(19) Евразийское патентное ведомство

(11) 035064 (13) B1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ

(45) Дата публикации и выдачи патента
2020.04.23

(21) Номер заявки
201890845

(22) Дата подачи заявки
2016.10.07

(51) Int. Cl. *G10L 19/008* (2013.01)

(54) МНОГОУРОВНЕВОЕ КОДИРОВАНИЕ И СТРУКТУРА ДАННЫХ ДЛЯ СЖАТЫХ ПРЕДСТАВЛЕНИЙ ЗВУКА ИЛИ ЗВУКОВОГО ПОЛЯ СИСТЕМЫ АМБИСОНИК ВЫСШЕГО ПОРЯДКА

(31) **15306591.7; 62/361,863**
(32) **2015.10.08; 2016.07.13**
(33) **EP; US**
(43) **2018.10.31**
(86) **PCT/EP2016/073971**
(87) **WO 2017/060412 2017.04.13**

(71)(73) Заявитель и патентовладелец:
ДОЛБИ ИНТЕРНЭШНЛ АБ (NL)

(72) Изобретатель:
Кордон Свен, Крюгер Александр (DE)

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

(56) EP-A1-2922057
Anonymous: "ISO/IEC JTC 1/SC 29 N ISO/IEC 23008-3:2015/PDAM 3 Information technology - High efficiency coding and media delivery in heterogeneous environments - Part 3: Part 3: 3D audio, AMENDMENT 3: MPEG-H 3D Audio Phase 2", 25 July 2015 (2015-07-25), pages 1-202, XP055329832, Retrieved from the Internet: URL:http://mpeg.chiariglione.org/standards/mpeg-h/3d-audio/text-isoiec-23008-3201xpd am-3-mpeg-h-3d-audio-phase-2 [retrieved on 2016-12-16], page 4, line 8 - page 4, line 18 Codebkldx and NumVvecIndices as disclosed at the end of Page 18 related to clause 12.3.2.2

(57) Настоящее изобретение относится к способу и устройству декодирования сжатого представления звука или звукового поля системы Амбисоник высшего порядка (HOA). Сжатое представление HOA содержит множество транспортных сигналов. Способ содержит этапы, на которых принимают битовый поток, содержащий сжатое представление HOA, соответствующее множеству иерархических уровней, которые включают в себя базовый уровень и два или более иерархических улучшающих уровней, причем битовый поток содержит для каждого уровня соответствующую полезную нагрузку расширения HOA, включающую в себя вспомогательную информацию для параметрического расширения воссозданного представления HOA, получаемого из транспортных сигналов, присвоенных соответствующему уровню и любым уровням ниже соответствующего уровня; определяют наиболее высокий применимый уровень среди множества уровней для декодирования, причем наиболее высокий применимый уровень представляет собой уровень ниже наиболее низкого уровня, который не был принят корректно; извлекают полезную нагрузку расширения HOA, присвоенную наиболее высокому применимому уровню, причем полезная нагрузка расширения HOA включает в себя вспомогательную информацию для параметрического улучшения воссозданного представления HOA, соответствующего наиболее высокому применимому уровню, причем воссозданное представление HOA, соответствующее наиболее высокому применимому уровню, может быть получено на основе транспортных сигналов, присвоенных наиболее высокому применимому уровню и любым уровням ниже наиболее высокого применимого уровня; декодируют сжатое представление HOA, соответствующее наиболее высокому применимому уровню, на основе информации уровня, транспортных сигналов, присвоенных наиболее высокому применимому уровню и любым уровням ниже наиболее высокого применимого уровня; и выполняют параметрическое улучшение декодированного представления HOA с использованием вспомогательной информации, включенной в полезную нагрузку расширения HOA, присвоенную наиболее высокому применимому уровню. Технический результат заключается, по меньшей мере, в повышении качества звука при низких битрейтах.

035064 B1

035064 B1

Перекрестная ссылка на родственные заявки

По настоящей заявке испрашивается приоритет по европейской патентной заявке № 15306653.5, поданной 15 октября 2015 года, содержание которой полностью включено в настоящую заявку посредством ссылки.

Область техники, к которой относится изобретение

Настоящий документ относится к способам и устройству для многоуровневого аудиокодирования. В частности, настоящий документ относится к способам и устройству для многоуровневого аудиокодирования кадров сжатых представлений звука (или звукового поля) системы Амбисоник высшего порядка (Higher-Order Ambisonics, HOA). Настоящий документ также относится к структурам данных (например, к битовым потокам) для представления кадров сжатых представлений HOA звука (или звукового поля).

Уровень техники

В текущем определении многоуровневого кодирования HOA вспомогательная информация для инструментов декодирования HOA пространственного предсказания сигналов, синтеза направленных подполосных сигналов и декодера параметрического дублирования звукового окружения (PAR) создана для улучшения специфического представления HOA. А именно в текущем определении многоуровневого кодирования HOA обеспеченные данные правильным образом расширяют представление HOA только наиболее высокого уровня (например, наиболее высокого улучшающего уровня). Для более низких уровней, в том числе для базового уровня, эти инструменты не улучшают частично воссозданное представление HOA правильным образом.

Инструменты синтеза направленных подполосных сигналов и декодера параметрического дублирования звукового окружения специально предназначены для низких скоростей передачи данных, когда доступны только несколько транспортных сигналов. Однако в многоуровневом кодировании HOA надлежащее улучшение (частично) воссозданных представлений HOA не является возможным особенно для уровней с низким битрейтом, например для базового уровня. Это явно нежелательно с точки обзора качества звука при низких битрейтах.

Кроме того, было обнаружено, что традиционный способ обработки закодированных V-векторных элементов для основанных на векторах сигналов не приводит к подходящему декодированию, если в HOADecoderConfig() сигнализируется CodedVVecLength == 1 (т.е. если активен векторный режим кодирования). В этом векторном режиме кодирования V-векторные элементы не передаются для индексов коэффициентов HOA, которые включены во множество ContAddHoaCoeff. Это множество включает в себя все индексы коэффициентов HOA AmbCoeffIdx[i], которые имеют AmbCoeffTransitionState == 0. Традиционно также нет необходимости добавлять взвешенный V-векторный сигнал, поскольку первоначальная последовательность коэффициентов HOA для этих индексов явно отправляется (сигнализируется). Таким образом, V-векторный элемент в традиционном подходе устанавливается равным нулю для этих индексов.

Однако в многоуровневом режиме кодирования множество непрерывных индексов коэффициентов HOA зависит от транспортных каналов, которые являются частью текущего активного уровня. Дополнительные индексы коэффициентов HOA, которые отправлены на более высоком уровне, могут отсутствовать на более низких уровнях. Тогда предположение, что векторный сигнал не должен вносить вклад в последовательности коэффициентов HOA, является неверным для индексов коэффициентов HOA, которые принадлежат последовательностям коэффициентов HOA, включенным в более высокие уровни.

Как следствие, V-вектор в многоуровневом кодировании HOA может не являться подходящим для декодирования каких-либо уровней ниже наиболее высокого уровня.

Таким образом, существует необходимость в схемах кодирования и битовых потоках, которые адаптированы к многоуровневому кодированию сжатых представлений HOA звука или звукового поля.

Настоящий документ решает упомянутые выше проблемы. В частности, описаны способы и кодеры/декодеры для многоуровневого кодирования кадров сжатых представлений звука или звукового поля HOA, а также структуры данных для представления кадров сжатых представлений звука или звукового поля HOA.

Сущность изобретения

В соответствии с аспектом описан способ многоуровневого кодирования кадра сжатого представления звука (или звукового поля) системы Амбисоник высшего порядка (Higher-Order Ambisonics, HOA). Сжатое представление HOA согласуется с проектом аудио стандарта MPEG-H 3D и любыми другими будущими принятыми стандартами или проектами стандартов. Сжатое представление HOA может включать в себя множество транспортных сигналов. Транспортные сигналы могут относиться к монауральным сигналам, например, представляющим либо преобладающие звуковые сигналы, либо последовательности коэффициентов представления HOA. Способ может включать в себя присвоение множества транспортных сигналов множеству иерархических уровней. Например, транспортные сигналы могут быть распределены множеству уровней. Множество уровней может включать в себя базовый уровень и один или более иерархических улучшающих уровней. Множество иерархических уровней может быть упорядочено от базового уровня через первый улучшающий уровень, второй улучшающий уровень и т.д. вплоть до общего наиболее высокого улучшающего уровня (общего наиболее высокого уровня). Способ

может дополнительно включать в себя формирование для каждого уровня соответствующей полезной нагрузки расширения HOA, включающей в себя вспомогательную информацию (например, улучшающую вспомогательную информацию) для параметрического расширения воссозданного представления HOA, которая может быть получена из транспортных сигналов, присвоенных соответствующему уровню и любым уровням ниже соответствующего уровня. Воссозданные представления HOA для более низких уровней могут упоминаться, как частично воссозданные представления HOA. Способ может дополнительно включать в себя присвоение сформированных полезных нагрузок расширения HOA их соответствующим уровням. Кроме того, способ может включать в себя сигнализацию сформированных полезных нагрузок расширения HOA в выдаваемом битовом потоке. Полезные нагрузки расширения HOA могут быть сигнализированы в полезной нагрузке HOAEnhFrame(). Таким образом, вспомогательная информация может быть перемещена из HOAFrame() в HOAEnhFrame().

Сконфигурированный, как упомянутый выше, предложенный способ применяет многоуровневое кодирование к (кадру) сжатому представлению HOA, чтобы обеспечить его высококачественное декодирование даже при низких битрейтах. В частности, предложенный способ гарантирует, что каждый уровень включает в себя подходящую полезную нагрузку расширения HOA (например, улучшающую вспомогательную информацию) для улучшения (частично) воссозданного представления звука, полученную из транспортных сигналов на любых уровнях вплоть до текущего уровня. При этом предполагается, что уровни вплоть до текущего уровня включают в себя, например, базовый уровень, первый улучшающий уровень, второй улучшающий уровень и т.д., вплоть до текущего уровня. При этом предполагается, что уровни вплоть до текущего уровня включают в себя, например, базовый уровень, первый улучшающий уровень, второй улучшающий уровень и т.д., вплоть до текущего уровня. Например, декодеру обеспечивается возможность улучшить (частично) воссозданное представление звука, полученное из базового уровня, со ссылкой на полезную нагрузку расширения HOA, присвоенную базовому уровню. В традиционном подходе воссозданное представление HOA только наиболее высокого улучшающего уровня может быть улучшено посредством полезной нагрузки расширения HOA. Таким образом, независимо от фактического наиболее высокого применимого уровня (например, уровня ниже наиболее низкого уровня, который не был принят корректно, и, таким образом, все уровни ниже наиболее высокого применимого уровня и сам наиболее высокий применимый уровень были приняты корректно), декодеру обеспечивается возможность улучшить или расширить воссозданное представление звука, даже если (частично) воссозданное представление звука может отличаться от полного представления звука. В частности, независимо от фактического наиболее высокого применимого уровня для декодера достаточно декодировать полезную нагрузку расширения HOA только для одного слоя (т.е. для наиболее высокого применимого уровня), чтобы улучшить или расширить (частично) воссозданное представление звука, которое может быть получено на основе всех транспортных сигналов, включенных в уровни вплоть до фактического наиболее высокого применимого уровня. Декодирование полезных нагрузок расширения HOA более высоких или более низких уровней не требуется. С другой стороны, предложенный способ позволяет полностью использовать преимущества сокращения требуемой ширины полосы, которое может быть достигнуто при применении многоуровневого кодирования.

В вариантах осуществления способ может дополнительно включать в себя передачу полезных нагрузок данных для множества уровней с соответствующими уровнями защиты от ошибок. Полезные нагрузки данных могут включать в себя соответствующие полезные нагрузки расширения HOA. Базовый уровень может иметь наиболее высокую защиту от ошибок, и один или более улучшающих уровней могут иметь последовательно уменьшающуюся защиту от ошибок. Тем самым может быть обеспечено, что, по меньшей мере, некоторое количество более низких уровней передаются достоверно, с другой стороны, сокращая полную требуемую ширину полосы без применения чрезмерной защиты от ошибок к более высоким уровням.

В вариантах осуществления полезные нагрузки расширения HOA могут включать в себя элементы битового потока для инструмента декодирования с пространственным предсказанием сигналов HOA. Дополнительно или в качестве альтернативы полезные нагрузки расширения HOA могут включать в себя элементы битового потока для инструмента декодирования с синтезом направленных подполосных сигналов HOA. Дополнительно или в качестве альтернативы полезные нагрузки расширения HOA могут включать в себя элементы битового потока для инструмента декодирования с параметрическим дублированием звукового окружения HOA.

В вариантах осуществления полезные нагрузки расширения HOA могут иметь тип usacExtElementType ID_EXT_ELE_HOA_ENH_LAYER.

В вариантах осуществления способ может дополнительно включать в себя формирование полезной нагрузки расширения конфигурации HOA, включающей в себя элементы битового потока для конфигурирования инструмента декодирования с пространственным предсказанием сигналов HOA, инструмента декодирования с синтезом направленных подполосных сигналов HOA и/или инструмента декодирования с параметрическим дублированием звукового окружения HOA. Полезная нагрузка расширения конфигурации HOA может быть включена в HOADecoderEnhConfig(). Способ может дополнительно включать в себя сигнализацию полезной нагрузки расширения конфигурации HOA в выдаваемом битовом потоке.

В вариантах осуществления способ может дополнительно включать в себя формирование полезной нагрузки конфигурации декодера HOA, включающей в себя информацию, указывающую присвоение полезных нагрузок расширения HOA множеству уровней. Способ может дополнительно включать в себя сигнализацию полезной нагрузки конфигурации декодера HOA в выдаваемом битовом потоке.

В вариантах осуществления способ может дополнительно включать в себя определение, является ли активным векторный режим кодирования. Если векторный режим кодирования активен, способ может дополнительно включать в себя определение для каждого уровня множества непрерывных индексов коэффициентов HOA на основе транспортных сигналов, присвоенных соответствующему уровню. Индексы коэффициентов HOA в множестве непрерывных индексов коэффициентов HOA могут являться индексами коэффициентов HOA, включенными в множество ContAddHOACoeff. Способ может дополнительно включать в себя формирование для каждого транспортного сигнала V-вектора на основе определенного множества непрерывных индексов коэффициентов HOA для уровня, которому присвоен соответствующий транспортный сигнал, таким образом, что сформированный V-вектор включает в себя элементы для любых транспортных сигналов, присвоенных уровням выше, чем уровень, которому присвоен соответствующий транспортный сигнал. Способ может дополнительно включать в себя сигнализацию сформированных V-векторов в выдаваемом битовом потоке.

В соответствии с другим аспектом описан способ многоуровневого кодирования кадра сжатого представления звука или звукового поля системы Амбисоник высшего порядка (Higher-Order Ambisonics, HOA). Сжатое представление HOA может включать в себя множество транспортных сигналов. Транспортные сигналы могут относиться к монауральным сигналам, например, представляющим либо преобладающие звуковые сигналы, либо последовательности коэффициентов представления HOA. Способ может включать в себя присвоение множества транспортных сигналов множеству иерархических уровней. Например, транспортные сигналы могут быть распределены множеству уровней. Множество уровней может включать в себя базовый уровень и один или более иерархических улучшающих уровней. Способ может дополнительно включать в себя определение, является ли активным векторный режим кодирования. Если векторный режим активен, способ может дополнительно включать в себя определение для каждого уровня множества непрерывных индексов коэффициентов HOA на основе транспортных сигналов, присвоенных соответствующему уровню. Индексы коэффициентов HOA в множестве непрерывных индексов коэффициента HOA могут являться индексами коэффициентов HOA, включенными в множество ContAddHOACoeff. Способ может дополнительно включать в себя формирование для каждого транспортного сигнала V-вектора на основе определенного множества непрерывных индексов коэффициентов HOA для уровня, которому присвоен соответствующий транспортный сигнал, таким образом, что сформированный V-вектор включает в себя элементы для любых транспортных сигналов, присвоенных уровням выше, чем уровень, которому присвоен соответствующий транспортный сигнал. Способ может дополнительно включать в себя сигнализацию сформированных V-векторов в выдаваемом битовом потоке.

Сконфигурированный таким образом предложенный способ гарантирует, что в векторном режиме кодирования подходящий V-вектор доступен для каждого транспортного сигнала, принадлежащего уровням вплоть до наиболее высокого применимого уровня. В частности, предложенный способ исключает случай, когда элементы V-вектора, соответствующего транспортным сигналам на более высоких уровнях, явно не сигнализируются. В соответствии с этим информация, включенная в уровни вплоть до наиболее высокого применимого уровня, является достаточной для декодирования любых транспортных сигналов, принадлежащих уровням вплоть до наиболее высокого применимого уровня. Таким образом, существует подходящее восстановление соответствующих воссозданных представлений HOA для более низких уровней (уровней с низким битрейтом), даже если более высокие уровни не могут быть корректно приняты декодером. С другой стороны, предложенный способ позволяет полностью использовать преимущества сокращения требуемой ширины полосы, которое может быть достигнуто при применении многоуровневого кодирования.

В соответствии с другим аспектом описан способ декодирования кадра сжатого представления звука или звукового поля системы Амбисоник высшего порядка (Higher-Order Ambisonics, HOA). Сжатое представление HOA может быть закодировано в множестве иерархических уровней. Множество иерархических уровней может включать в себя базовый уровень и один или более иерархических улучшающих уровней. Способ может включать в себя прием битового потока, относящегося к кадру сжатого представления HOA. Способ может дополнительно включать в себя извлечение полезных нагрузок для множества уровней. Каждая полезная нагрузка может включать в себя транспортные сигналы, присвоенные соответствующему уровню. Способ может дополнительно включать в себя определение наиболее высокого применимого уровня среди множества уровней для декодирования. Способ может дополнительно включать в себя извлечение полезной нагрузки расширения HOA, присвоенной наиболее высокому применимому уровню. Эта полезная нагрузка расширения HOA может включать в себя вспомогательную информацию для параметрического улучшения (частично) воссозданного представления HOA, соответствующего наиболее высокому применимому уровню. (Частично) воссозданное представление HOA, соответствующее наиболее высокому применимому уровню, может быть получено на основе

транспортных сигналов, присвоенных наиболее высокому применимому уровню и любым уровням ниже наиболее высокого применимого уровня. Способ может дополнительно включать в себя формирование (частично) воссозданного представления HOA, соответствующего наиболее высокому применимому уровню, на основе транспортных сигналов, присвоенных наиболее высокому применимому уровню и любым уровням ниже наиболее высокого применимого уровня. Кроме того, способ может включать в себя улучшение (например, параметрическое улучшение) (частично) воссозданного представления HOA с использованием вспомогательной информации, включенной в полезную нагрузку расширения HOA, присвоенную наиболее высокому применимому уровню. В результате может быть получено улучшенное воссозданное представление HOA.

Сконфигурированный таким образом предложенный способ гарантирует, что окончательное (например, расширенное) воссозданное представление HOA имеет оптимальное качество с использованием доступной (например, корректно принятой) информации в наилучшей возможной степени.

В вариантах осуществления полезные нагрузки расширения HOA могут включать в себя элементы битового потока для инструмента декодирования с пространственным предсказанием сигналов HOA. Дополнительно или в качестве альтернативы полезные нагрузки расширения HOA могут включать в себя элементы битового потока для инструмента декодирования с синтезом направленных подполосных сигналов HOA. Дополнительно или в качестве альтернативы полезные нагрузки расширения HOA могут включать в себя элементы битового потока для инструмента декодирования с параметрическим дублированием звукового окружения HOA.

В вариантах осуществления полезные нагрузки расширения HOA могут иметь тип usacExtElementType ID_EXT_ELE_HOA_ENH_LAYER.

В вариантах осуществления способ может дополнительно включать в себя извлечение полезной нагрузки расширения конфигурации HOA посредством анализа битового потока. Полезная нагрузка расширения конфигурации HOA может включать в себя элементы битового потока для конфигурирования инструмента декодирования с пространственным предсказанием сигналов HOA, инструмента декодирования с синтезом направленных подполосных сигналов HOA и/или инструмента декодирования с параметрическим дублированием звукового окружения HOA.

В вариантах осуществления способ может дополнительно включать в себя извлечение полезных нагрузок расширения HOA, соответственно присвоенных множеству уровней. Каждая полезная нагрузка расширения HOA может включать в себя вспомогательную информацию для параметрического расширения (частично) воссозданного представления HOA, соответствующего своему соответствующему присвоенному уровню. (Частично) воссозданное представление HOA, соответствующее своему соответствующему присвоенному уровню, может быть получено из транспортных сигналов, присвоенных этому уровню и любым уровням ниже этого уровня. Присвоение полезных нагрузок расширения HOA соответствующим уровням может быть известно из информации конфигурации, включенной в битовый поток.

В вариантах осуществления определение наиболее высокого применимого уровня может включать в себя определение множества индексов некорректных уровней, указывающего уровни, которые не были приняты корректно. Это может дополнительно включать в себя определение наиболее высокого применимого уровня как уровня, который является уровнем ниже уровня, указанного наименьшим (наиболее низким) индексом в множестве индексов некорректных уровней. Базовый уровень может иметь наиболее низкий индекс (например, индекс 1 уровня) и иерархические улучшающие уровни могут иметь последовательно более высокие индексы уровней. Тем самым предложенный способ гарантирует, что наиболее высокий применимый уровень выбран таким образом, что вся информация, требуемая для декодирования (частично) воссозданного представления HOA из наиболее высоких применимых уровней и любых уровней ниже наиболее высокого применимого уровня, является доступной.

В вариантах осуществления определение наиболее высокого применимого уровня может включать в себя определение множества индексов некорректных уровней, указывающего уровни, которые не были приняты корректно. Это может дополнительно включать в себя определение наиболее высокого применимого уровня предыдущего кадра, предшествующего текущему кадру. Это может, кроме того, включать в себя определение наиболее высокого применимого уровня как более низкого из наиболее высокого применимого уровня предыдущего кадра и уровня, который является уровнем ниже уровня, указанного самым малым индексом в множестве индексов некорректных уровней. Таким образом, наиболее высокий применимый уровень для текущего кадра выбирается таким образом, что вся информация, требуемая для декодирования (частично) воссозданного представления HOA из наиболее высокого применимого уровня и любых уровней ниже наиболее высокого применимого уровня, является доступной, даже если текущий кадр был закодирован дифференциально относительно предыдущего кадра.

В вариантах осуществления способ может дополнительно включать в себя принятие решения не выполнить параметрическое улучшение (частично) воссозданного представления HOA с использованием вспомогательной информации, включенной в полезную нагрузку расширения HOA, присвоенную наиболее высокому применимому уровню, если наиболее высокий применимый уровень текущего кадра ниже, чем наиболее высокий применимый уровень предыдущего кадра, и если текущий кадр был закодирован дифференциально относительно предыдущего кадра. Таким образом, воссозданное представление HOA

может быть декодировано без ошибки в случаях, в которых текущий кадр (включающий в себя вспомогательную информацию, включенную в полезную нагрузку расширения HOA, присвоенную наиболее высокому применимому уровню) был закодирован дифференциально относительно предыдущего кадра.

В вариантах осуществления множество индексов некорректных уровней может быть определено посредством оценки флагов корректности соответствующих полезных нагрузок расширения HOA. Индекс уровня данного уровня может быть добавлен к множеству индексов некорректных уровней, если флаг корректности для полезной нагрузки расширения HOA, присвоенной соответствующему уровню, не установлен. Тем самым множество индексов некорректных уровней может быть определено эффективным образом.

В соответствии с другим аспектом описана структура данных (например, битовый поток), представляющая кадр сжатого представления звука или звукового поля системы Амбисоник высшего порядка (Higher-Order Ambisonics, HOA). Сжатое представление HOA может включать в себя множество транспортных сигналов. Структура данных может включать в себя множество полезных нагрузок кадра HOA, связанных с соответствующими уровнями из множества иерархических уровней. Полезные нагрузки кадра HOA могут включать в себя соответствующие транспортные сигналы. Множество транспортных сигналов может быть присвоено (например, распределено) множеству уровней. Множество уровней может включать в себя базовый уровень и один или более иерархических улучшающих уровней. Структура данных может дополнительно включать в себя для каждого уровня соответствующую полезную нагрузку расширения HOA, включающую в себя вспомогательную информацию для параметрического улучшения (частично) воссозданного представления HOA, которая может быть получена из транспортных сигналов, присвоенных соответствующему уровню и любым уровням ниже соответствующего уровня.

В вариантах осуществления полезные нагрузки кадра HOA и полезные нагрузки расширения HOA для множества уровней могут быть снабжены соответствующими уровнями защиты от ошибок. Базовый уровень может иметь наиболее высокую защиту от ошибок, и один или более улучшающих уровней могут иметь последовательно уменьшающуюся защиту от ошибок.

В вариантах осуществления полезные нагрузки расширения HOA могут включать в себя элементы битового потока для инструмента декодирования с пространственным предсказанием сигналов HOA. Дополнительно или в качестве альтернативы полезные нагрузки расширения HOA могут включать в себя элементы битового потока для инструмента декодирования с синтезом направленных подполосных сигналов HOA. Дополнительно или в качестве альтернативы полезные нагрузки расширения HOA могут включать в себя элементы битового потока для инструмента декодирования с параметрическим дублированием звукового окружения HOA.

В вариантах осуществления полезные нагрузки расширения HOA могут иметь тип usacExtElementType ID_EXT_ELE_HOA_ENH_LAYER.

В вариантах осуществления структура данных может дополнительно включать в себя полезную нагрузку расширения конфигурации HOA, включающую в себя элементы битового потока для конфигурирования инструмента декодирования с пространственным предсказанием сигналов HOA, инструмента декодирования с синтезом направленных подполосных сигналов HOA и/или инструмента декодирования с параметрическим дублированием звукового окружения HOA.

В вариантах осуществления структура данных может дополнительно включать в себя полезную нагрузку конфигурации декодера HOA, включающую в себя информацию, указывающую присвоение полезных нагрузок расширения HOA множеству уровней.

В вариантах осуществления способы и устройства относятся к декодированию сжатого представления звука или звукового поля системы Амбисоник высшего порядка (Higher-Order Ambisonics, HOA). Устройство может быть выполнено с возможность выполнять или способ может включать в себя прием битового потока, содержащего сжатое представление HOA, соответствующее множеству иерархических уровней, которые включают в себя базовый уровень и один или более иерархических улучшающих уровней, причем множество уровней имеют присвоенные им компоненты базового сжатого представления звука или звукового поля, компоненты присвоены соответствующим уровням в соответствующих группах компонентов, определение наиболее высокого применимого уровня среди множества уровней для декодирования; извлечение полезной нагрузки расширения HOA, присвоенной наиболее высокому применимому уровню, причем полезная нагрузка расширения HOA включает в себя вспомогательную информацию для параметрического улучшения воссозданного представления HOA, соответствующего наиболее высокому применимому уровню, причем воссозданное представление HOA, соответствующее наиболее высокому применимому уровню, может быть получено на основе транспортных сигналов, присвоенных наиболее высокому применимому уровню и любым уровням ниже наиболее высокого применимого уровня; декодирование сжатого представления HOA, соответствующего наиболее высокому применимому уровню, на основе информации уровня, транспортных сигналов, присвоенных наиболее высокому применимому уровню и любым уровням ниже наиболее высокого применимого уровня; и параметрическое улучшение декодированного представления HOA с использованием вспомогательной информации, включенной в полезную нагрузку расширения HOA, присвоенную наиболее высокому применимому уровню.

Полезная нагрузка расширения HOA может включать в себя элементы битового потока для инструмента декодирования с пространственным предсказанием сигналов HOA. Информация уровня может указывать количество активных направленных сигналов в текущем кадре улучшающего уровня.

Информация уровня может указывать общее количество дополнительных коэффициентов окружения HOA для улучшающего уровня. Информация уровня может включать в себя индексы коэффициентов HOA для каждого дополнительного коэффициента окружения HOA для улучшающего уровня. Информация уровня может включать в себя улучшающую информацию, которая включает в себя по меньшей мере одно из пространственного предсказания сигналов, синтеза направленных подполосных сигналов и декодера параметрического дублирования звукового окружения. Сжатое представление HOA адаптировано для режима многоуровневого кодирования для контента на основе HOA, если в HOADecoderConfig() сигнализирован CodedVVecLength == 1. Кроме того, v-векторные элементы могут не передаваться для индексов, которые равны индексам дополнительных коэффициентов HOA, включенных в множество ContAddHoaCoeff. Множество ContAddHoaCoeff может быть определено отдельно для каждого множества иерархических уровней. Информация уровня включает в себя элементы NumLayers, причем каждый элемент указывает количество транспортных сигналов, включенных во все уровни вплоть до i-го уровня. Информация уровня может включать в себя индикатор всех фактически используемых уровней для k-го кадра. Информация уровня также может указывать, что определены все коэффициенты для преобладающих векторов. Информация уровня может указывать, что определены коэффициенты преобладающих векторов, соответствующих номеру больше MinNumOfCoeffsForAmbHOA.

Информация уровня может указывать, что MinNumOfCoeffsForAmbHOA и все элементы, определенные в ContAddHoaCoeff[lay], не передаются, где lay - индекс уровня, содержащего основанный на векторе сигнал, соответствующий вектору.

В соответствии с другим аспектом описан кодер для многоуровневого кодирования кадра сжатого представления звука или звукового поля системы Амбисоник высшего порядка (Higher-Order Ambisonics, HOA). Сжатое представление HOA может включать в себя несколько транспортных сигналов. Кодер может включать в себя процессор, выполненный с возможностью выполнять некоторые или все этапы способов в соответствии с первым упомянутым выше аспектом и вторым упомянутым выше аспектом.

В соответствии с другим аспектом описан декодер для декодирования кадра сжатого представления звука или звукового поля системы Амбисоник высшего порядка (Higher-Order Ambisonics, HOA). Сжатое представление HOA может быть закодировано в множестве иерархических уровней, которые включают в себя базовый уровень и один или более иерархических улучшающих уровней. Декодер может включать в себя процессор, выполненный с возможностью выполнять некоторые или все из этапов способов в соответствии с третьим упомянутым выше аспектом.

В соответствии с другим аспектом описана программа. Программа может быть адаптирована для исполнения на процессоре и для выполнения некоторых или всех этапов способа, изложенных в настоящем документе, при ее исполнении на вычислительном устройстве.

В соответствии с еще одним аспектом описан запоминающий носитель. Запоминающий носитель может включать в себя программу, адаптированную для исполнения на процессоре и для выполнения некоторых или всех этапов способа, изложенных в настоящем документе, при ее исполнении на вычислительном устройстве.

Согласно предпочтительному варианту осуществления предложен способ декодирования сжатого представления звука или звукового поля HOA, содержащий этапы, на которых принимают битовый поток, содержащий сжатое представление HOA, соответствующее множеству иерархических уровней, которые включают в себя базовый уровень и два или более иерархических улучшающих уровней, причем битовый поток содержит для каждого уровня соответствующую полезную нагрузку расширения HOA, включающую в себя вспомогательную информацию для параметрического улучшения воссозданного представления HOA, получаемого из транспортных сигналов, присвоенных соответствующему уровню и любым уровням ниже соответствующего уровня; определяют наиболее высокий применимый уровень среди множества уровней для декодирования, причем наиболее высокий применимый уровень представляет собой уровень ниже наиболее низкого уровня, который не был принят корректно; извлекают полезную нагрузку расширения HOA, присвоенную наиболее высокому применимому уровню, причем полезная нагрузка расширения HOA включает в себя вспомогательную информацию для параметрического улучшения воссозданного представления HOA, соответствующего наиболее высокому применимому уровню, причем воссозданное представление HOA, соответствующее наиболее высокому применимому уровню, может быть получено на основе транспортных сигналов, присвоенных наиболее высокому применимому уровню и любым уровням ниже наиболее высокого применимого уровня; декодируют сжатое представление HOA, соответствующее наиболее высокому применимому уровню, на основе информации уровня, транспортных сигналов, присвоенных наиболее высокому применимому уровню и любым уровням ниже наиболее высокого применимого уровня; и выполняют параметрическое улучшение декодированного представления HOA с использованием вспомогательной информации, включенной в полезную нагрузку расширения HOA, присвоенную наиболее высокому применимому уровню.

Согласно предпочтительному варианту осуществления, предложено устройство для декодирования

сжатого представления звука или звукового поля HOA, содержащее приемник, выполненный с возможностью принимать битовый поток, содержащий сжатое представление HOA, соответствующее множеству иерархических уровней, которые включают в себя базовый уровень и два или более иерархических улучшающих уровней, причем битовый поток содержит для каждого уровня соответствующую полезную нагрузку расширения HOA, включающую в себя вспомогательную информацию для параметрического улучшения воссозданного представления HOA, получаемого из транспортных сигналов, присвоенных соответствующему уровню и любым уровням ниже соответствующего уровня; декодер, выполненный с возможностью определять наиболее высокий применимый уровень среди множества уровней для декодирования, причем наиболее высокий применимый уровень представляет собой уровень ниже наиболее низкого уровня, который не был принят корректно, извлекать полезную нагрузку расширения HOA, присвоенную наиболее высокому применимому уровню, причем полезная нагрузка расширения HOA включает в себя вспомогательную информацию для параметрического улучшения воссозданного представления HOA, соответствующего наиболее высокому применимому уровню, причем воссозданное представление HOA, соответствующее наиболее высокому применимому уровню, может быть получено на основе транспортных сигналов, присвоенных наиболее высокому применимому уровню и любым уровням ниже наиболее высокого применимого уровня, декодировать сжатое представление HOA, соответствующее наиболее высокому применимому уровню, на основе информации уровня, транспортных сигналов, присвоенных наиболее высокому применимому уровню и любым уровням ниже наиболее высокого применимого уровня, и выполнять параметрическое улучшение декодированного представления HOA с использованием вспомогательной информации, включенной в полезную нагрузку расширения HOA, присвоенную наиболее высокому применимому уровню.

Полезная нагрузка расширения HOA может включать в себя элементы битового потока для инструмента декодирования с пространственным предсказанием сигналов HOA.

Информация уровня может указывать количество активных направленных сигналов в текущем кадре улучшающего уровня.

Информация уровня может указывать общее количество дополнительных коэффициентов окружения HOA для улучшающего уровня.

Информация уровня может включать в себя индексы коэффициентов HOA для каждого дополнительного коэффициента окружения HOA для улучшающего уровня.

Информация уровня может включать в себя улучшающую информацию, которая включает в себя по меньшей мере одно из пространственного предсказания сигналов, синтеза направленных подполосных сигналов и декодера параметрического дублирования звукового окружения.

Сжатое представление HOA может быть адаптировано для режима многоуровневого кодирования для контента на основе HOA, если в HOADecoderConfig() сигнализирован CodedVVecLength, равный 1.

Множество ContAddHoaCoeff может быть определено отдельно для каждого из множества иерархических уровней.

Информация уровня может включать в себя элементы NumLayers, причем каждый элемент указывает количество транспортных сигналов, включенных во все уровни вплоть до i-го уровня.

Информация уровня может включать в себя индикатор всех фактически используемых уровней для k-го кадра.

Информация уровня может указывать, что определены все коэффициенты для преобладающих векторов.

Информация уровня может указывать, что определены коэффициенты преобладающих векторов, соответствующих номеру, который больше MinNumOfCoeffsForAmbHOA.

Информация уровня может указывать MinNumOfCoeffsForAmbHOA, и все элементы, определенные в ContAddHoaCoeff[lay], не передаются, где lay - индекс уровня, содержащего основанный на векторе сигнал, соответствующий вектору.

Следует понимать, что утверждения, сделанные в отношении любого из упомянутых выше аспектов или их вариантов осуществления, также относятся к соответствующим другим аспектам или их вариантам осуществления, как поймет специалист в области техники. Повторение этих утверждений для каждого аспекта или варианта осуществления было опущено для краткости.

Следует отметить, что способы и устройство, включающие в себя предпочтительные варианты осуществления, изложенные в настоящем документе, могут использоваться автономно или в сочетании с другими способами и системами, раскрытыми в этом документе. Кроме того, все аспекты способов и устройства, изложенные в настоящем документе, могут быть произвольным образом объединены. В частности, признаки пунктов формулы изобретения могут быть объединены друг с другом произвольным образом.

Также следует отметить, что этапы способов и признаки устройств могут являться взаимозаменяемыми различным образом. В частности, подробности раскрытого способа могут быть реализованы как устройство, выполненное с возможностью исполнять некоторые или все этапы способа и наоборот, как поймет специалист в области техники.

Краткое описание чертежей

Изобретение разъяснено ниже иллюстративным образом со ссылкой на прилагаемые чертежи.

Фиг. 1 - блок-схема, иллюстрирующая присвоение полезных нагрузок базовому уровню и M-1 улучшающим уровням на стороне кодера.

Фиг. 2 - блок-схема, иллюстрирующая пример модуля приема и восстановления.

Фиг. 3 - блок-схема последовательности этапов, иллюстрирующая пример способа многоуровневого кодирования кадра сжатого представления HOA в соответствии с вариантами осуществления раскрытия.

Фиг. 4 - блок-схема последовательности этапов, иллюстрирующая другой пример способа многоуровневого кодирования кадра сжатого представления HOA в соответствии с вариантами осуществления раскрытия.

Фиг. 5 - блок-схема последовательности этапов, иллюстрирующая пример способа декодирования кадра сжатого представления HOA в соответствии с вариантами осуществления раскрытия.

Фиг. 6 - блок-схема, иллюстрирующая пример аппаратной реализации кодера в соответствии с вариантами осуществления раскрытия.

Фиг. 7 - блок-схема, иллюстрирующая пример аппаратной реализации декодера в соответствии с вариантами осуществления раскрытия.

Осуществление изобретения

Сначала будет описано сжатое представление звука (или звукового поля), к которому могут быть применены способы и кодеры/декодеры в соответствии с настоящим раскрытием.

Для потокового сжатого представления звука (или звукового поля) по каналу передачи с изменяющимися по времени условиями многоуровневое кодирование является средством, чтобы адаптировать качество принятого представления звука к условиям передачи, и в частности избежать нежелательных пропаданий сигнала.

Для многоуровневого кодирования сжатое представление звука (или звукового поля) обычно подразделяется на высокоприоритетный базовый уровень относительно небольшого размера и дополнительные улучшающие уровни с убывающими приоритетами и произвольными размерами. Каждый улучшающий уровень, как обычно предполагается, содержит нарастающую информацию, чтобы дополнить все из более низких уровней для улучшения качества сжатого представления звука (или звукового поля). Идея состоит в том, чтобы управлять величиной защиты от ошибок для передачи индивидуальных уровней в соответствии с их приоритетом. В частности, базовому уровню предоставляется высокая защита от ошибок, которая является разумной и приемлемой вследствие ее малого размера.

Далее предполагается, что полное сжатое представление звука (или звукового поля) в общем случае состоит из трех следующих компонентов.

1. Базовое сжатое представление звука (или звукового поля), само состоящее из нескольких взаимодополняющих компонентов, которое составляет определенно наибольший процент полного сжатого представления звука (или звукового поля).

2. Базовая вспомогательная информация, необходимая для декодирования базового сжатого представления звука, которая, как предполагается, имеет намного меньший размер по сравнению с базовым сжатым представлением звука (или звукового поля). Также предполагается, что она по наибольшей части состоит из двух следующих компонентов, оба из которых определяют восстановление только одного конкретного компонента базового сжатого представления звука.

a) Первый компонент содержит вспомогательную информацию, описывающую индивидуальные взаимодополняющие компоненты базового сжатого представления звука (или звукового поля) независимо от других взаимодополняющих компонентов.

b) Второй (факультативный) компонент содержит вспомогательную информацию, описывающую индивидуальные взаимодополняющие компоненты базового сжатого представления звука (или звукового поля) в зависимости от других взаимодополняющих компонентов. В частности, зависимость имеет следующие свойства:

зависимая вспомогательная информация для каждого индивидуального взаимодополняющего компонента базового сжатого представления звука (или звукового поля) достигает своей самой большой степени в случае, если другие определенные взаимодополняющие компоненты не содержатся в базовом сжатом представлении звука (или звукового поля);

в случае если дополнительные определенные взаимодополняющие компоненты добавлены к базовому сжатому представлению звука (или звукового поля), зависимая вспомогательная информация для рассматриваемого индивидуального взаимодополняющего компонента становится подмножеством первоначальной информации, тем самым сокращая ее размер.

3. Факультативная улучшающая вспомогательная информация для улучшения базового сжатого представления звука (или звукового поля). Также предполагается, что ее размер намного меньше, чем у базового сжатого представления звука (или звукового поля).

Один известный пример такого типа полного сжатого представления звука (или звукового поля) дан посредством сжатого представления HOA звукового поля, определенный предварительной версией аудио стандарта MPEG-H 3D.

1. Его базовое сжатое представление звукового поля может быть идентифицировано с помощью нескольких квантованных монауральных сигналов, представляющих либо так называемые преобладающие звуковые сигналы, либо последовательности коэффициентов так называемого компонента окружения HOA звукового поля.

2. Базовая вспомогательная информация среди прочего описывает для каждого из этих монауральных сигналов, каким образом он вносит пространственный вклад в звуковое поле. Эта информация может быть далее разделена на следующие два разных компонента.

(a) Вспомогательная информация, относящаяся к конкретным индивидуальным монауральным сигналам, которые являются независимыми от существования других монауральных сигналов. Такая вспомогательная информация может, например, определять монауральный сигнал для представления направленного сигнала (означающего общую плоскую волну) с некоторым направлением падения. В качестве альтернативы монауральный сигнал может быть определен как последовательность коэффициентов первоначального представления HOA, имеющую некоторый индекс.

(b) Вспомогательная информация, относящаяся к конкретным индивидуальным монауральным сигналам, которые зависят от существования других монауральных сигналов. Такая вспомогательная информация имеет место, например, если монауральные сигналы определены как так называемые основанные на векторах сигналы, что означает, что они распределены по направлениям в звуковом поле, причем распределение по направлениям определено посредством вектора. В некотором режиме (т.е. CodedVVecLength=1) конкретные компоненты этого вектора неявно установлены равными нулю и не являются частью сжатого векторного представления. Этими компонентами являются компоненты с индексами, равными индексам последовательностей коэффициентов первоначального представления HOA, которые являются частью базового сжатого представления звукового поля. Это означает, что, если индивидуальные компоненты вектора закодированы, их общее количество зависит от базового сжатого представления звукового поля, в частности от того, какие последовательности коэффициентов первоначального представления HOA оно содержит.

Если последовательности коэффициентов первоначального представления HOA не содержатся в базовом сжатом представлении звукового поля, зависимая базовая вспомогательная информация для каждого основанного на векторе сигнала состоит из всех векторных компонентов и имеет свой наибольший размер. В случае если последовательности коэффициентов первоначального представления HOA с некоторыми индексами добавляются к базовому сжатому представлению звукового поля, векторные компоненты с этими индексами удаляются из вспомогательной информации для каждого основанного на векторе сигнала, тем самым сокращая размер зависимой базовой вспомогательной информации для основанных на векторах сигналов.

3. Улучшающая вспомогательная информация состоит из следующих компонентов.

Параметры, относящиеся к так называемому (широкополосному) пространственному предсказанию для (линейного) предсказания недостающих частей звукового поля на основе направленных сигналов.

Параметры, относящиеся к так называемому синтезу направленных подполосных сигналов и параметрическому дублированию звукового окружения, которые являются инструментами сжатия, позволяющими пространственно распределять зависимое от частоты параметрическое предсказание дополнительных монауральных сигналов, чтобы дополнить до сих пор пространственно неполное или несовершенным образом сжатое представление HOA. Предсказание основано на последовательностях коэффициентов базового сжатого представления звукового поля. Важным аспектом является то, что упомянутый взаимодополняющий вклад в звуковое поле представлен в сжатом представлении HOA не посредством дополнительных квантованных сигналов, а посредством дополнительной вспомогательной информации сравнительно намного меньшего размера. Следовательно, два упомянутых инструмента кодирования особенно подходят для сжатия представлений HOA на низких скоростях передачи данных.

Второй пример сжатого представления монаурального сигнала с упомянутой выше структурой может состоять из следующих компонентов.

1. Некоторая закодированная спектральная информация для несвязных частотных полос вплоть до некоторой верхней частоты, которая может рассматриваться как базовое сжатое представление.

2. Некоторая базовая вспомогательная информация, определяющая закодированную спектральную информацию (например, количество и ширину закодированных частотных полос).

3. Некоторая улучшающая вспомогательная информация, состоящая из параметров так называемого копирования спектральной полосы (SBR), описывающая, каким образом параметрически воссоздать из базового сжатого представления спектральную информацию для полос более высокой частоты, которые не рассматриваются в базовом сжатом представлении.

Далее будет описан способ многоуровневого кодирования полного сжатого представления звука (или звукового поля), имеющего упомянутую выше структуру.

Предполагается, что сжатие основано на кадрах в том смысле, что оно обеспечивает сжатые представления (например, в форме пакетов данных или эквивалентно полезных нагрузок кадров) для последовательных временных интервалов, например временных интервалов равного размера. Предполагается, что эти пакеты данных содержат флаг корректности, значение, указывающее их размер, а также фактиче-

ские сжатые данные представления. Везде в последующем описании будет сделан фокус главным образом на обработке одного кадра, и поэтому индекс кадра будет опущен.

Предполагается, что каждая полезная нагрузка кадра рассматриваемого полного сжатого представления 1100 звука (или звукового поля) содержит J пакетов данных, каждый для одного компонента 1110-1, ..., 1110-J базового сжатого представления звука (или звукового поля), которые обозначены как $BSRC_j$, j=l,..., J. Кроме того, предполагается, что она содержит пакет с независимой базовой вспомогательной информацией 1120, обозначенной как BSI_I, определяющей отдельные компоненты $BSRC_j$ базового сжатого представления звука независимо от других компонентов. Факультативно предполагается, что она содержит пакет с зависимой базовой вспомогательной информацией, обозначенной как BSI_D, определяющей отдельные компоненты BSRCj базового сжатого представления звука в зависимости от других компонентов. Информация, содержащаяся в двух пакетах данных BSI_I и BSI_D, может быть факультативно сгруппирована в единый пакет данных BSI.

В конечном счете, она включает в себя полезную нагрузку улучшающей вспомогательной информации, обозначенной как ESI, с описанием того, каким образом улучшить воссозданный звук (или звуковое поле) из полного базового сжатого представления.

Описанная схема многоуровневого кодирования направлена на этапы, требующиеся для обеспечения возможности как для части сжатия, включающей в себя упаковку пакетов данных для передачи, а также для части приема и восстановления. Каждая часть будет подробно описана далее.

Далее будут описаны сжатие и упаковка для передачи. В случае многоуровневого кодирования (предполагая в целом M уровней, т.е. один базовый уровень и M-1 улучшающих уровней) каждый компонент полного сжатого представления 1100 звука (или звукового поля) обрабатывается следующим образом.

Базовое сжатое представление звука (или звукового поля) подразделяется на части, которые будут присвоены индивидуальным уровням. Без потери общности группировка может быть описана M+1 номерами J_m, m=0, ..., M, где J_0=l и J_M=J+1, в результате чего $BSRC_j$ присваивается m-му уровню для J_m-1≤j<J_m.

Вследствие своего небольшого размера целесообразно присвоить полную базовую вспомогательную информацию базовому уровню, чтобы избежать ее ненужной фрагментации. В то время как независимая базовая вспомогательная информация BSI_I остается без изменений для назначения, зависимая базовая вспомогательная информация должна быть обработана специальным образом для многоуровневого кодирования, чтобы позволить правильное декодирование на стороне приемника, с одной стороны, и сократить размер зависимой вспомогательной информации для передачи, с другой стороны. Предлагается выполнить ее декомпозицию на M частей 1130-1, ..., 1130-M, обозначенных как $BSI_{D,m}$, m=1,..., M, где m-я часть содержит зависимую вспомогательную информацию для каждого из компонентов $BSRC_j$, J_{m-1}≤j<J_m, базового сжатого представления звука, присвоенного m-му уровню, если соответствующая зависимая вспомогательная информация существует.

В случае если соответствующая зависимая вспомогательная информация не существует, предполагается, что $BSI_{D,m}$ является пустой. Вспомогательная информация $BSI_{D,m}$ зависит от всех компонентов $BSRC_j$, 1≤j<J_m, содержавшихся на всех уровнях вплоть до m-го.

В случае многоуровневого кодирования важно понимать, что улучшающая вспомогательная информация должна быть вычислена для каждого дополнительного уровня, поскольку она предназначается для улучшения предварительного восстановленного звука (или звукового поля), что, однако, зависит от доступных уровней для восстановления. Следовательно, сжатие должно обеспечить M индивидуальных пакетов данных 1140-1,..., 1140-M улучшающей вспомогательной информации, обозначенных как ESI_m, m=1,..., M, причем улучшающая вспомогательная информация в m-ом пакете данных ESI_m вычисляется, например, для улучшения представления звука (или звукового поля), полученного из всех данных, содержащихся на базовом уровне и улучшающих уровнях с индексами ниже m.

Подводя итог, на стадии сжатия должен быть обеспечен пакет данных кадра, обозначенный как FRAME, имеющий следующий состав:

$$FRAME=[BSRC_1 \ldots\ BSRC_J\ BSI_I\ BSI_{D,1} \ldots\ BSI_{D,M}\ ESI_1 \ldots\ ESI_M] \quad (1)$$

Подразумевается, что порядок индивидуальных полезных нагрузок с пакетом данных кадра в общем случае является произвольным.

Уже описанное назначение индивидуальных полезных нагрузок базовому и улучшающим уровням достигается посредством так называемого упаковщика транспортных уровней и схематично проиллюстрировано на фиг. 1.

Далее будут описаны прием и восстановление. Соответствующий модуль приема и восстановления проиллюстрирован на фиг. 2.

Сначала пакеты 1200, 1300-1,..., 1300-(M-1) индивидуальных уровней мультиплексируются, чтобы обеспечить принятый пакет кадра

$$[BSI_I\ \ BSI_{D,1}\ \ \ldots\ \ BSI_{D,M}\ \ ESI_1\ \ BSRC_1\ \ \ldots\ \ BSRC_{(J_1)-1}\ \ \ldots\ \ ESI_M \quad (2)$$

полного сжатого представления звука (или звукового поля), который затем передается на декомпрессор 2100. Предполагается, что, если передача индивидуального уровня была безошибочной, флаг

корректности, по меньшей мере, содержащейся полезной нагрузки улучшающей вспомогательной информации установлен равным значению "истинный". В случае ошибки вследствие передачи индивидуального уровня флаг корректности, по меньшей мере, в полезной нагрузке улучшающей вспомогательной информации на этом уровне установлен равным значению "ложный". Следовательно, корректность пакета уровня может быть определена на основе корректности содержащейся полезной нагрузки улучшающей вспомогательной информации.

В декомпрессоре 2100 принятый пакет кадра сначала подвергается демультиплексированию. С этой целью может использоваться информация размера каждой полезной нагрузки, чтобы избежать ненужного анализа данных индивидуальных полезных нагрузок.

На следующем этапе выбирается номер N_B наиболее высокого уровня, который будет фактически использоваться для восстановления базового представления звука. Наиболее высокий улучшающий уровень, который будет фактически использоваться для восстановления базового представления звука, задан N_B-1. Поскольку каждый уровень содержит точно одну полезную нагрузку улучшающей вспомогательной информации, из каждой полезной нагрузки улучшающей вспомогательной информации известно, является ли содержание уровня корректным или нет. Следовательно, выбор может быть достигнут с использованием всей полезной нагрузки улучшающей вспомогательной информации ESI_m, m=1,..., M. Кроме того, определяется индекс N_E полезной нагрузки улучшающей вспомогательной информации, которая будет использоваться для восстановления, который всегда либо равен N_B, либо равен нулю. Это означает, что улучшение достигается либо всегда в соответствии с базовым представлением звука, либо вообще не достигается. Более подробное описание выбора дано далее ниже.

Последовательно полезные нагрузки компонентов $BSRC_1$,..., $BSRC_j$ базового сжатого представления звука вместе со всеми полезными нагрузками базовой вспомогательной информации (т.е. BSI_I и $BSI_{D,m}$, m=1,..., M) и значением N_B передаются процессору 220 восстановления базового представления, который воссоздает базовое представление звука (или звукового поля) с использованием только этих компонентов базового сжатого представления звука, содержащихся на наиболее низких N_B уровнях (т.е. на базовом уровне и N_B-1 улучшающих уровнях). Предполагается, что требуемая информация о том, какие компоненты базового сжатого представления звука (или звукового поля) содержатся на индивидуальных уровнях, известна декомпрессору 2100 из пакета данных с информацией конфигурации, которая предполагается отправленной и принятой перед пакетами данных кадра. Фактическое декодирование каждой индивидуальной зависимой полезной нагрузки $BSI_{D,m}$, m=l,..., N_B, базовой вспомогательной информации может быть разбито на две части следующим образом.

1. Предварительное декодирование каждой полезной нагрузки $BSI_{D,m}$, m=1,..., N_B, с использованием ее зависимости от первых J_m-1 компонентов

$$BSRC_1,...,BSRC_{(J_m)-1}$$

базового сжатого представления звука, содержащихся на первых m уровнях, что предполагалось на стадии кодирования.

2. Последовательная коррекция каждой полезной нагрузки $BSI_{D,m}$, m=1,..., N_B, с учетом того, что базовый компонент звука, наконец, воссоздан из первых

$$J_{N_B}-1$$

компонентов

$$BSRC_1,...,BSRC_{(J_{N_B})-1}$$

базового сжатого представления звука, содержащихся на первых N_B>m уровнях, что является большим количеством компонентов, чем предполагалось для предварительного декодирования. Следовательно, коррекция может быть достигнута посредством отбрасывания неадекватной информации, что возможно вследствие первоначально принятого свойства зависимой базовой вспомогательной информации, состоящего в том, что, если некоторые взаимодополняющие компоненты добавляются к базовому сжатому представлению звука (или звукового поля), зависимая базовая вспомогательная информация для каждого индивидуального взаимодополняющего компонента становится подмножеством первоначальной.

В конечном счете воссозданное базовое представление звука (или звукового поля) вместе со всеми полезными нагрузками ESI_1, ..., ESI_M улучшающей вспомогательной информации, полезными нагрузками BSI_I и $BSI_{D,m}$, m=1,..., M, базовой вспомогательной информации и значение N_E обеспечиваются процессору 2300 восстановления базового представления, который вычисляет окончательное улучшенное представление звука (или звукового поля) с использованием только полезной нагрузки

$$ESI_{N_E}$$

улучшающей вспомогательной информации и с отбрасыванием всех других полезных нагрузок улучшающей вспомогательной информации. Если значение N_E равно нулю, отбрасываются все полезные нагрузи улучшающей вспомогательной информации, и воссозданное окончательное улучшенное представление звука (или звукового поля) является эквивалентным воссозданному базовому представлению звука (или звукового поля).

Далее будет описан выбор уровня. В случае если все пакеты данных кадра могут быть восстановле-

ны независимо друг от друга, номер N_B наиболее высокого уровня, который будет фактически использоваться для восстановления базового представления звука, и индекс N_E полезной нагрузки улучшающей вспомогательной информации, которая будет использоваться для восстановления, устанавливаются равными наибольшему номеру L корректной полезной нагрузки улучшающей вспомогательной информации, которая сама может быть определена посредством оценки флага корректности в полезных нагрузках улучшающей вспомогательной информации. Посредством использования знания размера каждой полезной нагрузки улучшающей вспомогательной информации, можно избежать сложного анализа фактических данных полезных нагрузок для определения их корректности.

В случае если используется это дифференциальное восстановление с межкадровыми зависимостями, дополнительно следует рассматривать решение от предыдущего кадра. С дифференциальным восстановлением независимые пакеты данных кадра передаются с регулярными временными интервалами, чтобы позволить начинать восстановление с тех моментов времени, когда определение значений N_B и N_E становятся независимыми от кадров, и оно выполняется, как описано выше.

Чтобы подробно разъяснить зависимое от кадров решение, сначала обозначаем для k-го кадра наибольший номер корректной полезной нагрузки улучшающей вспомогательной информации как L(k);

номер наиболее высокого уровня для выбора и использования для восстановления базового представления звука как $N_B(k)$;

номер полезной нагрузки улучшающей вспомогательной информации для использования для восстановления как $N_E(k)$.

Используя эти обозначения, номер наиболее высокого уровня для использования для восстановления базового представления звука как $N_B(k)$ вычисляется в соответствии с

$$N_B(k)=\min(N_B(k-1),\ L(k)). \quad (3)$$

Посредством выбора $N_B(k)$ не больше, чем $N_B(k-1)$ и L(k), гарантируется, что вся информация, требуемая для дифференциального восстановления базового представления звука, является доступной.

Номер $N_E(k)$ полезной нагрузки улучшающей вспомогательной информации для использования для восстановления определяется в соответствии с

$$N_E(k)=\begin{cases} N_B(k) & \text{if } N_B(k)=N_B(k-1) \\ 0 & \text{else} \end{cases}. \quad (4)$$

Это означает, в частности, что при условии, что номер $N_B(k)$ наиболее высокого уровня для использования для восстановления базового представления звука не изменяется, выбирается тот же самый соответствующий номер улучшающего уровня. Однако в случае изменения $N_B(k)$ улучшение запрещается посредством установки $N_E(k)$ равным нулю. Вследствие предполагаемого дифференциального восстановления улучшающей вспомогательной информации ее изменение в соответствии с $N_B(k)$ невозможно, поскольку это потребовало бы восстановления соответствующего уровня улучшающей вспомогательной информации в предыдущем кадре, которая, как предполагается, не была выполнена.

В качестве альтернативы, если при восстановлении все полезные нагрузки улучшающей вспомогательной информации с номером вплоть до $N_E(k)$ восстановлены параллельно, правило выбора (4) может быть заменено на

$$N_E(k)=N_B(k). \quad (5)$$

Наконец, следует отметить, что для дифференциального восстановления номер наиболее высокого используемого уровня может увеличиваться только в независимых пакетах данных кадра, тогда как уменьшение возможно в каждом кадре.

Далее буду описаны варианты осуществления раскрытия, относящиеся к многоуровневому кодированию кадра сжатого представления звука и к структуре данных (например, к битовому потоку), представляющей кадр закодированного сжатого представления звука, для случая сжатого представления HOA. В частности, будут описаны предложенные изменения схемы многоуровневого кодирования сжатого представления HOA.

В качестве коррекции режима многоуровневого кодирования для контента на основе HOA определен новый usacExtElementType, чтобы лучше адаптировать конфигурацию и полезные нагрузки инструментов декодирования, пространственного предсказания сигналов, синтеза направленных подполосных сигналов и декодера параметрического дублирования звукового окружения (PAR) к соответствующему улучшающему уровню HOA. Если активирован режим многоуровневого кодирования для контента на основе HOA, который сигнализирован посредством SingleLayer == 0, предлагается переместить соответствующие элементы битового потока этих инструментов в одну дополнительную полезную нагрузку расширения HOA нового типа для каждого уровня (в том числе базового уровня и одного или более улучшающих уровней).

Расширение должно быть сделано потому, что вспомогательная информация для этих инструментов создана для улучшения специфического представления HOA. В текущем определении многоуровневого кодирования HOA обеспеченные данные правильным образом расширяют только представление HOA наиболее высокого уровня. Для более низких уровней эти инструменты не улучшают частично воссозданное представление HOA правильным образом.

Таким образом, было бы лучше обеспечить вспомогательную информацию этих инструментов для каждого уровня, чтобы лучше адаптировать их к воссозданному представлению HOA соответствующего уровня.

Кроме того, инструменты синтеза направленных подполосных сигналов инструментов и декодера параметрического дублирования звукового окружения специально предназначены для низких скоростей передачи данных, когда доступны только несколько транспортных сигналов. Предложенное расширение, таким образом, предложило бы возможность оптимально адаптировать

вспомогательную информацию этих инструментов к количеству транспортных сигналов на уровне. В соответствии с этим качество звука воссозданного представления HOA для уровней с низким битрейтом, например для базового уровня, может быть значительно увеличено по сравнению с существующим многоуровневым подходом.

Кроме того, синтаксис битового потока для закодированных V-векторных элементов для основанных на векторах сигналов должен быть адаптирован для многоуровневого кодирования HOA, если CodedVVecLength == 1 сигнализирован в HOADecoderConfig(). В этом векторном режиме кодирования V-векторные элементы не передаются для индексов коэффициента HOA, которые включены в множество ContAddHoaCoeff. Это множество включает в себя все индексы коэффициентов HOA AmbCoeffIdx[i], которые имеют AmbCoeffTransitionState == 0. Также нет необходимости добавлять взвешенный V-векторный сигнал, поскольку первоначальная последовательность коэффициентов HOA для этих индексов явно отправляется. Таким образом, V-векторный элемент в традиционном подходе для этих индексов устанавливается равным нулю.

Однако в многоуровневом режиме кодирования множество непрерывных индексов коэффициентов HOA зависит от транспортных каналов, которые являются частью текущего активного уровня. Это означает, что дополнительные индексы коэффициентов HOA, отправленные на более высоком уровне, отсутствуют на более низких уровнях. Тогда предположение, что векторный сигнал не должен вносить вклад в последовательность коэффициентов HOA, является неверным для индексов коэффициентов HOA, которые принадлежат последовательностям коэффициентов HOA, включенным в более высокие уровни. Таким образом, предложено (явно) сигнализировать V-векторные элементы для этих недостающих индексов коэффициентов.

Как следствие предложено определить множество ContAddHoaCoeff для каждого уровня и использовать множество уровня, где V-векторный сигнал добавлен для выбора активных V-векторных элементов (транспортный сигнал V-векторного сигнала принадлежит им). Тем не менее, предложено, чтобы V-векторные данные остались в HOAFrame() и не были перемещены в HOAEnhFrame().

Далее будет описана интеграция в синтаксис битового потока MPEG-H. Соответствующий способ кодирования (например, способ многоуровневого кодирования кадра сжатого представления HOA звука или звукового поля) в соответствии с вариантами осуществления раскрытия будет описан со ссылкой на фиг. 3. Предложенные изменения битового потока MPEG-H 3D будут описаны ниже в приложении.

В режиме многоуровневого кодирования флаг SingleLayer в HOADecoderConfig() не активен (SingleLayer == 0), и количество уровней и их соответствующее количество присвоенных транспортных сигналов HOA определены. В общем случае сжатое представление HOA может содержать множество транспортных сигналов.

В соответствии с этим на этапе S3010 на фиг. 3 множество транспортных сигналов присваивается множеству иерархических уровней. Другими словами, транспортные сигналы распределяются множеству уровней. Можно сказать, что каждый уровень включает в себя соответствующие транспортные сигналы, присвоенные этому уровню. Каждый уровень может иметь более чем один транспортный сигнал, присвоенный ему. Множество уровней может включать в себя базовый уровень и один или более иерархических улучшающих уровней. Уровни могут быть упорядочены от базового уровня, через улучшающие уровни, вплоть до полного наиболее высокого улучшающего уровня (полного наиболее высокого уровня).

Предложено добавить дополнительную полезную нагрузку расширения конфигурации HOA и полезную нагрузку расширения кадра HOA с новым определенным типом usacExtElementType ID_EXT_ELE_HOA_ENH_LAYER в битовый поток MPEG-H, чтобы передавать одну полезную нагрузку пространственного предсказания сигналов, синтеза направленных подполосных сигналов и данных декодера PAR для каждого улучшающего уровня HOA (в том числе базового уровня). Эти дополнительные полезные нагрузки будут непосредственно следовать за полезной нагрузкой типа ID_EXT_ELE_HOA в mpegh3daExtElementConfig() и соответственно в mpegh3daFrame().

Таким образом, предложено переместить в случае SingleLayer == 0 элементы конфигурации для пространственного предсказания сигналов, синтеза направленных подполосных сигналов и декодера PAR из HOADecoderConfig() в новый определенный HOADecoderEnhConfig() и соответственно HOAPredictionInfo(), HOADirectionalPredictionInfo() и HOAParInfo() из HOAFrame() в новый определенный HOAEnhFrame().

В соответствии с этим на этапе S3020 для каждого уровня формируется соответствующая полезная нагрузка расширения HOA. Сформированная полезная нагрузка расширения HOA может включать в се-

бя вспомогательную информацию для параметрического улучшения воссозданного представления HOA, которая может быть получена из транспортных сигналов, присвоенных (например, включенных) соответствующему уровню и любым уровням ниже соответствующего уровня. Как указано выше, полезные нагрузки расширения HOA могут включать в себя элементы битового потока для одного или более из инструмента декодирования с пространственным предсказанием сигналов HOA, инструмента декодирования с синтезом направленных подполосных сигналов HOA и инструмента декодирования с параметрическим дублированием звукового окружения HOA. Кроме того, полезные нагрузки расширения HOA могут иметь тип usacExtElementType ID_EXT_ELE_HOA_ENH_LAYER.

На этапе S3030 сформированные полезные нагрузки расширения HOA присваиваются их соответствующим уровням.

Далее (не показано на фиг. 3) может быть сформирована полезная нагрузка расширения конфигурации HOA, включающая в себя элементы битового потока для конфигурирования инструмента декодирования с пространственным предсказанием сигналов HOA, инструмента декодирования с синтезом направленных подполосных сигналов HOA и/или инструмента декодирования с параметрическим дублированием звукового окружения HOA.

Далее (не показано на фиг. 3) может быть сформирована полезная нагрузка конфигурации декодера HOA, включающая в себя информацию, указывающую присвоения полезных нагрузок расширения HOA множеству уровней.

Далее будет описана передача многоуровневого битового потока (например, битового потока MPEG-H). Поскольку все полезные нагрузки расширения битового потока MPEG-H выровнены по байту и их размеры явно сигнализируются, подразумевается, что флаг elementLengthPresent == 1, декомпрессор может выполнить анализ битового потока MPEG-H, извлечь полезные нагрузки для уровней выше одного и передать их отдельно по разным каналам передачи. Базовый уровень содержит битовый поток MPEG-H (например, состоит из него), исключающий из себя данные для более высоких уровней. Недостающие полезные нагрузки расширения сигнализируются как пустые или не активные. Для полезных нагрузок типа ID_USAC_SCE ID_USAC_CPE и ID_USAC_LFE пустая полезная нагрузка сигнализируется посредством elementLength == 0, причем должен быть установлен elementLengthPresent == 1. Пустая полезная нагрузка типа ID_USAC_EXT может быть сигнализирована посредством установки флага usacExtElementPresent == 0 ("ложный").

В соответствии с этим на этапе S3040 сформированные полезные нагрузки расширения HOA сигнализируются (например, передаются или выдаются) в выдаваемом битовом потоке. В общем случае множество уровней и полезных нагрузок, присвоенных им, сигнализируются (например, передаются или выдаются) в выдаваемом битовом потоке. Кроме того, полезная нагрузка конфигурации декодера HOA и/или полезная нагрузка расширения конфигурации HOA могут быть сигнализированы (например, переданы или выданы) в выдаваемом битовом потоке.

Предполагается, что базовый уровень HOA (индекс уровня, равный одному) передается с наиболее высокой защитой от ошибок и имеет относительно малую битовую скорость. Защита от ошибок для следующих уровней (одного или более улучшающих уровней HOA) постоянно уменьшается в соответствии с увеличивающимся битрейтом улучшающих уровней. Вследствие плохих условий передачи и более низкой защиты от ошибок передача более высоких уровней может потерпеть неудачу, и в наихудшем случае корректно передается только базовый уровень. Предполагается, что применяется объединенная защита от ошибок для всех полезных нагрузок одного уровня. Таким образом, если передача уровня терпит неудачу, все полезные нагрузки соответствующего уровня отсутствуют.

Другими словами, полезные нагрузки данных для множества уровней могут быть переданы с соответствующими уровнями защиты от ошибок, причем базовый уровень имеет наиболее высокую защиту от ошибок и один или более улучшающих уровней имеют последовательно уменьшающуюся защиту от ошибок.

Если этапы не требуют некоторых других этапов в качестве предварительных условий, упомянутые выше этапы могут выполняться в любом порядке, и предполагается, что иллюстративный порядок, показанный на фиг. 3, не имеет ограничительного характера.

Как указано выше, синтаксис битового потока для закодированных V-векторных элементов для основанных на векторах сигналов должен быть адаптирован к многоуровневому кодированию HOA, если в HOADecoderConfig() сигнализирован CodedVVecLength == 1. Соответствующий способ кодирования (например, способ многоуровневого кодирования кадра сжатого представления HOA звука или звукового поля) в соответствии с вариантами осуществления раскрытия будет описан со ссылкой на фиг. 4.

На этапе S4010 на фиг. 4 множество транспортных сигналов присваиваются множеству иерархических уровней. Этот этап может быть выполнен таким же образом, как описанный выше этап S3010.

На этапе S4020 определяется, активен ли векторный режим кодирования. Это может включать в себя определение, действительно ли CodedVVecLength == 1.

Как указано выше, в традиционном подходе в векторном режиме кодирования V-векторные элементы не передаются для индексов коэффициента HOA, которые включены в множество ContAddHoaCoeff. Это множество включает в себя все индексы коэффициентов HOA

AmbCoeffIdx[i], которые имеют AmbCoeffTransitionState == 0. Также нет необходимости добавлять взвешенный V-векторный сигнал, поскольку первоначальная последовательность коэффициентов HOA для этих индексов явно отправляется. Таким образом, V-векторный элемент в традиционном подходе для этих индексов устанавливается равным нулю.

Однако в многоуровневом режиме кодирования множество непрерывных индексов коэффициентов HOA зависит от транспортных каналов, которые являются частью текущего активного уровня. Это означает, что дополнительные индексы коэффициентов HOA, отправленные на более высоком уровне, отсутствуют на более низких уровнях. Тогда предположение, что векторный сигнал не должен вносить вклад в последовательность коэффициентов HOA, является неверным для индексов коэффициентов HOA, которые принадлежат последовательностям коэффициентов HOA, включенным в более высокие уровни.

Таким образом, если векторный режим кодирования является активным, на этапе S4030 множество непрерывных индексов коэффициентов HOA (например, ContAddHoaCoeff) определяется (например, задается) для каждого уровня на основе транспортных сигналов, присвоенных соответствующему уровню.

Если векторный режим кодирования является активным, на этапе S4040 для каждого транспортного сигнала формируется V-вектор на основе определенного множества непрерывных индексов коэффициентов HOA для уровня, которому присвоен соответствующий транспортный сигнал. Каждый сформированный V-вектор может включать в себя элементы для любых транспортных сигналов, присвоенных уровням выше, чем уровень, которому присвоен соответствующий транспортный сигнал. Этот этап может включать в себя использование множества непрерывных индексов коэффициентов HOA, которое было определено для уровня, на котором добавлен V-векторный сигнал (уровня, которому принадлежит транспортный сигнал V-векторного сигнала), для выбора активных V-векторных элементов. Тем не менее предлагается, чтобы V-векторные данные оставались в HOAFrame() и не были перемещены в HOAEnhFrame().

Затем на этапе S4050 сформированные V-векторы (V-векторные сигналы) сигнализируются в выдаваемом битовом потоке. Это может включать в себя (явную) сигнализацию V-векторных элементов для упомянутых выше недостающих индексов коэффициентов.

Этапы S4020-S4050 на фиг. 4 также могут использоваться в контексте метода кодирования, проиллюстрированного на фиг. 3, например после этапа S3010. В этом случае этапы S3040 и S4050 могут быть объединены в один этап сигнализации.

Если этапы не требуют некоторых других этапов в качестве предварительных условий, упомянутые выше этапы могут выполняться в любом порядке, и предполагается, что иллюстративный порядок, показанный на фиг. 4, не имеет ограничительного характера.

На стороне приемника упаковщик битового потока MPEG-H может повторно вставить корректно принятые полезные нагрузки в базовый уровень битового потока MPEG-H и передать его к аудиодекодеру MPEG-H 3D.

Далее будет описана инициализация декодирования (конфигурация) HOA. Полезные нагрузки конфигурации HOA типа ID_EXT_ELE_HOA и ID_EXT_ELE_HOA_ENH_LAYER с их соответствующими размерами в байте вводятся в декодер HOA для его инициализации. Инструменты кодирование HOA конфигурируются в соответствии с элементами битового потока, определенными в HOAConfig(), которые анализируются из полезной нагрузки типа ID_EXT_ELE_HOA. Кроме того, эта полезная нагрузка содержит использование режима многоуровневого кодирования, количество уровней и соответствующее количество транспортных сигналов на уровень. Затем, если многоуровневое кодирование активировано (SingleLayer == 0), анализируются элементы HOAEnhConfig() из полезных нагрузок типа ID_EXT_ELE_HOA_ENH_LAYER, чтобы сконфигурировать соответствующие инструменты пространственного предсказания сигналов, синтеза направленных подполосных сигналов и декодера параметрического дублирования звукового окружения каждого уровня.

Элемент LayerIdx из HOAEnhConfig() вместе с порядком полезных нагрузок конфигурации улучшающего уровня HOA в mpegh3daExtElementConfig() указывают порядок улучшающих уровней HOA. Порядок полезных нагрузок кадра улучшающего уровня HOA типа ID_EXT_ELE_HOA_ENH_LAYER в mpegh3daFrame() идентичен порядку полезных нагрузок конфигурации в mpegh3daExtElementConfig(), чтобы явно присвоить полезные нагрузки кадра соответствующим уровням.

В случае SingleLayer == 1 (одноуровневое кодирование) полезные нагрузки типа ID_EXT_ELE_HOA_ENH_LAYER игнорируются, и пространственное предсказание сигналов, синтез направленных подполосных сигналов и декодер параметрического дублирования звукового окружения используют соответствующие данные из HOADecoderConfig() для своей конфигурации.

Далее будет описано декодирование кадра HOA в многоуровневом режиме. Соответствующий способ декодирования (например, способ декодирования кадра сжатого представления HOA звука или звукового поля) в соответствии с вариантами осуществления раскрытия будет описан со ссылкой на фиг. 5. Подразумевается, что сжатое представление HOA (например, выходные данные описанных выше способов на фиг. 3 или фиг. 4) было закодировано в множестве иерархических уровней, в том числе в базовом уровне и одном или более улучшающих уровнях.

На этапе S5010 на фиг. 5 принимается битовый поток, относящийся к кадру сжатого представ-

ления HOA.

Базовый декодер 3D-аудио декодирует корректно переданные транспортные сигналы HOA и создает транспортные сигналы со всеми отсчетами, равными нулю, для соответствующих некорректных полезных нагрузок. Декодированные транспортные сигналы вместе с флагами usacExtElementPresent, данными и размерами полезных нагрузок HOA типа ID_EXT_ELE_HOA и ID_EXT_ELE_HOA_ENH_LAYER вводятся в декодеру HOA. Полезные нагрузки расширения из типа ID_USAC_EXT с флагом usacExtElementPresent, установленным как "ложный", должны быть сигнализированы как недостающие полезные нагрузки декодеру HOA, чтобы гарантировать присвоение полезных нагрузок соответствующим уровням.

На этапе S5020 извлекаются полезные нагрузки для множества уровней. Каждая полезная нагрузка может включать в себя транспортные сигналы, присвоенные соответствующему уровню.

На этом этапе декодер HOA может проанализировать HOAFrame() из полезной нагрузки типа ID_EXT_ELE_HOA.

Затем корректные полезные нагрузки типа ID_EXT_ELE_HOA_ENH_LAYER и некорректные полезные нагрузки типа ID_EXT_ELE_HOA_ENH_LAYER определяются посредством, оценки соответствующего флага usacExtElementPresent полезных нагрузок, где некорректная полезная нагрузка обозначена флагом usacExtElementPresent со значением "ложный" и присвоение полезных нагрузок улучшения HOA индексам улучшающих уровней известно из конфигурации декодера HOA.

На этапе S5030 определяется наиболее высокий применимый уровень среди множества уровней для декодирования.

Поскольку уровни зависят друг от друга с точки зрения транспортных сигналов, декодер HOA может декодировать уровень только тогда, когда корректно приняты все уровни с более низким индексом. Наиболее высокий применимый уровень может быть выбран на этом этапе таким образом, чтобы все уровни вплоть до наиболее высокого применимого уровня были корректно приняты. Подробности этого этапа будут описаны ниже.

На этапе S5040 извлекается полезная нагрузка расширения HOA, присвоенная наиболее высокому применимому уровню. Как указано выше, полезная нагрузка расширения HOA может включать в себя вспомогательную информацию для параметрического улучшения воссозданного представления HOA, соответствующего наиболее высокому применимому уровню. При этом воссозданное представление HOA, соответствующее наиболее высокому применимому уровню, может быть получено на основе транспортных сигналов, присвоенных наиболее высокому применимому уровню и любым уровням ниже наиболее высокого применимого уровня.

Кроме того, могут быть извлечены полезные нагрузки расширения HOA, соответственно присвоенные оставшимся из множества уровней. Каждая полезная нагрузка расширения HOA может включать в себя вспомогательную информацию для параметрического улучшения воссозданного представления HOA, связанного с его соответствующим присвоенным уровнем. Воссозданное представление HOA, связанное с его соответствующим присвоенным уровнем, может быть получено из транспортных сигналов, присвоенных этому уровню и любым уровням ниже этого уровня.

Далее (не показано на фиг. 5) способ декодирования может содержать этап извлечения полезной нагрузки расширения конфигурации HOA. Это может быть сделано посредством анализа битового потока. Полезная нагрузка расширения конфигурации HOA может включать в себя элементы битового потока для конфигурирования инструмента декодирования с пространственным предсказанием сигналов HOA, инструмента декодирования с синтезом направленных подполосных сигналов и/или инструмента декодирования с параметрическим дублированием звукового окружения HOA.

На этапе S5050 (частично) воссозданное представление HOA, соответствующее наиболее высокому применимому уровню, формируется на основе транспортных сигналов, присвоенных наиболее высокому применимому уровню и любым уровням ниже наиболее высокого применимого уровня.

Количество фактически используемых транспортных сигналов $I_{ADD,LAY}(k)$ устанавливается в соответствии с (индексом $M_{LAY}(k)$) наиболее высоким применимым уровнем, и первое предварительное представление HOA декодируется из HOAFrame() и из соответствующих транспортных сигналов уровня и любых более низких уровней.

Затем на этапе S5060 воссозданное представление HOA улучшается (например, параметрически улучшается) с использованием вспомогательной информации, включенной в полезную нагрузку расширения HOA, присвоенную наиболее высокому применимому уровню.

Таким образом, представление HOA, полученное на этапе S5050, затем улучшается посредством пространственного

предсказания сигналов, синтеза направленных подполосных сигналов и декодера параметрического дублирования звукового окружения с использованием данных HOAEnhFrame(), проанализированных из полезной нагрузки расширения улучшающего уровня HOA типа ID_EXT_ELE_HOA_ENH_LAYER на текущем активном уровне $M_{LAY}(k)$, т.е., на наиболее высоком применимом уровне.

Информация, используемая на этапах S5020-S5060, может быть известна как информация уровня.

Если этапы не требуют некоторых других этапов в качестве предварительных условий, упомянутые

выше этапы могут выполняться в любом порядке, и предполагается, что иллюстративный порядок, показанный на фиг. 5, не имеет ограничительного характера.

Далее будут описаны подробности определения (например, выбора) наиболее высокого применимого уровня на этапе S5030.

Как указано выше, декодер HOA может декодировать уровень только тогда, когда корректно приняты все уровни с более низким индексом, поскольку уровни зависят друг от друга с точки зрения транспортных сигналов.

Для выбора наиболее высокого декодируемого уровня декодер HOA может создать множество индексов некорректных уровней, причем самый малый индекс из этого множества минус один дает в результате индекс M_{LAY} наиболее высокого декодируемогое улучшающего уровня. Множество индексов некорректных уровней может быть определено посредством оценки флагов корректности соответствующих полезных нагрузок расширения HOA.

Другими словами, определение наиболее высокого применимого уровня может включать в себя определение множества индексов некорректных уровней, указывающего уровни, которые не были приняты корректно. Это может дополнительно включать в себя определение наиболее высокого применимого уровня как уровня, который является уровнем ниже уровня, указанного наименьшим индексом в множестве индексов некорректных уровней. Тем самым гарантируется, что все уровни ниже наиболее высокого применимого уровня были приняты корректно.

В случае дифференциального кодирования кадров будет принят во внимание индекс наиболее высокого применимого уровня предыдущего (например, непосредственно предшествующего) кадра. Сначала будет описана ситуация, в которой сохраняется индекс наиболее высокого применимого уровня предыдущего (например, предшествующего) кадра.

Если индекс наиболее высокого применимого уровня (например, наиболее высокого декодируемого уровня) для текущего кадра равен индексу уровня предыдущего кадра $M_{LAY}(k)$-1, индекс уровня текущего кадра $M_{LAY}(k)$ устанавливается равным $M_{LAY}(k)$-1.

Затем количество фактически используемых транспортных сигналов $I_{ADD,LAY}(k)$ устанавливается в соответствии с $M_{LAY}(k)$, и первое предварительное представление HOA декодируется из HOAFrame() и из соответствующих транспортных сигналов уровня и любых более низких уровней, как указано выше. Это представление HOA затем улучшается посредством пространственного предсказания сигналов, синтеза направленных подполосных сигналов и декодера параметрического дублирования звукового окружения с использованием данных HOAEnhFrame(), проанализированных из полезной нагрузки расширения улучшающего уровня HOA типа ID_EXT_ELE_HOA_ENH_LAYER текущего активного уровня, $M_{LAY}(k)$, как указано выше.

Далее будет описана ситуация, в которой переключаются на индекс ниже, чем индекс наиболее высокого применимого уровня предыдущего (например, предшествующего) кадра. А именно в случае когда индекс наиболее высокого декодируемого уровня для текущего кадра меньше, чем индекс уровня предыдущего кадра $M_{LAY}(k)$-1, декодер HOA устанавливает $M_{LAY}(k)$ равным индексу наиболее высокого декодируемого уровня для текущего кадра. Декодирование полезных нагрузок для пространственного предсказания сигналов, синтеза направленных подполосных сигналов и декодера параметрического дублирования звукового окружения для нового уровня может начаться только в следующем кадре HOA с флагом hoaIndependencyFlag $==1$. Пока не будет принят такой HOAFrame(), представление HOA уровня с индексом $M_{LAY}(k)$ воссоздается без выполнения пространственного предсказания сигналов, синтеза направленных подполосных сигналов и декодера параметрического дублирования звукового окружения. Это означает, что количество фактически используемых транспортных сигналов $I_{ADD,LAY}(k)$ устанавливается в соответствии с $M_{LAY}(k)$, и только первое предварительное представление HOA декодируется из HOAFrame() и из соответствующих транспортных сигналов уровня и любых более низких уровней. Затем, если был принят HOAFrame() с флагом hoaIndependencyFlag $==1$, полезные нагрузки для пространственного предсказания сигналов, синтеза направленных подполосных сигналов и декодера параметрического дублирования звукового окружения анализируются и декодируются для улучшения предварительного представление HOA, и, таким образом, обеспечивается полное качество текущего активного уровня для этого кадра.

Таким образом, предложенный способ может содержать (не показано на фиг. 5) принятие решения не выполнять параметрическое улучшение воссозданного представления HOA с использованием вспомогательной информации, включенной в полезную нагрузку расширения HOA, присвоенную наиболее высокому применимому уровню, если наиболее высокий применимый уровень текущего кадра ниже, чем наиболее высокий применимый уровень предыдущего кадра (если текущий кадр был закодирован дифференциально относительно предыдущего кадра).

В общем случае определение наиболее высокого применимого уровня для текущего кадра может включать в себя определение множества индексов некорректных уровней, указывающего уровни, которые не были корректно приняты для текущего кадра. Это может дополнительно содержать определение наиболее высокого применимого уровня предыдущего кадра, предшествующего текущему кадру. Это может, кроме того, содержать определение наиболее высокого применимого уровня как более низкого из

наиболее высокого применимого уровня предыдущего кадра и уровня, который является уровнем ниже уровня, указанного самым малым индексом в множестве индексов некорректных уровней (если текущий кадр был закодирован дифференциально относительно предыдущего кадра).

Альтернативное решение может всегда анализировать все корректные полезные нагрузки улучшающего уровня (например, полезные нагрузки расширения HOA) параллельно, даже если они в настоящее время не активны. Это позволит напрямую переключиться на уровень с более низким индексом с полным качеством, причем пространственное предсказание сигналов, синтез направленных подполосных сигналов и декодер параметрического дублирования звукового окружения (PAR) могут быть непосредственно применены в переключенном кадре.

Далее будет описана ситуация, в которой переключаются на индекс выше, чем индекс наиболее высокого применимого уровня предыдущего (например, предшествующего) кадра. Это переключение на уровень с более высоким индексом может быть применено, только если mpegh3daFrame() имеет флаг usacIndependencyFlag == 1 (например, если кадр является независимым кадром), поскольку все соответствующие полезные нагрузки или состояния декодирования предыдущих кадров отсутствуют. Таким образом, декодер HOA сохраняет индекс уровня HOA $M_{LAY}(k)$ равным $M_{LAY}(k)$-l, пока не будет принят mpegh3daFrame() с флагом usacIndependencyFlag == 1 (например, независимый кадр), который содержит корректные данные для более высокого декодируемого уровня. Затем $M_{LAY}(k)$ устанавливается равным индексу наиболее высокого декодируемого уровня для текущего кадра и соответственно определяется количество фактически используемых транспортных сигналов $I_{ADD,LAY}(k)$. Предварительное представление HOA этого уровня декодируется из HOAFrame() и соответствующих транспортных сигналов и улучшается посредством пространственного предсказания сигналов, синтеза направленных подполосных сигналов и декодером параметрического дублирования звукового окружения с использованием HOAEnhFrame(), проанализированного из полезной нагрузки расширения улучшающего уровня HOA типа ID_EXT_ELE_HOA_ENH_LAYER текущего активного уровня $M_{LAY}(k)$.

Подразумевается, что предложенный способ многоуровневого кодирования сжатого представления звука может быть реализован кодером для многоуровневого кодирования сжатого представления звука. Такой кодер может содержать соответствующие блоки, выполненные с возможностью выполнять соответствующие описанные выше этапы. Пример такого кодера 6000 схематично проиллюстрирован на фиг. 6. Например, такой кодер 6000 может содержать блок 6010 присвоения транспортных сигналов, выполненный с возможностью выполнять упомянутый выше этап S3010, блок 6020 формирования полезной нагрузки уровня расширения HOA, выполненный с возможностью выполнять упомянутый выше этап S3020, блок 6030 присвоения полезной нагрузки расширения, выполненный с возможностью выполнять упомянутый выше этап S3030, и блок сигнализации или блок 6040 вывода, выполненный с возможностью выполнять упомянутый выше этап S3040. Далее подразумевается, что соответствующие блоки такого кодера могут быть воплощены посредством процессора 6100 вычислительного устройства, который выполнен с возможностью выполнять обработку, выполняемую каждым из упомянутых соответствующих блоков, т.е. выполнен с возможностью выполнять некоторые или все из упомянутых выше этапов предложенного способа кодирования, схематично проиллюстрированного на фиг. 3. Дополнительно или в качестве альтернативы процессор 6100 может быть выполнен с возможностью выполнять каждый из этапов способа кодирования, схематично проиллюстрированного в фиг. 4. С этой целью процессор 6100 может быть выполнен с возможностью реализовывать соответствующие блоки кодера. Кодер или вычислительное устройство могут дополнительно содержать память 6200, к которой может осуществлять доступ процессор 6100.

Далее подразумевается, что предложенный способ декодирования сжатого представления звука, которое закодировано в множестве иерархических уровней, может быть реализован декодером для декодирования сжатого представления звука, которое закодировано в множестве иерархических уровней. Такой декодер может содержать соответствующие блоки, выполненные с возможностью выполнять соответствующие описанные выше этапы. Пример такого декодера 7000 схематично проиллюстрирован на фиг. 7. Например, такой декодер 7000 может содержать блок 7010 приема, выполненный с возможностью выполнять упомянутый выше этап S5010, блок 7020 извлечения полезной нагрузки, выполненный с возможностью выполнять упомянутый выше этап S5020, блок 7030 определения наиболее высокого применимого уровня, выполненный с возможностью выполнять упомянутый выше этап S5030, блок 7040 извлечения полезной нагрузки расширения HOA, выполненный с возможностью выполнять упомянутый выше этап S5040, блок 7050 формирования воссозданного представления HOA, выполненный с возможностью выполнять упомянутый выше этап S5050, и блок 7060 улучшения, выполненный с возможностью выполнять упомянутый выше этап S5060. Далее подразумевается, что соответствующие блоки такого декодера могут быть воплощены посредством процессора 7100 вычислительного устройства, который выполнен с возможностью выполнять обработку, выполняемую каждым из упомянутых соответствующих блоков, т.е. выполнен с возможностью выполнять некоторые или все из упомянутых выше этапов предложенного способа декодирования. Декодер или вычислительное устройство могут дополнительно содержать память 7200, к которой может осуществлять доступ процессор 7100.

Далее будет описана структура данных (например, битовый поток) для размещения (например,

представления) сжатого представления HOA в режиме многоуровневого кодирования. Такая структура данных может являться результатом использования предложенных способов кодирования и может быть декодирована (например, восстановлена), посредством использования предложенного способа декодирования.

Структура данных может содержать множество полезных нагрузок кадра HOA, связанных с соответствующими множествами иерархических уровней. Множество транспортных сигналов может быть присвоено (например, может принадлежать) соответствующим уровням из множества уровней. Структура данных может содержать соответствующую полезную нагрузку расширения HOA, включающую в себя вспомогательную информацию для параметрического улучшения воссозданного представления HOA, которая может быть получена из транспортных сигналов, присвоенных соответствующему уровню и любым уровням ниже соответствующего уровня. Полезные нагрузки кадра HOA и полезные нагрузки расширения HOA для множества уровней могут быть снабжены соответствующими уровнями защиты от ошибок, как указано выше. Кроме того, полезные нагрузки расширения HOA могут содержать указанные выше элементы битового потока и могут иметь тип usacExtElementType ID_EXT_ELE_HOA_ENH_LAYER. Кроме того, структура данных может содержать полезную нагрузку расширения конфигурации HOA и/или полезную нагрузку конфигурации декодера HOA, включающую в себя указанные выше элементы битового потока.

Следует отметить, что описание и чертежи лишь иллюстрируют принципы предложенных способов и устройств. Таким образом, будет очевидно, что специалисты в области техники смогут создавать различные структуры, которые, хотя явно не описаны и не показаны в настоящем документе, воплощают принципы изобретения и включены в пределы его сущности и объема. Кроме того, все примеры, приведенные в настоящем документе, преимущественно явно предназначены лишь для обучения, чтобы помочь читателю в понимании принципов предложенных способов,и устройств, и концепций, внесенных изобретателями в развитие области техники, и должны быть истолкованы как не являющиеся ограничениями для таких специальным образом приведенных примеров и условий. Кроме того, предполагается, что все утверждения в настоящем документе, излагающие принципы, аспекты и варианты осуществления изобретения, а также их конкретные примеры, охватывают его эквиваленты.

Способы и устройство, описанные в настоящем документе, могут быть реализованы как программное обеспечение, программно-аппаратное обеспечение и/или аппаратные средства. Некоторые компоненты, например, могут быть реализованы как программное обеспечение, работающее на процессоре цифровой обработки сигналов или микропроцессоре. Другие компоненты, например, могут быть реализованы как аппаратные средства и/или как специализированные интегральные схемы. Сигналы, встречающиеся в описанных способах и устройстве, могут быть сохранены на носителях, таких как оперативное запоминающее устройство или оптические запоминающие носители. Они могут быть перенесены через сети, такие как радиосети, спутниковые сети, беспроводные сети или проводные сети, например Интернет.

Приложение

Предложенные изменения битового потока MPEG-H 3D

Изменения помечены посредством выделения серым цветом.

Таблица 1. Синтаксис mpegh3daExtElementConfig()

Синтаксис	К-во бит	Обозначение
mpegh3daExtElementConfig()		
{		
usacExtElementType = escapedValue(4, 8, 16);		
usacExtElementConfigLength = escapedValue(4, 8, 16);		
if (**usacExtElementDefaultLengthPresent**) {	1	uimsbf
usacExtElementDefaultLength = escapedValue(8, 16, 0) + 1;		
} else {		
usacExtElementDefaultLength = 0;		
}		
usacExtElementPayloadFrag;	1	uimsbf
switch (usacExtElementType) {		
case ID_EXT_ELE_FILL:		
/* Нет элемента конфигурации */		
break;		
case ID_EXT_ELE_MPEGS:		
SpatialSpecificConfig();		
break;		
case ID_EXT_ELE_SAOC:		
SAOCSpecificConfig();		
break;		
case ID_EXT_ELE_AUDIOPREROLL:		
/* Нет элемента конфигурации */		
break;		
case ID_EXT_ELE_UNI_DRC:		
mpegh3daUniDrcConfig();		
break;		
case ID_EXT_ELE_OBJ_METADATA:		
ObjectMetadataConfig();		
break;		
case ID_EXT_ELE_SAOC_3D:		
SAOC3DSpecificConfig();		
break;		
case ID_EXT_ELE_HOA:		
HOAConfig();		

```
        break;
    case ID_EXT_ELE_HOA_ENH_LAYER:
        HOAEnhConfig();
        break;
    case ID_EXT_ELE_FMT_CNVRTR
        /* Нет элемента конфигурации */
        break;
    default:                                                    ПРИМ
        while (usacExtElementConfigLength–) {
            tmp;                                                8          uimsbf
        }
        break;
    }
}
```

Примечание: запись по умолчанию для usacExtElementType используется для неизвестных extElementTypes, чтобы декодеры прежних версий могли справляться с будущими расширениями

Таблица 2. Значение usacExtElementType

usacExtElementType	Значение
ID_EXT_ELE_FILL	0
ID_EXT_ELE_MPEGS	1
ID_EXT_ELE_SAOC	2
ID_EXT_ELE_AUDIOPREROLL	3
ID_EXT_ELE_UNI_DRC	4
ID_EXT_ELE_OBJ_METADATA	5
ID_EXT_ELE_SAOC_3D	6
ID_EXT_ELE_HOA	7
ID_EXT_ELE_FMT_CNVRTR	8
ID_EXT_ELE_HOA_ENH_LAYER	9
/* зарезервировано для ISO */	10-127
/* зарезервировано для использования, кроме ISO */	128 и больше

Примечание: специфические для приложений значений usacExtElementType должны находиться в пространстве, зарезервированном для использования, кроме ISO. Они пропускаются декодером, поскольку минимальная структура требуется декодеру, чтобы пропустить эти расширения

Таблица 3. Интерпретация блоков данных для декодирования полезной нагрузки расширения

usacExtElementType	Соединенный usacExtElementSegmentData представляет:
ID_EXT_ELE_FILL	Последовательность fill_byte
ID_EXT_ELE_MPEGS	SpatialFrame(), см. ISO/IEC 23003-1
ID_EXT_ELE_SAOC	SAOCFrame(), см. ISO/IEC 23003-2
ID_EXT_ELE_AUDIOPREROLL	AudioPreRoll()
ID_EXT_ELE_UNI_DRC	uniDrcGain(), см. ISO/IEC 23003-4
ID_EXT_ELE_OBJ_METADATA	object_metadata()
ID_EXT_ELE_SAOC_3D	Saoc3DFrame()
ID_EXT_ELE_HOA	HOAFrame()
ID_EXT_ELE_HOA_ENH_LAYER	HOAEnhFrame()
ID_EXT_ELE_FMT_CNVRTR	FormatConverterFrame()
unknown	Неизвестные данные. Блок данных будет отброшен.

Таблица 4. Синтаксис HOADecoderConfig()

Синткасис	К-во бит	Обозначение
HOADecoderConfig(numHOATransportChannels)		
{		
MinAmbHoaOrder = escapedValue(3,5,0) – 1;	3,8	uimsbf
MinNumOfCoeffsForAmbHOA = (MinAmbHoaOrder + 1)^2;		
NumOfAdditionalCoders =		
numHOATransportChannels - MinNumOfCoeffsForAmbHOA;		
if(SingleLayer == 0){	1	bslbf
HOALayerChBits = ceil(log2(NumOfAdditionalCoders));		
NumHOAChannelsLayer[0] = codedLayerCh + MinNumOfCoeffsForAmbHOA;	HOALayerChBits	uimsbf
remainingCh = numHOATransportChannels – NumHOAChannelsLayer[0];		
NumLayers = 1;		
while (remainingCh>1) {		
HOALayerChBits = ceil(log2(remainingCh));		
NumHOAChannelsLayer[NumLayers] = NumHOAChannelsLayer[NumLayers-1] + codedLayerCh + 1;	HOALayerChBits	uimsbf
remainingCh = remainingCh – NumHOAChannelsLayer[NumLayers];		
NumLayers++;		
}		
if (remainingCh) {		
NumHOAChannelsLayer[NumLayers] = NumOfAdditionalCoders;		
NumLayers++;		
}		
}		
~~MaxNoOfDirSigsForPrediction = MaxNoOfDirSigsForPrediction + 1;~~	~~2~~	~~uimsbf~~
~~NoOfBitsPerScalefactor = NoOfBitsPerScalefactor + 1;~~	~~4~~	~~uimsbf~~
CodedSpatialInterpolationTime;	3	uimsbf
SpatialInterpolationMethod;	1	bslbf
CodedVVecLength;	2	uimsbf
MaxGainCorrAmpExp;	3	uimsbf
HOAFrameLengthIndicator;	2	uimsbf

```
    MaxHOAOrderToBeTransmitted = escapedValue(2,5,0) +
                                        MinAmbHoaOrder;
    MaxNumOfCoeffsToBeTransmitted =
                        (MaxHOAOrderToBeTransmitted + 1)^2;
    MaxNumAddActiveAmbCoeffs =
                        MaxNumOfCoeffsToBeTransmitted
                         - MinNumOfCoeffsForAmbHOA;
    VqConfBits = ceil( log2( ceil( log2(NumOfHoaCoeffs))))
    NumVVecVqElementsBits++;                          VqCon   uimsbf
                                                      fBits
    if( MinAmbHoaOrder == 1) {
        UsePhaseShiftDecorr;                          1       bslbf
    }
    if(SingleLayer==1) {
        HOADecoderEnhConfig();                        2       uimsbf
    }
    AmbAsignmBits = ceil( log2( MaxNumAddActiveAmbCoeffs ) );
    ActivePredIdsBits = ceil( log2( NumOfHoaCoeffs ) );
    i = 1;
    while( i * ActivePredIdsBits + ceil( log2( i ) ) < NumOfHoaCoeffs ){
        i++;
    }
    NumActivePredIdsBits = ceil( log2( max( 1, i – 1 ) ) );
    GainCorrPrevAmpExpBits = ceil( log2( ceil( log2(
                                        1.5 * NumOfHoaCoeffs ) )
                                      + MaxGainCorrAmpExp + 1 ) );
    for (i=0; i<NumOfAdditionalCoders; ++i){
        AmbCoeffTransitionState[i] = 3;
    }
}
```

Примечание: MinAmbHoaOrder=30 ... 37 зарезервированы. HOAFrameLengthIndicator=3 зарезервировано

Новая таблица. Синтаксис HOAEnhConfig()

Синткасис	К-во бит	Обозначение
HOAEnhConfig() { LayerIdx	HOALayerChBits	uimsbf
HOADecoderEnhConfig(); }		

Новая таблица. Синтаксис HOADecoderEnhConfig()

Синткасис	К-во бит	Обозначение
HOADecoderEnhConfig()		
{		
MaxNoOfDirSigsForPrediction = **MaxNoOfDirSigsForPrediction** + 1;	2	**uimsbf**
NoOfBitsPerScalefactor = **NoOfBitsPerScalefactor** + 1;	**4**	**uimsbf**
if(PredSubbandsIdx < 3) {	2	uimsbf
NumOfPredSubbands = NumOfPredSubbandsTable[PredSubbandsIdx];		
PredSubbandWidths = PredSubbandWidthTable[PredSubbandsIdx];		
}		
else {		
CodedNumberOfSubbands	5	uimsbf
NumOfPredSubbands = CodedNumberOfSubbands+1;		
PredSubbandWidths = getSubbandWidths(NumOfPredSubbands);		
}		
if (NumOfPredSubbands > 0) {		
FirstSBRSubbandIdxBits = ceil(log2 (NumOfPredSubbands+1));		
FirstSBRSubbandIdx;	FirstSBRSubbandIdxBits	**uimsbf**
MaxNumOfPredDirs = 2^(**MaxNumOfPredDirsLog2**);	3	**uimsbf**
MaxNumOfPredDirsPerBand = escapedValue(3,2,5) + 1;		
NumOfBitsPerDirIdx = NumOfBitsPerDirIdxTable[DirGridTableIdx];	2	**uimsbf**
}		
if(ParSubbandTableIdx < 3) {	2	**uimsbf**
NumOfParSubbands = NumOfParSubbandsTable[ParSubbandTableIdx];		
ParSubbandWidths = ParSubbandWidthTable[ParSubbandTableIdx];		
}		

```
    else {
        CodedNumberOfSubbands                                   5       uimsbf
        NumOfParSubbands = CodedNumberOfSubbands+1;
        ParSubbandWidths =
            getSubbandWidths(NumOfParSubbands);
    }
    if( NumOfParSubbands > 0 ) {
        LastFirstOrderSubbandIdxBits =
            ceil( log2(NumOfParSubbands + 1) );
        LastFirstOrderSubbandIdx;                               LastFi  uimsbf
                                                                rstOrd
                                                                erSub
                                                                bandI
                                                                dxBits
        for ( idx = 0; idx < NumOfParSubbands; idx++) {
            UseRealCoeffsPerParSubband[idx];                    1       bslbf
        }
        for ( idx = 0; idx < LastFirstOrderSubBandIdx; idx++) {
            UpmixHoaOrderPerParSubband[idx] = 1;
            MaxNumOfDecoSigs[idx] =
                (UpmixHoaOrderPerParSubband[idx] + 1)^2;
        }
        for ( idx = LastFirstOrderSubBandIdx;
            idx < NumOfParSubbands; idx++) {
            UpmixHoaOrderPerParSubband[idx] = 2;
            MaxNumOfDecoSigs[idx] =
                (UpmixHoaOrderPerParSubband[idx] + 1)^2;
        }
    }
}
```

Таблица 5. Синтаксис HOAFrame

Синткасис	К-во бит	Обозначение
HOAFrame()		
{		
NumOfDirSigs = 0;		
for(lay=0; (lay< NumLayers) & !SingleLayer; ++lay){		
NumOfDirSigsPerLayer[lay] = 0;		
NumOfAddHoaChansPerLayer[lay] = 0;		
NumOfContAddHoaChans[lay] = 0;		
}		
NumOfVecSigs = 0;		
NumOfAddHoaChans = 0;		
NumOfAddVVecValCoeffIdx = 0;		
hoaIndependencyFlag;	1	**bslbf**
for(i=0; i< NumOfAdditionalCoders; ++i){		
ChannelSideInfoData(i);		
HOAGainCorrectionData(i);		
switch ChannelType[i] {		
case 0:		
DirSigChannelIds[NumOfDirSigs] = i + 1;		
NumOfDirSigs++;		
for(lay=0; (lay< NumLayers) & !SingleLayer; ++lay){		
if((MinNumOfCoeffsForAmbHOA + i) <		
NumHOAChannelsLayer[lay]){		
NumOfDirSigsPerLayer[lay]++;		
}		
}		
break;		
case 1:		
VecSigChannelIds[NumOfVecSigs] = i + 1;		
VecSigLayerIdx[NumOfVecSigs] = 0;		
if (SingleLayer == 0) {		
lay = 0;		
while((MinNumOfCoeffsForAmbHOA + i)		
≥ NumHOAChannelsLayer[lay]) {		
lay ++;		
}		
VecSigLayerIdx[NumOfVecSigs] = lay;		
}		

```
            NumOfVecSigs++;
            break;
        case 2:
            if (AmbCoeffTransitionState[i] == 0) {
                for(lay=0; (lay< NumLayers); ++lay){
                    if( (MinNumOfCoeffsForAmbHOA + i ) <
                            NumHOAChannelsLayer[lay]){
                        ContAddHoaCoeff[lay]
                            [NumOfContAddHoaChans[lay]] =
                                = AmbCoeffIdx[i];
                        NumOfContAddHoaChans[lay]++;
                    }
                }
                AddHoaCoeff[NumOfAddHoaChans] = AmbCoeffIdx[i];
                for(lay=0; (lay< NumLayers) & (SingleLayer == 0); ++lay){
                    if( (MinNumOfCoeffsForAmbHOA + i ) <
                            NumHOAChannelsLayer[lay]){
                        AddHoaCoeffPerLayer[lay][NumOfAddHoaChans]
                            = AmbCoeffIdx[i];
                        NumOfAddHoaChansPerLayer[lay]++;
                    }
                }
                NumOfAddHoaChans++;
                break;
            }
    }

    for ( i= NumOfAdditionalCoders;
            i< NumHOATransportChannels; ++i){
        HOAGainCorrectionData(i);
    }

    for(i=0; i< NumOfVecSigs; ++i){
        VVectorData ( VecSigChannelIds(i) );
    }

    if(SingleLayer==1) {
        HOAEnhFrame();
    }
}
```

Примечание: кодер должен установить hoaIndependencyFlag равным 1, если usacIndependency-Flag (см mpegh3daFrame()) установлен равным 1

Примечание: если SingleLayer == 1, установить NumLayers=1

NumOfDirSigsPerLayer[lay] - эти элементы определяют количество активных направленных сигналов в текущем HOAFrame(), фактически используемых на улучшающем уровне lay HOA.

AddHoaCoeffPerLayer[lay] - этот массив содержит индекс коэффициента HOA для каждого дополнительного коэффициента окружения HOA, фактически используемого на улучшающем уровне lay HOA.

NumOfAddHoaChansPerLayer[lay] - этот элемент сигнализирует общее количество дополнительных коэффициентов окружения HOA, фактически используемых на улучшающем уровне lay HOA.

Новая таблица. Синтаксис HOAEnhFrame

Синткасис	К-во бит	Обозначение
HOAEnhFrame()		
{		
if(((SingleLayer==1) & (NumOfDirSigs > 0)) |		
((SingleLayer==0) & (NumOfDirSigsPerLayer[lay]) > 0)){		
HOAPredictionInfo()		
}		
if(NumOfPredSubbands > 0) {		
HOADirectionalPredictionInfo();		
}		
if(NumOfParSubbands > 0) {		
HOAParInfo();		
}		
}		

Примечание: lay - индекс текущего активного улучшающего уровня HOA

Таблица 6. Синтаксис VVectorData ()

Синткасис	К-во бит	Обозначение
VVectorData(i)		
{		
if (CodedVVecLength == 1) {		
VVecLengthUsed = VVecLength[i];		
VVecCoeffIdUsed = VVecCoeffId[i];		
} else {		
VVecLengthUsed = VVecLength;		
VVecCoeffIdUsed = VVecCoeffId;		
}		
if (NbitsQ(k)[i] == 4){		
if (NumVvecIndices(k)[i] == 1) {		
VecIdx[0] = **VecIdx** + 1;	10	uimsbf
WeightVal[0] = ((**SgnVal***2)-1);	1	uimsbf
} else {		
WeightIdx;	8	**uimsbf**
nbitsIdx = ceil(log2(NumOfHoaCoeffs));		
for (j=0; j< NumVvecIndices(k)[i]; ++j) {		
VvecIdx[j] = **VvecIdx** + 1;	**nbitsIdx**	**uimsbf**
WeightVal[j] = ((**SgnVal***2)-1) *	1	**uimsbf**
WeightValCdbk[CodebkIdx(k)[i]][WeightIdx][j];		
}		
}		
}		
else if (NbitsQ(k)[i] == 5) {		
for (m=0; m< VVecLengthUsed; ++m){		
aVal[i][m] = (**VecVal** / 128.0) – 1.0;	8	**uimsbf**
}		
else if(NbitsQ(k)[i] >= 6) {		
for (m=0; m< VVecLengthUsed; ++m){		
huffIdx = *huffSelect*(VVecCoeffIdUsed[m], PFlag[i],		
CbFlag[i]);		
cid = *huffDecode*(NbitsQ[i], huffIdx, **huffVal**);	**dynamic**	**huffDecode**
aVal[i][m] = 0.0;		
if (cid > 0) {		
aVal[i][m] = sgn = (**sgnVal** * 2) - 1;	1	**bslbf**
if (cid > 1) {		
aVal[i][m] = sgn * (2.0^(cid -1) + **intAddVal**);	**cid-1**	**uimsbf**

```
                }
            }
        }
    }
}
```

Примечание: см. 0 для вычисления VVecLength

Таблица 7. Синтаксис HOAPredictionInfo(DirSigChannelIds, NumOfDirSigs)

Синткасис	К-во бит	Обозначение
HOAPredictionInfo()		
{		
if(SingleLayer==1){		
PredIdsBits = ceil(log2(NumOfDirSigs + 1));		
}		
else{		
PredIdsBits = ceil(log2(NumOfDirSigsPerLayer[lay] + 1));		
}		
if(**PSPredictionActive**){	1	bslbf
NumActivePred = 0;		
if(**KindOfCodedPredIds**){	1	bslbf
NumActivePred = **NumActivePredIds** + 1;	**NumActivePredIdsBits**	uimsbf
i=0;		
while(i < NumActivePred){		
PredIds[i] = **PredIds**[i] + 1;	**ActivePredIdsBits**	uimsbf
i++;		
}		
}		
else{		
for (i=0; i<(HoaOrder +1)^2; i++) {		
if(**ActivePred**[i]) {	1	bslbf
NumActivePred ++;		
}		
}		
}		
NumOfGains=0;		
for (i=0; i<NumActivePred * MaxNoOfDirSigsForPrediction; i++) {		
if (**PredDirSigIds**[i] > 0) {	**PredIdsBits**	uimsbf
PredDirSigIds[i] = DirSigChannelIds[PredDirSigIds[i] - 1];		
NumOfGains++;		
}		
}		
n=0;		
for (i=0; i< NumOfGains; i++) {		
if (PredDirSigIds[i]>0) {		

PredGains[n];	NoOfBitsPerScalefactor	bslbf
n++;		
}		
}		
}		
}		

Примечание: lay - индекс текущего активного улучшающего уровня HOA

Таблица AMD1.2. Синтаксис HOADirectionalPredictionInfo()

Синткасис	К-во бит	Обозначение
HOADirectionalPredictionInfo ()		
{		
if(UseDirectionalPrediction) {	1	bslbf
if (!hoaIndependencyFlag) {		
KeepPreviousGlobalPredDirsFlag;	1	bslbf
}		
else{		
KeepPreviousGlobalPredDirsFlag = 0;		
}		
if(!KeepPreviousGlobalPredDirsFlag) {		
NumOfGlobalPredDirs = NumOfGlobalPredDirs + 1;	MaxNumOfPredDirsLog2	bslbf
NumBitsForRelDirGridIdx = ceil(log2(NumOfGlobalPredDirs));		
for (idx=0; idx < NumOfGlobalPredDirs; idx++) {		
GlobalPredDirsIds[idx];	NumOfBitsPerDirIdx	uimsbf
}		
}		
else{		
/* Сохранить значения из предыдущей HOADirectionalPredictionInfo полезной нагрузки для NumOfGlobalPredDirs и GlobalPredDirsIds. */		
}		
if(SingleLayer==1){		
SortedAddHoaCoeff = sort(AddHoaCoeff, ‘ascend’);		
NumOfAddHoaChansUsed = NumOfAddHoaChans;		
}		
else{		
SortedAddHoaCoeff = sort(AddHoaCoeffPerLayer[lay], ‘ascend’);		

NumOfAddHoaChansUsed =			
NumOfAddHoaChansPerLayer[lay];			
}			
for (band = 0; band < NumOfPredSubbands; band++) {			
for (dir = 0; dir < MaxNumOfPredDirsPerBand; dir++) {			
for (hoaIdx = 0;			
hoaIdx < MinNumOfCoeffsForAmbHOA;			
hoaIdx++) {			
DecodedMagDiff[band][dir][hoaIdx] = 0;			
DecodedAngleDiff[band][dir][hoaIdx] = 0;			
}			
}			
}			
for (band = 0; band < NumOfPredSubbands; band++) {			
if (!hoaIndependencyFlag) {			
KeepPreviousDirPredMatrixFlag[band];	1	bslbf	
}			
else{			
KeepPreviousDirPredMatrixFlag[band] = 0;			
}			
if (!KeepPreviousDirPredMatrixFlag[band]) {			
UseHuffmanCodingDiffMag;	1	bslbf	
if(band < FirstSBRSubbandIdx) {			
UseHuffmanCodingDiffAngle;	1	bslbf	
for (dir = 0; dir < MaxNumOfPredDirsPerBand; dir++) {			
if (**DirIsActive[band][dir]**) {	1	bslbf	
RelDirGridIdx;	NumBitsForRelDirGridIdx	uimsbf	
PredDirGridIdx[band][dir] = GlobalPredDirsIds[RelDirGridIdx];			
for (hoaIdx = 0;			
hoaIdx < MinNumOfCoeffsForAmbHOA;			
hoaIdx++) {			

```
                        readDirPredDiffValues (band, dir, hoaIdx,
                            UseHuffmanCodingDiffAbs,
                            UseHuffmanCodingDiffAngle,
                            FirstSBRSubbandIdx);
                    }
                    for ( idx = 0;
                        idx < NumOfAddHoaChansUsed;
                        idx++ ) {
                        readDirPredDiffValues (band, dir,
                            SortedAddHoaCoeff[idx] -1,
                            UseHuffmanCodingDiffAbs,
                            UseHuffmanCodingDiffAngle,
                            FirstSBRSubbandIdx);
                    }
                }
            }
        }
    }
}
```

Примечание: lay - индекс текущего активного улучшающего уровня HOA

Таблица 8. Определение SingleLayer

Значение	Описание
0	Сигнал HOA обеспечивается на нескольких уровнях; разрешает сигнализацию распределения транспортных каналов HOA по разным уровням
1	Сигнал HOA обеспечивается на одном уровне

codedLayerCh - этот элемент указывает для первого (т.е. базового) уровня количество включенных транспортных сигналов, которое задано как codedLayerCh+MinNumOfCoeffsForAmbHOA. Для более высоких (т.е. улучшающих) уровней этот элемент указывает количество дополнительных сигналов, включенных в улучшающий уровень, по сравнению со следующим более низким уровнем, которое задается как codedLayerCh+1.

HOALayerChBits - этот элемент указывает количество битов для считывания codedLayerCh.

NumLayers - этот элемент указывает (после считывания HOADecoderConfig()) общее количество уровней в битовом потоке.

NumHOAChannelsLayer - этот элемент является массивом, состоящим из элементов NumLayers, в которых i-ый элемент указывает количество транспортных сигналов, включенных во все уровни вплоть до i-го уровня.

12.4.1.x Параметры, зависимые от кадра и пользователя

$M_{LAY}(k)$ - количество всех фактически используемых уровней для k-го кадра (для определения) на стороне декодера. В случае многоуровневого кодирования (указанного посредством SingleLayer == 0) это количество должно быть меньше или равно общему количеству уровней, присутствующих в битовом потоке, т.е. $M_{LAY} \leq$ NumLayers. В случае одноуровневого кодирования (указанного посредством SingleLayer == 1) M_{LAY} устанавливается равным 1.

В зависимости от выбора $M_{LAY}(k)$ количество $I_{ADD,LAY}(k)$ дополнительных транспортных каналов, фактически используемых для пространственного декодирования HOA (т.е. дополнительно к O_{MIN} каналам, которые подразумеваются всегда используемыми), вычисляется следующим образом:

```
if(SingleLayer | (!SingleLayer & M_LAY(k) == NumLayers))
{
        I_ADD,LAY (k) = NumOfAdditionalCoders;
}
else
{
        I_ADD,LAY (k)  = NumHOACannelsLayer[M_LAY(k) − 1] - MinNumOfCoeffsForAmbHOA;
}
```

VVecLength и VVecCoeffId

Слово codedVVecLength указывает:

0) полную длину вектора (элементы NumOfHoaCoeffs), указывает, что все коэффициенты для преобладающих векторов (NumOfHoaCoeffs) определены;

1) векторные элементы от 1 до MinNumOfCoeffsForAmbHOA и все элементы, заданные в определенном ContAddHoaCoeff[lay] текущего активного уровня с индексом lay=0...NumLayers-1, не передаются; для одноуровневого режима SingleLayer == 1 переменная NumLayers должна быть установлена равной 1; указывает, что определены только коэффициенты преобладающего вектора, соответствующие номеру, больше MinNumOfCoeffsForAmbHOA, далее вычитаются NumOfContAddAmbHoaChan[lay] коэффициентов, идентифицированных в ContAddAmbHoaChan[lay], список ContAddAmbHoaChan[lay] определяет дополнительные каналы, соответствующие порядку, который превышает порядок MinAmbHoaOrder;

2) векторные элементы от 1 до MinNumOfCoeffsForAmbHOA не передаются; указывает, что определены коэффициенты преобладающих векторов, соответствующие номеру больше MinNumOfCoeffsForAmbHOA.

В случае codedVVecLength == 1 массив VVecLength[i], а также двухмерный массив VVecCoeffId[i] являются корректными для VVector с индексом i, в других случаях элемент VVecLength, а также массив VVecCoeffId[m] являются корректными для всего VVector в пределах HOAFrame. Для приведенного ниже алгоритма присвоения функции помощника определена следующим образом.

```
switch CodedVVecLength{
        case 0:
                VVecLength = NumOfHoaCoeffs;
                for (m=0; m<VVecLength; ++m) {
                        VVecCoeffId[m] = m;
                }
                break;
        case 1:
                for (i=0; i < NumOfVecSigs; ++i) {
                        lay = VecSigLayerIdx[i];
                        VVecLength[i] = NumOfHoaCoeffs
                                -.MinNumOfCoeffsForAmbHOA
                                - NumOfContAddHoaChans[lay];
                        CoeffIdx = MinNumOfCoeffsForAmbHOA+1;
                        for (m=0; m<VVecLength[i]; ++m) {
                                bIsInArray = isMemberOf(CoeffIdx,
                                        ContAddHoaCoeff[lay],
                                        NumOfContAddHoaChans[lay]);
                                while (bIsInArray) {
                                        CoeffIdx++;
                                        bIsInArray = isMemberOf(CoeffIdx,
                                                ContAddHoaCoeff[lay],
                                                NumOfContAddHoaChans[lay]);
                                }
                                VVecCoeffId[i][m] = CoeffIdx-1;
                        }
                }
                break;
        case 2:
                VVecLength = NumOfHoaCoeffs – MinNumOfCoeffsForAmbHOA;
                for (m=0; m< VVecLength; ++m) {
                        VVecCoeffId[m] = m + MinNumOfCoeffsForAmbHOA;
                }
}
```

Первый оператор switch с тремя case (0-2), таким образом, обеспечивает метод, посредством которого можно определить длину преобладающего вектора с точки зрения количества (VVecLength) и индексов коэффициентов (VVecCoeffId).

12.4.1.X Преобразование в элемент VVec

Вид обратного квантования V-вектора сигнализируется посредством слова NbitsQ. Значение 4 для NbitsQ указывает векторное квантование. Когда NbitsQ == 5, выполняется однородное 8-битное скалярное обратное квантование. Напротив, значение NbitsQ ≥6 указывает применение декодирования Хафмана скалярным образом квантованного V-вектора. Режим предсказания обозначен как PFlag, в то время как CbFlag представляет информационный бит таблицы Хафмана.

```
if (CodedVVecLength == 1) {
    VVecLengthUsed = VVecLength[i];
    VVecCoeffIdUsed = VVecCoeffId[i];
} else {
    VVecLengthUsed = VVecLength;
    VVecCoeffIdUsed = VVecCoeffId;
}
if (NbitsQ(k)[i] == 4) {
    if (NumVvecIndices == 1) {
        for (m=0; m< VVecLengthUsed; ++m) {
            idx = VVecCoeffIdUsed[m];
            v^(i)_idx(k) = WeightVal[0] * VecDict[900].[VvecIdx[0]][idx];
        }
    } else {
        cdbLen = O;
        if (N==4) {
            cdbLen = 32;
        }
        for (m=0; m<O; ++m) {
            TmpVVec[m] = 0;
            for (j=0; j< NumVvecIndices; ++j) {
                TmpVVec[m] += WeightVal[j] * VecDict[cdbLen].[VvecIdx[j]][m];
            }
        }
        FNorm = 0.0;
        for (m=0; m<O; ++m) {
            FNorm += TmpVVec[m] * TmpVVec[m];
        }
        FNorm = (N+1)/sqrt(FNorm);
        for (m=0; m< VVecLengthUsed; ++m) {
                idx = VVecCoeffIdUsed[m];
                v^(i)_idx(k)= TmpVVec[idx] * FNorm;
        }
    }
}
elseif (NbitsQ(k)[i] == 5) {
    for (m=0; m< VVecLengthUsed; ++m) {
        v^(i)_VVecCoeffIdUsed[m](k) =  (N+1)*aVal[i][m];
```

```
        }
    }
    elseif (NbitsQ(k)[i] >= 6) {
        for (m=0; m< VVecLengthUsed; ++m) {
            v^(i)_VVecCoeffIdUsed[m](k) = (N+1) * (2^(16 – NbitsQ(k)[i])*aVal[i][m])/2^15;
            if (PFlag(k)[i] == 1) {
                v^(i)_VVecCoeffIdUsed[m](k) += v^(i)_VVecCoeffIdUsed[m](k − 1);
            }
        }
    }
```

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ декодирования сжатого представления (2100) звука или звукового поля системы Амбисоник высшего порядка (HOA), причем способ содержит этапы, на которых

принимают (S5010) битовый поток, содержащий сжатое представление HOA, соответствующее множеству иерархических уровней, которые включают в себя базовый уровень (1200) и два или более иерархических улучшающих уровней (1300-1, 1300-(M-1)), причем битовый поток содержит для каждого уровня соответствующую полезную нагрузку расширения HOA, включающую в себя вспомогательную информацию для параметрического расширения воссозданного представления HOA, получаемого из транспортных сигналов, присвоенных соответствующему уровню и любым уровням ниже соответствующего уровня;

определяют (S5030) наиболее высокий применимый уровень среди множества уровней для декодирования, причем наиболее высокий применимый уровень представляет собой уровень ниже наиболее низкого уровня, который не был принят корректно;

извлекают (S5040) полезную нагрузку расширения HOA, присвоенную наиболее высокому применимому уровню, причем полезная нагрузка расширения HOA включает в себя вспомогательную информацию для параметрического улучшения воссозданного представления HOA, соответствующего наиболее высокому применимому уровню, причем воссозданное представление HOA, соответствующее наиболее высокому применимому уровню, может быть получено на основе транспортных сигналов, присвоенных наиболее высокому применимому уровню и любым уровням ниже наиболее высокого применимого уровня;

декодируют (S5050) сжатое представление HOA, соответствующее наиболее высокому применимому уровню, на основе информации уровня, транспортных сигналов, присвоенных наиболее высокому применимому уровню и любым уровням ниже наиболее высокого применимого уровня; и

выполняют (S5060) параметрическое улучшение декодированного представления HOA с использованием вспомогательной информации, включенной в полезную нагрузку расширения HOA, присвоенную наиболее высокому применимому уровню.

2. Способ по п.1, в котором полезная нагрузка расширения HOA включает в себя элементы битового потока для инструмента декодирования с пространственным предсказанием сигналов HOA.

3. Способ по п.1, в котором информация уровня указывает количество активных направленных сигналов в текущем кадре улучшающего уровня (1300-1, 1300-(M-1)).

4. Способ по п.1, в котором информация уровня указывает общее количество дополнительных коэффициентов окружения HOA для улучшающего уровня (1300-1, 1300-(M-1)).

5. Способ по п.1, в котором информация уровня включает в себя индексы коэффициентов HOA для каждого дополнительного окружающего коэффициента HOA для улучшающего уровня (1300-1, 1300-(M-1)).

6. Способ по п.1, в котором информация уровня включает в себя улучшающую информацию, которая включает в себя по меньшей мере одно из пространственного предсказания сигналов, синтеза направленных подполосных сигналов и декодера параметрического дублирования звукового окружения.

7. Способ по п.1, в котором сжатое представление HOA адаптировано для режима многоуровневого кодирования для контента на основе HOA, если в HOADecoderConfig() сигнализирован CodedVVecLength, равный 1.

8. Способ по п.1, в котором множество ContAddHoaCoeff определено отдельно для каждого из множества иерархических уровней.

9. Способ по п.1, в котором информация уровня включает в себя элементы NumLayers, причем каждый элемент указывает количество транспортных сигналов, включенных во все уровни вплоть до i-го уровня.

10. Способ по п.1, в котором информация уровня включает в себя индикатор всех фактически используемых уровней для k-го кадра.

11. Способ по п.1, в котором информация уровня указывает, что определены все коэффициенты для преобладающих векторов.

12. Способ по п.1, в котором информация уровня указывает, что определены коэффициенты преобладающих векторов, соответствующих номеру, который больше MinNumOfCoeffsForAmbHOA.

13. Способ по п.1, в котором информация уровня указывает MinNumOfCoeffsForAmbHOA, и все элементы, определенные в ContAddHoaCoeff[lay], не передаются, где lay - индекс уровня, содержащего основанный на векторе сигнал, соответствующий вектору.

14. Устройство (2100) для декодирования сжатого представления звука или звукового поля системы Амбисоник высшего порядка (HOA), содержащее

приемник, выполненный с возможностью принимать (S5010) битовый поток, содержащий сжатое представление HOA, соответствующее множеству иерархических уровней, которые включают в себя базовый уровень (1200) и два или более иерархических улучшающих уровней (1300-1, 1300-(M-1)), причем битовый поток содержит для каждого уровня соответствующую полезную нагрузку расширения HOA, включающую в себя вспомогательную информацию для параметрического расширения воссозданного представления HOA, получаемого из транспортных сигналов, присвоенных соответствующему уровню и любым уровням ниже соответствующего уровня,

декодер, выполненный с возможностью

определять (S5030) наиболее высокий применимый уровень среди множества уровней для декодирования, причем наиболее высокий применимый уровень представляет собой уровень ниже наиболее низкого уровня, который не был принят корректно;

извлекать (S5040) полезную нагрузку расширения HOA, присвоенную наиболее высокому применимому уровню, причем полезная нагрузка расширения HOA включает в себя вспомогательную информацию для параметрического улучшения воссозданного представления HOA, соответствующего наиболее высокому применимому уровню, причем воссозданное представление HOA, соответствующее наиболее высокому применимому уровню, может быть получено на основе транспортных сигналов, присвоенных наиболее высокому применимому уровню и любым уровням ниже наиболее высокого применимого уровня;

декодировать (S5050) сжатое представление HOA, соответствующее наиболее высокому применимому уровню, на основе информации уровня, транспортных сигналов, присвоенных наиболее высокому применимому уровню и любым уровням ниже наиболее высокого применимого уровня; и

выполнять (S5060) параметрическое улучшение декодированного представления HOA с использованием вспомогательной информации, включенной в полезную нагрузку расширения HOA, присвоенную наиболее высокому применимому уровню.

15. Устройство по п.14, в котором полезная нагрузка расширения HOA включает в себя элементы битового потока для инструмента декодирования с пространственным предсказанием сигналов HOA.

16. Устройство по п.14, в котором информация уровня указывает количество активных направленных сигналов в текущем кадре улучшающего уровня (1300-1, 1300-(M-1)).

17. Устройство по п.14, в котором информация уровня указывает общее количество дополнительных коэффициентов окружения HOA для улучшающего уровня (1300-1, 1300-(M-1)).

18. Устройство по п.14, в котором информация уровня включает в себя индексы коэффициентов HOA для каждого дополнительного окружающего коэффициента HOA для улучшающего уровня (1300-1, 1300-(M-1)).

19. Устройство по п.14, в котором информация уровня включает в себя улучшающую информацию, которая включает в себя по меньшей мере одно из пространственного предсказания сигналов, синтеза направленных подполосных сигналов и декодера параметрического дублирования звукового окружения.

20. Устройство по п.14, в котором сжатое представление HOA адаптировано для режима многоуровневого кодирования для контента на основе HOA, если в HOADecoderConfig() сигнализирован CodedVVecLength, равный 1.

21. Устройство по п.14, в котором множество ContAddHoaCoeff определено отдельно для каждого из множества иерархических уровней.

22. Устройство по п.14, в котором информация уровня включает в себя элементы NumLayers, причем каждый элемент указывает количество транспортных сигналов, включенных во все уровни вплоть до i-го уровня.

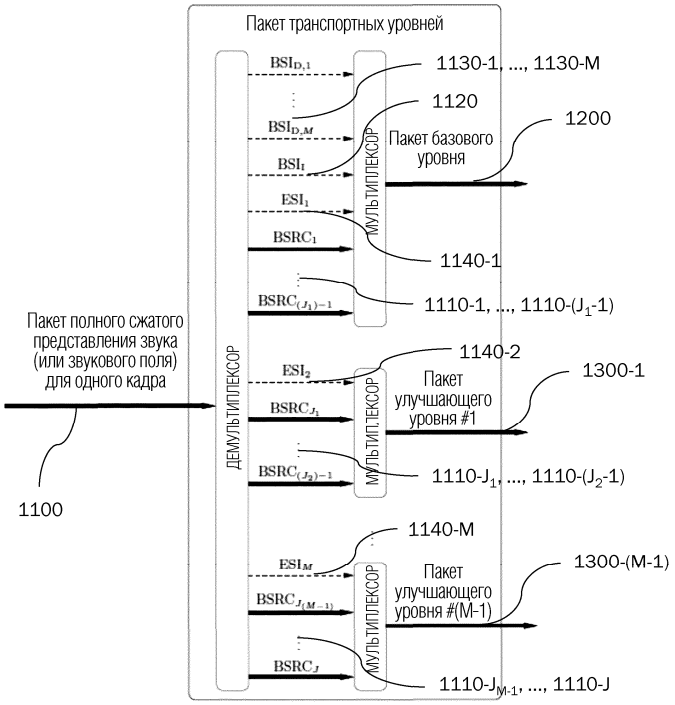
23. Устройство по п.14, в котором информация уровня включает в себя индикатор всех фактически используемых уровней для k-го кадра.

24. Устройство по п.14, в котором информация уровня указывает, что определены все коэффициенты для преобладающих векторов.

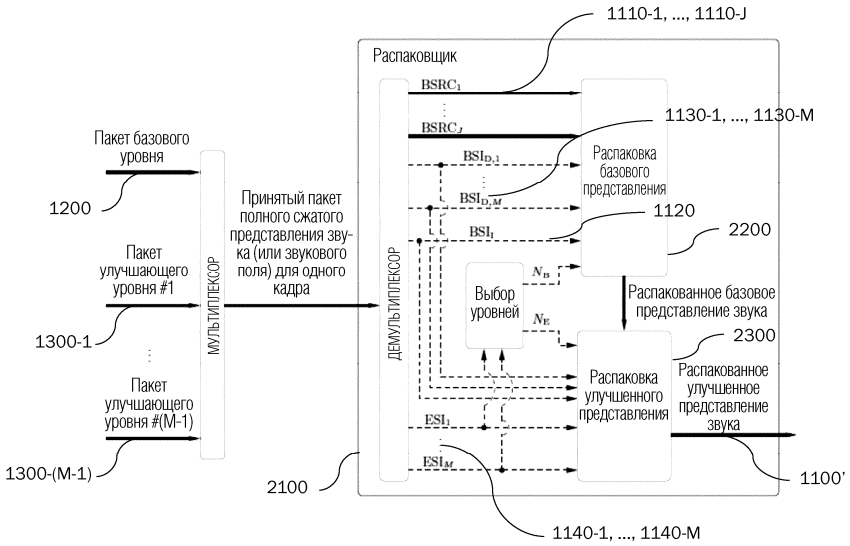
25. Устройство по п.14, в котором информация уровня указывает, что определены коэффициенты преобладающих векторов, соответствующих номеру, который больше MinNumOfCoeffsForAmbHOA.

26. Устройство по п.14, в котором информация уровня указывает MinNumOfCoeffsForAmbHOA, и

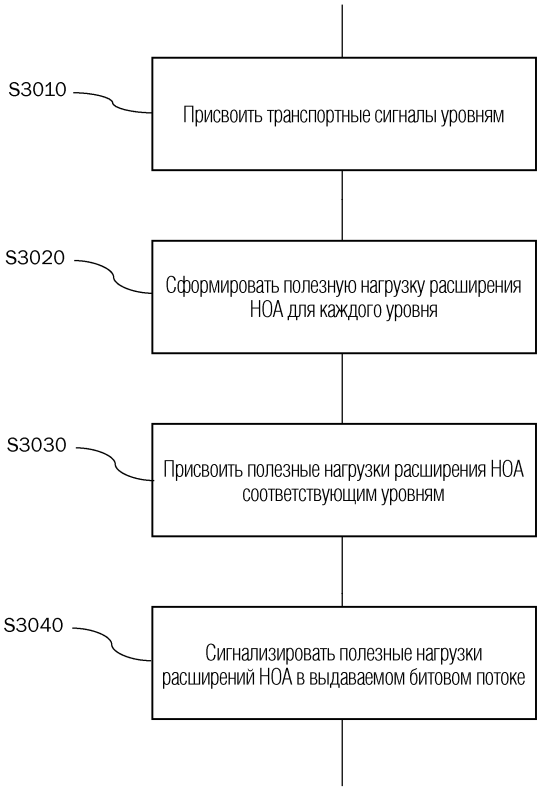
все элементы, определенные в ContAddHoaCoeff[lay], не передаются, где lay - индекс уровня, содержащего основанный на векторе сигнал, соответствующий вектору.



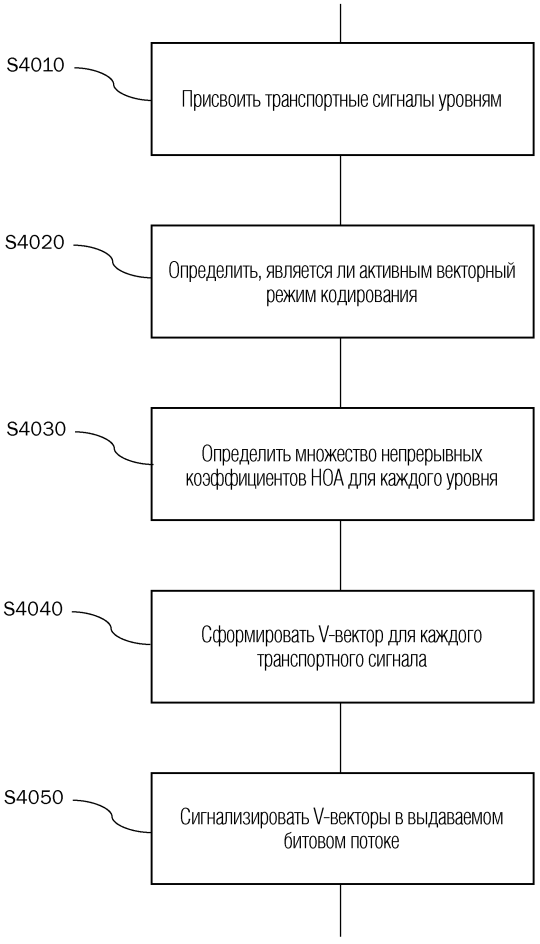
Фиг. 1



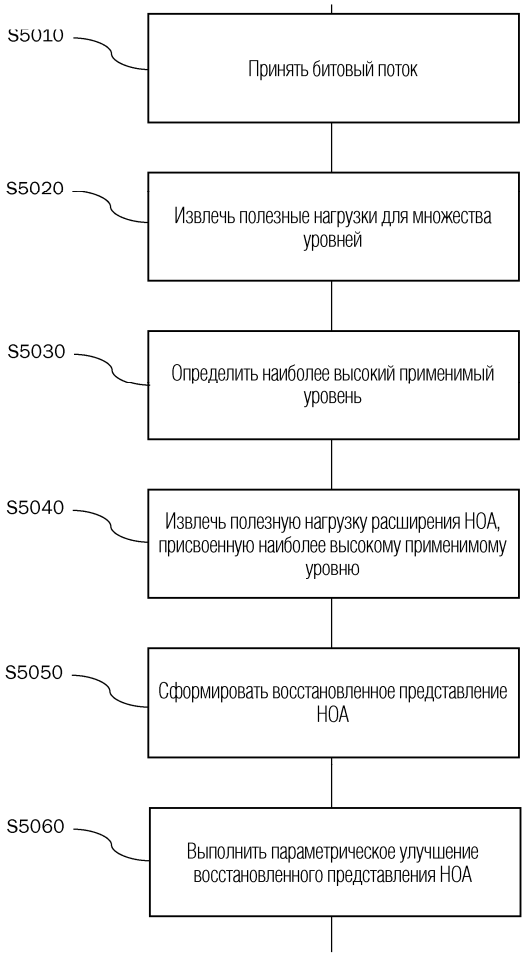
Фиг. 2


S3010
Присвоить транспортные сигналы уровням
S3020
Сформировать полезную нагрузку расширения HOA для каждого уровня
S3030
Присвоить полезные нагрузки расширения HOA соответствующим уровням
S3040
Сигнализировать полезные нагрузки расширений HOA в выдаваемом битовом потоке

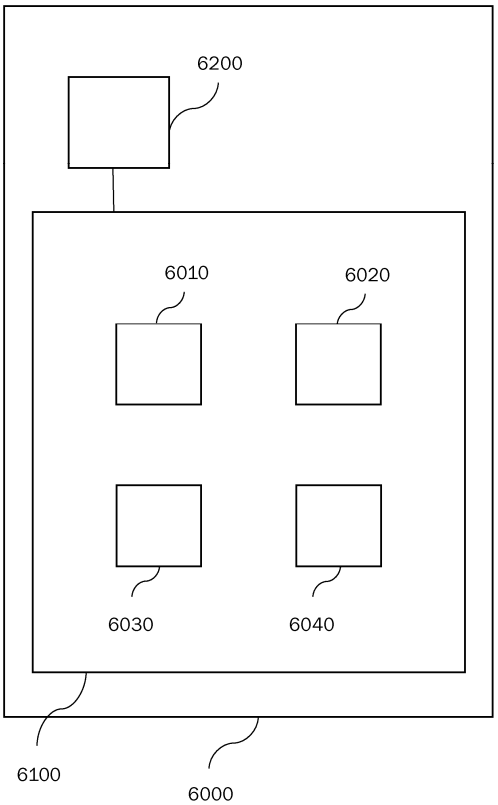
Фиг. 3


S4010
Присвоить транспортные сигналы уровням
S4020
Определить, является ли активным векторный режим кодирования
S4030
Определить множество непрерывных коэффициентов HOA для каждого уровня
S4040
Сформировать V-вектор для каждого транспортного сигнала
S4050
Сигнализировать V-векторы в выдаваемом битовом потоке

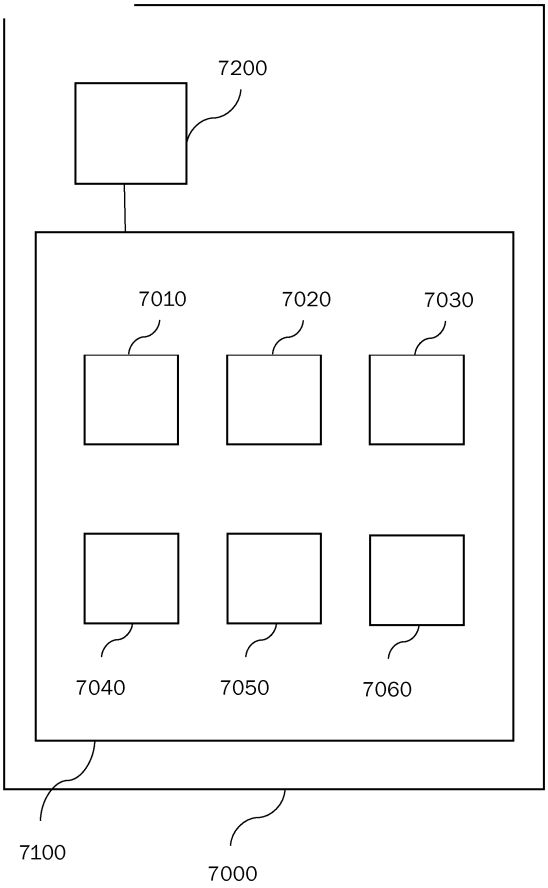
Фиг. 4


S5010
Принять битовый поток
S5020
Извлечь полезные нагрузки для множества уровней
S5030
Определить наиболее высокий применимый уровень
S5040
Извлечь полезную нагрузку расширения HOA, присвоенную наиболее высокому применимому уровню
S5050
Сформировать восстановленное представление HOA
S5060
Выполнить параметрическое улучшение восстановленного представления HOA

Фиг. 5


6200
6010
6020
6030
6040
6100
6000

Фиг. 6



Фиг. 7

Евразийская патентная организация, ЕАПВ
Россия, 109012, Москва, Малый Черкасский пер., 2