



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК
G10L 21/038 (2021.02)

(21)(22) Заявка: 2017144521, 04.02.2015

(24) Дата начала отсчета срока действия патента:
04.02.2015

Дата регистрации:
11.01.2022

Приоритет(ы):

(30) Конвенционный приоритет:
07.02.2014 FR 1450969

Номер и дата приоритета первоначальной заявки,
из которой данная заявка выделена:
2016136008 07.02.2014

(43) Дата публикации заявки: 18.02.2019 Бюл. № 5

(45) Опубликовано: 11.01.2022 Бюл. № 2

Адрес для переписки:
129090, Москва, ул. Б. Спасская, 25, стр. 3, ООО
"Юридическая фирма Городисский и
Партнеры"

(72) Автор(ы):

КАНЕВСКА Магдалена (FR),
РАГО Стефан (FR)

(73) Патентообладатель(и):

КОНИНКЛЕЙКЕ ФИЛИПС Н.В. (NL)

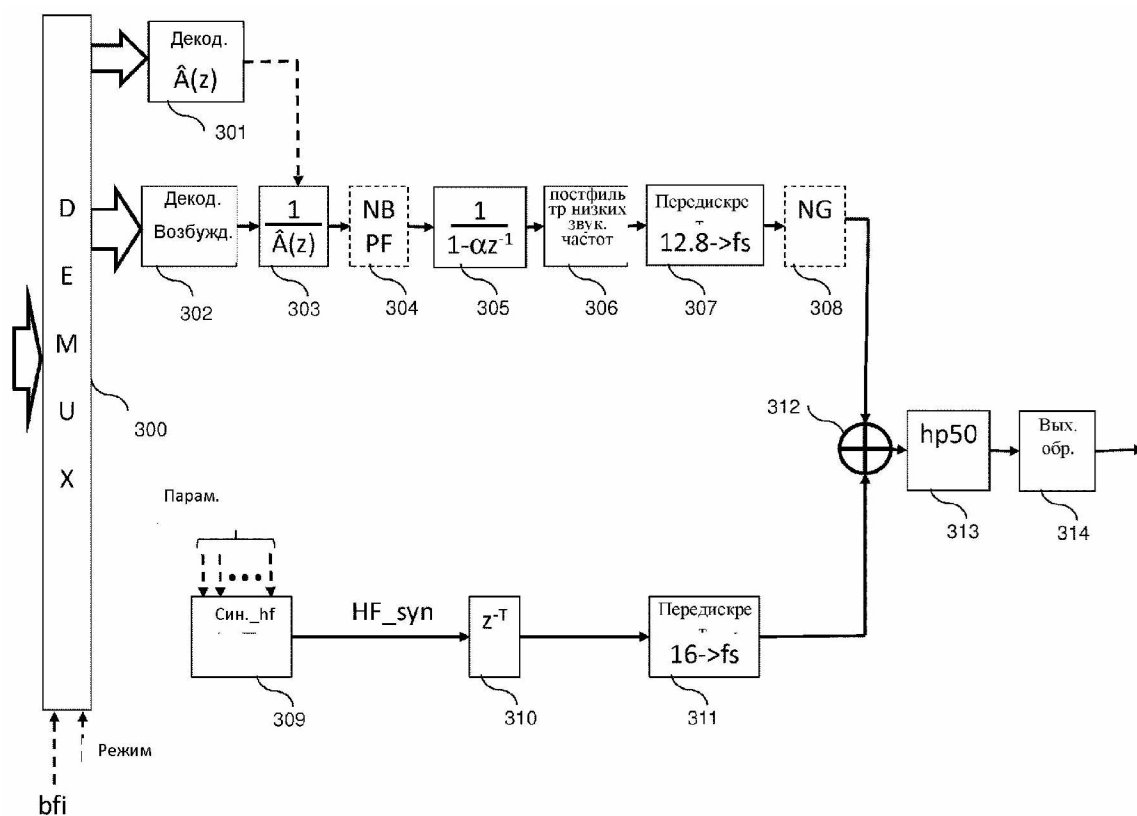
(56) Список документов, цитированных в отчете
о поиске: US 2009/0048846 A1, 19.02.2009. US
2010/0292994 A1, 18.11.2010. US 2010/0063824
A1, 11.03.2010. US 2012/0128177 A1, 24.05.2012.
RU 2452044 C1, 27.05.2012.

(54) УЛУЧШЕННОЕ РАСШИРЕНИЕ ДИАПАЗОНА ЧАСТОТ В ДЕКОДЕРЕ ЗВУКОВОГО СИГНАЛА

(57) Реферат:

Изобретение относится к области вычислительной техники для обработки аудиоданных. Технический результат заключается в повышении точности декодирования сигнала звуковой частоты. Технический результат достигается за счет получения сигнала, декодированного в первом диапазоне частот, называемом нижним диапазоном; извлечения тональных составляющих и сигнала окружения из сигнала, возникающего из декодированного сигнала нижнего диапазона; объединения тональных составляющих и сигнала окружения посредством адаптивного микширования с использованием коэффициентов регулирования уровня энергии для получения звукового сигнала, называемого объединенным сигналом; и

расширения декодированного сигнала нижнего диапазона по меньшей мере на по меньшей мере одном втором диапазоне частот, находящемся выше, чем первый диапазон частот перед этапом извлечения, формирующего расширенный декодированный сигнал нижнего диапазона $U_{НВ1}(k)$, причем извлечение содержит следующие операции: (а) вычисление тональной энергии расширенного декодированного сигнала нижнего диапазона; (б) вычисление сигнала окружения в абсолютном значении, соответствующем среднему значению уровня спектра спектральная линия за спектральной линией, и вычисление энергии преобладающих тональных составляющих декодированного в верхнем диапазоне спектра. 3 н. и 6 з.п. ф-лы, 6 ил.



Фиг. 3



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(12) **ABSTRACT OF INVENTION**

(52) CPC
G10L 21/038 (2021.02)

(21)(22) Application: **2017144521, 04.02.2015**

(24) Effective date for property rights:
04.02.2015

Registration date:
11.01.2022

Priority:

(30) Convention priority:
07.02.2014 FR 1450969

Number and date of priority of the initial application,
from which the given application is allocated:
2016136008 07.02.2014

(43) Application published: **18.02.2019 Bull. № 5**

(45) Date of publication: **11.01.2022 Bull. № 2**

Mail address:
**129090, Moskva, ul. B. Spasskaya, 25, str. 3, OOO
"Yuridicheskaya firma Gorodisskij i Partnery"**

(72) Inventor(s):

**KANIEWSKA, Magdalena (FR),
RAGOT, Stephane (FR)**

(73) Proprietor(s):

Koninklijke Philips N.V. (NL)

(54) **IMPROVED FREQUENCY RANGE EXTENSION IN SOUND SIGNAL DECODER**

(57) Abstract:

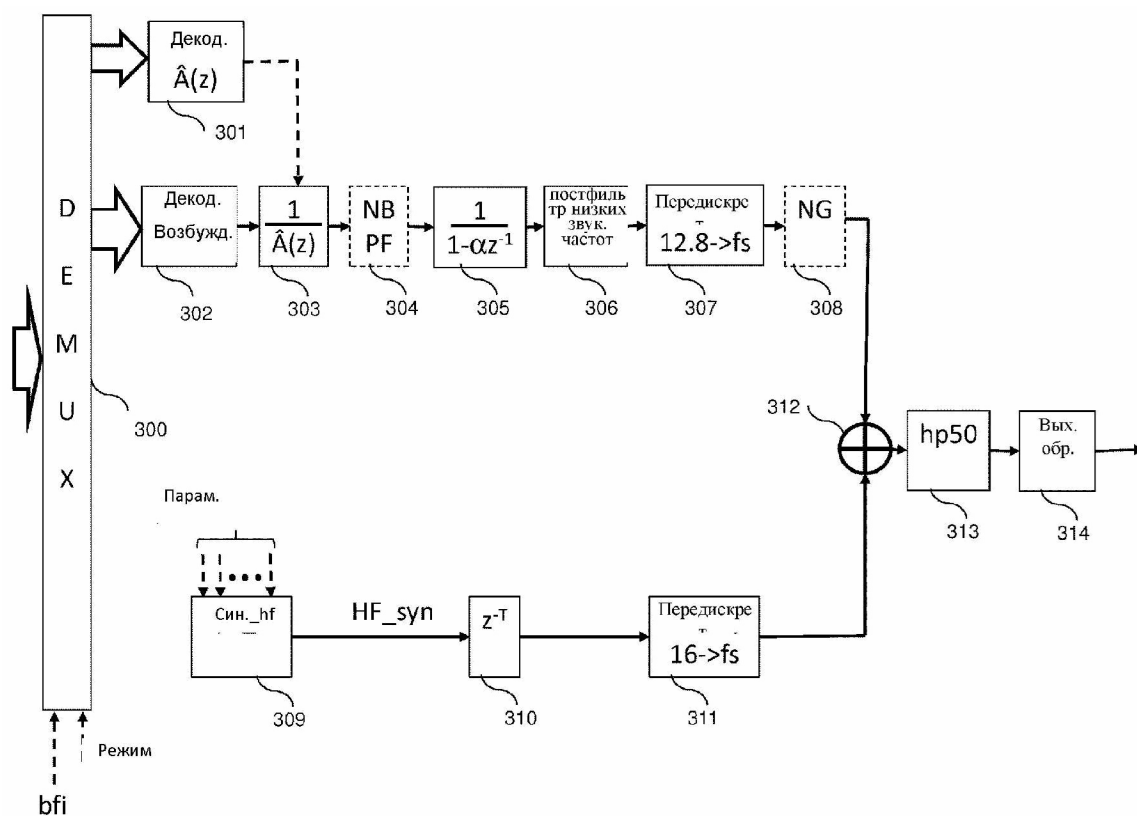
FIELD: computer technology.

SUBSTANCE: invention relates to the field of computer technology for audio data processing. A signal decoded in the first frequency range called the lower range is received; tonal components and an environmental signal are extracted from a signal arising from the decoded lower range signal; tonal components and the environmental signal are combined by means of adaptive mixing using coefficients of power level adjustment to obtain a sound signal called a combined signal; and the decoded lower range signal is extended on at least one second frequency range, which is higher than the first frequency range before the extraction

stage, forming the extended decoded lower range signal $U_{HB1}(k)$, wherein the extraction contains following operations: (a) calculating tonal energy of the extended decoded lower range signal; (b) calculating the environmental signal in the absolute value corresponding to the average value of a spectrum level of a spectral line by a spectral line, and calculating energy of prevailing tonal components of the decoded lower range spectrum.

EFFECT: increase in the accuracy of decoding of the sound frequency signal.

9 cl, 6 dwg



Фиг. 3

Настоящее изобретение относится к области кодирования/декодирования и обработки сигналов звуковой частоты (например, речи, музыки или других подобных сигналов) для их передачи или их хранения.

Более конкретно, настоящее изобретение относится к способу расширения диапазона частот и устройству в декодере или процессоре, производящему выделение сигнала звуковой частоты из шумов.

Существуют многочисленные методы сжатия (с потерей) сигнала звуковой частоты, такого как речь или музыка.

Традиционные способы кодирования для разговорных приложений, как правило, классифицируются как *кодирование формы сигналов* (PCM - "импульсно-кодовая модуляция", ADPCM - "адаптивная дифференциальная импульсно-кодовая модуляция", кодирование с преобразованием и т.д.), *параметрическое кодирование* (LPC - "кодирования с линейным предсказанием", синусоидальное кодирование и т.д.) и параметрическое гибридное кодирование с квантованием параметров посредством "анализа через синтез", из которого кодирование CELP ("линейное предсказание с кодовым возбуждением") является наиболее известным примером.

Для не разговорных приложений известный уровень техники для кодирования (монофонического) звукового сигнала включает в себя перцептуальное кодирование посредством преобразования или в поддиапазонах с параметрическим кодированием высоких частот посредством репликации полосы (SBR - репликации спектральной полосы).

Обзор способов кодирования обычной речи и звука можно найти в работах W.B. Kleijn and K.K. Paliwal (eds.), *Speech Coding and Synthesis*, Elsevier, 1995; M. Bosi, R.E. Goldberg, *Introduction to Digital Audio Coding and Standards*, Springer 2002; J. Benesty, M.M. Sondhi, Y. Huang (eds.), *Handbook of Speech Processing*, Springer 2008.

В данном случае внимание сосредоточено, более конкретно, на стандартизированном 3GPP AMR-WB ("адаптивном многоскоростном при широкополосной передаче") кодеке (кодере и декодере), который работает на входной/выходной частоте 16 кГц, и в котором сигнал разделяется на два поддиапазона: нижний диапазон (0-6,4 кГц), который подвергается дискретизации на 12,8 кГц и кодируется посредством модели CELP, и верхний диапазон (6,4-7 кГц), который параметрически восстанавливается посредством "*расширения диапазона*" (или BWE - "расширения полосы частот") с или без дополнительной информации в зависимости от режима текущего кадра. Можно отметить, что ограничение кодированного диапазона кодека AMR-WB на частоте 7 кГц, по существу, связано с тем, что частотная характеристика при передаче широкополосных терминалов была аппроксимирована в момент стандартизации (ETSI/3GPP, в дальнейшем ITU-T) в соответствии с частотной маской, определенной в стандарте ITU-T P.341 и, более конкретно посредством так называемого фильтра "P341", определенного в стандарте ITU-T G.191, который отсекает частоты выше 7 кГц (данный фильтр соблюдает маску, определенную в P.341). Тем не менее, в теории хорошо известно, что сигнал, подвергнутый дискретизации с частотой 16 кГц, может иметь определенный звуковой диапазон от 0 до 8000 Гц; поэтому AMR-WB-кодек вводит ограничение верхнего диапазона по сравнению с теоретической полосой пропускания 8 кГц.

3GPP AMR-WB речевой кодек был стандартизирован в 2001 году, главным образом, для режима с коммутацией каналов (CS) телефонных приложений на GSM (2G) и UMTS (3G). Этот же кодек также был стандартизирован в 2003 году в стандарте ITU-T в виде рекомендации G.722.2 "Широкополосное кодирование речи при приблизительно 16

кбит/с с применением адаптивного многоскоростного широкополосного кодирования (AMR-WB)".

Он имеет девять скоростей цифрового потока, называемых режимами, от 6,6 до 23,85 кбит/с, и содержит механизмы непрерывной передачи (DTX "прерывистой передачи") с определением присутствия голосового сигнала (VAD) и генерацией комфортного шума (CNG) из кадров описания периода молчания (SID "дескриптор информации о паузе") и механизмы коррекции потерянных кадров (FEC "маскирование стирания кадров", иногда называемое PLC "маскирование потери пакетов").

Подробная информация о кодировании AMR-WB и алгоритме декодирования не повторяется в данной заявке; подробное описание этого кодека можно найти в спецификациях 3GPP (TS 26.190, 26.191, 26.192, 26.193, 26.194, 26.204) и в ITU-T-G.722.2 (и соответствующих дополнениях и приложениях), и в статье B. Bessette и другие, озаглавленной "The adaptive multirate wideband speech codec (AMR-WB)", IEEE Transactions on Speech and Audio Processing, vol. 10, no. 8, 2002, pp. 620-636, и исходных кодах связанных стандартов 3GPP и ITU-T.

Принцип расширения диапазона в кодеке AMR-WB довольно слабо развит. В действительности верхний диапазон (6,4-7 кГц) генерируется посредством формирования белого шума во времени (применяемого при формировании коэффициентов усиления в расчете на подкадр) и частотной огибающей (посредством применения синтезирующего фильтра линейного предсказания или LPC - "кодирования с линейным предсказанием"). Данный метод расширения диапазона показан на фиг. 1.

Белый шум $u_{HB1}(n)$, $n=0, L, 79$ генерируется с частотой 16 кГц для каждого подкадра длительностью 5 мс посредством линейного конгруэнтного генератора (блок 100).

Этот шум $u_{HB1}(n)$ формируется во времени посредством применения коэффициентов усиления для каждого подкадра; эта операция разбивается на два этапа обработки (блоки 102, 106 или 109):

- Первый коэффициент вычисляется (блок 101), чтобы установить белый шум $u_{HB1}(n)$ (блок 102) на том же уровне, что возбуждение $u(n)$, $n=0, L, 63$, декодируемое на 12,8 кГц в нижнем диапазоне:

$$u_{HB2}(n) = u_{HB1}(n) \sqrt{\frac{\sum_{l=0}^{63} u(l)^2}{\sum_{l=0}^{79} u_{HB1}(l)^2}}$$

Можно отметить, что нормализация энергий выполняется путем сравнения блоков разного размера (64 для $u(n)$ и 80 для $u_{HB1}(n)$) без компенсации отличий в частотах дискретизации (12,8 или 16 кГц).

- Затем получается возбуждение в верхнем диапазоне (блок 106 или 109) в виде:

$$u_{HB}(n) = \hat{g}_{HB} u_{HB2}(n)$$

где коэффициент усиления $\hat{g}_{HB}$ получают по-разному в зависимости от скорости цифрового потока. Если скорость цифрового потока текущего кадра < 23,85 кбит/с, коэффициент усиления $\hat{g}_{HB}$ оценивают "вслепую" (то есть без дополнительной информации); в этом случае блок 103 фильтрует сигнал, декодированный в нижнем диапазоне, посредством фильтра высоких частот с частотой среза 400 Гц для получения

сигнала $\hat{s}_{hp}(n)$, $n=0, L, 63$ - данный фильтр высоких частот устраняет влияние очень низких частот, которые могут исказить оценку, выполненную в блоке 104 - затем "наклон" (индикатор крутизны спектра), обозначенный e_{tilt} сигнала $\hat{s}_{hp}(n)$ вычисляют с помощью нормированной автокорреляции (блок 104):

$$e_{tilt} = \frac{\sum_{n=1}^{63} \hat{s}_{hp}(n) \hat{s}_{hp}(n-1)}{\sum_{n=0}^{63} \hat{s}_{hp}(n)^2}$$

и, наконец, $\hat{g}_{HB}$ вычисляют в виде:

$$\hat{g}_{HB} = w_{SP} g_{SP} + (1 - w_{SP}) g_{BG}$$

где $g_{SP} = 1 - e_{tilt}$ - коэффициент усиления, применяемый в активных речевых (SP) кадрах,

$g_{BG} = 1,25 g_{SP}$ - коэффициент усиления, применяемый в неактивных речевых кадрах,

относящихся к фоновому (BG) шуму, и w_{SP} - весовая функция, которая зависит от определения присутствия голосового сигнала (VAD). Понятно, что оценка наклона (e_{tilt}) позволяет адаптировать уровень верхнего диапазона в зависимости от

спектрального характера сигнала; эта оценка особенно важна, когда крутизна спектра CELP декодированного сигнала такова, что средняя энергия уменьшается, когда частота увеличивается (в случае вокализованного сигнала, где e_{tilt} близка к 1, поэтому $g_{SP} = 1 - e_{tilt}$ соответственно снижается). Следует также отметить, что коэффициент $\hat{g}_{HB}$ при AMR-WB декодировании ограничен принимать значения в интервале [0,1; 1,0]. Фактически, для сигналов, спектр которых имеет больше энергии на высоких частотах (e_{tilt} близко к -1, g_{SP} близко к 2), коэффициент усиления $\hat{g}_{HB}$, как правило, занижается.

При 23,85 кбит/с единица корректирующей информации передается на кодер AMR-WB и декодируется (блоки 107, 108) с целью уточнения коэффициента усиления, оцененного для каждого подкадра (4 бита каждые 5 мс или 0,8 кбит/с).

Искусственное возбуждение $u_{HB}(n)$ после этого фильтруется (блок 111) посредством синтезирующего фильтра синтеза LPC с передаточной функцией $1/A_{HB}(z)$ и работающего на частоте дискретизации 16 кГц. Структура данного фильтра зависит от скорости передачи текущего кадра:

При 6,6 кбит/с фильтр $1/A_{HB}(z)$ получается взвешиванием посредством коэффициента $\gamma=0,9$ фильтра LPC порядка 20, $1/\hat{A}^{ex}(z)$, который "экстраполирует" фильтр LPC порядка 16, $1/\hat{A}(z)$, декодированный в нижнем диапазоне (на 12,8 кГц) - подробная информация об экстраполяции в области параметров ISF (спектральной частоты иммитанса) описана в стандарте G.722.2 в разделе 6.3.2.1; в этом случае

$$1/A_{HB}(z) = 1/\hat{A}^{ex}(z/\gamma)$$

При скоростях цифрового потока $> 6,6$ кбит/с фильтр $1/A_{HB}(z)$ имеет порядок 16 и просто соответствует:

$$1/A_{HB}(z) = 1/\hat{A}(z/\gamma)$$

где $\gamma=0,6$. Следует отметить, что в этом случае фильтр $1/\hat{A}(z/\gamma)$ используется на частоте 16 кГц, что приводит к расширению (посредством пропорционального преобразования) частотной характеристики этого фильтра от [0; 6,4 кГц] до [0; 8 кГц].

В заключение, результат $s_{HB}(n)$ обрабатывается посредством полосового фильтра (блок 112) типа FIR ("конечной импульсной характеристики"), чтобы сохранить только диапазон 6-7 кГц; при скорости 23,85 кбит/с низкочастотный фильтр также типа FIR (блок 113) добавляется к обработке для дополнительного подавления частот выше 7 кГц. В заключение, высокочастотный (HF) синтез добавляется (блок 130) к низкочастотному (LF) синтезу, полученному посредством блоков 120-123 и подвергнутому передискретизации с частотой 16 кГц (блок 123). Таким образом, даже если верхний диапазон расширяется в теории от 6,4 до 7 кГц в кодеке AMR-WB, HF синтез скорее содержится в полосе 6-7 кГц перед суммированием с LF синтезом.

В методе расширения диапазона кодека AMR-WB можно выделить ряд недостатков:

Сигнал в верхнем диапазоне представляет собой сформированный белый шум (сформированный посредством временных коэффициентов усиления для каждого подкадра путем фильтрации посредством $1/A_{HB}(z)$ и полосовой фильтрации), который не является хорошей общей моделью сигнала в диапазоне 6,4-7 кГц. Например, существуют чрезвычайно гармонические музыкальные сигналы, для которых диапазон 6,4-7 кГц содержит синусоидальные составляющие (или тональные сигналы) и не содержит шума (или содержит мало шума); для этих сигналов расширение диапазона кодека AMR-WB значительно ухудшает качество.

Фильтр низких частот на 7 кГц (блок 113) вносит сдвиг почти 1 мс между нижним и верхним диапазонами, которые могут потенциально ухудшить качество определенных сигналов посредством незначительной десинхронизации двух диапазонов при 23,85 кбит/с - данная десинхронизация может также представлять проблемы при переключении скорости цифрового потока с 23,85 кбит/с на другие режимы.

Оценка коэффициентов усиления для каждого подкадра (блок 101, 103-105) не является оптимальной. Отчасти она основывается на выравнивании "абсолютной" энергии в расчете на один подкадр (блок 101) между сигналами на различных частотах: искусственным возбуждением на частоте 16 кГц (белый шум) и сигналом на частоте 12,8 кГц (декодированное возбуждение ACELP). Следует отметить, в частности, что этот подход в неявном виде вызывает затухание возбуждения в верхнем диапазоне (в соотношении $12,8/16=0,8$); собственно также следует отметить, что компенсация предискажений не выполняется в верхнем диапазоне в кодеке AMR-WB, который в неявном виде порождает усиление относительно близкое к 0,6 (что соответствует значению частотной характеристики $1/(1-0,68z^{-1})$ на частоте 6400 Гц). В действительности, коэффициенты 1/0,8 и 0,6 компенсируются приблизительно.

Что касается речи, характеристические тесты кодека 3GPP AMR-WB, задокументированные в 3GPP отчете TR 26.976, показали, что режим при 23,85 кбит/с имеет не такое хорошее качество, как при 23,05 кбит/с, при этом его качество в действительности подобно режиму при 15,85 кбит/с. Это показывает, в частности, что уровень искусственного HF сигнала необходимо контролировать очень аккуратно, так как качество ухудшается при скорости 23,85 кбит/с, при этом считается, что 4 бита на кадр наилучшим образом обеспечивают возможность аппроксимации энергии исходных высоких частот.

Ограничение кодированного диапазона на 7 кГц является следствием строгой модели

передаточной функции акустических терминалов (фильтр P.341 в стандарте ITU-T G.191). В настоящее время для частоты дискретизации 16 кГц остаются важными частоты в диапазоне 7-8 кГц, особенно для музыкальных сигналов, чтобы обеспечивать хороший уровень качества.

5 Алгоритм декодирования AMR-WB был частично улучшен с разработкой масштабируемого кодека ITU-T G.718, который был стандартизован в 2008 году.

Стандарт ITU-T G.718 содержит так называемый режим функциональной совместимости, для которого ядро кодирования совместимо с кодированием G.722.2 (AMR-WB) при 12,65 кбит/с; кроме того, декодер G.718 имеет конкретный признак того, чтобы быть в состоянии декодировать битовый поток AMR-WB/G.722.2 на всех возможных скоростях цифрового потока в кодеке AMR-WB- (от 6,6 до 23,85 кбит/с).

10 Функционально совместимый декодер G.718 в режиме малой задержки (G.718-LD) показан на фиг. 2. Ниже приведен список улучшений, предусмотренных функциональными возможностями декодирования AMR-WB битового потока в декодере G.718, со ссылками на фиг. 1 при необходимости:

Расширение диапазона (описано, например, в пункте 7.13.1 Рекомендации G.718, блок 206) идентично тому, что в декодере AMR-WB, за исключением того, что полосовой фильтр 6-7 кГц и синтезирующий фильтр $1/A_{NB}(z)$ (блоки 111 и 112) находятся в обратном порядке. Кроме того, при 23,85 кбит/с 4 бита, передаваемых в расчете на один подкадр кодером AMR-WB, не используются в функционально совместимом декодере G.718; следовательно, синтез высоких частот (HF) при 23,85 кбит/с идентичен синтезу при 23,05 кбит/с, что позволяет избежать известной проблемы качества декодирования AMR-WB при 23,85 кбит/с. Фильтр низких частот 7 кГц (блок 113) заведомо не используется, а специфическое декодирование режима 23,85 кбит/с не совершается (блоки 107-109).

25 Пост-обработка синтеза на 16 кГц (смотри пункт 7.14 G.718) реализуется в G.718 посредством "порогового шумоподавителя" в блоке 208 (для "улучшения" качества периодов молчания посредством снижения уровня), при этом высокочастотная фильтрация (блок 209), постфильтр низких частот (так называемый "постфильтр низких звуковых частот") в блоке 210 подавляют перекрестные гармонические помехи на низких частотах и преобразование в 16-битные целые числа с контролем насыщения (с управлением усиления или AGC) в блоке 211.

Тем не менее расширение диапазона в кодеках AMR-WB и/или G.718 (функционально совместимый режим) по-прежнему ограничивается рядом аспектов.

35 В частности, синтез высоких частот посредством формируемого белого шума (посредством временного подхода типа входного фильтра LPC) является весьма ограниченной моделью сигнала в диапазоне частот выше 6,4 кГц.

Только диапазон 6,4-7 кГц искусственно повторно синтезируется, в то время как на практике более широкий диапазон (до 8 кГц) теоретически возможен при частоте дискретизации 16 кГц, что потенциально может способствовать повышению качества сигналов, если они не являются предварительно обработанными посредством фильтра типа P.341 (50-7000 Гц), как определено в *Software Tool Library* (стандарт G.191) ITU-T.

40 Следовательно, существует необходимость улучшить расширение диапазона в кодеке типа AMR-WB или функционально совместимой версии этого кодека или, в более общем смысле, улучшить расширение диапазона звукового сигнала, в частности, для того чтобы улучшить частотный состав расширения диапазона.

Настоящее изобретение улучшает ситуацию.

Изобретение предлагает для этой цели способ расширения диапазона частот сигнала звуковой частоты во время процесса декодирования или улучшения, включающего

этап получения сигнала, декодированного в первом диапазоне частот, называемом нижним диапазоном. Способ включает следующие этапы:

- извлечение тональных составляющих и сигнала окружения из сигнала, возникающего из декодированного сигнала нижнего диапазона;

5 - объединение тональных составляющих и сигнала окружения посредством адаптивного микширования с использованием коэффициентов регулирования уровня энергии для получения звукового сигнала, называемого объединенный сигнал;

10 - расширение по меньшей мере одного второго диапазона частот, находящегося выше, чем первый диапазон частот декодированного сигнала нижнего диапазона, до этапа извлечения или объединенного сигнала после этапа объединения.

Следует отметить, что в дальнейшем термин "расширение диапазона" будет пониматься в широком смысле и будет включать не только случай расширения поддиапазона на высоких частотах, но и случай замены поддиапазонов, которые установлены в ноль (типа "шумового наполнения" в кодировании с преобразованием).

15 Таким образом, в одно и то же время с учетом тональных составляющих и сигнала окружения, извлеченного из сигнала, возникающего из декодирования нижнего диапазона, можно выполнить расширение диапазона с моделью сигнала, подходящей к истинной природе сигнала в противоположность использованию искусственного шума. Качество расширения диапазона, таким образом, улучшается, в частности, для
20 определенных типов сигналов, таких как музыкальные сигналы.

Действительно, сигнал, декодированный в нижнем диапазоне частот, содержит часть, соответствующую звуковому окружению, которая может быть перенесена в высокую частоту таким образом, что микширование гармонических составляющих и существующего окружения позволяет обеспечить целостный восстановленный верхний
25 диапазон.

Следует отметить, что даже если изобретение продиктовано улучшением качества расширения диапазона в контексте функционально совместимого кодирования AMR-WB, различные варианты осуществления применяются к более общему случаю расширения диапазона звукового сигнала, в частности, в устройстве улучшения качества,
30 выполняющем анализ звукового сигнала, чтобы извлечь параметры, необходимые для расширения диапазона.

Различные конкретные варианты осуществления, указанные ниже, могут быть добавлены по отдельности или в сочетании друг с другом к этапам способа расширения, определенного выше.

35 В одном варианте осуществления расширение диапазона выполняется в области возбуждения и декодированный сигнал нижнего диапазона представляет собой декодированный сигнал возбуждения нижнего диапазона.

Преимущество этого варианта осуществления заключается в том, что в области возбуждения возможно преобразование без обработки методом окна (или, что то же
40 самое, с неявным прямоугольным окном длины кадра). Тогда в этом случае не слышен артефакт (блок эффектов).

В первом варианте осуществления извлечение тональных составляющих и сигнала окружения выполняется в соответствии со следующими этапами:

- обнаружение в частотной области преобладающих тональных составляющих декодированного или декодированного и расширенного сигнала нижнего диапазона;
- вычисление остаточного сигнала посредством извлечения преобладающих тональных составляющих, чтобы получить сигнал окружения.

Данный вариант осуществления обеспечивает точное обнаружение тональных

составляющих.

Во втором варианте осуществления, низкой сложности, извлечение тональных составляющих и сигнала окружения выполняется в соответствии со следующими этапами:

- получение сигнала окружения посредством вычисления среднего значения спектра декодированного или декодированного и расширенного сигнала нижнего диапазона;
- получение тональных составляющих посредством вычитания вычисленного сигнала окружения из декодированного или декодированного и расширенного сигнала нижнего диапазона.

В одном варианте осуществления на этапе объединения коэффициент регулирования уровня энергии, применяемый для адаптивного микширования, вычисляется в зависимости от полной энергии декодированного или декодированного и расширенного сигнала нижнего диапазона и тональных составляющих.

Применение этого коэффициента регулирования позволяет адаптировать этап объединения к характеристикам сигнала таким образом, чтобы оптимизировать относительную долю сигнала окружения в микшированном сигнале. Уровень энергии, таким образом, регулируется так, чтобы избежать звуковых артефактов.

В предпочтительном варианте осуществления декодированный сигнал нижнего диапазона проходит этап преобразования или декомпозиции поддиапазона на основе блока фильтров, затем в частотной области или области поддиапазона выполняются этапы извлечения и объединения.

Реализация расширения диапазона в частотной области позволяет получить высокое качество частотного анализа, которое не доступно при временном подходе, и также позволяет иметь разрешение по частоте, являющееся достаточным для обнаружения тональных составляющих.

В подробном варианте осуществления декодированный и расширенный сигнал нижнего диапазона получается в соответствии со следующим уравнением:

$$U_{\text{HBI}}(k) = \begin{cases} 0 & k = 0, L, 199 \\ U(k) & k = 200, L, 239 \\ U(k + \text{start_band} - 240) & k = 240, L, 319 \end{cases}$$

где k - индекс дискретного значения, $U(k)$ - спектр сигнала, полученного после этапа преобразования, $U_{\text{HBI}}(k)$ - спектр расширенного сигнала, и start_band - предопределенная переменная.

Таким образом, эта функция включает передискретизацию сигнала посредством добавления дискретных значений к спектру этого сигнала. Однако возможны и другие способы расширения сигнала, например, посредством транспонирования в обработке поддиапазона.

Настоящее изобретение также предусматривает устройство для расширения диапазона частот сигнала звуковой частоты, при этом сигнал декодирован в первом диапазоне частот, называемом нижним диапазоном. Устройство содержит:

- модуль для извлечения тональных составляющих и сигнала окружения на основе сигнала, возникающего из декодированного сигнала нижнего диапазона;
- модуль для объединения тональных составляющих и сигнала окружения посредством адаптивного микширования с использованием коэффициентов регулирования уровня энергии для получения звукового сигнала, называемого объединенный сигнал;
- модуль для расширения на по меньшей мере одном втором диапазоне частот,

находящемся выше, чем первый диапазон частот, реализованный на декодированном сигнале нижнего диапазона перед модулем извлечения или на объединенном сигнале после модуля объединения.

Данное устройство демонстрирует те же преимущества, что и описанный ранее способ, который оно реализует.

Целью настоящего изобретения является декодер, содержащий описанное устройство.

Целью является компьютерная программа, содержащая команды программного кода для реализации этапов способа расширения диапазона, описанного выше, при выполнении этих команд процессором.

Наконец, настоящее изобретение относится к носителю данных, который может считываться процессором, встроенным или нет в устройство расширения диапазона, по возможности съемному, хранящему компьютерную программу, реализующую способ расширения диапазона, как описано выше.

Другие признаки и преимущества настоящего изобретения станут более очевидными из нижеследующего описания, приведенного только в качестве неограничивающего примера и со ссылкой на прилагаемые графические материалы, на которых:

на фиг. 1 показана часть декодера типа AMR-WB, реализующая этапы расширения диапазона частот согласно известному уровню техники и описанная выше;

на фиг. 2 показан декодер функционально совместимого типа 16 кГц G.718-LD

согласно известному уровню техники и описанный выше;

на фиг. 3 показан декодер, который является функционально совместимым с кодированием AMR-WB, имеющий в своем составе устройство расширения диапазона в соответствии с вариантом осуществления настоящего изобретения;

на фиг. 4 в виде блок-схемы показаны основные этапы способа расширения диапазона в соответствии с вариантом осуществления настоящего изобретения;

на фиг. 5 показан вариант осуществления в частотной области встроенного в декодер устройства расширения диапазона в соответствии с настоящим изобретением; и

на фиг. 6 показана аппаратная реализация устройства расширения диапазона в соответствии с настоящим изобретением.

На фиг. 3 показан примерный декодер, совместимый со стандартом AMR-WB/G.722.2, в котором присутствует пост-обработка, подобная той, которая представлена в G.718 и описана со ссылкой на фиг. 2, и улучшенное расширение диапазона в соответствии со способом расширения согласно настоящему изобретению, реализованный посредством устройства расширения диапазона, проиллюстрированным блоком 309.

В отличие от декодирования AMR-WB, которое работает с частотой дискретизации выходного сигнала 16 кГц, и G.718 декодера, который работает на 8 или 16 кГц, в данной заявке рассматривается декодер, который может работать с выходным сигналом (синтеза) на частоте $f_s=8, 16, 32$ или 48 кГц. Следует отметить, что при этом предполагается, что кодирование было выполнено в соответствии с алгоритмом AMR-

WB с внутренней частотой 12,8 кГц для кодирования CELP нижнего диапазона и при 23,85 кбит/с интервального кодирования подкадра на частоте 16 кГц, но также возможны функционально совместимые варианты кодера AMR-WB; несмотря на то, что изобретение описывается в данном случае на уровне декодирования, при этом предполагается, что кодирование может также работать с входным сигналом на частоте

$f_s=8, 16, 32$ или 48 кГц, и соответствующие операции передискретизации, выходящие за объем настоящего изобретения, реализуются в кодировании в зависимости от значения f_s . Следует отметить, что при $f_s=8$ кГц в декодере, в случае декодирования, который совместим с AMR-WB, нет необходимости расширять нижний диапазон 0-6,4 кГц,

поскольку восстановленный звуковой диапазон на частоте f_s ограничивается 0-4000 Гц.

На фиг. 3 декодирование CELP (LF - низких частот) по-прежнему работает на внутренней частоте 12,8 кГц, как в AMR-WB и G.718, и расширение диапазона (HF - высоких частот), которое является предметом настоящего изобретения работает на частоте 16 кГц, и синтезы LF и HF объединяются (блок 312) на частоте f_s после надлежащей передискретизации (блоки 307 и 311). В вариантах изобретения объединение нижнего и верхнего диапазонов может быть выполнено на частоте 16 кГц после передискретизации нижнего диапазона из 12,8 в 16 кГц до передискретизации объединенного сигнала с частотой f_s .

Декодирование в соответствии с фиг. 3 зависит от режима AMR-WB (или скорости цифрового потока), связанного с текущим принятым кадром. В качестве индикатора, и без воздействия на блок 309, декодирование части CELP в нижнем диапазоне включает следующие этапы:

демультиплексирование кодированных параметров (блок 300) в случае правильно принятого кадра ($bfi=0$, где bfi - "индикатор плохого кадра" со значением 0 для принятого кадра и 1 для потерянного кадра);

декодирование параметров ISF с интерполяцией и преобразованием в коэффициенты LPC (блок 301), как описано в пункте 6.1 стандарта G.722.2;

декодирование возбуждения CELP (блок 302) с адаптивной и фиксированной частью для восстановления возбуждения (exc или $u'(n)$) в каждом подкадре длиной 64 на частоте 12,8 кГц:

$$u'(n) = \hat{g}_p v(n) + \hat{g}_c c(n), \quad n = 0, L, 63$$

следуя обозначениям пункта 7.1.2.1 G.718 относительно декодирования CELP, где $v(n)$ и $c(n)$ - кодовые слова адаптивных и фиксированных словарей соответственно и $\hat{g}_p$ и $\hat{g}_c$ - связанные декодированные коэффициенты усиления. Данное возбуждение $u'(n)$ используется в адаптивном словаре следующего подкадра; затем он

обрабатывается и, как в G.718, возбуждение $u'(n)$ (также обозначаемое exc) выделяется из своей модифицированной постобработанной версии $u(n)$ (также обозначаемой $exc2$), которая служит в качестве входных данных для синтезирующего фильтра $1/\hat{A}(z)$ в блоке 303. В вариантах, которые могут быть реализованы для изобретения, операции пост-обработки, применяемые к возбуждению, могут быть изменены (например, дисперсия фазы может быть увеличена) или данные операции пост-обработки могут быть расширены (например, может быть реализовано снижение перекрестных гармонических помех), не влияя на сущность способа расширения диапазона в соответствии с настоящим изобретением;

синтезирующую фильтрацию посредством $1/\hat{A}(z)$ (блок 303), где декодированный фильтр LPC $\hat{A}(z)$ имеет 16 порядок;

узкополосную пост-обработку (блок 304) в соответствии с пунктом 7.3 G.718, если $f_s=8$ кГц;

компенсацию предискажений (блок 305) посредством фильтра $1/(1-0,68z^{-1})$;

пост-обработку низких частот (блок 306), как описано в пункте 7.14.1.1 G.718. Данная обработка вводит задержку, которая учитывается при декодировании верхнего диапазона ($> 6,4$ кГц);

передискретизацию внутренней частоты 12,8 кГц на выходную частоту f_s (блок 307). Возможен ряд вариантов осуществления. Без потери общности, в данном случае в качестве примера считается, что если $f_s=8$ или 16 кГц, то передискретизация, описанная в пункте 7.6 G.718, повторяется, и если $f_s=32$ или 48 кГц, то используются

дополнительные фильтры с конечной импульсной характеристикой (FIR);
вычисление параметров "порогового шумоподавителя" (этап 308), которое предпочтительно выполняется, как описано в пункте 7.14.3 G.718.

В вариантах, которые могут реализовываться для изобретения, операции пост-обработки, применяемые к возбуждению могут изменяться (например, дисперсия фазы может увеличиваться) или данные операции пост-обработки могут быть расширены (например, может реализовываться снижение перекрестных гармонических помех), не влияя на характер расширения диапазона. В данной заявке не описывается случай декодирования нижнего диапазона, когда теряется текущий кадр ($bfi=1$), который является информативным в стандарте 3GPP AMR-WB; в целом, либо имеют дело с декодером AMR-WB, либо декодером общего вида, опирающимся на модель входного фильтра, он, как правило, связан с наилучшими оценками возбуждения LPC и коэффициентами синтезирующего фильтра LPC таким образом, чтобы восстанавливать потерянный сигнал при сохранении модели входного фильтра. При $bfi=1$ считается, что расширение диапазона (блок 309) может работать как в случае $bfi=0$ и скорости цифрового потока <23,85 кбит/с; таким образом, описание настоящего изобретения будет в дальнейшем предполагать, без потери общности, что $bfi=0$.

Следует отметить, что использование блоков 306, 308, 314 не является обязательным.

Кроме того, следует отметить, что декодирование нижнего диапазона, описанное выше, предполагает так называемый "активный" текущий кадр со скоростью цифрового потока от 6,6 до 23,85 кбит/с. Фактически, когда режим DTX активируется, определенные кадры могут кодироваться как "неактивные", и в этом случае можно либо передать дескриптор паузы (по 35 бит), либо не передавать ничего. В частности, следует напомнить о том, что кадр SID кодера AMR-WB описывает несколько параметров: параметры ISF, усредненные по 8 кадрам, среднюю энергию по 8 кадрам, "флаг сглаживания" для восстановления нестационарного шума. Во всех случаях в декодере используется такая же модель декодирования, как и для активного кадра, с восстановлением возбуждения и фильтра LPC для текущего кадра, что дает возможность применять настоящее изобретение даже к неактивным кадрам. То же самое замечание относится к декодированию "потерянных кадров" (или FEC, PLC), в которых применяется LPC модель.

Данный примерный декодер работает в области возбуждения и, следовательно, содержит этап декодирования сигнала возбуждения нижнего диапазона. Устройство расширения диапазона и способ расширения диапазона в понимании настоящего изобретения также работают в области, отличной от области возбуждения и, в частности, с декодированным прямым сигналом нижнего диапазона или сигналом, взвешенным фильтром, учитывающим восприятие.

В отличие от декодирования AMR-WB или G.718, описанный декодер позволяет расширить декодированный нижний диапазон (50-6400 Гц с учетом 50 Гц фильтрации верхних частот в декодере, 0-6400 Гц в общем случае) до расширенного диапазона, ширина которого изменяется в диапазоне приблизительно от 50-6900 Гц до 50-7700 Гц, в зависимости от режима, реализованного в текущем кадре. Таким образом, его можно отнести к первому диапазону частот от 0 до 6400 Гц и второму диапазону частот от 6400 до 8000 Гц. Действительно, в пользуемого преимуществе варианте

осуществления возбуждение для высоких частот и генерируется в частотной области в диапазоне от 5000 до 8000 Гц, чтобы обеспечить полосовую фильтрацию шириной от 6000 до 6900 или 7700 Гц, наклон которой не слишком крутой в заграждаемом верхнем диапазоне.

5 Синтезирующую часть верхнего диапазона получают в блоке 309, представляющем устройство расширения диапазона в соответствии с настоящим изобретением, и которое подробно изображено на фиг. 5 в варианте осуществления.

В целях совмещения декодированных нижних и верхних диапазонов, вводится задержка (блок 310) для синхронизации выходных сигналов блоков 306 и 309, и верхний
10 диапазон, синтезируемый на частоте 16 кГц, подвергается передискретизации из 16 кГц в частоту f_s (выходной сигнал блока 311). Значение задержки T должно быть адаптировано для других случаев ($f_s=32, 48$ кГц) в зависимости от реализуемых операций обработки. Следует напомнить, что при $f_s=8$ кГц не нужно применять блоки 309-311, потому что диапазон сигнала на выходе декодера ограничивается значениями 0-4000
15 Гц.

Следует отметить, что способ расширения согласно изобретению, реализованный в блоке 309 в соответствии с первым вариантом осуществления, предпочтительно не вносит никакой дополнительной задержки относительно нижнего диапазона восстановленного на 12,8 кГц; тем не менее, в вариантах осуществления настоящего
20 изобретения (например, за счет использования преобразования время/частота с перекрытием), может быть внесена задержка. Таким образом, в целом, значение T в блоке 310 должно регулироваться в соответствии с конкретной реализацией. Например, в том случае, когда последующая обработка низких частот (блок 306) не используется, задержка, вносимая для $f_s=16$ кГц, может быть зафиксирована на $T=15$.

25 Затем нижние и верхние диапазоны объединяются (складываются) в блоке 312 и полученный синтез подвергается пост-обработке посредством высокочастотной фильтрации (типа IIR) 50 Гц порядка 2, коэффициенты которой зависят от частоты f_s (блок 313) и выходной пост-обработки с необязательным применением "*порогового шумоподавителя*" способом, подобным G.718 (блок 314).

30 Устройство расширения диапазона согласно настоящему изобретению, проиллюстрированное блоком 309 в соответствии с вариантом осуществления декодера на фиг. 5, реализует способ расширения диапазона (в широком смысле), описанный теперь со ссылкой на фиг. 4.

Данное устройство расширения также может быть независимым от декодера и может
35 реализовывать способ, описанный на фиг. 4, чтобы выполнить расширение диапазона существующего звукового сигнала, хранящегося или передающегося в устройство с анализом звукового сигнала, чтобы извлечь из него, например, возбуждение и фильтр LPC.

Данное устройство принимает декодированный сигнал в качестве входного в первом
40 диапазоне частот, называемым нижним диапазоном $u(n)$, который может находиться в области возбуждения либо сигнала. В варианте осуществления, описанном в данной заявке, этап декомпозиции поддиапазона (E401b) посредством преобразования время-частота или блока фильтров применяется к декодированному сигналу нижнего
45 диапазона для получения спектра декодированного сигнала $U(k)$ нижнего диапазона для реализации в частотной области.

Этап E401a расширения декодированного сигнала нижнего диапазона во втором диапазоне частот, находящемся выше, чем первый диапазон частот, для того чтобы

получить расширенный декодированный сигнал $U_{HBA}(k)$ нижнего диапазона, может быть выполнен на этом декодированном сигнале нижнего диапазона до или после этапа анализа (декомпозиции на поддиапазоны). Данный этап расширения может включать одновременно этап передискретизации и этап расширения или просто этап транспонирования или транспозиции частоты в зависимости от сигнала, полученного на входе. Следует отметить, что в вариантах этап E401a может быть выполнен в конце обработки, описанной на фиг. 4, то есть на объединенном сигнале, при этом данная обработка осуществляется в основном на сигнале нижнего диапазона перед расширением, причем результат является равноценным.

Данный этап подробно описан далее в варианте осуществления, описанном со ссылкой на фиг. 5.

Этап E402 извлечения сигнала окружения ($U_{HBA}(k)$) и тональных составляющих ($y(k)$) осуществляется на основании декодированного сигнала нижнего диапазона ($U(k)$) или декодированного и расширенного сигнала нижнего диапазона ($U_{HBA}(k)$). При этом окружение определяется как остаточный сигнал, который получается посредством удаления основных (или преобладающих) гармоник (или тональных составляющих) из существующего сигнала.

В большинстве широкополосных сигналов (дискретизированных на частоте 16 кГц), высокая частота (> 6 кГц) содержит информацию, которая в целом аналогична той, что присутствует в нижнем диапазоне.

Этап извлечения тональных составляющих и сигнала окружения включает, например, следующие этапы:

- обнаружение преобладающих тональных составляющих декодированного (или декодированного и расширенного) сигнала нижнего диапазона в частотной области;

- и
- вычисление остаточного сигнала посредством извлечения преобладающих тональных составляющих, чтобы получить сигнал окружения.

Этот этап также может быть получен посредством:

- получения сигнала окружения посредством вычисления среднего значения декодированного (или декодированного и расширенного) сигнала нижнего диапазона;

- и
- получения тональных составляющих посредством вычитания вычисленного сигнала окружения из декодированного или декодированного и расширенного сигнала нижнего диапазона.

Тональные составляющие и сигнал окружения затем объединяются адаптивным образом посредством коэффициентов регулирования уровня энергии на этапе E403, чтобы получить так называемый объединенный сигнал ($U_{HB2}(k)$). Затем может быть реализован этап E401a расширения, если он еще не был выполнен на декодированном сигнале нижнего диапазона.

Таким образом, объединение этих двух типов сигналов позволяет получить объединенный сигнал с характеристиками, которые являются более подходящими для определенных типов сигналов, таких как музыкальные сигналы, и более богатых по частотному составу и в расширенном диапазоне частот, соответствующем всему диапазону частот, включающему первый и второй диапазон частот.

Расширение диапазона в соответствии со способом, улучшает качество сигналов этого типа в отношении расширения, описанного в стандарте AMR-WB.

Использование объединения сигнала окружения и тональных составляющих позволяет обогатить данный сигнал расширения, для того чтобы сделать его ближе к характеристикам подлинного сигнала, а не искусственного сигнала.

Данный этап объединения будет детально описан далее со ссылкой на фиг. 5.

5 Этап синтеза, который соответствует анализу в 401b, выполняется в E404b для восстановления сигнала во временной области.

При необходимости этап регулирования уровня энергии сигнала верхнего диапазона может быть выполнен в E404a до и/или после этапа синтеза посредством применения коэффициента усиления и/или посредством соответствующей фильтрации. Этот этап
10 будет пояснен более подробно в варианте осуществления, описанном на фиг. 5, для блоков 501-507.

В примерном варианте осуществления устройство 500 расширения диапазона в данный момент описывается со ссылкой на фиг. 5, иллюстрирующей одновременно и это устройство, но также и модули обработки, подходящие для реализации в декодере
15 функционально совместимого типа с кодированием AMR-WB. Данное устройство 500 реализует способ расширения диапазона, описанный выше со ссылкой на фиг. 4.

Таким образом, блок 510 обработки принимает декодированный сигнал нижнего диапазона ($u(n)$). В конкретном варианте осуществления расширение диапазона
20 использует декодированное возбуждение на 12,8 кГц (exc2 или $u(n)$) в качестве выходного сигнала посредством блока 302, показанного на фиг. 3.

Этот сигнал подвергается декомпозиции на частотные поддиапазоны модулем 510 декомпозиции поддиапазона (который реализует этап E401b на фиг. 4), который в общем случае осуществляет преобразование или применяет блок фильтров, чтобы
25 получить декомпозицию на поддиапазоны $U(k)$ сигнала $u(n)$.

В конкретном варианте осуществления преобразование типа DCT-IV ("дискретное косинусное преобразование" -IV типа) (блок 510) применяется к текущему кадру, оставляющему 20 мс (256 дискретных значений), без обработки методом окна, что сводится к непосредственному преобразованию $u(n)$ при $n = 0, L, 255$ в соответствии
30 со следующей формулой:

$$U(k) = \sum_{n=0}^{N-1} u(n) \cos \left(\frac{\pi}{N} \left(n + \frac{1}{2} \right) \left(k + \frac{1}{2} \right) \right)$$

где $N = 256$ и $k = 0, L, 255$.

35 Возможно преобразование без обработки методом окна (или, что тоже самое, с неявным прямоугольным окном длины кадра), когда обработка выполняется в области возбуждения, а не в области сигнала. В этом случае не слышен артефакт (межблоковые эффекты), что представляет собой значительное преимущество данного варианта осуществления настоящего изобретения.

40 В данном варианте осуществления преобразование DCT-IV реализуется посредством FFT в соответствии с так называемым "Evolved DCT (EDCT)" алгоритмом, описанным в статье D.M. Zhang, H.T. Li, A Low Complexity Transform - Evolved DCT, IEEE 14th International Conference on Computational Science and Engineering (CSE), Aug. 2011, pp. 144-149, и реализованным в стандартах ITU-T G.718 Приложение В и G.729.1 Приложение Е.
45 Е.

В вариантах изобретения, и без потери общности, преобразование DCT-IV сможет быть заменено другими кратковременными преобразованиями время-частота той же длины в области возбуждения или в области сигнала, такими как FFT ("быстрое

преобразование Фурье") или DCT-II (дискретное косинусное преобразование II типа). В альтернативном варианте возможна замена DCT-IV в кадре на преобразование с перекрытием-сложением и применением окна с длиной, большей, чем длина текущего кадра, например с помощью MDCT ("модифицированного дискретного косинусного преобразования"). В этом случае задержка T в блоке 310, показанном на фиг. 3, должна быть должным образом отрегулирована (уменьшена) в зависимости от дополнительной задержки из-за анализа/синтеза посредством данного преобразования.

В другом варианте осуществления изобретения декомпозиция поддиапазона выполняется посредством применения блока действительных или комплексных фильтров, например, типа PQMF (псевдо-QMF). Для некоторых блоков фильтров для каждого поддиапазона в данном кадре, получается не спектральное значение, а ряд временных значений, связанных с поддиапазоном; в данном случае вариант осуществления, пользующийся преимуществом в данном изобретении, может быть применен при проведении, например, преобразования каждого поддиапазона и посредством вычисления сигнала окружения в области абсолютных значений, при этом тональные составляющие, по-прежнему получают посредством вычисления разности между сигналом (по абсолютной величине) и сигнала окружения. В случае блока комплексных фильтров, комплексный модуль дискретных значений заменит абсолютное значение.

В других вариантах осуществления настоящее изобретение будет применяться в системе, использующей два поддиапазона, при этом нижний диапазон анализируется посредством преобразования или посредством блока фильтров.

После этого в случае DCT спектр DCT $U(k)$ 256 дискретных значений, охватывающих диапазон 0-6400 Гц (на 12,8 кГц), расширяется (блок 511) в спектр 320 дискретных значений, охватывающих диапазон 0-8000 Гц (на 16 кГц) в следующем виде:

$$U_{\text{нвл}}(k) = \begin{cases} 0 & k = 0, L, 199 \\ U(k) & k = 200, L, 239 \\ U(k + \text{start_band} - 240) & k = 240, L, 319 \end{cases}$$

где предпочтительно принимается, что $\text{start_band} = 160$.

Блок 511 реализует этап E401a на фиг. 4, то есть расширение декодированного сигнала нижнего диапазона. Этот этап может также содержать передискретизацию из 12,8 в 16 кГц в частотной области посредством добавления $\frac{1}{4}$ дискретных значений ($k = 240, L, 319$) к спектру, при этом отношение 16 к 12,8 составляет 5/4.

В диапазоне частот, соответствующем дискретным значениям, лежащим от индексов 200 до 239, исходный спектр сохраняется для возможности применения к нему характеристики монотонно возрастающего затухания фильтра высоких частот в этом диапазоне частот, а также, чтобы не вносить слышимые дефекты на этапе добавления низкочастотного синтеза к высокочастотному синтезу.

Следует отметить, что в этом варианте осуществления генерация расширенного спектра с избыточной частотой дискретизации выполняется в диапазоне частот, находящемся от 5 до 8 кГц, следовательно, включающем второй диапазон частот (6,4-8 кГц), лежащий выше первого диапазона частот (0- 6,4 кГц).

Таким образом, расширение декодированного сигнала нижнего диапазона выполняется по меньшей мере на втором диапазоне частот, но и на части первого диапазона частот.

Очевидно, что значения, определяющие эти диапазоны частот, могут быть разными в зависимости от декодера или устройства обработки, в котором применяется

изобретение.

Кроме того, блок 511 выполняет неявную высокочастотную фильтрацию в диапазоне 0-5000 Гц, так как первые 200 дискретных значений $U_{HBI}(k)$ устанавливаются в нуль; как описано ниже, данная высокочастотная фильтрация также может быть дополнена частью монотонно возрастающего затухания спектральных значений индексов $k = 200, L, 255$ в диапазоне 5000-6400 Гц; это монотонно возрастающее затухание реализуется в блоке 501, но может быть выполнено отдельно за пределами блока 501. Равносильно тому, что и в вариантах настоящего изобретения, внедрение высокочастотной фильтрации, разделенной на блоки коэффициентов индекса $k = 0, L, 199$, установленных в ноль, затухающих коэффициентов $k = 200, L, 255$ в преобразованной области, можно будет, в силу вышесказанного, выполнять в ходе одного этапа.

В этом примерном варианте осуществления и в соответствии с определением $U_{HBI}(k)$, следует отметить, что диапазон 5000-6000 Гц $U_{HBI}(k)$ (что соответствует индексам $k = 200, L, 239$) копируется из диапазона 5000-6000 Гц $U(k)$. Данный подход позволяет сохранить исходный спектр в этом диапазоне и позволяет избежать введения искажения в диапазоне 5000-6000 Гц при сложении синтеза HF с синтезом LF, в частности, фаза сигнала (неявно представленная в области DCT-IV) в данном диапазоне сохраняется.

Диапазон 6000-8000 Гц $U_{HBI}(k)$ в данном случае определяется посредством копирования диапазона 4000-6000 Гц $U(k)$, поскольку значение *start_band* предпочтительно устанавливается равным 160.

В одном из вариантов варианта осуществления значение *start_band* будет иметь возможность быть адаптивным вблизи значения 160, не изменяя сущность изобретения. Подробности подстройки значения *start_band* в данном документе не описываются, потому что они выходят за рамки изобретения без изменения его объема.

В большинстве широкополосных сигналов (подвергнутых дискретизации с частотой 16 кГц) верхний диапазон (>6 кГц) содержит информацию окружения, которая естественным образом подобна той, что присутствует в нижнем диапазоне. При этом окружение определяется как остаточный сигнал, который получается посредством удаления основных (или преобладающих) гармоник из существующего сигнала. Уровень гармоничности в диапазоне 6000-8000 Гц, как правило, коррелирует с уровнем в более низкочастотных диапазонах.

Этот декодированный и расширенный сигнал нижнего диапазона предусматривается в качестве входного сигнала для устройства 500 расширения и, в частности, в качестве входного сигнала для модуля 512. Таким образом, блок 512 для извлечения тональных составляющих и сигнала окружения реализует этап E402, приведенный на фиг. 4, в частотной области. Таким образом, получается сигнал окружения $U_{HBA}(k)$ для $k = 240, L, 319$ (80 дискретных значений), для второго диапазона частот, так называемого высокочастотным, для того чтобы объединить его в дальнейшем адаптивным способом с извлеченными тональными составляющими $y(k)$ в блоке 513 объединения.

В конкретном варианте осуществления извлечение тональных составляющих и сигнала окружения (в диапазоне 6000-8000 Гц) выполняется в соответствии со следующими операциями:

Вычисление полной энергии расширенного декодированного сигнала $ener_{HB}$ нижнего диапазона:

$$5 \quad ener_{HB} = \sum_{k=240}^{319} U_{HB}(k)^2 + \varepsilon$$

где $\varepsilon=0,1$ (это значение может быть разным, оно определено в данном случае в качестве примера).

Вычисление окружения (по абсолютной величине), которое соответствует в данном случае среднему уровню спектра $lev(i)$ (спектральная линия за спектральной линией) и вычисление энергии $ener_{tonal}$ преобладающих тональных частей (в высокочастотном спектре)

Для $i=0...L-1$ этот средний уровень получается с помощью следующего уравнения:

$$15 \quad lev(i) = \frac{1}{fn(i) - fb(i) + 1} \sum_{j=fb(i)}^{fn(i)} |U_{HB}(j+240)|$$

Это соответствует среднему уровню (по абсолютной величине) и, следовательно, представляет собой своего рода огибающую спектра. В данном варианте осуществления $L=80$ и представляет длину спектра и индекс i от 0 до $L-1$ соответствует индексам $j+240$ от 240 до 319, то есть спектру от 6 до 8 кГц.

В целом, $fb(i)=i-7$ и $fn(i)=i+7$, однако первый и последний 7 индексы ($i=0, L, 6$ и $i=L-7, L, L-1$) требуют специальной обработки, и без потери общности, тогда определяем:

$$25 \quad \begin{aligned} fb(i) &= 0 \text{ и } fn(i) = i+7 \text{ для } i=0, L, 6 \\ fb(i) &= i-7 \text{ и } fn(i) = L-1 \text{ для } i=L-7, L, L-1 \end{aligned}$$

В вариантах осуществления изобретения среднее из $|U_{HB}(j+240)|$, $j=fb(i), \dots, fn(i)$ может быть заменено на медианное значение по тому же множеству значений, т.е.

30 $lev(i) = median_{j=fb(i), \dots, fn(i)} (|U_{HB}(j+240)|)$ Данный вариант имеет недостаток, являясь более сложным (с точки зрения количества вычислений), чем скользящее среднее. В других вариантах неравномерное взвешивание может быть применено к усредненным членам, или медианная фильтрация может быть заменена, например, другими

35 нелинейными фильтрами типа "фильтров стека".

Также вычисляется остаточный сигнал:

$$y(i) = |U_{HB}(i+240)| - lev(i), \quad i=0, K, L-1$$

40 что соответствует (приблизительно) тональным составляющим, если значение $y(i)$ в данной спектральной линии i является положительным ($y(i) > 0$).

Таким образом, это вычисление предполагает неявное обнаружение тональных составляющих. Тональные части, таким образом, обнаруживаются в неявном виде посредством промежуточного члена $y(i)$, представляющего адаптивный порог. Причем условием обнаружения является то, что $y(i) > 0$. В вариантах изобретения это условие

45 может быть изменено, например, посредством определения адаптивного порога в зависимости от локальной огибающей сигнала или в форме $y(i) > lev(i) + x \text{ dB}$, где x имеет predetermined значение (например, $x=10$ дБ).

Энергия преобладающих тональных частей определяется следующим уравнением:

$$ener_{total} = \sum_{i=0, \dots, N-1} y(i)^2$$

Конечно, могут быть предусмотрены другие схемы для выделения сигнала окружения. Например, данный сигнал окружения может быть извлечен из низкочастотного сигнала или, при необходимости, из другого диапазона частот (или нескольких диапазонов частот).

Обнаружение тональных всплесков или составляющих может быть выполнено по-разному.

Извлечение данного сигнала окружения также может быть выполнено на декодированном, но не расширенном возбуждении, то есть до спектрального расширения или этапа транспонирования, то есть, например, на участке низкочастотного сигнала, а не непосредственно на высокочастотном сигнале.

В альтернативном варианте осуществления извлечение тональных составляющих и сигнала окружения выполняется в другом порядке и в соответствии со следующими этапами:

- обнаружение преобладающих тональных составляющих декодированного (или декодированного и расширенного) сигнала нижнего диапазона в частотной области;
- вычисление остаточного сигнала посредством извлечения преобладающих тональных составляющих, чтобы получить сигнал окружения.

Этот вариант может, например, быть осуществлен следующим образом: Всплеск (или тональная составляющая) обнаруживается в спектральной линии индекса i в спектре амплитуды $|U_{HBI}(i+240)|$, если удовлетворяется следующий критерий:

$$|U_{HBI}(i+240)| > |U_{HBI}(i+240-1)| \text{ и } |U_{HBI}(i+240)| > |U_{HBI}(i+240+1)|,$$

для $i = 0, K, L-1$. Как только обнаруживается всплеск в спектральной линии индекса i , применяется синусоидальная модель для того, чтобы оценить амплитуду, частоту и, при необходимости, фазовые параметры тональной составляющей, связанной с этим всплеском. Подробное описание этой оценки не представлено в данной заявке, но оценка частоты обычно может обратиться к параболической интерполяции по 3 точкам, для того чтобы найти максимум параболы, аппроксимирующей 3 точки амплитуды $|U_{HBI}(i+240)|$ (выраженной в дБ), при этом оценка амплитуды получается посредством этой же интерполяции. Поскольку область преобразования, используемая в данном случае (DCT-IV), не позволяет получить фазу непосредственно, то в одном варианте осуществления можно будет пренебречь этим членом, но в вариантах можно будет применить квадратурное преобразования типа DST для оценки фазовой составляющей. Начальное значение $y(i)$ устанавливается равным нулю для $i = 0, K, L-1$. При этом оцениваются синусоидальные параметры (частота, амплитуда и, при необходимости, фаза) каждой тональной составляющей, затем вычисляется член $y(i)$ как сумма предопределенных прототипов (спектров) чистых синусоид, преобразованных в область DCT-IV (или другую область, если используется какая-либо другая декомпозиция поддиапазона) в соответствии с оцененными синусоидальными параметрами. Наконец, применяется абсолютное значение к членам $y(i)$ для выражения области амплитудного спектра в виде абсолютных значений.

Возможны другие схемы определения тональных составляющих, например, также было бы можно вычислить огибающую сигнала $env(i)$ посредством сплайн-

интерполяции локальных максимальных значений (обнаруженных всплесков)

$|U_{HB1}(i+240)|$, чтобы понизить эту огибающую на определенный уровень в дБ с целью обнаружить тональные составляющие в виде всплесков, которые превышают эту огибающую и определить $y(i)$ как

$$y(i) = \max(|U_{HB1}(i+240)| - \text{env}(i), 0)$$

В данном варианте вследствие этого получается окружение с помощью уравнения:

$$\text{lev}(i) = |U_{HB1}(i+240)| - y(i), \quad i = 0, K, L-1$$

В других вариантах изобретения абсолютное значение спектральных значений будут заменены, например, квадратами спектральных значений, не изменяя принцип изобретения; в этом случае будет необходим квадратный корень для того, чтобы вернуться к области сигнала, причем это является более сложным для осуществления.

Модуль 513 объединения выполняет этап объединения посредством адаптивного микширования сигнала окружения и тональных составляющих. Соответственно, коэффициент регулирования Γ уровня окружения определяется следующим уравнением:

$$\Gamma = \beta \frac{\text{ener}_{HB} - \text{ener}_{\text{tonal}}}{\text{ener}_{HB} - \beta \text{ener}_{\text{tonal}}}$$

β - коэффициент, примерное вычисление которого приведено ниже.

Для получения расширенного сигнала сначала получают объединенный сигнал в абсолютных значениях для $i = 0 \dots L-1$:

$$y'(i) = \begin{cases} \Gamma y(i) + \frac{1}{\Gamma} \text{lev}(i) & y(i) > 0 \\ y(i) + \frac{1}{\Gamma} \text{lev}(i) & y(i) \leq 0 \end{cases}$$

к которому применяются знаки $U_{HB1}(k)$:

$$y''(i) = \text{sgn}(U_{HB1}(i+240)) \cdot y'(i)$$

где функция $\text{sgn}(\cdot)$ дает знак:

$$\text{sgn}(x) = \begin{cases} 1 & x \geq 0 \\ -1 & x < 0 \end{cases}$$

По определению коэффициент $\Gamma > 1$. Тональные составляющие, обнаруженная спектральная линия за спектральной линией посредством условия $y(i) > 0$, уменьшаются коэффициентом Γ ; средний уровень усиливается с коэффициентом $1/\Gamma$.

В блоке 513 адаптивного микширования, коэффициент регулирования уровня энергии вычисляется в зависимости от полной энергии декодированного (или декодированного и расширенного) сигнала нижнего диапазона и тональных составляющих.

В предпочтительном варианте осуществления адаптивного микширования, регулирование энергии осуществляется следующим образом:

$$U_{HB2}(k) = \text{fac} \cdot y''(k-240), \quad k = 240, L, 319$$

$U_{HB2}(k)$ - объединенный сигнал расширения диапазона.

Коэффициент регулирования определяется следующим уравнением:

$$fac = \gamma \sqrt{\frac{ener_{HB}}{\sum_{i=0}^{L-1} y''(i)}}$$

где γ позволяет избежать завышения энергии. В примерном варианте осуществления, вычисляют β таким образом, чтобы сохранить тот же уровень сигнала окружения по отношению к энергии тональных составляющих в следующих друг за другом диапазонах сигнала. Вычисляют энергию тональных составляющих в трех диапазонах: 2000-4000 Гц, 4000-6000 Гц и 6000-8000 Гц, с

$$E_{N2-4} = \sum_{k \in N(80,159)} U^2(k)$$

$$E_{N4-6} = \sum_{k \in N(160,239)} U^2(k)$$

$$E_{N4-6} = \sum_{k \in N(240,319)} U^2(k)$$

где

$$U'(k) = \begin{cases} \sqrt{\frac{\sum_{k=160}^{239} U^2(k)}{\sum_{k=80}^{159} U^2(k)}} U(k) & k = 80, \dots, 159 \\ U(k) & k = 160, \dots, 239 \\ \sqrt{\frac{\sum_{k=160}^{239} U^2(k)}{\sum_{k=240}^{319} U_{HB1}^2(k)}} U_{HB1}(k) & k = 240, \dots, 319 \end{cases}$$

и где $N(k_1, k_2)$ - множество индексов k , для которых коэффициент индекса k классифицируется как связанный с тональными составляющими. Это множество может быть получено, например, посредством определения локальных всплесков в $U'(k)$, удовлетворяющих $|U'(k)| > lev(k)$, или $lev(k)$ вычисляется как среднее значение уровня спектра спектральная линия за спектральной линией.

Следует отметить, что возможны другие схемы для вычисления энергии тональных составляющих, например, посредством принятия медианного значения спектра в рассматриваемом диапазоне.

Зафиксируем β таким образом, чтобы отношение энергии тональных составляющих в диапазонах 4-6 кГц и 6-8 кГц было таким же, как между диапазонами 2-4 кГц и 4-6 кГц:

$$\beta = \frac{\rho - E_{N6-8}}{\sum_{k=160}^{239} U^2(k) - E_{N6-8}}$$

где

$$E_{N4-6} = \max(E_{N4-6}, E_{N2-4}), \quad \rho = \frac{E_{N4-6}^2}{E_{N2-4}}, \quad \rho = \max(\rho, E_{N6-8})$$

и $\max(\dots)$ -функция, которая дает максимум двух аргументов.

В вариантах изобретения вычисление β может быть заменено другими схемами. Например, в варианте можно будет извлечь (вычислить) различные параметры (или "признаки"), характеризующие сигнал нижнего диапазона, включая параметр "наклона", подобный тому, который вычисляется в кодеке AMR-WB, и коэффициент β будет оцениваться в зависимости от линейной регрессии на основе этих различных параметров, ограничивая его значение в диапазоне от 0 до 1. Линейная регрессия, например, сможет быть оценена контролируемым образом посредством оценки коэффициента β , учитывая исходный верхний диапазон в обучающей базе. Следует отметить, что то, каким образом вычисляется β , не ограничивает сущность изобретения.

После этого параметр β может быть использован для вычисления γ , принимая во внимание тот факт, что сигнал с сигналом окружения, добавленным в заданном диапазоне, в общем случае воспринимается как более сильный, чем гармонический сигнал с той же энергией в том же диапазоне. Если определяют α , часть сигналов окружения добавляется к гармоническому сигналу:

$$\alpha = \sqrt{1 - \beta}$$

можно будет вычислить γ как убывающую функцию α , например $\gamma = b - a\sqrt{\alpha}$,

$b = 1,1$, $a = 1,2$ и γ , ограниченную от 0,3 до 1. При этом в рамках данного изобретения опять же возможны другие определения α и γ .

На выходе устройства 500 расширения диапазона блок 501 в конкретном варианте осуществления осуществляет при необходимости двойственную операцию применения частотной характеристики полосового фильтра и фильтрации предискажений (или частотной пост-коррекции) в частотной области.

В одном варианте изобретения фильтрация предискажений может выполняться во временной области после блока 502 до блока 510; однако, в этом случае полосовая фильтрация, выполняемая в блоке 501, может оставить определенные низкочастотные компоненты очень низких уровней, которые усиливаются посредством компенсации предискажений, что может несущественно изменять декодированный нижний диапазон. По этой причине, в данном случае предпочтительно выполнять компенсацию предискажений в частотной области. В предпочтительном варианте осуществления коэффициенты индекса $k = 0, L, 199$ устанавливаются равными нулю, таким образом, компенсация предискажений ограничивается старшими коэффициентами.

Сначала компенсируются предискажения возбуждения в соответствии со следующим уравнением:

$$U_{HB2}(k) = \begin{cases} 0 & k = 0, L, 199 \\ G_{deamp}(k-200)U_{HB2}(k) & k = 200, L, 255 \\ G_{deamp}(55)U_{HB2}(k) & k = 256, L, 319 \end{cases}$$

где $G_{deamp}(k)$ - частотная характеристика фильтра $1/(1-0,68z^{-1})$ в ограниченном дискретном диапазоне частот. Принимая во внимание дискретные (нечетные) частоты DCT-IV, $G_{deamp}(k)$ определяется в данном случае как:

$$G_{deemph}(k) = \frac{1}{|e^{j\theta_k} - 0,68|}, \quad k = 0, L, 255$$

где

$$\theta_k = \frac{256 - 80 + k + \frac{1}{2}}{256}.$$

В случае, когда применяется преобразование, отличное от DCT-IV, определение θ_k можно регулировать (например, для четных частот).

Следует отметить, что компенсация предсказаний применяется в два этапа для $k = 200, L, 255$, соответствующих диапазону частот 5000-6400 Гц, где характеристика $1/(1 - 0,68z^{-1})$ применяется как на частоте 12,8 кГц, и $k = 256, L, 319$, соответствующий диапазону частот 6400-8000 Гц, где характеристика расширяется от частоты 16 кГц до постоянного значения в диапазоне 6,4-8 кГц.

Следует отметить, что в кодеке AMR-WB для HF синтеза не компенсируются предсказания.

В варианте осуществления, представленном здесь, для высокочастотного сигнала, наоборот, компенсируются предсказания для того, чтобы вернуть его в область, согласованную с низкочастотным сигналом (0-6,4 кГц), который выходит из блока 305, приведенного на фиг. 3. Это важно для оценки и последующей регулировки энергии синтеза HF.

В одном варианте осуществления с целью уменьшения сложности можно установить $G_{deemph}(k)$ равным постоянному значению, независимому от k , принимая, например, $G_{deemph}(k) = 0,6$, что примерно соответствует среднему значению $G_{deemph}(k)$ для $k = 200, L, 319$ в условиях описанного выше варианта осуществления.

В другом варианте осуществления декодера компенсация предсказаний может быть проведена аналогичным образом во временной области после обратного DCT.

В дополнение к компенсации предсказаний применяется полосовая фильтрация с двумя отдельными частями: одна - фиксированная высокочастотная, другая - адаптивная низкочастотная (функция скорости цифрового потока).

Данная фильтрация выполняется в частотной области.

В предпочтительном варианте осуществления частичная характеристика фильтра низких частот вычисляется в частотной области следующим образом:

$$G_{lp}(k) = 1 - 0,999 \frac{k}{N_{lp} - 1}$$

где $N_{lp} = 60$ при скорости 6,6 кбит/с, 40 при скорости 8,85 кбит/с и 20 при скоростях цифрового потока $> 8,85$ бит/с.

Затем полосовой фильтр применяется в виде:

$$U_{HB3}(k) = \begin{cases} 0 & k = 0, L, 199 \\ G_{hp}(k-200)U_{HB2}(k) & k = 200, L, 255 \\ U_{HB2}(k) & k = 256, L, 319 - N_p \\ G_p(k-320 - N_p)U_{HB2}(k) & k = 320 - N_p, L, 319 \end{cases}$$

Определение $G_{hp}(k)$, $k = 0, L, 55$, дается, например, в таблице 1 ниже.

Таблица 1

K	$g_{hp}(k)$	K	$g_{hp}(k)$	K	$g_{hp}(k)$	K	$g_{hp}(k)$
0	0,001622428	14	0,114057967	28	0,403990611	42	0,776551214
1	0,004717458	15	0,128865425	29	0,430149896	43	0,800503267
2	0,008410494	16	0,144662643	30	0,456722014	44	0,823611104
3	0,012747280	17	0,161445005	31	0,483628433	45	0,845788355
4	0,017772424	18	0,179202219	32	0,510787115	46	0,866951597
5	0,023528982	19	0,197918220	33	0,538112915	47	0,887020781
6	0,030058032	20	0,217571104	34	0,565518011	48	0,905919644
7	0,037398264	21	0,238133114	35	0,592912340	49	0,923576092
8	0,045585564	22	0,259570657	36	0,620204057	50	0,939922577
9	0,054652620	23	0,281844373	37	0,647300005	51	0,954896429
10	0,064628539	24	0,304909235	38	0,674106188	52	0,968440179
11	0,075538482	25	0,328714699	39	0,700528260	53	0,980501849
12	0,087403328	26	0,353204886	40	0,726472003	54	0,991035206
13	0,100239356	27	0,378318805	41	0,751843820	55	1,000000000

Следует отметить, что в вариантах изобретения значения $G_{hp}(k)$ могут изменяться при сохранении монотонно возрастающего затухания. Аналогичным образом, низкочастотная фильтрация с переменной полосой пропускания $G_p(k)$ может регулироваться с помощью значений или частотной поддержки, которые отличаются, не изменяя принцип этого этапа фильтрации.

Кроме того, следует отметить, что полосовая фильтрация может адаптироваться посредством определения единого этапа фильтрации, объединяющего высокочастотную и низкочастотную фильтрацию.

В другом варианте осуществления полосовая фильтрация может выполняться равноценным образом во временной области (как и в блоке 112 на фиг. 1) с различными коэффициентами фильтра в соответствии со скоростью цифрового потока после этапа обратного DCT. Тем не менее, следует отметить, что предпочтительно выполнить этот этап непосредственно в частотной области, так как фильтрация выполняется в области возбуждения LPC и, следовательно, проблемы циклической свертки и краевых эффектов очень ограничены в этой области.

Блок 502 обратного преобразования выполняет обратное DCT на 320 дискретных значениях, чтобы найти высокочастотное возбуждение, подвергнутое дискретизации с частотой 16 кГц. Его реализация идентична блоку 510, потому что DCT-IV является ортонормированным, за исключением того, что длина преобразования составляет 320 вместо 256, и получается следующее:

$$u_{HB}(n) = \sum_{k=0}^{N_{16k}-1} U_{HB3}(k) \cos\left(\frac{\pi}{N_{16k}}\left(k + \frac{1}{2}\right)\left(n + \frac{1}{2}\right)\right)$$

где $N_{16k} = 320$ и $k = 0, L, 319$.

В том случае, когда блок 510 не является DCT, а некоторым другим преобразованием

или декомпозицией на поддиапазоны, блок 502 осуществляет синтез, соответствующий анализу, осуществленному в блоке 510.

Сигнал, подвергнутый дискретизации с частотой 16 кГц, после этого при необходимости масштабируется коэффициентами усиления, определенными в расчете на подкадр из 80 дискретных значений (блок 504).

В предпочтительном варианте осуществления коэффициент усиления $g_{HB1}(m)$ сначала вычисляется (блок 503) в расчете на подкадр посредством отношения энергии подкадров таким образом, что в каждом подкадре индекса $m=0, 1, 2$ или 3 текущего кадра:

$$g_{HB1}(m) = \sqrt{\frac{e_3(m)}{e_2(m)}}$$

где

$$e_1(m) = \sum_{n=0}^{63} u(n+64m)^2 + \varepsilon$$

$$e_2(m) = \sum_{n=0}^{79} u_{HB}(n+80m)^2 + \varepsilon$$

$$e_3(m) = e_1(m) \frac{\sum_{n=0}^{319} u_{HB}(n)^2 + \varepsilon}{\sum_{n=0}^{255} u(n)^2 + \varepsilon}$$

при $\varepsilon = 0,01$. Коэффициент усиления в расчете на подкадр $g_{HB1}(m)$ может быть записан в виде:

$$g_{HB1}(m) = \sqrt{\frac{\sum_{n=0}^{63} u(n+64m)^2 + \varepsilon}{\sum_{n=0}^{255} u(n)^2 + \varepsilon} \cdot \frac{\sum_{n=0}^{79} u_{HB}(n+80m)^2 + \varepsilon}{\sum_{n=0}^{319} u_{HB}(n)^2 + \varepsilon}}$$

что показывает, что в сигнале u_{HB} обеспечивается такое же отношение энергии на подкадр и энергии на кадр, как в сигнале $u(n)$.

Блок 504 выполняет масштабирование объединенного сигнала (включенного на этапе E404a фиг.4) в соответствии со следующим уравнением:

$$u_{HB}'(n) = g_{HB1}(m)u_{HB}(n), \quad n = 80m, L, 80(m+1)-1$$

Следует отметить, что реализация блока 503 отличается от блока 101, приведенного на фиг.1, так как энергия на текущем уровне кадра учитывается в дополнение к этому подкадру. Это обеспечивает отношение энергии каждого подкадра в зависимости от энергии кадра. Следовательно, сравниваются отношения энергии (или относительные энергии), а не абсолютные энергии нижнего диапазона и верхнего диапазона.

Таким образом, данный этап масштабирования позволяет сохранять в верхнем диапазоне отношение энергии подкадра и кадра таким же образом, как в нижнем диапазоне.

Затем при необходимости блок 506 выполняет масштабирование сигнала

(включенного на этапе E404a фиг.4) в соответствии со следующим уравнением:

$$u_{HB}''(n) = g_{HB2}(m)u_{HB}'(n), \quad n = 80m, L, 80(m+1) - 1$$

где коэффициент усиления $g_{HB2}(m)$ получается из блока 505 путем выполнения блоков 103, 104 и 105 кодека AMR-WB (при этом входной сигнал блока 103 представляет собой возбуждение, декодированное в нижнем диапазоне, $u(n)$). Блоки 505 и 506 могут быть использованы для регулирования уровня синтезирующего фильтра LPC (блок 507), в данном случае в зависимости от наклона сигнала. Возможны другие схемы для вычисления коэффициента усиления $g_{HB2}(m)$, не изменяя сущность изобретения.

В заключение, сигнал $u_{HB}'(n)$ или $u_{HB}''(n)$ фильтруется посредством модуля 507 фильтрации, который может быть воплощен в данном случае, посредством принятия в качестве передаточной функции $1/\hat{A}(z/\gamma)$, где $\gamma=0,9$ при 6,6 кбит/с и $\gamma=0,6$ при других скоростях цифрового потока, тем самым ограничивая порядок фильтра до порядка 16.

В одном варианте этот фильтр можно будет выполнять таким же образом, как описано для блока 111 на фиг. 1 AMR-WB декодера, но порядок фильтра изменятся до 20 при скорости цифрового потока 6,6, что не существенно меняет качество синтезированного сигнала. В другом варианте возможно выполнение синтезирующей фильтрации LPC в частотной области после вычисления частотной характеристики фильтра, реализованного в блоке 507.

В альтернативных вариантах осуществления настоящего изобретения кодирование нижнего диапазона (0-6,4 кГц) может заменяться кодером CELP, отличающимся от того, что используется в AMR-WB, таким как, например, кодер CELP в G.718 при скорости 8 кбит/с. Без потери общности могли бы применяться другие широкополосные кодеры или кодеры, работающие на частотах выше 16 кГц, в которых кодирование нижнего диапазона производится на внутренней частоте 12,8 кГц. Кроме того, очевидно, что изобретение может быть адаптировано к частотами дискретизации, отличающимся от частоты 12,8 кГц, когда низкочастотный кодер работает с частотой дискретизации ниже, чем у исходного или восстановленного сигнала. Когда при низкочастотном декодировании не используется линейное предсказание, отсутствует сигнал возбуждения для расширения, в этом случае возможно выполнение анализа LPC сигнала, восстановленного в текущем кадре, и возбуждение LPC будет вычисляться таким образом, чтобы была возможность применения изобретения.

Наконец, в другом варианте изобретения возбуждение или сигнал нижнего диапазона ($u(n)$) подвергается передискретизации, например посредством линейной интерполяции или кубической "сплайн"-интерполяции из 12,8 в 16 кГц перед преобразованием (например, DCT-IV) длины 320. Этот вариант имеет недостаток, что является более сложным, так как преобразование (DCT-IV) возбуждения или сигнала вычисляется в дальнейшем по большей длине и передискретизация не выполняется в области преобразования.

Кроме того, в вариантах изобретения все вычисления, необходимые для оценки коэффициентов усиления (G_{HB1} , $g_{HB1}(m)$, $g_{HB2}(m)$, g_{HB3} , ...) могут выполняться в логарифмической области.

На фиг. 6 представлен примерный физический вариант осуществления устройства 600 расширения диапазона в соответствии с настоящим изобретением. Последнее может быть неотъемлемой частью декодера сигнала звуковой частоты или единицы оборудования, принимающей сигналы звуковой частоты, декодируемые или нет.

Данный тип устройства содержит процессор PROC, взаимодействующий с блоком памяти ВМ, содержащим запоминающее и/или оперативное запоминающее устройство MEM.

Такое устройство содержит модуль ввода E, способный принимать декодированный или извлеченный звуковой сигнал в первом диапазоне частот, называемом нижним диапазоном частот, восстановленным в частотной области ($U(k)$). Оно содержит модуль вывода S, способный передавать сигнал расширения во втором диапазоне частот ($U_{HB2}(k)$), например, к модулю 501 фильтрации, приведенному на фиг. 5.

Блок памяти может преимущественно содержать компьютерную программу, содержащую команды программного кода для выполнения этапов способа расширения диапазона в понимании настоящего изобретения, при выполнении этих команд процессором PROC, и, в частности, этапов извлечения (E402) тональных составляющих и сигнала окружения из сигнала, возникающего из декодированного сигнала нижнего диапазона ($U(k)$), объединения (E403) тональных составляющих ($y(k)$) и сигнала окружения ($U_{HBA}(k)$) посредством адаптивного микширования с использованием коэффициентов регулирования уровня энергии для получения звукового сигнала, называемого объединенным сигналом ($U_{HB2}(k)$), расширения (E401a) на по меньшей мере одном втором диапазоне частот, находящемся выше, чем первый диапазон частот, декодированного сигнала нижнего диапазона перед этапом извлечения или объединенного сигнала после этапа объединения.

Как правило, описание на фиг. 4 повторяет этапы алгоритма такой компьютерной программы. Компьютерная программа также может храниться в системе памяти, которая может считываться считывателем устройства или которая может загружаться в его пространство памяти.

Память MEM хранит, как правило, все данные, необходимые для реализации способа.

В одном возможном варианте осуществления устройство, описанное таким образом, может также содержать функции низкочастотного декодирования и другие функции обработки, описанные, например, на фиг. 5 и 3, в дополнение к функциям расширения диапазона в соответствии с изобретением.

(57) Формула изобретения

1. Способ расширения диапазона частот сигнала звуковой частоты во время процесса декодирования, включающий этапы:

получения сигнала, декодированного в первом диапазоне частот, называемом нижним диапазоном;

извлечения тональных составляющих и сигнала окружения из сигнала, возникающего из декодированного сигнала нижнего диапазона;

объединение тональных составляющих и сигнала окружения посредством адаптивного микширования с использованием коэффициентов регулирования уровня энергии для получения звукового сигнала, называемого объединенным сигналом; и

расширение декодированного сигнала нижнего диапазона по меньшей мере на по меньшей мере одном втором диапазоне частот, находящемся выше, чем первый диапазон частот перед этапом извлечения, формирующего расширенный декодированный сигнал нижнего диапазона $U_{HB1}(k)$,

причем извлечение содержит следующие операции:

(а) вычисление тональной энергии расширенного декодированного сигнала нижнего диапазона;

(b) вычисление сигнала окружения в абсолютном значении, соответствующем среднему значению уровня спектра спектральной линии за спектральной линией, и вычисление энергии преобладающих тональных составляющих декодированного в верхнем диапазоне спектра.

- 5 2. Способ по п. 1, в котором операция вычисления (а) тональной энергии расширенного декодированного сигнала нижнего диапазона включает в себя вычисление:

$$ener_{HB} = \sum_{k=240}^{319} U_{HB1}(k)^2 + \varepsilon, \text{ где } \varepsilon=0,1.$$

- 10 3. Способ по одному из пп. 1 или 2, отличающийся тем, что среднее значение уровня спектра спектральная линия за спектральной линией получают с использованием уравнения:

$$15 \quad lev(i) = \frac{1}{fn(i) - fb(i) + 1} \sum_{j=fb(i)}^{fn(i)} |U_{HB1}(j + 240)|,$$

$$fb(i) = 0 \text{ и } fn(i) = i + 7 \text{ для } i = 7, \dots, L-8$$

$$fb(i) = i - 7 \text{ и } fn(i) = L - 1 \text{ для } i = L - 7, L, L - 1.$$

- 20 4. Способ по одному из пп. 1, или 2, или 3, отличающийся тем, что вычисление энергии преобладающих тональных составляющих включает в себя вычисление остаточного сигнала:

$$y(i) = |U_{HB1}(i + 240)| - lev(i), \quad i = 0, \dots, L - 1$$

- 25 5. Способ по п.4, дополнительно содержащий этап обнаружения тональных составляющих на основе условия обнаружения остаточного сигнала.

6. Способ по п. 5, в котором условием обнаружения является $y(i) > 0$.

- 30 7. Способ по п. 6, в котором энергия преобладающих тональных частей определяется следующим уравнением:

$$ener_{total} = \sum_{i=0, \dots, L-1, y(i)>0} y(i)^2.$$

- 35 8. Устройство для расширения диапазона частот сигнала звуковой частоты, при этом сигнал декодирован в первом диапазоне частот, называемом нижним диапазоном частот, отличающееся тем, что содержит:

считываемую компьютером память, содержащую сохраненные на ней команды; процессор, выполненный с возможностью выполнения действий под управлением команд, содержащих:

- 40 получение сигнала, декодированного в первом диапазоне частот, называемом нижним диапазоном;

извлечение тональных составляющих и сигнала окружения из сигнала, возникающего из декодированного сигнала нижнего диапазона;

- 45 объединение тональных составляющих и сигнала окружения посредством адаптивного микширования с использованием коэффициентов регулирования уровня энергии для получения звукового сигнала, называемого объединенным сигналом; и

расширение декодированного сигнала нижнего диапазона по меньшей мере на по меньшей мере одном втором диапазоне частот, находящемся выше, чем первый диапазон частот перед этапом извлечения, формирующего расширенный декодированный сигнал

нижнего диапазона UHB1(k),

причем извлечение содержит следующие операции:

(а) вычисление тональной энергии расширенного декодированного сигнала нижнего диапазона;

5 (b) вычисление сигнала окружения в абсолютном значении, соответствующем среднему значению уровня спектра спектральная линия за спектральной линией, и вычисление энергии преобладающих тональных составляющих декодированного в верхнем диапазоне спектра.

10 9. Декодер сигнала звуковой частоты, отличающийся тем, что содержит устройство для расширения диапазона частот по п. 8.

15

20

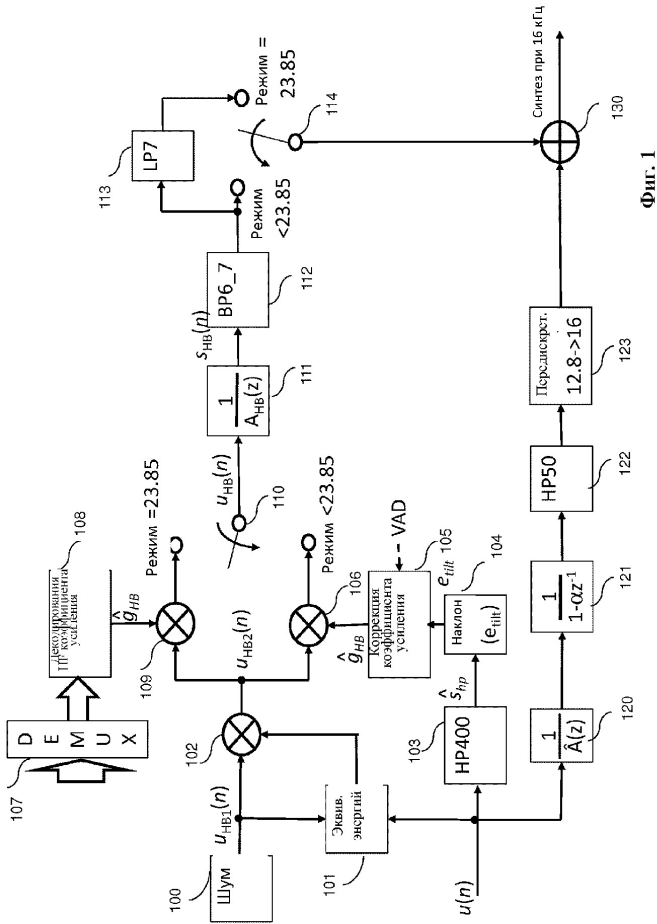
25

30

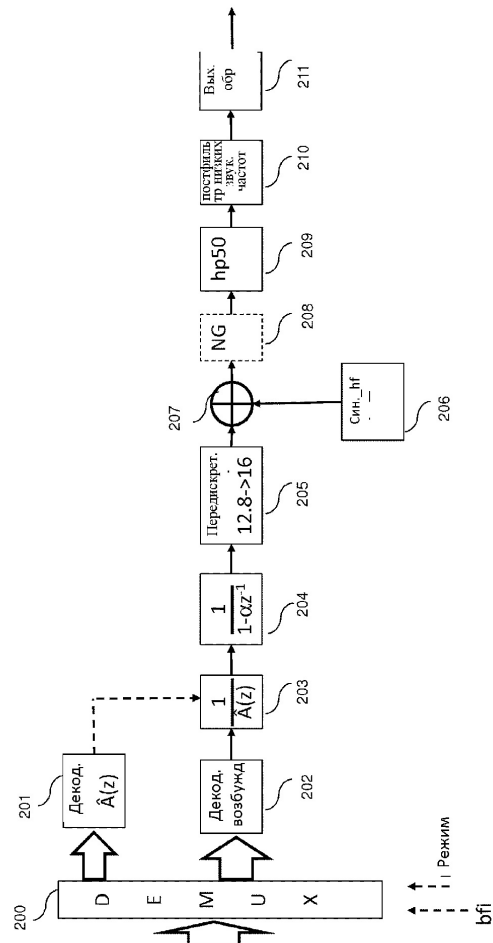
35

40

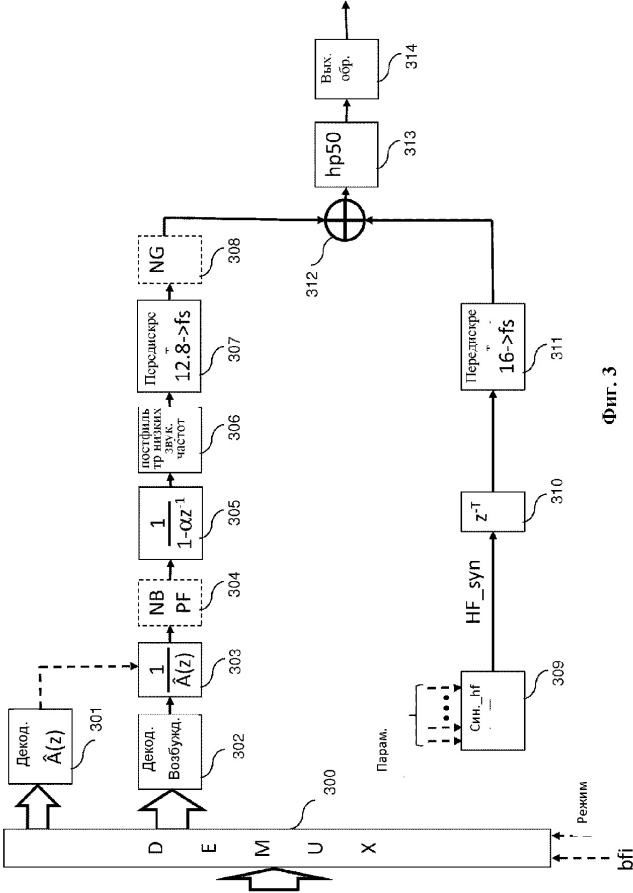
45



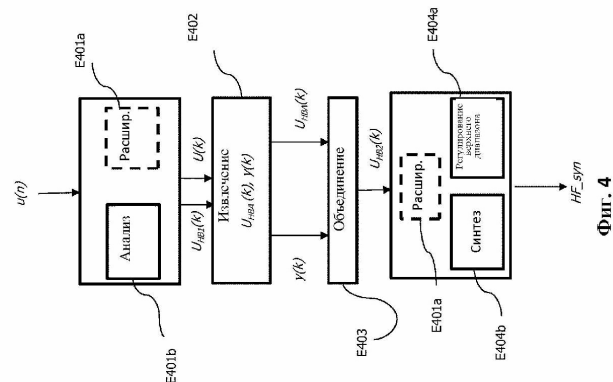
Фиг. 1

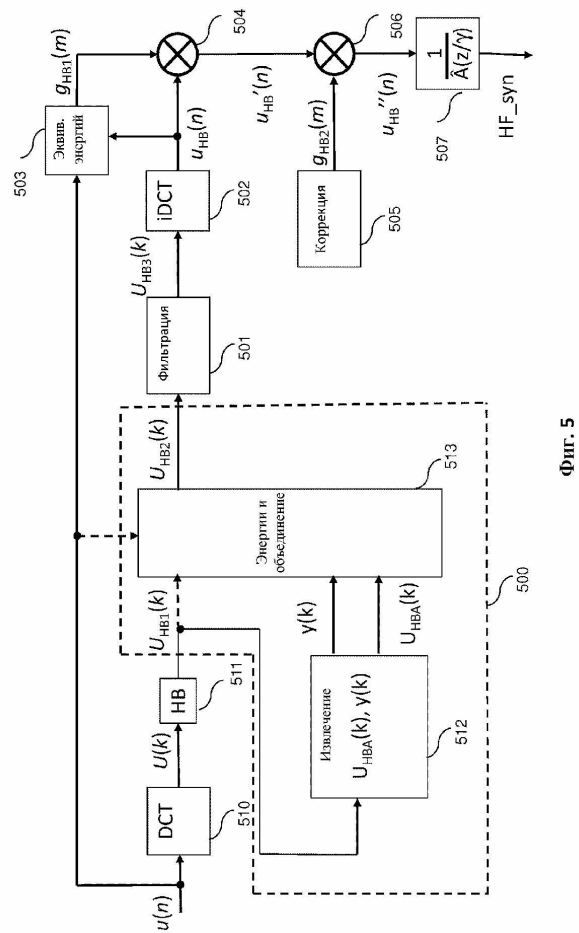


Фиг. 2

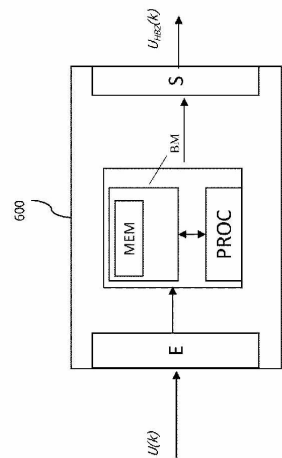


Фиг. 3





Фиг. 5



Фиг. 6