



УКРАЇНА

(19) **UA** (11) **129049** (13) **C2**
(51) МПК
G10L 21/0388 (2013.01)
G10L 19/02 (2013.01)
G10L 19/24 (2013.01)

НАЦІОНАЛЬНИЙ ОРГАН
ІНТЕЛЕКТУАЛЬНОЇ ВЛАСНОСТІ
ДЕРЖАВНА ОРГАНІЗАЦІЯ
"УКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ
ОФІС ІНТЕЛЕКТУАЛЬНОЇ
ВЛАСНОСТІ ТА ІННОВАЦІЙ"

(12) ОПИС ДО ПАТЕНТУ НА ВИНАХІД

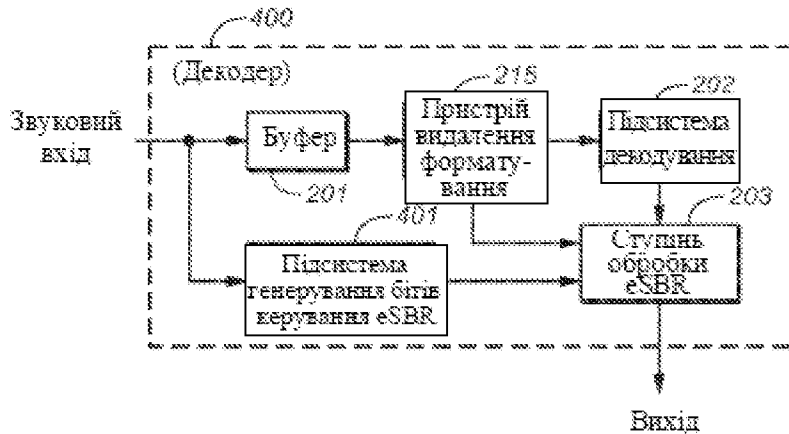
(21) Номер заявки:	a 2020 07393	(72) Винахідник(и):	Чоерлінг Крістофер (SE), Віллемоес Ларс (SE), Пурнхаген Хейко (SE), Екstrand Пер (SE)
(22) Дата подання заявки:	25.04.2019	(73) Володілець (володільці):	ДОЛБІ ІНТЕРНЕТНЛ АБ, Apollo Building, 3E, Herikerbergweg 1- 35, 1101 CN Amsterdam Zuidoost, the Netherlands (NL)
(24) Дата, з якої є чинними права інтелектуальної власності:	02.01.2025	(74) Представник:	Михайлюк Ганна Валентинівна, реєстр. №184
(31) Номер попередньої заявки відповідно до Паризької конвенції:	18169156.9	(56) Перелік документів, взятих до уваги експертизою:	US 2018025737 A1, 25.01.2018 ANONYMOUS: "ISO/IEC 23003- 3:201x/DIS of Unified Speech and Audio Coding", 95. MPEG MEETING; 24-1- 2011 - 28-1-2011; DAEGU; (MOTION PICTURE EXPERT GROUP OR ISO/IEC JTC1/SC29/WG11), no. N11863, 09.02.2011, XP030018356, section 7.5 figures 5, 16 GAYER MARC ET AL: "A Guideline to Audio Codec Delay", AES CONVENTION 116; MAY 2004, AES, 60 EAST 42ND STREET, ROOM 2520 NEW YORK 10165-2520, USA, 01.05.2004, XP040506870, sections 3.1.5, 3.1.6
(32) Дата подання попередньої заявки відповідно до Паризької конвенції:	25.04.2018		
(33) Код держави- учасниці Паризької конвенції, до якої подано попередню заявку:	EP		
(41) Публікація відомостей про заявку:	28.12.2020, Бюл.№ 24		
(46) Публікація відомостей про державну реєстрацію:	01.01.2025, Бюл.№ 1		
(86) Номер та дата подання міжнародної заявки, поданої відповідно до Договору РСТ	PCT/EP2019/060600, 25.04.2019		

(54) ІНТЕГРАЦІЯ МЕТОДИК РЕКОНСТРУКЦІЇ ВИСОКИХ ЧАСТОТ ЗВУКУ

UA 129049 C2

(57) Реферат:

Розкритий спосіб декодування кодованого звукового бітового потоку. Спосіб включає приймання кодованого звукового бітового потоку і декодування звукових даних для генерування декодованого звукового сигналу нижнього діапазону. Спосіб додатково включає вилучення метаданих реконструкції високих частот і фільтрацію декодованого звукового сигналу нижнього діапазону за допомогою блока фільтрів аналізу для генерування фільтрованого звукового сигналу нижнього діапазону. Спосіб також включає вилучення прапора, який вказує, чи необхідно виконувати або спектральний перенос, або гармонічну транспозицію на звукових даних, і відновлення частини, що стосується верхнього діапазону, звукового сигналу з використанням фільтрованого звукового сигналу нижнього діапазону і метаданих реконструкції високих частот згідно з прапором. Відновлення високих частот виконують як операцію постобробки із затримкою 3010 дискретних значень на звуковий канал.



Фіг. 5

Перехресне посилання на споріднені заявки

Дана заявка заявляє пріоритет європейської заявки на патент EP18169156.9, поданої 25 квітня 2018 р., яка шляхом посилання включена в даний опис.

Галузь техніки

5 Варіанти здійснення даного винаходу стосуються обробки звукових сигналів, і зокрема кодування, декодування або транскодування звукових бітових потоків із даними керування, які задають необхідність виконання на звукових даних або базової форми реконструкції високих частот ("HFR"), або покращеної форми HFR.

Передумови винаходу

10 Звичайний звуковий бітовий потік містить як звукові дані (наприклад, кодовані звукові дані), що характеризують один або більше каналів звукового вмісту, так і метадані, що вказують щонайменше одну характеристику звукових даних або звукового вмісту. Одним добре відомим форматом для генерування кодованого звукового бітового потоку є формат MPEG-4 "Перспективне звукове кодування" (AAC), описаний у стандарті MPEG ISO/IEC 14496-3:2009. У
15 стандарті MPEG-4 AAC означає "перспективне звукове кодування", а HE-AAC означає "високоєфективне перспективне звукове кодування".

Стандарт MPEG-4 AAC визначає декілька звукових профілів, які визначають, які об'єкти й інструменти кодування знаходяться у сумісному кодері або декодері. Три з цих звукових профілів являються собою (1) профіль AAC, (2) профіль HE-AAC і (3) профіль HE-AAC v2.
20 Профіль AAC містить тип об'єкта AAC низької складності (або "AAC-LC"). Об'єкт AAC-LC являє собою аналог профілю MPEG-2 AAC низької складності з деякими вдосконаленнями і не містить ні тип об'єкта копіювання спектральної смуги ("SBR"), ні тип об'єкта параметричного стерео ("PS"). Профіль HE-AAC являє собою надмножину профілю AAC і додатково містить тип об'єкта SBR. Профіль HE-AAC v2 являє собою надмножину профілю HE-AAC і додатково містить тип
25 об'єкта PS.

Тип об'єкта SBR містить інструмент копіювання спектральної смуги, який являє собою важливий інструмент кодування реконструкції високих частот ("HFR"), який значно підвищує ефективність стиснення кодеків сприйманих звукових сигналів. SBR реконструює
30 високочастотні складові звукового сигналу на стороні приймача (наприклад, у декодері). Таким чином, від кодера вимагається тільки кодувати і передавати низькочастотні складові, що робить можливою набагато більш високу якість звуку при низьких швидкостях передачі даних. SBR базується на копіюванні послідовностей гармонік, раніше усічених із метою зменшення швидкості передачі даних, із доступного сигналу з обмеженою смугою пропускання і даних керування, отриманих із кодера. Співвідношення між тональними і шумоподібними складовими зберігається за допомогою адаптивної зворотної фільтрації, а також необов'язкового додавання шуму і синусоїд. У стандарті MPEG-4 AAC інструмент SBR виконує спектральне вставлення (яке
35 також називається "лінійним переносом" або "спектральним переносом"), у разі якого ряд послідовних піддіапазонів квадратурного дзеркального фільтра (QMF) копіюють (або "вставляють") із переданої частини, що стосується нижнього діапазону, звукового сигналу в частину, що стосується верхнього діапазону, звукового сигналу, яка генерується в декодері.
40

Спектральне вставлення, або лінійний перенос, може не бути ідеальним для певних типів звуку, таких як музичний вміст із відносно низькими частотами розділення. Тому є необхідними методи для вдосконалення копіювання спектральної смуги.

Стислий опис варіантів здійснення винаходу

45 Перший клас варіантів здійснення стосується розкриття способу декодування кодованого звукового бітового потоку. Спосіб включає приймання кодованого звукового бітового потоку і декодування звукових даних для генерування декодованого звукового сигналу нижнього діапазону. Спосіб додатково включає вилучення метаданих реконструкції високих частот і фільтрацію декодованого звукового сигналу нижнього діапазону за допомогою блока фільтрів аналізу для генерування фільтрованого звукового сигналу нижнього діапазону. Спосіб
50 додатково включає вилучення прапора, який вказує, чи необхідно виконувати або спектральний перенос, або гармонічну транспозицію на звукових даних, і відновлення частини, що стосується верхнього діапазону, звукового сигналу з використанням фільтрованого звукового сигналу нижнього діапазону і метаданих реконструкції високих частот згідно з прапором. Зрештою, спосіб
55 включає об'єднання фільтрованого звукового сигналу нижнього діапазону і відновленої частини, що стосується верхнього діапазону, для формування широкосмугового звукового сигналу.

Другий клас варіантів здійснення стосується декодера звуку для декодування кодованого звукового бітового потоку. Декодер містить інтерфейс введення для приймання кодованого
60 звукового бітового потоку, де кодований звуковий бітовий потік містить звукові дані, які

представляють частину, що стосується нижнього діапазону, звукового сигналу, й основний декодер для декодування звукових даних для генерування декодованого звукового сигналу нижнього діапазону. Декодер також містить демультимплексор для вилучення з кодованого звукового бітового потоку метаданих реконструкції високих частот, де метадані реконструкції високих частот містять робочі параметри для процесу реконструкції високих частот, який виконує лінійний перенос послідовної кількості піддіапазонів із частини, що стосується нижнього діапазону, звукового сигналу в частину, що стосується верхнього діапазону, звукового сигналу, і блок фільтрів аналізу для фільтрації декодованого звукового сигналу нижнього діапазону для генерування фільтрованого звукового сигналу нижнього діапазону. Декодер додатково містить демультимплексор для вилучення з кодованого звукового бітового потоку прапора, який вказує, чи необхідно виконувати або лінійний перенос, або гармонічну транспозицію на звукових даних, і пристрій відновлення високих частот для відновлення частини, що стосується верхнього діапазону, звукового сигналу з використанням фільтрованого звукового сигналу нижнього діапазону і метаданих реконструкції високих частот згідно з прапором. Зрештою, декодер містить блок синтезувальних фільтрів для об'єднання фільтрованого звукового сигналу нижнього діапазону і відновленої частини, що стосується верхнього діапазону, для формування широкосмугового звукового сигналу.

Інші класи варіантів здійснення стосуються кодування і транскодування звукових бітових потоків, які містять метадані, що ідентифікують необхідність виконання обробки покращеного копіювання спектральної смуги (eSBR).

Стислий опис графічних матеріалів

На фіг. 1 показана структурна схема одного варіанта здійснення системи, яка може бути пристосована для виконання одного варіанта здійснення способу відповідно до винаходу.

На фіг. 2 показана структурна схема кодера, який являє собою один варіант здійснення блока обробки звуку відповідно до винаходу.

На фіг. 3 показана структурна схема системи, яка містить декодер, який являє собою один варіант здійснення блока обробки звуку відповідно до винаходу, і необов'язково також з'єднаний із ним постпроцесор.

На фіг. 4 показана структурна схема декодера, який являє собою один варіант здійснення блока обробки звуку відповідно до винаходу.

На фіг. 5 показана структурна схема декодера, який являє собою інший варіант здійснення блока обробки звуку відповідно до винаходу.

На фіг. 6 показана структурна схема іншого варіанта здійснення блока обробки звуку відповідно до винаходу.

На фіг. 7 показана схема блока даних бітового потоку MPEG-4 AAC, у тому числі сегментів, на які він розділений.

Умовні позначки й термінологія

Усюди в даному документі, у тому числі формулі винаходу, вираження виконання операції "відносно" сигналу або даних (наприклад, фільтрація, масштабування, перетворення або застосування коефіцієнта підсилення до сигналів або даних) використовується в широкому значенні для позначення виконання операції безпосередньо відносно сигналу або даних, або відносно обробленої версії сигналу або даних (наприклад, відносно версії сигналу, який був підданий попередній фільтрації або попередній обробці перед виконанням операції в його відношенні).

Усюди в даному документі, у тому числі формулі винаходу, вираження "блок обробки звуку" або "звуковий процесор" використовуються в широкому значенні, для позначення системи, пристрою або апарата, пристосованих для обробки звукових даних. Приклади блоків обробки звуку включають, але без обмеження, кодери, транскодери, декодери, кодеки, системи попередньої обробки, системи постобробки і системи обробки бітового потоку (які іноді називаються "інструментами обробки бітового потоку"). Практично вся споживча електроніка, така як мобільні телефони, телевізори, ноутбуки і планшетні комп'ютери, містить блок обробки звуку або звуковий процесор.

Усюди в даному документі, у тому числі формулі винаходу, термін "з'єднує", або "з'єднаний", використовується в широкому значенні для позначення або безпосереднього, або непрямого з'єднання. Таким чином, якщо перший пристрій з'єднаний із другим пристроєм, дане з'єднання може бути здійснене за допомогою безпосереднього з'єднання або за допомогою непрямого з'єднання через інші пристрої або з'єднання. Більше того, компоненти, які вбудовані в інші компоненти або об'єднані з ними, також пов'язані один з одним.

Докладний опис варіантів здійснення винаходу

Стандарт MPEG-4 AAC передбачає, що кодований бітовий потік MPEG-4 AAC містить

метадані, які вказують кожний тип обробки реконструкції високих частот ("HFR"), яку необхідно застосовувати (якщо який-небудь треба застосовувати) декодеру для декодування звукового вмісту бітового потоку, й/або які керують цією обробкою HFR, й/або вказують щонайменше одну характеристику або параметр щонайменше одного інструмента HFR, який необхідно застосовувати для декодування звукового вмісту бітового потоку. У даному документі вираження "метадані SBR" використовується для позначення метаданих, що стосуються типу, описаного або згаданого в стандарті MPEG-4 AAC для використання при копіюванні спектральної смуги ("SBR"). Спеціалістам у даній галузі техніки зрозуміло, що SBR являє собою форму HFR.

SBR переважно використовують як двошвидкісну систему, в якій базовий кодек діє на половині початкової частоти дискретизації, тоді як SBR діє на початковій частоті дискретизації. Кодер SBR діє паралельно з базовим основним кодеком, хоча і на більш високій частоті дискретизації. Хоча SBR являє собою, головним чином, прикінцеву обробку в декодері, для забезпечення найбільш точної реконструкції високих частот у декодері важливі параметри вилучаються в кодері. Кодер оцінює обвідну спектра діапазону SBR для часового і частотного діапазону/розділення, що підходить для характеристик сегментів поточного вхідного сигналу. Обвідна спектра оцінюється за допомогою комплексного QMF-аналізу і наступного обчислення енергії. Часове і частотне розділення обвідних спектра можна вибрати з більш високим ступенем свободи, щоб забезпечити найкраще підходяще частотно-часове розділення для даного вхідного сегмента. При оцінці обвідної необхідно враховувати, що перехідний сигнал, спершу розташований, головним чином, у високочастотній ділянці (наприклад, сигнал педальної тарілки), буде лише в невеликому ступені присутнім у генерованому SBR верхньому діапазоні до корекції обвідної, оскільки верхній діапазон у декодері базується на нижньому діапазоні, де перехідний сигнал набагато менше виражений у порівнянні з верхнім діапазоном. Даний аспект вводить різні вимоги для частотно-часового розділення даних обвідної спектра в порівнянні зі звичайною оцінкою обвідної спектра, яка використовується в алгоритмах кодування звуку.

Окрім обвідної спектра вилучається декілька додаткових параметрів, що представляють спектральні характеристики вхідного сигналу для різних часових і частотних ділянок. Оскільки кодер відповідним чином має доступ до початкового сигналу, а також до інформації про те, як блок SBR у декодері буде створювати верхній діапазон, за наявності конкретного набору параметрів керування, система може обробляти ситуації, в яких нижній діапазон становить строгий гармонічний ряд, а верхній діапазон, який необхідно відтворити, становить здебільшого випадкові складові сигналу, а також ситуації, в яких у початковому верхньому діапазоні є сильні тональні складові, які не мають аналогів у нижньому діапазоні, на якому базується верхній діапазон. Крім того, кодер SBR діє в щільному зв'язку з базовим основним кодеком для оцінки того, який частотний діапазон слід охоплювати за допомогою SBR у даний момент часу. Дані SBR ефективно кодуються перед передачею з використанням ентропійного кодування, а також каналних залежностей даних керування, у випадку стереофонічних сигналів.

Алгоритми вилучення параметрів керування зазвичай необхідно ретельно підлашувати відносно базового кодеку при заданій бітовій швидкості і заданій частоті дискретизації. Це викликано тим, що більш низька бітова швидкість зазвичай передбачає більш широкий діапазон SBR у порівнянні з високою бітовою швидкістю, а різні частоти дискретизації відповідають різним часовим розділенням кадрів SBR.

Декодер SBR зазвичай містить декілька різних частин. Він містить модуль декодування бітового потоку, модуль реконструкції високих частот (HFR), модуль додаткових високочастотних складових і модуль коректора обвідної. Система основана на блоці комплекснозначних QMF-фільтрів (для високоякісного SBR) або блоці дійснозначних QMF-фільтрів (для SBR малої потужності). Варіанти здійснення даного винаходу можуть застосовуватися і до високоякісного SBR, і до SBR малої потужності. У модулі вилучення бітового потоку дані керування зчитують із бітового потоку і декодують. Перед зчитуванням даних обвідної з бітового потоку для поточного кадру отримують частотно-часову сітку. Базовий основний декодер декодує звуковий сигнал поточного кадру (хоча і з більш низькою частотою дискретизації) для отримання звукових дискретних значень у часовій ділянці. Отримуваний кадр звукових даних використовується модулем HFR для реконструкції високих частот. Декодований сигнал нижнього діапазону потім піддають аналізу з використанням блока QMF-фільтрів. Потім на дискретних значеннях піддіапазонів блока QMF-фільтрів виконують реконструкцію високих частот і корекцію обвідної. Високі частоти реконструюють із нижнього діапазону гнучким чином на основі заданих параметрів керування. Крім того, для забезпечення належних спектральних характеристик заданої часової/частотної ділянки реконструйований верхній діапазон піддають адаптивній фільтрації на основі каналів піддіапазонів згідно з даними керування.

Вищий рівень бітового потоку MPEG-4 AAC являє собою послідовність блоків даних (елементи "raw_data_block"), кожний із яких являє собою сегмент даних (який у даному документі називається "блоком"), який містить звукові дані (зазвичай для проміжку часу в 1024 або 960 дискретних значень), і пов'язану інформацію й/або інші дані. У даному документі термін "блок" використовується для позначення сегмента бітового потоку MPEG-4 AAC, який містить звукові дані (і відповідні метадані, а також необов'язково інші пов'язані дані), які визначають або вказують один (але не більше одного) елемент "raw_data_block".

Кожний блок бітового потоку MPEG-4 AAC може містити деяку кількість синтаксичних елементів (кожний із яких також реалізований у бітовому потоці як сегмент даних). У стандарті MPEG-4 AAC визначено сім типів таких синтаксичних елементів. Кожний синтаксичний елемент ідентифікують за допомогою відрізного значення елемента даних "id_syn_ele". Приклади синтаксичних елементів включають "single_channel_element()", "channel_pair_element()" і "fill_element()». Елемент одиничного каналу single_channel_element являє собою контейнер, який містить звукові дані одного звукового каналу (монофонічний звуковий сигнал). Елемент пари каналів channel_pair_element містить звукові дані двох звукових каналів (тобто стереофонічний звуковий сигнал).

Заповнювальний елемент fill_element являє собою контейнер інформації, який містить ідентифікатор (наприклад, значення відзначеного вище елемента "id_syn_ele"), за яким йдуть дані, що називаються "заповнювальними даними". Заповнювальні елементи історично використовували для корекції миттєвої бітової швидкості бітових потоків, які підлягають передачі по каналу з постійною швидкістю. Завдяки додаванню відповідної кількості заповнювальних даних до кожного блока можна досягнути постійної швидкості передачі даних.

Згідно з варіантами здійснення даного винаходу заповнювальні дані можуть містити один або більше елементів корисних даних розширення, які розширяють тип даних (наприклад, метадані) і можуть бути передані в бітовому потоці. Декодер, який приймає бітові потоки із заповнювальними даними, що містять новий тип даних, може необов'язково бути використаний пристроєм, який приймає бітовий потік (наприклад, декодером), для розширення функціональних можливостей пристрою. Таким чином, як може бути зрозуміло спеціалісту в даній галузі техніки, заповнювальні елементи являють собою спеціальний тип структури даних і відрізняються від структур даних, які зазвичай використовуються для передачі звукових даних (наприклад, корисних звукових даних, які містять дані каналів).

У деяких варіантах здійснення даного винаходу ідентифікатор, який використовується для ідентифікації заповнювального елемента, може складатися з трибітного цілого числа без знака, в якого спочатку передають старший значущий біт ("uimsbf"), що має значення 0×6 . В одному блоці може існувати декілька екземплярів синтаксичного елемента одного типу (наприклад, декілька заповнювальних елементів).

Іншим стандартом кодування звукових бітових потоків є стандарт MPEG "Уніфіковане кодування мови і звуку" (USAC) (ISO/IEC 23003-3:2012). У стандарті MPEG USAC описано кодування і декодування звукового вмісту з використанням обробки копіювання спектральної смуги (у тому числі обробки SBR, як описано в стандарті MPEG-4 AAC, а також інших покращених форм обробки копіювання спектральної смуги). У випадку даної обробки застосовують інструменти копіювання спектральної смуги (які іноді називаються в даному документі "інструментами покращеного SBR" або "інструментами eSBR") із розширеної і покращеної версії набору інструментів SBR, описаних у стандарті MPEG-4 AAC. Таким чином, eSBR (визначене в стандарті USAC) являє собою вдосконалення SBR (визначеного в стандарті MPEG-4 AAC).

У даному документі вираження "обробка покращеного SBR" (або "обробка eSBR") використовується для позначення обробки копіювання спектральної смуги з використанням щонайменше одного інструмента eSBR (наприклад, щонайменше одного інструмента eSBR, описаного або згаданого в стандарті MPEG USAC), який не описаний або не згаданий у стандарті MPEG-4 AAC. Прикладами таких інструментів eSBR є гармонічна транспозиція і попередня обробка, або "попереднє згладжування", QMF-вставки.

Гармонічний транспозер цілочислового порядку T відображає синусоїду з частотою ω в синусоїду з частотою $T\omega$ з одночасним збереженням тривалості сигналу. Зазвичай для отримання кожної частини потрібного вихідного діапазону частот використовується в послідовності три порядки, $T=2, 3, 4$, з використанням найменшого можливого порядку транспозиції. Якщо потрібне виведення вище четвертого порядку діапазону транспозиції, його можна отримати шляхом зсувів частоти. З метою доведення до мінімуму обчислювальної складності, коли можливо, для обробки створюють дискретизовані часові ділянки основної смуги, близькі до критичних.

Гармонічний транспозер може базуватися або на QMF, або на DFT. При використанні гармонічного транспозера на основі QMF розширення смуги пропускання сигналу основного кодера в часовій ділянці повністю здійснюється в QMF-ділянці із використанням конструкції модифікованого фазового вокодера, яка виконує проріджування, за яким іде розтягнення за часом для кожного QMF-піддіапазону. Транспозицію із використанням декількох коефіцієнтів транспозиції (наприклад, $T=2, 3, 4$) здійснюють на загальному ступені перетворення аналізу/синтезу QMF. Оскільки гармонічний транспозер на основі QMF не володіє адаптивною до сигналу передискретизацією в частотній ділянці, відповідний прапор у бітовому потоці ("sbrOversamplingFlag[ch]») може бути проігнорований.

Для зниження складності при використанні гармонічного транспозера на основі DFT транспозери з коефіцієнтами 3 і 4 (транспозери 3-го і 4-го порядків) переважно вбудовані в транспозер із коефіцієнтом 2 (транспозер 2-го порядку) за допомогою інтерполяції. Для кожного кадру (який відповідає дискретним значенням основного кодера coreCoderFrameLength) номінальний "повний розмір" перетворення транспозера в першу чергу визначається в бітовому потоці прапором ("sbrOversamplingFlag[ch]») адаптивною до сигналу передискретизації в частотній ділянці.

Якщо sbrPatchingMode==1, із вказівкою, що для генерування верхнього діапазону необхідно використовувати лінійну транспозицію, то може бути введений додатковий етап, щоб уникнути порушень безперервності в формі обвідної спектра високочастотного сигналу, який вводиться в наступний коректор обвідної. Це покращує роботу наступного ступеня корекції обвідної, що призводить до більшої стабільності сприйманого сигналу верхнього діапазону. Операція додаткової попередньої обробки є корисною для типів сигналів, у яких груба обвідна спектра сигналу нижнього діапазону, що використовується для реконструкції високих частот, проявляє великі варіації рівня. Однак значення елемента бітового потоку можна визначити в кодері шляхом застосування будь-якого типу класифікації, яка залежить від сигналу. Додаткову попередню обробку переважно активують за допомогою однобітного елемента бітового потоку, "bs_sbr_preprocessing". Якщо "bs_sbr_preprocessing" встановлений на одиницю, додаткова обробка включена. Якщо "bs_sbr_preprocessing" встановлена на нуль, додаткова попередня обробка відключена. У додатковій обробці переважно використовується крива попереднього посилення, яка використовується генератором високих частот для масштабування нижнього діапазону, X_{Low} , для кожної вставки. Наприклад, криву попереднього посилення можна обчислити за формулою:

$$preGain(k) = 10^{(\text{meanNrg} - \text{lowEnvSlope}(k))/20}, 0 \leq k < k_0$$

де k_0 – перший QMF-піддіапазон у таблиці діапазонів головних частот, а lowEnvSlope обчислюють із використанням функції, яка обчислює коефіцієнти найкращого підходящого полінома (у значенні найменших квадратів), такої як polyfit(). Наприклад,

$$\text{polyfit}(3, k_0, x_lowband, \text{lowEnv}, \text{lowEnvSlope}),$$

може бути використана (з використанням полінома третього ступеня) і де

$$\text{lowEnv}(k) = 10 \log_{10} \frac{\varphi_k(0,0)}{\text{numTimeSlots} \cdot \text{RATE} + 6}, 0 \leq k < k_0$$

де $x_lowband(k) = [0 \dots k_0 - 1]$, numTimeSlot – кількість часових проміжків обвідної SBR, які існують у рамках кадру, RATE – константа, яка вказує кількість дискретних значень QMF-піддіапазонів на часовий проміжок (наприклад, 2), φ_k – коефіцієнт фільтра лінійного передбачення (потенційно отримуваний за допомогою коваріаційного методу), і де

$$\text{meanNrg} = \frac{\sum_{k=0}^{k_0-1} \text{lowEnv}(k)}{k_0}$$

Бітовий потік, згенерований відповідно до стандарту MPEG USAC (який іноді називається в даному документі "бітовим потоком USAC"), містить кодований звуковий вміст і зазвичай містить метадані, які вказують кожний тип обробки копіювання спектральної смуги, яку необхідно застосовувати декодеру для декодування звукового вмісту бітового потоку USAC, й/або метадані, які керують такою обробкою копіювання спектральної смуги, й/або які вказують щонайменше одну характеристику або параметр щонайменше одного інструмента SBR й/або eSBR, який необхідно застосовувати для декодування звукового вмісту бітового потоку USAC.

У даному документі вираження "метадані покращеного SBR" (або "метадані eSBR") використовуються для позначення метаданих, які вказують кожний тип обробки копіювання спектральної смуги, яку необхідно застосовувати декодеру для декодування звукового вмісту кодованого звукового бітового потоку (наприклад, бітового потоку USAC), й/або які керують такою обробкою копіювання спектральної смуги, й/або які вказують щонайменше одну

характеристику або параметр щонайменше одного інструмента SBR й/або eSBR, який необхідно застосовувати для декодування такого звукового вмісту, але який не описаний або не згаданий у стандарті MPEG-4 AAC. Прикладом метаданих eSBR є метадані (які вказують обробку копіювання спектральної смуги або які керують нею), які описані або згадані в стандарті MPEG USAC, але не в стандарті MPEG-4 AAC. Таким чином, метадані eSBR у даному документі означають метадані, які не є метаданими SBR, і метадані SBR у даному документі означають метадані, які не є метаданими eSBR.

Бітовий потік USAC може містити і метадані SBR, і метадані eSBR. Більш конкретно, бітовий потік USAC може містити метадані eSBR, які керують виконанням обробки eSBR декодером, і метадані SBR, які керують виконанням обробки SBR декодером. Згідно зі звичайними варіантами здійснення даного винаходу метадані eSBR (наприклад, конфігураційні дані, специфічні для eSBR) включені (відповідно до даного винаходу) в бітовий потік MPEG-4 AAC (наприклад, у контейнер `sbr_extension()` у кінці корисних даних SBR).

Під час виконання обробки eSBR, в ході декодування кодованого бітового потоку з використанням набору інструментів eSBR (який містить щонайменше один інструмент eSBR), високочастотний діапазон звукового сигналу відновлюють декодером на основі копіювання послідовностей гармонік, які були усічені під час кодування. Така обробка eSBR зазвичай коректує обвідну спектра згенерованого високочастотного діапазону, а також застосовує зворотну фільтрацію і додає шумові і синусоїдні складові для відтворення спектральних характеристик початкового звукового сигналу.

Згідно зі звичайними варіантами здійснення даного винаходу метадані eSBR включають (наприклад, включають невелику кількість бітів керування, які являють собою метадані eSBR) в один або більше сегментів метаданих кодованого звукового бітового потоку (наприклад, бітового потоку MPEG-4 AAC), який також містить кодовані звукові дані в інших сегментах (сегментах звукових даних). Зазвичай щонайменше один такий сегмент метаданих кожного блока бітового потоку являє собою (або містить) заповнювальний елемент (який містить ідентифікатор, який вказує початок заповнювального елемента), і метадані eSBR включають у заповнювальний елемент після цього ідентифікатора.

На фіг. 1 показана структурна схема наведеного як приклад ланцюжка обробки звуку (системи обробки звукових даних), у якому один або більше елементів системи можуть бути виконані згідно з одним варіантом здійснення даного винаходу. Система містить наступні елементи, з'єднані разом, як показано: кодер 1, підсистему 2 доставки, декодер 3 і блок 4 постобробки. У варіаціях показаної системи опускають один або більше елементів або включають додаткові блоки обробки звукових даних.

У деяких реалізаціях кодер 1 (який необов'язково містить блок попередньої обробки) виконаний із можливістю приймання як введення дискретних значень PCM (у часовій ділянці), які містять звуковий вміст, і виведення кодованого звукового бітового потоку (з форматом, сумісним зі стандартом MPEG-4 AAC), який характеризує звуковий вміст. Дані бітового потоку, які характеризують звуковий вміст, іноді називають "звуковими даними" або "кодованими звуковими даними". Якщо кодер виконаний згідно зі звичайним варіантом здійснення даного винаходу, звуковий бітовий потік, який виводиться з кодера, містить метадані eSBR (і зазвичай інші метадані), а також звукові дані.

Один або більше кодованих звукових бітових потоків, які виводяться з кодера 1, можуть бути передані в підсистему 2 доставки кодованого звуку. Підсистема 2 виконана з можливістю зберігання й/або доставки кожного кодованого бітового потоку, який виводиться з кодера 1. Кодований звуковий бітовий потік, який виводиться з кодера 1, може бути збережений у підсистемі 2 (наприклад, у формі диска DVD або Blu ray), або переданий підсистемою 2 (яка може реалізовувати канал або мережу зв'язку), або може бути і збережений, і переданий підсистемою 2.

Декодер 3 виконаний із можливістю декодування кодованого звукового бітового потоку MPEG-4 AAC (який генерується кодером 1), який він приймає через підсистему 2. У деяких варіантах здійснення декодер 3 виконаний із можливістю вилучення метаданих eSBR із кожного блока бітового потоку і декодування бітового потоку (у тому числі за допомогою виконання обробки eSBR з використанням вилучених метаданих eSBR) для генерування декодованих звукових даних (наприклад, потоків декодованих звукових дискретних значень PCM). У деяких варіантах здійснення декодер 3 виконаний із можливістю вилучення метаданих SBR із бітового потоку (однак з ігноруванням метаданих eSBR, включених у бітовий потік) і декодування бітового потоку (у тому числі за допомогою виконання обробки SBR із використанням вилучених метаданих SBR) для генерування декодованих звукових даних (наприклад, потоків декодованих звукових дискретних значень PCM). Зазвичай декодер 3 містить буфер, у якому зберігаються

(наприклад, енергонезалежно) сегменти кодованого звукового бітового потоку, прийнятого з підсистеми 2.

Блок 4 постобробки за фіг. 1 виконаний із можливістю приймання потоку декодованих звукових даних із декодера 3 (наприклад, декодованих звукових дискретних значень PCM) і виконання постобробки відносно них. Блок постобробки також може бути виконаний із
5
10
15

можливістю представлення підданого постобробці звукового вмісту (або декодованого звуку, прийнятого з декодера 3) для відтворення одним або більше динаміками.
На фіг. 2 показана структурна схема кодера (100), який являє собою один варіант здійснення блока обробки звуку відповідно до винаходу. Будь-який із компонентів або елементів кодера 100 може бути реалізований як один або більше процесів й/або одна або більше схем (наприклад, ASIC, FPGA або інших інтегральних схем), в апаратному забезпеченні, програмному забезпеченні або в поєднанні апаратного і програмного забезпечення. Кодер 100 містить кодер 105, ступінь 107 формувача швидкості передачі даних/пристрою форматування, ступінь 106 генерування метаданих і буферний запам'ятовувальний пристрій 109, з'єднані таким чином, як показано. Зазвичай кодер 100 також містить інші обробляючі елементи (не показані). Кодер 100 виконаний із можливістю перетворення вхідного звукового бітового потоку в кодований вихідний бітовий потік MPEG-4 AAC.

Генератор 106 метаданих з'єднаний і виконаний із можливістю генерування (й/або пропускання на ступінь 107) метаданих (у тому числі метаданих eSBR і метаданих SBR) із метою включення ступеня 107 в кодований бітовий потік для виведення з кодера 100.
20

Кодер 105 з'єднаний і виконаний із можливістю кодування вхідних звукових даних (наприклад, шляхом виконання їхнього стиснення) і передачі отриманого в результаті кодованого звуку на ступінь 107 із метою включення в кодований бітовий потік для виведення зі ступеня 107.
25

Ступінь 107 виконаний із можливістю мультиплексування кодованого звуку із кодера 105 і метаданих (у тому числі метаданих eSBR і метаданих SBR) із генератора 106 для генерування кодованого бітового потоку для виведення зі ступеня 107, переважно таким чином, що кодований бітовий потік має формат, який задається одним із варіантів здійснення даного винаходу.
30

Буферний запам'ятовувальний пристрій 109 виконаний із можливістю зберігання (наприклад, енергонезалежно) щонайменше одного блока кодованого звукового бітового потоку зі ступеня 107, і послідовність блоків кодованого звукового бітового потоку потім передається з буферного запам'ятовувального пристрою 109 як виведення з кодера 100 в систему доставки.
35

На фіг. 3 показана структурна схема системи, яка містить декодер (200), що являє собою один варіант здійснення блока обробки звуку відповідно до винаходу, і необов'язково також з'єднаний із ним постпроцесор (300). Будь-який із компонентів, або елементів, декодера 200 і постпроцесора 300 може бути реалізований як один або більше процесів й/або одна або більше схем (наприклад, ASIC, FPGA або інших інтегральних схем), в апаратному забезпеченні, програмному забезпеченні або в поєднанні апаратного і програмного забезпечення. Декодер 200 містить буферний запам'ятовувальний пристрій 201, пристрій 205 видалення форматування (синтаксичний аналізатор) корисних даних бітового потоку, підсистему 202 декодування звуку (яка іноді називається "основним" ступенем декодування або "основною" підсистемою декодування), ступінь 203 обробки eSBR і ступінь 204 генерування бітів керування, з'єднані таким чином, як показано. Зазвичай декодер 200 також містить інші обробляючі елементи (не показані).
40
45

Буферний запам'ятовувальний пристрій (буфер) 201 зберігає (наприклад, енергонезалежно) щонайменше один блок кодованого звукового бітового потоку MPEG-4 AAC, прийнятого декодером 200. У роботі декодера 200 послідовність блоків бітового потоку передається з буфера 201 в пристрій 205 видалення форматування.
50

У варіаціях варіанта здійснення за фіг. 3 (або варіанта здійснення за фіг. 4, який буде описаний), АРУ, який не є декодером (наприклад, АРУ 500 за фіг. 6), містить буферний запам'ятовувальний пристрій (наприклад, буферний запам'ятовувальний пристрій, ідентичний із буфером 201), який зберігає (наприклад, енергонезалежно) щонайменше один блок кодованого звукового бітового потоку (наприклад, звукового бітового потоку MPEG-4 AAC) такого ж типу, як прийнятий буфером 201 за фіг. 3 або фіг. 4 (тобто кодованого звукового бітового потоку, який містить метадані eSBR).
55

Також із посиланням на фіг. 3 пристрій 205 видалення форматування з'єднаний і виконаний із можливістю демультимплексування кожного блока бітового потоку для вилучення з нього метаданих SBR (у тому числі квантованих даних обвідної) і метаданих eSBR (і зазвичай інших метаданих) із метою передачі щонайменше метаданих eSBR і метаданих SBR на ступінь 203
60

обробки eSBR, а також зазвичай для передачі інших вилучених метаданих у підсистему 202 декодування (і необов'язково також у генератор 204 бітів керування). Пристрій 205 видалення форматування також з'єднаний і виконаний із можливістю вилучення звукових даних із кожного блока бітового потоку і передачі вилучених звукових даних у підсистему 202 декодування

5

(ступінь декодування). Система, показана на фіг. 3, необов'язково також містить постпроцесор 300. Постпроцесор 300 містить буферний запам'ятовувальний пристрій (буфер) 301 й інші обробляючі елементи (не показані), у тому числі щонайменше один обробляючий елемент, з'єднаний із буфером 301. Буфер 301 зберігає (наприклад, енергонезалежно) щонайменше один блок (або кадр) декодованих звукових даних, прийнятих постпроцесором 300 із декодера 200. Обробляючі елементи постпроцесора 300 з'єднані і виконані з можливістю приймання й адаптивної обробки послідовності блоків (або кадрів) виведення декодованого звуку з буфера 301 із використанням виведення метаданих із підсистеми 202 декодування (й/або пристрою 205 видалення форматування) й/або виведення бітів керування зі ступеня 204 декодера 200.

10

15

Підсистема 202 декодування звуку декодера 200 виконана з можливістю декодування звукових даних, вилучених синтаксичним аналізатором 205 (таке декодування можна назвати "основною" операцією декодування), для генерування декодованих звукових даних і передачі декодованих звукових даних на ступінь 203 обробки eSBR. Декодування виконується в частотній ділянці і зазвичай включає зворотне квантування, за яким йде спектральна обробка. Зазвичай останній ступінь обробки в підсистемі 202 застосовує перетворення із частотної ділянки в часову ділянку до декодованих звукових даних у частотній ділянці, таким чином виведення підсистеми являє собою декодовані звукові дані в часовій ділянці. Ступінь 203 виконаний із можливістю застосування інструментів SBR й інструментів eSBR, вказаних метаданими SBR і eSBR (вилученими синтаксичним аналізатором 205), до декодованих звукових даних (тобто виконання обробки SBR і eSBR відносно виведення підсистеми 202 декодування з використанням метаданих SBR і eSBR) для генерування повністю декодованих звукових даних, які виводяться (наприклад, у постпроцесор 300) із декодера 200. Зазвичай декодер 200 містить запам'ятовувальний пристрій (доступний для підсистеми 202 і ступеня 203), який зберігає звукові дані з видаленим форматуванням і метадані, виведені з пристрою 205 видалення форматування, і ступінь 203 виконаний із можливістю здійснення доступу до цих звукових даних і метаданих (у тому числі метаданих SBR і eSBR) в міру необхідності в ході обробки SBR і eSBR. Обробку SBR і обробку eSBR на ступені 203 можна вважати постобробкою відносно виведення основної підсистеми 202 декодування. Необов'язково декодер 200 також містить підсистему прикінцевого підвищувального міксування (яка може застосовувати інструменти параметричної стереофонії ("PS")), визначені в стандарті MPEG-4 AAC, із використанням метаданих PS, вилучених пристроєм 205 видалення форматування, й/або бітів керування, згенерованих у підсистемі 204), з'єднану і виконану з можливістю виконання підвищувального міксування відносно виведення ступеня 203 для генерування повністю декодованого і підданого підвищувальному міксуванню звуку, який виводиться з декодера 200. Альтернативно постпроцесор 300 виконаний із можливістю виконання підвищувального міксування відносно виведення декодера 200 (наприклад, із використанням метаданих PS, вилучених пристроєм 205 видалення форматування, й/або бітів керування, згенерованих у підсистемі 204).

20

25

30

35

40

45

50

55

У відповідь на метадані, вилучені пристроєм 205 видалення форматування, генератор 204 бітів керування може генерувати дані керування, і дані керування можуть бути використані в декодері 200 (наприклад, у підсистемі прикінцевого підвищувального міксування) й/або передані як виведення декодера 200 (наприклад, у постпроцесор 300 для використання при постобробці). У відповідь на метадані, вилучені із вхідного бітового потоку (і необов'язково також у відповідь на дані керування), ступінь 204 може генерувати (і передавати в постпроцесор 300) біти керування, які вказують, що декодовані звукові дані, які виводяться зі ступеня 203 обробки eSBR, слід піддати постобробці конкретного типу. У деяких реалізаціях декодер 200 виконаний із можливістю передачі метаданих, вилучених пристроєм 205 видалення форматування із вхідного бітового потоку, в постпроцесор 300, і постпроцесор 300 виконаний із можливістю виконання постобробки відносно декодованих звукових даних, які виводяться з декодера 200, з використанням метаданих.

60

На фіг. 4 показана структурна схема блока обробки звуку ("APU") (210), який являє собою інший варіант здійснення блока обробки звуку відповідно до винаходу. APU 210 являє собою традиційний декодер, виконаний без можливості виконання обробки eSBR. Будь-який із компонентів або елементів APU 210 може бути реалізований як один або більше процесів й/або одна або більше схем (наприклад, ASIC, FPGA або інших інтегральних схем), в апаратному

забезпеченні, програмному забезпеченні або в поєднанні апаратного і програмного забезпечення. APU 210 містить буферний запам'ятовувальний пристрій 201, пристрій 215 видалення форматування (синтаксичний аналізатор) корисних даних бітового потоку, підсистему 202 декодування звуку (яка іноді називається "основним" ступенем декодування або
5 "основною" підсистемою декодування) і ступінь 213 обробки SBR, з'єднані таким чином, як показано. Зазвичай APU 210 також містить інші обробляючі елементи (не показані). APU 210 може представляти, наприклад, кодер, декодер або транскодер звуку.

Елементи 201 і 202 APU 210 є ідентичними з ідентично пронумерованими елементами декодера 200 (за фіг. 3), і їхній опис, наведений вище, повторюватися не буде. У роботі APU 210
10 послідовність блоків кодованого звукового бітового потоку (бітового потоку MPEG-4 AAC), прийнята APU 210, передається з буфера 201 в пристрій 215 видалення форматування.

Пристрій 215 видалення форматування з'єднаний і виконаний із можливістю демультимплексування кожного блока бітового потоку з метою вилучення з нього метаданих SBR (у тому числі квантованих даних об'ємної) і зазвичай інших метаданих, але з ігноруванням
15 метаданих eSBR, які можуть бути включені в бітовий потік відповідно до будь-якого з варіантів здійснення даного винаходу. Пристрій 215 видалення форматування виконаний із можливістю передачі щонайменше метаданих SBR на ступінь 213 обробки SBR. Пристрій 215 видалення форматування також з'єднаний і виконаний із можливістю вилучення звукових даних із кожного блока бітового потоку і передачі вилучених звукових даних у підсистему 202 декодування
20 (ступінь декодування).

Підсистема 202 декодування звуку декодера 200 виконана з можливістю декодування звукових даних, вилучених пристроєм 215 видалення форматування (таке декодування можна назвати "основною" операцією декодування), для генерування декодованих звукових даних і
25 передачі декодованих звукових даних на ступінь 213 обробки SBR. Декодування виконується в частотній ділянці. Зазвичай останній ступінь обробки в підсистемі 202 застосовує перетворення із частотної ділянки в часову ділянку до декодованих звукових даних у частотній ділянці, таким чином виведення підсистеми являє собою декодовані звукові дані в часовій ділянці. Ступінь 213 виконана з можливістю застосування інструментів SBR (але не інструментів eSBR), вказаних метаданими SBR (вилученими пристроєм 215 видалення форматування), до декодованих
30 звукових даних (тобто виконання обробки SBR відносно виведення підсистеми 202 декодування з використанням метаданих SBR) для генерування повністю декодованих звукових даних, які виводяться (наприклад, у постпроцесор 300) із APU 210. Зазвичай APU 210 містить запам'ятовувальний пристрій (доступний для підсистеми 202 і ступеня 213), який зберігає звукові дані з видаленим форматуванням і метадані, виведені з пристрою 215 видалення
35 форматування, і ступінь 213 виконаний із можливістю здійснення доступу до звукових даних і метаданих (у тому числі метаданих SBR) в міру необхідності в ході обробки SBR. Обробку SBR на ступені 213 можна вважати постобробкою відносно виведення основної підсистеми 202 декодування. Необов'язково APU 210 також містить підсистему прикінцевого підвищувального мікшування (яка може застосовувати інструменти параметричної стереофонії ("PS"), визначені в
40 стандарті MPEG-4 AAC, із використанням метаданих PS, вилучених пристроєм 215 видалення форматування), з'єднану і виконану з можливістю виконання підвищувального мікшування відносно виведення ступеня 213 для генерування повністю декодованого і підданого підвищувальному мікшуванню звуку, який виводиться з APU 210. Альтернативно постпроцесор виконаний із можливістю виконання підвищувального мікшування відносно виведення APU 210
45 (наприклад, із використанням метаданих PS, вилучених пристроєм 215 видалення форматування, й/або бітів керування, згенерованих в APU 210).

Різні реалізації кодера 100, декодера 200 і APU 210 виконані з можливістю виконання різних варіантів здійснення способу відповідно до винаходу.

Згідно з деякими варіантами здійснення метадані eSBR включені (наприклад, включена невелика кількість бітів керування, які являють собою метадані eSBR) у кодований звуковий бітовий потік (наприклад, у бітовий потік MPEG-4 AAC), таким чином, що традиційні декодери (які виконані без можливості здійснення синтаксичного аналізу метаданих eSBR або
50 використання будь-якого інструмента eSBR, якого стосуються метадані eSBR) можуть ігнорувати метадані eSBR, але все таки декодувати бітовий потік у тій мірі, яка є можливою без використання метаданих eSBR або будь-якого інструмента eSBR, якого стосуються метадані eSBR, зазвичай без яких-небудь значущих втрат у якості декодованого звуку. Однак декодери eSBR, виконані з можливістю синтаксичного аналізу бітового потоку, з метою ідентифікації метаданих eSBR, і використання щонайменше одного інструмента eSBR у відповідь на метадані eSBR, будуть отримувати вигоду від використання щонайменше одного такого інструмента
60 eSBR. Тому варіанти здійснення даного винаходу надають засоби для ефективною передачі

даних керування або метаданих покращеного копіювання спектральної смуги (eSBR) із забезпеченням зворотної сумісності.

Зазвичай метадані eSBR у бітовому потоці вказують (наприклад, вказують щонайменше одну їхню характеристику або параметр) один або більше з наступних інструментів eSBR (які описані в стандарті MPEG USAC і які могли бути або могли не бути застосовані кодером під час генерування бітового потоку):

- гармонічна транспозиція; і
- додаткова попередня обробка (попереднє згладжування) QMF вставки.

Наприклад, метадані eSBR, включені в бітовий потік, можуть вказувати значення параметрів (описаних у стандарті MPEG USAC і даному винаході): `sbrPatchingMode[ch]`, `sbrOversamplingFlag[ch]`, `sbrPitchInBins[ch]` і `bs_sbr_preprocessing`.

У даному документі позначення "X[ch]", де X – деякий параметр, означає, що параметр стосується каналу ("ch") звукового вмісту кодованого бітового потоку, який підлягає декодуванню. Для простоти вираження [ch] іноді опускається і передбачається, що відповідний параметр стосується каналу звукового вмісту.

У даному документі позначення X[ch][env], де X – деякий параметр, означає, що параметр стосується обвідної ("env") SBR каналу ("ch") звукового вмісту кодованого бітового потоку, який підлягає декодуванню. Для простоти вираження [env] і [ch] іноді опускаються і передбачається, що відповідний параметр стосується обвідної SBR каналу звукового вмісту.

Під час декодування кодованого бітового потоку роботою гармонічної транспозиції на ступені обробки eSBR декодування (для кожного каналу, "ch", звукового вмісту, вказаного бітовим потоком) керують за допомогою наступних параметрів метаданих eSBR: `sbrPatchingMode[ch]`; `sbrOversamplingFlag[ch]`; `sbrPitchInBinsFlag[ch]`; і `sbrPitchInBins[ch]`.

Значення "`sbrPatchingMode[ch]`" вказує тип транспозера, який використовується в eSBR: `sbrPatchingMode[ch] = 1` вказує вставку шляхом лінійної транспозиції, як описано в розділі 4.6.18 стандарту MPEG-4 AAC (при використанні або з високоякісним SBR, або з SBR малої потужності); `sbrPatchingMode[ch] = 0` вказує на гармонічне вставляння SBR, як описано в розділі 7.5.3 або 7.5.4 стандарту MPEG USAC.

Значення "`sbrOversamplingFlag[ch]`" вказує на використання в eSBR адаптивної до сигналу передискретизації в частотній ділянці в поєднанні з гармонічним вставлянням SBR на основі DFT, як описано в розділі 7.5.3 стандарту MPEG USAC. Цей прапор керує розміром перетворень DFT, що використовуються в транспозері: 1 вказує, що адаптивна до сигналу передискретизація в частотній ділянці включена, як описано в розділі 7.5.3.1 стандарту MPEG USAC; 0 вказує, що адаптивна до сигналу передискретизація в частотній ділянці відключена, як описано в розділі 7.5.3.1 стандарту MPEG USAC.

Значення "`sbrPitchInBinsFlag[ch]`" керує інтерпретацією параметра `sbrPitchInBins[ch]`: 1 вказує, що значення в `sbrPitchInBins[ch]` є дійсним і більше нуля; 0 вказує, що значення `sbrPitchInBins[ch]` встановлено на нуль.

Значення "`sbrPitchInBins[ch]`" керує додаванням множників векторного добутку в гармонічному транспозері SBR. Значення `sbrPitchInBins[ch]` є цілим числом у діапазоні [0,127] і представляє відстань, виміряну в елементах розділення за частотою, для DFT з 1536 лініями, яке діє на частоту дискретизації основного кодера.

У випадку, коли бітовий потік MPEG-4 AAC характеризує пару каналів SBR, при цьому канали не з'єднані (а не єдиний канал SBR), бітовий потік характеризує два екземпляра вищеописаного синтаксису (для гармонічної або негармонічної транспозиції), по одному для кожного каналу `sbr_channel_pair_element()`.

Гармонічна транспозиція інструмента eSBR зазвичай підвищує якість декодованих музикальних сигналів при відносно низьких частотах розділення. Негармонічна транспозиція (тобто традиційне спектральне вставляння) зазвичай покращує мовні сигнали. Таким чином, початковою точкою в прийнятті рішення щодо того, транспозиція якого типу є переважною для кодування конкретного звукового вмісту, є вибір способу транспозиції залежно від виявлення мови/музики, при цьому гармонічну транспозицію застосовують до музикального вмісту, а спектральне вставляння – до мовного вмісту.

Роботою попереднього згладжування в ході обробки eSBR керують за допомогою значення одинбітного параметра метаданих eSBR, відомого як "`bs_sbr_preprocessing`", у тому значенні, що попереднє згладжування або виконують, або не виконують, залежно від значення цього єдиного біта. При використанні алгоритму QMF-вставки SBR, описаного в розділі 4.6.18.6.3 стандарту MPEG-4 AAC, етап попереднього згладжування може виконуватися (якщо вказано параметром "`bs_sbr_preprocessing`") у спробі уникнути порушень безперервності форми обвідної спектра високочастотного сигналу, який вводиться в коректор обвідної (коректор обвідної

виконує інший ступінь обробки eSBR). Попереднє згладжування зазвичай покращує роботу наступного ступеня корекції обвідної, що призводить до більшої стабільності сприйманого сигналу верхнього діапазону.

5 Очікується, що загальні вимоги до бітової швидкості для включення в бітовий потік MPEG-4 AAC метаданих eSBR, які вказують вищезгадані інструменти eSBR (гармонічну транспозицію і попереднє згладжування), мають порядок декількох сотень бітів у секунду, оскільки відповідно до деяких варіантів здійснення даного винаходу передаються тільки диференційні дані керування, необхідні для виконання обробки eSBR. Традиційні декодери можуть ігнорувати цю інформацію, оскільки вона включена із забезпеченням зворотної сумісності (як буде роз'яснено пізніше). Тому негативний вплив на бітову швидкість, пов'язаний із включенням метаданих eSBR, є нехтовно малим, із багатьох причин, у тому числі наступних:

10 - втрати бітової швидкості передачі даних (через включення метаданих eSBR) становлять дуже невелику долю загальної бітової швидкості, оскільки передаються тільки диференційні дані керування, необхідні для виконання обробки eSBR (і не виконується одночасна передача даних керування SBR); і

15 - налаштування інформації керування, яка стосується SBR, зазвичай не залежить від деталей транспозиції. Приклади того, як дані керування в дійсності залежать від роботи транспозера, обговорені в даній заявці нижче.

Отже, варіанти здійснення даного винаходу надають засоби для ефективної передачі даних керування або метаданих покращеного копіювання спектральної смуги (eSBR) із забезпеченням зворотної сумісності. Ця ефективна передача даних керування eSBR знижує вимоги щодо пам'яті в декодерах, кодерах і транскодерах, у яких застосовуються аспекти даного винаходу, без відчутного негативного впливу при цьому на бітову швидкість. Крім того, складність і вимоги обробки, пов'язані з виконанням eSBR згідно з варіантами здійснення даного винаходу, також зменшується, оскільки дані SBR необхідно обробляти тільки один раз, а не передавати одночасно, як було б у випадку, якщо б eSBR оброблялось як повністю окремий тип об'єкта в MPEG-4 AAC замість інтеграції в кодек MPEG-4 AAC із забезпеченням зворотної сумісності.

Далі, з посиланням на фіг. 7, описані елементи блока ("raw_data_block") бітового потоку MPEG-4 AAC, у який метадані eSBR включені згідно з деякими варіантами здійснення даного винаходу. На фіг. 7 представлена схема блока ("raw_data_block") бітового потоку MPEG-4 AAC, на якій показані деякі його сегменти.

Блок бітового потоку MPEG-4 AAC може містити щонайменше один елемент "single_channel_element()" (наприклад, елемент одного каналу, показаний на фіг. 7) й/або щонайменше один елемент "channel_pair_element()" (конкретно не показаний на фіг. 7, хоча він може бути), який містить звукові дані для звукової програми. Блок також може містити ряд елементів "fill_elements" (наприклад, заповнювальний елемент 1 й/або заповнювальний елемент 2 на фіг. 7), які містять дані (наприклад, метадані), що стосуються програми. Кожний "single_channel_element()" містить ідентифікатор (наприклад, "ID1" на фіг. 7), який вказує початок елемента одного каналу, і може містити звукові дані, які характеризують окремий канал багатоканальної звукової програми. Кожний "channel_pair_element" містить ідентифікатор (не показаний на фіг. 7), який вказує початок елемента пари каналів, і може містити звукові дані, які характеризують два канали програми.

Елемент fill_element (який називається в даному документі заповнювальним елементом) бітового потоку MPEG-4 AAC містить ідентифікатор ("ID2" на фіг. 7), який вказує початок заповнювального елемента, і заповнювальні дані після ідентифікатора. Ідентифікатор ID2 може складатися з трибітного цілого числа без знака, в якого спочатку передається старший значущий біт ("uimsbf"), що має значення 0×6 . Заповнювальні дані можуть містити елемент extension_payload() (який іноді називається в даному документі корисними даними розширення), синтаксис якого показаний у таблиці 4.57 стандарту MPEG-4 AAC. Існує декілька типів корисних даних розширення, які ідентифікуються за допомогою параметра "extension_type", який являє собою чотирибітне ціле число без знака, в якого спочатку передають старший значущий біт ("uimsbf").

Заповнювальні дані (наприклад, їхні корисні дані розширення) можуть містити заголовок, або ідентифікатор (наприклад, "заголовок 1" на фіг. 7), який вказує сегмент заповнювальних даних, який вказує об'єкт SBR (тобто заголовок ініціалізує тип "об'єкт SBR", який у стандарті MPEG-4 AAC називається sbr_extension_data()). Наприклад, корисні дані розширення копіювання спектральної смуги (SBR) ідентифікуються значенням "1101" або "1110" для поля extension_type в заголовку, при цьому ідентифікатор "1101" ідентифікує корисні дані розширення з даними SBR, а "1110" ідентифікує корисні дані розширення з даними SBR і циклічним контролем надмірності (CRC) для перевірки правильності даних SBR.

Якщо заголовок (наприклад, поле `extension_type`) ініціалізує тип об'єкта SBR, метадані SBR (які іноді називаються в даному документі "даними копіювання спектральної смуги" і називаються `sbr_data()` в стандарті MPEG-4 AAC) йдуть за заголовком, а за метаданими SBR може йти щонайменше один елемент розширення копіювання спектральної смуги (наприклад, "елемент розширення SBR" заповнювального елемента 1 на фіг. 7). Такий елемент розширення копіювання спектральної смуги (сегмент бітового потоку) в стандарті MPEG-4 AAC називається контейнером "`sbr_extension()`". Елемент розширення копіювання спектральної смуги обов'язково містить заголовок (наприклад, "заголовок розширення SBR" заповнювального елемента 1 на фіг. 7).

У стандарті MPEG-4 AAC передбачено, що елемент розширення копіювання спектральної смуги може містити PS (параметричні стереофонічні) дані для звукових даних програми. У стандарті MPEG-4 AAC передбачено, що якщо заголовок заповнювального елемента (наприклад, його корисних даних розширення) ініціалізує тип об'єкта SBR (як робить "заголовок 1" на фіг. 7) і елемент розширення копіювання спектральної смуги заповнювального елемента містить PS-дані, заповнювальний елемент (наприклад, його корисні дані розширення) містить дані копіювання спектральної смуги і параметр "`bs_extension_id`", значення якого (тобто "`bs_extension_id`" = 2) вказує, що PS-дані включені в елемент розширення копіювання спектральної смуги заповнювального елемента.

Згідно з деякими варіантами здійснення даного винаходу метадані eSBR (наприклад, прапор, який вказує, чи необхідно виконувати обробку покращеного копіювання спектральної смуги (eSBR) на звуковому вмісті блока) включені в елемент розширення копіювання спектральної смуги заповнювального елемента. Наприклад, такий прапор вказаний у заповнювальному елементі 1 на фіг. 7, при цьому прапор з'являється після заголовка ("заголовка розширення SBR" заповнювального елемента 1) "елемента розширення SBR" заповнювального елемента 1. Необов'язково такий прапор і додаткові метадані eSBR включені в елемент розширення копіювання спектральної смуги після заголовка елемента розширення копіювання спектральної смуги (наприклад, в елемент розширення SBR заповнювального елемента 1 на фіг. 7 після заголовка розширення SBR). Згідно з деякими варіантами здійснення даного винаходу заповнювальний елемент, який містить метадані eSBR, також містить параметр "`bs_extension_id`", значення якого (наприклад, `bs_extension_id=3`) вказує, що метадані eSBR включені в заповнювальний елемент і що на звуковому вмісті відповідного блока необхідно виконати обробку eSBR.

Згідно з деякими варіантами здійснення даного винаходу метадані eSBR включені в заповнювальний елемент (наприклад, заповнювальний елемент 2 на фіг. 7) бітового потоку MPEG-4 AAC, який відрізняється від елемента розширення копіювання спектральної смуги (елемента розширення SBR) заповнювального елемента. Це пояснюється тим, що заповнювальні елементи, які містять `extension_payload()` з даними SBR або даними SBR з CRC, не містять ніяких інших корисних даних розширення будь-якого іншого типу розширення. Тому в варіантах здійснення, в яких метадані eSBR зберігають свої власні корисні дані розширення, для зберігання метаданих eSBR використовується окремий заповнювальний елемент. Такий заповнювальний елемент містить ідентифікатор (наприклад, "ID2" на фіг. 7), який вказує початок заповнювального елемента, і заповнювальні дані після ідентифікатора. Заповнювальні дані можуть містити елемент `extension_payload()` (який іноді називається в даному документі корисними даними розширення), синтаксис якого показаний у таблиці 4.57 стандарту MPEG-4 AAC. Заповнювальні дані (наприклад, їхні корисні дані розширення) містять заголовок (наприклад, "заголовок 2" заповнювального елемента 2 на фіг. 7), який вказує об'єкт eSBR (тобто заголовок ініціалізує тип об'єкта покращеного копіювання спектральної смуги (eSBR)), і заповнювальні дані (наприклад, їхні корисні дані розширення) містять метадані eSBR після заголовка. Наприклад, заповнювальний елемент 2 на фіг. 7 містить такий заголовок ("заголовок 2"), а також містить, після заголовка, метадані eSBR (тобто "прапор" у заповнювальному елементі 2, який вказує, чи необхідно виконувати обробку покращеного копіювання спектральної смуги (eSBR) на звуковому вмісті блока). Необов'язково додаткові метадані eSBR також включені в заповнювальні дані заповнювального елемента 2 на фіг. 7, після "заголовка 2". У варіантах здійснення, що описуються в даному абзаці, заголовок (наприклад, "заголовок 2" на фіг. 7) має ідентифікаційне значення, яке не є одним зі звичайних значень, заданих у таблиці 4.57 стандарту MPEG-4 AAC, і замість цього вказує корисні дані розширення eSBR (таким чином поле `extension_type` заголовка вказує, що заповнювальні дані містять метадані eSBR).

У першому класі варіантів здійснення даний винахід являє собою блок обробки звуку (наприклад, декодер), який містить:

запам'ятовувальний пристрій (наприклад, буфер 201, представлений на фіг. 3 або 4),

виконаний із можливістю зберігання щонайменше одного блока кодованого звукового бітового потоку (наприклад, щонайменше одного блока бітового потоку MPEG-4 AAC);

5 пристрій видалення форматування корисних даних бітового потоку (наприклад, елемент 205, представлений на фіг. 3, або елемент 215, представлений на фіг. 4), з'єднаний із запам'ятовувальним пристроєм і виконаний із можливістю демультимплексування щонайменше однієї частини вказаного блока бітового потоку; і

10 підсистему декодування (наприклад, елементи 202 і 203, представлені на фіг. 3, або елементи 202 і 213, представлені на фіг. 4), з'єднану і виконану з можливістю декодування щонайменше однієї частини звукового вмісту вказаного блока бітового потоку, при цьому блок містить:

заповнювальний елемент, який містить ідентифікатор, який вказує початок заповнювального елемента (наприклад, ідентифікатор "id_syn_ele", який має значення 0×6 , згідно з таблицею 4.85 стандарту MPEG-4 AAC), і заповнювальні дані після ідентифікатора, при цьому заповнювальні дані містять:

15 щонайменше один прапор, що ідентифікує, чи необхідно виконувати обробку покращеного копіювання спектральної смуги (eSBR) на звуковому вмісті блока (наприклад, із використанням даних копіювання спектральної смуги і метаданих eSBR, включених у блок).

20 Прапор являє собою метадані eSBR, і прикладом прапора є прапор "sbrPatchingMode". Іншим прикладом прапора є прапор "harmonicSBR". Обидва з цих прапорів вказують, чи необхідно виконувати на звукових даних блока базову форму копіювання спектральної смуги або покращену форму спектрального копіювання. Базова форма спектрального копіювання являє собою спектральне вставлення, а покращена форма копіювання спектральної смуги являє собою гармонічну транспозицію.

25 У деяких варіантах здійснення заповнювальні дані також містять додаткові метадані eSBR (тобто метадані eSBR, відмінні від прапора).

Запам'ятовувальний пристрій може являти собою буферний запам'ятовувальний пристрій (наприклад, реалізацію буфера 201, представленого на фіг. 4), який зберігає (наприклад, енергонезалежно) щонайменше один блок кодованого звукового бітового потоку.

30 Оцінюється, що складність виконання обробки eSBR (з використанням гармонічної транспозиції і попереднього згладжування eSBR) декодером eSBR у ході декодування бітового потоку MPEG-4 AAC, який містить метадані eSBR (які вказують ці інструменти eSBR), буде наступною (для звичайного декодування із вказаними параметрами):

- гармонічна транспозиція (16 кбіт/с, 14400/28800 Гц)
- на основі DFT: 3,68 WMOPS (зважених мільйонів операцій у секунду);
- 35 - на основі QMF: 0,98 WMOPS;
- попередня обробка (попереднє згладжування) QMF-вставки: 0,1 WMOPS.

Відомо, що для перехідних сигналів транспозиція на основі DFT зазвичай виконується краще, ніж транспозиція на основі QMF.

40 Згідно з деякими варіантами здійснення даного винаходу заповнювальний елемент (кодованого звукового бітового потоку), який містить метадані eSBR, також містить параметр (наприклад, параметр "bs_extension_id"), значення якого (наприклад, bs_extension_id=3) сигналізує про те, що метадані eSBR включені в заповнювальний елемент і що обробку eSBR необхідно виконувати на звуковому вмісті відповідного блока, й/або параметр (наприклад, той же параметр "bs_extension_id"), значення якого (наприклад, bs_extension_id=2) сигналізує про те, що контейнер sbr_extension() заповнювального елемента містить PS-дані. Наприклад, як вказано нижче в таблиці 1, такий параметр, що має значення bs_extension_id=2, може сигналізувати про те, що контейнер sbr_extension() заповнювального елемента містить PS-дані, а такий параметр, що має значення bs_extension_id=3, може сигналізувати про те, що контейнер sbr_extension() заповнювального елемента містить метадані eSBR:

50

Таблиця 1

bs_extension_id	Значення
0	Зарезервовано
1	Зарезервовано
2	EXTENSION_ID_PS
3	EXTENSION_ID_ESBR

Згідно з деякими варіантами здійснення даного винаходу синтаксис кожного елемента розширення копіювання спектральної смуги, який містить метадані eSBR й/або PS-дані, як

вказано нижче в таблиці 2 (у якій "sbr_extension()» означає контейнер, який являє собою елемент розширення копіювання спектральної смуги, "bs_extension_id" є таким, як описано в таблиці 1 вище, "ps_data" означає PS-дані і "esbr_data" означає метадані eSBR).

Таблиця 2

sbr_extension(bs_extension_id, num_bits_left)	
{	
switch (bs_extension_id) {	
case EXTENSION_ID_PS:	
num_bits_left -= ps_data();	Примітка 1
break;	
case EXTENSION_ID_ESBR:	
num_bits_left -= esbr_data();	Примітка 2
break;	
default:	
bs_fill_bits;	
num_bits_left=0;	
break;	
}	
}	
Примітка 1: ps_data() повертає кількість зчитаних бітів.	
Примітка 2: esbr_data() повертає кількість зчитаних бітів.	

5

У наведеному як приклад варіанті здійснення esbr_data(), на який є посилання в таблиці 2 вище, вказує значення наступних параметрів метаданих:

1. однобітний параметр метаданих "bs_sbr_preprocessing"; і

10

2. для кожного каналу ("ch") звукового вмісту кодованого бітового потоку, який підлягає декодуванню, кожний із вищеописаних параметрів: "sbrPatchingMode[ch]»; "sbrOversamplingFlag[ch]»; "sbrPitchInBinsFlag[ch]»; і "sbrPitchInBins[ch]».

Наприклад, для вказання цих параметрів метаданих у деяких варіантах здійснення esbr_data() може мати синтаксис, вказаний у таблиці 3.

Таблиця 3

Синтаксис	Кількість бітів
esbr_data(id_aac, bs_coupling)	
{	
bs_sbr_preprocessing;	1
if (id_aac == ID_SCE) {	
if (sbrPatchingMode[0] == 0) {	1
sbrOversamplingFlag[0];	1
if (sbrPitchInBinsFlag[0])	1
sbrPitchInBins[0];	7
else	
sbrPitchInBins[0] = 0;	
} else {	
sbrOversamplingFlag[0] = 0;	
sbrPitchInBins[0] = 0;	
}	
} else if (id_aac == ID_CPE) {	
If (bs_coupling) {	
if (sbrPatchingMode[0,1] == 0) {	1
sbrOversamplingFlag[0,1];	1
if (sbrPitchInBinsFlag[0,1])	1
sbrPitchInBins[0,1];	7
else	
sbrPitchInBins[0,1] = 0;	
} else {	

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

Примітка: bs_sbr_preprocessing визначений, як описано в розділі 6.2.12 ISO/IEC 23003-3:2012. sbrPatchingMode[ch], sbrOversamplingFlag[ch], sbrPitchInBinsFlag[ch] і sbrPitchInBins[ch] визначені, як описано в розділі 7.5 ISO/IEC 23003-3:2012.

Вищеописаний синтаксис забезпечує можливість ефективної реалізації покращеної форми копіювання спектральної смуги, такої як гармонічна транспозиція, як розширення для традиційного декодера. Зокрема, дані eSBR у таблиці 3 містять тільки ті параметри, які є необхідними для виконання покращеної форми копіювання спектральної смуги, які або вже не підтримуються в бітовому потоці, або не є такими, що виводяться прямо з параметрів, що вже підтримуються в бітовому потоці. Всю решту параметрів і даних обробки, необхідних для виконання покращеної форми копіювання спектральної смуги, вилучають із параметрів, які існували раніше, у вже визначених місцях бітового потоку.

Наприклад, декодер, який відповідає вимогам MPEG-4 HE-AAC або HE-AAC v2, можна розширити таким чином, щоб він включав покращену форму копіювання спектральної смуги, таку як гармонічна транспозиція. Ця покращена форма копіювання спектральної смуги являє собою доповнення до базової форми копіювання спектральної смуги, яка вже підтримується декодером. У контексті декодера, який відповідає вимогам MPEG-4 HE-AAC або HE-AAC v2, ця базова форма копіювання спектральної смуги являє собою інструмент спектральної QMF-вставки SBR, як визначено в розділі 4.6.18 стандарту MPEG-4 AAC.

При виконанні покращеної форми копіювання спектральної смуги розширений декодер HE-AAC може знову використовувати багато з параметрів бітового потоку, вже включених у корисні дані розширення SBR бітового потоку. Конкретні параметри, які можуть бути знову використані, включають, наприклад, різні параметри, які визначають таблицю діапазонів головних частот. Ці параметри включають bs_start_freq (параметр, який визначає початок параметра таблиці головних частот), "bs_stop_freq" (параметр, який визначає кінець таблиці головних частот), "bs_freq_scale" (параметр, який визначає кількість діапазонів частот на октаву) і "bs_alter_scale" (параметр, який змінює масштаб діапазонів частот). Параметри, які можуть бути знову використані, також включають параметри, які визначають таблицю смуги шумів ("bs_noise_bands"), і параметри таблиці смуги обмежувача ("bs_limiter_bands"). Відповідно в різних варіантах здійснення щонайменше деякі з еквівалентних параметрів, заданих у стандарті USAC, не включаються в бітовий потік, завдяки чому зменшується переповнення бітового

поток інформацією керування. Зазвичай, якщо параметр, заданий у стандарті AAC, має еквівалентний параметр, заданий у стандарті USAC, то еквівалентний параметр, заданий у стандарті USAC, має таку ж назву, як параметр, заданий у стандарті AAC, наприклад масштабний коефіцієнт $E_{\text{OrigMapped}}$ обвідної. Однак еквівалентний параметр, заданий у стандарті USAC, зазвичай має інше значення, яке "налаштовано" для обробки покращеного SBR, визначеного у стандарті USAC, а не для обробки SBR, визначеного у стандарті AAC.

Для підвищення суб'єктивної якості звукового вмісту з гармонічною частотною структурою і сильними тональними характеристиками, зокрема при низьких бітових швидкостях, рекомендується активація покращеного SBR. Значення відповідного елемента бітового потоку (тобто `esbr_data()`), який керує цими інструментами, можуть бути визначені в кодері шляхом застосування механізму класифікації, залежного від сигналу. Як правило, використання способу гармонічного вставляння (`sbrPatchingMode == 1`) є переважним для кодування музикальних сигналів із дуже низькими бітовими швидкостями, коли основний кодек може бути значно обмежений у смузі пропускання звуку. Це є особливо вірним, якщо ці сигнали мають виражену гармонічну структуру. Навпаки, використання звичайного способу вставляння SBR є переважним для мовних і змішаних сигналів, оскільки він забезпечує краще зберігання тимчасової структури мови.

Для покращення роботи гармонічного транспозера можна активувати етап попередньої обробки (`bs_sbr_preprocessing == 1`), який намагається уникнути внесення порушень однорідності спектра сигналу, який надходить у подальший коректор обвідної. Робота інструмента є корисною для типів сигналів, у яких груба обвідна спектра сигналу нижнього діапазону, що використовується для реконструкції високих частот, проявляє великі варіації рівня.

Для покращення перехідної характеристики гармонічного вставляння SBR можна застосовувати адаптивну до сигналу передискретизацію в частотній ділянці (`sbrOversamplingFlag == 1`). Оскільки адаптивна до сигналу передискретизація в частотній ділянці збільшує обчислювальну складність транспозера, але дає користь тільки для кадрів, які містять перехідні сигнали, використанням цього інструмента керують за допомогою елемента бітового потоку, який передають один раз на кадр і на незалежний канал SBR.

Декодеру, який діє в запропонованому режимі покращеного SBR, зазвичай необхідно мати можливість переключення між традиційним і покращеним вставлянням SBR. Тому може вноситися затримка, яка може бути настільки тривалою, як тривалість одного базового кадру звуку, залежно від налаштувань декодера. Зазвичай затримка буде однаковою і для традиційного, і для покращеного вставляння SBR.

На додаток до численних параметрів, розширений декодер HE-AAC при виконанні покращеної форми копіювання спектральної смуги згідно з варіантами здійснення даного винаходу може знову використовувати й інші елементи даних. Наприклад, дані обвідної і дані шумового порога також можуть бути вилучені з даних `bs_data_env` (масштабних коефіцієнтів обвідної) і `bs_noise_env` (масштабних коефіцієнтів шумового порога) і використані в ході покращеної форми копіювання спектральної смуги.

По суті ці варіанти здійснення використовують конфігураційні параметри і дані обвідної, які вже підтримуються традиційним декодером HE-AAC або HE-AAC v2, в корисних даних розширення SBR для забезпечення покращеної форми копіювання спектральної смуги, яка вимагає як можна менше додаткових даних, які передаються. Метадані були спершу налаштовані для базової форми HFR (наприклад, для операції спектрального переносу SBR), але згідно з варіантами здійснення використовуються для покращеної форми HFR (наприклад, гармонічної транспозиції eSBR). Як обговорено раніше, метадані, як правило, представляють робочі параметри (наприклад, масштабні коефіцієнти обвідної, масштабні коефіцієнти шумового порога, параметри часової/частотної сітки, інформацію про додавання синусоїд, змінну частоту/діапазон розділення, режим зворотної фільтрації, розділення обвідної, режим вирівнювання, режим частотної інтерполяції), налаштовані і призначені для використання з базовою формою HFR (наприклад, із лінійним спектральним переносом). Однак ці метадані в поєднанні з додатковими параметрами метаданих, специфічними для покращеної форми HFR (наприклад, гармонічної транспозиції), можуть бути використані для раціональної й ефективної обробки звукових даних із використанням покращеної форми HFR.

Відповідно розширені декодери, які підтримують покращену форму копіювання спектральної смуги, можна створювати надзвичайно ефективним чином, покладаючись на вже визначені елементи бітового потоку (наприклад, елементи в корисних даних розширення SBR) і додаючи тільки ті параметри, які необхідні для підтримки покращеної форми копіювання спектральної смуги (в корисних даних розширення заповнювального елемента). Ця ознака ущільнення даних

у поєднанні з розміщенням параметрів, які знову додаються, в зарезервованому полі даних, такому як контейнер розширення, суттєво знижує бар'єри для створення декодера, який підтримує покращену форму копіювання спектральної смуги, завдяки забезпеченню зворотної сумісності бітового потоку з традиційним декодером, який не підтримує покращену форму копіювання спектральної смуги.

У таблиці 3 число в правій колонці вказує кількість бітів відповідного параметра з лівої колонки.

У деяких варіантах здійснення тип об'єкта SBR, визначений у MPEG-4 AAC, оновлюють таким чином, щоб він містив інструмент SBR і аспекти інструмента покращеного SBR (eSBR), як вказано сигналами в елементі розширення SBR (`bs_extension_id == EXTENSION_ID_ESBR`). Якщо декодер виявляє і підтримує цей елемент розширення SBR, то декодер використовує вказані сигналами аспекти інструмента покращеного SBR. Тип об'єкта SBR, оновлений таким чином, називається "покращеннями SBR".

У деяких варіантах здійснення даний винахід являє собою спосіб, який включає етап кодування звукових даних із генеруванням кодованого бітового потоку (наприклад, бітового потоку MPEG-4 AAC), який містить метадані eSBR в щонайменше одному сегменті щонайменше одного блока кодованого бітового потоку і звукові дані в щонайменше одному іншому сегменті блока. У звичайних варіантах здійснення спосіб включає етап мультиплексування звукових даних із метаданими eSBR у кожному блоці кодованого бітового потоку. При звичайному декодуванні кодованого бітового потоку в декодері eSBR декодер вилучає метадані eSBR з бітового потоку (у тому числі за допомогою синтаксичного аналізу і демультіплексування метаданих eSBR і звукових даних) і використовує метадані eSBR для обробки звукових даних, щоб генерувати потік декодованих звукових даних.

Інший аспект даного винаходу являє собою декодер eSBR, виконаний із можливістю виконання обробки eSBR (наприклад, із використанням щонайменше одного з інструментів eSBR, відомих як гармонічна транспозиція або попереднє згладжування) в ході декодування кодованого звукового бітового потоку (наприклад, бітового потоку MPEG-4 AAC), який не включає метадані eSBR. Приклад такого декодера буде описаний із посиланням на фіг. 5.

Декодер (400) eSBR, представлений на фіг. 5, містить буферний запам'ятовувальний пристрій 201 (ідентичний із запам'ятовувальним пристроєм 201, представленим на фіг. 3 і 4), пристрій 215 видалення форматування корисних даних бітового потоку (ідентичний із пристроєм 215 видалення форматування, представленим на фіг. 4), підсистему 202 декодування звуку (яка іноді називається "основним" ступенем декодування або "основною" підсистемою декодування, яка ідентична з основною підсистемою 202 декодування, представленою на фіг. 3), підсистему 401 генерування даних керування eSBR і ступінь 203 обробки eSBR (ідентичний зі ступенем 203, представленим на фіг. 3), з'єднані таким чином, як показано. Зазвичай декодер 400 також містить інші обробляючі елементи (не показані).

У роботі декодера 400 послідовність блоків кодованого звукового бітового потоку (бітового потоку MPEG-4 AAC), прийнята декодером 400, передається з буфера 201 в пристрій 215 видалення форматування.

Пристрій 215 видалення форматування з'єднаний і виконаний із можливістю демультіплексування кожного блока бітового потоку для вилучення із нього метаданих SBR (у тому числі квантованих даних обвідної) і зазвичай також інших метаданих. Пристрій 215 видалення форматування виконаний із можливістю передачі щонайменше метаданих SBR на ступінь 203 обробки eSBR. Пристрій 215 видалення форматування також з'єднаний і виконаний із можливістю вилучення звукових даних із кожного блока бітового потоку і передачі вилучених звукових даних у підсистему 202 декодування (ступінь декодування).

Підсистема 202 декодування звуку декодера 400 виконана з можливістю декодування звукових даних, вилучених пристроєм 215 видалення форматування (таке декодування можна назвати "основною" операцією декодування), для генерування декодованих звукових даних і передачі декодованих звукових даних на ступінь 203 обробки eSBR. Декодування виконується в частотній ділянці. Зазвичай останній ступінь обробки в підсистемі 202 застосовує перетворення із частотної ділянки в часову ділянку до декодованих звукових даних у частотній ділянці, таким чином виведення підсистеми являє собою декодовані звукові дані в часовій ділянці. Ступінь 203 виконаний із можливістю застосування інструментів SBR (й інструментів eSBR), вказаних метаданими SBR (вилученими пристроєм 215 видалення форматування) і метаданими eSBR, згенерованими в підсистемі 401, до декодованих звукових даних (тобто виконання обробки SBR і eSBR відносно виведення підсистеми 202 декодування з використанням метаданих SBR і eSBR), щоб генерувати повністю декодовані звукові дані, які виводяться з декодера 400. Зазвичай декодер 400 містить запам'ятовувальний пристрій (доступний для підсистеми 202 і

ступеня 203), який зберігає звукові дані з видаленим форматуванням і метадані, виведені з пристрою 215 видалення форматування (і необов'язково також із підсистеми 401), і ступінь 203 виконаний із можливістю здійснення доступу до звукових даних і метаданих у міру необхідності в ході обробки SBR і eSBR. Обробку SBR на ступені 203 можна вважати постобробкою відносно виведення основної підсистеми 202 декодування. Необов'язково декодер 400 також містить підсистему прикінцевого підвищувального мікшування (яка може застосовувати інструменти параметричної стереофонії ("PS"), визначені в стандарті MPEG-4 AAC, із використанням метаданих PS, вилучених пристроєм 215 видалення форматування), з'єднану і виконану з можливістю виконання підвищувального мікшування відносно виведення ступеня 203 для генерування повністю декодованого і підданого підвищувальному мікшуванню звуку, який виводиться з APU 210.

Параметрична стереофонія є інструментом кодування, який представляє стереофонічний сигнал із використанням лінійного знижувального мікшування лівого і правого каналів стереофонічного сигналу і набору просторових параметрів, що описують стереозображення. У параметричній стереофонії зазвичай використовують три типи просторових параметрів: (1) міжканальні різниці інтенсивності (IID), що описують різниці інтенсивностей між каналами; (2) міжканальні різниці фази (IPD), що описують різниці фази між каналами; і (3) міжканальна когерентність (ICC), що описує когерентність (або подібність) між каналами. Когерентність можна виміряти як максимум взаємної кореляції залежно від часу або фази. Ці три параметра, як правило, забезпечують можливість високоякісної реконструкції стереозображення. Однак параметри IPD задають лише відносні різниці фази між каналами стереофонічного вхідного сигналу і не вказують розподіл цих різниць фази по лівому і правому каналам. Тому додатково можна використовувати параметри четвертого типу, що описують загальний зсув фази або загальну різницю фаз (OPD). У процесі реконструкції стереофонічного сигналу послідовні сегменти, що обробляються методом вікна, як прийнятого сигналу знижувального мікшування, $s[n]$, так і декорельованої версії прийнятого знижувального мікшування, $d[n]$, обробляються разом із просторовими параметрами для генерування лівого ($l_k(n)$) і правого ($r_k(n)$) реконструйованих сигналів відповідно до наступної формули:

$$l_k(n) = H_{11}(k, n)s_k(n) + H_{21}(k, n)d_k(n)$$

$$r_k(n) = H_{12}(k, n)s_k(n) + H_{22}(k, n)d_k(n)$$

де H_{11} , H_{12} , H_{21} і H_{22} визначаються стереофонічними параметрами. Сигнали $l_k(n)$ і $r_k(n)$ зрештою перетворюються назад у часову ділянку за допомогою частотно-часового перетворення.

Підсистема 401 генерування даних керування, представлена на фіг. 5, з'єднана і виконана з можливістю виявлення щонайменше однієї властивості кодованого звукового бітового потоку, який підлягає декодуванню, і генерування даних керування eSBR (які можуть являти собою або включають метадані eSBR будь-якого з типів, включені в кодовані звукові бітові потоки згідно з іншими варіантами здійснення даного винаходу) у відповідь на щонайменше один результат етапу виявлення. Дані керування eSBR передаються на ступінь 203, щоб запускати застосування окремих інструментів eSBR або комбінацій інструментів eSBR при виявленні конкретної властивості (або комбінації властивостей) бітового потоку й/або керувати застосуванням таких інструментів eSBR. Наприклад, для керування виконанням обробки eSBR з використанням гармонічної транспозиції деякі варіанти здійснення підсистеми 401 генерування даних керування будуть містити: детектор музики (наприклад, спрощену версію традиційного детектора музики) для встановлення параметра `sbrPatchingMode[ch]` (і передачі встановленого параметра на ступінь 203) у відповідь на виявлення того, що бітовий потік характеризує або не характеризує музику; детектор перехідних сигналів для встановлення параметра `sbrOversamplingFlag[ch]` (і передачі встановленого параметра на ступінь 203) у відповідь на виявлення присутності або відсутності перехідних сигналів у звуковому вмісті, вказаному бітовим потоком; й/або детектор основного тону для встановлення параметрів `sbrPitchInBinsFlag[ch]` і `sbrPitchInBins[ch]` (і передачі встановлених параметрів на ступінь 203) у відповідь на виявлення основного тону звукового вмісту, вказаного бітовим потоком. Іншими аспектами даного винаходу є способи декодування звукового бітового потоку, які виконуються за допомогою будь-якого варіанта здійснення декодера відповідно до винаходу, описаного в цьому абзаці і попередньому абзаці.

Аспекти даного винаходу включають спосіб кодування або декодування, що стосується типу, з можливістю виконання якого виконаний (наприклад, запрограмований) будь-який варіант здійснення APU, системи або пристрою відповідно до винаходу. Інші аспекти даного винаходу включають систему або пристрій, виконані з можливістю (наприклад, запрограмовані для) виконання будь-якого варіанта здійснення способу відповідно до винаходу, і

машинопрочитуваний носій (наприклад, диск), на якому зберігається (наприклад, енергонезалежно) код, призначений для реалізації будь-якого варіанта здійснення способу відповідно до винаходу або його етапів. Наприклад, система відповідно до винаходу може являти собою або включати програмований процесор загального призначення, процесор цифрової обробки сигналів або мікропроцесор, запрограмований із використанням програмного забезпечення або програмно-апаратного забезпечення й/або іншим чином виконаний із можливістю виконання будь-якої з множини операцій відносно даних, у тому числі варіанта здійснення способу відповідно до винаходу або його етапів. Такий процесор загального призначення може являти собою або включати комп'ютерну систему, яка містить пристрій введення, запам'ятовувальний пристрій і схему обробки, запрограмовану для (й/або іншим чином виконану з можливістю) виконання варіанта здійснення способу відповідно до винаходу (або його етапів) у відповідь на дані, які передаються до неї.

Варіанти здійснення даного винаходу можуть бути реалізовані в апаратному забезпеченні, програмно-апаратному забезпеченні, програмному забезпеченні або їхній комбінації (наприклад, як програмована логічна матриця). Якщо не зазначене інше, алгоритми або процеси, включені як частини даного винаходу, по своїй суті не стосуються якого-небудь конкретного комп'ютера або іншого пристрою. Зокрема, різні машини загального призначення можуть бути використані разом із програмами, написаними згідно з ідеями, викладеними в даному документі, або може бути зручніше сконструювати більш спеціалізований пристрій (наприклад, інтегральні мікросхеми) для виконання необхідних етапів способу. Таким чином, даний винахід може бути реалізований в одній або більше комп'ютерних програмах, які здійснюються на одній або більше програмованих комп'ютерних системах (наприклад, на реалізації будь-якого з елементів, представлених на фіг. 1, або кодері 100, представленому на фіг. 2, (або його елементі), або декодері 200, представленому на фіг. 3, (або його елементі), або декодері 210, представленому на фіг. 4, (або його елементі), або декодері 400, представленому на фіг. 5, (або його елементі)), кожна з яких містить щонайменше один процесор, щонайменше одну систему зберігання даних (у тому числі енергозалежні й енергонезалежні запам'ятовувальні пристрої й/або елементи зберігання), щонайменше один пристрій або порт введення і щонайменше один пристрій або порт виведення. Програмний код застосовується до вхідних даних для виконання функцій, описаних у даному документі, і генерування вихідної інформації. Вихідна інформація відомим способом застосовується до одного або більше пристроїв виведення.

Кожна така програма може бути реалізована на будь-якій необхідній комп'ютерній мові (у тому числі машинній, асемблерній або процедурній високого рівня, логічній або об'єктно-орієнтованій мовах програмування) для здійснення зв'язку з комп'ютерною системою. У кожному разі мова може бути компільовуваною або інтерпретовуваною мовою.

Наприклад, при реалізації за допомогою послідовностей команд комп'ютерного програмного забезпечення різні функції й етапи варіантів здійснення даного винаходу можуть бути реалізовані за допомогою багатопотокових послідовностей команд програмного забезпечення, запущених на підходящому апаратному забезпеченні цифрової обробки сигналів, і в цьому випадку різні пристрої, етапи і функції варіантів здійснення можуть відповідати частинам команд програмного забезпечення.

Кожну таку комп'ютерну програму переважно зберігають або завантажують на інформаційні носії або пристрій (наприклад, твердотільний запам'ятовувальний пристрій або носії, або магнітний чи оптичний носії), що зчитуються програмованим комп'ютером загального або спеціального призначення, для налаштування конфігурації і роботи комп'ютера, коли інформаційні носії або пристрій зчитуються комп'ютерною системою для виконання процедур, описаних у даному документі. Система відповідно до винаходу також може бути реалізована у вигляді машинопрочитуваного інформаційного носія, оснащеного комп'ютерною програмою (тобто який зберігає її), де оснащений таким чином інформаційний носій викликає роботу комп'ютерної системи особливим і попередньо визначеним чином для виконання функцій, описаних у даному документі.

Був описаний ряд варіантів здійснення даного винаходу. Однак необхідно розуміти, що різні модифікації можуть бути здійснені без відступу від об'єму формули винаходу. У світі вищевикладених ідей є можливими численні модифікації і зміни даного винаходу. Наприклад, для полегшення ефективних реалізацій зсуви фази можна використовувати в комбінації з аналізом за допомогою блоків комплексних QMF-фільтрів аналізу і синтезу. Блок фільтрів аналізу відповідає за фільтрацію сигналу нижнього діапазону у часовій ділянці, який генерується основним декодером, в множину піддіапазонів (наприклад, QMF-піддіапазонів). Блок синтезувальних фільтрів відповідає за об'єднання відновленого верхнього діапазону,

отриманого за допомогою вибраної методики HFR (як вказано прийнятим параметром sbrPatchingMode), з декодованим нижнім діапазоном для отримання широкосмугового вихідного звукового сигналу. Дана реалізація блока фільтрів, що діє у певному режимі частоти дискретизації, наприклад в режимі звичайної двошвидкісної дії або в режимі SBR зі зниженою дискретизацією, не повинна, однак, мати зсуви фази, що залежать від бітового потоку. QMF-блоки, що використовуються в SBR, являють собою комплексно-експоненціальне розширення теорії косинусно-модульованих блоків фільтрів. Можна показати, що при розширенні косинусно-модульованого блока фільтрів за допомогою комплексно-експоненціальної модуляції обмеження, пов'язані з виключенням паразитних сигналів, стають неактуальними. Таким чином, для QMF-блоків SBR як фільтри аналізу, $h_k(n)$, так і синтезувальні фільтри, $f_k(n)$, можна визначити як:

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

де $p_0(n)$ – це дійснозначний симетричний або асиметричний фільтр-прототип (зазвичай фільтр-прототип нижніх частот), M означає кількість каналів, а N – порядок фільтра-прототипу. Кількість каналів, які використовуються в блоці фільтрів аналізу, можуть відрізнятися від кількості каналів, які використовуються в блоці синтезувальних фільтрів. Наприклад, блок фільтрів аналізу може мати 32 канали, а блок синтезувальних фільтрів може мати 64 канали. При роботі блока синтезувальних фільтрів у режимі зниженої дискретизації блок синтезувальних фільтрів може мати тільки 32 канали. Оскільки дискретні значення піддіапазонів із блока фільтрів є комплекснозначними, до блока фільтрів аналізу може бути прикладений додатковий, який можливо залежить від каналів, етап зсуву фази. Ці додаткові зсуви фази необхідно компенсувати перед блоком синтезувальних фільтрів. Хоча складові зсуви фаз здебільшого можуть являти собою довільні значення, без погіршення роботи QMF-ланцюжку аналізу/синтезу, їх також можна обмежити певними значеннями для перевірки відповідності. На сигнал SBR буде впливати вибір фазових коефіцієнтів, а сигнал нижніх частот, який надходить із основного декодера, впливати не буде. Якість звуку вихідного сигналу не зазнає негативного впливу.

Коефіцієнти фільтра-прототипу, $p_0(n)$, можна визначити з довжиною L , 640, як показано нижче в таблиці 4.

Таблиця 4

n	$p_0(n)$	n	$p_0(n)$	n	$p_0(n)$
0	0,0000000000	214	0,0019765601	428	0,0117623832
1	-0,0005525286	215	-0,0032086896	429	0,0163701258
2	-0,0005617692	216	-0,0085711749	430	0,0207997072
3	-0,0004947518	217	-0,0141288827	431	0,0250307561
4	-0,0004875227	218	-0,0198834129	432	0,0290824006
5	-0,0004893791	219	-0,0258227288	433	0,0329583930
6	-0,0005040714	220	-0,0319531274	434	0,0366418116
7	-0,0005226564	221	-0,0382776572	435	0,0401458278
8	-0,0005466565	222	-0,0447806821	436	0,0434768782
9	-0,0005677802	223	-0,0514804176	437	0,0466303305
10	-0,0005870930	224	-0,0583705326	438	0,0495978676
11	-0,0006132747	225	-0,0654409853	439	0,0524093821
12	-0,0006312493	226	-0,0726943300	440	0,0550460034
13	-0,0006540333	227	-0,0801372934	441	0,0575152691
14	-0,0006777690	228	-0,0877547536	442	0,0598166570
15	-0,0006941614	229	-0,0955533352	443	0,0619602779
16	-0,0007157736	230	-0,1035329531	444	0,0639444805
17	-0,0007255043	231	-0,1116826931	445	0,0657690668
18	-0,0007440941	232	-0,1200077984	446	0,0674525021
19	-0,0007490598	233	-0,1285002850	447	0,0689664013
20	-0,0007681371	234	-0,1371551761	448	0,0703533073
21	-0,0007724848	235	-0,1459766491	449	0,0715826364
22	-0,0007834332	236	-0,1549607071	450	0,0726774642
23	-0,0007779869	237	-0,1640958855	451	0,0736406005
24	-0,0007803664	238	-0,1733808172	452	0,0744664394

25	-0,0007801449	239	-0,1828172548	453	0,0751576255
26	-0,0007757977	240	-0,1923966745	454	0,0757305756
27	-0,0007630793	241	-0,2021250176	455	0,0761748321
28	-0,0007530001	242	-0,2119735853	456	0,0765050718
29	-0,0007319357	243	-0,2219652696	457	0,0767204924
30	-0,0007215391	244	-0,2320690870	458	0,0768230011
31	-0,0006917937	245	-0,2423016884	459	0,0768173975
32	-0,0006650415	246	-0,2526480309	460	0,0767093490
33	-0,0006341594	247	-0,2631053299	461	0,0764992170
34	-0,0005946118	248	-0,2736634040	462	0,0761992479
35	-0,0005564576	249	-0,2843214189	463	0,0758008358
36	-0,0005145572	250	-0,2950716717	464	0,0753137336
37	-0,0004606325	251	-0,3059098575	465	0,0747452558
38	-0,0004095121	252	-0,3168278913	466	0,0741003642
39	-0,0003501175	253	-0,3278113727	467	0,0733620255
40	-0,0002896981	254	-0,3388722693	468	0,0725682583
41	-0,0002098337	255	-0,3499914122	469	0,0717002673
42	-0,0001446380	256	0,3611589903	470	0,0707628710
43	-0,0000617334	257	0,3723795546	471	0,0697630244
44	0,0000134949	258	0,3836350013	472	0,0687043828
45	0,0001094383	259	0,3949211761	473	0,0676075985
46	0,0002043017	260	0,4062317676	474	0,0664367512
47	0,0002949531	261	0,4175696896	475	0,0652247106
48	0,0004026540	262	0,4289119920	476	0,0639715898
49	0,0005107388	263	0,4402553754	477	0,0626857808
50	0,0006239376	264	0,4515996535	478	0,0613455171
51	0,0007458025	265	0,4629308085	479	0,0599837480
52	0,0008608443	266	0,4742453214	480	0,0585915683
53	0,0009885988	267	0,4855253091	481	0,0571616450
54	0,0011250155	268	0,4967708254	482	0,0557173648
55	0,0012577884	269	0,5079817500	483	0,0542452768
56	0,0013902494	270	0,5191234970	484	0,0527630746
57	0,0015443219	271	0,5302240895	485	0,0512556155
58	0,0016868083	272	0,5412553448	486	0,0497385755
59	0,0018348265	273	0,5522051258	487	0,0482165720
60	0,0019841140	274	0,5630789140	488	0,0466843027
61	0,0021461583	275	0,5738524131	489	0,0451488405
62	0,0023017254	276	0,5845403235	490	0,0436097542
63	0,0024625616	277	0,5951123086	491	0,0420649094
64	0,0026201758	278	0,6055783538	492	0,0405349170
65	0,0027870464	279	0,6159109932	493	0,0390053679
66	0,0029469447	280	0,6261242695	494	0,0374812850
67	0,0031125420	281	0,6361980107	495	0,0359697560
68	0,0032739613	282	0,6461269695	496	0,0344620948
69	0,0034418874	283	0,6559016302	497	0,0329754081
70	0,0036008268	284	0,6655139880	498	0,0315017608
71	0,0037603922	285	0,6749663190	499	0,0300502657
72	0,0039207432	286	0,6842353293	500	0,0286072173
73	0,0040819753	287	0,6933282376	501	0,0271859429
74	0,0042264269	288	0,7022388719	502	0,0257875847
75	0,0043730719	289	0,71109410426	503	0,0244160992
76	0,0045209852	290	0,7194462634	504	0,0230680169
77	0,0046606460	291	0,7277448900	505	0,0217467550
78	0,0047932560	292	0,7358211758	506	0,0204531793
79	0,0049137603	293	0,7436827863	507	0,0191872431
80	0,0050393022	294	0,7513137456	508	0,0179433381
81	0,0051407353	295	0,7587080760	509	0,0167324712
82	0,0052461166	296	0,7658674865	510	0,0155405553

83	0,0053471681	297	0,7727780881	511	0,0143904666
84	0,0054196775	298	0,7794287519	512	-0,0132718220
85	0,0054876040	299	0,7858353120	513	-0,0121849995
86	0,0055475714	300	0,7919735841	514	-0,0111315548
87	0,0055938023	301	0,7978466413	515	-0,0101150215
88	0,0056220643	302	0,8034485751	516	-0,0091325329
89	0,0056455196	303	0,8087695004	517	-0,0081798233
90	0,0056389199	304	0,8138191270	518	-0,0072615816
91	0,0056266114	305	0,8185776004	519	-0,0063792293
92	0,0055917128	306	0,8230419890	520	-0,0055337211
93	0,0055404363	307	0,8272275347	521	-0,0047222596
94	0,0054753783	308	0,8311038457	522	-0,0039401124
95	0,0053838975	309	0,8346937361	523	-0,0031933778
96	0,0052715758	310	0,8379717337	524	-0,0024826723
97	0,0051382275	311	0,8409541392	525	-0,0018039472
98	0,0049839687	312	0,8436238281	526	-0,0011568135
99	0,0048109469	313	0,8459818469	527	-0,0005464280
100	0,0046039530	314	0,8480315777	528	0,0000276045
101	0,0043801861	315	0,8497805198	529	0,0005832264
102	0,0041251642	316	0,8511971524	530	0,0010902329
103	0,0038456408	317	0,8523047035	531	0,0015784682
104	0,0035401246	318	0,8531020949	532	0,0020274176
105	0,0032091885	319	0,8535720573	533	0,0024508540
106	0,0028446757	320	0,8537385600	534	0,0028446757
107	0,0024508540	321	0,8535720573	535	0,0032091885
108	0,0020274176	322	0,8531020949	536	0,0035401246
109	0,0015784682	323	0,8523047035	537	0,0038456408
110	0,0010902329	324	0,8511971524	538	0,0041251642
111	0,0005832264	325	0,8497805198	539	0,0043801861
112	0,0000276045	326	0,8480315777	540	0,0046039530
113	-0,0005464280	327	0,8459818469	541	0,0048109469
114	-0,0011568135	328	0,8436238281	542	0,0049839687
115	-0,0018039472	329	0,8409541392	543	0,0051382275
116	-0,0024826723	330	0,8379717337	544	0,0052715758
117	-0,0031933778	331	0,8346937361	545	0,0053838975
118	-0,0039401124	332	0,8311038457	546	0,0054753783
119	-0,0047222596	333	0,8272275347	547	0,0055404363
120	-0,0055337211	334	0,8230419890	548	0,0055917128
121	-0,0063792293	335	0,8185776004	549	0,0056266114
122	-0,0072615816	336	0,8138191270	550	0,0056389199
123	-0,0081798233	337	0,8087695004	551	0,0056455196
124	-0,0091325329	338	0,8034485751	552	0,0056220643
125	-0,0101150215	339	0,7978466413	553	0,0055938023
126	-0,0111315548	340	0,7919735841	554	0,0055475714
127	-0,0121849995	341	0,7858353120	555	0,0054876040
128	0,0132718220	342	0,7794287519	556	0,0054196775
129	0,0143904666	343	0,7727780881	557	0,0053471681
130	0,0155405553	344	0,7658674865	558	0,0052461166
131	0,0167324712	345	0,7587080760	559	0,0051407353
132	0,0179433381	346	0,7513137456	560	0,0050393022
133	0,0191872431	347	0,7436827863	561	0,0049137603
134	0,0204531793	348	0,7358211758	562	0,0047932560
135	0,0217467550	349	0,7277448900	563	0,0046606460
136	0,0230680169	350	0,7194462634	564	0,0045209852
137	0,0244160992	351	0,7109410426	565	0,0043730719
138	0,0257875847	352	0,7022388719	566	0,0042264269
139	0,0271859429	353	0,6933282376	567	0,0040819753
140	0,0286072173	354	0,6842353293	568	0,0039207432

141	0,0300502657	355	0,6749663190	569	0,0037603922
142	0,0315017608	356	0,6655139880	570	0,0036008268
143	0,0329754081	357	0,6559016302	571	0,0034418874
144	0,0344620948	358	0,6461269695	572	0,0032739613
145	0,0359697560	359	0,6361980107	573	0,0031125420
146	0,0374812850	360	0,6261242695	574	0,0029469447
147	0,0390053679	361	0,6159109932	575	0,0027870464
148	0,0405349170	362	0,6055783538	576	0,0026201758
149	0,0420649094	363	0,5951123086	577	0,0024625616
150	0,0436097542	364	0,5845403235	578	0,0023017254
151	0,0451488405	365	0,5738524131	579	0,0021461583
152	0,0466843027	366	0,5630789140	580	0,0019841140
153	0,0482165720	367	0,5522051258	581	0,0018348265
154	0,0497385755	368	0,5412553448	582	0,0016868083
155	0,0512556155	369	0,5302240895	583	0,0015443219
156	0,0527630746	370	0,5191234970	584	0,0013902494
157	0,0542452768	371	0,5079817500	585	0,0012577884
158	0,0557173648	372	0,4967708254	586	0,0011250155
159	0,0571616450	373	0,4855253091	587	0,0009885988
160	0,0585915683	374	0,4742453214	588	0,0008608443
161	0,0599837480	375	0,4629308085	589	0,0007458025
162	0,0613455171	376	0,4515996535	590	0,0006239376
163	0,0626857808	377	0,4402553754	591	0,0005107388
164	0,0639715898	378	0,4289119920	592	0,0004026540
165	0,0652247106	379	0,4175696896	593	0,0002949531
166	0,0664367512	380	0,4062317676	594	0,0002043017
167	0,0676075985	381	0,3949211761	595	0,0001094383
168	0,0687043828	382	0,3836350013	596	0,0000134949
169	0,0697630244	383	0,3723795546	597	-0,0000617334
170	0,0707628710	384	-0,3611589903	598	-0,0001446380
171	0,0717002673	385	-0,3499914122	599	-0,0002098337
172	0,0725682583	386	-0,3388722693	600	-0,0002896981
173	0,0733620255	387	-0,3278113727	601	-0,0003501175
174	0,0741003642	388	-0,3168278913	602	-0,0004095121
175	0,0747452558	389	-0,3059098575	603	-0,0004606325
176	0,0753137336	390	-0,2950716717	604	-0,0005145572
177	0,0758008358	391	-0,2843214189	605	-0,0005564576
178	0,0761992479	392	-0,2736634040	606	-0,0005946118
179	0,0764992170	393	-0,2631053299	607	-0,0006341594
180	0,0767093490	394	-0,2526480309	608	-0,0006650415
181	0,0768173975	395	-0,2423016884	609	-0,0006917937
182	0,0768230011	396	-0,2320690870	610	-0,0007215391
183	0,0767204924	397	-0,2219652696	611	-0,0007319357
184	0,0765050718	398	-0,2119735853	612	-0,0007530001
185	0,0761748321	399	-0,2021250176	613	-0,0007630793
186	0,0757305756	400	-0,1923966745	614	-0,0007757977
187	0,0751576255	401	-0,1828172548	615	-0,0007801449
188	0,0744664394	402	-0,1733808172	616	-0,0007803664
189	0,0736406005	403	-0,1640958855	617	-0,0007779869
190	0,0726774642	404	-0,1549607071	618	-0,0007834332
191	0,0715826364	405	-0,1459766491	619	-0,0007724848
192	0,0703533073	406	-0,1371551761	620	-0,0007681371
193	0,0689664013	407	-0,1285002850	621	-0,0007490598
194	0,0674525021	408	-0,1200077984	622	-0,0007440941
195	0,0657690668	409	-0,1116826931	623	-0,0007255043
196	0,0639444805	410	-0,1035329531	624	-0,0007157736
197	0,0619602779	411	-0,0955533352	625	-0,0006941614
198	0,0598166570	412	-0,0877547536	626	-0,0006777690

199	0,0575152691	413	-0,0801372934	627	-0,0006540333
200	0,0550460034	414	-0,0726943300	628	-0,0006312493
201	0,0524093821	415	-0,0654409853	629	-0,0006132747
202	0,0495978676	416	-0,0583705326	630	-0,0005870930
203	0,0466303305	417	-0,0514804176	631	-0,0005677802
204	0,0434768782	418	-0,0447806821	632	-0,0005466565
205	0,0401458278	419	-0,0382776572	633	-0,0005226564
206	0,0366418116	420	-0,0319531274	634	-0,0005040714
207	0,0329583930	421	-0,0258227288	635	-0,0004893791
208	0,0290824006	422	-0,0198834129	636	-0,0004875227
209	0,0250307561	423	-0,0141288827	637	-0,0004947518
210	0,0207997072	424	-0,0085711749	638	-0,0005617692
211	0,0163701258	425	-0,0032086896	639	-0,0005525280
212	0,0117623832	426	0,0019765601		
213	0,0069636862	427	0,0069636862		

Фільтр-прототип, $p_0(n)$, також можна отримати з таблиці 4 за допомогою однієї або більше математичних операцій, таких як округлення, субдискретизація, інтерполяція і проріджування.

5 Попри те, що налаштування інформації керування, яка стосується SBR, зазвичай не залежить від подробиць транспозиції (як обговорено раніше), в деяких варіантах здійснення певні елементи даних керування можуть передаватися одночасно в контейнері розширення eSBR (`bs_extension_id == EXTENSION_ID_ESBR`) з метою підвищення якості відновленого сигналу. Деякі з елементів, які передаються одночасно, можуть включати дані шумового порога (наприклад, масштабні коефіцієнти шумового порога і параметр, який вказує напрямок, у напрямку або частоти, або часу, різницевого кодування для кожного шумового порогу), дані зворотної фільтрації (наприклад, параметр, який вказує режим зворотної фільтрації, вибраний із відсутності зворотної фільтрації, низького рівня зворотної фільтрації, проміжного рівня зворотної фільтрації і сильного рівня зворотної фільтрації) і дані пропущених гармонік (наприклад, параметр, який вказує, чи варто додати синусоїду в конкретну смугу частот відновлюваного верхнього діапазону). Все ці елементи базуються на виконанні в кодері синтетичній емуляції транспозера декодера, і тому при правильному налаштуванні для вибраного транспозера можуть підвищувати якість відновленого сигналу.

10 Зокрема, в деяких варіантах здійснення дані пропущених гармонік і дані керування зворотної фільтрації передають у контейнері розширення eSBR (разом з іншими параметрами бітового потоку з таблиці 3) і налаштовують для гармонічного транспозера eSBR. Додаткова бітова швидкість, необхідна для передачі цих двох класів метаданих для гармонічного транспозера eSBR, є відносно низькою. Тому відправлення налаштованих даних пропущених гармонік й/або даних керування зворотної фільтрації в контейнері розширення eSBR буде підвищувати якість звуку, який отримується за допомогою транспозера, лише з мінімальним при цьому впливом на бітову швидкість. Для забезпечення зворотної сумісності з традиційними декодерами параметри, налаштовані для операції спектрального переносу SBR, також можуть бути відправлені в бітовому потоці як частина даних керування SBR з використанням або прихованої, або явної сигналізації.

30 Складність декодера з покращеннями SBR, як описано в даній заявці, повинна бути обмежена таким чином, щоб значно не підвищувати загальну обчислювальну складність реалізації. Переважно PCU (MOP) для типу об'єкта SBR, при використанні інструмента eSBR, є меншою або дорівнює 4,5, і RCU для типу об'єкта SBR, при використанні інструмента eSBR, є меншою або дорівнює 3. Наближена обчислювальна потужність дана в одиницях складності процесора (Processor Complexity Unit, PCU), виражених у цілочислових кількостях MOPS. Наближене завантаження RAM дано в одиницях складності RAM (RAM Complexity Units, RCU), виражених у цілочислових кількостях кілослів (1000 слів). Кількості RCU не включають робочі буфери, які можуть сумісно використовуватися різними об'єктами й/або каналами. Крім того, PCU є пропорційною частоті дискретизації. Значення PCU дані в MOPS (мільйони операцій у секунду) на канал, а значення RCU – в кіло словах на канал.

40 До стиснутих даних, таких як кодований звук HE-AAC, які можуть бути декодовані за допомогою різних конфігурацій декодера, необхідна особлива увага. У цьому випадку декодування можна здійснювати із забезпеченням зворотної сумісності (тільки AAC), а також у покращеному режимі (AAC+SBR). Якщо стиснуті дані допускають і зворотно сумісне, і розширене декодування, і якщо декодер діє в покращеному режимі таким чином, що використовує постпроцесор, який вносить деяку додаткову затримку (наприклад, постпроцесор

SBR в HE-AAC), то необхідно забезпечити, щоб ця додаткова часова затримка, що вноситься відносно режиму зворотної сумісності, як описано відповідним значенням n , враховувалась при представленні блока сумісності. Для забезпечення правильної обробки відміток часу сумісності (щоб звук залишався синхронізованим з іншими медіаданими), додаткова затримка, яка

5 внесена постобробкою і дана в кількості дискретних значень (на звуковий канал) на вихідній частоті дискретизації, дорівнює 3010, якщо режим роботи декодера включає покращення SBR (у тому числі eSBR), як описано в цій заявці. Тому для блока сумісності звуку час сумісності застосовується до 3011-го звукового дискретного значення в блоці сумісності, якщо режим роботи декодера включає покращення SBR, як описано в цій заявці.

10 Для підвищення суб'єктивної якості звукового вмісту з гармонічною частотною структурою і сильними тональними характеристиками, зокрема при низьких бітових швидкостях, слід активувати покращення SBR. Значення відповідного елемента бітового потоку (тобто `esbr_data()`), який керує цими інструментами, можуть бути визначені в кодері шляхом застосування механізму класифікації, залежного від сигналу.

15 Як правило, використання способу гармонічного вставлення (`sbrPatchingMode == 0`) є переважним для кодування музикальних сигналів із дуже низькими бітовими швидкостями, коли основний кодек може бути значно обмежений у смузі пропускання звуку. Це є особливо вірним, якщо ці сигнали мають виражену гармонічну структуру. Навпаки, використання звичайного способу вставлення SBR є переважним для мовних і змішаних сигналів, оскільки він забезпечує

20 краще зберігання тимчасової структури мови.

Для покращення роботи транспозера MPEG-4 SBR можна активувати етап попередньої обробки (`bs_sbr_preprocessing == 1`), який усуває внесення порушень однорідності спектра сигналу, який іде у подальший коректор обвідної. Робота інструмента є корисною для типів сигналів, у яких груба обвідна спектра сигналу нижнього діапазону, що використовується для

25 реконструкції високих частот, проявляє великі варіації рівня.

Для покращення перехідної характеристики гармонічного вставлення SBR (`sbrPatchingMode == 0`) можна застосовувати адаптивну до сигналу передискретизацію в частотній ділянці (`sbrOversamplingFlag == 1`). Оскільки адаптивна до сигналу передискретизація в частотній ділянці збільшує обчислювальну складність транспозера, але дає користь тільки для кадрів, які

30 містять перехідні сигнали, використанням цього інструмента керують за допомогою елемента бітового потоку, який передають один раз на кадр і на незалежний канал SBR.

Звичайні рекомендовані налаштування бітової швидкості для HE-AACv2 з покращеннями SBR (тобто із включенням гармонічного транспозера інструмента eSBR) відповідають 20–32 кбіт/с для стереофонічного звукового вмісту при частотах дискретизації 44,1 кГц або 48 кГц. Відносний суб'єктивний виграв у якості покращень SBR збільшується в напрямку до нижньої межі бітової швидкості, і належним чином виконаний декодер дозволяє розширити цей діапазон до ще більш низьких бітових швидкостей. Представлені вище бітові швидкості являють собою

35 лише рекомендації і можуть бути адаптовані для конкретних експлуатаційних вимог.

Декодеру, який діє в запропонованому режимі покращеного SBR, зазвичай необхідно мати можливість переключення між традиційним і покращеним вставленням SBR. Тому може вноситися затримка, яка може бути настільки тривалою, як тривалість одного базового кадру звуку, залежно від налаштувань декодера. Зазвичай затримка буде однаковою і для традиційного, і для покращеного вставлення SBR.

Необхідно розуміти, що в межах об'єму прикладеної формули винаходу даний винахід може бути реалізований на практиці інакше, у порівнянні з тим, що конкретно описано в даному документі. Будь-які посилальні позначення, що містяться в представленій нижче формулі винаходу, наведені тільки для ілюстрації, і їх ніяким чином не слід використовувати для тлумачення або обмеження формули винаходу.

45

Різні аспекти даного винаходу можна зрозуміти з наступних пронумерованих наведених як

50 приклад варіантів здійснення (ПНВЗ).

ПНВЗ 1. Спосіб виконання реконструкції високих частот звукового сигналу, який включає:

приймання кодованого звукового бітового потоку, при цьому кодований звуковий бітовий потік містить звукові дані, які представляють частину, що стосується нижнього діапазону, звукового сигналу, і метадані реконструкції високих частот;

55 декодування звукових даних для генерування декодованого звукового сигналу нижнього діапазону;

вилучення з кодованого звукового бітового потоку метаданих реконструкції високих частот, при цьому метадані реконструкції високих частот містять робочі параметри для процесу реконструкції високих частот, робочі параметри включають параметр режиму вставлення,

60 розташований у зворотно сумісному контейнері розширення кодованого звукового бітового

поток, при цьому перше значення параметра режиму вставляння вказує на спектральний перенос, а друге значення параметра режиму вставляння вказує на гармонічну транспозицію з використанням розтягнення діапазону частот фазовим вокодером;

5 фільтрацію декодованого звукового сигналу нижнього діапазону для генерування фільтрованого звукового сигналу нижнього діапазону;

відновлення частини, що стосується верхнього діапазону, звукового сигналу з використанням фільтрованого звукового сигналу нижнього діапазону і метаданих реконструкції високих частот, при цьому відновлення включає спектральний перенос, якщо параметр режиму вставляння має перше значення, і відновлення включає гармонічну транспозицію з використанням розтягнення діапазону частот фазовим вокодером, якщо параметр режиму вставляння має друге значення; і

об'єднання фільтрованого звукового сигналу нижнього діапазону з відновленою частиною, що стосується верхнього діапазону, для формування широкосмугового звукового сигналу,

10 при цьому фільтрацію, відновлення й об'єднання виконують як операцію постобробки із затримкою в 3010 дискретних значень на звуковий канал або менше.

ПНВЗ 2. Спосіб відповідно до ПНВЗ 1, в якому кодований звуковий бітовий потік додатково включає заповнювальний елемент з ідентифікатором, який вказує початок заповнювального елемента, і заповнювальними даними після ідентифікатора, при цьому заповнювальні дані містять зворотно сумісний контейнер розширення.

20 ПНВЗ 3. Спосіб відповідно до ПНВЗ 2, в якому ідентифікатор являє собою трибітне ціле число без знака, у якого спочатку передають старший значущий біт, що має значення 0×6 .

ПНВЗ 4. Спосіб відповідно до ПНВЗ 2 або ПНВЗ 3, в якому заповнювальні дані містять корисні дані розширення, при цьому корисні дані розширення містять дані розширення копіювання спектральної смуги, і корисні дані розширення ідентифікують за допомогою 25 чотирибітного цілого числа без знака, у якого спочатку передають старший значущий біт, що має значення "1101" або "1110", і необов'язково

при цьому дані розширення копіювання спектральної смуги містять:

необов'язковий заголовок копіювання спектральної смуги,

дані копіювання спектральної смуги після заголовка та

30 елемент розширення копіювання спектральної смуги після даних копіювання спектральної смуги, і при цьому в елемент розширення копіювання спектральної смуги включають прапор.

ПНВЗ 5. Спосіб відповідно до будь-якого з ПНВЗ 1–4, в якому метадані реконструкції високих частот містять масштабні коефіцієнти обвідної, масштабні коефіцієнти шумового порога, інформацію часової/частотної сітки або параметр, який вказує частоту розділення.

35 ПНВЗ 6. Спосіб відповідно до будь-якого з ПНВЗ 1–5, в якому зворотно сумісний контейнер розширення додатково містить прапор, який вказує, чи використовується додаткова попередня обробка для уникання порушень безперервності у формі обвідної спектра частини, що стосується верхнього діапазону, коли параметр режиму вставляння дорівнює першому значенню, при цьому перше значення прапора включає додаткову попередню обробку, а друге значення прапора відключає додаткову попередню обробку.

40 ПНВЗ 7. Спосіб відповідно до ПНВЗ 6, в якому додаткова попередня обробка включає обчислення кривої попереднього посилення з використанням коефіцієнта фільтра лінійного передбачення.

45 ПНВЗ 8. Спосіб відповідно до будь-якого з ПНВЗ 1–5, в якому зворотно сумісний контейнер розширення додатково містить прапор, який вказує, чи необхідно застосовувати адаптивну до сигналу передискретизацію в частотній ділянці, коли параметр режиму вставляння дорівнює другому значенню, при цьому перше значення прапора включає адаптивну до сигналу передискретизацію в частотній ділянці, а друге значення прапора відключає адаптивну до сигналу передискретизацію в частотній ділянці.

50 ПНВЗ 9. Спосіб відповідно до ПНВЗ 8, в якому адаптивну до сигналу передискретизацію в частотній ділянці застосовують тільки для кадрів, які містять перехідний сигнал.

ПНВЗ 10. Спосіб відповідно до будь-якого з попередніх ПНВЗ, в якому гармонічну транспозицію з використанням розтягнення діапазону частот фазовим вокодером виконують з оцінною складністю 4,5 мільйона або менше операцій у секунду і 3 або менше кілобайт пам'яті.

55 ПНВЗ 11. Енергонезалежний машинопрочитуваний носій, який містить команди, які при здійсненні процесором виконують спосіб відповідно до будь-якого з ПНВЗ 1–10.

ПНВЗ 12. Комп'ютерний програмний продукт, який містить команди, які при здійсненні обчислювальним пристроєм або системою викликають здійснення вказаним обчислювальним пристроєм або системою способу відповідно до будь-якого з ПНВЗ 1–10.

60 ПНВЗ 13. Блок обробки звуку для виконання реконструкції високих частот звукового сигналу,

який містить:

інтерфейс введення для приймання кодованого звукового бітового потоку, при цьому кодований звуковий бітовий потік містить звукові дані, які представляють частину, що стосується нижнього діапазону, звукового сигналу, і метадані реконструкції високих частот;

5 основний декодер звуку для декодування звукових даних для генерування декодованого звукового сигналу нижнього діапазону;

пристрій видалення форматування для вилучення з кодованого звукового бітового потоку метаданих реконструкції високих частот, при цьому метадані реконструкції високих частот містять робочі параметри для процесу реконструкції високих частот, робочі параметри включають параметр режиму вставляння, розташований у зворотно сумісному контейнері розширення кодованого звукового бітового потоку, при цьому перше значення параметра режиму вставляння вказує на спектральний перенос, а друге значення параметра режиму вставляння вказує на гармонічну транспозицію з використанням розтягнення діапазону частот фазовим вокодером;

15 блок фільтрів аналізу для фільтрації декодованого звукового сигналу нижнього діапазону для генерування фільтрованого звукового сигналу нижнього діапазону;

пристрій відновлення високих частот для реконструкції частини, що стосується верхнього діапазону, звукового сигналу з використанням фільтрованого звукового сигналу нижнього діапазону і метаданих реконструкції високих частот, при цьому реконструкція включає спектральний перенос, якщо параметр режиму вставляння має перше значення, і реконструкція включає гармонічну транспозицію з використанням розтягнення діапазону частот фазовим вокодером, якщо параметр режиму вставляння має друге значення; і

20 блок синтезувальних фільтрів для об'єднання фільтрованого звукового сигналу нижнього діапазону з відновленої частиною, що стосується верхнього діапазону, для формування широкосмугового звукового сигналу,

при цьому блок фільтрів аналізу, пристрій відновлення високих частот і блок синтезувальних фільтрів виконуються в постпроцесорі із затримкою в 3010 дискретних значень на канал звуку або менше.

30 ПНВЗ 14. Блок обробки звуку відповідно до ПНВЗ 13, в якому гармонічна транспозиція з використанням розтягнення діапазону частот фазовим вокодером виконується з оцінкою складністю 4,5 мільйона або менше операцій у секунду і 3 або менше кілобайт пам'яті.

ФОРМУЛА ВИНАХОДУ

35 1. Спосіб виконання реконструкції високих частот звукового сигналу, який включає:

приймання кодованого звукового бітового потоку, при цьому кодований звуковий бітовий потік містить звукові дані, які представляють частину, що стосується нижнього діапазону, звукового сигналу, і метадані реконструкції високих частот;

40 декодування звукових даних для генерування декодованого звукового сигналу нижнього діапазону;

вилучення з кодованого звукового бітового потоку метаданих реконструкції високих частот, при цьому метадані реконструкції високих частот містять робочі параметри для процесу реконструкції високих частот, робочі параметри включають параметр режиму вставляння, розташований у зворотно-сумісному контейнері розширення кодованого звукового бітового потоку, при цьому перше значення параметра режиму вставляння вказує на спектральний перенос, а друге значення параметра режиму вставляння вказує на гармонічну транспозицію з використанням розтягнення діапазону частот фазовим вокодером;

45 фільтрацію декодованого звукового сигналу нижнього діапазону в множину піддіапазонів із використанням блока комплексних QMF-фільтрів аналізу для генерування фільтрованого звукового сигналу нижнього діапазону;

відновлення частини, що стосується верхнього діапазону, звукового сигналу з використанням фільтрованого звукового сигналу нижнього діапазону і метаданих реконструкції високих частот, при цьому відновлення включає спектральний перенос, якщо параметр режиму вставляння має перше значення, і відновлення включає гармонічну транспозицію з використанням розтягнення

55 діапазону частот фазовим вокодером, якщо параметр режиму вставляння має друге значення; і об'єднання фільтрованого звукового сигналу нижнього діапазону з відновленою частиною, що стосується верхнього діапазону, із використанням блока комплексних синтезувальних QMF-фільтрів для формування широкосмугового звукового сигналу,

при цьому фільтри аналізу $h_k(n)$ блока комплексних QMF-фільтрів аналізу і синтезувальні

60 фільтри $f_k(n)$ блока комплексних синтезувальних QMF-фільтрів визначають за формулою:

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

де $p_0(n)$ являє собою дійснозначний фільтр-прототип, M означає кількість каналів, а N являє собою порядок фільтра-прототипу; і

при цьому фільтрацію, відновлення й об'єднання виконують як операцію постобробки із затримкою в 3010 дискретних значень на звуковий канал.

2. Спосіб за п. 1, який **відрізняється** тим, що кодований звуковий бітовий потік додатково включає заповнювальний елемент з ідентифікатором, який вказує початок заповнювального елемента, і заповнювальними даними після ідентифікатора, при цьому заповнювальні дані містять зворотно-сумісний контейнер розширення.

3. Спосіб за п. 2, який **відрізняється** тим, що ідентифікатор являє собою трибітне ціле число без знака, у якого спочатку передають старший значущий біт, що має значення 0×6.

4. Спосіб за п. 2 або 3, який **відрізняється** тим, що заповнювальні дані містять корисні дані розширення, при цьому корисні дані розширення містять дані розширення копіювання спектральної смуги і корисні дані розширення ідентифікують за допомогою чотирибітного цілого числа без знака, у якого спочатку передають старший значущий біт, що має значення "1101" або "1110", і необов'язково

при цьому дані розширення копіювання спектральної смуги містять:

необов'язковий заголовок копіювання спектральної смуги,

дані копіювання спектральної смуги після заголовка, та

елемент розширення копіювання спектральної смуги після даних копіювання спектральної смуги, і при цьому в елемент розширення копіювання спектральної смуги включають прапор.

5. Спосіб за будь-яким із пп. 1-4, який **відрізняється** тим, що метадані реконструкції високих частот містять масштабні коефіцієнти обвідної, масштабні коефіцієнти шумового порога, інформацію часової/частотної сітки або параметр, який вказує частоту розділення.

6. Спосіб за будь-яким із пп. 1-5, який **відрізняється** тим, що зворотно-сумісний контейнер розширення додатково містить прапор, який вказує, чи використовується додаткова попередня обробка для уникання порушень безперервності у формі обвідної спектра частини, що стосується верхнього діапазону, коли параметр режиму вставляння дорівнює першому значенню, при цьому перше значення прапора включає додаткову попередню обробку, а друге значення прапора відключає додаткову попередню обробку.

7. Спосіб за п. 6, який **відрізняється** тим, що додаткова попередня обробка включає обчислення кривої попереднього посилення з використанням коефіцієнта фільтра лінійного передбачення.

8. Спосіб за будь-яким із пп. 1-5, який **відрізняється** тим, що зворотно-сумісний контейнер розширення додатково містить прапор, який вказує, чи необхідно застосовувати адаптивну до сигналу передискретизацію в частотній ділянці, коли параметр режиму вставляння дорівнює другому значенню, при цьому перше значення прапора включає адаптивну до сигналу передискретизацію в частотній ділянці, а друге значення прапора відключає адаптивну до сигналу передискретизацію в частотній ділянці.

9. Спосіб за п. 8, який **відрізняється** тим, що адаптивну до сигналу передискретизацію в частотній ділянці застосовують тільки для кадрів, які містять перехідний сигнал.

10. Спосіб за будь-яким із попередніх пунктів, який **відрізняється** тим, що гармонічну транспозицію з використанням розтягнення діапазону частот фазовим вокодером виконують з оцінною складністю 4,5 мільйона або менше операцій у секунду і 3 або менше кілобайт пам'яті.

11. Машинозчитуваний носій інформації, який містить команди, які при здійсненні обчислювальним пристроєм або системою викликають здійснення вказаним обчислювальним пристроєм або системою способу за будь-яким із пп. 1-10.

12. Блок обробки звуку для виконання реконструкції високих частот звукового сигналу, який містить:

інтерфейс введення для приймання кодованого звукового бітового потоку, при цьому кодований звуковий бітовий потік містить звукові дані, які представляють частину, що стосується нижнього діапазону, звукового сигналу, і метадані реконструкції високих частот;

основний декодер звуку для декодування звукових даних для генерування декодованого звукового сигналу нижнього діапазону;

пристрій видалення форматування для вилучення з кодованого звукового бітового потоку метаданих реконструкції високих частот, при цьому метадані реконструкції високих частот містять робочі параметри для процесу реконструкції високих частот, робочі параметри включають параметр режиму вставляння, розташований у зворотно-сумісному контейнері

розширення кодованого звукового бітового потоку, при цьому перше значення параметра режиму вставляння вказує на спектральний перенос, а друге значення параметра режиму вставляння вказує на гармонічну транспозицію з використанням розтягнення діапазону частот фазовим вокодером;

5 блок фільтрів аналізу для фільтрації декодованого звукового сигналу нижнього діапазону в множині піддіапазонів із використанням блока комплексних QMF-фільтрів аналізу для генерування фільтрованого звукового сигналу нижнього діапазону;

пристрій відновлення високих частот для реконструкції частини, що стосується верхнього діапазону, звукового сигналу з використанням фільтрованого звукового сигналу нижнього діапазону і метаданих реконструкції високих частот, при цьому реконструкція включає спектральний перенос, якщо параметр режиму вставляння має перше значення, і реконструкція включає гармонічну транспозицію з використанням розтягнення діапазону частот фазовим вокодером, якщо параметр режиму вставляння має друге значення; і

10 блок синтезувальних фільтрів для об'єднання фільтрованого звукового сигналу нижнього діапазону з відновленою частиною, що стосується верхнього діапазону, із використанням блока комплексних синтезувальних QMF-фільтрів для формування широкосмугового звукового сигналу,

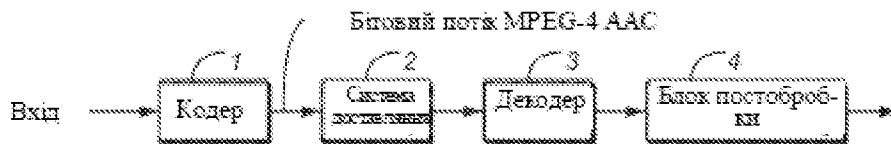
при цьому блок фільтрів аналізу, пристрій відновлення високих частот і блок синтезувальних фільтрів виконуються в постпроцесорі із затримкою в 3010 дискретних значень на канал звуку; і

20 при цьому фільтри аналізу $h_k(n)$ блока комплексних QMF-фільтрів аналізу і синтезувальні фільтри $f_k(n)$ блока комплексних синтезувальних QMF-фільтрів визначають за формулою:

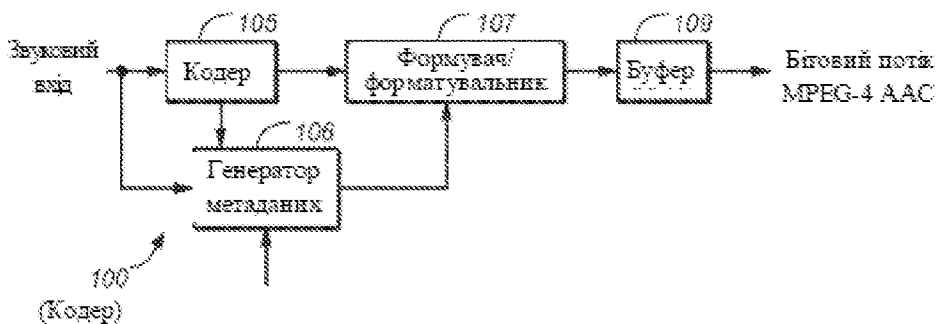
$$h_k(n) = f_k(n) = p_0(n) \exp\left\{i \frac{\pi}{M} \left(k + \frac{1}{2}\right) \left(n - \frac{1}{2}\right)\right\}, 0 \leq n \leq N; 0 \leq k < M$$

де $p_0(n)$ являє собою дійснозначний фільтр-прототип, M означає кількість каналів, а N являє собою порядок фільтра-прототипу.

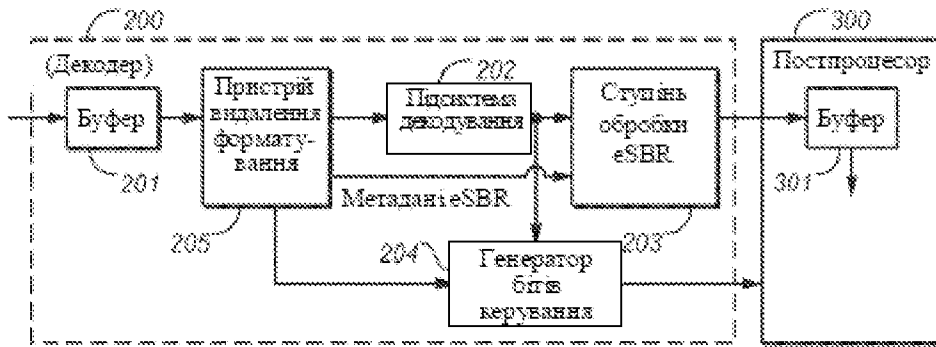
25 13. Блок обробки звуку за п. 12, який **відрізняється** тим, що гармонічна транспозиція з використанням розтягнення діапазону частот фазовим вокодером виконується з оцінкою складності 4,5 мільйона або менше операцій у секунду і 3 або менше кілобайт пам'яті.



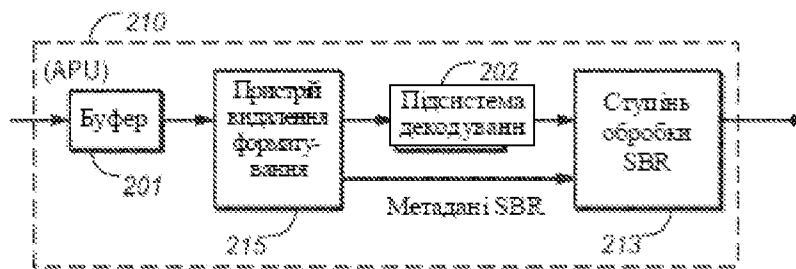
Фіг. 1



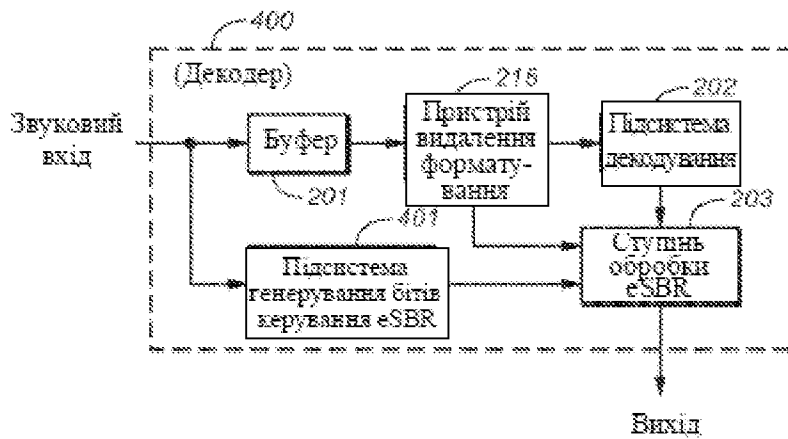
Фіг. 2



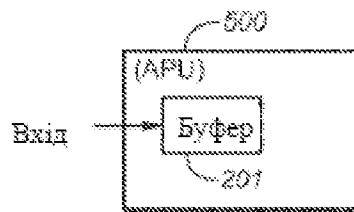
Фіг. 3



Фіг. 4



Фіг. 5



Фіг. 6

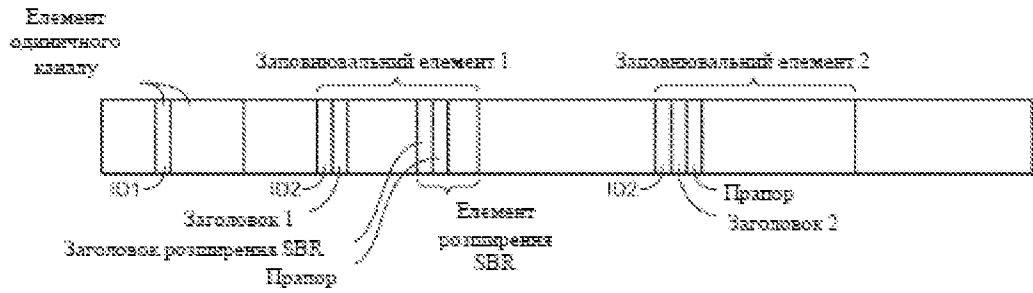


Fig. 7