

ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(21)(22) Заявка: 2012137189/08, 26.01.2011

(24) Дата начала отсчета срока действия патента:
26.01.2011

Приоритет(ы):

(30) Конвенционный приоритет:
02.02.2010 ЕР 10152388.4

(43) Дата публикации заявки: 10.03.2014 Бюл. № 7

(45) Опубликовано: 10.08.2015 Бюл. № 22

(56) Список документов, цитированных в отчете о
поиске: US 2009136048 A1, 28.05.2009. WO
2006027717 A1, 16.03.2006. WO 0207481 A2,
24.01.2002. WO 2006056910 A1, 01.06.2006. (см.
прод.)(85) Дата начала рассмотрения заявки РСТ на
национальной фазе: 03.09.2012(86) Заявка РСТ:
IB 2011/050334 (26.01.2011)(87) Публикация заявки РСТ:
WO 2011/095913 (11.08.2011)

Адрес для переписки:

129090, Москва, ул. Б. Спасская, 25, строение 3,
ООО "Юридическая фирма Городисский и
Партнеры"

(72) Автор(ы):

ХЯРМЯ Аки Сакари (NL),
ДЕ БРЮЭЙН Вернер Паулус Йосефус (NL)

(73) Патентообладатель(и):

КОНИКЛЕЙКЕ ФИЛИПС
ЭЛЕКТРОНИКС Н.В. (NL)

R U 2 5 5 9 7 1 3 C 2

(54) ПРОСТРАНСТВЕННОЕ ВОСПРОИЗВЕДЕНИЕ ЗВУКА

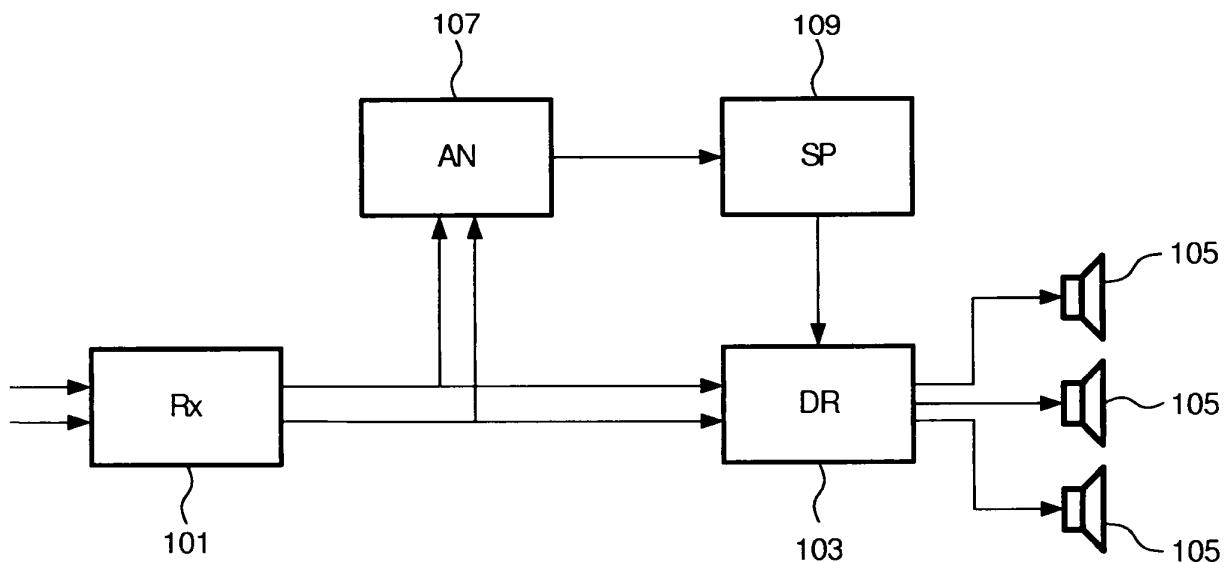
(57) Реферат:

Изобретение относится к средствам для пространственного воспроизведения звука. Технический результат заключается в улучшении пространственного восприятия при прослушивании. Устройство для пространственного воспроизведения звука содержит приемное устройство для приема многоканального аудиосигнала. Анализатор определяет пространственное свойство многоканального аудиосигнала, такое как пространственная сложность или организация. Процессор выбора затем выбирает режим воспроизведения из множества режимов

воспроизведения звука, при этом режимы многоканального воспроизведения звука используют различные технологии пространственного рендеринга. Схема воспроизведения затем возбуждает набор громкоговорителей, чтобы воспроизводить многоканальный аудиосигнал с использованием выбранного режима воспроизведения. Переключение между режимами воспроизведения может быть быстрым (например, в порядке от 100 мс до 10 с), тем самым обеспечивая краткосрочную адаптацию режима воспроизведения к характеристикам сигналов.

Подход может, в частности, предоставлять улучшенное пространственное восприятие для

слушателя. 2 н. и 12 з.п. ф-лы, 3 ил.



ФИГ. 1

(56) (продолжение):

WO 0149074 A2, 05.07.2001. US 2008160943 A1, 03.07.2008. US 6198827 B1, 06.03.2001. RU 2145446 C1, 10.02.2000.

FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(12) ABSTRACT OF INVENTION

(21)(22) Application: 2012137189/08, 26.01.2011

(24) Effective date for property rights:
26.01.2011

Priority:

(30) Convention priority:
02.02.2010 EP 10152388.4

(43) Application published: 10.03.2014 Bull. № 7

(45) Date of publication: 10.08.2015 Bull. № 22

(85) Commencement of national phase: 03.09.2012

(86) PCT application:
IB 2011/050334 (26.01.2011)(87) PCT publication:
WO 2011/095913 (11.08.2011)

Mail address:

129090, Moskva, ul. B. Spasskaja, 25, stroenie 3,
OOO "Juridicheskaja firma Gorodisskij Partnery"

(72) Inventor(s):

KhJaRMJa Aki Sakari (NL),
DE BRJuEhJN Verner Paulus Josefus (NL)

(73) Proprietor(s):

KONINKLEJKE FILIPS EhLEKTRONIKS
N.V. (NL)

R U 2 5 5 9 7 1 3 C 2

(54) SPATIAL REPRODUCTION OF SOUND

(57) Abstract:

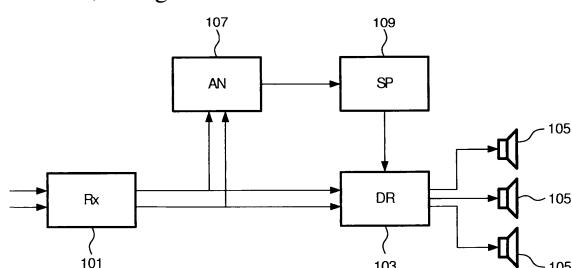
FIELD: physics, acoustics.

SUBSTANCE: invention relates to means for spatial reproduction of sound. An apparatus for spatial reproduction of sound has a receiving device for receiving a multichannel audio signal. An analyser determines the spatial property of the multichannel audio signal, such as space complexity or organisation. A selection process then selects a reproduction mode from a plurality of sound reproduction modes, wherein multichannel sound reproduction modes employ different spatial rendering techniques. A reproduction circuit then drives a set of loudspeakers in order to reproduce the multichannel audio signal using the selected reproduction mode. Switching between reproduction modes can be fast (for example, in the order of 100 ms to 10 s), thereby facilitating short-term

adaptation of the reproduction mode to signal characteristics. The approach can particularly improve spatial perception for the listener.

EFFECT: improved spatial perception when listening.

14 cl, 3 dwg



Фиг.1

Область техники, к которой относится изобретение

Изобретение относится к пространственному воспроизведению звука и, в частности, но не только, к пространственному воспроизведению звука, включающему в себя микширование с увеличением частоты многоканального аудиосигнала.

5 Уровень техники

Пространственное воспроизведение звука в форме стереозаписей и стереовоспроизведения известно в течение нескольких десятилетий. В последние десятилетия используются более совершенные компоновки и обработка сигналов для того, чтобы предоставлять улучшенное пространственное восприятие при прослушивании. В частности, стало распространенным использование объемного звука с использованием, например, 5 или 7 пространственных динамиков, чтобы предоставлять расширенное восприятие в связи, например, с просмотром фильмов или телевизионных передач. Помимо этого, компактные системы громкоговорителей с несколькими режимами возбуждения, к примеру, "звуковые жалюзи" стали популярным вариантом для традиционных стереосистем и систем 5.1. Эти устройства предоставляют восприятие широкой пространственной аудиокартины для слушателя даже из небольшого устройства. Это основано на цифровой обработке сигналов и специальной физической компоновке устройства.

Пространственная звуковая обработка все в большей степени использует 20 усовершенствованную обработку сигналов в качестве части воспроизведения звука, чтобы предоставлять улучшенное пространственное восприятие. Например, сложные алгоритмы могут быть использованы для того, чтобы микшировать с увеличением частоты аудиосигнала к более высокому числу каналов. Например, 5-канальный сигнал объемного звучания на передающей стороне может быть смикширован с понижением 25 частоты к стерео- или моносигналу. Этот сигнал затем распределяется, и воспроизведение звука включает в себя микширование с увеличением частоты принимаемого сигнала к исходному 5-канальному сигналу.

В качестве другого примера, обработка сигналов может быть использована для 30 того, чтобы предоставлять эффект расширения звука в стереосигнал, приводящий к восприятию слушателем более широкой звуковой сцены. Типично способы основаны на операциях обработки сигналов, которые уменьшают корреляцию между каналами. Эти технологии, в частности, являются популярными в компактных упомянутых выше акустических системах.

В качестве другого примера, воспроизведение пространственного сигнала может 35 включать в себя извлечение доминирующего источника звука, например, в стереосигнале. Оставшийся остаточный сигнал типично должен соответствовать окружающему стереоизображению, которое является более рассеянным. Доминирующий сигнал и окружающий сигнал затем могут быть воспроизведены по-разному, так что характеристики воспроизведения оптимизируются для каждого сигнала.

40 Тем не менее, хотя такие технологии пространственного воспроизведения звука улучшают восприятие при прослушивании, зачастую возникают некоторые ассоциированные недостатки. В частности, воспроизведение может не предоставлять оптимальное пространственное восприятие во всех случаях, и обработка сигналов в некоторых случаях может фактически приводить к ухудшенному пространственному восприятию.

Следовательно, улучшенная система для пространственного воспроизведения звука имеет преимущество, в частности, имеет преимущество система, предоставляющая повышенную гибкость, упрощенную работу, упрощенную реализацию, улучшенное

пространственное восприятие при прослушивании и/или повышенную производительность.

Сущность изобретения

Следовательно, изобретение предпочтительно нацелено на уменьшение, облегчение

или устранение одного или более вышеуказанных недостатков по отдельности или в любом сочетании.

Согласно аспекту изобретения, предусмотрена устройство для пространственного воспроизведения звука, причем устройство содержит: приемное устройство для приема многоканального аудиосигнала; схему для определения пространственного свойства

многоканального аудиосигнала; схему для выбора выбранного режима воспроизведения из множества режимов воспроизведения звука в ответ на пространственное свойство, причем режимы многоканального воспроизведения звука используют различные технологии пространственного рендеринга; и схему воспроизведения для возбуждения набора пространственных каналов, предоставленных посредством набора

громкоговорителей, чтобы воспроизводить многоканальный аудиосигнал с использованием выбранного режима воспроизведения; при этом множество режимов воспроизведения звука содержит, по меньшей мере, два из следующего: режим монофонического воспроизведения; режим воспроизведения, поддерживающий пространственные характеристики многоканального сигнала; режим воспроизведения, содержащий обработку пространственного расширения; и режим воспроизведения, содержащий разделение, по меньшей мере, на один доминирующий исходный сигнал и окружающий сигнал и применение различного пространственного воспроизведения, по меньшей мере, одного первичного исходного сигнала и окружающего сигнала.

Изобретение может предоставлять улучшенное воспроизведение звука во многих вариантах осуществления. В частности, улучшенное пространственное восприятие может предоставляться во многих сценариях. Типично, пространственное воспроизведение может быть улучшено для конкретного аудиосигнала. Подход дополнительно может обеспечивать реализацию с низкой сложностью и упрощенную работу во многих вариантах осуществления.

Выбор надлежащего способа воспроизведения может быть оптимизирован для испытываемых конкретных условий при поддержании низкой сложности.

Пространственное свойство может служить признаком пространственной организации и/или пространственной сложности сигнала. Например, пространственное свойство может служить признаком присутствия одного или более доминирующих источников звука в соответствии с подходящим критерием или процессом для извлечения доминирующих источников звука. В некоторых вариантах осуществления, пространственное свойство может служить признаком пространственного распределения источников звука в звуковом изображении, предоставленном посредством многоканального сигнала.

Набором громкоговорителей, в частности, могут быть громкоговорителями установки объемного звука, содержащей, например, 3, 5 или 7 пространственных динамиков (в дополнение, возможно, к динамику с непространственным низкочастотным эффектом или сабвуферу). Набором громкоговорителей могут быть системы громкоговорителей с несколькими режимами возбуждения типично с тремя или более по отдельности возбуждаемыми громкоговорителями (или массивами громкоговорителей) в одном физическом устройстве. Набор громкоговорителей также может содержать множество таких устройств.

В соответствии с необязательным признаком изобретения, по меньшей мере, один

из режимов воспроизведения звука содержит, по меньшей мере, одно из следующего: микширование с увеличением частоты к более высокому числу пространственных каналов, чем число каналов многоканального аудиосигнала; и микширование с понижением частоты к меньшему числу пространственных каналов, чем число каналов многоканального аудиосигнала.

Изобретение может предоставлять улучшенное пространственное восприятие.

Например, некоторые звуковые изображения стереосигнала могут предоставлять улучшенное пространственное восприятие при воспроизведении в качестве моносигнала. Другие звуковые изображения стереосигнала могут предоставлять улучшенное

пространственное восприятие при воспроизведении в качестве расширенного стереосигнала в комбинации с центральным сигналом, т.е. при воспроизведении с использованием трех пространственных каналов.

В соответствии с необязательным признаком изобретения, набор пространственных каналов содержит различное число каналов, чем многоканальный аудиосигнал.

Изобретение может предоставлять улучшенное пространственное восприятие для системы воспроизведения звука и может, в частности, обеспечивать дополнительные степени свободы в адаптации воспроизведения звука к конкретному звуковому изображению и пространственным характеристикам.

В соответствии с необязательным признаком изобретения, максимальная частота переключения для переключения между режимами воспроизведения звука превышает 1 Гц.

Это позволяет предоставлять динамическую адаптацию и оптимизацию, которая может близко совпадать с варьирующими характеристиками аудио, тем самым предоставляя улучшенное восприятие при прослушивании.

Признак может обеспечивать повышенную производительность и улучшенную адаптацию режима воспроизведения к аудиосигналу, тем самым предоставляя расширенное восприятие при прослушивании. Подход может обеспечивать краткосрочную адаптацию воспроизведения к характеристикам сигналов.

В некоторых вариантах осуществления, схема для определения пространственного свойства преимущественно может быть выполнена с возможностью определять пространственное свойство с постоянной времени меньше 500 секунд, 100 секунд, 1 секунды, 500 мс, 100 мс или даже 50 мс.

Постоянная времени представляет время, которое требуется для пространственного свойства, чтобы достигать $1-1/e \approx 63\%$ от своего конечного (асимптотического) значения после ступенчатого изменения.

В некоторых вариантах осуществления, схема для определения пространственного свойства выполнена с возможностью включать в себя фильтрацию низких частот пространственного свойства, причем фильтрация низких частот имеет частоту отсечки при 3 дБ, превышающую 0,001 Гц, 0,01 Гц, 0,1 Гц, 1 Гц, 10 Гц или 50 Гц.

Множество режимов воспроизведения звука содержит, по меньшей мере, два из следующего: режим монофонического воспроизведения; режим воспроизведения, поддерживающий пространственные характеристики многоканального сигнала; режим воспроизведения, содержащий обработку пространственного расширения; и режим воспроизведения, содержащий разделение, по меньшей мере, на один доминирующий исходный сигнал и окружающий сигнал и применение различного пространственного воспроизведения, по меньшей мере, одного первичного исходного сигнала и окружающего сигнала.

Эти технологии воспроизведения могут быть, в частности, преимущественными и

подходящими для того, чтобы предоставлять улучшенные характеристики прослушивания для различных характеристик аудио. Во многих вариантах осуществления, множество режимов воспроизведения звука преимущественно может содержать два, три, или все четыре режима воспроизведения, поскольку они, в частности,

5 подходит для различных характеристик и тем самым вместе предоставляют набор режимов, которые предоставляют улучшенное воспроизведение для большого спектра характеристик аудио. Технологии могут, в частности, вместе предоставлять подходящие характеристики воспроизведения для широкого диапазона аудиосигналов.

В соответствии с необязательным признаком изобретения, устройство дополнительно 10 содержит: схему для определения характеристики содержимого для многоканального аудиосигнала; при этом схема для выбора выполнена с возможностью дополнительно выбирать выбранный алгоритм воспроизведения в ответ на характеристику содержимого.

Это позволяет дополнительно улучшать адаптацию воспроизведения и предоставлять 15 улучшенное пространственное восприятие во многих вариантах осуществления. Характеристика содержимого, например, может быть определена посредством анализа содержимого многоканального аудиосигнала и/или ассоциированного видеосигнала.

В соответствии с необязательным признаком изобретения, схема для определения 20 характеристики содержимого выполнена с возможностью определять характеристику содержимого в ответ на метаданные, ассоциированные с многоканальным аудиосигналом.

Это позволяет предоставлять, в частности, точный подход с низкой сложностью, который может быть преимущественным во многих вариантах осуществления.

В соответствии с необязательным признаком изобретения, схема для воспроизведения 25 многоканального аудиосигнала выполнена с возможностью адаптировать характеристику технологии пространственного рендеринга выбранного режима воспроизведения в ответ на характеристику содержимого.

Это позволяет дополнительно улучшать адаптацию воспроизведения и предоставлять улучшенное пространственное восприятие во многих вариантах осуществления.

30 В соответствии с необязательным признаком изобретения, схема для воспроизведения многоканального аудиосигнала выполнена с возможностью адаптировать характеристику технологии пространственного рендеринга выбранного режима воспроизведения в ответ на пространственное свойство.

Это позволяет дополнительно улучшать адаптацию воспроизведения и предоставлять 35 улучшенное пространственное восприятие во многих вариантах осуществления.

В соответствии с необязательным признаком изобретения, характеристикой пространственной обработки является степень пространственного расширения, применяемая, по меньшей мере, к двум каналам многоканального аудиосигнала.

40 Это позволяет предоставлять, в частности, преимущественную оптимизацию, поскольку пространственное расширение позволяет предоставлять значительно расширенное пространственное восприятие для некоторых характеристик аудио, но может ухудшать пространственное восприятие для других характеристик аудио. Соответственно, оптимизация пространственного расширения к характеристикам аудио позволяет предоставлять, в частности, преимущественную производительность.

45 В соответствии с необязательным признаком изобретения, схема для воспроизведения многоканального аудиосигнала выполнена с возможностью постепенно переходить от первого выбранного алгоритма воспроизведения ко второму выбранному алгоритму воспроизведения.

Это позволяет предоставлять повышенную производительность и позволяет, в частности, уменьшать заметность изменения между различными режимами воспроизведения. Устройство, в частности, может быть выполнено с возможностью, во время интервала перехода, формировать сигналы возбуждения для набора

5 громкоговорителей с использованием, как первого выбранного алгоритма воспроизведения, так и второго выбранного алгоритма воспроизведения и возбуждать набор громкоговорителей посредством сигналов, сформированных в качестве взвешенной комбинации сигналов возбуждения, причем взвешивание динамически изменяется во время интервала перехода.

10 В соответствии с необязательным признаком изобретения, схема для определения пространственного свойства выполнена с возможностью определять пространственное свойство в ответ на энергетический индикатор для комбинированного сигнала, по меньшей мере, двух каналов многоканального аудиосигнала относительно энергетического индикатора для разностного сигнала, по меньшей мере, двух каналов.

15 Это может быть, в частности, преимущественным пространственным свойством для адаптации пространственного воспроизведения. В частности, это позволяет предоставлять преимущественный компромисс между точностью и сложностью для многих сценариев.

20 В соответствии с необязательным признаком изобретения, схема для определения пространственного свойства выполнена с возможностью раскладывать многоканальный аудиосигнал, по меньшей мере, на один сигнал доминирующего источника звука и остаточный сигнал и определять пространственное свойство в ответ на энергетический индикатор для сигнала доминирующего источника звука относительно энергетического индикатора для остаточного сигнала.

25 Это может быть, в частности, преимущественным пространственным свойством для адаптации пространственного воспроизведения. В частности, это позволяет предоставлять преимущественный компромисс между точностью и сложностью для многих сценариев.

Согласно аспекту изобретения, предусмотрен способ пространственного воспроизведения звука, при этом способ содержит: прием многоканального аудиосигнала; определение пространственного свойства многоканального аудиосигнала; выбор выбранного режима воспроизведения из множества режимов воспроизведения звука в ответ на пространственное свойство, причем режимы многоканального воспроизведения звука используют различные технологии пространственного 35 рендеринга; и возбуждение набора громкоговорителей, чтобы воспроизводить многоканальный аудиосигнал с использованием выбранного режима воспроизведения; при этом множество режимов воспроизведения звука содержит, по меньшей мере, два из следующего: режим монофонического воспроизведения; режим воспроизведения, поддерживающий пространственные характеристики многоканального сигнала; режим 40 воспроизведения, содержащий обработку пространственного расширения; и режим воспроизведения, содержащий разделение, по меньшей мере, на один доминирующий исходный сигнал и окружающий сигнал и применение различного пространственного воспроизведения, по меньшей мере, одного первичного исходного сигнала и окружающего сигнала.

45 Эти и другие аспекты, признаки и преимущества изобретения должны становиться очевидными и должны истолковываться со ссылкой на описанные далее варианты осуществления.

Краткое описание чертежей

Варианты осуществления изобретения описаны далее только в качестве примера со ссылкой на чертежи, из которых:

Фиг.1 является иллюстрацией примера системы для пространственного воспроизведения звука в соответствии с некоторыми вариантами осуществления изобретения;

Фиг.2 является иллюстрацией примера элементов системы для пространственного воспроизведения звука в соответствии с некоторыми вариантами осуществления изобретения; и

Фиг.3 является иллюстрацией примера системы для пространственного воспроизведения звука в соответствии с некоторыми вариантами осуществления изобретения.

Подробное описание вариантов осуществления изобретения

Последующее описание сосредоточено на вариантах осуществления изобретения, применимых к пространственному воспроизведению звука стереосигнала с использованием микширования с увеличением частоты к трем каналам. Тем не менее, следует принимать во внимание, что изобретение не ограничено этим вариантом применения, а может быть применено ко многим другим аудиосигналам и способам воспроизведения.

Фиг.1 иллюстрирует пример системы для воспроизведения звука в соответствии с некоторыми вариантами осуществления изобретения. Система содержит приемное устройство 101, которое принимает пространственный аудиосигнал, содержащий множество аудиоканалов. В примере, входной сигнал является стереосигналом, но следует принимать во внимание, что в других вариантах осуществления другие числа каналов могут использоваться. Например, входной сигнал может быть пятиканальным входным сигналом объемного звучания. В некоторых сценариях, входной сигнал может быть кодированным сигналом, и приемное устройство 101 может быть выполнено с возможностью частично или полностью декодировать входной сигнал для последующей обработки посредством системы. Например, для каждого сегмента кодирования, частотное представление входного сигнала может быть сформировано в качестве промежуточного частотного представления, используемого посредством схемы кодирования. Также следует принимать во внимание, что множество каналов входного сигнала может быть представлено посредством одного кодированного аудиосигнала и ассоциированных параметрических данных. Например, многоканальный входной сигнал может быть кодированным моносигналом и пространственными параметрическими данными. В качестве конкретного примера, входной сигнал может быть параметрическим стереосигналом.

Входной многоканальный аудиосигнал может быть принят из любого внутреннего или внешнего источника.

Приемное устройство 101 соединяется со схемой 103 формирователя сигналов возбуждения, которая принимает многоканальный (в конкретном примере стереосигнал) из приемного устройства 101. Схема 103 формирователя сигналов возбуждения формирует сигналы возбуждения для набора громкоговорителей 105. Набор громкоговорителей предоставляет число пространственных каналов. В примере, громкоговорители предоставляют левый канал, правый канал и центральный канал, но следует принимать во внимание, что в других вариантах осуществления больше (или меньше) пространственных каналов могут предоставляться. Например, в некоторых вариантах осуществления, громкоговорители могут предоставлять только левый и правый канал. В других вариантах осуществления, полная система объемного звучания

снабжена например, пятью или семью пространственными каналами.

В некоторых примерах, число пространственных каналов, предоставляемых посредством динамиков в наборе громкоговорителей 105, может быть равным числу каналов в многоканальном сигнале. Тем не менее, в примере, число пространственных

5 каналов, предоставляемых посредством набора громкоговорителей 105, больше числа каналов в многоканальном сигнале. В примере, схема 103 формирователя сигналов возбуждения может работать в некоторых режимах воспроизведения, которые включают в себя микширование с увеличением частоты каналов многоканального сигнала к числу пространственных каналов. Альтернативно или дополнительно, схема 103

10 формирует сигналов возбуждения может включать в себя функциональность для выбора поднабора доступных каналов, по меньшей мере, в некоторых режимах воспроизведения, причем поднабор отличается в различных режимах воспроизведения. Один или более этих режимов дополнительно может включать в себя микширование с понижением частоты входных каналов. Например, для входного стереосигнала, один

15 режим воспроизведения может предоставлять вывод с использованием двух из пространственных каналов (например, левого и правого), другой режим воспроизведения может использовать только один пространственный канал (например, центральный канал), и еще один другой режим воспроизведения может использовать три пространственных канала (например, левый, правый и центральный канал).

20 В конкретном примере, набор громкоговорителей 105 содержит три громкоговорителя в пространственной компоновке, тем самым предоставляя три пространственных канала. Таким образом, динамики набора громкоговорителей 105 соответствуют левому, правому и среднему динамику.

Таким образом, набор громкоговорителей выполнен с возможностью предоставлять 25 пространственное восприятие. В некоторых вариантах осуществления, схема 103 формирователя сигналов возбуждения может знать точное размещение громкоговорителей относительно позиции прослушивания, но типично это не происходит, и пространственное воспроизведение звука основано на допускаемом размещении громкоговорителей, как известно из традиционных систем объемного 30 звучания и стереосистем. Набор громкоговорителей предоставляет множество пространственных каналов, например, они могут предоставлять левый, правый и центральный пространственный канал, которые используются для того, чтобы предоставлять пространственное восприятие для слушателя. Тем не менее, набор громкоговорителей не должен иметь один отдельный громкоговоритель для каждого 35 канала. Например, набор громкоговорителей может содержать массив громкоговорителей и ассоциированную функциональность возбуждения для предоставления пространственных каналов с использованием технологий формирования диаграммы направленности аудио. Таким образом, громкоговорители из набора громкоговорителей 105 по фиг.1 могут восприниматься в качестве виртуальных 40 громкоговорителей, которые соответствуют данному пространственному местоположению или каналу. В некоторых вариантах осуществления, каждый виртуальный громкоговоритель может соответствовать физическому громкоговорителю, но это не необходимо во всех вариантах осуществления.

Схема 103 формирователя сигналов возбуждения выполнена с возможностью 45 использовать различные режимы воспроизведения звука при возбуждении громкоговорителей 105. Различные режимы воспроизведения звука используют различные технологии пространственного рендеринга. Таким образом, различные режимы воспроизведения звука могут применять различные алгоритмы

пространственной обработки, и тем самым различные режимы воспроизведения звука имеют различные пространственные характеристики аудио. Например, один режим воспроизведения звука может представлять многоканальный сигнал с использованием только одного громкоговорителя 105 (т.е. в качестве моновоспроизведения), другой 5 режим воспроизведения может просто возбуждать каждый громкоговоритель с помощью сигнала соответствующего пространственного канала вообще без пространственной обработки, тем самым поддерживая пространственные характеристики входного сигнала. Еще один другой режим воспроизведения может расширять входные каналы по всем громкоговорителям и вводить пространственное 10 расширение. Таким образом, схема 103 формирователя сигналов возбуждения спроектирована с возможностью позволять предоставлять существенно отличающуюся пространственную обработку и возбуждать набор громкоговорителей 105 с существенно 15 отличающимися свойствами. Фактически, различные режимы воспроизведения не просто используют различные настройки параметров для данной пространственной обработки, но применяют различные базовые принципы и в конкретном использовании различные алгоритмы и способы пространственной обработки.

Такое множество режимов воспроизведения может обеспечивать предоставление существенно отличающихся эффектов посредством системы и может обеспечивать высокую изменчивость в пространственном восприятии слушателя. Тем не менее, 20 авторы изобретения выяснили, что хотя пространственная обработка сигналов может предоставлять расширенное восприятие, она также может в некоторых случаях приводить к уменьшенному пространственному восприятию. Например, эффект алгоритма преобразования аудиоформата (к примеру, пространственное расширение, 25 микширование с увеличением частоты, преобразование в моносигнал и т.д.) на воспринимаемое стереоизображение может различаться для различного содержимого и характеристик сигналов.

Например, способ может предоставлять широкое пространственное изображение, 30 которое является подходящим для сцены боевика, но этот способ может восприниматься как беспокоящий и нечеткий в случае программы новостей или музыки с одним инструментом. Т.е. микширование с увеличением частоты или стереорасширение, 35 которое может быть подходящим для одного типа содержимого, может формировать нежелательный эффект при использовании для другого типа содержимого.

В качестве другого примера, алгоритмы микширования с увеличением частоты, 40 которые нацелены на извлечение центрального канала из стереосигнала, могут не всегда работать оптимально, когда нет четкого центрального источника звука в стереосмешении. Если способ извлечения центрального канала используется для такого содержимого, это может приводить к уменьшению ширине стереоизображения.

Предоставление возможности конечному пользователю вручную выбирать или 45 регулировать режим воспроизведения может обеспечивать снижение этой чувствительности, поскольку пользователь может выбирать режим, предоставляющий самое приятное пространственное восприятие. Тем не менее, авторы изобретения выяснили, что это решение зачастую может быть непрактичным, поскольку оно обеспечивает только медленную и очень трудоемкую адаптацию.

Решение может состоять в том, чтобы задавать режим воспроизведения для каждого 50 возможного типа аудио. Например, для программы новостей используется один конкретный режим воспроизведения, для фильма используется другой конкретный режим воспроизведения и т.д. Тем не менее, авторы изобретения выяснили, что такой подход с большой вероятностью является неточным, поскольку предпочтительное

пространственное воспроизведение может не быть связано напрямую с определенным типом аудио.

Фактически, авторы изобретения выяснили, что существенно улучшенное восприятие зачастую может достигаться посредством реализации динамического выбора в реальном

5 времени подходящего режима воспроизведения. Авторы изобретения дополнительно выяснили, что преимущественная производительность может достигаться посредством реализации такого динамического выбора на основе пространственного свойства входного сигнала. Таким образом, в системе по фиг.1, режим воспроизведения динамически выбирается на основе пространственного свойства входного сигнала.

10 Тем самым, достигается быстрая адаптация в реальном времени режима воспроизведения к конкретным изменениям во входном сигнале.

Такой подход обеспечивает автоматическое и динамическое приспособление воспроизведения звука к текущим характеристикам сигнала, тем самым обеспечивая расширенное восприятие при прослушивании. Кроме того, подход обеспечивает очень

15 быструю адаптацию, что разрешает оптимизацию режима воспроизведения для текущих характеристик и настроек вместо средних или ожидаемых характеристик, например, для определенного типа аудио или конкретного типа программы, который представляет аудио. Например, подход обеспечивает динамическое и автоматическое изменение режима воспроизведения во время звуковой дорожки фильма, так что, например, звуки

20 диалогов и действий воспроизводятся посредством наиболее подходящего алгоритма воспроизведения для этого конкретного звука. Например, известно, что пространственное изображение зачастую изменяется непрерывно в течение длительности мультимедийного элемента. Например, аудиосцена фильма может содержать чередование между широкими сценами стереофонического аудио и моментами, когда

25 только один источник звука, к примеру, речь актера, является слышимым. В первом случае требуется, чтобы стереоизображение было широким и иммерсивным, в то время как во втором случае естественно иметь четко локализованное пространственное местоположение для речи. Система по фиг.1 предусматривает автоматическое регулирование режима воспроизведения, чтобы отражать такие настройки.

30 В частности, система по фиг.1 содержит анализатор 107, который выполнен с возможностью определять пространственное свойство многоканального аудиосигнала. Пространственное свойство может, в частности, быть индикатором относительно степени пространственной организации или сложности, которая присутствует во входном сигнале. Пространственное свойство может служить признаком степени

35 пространственного расширения и может, в частности, служить признаком того, отличается входной сигнал посредством одного или более одиночных четко определенных источников звука или в большей степени отличается посредством окружающего звука без сильных направленных меток.

Анализатор 107 соединяется с процессором 109 выбора, в который подается

40 пространственное свойство и который выполнен с возможностью выбирать режим воспроизведения из множества режимов воспроизведения звука, которые могут быть использованы посредством схемы 103 формирователя сигналов возбуждения. Процессор 109 выбора дополнительно соединен со схемой 103 формирователя сигналов возбуждения и управляет ей, чтобы использовать выбранный режим воспроизведения.

45 Таким образом, по мере того как пространственное свойство варьируется, процессор 109 выбора динамически и автоматически переключается между режимами воспроизведения, чтобы предоставлять оптимальную обработку воспроизведения для текущих характеристик. Таким образом, достигается улучшенное пространственное

восприятие.

Система, в частности, выполнена с возможностью обеспечивать краткосрочную адаптацию режима воспроизведения к характеристикам сигналов. Таким образом, может разрешаться быстрое переключение, тем самым позволяя пространственному воспроизведению не только быть оптимизированным в среднем (за длительный период), но также и совпадать с более мгновенными изменениями сигнала.

Соответственно, анализатор 107 выполнен с возможностью формировать оценку в форме пространственного свойства, которое подвергается фильтрации нижних частот или усредняется, но с относительно высокой частотой. Аналогично, фактическое переключение между режимами воспроизведения может выполняться с относительно высокой частотой. Таким образом, вместо того, чтобы выбирать режим воспроизведения и использовать его, например, для всей программы, система по фиг.1 динамически адаптирует режим воспроизведения так, что он совпадает с краткосрочными изменениями в характеристиках сигналов.

Предпочтительные динамические характеристики системы могут зависеть от конкретных характеристик и настроек отдельного варианта осуществления.

Тем не менее, во многих вариантах осуществления, в частности, преимущественная производительность может достигаться с помощью системы, которая обеспечивает обновления режима воспроизведения с интервалами, которые варьируются в диапазоне типично приблизительно от 50 мс до 5 минут. Точный динамический характер может быть выбран на основе компромисса между точностью адаптации к текущим характеристикам сигнала и надежностью системы и степенью артефактов, ассоциированных с переключением между различными режимами.

Во многих вариантах осуществления, фильтрация нижних частот, включенная при определении пространственного свойства, преимущественно имеет частоту отсечки при 3 дБ, превышающую 0,001 Гц, 0,01 Гц, 0,1 Гц, 1 Гц, 10 Гц или 50 Гц, в зависимости от конкретных настроек отдельного варианта осуществления. Соответственно, пространственное свойство преимущественно может быть определено с постоянной времени меньше 500 секунд, 100 секунд, 10 секунд, 1 секунды, 500 мс, 100 мс или даже 50 мс. Постоянная времени может задаваться как время, которое требуется для пространственного свойства, чтобы достигать $1-1/e \approx 63\%$ от своего конечного (асимптотического) значения после ступенчатого изменения. Например, пространственное свойство может отслеживать или зависеть от одной или более пространственных характеристик многоканального сигнала. Ступенчатое изменение в этой пространственной характеристике при поддержании всех остальных параметров постоянными должно приводить к изменению в пространственном свойстве. Постоянная времени для определения пространственного свойства затем может измеряться в качестве времени, которое требуется для этого изменения, чтобы достигать $1-1/e \approx 63\%$ от своего конечного (асимптотического) значения.

Аналогично, переключение может компоноваться в соответствии с аналогичной динамикой. В частности, максимальная частота переключения для переключения между режимами воспроизведения может превышать 0,01 Гц; 0,1 Гц, 1 Гц или даже 10 Гц. Максимальная частота может быть наиболее быстрым возможным переключением вследствие определения пространственного свойства и/или фактической операции переключения. Таким образом, максимальная частота переключения может быть наибольшим изменением частоты в базовых пространственных характеристиках аудиосигнала, которому может следовать система.

В конкретном варианте осуществления, схема 103 формирователя сигналов

возбуждения выполнена с возможностью переключаться между четырьмя различными режимами воспроизведения.

В первом режиме воспроизведения схема 103 формирователя сигналов возбуждения просто поддерживает исходный стереосигнал и не вводит пространственной

5 модификации. Таким образом, этот режим работы поддерживает пространственные характеристики многоканального входного сигнала. В конкретном примере, входной стереосигнал просто воспроизводится в качестве стереосигнала, т.е. левый входной канал подается в левый громкоговоритель, а правый входной канал подается в правый громкоговоритель.

10 Таким образом, в этом режиме воспроизведения схема 103 формирователя сигналов возбуждения предоставляет стереофоническое воспроизведение исходных аудиоканалов.

Во втором режиме воспроизведения схема 103 формирователя сигналов возбуждения воспроизводит входной сигнал в качестве моносигнала. Например, эти два стереоканала могут быть комбинированы (например, посредством простого суммирования), и

15 результирующий моносигнал может быть подан в центральный громкоговоритель без подачи сигнала в левый либо в правый громкоговоритель. Таким образом, второй режим воспроизведения схемы 103 формирователя сигналов возбуждения включает в себя микширование с понижением частоты входного сигнала и является режимом монофонического воспроизведения. Такой режим воспроизведения может, в частности,

20 являться преимущественным и т.д. в сценариях, в которых аудио соответствует одному центрально размещенному источнику звука, к примеру, в сценарии диктора новостей для программы новостей.

В третьем режиме воспроизведения схема 103 формирователя сигналов возбуждения выполнена с возможностью вводить обработку пространственного расширения. В

25 конкретном примере, третий режим воспроизведения содержит применение алгоритма стереорасширения к входному стереосигналу. Такое стереорасширение имеет тенденцию предоставлять декорреляцию стереоканалов, так что достигается восприятие

30 укрупненного пространственного изображения. Следует принимать во внимание, что различные технологии пространственного расширения должны быть известными для специалистов в данной области техники, и что любой подходящий алгоритм может быть использован без отступления от изобретения.

Такая обработка может, в частности, являться преимущественной, когда звуковое изображение управляет посредством окружающих звуков, а не конкретных локализованных источников звука. Например, это позволяет предоставлять

35 расширенное восприятие при воспроизведении музыки, созданной посредством большого оркестра с множеством инструментов.

В четвертом режиме воспроизведения схема 103 формирователя сигналов возбуждения разделяет входной сигнал на один или более первичных исходных сигналов, при этом каждый первичный сигнал пытается содержать звук только из конкретного

40 доминирующего источника звука. Следует принимать во внимание, что специалисты в данной области техники должны знать различные алгоритмы для обнаружения и извлечения доминирующих источников звука, и что любой подходящий алгоритм может быть использован без отступления от изобретения. Схема 103 формирователя сигналов возбуждения дополнительно формирует остаточный сигнал, соответствующий сигналу

45 после извлечения доминирующего источника(ов) звука. В четвертом режиме воспроизведения входной стереосигнал тем самым раскладывается на один или более сигналов первичного источника звука и окружающие стереосигналы или сигналы объемного звучания.

Сигнал доминирующего источника звука и остаточный сигнал затем обрабатываются по-разному, так что различная пространственная обработка применяется к сигналам. В качестве простого примера, пространственное расширение может применяться к остаточному сигналу, а не к сигналам доминирующего источника звука. Таким образом, 5 пространственно четко определенное размещение доминирующих источников звука не модифицируется, тогда как расширенное звуковое изображение достигается для остаточного сигнала, который типично соответствует окружающему звуковому окружению. Кроме того, сигнал доминирующего источника звука, например, может 10 быть представлен в центральном пространственном канале, и остаточный сигнал может быть представлен в правом и левом пространственных каналах. Таким образом, в этом режиме воспроизведения используются все пространственные каналы, предоставляемые посредством набора громкоговорителей, и режим содержит микширование с увеличением частоты входного сигнала.

Предложены способы для того, чтобы оценивать пространственное исходное

15 распределение из аудиоканалов. Например, способ для определения направления заметного источника звука из многоканальных аудиоданных и оценки уровня окружающего звука предложен в работе M. Goodwin и J-M. Jot, "Multichannel surround format conversion and generalized upmix", AES 30th int. Conference, Финляндия, март 2007 года. Два других способа для оценки распределения нескольких источников звука в 20 стереосмешении изучены, например, в работе A. Härmä and C. Faller "Spatial decomposition of time-frequency regions: subbands or sinusoids", AES 116th Convention, Берлин, Германия, 8-11 мая 2004 года.

25 Четвертый режим воспроизведения может быть, в частности, подходящим, например, для сигналов, которые являются смешением между конкретными источниками звука и окружающим звуком или шумом.

Анализ пространственного распределения источников звука во входном сигнале 30 посредством анализатора 107, например, может быть основан на частотно-избирательном анализе энергии аудио в каждом канале и/или частотно-избирательном анализе изменения некоторых подходящих числовых показателей, которые представляют подобия между каналами. Например, анализатор 107 может использовать аналитические 35 способы, аналогичные способам, используемым в стандарте объемного звучания MPEG. Таким образом, они могут быть основаны на разложении на подполосы частот входных сигналов и вычислении энергии и значений ковариации между подполосами частот в различных каналах. Тем не менее, следует принимать во внимание, что может быть использовано множество других подходов, такие как показатели корреляции, связанные 40 с параметрическими представлениями сигналов, и/или взаимная информация, характеризующая подобие между различными каналами.

Фиг.2 иллюстрирует конкретный подход, который может быть использован в системе по фиг.1.

40 В примере, анализатор 107 содержит сумматор 201 и модуль 203 вычитания, в которые подаются входные левый и правый сигналы. Сумматор суммирует эти два сигнала, а модуль 203 вычитания вычитает один из другого. Сумматор 201 подается в первый модуль 205 оценки энергии, который вычисляет энергию сигналов суммирующего сигнала, сформированного посредством сумматора 201. Модуль 203 вычитания подается 45 во второй модуль 207 оценки энергии, который измеряет энергию сигналов разностного сигнала, сформированного посредством модуля вычитания. Первый и второй модули 205, 207 оценки энергии соединяются с процессором 109 выбора, который выбирает режим воспроизведения на основе индикатора пространственного свойства относительно

суммирующих и разностных энергий.

Таким образом, в примере, выбор режима воспроизведения основан на вычислении суммирующих и разностных сигналов между сигналами левого и правого канала и сравнении кратковременных энергий сигналов. Когда энергия суммирующего сигнала значительно больше, чем разностного сигнала, оценивается то, что входной стереосигнал является практически монофоническим. Когда энергии суммирующего и разностного сигнала находятся на идентичном уровне, или энергия разностного сигнала больше энергии суммирующего сигнала, входной сигнал считается обычным стереоаудиосигналом.

Таким образом, значение обнаружения в каждый период анализа энергии может задаваться следующим образом:

$$\rho = \begin{cases} 1, & \text{if } E_{\text{sum}} > AE_{\text{diff}} \\ 0, & \text{if } E_{\text{sum}} \leq AE_{\text{diff}}, \end{cases}$$

где E_{sum} и E_{diff} являются кратковременными энергиями суммирующих и разностных сигналов, соответственно, и A является скалярным коэффициентом, который typically значительно больше единицы (например, $A=100$).

Работа схемы 103 формирователя сигналов возбуждения и, в частности, переключение между различными режимами воспроизведения может быть реализовано как динамическая матричная операция:

$$\begin{pmatrix} y_l(n) \\ y_r(n) \\ y_c(n) \end{pmatrix} = \begin{pmatrix} 1 - p(n) & 0 \\ 0 & 1 - p(n) \\ \frac{p(n)}{2} & \frac{p(n)}{2} \end{pmatrix} \begin{pmatrix} x_l(n) \\ x_r(n) \end{pmatrix},$$

где $x_l(n)$ и $x_r(n)$ являются исходным левым и правым стереосигналом, p является индексом для цифровых сигналов выборок. Выводы $y_l(n)$, $x_r(n)$ и $y_c(n)$ являются значениями возбуждения для левого, правого и центрального динамиков, соответственно.

Таким образом, в примере, энергии сигналов суммирующих и разностных сигналов используются для того, чтобы переключаться между практически монофоническим воспроизведением с использованием центрального динамика и стереовоспроизведением с использованием левого и правого динамиков.

В качестве другого примера, суммирующие и разностные операции могут заменяться посредством более обобщенных операций. Например, направление доминирующего источника звука может быть оценено посредством анализа главных компонентов (PCA) (или других аналогичных способов, таких как адаптивное разложение по собственным значениям). Дополнительно, могут быть использованы взвешенные суммы и разности, так что доминирующий источник звука исключается из разностного сигнала. Это может приводить к структурно очень похожему, но более обобщенному решению, чем пример по фиг.2.

Описанный подход, например, может применяться независимо в различных частотных интервалах, к примеру, в отдельных элементах разрешения по частоте, сформированных посредством преобразования Фурье, или в подполосах частот гребенки фильтров.

В конкретном примере, вышеуказанный подход сначала используется для того, чтобы определять то, где входной сигнал имеет практически монофонический характер. Если да, используется второй режим воспроизведения (монофоническое представление). Если нет, т.е. если $\rho = 0$, последующая обработка выполняется для того, чтобы выбирать,

какой из других режимов воспроизведения должен быть использован. Эти способы воспроизведения, в частности, могут переключаться посредством надлежащего переключения обработки, которая применяется к $x_l(n)$ и $x_r(n)$. Например, для первого режима воспроизведения (поддерживающего пространственные характеристики 5 входного сигнала) входные каналы используются непосредственно как $x_l(n)$ и $x_r(n)$ (и таким образом, $y_l(n)$ и $y_r(n)$), тогда как для третьего режима воспроизведения (расширения) пространственное расширение сначала применяется к входным сигналам до того, как они используются как $x_l(n)$ и $x_r(n)$ (и таким образом, $y_l(n)$ и $y_r(n)$) и подаются 10 в громкоговорители.

В некоторых вариантах осуществления, анализатор 107 может определять сигнал доминирующего источника звука, содержащий один или более доминирующих источников звука. Затем может быть сформирован остаточный сигнал, представляющий сигнал, остающийся после того, как доминирующий источник(и) звука извлечен. В 15 завершение, пространственное свойство может быть определено в ответ на энергетический индикатор для сигнала доминирующего источника звука относительно энергетического индикатора для остаточного сигнала.

Например, технологии направленной фильтрации могут быть использованы для 20 того, чтобы извлекать доминирующий источник из стереосмешения входного сигнала. Это извлечение может использовать любую подходящую технологию для разложения многоканальных сигналов, включающую в себя алгоритмы формирования диаграммы направленности, адаптивные алгоритмы формирования диаграммы направленности, алгоритмы и способы разделения источников вслепую для подавлений многоканального шума, как должно быть известно специалистам в данной области техники.

25 После извлечения доминирующего (или первичного) источника из смешения определяется многоканальный остаточный сигнал, в котором доминирующий источник звука исключен или подавлен.

В этом случае, значение обнаружения может вычисляться следующим образом:

$$30 \quad \rho = \begin{cases} 1, & \text{if } E_{\text{prim}} > BE_{\text{res}}, \\ 0, & \text{if } E_{\text{prim}} \leq BE_{\text{res}}, \end{cases}$$

где E_{prim} является энергетическим показателем для сигнала доминирующего или первичного источника звука, а E_{res} является энергетическим показателем для 35 остаточного сигнала. Значение параметра B типично составляет приблизительно единицу в зависимости от конкретных характеристик извлечения первичного сигнала. Если энергия извлеченного доминирующего источника является низкой по сравнению с остатком, система определяет то, что смешение не содержит доминирующий/первичный источник звука. В этом случае, третий способ воспроизведения может быть выбран для того, чтобы предоставлять расширенное пространственное изображение.

40 В противном случае устройство может переходить к оценке того, содержит или нет остаточный сигнал другой доминирующий источник звука. Это, например, может выполняться посредством итеративного применения разделения первичных источников к остаточному сигналу. В качестве другого примера, определение может быть основано на вычислении показателей подобия между многоканальными сигналами. Типичными 45 показателями подобия являются различные типы взвешенных показателей корреляции, такие как корреляция Пирсона, оценки для максимального значения корреляционной функции или нормализованная корреляционная функция. Также можно использовать различные типы разностных функций величин или теоретико-информационных

показателей, такие как взаимная информация. Если показатель демонстрирует низкое подобие между двумя остаточными сигналами, это служит признаком присутствия одного доминирующего источника звука с некоторым окружающим сигналом (поскольку ранее обнаружено, что сигнал не является практически монофоническим).

- 5 Соответственно, четвертый режим воспроизведения может быть использован при воспроизведении сигнала доминирующего или первичного источника без пространственного расширения (например, в качестве моносигнала, подаваемого в центральный канал), тогда как пространственное расширение применяется к остаточному стереосигналу, который затем подается в левый и правый
- 10 громкоговорители.

Тем не менее, если обнаружено, что каналы остаточного сигнала имеют высокое подобие, то это, вероятно, отражает то, что входной сигнал, возможно, состоит из двух доминирующих источников, которые могут лучше воспроизводиться посредством первого способа воспроизведения, и, соответственно, он выбирается.

- 15 Переключение между различными режимами воспроизведения во многих вариантах осуществления преимущественно может быть плавным и постепенным переходом. Это может уменьшать и сокращать артефакты, возникающие в результате различных пространственных характеристик различных режимов воспроизведения.

- 20 В качестве примера, переключение из монорежима в режим стереовоспроизведения может выполняться согласно следующему:

$$\begin{pmatrix} y_l(n) \\ y_r(n) \\ y_c(n) \end{pmatrix} = \begin{pmatrix} 1 - p(n) & 0 & 0 \\ 0 & 1 - p(n) & 0 \\ p(n) & p(n) & \frac{p(n)}{2} \end{pmatrix} \begin{pmatrix} x_l(n) \\ x_r(n) \\ x_c(n) \end{pmatrix},$$

- 25 где:

$$p(n) = \alpha p(n-1) + (1 - \alpha) \rho,$$

где коэффициент α временного интегрирования является значением в интервале $[0,1]$. Типичное значение, например, может составлять $\alpha=0,95$.

- 30 В качестве более общего примера, устройство может быть выполнено с возможностью работать в двух (или более) режимах воспроизведения одновременно. Сигналы, сформированные из двух режимов воспроизведения, между которыми переключается система, затем могут смешиваться вместе, при этом взвешивание двух режимов постепенно изменяется с предыдущего режима воспроизведения на новый режим воспроизведения. Например, для каждого громкоговорителя соответствующие сигналы, сформированные посредством двух режимов воспроизведения, могут быть суммированы согласно следующему:

$$y(n) = \beta(n) \cdot x_p(n) + (1 - \beta(n)) \cdot x_n(n),$$

- 35 где $y(n)$ является сигналом возбуждения для динамика, x_p является выборкой, сформированной посредством предыдущего режима воспроизведения, x_n является выборкой, сформированной посредством нового режима воспроизведения, n является индексом выборки, а β является значением, которое постепенно изменяется от 1 до 0 с подходящей временной характеристикой.

- 40 Во многих вариантах осуществления, время перехода в интервале от 10 мс до 1 секунды имеет тенденцию предоставлять преимущественную производительность.

Время перехода может измеряться в качестве времени, когда новый режим воспроизведения изменяется со взвешивания в 10% до взвешивания в 90% от результирующего комбинированного сигнала.

В некоторых вариантах осуществления, возбуждающая схема 103 дополнительно выполнена с возможностью адаптировать характеристику технологии пространственного рендеринга выбранного режима воспроизведения в ответ на пространственное свойство. Например, для третьего режима воспроизведения, степень 5 применяемого пространственного расширения может регулироваться в зависимости от пространственного приоритета. Таким образом, в данном примере, анализ пространственного смешения входного сигнала также используется для того, чтобы управлять величиной декорреляции или "параметром стереорасширения" алгоритма пространственного расширения. Например, если пространственное свойство указывает, 10 что входной сигнал содержит наполненное и широкое пространственное изображение с несколькими источниками или, например, рассеянный сигнал без различимого источника звука, большее стереорасширение может применяться в воспроизведении, чем тогда, когда существует по существу идентичное содержимое в обоих каналах. Первый случай может отличаться от второго случая посредством оценки величины 15 корреляции между двумя аудиоканалами.

В качестве еще одного примера, может рассматриваться сигнал, в котором два отдельных источника доминируют над левым и правым каналом, соответственно. В этом случае, целевое пространственное изображение состоит из двух четко локализованных разделенных источников в стереоизображении (например, дуэт из 20 певца слева и гитары справа). В этом случае, корреляция между каналами является низкой. Если стереорасширение применяется к сигналам вследствие корреляции для сигналов, то сформированное пространственное изображение является широким. Тем не менее, в этом случае стереоизображение должно становиться размытым в связи с отсутствием четко локализованного характера двух целевых стереоизображений.

25 Следовательно, вероятно, лучше использовать прямое (нерасширенное) стереовоспроизведение для этого типа содержимого, чтобы сохранять четко локализованные источники в изображении. Можно обнаруживать, имеет стереоизображение простое смешение небольшого числа некоррелированных источников или оно является сложным смешением нескольких источников звука. Простой способ 30 осуществлять это состоит в том, чтобы анализировать нормализованную взаимную корреляцию C между левым и правым каналом. На основе такого обоснования, выбор режима воспроизведения в некоторых вариантах осуществления может быть основан на следующей логике:

35 Если $C < T_{low}$, считается, что содержимое состоит из двух некоррелированных источников слева и права, и стандартное (нерасширенное) стереовоспроизведение выбирается так, что оно сохраняет локализацию двух источников

40 Если $T_{low} < C < T_{high}$, считается, что содержимое является обычным сложным стереоматериалом. Подход стереорасширения, соответственно, используется для воспроизведения для этого типа содержимого.

Если $T_{high} < C$, считается, что содержимое имеет один различный источник. Способ стереовоспроизведения или конкретное воспроизведение для монофонического содержимого, следовательно, выбирается для этого типа ввода.

45 Нормализованная корреляционная функция, например, может быть корреляцией Пирсона, представленной посредством:

$$C = E[x_l(n)x_n(n)] / \sqrt{(E[x_l(n)x_l(n)]E[x_r(n)x_r(n)])},$$

или показателем нормализованной корреляции, предложенным автором Avendado

(C. Avendado, "Frequency-domain source identification and manipulation in stereo mixes for enhancement, suppression and re-panning applications", IEEE Proc. WASPAA, Нью-Йорк, США, 2003 года), который задается следующим образом:

$$C = 2E[x_l(n)x_n(n)]/(E[x_l(n)x_l(n)] + E[x_r(n)x_r(n)]).$$

Обнаружение также может быть основано на статистике корреляции и разностей уровня между каналами в небольших частотно-временных сегментах входных сигналов.

Система по фиг.1 может предоставлять улучшенное восприятие при прослушивании во многих сценариях и для многих реальных сигналов. В частности, пространственное восприятие для систем на основе микширования с увеличением частоты может быть улучшено во многих сценариях. Например, алгоритмы микширования с увеличением частоты, которые выполнены с возможностью извлекать центральный канал из стереосигнала, могут предоставлять очень хорошую производительность, когда центральный источник звука присутствует в звуковом изображении, но могут не всегда работать идеально в случае, когда нет четкого центрального изображения в стереосмещении. Фактически, если способ извлечения центрального канала используется для такого содержимого, это может приводить к уменьшению ширины стереоизображения. Описанный подход обеспечивает динамическую адаптацию воспроизведения входного сигнала к использованию подходящего подхода микширования с увеличением частоты.

В некоторых вариантах осуществления, выбор режима воспроизведения дополнительно может учитывать свойство содержимого для входного сигнала. Пример такого проиллюстрирован на фиг.3, который показывает систему по фиг.1, модифицированную так, что она включает в себя процессор 301 содержимого, который выполнен с возможностью определять характеристику содержимого для сигнала.

Характеристика содержимого, например, может указывать жанр, тип программы, ассоциированный с аудиосигналом (например, если аудио ассоциировано с мультимедийным элементом, таким как, например, телевизионная или радиопередача), исполнителя, ассоциированного с аудио, и т.д.

Характеристика содержимого, например, может быть определена из метаданных, ассоциированных с входным сигналом. Таким образом, в некоторых сценариях метаданные могут приниматься отдельно или, например, встраиваться в аудиосигнал. Процессор 301 содержимого может быть выполнен с возможностью извлекать данные, описывающие содержимое входного сигнала.

В других вариантах осуществления, процессор 301 содержимого может быть выполнен с возможностью осуществлять анализ содержимого принимаемого входного сигнала и определять характеристику содержимого на основе такого анализа содержимого. Например, процессор 301 содержимого может анализировать сигнал, чтобы определять то, содержит он преимущественно речь, музыку или, например, громкие взрывы. Он затем может оценивать соответствующий тип содержимого, к примеру, выбирать между программой новостей, музыкальной программой и боевиком на основе анализа. Следует принимать во внимание, что различные подходы анализа содержимого должны быть известными для специалистов в данной области техники, и что любой подходящий алгоритм может быть использован. Для аудиовизуальных сигналов (т.е. тех, в которых входной аудиосигнал соединяется с видеосигналом), анализ содержимого альтернативно или дополнительно может быть основан на видеосигнале, ассоциированном с входным сигналом.

Характеристика содержимого подается в процессор 109 выбора, который продолжает включать ее в выбор режима воспроизведения, который следует использовать. В

частности, краткосрочное переключение между различными режимами воспроизведения по-прежнему может быть определено на основе краткосрочных изменений пространственного свойства, но точные критерии переключения могут быть модифицированы в зависимости от того, каково содержимое. Например, система может с большей вероятностью переключаться на подход пространственного расширения для боевика, чем для программы новостей.

Таким образом, данные, указывающие тип содержимого, могут быть использованы при выборе оптимального способа пространственного воспроизведения, который следует использовать. В частности, характеристика содержимого может быть использована для того, чтобы повышать надежность стратегии выбора режима воспроизведения. Включение характеристики содержимого в решение может снижать риск выбора несоответствующего режима воспроизведения.

Например, в некоторых случаях пространственный анализ сигнала может приводить к пространственному свойству, которое не указывает четко подходящий режим воспроизведения. В этом случае, может быть желательным учитывать содержимое при выборе режима воспроизведения. Таким образом, характеристика содержимого может учитываться в случаях, если пространственный анализ сигналов не классифицирует четко пространственное смешение сигнала на один из четырех классов воспроизведения, а находится в неопределенной "серой" области между двумя или более из них. В некоторых вариантах осуществления, интервалы пространственного свойства, которые соответствуют каждому из режимов воспроизведения, например, могут зависеть от конкретного свойства. Это, например, может приводить к выбору между режимом немодифицированного стереовоспроизведения и режимом расширенного стереовоспроизведения, отличающемуся, например, для программы новостей и боевика.

Таким образом, расширение может быть использовано меньше для программы новостей, чем для боевика.

В некоторых вариантах осуществления, схема 103 формирователя сигналов возбуждения может адаптировать характеристику технологии пространственного рендеринга выбранного режима воспроизведения в ответ на характеристику содержимого. Таким образом, характеристика содержимого, отражающая информацию о типе содержимого входного сигнала, может быть использована для того, чтобы управлять параметрами выбранного режима пространственного воспроизведения. Например, величина расширения, которое применяется, когда система определяет, что стереорасширение является оптимальным способом воспроизведения, может регулироваться в зависимости от типа содержимого. С этой целью, классификация типа содержимого может выполняться на высоком уровне, например, с различием между такими классами, как "новости", "фильм", "музыка", "документальный фильм" и т.д. Тем не менее, также может быть выгодным выполнять классификацию на подтипы, например, различные жанры музыки или различные типы фильмов. Например, определенные жанры музыки типично ассоциированы с довольно глубокой звуковой сценой и акустической атмосферой (например, авторское исполнение или камерная музыка), в то время как другие жанры ассоциированы с широкой звуковой сценой и очень обширной архитектурной акустикой (например, хоровая музыка). Знание музыкального жанра может, в дополнение к анализу пространственного смешения аудиосигнала, помочь выбирать надлежащий режим воспроизведения и/или задавать параметры режима пространственного воспроизведения.

Вышеприведенное описание сосредоточено на вариантах осуществления, в которых набор громкоговорителей предоставляет больше пространственных каналов (в

частности, три пространственных канала), чем входной сигнал (в частности, два канала). Тем не менее, следует принимать во внимание, что в других вариантах осуществления набор громкоговорителей может предоставлять не больше пространственных каналов, чем входной сигнал.

- 5 Фактически, во многих вариантах осуществления, может быть преимущественным для набора громкоговорителей предоставлять меньше пространственных каналов, чем входной сигнал. Например, семиканальный входной сигнал объемного звучания может быть воспроизведен в трех пространственных каналах. В таких вариантах осуществления, потенциально сложная пространственная обработка может быть
- 10 использована для того, чтобы предоставлять преимущественную производительность, и описанные принципы могут быть использованы для того, чтобы выбирать то, какой режим воспроизведения следует применять к конкретным пространственным характеристикам входного сигнала. Таким образом, различные алгоритмы микширования с понижением частоты могут быть использованы в зависимости от
- 15 пространственной характеристики входного сигнала.

Следует принимать во внимание, что вышеприведенное описание для понятности имеет описанные варианты осуществления со ссылкой на различные функциональные схемы, блоки и процессоры. Тем не менее, должно быть очевидным, что любое надлежащее распределение функциональности между различными функциональными

20 схемами, блоками или процессорами может быть использовано без отступления от изобретения. Например, функциональность, проиллюстрированная как выполняемая посредством отдельных процессоров или контроллеров, может быть выполнена посредством одного процессора или контроллера. Следовательно, ссылки на конкретные функциональные блоки или схемы должны рассматриваться только как ссылки на

25 надлежащее средство предоставления описанной функциональности, а не обозначать точную логическую или физическую структуру либо организацию.

Изобретение может быть реализовано в любой надлежащей форме, включающей в себя аппаратные средства, программное обеспечение, микропрограммное обеспечение или любую комбинацию вышеуказанного. Необязательно, изобретение может быть

30 реализовано, по меньшей мере, частично как вычислительное программное обеспечение, выполняемое на одном или более процессоров данных и/или процессоров цифровых сигналов. Элементы и компоненты варианта осуществления изобретения могут быть физически, функционально и логически реализованы любым надлежащим образом. Фактически, функциональность может быть реализована в одном блоке, множестве

35 блоков или как часть других функциональных блоков. По существу, изобретение может быть реализовано в одном блоке или может быть физически и функционально распределено между различными блоками, схемами и процессорами.

Хотя настояще изобретение описано в связи с некоторыми вариантами осуществления, оно не имеет намерение быть ограниченным конкретной изложенной

40 в данном документе формой. Вместо этого, объем настоящего изобретения ограничен только посредством прилагаемой формулы изобретения. Дополнительно, хотя предположительно признак описывается в данном документе в связи с конкретными вариантами осуществления, специалисты в данной области техники должны признавать, что различные признаки описанных вариантов осуществления могут быть

45 комбинированы в соответствии с изобретением. В формуле изобретения, термин "содержащий" не исключает наличия других элементов или этапов.

Более того, хотя перечислены по отдельности, множество средств, элементов, схем или этапов способа может быть реализовано посредством, к примеру, одной схемы,

блока или процессора. Дополнительно, хотя отдельные признаки могут быть включены в различные пункты формулы изобретения, они могут быть преимущественно комбинированы, и их включение в различные пункты формулы изобретения не подразумевает, что комбинация признаков является невыполнимой и/или

- 5 преимущественной. Так же, включение признака в одну категорию пунктов формулы изобретения не налагает ограничение на эту категорию, а вместо этого указывает то, что признак в равной степени применим к другим категориям пунктов формулы изобретения по мере необходимости. Более того, порядок признаков в пунктах формулы изобретения не налагает какой-либо конкретный порядок, в котором признаки должны 10 осуществляться, и, в частности, порядок отдельных этапов в пункте формулы изобретения на способ не подразумевает, что этапы должны выполняться в этом порядке. Вместо этого, этапы могут выполняться в любом надлежащем порядке. Кроме того, ссылки в единственном числе не исключают множественность. Таким образом, 15 ссылки на "первый", "второй" и т.д. не исключают множественность. Ссылки с номерами в формуле изобретения предоставлены просто в качестве поясняющего примера, и они не должны истолковываться как каким-либо образом ограничивающие объем формулы изобретения.

Формула изобретения

- 20 1. Устройство для пространственного воспроизведения звука, содержащее: приемное устройство (101) для приема многоканального аудиосигнала; схему (107) для определения пространственного свойства многоканального аудиосигнала, при этом пространственное свойство служит показателем пространственной организации и/или пространственной сложности данного 25 многоканального аудиосигнала;

схему (109) для выбора выбранного режима воспроизведения из множества режимов воспроизведения звука в ответ на упомянутое пространственное свойство, причем режимы многоканального воспроизведения звука используют разные технологии пространственного рендеринга; и

30 схему (103) воспроизведения для возбуждения набора пространственных каналов, обеспечиваемых посредством набора громкоговорителей (105), чтобы воспроизводить многоканальный аудиосигнал с использованием выбранного режима воспроизведения, при этом упомянутое множество режимов воспроизведения звука содержит по меньшей мере два из следующего:

35 режим монофонического воспроизведения;

режим воспроизведения, в котором сохраняются пространственные характеристики многоканального сигнала;

режим воспроизведения, содержащий обработку по пространственному расширению; и

40 режим воспроизведения, содержащий разделение на по меньшей мере один доминирующий исходный сигнал и окружающий сигнал и применение разного пространственного воспроизведения в отношении данного по меньшей мере одного основного исходного сигнала и окружающего сигнала.

2. Устройство по п. 1, в котором по меньшей мере один из режимов воспроизведения

- 45 звука содержит по меньшей мере одно из повышающего микширования до большего количества пространственных каналов, чем количество каналов многоканального аудиосигнала, и понижающего микширования до меньшего количества пространственных каналов, чем количество каналов многоканального аудиосигнала.

3. Устройство по п. 1, в котором упомянутый набор пространственных каналов содержит другое количество каналов, нежели многоканальный аудиосигнал.

4. Устройство по п. 1, в котором максимальная частота переключения для переключения между режимами воспроизведения звука превышает 1 Гц.

5. Устройство по п. 1, в котором схема (107) для определения пространственного свойства выполнена с возможностью определять пространственное свойство с постоянной времени не больше 10 секунд.

6. Устройство по п. 1, дополнительно содержащее схему (301) для определения характеристики содержимого для многоканального аудиосигнала, при этом схема (109) для выбора выполнена с возможностью дополнительно выбирать выбираемый алгоритм воспроизведения в ответ на характеристику содержимого.

7. Устройство по п. 6, в котором схема (301) для определения характеристики содержимого выполнена с возможностью определять характеристику содержимого в ответ на метаданные, относящиеся к многоканальному аудиосигналу.

8. Устройство по п. 6, в котором схема (103) воспроизведения для воспроизведения многоканального аудиосигнала выполнена с возможностью адаптировать характеристику технологии пространственного рендеринга выбранного режима воспроизведения в ответ на характеристику содержимого.

9. Устройство по п. 1, в котором схема (103) воспроизведения для воспроизведения многоканального аудиосигнала выполнена с возможностью адаптировать характеристику технологии пространственного рендеринга выбранного режима воспроизведения в ответ на упомянутое пространственное свойство.

10. Устройство по п. 9, в котором упомянутой характеристикой является степень пространственного расширения, применяемая к по меньшей мере двум каналам многоканального аудиосигнала.

11. Устройство по п. 1, в котором схема (103) воспроизведения для воспроизведения многоканального аудиосигнала выполнена с возможностью постепенно переходить с первого выбранного алгоритма воспроизведения на второй выбранный алгоритм воспроизведения.

12. Устройство по п. 1, в котором схема (107) для определения пространственного свойства выполнена с возможностью определять пространственное свойство в ответ на энергетический показатель для комбинированного сигнала по меньшей мере двух каналов многоканального аудиосигнала относительно энергетического показателя для разностного сигнала этих по меньшей мере двух каналов.

13. Устройство по п. 1, в котором схема (107) для определения пространственного свойства выполнена с возможностью выполнять разложение многоканального аудиосигнала на по меньшей мере один сигнал доминирующего источника звука и остаточный сигнал и определять пространственное свойство в ответ на энергетический показатель для этого сигнала доминирующего источника звука относительно энергетического показателя для остаточного сигнала.

14. Способ пространственного воспроизведения звука, содержащий этапы, на которых:

принимают многоканальный аудиосигнал;

45 определяют пространственное свойство многоканального аудиосигнала, при этом пространственное свойство служит показателем пространственной организации и/или пространственной сложности данного многоканального аудиосигнала;

выбирают выбираемый режим воспроизведения из множества режимов воспроизведения звука в ответ на упомянутое пространственное свойство, причем

режимы многоканального воспроизведения звука используют разные технологии пространственного рендеринга; и

возбуждают набор громкоговорителей (105) для воспроизведения многоканального аудиосигнала с использованием выбранного режима воспроизведения, при этом

5 упомянутое множество режимов воспроизведения звука содержит по меньшей мере два из следующего:

режим монофонического воспроизведения;

режим воспроизведения, в котором сохраняются пространственные характеристики многоканального сигнала;

10 режим воспроизведения, содержащий обработку по пространственному расширению; и

режим воспроизведения, содержащий разделение на по меньшей мере один доминирующий исходный сигнал и окружающий сигнал и применение разного пространственного воспроизведения в отношении данного по меньшей мере одного

15 основного исходного сигнала и окружающего сигнала.

20

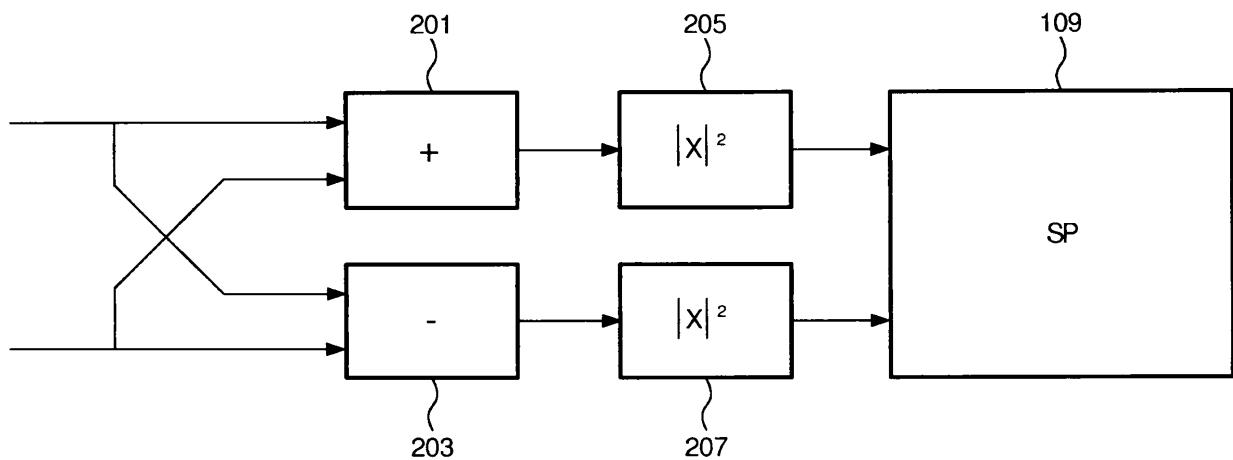
25

30

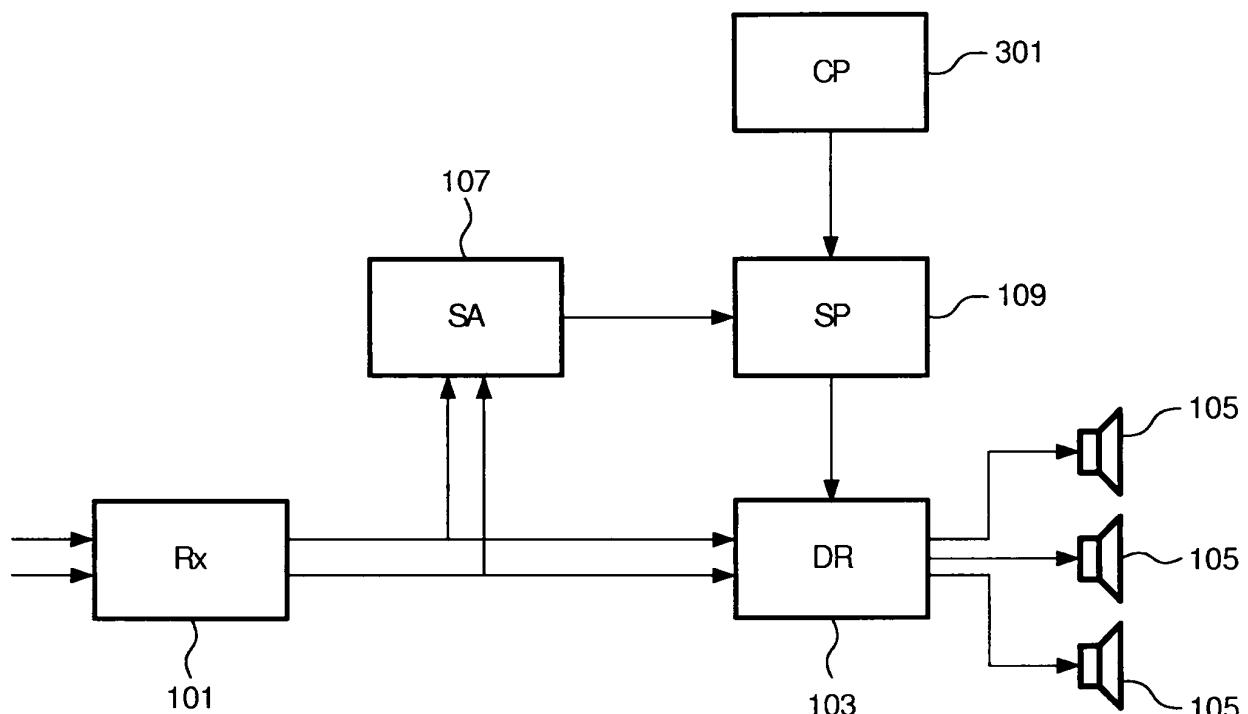
35

40

45



Фиг.2



Фиг.3