РОССИЙСКАЯ ФЕДЕРАЦИЯ

(19) **RU** (11) **2 554 844**(13) **C2**

(51) МПК
G10L 19/008 (2013.01)

ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(21)(22) Заявка: **2012144366/08, 06.04.2011**

(24) Дата начала отсчета срока действия патента:
06.04.2011

Приоритет(ы):
(30) Конвенционный приоритет:
09.04.2010 US 61/322,458

(43) Дата публикации заявки: **27.04.2014** Бюл. № 12

(45) Опубликовано: **27.06.2015** Бюл. № 18

(56) Список документов, цитированных в отчете о поиске: **US 2005197831 A1, 08.09.2005** (см. прод.)

(85) Дата начала рассмотрения заявки PCT на национальной фазе: **18.10.2012**

(86) Заявка PCT:
EP 2011/055369 (06.04.2011)

(87) Публикация заявки PCT:
WO 2011/124616 (13.10.2011)

Адрес для переписки:
123242, Москва, Кудринская пл., 1, а/я 35, "Михайлюк, Сороколат и партнеры-патентные поверенные"

(72) Автор(ы):
ПУРНХАГЕН Хейко (SE),
КАРЛЬССОН Понтус (SE),
ВИЛЛЕМОЕС Ларс (SE)

(73) Патентообладатель(и):
ДОЛБИ ИНТЕРНЕШНЛ АБ (NL)

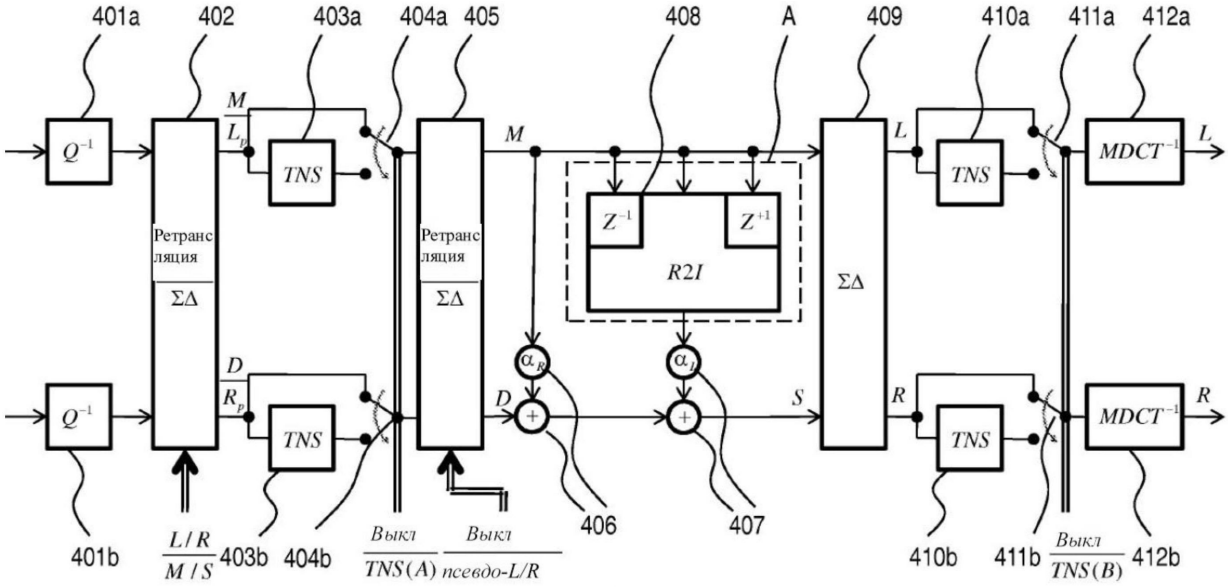
(54) СТЕРЕОФОНИЧЕСКОЕ КОДИРОВАНИЕ НА ОСНОВЕ MDCT С КОМПЛЕКСНЫМ ПРЕДСКАЗАНИЕМ

(57) Реферат:

Изобретение относится к средствам для стереофонического кодирования и декодирования с использованием комплексного предсказания в частотной области. Технический результат заключается в обеспечении высокого качества звука при уменьшении вычислительных затрат. В одном из вариантов осуществления изобретения способ декодирования, предназначенный для получения выходного стереофонического сигнала из входного стереофонического сигнала, закодированного посредством стереофонического кодирования с комплексным предсказанием и включающего первые представления двух входных каналов в частотной области, включает следующие этапы повышающего микширования: вычисление второго представления первого входного канала в частотной области и вычисление выходного канала на основе первого и второго представлений первого входного канала в частотной области, первого представления второго входного канала в частотной области и коэффициента комплексного предсказания. Способ включает выполнение модификаций в частотной области селективно перед повышающим микшированием или после повышающего микширования. 3 н. и 12 з.п. ф-лы, 19 ил.

RU 2554844 C2

RU 2554844 C2



Фиг. 4

(56) (продолжение):
US 2005165587 A1, 28.07.2005 US 2009216542 A1, 27.08.2009 RU 2380766 C2, 27.01.2010EP 2137725 A1, 30.12.2009

RU 2554844 C2

RU 2554844 C2

RUSSIAN FEDERATION

FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(19) **RU** (11) **2 554 844**(13) **C2**

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

(12) ABSTRACT OF INVENTION

(21)(22) Application: **2012144366/08, 06.04.2011**

(24) Effective date for property rights:
06.04.2011

Priority:
(30) Convention priority:
09.04.2010 US 61/322,458

(43) Application published: **27.04.2014** Bull. № **12**

(45) Date of publication: **27.06.2015** Bull. № **18**

(85) Commencement of national phase: **18.10.2012**

(86) PCT application:
EP 2011/055369 (06.04.2011)

(87) PCT publication:
WO 2011/124616 (13.10.2011)

Mail address:
123242, Moskva, Kudrinskaja pl., 1, a/ja 35, "Mikhajljuk, Sorokolat i partnery-patentnye poverennye"

(72) Inventor(s):
PURNKhAGEN Khejko (SE),
KARL'SSON Pontus (SE),
VILLEMOES Lars (SE)

(73) Proprietor(s):
DOLBI INTERNEShNL AB (NL)

RU 2554844 C2

(54) MDCT-BASED COMPLEX PREDICTION STEREO CODING

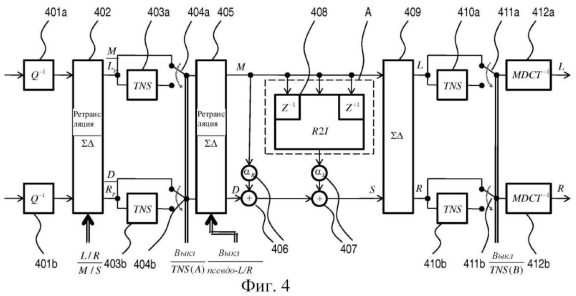
(57) Abstract:

FIELD: radio engineering, communication.

SUBSTANCE: invention relates to means of stereo encoding and decoding using complex prediction in the frequency domain. In one of the versions of the invention, a decoding method for obtaining an output stereo signal from an input stereo signal encoded by complex prediction coding and comprising first frequency-domain representations of two input channels comprises upmixing steps of: computing a second frequency-domain representation of a first input channel; and computing an output channel based on the first and second frequency-domain representations of the first input channel, the first frequency-domain representation of the second input channel and a complex prediction coefficient. The method includes performing frequency-domain modifications selectively before or after upmixing.

EFFECT: providing high audio quality while reducing computational costs.

15 cl, 19 dwg



Фиг. 4

RU 2554844 C2

Область технического применения

Изобретение, раскрытое в данном документе, в общем, относится к стереофоническому кодированию и, точнее, к способам стереофонического кодирования с использованием комплексного предсказания в частотной области.

Предпосылки изобретения

Совместное кодирование левого (L) и правого (R) каналов стереофонического сигнала делает возможным более эффективное кодирование по сравнению с независимым кодированием L и R. Общий подход для совместного стереофонического кодирования представляет собой среднее/побочное (M/S) кодирование. Здесь средний (M) сигнал формируется путем сложения сигналов L и R, например, сигнал M может иметь форму

M=(L+R)/2.

Также путем вычитания двух каналов L и R формируется побочный сигнал (S), например, сигнал S может иметь форму

S=(L-R)/2.

В случае M/S-кодирования вместо сигналов L и R кодируются сигналы M и S.

В стандарте MPEG (Экспертная группа по вопросам движущегося изображения) AAC (перспективное звуковое кодирование) (см. документ стандарта ISO/IEC 13818-7) L/R-стереофоническое кодирование и M/S-стереофоническое кодирование могут выбираться изменяющимся в зависимости от времени или изменяющимся в зависимости от частоты образом. Так, стереофонический кодер может применять L/R-кодирование для некоторых частотных полос стереофонического сигнала, в то время как для кодирования других частотных полос стереофонического сигнала используется M/S-кодирование (изменение в зависимости от частоты). Кроме того, кодер может переключаться между L/R- и M/S-кодированием с течением времени (изменение в зависимости от времени). В стандарте MPEG AAC стереофоническое кодирование осуществляется в частотной области, конкретнее, в области MDCT (модифицированного дискретного косинусного преобразования). Это позволяет адаптивно выбирать или L/R-, или M/S-кодирования изменяющимся в зависимости от частоты, а также изменяющимся в зависимости от времени образом.

Параметрическое стереофоническое кодирование представляет собой способ эффективного кодирования стереофонического звукового сигнала как монофонического сигнала плюс небольшое количество дополнительной информации для стереофонических параметров. Оно составляет часть стандарта MPEG-4 Audio (см. документ стандарта ISO/IEC 14496-3). Монофонический сигнал может кодироваться с использованием любого кодера звука. Стереофонические параметры могут внедряться во вспомогательную часть монофонического битового потока, и, таким образом, достигается полная прямая и обратная совместимость. В декодере в первую очередь декодируется монофонический сигнал, после чего стереофонический сигнал реконструируется при помощи стереофонических параметров. Декоррелированная версия декодированного монофонического сигнала, которая имеет нулевую взаимную корреляцию с монофоническим сигналом, генерируется посредством декоррелятора, например, соответствующего фазового фильтра, который может включать одну или несколько линий задержки. По существу, декоррелированный сигнал имеет такое же спектральное и временное распределение энергии, как и монофонический сигнал. Монофонический сигнал совместно с декоррелированным сигналом являются входными в процесс повышающего микширования, который управляется стереофоническими параметрами и который реконструирует стереофонический сигнал. Для получения дополнительной информации см. статью "Low Complexity Parametric Stereo Coding in

MPEG-4", H. Purnhagen, Proc. of the 7th Int. Conference on Digital Audio Effects (DAFx'04), Naples, Italy, October 5-8, 2004, pages 163-168.

MPEG Surround (MPS; см. ISO/IEC 23003-1 и статью "MPEG Surround - The ISO/MPEG Standard for Efficient and Compatible Multi-Channel Audio Coding", J. Herre et al., Audio Engineering Convention Paper 7084, 122nd Convention, May 5-8, 2007) позволяет объединять принципы параметрического стереофонического кодирования с остаточным кодированием, замещающим декоррелированный сигнал передаваемым остатком и, таким образом, улучшающим воспринимаемое качество. Остаточное кодирование может выполняться путем понижающего микширования многоканального сигнала и, необязательно, путем извлечения пространственных меток. В ходе процесса понижающего микширования вычисляются, а затем кодируются и передаются остаточные сигналы, представляющие сигнал ошибки. В декодере они могут замещать декоррелированные сигналы. При гибридном подходе они могут замещать декоррелированные сигналы в некоторых полосах частот, предпочтительно, в относительно низкочастотных полосах.

В соответствии с современной системой MPEG унифицированного кодирования речи и звука (USAC), два примера которой приведены на Фиг. 1, декодер включает блок комплекснозначных квадратурных зеркальных фильтров (QMF), расположенный в нисходящем направлении относительно базового декодера. QMF-представление, получаемое на выходе блока фильтров, является комплексным - и, таким образом, передискретизированным в два раза - и может быть организовано как сигнал понижающего микширования (или, эквивалентно, средний сигнал) M и остаточный сигнал D, к которым применяется матрица повышающего микширования с комплексными элементами. Сигналы L и R (в области QMF) получаются как:

$$\begin{bmatrix} L \\ R \end{bmatrix} = g \begin{bmatrix} 1-\alpha & 1 \\ 1+\alpha & -1 \end{bmatrix} \begin{bmatrix} M \\ D \end{bmatrix},$$

где g - действительнозначный коэффициент усиления, и α - комплекснозначный коэффициент предсказания. Предпочтительно, α выбирается так, чтобы минимизировать энергию остаточного сигнала D. Коэффициент усиления может определяться путем нормализации, т.е. так, чтобы обеспечить то, что мощность суммарного сигнала будет равна сумме мощностей левого и правого сигналов. Действительные и мнимые части каждого из сигналов L и R являются взаимно избыточными - в принципе, каждая из них может быть вычислена на основе другой, - но они являются полезными для того, чтобы сделать возможным последующее применение декодера с репликацией спектральной полосы (SBR) без возникновения слышимых артефактов из-за наложения спектров. По похожим причинам, использование передискретизированного представления сигнала также может выбираться с целью предотвращения появления артефактов, связанных с другой обработкой сигнала, адаптивной ко времени или к частоте (не показана), такой как, например, повышающее микширование монофонического сигнала в стереофонический сигнал. Последним этапом обработки в декодере является обратная QMF-фильтрация. Отмечается, что QMF-представление сигнала с ограниченной полосой допускает разностные способы с ограниченной полосой и способы «остаточного заполнения», которые могут интегрироваться в декодеры этого типа.

Приведенная выше структура кодирования хорошо подходит для низких битовых скоростей передачи данных, как правило, ниже 80 Кбит/с, но не является оптимальной для более высоких битовых скоростей передачи данных в том, что касается

вычислительной сложности. Точнее, при более высоких битовых скоростях передачи данных инструмент SBR, как правило, не применяется (поскольку он не будет повышать эффективность кодирования). Поэтому в декодере без ступени SBR только наличие комплекснозначной матрицы повышающего микширования оправдывает присутствие блока QMF-фильтров, который требует больших вычислительных ресурсов и вносит задержку (при длине кадра 1024 дискретных значений блок анализирующих/синтезирующих QMF-фильтров вносит задержку в 961 дискретных значений). Это ясно указывает на потребность в более эффективной структуре кодирования.

Краткое описание изобретения

Целью настоящего изобретения является создание способов и устройства для стереофонического кодирования, которые являются эффективными в вычислительном отношении также и в диапазоне высоких битовых скоростей передачи данных.

Изобретение исполняет указанную цель, предусматривая кодер и декодер, способы кодирования и декодирования и, соответственно, компьютерные программные продукты, предназначенные для кодирования и декодирования, что определено независимыми пунктами формулы изобретения. Зависимые пункты формулы изобретения определяют варианты осуществления изобретения.

В первой особенности изобретение предусматривает систему декодера, предназначенную для создания стереофонического сигнала путем стереофонического кодирования с комплексным предсказанием, и система декодера включает:

- повышающее микширование, адаптированное для генерирования стереофонического сигнала на основе первых представлений сигнала понижающего микширования (M) понижающего микширования и остаточного сигнала (D) в частотной области, где каждое из первых представлений в частотной области включает первые спектральные составляющие, представляющие спектральный состав соответствующего сигнала, выраженного в первом подпространстве многомерного пространства, и ступень повышающего микширования включает:

- - модуль, предназначенный для вычисления второго представления сигнала понижающего микширования в частотной области на основе его первого представления в частотной области, где второе представление в частотной области включает вторые спектральные составляющие, представляющие спектральный состав сигнала, выраженного во втором подпространстве многомерного пространства, которое включает часть многомерного пространства, не включенную в первое подпространство;

- - взвешенный сумматор, предназначенный для вычисления побочного сигнала (S) на основе первого и второго представлений сигнала понижающего микширования в частотной области, первого представления остаточного сигнала в частотной области и коэффициента (α) комплексного предсказания, закодированного в сигнале битового потока; и

- суммарно-разностную ступень, предназначенную для вычисления стереофонического сигнала на основе первого представления сигнала понижающего микширования в частотной области и побочного сигнала,

где ступень повышающего микширования также может действовать в режиме ретрансляции, в котором указанные сигнал понижающего микширования и остаточный сигнал подаются непосредственно на суммарно-разностную ступень.

Во второй особенности изобретение предусматривает систему кодера, предназначенную для кодирования стереофонического сигнала посредством сигнала битового потока путем стереофонического кодирования с комплексным предсказанием, которая включает:

- оцениватель, предназначенный для оценки коэффициента комплексного предсказания;

- ступень кодирования, действующая для:

(a) преобразования стереофонического сигнала в представлений в частотной области сигнала понижающего микширования и остаточного сигнала во взаимосвязи, определяемой значением коэффициента комплексного предсказания;

и

мультиплексор, предназначенный для приема выходных данных ступени кодирования и оценивателя и для их кодирования посредством указанного сигнала битового потока.

В третьей и четвертой особенностях изобретения предусматриваются способы кодирования стереофонического сигнала в битовый поток и декодирования битового потока в, по меньшей мере, один стереофонический сигнал. Технические признаки каждого из способов аналогичны таковым, соответственно, для системы кодера и системы декодера. В пятой и шестой особенностях изобретение также предусматривает компьютерный программный продукт, содержащий команды, предназначенные для исполнения каждого из способов на компьютере.

Изобретение извлекает выгоду из преимуществ унифицированного стереофонического кодирования в системе MPEG USAC. Эти преимущества сохраняются и при более высоких битовых скоростях передачи данных, когда SBR, как правило, не используется, без значительного увеличения вычислительной сложности, которым мог бы сопровождаться подход на основе QMF. Это является возможным, поскольку критически дискретизированное преобразование MDCT, которое находится в основе кодирования с преобразованием в системе MPEG USAC, в соответствии с изобретением может применяться для стереофонического кодирования с комплексным предсказанием, по меньшей мере, в тех случаях, когда полосы пропускания кодированных звуковых сигналов канала понижающего микширования и остаточного канала одинаковы, и процесс повышающего микширования не включает декорреляцию. Это означает, что дополнительное QMF-преобразование больше не требуется. Показательная реализация стереофонического кодирования с комплексным предсказанием в области QMF фактически значительно увеличивала бы количество операций в единицу времени по сравнению с традиционным L/R- или M/S-стерео. Поэтому кодирующее устройство согласно изобретению оказывается конкурентоспособным при указанных битовых скоростях передачи данных, обеспечивая высокое качество звука при умеренных вычислительных затратах.

Как понятно специалистам, то, что ступень повышающего микширования также может действовать и в режиме ретрансляции, позволяет декодеру адаптивно выполнять декодирование в соответствии с традиционным прямым, или совместным, кодированием и кодированием с комплексным предсказанием, что определяется на стороне кодера. Поэтому в тех случаях, когда декодер не может положительно повысить уровень качества выше уровня традиционного прямого L/R-стереофонического кодирования или совместного M/S-стереофонического кодирования, он может, по меньшей мере, гарантировать, что будет поддерживаться тот же уровень. Таким образом, декодер согласно данной особенности изобретения может с функциональной точки зрения считаться расширенным относительно предпосылок.

Как преимущество перед стереофоническим сигналом, кодированным с предсказанием на основе QMF, возможна совершенная реконструкция сигнала (не считая ошибок квантования, которые могут быть сделаны сколь угодно малыми).

Таким образом, изобретение предусматривает кодирующее устройство для

стереофонического кодирования сигнала на основе преобразования путем комплексного предсказания. Предпочтительно, устройство согласно изобретению не ограничивается стереофоническим кодированием с комплексным предсказанием, но также может действовать и в режиме прямого L/R-стереофонического кодирования или совместного M/S-стереофонического кодирования в соответствии с предпосылками так, чтобы можно было выбирать наиболее подходящий способ кодирования для конкретного применения или в ходе отдельного промежутка времени.

В качестве основы для комплексного предсказания согласно изобретению используется передискретизированное (например, комплексное) представление сигнала, включающее указанные первую и вторую спектральные составляющие, и поэтому модули, предназначенные для вычисления указанного передискретизированного представления, располагаются в системе кодера и в системе декодера согласно изобретению. Спектральные составляющие относятся к первому и второму подпространствам многомерного пространства, которое может представлять собой множество функций, зависящих от времени, на интервале заданной длины (например, предварительно определенной длины временного кадра), дискретизированном с конечной частотой дискретизации. Хорошо известно, что функции в таком особом многомерном пространстве могут аппроксимироваться конечной взвешенной суммой базисных функций.

Как понятно специалистам, кодер, адаптированный для совместного действия с декодером, оснащается эквивалентными модулями для создания передискретизированного представления, на котором основывается кодирование с предсказанием, так, чтобы сделать возможным достоверное воспроизведение кодированного сигнала. Указанные эквивалентные модули могут быть идентичными или сходными модулями, имеющими идентичные, или сходные, характеристики передачи. В частности, модули кодера и декодера, соответственно, могут представлять собой похожие или непохожие обрабатывающие блоки, исполняющие соответствующие компьютерные программы, которые выполняют эквивалентные наборы математических операций.

В некоторых вариантах осуществления системы декодера или системы кодера первые спектральные составляющие имеют действительные значения, выраженные в первом подпространстве, а вторые спектральные составляющие имеют мнимые значения, выраженные во втором подпространстве. Первые и вторые спектральные составляющие совместно образуют комплексное спектральное представление сигнала. Первое подпространство может представлять собой линейную оболочку первого набора базисных функций, в то время как второе подпространство может представлять собой оболочку набора вторых базисных функций, некоторые из которых линейно независимы от первого множества базисных функций.

В одном из вариантов осуществления изобретения модуль, предназначенный для вычисления комплексного представления, представляет собой преобразование действительного в мнимое, т.е. модуль для вычисления мнимых частей спектра дискретного сигнала на основании действительного спектрального представления этого сигнала. Преобразование может основываться на точных или приближенных математических зависимостях, таких как формулы из гармонического анализа или эвристические зависимости.

В некоторых вариантах осуществления системы декодера или системы кодера первые спектральные составляющие могут быть получены посредством преобразования дискретного сигнала во временной области из временной области в частотную,

предпочтительно, посредством преобразование Фурье, такого как дискретное косинусное преобразование (DCT), модифицированное дискретное косинусное преобразование (MDCT), дискретное синусное преобразование (DST), модифицированное дискретное синусное преобразование (MDST), быстрое преобразование Фурье (FFT), алгоритм Фурье на основе простых множителей и т.п. В первых четырех случаях вторые спектральные составляющие могут быть затем получены, соответственно, путем DST, MDST, DCT и MDCT. Как хорошо известно, линейная оболочка косинусов, которые являются периодическими на единичном интервале, образует подпространство, которое не полностью помещается в линейной оболочке синусов, периодических на том же интервале. Предпочтительно, первые спектральные составляющие могут быть получены посредством MDCT и вторые спектральные составляющие могут быть получены посредством MDST.

В одном из вариантов осуществления изобретения система декодера включает, по меньшей мере, один модуль временного ограничения шума (модуль TNS, или TNS-фильтр), который располагается в восходящем направлении относительно ступени повышающего микширования. Вообще говоря, использование TNS повышает воспринимаемое качество звука для сигналов с составляющими переходного типа, и это также применимо к вариантам осуществления системы декодера согласно изобретению, содержащим TNS в качестве характерного признака. В традиционном L/R- или M/S-стереофоническом кодировании TNS-фильтр может применяться как последний этап обработки в частотной области непосредственно перед обратным преобразованием. В случае стереофонического кодирования с комплексным предсказанием, однако, часто является более преимущественным применение TNS-фильтра на сигнале понижающего микширования и остаточном сигнале, т.е. перед матрицей повышающего микширования. Иными словами, TNS применяется к линейным комбинациям левого и правого каналов, что имеет некоторые преимущества. Во-первых, может оказаться, что в данной ситуации TNS полезно только для, скажем, сигнала понижающего микширования. Тогда для остаточного сигнала TNS-фильтрация может подавляться, или пропускаться, и необходимо передавать коэффициенты TNS-фильтра только для сигнала понижающего микширования, что может означать более экономное использование доступной полосы пропускания. Во-вторых, вычисление передискретизированного представления сигнала понижающего микширования (например, данных MDST, получаемых исходя из данных MDCT так, чтобы можно было сформировать комплексное представление в частотной области), которое необходимо для кодирования с комплексным предсказанием, может потребовать того, чтобы можно было вычислить представление сигнала понижающего во временной области. В свою очередь, это означает, что сигнал понижающего микширования предпочтительно должен быть доступен как временная последовательность спектров MDCT, полученных единообразным образом. Если TNS-фильтр применялся в декодере после матрицы повышающего микширования, которая преобразовывает представление понижающего микширования/остаточное представление в левое/правое представление, будет доступна только последовательность оставшихся после TNS спектров MDCT сигнала понижающего микширования. Это может сделать эффективное вычисление соответствующих спектров MDST весьма затруднительным, особенно если левый и правый каналы использовали TNS-фильтры с отличающимися характеристиками.

Следует подчеркнуть, что доступность временной последовательности спектров MDCT не является абсолютным критерием для получения MDST-представления, подходящего для того, чтобы оно служило основой для кодирования с комплексным

предсказанием. В дополнение к экспериментальным свидетельствам этот факт можно объяснить тем, что TNS в общем применимо только к более высоким частотам, таким как частоты выше нескольких килогерц, поэтому остаточный сигнал, фильтрованный посредством TNS, приблизительно соответствует нефильтрованному остаточному сигналу для менее высоких частот. Таким образом, изобретение может быть осуществлено как декодер для стереофонического кодирования с комплексным предсказанием, в котором, как указывается ниже, TNS-фильтры имеют и иное размещение, чем в восходящем направлении относительно ступени повышающего микширования.

В одном из вариантов осуществления изобретения система декодера включает, по меньшей мере, один дополнительный модуль TNS, расположенный в нисходящем направлении относительно ступени повышающего микширования. Посредством положения селектора выбирается или модуль (модули) TNS в восходящем направлении относительно ступени повышающего микширования, или модуль (модули) TNS в нисходящем направлении относительно ступени повышающего микширования. В некоторых обстоятельствах вычисление комплексного представления в частотной области не требует того, чтобы могло вычисляться представление сигнала понижающего микширования во временной области. Кроме того, как излагается ниже, декодер может селективно действовать в режиме прямого или совместного кодирования, не применяя кодирование с комплексным предсказанием, и тогда может оказаться более подходящим применение традиционного местоположения модулей TNS, т.е. как одного из последних этапов обработки в частотной области.

В одном из вариантов осуществления изобретения система декодера адаптирована для экономии ресурсов обработки и, возможно, энергии путем отключения модуля, предназначенного для вычисления второго представления сигнала понижающего микширования в частотной области, когда последнее не является необходимым. Предполагается, что сигнал понижающего микширования является разделенным на последовательные временные блоки, каждый из которых связан со значением коэффициента комплексного предсказания. Это значение может определяться посредством решения, принимаемого для каждого временного блока кодером во взаимодействии с декодером. Кроме того, в данном варианте осуществления изобретения модуль, предназначенный для вычисления второго представления сигнала понижающего микширования в частотной области, адаптируется для самостоятельного отключения, если для данного временного блока абсолютное значение мнимой части коэффициента комплексного предсказания равна нулю или не превышает предварительно определенный допуск. Отключение модуля может подразумевать, что для данного временного блока второе представление сигнала понижающего микширования в частотной области не вычисляется. Если отключение не происходит, второе представление в частотной области (например, набор коэффициентов MDST) может умножаться на нуль или на число, имеющее, в значительной мере, тот же порядок величины, что и машинное эпсилон (единица округления) декодера, или какое-либо другое подходящее пороговое значение.

В дальнейшем развитии предшествующего варианта осуществления изобретения экономия ресурсов обработки достигается на подуровне временного блока, на которые разделяется сигнал понижающего микширования. Например, таким подуровнем в пределах временного блока может быть полоса частот, где декодер определяет значение коэффициента комплексного предсказания для каждой полосы частот в пределах временного блока. Сходным образом модуль, предназначенный для генерирования

второго представления в частотной области, адаптируется для подавления его действия для полосы частот в пределах временного блока, где коэффициент комплексного предсказания равен нулю или имеет абсолютное значение меньше допуска.

В одном из вариантов осуществления изобретения первые спектральные составляющие представляют собой коэффициенты преобразования, расположенные в одном или нескольких временных блоках коэффициентов преобразования, где каждый блок генерируется путем применения преобразования к временному отрезку сигнала во временной области. Кроме того модуль, предназначенный для вычисления второго представления сигнала понижающего микширования в частотной области, адаптирован для

• получения одной или нескольких первых промежуточных составляющих из, по меньшей мере, некоторых первых спектральных составляющих;

• формирования комбинации указанного одного или нескольких первых спектральных составляющих в соответствии с, по меньшей мере, частью одной или нескольких импульсных характеристик с целью получения одной или нескольких вторых промежуточных составляющих; и

• получения указанной одной или нескольких вторых спектральных составляющих из указанной одной или нескольких вторых промежуточных составляющих.

Данная процедура выполняет вычисление второго представления в частотной области непосредственно из первого представления в частотной области, как более подробно описано в патенте США №6980933 B2, в особенности в разделах 8-28 и, в частности, в уравнении 41. Как понятно специалистам, вычисление не выполняется через временную область в отличие, например, от обратного преобразования, за которым следует другое преобразование.

Для примера реализации стереофонического кодирования с комплексным предсказанием согласно изобретению было оценено, что вычислительная сложность возрастает лишь незначительно (значительно меньше, чем возрастание, вызываемое стереофоническим кодированием с комплексным предсказанием в области QMF) по сравнению с традиционным L/R- или M/S-стерео. Один из вариантов осуществления изобретения этого типа, включающий точное вычисление вторых спектральных составляющих, вносит задержку, которая, как правило, лишь на несколько процентов длительнее, чем задержка, вносимая реализацией на основе QMF (полагая длину временного блока равной 1024 дискретных значений и сравнивая ее с задержкой гибридного блока анализирующих/синтезирующих QMF-фильтров, которая равна 961 дискретных значений).

Соответственно, по меньшей мере, в некоторых из предыдущих вариантов осуществления изобретения импульсные характеристики адаптируются к преобразованию, посредством которого можно получить первое представление в частотной области, и, точнее, адаптируются согласно его частотным характеристикам.

В некоторых вариантах осуществления изобретения первое представление сигнала понижающего микширования в частотной области получается путем преобразования, которое применяется в связи с одной или несколькими анализирующими оконными функциями (или обрезными функциями, например, прямоугольным окном, синусным окном, окном, производным от окна Кайзера-Бесселя, и т.д.), одной из целей которых является временная сегментация без внесения пагубного количества шума или нежелательного изменения спектра. Возможно, указанные оконные функции являются частично перекрывающимися. В таком случае, предпочтительно, частотные характеристики преобразования зависят от характеристик указанной одной или

нескольких анализирующих оконных функций.

Продолжая обращаться к вариантам осуществления изобретения, включающим в качестве характерного признака вычисление второго представления в частотной области в пределах частотной области, можно понизить вычислительную нагрузку, что заключается в использовании приближенного второго представления в частотной области. Указанное приближение может выполняться путем отказа от требования полной информации, на которой основывается вычисление. Согласно идеям патента США №6980933 B2, например, для точного вычисления второго представления сигнала понижающего микширования в частотной области в одном блоке, в первую очередь, требуются данные в частотной области из трех временных блоков, а именно: блока, одновременного с выходным блоком, предшествующего блока и последующего блока. Для целей кодирования с комплексным предсказанием согласно настоящему изобретению пригодные приближения могут быть получены путем пропуска - или замещения на нуль - данных, происходящих из последующего блока (посредством чего действие модуля может стать причинным, т.е. он не будет вносить задержку) и/или из предшествующего блока так, чтобы вычисление второго представления в частотной области основывалось на данных только из одного или двух временных блоков. Следует отметить, что даже если пропуск входных данных может подразумевать изменение масштаба второго представления в частотной области - в том смысле, что, например, оно больше не будет представлять равную мощность, - оно все еще может быть использовано в качестве основы для кодирования с комплексным предсказанием, поскольку оно, как указывалось выше, вычисляется на концах кодера и декодера эквивалентным образом. В действительности возможное изменение масштаба такого типа будет компенсироваться путем соответствующего изменения значения коэффициента предсказания.

Еще один приближенный способ вычисления части второго представления сигнала понижающего микширования в частотной области, образующей спектральные составляющие, может включать объединение, по меньшей мере, двух составляющих из первого представления в частотной области. Последние составляющие могут быть смежными во времени и/или по частоте. В качестве альтернативы, они могут объединяться посредством фильтрации с импульсной характеристикой конечной длительности (FIR) с относительно небольшим количеством звеньев. Например, в системе, применяющей размер временного блока 1024, указанные фильтры могут включать 2, 3, 4 и т.д. звеньев. Описания способов приближенного вычисления такого рода можно найти, например, в заявке на патент США №2005/0197831 A1. Если используется оконная функция, такая как, например, непрямоугольная функция, придающая относительно малые веса окружению каждой из границ временного блока, может оказаться целесообразным основывать вторые спектральные составляющие во временном блоке только на комбинациях первых спектральных составляющих в том же временном блоке, подразумевая, что для краевых составляющих доступно не такое же количество информации. Ошибка приближения, возможно, вносимая при такой практической реализации, в некоторой степени подавляется, или скрывается, посредством формы оконной функции.

В одном из вариантов осуществления декодера, который сконструирован для вывода стереофонического сигнала во временной области, в него включена возможность переключения между прямым, или совместным, стереофоническим кодированием и кодированием с комплексным предсказанием. Это достигается путем снабжения:

• переключателем, который может селективно действовать или как ретранслирующая

ступень (не модифицирующая сигналы), или как суммарно-разностное преобразование;

• ступенью обратного преобразования, предназначенной для выполнения преобразования «частота-время»; и

• селекторной схемой, предназначенной для подачи на ступень обратного преобразования или прямо (или совместно) кодированного сигнала, или сигнала, кодированного с комплексным предсказанием.

Как понятно специалистам, указанная гибкость части декодера предоставляет кодеру свободу выбора между традиционным прямым, или совместным, кодированием и кодированием с комплексным предсказанием. Поэтому в тех случаях, когда уровень качества традиционного L/R-стереофонического кодирования, или совместного M/S-стереофонического кодирования, не может быть повышен, данный вариант осуществления изобретения может, по меньшей мере, гарантировать, что будет поддерживаться тот же уровень. Таким образом, декодер согласно изобретению можно считать расширенным по отношению к текущему уровню техники.

Другая группа вариантов осуществления изобретения системы декодера выполняет вычисление вторых спектральных составляющих во втором представлении в частотной области через временную область. Точнее, применяется преобразование, обратное тому, посредством которого получаются (или могут быть получены) первые спектральные составляющие, за которым следует другое преобразование, дающее на выходе вторые спектральные составляющие. В частности, за обратным MDCT может следовать MDST. Для того чтобы уменьшить количество преобразований и обратных преобразований, выходной сигнал обратного MDCT в таком варианте осуществления изобретения может подаваться и к MDST, и к терминалам вывода данных системы декодирования (которым, возможно, предшествуют дополнительные этапы обработки).

Для примера реализации кодирования с комплексным предсказанием согласно изобретению было оценено, что вычислительная сложность по сравнению с традиционным L/R- или M/S-стереофоническим кодированием возрастает лишь незначительно (в еще меньшей степени, чем возрастание, вызываемое кодированием с комплексным предсказанием в области QMF).

Как дополнительное развитие варианта осуществления изобретения, относящегося к предшествующему параграфу, ступень повышающего микширования может включать дополнительную ступень обратного преобразования, предназначенную для обработки побочного сигнала. В этом случае суммарно-разностная ступень снабжается представлением побочного сигнала во временной области, которое генерируется указанной дополнительной ступенью обратного преобразования, и представлением сигнала понижающего микширования во временной области, которое генерируется уже упоминавшейся ступенью обратного преобразования. Следует повторно отметить, что преимущественной с точки зрения вычислительной сложности является подача последнего сигнала и на суммарно-разностную ступень, и на указанную другую ступень преобразования, которая упоминалась выше.

В одном из вариантов осуществления изобретения декодер, сконструированный для вывода стереофонического сигнала во временной области, включает возможность переключения между L/R-стереофоническим кодированием, или совместным M/S-стереофоническим кодированием, и кодированием с комплексным предсказанием. Это достигается путем оснащения:

• переключателем, способным действовать или как ретранслирующая ступень, или как суммарно-разностная ступень;

• дополнительной ступенью обратного преобразования, предназначенной для

вычисления представления побочного сигнала во временной области;

• селекторной схемой, предназначенной для связывания ступеней обратного преобразования или с суммарно-разностной ступенью, связанной с точкой в восходящем направлении относительно ступени повышающего микширования и в нисходящем направлении относительно переключателя (предпочтительно, когда переключатель приводится в действие для того, чтобы выполнять функцию полосового фильтра, как может быть в случае декодирования стереофонического сигнала, генерируемого путем кодирования с комплексным предсказанием), или для объединения сигнала понижающего микширования из переключателя с побочным сигналом из взвешенного сумматора (предпочтительно, когда переключатель приводится в действие для выполнения функции суммарно-разностной ступени, как может быть в случае декодирования стереофонического сигнала, закодированного напрямую).

Как понятно специалистам, это предоставляет кодеру свободу выбора между традиционным прямым, или совместным, кодированием и кодированием с комплексным предсказанием, что подразумевает возможность гарантии того, что уровень качества будет, по меньшей мере, эквивалентен уровню при прямом, или совместном, стереофоническом кодировании.

В одном из вариантов осуществления изобретения система кодера согласно второй особенности изобретения может включать оцениватель, предназначенный для оценки коэффициента комплексного предсказания с целью снижения, или минимизации, мощности сигнала, или средней мощности сигнала, для остаточного сигнала. Минимизация может происходить по промежутку времени, предпочтительно, временному отрезку, или временному блоку, или временному кадру сигнала, который подвергается кодированию. В качестве критерия мгновенной мощности сигнала может быть выбран квадрат амплитуды, а интеграл квадрата амплитуды (формы сигнала) по промежутку времени может быть выбран в качестве критерия средней мощности сигнала в этом промежутке. Соответственно, коэффициент комплексного предсказания определяется на основе временного блока и полосы частот, т.е. его значение устанавливается таким образом, чтобы он понижал среднюю мощность (т.е. полную энергию) остаточного сигнала в этом временном блоке и полосе частот. В частности, выходной сигнал, на котором в соответствии с математическими зависимостями, известными специалистам, может вычисляться коэффициент комплексного предсказания, могут создавать такие модули, предназначенные для оценки параметров параметрического стереофонического кодирования, как IID, ICC и IPD или сходные с ними.

В одном из вариантов осуществления изобретения ступень кодирования в системе кодера также может действовать для выполнения функции ретранслирующей ступени так, чтобы делать возможным прямое стереофоническое кодирование. Выбирая прямое стереофоническое кодирование в ситуациях, когда ожидается обеспечение более высокого качества, система кодера может гарантировать, что закодированный стереофонический сигнал будет иметь, по меньшей мере, такое же качество, как при прямом кодировании. Сходным образом, в ситуациях, когда бóльшая вычислительная трудоемкость, привносимая кодированием с комплексным предсказанием, не мотивирована значительным повышением качества, для системы кодера, таким образом, является легкодоступной возможность экономии вычислительных ресурсов. Принятие решения о выборе между совместным кодированием, прямым кодированием, кодированием с действительным предсказанием и кодированием с комплексным предсказанием в кодере, главным образом, основывается на соображениях оптимизации

соотношения скорость/искажения.

В одном из вариантов осуществления изобретения система кодера может включать модуль, предназначенный для вычисления второго представления в частотной области непосредственно (т.е. без применения обратного преобразования во временную область и без использования данных сигнала во временной области) на основе первых спектральных составляющих. Относительно соответствующих вариантов осуществления системы декодера, описанной выше, данный модуль может иметь аналогичную конструкцию, а именно: включать аналогичные операции обработки, но в другом порядке так, чтобы кодер был адаптирован для вывода данных, пригодных в качестве входных на стороне декодера. С целью иллюстрации данного варианта осуществления изобретения предполагается, что стереофонический сигнал, который подвергается кодированию, включает средний и побочный каналы или был преобразован в такую конструкцию, и ступень кодирования адаптирована для приема первого представления в частотной области. Ступень кодирования включает модуль, предназначенный для вычисления второго представления среднего сигнала в частотной области. (Упоминаемые здесь первое и второе представления в частотной области аналогичны определенным выше; в частности, первые представления в частотной области могут представлять собой MDCT- представления, а второе представление в частотной области может представлять собой MDST-представление.) Ступень кодирования также включает взвешенный сумматор, предназначенный для вычисления остаточного сигнала как линейной комбинации, образованной из побочного сигнала и двух представлений среднего сигнала в частотной области, взвешенных, соответственно, по действительным и мнимым частям коэффициента комплексного предсказания. Средний сигнал, или его пригодное первое представление в частотной области, может использоваться непосредственно в качестве сигнала понижающего микширования. Кроме того, в данном варианте осуществления изобретения оцениватель определяет значение коэффициента комплексного предсказания с целью минимизации мощности, или средней мощности, остаточного сигнала. Последняя операция (оптимизации) может выполняться или посредством управления с обратной связью, где оцениватель может принимать остаточный сигнал, полученный посредством текущих значений коэффициента предсказания и предназначенный для дальнейшей регулировки в случае необходимости, или, по способу прямой связи - путем вычислений, выполняемых непосредственно на левом/правом каналах оригинального стереофонического сигнала или на среднем/побочном каналах. Предпочтительным является способ прямой связи, по которому коэффициент комплексного предсказания определяется непосредственно (в частности, не итеративно и без обратной связи) на основе первого и второго представлений среднего сигнала в частотной области и первого представления побочного сигнала в частотной области. Следует отметить, что за определением коэффициента комплексного предсказания может следовать принятие решения о том, какое применять кодирование: прямое, совместное, кодирование с действительным или с комплексным предсказанием, - где учитывается результирующее качество (предпочтительно, воспринимаемое качество с учетом, например, сигнал/маска) для каждой из доступных возможностей; поэтому приведенные выше утверждения не следует толковать как утверждения того, что в кодере не существует механизма обратной связи.

В одном из вариантов осуществления изобретения система кодера включает модули, предназначенные для вычисления второго представления среднего (или понижающего микширования) сигнала в частотной области через временную область. Следует понимать, что подробности реализации, относящиеся к этому варианту осуществления

изобретения, по меньшей мере, в той мере, в какой рассматривается вычисление второго представления в частотной области, сходны или могут отрабатываться аналогично соответствующим вариантам осуществления декодера. В данном варианте осуществления изобретения ступень кодирования включает:

• суммарно-разностную ступень, предназначенную для преобразования стереофонического сигнала в форму, включающую средний и побочный каналы;

• ступень преобразования, предназначенную для создания представления побочного канала в частотной области и комплекснозначного (и поэтому передискретизированного) представления среднего канала в частотной области; и

• взвешенный сумматор, предназначенный для вычисления остаточного сигнала, где в качестве весового коэффициента используется коэффициент комплексного предсказания.

Здесь оцениватель может принимать остаточный сигнал и определять, возможно, по способу управления с обратной связью, коэффициент комплексного предсказания так, чтобы понижать, или минимизировать, мощность, или среднюю мощность, остаточного сигнала. Предпочтительно, однако, оцениватель принимает стереофонический сигнал, который подвергается кодированию, и определяет на его основе коэффициент предсказания. С точки зрения вычислительной экономии преимущественным является использование критически дискретизированного представления побочного канала в частотной области, поскольку последнее в данном варианте осуществления изобретения не будет подвергаться умножению на комплексное число. Соответственно, ступень преобразования может включать ступень MDCT и расположенную параллельно ступень MDST, где обе ступени принимают в качестве входного сигнала представление среднего канала во временной области. Таким образом, генерируется передискретизированное представление среднего канала в частотной области и критически дискретизированное представление побочного канала в частотной области.

Следует отметить, что способы и устройство, раскрытые в данном разделе, после соответствующих модификаций в пределах возможностей специалистов, включая типовые эксперименты, могут применяться для кодирования сигналов, содержащих больше двух каналов. Модификации, предназначенные для обеспечения пригодности к указанной многоканальной эксплуатации, могут следовать, например, по пути направлений, описанных в разделах 4 и 5 процитированной выше статьи J. Herre и др.

Характерные признаки из двух или большего количества описанных выше вариантов осуществления изобретения могут комбинироваться в дальнейшие варианты осуществления изобретения, если они не дополнительны в явном виде. Тот факт, что два характерных признака излагаются в различных пунктах формулы изобретения, не препятствует тому, чтобы они могли выигрышно комбинироваться. Аналогично, дальнейшие варианты осуществления изобретения также могут предусматриваться с пропуском некоторых характерных признаков, которые не являются необходимыми или несущественны для намеченной цели. Например, система декодирования согласно изобретению может осуществляться без ступени деквантования в тех случаях, когда кодированный сигнал, подвергаемый обработке, не является квантованным или уже доступен в форме, пригодной для обработки ступенью повышающего микширования.

Краткое описание чертежей

Ниже изобретение будет дополнительно проиллюстрировано посредством вариантов осуществления изобретения, описываемых в следующем разделе, с отсылкой к сопроводительным чертежам, на которых:

фигура 1 состоит из двух обобщенных блок-схем (1A и 1B), показывающих декодеры на основе QMF согласно текущему уровню техники;

фигура 2 - обобщенная блок-схема системы стереофонического декодера на основе MDCT с комплексным предсказанием в соответствии с одним из вариантов осуществления настоящего изобретения, где комплексное представление канала сигнала, подвергаемого декодированию, вычисляется в частотной области;

фигура 3 - обобщенная блок-схема системы стереофонического декодера на основе MDCT с комплексным предсказанием в соответствии с одним из вариантов осуществления настоящего изобретения, где комплексное представление канала сигнала, подвергаемого декодированию, вычисляется во временной области;

фигура 4 показывает альтернативный вариант осуществления системы декодера по фигуре 2, где может быть выбрано положение активной ступени TNS;

фигура 5 включает обобщенные блок-схемы, показывающие системы стереофонического кодера на основе MDCT с комплексным предсказанием в соответствии с вариантами осуществления другой особенности настоящего изобретения;

фигура 6 - обобщенная блок-схема системы стереофонического кодера на основе MDCT с комплексным предсказанием в соответствии с одним из вариантов осуществления настоящего изобретения, где комплексное представление канала сигнала, подвергаемого кодированию, вычисляется на основе его представления во временной области;

фигура 7 показывает альтернативный вариант осуществления системы кодера по фигуре 6, который может действовать также и в режиме прямого L/R-кодирования;

фигура 8 - обобщенная блок-схема системы стереофонического кодера на основе MDCT с комплексным предсказанием в соответствии с одним из вариантов осуществления настоящего изобретения, где комплексное представление канала сигнала, подвергаемого декодированию, вычисляется на основе его первого представления в частотной области, система декодера которого может действовать также и в режиме прямого L/R-кодирования;

фигура 9 показывает альтернативный вариант осуществления системы кодера по фигуре 7, который также включает ступень TNS, расположенную в нисходящем направлении относительно ступени кодирования;

фигура 10 показывает альтернативные варианты осуществления части, отмеченной на фигурах 2-8 как A;

фигура 11 показывает альтернативный вариант осуществления системы кодера по фигуре 8, который также включает два устройства модификации в частотной области, расположенные, соответственно, в нисходящем и восходящем направлениях относительно ступени кодирования;

фигура 12 - графическое представление результатов испытательного прослушивания при 96 Кбит/с для шести объектов, показывающее различные возможности компромисса между сложностью и качеством при вычислении, или приближении, спектра MDST, где точки данных, отмеченные "+", относятся к скрытому эталону, "x" относится к якорю с полосой, ограниченной 3,5 КГц, "*" относится к традиционному стереофоническому кодированию USAC (M/S или L/R), "□" относится к унифицированному стереофоническому кодированию в области MDCT посредством комплексного предсказания с заблокированной мнимой частью коэффициента предсказания (т.е. с действительным предсказанием, не требующим MDST), "■" относится к унифицированному стереофоническому кодированию в области MDCT посредством комплексного предсказания с использованием для вычисления приближения MDST

текущего кадра MDCT, "◦" относится к унифицированному стереофоническому кодированию в области MDCT посредством комплексного предсказания с использованием для вычисления приближения MDST текущего и предыдущего кадров MDCT, и "●" относится к унифицированному стереофоническому кодированию в области MDCT посредством комплексного предсказания с использованием для вычисления MDST текущего, предыдущего и следующего кадров MDCT;

фигура 13 представляет данные по фигуре 12, но как дифференциальную оценку в отношении унифицированного стереофонического кодирования в области MDCT посредством комплексного предсказания с использованием для вычисления приближения MDST текущего кадра MDCT;

фигура 14 включает обобщенные блок-схемы (14A-14C), показывающие три варианта осуществления системы декодера в соответствии с вариантами осуществления изобретения;

фигура 15 - схема последовательности операций, показывающая способ декодирования в соответствии с одним из вариантов осуществления изобретения; и

фигура 16 - схема последовательности операций, показывающая способ кодирования в соответствии с одним из вариантов осуществления изобретения.

Подробное описание вариантов осуществления изобретения

I. Системы декодера

Фигура 2 в форме обобщенной блок-схемы показывает систему декодирования, предназначенную для декодирования битового потока, включающего, по меньшей мере, одно значение коэффициента комплексного предсказания $\alpha = \alpha_R + i\alpha_I$ и MDCT-представление стереофонического сигнала, содержащего канал M понижающего микширования и остаточный канал D. Действительная и мнимая части α_R, α_I коэффициента предсказания могут квантоваться и/или кодироваться совместно. Предпочтительным, однако, является, чтобы действительная и мнимая части квантовались независимо и единообразно, как правило, с величиной шага 0,1 (безразмерное число). В соответствии со стандартом MPEG разрешающая способность спектральной полосы, используемая для коэффициента комплексного предсказания, необязательно равна разрешающей способности для полос масштабных коэффициентов (sfb; т.е. группы линий MDCT, которая использует такую же величину шага квантования и диапазон квантования). В частности, разрешающая способность полосы частот для коэффициента предсказания может быть обоснованной психоакустически, как, например, шкала Барка. Демультиплексор 201 адаптирован для извлечения указанных MDCT-представлений и коэффициента предсказания (части показанной на фигуре Управляющей информации) из битового потока, который в него подается. На самом деле, в битовом потоке может быть закодировано больше управляющей информации, чем только информация коэффициента комплексного предсказания, например, команды о том, следует декодировать битовый поток в режиме с предсказанием или без предсказания, информация TNS и т.д. Информация TNS может включать значения параметров TNS, предназначенных для применения (синтезирующими) TNS-фильтрами системы декодера. Если для нескольких TNS-фильтров предполагается использование одинаковых наборов параметров TNS, более экономным является прием этой информации в форме бита, указывающего такую идентичность наборов параметров, а не независимый прием двух наборов параметров. Также может быть включена информация о том, применять TNS перед ступенью или после ступени повышающего микширования, по необходимости на основе, например, психоакустической оценки этих двух возможностей. Кроме того, тогда управляющая информация может указывать ограниченные по отдельности полосы

пропускания для сигнала понижающего микширования и остаточного сигнала. Для каждого канала полосы частот выше границы полосы пропускания не будут декодироваться, но будут приравниваться нулю. В некоторых случаях самые высокочастотные полосы обладают настолько малым запасом энергии, что они уже являются квантованными в нуль. В стандартной практике (ср. с параметром max_sfb в стандарте MPEG) обычно используется одинаковое ограничение полосы пропускания и для сигнала понижающего микширования, и для остаточного сигнала. Однако остаточный сигнал в большей степени, чем сигнал понижающего микширования, содержит свой запас энергии локализованным в полосах менее высоких частот. Поэтому, размещая специально предназначенный верхний предел полосы пропускания на остаточном сигнале, можно понизить битовую скорость передачи данных без значительного снижения качества. Например, это может управляться двумя независимыми параметрами max_sfb, закодированными в битовом потоке: одним - для сигнала понижающего микширования, и одним - для остаточного сигнала.

В данном варианте осуществления изобретения MDCT-представление стереофонического сигнала сегментируется на последовательные временные кадры (временные блоки), включающие фиксированное количество точек данных (например, 1024 точек), одно из нескольких фиксированных количеств точек данных (например, 128 или 1024 точек) или переменное количество точек. Как известно специалистам, MDCT является критически дискретизированным. Выходной сигнал системы декодирования, указанный в правой части иллюстрации, представляет собой стереофонический сигнал во временной области, содержащий левый L и правый R каналы. Модули 202 деквантования адаптированы для манипуляций с битовым потоком, входящим в систему декодирования, или там, где это необходимо, с двумя битовыми потоками, получаемыми после демультиплексирования оригинального битового потока и соответствующими каждому из каналов, понижающего микширования и остаточному. Деквантованные сигналы каналов подаются в узел 203 переключения, действующий или в режиме ретрансляции, или в суммарно-разностном режиме, которые относятся к соответствующим матрицам преобразования

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

и

$$\frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Как будет дополнительно разъясняться в следующем параграфе, система декодера включает второй узел 205 переключения. Оба узла 203, 205 переключения, как и большинство других переключателей и узлов переключения в данном варианте осуществления изобретения и в вариантах осуществления изобретения, которые будут описаны ниже, действуют частотно-избирательным образом. Это делает возможным декодирование большого разнообразия режимов декодирования, например, декодирование зависящего от частоты L/R- или M/S-декодирования, как известно на текущем уровне техники. Поэтому декодер согласно изобретению можно считать расширенным по отношению к текущему уровню техники.

Пока что полагая, что узел 203 переключения находится в режиме ретрансляции, в данном варианте осуществления изобретения деквантованные сигналы проходят через соответствующие TNS-фильтры 204. TNS-фильтры 204 не существенны для действия системы декодирования и могут замещаться ретранслирующими элементами. После этого сигнал подается во второй узел 205 переключения, выполняющий ту же функцию,

что и узел 203 переключения, расположенный в восходящем направлении. При условии, что входные сигналы соответствуют описанным выше, и второй узел 205 переключения установлен в режим ретрансляции, выходной сигнал первого представляет собой сигнал канала понижающего микширования и сигнал остаточного канала. Сигнал понижающего микширования, по-прежнему представленный его последовательным во времени спектром MDCT, подается в преобразование 206 действительного в мнимое, адаптированное для вычисления на его основе спектра MDST сигнала понижающего микширования. В данном варианте осуществления изобретения один кадр MDST основывается на трех кадрах MDCT: одном предыдущем кадре, одном текущем (или одновременном) кадре и одном последующем кадре. Символически указывается - (Z^{-1}, Z), - что входная сторона преобразования 206 действительного в мнимое включает элементы задержки.

MDST-представление сигнала понижающего микширования, полученное из преобразования 206 действительного в мнимое, взвешивается по мнимой части α_I коэффициента предсказания и добавляется к MDCT-представлению сигнала понижающего микширования, взвешенному по действительной части α_R коэффициента предсказания, и MDCT-представлению остаточного сигнала. Два сложения и умножения выполняются умножителями и сумматорами 210, 211, совместно образующими (функционально) взвешенный сумматор, которые снабжаются значением коэффициента комплексного предсказания α, закодированным в битовом потоке, изначально принимаемом системой декодера. Коэффициент комплексного предсказания может определяться один раз для каждого временного кадра. Также он может определяться чаще, как, например, один раз для каждой полосы частот в пределах кадра, где полосы частот представляет собой психоакустически мотивированное разделение. Также он может определяться реже, как будет описываться ниже в связи с системами кодирования согласно изобретению. Преобразование 206 действительного в мнимое синхронизируется со взвешенным сумматором так, чтобы текущий кадр MDST сигнала канала понижающего микширования объединялся с одновременными кадрами MDCT для каждого из сигналов, сигнала канала понижающего микширования и сигнала остаточного канала. Сумма трех этих сигналов представляет собой побочный сигнал $S = \mathrm{Re}\{\alpha M\} + D$. В данном выражении M включает и MDCT-, и MDST- представления сигнала понижающего микширования, а именно: $M = M_{MDCT} - iM_{MDST}$, где $D = D_{MDCT}$ является действительнозначным. Таким образом, получается стереофонический сигнал, содержащий канал понижающего микширования и побочный канал, из которого суммарно-разностное преобразование 207 реконструирует левый и правый каналы следующим образом:

$$\begin{cases} L = M + S \\ R = M - S \end{cases}.$$

Эти сигналы представлены в области MDCT. Последним этапом системы декодирования является применение обратного MDCT 209 к каждому из каналов, посредством чего получается представление левого/правого стереофонического сигнала во временной области.

Возможная реализация преобразования 206 действительного в мнимое дополнительно описана в патенте США №6980933 B2 заявителя, как указано выше. По формуле 41 в этом патенте преобразование может быть выражено как фильтр с импульсной характеристикой конечной длительности, например, для четных точек:

$$S(2v) = \frac{2}{N}\sum_{p=0}^{N-1}|(-1)^{p+1}X_I(p) + X_{III}(p)|h_{I,III}(2v-p)$$
$$+\frac{4}{N}\sum_{l=0}^{N-1}X_{II}(2l+1)h_{II}(2v-2l-1),$$

где $S(2v)$ - $2v$-я точка данных MDST, X_I, X_{II}, X_{III} - данные MDCT для каждого из кадров, и N - длина кадра. Кроме того, $h_{I,III}, h_{II}$ - это импульсные характеристики, зависящие от применяемой оконной функции и поэтому определяемые для каждого выбора оконной функции, такой как прямоугольная, синусоидальная или производная от окна Кайзера-Бесселя, и для каждой длины кадра. Сложность этого вычисления можно понизить путем пропуска тех импульсных характеристик, которые имеют относительно меньший запас энергии и вносят относительно меньший вклад в данные MDST. В качестве альтернативы или расширения данной реализации сами импульсные характеристики могут укорачиваться, например, от полной длины кадра N до меньшего количества точек. Например, длина импульсной характеристики может уменьшаться от 1024 точек (звеньев) до 10 точек. Самым предельным усечением, которое еще можно считать значимым, является

$$S(v) = X_{II}(v+1) - X_{II}(v-1).$$

Другие прямые приближения можно найти в заявке на патент США №2005/0197831 A1.

Также можно уменьшить количество входных данных, на которых основывается вычисление. Для иллюстрации, преобразование 206 действительного в мнимое и его соединения в восходящем направлении, которые указываются участком, обозначенным на иллюстрации как «A», могут замещаться упрощенными вариантами, два из которых, A' и A", показаны на фигуре 10. Вариант A' создает приближенное мнимое представление сигнала. В этом случае вычисление MDST учитывает только текущий и предыдущий кадры. С отсылкой к формуле, приведенной выше в данном параграфе, это можно осуществить, задавая $X_{III}(p) = 0$ при p=0,...,N-1 (индекс III обозначает последующий временной кадр). Поскольку вариант A' не требует спектра MDCT для последующего кадра в качестве входных данных, вычисление MDST не подвергается какой-либо временной задержке. Очевидно, данное приближение несколько снижает точность полученного сигнала MDST, но также может подразумевать, что энергия этого сигнала станет пониженной; последний факт может быть полностью компенсирован путем увеличения α_l как результата сущности кодирования с предсказанием.

Также на фигуре 10 показан вариант A", который в качестве входного сигнала использует только данные MDCT для текущего кадра. Вполне возможно, вариант A" генерирует менее точное представление MDST, чем вариант A'. С другой стороны, он, как и вариант A', действует с нулевой задержкой и обладает менее высокой вычислительной сложностью. Как уже упоминалось, свойства кодирования формы сигнала не затрагиваются, поскольку системой кодера и системой декодера используется одно и то же приближение.

Следует отметить, что независимо от того, применяются варианты A, A' или A", или какое-либо их усовершенствование, необходимо вычислять только те участки спектра MDST, для которых мнимая часть коэффициента комплексного предсказания является ненулевой, $\alpha_l \neq 0$. В условиях практического применения это означает, что абсолютное значение $|\alpha_l|$ мнимой части должно быть больше предварительно определенного

порогового значения, которое может быть отнесено к единице округления используемого аппаратного обеспечения. В случае, когда мнимая часть коэффициента равна нулю для всех частотных полос в пределах временного кадра, для этого кадра нет необходимости в вычислении каких-либо данных MDST. Таким образом, соответственно, преобразование 206 вещественного в мнимое адаптируется к отклику на появление слишком малых значений $|\alpha_i|$ путем отсутствия генерирования выходного сигнала MDST, посредством чего можно сэкономить вычислительные ресурсы. Однако в вариантах осуществления изобретения, где для генерирования одного кадра данных MDST используется больше кадров, чем текущий кадр, любые элементы в восходящем направлении относительно преобразования 206 должны продолжать действовать соответственно даже тогда, когда спектр MDST не требуется - в частности, второй узел 205 переключения должен продолжать пересылку спектров MDCT - так, чтобы достаточное количество входных данных было уже доступно для преобразования 206 действительного в мнимое уже тогда, когда появится следующий временной кадр, связанный с ненулевым коэффициентом предсказания; разумеется, этим кадром может быть следующий кадр.

Возвращаясь к фигуре 2, функционирование системы декодирования описано в предположении, что оба узла 203, 205 переключения установлены в соответствующие режимы ретрансляции. Как будет описано ниже, система декодера также может декодировать сигналы, которые не являются кодированными с предсказанием. Для такого применения второй узел 205 переключения устанавливается в суммарно-разностный режим и, соответственно, как указано на иллюстрации, селекторная схема 208 будет установлена в нижнее положение, таким образом, обеспечивая подачу сигналов непосредственно к обратному преобразованию 209 из точки расположения источника между TNS-фильтрами 204 и вторым узлом 205 переключения. Для обеспечения корректного декодирования сигнал, соответственно, имеет в точке расположения источника L/R-форму. Поэтому для того, чтобы обеспечить подачу в преобразование действительного в мнимое корректного среднего сигнала (т.е. сигнала понижающего микширования) во все моменты времени (а не, скажем, периодически - левого сигнала), второй узел 205 переключения в ходе декодирования стереофонического сигнала, не подвергавшегося кодированию с предсказанием, предпочтительно устанавливается в суммарно-разностный режим. Как указывалось выше, кодирование с предсказанием можно для некоторых кадров заменить традиционным прямым, или совместным, кодированием, основываясь, например, на данных решения о соотношении скорости и качества звука. Итог этого решения может сообщаться из кодера в декодер различными способами, например, посредством значения специально предназначенного бита-индикатора в каждом кадре, или по отсутствию, или наличию, значения коэффициента предсказания. При установлении указанных фактов роль первого узла 203 переключения становится легко понятной. Действительно, в режиме кодирования без предсказания система декодера может обрабатывать сигналы и в соответствии с прямым (L/R) стереофоническим кодированием, и в соответствии с совместным (M/S) стереофоническим кодированием, и при действии первого узла 203 переключения или в режиме ретрансляции, или в суммарно-разностном режиме можно обеспечивать то, что точка расположения источника будет всегда обеспечена сигналом, кодированным напрямую. Очевидно, узел 203 переключения при функционировании в качестве суммарно-разностной ступени будет преобразовывать входной сигнал в M/S-форме во входной сигнал (подаваемый в необязательные TNS-фильтры 204) в L/R-форме.

Система декодера принимает сигнал о том, следует системе декодера декодировать

конкретный временной кадр в режиме кодирования с предсказанием или в режиме кодирования без предсказания. Сигнал о режиме без предсказания может передаваться посредством значения специально предназначенного бита-индикатора в каждом кадре или посредством отсутствия (или нулевого значения) коэффициента предсказания. Режим предсказания может сообщаться аналогично. Особенно благоприятная реализация, которая делает возможным обратный переход на пониженный уровень без каких-либо накладных затрат, использует зарезервированное четвертое значение двухбитного поля ms_mask_present (см. MPEG-2 AAC, документ ISO/IEC 13818-7), которое передается в каждом кадре и определяется следующим образом:

Определение ms_mask_present в USAC	
Значение	Смысл
00	L/R-кодирование для всех частотных полос
01	Для указания L/R или M/S используется один сигнальный бит на полосу частот
10	M/S-кодирование для всех частотных полос
11	Зарезервированный

При переопределении значения 11 как имеющего смысл «кодирование с комплексным предсказанием» декодер может действовать во всех режимах совместимости, в частности, в режимах M/S- и L/R-кодирования, без каких-либо потерь битовой скорости передачи данных и по-прежнему является доступным для приема сигнала, указывающего режим кодирования с комплексным предсказанием для соответствующих кадров.

Фигура 4 показывает систему декодера той же общей конструкции, что и показанная на фигуре 2, но включающую, однако, по меньшей мере, две другие конструкции. Во-первых, система по фигуре 4 включает переключатели 404, 411, которые делают возможным применение какой-либо ступени обработки, включающей модификацию в частотной области, в восходящем и/или нисходящем направлении относительно ступени повышающего микширования. Это выполняется, с одной стороны, посредством первого набора модификаторов 403 в частотной области (изображенных на данной фигуре как синтезирующие TNS-фильтры), предусматриваемых совместно с первым переключателем 404 в нисходящем направлении относительно модулей 401 деквантования и первого узла 402 переключения, но в нисходящем направлении относительно второго узла 405 переключения, расположенного в восходящем направлении непосредственно перед ступенью 406, 407, 408, 409 повышающего микширования. С другой стороны, система декодера включает второй набор модификаторов 410 в частотной области, которые предусматриваются совместно со вторым переключателем 411 в нисходящем направлении относительно ступени 406, 407, 408, 409 повышающего микширования, но в восходящем направлении относительно ступени 412 обратного преобразования. Преимущественно, как показано на иллюстрации, каждый модификатор в частотной области расположен параллельно с ретранслирующей линией, которая подключается в восходящем направлении относительно входной стороны модификатора в частотной области и в нисходящем направлении относительно связанного с ней переключателя. В силу такой конструкции модификатор в частотной области снабжается данными сигнала во все моменты времени, что делает возможной обработку в частотной области на основе большего количества кадров, чем только текущий кадр. Решение о том, применять первый 403 или второй 410 набор модификаторов в частотной области, может приниматься кодером (и передаваться в битовом потоке) или может основываться на том, применяется ли кодирование с предсказанием, или может основываться на каких-либо иных критериях, найденных подходящими в условиях практического применения. Например, если

модификаторами в частотной области являются TNS-фильтры, то для некоторых видов сигналов преимущественным является использование первого набора 403, в то время как использование второго набора 410 может оказаться преимущественным для других видов сигналов. Если итог этого выбора закодирован в битовом потоке, то система декодера будет соответственно приводить в действие соответствующий набор TNS-фильтров.

Для облегчения понимания системы декодера, показанной на фигуре 4, следует недвусмысленно отметить, что декодирование сигнала, кодированного напрямую (L/R), происходит тогда, когда $\alpha=0$ (это подразумевает, что псевдо-L/R и L/R идентичны, и что побочный и остаточный сигналы не отличаются), первый узел 402 переключения находится в режиме ретрансляции, второй узел переключения находится в суммарно-разностном режиме, что, таким образом, приводит к тому, что сигнал между вторым узлом 405 переключения и суммарно-разностной ступенью 409 ступени повышающего микширования имеет M/S-форму. Тогда, поскольку ступень повышающего микширования будет эффективно представлять собой ретранслирующую ступень, неважно, первый или второй набор модификаторов будет приводиться в действие (с использованием соответствующих переключателей 404, 411).

Фигура 3 иллюстрирует систему декодера согласно одному из вариантов осуществления изобретения, который, по отношению к вариантам по фигурам 2 и 4, представляет другой подход к созданию данных MDST, необходимых для повышающего микширования. Как и уже описанные системы декодера, система по фигуре 3 включает модули 301 деквантования, первый узел 302 переключения, действующий или в режиме ретрансляции, или в суммарно-разностном режиме, и (синтезирующие) TNS-фильтры 303, которые последовательно расположены начиная от входного конца системы декодера. Модули в нисходящем направлении от этой точки селективно используются посредством двух вторых переключателей 305, 310, которые предпочтительно действуют совместно так, чтобы они оба находились или в верхних положениях, или в нижних положениях, как указано на фигуре. На выходном конце системы декодера находится суммарно-разностная ступень 312 и, непосредственно перед ней в восходящем направлении, два модуля 306, 311 обратного MDCT, предназначенные для преобразования представления каждого из каналов в области MDCT в представление во временной области.

При декодировании с комплексным предсказанием, где система декодера снабжается битовым потоком, в котором закодированы сигнал понижающего микширования /остаточный стереофонический сигнал и значения коэффициента комплексного предсказания, первый узел 302 переключения устанавливается в режим ретрансляции, а вторые переключатели 305, 310 устанавливаются в верхнее положение. В нисходящем направлении относительно TNS-фильтров два канала (деквантованного, TNS-фильтрованного, MDCT) звукового сигнала обрабатываются по-разному. Сигнал понижающего микширования подается, с одной стороны, в умножитель и сумматор 308, который добавляет MDCT-представление канала понижающего микширования, взвешенное по действительной части α_R коэффициента предсказания, к MDCT-представлению остаточного канала и, с другой стороны, к одному из модулей 306 обратного MDCT-преобразования. Представление канала понижающего микширования M во временной области, которое является выходным сигналом из модуля 306 обратного MDCT-преобразования, подается и на конечную суммарно-разностную ступень 312, и в модуль 307 MDST-преобразования. Указанное двойное использование представления канала понижающего микширования во временной области является преимущественным

с точки зрения вычислительной сложности. Полученное таким образом MDST-представление канала понижающего микширования подается в следующий умножитель и сумматор 309, который после взвешивания по мнимой части α_I коэффициента предсказания добавляет этот сигнал к линейной комбинации на выходе сумматора 308; таким образом, выходной сигнал сумматора 309 представляет собой сигнал побочного канала, $S = \mathrm{Re}\{\alpha M\} + D$. Сходным образом с системой декодера, показанной на фигуре 2, умножители и сумматоры 308, 309 легко могут быть объединены с образованием взвешенного мультисигнального сумматора, который выводит MDCT- и MDST-представления сигнала понижающего микширования, MDCT-представление остаточного сигнала и значение коэффициента комплексного предсказания. В нисходящем направлении от этой точки в настоящем варианте осуществления изобретения перед тем, как сигнал побочного канала подается на конечную суммарно-разностную ступень 312, остается только прохождение через модуль 311 обратного MDCT-преобразования.

Необходимая синхронность в системе декодера может достигаться путем применения одинаковых длин преобразования и форм окон для обоих модулей 306, 311 обратного MDCT-преобразования, как уже применяется на практике в частотно-избирательном M/S- и L/R-кодировании. Комбинацией некоторых вариантов осуществления модуля 306 обратного MDCT-преобразования и некоторых вариантов осуществления модуля 307 MDST вводится задержка на один кадр. Поэтому предусматривается пять необязательных блоков 313 задержки (или команд программного обеспечения - для осуществления этого действия в компьютерной реализации) так, чтобы часть системы, расположенная справа от пунктирной линии при необходимости могла быть задержана на один кадр относительно левой части. Очевидно, все пересечения между пунктирной линией и линиями связи снабжаются блоками задержки за исключением линии связи между модулем 306 обратного MDCT и модулем 307 MDST-преобразования, которая находится там, где возникающая задержка нуждается в компенсации.

Вычисление данных MDST для одного временного кадра требует данных одного кадра из представления во временной области. Однако обратное MDCT-преобразование основывается на одном (текущем), двух (предпочтительно - предыдущем и текущем) или трех (предпочтительно - предыдущем, текущем и следующем) последовательных кадрах. В силу хорошо известного подавления побочных низкочастотных составляющих (TDAC), связанного с MDCT, трехкадровая возможность выполняет полное перекрывание входных кадров и, таким образом, обеспечивает наилучшую (и, вероятно, совершенную) точность, по меньшей мере, для кадров, содержащих побочные низкочастотные составляющие во временной области. Ясно, что трехкадровое обратное MDCT действует с задержкой на один кадр. Допуская использование приближенного представления во временной области в качестве входного сигнала в MDST-преобразование, можно избежать этой задержки и, таким образом, необходимости в компенсации задержек между различными частями системы декодера. В двухкадровой возможности в более ранней половине кадра возникает наложение/сложение, делающее возможным TDAC, и побочные низкочастотные составляющие могут присутствовать только в более поздней половине кадра. В однокадровой возможности отсутствие TDAC подразумевает, что в кадре могут появляться побочные низкочастотные составляющие; однако MDST-представление, достигаемое таким образом используемое в качестве промежуточного сигнала в кодировании с комплексным предсказанием, может по-прежнему обеспечивать удовлетворительное качество.

Система декодирования, проиллюстрированная на фигуре 3, также может действовать в двух режимах декодирования без предсказания. Для декодирования L/R-

стереофонического сигнала, кодированного напрямую, вторые переключатели 305, 310 устанавливаются в нижнее положение, и первый узел 302 переключения устанавливается в режим ретрансляции. Таким образом, сигнал имеет L/R-форму в восходящем направлении относительно суммарно-разностной ступени 304, которая преобразовывает его в M/S-форму, на которой происходит обратное MDCT-преобразование и суммарно-разностная операция. Для декодирования стереофонического сигнала, представленного в совместно кодированной M/S-форме, первый узел 302 переключения вместо этого устанавливается в суммарно-разностный режим так, чтобы сигнал между первым узлом 302 переключения и суммарно-разностной ступенью 304 имел L/R-форму, которая часто является более подходящей с точки зрения TNS-фильтрации, чем могла бы быть M/S-форма.

Обработка в нисходящем направлении относительно суммарно-разностной ступени 304 идентична обработке в случае прямого L/R-декодирования.

Фигура 14 состоит из трех обобщенных блок-схем декодеров согласно вариантам осуществления изобретения. В отличие от некоторых других блок-схем, сопровождающих данную заявку, линия связи на фигуре 14 может символизировать многоканальный сигнал. В частности, указанная линия связи может располагаться для передачи стереофонического сигнала, включающего левый/правый, средний/побочный, понижающего микширования/остаточный, псевдолевый/псевдоправый каналы и другие комбинации.

Фигура 14A показывает систему декодера, предназначенную для декодирования представления входного сигнала в частотной области (указываемого для целей данной фигуры как MDCT-представление). Система декодера адаптирована для передачи в качестве выходного сигнала представления стереофонического сигнала во временной области, которое генерируется на основе входного сигнала. Для того чтобы иметь возможность декодировать входной сигнал, закодированный посредством стереофонического кодирования с комплексным предсказанием, система декодера оснащается ступенью 1410 повышающего микширования. Однако она также способна манипулировать с входным сигналом, закодированным в других форматах, и, возможно, с сигналом, который чередуется во времени между несколькими форматами кодирования, например, за последовательностью временных кадров, закодированной посредством кодирования с комплексным предсказанием, может следовать временной участок, закодированный посредством прямого левого/правого кодирования. Способность системы декодера манипулировать с различными форматами кодирования достигается путем создания линии связи (ретрансляции), расположенной параллельно указанной ступени 1410 повышающего микширования. Посредством переключателя 1411 можно выбрать, выходной сигнал ступени 1410 повышающего микширования (нижнее положение переключателя на фигуре) или необработанный сигнал, доступный над линией связи (верхнее положение переключателя на фигуре), будет подаваться в модули декодера, расположенные дальше в нисходящем направлении. В данном варианте осуществления изобретения модуль 1412 обратного MDCT, который преобразовывает MDCT-представление сигнала в представление во временной области, располагается в нисходящем направлении относительно переключателя. Например, сигнал, подаваемый на ступень 1410 повышающего микширования, может представлять собой стереофонический сигнал в форме понижающего микширования /остаточной форме. Ступень 1410 повышающего микширования адаптирована для получения побочного сигнала и для выполнения суммарно-разностной операции так, чтобы выходным сигналом являлся левый/правый стереофонический сигнал (в области MDCT).

Фигура 14B показывает систему декодера, сходную с системой по фигуре 14A. Настоящая система адаптирована для приема битового потока в качестве входного сигнала. Вначале битовый поток обрабатывается комбинированным модулем 1420 демультиплексора и деквантования, который создает в качестве первого выходного сигнала MDCT-представление многоканального стереофонического сигнала для дальнейшей обработки, которая определяется положением переключателя 1422, имеющего ту же функцию, что и переключатель 1411 по фигуре 14A. Точнее, переключатель 1422 определяет, должен первый выходной сигнал из демультиплексора и деквантования обрабатываться ступенью 1421 повышающего микширования и модулем 1423 обратного MDCT (нижнее положение) или только модулем 1423 обратного MDCT (верхнее положение). Комбинированный модуль 1420 демультиплексора и деквантования также выводит управляющую информацию. В настоящем случае управляющая информация, связанная со стереофоническим сигналом, может включать данные, указывающие на то, какое из положений переключателя 1422, нижнее или верхнее, подходит для декодирования сигнала, или, более абстрактно, согласно какому формату кодирования должен декодироваться стереофонический сигнал. Управляющая информация также может включать параметры для регулировки свойств ступени 1421 повышающего микширования, например, значение коэффициента комплексного предсказания α, использованного при кодировании с комплексным предсказанием, как уже описано выше.

Фигура 14C показывает систему декодера, которая, в дополнение к объектам, аналогичным таковым на фигуре 14 B, включает, первое и второе устройства 1431, 1435 модификации в частотной области, соответственно, расположенные в восходящем и нисходящем направлениях относительно ступени 1433 повышающего микширования. Для целей данной фигуры каждое устройство модификации в частотной области проиллюстрировано TNS-фильтром. Однако под термином «устройство модификации в частотной области» также могут пониматься и иные процессы, чем TNS-фильтрация, для которых допустимо применение или перед ступенью, или после ступени повышающего микширования. Примеры модификаций в частотной области включают предсказание, накопление помех, растягивание полосы пропускания и нелинейную обработку. Психоакустические соображения и похожие причины, которые, возможно, включают свойства сигнала, подвергаемого обработке, и/или конфигурацию или настройки устройства модификации в частотной области, иногда указывают, что преимущественным является применение указанной модификации в частотной области в восходящем, а не в нисходящем направлении относительно ступени 1433 повышающего микширования. В других случаях путем похожих соображений может быть установлено, что расположение модификации в частотной области в нисходящем направлении более предпочтительно, чем расположение в восходящем направлении. Посредством переключателей 1432, 1436 устройства 1431, 1435 модификации в частотной области могут селективно приводиться в действие так, чтобы в ответ на управляющую информацию система декодера могла выбирать требуемую конфигурацию. Например, фигура 14C показывает конфигурацию, в которой стереофонический сигнал из комбинированного модуля 1430 демультиплексора и деквантования вначале обрабатывается первым устройством 1431 модификации в частотной области, затем подается к ступени 1433 повышающего микширования и в конце направляется прямо в модуль 1437 обратного MDCT без прохождения через второе устройство 1435 модификации в частотной области. Как разъясняется в разделе «Краткое описание изобретения», в кодировании с комплексным предсказанием данная конфигурация

является предпочтительной перед возможностью выполнения TNS после повышающего микширования.

II. Системы кодера

Система кодера согласно изобретению будет описана ниже с отсылкой к фигуре 5, которая представляет собой обобщенную блок-схему системы кодера, предназначенной для кодирования левого/правого (L/R) стереофонического сигнала как выходного битового потока посредством кодирования с комплексным предсказанием. Система кодера принимает представление сигнала во временной или в частотной области и подает его на ступень понижающего микширования и в оцениватель коэффициента предсказания. Действительные и мнимые части коэффициентов предсказания передаются на ступень понижающего микширования для управления преобразованием левого и правого каналов в канал понижающего микширования и остаточный канал, которые затем подаются в конечный мультиплексор MUX. Если сигнал не подается в кодер как представление в частотной области, то он преобразовывается в такое представление на ступени понижающего микширования или в мультиплексоре.

Одним из принципов кодирования с предсказанием является преобразование левого/правого сигнала в среднюю/побочную форму, т.е.:

$$\begin{cases} M = (L + R)/2 \\ S = (L - R)/2 \end{cases},$$

а затем использование остаточной корреляции между этими каналами, т.е. задание

$$S = \mathrm{Re}\{\alpha M\} + D,$$

где α - коэффициент комплексного предсказания, который требуется определить, и D - остаточный сигнал. Можно выбрать α так, чтобы энергия остаточного сигнала $D = S - \mathrm{Re}\{\alpha M\}$ минимизировалась. Минимизация энергии может производиться относительно мгновенной мощности, кратко- или долговременной энергии (средней мощности), что в случае дискретного сигнала равносильно оптимизации в среднеквадратическом значении.

Действительная и мнимая части α_R, α_I коэффициента предсказания могут квантоваться и/или кодироваться совместно. Предпочтительно, однако, чтобы действительная и мнимая части квантовались независимо и единообразно, как правило, с величиной шага 0,1 (безразмерное число). В соответствии со стандартом MPEG, разрешающая способность полосы частот, используемая для коэффициента комплексного предсказания, необязательно аналогична разрешающей способности для полос масштабных коэффициентов (sfb; т.е. группы линий MDCT, которые используют одну и ту же величину шага MDCT-квантования и диапазон квантования). В частности, разрешающая способность полосы частот для коэффициента предсказания может быть обоснована психоакустически, как, например, шкала Барка. Следует отметить, что разрешающая способность полосы частот может изменяться в случае изменения длины преобразования.

Как уже отмечалось, система кодера согласно изобретению может иметь свободу в том, применять стереофоническое кодирование с предсказанием или нет, где последний случай подразумевает обратный переход на пониженный уровень - к L/R- или M/S-кодированию. Указанное решение может приниматься на основе временного кадра или, тоньше, на основе полосы частот в пределах временного кадра. Как отмечалось выше, отрицательный результат принятия решения может сообщаться декодирующему объекту различными способами, например, посредством значения специально предназначенного бита-индикатора в каждом кадре, или по отсутствию (или нулевой величине) значения коэффициента предсказания. Положительное решение может

передаваться аналогично. Особенно преимущественная реализация, которая делает возможным обратный переход на пониженный уровень без каких-либо накладных затрат, использует зарезервированное четвертое значение двухбитного поля ms_mask_present (см. MPEG-2 AAC, документ ISO/IEC 131818-7), которое передается в каждом кадре и определяется следующим образом:

Определение ms_mask_present в USAC	
Значение	Смысл
00	L/R-кодирование для всех частотных полос
01	Для указания L/R- или M/S используется один сигнальный бит на полосу
10	M/S-кодирование для всех частотных полос
11	Зарезервированное

При переопределении значения 11 так, чтобы оно означало «кодирование с комплексным предсказанием», кодер может действовать во всех режимах совместимости, в частности, в режимах M/S- и L/R-кодирования, без какой-либо потери битовой скорости передачи данных и по-прежнему способен передавать сигнал о кодировании с комплексным предсказанием для тех кадров, где оно является преимущественным.

Соответствующее решение может основываться на соображениях о соотношении скорости передачи данных и качества звука. В качестве меры качества могут использоваться данные, полученные с использованием психоакустической модели, заключенной в кодере (как часто бывает в случае имеющихся в наличии кодеров звука на основе MDCT). В частности, некоторые варианты осуществления кодера предусматривают выбор коэффициента предсказания, оптимизированного на основе соотношения скорости и искажений. Соответственно, в таких вариантах осуществления изобретения мнимая часть, - а, возможно, также и действительная часть, - коэффициента предсказания приравнивается нулю в случае, если выигрыш от предсказания не экономит достаточно битов для кодирования остаточного сигнала так, чтобы это оправдывало затраты битов, необходимых для кодирования коэффициента предсказания.

Варианты осуществления кодера могут кодировать информацию, относящуюся к TNS, в битовом потоке. Эта информация может включать значения параметров TNS, которые применяются (синтезирующими) TNS-фильтрами на стороне декодера. Если для обоих каналов используются идентичные наборы параметров TNS, более экономным является включение сигнального бита, указывающего на эту идентичность наборов параметров, а не независимая передача двух наборов параметров. Также может быть включена информация о том, следует применять TNS перед или после ступени повышающего микширования, как должно на основании, например, психоакустической оценки обеих доступных возможностей.

В качестве еще одного необязательного характерного признака, который является потенциально полезным с точки зрения сложности и битовой скорости передачи данных, кодер может адаптироваться для использования отдельно ограниченной полосы пропускания при кодировании остаточного сигнала. Полосы частот выше этого предела не будут передаваться в декодер, но будут приравнены нулю. В некоторых случаях самые высокочастотные полосы имеют настолько низкое энергосодержание, что они квантуются в нуль. В стандартной практике (ср. с параметром max_sfb в стандарте MPEG) обычно используется одинаковое ограничение полосы пропускания и для сигнала понижающего микширования, и для остаточного сигнала. К настоящему времени авторы изобретения экспериментально обнаружили, что остаточный сигнал в большей степени, чем сигнал понижающего микширования, имеет энергосодержание, локализованное в полосах менее высоких частот. Поэтому, помещая специально

предназначенный верхний предел ширины полосы пропускания на остаточный сигнал, возможно снижение битовой скорости передачи данных без существенной потери качества. Например, это может выполняться путем передачи двух независимых параметров max_sfb: одного - для сигнала понижающего микширования, и одного - для остаточного сигнала.

Следует отметить, что, хотя вопросы оптимального определения коэффициента предсказания, его квантования и кодирования, обратного перехода на пониженный уровень в режим M/S или L/R, TNS-фильтрации, верхнего ограничения полосы пропускания и т.д. обсуждались в отношении системы декодера, показанной на фигуре 5, те же факты в равной степени применимы и к вариантам осуществления изобретения, который будут раскрыты ниже с отсылкой к следующим фигурам.

Фигура 6 показывает другую систему кодера согласно изобретению, адаптированную для выполнения стереофонического кодирования с комплексным предсказанием. Система принимает в качестве входного сигнала представление стереофонического сигнала во временной области, сегментированного на последовательные, возможно, перекрывающиеся временные кадры и включающего левый и правый каналы. Суммарно-разностная ступень 601 преобразовывает сигнал в средний и побочный каналы. Средний канал подается и в модуль 602 MDCT, и в модуль 603 MDST, в то время как побочный канал подается только в модуль 604 MDCT. Оцениватель 605 коэффициента предсказания оценивает для каждого временного кадра - и, возможно, для отдельных частотных полос в пределах кадра - значение коэффициента комплексного предсказания α, как это разъяснено выше. Значение коэффициента α подается в качестве весового коэффициента во взвешенные сумматоры 606, 607, которые формируют остаточный сигнал D как линейную комбинацию MDCT- и MDST-представлений среднего сигнала и MDCT-представления побочного сигнала. Предпочтительно, коэффициент комплексного предсказания подается во взвешенные сумматоры 606, 607 представленным в той же схеме квантования, которая будет использоваться при его кодировании в битовый поток; это, очевидно, обеспечивает более достоверную реконструкцию, поскольку и кодер, и декодер применяют одно и то же значение коэффициента предсказания. Остаточный сигнал, средний сигнал (который, более соответственно, называется сигналом понижающего микширования, когда он возникает в сочетании с остаточным сигналом) и коэффициент предсказания подаются на комбинированную ступень 608 квантования и мультиплексора, которая кодирует их и возможную дополнительную информацию в выходной битовый поток.

Фигура 7 показывает изменение системы кодера по фигуре 6. Как видно из сходства символов на фигуре, она имеет сходную конструкцию, но также содержит и дополнительную функцию действия в режиме обратного перехода на пониженный уровень прямого L/R-кодирования. Система кодера приводится в действие между режимом кодирования с комплексным предсказанием и режимом обратного перехода на пониженный уровень посредством переключателя 710, предусмотренного в восходящем направлении непосредственно перед комбинированной ступенью 709 квантования и мультиплексора. В верхнем положении, как показано на фигуре, переключатель 710 будет приводить к действию кодера в режиме обратного перехода на пониженный уровень. Из точек в нисходящем направлении непосредственно за модулями 702, 704 MDCT средний/побочный сигнал подается на суммарно-разностную ступень 705, которая после преобразования в левую/правую форму пропускает его к переключателю 710, который связывает его с комбинированной ступенью 709 квантования и мультиплексора.

Фигура 8 показывает систему кодера согласно настоящему изобретению. В отличие от систем кодера по фигурам 6 и 7, данный вариант осуществления изобретения получает данные MDST, необходимые для кодирования с комплексным предсказанием, непосредственно из данных MDCT, т.е. путем преобразования действительного в мнимое в частотной области. Преобразование действительного в мнимое применяет какой-либо из подходов, описанных в связи с системами декодера по фигурам 2 и 4. Важно, чтобы способ вычисления в декодере совпадал со способом вычисления в кодере так, чтобы могло выполняться достоверное декодирование; предпочтительно, на стороне кодера и на стороне декодера используются одинаковые способы преобразования действительного в мнимое. Как и для вариантов осуществления декодера, участок A, окруженный пунктирной линией и включающий преобразование 804 действительного в мнимое, может быть замещен приближенными вариантами с использованием в качестве входных данных меньшего количества входных временных кадров. Аналогично, кодирование может быть упрощено с использованием любого из описанных выше, или какого-либо другого, приближенного метода.

На более высоком уровне система 8 имеет конструкцию, которая отличается от конструкции, которая могла бы следовать за прямым действием по замещению модуля MDST на фигуре 7 (надлежащим образом подключенным) модулем преобразования действительного в мнимое. Настоящая архитектура понятна и выполняет функцию переключения между кодированием с предсказанием и прямым L/R-кодированием устойчивым и вычислительно экономичным способом. Входной стереофонический сигнал подается в модули 801 MDCT-преобразования, которые выводят представление каждого из каналов в частотной области. Оно подается и в конечный переключатель 808 для приведения в действие системы кодера между режимами кодирования с предсказанием и прямого кодирования, и на суммарно-разностную ступень 802. В прямом L/R-кодировании, или совместном M/S-кодировании, которое осуществляется во временном кадре, для которого коэффициент предсказания приравнен нулю, - данный вариант осуществления изобретения подвергает входной сигнал только MDCT-преобразованию, квантованию и мультиплексированию, где последние два этапа выполняются комбинированной ступенью 807 квантования и мультиплексора, расположенной на выходном конце системы, куда подается битовый поток. При кодировании с предсказанием каждый из каналов претерпевает дальнейшую обработку между суммарно-разностной ступенью 802 и переключателем 808. Из MDCT-представления среднего сигнала преобразование 804 действительного в мнимое получает данные MDST и направляет их и в оцениватель 803 коэффициента предсказания, и во взвешенный сумматор 806. Как и в системах кодера по фигурам 6 и 7, следующий взвешенный сумматор 805 используется для объединения побочного сигнала с взвешенными MDCT- и MDST-представлениями среднего сигнала с целью формирования сигнала остаточного канала, который кодируется совместно с сигналом среднего канала (т.е. канала понижающего микширования) и с коэффициентом предсказания при помощи комбинированного модуля 807 квантования и мультиплексора.

С отсылкой к фигуре 9 будет проиллюстрировано, что каждый из вариантов осуществления системы кодера может объединяться с одним или несколькими (анализирующими) TNS-фильтрами. В соответствии с предыдущими обсуждениями, часто более преимущественным является применение TNS-фильтрации к сигналу в его форме понижающего микширования. Поэтому, как показано на фигуре 9, адаптация системы кодера по фигуре 7 для включения в нее TNS выполняется путем добавления TNS-фильтров 911 в восходящем направлении непосредственно перед комбинированным

модулем 909 квантования и мультиплексора.

Вместо правого/остаточного TNS-фильтра 911b в нисходящем направлении непосредственно за участком переключателя 910, адаптированного для манипуляций с правым, или остаточным, каналом, могут предусматриваться два отдельных TNS-фильтра (не показаны). Таким образом, каждый из двух TNS-фильтров будет снабжаться данными сигнала соответствующего канала во все моменты времени, что делает возможной TNS-фильтрацию на основе одного или нескольких временных кадров, а не только текущего кадра. Как уже было отмечено, TNS-фильтры являются всего лишь примером устройств модификации в частотной области, в особенности устройств, основывающих обработку на большем количестве кадров, чем только текущий кадр, которые при таком расположении могут извлекать такую же или бóльшую выгоду, чем извлекает TNS-фильтр.

В качестве другой возможной альтернативы варианту осуществления изобретения, показанному на фигуре 9, TNS-фильтры, предназначенные для селективного приведения в действие, могут размещаться более чем в одной точке для каждого канала. Эта конструкция сходна с конструкцией системы декодера, показанной на фигуре 4, где разные наборы TNS-фильтров могут быть связаны посредством переключателей. Это позволяет делать выбор наиболее подходящей ступени при TNS-фильтрации для каждого временного кадра. В частности, может оказаться преимущественным переключение между различными положениями TNS в связи с переключением между стереофоническим кодированием с комплексным предсказанием и другими режимами кодирования.

Фигура 11 показывает изменение, основанное на системе кодера по фигуре 8, в котором посредством преобразования 1105 действительного в мнимое получается второе представление сигнала понижающего микширования в частотной области. Аналогично системе декодера, показанной на фигуре 4, данная система кодера также включает селективно приводимые в действие модули модификаторов в частотной области: один - 1102 - предусматриваемый в восходящем направлении относительно ступени понижающего микширования, и один - 1109 - предусматриваемый в нисходящем направлении относительно этой ступени. Модули 1102, 1109 в частотной области, которые на данной фигуре представлены на примере TNS-фильтров, могут связываться с каждым из трактов сигнала с использованием четырех переключателей 1103a, 1103b, 1109a и 1109b.

III. Неаппаратные варианты осуществления изобретения

Варианты осуществления третьей и четвертой особенностей изобретения показаны на фигурах 15 и 16. Фигура 15 показывает способ декодирования битового потока в стереофонический сигнал, который включает следующие этапы:

1. Входным сигналом является битовый поток.

2. Битовый поток деквантуется, посредством чего получается представление канала понижающего микширования и остаточного канала в частотной области.

3. Вычисляется второе представление канала понижающего микширования в частотной области.

4. На основе трех представлений каналов в частотной области вычисляется сигнал побочного канала.

5. На основе побочного канала и канала понижающего микширования вычисляется стереофонический сигнал, преимущественно, в левой/правой форме.

6. Выходным сигналом является полученный таким образом стереофонический сигнал.

Этапы 3-5 можно рассматривать как процесс повышающего микширования. Каждый

из этапов 1-6 аналогичен соответствующей функции в любой из систем декодера, раскрытых в предыдущих частях данного текста, и в тех же частях могут быть отысканы остальные подробности, относящиеся к их реализации.

Фигура 16 показывает способ кодирования стереофонического сигнала как сигнала битового потока, который включает следующие этапы:

1. Входным сигналом является стереофонический сигнал.

2. Стереофонический сигнал преобразовывается в первое представление в частотной области.

3. Определяется коэффициент комплексного предсказания.

4. Представление в частотной области подвергается понижающему микшированию.

5. Сигнал понижающего микширования и остаточный сигнал кодируются как битовый поток совместно с коэффициентом комплексного предсказания.

6. Выходным сигналом является битовый поток.

Каждый из этапов 1-5 аналогичен соответствующей функции в любой из систем кодера, раскрытых в предыдущих частях данного текста, и в тех же частях могут быть отысканы остальные подробности, относящиеся к их реализации.

Оба способа могут быть выражены как машиночитаемые команды в форме программ, реализованных программно, и могут исполняться компьютером. Объем защиты настоящего изобретения распространяется на указанное программное обеспечение и компьютерные программные продукты, предназначенные для распространения указанного программного обеспечения.

IV. Экспериментальная оценка

Некоторые из раскрытых в данном документе вариантов осуществления изобретения были количественно оценены экспериментально. В данном подразделе будут суммированы наиболее важные части полученного в этом процессе экспериментального материала.

Вариант осуществления изобретения, использованный для экспериментов, имел следующие характеристики:

(i) Каждый спектр MDST (для временного кадра) вычислялся посредством фильтрации с двумерной импульсной характеристикой конечной длительности из текущего, предыдущего и следующего спектров MDCT.

(ii) Использовалась психоакустическая модель из стереофонического кодера USAC.

(iii) Вместо параметров PS, ICC, CLD и IPD, передавались действительная и мнимая части коэффициента комплексного предсказания. Действительная и мнимая части обрабатывались независимо, ограничивались интервалом [-3,0, 3,0] и квантовались с использованием величины шага 0,1. Затем они кодировались с дифференцированием по времени и в конце кодировались по Хаффману с использованием кодового словаря масштабных коэффициентов USAC. Коэффициенты предсказания обновлялись в каждой второй полосе масштабных коэффициентов, что приводило к разрешающей способности по частоте, которая близка к таковой для MPEG Surround (см., например, ISO/IEC 230031). Данная схема квантования и кодирования приводит к средней битовой скорости передачи данных, приблизительно, 2 Кбит/с для указанной стереофонической вспомогательной информации в типичной конфигурации с целевой битовой скоростью передачи данных 96 Кбит/с.

(iv) Формат битового потока модифицируется без разрыва текущих битовых потоков USAC, поскольку 2-битный элемент битового потока ms_mask_present в текущий момент времени имеет только три возможных значения. Использование для указания комплексного предсказания четвертого значения допускает режим обратного перехода

на пониженный уровень среднего/побочного кодирования без потери бит (дальнейшие подробности по этой теме см. в предыдущем подразделе данного раскрытия).

Испытательные прослушивания выполнялись в соответствии с методом MUSHRA, который заключается в индивидуальном проигрывании через наушники и в использовании 8 испытуемых элементов с частотой дискретизацией 48 КГц. В данном испытании принимали участие три, пять или шесть испытуемых объектов. Влияние различных приближений MDST оценивалось с целью иллюстрации практического компромисса между сложностью и качеством, который существует между указанными возможностями. Результаты находятся на фигурах 12 и 13, где первая показывает полученные абсолютные оценки, и последняя показывает дифференциальные оценки относительно 96s USAC cp1f, т.е. относительно унифицированного стереофонического кодирования в области MDCT с комплексным предсказанием с использованием текущего кадра MDCT для вычисления приближения MDST. Как можно видеть, выигрыш в качестве, достигаемый унифицированным стереофоническим кодированием на основе MDCT, увеличивается, когда применяются вычислительно более сложные подходы к вычислению спектра MDST. В отношении среднего общего испытания система 96s USAC cp1f на основе одного кадра обеспечивает значительное увеличение эффективности кодирования по сравнению с традиционным стереофоническим кодированием. В свою очередь, даже еще лучшие результаты получаются для 96s USAC cp3f, а именно: унифицированного стереофонического кодирования в области MDCT с комплексным предсказанием с использованием для вычисления MDST текущего, предыдущего и следующего кадров MDCT.

V. Заключительные замечания

Дальнейшие варианты осуществления настоящего изобретения будут ясны специалистам в данной области после прочтения приведенного выше описания. И хотя настоящее описание и графические материалы раскрывают варианты осуществления изобретения и примеры, изобретение не ограничивается этими конкретными примерами. Многочисленные модификации и изменения могут вноситься без отступления от объема настоящего изобретения, который определяется сопроводительной формулой изобретения.

Следует отметить, что способы и устройство, раскрытые в данной заявке после соответствующих модификаций могут применяться в пределах возможностей специалиста, включая типовые эксперименты, для кодирования сигналов, содержащих больше двух каналов. Особенно следует подчеркнуть, что любые сигналы, параметры и матрицы, упоминаемые в связи с описанными вариантами осуществления изобретения, могут быть зависящими от частоты или независящими от частоты и/или зависящими от времени или независящими от времени. Описанные этапы вычислений могут осуществляться почастотно или для всех частотных полос в момент времени, и все объекты могут осуществляться так, чтобы они обладали частотно-избирательным действием. Для целей заявки любые схемы квантования могут адаптироваться в соответствии с психоакустическими моделями. Кроме того, следует отметить, что все различные суммарно-разностные преобразования, т.е. преобразования из низведенной/остаточной формы в псевдо-L/R-форму, а также преобразование L/R в M/S и преобразование M/S в L/R, имеют форму

$$g\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix},$$

где может изменяться только коэффициент g усиления. Таким образом, регулируя коэффициенты усиления по отдельности, определенный коэффициент усиления при

кодировании можно компенсировать путем соответствующего выбора коэффициента усиления при декодировании. Кроме того, как понятно специалистам в данной области, четное количество последовательно расположенных суммарно-разностных преобразований дает эффект ретранслирующей ступени, возможно, с неединичным коэффициентом усиления.

Системы и способы, описанные выше в данном описании, могут быть реализованы как программное обеспечение, встроенное программное обеспечение, аппаратное обеспечение или их комбинация. Некоторые компоненты или все компоненты могут реализовываться как программное обеспечение, исполняемое процессором обработки цифровых сигналов или микропроцессором, или реализовываться как аппаратное обеспечение, или как интегральная микросхема специального назначения. Указанное программное обеспечение может распространяться на машиночитаемых носителях, которые могут включать компьютерные носители данных или средства связи. Как хорошо известно специалистам в данной области, компьютерные носители данных включают энергозависимые и энергонезависимые, съемные и несъемные носители, реализованные любым способом или посредством технологии, предназначенной для хранения в памяти информации, такой как машиночитаемые команды, структуры данных, программные модули или другие данные. Компьютерные носители данных в качестве неограничивающих примеров включают RAM, ROM, EEPROM, флэш-память или другие технологии памяти, CD-ROM, компакт-диск формата DVD или другие оптические дисковые носители, кассеты с магнитной лентой, магнитную ленту, магнитные дисковые носители или другие магнитные устройства хранения информации, или любой другой носитель, который можно использовать для хранения в памяти нужной информации и который может быть доступен для компьютера. Кроме того, специалистам в данной области известно, что средства связи, как правило, заключают в себе машиночитаемые команды, структуры данных, программные модули или другие данные в модулированном сигнале данных, таком как несущая волна или другой механизм передачи, и включают любые средства доставки информации.

Формула изобретения

1. Система декодера, предназначенная для создания стереофонического сигнала посредством стереофонического кодирования с комплексным предсказанием, где система декодера включает:

ступень (406, 407, 408, 409; 1433) повышающего микширования, адаптированную для генерирования стереофонического сигнала на основе первых представлений сигнала (M) понижающего микширования и остаточного сигнала (D) в частотной области, где каждое из первых представлений в частотной области включает первые спектральные составляющие, которые представляют спектральный состав соответствующего сигнала, выраженного в первом подпространстве многомерного пространства, где ступень повышающего микширования включает:

модуль (408), предназначенный для вычисления второго представления сигнала понижающего микширования в частотной области на основе его первого представления в частотной области, где второе представление в частотной области включает вторые спектральные составляющие, представляющие спектральный состав сигнала, выраженного во втором подпространстве многомерного пространства, которое включает часть многомерного пространства, не включенную в первое подпространство;

взвешенный сумматор (406, 407), предназначенный для вычисления побочного сигнала (S) на основе первого и второго представлений сигнала понижающего

микширования в частотной области, первого представления остаточного сигнала в частотной области и коэффициента комплексного предсказания (α), закодированного в сигнале битового потока; и

суммарно-разностную ступень (409), предназначенную для вычисления стереофонического сигнала на основе первого представления сигнала понижающего микширования в частотной области и побочного сигнала;

ступень (403; 1431) первого модификатора в частотной области, расположенную в восходящем направлении относительно ступени повышающего микширования и способную действовать в активном режиме, в котором она обрабатывает представление в частотной области, по меньшей мере, одного сигнала, и в пассивном режиме, в котором она играет роль ретранслятора; и

ступень (410; 1435) второго модификатора в частотной области, расположенную в нисходящем направлении относительно ступени повышающего микширования и способную действовать в активном режиме, в котором она обрабатывает представление в частотной области, по меньшей мере, одного сигнала, и в пассивном режиме, в котором она играет роль ретранслятора.

2. Система декодера по п. 1, отличающаяся тем, что по меньшей мере одна из указанных ступеней модификатора в частотной области представляет собой ступень временного ограничения шума, TNS.

3. Система декодера по п. 2, отличающаяся тем, что также адаптирована для приема для каждого временного кадра поля данных, связанного с этим кадром, и для действия, в ответ на значение поля данных, ступени первого модификатора в частотной области в ее активном режиме или в ее режиме ретрансляции и действия ступени второго модификатора в частотной области в ее активном режиме или в ее режиме ретрансляции.

4. Система декодера по п. 1, отличающаяся тем, что также включает:

ступень деквантования (401), расположенную в восходящем направлении относительно ступени повышающего микширования, предназначенную для создания указанных первых представлений сигнала понижающего микширования (M) и остаточного сигнала (D) в частотной области на основе сигнала битового потока.

5. Система декодера по п. 1, отличающаяся тем, что:

первые спектральные составляющие имеют действительные значения, выраженные в первом подпространстве; и

вторые спектральные составляющие имеют мнимые значения, выраженные во втором подпространстве;

необязательно, первые спектральные составляющие могут быть получены одной из следующих операций:

дискретное косинусное преобразование, DCT, или

модифицированное дискретное косинусное преобразование, MDCT, и

необязательно, вторые спектральные составляющие могут быть получены одной из следующих операций:

дискретное синусное преобразование, DST, или

модифицированное дискретное синусное преобразование, MDST.

6. Система декодера по п. 5, отличающаяся тем, что:

сигнал понижающего микширования сегментирован на последовательные временные кадры, каждый из которых связан со значением коэффициента комплексного предсказания; и

модуль для вычисления второго представления сигнала понижающего микширования в частотной области адаптирован для самостоятельного отключения в ответ на то, что

абсолютное значение мнимой части коэффициента комплексного предсказания меньше предварительно определенного допуска для временного кадра, и, таким образом, он не генерирует выходной сигнал для этого временного кадра.

7. Система декодера по п. 1, отличающаяся тем, что:

первые спектральные составляющие представляют собой коэффициенты преобразования, расположенные в одном или нескольких временных кадрах коэффициентов преобразования, и каждый блок генерируется путем применения преобразования к временному отрезку сигнала во временной области; и

модуль, предназначенный для вычисления второго представления сигнала понижающего микширования в частотной области адаптирован для:

получения одной или нескольких первых промежуточных составляющих из, по меньшей мере, некоторых из первых спектральных составляющих;

формирования комбинации указанной одной или нескольких первых спектральных составляющих в соответствии с, по меньшей мере, частью одной или нескольких импульсных характеристик с целью получения одной или нескольких вторых промежуточных составляющих; и

получения указанной одной или нескольких вторых спектральных составляющих из указанной одной или нескольких вторых промежуточных составляющих.

8. Система декодера по п. 7, отличающаяся тем, что части одной или нескольких импульсных характеристик основываются на частотных характеристиках одного или нескольких преобразований, где, необязательно, частотные характеристики преобразования зависят от характеристик одной или нескольких анализирующих оконных функций, которые применяются вместе с преобразованием к временному отрезку сигнала.

9. Система декодера по одному из пп. 7 и 8, отличающаяся тем, что модуль, предназначенный для вычисления второго представления сигнала понижающего микширования в частотной области, адаптирован для получения каждого временного кадра вторых спектральных составляющих на основе одного из следующих вариантов:

(a) одновременного временного кадра первых спектральных составляющих;

(b) одновременного и предшествующего временных кадров первых спектральных составляющих; и

(c) одновременного, предшествующего и последующего временных кадров первых спектральных составляющих.

10. Система декодера по п. 1, отличающаяся тем, что модуль, предназначенный для вычисления второго представления сигнала понижающего микширования в частотной области, адаптирован для вычисления приближенного второго спектрального представления, включающего приближенные вторые спектральные составляющие, определяемые при помощи комбинации из, по меньшей мере, двух смежных во времени и/или смежных по частоте первых спектральных составляющих.

11. Система декодера по п. 1, отличающаяся тем, что указанный стереофонический сигнал представлен во временной области, и система декодера также включает:

узел (203) переключения, расположенный между указанной ступенью деквантования и указанной ступенью повышающего микширования, действующий для выполнения следующих функций:

(a) ретранслирующей ступени, или

(b) суммарно-разностной ступени,

таким образом, делающий возможным переключение между входными стереофоническими сигналами, кодированными напрямую и совместно;

ступень (209) обратного преобразования, адаптированную для вычисления представления стереофонического сигнала во временной области; и

селекторную схему (208), расположенную в восходящем направлении относительно ступени обратного преобразования, адаптированную для ее селективного связывания с:

(a) точкой в нисходящем направлении относительно ступени повышающего микширования, посредством чего стереофонический сигнал, полученный посредством комплексного предсказания, подается на ступень обратного преобразования; или

(b) точкой в нисходящем направлении относительно узла (203) переключения и в восходящем направлении относительно ступени повышающего микширования, посредством чего на ступень обратного преобразования подается стереофонический сигнал, полученный путем прямого стереофонического кодирования.

12. Система декодера по п. 1, отличающаяся тем, что модуль, предназначенный для вычисления второго представления сигнала понижающего микширования в частотной области, включает:

ступень (306) обратного преобразования, предназначенную для вычисления представления низведенного и/или побочного сигнала во временной области на основе первого представления в частотной области соответствующего сигнала в первом подпространстве многомерного пространства; и

ступень (307) преобразования, предназначенную для вычисления второго представления соответствующего сигнала в частотной области на основе представления этого сигнала во временной области,

где, предпочтительно, ступень (306) обратного преобразования выполняет обратное модифицированное дискретное косинусное преобразование, MDCT, и ступень преобразования выполняет модифицированное дискретное синусное преобразование, MDST.

13. Система декодера по п. 12, отличающаяся тем, что указанный стереофонический сигнал представлен во временной области, и система декодера также включает:

узел (302) переключения, расположенный между указанной ступенью деквантования и указанной ступенью повышающего микширования, способный действовать как:

(a) ретранслирующая ступень для применения в совместном стереофоническом кодировании; или

(b) суммарно-разностная ступень для применения в прямом стереофоническом кодировании;

дополнительную ступень (311) обратного преобразования, расположенную на ступени повышающего микширования, предназначенную для вычисления представления побочного сигнала во временной области;

селекторную схему (305, 310), расположенную в восходящем направлении относительно ступеней (306, 301) обратного преобразования, адаптированную для их селективного связывания с:

(a) дополнительной суммарно-разностной ступенью (304), которая, в свою очередь, связана с точкой в нисходящем направлении относительно узла (302) переключения и в восходящем направлении относительно ступени повышающего микширования; или

(b) сигналом понижающего микширования, получаемым из узла (302) переключения, и побочным сигналом, получаемым из взвешенного сумматора (308, 309).

14. Способ декодирования, предназначенный для повышающего микширования входного стереофонического сигнала посредством стереофонического кодирования с комплексным предсказанием в выходной стереофонический сигнал, где:

указанный входной стереофонический сигнал включает первые представления канала (M) понижающего микширования и остаточного канала (D) в частотной области и коэффициент комплексного предсказания (α); и

каждое из первых представлений в частотной области включает первые спектральные составляющие, которые представляют спектральный состав соответствующего сигнала, выраженного в первом подпространстве многомерного пространства,

способ выполняется ступенью повышающего микширования и включает следующие этапы:

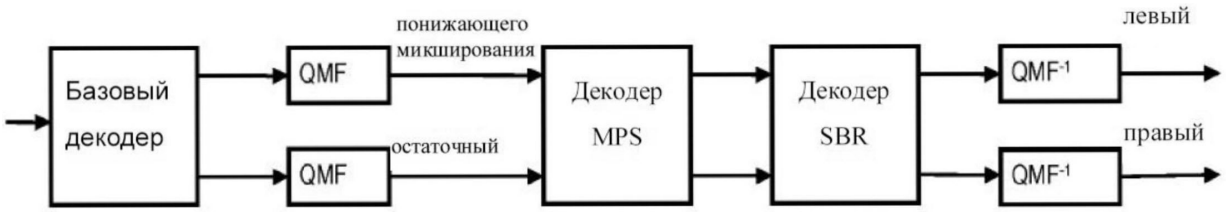
вычисление второго представления канала понижающего микширования в частотной области на основе его первого представления в частотной области, где второе представление в частотной области включает вторые спектральные составляющие, представляющие спектральный состав сигнала, выраженного во втором подпространстве многомерного пространства, которое включает часть многомерного пространства, не включенную в первое подпространство;

вычисление побочного канала на основе первого и второго представлений сигнала понижающего микширования в частотной области, первого представления остаточного сигнала в частотной области и коэффициента комплексного предсказания;

и также включает или выполняемый перед этапом повышающего микширования этап временного ограничения шума, TNS, для указанного первого представления сигнала понижающего микширования в частотной области и/или указанного первого представления остаточного сигнала в частотной области;

или выполняемый после этапа повышающего микширования этап применения TNS к, по меньшей мере, одному каналу указанного стереофонического сигнала.

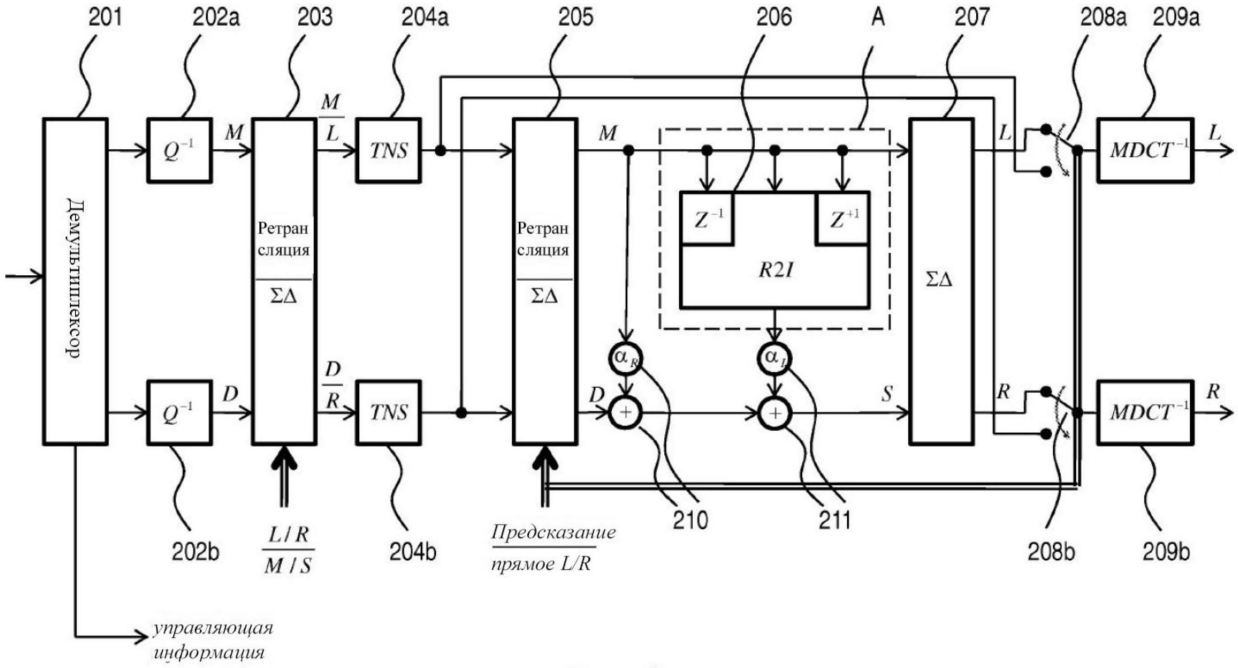
15. Машиночитаемый носитель данных, в памяти которого хранятся команды, которые при их исполнении компьютером выполняют способ, изложенный в п. 14.



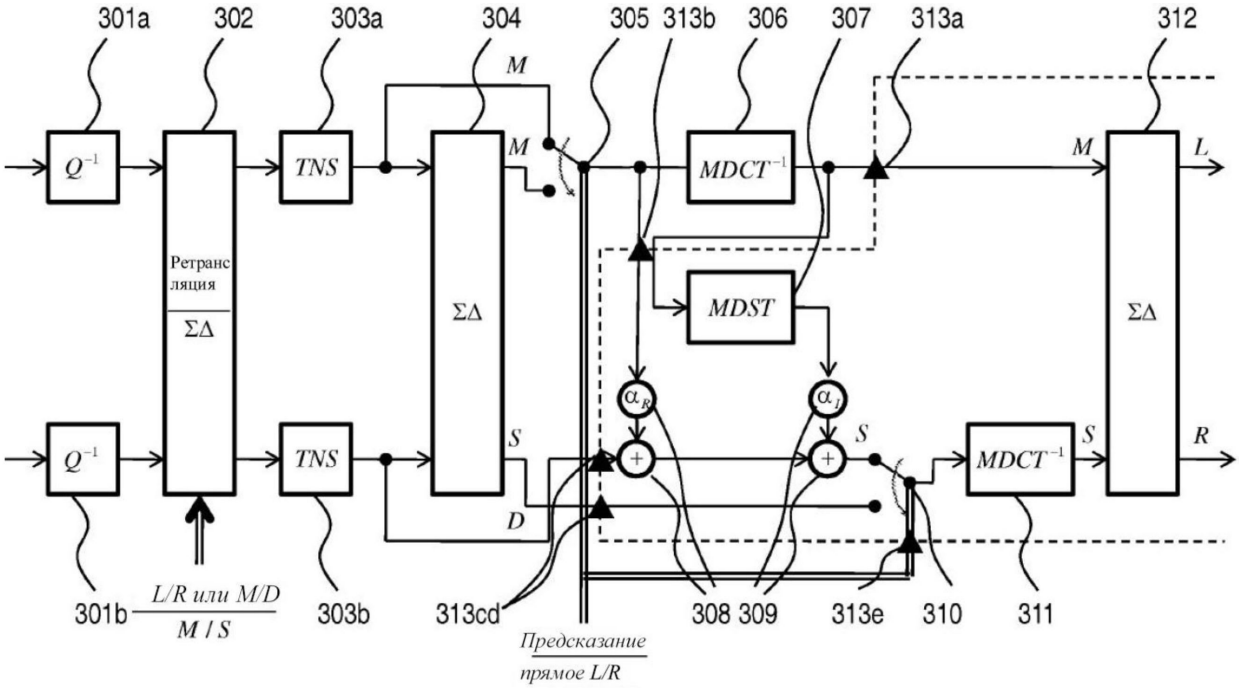
Фиг. 1A



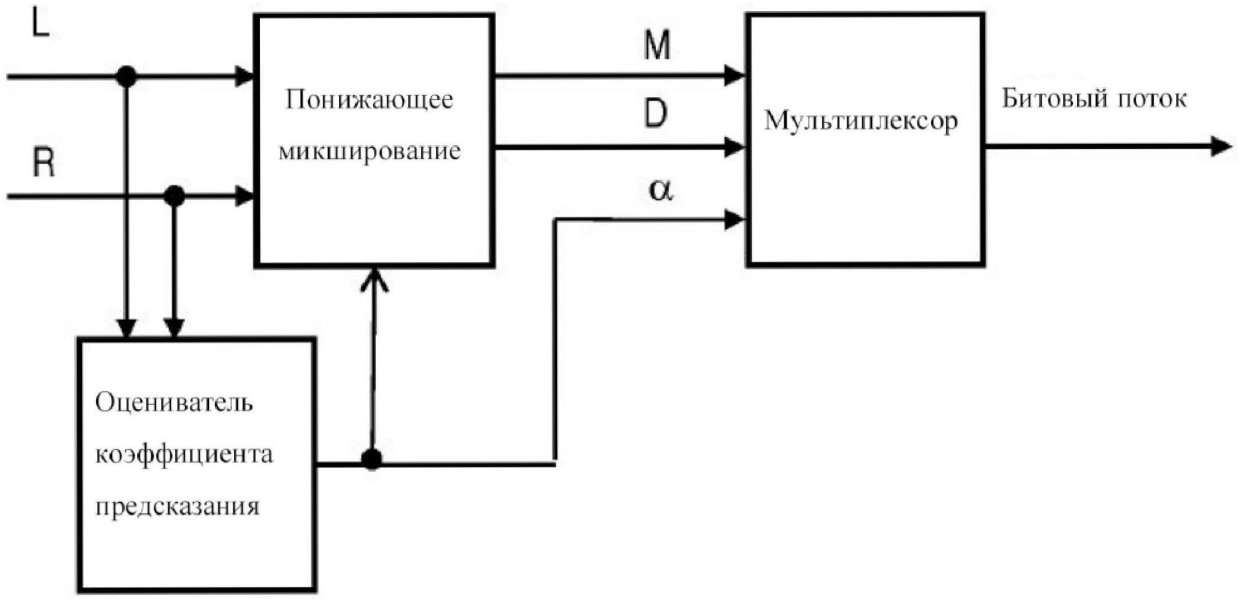
Фиг. 1B



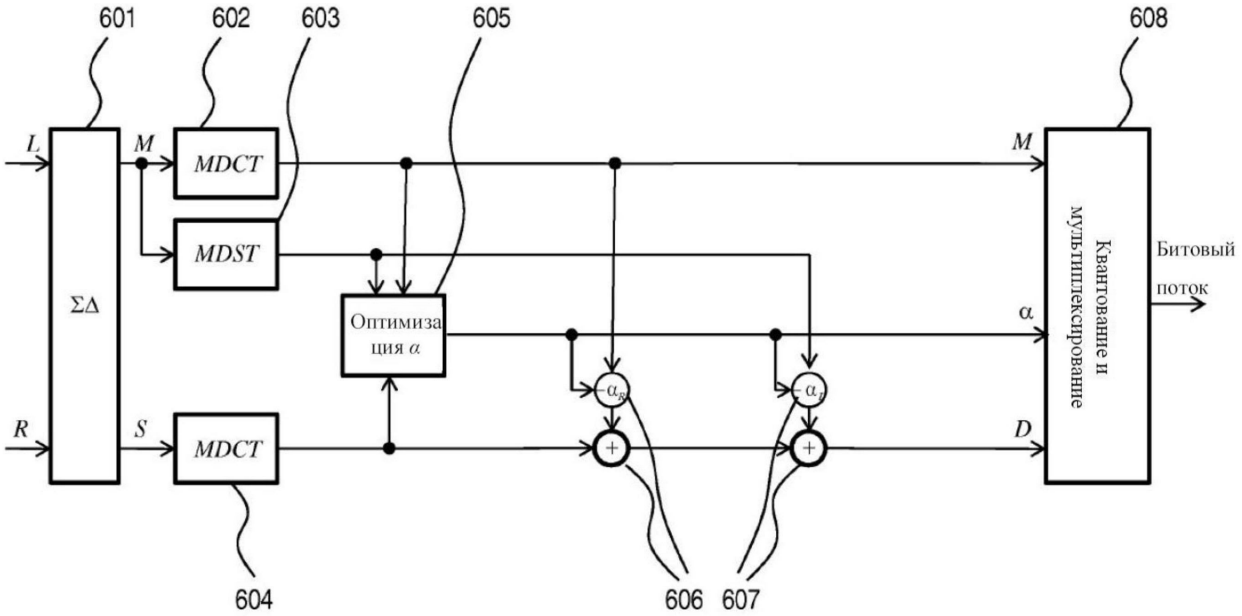
Фиг. 2


301a
302
303a
304
305
313b
306
307
313a
312
Q^{-1}
TNS
M
MDCT^{-1}
L
Ретранс ляция
ΣΔ
MDST
α_R
α_I
S
R
D
301b
L/R или M/D
M / S
303b
313cd
308 309
313e
310
311
Предсказание
прямое L/R

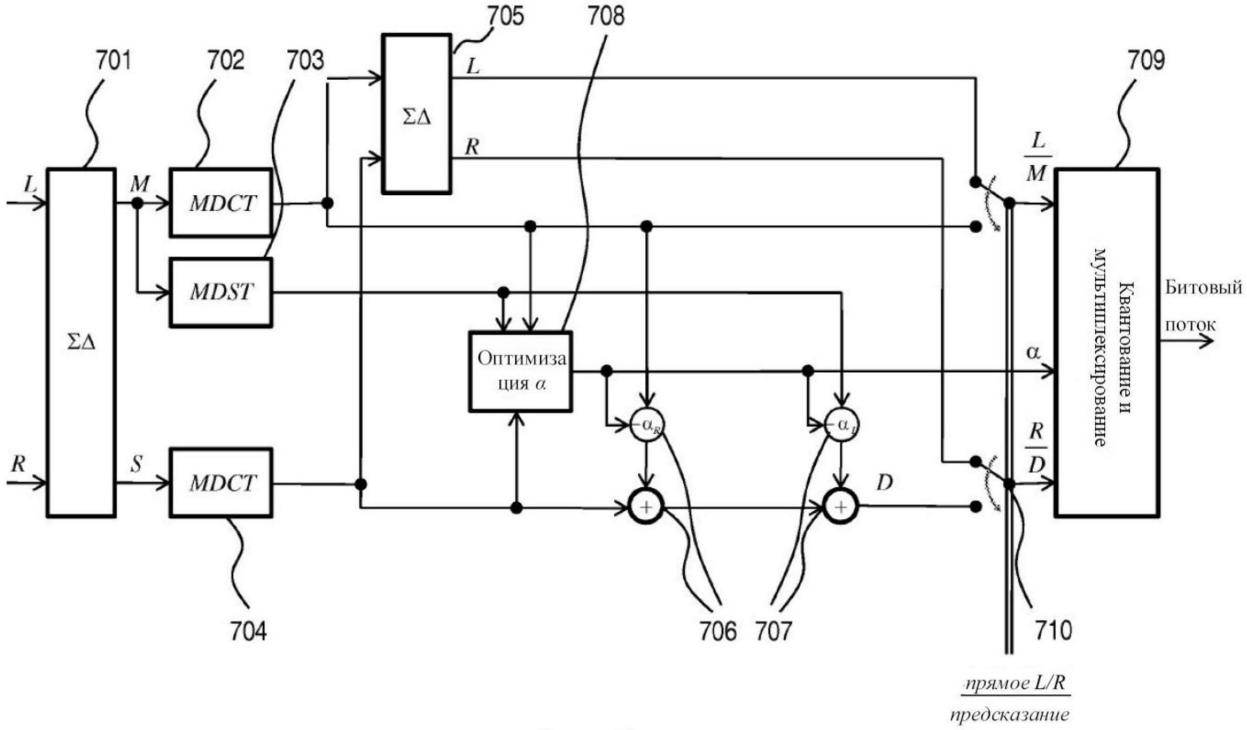
Фиг. 3


L
R
Понижающее микширование
M
D
α
Мультиплексор
Битовый поток
Оцениватель коэффициента предсказания

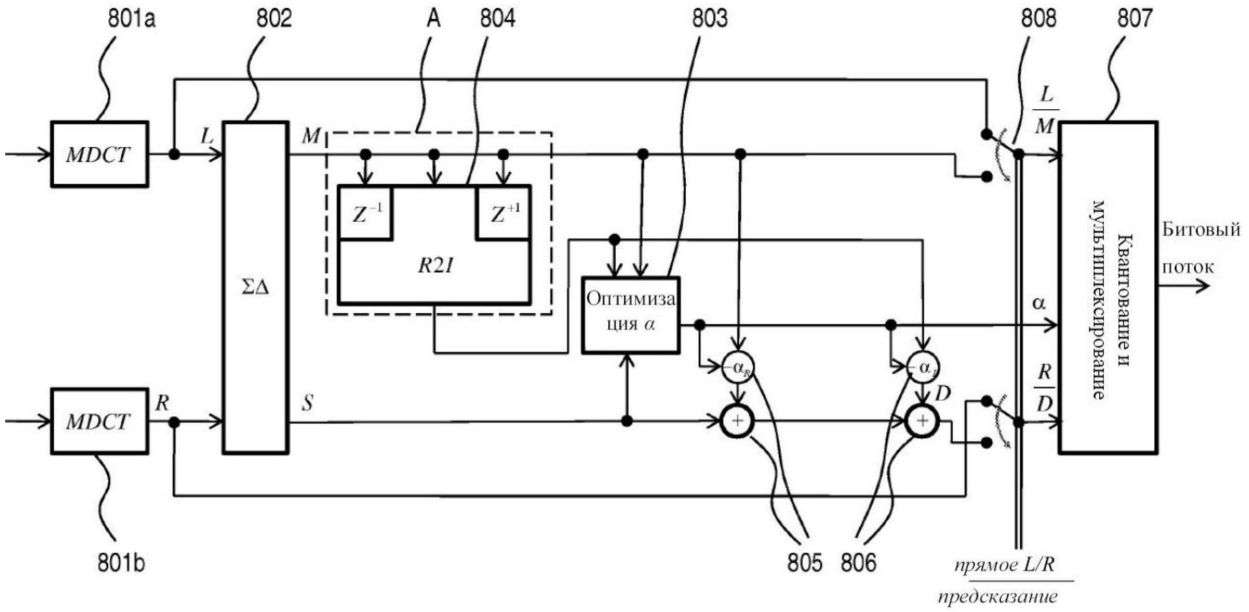
Фиг. 5


601
602
603
605
608
L
M
MDCT
MDST
ΣΔ
Оптимизация α
R
S
MDCT
M
α
D
Квантование и мультиплексирование
Битовый поток
604
606
607

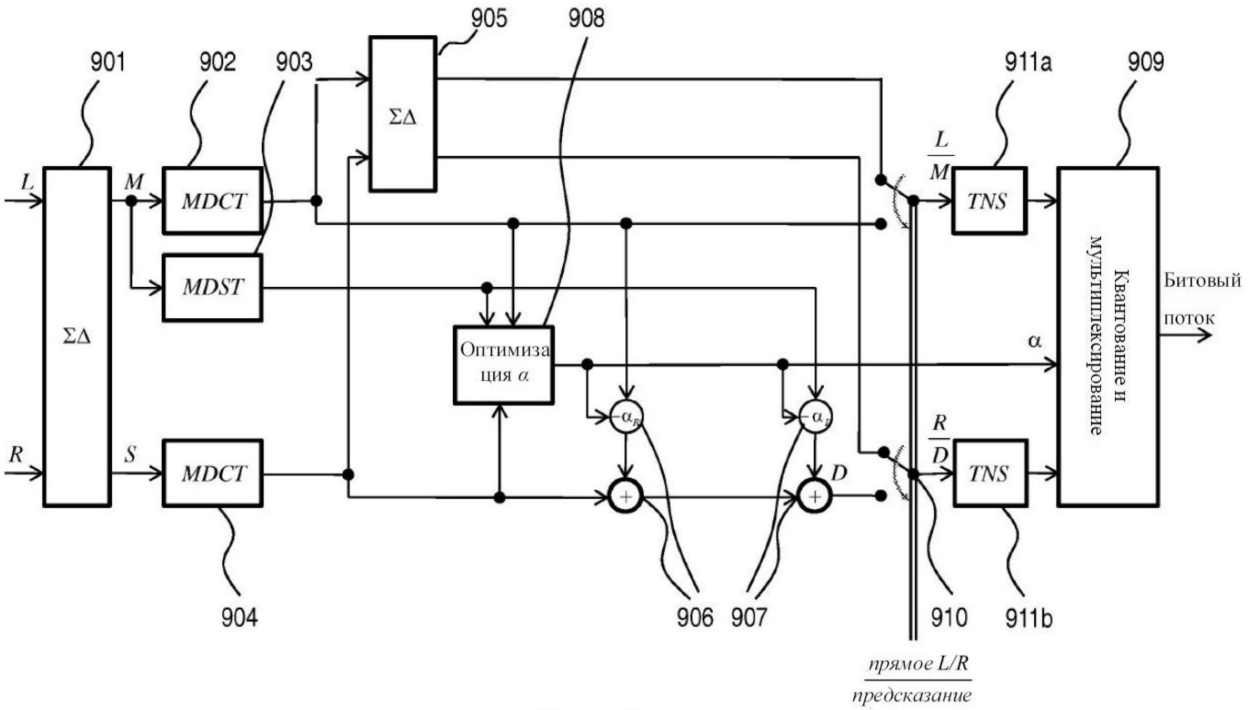
Фиг. 6


705
708
701
702
703
709
L
R
ΣΔ
L
M
MDCT
MDST
Оптимизация α
α
R
D
S
MDCT
Квантование и мультиплексирование
Битовый поток
704
706
707
710
прямое L/R
предсказание

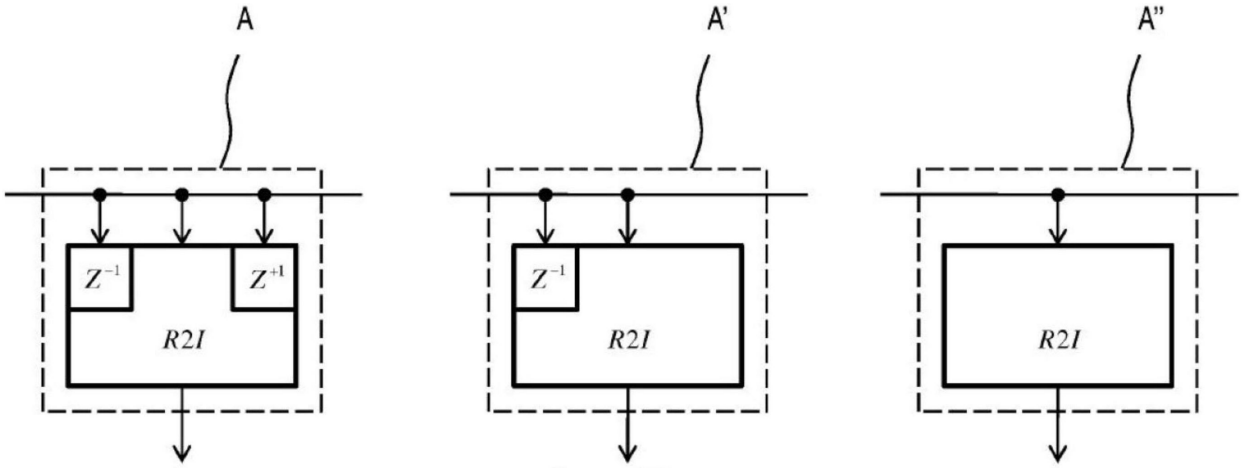
Фиг. 7


801a
802
A
804
803
808
807
MDCT
L
M
Z⁻¹
Z⁺¹
R21
ΣΔ
Оптимизация α
L/M
α
R/D
D
Квантование и мультиплексирование
Битовый поток
MDCT
R
S
801b
805
806
прямое L/R предсказание

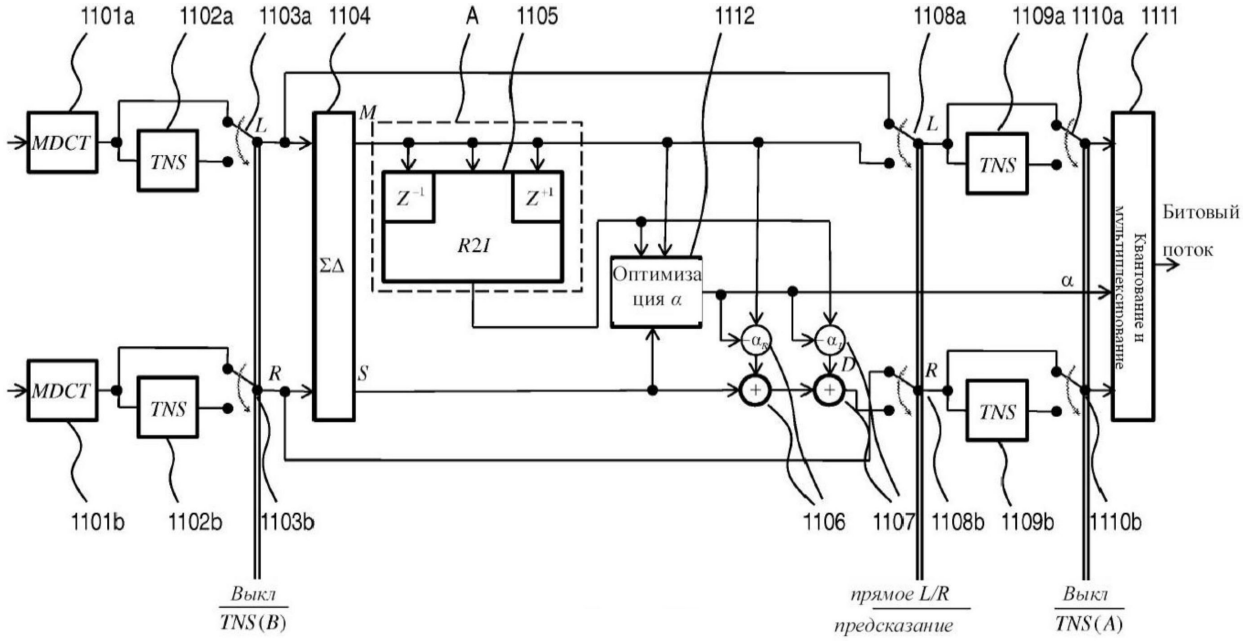
Фиг. 8


905
908
901
902
903
911a
909
ΣΔ
L
M
MDCT
MDST
TNS
Оптимизация α
α
R
S
D
Квантование и мультиплексирование
Битовый поток
904
906
907
910
911b
прямое L/R предсказание

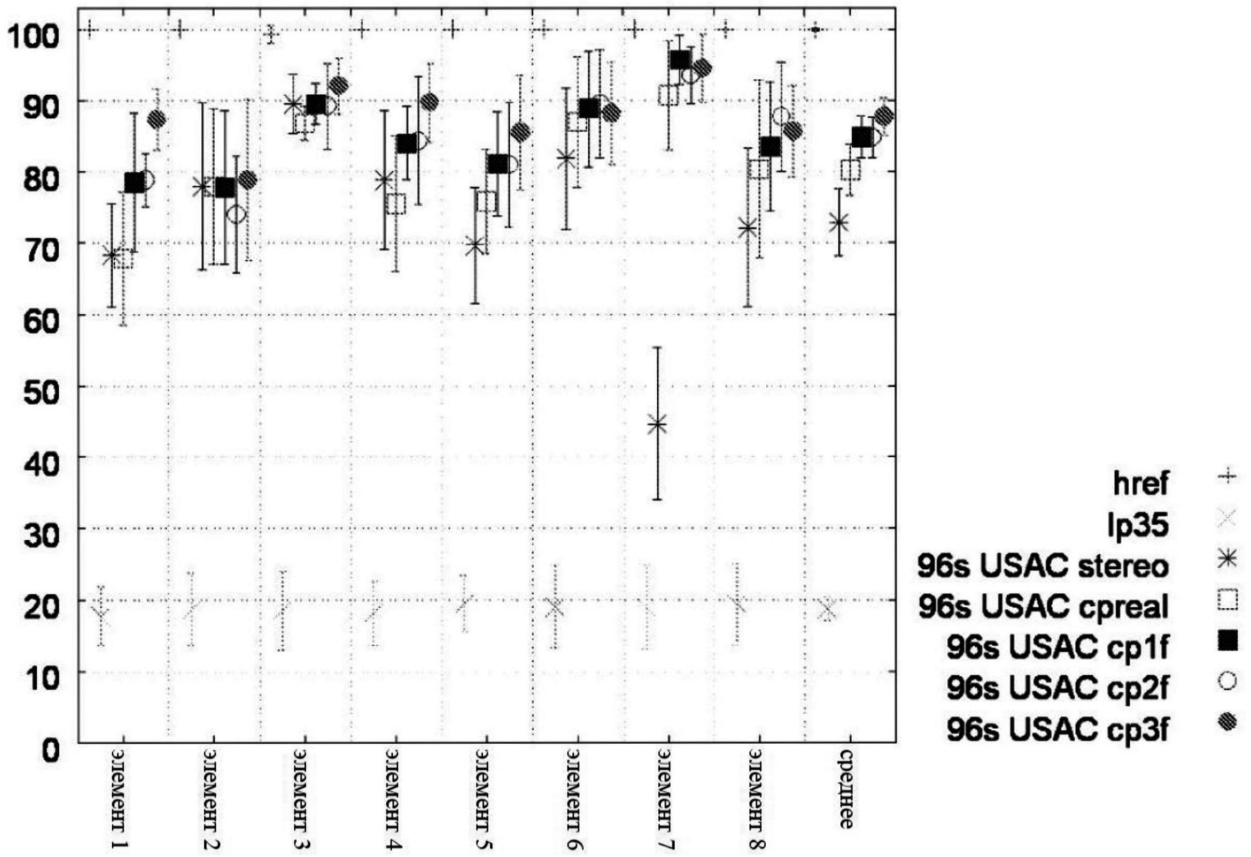
Фиг. 9


A
A'
A"
Z⁻¹
Z⁺¹
R2I

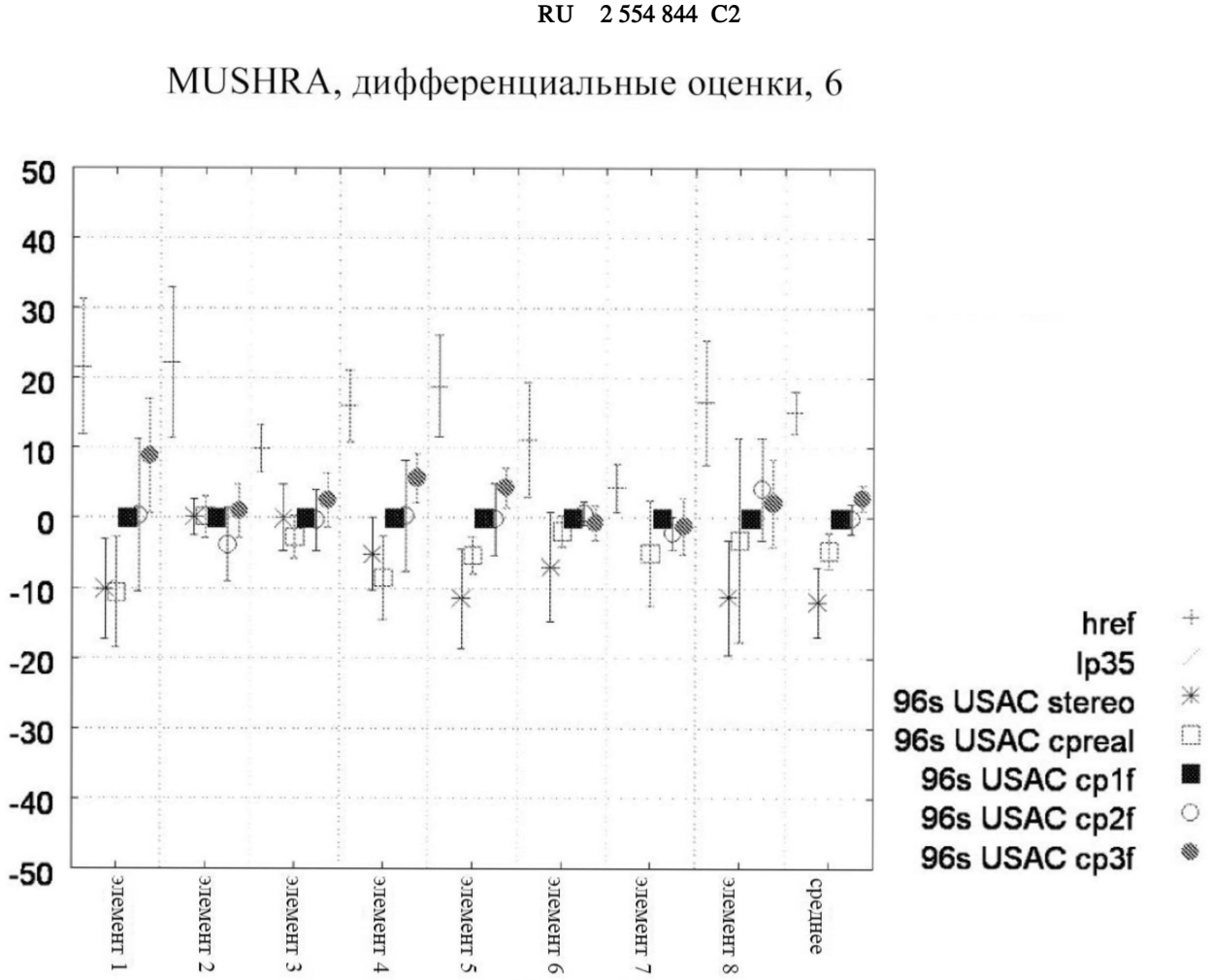
Фиг. 10


1101a
1102a
1103a
1104
A
1105
1112
1108a
1109a
1110a
1111
MDCT
TNS
L
M
ΣΔ
Z⁻¹
Z⁺¹
R2I
Оптимизация α
α
Битовый поток
Квантование и мультиплексирование
R
S
D
1101b
1102b
1103b
1106
1107
1108b
1109b
1110b
Выкл TNS(B)
прямое L/R предсказание
Выкл TNS(A)

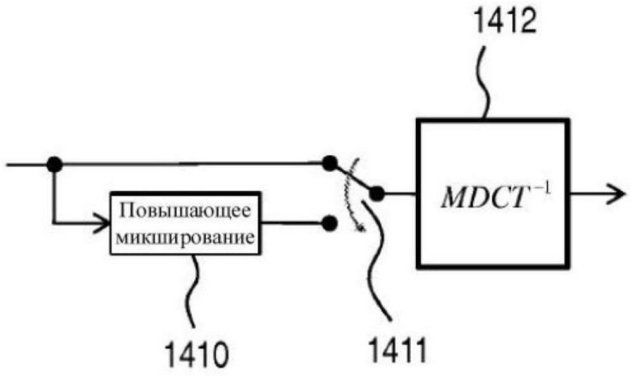
Фиг. 11


MUSHRA, абсолютные оценки, 6 объектов
100
90
80
70
60
50
40
30
20
10
0
элемент 1
элемент 2
элемент 3
элемент 4
элемент 5
элемент 6
элемент 7
элемент 8
среднее
href
lp35
96s USAC stereo
96s USAC cpreal
96s USAC cp1f
96s USAC cp2f
96s USAC cp3f

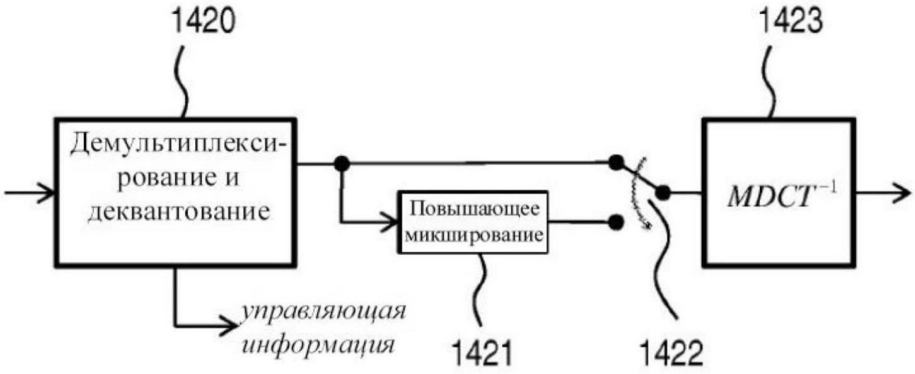
Фиг. 12



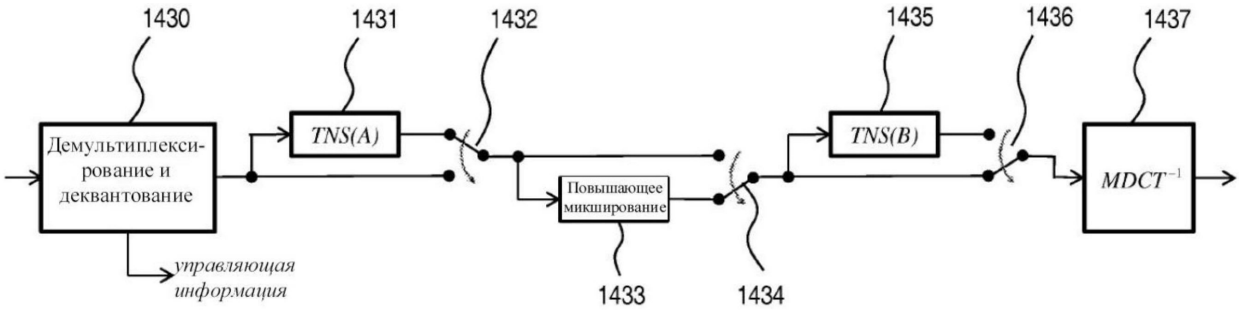
Фиг. 13



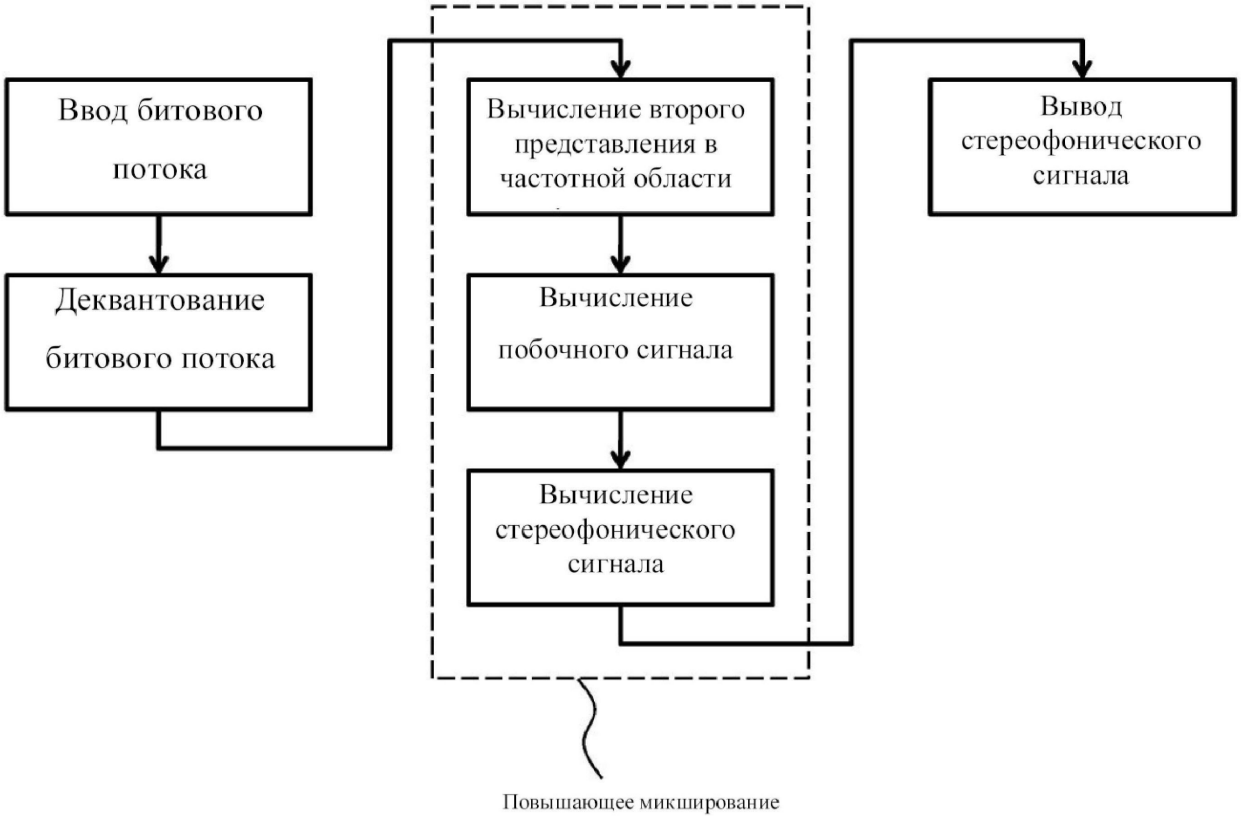
Фиг. 14A



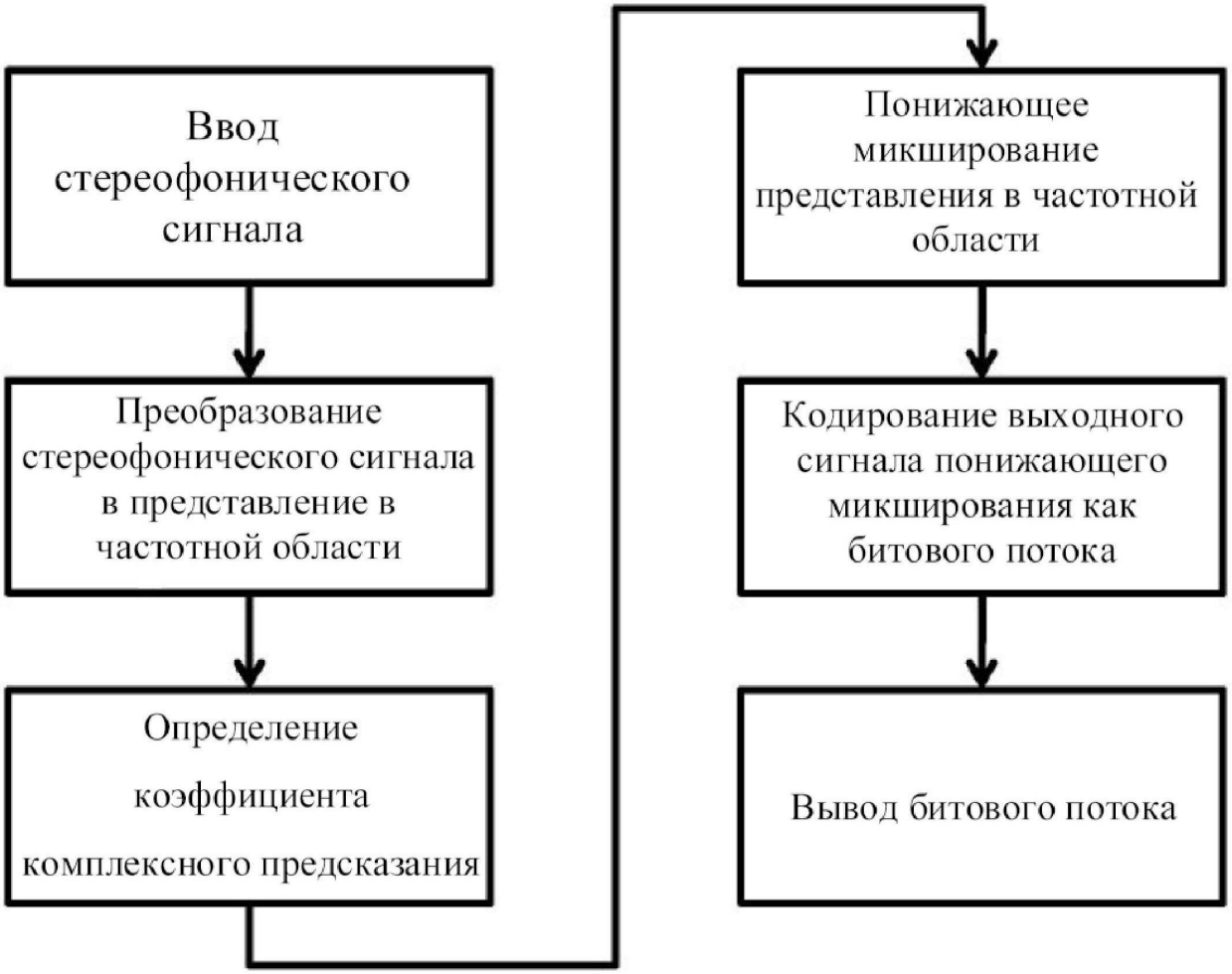
Фиг. 14B



Фиг. 14C



Фиг. 15


Ввод стереофонического сигнала
Преобразование стереофонического сигнала в представление в частотной области
Определение коэффициента комплексного предсказания
Понижающее микширование представления в частотной области
Кодирование выходного сигнала понижающего микширования как битового потока
Вывод битового потока

Фиг. 16