

(19)



**Евразийское
патентное
ведомство**

(11) **046317**

(13) **B1**

(12) **ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ**

(45) Дата публикации и выдачи патента
2024.02.28

(21) Номер заявки
202292258

(22) Дата подачи заявки
2020.11.19

(51) Int. Cl. **H04N 19/186** (2014.01)
H04N 19/105 (2014.01)
H04N 19/593 (2014.01)

(54) **ВНУТРЕННЕЕ ПРОГНОЗИРОВАНИЕ ЦВЕТНОСТИ В КОДИРОВАНИИ И
ДЕКОДИРОВАНИИ ВИДЕО**

(31) **2001722.4**

(32) **2020.02.07**

(33) **GB**

(43) **2022.10.07**

(86) **PCT/GB2020/052963**

(87) **WO 2021/156587 2021.08.12**

(71)(73) Заявитель и патентовладелец:
**БРИТИШ БРОДКАСТИНГ
КОРПОРЕЙШН (GB)**

(72) Изобретатель:
**Бланш Марк Горриз, Мрак Марта,
Блази Саверио (GB)**

(74) Представитель:
**Сеитова А.Н. (KZ), Микуцкая Т.Ю.
(RU)**

(56) **US-A1-2019342546**

**LI YUE ET AL: "A Hybrid Neural Network
for Chroma Intra Prediction", 2018 25TH IEEE
INTERNATIONAL CONFERENCE ON IMAGE
PROCESSING (ICIP), IEEE, 7 October 2018
(2018-10-07), pages 1797-1801, XP033454967,
DOI: 10.1109/ICIP.2018.8451396 [retrieved on
2018-08-29] cited in the application abstract page
1798, left-hand column, line 4 - line 16 sections 2, 3.2
and 5**

**EP-A1-2717578
WO-A1-2019201232**

(57) Предложен режим внутреннего прогнозирования для данных цветности, где реализованный нейронной сетью модуль внимания управляет кодированием блока данных цветности в отношении данных яркости для смещенного блока яркости.

046317

B1

046317
B1

Область техники, к которой относится изобретение

Настоящее изобретение относится к кодированию и декодированию видео.

Предшествующий уровень техники

Цифровое видео может быть закодировано для передачи и/или хранения. Такая кодировка имеет преимущество в эффективности, с точки зрения спроса на полосу пропускания канала, требований к памяти, а также обнаружения и/или исправления ошибок.

Методы сжатия видео включают в себя получение блоков видеоданных и использование различных подходов для использования преимущества избыточности. Например, внутреннее прогнозирование представляет собой пространственный подход, который включает в себя кодирование видеоблока со ссылкой на другой (прогнозирующий) видеоблок в том же кадре видеоданных. С другой стороны, внутреннее прогнозирование - это временной подход, включающий кодирование видеоблока видеокadra со ссылкой на соответствующий (прогнозирующий) блок другого кадра видеоданных.

Как правило, видеоданные содержат данные яркости и данные цветности. Данные яркости представляют собой яркость изображения без привязки к цвету. По сути, данные яркости описывают черно-белое изображение, и фактически такое изображение может быть реконструировано только из данных яркости. Данные цветности представляют информацию о цвете, которую можно использовать для восстановления цвета на черно-белом изображении.

При кодировании данных яркости и цветности можно использовать различные подходы. Человеческий глаз гораздо более чувствителен к яркости, чем к цветности; это дает возможность сжимать данные цветности больше, чем данные яркости.

Перечень фигур чертежей

На фиг. 1 показано схематическое представление сети связи в соответствии с вариантом осуществления изобретения;

на фиг. 2 показано схематическое изображение излучателя сети связи с фиг. 1;

на фиг. 3 показана схема, иллюстрирующая кодер, реализованный на излучателе с фиг. 2;

на фиг. 4 показана технологическая схема структуры и функций модуля прогнозирования кодера с фиг. 3;

на фиг. 5 показано схематическое изображение приемника сети связи с фиг. 1;

на фиг. 6 показана схема, иллюстрирующая декодер, реализованный на приемнике с фиг. 4; и;

на фиг. 7 показана технологическая схема структуры и функций модуля прогнозирования декодера с фиг. 6.

Сведения, подтверждающие возможность осуществления изобретения

Универсальное кодирование видео (УКВ) (MPEG-I часть 3) - это стандарт сжатия видео, который в настоящее время разрабатывается Объединенной группой экспертов по видеоматериалам (JVET), объединенной группой экспертов по видеоматериалам MPEG рабочей группы ISO/IEC JTC 1 и VCEG рабочей группы ITU-T.

Внутреннее прогнозирование обычно используется для использования пространственной избыточности внутри кадра путем вычисления прогнозов на основе уже реконструированных выборок в непосредственной близости от текущего блока. Последний проект УКВ (далее по тексту - "УКВ") позволяет использовать большое количество возможных режимов внутреннего прогнозирования (далее - "внутренние режимы") для прогнозирования компонента яркости, включая до 65 угловых (направленных) режимов, планарное прогнозирование и DC-прогнозирование, а также другие передовые методы прогнозирования.

Это большое количество опций достигается за счет значительного объема сигнализирующих данных, чтобы указать, какой из внутренних режимов был использован в блоке данных яркости. И, наоборот, чтобы ограничить влияние сигнализирующих внутренних режимов, используется уменьшенное количество внутренних режимов для внутреннего прогнозирования компонентов цветности. Сюда входит подмножество режимов, доступных для прогнозирования яркости, включая, например, Производный режим (DM, соответствующий использованию режима, используемого для прогнозирования яркости), а также небольшой набор фиксированных кандидатов, включая планарный, DC, чисто горизонтальный и чисто вертикальный режимы.

Кроме того, УКВ определяет использование внутренних режимов Межкомпонентной линейной модели (МКЛМ). При использовании МКЛМ компонент цветности прогнозируется на основе уже реконструированных выборок яркости с использованием линейной модели. Параметры модели выводятся с помощью простой линейной регрессии, применяемой к реконструированным соседним выборкам яркости и цветности, извлеченным из верхних левых границ блока кодирования. В настоящее время в УКВ используются три режима МКЛМ, в том числе общий режим ЛМ, использующий доступные выборки в верхнем левом углу, и два направленных режима ЛМ, использующие расширенные массивы эталонных выборок, извлеченных исключительно сверху или слева от текущего блока соответственно.

Было установлено, что использование ЛМ-прогнозирования эффективно для повышения эффективности внутреннего прогнозирования цветности. Однако варианты осуществления изобретения, описанные в настоящем документе, направлены на получение дополнительных преимуществ от других подходов.

Здесь замечено, что использование простых линейных прогнозов может быть ограничивающим. Варианты осуществления изобретения, описанные в настоящем документе, могут обеспечить улучшенную производительность по сравнению с существующими технологиями за счет использования более сложных архитектур, основанных на механизмах Машинного обучения (МО).

Гибридная нейронная сеть для внутреннего прогнозирования цветности описана в разделе "Гибридная нейронная сеть для внутреннего прогнозирования цветности" (Ли, Юэ, Ли Ли, Чжу Ли, Цзяньчао Ян, Нин Сюй, Донг Лю и Хоуцян Ли; 2018 25-я Международная конференция IEEE по обработке изображений (ICIP), стр. 1797-1801. IEEE, 2018)], в котором конволюционная нейронная сеть (CNN) для извлечения признаков из реконструированных выборок яркости текущего блока была объединена с существующими полносвязными архитектурами для извлечения межкомпонентных корреляций между соседними выборками яркости и цветности. Используя такую архитектуру, можно получить сложные нелинейные отображения для сквозного прогнозирования каналов Сb и Сr. Однако такая архитектура обычно игнорирует пространственную корреляцию граничащих выборок при прогнозировании соответствующих местоположений в прогнозируемом блоке.

Варианты осуществления изобретения, описанные в настоящем документе, обеспечивают архитектуру нейронной сети, которая включает в себя модуль внимания для управления вкладом каждой эталонной соседней выборки при вычислении прогноза для местоположения каждой выборки, гарантируя, что пространственная информация учитывается при вычислении прогноза. Подобно вышеупомянутой гибридной нейронной сети, способ согласно раскрытому здесь варианту осуществления изобретения использует схему, основанную на трех ветвях сети, которые объединяются для создания выборок прогнозирования. Первые две ветви работают одновременно для извлечения признаков из доступных реконструированных выборок, включая уже реконструированный блок яркости, а также соседние эталонные выборки яркости и цветности. Первая ветвь (называемая межкомпонентной граничной ветвью) предназначена для извлечения межкомпонентной информации из соседних реконструированных выборок с использованием расширенного эталонного массива слева от текущего блока и над ним. Вторая ветвь (называемая конволюционной ветвью яркости) извлекает пространственные шаблоны над совмещенным реконструированным блоком яркости, применяя конволюционные операции. Варианты осуществления изобретения, описанные в настоящем документе, представляют процесс объединения признаков, выдаваемых первой и второй ветвями, с использованием модуля внимания. Примеры модуля внимания будут описаны ниже. Наконец, выходные данные модуля внимания подаются в третью ветвь сети для получения результирующих прогнозов выходной цветности Сb и Сr.

Данный подход, где нейронные сети используются для реализации прогнозирования цветности, можно рассматривать как расширение существующих режимов прогнозирования цветности, определенных в УКВ. Таким образом, в варианте осуществления изобретения может быть реализована архитектура на основе нейронной сети, чтобы конкурировать с установленными режимами для поддерживаемых настроек 4×4, 8×8, 16×16. Затем для каждой единицы прогнозирования кодер выбирает между традиционными угловыми режимами, моделями ЛМ или раскрытым режимом нейронной сети; путем минимизации критерия стоимости искажения скорости.

Вышеупомянутая компоновка может быть реализована в сети видеосвязи, предназначенной для обработки видеопрезентации с помощью методов кодирования, позволяющих передавать (или сохранять) его для декодирования устройством воспроизведения.

В общих чертах, аспект настоящего изобретения обеспечивает внутреннее прогнозирование выборок цветности на основе выборок яркости и маски внимания, при этом маска внимания настраивается посредством конволюционной нейронной сети.

Как проиллюстрировано на фиг. 1, устройство проиллюстрировано как включающее схематическую сеть видеосвязи 10, в которой излучатель 20 и приемник 30 обмениваются данными через канал связи 40. На практике канал связи 40 может включать спутниковый канал связи, кабельную сеть, наземную сеть радиовещания, канал связи, реализованный на основе POTS, например, используемый для предоставления интернет-услуг в жилых помещениях и помещениях малого бизнеса, волоконно-оптические системы связи или комбинация любого из вышеперечисленного и любого другого средства связи.

Кроме того, изобретение также распространяется на передачу посредством физической передачи носителя данных, на котором хранится машиночитаемая запись кодированного битового потока, для прохождения к соответствующим образом сконфигурированному приемнику, способному считывать носитель и получать из него битовый поток. Примером этого является предоставление цифрового универсального диска (DVD) или его эквивалента. Следующее описание сосредоточено на передаче сигнала, например, с помощью носителя электронного или электромагнитного сигнала, но не должно рассматриваться как исключаящее вышеупомянутый подход, включающий носители данных.

Как показано на фиг. 2, излучатель 20 по структуре и функции представляет собой компьютерное устройство. Он может иметь некоторые признаки, общие с компьютерным устройством общего назначения, но некоторые признаки могут быть специфичными для реализации, учитывая специализированную функцию, для которой должен быть использован излучатель 20. Читатель поймет, какие функции могут быть обусловлены общим назначением, а для каких может потребоваться конфигурация, специальная

для использования в видеоизлучателе.

Таким образом, излучатель 20 содержит процессор графического вывода (GPU) 202, сконфигурированный для конкретного использования в обработке графики и аналогичных операциях. Излучатель 20 также содержит один или более других процессоров 204, либо в целом подготовленных, либо сконфигурированных для других целей, таких как математические операции, обработка звука, управление каналом связи и так далее.

Входной интерфейс 206 обеспечивает средство для приема действий пользователя по вводу. Такие действия пользователя могут, например, быть вызваны взаимодействием пользователя с конкретным блоком ввода, содержащим одну или более кнопок управления и/или переключателей, клавиатурой, мышью или другим указывающим устройством, блоком распознавания речи, выполненным с возможностью приема и обработки речи в команды управления, сигнальным процессором, выполненным с возможностью приема и управления процессами от другого устройства, такого как планшет или смартфон, или приемником дистанционного управления. Этот перечень будет воспринят как неисчерпывающий, и читатель может предусмотреть другие формы ввода, будь то инициированные пользователем или автоматизированные.

Аналогично, выходной интерфейс 214 выполнен с возможностью обеспечения средства для вывода сигналов пользователю или другому устройству. Такой вывод может включать в себя сигнал дисплея для управления локальным блоком отображения видеосигнала (VDU) или любым другим устройством.

Интерфейс 208 передачи данных реализует канал связи, будь то широкополосный или сквозной, с одним или более получателями сигналов. В контексте настоящего варианта осуществления интерфейс передачи данных выполнен с возможностью вызывать излучение сигнала, несущего битовый поток, определяющий видеосигнал, кодируемый излучателем 20.

Процессоры 204, и, в частности, для пользы настоящего изобретения, GPU 202, выполнены с возможностью выполнения компьютерных программ при работе кодера. При этом используются средства хранения данных, предоставляемые запоминающим устройством большой емкости 208, которое выполнено с возможностью обеспечения крупномасштабного хранения данных, хотя и на основе относительно медленного доступа, и на практике будет хранить компьютерные программы и, в текущем контексте, данные видеопрезентации при подготовке к выполнению процесса кодирования.

Постоянное запоминающее устройство (ПЗУ) 210 предварительно сконфигурировано с исполняемыми программами, предназначенными для обеспечения ядра функциональности излучателя 20, а Оперативное запоминающее устройство (ОЗУ) 212 предусмотрено для быстрого доступа и хранения данных и программных инструкций в процессе выполнения компьютерной программы.

Далее будет описана функция излучателя 20 со ссылкой на фиг. 3. На фиг. 3 показан процесс обработки, выполняемый кодером, реализованным на излучателе 20 с помощью исполняемых команд, на файле данных, представляющем собой видеопрезентацию, содержащую множество кадров для последовательного отображения в виде последовательности изображений.

Файл данных может также содержать информацию о воспроизведении звука, сопровождающую видеопрезентацию, и прочую дополнительную информацию, такую как электронная информация о руководстве по программе, субтитры или метаданные для обеспечения возможности каталогизации презентации. Обработка этих аспектов файла данных не имеет отношения к настоящему изобретению.

Как показано на фиг. 3, текущее изображение или кадр в последовательности изображений передается в модуль 230 разделения, где он разбивается на прямоугольные блоки заданного размера для обработки кодером. Эта обработка может быть последовательной или параллельной. Подход может зависеть от возможностей обработки конкретной реализации.

Каждый блок затем вводится в модуль 232 прогнозирования, который стремится устранить временное и пространственное избыточное количество данных, присутствующее в последовательности, и получить сигнал прогнозирования с использованием ранее закодированного содержимого. Информация, позволяющая вычислять такое прогнозирование, кодируется в битовом потоке. Эта информация должна быть достаточной для выполнения расчетов, включая возможность извлечения посредством анализа на приемнике другой информации, необходимой для завершения прогнозирования. Дальнейшие подробности модуля прогнозирования данного варианта осуществления изобретения будут предоставлены ниже.

Сигнал прогнозирования вычитается из исходного сигнала для получения остаточного сигнала. Затем его вводят в модуль 234 преобразования, который пытается дополнительно уменьшить пространственное избыточное количество данных внутри блока путем использования более подходящего представления данных. Как отмечено выше, в данном варианте осуществления изобретения предполагается, что трансформация домена не может быть реализована для каждого набора остатков, и в зависимости от вероятной эффективности выполнения этого, вместо этого может быть реализован пропуск преобразования. Использование пропуска преобразования может сигнализироваться в битовом потоке.

Затем результирующий сигнал обычно квантуется модулем 236 квантования, и, наконец, результирующие данные, сформированные из коэффициентов и информации, необходимой для вычисления прогнозирования для текущего блока, вводятся в модуль 238 энтропийного кодирования, который использует статистическую избыточность для представления сигнала в компактной форме с помощью коротких

двоичных кодов. Читатель заметит, что если использовался пропуск с преобразованием, используется остаточное кодирование с пропуском преобразования (TSRC).

При повторном воздействии средства кодирования излучателя 20 может быть построен битовый поток информационных элементов блока для передачи приемнику или множеству приемников, в зависимости от обстоятельств. Поток битов может также содержать информационные элементы, которые применяются к множеству информационных элементов блока и, таким образом, удерживаются в синтаксисе потока битов, независимо от информационных элементов блока. Примеры таких информационных элементов включают в себя варианты конфигурации, параметры, применимые к последовательности кадров, и параметры, относящиеся к видеопрезентации в целом.

Далее модуль 232 прогнозирования будет описан более подробно со ссылкой на фиг. 4. Как будет понятно, это всего лишь пример, и могут быть рассмотрены другие подходы в рамках настоящего изобретения и прилагаемой формулы изобретения.

Следующий процесс выполняется на каждом блоке в кадре.

Модуль 232 прогнозирования сконфигурирован для обработки входных выборок яркости и выборок цветности для получения данных прогнозирования, которые затем можно использовать для создания остатков для дальнейшей обработки, как описано выше.

Теперь опишем структуру и функции модуля 232 прогнозирования со ссылкой на фиг. 4.

В целом, модуль 232 прогнозирования содержит архитектуру нейронной сети, которая включает в себя модуль внимания 250 для управления вкладом каждой эталонной соседней выборки при вычислении прогноза для местоположения каждой выборки, обеспечивая учет пространственной информации при вычислении прогноза. Модуль 232 прогнозирования имеет три ветви сети, которые объединяются для создания выборок прогнозирования. Как показано, первая и вторая ветви реализуются параллельно для извлечения признаков из доступных выборок, включая блок яркости, а также соседние эталонные выборки яркости и цветности. Первая ветвь (называемая межкомпонентной граничной ветвью) предназначена для извлечения межкомпонентной информации из соседних реконструированных выборок с использованием расширенного эталонного массива слева от текущего блока и над ним. Вторая ветвь (называемая конволюционной ветвью яркости) извлекает пространственные шаблоны над совмещенным блоком яркости, применяя конволюционные операции. Модуль внимания объединяет функции, выводимые первой и второй ветвями.

Вывод модуля 250 внимания, наконец, подается в третью ветвь сети для получения результирующих прогнозов цветности C_b и C_r на выходе.

Обучение "на основе внимания", используемое в рамках глубокого обучения, используется для повышения производительности обученных сетей в сложных задачах прогнозирования. Модели внимания могут уменьшить количество сложных задач, прогнозируя более мелкие "области внимания", которые обрабатываются последовательно, чтобы способствовать более эффективному обучению. В частности, "собственное внимание" (или "внутреннее внимание") используется для оценки влияния определенных входных переменных на выходные данные, при этом прогноз вычисляется с упором на наиболее релевантные элементы той же последовательности. Модуль 250 внимания описанного варианта осуществления изобретения пытается объединить признаки из первой и второй ветвей сети, чтобы оценить влияние каждой входной переменной в отношении их пространственного положения.

Этот подход может устранить важное ограничение аналогичных методов для вычисления выборок с внутренним прогнозированием на основе МО, которые обычно полностью отбрасывают информацию, содержащуюся в пространственном расположении входных и выходных выборок.

В примере использования варианта осуществления изобретения предполагается, что внутреннее прогнозирование уже было выполнено для компонента яркости и что окончательные реконструированные выборки яркости доступны. В зависимости от используемой подвыборки цветности этот блок затем подвергается понижению дискретизации с использованием обычных фильтров, доступных в УКВ, для получения блока совмещенных выборок яркости с теми же размерами $N \times N$ в пространстве, что и соответствующие компоненты цветности. Для простоты в этой версии алгоритма рассматриваются только квадратные блоки. Можно использовать блоки большего размера, а также усовершенствованные фильтры понижения частоты дискретизации без потери общности.

Совмещенный блок яркости обозначается $X \in \mathbb{R}^{N \times N}$. Массив эталонных выборок в левом верхнем углу текущего блока обозначается $V_c \in \mathbb{R}^{2N+1}$, $b = 2N+1$, где $c = Y, C_b, C_r$ для трех компонентов соответственно, где V построен следующим образом.

Во-первых, выборки на левой границе рассматриваются, начиная с самой нижней выборки; затем считается угол; затем рассматриваются выборки сверху, начиная с самой левой выборки. Если некоторые эталонные выборки недоступны, они дополняются заранее определенным значением.

Наконец, межкомпонентный объем, полученный конкатенацией трех эталонных массивов V_y, V_{C_b}, V_{C_r} , обозначается $S \in \mathbb{R}^{3 \times b}$. После применения межкомпонентной граничной ветви над S и конволюционной ветви яркости над X получают преобразованные объемы признаков $\tilde{X} \in \mathbb{R}^{N \times N \times D = N^2 \times C}$ и $\tilde{S} \in \mathbb{R}^{b \times D}$ соответственно.

Рисунок 4 иллюстрирует архитектуру подсети для обеих ветвей (для частного случая блока 8×8). В таблице представлено раскрытие выбранных сетевых гиперпараметров для блоков других размеров. Подробные сведения о построении и реализации обеих ветвей не приводятся из-за использования структуры, аналогичной вышеупомянутому современному гибриднему методу на основе нейронной сети.

Таблица - гиперпараметры сети в зависимости от размера блока. В частности, глубина конволюционных ядер применяется к различным ветвям предлагаемой архитектуры.

Ветка	4 x 4	8 x 8	16 x 16
Межкомпонентная границная ветвь	16, 32	32, 64	64, 96
Конволюционная ветвь яркости	32, 32	64, 64	96, 96
Модуль внимания	16, 16, 32	16, 16, 64	16, 16, 96
Выходная ветвь	32, 2	64, 2	96, 2

Уже в рамках предлагаемого модуля слияния на основе внимания каждая из этих двух карт признаков свернута с использованием ядра 1×1 , чтобы спроецировать их на два соответствующих уменьшенных пространства признаков. В частности, \tilde{S} свернут с фильтром $W_F \in R^{h \times D}$ для получения h -мерной матрицы признаков F . Аналогичным образом \tilde{X} свернут с фильтром $W_G \in R^{h \times C}$ для получения h -мерной матрицы признаков G . Две матрицы перемножаются, чтобы получить карту предварительного внимания $M = G^T F$. Наконец, получается матрица внимания $A \in R^{N^2 \times b}$, применяя операцию многопеременной логистической функции к каждому элементу M для произведения вероятности того, что каждое местоположение границы сможет прогнозировать местоположение каждой выборки в блоке. Формально, для каждого элемента $\alpha_{i,j}$ в A , где $j = 0 \dots N^2 - 1$ представляет собой местоположение выборки в прогнозируемом блоке, а $j = 0 \dots b - 1$ представляет местоположение эталонной выборки, применяется следующее:

$$\alpha_{i,j} = \frac{e^{\frac{m_{i,j}}{T}}}{\sum_{i=0}^{N^2-1} e^{\frac{m_{i,j}}{T}}}$$

где T - температурный параметр многопеременной логистической функции, управляющий гладкостью генерируемых вероятностей, при $0 \leq T \leq 1$. Обратите внимание, что чем меньше значение T , тем более локализованными являются полученные области внимания, что приводит к соответственно меньшему количеству граничных выборок, вносящих вклад в данное прогнозируемое местоположение.

Весовая сумма вклада каждой эталонной выборки в прогнозирование заданного выходного местоположения получается путем вычисления скалярного произведения межкомпонентных граничных признаков \tilde{S} и матрицы A внимания, или формально $\langle \tilde{S}, A \rangle$, где $\langle \cdot \rangle$ - скалярное произведение. Для дальнейшего уточнения $\langle \tilde{S}, A \rangle$ эту взвешенную сумму можно умножить на выход ветви яркости. Для этого выход ветви яркости должен быть преобразован для изменения его размеров посредством конволюции 1×1 с использованием матрицы $W_{\tilde{X}} \in R^{D \times C}$, чтобы получить преобразованное представление $\tilde{X} \in R^{N^2 \times D}$. Эта операция применяется для обеспечения того, чтобы размеры выходных данных конволюционной ветви яркости были совместимы с $\langle \tilde{S}, A \rangle$. Наконец, выход модели внимания $\theta \in R^{N^2 \times D}$ получается следующим образом: $\tilde{X} \circ \langle \tilde{S}, A \rangle$, где \circ - поэлементное произведение.

Например, для параметра измерения h в модуле внимания установлено значение 16 для всех обученных моделей, что является оптимальным компромиссом между производительностью и сложностью. Значение $T = 0,5$ является перекрестным, обеспечивая оптимальный баланс между информативными выборками и зашумленными от остальных граничных местоположений.

Дополнительно следует отметить, что модуль 232 прогнозирования может быть сконфигурирован для использования одного из множества predetermined режимов взаимного прогнозирования, но взаимное прогнозирования не имеет отношения к объему настоящего изобретения. Более того, приведенное выше изобретение не исключает возможности реализации существующих режимов внутреннего прогнозирования, указанных, например, в УКВ.

Структурная архитектура приемника показана на фиг. 5. Он имеет элементы, представляющие собой компьютерное устройство. Таким образом, приемник 30 содержит графический процессор 302, сконфигурированный для конкретного использования в обработке графики и аналогичных операциях. Приемник 30 также содержит один или более других процессоров 304, либо в целом подготовленных, либо сконфигурированных для других целей, таких как математические операции, обработка звука, управление каналом связи и так далее.

Как будет понятно читателю, приемник 30 может быть реализован в виде приставки, переносного персонального электронного устройства, персонального компьютера или любого другого устройства, подходящего для воспроизведения видеопрезентаций.

Входной интерфейс 306 обеспечивает средство для приема действий пользователя по вводу. Такие действия пользователя по вводу могут, например, быть вызваны взаимодействием пользователя с конкретным блоком ввода, содержащим одну или более кнопок управления и/или переключателей, клавиатурой, мышью или другим указывающим устройством, блоком распознавания речи, выполненным с воз-

возможностью приема и обработки речи в команды управления, сигнальным процессором, выполненным с возможностью приема и управления процессами от другого устройства, такого как планшет или смартфон, или приемником дистанционного управления. Этот перечень будет воспринят как неисчерпывающий, и могут быть предусмотрены другие формы ввода, будь то инициированные пользователем или автоматизированные.

Аналогичным образом, выходной интерфейс 314 выполнен с возможностью обеспечения средства для вывода сигналов пользователю или другому устройству. Такой вывод может включать в себя телевизионный сигнал в подходящем формате для управления местным телевизионным устройством.

Интерфейс 308 передачи данных реализует канал связи, будь то широковещательный или сквозной, с одним или более получателями сигналов. В контексте настоящего варианта осуществления интерфейс передачи данных выполнен с возможностью вызывать излучение сигнала, несущего битовый поток, определяющий видеосигнал, кодируемый приемником 30.

Процессоры 304, и, в частности, для пользы настоящего изобретения, графический процессор GPU 302, выполнены с возможностью выполнения компьютерных программ при работе приемника. При этом используются средства хранения данных, предоставляемые запоминающим устройством 308 большой емкости, которое выполнено с возможностью обеспечения крупномасштабного хранения данных, хотя и на основе относительно медленного доступа, и на практике будет хранить компьютерные программы и, в текущем контексте, данные видеопрезентации, полученные в результате выполнения процесса приема.

Постоянное запоминающее устройство (ПЗУ) 310 предварительно сконфигурировано с исполняемыми программами, предназначенными для обеспечения ядра функциональности приемника 30, а Оперативное запоминающее устройство 312 предусмотрено для быстрого доступа и хранения данных и программных инструкций в процессе выполнения компьютерной программы.

Далее будет описана функция приемника 30 со ссылкой на фиг. 6. На фиг. 6 показан процесс обработки, выполняемый декодером, реализованным на приемнике 20 с помощью исполняемых инструкций, на битовом потоке, принятом в приемнике 30, содержащем структурированную информацию, из которой может быть получена видеопрезентация, содержащая реконструкцию кадров, закодированных функциональностью кодера излучателя 20.

Процесс декодирования, показанный на фиг. 6, направлен на то, чтобы обратить процесс, выполняемый на кодере. Читателю будет понятно, что это не означает, что процесс декодирования является строго обратным процессу кодирования.

Принятый битовый поток содержит последовательность закодированных информационных элементов, причем каждый элемент связан с некоторым блоком. Информационный элемент блока декодируют в модуле 330 энтропийного декодирования для получения блока коэффициентов и информации, необходимой для вычисления прогнозирования для текущего блока. Блок коэффициентов, как правило, деквантизируется в модуле 332 деквантизации и, как правило, обратно преобразовывается в пространственную область с помощью модуля 334 преобразования, если только декодеру не сообщается о пропуске преобразования.

Как отмечалось выше, читатель поймет, что декодирование энтропии, деквантизация и обратное преобразование должны использоваться в приемнике только в том случае, если на излучателе использовались кодирование энтропии, квантование и преобразование, соответственно.

Сигнал прогнозирования генерируется, как и раньше, из ранее декодированных выборок из текущих или предыдущих кадров и с использованием информации, декодированной из битового потока, модулем 336 прогнозирования. Затем получают реконструкцию исходного блока изображения из декодированного остаточного сигнала и вычисленного блока прогнозирования в блоке 338 реконструкции. В вариантах осуществления, модуль 336 прогнозирования реагирует на информацию о потоке битов, сигнализирующую об использовании внутреннего прогнозирования и, если такая информация присутствует, на показания из этой информации о потоке битов, которые позволяют декодеру определить, какой режим внутреннего прогнозирования был использован и, таким образом, какой метод прогнозирования должен быть использован при реконструкции выборки данных блока.

При повторном воздействии функциональности декодирования на последовательно принятые информационные элементы блока, графические блоки могут быть реконструированы в кадры, которые затем могут быть собраны для создания видеопрезентации для воспроизведения.

Пример алгоритма декодера, дополняющий алгоритм кодера, описанный ранее, показан на фиг. 7. Этот алгоритм относится к использованию режима прогнозирования цветности, реализованного нейронной сетью, если реализация позволяет также реализовать другие режимы прогнозирования цветности, как такие режимы в настоящее время определены в УКВ, то это также может быть сообщено декодеру.

Как отмечалось ранее, декодирующая функциональность приемника 30 извлекает из битового потока последовательность информационных элементов блока, закодированных средством кодирования излучателя 20, определяющую информацию блока и сопутствующую информацию конфигурации.

Следующий процесс выполняется для каждого закодированного блока в кадре, подлежащим декодированию.

Модуль 332 прогнозирования выполнен с возможностью обработки входных закодированных дан-

ных яркости и закодированных данных цветности для получения реконструированных выборочных данных яркости и цветности, которые затем можно использовать для дальнейшей обработки, как описано выше.

Теперь опишем структуру и функцию модуля 332 прогнозирования со ссылкой на фиг. 7.

В целом, модуль 332 прогнозирования содержит архитектуру нейронной сети, которая включает в себя модуль 350 внимания для управления вкладом каждой эталонной соседней выборки при вычислении прогноза для местоположения каждой выборки, обеспечивая учет пространственной информации при вычислении прогноза. Модуль прогнозирования 332 имеет три ветви сети, которые объединяются для создания выборок прогнозирования. Как показано, первая и вторая ветви реализуются параллельно для извлечения признаков из доступных реконструированных выборок, включая уже реконструированный блок яркости, а также соседние эталонные выборки яркости и цветности. Первая ветвь (называемая межкомпонентной граничной ветвью) предназначена для извлечения межкомпонентной информации из соседних реконструированных выборок с использованием расширенного эталонного массива слева от текущего блока и над ним. Вторая ветвь (называемая конволюционной ветвью яркости) извлекает пространственные шаблоны над совмещенным блоком яркости, применяя конволюционные операции. Модуль внимания объединяет функции, выводимые первой и второй ветвями.

Вывод модуля 350 внимания, наконец, подается в третью ветвь сети для получения результирующих прогнозирование цветности C_b и C_t на выходе.

В декодере используется тот же подход к работе модуля внимания, что и в кодере. Ожидается, что модуль внимания в декодере обучен таким же образом, как и в кодере, для воспроизведения решений внутреннего прогнозирования, принятых в кодере. Это обеспечивает восстановление данных цветности в декодере.

Следует понимать, что изобретение не ограничивается описанными выше вариантами, и различные модификации и усовершенствования могут быть сделаны без отступления от описанных концепций. Кроме случаев взаимоисключения, любой из признаков может быть использован отдельно или в комбинации с любыми другими признаками, и настоящее изобретение распространяется и включает в себя все комбинации и подкомбинации одного или более признаков, описанных в настоящем документе.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ декодирования видеоданных, который включает в себя:
 - извлечение эталонных выборок из реконструированных выборок яркости и реконструированных выборок цветности; и
 - построение по меньшей мере одного блока выборок прогнозирования цветности из эталонных выборок, где построение выборок прогнозирования цветности зависит от пространственного расположения эталонных выборок, и в котором построение выборок прогнозирования цветности зависит от использования модуля внимания.
2. Способ по п.1, в котором эталонные выборки включают реконструированные выборки цветности из соседних блоков указанного блока выборок прогнозирования цветности.
3. Способ по п.1 или 2, в котором эталонные выборки включают реконструированные выборки яркости, расположенные рядом с блоком выборок прогнозирования цветности.
4. Способ по п.1, 2 или 3, в котором эталонные выборки включают реконструированные выборки яркости из соседних блоков блока расположенных рядом реконструированных выборок яркости.
5. Способ по любому из предшествующих пунктов, в котором модуль внимания сконфигурирован как глубокая нейронная сеть.
6. Способ по любому из предшествующих пунктов, в котором построение выборок прогнозирования цветности зависит от вычисления межкомпонентной информации.
7. Способ по любому из предшествующих пунктов, в котором построение выборок прогнозирования цветности зависит от извлечения данных пространственного изображения над блоком данных яркости с помощью по меньшей мере одной конволюционной операции.
8. Способ по любому из предшествующих пунктов, включающий управление, с помощью модуля внимания, вкладом каждой эталонной соседней выборки в вычислении прогнозирования для местоположения выборки.
9. Способ по любому из предшествующих пунктов, в котором дополнительно реализуют один или несколько других режимов построения по меньшей мере одного блока данных цветности, и в котором режим построения по меньшей мере одного блока данных цветности определяется на основе принятого сигнала.
10. Декодер для декодирования видеоданных, включающий в себя:
 - экстрактор эталонных выборок для извлечения эталонных выборок из реконструированных выборок яркости и реконструированных выборок цветности; и
 - конструктор выборок прогнозирования цветности для построения по меньшей мере одного блока выборок прогнозирования цветности из эталонных выборок, где конструктор выборок прогнозирования

цветности выполнен с возможностью выборки прогнозирования цветности в зависимости от пространственного расположения эталонных выборок, в котором конструктор выборок прогнозирования цветности выполнен с возможностью построения выборки прогнозирования цветности в зависимости от использования модуля внимания.

11. Декодер по п.10, в котором эталонные выборки включают реконструированные выборки цветности из соседних блоков блока выборок прогнозирования цветности.

12. Декодер по п.10 или 11, в котором эталонные выборки включают реконструированные выборки яркости, расположенные рядом с блоком выборок прогнозирования цветности.

13. Декодер по п.10, 11 или 12, в котором эталонные выборки включают реконструированные выборки яркости из соседних блоков блока расположенных рядом реконструированных выборок яркости.

14. Декодер по любому из пп.10-13, в котором модуль внимания сконфигурирован как глубокая нейронная сеть.

15. Декодер по любому из пп.10-14, в котором построение выборок прогнозирования цветности зависит от вычисления межкомпонентной информации.

16. Декодер по любому из пп.10-15, в котором построение выборок прогнозирования цветности зависит от извлечения данных пространственного изображения над блоком данных яркости с помощью по меньшей мере одной конволюционной операции.

17. Декодер по любому из пп.10-16, в котором модуль внимания выполнен с возможностью управлять вкладом каждой эталонной соседней выборки в вычислении прогнозирования для местоположения выборки.

18. Декодер по любому из пп.10-17, в котором конструктор выборок прогнозирования цветности выполнен с возможностью работать в одном или более других режимах построения по меньшей мере одного блока данных цветности, и где режим построения по меньшей мере одного блока данных цветности определен на основе принятого сигнала.

19. Способ кодирования видеоданных, содержащих данные яркости и данные цветности, который включает в себя:

кодирование по меньшей мере одного блока данных яркости, чтобы генерировать данные остатков яркости по отношению к данным эталонной выборки яркости; и

кодирование по меньшей мере одного блока данных цветности, расположенного рядом с блоком данных яркости, для генерирования данных остатков цветности относительно данных эталонной выборки цветности, причем кодирование блока данных цветности включает в себя:

построение по меньшей мере одного блока выборок прогнозирования цветности из эталонных выборок в зависимости от пространственного расположения эталонных выборок, и в котором построение выборок прогнозирования цветности зависит от использования модуля внимания.

20. Кодер, предназначенный для осуществления способа по п.19.

21. Компьютерный программный продукт, содержащий программные инструкции, которые при выполнении компьютером обеспечивают осуществление способа по пп.1-9 или п.19.

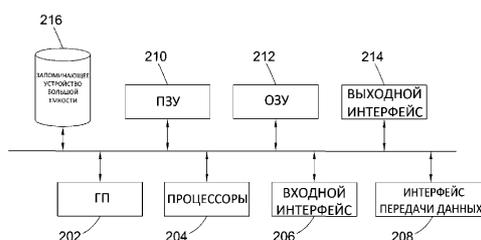
22. Сигнал, несущий информацию, закодированную способом по п.19.

23. Сигнал, несущий информацию, которая может быть декодирована способом по любому из пп.1-9.



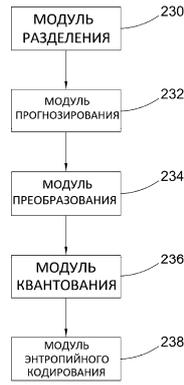
10 ↗

Фиг. 1

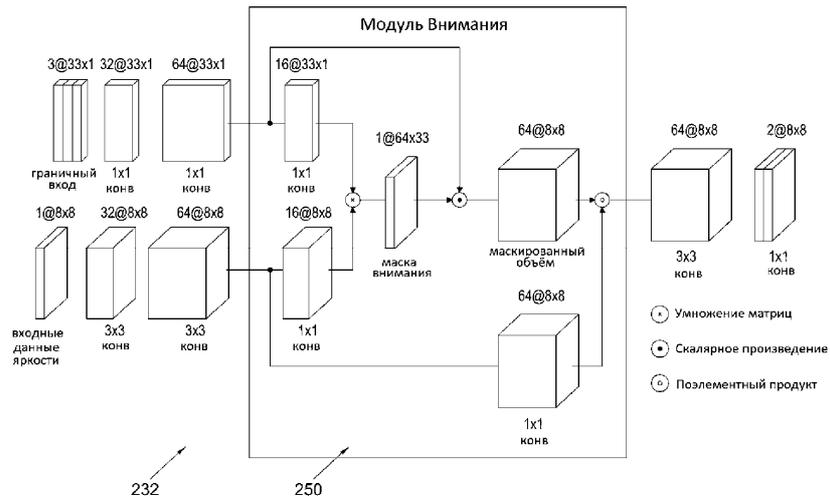


20 ↗

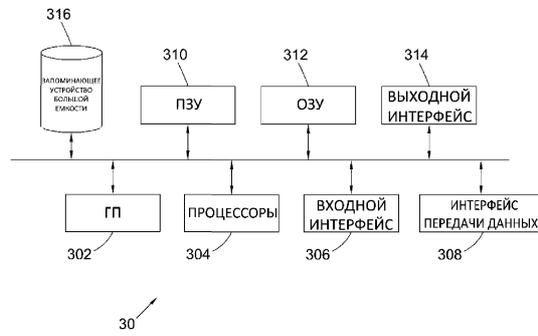
Фиг. 2



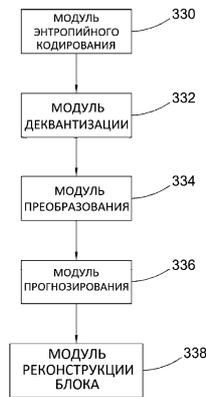
Фиг. 3



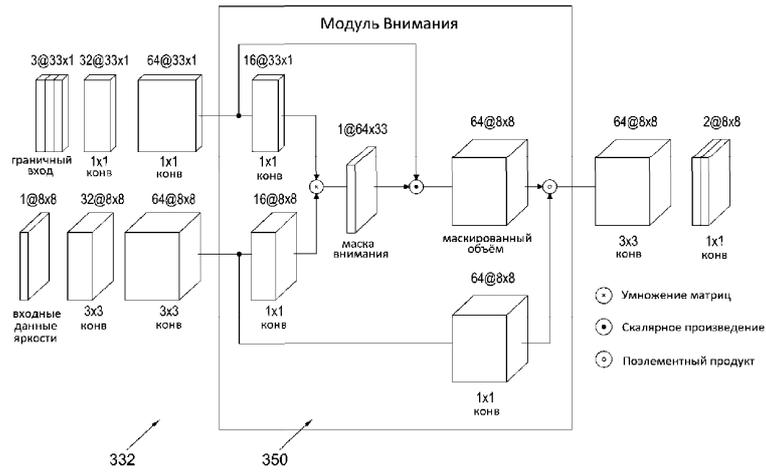
Фиг. 4



Фиг. 5



Фиг. 6



Фиг. 7

