выполнен с возможностью выполнять повышающее микширование над выходным сигналом декодера 200 (например, с использованием метаданных PS, извлеченных деформатером 205, и/или управляющих битов, сформированных в подсистеме 204).

В ответ на метаданные, извлеченные деформатером 205, генератор 204 управляющих битов может формировать управляющие данные, и управляющие данные могут использоваться в пределах декодера 200 (например, в подсистеме завершающего повышающего микширования) и/или предъявляться в качестве выходного сигнала декодера 200 (например, в постпроцессор 300 для использования при постобработке). В ответ на метаданные, извлеченные из входного битового потока (и, по выбору, также в ответ на управляющие данные), каскад 204 может формировать (и предъявлять в постпроцессор 300) управляющие биты, указывающие, что декодированные звуковые данные, выдаваемые из каскада 203 обработки eSBR, должны подвергаться конкретному типу постобработки. В некоторых вариантах осуществления, декодер 200 выполнен с возможностью предъявлять метаданные, извлеченные деформатером 205 из входного битового потока, в постпроцессор 300, а постпроцессор 300 выполнен с возможностью выполнять постобработку над декодированными звуковыми данными, выведенными из декодера 200, с использованием метаданных.

Фиг. 4 – структурная схема блока (210) обработки звукового сигнала («APU»), который является еще одним вариантом осуществления обладающего признаками изобретения блока обработки

звукового сигнала. APU 210 является унаследованным декодером, который не выполнен с возможностью выполнять обработку eSBR. Любые из компонентов или элементов APU 210 могут быть реализованы в виде одного или более процессов и/или одной или более схем (например, ASIC (специализированных интегральных схем), FPGA (программируемых пользователем вентильных матриц) или других интегральных схем) в аппаратных средствах, программном обеспечении или комбинации аппаратных средств и программного обеспечения. APU 210 содержит буферную память 201, деформатер 215 (синтаксический анализатор) полезной нагрузки битового потока, подсистему 202 декодирования звукового сигнала (иногда упоминаемую как каскад «основного» декодирования или подсистема «основного» декодирования), и каскад 213 обработки SBR, соединенные, как показано. К тому же, типично, APU 210 включает в себя другие элементы обработки (не показаны). APU 210, например, может представлять собой кодировщик, декодер или перекодировщик звукового сигнала.

Элементы 201 и 202 APU 210 идентичны одинаково пронумерованным элементам декодера 200 (по фиг. 3), и вышеприведенное их описание повторяться не будет. При эксплуатации APU 210, последовательность блоков кодированного звукового битового потока (битового потока AAC MPEG-4), принимаемого APU 210, предъявляется из буфера 201 в деформатер 215.

Деформатер 215 присоединен и выполнен с возможностью демультимплексировать каждый блок битового потока, чтобы извлекать метаданные SBR (в том числе, данные квантованной огибающей) и также типично другие метаданные из него, но игнорировать метаданные eSBR, которые могут быть включены в битовый поток, в соответствии с любым вариантом осуществления настоящего изобретения. Деформатер 215 выполнен с возможностью добавлять по меньшей мере метаданные SBR в каскад 213 обработки SBR. Деформатер 215 также присоединен и выполнен с возможностью извлекать звуковые данные из каждого блока битового потока, и предъявлять извлеченные звуковые данные в подсистему 202 декодирования (каскад декодирования).

Подсистема 202 декодирования звука декодера 200 выполнена с возможностью декодировать звуковые данные, извлеченные деформатером 215 (такое декодирование может упоминаться как операция «основного» декодирования), для формирования декодированных звуковых данных и предъявлять декодированные звуковые данные в каскад 213 обработки SBR. Декодирование выполняется в частотной области. Типично, завершающий каскад обработки в подсистеме 202 применяет преобразование из частотной области во временную область к декодированным звуковым данным частотной области, так чтобы выходными данными подсистемы были декодированные звуковые данные временной области. Каскад 213 выполнен с возможностью применять средства SBR (но не средства eSBR), указанные метаданными SBR (извлеченными деформатером 215), к декодированным звуковым данным (то есть, выполнять обработку SBR над выходными данными подсистемы 202 декодирования с использованием метаданных SBR) для формирования полностью декодированных звуковых данных, которые выводятся (например, в постпроцессор 300) из APU 210. Типично, APU 210 включает в себя память (доступную из подсистемы 202 и каскада 213), которая хранит деформатированные звуковые данные и метаданные, выданные из деформатера 215, а каскад 213 выполнен с возможностью осуществлять доступ к звуковым данным и метаданным (в том числе, метаданным SBR) по мере надобности во время обработки SBR. Обработка SBR в каскаде 213 может считаться последующей обработкой на выходе из подсистемы 202 основного декодирования. По выбору, APU 210 также включает в себя подсистему завершающего повышающего микширования (которая может применять средства параметрической стереофонии («PS»), определенные в стандарте AAC MPEG-4, с использованием метаданных PS, извлеченных деформатером 215), которая присоединена и выполнена с возможностью выполнять повышающее микширование над выходным сигналом каскада 213, чтобы формировать полностью декодированный подвергнутый повышающему микшированию звуковой сигнал, который выводится из APU 210. В качестве альтернативы, постпроцессор выполнен с возможностью выполнять повышающее микширование над выходным сигналом APU 210 (например, с использованием метаданных PS, извлеченных

деформатером 215, и/или управляющих битов, сформированных в АРУ 210).

Различные реализации кодировщика 100, декодера 200 и АРУ 210 выполнены с возможностью выполнять разные варианты осуществления обладающего признаками изобретения способа.

В соответствии с некоторыми вариантами осуществления, метаданные eSBR включены (например, небольшое количество управляющих битов, которые являются метаданными eSBR, включены) в кодированный звуковой битовый поток (например, битовый поток AAC MPEG-4), так чтобы унаследованные декодеры (которые не выполнены с возможностью синтаксически анализировать метаданные eSBR и использовать какое бы то ни было средство eSBR, к которому относятся метаданные eSBR), могут игнорировать метаданные eSBR, но, тем не менее, декодировать битовый поток, насколько возможно без использования метаданных eSBR или какого бы то ни было средства eSBR, к которому относятся метаданные eSBR, типично без какого бы то ни было значительного ухудшения качества декодированного звукового сигнала. Однако, декодеры eSBR, выполненные с возможностью синтаксически анализировать битовый поток для идентификации метаданных eSBR и для использования по меньшей мере одного средства eSBR в ответ на метаданные eSBR, будут обладать выгодами использования по меньшей мере одного такого средства eSBR. Поэтому, варианты осуществления изобретения предусматривают средство для рациональной передачи управляющих данных усовершенствованной репликации полос спектра (eSBR) или метаданных обратно совместимым образом.

Типично, метаданные eSBR в битовом потоке являются указывающими (например служат признаком по меньшей мере одной характеристики или параметра) одного или более из следующих средств eSBR (которые описаны в стандарте MPEG USAC, и которые могли быть или могли не быть примененными кодировщиком во время формирования битового потока):

- Гармоническая транспозиция; и
- Дополнительная предварительная обработка вклейкой с помощью КЗФ (предварительное выравнивание).

Например, метаданные eSBR, включенные в битовый поток, могут быть указывающими значения параметров (описанных в стандарте USAC MPEG и в настоящем изобретении): sbrPitchingMode[ch], sbrOversamplingFlag[ch], sbrPitchInBins[ch], sbrPitchInBins[ch] и bs_sbr_preprocessing.

В материалах настоящей заявки, обозначение X[ch], где X - некоторый параметр, обозначает, что параметр относится к каналу («ch») звукового контента кодированного битового потока, подлежащего декодированию. Ради простоты, мы иногда опускаем выражение [ch] и допускаем, что соответствующий параметр относится к каналу звукового контента.

В материалах настоящей заявки, обозначение X[ch][env], где X - некоторый параметр, обозначает, что параметр относится к огибающей («env») SBR канала («ch») звукового контента кодированного битового потока, подлежащего декодированию. Ради простоты, мы иногда опускаем выражения [env] и [ch], и допускаем, что соответствующий параметр относится к огибающей SBR канала звукового контента.

Во время декодирования кодированного битового потока, выполнение гармонической транспозиции во время стадии обработки eSBR декодирования (для каждого канала, «ch», звукового контента, указанного битовым потоком), управляется следующими параметрами метаданных eSBR: sbrPitchingMode[ch]: sbrOversamplingFlag[ch]; sbrPitchInBinsFlag[ch]; и sbrPitchInBins[ch].

Значение «sbrPitchingMode[ch]» указывает тип транспозера, используемый в eSBR: sbrPitchingMode[ch]=1 указывает негармоническую вклейку, как описано в разделе 4.6.18.6.3 стандарта AAC MPEG-4; sbrPitchingMode[ch]=0 указывает гармоническую вклейку SBR, как указанная в разделе 7.5.3 или 7.5.4 стандарта USAC MPEG USAC.

Значение «sbrOversamplingFlag[ch]» указывает использование адаптивной по отношению к сигналу передискретизации частотной области в eSBR в комбинации основанной на ДПФ гармонической вклейки SBR, как описано в разделе 7.5.3 стандарта USAC MPEG. Этот флажковый признак управляет размером

ДПФ, которые используются в транспозере: 1 указывает

активированную адаптивную по отношению к сигналу передискретизацию частотной области, как описано в разделе 7.5.3.1 стандарта USAC MPEG; 0 указывает деактивированную адаптивную по отношению к сигналу передискретизацию частотной области, как описано в разделе 7.5.3.1 стандарта USAC MPEG.

Значение «sbrPitchInBinsFlag[ch]» управляет интерпретацией параметра sbrPitchInBins[ch]: 1 указывает, что значение в sbrPitchInBins[ch] действительно и больше нуля; 0 указывает, что значение sbrPitchInBins[ch] установлено в ноль.

Значение «sbrPitchInBins[ch]» управляет добавлением членов векторного произведения в гармонический транспозер SBR. Значение sbrPitchInBins[ch] является целочисленным значением в диапазоне [0,127] и представляет собой расстояние, измеренное в элементах разрешения по частоте для 1536-точечное ДПФ, действующее на частоте дискретизации основного кодера.

В случае, если битовый поток AAC MPEG-4 указывает пару каналов SBR, чьи каналы не соединены (вместо одиночного канала SBR), битовый поток является указывающим два экземпляра вышеприведенного синтаксиса (для гармонической или негармонической транспозиции), один для каждого канала sbr_channel_pair_element().

Гармоническая транспозиция средства eSBR типично улучшает качество декодированных музыкальных сигналов с относительно низкими частотами разделения. Негармоническая транспозиция (то есть, унаследованная спектральная вклейка) типично улучшает речевые сигналы. Отсюда, отправная точка при принятии решения в отношении того, какой тип транспозиции предпочтителен для кодирования конкретного звукового контента, состоит в том, чтобы выбирать способ транспозиции в зависимости от выявления речи/музыки, причем, гармоническая транспозиция выполняется над музыкальным контентом, и спектральная вклейка над речевым контентом.

Выполнение предварительного выравнивания во время обработки eSBR управляется значением однобитного параметра метаданных eSBR, известного как «bs_sbr_preprocessing», в том смысле, что предварительное выравнивание выполняется или не выполняется в

зависимости от значения этого одиночного бита. Когда используется алгоритм вклейки с помощью КЗФ SBR, как описанный в разделе 4.6.18.6.3 стандарта AAC MPEG-4, этап предварительного выравнивания может выполняться (когда указан параметром «bs_sbr_preprocessing») в попытке избежать отсутствия непрерывности профиля огибающей спектра высокочастотного сигнала, подаваемого в последующий корректор огибающей (корректор огибающей играет роль еще одной стадии обработки eSBR). Предварительное выравнивание типично улучшает действие последующего каскада коррекции огибающей, давая в результате сигнал в полосе высоких частот, который воспринимается более устойчивым.

Ожидается, что общее требование к скорости передачи в битах для включения в метаданные eSBR битового потока AAC MPEG-4, указывающие упомянутые выше средства eSBR (гармоническую транспозицию и предварительное выравнивание) будет порядка нескольких сотен бит в секунду, так как только отличительные управляющие данные, необходимые для выполнения обработки eSBR, передаются в соответствии с некоторыми вариантами осуществления изобретения. Унаследованные декодеры могут игнорировать эту информацию, так как она включена обратно совместимым образом (как будет пояснено позже). Поэтому, пагубное влияние на скорости передачи битов, связанную с включением в состав метаданных eSBR, незначительно по ряду причин, в том числе, следующим:

- Ухудшение скорости передачи битов (обусловленное включением в состав метаданных eSBR) является весьма небольшой долей полной скорости передачи битов, так как передаются только отличительные управляющие данные, необходимые для выполнения обработки eSBR (а не одновременная передача управляющих данных SBR); и
- Настройка связанной с SBR управляющей информации типично не зависит от подробностей транспозиции.

Таким образом, варианты осуществления изобретения предусматривают средство для рациональной передачи управляющих данных усовершенствованной репликации полос спектра (eSBR) или

метаданных обратно совместимым образом. Рациональная передача управляющих данных eSBR уменьшает требования к памяти в декодерах, кодировщиках и перекодировщиках, применяющих аспекты изобретения, тем временем не подвергаясь ощутимому неблагоприятному воздействию на скорость передачи битов. Более того, сложность и требования к обработке, связанные с выполнением eSBR в соответствии с вариантами осуществления изобретения, также уменьшаются, так как данные SBR необходимо обрабатывать только один раз, и нет одновременной передачи, которая имела бы место, если бы eSBR обрабатывался в виде полностью отдельного типа объекта в AAC MPEG-4 вместо встраивания в кодек AAC MPEG-4 обратно совместимым образом.

Затем, со ссылкой на фиг. 7, опишем элементы блока («raw_data_block») битового потока AAC MPEG-4, в котором метаданные eSBR включены в состав в соответствии с некоторыми вариантами осуществления настоящего изобретения. Фиг. 7 – схема блока («raw_data_block») битового потока AAC MPEG-4, показывающая некоторые его сегменты.

Блок битового потока AAC MPEG-4 может включать в себя по меньшей мере один «single_channel_element()» (например, элемент одиночного канала, показанный на фиг. 7), и/или по меньшей мере один «channel_pair_element()» (особо не показанный на фиг. 7, хотя он может присутствовать), в том числе, звуковые данные для звуковой программы. Блок также может включать в себя некоторое количество «fill_elements» (например, элемент 1 заполнения и/или элемент 2 заполнения по фиг. 7), включающих в себя данные (например, метаданные), имеющие отношение к программе. Каждый «single_channel_element()» включает в себя идентификатор (например, «ID1» по фиг. 7), указывающий начало элемента одиночного канала, и может включать в себя звуковые данные, указывающие другой канал многоканальной звуковой программы. Каждый «channel_pair_element» включает в себя идентификатор (не показан на фиг. 7), указывающий начало элемента пары каналов, и может включать в себя звуковые данные, указывающие два канала программы.

fill_element (в материалах настоящей заявки упоминаемый как

элемент заполнения) битового потока AAC MPEG-4 включает в себя идентификатор («ID2» по фиг. 7), указывающий начало элемента заполнения, и данные заполнения после идентификатора. Идентификатор ID2 может состоять из трехбитного беззнакового целого числа, передаваемого старшим битом вперед («uimsbf»), имеющего значение 0×6. Данные заполнения могут включать в себя элемент `extension_payload()` (иногда в материалах настоящей заявки упоминаемый как полезная нагрузка расширения), чей синтаксис показан в Таблице 4.57 стандарта AAC MPEG-4. Несколько типов полезной нагрузки расширения существуют и идентифицируются посредством параметра «`extension_type`», который является четырехбитным беззнаковым целым числом, передаваемым старшим битом вперед («uimsbf»).

Данные заполнения (например, их полезная нагрузка расширения) могут включать в себя заголовок или идентификатор (например, «заголовок 1» по фиг. 7), который указывает сегмент данных заполнения, которые являются указывающими объект SBR (то есть, заголовок инициализирует тип «объекта SBR», указываемый ссылкой как `sbr_extension_data()` в стандарте AAC MPEG-4). Например, полезная нагрузка расширения репликации полос спектра (SBR) идентифицируется значением '1101' или '1110', что касается поля `extension_type` в заголовке, причем, '1101' идентифицирует полезную нагрузку расширения с данными SBR, а '1110' идентифицирует и полезную нагрузку расширения с данными SBR с циклическим избыточным кодом (CRC) для проверки правильности данных SBR.

Когда заголовок (например, поле `extension_type`) инициализирует тип объекта SBR, метаданные SBR (иногда упоминаемые как «данные репликации полос спектра» и указываемые ссылкой как `sbr_data()` в стандарте AAC MPEG-4) следуют за заголовком, и по меньшей мере один элемент расширения репликации полос спектра (например, «элемент расширения SBR» элемента 1 заполнения по фиг. 7) может следовать за метаданными SBR. Такой элемент расширения репликации полос спектра (сегмент битового потока) указывается ссылкой как контейнер «`sbr_extension()`» в

стандарте AAC MPEG-4. Элемент расширения репликации полос спектра по выбору включает в себя заголовок (например, «заголовок расширения SBR» элемента 1 заполнения по фиг. 7).

Стандарт AAC MPEG-4 предполагает, что элемент расширения репликации полос спектра может включать в себя данные PS (параметрической стереофонии) для звуковых данных программы. Стандарт AAC MPEG-4 предполагает, что, когда заголовок элемента заполнения (например, его полезной нагрузки расширения) инициализирует тип объекта SBR (как это делает «header1» по фиг. 7), и элемент расширения репликации полос спектра элемента заполнения включает в себя данные PS, элемент заполнения (например, его полезная нагрузка расширения) включает в себя данные репликации полос спектра, и параметр «bs_extension_id», чье значение (то есть, bs_extension_id=2) указывает, что данные PS включены в элемент расширения репликации полос спектра элемента заполнения.

В соответствии с некоторыми вариантами осуществления настоящего изобретения, метаданные eSBR (например, флажковый признак, указывающий, должна ли обработка усовершенствованной репликации полос спектра (eSBR) выполняться над звуковым контентом блока). Например, такой флажковый признак указывается в элементе 1 заполнения по фиг. 7, где флажковый признак встречается после заголовка («заголовка расширения SBR» элемента 1 заполнения) «элемента расширения SBR» элемента 1 заполнения. По выбору, такой флажковый признак и дополнительные данные eSBR включены в элемент расширения репликации полос спектра после заголовка элемента расширения репликации полос спектра (например, в элемент расширения SBR элемента 1 заполнения на фиг. 7, после заголовка расширения SBR). В соответствии с некоторыми вариантами осуществления настоящего изобретения, элемент заполнения, который включает в себя метаданные eSBR, также включает в себя параметр «bs_extension_id», чье значение (например, bs_extension_id=3) указывает, что метаданные eSBR включены в элемент заполнения, и что обработка eSBR должна быть выполнена над звуковым контентом уместного блока.

В соответствии с некоторыми вариантами осуществления

изобретения, метаданные, eSBR включены в элемент заполнения (например, элемент 2 заполнения по фиг. 7) битового потока AAC MPEG-4, иные чем в элементе расширения репликации полос спектра (элементе расширения SBR) элемента заполнения. Причина в том, что элементы заполнения, содержащие в себе `extension_payload()` с данными SBR или данные SBR с CRC, не содержат в себе никакой другой полезной нагрузки расширения любого другого типа расширения. Поэтому, в вариантах осуществления, где метаданные eSBR хранятся своей собственной полезной нагрузкой расширения, отдельный элемент заполнения, используется для хранения метаданных eSBR. Такой элемент заполнения включает в себя идентификатор (например, «ID2» по фиг. 7), указывающий начало элемента заполнения, и данные заполнения после идентификатора. Данные заполнения могут включать в себя элемент `extension_payload()` (иногда в материалах настоящей заявки упоминаемый как полезная нагрузка расширения), чей синтаксис показан в Таблице 4.57 стандарта AAC MPEG-4. Данные заполнения (например, их полезная нагрузка расширения) включает в себя заголовок (например, «header2» элемента 2 заполнения по фиг. 7), который является указывающим объект eSBR (то есть, заголовок инициализирует тип объекта усовершенствованной репликации полос спектра (eSBR)), и данные заполнения (например, его полезная нагрузка расширения) включают в себя метаданные eSBR после заголовка. Например, элемент 2 заполнения по фиг. 7 включает в себя такой заголовок («header2») и также включает в себя, после заголовка, метаданные eSBR (то есть, «флажковый признак» в элементе 2 заполнения, который указывает, должна ли обработка усовершенствованной репликации полос спектра (eSBR) выполняться над звуковым контентом блока). По выбору, дополнительные метаданные eSBR также включены в данные заполнения элемента 2 заполнения по фиг. 7 после заголовка 2. В вариантах осуществления, описываемым в настоящем параграфе, заголовок (например, заголовок 2 по фиг. 7) имеет идентификационное значение, которое не является одним из традиционных значений, заданных в Таблице 4.57 стандарта AAC MPEG-4, и, взамен, является указывающим полезную нагрузку расширения eSBR (так что

поле `extension_type` указывает, что данные заполнения включают в себя метаданные eSBR).

В первой группе вариантов осуществления, изобретение является блоком обработки звукового сигнала (например, декодером), содержащим:

память (например, буфер 201 по фиг. 3 или 4), выполненную с возможностью хранить по меньшей мере один блок кодированного звукового битового потока (например, по меньшей мере один блок битового потока AAC MPEG-4);

деформатер полезной нагрузки битового потока (например, элемент 205 по фиг. 3 или элемент 215 по фиг. 4), присоединенный к памяти и выполненный с возможностью демультиплексировать по меньшей мере одну часть упомянутого блока битового потока; и

подсистему декодирования (например, элементы 202 и 203 по фиг. 3 или элементы 202 и 213 по фиг. 4), присоединенную и выполненную с возможностью декодировать по меньшей мере одну часть звукового контента упомянутого блока битового потока, при этом, блок включает в себя:

элемент заполнения, включающий в себя идентификатор, указывающий начало элемента заполнения (например, идентификатор «`id_syn_ele`», имеющий значение 0x6, по Таблице 4.85 стандарта AAC MPEG-4), и данные заполнения после идентификатора, при этом, данные заполнения включают в себя:

по меньшей мере один флажковый признак, идентифицирующий, должна ли обработка усовершенствованной репликации полос спектра (eSBR) выполняться над звуковым контентом блока (например, с использованием данных репликации полос спектра и метаданных eSBR, включенных в блок).

Флажковым признаком являются метаданные eSBR, и примером флажкового признака является флажковый признак `sbrPatchingMode`. Еще одним примером флажкового признака является флажковый признак гармонической SBR. Оба этих флажковых признака указывают, базовая форма репликации полос спектра или усовершенствованная форма репликации спектра должна выполняться над звуковыми данными блока. Базовой формой репликации спектра

является спектральная вклейка, а усовершенствованной формой репликации полос спектра является гармоническая транспозиция.

В некоторых вариантах осуществления, данные заполнения также включают в себя дополнительные метаданные eSBR (то есть, метаданные eSBR, иные чем флажковый признак).

Память может быть буферной памятью (например, реализацией буфера 201 по фиг. 4), которая хранит (например, неэфемерным образом) по меньшей мере один блок кодированного звукового битового потока.

Оценивается, что сложность выполнения обработки eSBR (с использованием гармонической транспозиции и предварительного выравнивания eSBR) посредством декодера eSBR во время декодирования битового потока AAC MPEG-4, который включает в себя метаданные eSBR (указывающие эти средства eSBR), была бы следующей (для типичного декодирования с указанными параметрами):

- Гармоническая транспозиция (16 кбит/с, 14400/28800 Гц)
 - о на основании ДПФ: 3,68 WMOPS (взвешенных миллионов операций в секунду);
 - о на основании КЗФ: 0,98 WMOPS;
- предварительная обработка вклейки с помощью КЗФ (предварительное выравнивание): 0,1 WMOPS. Известно, что основанная на ДПФ транспозиция типично работает лучше, чем основанная на КЗФ транспозиция, применительно к переходным процессам.

В соответствии с некоторыми вариантами осуществления настоящего изобретения, элемент заполнения (кодированного звукового битового потока), который включает в себя метаданные eSBR, также включает в себя параметр (например, параметр «bs_extension_id»), чье значение (например, bs_extension_id=3) сигнализирует, что метаданные eSBR включены в элемент заполнения, и что обработка eSBR должна быть выполнена над звуковым контентом уместного блока, и/или параметр (например, тот же самый параметр «bs_extension_id», чье значение (например, bs_extension_id=2) сигнализирует, что контейнер sbr_extension() элемента заполнения включает в себя данные PS. Например, как

указано в Таблице 1, приведенной ниже, такой параметр, имеющий значение `bs_extension_id=2`, может сигнализировать, что контейнер `sbr_extension()` элемента заполнения включает в себя данные PS, и такой параметр, имеющий значение `bs_extension_id=3`, может сигнализировать, что контейнер `sbr_extension()` элемента заполнения включает в себя метаданные eSBR:

Таблица 1

bs_extension_id	Значение
0	Зарезервировано
1	Зарезервировано
2	EXTENSION_ID_PS
3	EXTENSION_ID_ESBR

В соответствии с некоторыми вариантами осуществления изобретения, синтаксис каждого элемента расширения репликации полос спектра, который включает в себя метаданные eSBR и/или данные PS является таким, как указано в Таблице 2, приведенной выше (в которой «`sbr_extension()`» обозначает контейнер, который является элементом расширения репликации полос спектра, «`bs_extension_id`» является таким, как описано в Таблице 1, приведенной выше, «`ps_data`» обозначает данные PS, а «`esbr_data`» обозначает метаданные eSBR):

Таблица 2

<code>sbr_extension(bs_extension_id, num_bits_left)</code>	
<code>{</code>	
<code>switch (bs_extension_id) {</code>	
<code>case EXTENSION_ID_PS:</code>	
<code>num_bits_left -= ps_data();</code>	Примечание 1
<code>break;</code>	
<code>case EXTENSION_ID_ESBR:</code>	
<code>num_bits_left -= esbr_data();</code>	Примечание 2
<code>break;</code>	
<code>default:</code>	
<code>bs_fill_bits;</code>	
<code>num_bits_left=0;</code>	
<code>break;</code>	

}	
}	
Примечание 1: <code>ps_data()</code> возвращает количество считанных бит.	
Примечание 2: <code>esbr_data()</code> возвращает количество считанных бит.	

В примерном варианте осуществления, `esbr_data()`, упомянутый в Таблице 2, приведенной выше, является указывающим значения следующих параметров метаданных:

1. однобитный параметр метаданных, «`bs_sbr_preprocessing`»; и

2. для каждого канала («`ch`») звукового контента кодированного битового потока, подлежащего декодированию, каждый из описанных выше параметров: «`sbrPatchingMode[ch]`»; «`sbrOversamplingFlag[ch]`»; «`sbrPitchInBinsFlag[ch]`»; и «`sbrPitchInBins[ch]`».

Например, в некоторых вариантах осуществления, `esbr_data()` может иметь синтаксис, указанный в Таблице 3, для указания этих параметров метаданных:

Таблица 3

Синтаксис	Количество битов
<code>esbr_data(id_aac, bs_coupling)</code>	
{	
<code>bs_sbr_preprocessing</code>;	1
<code>if (id_aac == ID_SCE) {</code>	
<code>if (<code>sbrPatchingMode</code>[0] == 0) {</code>	1
<code>sbrOversamplingFlag</code> [0];	1
<code>if (<code>sbrPitchInBinsFlag</code>[0])</code>	1
<code>sbrPitchInBins</code> [0];	7
<code>else</code>	
<code>sbrPitchInBins[0]=0;</code>	
<code>} else {</code>	
<code>sbrOversamplingFlag[0]=0;</code>	
<code>sbrPitchInBins[0]=0;</code>	
}	

} else if (id_aac == ID_CPE) {	
If (bs_coupling) {	
if (sbrPatchingMode[0,1] == 0) {	1
sbrOversamplingFlag[0,1];	1
if (sbrPitchInBinsFlag[0,1])	1
sbrPitchInBins[0,1];	7
else	
sbrPitchInBins[0,1]=0;	
} else {	
sbrOversamplingFlag[0,1]=0;	
sbrPitchInBins[0,1]=0;	
}	
} else { /* bs_coupling == 0 */	
if (sbrPatchingMode[0] == 0) {	1
sbrOversamplingFlag[0];	1
if (sbrPitchInBinsFlag[0])	1
sbrPitchInBins[0];	7
else	
sbrPitchInBins[0]=0;	
} else {	
sbrOversamplingFlag[0]=0;	
sbrPitchInBins[0]=0;	
}	
if (sbrPatchingMode[1] == 0) {	1
sbrOversamplingFlag[1];	1
if (sbrPitchInBinsFlag[1])	1
sbrPitchInBins[1];	7
else	
sbrPitchInBins[1]=0;	
} else {	
sbrOversamplingFlag[1]=0;	
sbrPitchInBins[1]=0;	
}	
}	

}	
}	
Примечание: bs_sbr_preprocessing определен, как описано в разделе 6.2.12 ISO/IEC 23003-3:2012. sbrPatchingMode[ch], sbrOversamplingFlag[ch], sbrPitchInBinsFlag[ch] и sbrPitchInBins[ch] определены, как описано в разделе 7.5 ISO/IEC 23003-3:2012.	

Приведенный выше синтаксис дает возможность рациональной реализации усовершенствованной формы репликации полос спектра, такой как гармоническая транспозиция, в качестве расширения унаследованного декодера. Более точно, данные eSBR из Таблицы 3 включают в себя только те параметры, необходимые для выполнения усовершенствованной формы репликации полос спектра, которые еще не поддерживаются в битовом потоке и не являются выводимыми непосредственно из параметров, уже поддерживаемых в битовом потоке. Все другие параметры и данные обработки, необходимые для выполнения усовершенствованной формы репликации полос спектра извлекаются из существующих ранее параметров в уже определенных расположениях в битовом потоке.

Например, совместимый с HE-AAC MPEG-4 или с HE-AAC v2 декодер может быть расширен, чтобы включать в себя усовершенствованную форму репликации полос спектра, такую как гармоническая транспозиция. Эта усовершенствованная форма репликации полос спектра является дополнением к базовой форме репликации полос спектра, уже поддерживаемой декодером. В контексте совместимого с HE-AAC или HE-AAC v2 MPEG-4 декодера, эта базовая форма репликации полос спектра является средством SBR спектральной вклейки с помощью КЭФ, как определено в разделе 4.6.18 стандарта AAC MPEG-4.

При выполнении усовершенствованной формы репликации полос спектра, декодер расширенного HE-AAC может повторно использовать многие параметры битового потока, уже включенные в полезную нагрузку расширения SBR битового потока. Конкретные параметры, которые могут повторно использоваться, например, включают в себя различные параметры, которые определяют сводную таблицу полос

частот. Эти параметры включают в себя `bs_start_freq` (параметр, который определяет начало параметра сводной таблицы частот), `bs_stop_freq` (параметр, который определяет окончание сводной таблицы частот), `bs_freq_scale` (параметр, который определяет количество полос частот на октаву) и `bs_alter_scale` (параметр, который изменяет шкалу полос частот). Параметры, которые могут повторно использоваться, также включают в себя параметры, которые определяют таблицу шумовых полос (`bs_noise_bands`) и параметры таблицы ограничительных полос (`bs_limiter_bands`). Соответственно, в различных вариантах осуществления, по меньшей мере некоторые из эквивалентных параметров, заданных в стандарте USAC, не включены в битовый поток, тем самым, сокращая служебные сигналы и данные управления в битовом потоке. Типично, в тех случаях, когда параметр, заданный в стандарте AAC, имеет эквивалентный параметр, заданный в USAC, эквивалентный параметр, заданный в USAC, имеет такое же наименование, как параметр, заданный в стандарте AAC, например, масштабный коэффициент огибающей, `EorigMapped`. Однако, эквивалентный параметр, заданный в стандарте USAC, типично имеет другое значение, которое «приспособлено» скорее для обработки усовершенствованной SBR, определенной в стандарте USAC, нежели для обработки SBR, определенной в стандарте AAC.

В дополнение к многочисленным параметрам, другие элементы данных также могут повторно использоваться декодером расширенного HE-AAC при выполнении усовершенствованной формы репликации полос спектра в соответствии с вариантами осуществления изобретения. Например, данные огибающей и данные уровня собственных шумов также могут извлекаться из данных `bs_data_env` (масштабных коэффициентов огибающей) и `bs_noise_env` (масштабных коэффициентов уровня собственных шумов) и использоваться во время усовершенствованной формы репликации полос спектра.

По сути, эти варианты осуществления используют конфигурационные параметры и данные огибающей, уже поддерживаемые унаследованным декодером HE-AAC или HE-AAC v2 в полезной нагрузке расширения SBR, чтобы сделать возможной

усовершенствованную форму репликации полос спектра, требуя как можно меньшего количества дополнительных передаваемых данных. Метаданные изначально были приспособлены для базовой формы HFR (например, спектральной вклейки SBR), но, в соответствии с вариантами осуществления, используется для усовершенствованной формы HFR (например, гармонической транспозиции eSBR). Как обсуждено ранее, метаданные как правило представляют собой рабочие параметры (например масштабные коэффициенты огибающей, масштабные коэффициенты уровня собственных шумов, параметры время/частотной сетки, информацию о синусоидальном дополнении, переменную частоту/полосу разделения, режим обратной фильтрации, разрешение огибающей, режим сглаживания, режим частотной интерполяции). приспособленные и предназначенные для использования с базовой формой HFR (например, линейным преобразованием). Однако, эти метаданные, комбинированные с дополнительными параметрами метаданных, специфичными усовершенствованной форме HFR (например, гармонической транспозиции), могут использоваться для рациональной и эффективной обработки звуковых данных с использованием усовершенствованной формы HFR.

Соответственно, расширенные декодеры, которые поддерживают усовершенствованную форму репликации полос спектра, могут быть созданы очень эффективным образом, полагаясь на уже определенные элементы битового потока (например, таковые в полезной нагрузке расширения SBR) и добавляя только те параметры, которые необходимы для поддержки усовершенствованной формы репликации полос спектра (в полезной нагрузке расширения элемента заполнения). Этот признак сокращения данных в комбинации с размещением вновь добавленных параметров в зарезервированных полях данных, таких как контейнер расширения, существенно уменьшает барьеры к созданию декодера, который поддерживает усовершенствованную репликацию полос спектра, гарантируя, что битовый поток обратно совместим с унаследованным декодером, не поддерживающим усовершенствованную форму репликации полос спектра.

В таблице 3, номер в правом столбце указывает количество

битов соответствующего параметра в левом столбце.

В некоторых вариантах осуществления, тип объекта SBR, определенный в AAC MPEG-4, обновлен, чтобы содержать в себе средство SBR или аспекты средства усовершенствованной SBR (eSBR), что сигнализируется в элементе расширения SBR (`bs_extension_id== EXTENSION_ID_ESBR`).

В некоторых вариантах осуществления, изобретение является способом, включающим в себя этап кодирования звуковых данных для формирования кодированного битового потока (например, битового потока AAC MPEG-4), в том числе, посредством включения метаданных eSBR в по меньшей мере один сегмент по меньшей мере одного блока кодированного битового потока и звуковых данных в по меньшей мере один другой сегмент блока. В типичных вариантах осуществления, способ включает в себя этап мультиплексирования звуковых данных с метаданными eSBR в каждом блоке кодированного битового потока. При типичном декодировании кодированного битового потока в декодере eSBR, декодер извлекает метаданные eSBR из битового потока (в том числе, посредством синтаксического разбора и демультимплексирования метаданных eSBR и звуковых данных) и использует метаданные eSBR для обработки звуковых данных, чтобы формировать поток декодированных звуковых данных.

Еще одним аспектом изобретения является декодер eSBR, выполненный с возможностью выполнять обработку eSBR (например, с использованием по меньшей мере одного из средств eSBR, известных как гармоническая транспозиция или предварительное выравнивание) во время декодирования кодированного звукового потока (например, битового потока AAC MPEG-4), который не включает в себя метаданные eSBR. Пример такого декодера будет описан со ссылкой на фиг. 5.

Декодер (400) eSBR по фиг. 5 включает в себя буферную память 201 (которая идентична памяти 201 по фиг. 3 и 4), деформатор 215 полезной нагрузки битового потока (который идентичен деформатору 215 по фиг. 4), подсистему 202 декодирования звукового сигнала (иногда упоминаемую как каскад «основного» декодирования или подсистема «основного»

декодирования, и которая идентична подсистеме 202 основного декодирования по фиг. 3), подсистему 401 формирования управляющих данных eSBR и каскад 203 обработки eSBR (который идентичен каскаду 203 по фиг. 3, соединенные как показано. К тому же, типично, декодер 400 включает в себя другие элементы обработки (не показаны).

При эксплуатации декодера 400, последовательность блоков кодированного звукового битового потока (битового потока AAC MPEG-4), принимаемого декодером 400, предъявляется из буфера 201 в деформатер 215.

Деформатер 215 присоединен и выполнен с возможностью демультимплексировать каждый блок битового потока для извлечения метаданных SBR (в том числе квантованных данных огибающей) и также типично других метаданных из него. Деформатер 215 выполнен с возможностью добавлять по меньшей мере метаданные SBR в каскад 203 обработки eSBR. Деформатер 215 также присоединен и выполнен с возможностью извлекать звуковые данные из каждого блока битового потока, и предъявлять извлеченные звуковые данные в подсистему 202 декодирования (каскад декодирования).

Подсистема 202 декодирования звука декодера 400 выполнена с возможностью декодировать звуковые данные, извлеченные деформатером 215 (такое декодирование может упоминаться как операция «основного» декодирования), для формирования декодированных звуковых данных и предъявлять декодированные звуковые данные в каскад 203 обработки eSBR. Декодирование выполняется в частотной области. Типично, завершающий каскад обработки в подсистеме 202 применяет преобразование из частотной области во временную область к декодированным звуковым данным частотной области, так чтобы выходными данными подсистемы были декодированные звуковые данные временной области. Каскад 203 выполнен с возможностью применять средства SBR (и средства eSBR), указанные метаданными SBR (извлеченными деформатером 215) и метаданными eSBR, сформированными в подсистеме 401, к декодированным звуковым данным (то есть, выполнять обработку SBR и eSBR над выходными данными подсистемы 202 декодирования с использованием метаданных SBR и eSBR) для формирования полностью

декодированных звуковых данных, которые выводятся из декодера 400. Типично, декодер 400 включает в себя память (доступную из подсистемы 202 и каскада 203), которая хранит деформированные звуковые данные и метаданные, выданные из деформатора 215 (и, по выбору, также подсистемы 401), а каскад 203 выполнен с возможностью осуществлять доступ к звуковым данным и метаданным по мере необходимости во время обработки SBR и eSBR. Обработка SBR в каскаде 203 может считаться последующей обработкой на выходе из подсистемы 202 основного декодирования. По выбору, декодер 400 также включает в себя подсистему завершающего повышающего микширования (которая может применять средства параметрической стереофонии («PS»), определенные в стандарте AAC MPEG-4, с использованием метаданных PS, извлеченных деформатором 215), которая присоединена и выполнена с возможностью выполнять повышающее микширование над выходным сигналом каскада 203, чтобы формировать полностью декодированный подвергнутый повышающему микшированию звуковой сигнал, который выводится из APU 210.

Подсистема 401 формирования управляющих данных по фиг. 5 присоединена и выполнена с возможностью выявлять по меньшей мере одно свойство кодированного звукового битового потока, подлежащего декодированию, и формировать управляющие данные eSBR (которые могут быть или включать в себя метаданные eSBR любого из типов, включенных в кодированные звуковые битовые потоки, в соответствии с другими вариантами осуществления изобретения) в ответ на по меньшей мере один результат этапа выявления. Управляющие данные eSBR добавляются в каскад 203, чтобы приводить в действие применение отдельных средств eSBR или комбинации средств eSBR при выявлении конкретного свойства (или комбинации свойств) битового потока и/или управлять применением таких средств eSBR. Например, для того чтобы контролировать выполнение обработки eSBR с использованием гармонической транспозиции, некоторые варианты осуществления подсистемы 401 формирования управляющих данных включали бы в себя: детектор музыки (например, упрощенный вариант традиционного детектора музыки) для установки параметра `sbrPatchingMode[ch]` (и предъявления установленного параметра в каскад 203) в ответ на

выявление, что битовый поток является или не является указывающим музыку; детектор переходных процессов для установки параметра `sbrOversamplingFlag[ch]` (и предъявления установленного параметра в каскад 203) в ответ на выявление наличия или отсутствия переходных процессов в звуковом контенте, указываемом битовым потоком; и/или детектор основного тона для установки параметров `sbrPitchInBinsFlag[ch]` и `sbrPitchInBins[ch]` (и предъявления установленных параметров в каскад 203) в ответ на выявление основного тона звукового контента, указываемого битовым потоком. Другими аспектами изобретения являются способы декодирования звукового битового потока, выполняемые любым вариантом осуществления обладающего признаками изобретения декодером, описанным в этом параграфе и предыдущем параграфе.

Аспекты изобретения включают в себя способ кодирования или декодирования типа, который любой вариант осуществления обладающих признаками изобретения APU, системы или устройства выполнен с возможностью (или запрограммирован) выполнять. Другие аспекты изобретения включают в себя систему или устройство выполненные с возможностью (например, запрограммированные) выполнять любой вариант осуществления обладающего признаками изобретения способа, и машинно-читаемый носитель (например, диск), который хранит машинную программу (например, неэфемерным образом) для реализации любого варианта осуществления обладающего признаками изобретения способа или его этапов. Например, обладающая признаками изобретения система может быть или включать в себя программируемый процессор общего применения, цифровой сигнальный процессор или микропроцессор, запрограммированный программным обеспечением или микропрограммным обеспечением и/или иным образом выполненный с возможностью выполнять любые из многообразия операций над данными, в том числе, вариант осуществления обладающего признаками изобретения способа или его этапов. Такой процессор общего применения может быть или включать в себя компьютерную систему, включающую в себя устройство ввода, память, схему обработки, запрограммированную (и/или иным образом выполненную с возможностью) выполнять вариант осуществления обладающего

признаками изобретения способа (или его этапов) в ответ на данные, предъявленные ему.

Варианты осуществления настоящего изобретения могут быть реализованы в аппаратных средствах, микропрограммном обеспечении или программном обеспечении, либо комбинацией того и другого (например, в виде программируемой логической матрицы). Если не указан иной способ действий, алгоритмы и процессы, включенные в состав в качестве части изобретения, по своей природе, не имеют отношения к какому бы то ни было конкретному компьютеру или другому устройству. В частности, различные машины общего применения могут использоваться с программами, написанными в соответствии с доктринами, приведенными в материалах настоящей заявки, или может быть более удобным сконструировать более специализированное устройство (например, интегральные схемы) для выполнения требуемых этапов способа. Таким образом, изобретение может быть реализовано в одной или более компьютерных программ, исполняющихся в одной или более программируемых компьютерных систем (например, реализации любого из элементов по фиг. 1 или кодировщика 100 по фиг. 2 (или его элемента), или декодера 200 по фиг. 3 (или его элемента), или декодера 210 по фиг. 4 (или его элемента) или декодера 400 по фиг. 5 (или его элемента)), каждая из которых содержит по меньшей мере один процессор, по меньшей мере одну систему хранения данных (в том числе, энергозависимые или энергонезависимые элементы памяти и/или запоминающие элементы), по меньшей мере одно устройство или порт ввода и по меньшей мере одно устройство или порт вывода. Управляющая программа применяется к входным данным для выполнения функций, описанных в материалах настоящей заявки, и формирует выходную информацию. Выходная информация направляется в одно или более устройств вывода известным образом.

Каждая такая программа может быть реализована на любом желательном компьютерном языке (включая машинные, компоновочные или высокоуровневые процедурные, логические или объектно-ориентированные языки программирования) для обмена информацией с компьютерной системой. В любом случае, язык может быть компилируемым или интерпретируемым языком.

Например, когда реализуются последовательностями команд компьютерного программного обеспечения, различные функции и этапы вариантов осуществления изобретения могут реализовываться последовательностями команд многопоточного программного обеспечения, работающими на пригодных аппаратных средствах цифровой обработки сигналов, в каком случае, различные устройства, этапы и функции вариантов осуществления могут соответствовать порциям команд программного обеспечения.

Каждая такая компьютерная программа предпочтительно хранится на или загружается на запоминающие носители или в устройство (например, твердотельную память или носители, либо магнитные или оптические носители), читаемые программируемым компьютером общего применения или специального назначения, для конфигурирования и управления компьютером, когда запоминающие носители или устройство считываются компьютерной системой, чтобы выполнять процедуры, описанные в материалах настоящей заявки. Обладающая признаками изобретения система также может быть реализована в виде машинно-читаемого запоминающего носителя, сконфигурированного компьютерной программой (то есть, хранящего компьютерную программу), где запоминающий носитель, сконфигурированный таким образом, побуждает компьютерную систему действовать конкретным и predetermined образом для выполнения функций, описанных в материалах настоящей заявки.

Было описано некоторое количество вариантов осуществления изобретения. Тем не менее, будет понятно, что различные модификации могут быть произведены, не выходя из сущности и объема изобретения. Многие модификации и разновидности настоящего изобретения возможны в свете вышеприведенных доктрин. Например, для того чтобы содействовать эффективным реализациям, фазовые сдвиги могут использоваться в комбинацией с гребенками фильтров разложения сигнала и синтеза сигнала комплексным КЗФ. Гребенка фильтров разложения сигнала ответственна за фильтрацию сигнала в полосе низких частот временной области, сформированного основным декодером, в множество поддиапазонов (например, поддиапазонов КЗФ). Гребенка фильтров разложения сигнала ответственна за комбинирование восстановленной полосы

высоких частот, созданной посредством выбранной технологии HFR (которая указана принятым параметром `sbrPatchingMode`), с декодированной полосой низких частот, чтобы синтезировать широкополосный выходной звуковой сигнал. Данная реализация гребенки фильтров, работающая в определенном режиме частоты выборки, например, режиме нормальной работы на удвоенной частоте или SBR с понижающей дискретизацией, однако, не должна иметь фазовых сдвигов, которые зависят от битового потока. Гребенки КЗФ, используемые в SBR являются расширением сложной экспоненциальной функции по теории гребенок фильтров с косинусоидальной модуляцией. Может быть показано, что ограничения подавления паразитных сигналов становятся устаревшими при расширении гребенки фильтров с косинусоидальной модуляцией комплексной экспоненциальной модуляцией. Таким образом, что касается гребенок из КЗФ SBR, как фильтры разложения сигнала, $h_k(n)$, так и фильтры синтеза сигнала, $f_k(n)$, могут быть определены посредством:

$$h_k(n) = f_k(n) = p_0(n) \exp\left\{i \frac{\pi}{M} \left(k + \frac{1}{2}\right) \left(n - \frac{N}{2}\right)\right\}, \quad 0 \leq n \leq N; 0 \leq k < M \quad (1)$$

где $p_0(n)$ – вещественнозначный симметричный или несимметричный прототипный фильтр (типично, прототипный фильтр нижних частот), M обозначает количество каналов, а N – порядок прототипного фильтра. Количество каналов, используемых в гребенке фильтров разложения сигнала, может быть иным, чем количество каналов, используемых в гребенке фильтров синтеза сигнала. Например, гребенка фильтров разложения сигнала может иметь 32 канала, а гребенка фильтров синтеза сигнала может иметь 64 канала. При эксплуатации гребенки фильтров синтеза сигналов в режиме с понижающей дискретизацией, гребенка фильтров синтеза сигнала может иметь только 32 канала. Поскольку отсчеты поддиапазонов из гребенки фильтров являются комплекснозначными, ступень аддитивного возможно каналозависимого фазового сдвига может быть добавлена к гребенке фильтров разложения сигнала. Эти дополнительные фазовые сдвиги необходимо компенсировать до гребенки фильтров синтеза сигнала. Несмотря на то, что элементы

фазового сдвига в принципе могут быть произвольными значениями, не нарушающими работу цепочки разложения/синтеза сигнала с помощью КЗФ, они также могут быть ограничены определенными значениями для контроля соответствия. Сигнал SBR будет находиться под влиянием выбора фазовых множителей, в то время как сигнал нижних частот, приходящий из основного декодера, не будет. Качество звука выходного сигнала не нарушается.

Коэффициенты прототипного фильтра, $p_0(n)$, могут быть ограничены длиной, L в 640, как показано в Таблице 4, приведенной ниже.

Таблица 4

n	$p_0(n)$	n	$p_0(n)$	n	$p_0(n)$
0	0,0000000000	214	0,0019765601	428	0,0117623832
1	-0,0005525286	215	-0,0032086896	429	0,0163701258
2	-0,0005617692	216	-0,0085711749	430	0,0207997072
3	-0,0004947518	217	-0,0141288827	431	0,0250307561
4	-0,0004875227	218	-0,0198834129	432	0,0290824006
5	-0,0004893791	219	-0,0258227288	433	0,0329583930
6	-0,0005040714	220	-0,0319531274	434	0,0366418116
7	-0,0005226564	221	-0,0382776572	435	0,0401458278
8	-0,0005466565	222	-0,0447806821	436	0,0434768782
9	-0,0005677802	223	-0,0514804176	437	0,0466303305
10	-0,0005870930	224	-0,0583705326	438	0,0495978676
11	-0,0006132747	225	-0,0654409853	439	0,0524093821
12	-0,0006312493	226	-0,0726943300	440	0,0550460034
13	-0,0006540333	227	-0,0801372934	441	0,0575152691
14	-0,0006777690	228	-0,0877547536	442	0,0598166570
15	-0,0006941614	229	-0,0955533352	443	0,0619602779
16	-0,0007157736	230	-0,1035329531	444	0,0639444805
17	-0,0007255043	231	-0,1116826931	445	0,0657690668
18	-0,0007440941	232	-0,1200077984	446	0,0674525021
19	-0,0007490598	233	-0,1285002850	447	0,0689664013
20	-0,0007681371	234	-0,1371551761	448	0,0703533073
21	-0,0007724848	235	-0,1459766491	449	0,0715826364
22	-0,0007834332	236	-0,1549607071	450	0,0726774642

23	-0,0007779869	237	-0,1640958855	451	0,0736406005
24	-0,0007803664	238	-0,1733808172	452	0,0744664394
25	-0,0007801449	239	-0,1828172548	453	0,0751576255
26	-0,0007757977	240	-0,1923966745	454	0,0757305756
27	-0,0007630793	241	-0,2021250176	455	0,0761748321
28	-0,0007530001	242	-0,2119735853	456	0,0765050718
29	-0,0007319357	243	-0,2219652696	457	0,0767204924
30	-0,0007215391	244	-0,2320690870	458	0,0768230011
31	-0,0006917937	245	-0,2423016884	459	0,0768173975
32	-0,0006650415	246	-0,2526480309	460	0,0767093490
33	-0,0006341594	247	-0,2631053299	461	0,0764992170
34	-0,0005946118	248	-0,2736634040	462	0,0761992479
35	-0,0005564576	249	-0,2843214189	463	0,0758008358
36	-0,0005145572	250	-0,2950716717	464	0,0753137336
37	-0,0004606325	251	-0,3059098575	465	0,0747452558
38	-0,0004095121	252	-0,3168278913	466	0,0741003642
39	-0,0003501175	253	-0,3278113727	467	0,0733620255
40	-0,0002896981	254	-0,3388722693	468	0,0725682583
41	-0,0002098337	255	-0,3499914122	469	0,0717002673
42	-0,0001446380	256	0,3611589903	470	0,0707628710
43	-0,0000617334	257	0,3723795546	471	0,0697630244
44	0,0000134949	258	0,3836350013	472	0,0687043828
45	0,0001094383	259	0,3949211761	473	0,0676075985
46	0,0002043017	260	0,4062317676	474	0,0664367512
47	0,0002949531	261	0,4175696896	475	0,0652247106
48	0,0004026540	262	0,4289119920	476	0,0639715898
49	0,0005107388	263	0,4402553754	477	0,0626857808
50	0,0006239376	264	0,4515996535	478	0,0613455171
51	0,0007458025	265	0,4629308085	479	0,0599837480
52	0,0008608443	266	0,4742453214	480	0,0585915683
53	0,0009885988	267	0,4855253091	481	0,0571616450
54	0,0011250155	268	0,4967708254	482	0,0557173648
55	0,0012577884	269	0,5079817500	483	0,0542452768
56	0,0013902494	270	0,5191234970	484	0,0527630746

57	0,0015443219	271	0,5302240895	485	0,0512556155
58	0,0016868083	272	0,5412553448	486	0,0497385755
59	0,0018348265	273	0,5522051258	487	0,0482165720
60	0,0019841140	274	0,5630789140	488	0,0466843027
61	0,0021461583	275	0,5738524131	489	0,0451488405
62	0,0023017254	276	0,5845403235	490	0,0436097542
63	0,0024625616	277	0,5951123086	491	0,0420649094
64	0,0026201758	278	0,6055783538	492	0,0405349170
65	0,0027870464	279	0,6159109932	493	0,0390053679
66	0,0029469447	280	0,6261242695	494	0,0374812850
67	0,0031125420	281	0,6361980107	495	0,0359697560
68	0,0032739613	282	0,6461269695	496	0,0344620948
69	0,0034418874	283	0,6559016302	497	0,0329754081
70	0,0036008268	284	0,6655139880	498	0,0315017608
71	0,0037603922	285	0,6749663190	499	0,0300502657
72	0,0039207432	286	0,6842353293	500	0,0286072173
73	0,0040819753	287	0,6933282376	501	0,0271859429
74	0,0042264269	288	0,7022388719	502	0,0257875847
75	0,0043730719	289	0,7109410426	503	0,0244160992
76	0,0045209852	290	0,7194462634	504	0,0230680169
77	0,0046606460	291	0,7277448900	505	0,0217467550
78	0,0047932560	292	0,7358211758	506	0,0204531793
79	0,0049137603	293	0,7436827863	507	0,0191872431
80	0,0050393022	294	0,7513137456	508	0,0179433381
81	0,0051407353	295	0,7587080760	509	0,0167324712
82	0,0052461166	296	0,7658674865	510	0,0155405553
83	0,0053471681	297	0,7727780881	511	0,0143904666
84	0,0054196775	298	0,7794287519	512	-0,0132718220
85	0,0054876040	299	0,7858353120	513	-0,0121849995
86	0,0055475714	300	0,7919735841	514	-0,0111315548
87	0,0055938023	301	0,7978466413	515	-0,0101150215
88	0,0056220643	302	0,8034485751	516	-0,0091325329
89	0,0056455196	303	0,8087695004	517	-0,0081798233
90	0,0056389199	304	0,8138191270	518	-0,0072615816

91	0,0056266114	305	0,8185776004	519	-0,0063792293
92	0,0055917128	306	0,8230419890	520	-0,0055337211
93	0,0055404363	307	0,8272275347	521	-0,0047222596
94	0,0054753783	308	0,8311038457	522	-0,0039401124
95	0,0053838975	309	0,8346937361	523	-0,0031933778
96	0,0052715758	310	0,8379717337	524	-0,0024826723
97	0,0051382275	311	0,8409541392	525	-0,0018039472
98	0,0049839687	312	0,8436238281	526	-0,0011568135
99	0,0048109469	313	0,8459818469	527	-0,0005464280
100	0,0046039530	314	0,8480315777	528	0,0000276045
101	0,0043801861	315	0,8497805198	529	0,0005832264
102	0,0041251642	316	0,8511971524	530	0,0010902329
103	0,0038456408	317	0,8523047035	531	0,0015784682
104	0,0035401246	318	0,8531020949	532	0,0020274176
105	0,0032091885	319	0,8535720573	533	0,0024508540
106	0,0028446757	320	0,8537385600	534	0,0028446757
107	0,0024508540	321	0,8535720573	535	0,0032091885
108	0,0020274176	322	0,8531020949	536	0,0035401246
109	0,0015784682	323	0,8523047035	537	0,0038456408
110	0,0010902329	324	0,8511971524	538	0,0041251642
111	0,0005832264	325	0,8497805198	539	0,0043801861
112	0,0000276045	326	0,8480315777	540	0,0046039530
113	-0,0005464280	327	0,8459818469	541	0,0048109469
114	-0,0011568135	328	0,8436238281	542	0,0049839687
115	-0,0018039472	329	0,8409541392	543	0,0051382275
116	-0,0024826723	330	0,8379717337	544	0,0052715758
117	-0,0031933778	331	0,8346937361	545	0,0053838975
118	-0,0039401124	332	0,8311038457	546	0,0054753783
119	-0,0047222596	333	0,8272275347	547	0,0055404363
120	-0,0055337211	334	0,8230419890	548	0,0055917128
121	-0,0063792293	335	0,8185776004	549	0,0056266114
122	-0,0072615816	336	0,8138191270	550	0,0056389199
123	-0,0081798233	337	0,8087695004	551	0,0056455196
124	-0,0091325329	338	0,8034485751	552	0,0056220643

125	-0,0101150215	339	0,7978466413	553	0,0055938023
126	-0,0111315548	340	0,7919735841	554	0,0055475714
127	-0,0121849995	341	0,7858353120	555	0,0054876040
128	0,0132718220	342	0,7794287519	556	0,0054196775
129	0,0143904666	343	0,7727780881	557	0,0053471681
130	0,0155405553	344	0,7658674865	558	0,0052461166
131	0,0167324712	345	0,7587080760	559	0,0051407353
132	0,0179433381	346	0,7513137456	560	0,0050393022
133	0,0191872431	347	0,7436827863	561	0,0049137603
134	0,0204531793	348	0,7358211758	562	0,0047932560
135	0,0217467550	349	0,7277448900	563	0,0046606460
136	0,0230680169	350	0,7194462634	564	0,0045209852
137	0,0244160992	351	0,7109410426	565	0,0043730719
138	0,0257875847	352	0,7022388719	566	0,0042264269
139	0,0271859429	353	0,6933282376	567	0,0040819753
140	0,0286072173	354	0,6842353293	568	0,0039207432
141	0,0300502657	355	0,6749663190	569	0,0037603922
142	0,0315017608	356	0,6655139880	570	0,0036008268
143	0,0329754081	357	0,6559016302	571	0,0034418874
144	0,0344620948	358	0,6461269695	572	0,0032739613
145	0,0359697560	359	0,6361980107	573	0,0031125420
146	0,0374812850	360	0,6261242695	574	0,0029469447
147	0,0390053679	361	0,6159109932	575	0,0027870464
148	0,0405349170	362	0,6055783538	576	0,0026201758
149	0,0420649094	363	0,5951123086	577	0,0024625616
150	0,0436097542	364	0,5845403235	578	0,0023017254
151	0,0451488405	365	0,5738524131	579	0,0021461583
152	0,0466843027	366	0,5630789140	580	0,0019841140
153	0,0482165720	367	0,5522051258	581	0,0018348265
154	0,0497385755	368	0,5412553448	582	0,0016868083
155	0,0512556155	369	0,5302240895	583	0,0015443219
156	0,0527630746	370	0,5191234970	584	0,0013902494
157	0,0542452768	371	0,5079817500	585	0,0012577884
158	0,0557173648	372	0,4967708254	586	0,0011250155

159	0,0571616450	373	0,4855253091	587	0,0009885988
160	0,0585915683	374	0,4742453214	588	0,0008608443
161	0,0599837480	375	0,4629308085	589	0,0007458025
162	0,0613455171	376	0,4515996535	590	0,0006239376
163	0,0626857808	377	0,4402553754	591	0,0005107388
164	0,0639715898	378	0,4289119920	592	0,0004026540
165	0,0652247106	379	0,4175696896	593	0,0002949531
166	0,0664367512	380	0,4062317676	594	0,0002043017
167	0,0676075985	381	0,3949211761	595	0,0001094383
168	0,0687043828	382	0,3836350013	596	0,0000134949
169	0,0697630244	383	0,3723795546	597	-0,0000617334
170	0,0707628710	384	-0,3611589903	598	-0,0001446380
171	0,0717002673	385	-0,3499914122	599	-0,0002098337
172	0,0725682583	386	-0,3388722693	600	-0,0002896981
173	0,0733620255	387	-0,3278113727	601	-0,0003501175
174	0,0741003642	388	-0,3168278913	602	-0,0004095121
175	0,0747452558	389	-0,3059098575	603	-0,0004606325
176	0,0753137336	390	-0,2950716717	604	-0,0005145572
177	0,0758008358	391	-0,2843214189	605	-0,0005564576
178	0,0761992479	392	-0,2736634040	606	-0,0005946118
179	0,0764992170	393	-0,2631053299	607	-0,0006341594
180	0,0767093490	394	-0,2526480309	608	-0,0006650415
181	0,0768173975	395	-0,2423016884	609	-0,0006917937
182	0,0768230011	396	-0,2320690870	610	-0,0007215391
183	0,0767204924	397	-0,2219652696	611	-0,0007319357
184	0,0765050718	398	-0,2119735853	612	-0,0007530001
185	0,0761748321	399	-0,2021250176	613	-0,0007630793
186	0,0757305756	400	-0,1923966745	614	-0,0007757977
187	0,0751576255	401	-0,1828172548	615	-0,0007801449
188	0,0744664394	402	-0,1733808172	616	-0,0007803664
189	0,0736406005	403	-0,1640958855	617	-0,0007779869
190	0,0726774642	404	-0,1549607071	618	-0,0007834332
191	0,0715826364	405	-0,1459766491	619	-0,0007724848
192	0,0703533073	406	-0,1371551761	620	-0,0007681371

193	0,0689664013	407	-0,1285002850	621	-0,0007490598
194	0,0674525021	408	-0,1200077984	622	-0,0007440941
195	0,0657690668	409	-0,1116826931	623	-0,0007255043
196	0,0639444805	410	-0,1035329531	624	-0,0007157736
197	0,0619602779	411	-0,0955533352	625	-0,0006941614
198	0,0598166570	412	-0,0877547536	626	-0,0006777690
199	0,0575152691	413	-0,0801372934	627	-0,0006540333
200	0,0550460034	414	-0,0726943300	628	-0,0006312493
201	0,0524093821	415	-0,0654409853	629	-0,0006132747
202	0,0495978676	416	-0,0583705326	630	-0,0005870930
203	0,0466303305	417	-0,0514804176	631	-0,0005677802
204	0,0434768782	418	-0,0447806821	632	-0,0005466565
205	0,0401458278	419	-0,0382776572	633	-0,0005226564
206	0,0366418116	420	-0,0319531274	634	-0,0005040714
207	0,0329583930	421	-0,0258227288	635	-0,0004893791
208	0,0290824006	422	-0,0198834129	636	-0,0004875227
209	0,0250307561	423	-0,0141288827	637	-0,0004947518
210	0,0207997072	424	-0,0085711749	638	-0,0005617692
211	0,0163701258	425	-0,0032086896	639	-0,0005525280
212	0,0117623832	426	0,0019765601		
213	0,0069636862	427	0,0069636862		

Прототипный фильтр, $p_0(n)$, также может быть выведен из таблицы 4 посредством одной или более математических операций, таких как округление, понижающая дискретизация, интерполяция или прореживание.

Понятно, что, в пределах объема прилагаемой формулы изобретения, изобретение может быть осуществлено на практике иным образом, чем конкретно описанный в материалах настоящей заявки. Любые номера ссылок, содержащиеся в нижеследующей формуле изобретения, предназначены только для иллюстративных целей и никоим образом не должны использоваться для толкования или ограничения формулы изобретения.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ для декодирования кодированного звукового битового потока, способ содержит этапы, на которых:

принимают кодированный звуковой битовый поток, кодированный звуковой битовый поток включает в себя звуковые данные, представляющие собой часть полосы низких частот звукового сигнала;

декодируют звуковые данные для формирования декодированного звукового сигнала в полосе низких частот;

извлекают из кодированного звукового битового потока метаданные реконструкции высоких частот, метаданные реконструкции высоких частот включают в себя рабочие параметры для процесса реконструкции высоких частот, который преобразует некоторое количество поддиапазонов из части полосы низких частот звукового сигнала в часть полосы высоких частот звукового сигнала;

извлекают из кодированного звукового битового потока параметр, указывающий использовать или нет адаптивную по отношению к сигналу передискретизации частотную область;

фильтруют декодированный звуковой сигнал в полосе низких частот гребенкой фильтров разложения сигнала для формирования фильтрованного звукового сигнала в полосе низких частот;

извлекают из кодированного звукового битового потока флажковый указатель, указывающий, должны ли выполняться над звуковыми данными линейное преобразование или гармоническая транспозиция;

восстанавливают часть полосы высоких частот звукового сигнала с использованием фильтрованного звукового сигнала в полосе низких частот и метаданных реконструкции высоких частот в соответствии с флажковым указателем; и

комбинируют фильтрованный звуковой сигнал в полосе низких частот и восстановленную часть полосы высоких частот для формирования широкополосного звукового сигнала, используя гребенку фильтров синтеза сигнала,

при этом, гребенка фильтров разложения сигнала включает в себя фильтры разложения сигнала, $h_k(n)$, которые являются модулированными вариантами прототипного фильтра, $p_0(n)$, согласно:

$$h_k(n) = p_0(n) \exp\left\{i\frac{\pi}{M}\left(k + \frac{1}{2}\right)\left(n - \frac{N}{2}\right)\right\}, \quad 0 \leq n \leq N; 0 \leq k < M$$

где $p_0(n)$ – вещественнозначный симметричный или несимметричный прототипный фильтр, M – количество каналов в гребенке фильтров разложения сигнала, а N – порядок прототипного фильтра, и

при этом, количество каналов в гребенке фильтров разложения сигнала отличается от количества каналов в гребенке фильтров синтеза сигнала.

2. Способ по п. 1, в котором кодированный звуковой битовый поток также включает в себя элемент заполнения с идентификатором, указывающим начало элемента заполнения, и данными заполнения после идентификатора, при этом, данные заполнения включают в себя флажковый указатель.

3. Способ по п. 2, в котором идентификатор является трехбитным беззнаковым целым числом, передаваемым старшим битом вперед и имеющим значение 0×6 .

4. Способ по п. 2, в котором данные заполнения включают в себя полезную нагрузку расширения, полезная нагрузка расширения включает в себя данные расширения репликации полос спектра, и полезная нагрузка расширения идентифицируется четырехбитным беззнаковым целым числом, передаваемым старшим битом вперед и имеющим значение '1101' или '1110', и, по выбору,

при этом, данные расширения репликации полос спектра включают в себя:

необязательный заголовок репликации полос спектра,

данные репликации полос спектра после заголовка,

элемент расширения репликации полос спектра после данных репликации полос спектра, и при этом, флажковый указатель включен в элемент расширения репликации полос спектра.

5. Способ по любому одному из п. п. 1-4, в котором метаданные реконструкции высоких частот включают в себя рабочий

параметр, выбранный из группы, состоящей из масштабных коэффициентов огибающей, масштабных коэффициентов уровня собственных шумов, информации о синусоидальном дополнении, информации о временной/частотной сетке, частоту разделения и режима обратной фильтрации.

6. Способ по любому одному из п. п. 1-4, в котором прототипный фильтр, $p_0(n)$, выводится из коэффициентов по Таблице 4.

7. Способ по любому одному из п. п. 1-4, в котором прототипный фильтр, $p_0(n)$, выводится из коэффициентов по Таблице 4 посредством одной или более математических операций, выбранных из группы, состоящей из округления, понижающей дискретизации, интерполяции или прореживания.

8. Способ по любому одному из п. п. 1-4, в котором фазовый сдвиг добавляется в фильтрованный звуковой сигнал в полосе низких частот после фильтрации и компенсируется перед комбинированием.

9. Долговременный машинно-читаемый носитель, содержащий команды, которые, когда исполняются процессором, побуждают его выполнять способ по п. 1.

10. Декодер для декодирования кодированного звукового битового потока, декодер содержит:

входной интерфейс для приема кодированного звукового битового потока, кодированный звуковой битовый поток включает в себя звуковые данные, представляющие собой часть полосы низких частот звукового сигнала;

основной декодер для декодирования звуковых данных для формирования декодированного звукового сигнала в полосе низких частот;

синтаксический анализатор для извлечения из кодированного звукового битового потока метаданных реконструкции высоких частот и параметр, указывающий использовать или нет адаптивную по отношению к сигналу передискретизации частотную область, метаданные реконструкции высоких частот включают в себя рабочие параметры для процесса реконструкции высоких частот, который преобразует некоторое количество поддиапазонов из части полосы

низких частот звукового сигнала в часть полосы высоких частот звукового сигнала;

гребенку фильтров разложения сигнала для фильтрации декодированного звукового сигнала в полосе низких частот, чтобы формировать фильтрованный звуковой сигнал в полосе низких частот;

синтаксический анализатор для извлечения из кодированного звукового битового потока флажкового указателя, указывающего, должны ли выполняться над звуковыми данными линейное преобразование или гармоническая транспозиция;

восстановитель высоких частот для восстановления части полосы высоких частот звукового сигнала с использованием фильтрованного звукового сигнала в полосе низких частот и метаданных реконструкции высоких частот в соответствии с флажковым указателем; и

гребенку фильтров синтеза сигнала для комбинирования фильтрованного звукового сигнала в полосе низких частот и восстановленной части полосы высоких частот для формирования широкополосного звукового сигнала,

при этом, гребенка фильтров разложения сигнала включает в себя фильтры разложения сигнала, $h_k(n)$, которые являются модулированными вариантами прототипного фильтра, $p_0(n)$, согласно:

$$h_k(n) = p_0(n) \exp\left\{i \frac{\pi}{M} \left(k + \frac{1}{2}\right) \left(n - \frac{N}{2}\right)\right\}, \quad 0 \leq n \leq N; 0 \leq k < M$$

где $p_0(n)$ – вещественнозначный симметричный или несимметричный прототипный фильтр, M – количество каналов в гребенке фильтров разложения сигнала, а N – порядок прототипного фильтра, и

при этом количество каналов в гребенке фильтров разложения сигнала отличается от количества каналов в гребенке фильтров синтеза сигнала.

11. Декодер по п. 10, в котором кодированный звуковой битовый поток также включает в себя элемент заполнения с идентификатором, указывающим начало элемента заполнения, и

данными заполнения после идентификатора, при этом, данные заполнения включают в себя флажковый указатель.

12. Декодер по п. 11, в котором идентификатор является трехбитным беззнаковым целым числом, передаваемым старшим битом вперед и имеющим значение 0×6 .

13. Декодер по п. 11, в котором данные заполнения включают в себя полезную нагрузку расширения, полезная нагрузка расширения включает в себя данные расширения репликации полос спектра, и полезная нагрузка расширения идентифицируется четырехбитным беззнаковым целым числом, передаваемым старшим битом вперед и имеющим значение '1101' или '1110', и, по выбору,

при этом, данные расширения репликации полос спектра включают в себя:

необязательный заголовок репликации полос спектра,
данные репликации полос спектра после заголовка,
элемент расширения репликации полос спектра после данных репликации полос спектра, и при этом, флажковый указатель включен в элемент расширения репликации полос спектра.

14. Декодер по любому одному из п. п. 10-13, в котором метаданные реконструкции высоких частот включают в себя рабочий параметр, выбранный из группы, состоящей из масштабных коэффициентов огибающей, масштабных коэффициентов уровня собственных шумов, информации о синусоидальном дополнении, информации о временной/частотной сетке, частоту деления и режима обратной фильтрации.

15. Декодер по любому одному из п. п. 10-13, в котором прототипный фильтр, $p_0(n)$, выводится из коэффициентов по Таблице 4.

16. Декодер по любому одному из п. п. 10-13, в котором прототипный фильтр, $p_0(n)$, выводится из коэффициентов по Таблице 4 посредством одной или более математических операций, выбранных из группы, состоящей из округления, понижающей дискретизации, интерполяции или прореживания.

17. Декодер по любому одному из п. п. 10-13, в котором фазовый сдвиг добавляется в фильтрованный звуковой сигнал в

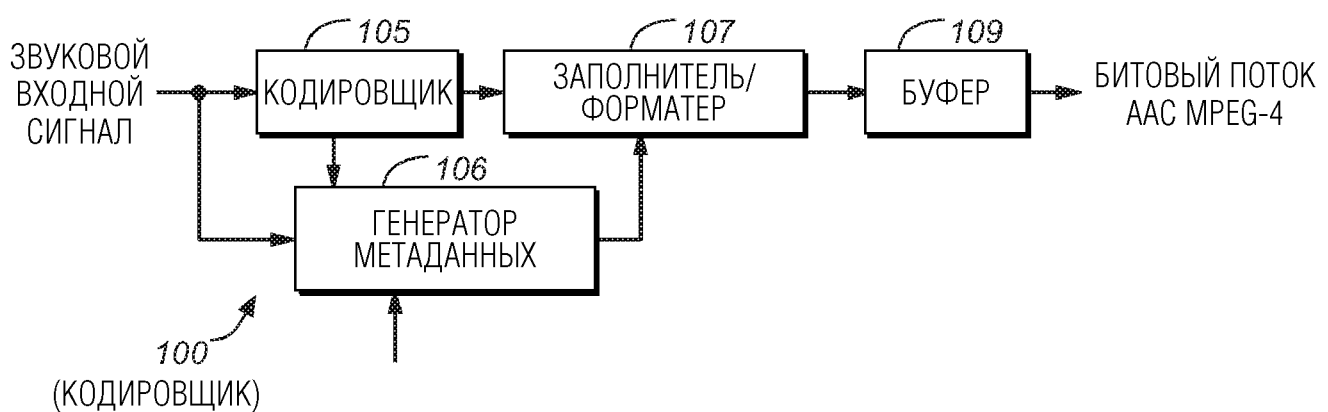
полосе низких частот после фильтрации и компенсируется перед комбинированием.

По доверенности

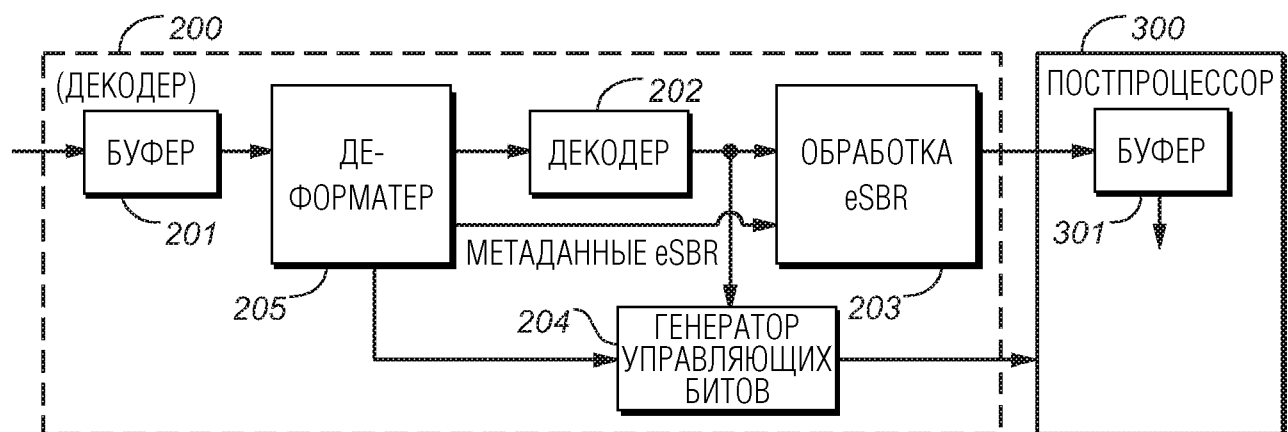
ФИГ.1



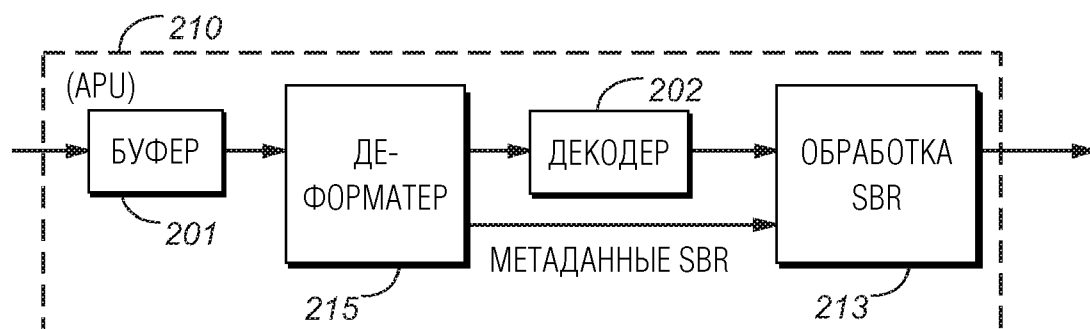
ФИГ.2

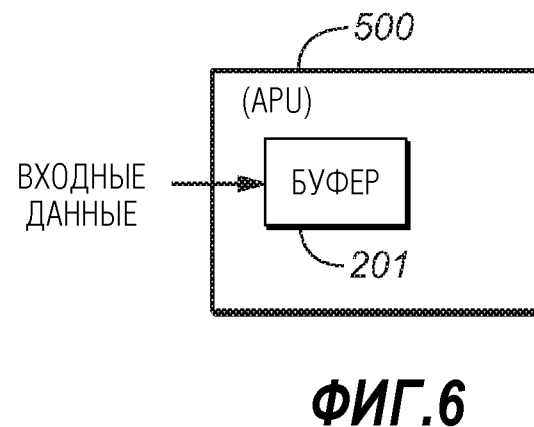
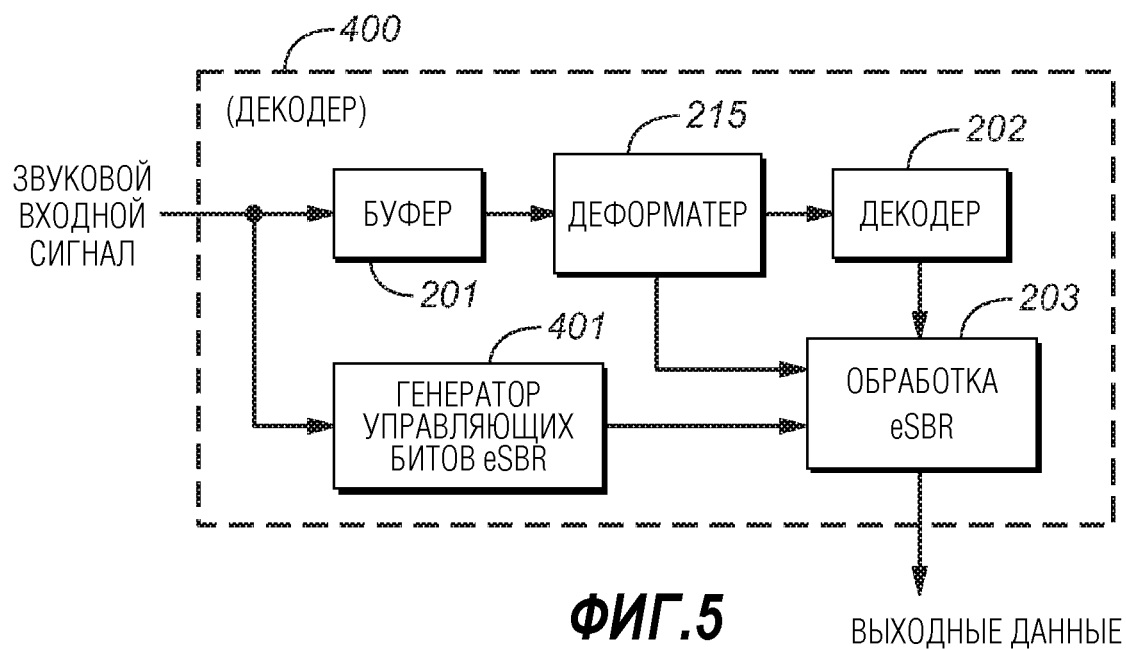


ФИГ.3

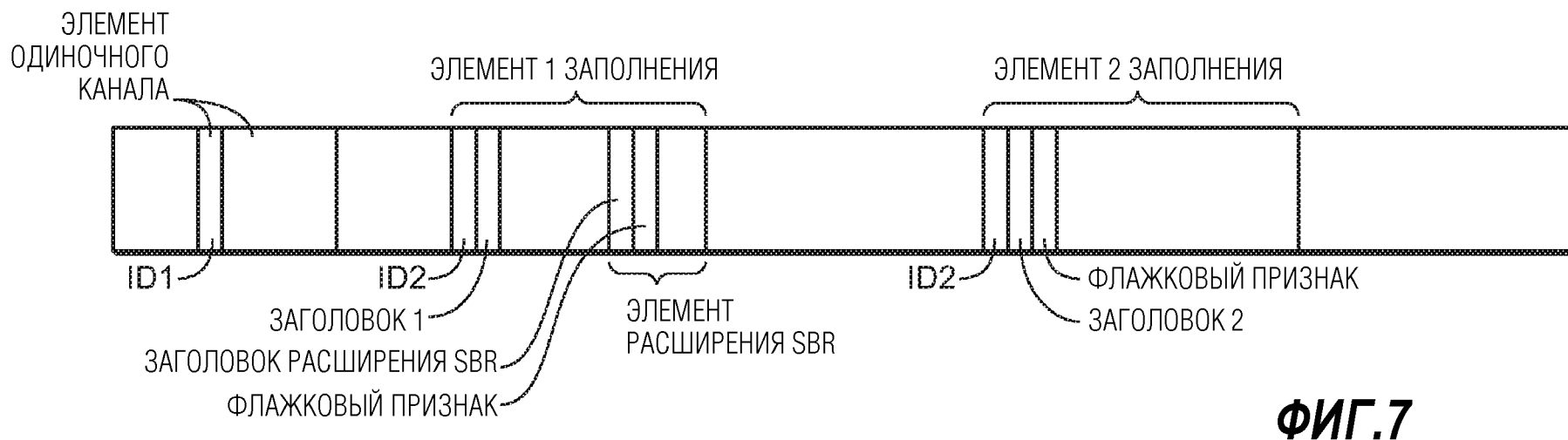


ФИГ.4





2/2



PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference D17023WO01	FOR FURTHER ACTION	see Form PCT/ISA/220 as well as, where applicable, item 5 below.
International application No. PCT/US18/23183	International filing date (<i>day/month/year</i>) 19 March 2018 (19.03.2018)	(Earliest) Priority Date (<i>day/month/year</i>) 23 March 2017 (23.03.2017)
Applicant DOLBY INTERNATIONAL AB		

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of 3 sheets.

☐ It is also accompanied by a copy of each prior art document cited in this report.

1. Basis of the report

a. With regard to the **language**, the international search was carried out on the basis of:

- ☒ the international application in the language in which it was filed.
☐ a translation of the international application into _____ which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).

b. ☐ This international search report has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43.6bis(a)).

c. ☐ With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, see Box No. I.

2. ☒ **Certain claims were found unsearchable** (see Box No. II).

3. ☐ **Unity of invention is lacking** (see Box No. III).

4. With regard to the **title**,

- ☒ the text is approved as submitted by the applicant.
☐ the text has been established by this Authority to read as follows:

5. With regard to the **abstract**,

- ☒ the text is approved as submitted by the applicant.
☐ the text has been established, according to Rule 38.2, by this Authority as it appears in Box No. IV. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. With regard to the **drawings**,

- a. the figure of the drawings to be published with the abstract is Figure No. 5
☒ as suggested by the applicant.
☐ as selected by this Authority, because the applicant failed to suggest a figure.
☐ as selected by this Authority, because this figure better characterizes the invention.
- b. ☐ none of the figures is to be published with the abstract.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US18/23183

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 9 & 19
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US18/23183

A. CLASSIFICATION OF SUBJECT MATTER

IPC - G06F 17/10, G10L 19/02, 19/22, 19/24, 19/26 (2018.01)

CPC - G06F 17/10, G10L 19/02, 19/22, 19/24, 19/26, 19/265

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

See Search History document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y --- A	US 2015/0317986 A1 (DOLBY INTERNATIONAL AB) November 5, 2015; paragraphs [0006], [0053], [0068], [0055] [0068]-[0071], [0081], & [0082] claims 19, 20	1, 5/1, & 10 --- 2-4, 5/2-5/4, 11-15 --- 6-8
Y --- A	WO 2016/146492 A1 (DOLBY INTERNATIONAL AB) September 22, 2016; page 1, lines 26-28, page 2, lines 1-10 page 9, lines 30-33, page 12, lines 9-31, page 20, lines 1-20, page 25, lines 7-19 claims 1, 16, 17, & 20-22, FIG. 3	2-4, 5/2-5/44, 11-14, & 15/11-15/14 --- 16-18
A	US 2011/0302230 A1 (EKSTRAND, P) December 8, 2011; paragraph [0069], claims 9 & 11	1-8 & 10-18

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

22 May 2018 (22.05.2018)

Date of mailing of the international search report

11 JUN 2018

Name and mailing address of the ISA/

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents

P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Shane Thomas

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774