



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ,
ПАТЕНТАМ И ТОВАРНЫМ ЗНАКАМ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(21), (22) Заявка: **2005111002/09**, **14.04.2005**

(24) Дата начала отсчета срока действия патента:
14.04.2005

(30) Конвенционный приоритет:
15.04.2004 US 10/826,842

(43) Дата публикации заявки: **20.10.2006**

(45) Опубликовано: **10.05.2009** Бюл. № 13

(56) Список документов, цитированных в отчете о поиске: **US 6654503 B1**, **25.11.2003. MUKERJEE K. Decaying tree lossless coding for image and video, Data Compression Conference, 2004. Proceedings. DCC 2004 Snowbird, March 23-25 2004. MEMON N. et al. Recent Developments in Context-Based Predictive Techniques for Lossless Image Compression, The Computer Journal, Oxford University Press, vol.40, no.2/3, 1997, (см. прод.)**

Адрес для переписки:
**129090, Москва, ул. Б.Спасская, 25, стр.3,
ООО "Юридическая фирма Городисский и
Партнеры", пат.пов. Ю.Д.Кузнецову,
рег.№ 595**

(72) Автор(ы):
МУКЕРДЖИ Кунал (US)

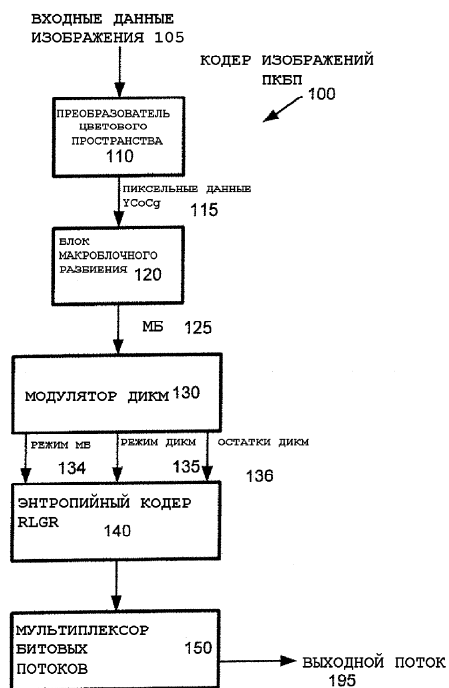
(73) Патентообладатель(и):
МАЙКРОСОФТ КОРПОРЕЙШН (US)

(54) ПРЕДСКАЗАТЕЛЬНОЕ КОДИРОВАНИЕ БЕЗ ПОТЕРЬ ДЛЯ ИЗОБРАЖЕНИЙ И ВИДЕО

(57) Реферат:

Изобретение относится к кодированию без потерь для изображений и видео. Техническим результатом является создание эффективного способа энтропийного кодирования без потерь, несмотря на различие характеристик содержимого фотографических и графических изображений. Указанный результат достигается тем, что кодирование с предсказанием без потерь действует на основе

макроблоков для совместимости с существующими кодеками изображения и видео и предусматривает выбор и применение одного из множественных доступных режимов дифференциальной импульсно-кодовой модуляции (ДИКМ) к отдельным макроблокам для создания остатков ДИКМ, имеющих более близкое к оптимальному распределение для энтропийного кодирования Голомба-Райса по длине серии. 5 н. и 14 з.п. ф-лы, 7 ил.



ФИГ. 1

(56) (продолжение):

с.128-136. RU 2093957 C2, 20.10.1997. US 5774590 A, 30.06.1998. JP 2001077700 A, 23.03.2001.



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY,
PATENTS AND TRADEMARKS

(51) Int. Cl.
H04N 7/44 (2006.01)

(12) ABSTRACT OF INVENTION

(21), (22) Application: **2005111002/09, 14.04.2005**

(24) Effective date for property rights:
14.04.2005

(30) Priority:
15.04.2004 US 10/826,842

(43) Application published: **20.10.2006**

(45) Date of publication: **10.05.2009 Bull. 13**

Mail address:

**129090, Moskva, ul. B.Spasskaja, 25, str.3, OOO
"Juridicheskaja firma Gorodisskij i Partnery",
pat.pov. Ju.D.Kuznetsovu, reg.№ 595**

(72) Inventor(s):

MUKERDZhi Kunal (US)

(73) Proprietor(s):

MAJKROSOFT KORPOREJShN (US)

(54) LOSSLESS PREDICTIVE ENCODING FOR IMAGES AND VIDEO

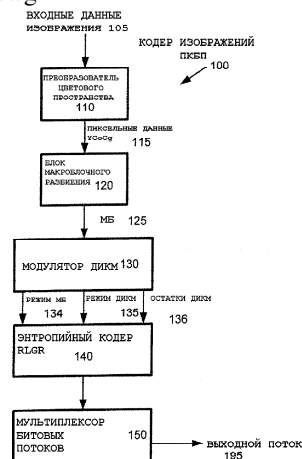
(57) Abstract:

FIELD: physics, image processing.

SUBSTANCE: invention concerns lossless encoding of images and video. Effect of invention is achieved by lossless predictive encoding based on macrounits for compatibility with existing image and video codecs, and provides for selection and application of on of available multiple modes of differential pulse-code modulation (DPCM) to separate macrounits to create DPCM residues with distribution close to optimum for entropy Golomb-Rice coding by series length.

EFFECT: efficient method of entropy lossless encoding despite different content properties of photos and graphic images.

19 cl, 7 dwg



ФИГ. 1

Область применения изобретения

Изобретение относится к кодированию без потерь для изображений и видео.

Предпосылки изобретения

Кодирование изображений без потерь имеет большое количество важных приложений, включая высококачественную цифровую фотографию, фильмографию, графику и т.д. Оно также применяется к профессиональному кодированию видео для кодирования кадров видео с максимально возможной настройкой качества, т.е. без потерь. Изображения в этих приложениях могут иметь различные характеристики, что представляет трудности при разработке кодека изображения, имеющего общее применение в этих приложениях. Например, изображения в графике обычно имеют резкие края или переходы цвета (например, между цветами текста и фона, и на границах соседних форм), тогда как фотографические изображения обычно имеют непрерывный тон (т.е. непрерывное изменение цвета (например, в виде градиента) по изображению).

В силу различий в характеристиках изображения, наиболее общие кодеки изображения предназначены для сжатия либо фотографических (например, JPEG), либо графических (GIF) изображений. При сжатии фотографических изображений обычно используется декоррелирующее преобразование наподобие ДКП (дискретное конусное преобразование) или «вейвлетов», тогда как для сжатия графических изображений обычно используются кодеки на основе строк, например, LZ77 или LZ78. В общем случае, фотографические кодеки плохо работают с графикой, поскольку основная гипотеза локальной гладкости или смещения постоянного тока, которая лежит в основе методов преобразования, обычно не действует в графике. Напротив, графические кодеки плохо работают на фотографических изображениях, поскольку алфавит слишком велик для построения хорошего словаря. В результате, существующие кодеки изображения для фотографических изображений не способны ни легко взаимодействовать с кодеками изображения и видео, ни эффективно оперировать графическим содержимым.

Например, CALIC [описанный в X. Wu, N. Memon и K. Sayood, "A Context-Based Adaptive Lossless/Nearly-Lossless Coding Scheme For Continuous-Tone Images," ISO, 1995], JPEG-LS [описанный в M.J. Weinberger и G. Seroussi, "The LOCO-I Lossless Image Compression Algorithm: Principles and Standardization into JPEG-LS," IEEE Trans. Image Processing, т. 9, стр. 1309-1324, август 2000] и SPIHT [описанный в Said и W.A. Pearlman, "A New Fast And Efficient Image Codec Based On Set Partitioning In Hierarchical Trees," IEEE Trans. On Circuits and Systems for Video Technology, т. 6, №. 6, стр. 243-250, июнь 1996] являются современными кодеками без потерь для фотографических изображений. Однако они не способны легко взаимодействовать с кодеками изображения и видео, и не работают эффективно с графическим содержимым. С другой стороны, GIF является современным кодеком графики без потерь. Однако, он тоже не работает с фотографическим содержимым, и его нелегко встроить в кодек видео. PTC [описанный в H.S. Malvar, "Fast Progressive Image Coding Without Wavelets," стр. 243-252, DCC 2000] является кодеком на основе макроблоков, который легко интегрировать в кодеки изображения и/или видео. Однако он плохо работает с графическим содержимым. Кроме того, ВТРС [описанный в J.A. Robinson, "Efficient General-Purpose Image Compression With Binary Tree Predictive Coding," IEEE Transactions on Image Processing, т. 6, №. 4, апрель 1997] способен обрабатывать фотографические и графические изображения единообразно и с оптимизацией по скорости, но его сжатие далеко от того, что требуется.

Раскрытие изобретения

Предсказательное кодирование без потерь (ПКБП) призвано решать рассмотренную выше проблему обеспечения сжатия изображения без потерь, которое, в общем случае, применимо к разнообразным изображениям и видео.

Проиллюстрированная здесь реализация ПКБП обеспечивает сжатие изображения без потерь для весьма разнообразного фотографического содержимого (изображений, видео, графики и пр.) с эффективностью сжатия, сравнимой с существующими кодеками изображения без потерь, и осуществляется сложностью во время среды выполнения, обеспечивающей большую скорость, чем большинство кодеков изображения без потерь.

Описанные здесь различные реализации предсказательного кодирования без потерь достигают этих результатов посредством комбинации, по меньшей мере, некоторых из следующих пунктов:

1. Действует в цветовом пространстве YCoCg. Это цветовое пространство повышает эффективность кодирования содержимого фотографического и графического изображения, и, кроме того, преобразование цветового пространства в/из распространенного представления цвета RGB (красный-зеленый-синий) может осуществляться с использованием быстрой, полностью целочисленной процедуры.

2. Действует на макроблоках. Это гарантирует, что предсказательное кодирование без потерь можно легко интегрировать в существующие кодеки изображения и видео, и что оно благоприятствует масштабируемым пространственно-временным реализациям в оборудовании и программном обеспечении, например, с использованием слоев.

3. Использует обширное множество предсказаний локальной, дифференциальной импульсно-кодовой модуляции (ДИКМ) на уровне макроблоков. Они предназначены для оптимальной декорреляции данных изображения из фотографических, а также графических источников, без помощи аномальных выборок или повторного сканирования изображения (см. J.A. Robinson, "Efficient General-Purpose Image Compression With Binary Tree Predictive Coding," IEEE Transactions on Image Processing, т. 6, №. 4, апрель 1997; и X. Wu, N. Memon и K. Sayood, "A Context-Based Adaptive Lossless/Nearly-Lossless Coding Scheme For Continuous-Tone Images," ISO, 1995). ДИКМ также способны порождать остатки, которые имеют двусторонние лапласовы распределения с нулевым смещением, поскольку они наилучшим образом кодируются методом энтропийного кодирования Голомба-Райса по длине серии (RLGR).

4. Использует метод Голомба-Райса длины серии (RLGR) для энтропийного кодирования различных распределений символов.

В одной реализации, где используется эта комбинация пунктов, предсказательное кодирование без потерь обеспечивает сжатие без потерь для всего фотографического содержимого (изображений, видео, а также графики), которое не только равно сжатию, обеспечиваемому CALIC, и больше, чем обеспечивают другие существующие форматы, включая JPEG-LS, PTC, ВTPC и т.д., но также приводит к удвоению или более сжатия для большинства графического содержимого по сравнению с этими кодеками фотографического типа и обеспечивает сложность во время выполнения, которая вдвое или более быстрее, чем обеспечивают самые передовые современные кодеки фотографического типа.

Дополнительные признаки и преимущества изобретения явствуют из нижеследующего подробного описания вариантов осуществления, приведенного совместно с прилагаемыми чертежами.

Краткое описание чертежей

Фиг.1 - блок-схема кодера изображений, использующего предсказательное кодирование без потерь.

Фиг.2 - схема, изображающая макроблок и слоистую структуру предсказательного кодирования без потерь, используемого в кодере, показанном на фиг.1.

Фиг.3 - схема, изображающая окружение пикселя в макроблоке, на котором базируется предсказание ДИКМ в кодере, показанном на фиг.1.

Фиг.4 - схема, изображающая режимы предсказания ДИКМ, применяемые в предсказательном кодировании без потерь, используемом в кодере, показанном на фиг.1.

Фиг.5 - текст псевдокода процедуры кодирования для предсказательного кодирования без потерь, используемого в кодере, показанном на фиг.1.

Фиг.6 - блок-схема декодера изображений, использующего предсказательное кодирование без потерь.

Фиг.7 - блок-схема вычислительной среды, пригодной для реализации кодера ПКБП, представленного на фиг.1 и 6.

Подробное описание

Нижеследующее описание посвящено реализациям предсказательного кодирования без потерь, которое объединяет в себе некоторые или все из энтропийного кодирования Голомба-Райса по длине серии (RLGR), множественных режимов ДИКМ и структуры кодирования цветового пространства YCoCg и макроблока (МБ) для обеспечения эффективного и быстрого кодера, применимого к разнообразному содержимому изображения, включая фотографию (непрерывный тон), графику и видео.

1. Кодер ПКБП

Согласно фиг.1, иллюстративный пример кодера 100 изображений, который базируется на предсказательном кодировании без потерь (ПКБП), осуществляет кодирование или сжатие данных 105 изображения. Данные изображения, вводимые в кодер ПКБП, могут представлять собой данные изображения в любом из разнообразных форматов без сжатия. Например, обычный формат, обрабатываемый иллюстрируемым кодером изображений, представляет собой формат «красный-зеленый-синий» (RGB) данных изображения, кадр видео и т.д. Эти данные изображения в формате RGB обычно структурированы в виде двумерной матрицы элементов изображения (пикселей), причем каждый пиксель представлен как цветовая выборка «красный-зеленый-синий» (RGB) изображения. Альтернативные реализации кодера изображений могут использовать другие входные форматы данных изображения. Следует также понимать, что этот кодер может входить в состав кодера видео для кодирования кадра в видеопоследовательности, с использованием предсказательного кодирования без потерь.

Кодер 100 изображений ПКБП обрабатывает эти данные изображения посредством набора процессов, который включает в себя преобразователь 110 цветового пространства, блок 120 макроблочного разбиения, модулятор 130 ДИКМ и энтропийный кодер 140 RLGR. Преобразователь 110 цветового пространства преобразует пиксели входных данных изображения из отображаемого представления цветового пространства в цветовое пространство YCoCg, что повышает эффективность кодирования. Блок 120 макроблочного разбиения разбивает изображение на макроблоки для совместимости с кодерами изображения и видео на макроблочной и слоистой основе. Кодер 130 ДИКМ выбирает и применяет один из

множества доступных режимов предсказания ДИКМ к каждому отдельному макроблоку, в результате чего возникают остатки предсказания, имеющие распределение, пригодное для энтропийного кодирования RGR. Затем энтропийный кодер 140 RLGR кодирует остатки предсказания макроблока. В результате возникает

1.1 Преобразователь цветового пространства YCoCg

В частности, преобразователь 110 цветового пространства преобразует цветовой формат входных данных изображения в цветовое пространство YCoCg. Пиксели входных данных изображения обычно представлены в легко отображаемом формате цветового пространства, например в цветовом пространстве «красный-зеленый-синий» (RGB). Цветовое пространство YCoCg лучше подходит для пространственно-эффективного кодирования изображения. Как выяснилось, формат YCoCg пригоден как для фотографических, так и для графических изображений, и превосходит другие цветовые преобразования в отношении эффективности кодирования. В частности, цветовое пространство YCoCg без потерь дает улучшение на ~15% по сравнению с цветовым пространством RGB в кодировании ПКБП.

Преобразование цветового пространства из RGB в YCoCg производится до любого кодирования входных данных изображения в формате RGB. В этой иллюстративной реализации кодера 100 ПКБП, преобразователь 110 цветового пространства использует процесс преобразования из RGB в YCoCg, более подробно описанный в H.S. Malvar и G.J. Sullivan, "YCoCg-R: A Color Space With RGB Reversibility and Low Dynamic Range," Joint Video Team of ISO/IEC MPEG & ITU-T VCEG Doc. JVT-1014, июль 2003, который обеспечивает возможность без потерь декоррелировать RGB в цветовое пространство YCoCg. Преобразование цветового пространства из RGB в YCoCg можно осуществить посредством прямого преобразования, заданного в следующем уравнении (1).

$$\begin{bmatrix} Y \\ C_0 \\ C_g \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

Декодер 600 (Фиг.6) может включать в себя блок преобразования в цветовое пространство RGB. Это преобразование использует обратную операцию, определенную в уравнении (2)

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} Y \\ C_0 \\ C_g \end{bmatrix} \quad (2)$$

Этапы подъема используются для получения прямого и обратного преобразования, что обеспечивает быстрые полностью целочисленные процедуры преобразования. Этапы подъема в прямом направлении заданы следующим образом:

$$C_0 = R - B;$$

$$x = B + C_0/2;$$

$$C_g = G - x;$$

$$Y = x + (C_g/2);$$

И в обратном направлении так:

$$x = Y - (C_g/2);$$

$$G=x+Cg;$$

$$B=x-(Co/2);$$

$$R=Co+B;$$

Можно видеть, что эти этапы подъема можно реализовать с полностью целочисленными сложениями, вычитаниями и сдвигами разряда, которые все выполняются чрезвычайно быстро.

1.2 Макроблочное кодирование

Блок 120 макроблочного разбиения разбивает изображение на макроблоки (МБ), как показано на фиг.2. В одной реализации кодера 100 изображений, МБ имеет размер 16x16 пикселей. В альтернативных реализациях можно использовать другие размеры макроблоков. Макроблочная структура облегчает практическое использование кодера изображений на основе ПКБП и облегчает его внедрение в популярные кодеки изображения, а также видео. Она также позволяет аппаратным и программным реализациям легко использовать слоистое кодирование, при котором закодированный битовый поток пакетизируется в слои. Слои обычно представляют собой некоторое целое число строк МБ. Это обеспечивает более гибкое пространство памяти и легкую приспособляемость к пространственно-временной масштабируемости (например, путем использования параллельных и многопоточных блоков выполнения).

Блок макроблочного разбиения разбивает каждую плоскость цветового пространства YCoCg на МБ и кодирует их по отдельности. Для каждого отдельного МБ кодер 100 изображений ПКБП кодирует следующие синтаксические элементы: режим ДИКМ, режим МБ и остатки ДИКМ.

Элемент «режим ДИКМ» указывает режим ДИКМ, выбранный модулятором 130 ДИКМ для декорреляции данных в этом МБ. В этой реализации модулятор ДИКМ выбирает из восьми возможных режимов ДИКМ, хотя в альтернативных реализациях кодера ПКБП можно использовать меньше или больше режимов ДИКМ. Блок макроблочного разбиения использует отдельный контекст RLGR для кодирования режима ДИКМ для МБ.

Для элемента «режим МБ» блок макроблочного разбиения использует два символьных алфавита с отдельным специализированным контекстом кодирования RLGR. Режим МБ сигнализирует одно из следующих двух событий: а) МБ кодирует значения ДИКМ; или б) МБ является «плоским» и потому пропускается. В последнем случае событие обрабатывается как ранний выход из кодирования/декодирования пикселей МБ как остатки ДИКМ, поскольку элемента «режим ДИКМ» достаточно для повторной генерации всех значений пикселя. Это кодирование плоского режима макроблока более подробно описано ниже.

Наконец, если МБ не пропущен, то остатки ДИКМ кодируются в сегменте МБ выходного потока, закодированного в режиме ПКБП.

1.3 Режимы ДИКМ

Модулятор 130 ДИКМ выбирает и применяет режим ДИКМ для текущего МБ, что позволяет более оптимально декоррелировать МБ для создания остатков ДИКМ, которые лучше сжимаются с помощью энтропийного кодирования RLGR. Энтропийное кодирование RLGR достигает лучшей эффективности кодирования, когда его входные значения имеют двустороннее лапласово распределение с нулевым смещением. Например, модулятор 130 ДИКМ в иллюстрируемом кодере на основе ПКБП может переключаться между восемью разными режимами ДИКМ для декорреляции каждого МБ. Альтернативные реализации могут включать в себя

больше или меньше режимов ДИКМ. Эти различные режимы ДИКМ способны создавать остатки, имеющие это оптимальное распределение для различных пиксельных шаблонов МБ. Это позволяет эффективно кодировать различные основные направления краев, которые могут попадать в МБ.

Согласно фиг.3 различные режимы ДИКМ указывают, какие из соседей 300 пикселя используются для предсказания значения пикселя. В частности, значение каждого пикселя 310 макроблока прогнозируется для некоторой комбинации из одного или нескольких соседних пикселей 320-323. Разность, полученная вычитанием фактического значения пикселя из его предсказанного значения, является остаточным значением ДИКМ для этого пикселя. Восемь режимов ДИКМ в иллюстрируемом модуляторе 130 ДИКМ используют прогнозы, основанные на комбинациях соседних левого пикселя 320, верхнего левого пикселя 321, верхнего пикселя 322 и верхнего правого пикселя 323. Это позволяет применять режим ДИКМ в единичном, одностороннем проходе или сканировании по МБ (т.е. сканировании каждой строки пикселей слева направо, с верхней строки к нижней в МБ).

На фиг.4 изображены восемь режимов предсказания ДИКМ, используемых в иллюстрируемом модуляторе 130 ДИКМ. Эти режимы способны декоррелировать различные обычные пиксельные шаблоны в набор остатков, распределение символов которых лучше подходит для энтропийного кодирования RLGR. В частности, эти режимы предсказания ДИКМ таковы:

Режим 0: Это «сырой» режим ДИКМ или отсутствия режима, когда отдельные пиксели кодируются напрямую без какого-либо вычитания. Этот режим полезен для МБ случайного или «пятнистого» типа без какого-либо однородно хорошего предсказателя по МБ.

Режим 1: В этом режиме ДИКМ, значение пикселя вычитается из его соседа непосредственно слева. Этот режим полезен, когда главные края лежат вдоль горизонтали.

Режим 2: В этом режиме ДИКМ, значение пикселя вычитается из его соседа непосредственно сверху. Этот режим полезен, когда главные края лежат вдоль вертикали.

Режим 3: В этом режиме ДИКМ, значение вычитается из минимального из его левого и верхнего соседей или, альтернативно, из максимального из его левого и верхнего соседей. Этот режим ДИКМ полезен для диагональных краев рампы, проходящих через позицию текущего пикселя.

Режим 4: В этом режиме ДИКМ, значение вычитается из среднего для его верхнего и верхнего правого соседей. Этот режим ДИКМ полезен для диагональных краев рампы с разной ориентацией.

Режим 5: В этом режиме ДИКМ, значение вычитается из своего верхнего левого соседа. Этот режим ДИКМ полезен для диагональных полос, например, в графическом содержимом.

Режим 6: В этом режиме ДИКМ, модулятор 130 вычитает одно и то же значение из его левого соседа как разность между его верхним и верхним левым соседями. Этот режим полезен для полосчатых горизонтальных рамп.

Режим 7: В этом режиме ДИКМ, модулятор 130 вычитает среднее для левого и верхнего соседей. Это также полезно, когда в МБ преобладают диагональные края.

Модулятор ДИКМ тестирует каждый из режимов предсказания ДИКМ 1-7 (т.е. отличных от отсутствия режима ДИКМ, которое является режимом 0), чтобы выбрать, какой режим ДИКМ создает более сжимаемые остатки ДИКМ. Модулятор

ДИКМ применяет соответствующие режимы ДИКМ и проверяет распределение символов для полученных остатков. Затем модулятор ДИКМ проверяет, какие остатки, порожденные режимом предсказания, имеют распределение, ближайшее к идеальному распределению для энтропийного кодирования RLGR. Модулятор ДИКМ также проверяет, достаточно ли близко ближайшее распределение к идеальному двустороннему лапласову распределению с нулевым смещением. Модулятор ДИКМ выбирает режим предсказания ДИКМ с ближайшим к идеалу распределением для макроблока, пока не будет достигнут порог достаточности. В противном случае, если режим предсказания ДИКМ с ближайшим к идеалу распределением не достигает порога достаточности, то модулятор ДИКМ выбирает отсутствие режима ДИКМ (режим 0) по умолчанию.

1.4 Кодирование плоского режима МБ

Согласно рассмотренному выше кодер 100 на основе ПКБП также может кодировать МБ в плоский режим МБ. Плоский режим МБ используется, когда все остатки ДИКМ, полученные в результате применения режима ДИКМ к МБ, равны нулю. При тестировании режимов ДИКМ для выбора режима ДИКМ для использования на МБ, модулятор ДИКМ также проверяет, создают ли тестируемый в данный момент режим ДИКМ только нулевые остатки ДИКМ для МБ. Определив, что режим ДИКМ порождает только нулевые остатки ДИКМ, кодер 100 на основе ПКБП кодирует МБ в плоский режим МБ - без необходимости тестировать другие режимы ДИКМ. При кодировании в плоский режим МБ кодер 100 на основе ПКБП может кодировать МБ в выходной битовый поток, как только режим МБ и режим ДИКМ (т.е. пропуская кодирование остатков ДИКМ). Кодирования плоского режима МБ и режима ДИКМ для МБ в выходной битовый поток достаточно для декодирования значений пикселей МБ. Поскольку остатки ДИКМ не кодируются, плоский режим МБ дает более высокую эффективность сжатия для кодирования МБ.

1.5 Множественные контексты кодирования Голомба-Райса по длине серии (RLGR)

Согласно фиг.1, режим МБ, режим ДИКМ и остатки ДИКМ, созданные модулятором 130 ДИКМ, подвергаются энтропийному кодированию с использованием кодирования RLGR в энтропийном кодере RLGR 140. Энтропийный кодер 140 RLGR в иллюстрируемом кодере 100 на основе ПКБП использует процесс кодирования Голомба-Райса по длине серии, описанный в Н. Malvar, "Fast Progressive wavelet coding," Proc. IEEE Data Compression Conference, Snowbird, UT, стр. 336-343, март-апрель 1999. Этот процесс кодирования RLGR не является истинно общим энтропийным кодером, например адаптивным арифметическим кодером. Кодирование RLGR делает предположение, что наиболее вероятный символ равен нулю. Поэтому если строка чисел с наиболее вероятным символом, не равным нулю, поступает в RLGR, то кодирование RLGR будет иметь низкую эффективность кодирования. Если его вход поступает из источника с распределением символов, близким к лапласову, то процесс кодирования RLGR будет кодировать эти данные очень хорошо, весьма близко к энтропии, и во многих случаях это будет лучше, чем адаптивное арифметическое кодирование. В кодере 100 на основе ПКБП режимы предсказания ДИКМ способны порождать двусторонние лапласовы распределения с нулевым смещением для целых чисел со знаком (для фотографических и графических изображений) обычного содержимого фотографических и графических изображений, на котором RLGR работает наилучшим образом.

Энтропийный кодер 140 RLGR в иллюстративном кодере на основе ПКБП использует отдельный контекст RLGR для каждого из: а) режима МБ (плоского или

нет); b) режима ДИКМ; c) остаточных значений ДИКМ (двусторонние лапласовы распределения с нулевым смещением для целых чисел). В каждом из этих контекстов, энтропийный кодер RLGR осуществляет адаптивное двоичное кодирование Голомба-Райса по длине серии двоичной строки, сформированной значащими битами, происходящими из символов, кодируемых с помощью отдельного контекста, например остатков ДИКМ для контекста остаточных значений ДИКМ.

Использование множественных контекстов RLGR для кодирования разных распределений символов повышает эффективность энтропийного кодирования. Это потому, что очень важно адаптироваться к каждому отдельному распределению и его уникальной асимметрии. Например, режимы МБ, скорее всего, не все являются плоскими. Соответственно, их распределение, скорее всего, отклоняется от плоского случая. Но было бы лучше, если бы содержимое было графическим, а не фотографическим. Выделение конкретного контекста RLGR позволяет адаптировать энтропийный кодер RLGR в кодере 100 на основе ПКБП к таким унимодальным распределениям с большей эффективностью. В альтернативных реализациях кодера на основе ПКБП можно использовать больше или меньше контекстов RLGR.

Использование дополнительных отдельных контекстов RLGR для кодирования остатков ДИКМ в таких реализациях может обеспечивать повышенную эффективность энтропийного кодирования и препятствовать размыванию контекста, но, на практике, в иллюстрируемой реализации используются три отдельных контекста.

После энтропийного кодирования энтропийного кодера RLGR мультиплексор 150 битовых потоков компонуют данные для МБ, закодированные в режиме RLGR, в выходной битовый поток 195. В реализациях, использующих слоистое кодирование, мультиплексор битовых потоков компонует или пакетизирует закодированные МБ в слои.

1.6 Текст псевдокода

Кодирование ПКБП, осуществляемое в вышеописанном кодере 100 на основе ПКБП, приведено в тексте 500 псевдокода, показанном на фиг.5. В этом тексте псевдокода, входной параметр “ImageBand” представляет данные изображения из одной их координат цветового пространства, т.е. Y, Co или Cg. Этот процесс кодирования ПКБП применяется после преобразования цветового пространства изображения в цветовое пространство YCoCg.

2. Декодер ПКБП

Согласно фиг.6, декодер 600 изображений, основанный на предсказательном кодировании без потерь (ПКБП), осуществляет декодирование выходного битового потока 195, созданного из кодирования ПКБП кодером 100 изображений на основе ПКБП. В этом декодере 600 изображений на основе ПКБП демультимплексор 610 битового потока сначала разделяет отдельные закодированные МБ в битовом потоке и закодированный режим МБ, режим ДИКМ и остатки ДИКМ для этого МБ. Демультимплексор битового потока обеспечивает отдельные данные для декодера 620 RLGR.

Декодер 620 RLGR декодирует закодированные по RLGR режим МБ, режим ДИКМ и остатки ДИКМ для каждого МБ. Декодер 620 RLGR использует процесс декодирования RLGR, описанный в Н. Malvar, “Fast Progressive wavelet coding,” Proc. IEEE Data Compression Conference, Snowbird, UT, стр. 336-343, март-апрель 1999. Затем декодер 620 RLGR выдает декодированные данные на демодулятор 630 ДИКМ.

Демодулятор 630 ДИКМ осуществляет обратный процесс над остатками ДИКМ

для режима предсказания ДИКМ, который использовался для МБ и, таким образом, восстанавливает данные МБ. Для МБ, закодированного в плоский режим МБ, демодулятор 630 ДИКМ осуществляет обратный процесс для всех нулевых остатков для декодированного режима предсказания ДИКМ.

После применения обратного предсказания ДИКМ восстановитель 640 изображений повторно компонует МБ для восстановления изображения. Затем преобразователь 650 цветового пространства осуществляет обратное преобразование цветового пространства YCoCg для преобразования этих данных изображения обратно в изображение RGB. В некоторых реализациях это преобразование можно опустить, и изображение остается в формате цветового пространства YCoCg.

3. Вычислительная среда

Вышеописанный кодер 100 на основе ПКБП и/или декодер 600 (кодек ПКБП) можно реализовать на любом из различных устройств обработки изображения и видео и вычислительных устройств, включая компьютеры той или иной мощности (персональные, рабочие станции, серверы, карманные, портативные, планшетные или другие мобильные устройства), в распределенных вычислительных сетях и других веб-службах, и в устройствах записи/воспроизведения/приема/просмотра изображений и видео, в качестве нескольких общих примеров. Кодек на основе ПКБП может быть реализован в виде аппаратной схемы, а также в виде программного обеспечения 780 кодера, выполняющегося в компьютере или в другой вычислительной среде, например, показанной на фиг.7.

На фиг.7 проиллюстрирован обобщенный пример подходящего устройства обработки изображения/видео в вычислительной среде 700 (например, компьютера), в которой можно реализовать описанные способы. Эта среда 700 не призвана накладывать какие-либо ограничения на объем использования функциональных возможностей изобретения, поскольку настоящее изобретение можно реализовать в различных средах обработки изображения/видео общего назначения или специального назначения.

Согласно фиг.7 вычислительная среда 700 включает в себя, по меньшей мере, один блок 710 обработки и память 720. На фиг.7 эта наиболее базовая конфигурация 730 обведена пунктирной линией. Блок обработки 710 выполняет компьютерно-выполняемые команды и может быть реальным или виртуальным процессором. В многопроцессорной системе, множественные блоки обработки выполняют компьютерно-выполняемые команды для повышения мощности обработки. Память 720 может представлять собой энергозависимую память (например, регистры, кэш, ОЗУ), энергонезависимую память (например, ПЗУ, ЭППЗУ, флэш-память и т.д.) или некоторую их комбинацию. В памяти 720 хранится программное обеспечение 780, реализующее кодек на основе ПКБП.

Вычислительная среда может иметь дополнительные особенности. Например, вычислительная среда 700 включает в себя запоминающее устройство 740, одно или несколько устройств 750 ввода, одно или несколько устройств 760 вывода и одно или несколько средств 770 связи. Механизм взаимного соединения (не показан), например, шина, контроллер или сеть, соединяет между собой компоненты вычислительной среды 700. Обычно, программное обеспечение операционной системы (не показано) обеспечивает операционную среду для другого программного обеспечения, выполняющегося в вычислительной среде 700, и координирует действия компонентов вычислительной среды 700.

Запоминающее устройство 740 может быть сменным и стационарным и включает в

себя магнитные диски, магнитные ленты или кассеты, CD-ROM, CD-RW, DVD или любой другой носитель, который можно использовать для хранения информации и к которым можно осуществлять доступ в вычислительной среде 700. В запоминающем устройстве 740 хранятся команды для программного обеспечения 780 кодека на основе ПКБП.

Устройство(а) 750 ввода (например, устройства, действующие как контрольная точка в архитектуре 100 взаимосвязи устройств) может представлять собой тактильное устройство ввода, например клавиатуру, мышь, перо или шаровой манипулятор, голосовое устройство ввода, сканирующее устройство или другое устройство, которое обеспечивает ввод в вычислительную среду 700. Для аудио устройством(ами) 750 ввода может быть звуковая карта или аналогичное устройство, которое принимает входной аудиосигнал в аналоговой или цифровой форме, или устройство чтения CD-ROM, которые подает выборки аудиосигнала в вычислительную среду. Устройство(а) 760 вывода может представлять собой дисплей, принтер, громкоговоритель, устройство записи CD или другое устройство, которое обеспечивает выходной сигнал вычислительной среды 700.

Средство(а) связи 770 обеспечивают связь по среде связи с другой вычислительной сущностью. Среда связи переносит информацию, например компьютерно-выполняемые команды, аудио/видеоинформацию или информацию другого типа или другие данные в виде сигнала, модулированного данными. Сигнал, модулированный данными, это сигнал, одна или несколько характеристик которого задается или изменяется таким образом, чтобы кодировать информацию в сигнале. В порядке примера, но не ограничения, среды связи включают в себя проводные или беспроводные технологии, реализованные посредством электрического, оптического, РЧ, инфракрасного, акустического или иного носителя.

Обработка макрорасширения и методы отображения могут быть описаны здесь в общем контексте компьютерно-считываемых сред. Компьютерно-считываемые среды могут представлять собой любые доступные среды, к которым можно осуществлять доступ в вычислительной среде. В порядке примера, но не ограничения, применительно к вычислительной среде 700, компьютерно-считываемые среды включают в себя память 720, запоминающее устройство 740, среды связи и комбинации любых вышеперечисленных элементов.

Способы могут быть описаны здесь в общем контексте компьютерно-выполняемых команд, например входящих в состав программных модулей, выполняющихся в вычислительной среде на целевом реальном или виртуальном процессоре. В общем случае, программные модули включают в себя процедуры, программы, библиотеки, объекты, классы, компоненты, структуры данных и т.д., которые выполняют определенные задачи или реализуют определенные абстрактные типы данных. Функциональные возможности программных модулей могут быть объединены или распределены между программными модулями в соответствии с нуждами различных вариантов осуществления. Компьютерно-выполняемые команды для программных модулей могут выполняться в локальной или распределенной вычислительной среде.

В целях представления, такие термины, как «определять», «генерировать», «регулировать» и «применять» используются в подробном описании для описания компьютерных операций в вычислительной среде. Эти термины являются абстракциями высокого уровня для операций, осуществляемых компьютером, и их не следует путать с действиями, осуществляемыми человеком. Фактические компьютерные операции, соответствующие этим терминам, варьируются в

зависимости от реализации.

Ввиду большого количества возможных вариантов осуществления, к которым можно применять принципы нашего изобретения, мы заявляем в качестве нашего изобретения все такие варианты осуществления, которые отвечают объему и сущности
5
нижеследующей формулы изобретения и ее эквивалентам.

Формула изобретения

1. Способ кодирования без потерь информационных сред изображения и видео,
10
содержащий этапы, на которых

разбивают входные данные изображения на блочные части,
для отдельной блочной части выбирают один из нескольких доступных режимов
предсказания дифференциальной импульсно-кодовой модуляции (ДИКМ) для
применения к блочной части, который, из доступных режимов предсказания ДИКМ,
15
дает распределение символов, являющееся наиболее близким к оптимальному
двустороннему несмещенному распределению символов энтропийного кодера
Голомба-Райса по длине серии,

применяют выбранный режим предсказания ДИКМ к блочной части, и
20
подвергают энтропийному кодированию остатки ДИКМ для блочной части
используя кодирование Голомба-Райса по длине серии; и
выводят закодированные остатки ДИКМ для блочной части в битовый поток.

2. Способ по п.1, дополнительно содержащий этап, на котором преобразуют
входные данные изображения в формат цветового пространства YCoCg.

3. Способ по п.1, дополнительно содержащий этап, на котором кодируют режим
предсказания ДИКМ и остатки ДИКМ с помощью отдельных контекстов
кодирования Голомба-Райса по длине серии.

4. Способ по п.1, дополнительно содержащий этапы, на которых определяют,
30
создает ли применение выбранного режима предсказания ДИКМ к блочной части
только остатки ДИКМ с нулевым значением, и кодируют указание того, что блочная
часть является плоской, вместо того, чтобы подвергать энтропийному кодированию
остатки ДИКМ блочной части.

5. Способ по п.1, в котором на этапе выбора режима предсказания ДИКМ
35
определяют, обеспечивает ли режим предсказания ДИКМ, наиболее близкое к
оптимальному распределение символов для энтропийного кодирования, к
оптимальному распределению символов для энтропийного кодирования, и если не
достаточно близок, то не применяют никакого ДИКМ к макроблоку перед
40
энтропийным кодированием.

6. Способ по п.1, в котором режимы предсказания ДИКМ содержат режимы,
способные создавать оптимальное распределение символов для энтропийного
кодирования для блочных частей, содержащее изображения которых, в основном,
является горизонтальным главным краем, вертикальным главным краем,
45
диагональными краями рампы, полосами и полосчатыми горизонтальными рампами.

7. Способ по п.1, в котором режимы предсказания ДИКМ содержат: первый режим,
в котором значение пикселя вычитают из его левого соседнего пикселя, второй режим,
в котором значение пикселя вычитают из его верхнего соседнего пикселя, третий
50
режим, в котором значение пикселя вычитают из минимума или максимума его левого
и верхнего соседних пикселей, четвертый режим, в котором значение пикселя
вычитают из среднего для его верхнего и верхнего правого соседних пикселей, пятый
режим, в котором значение пикселя вычитают из его верхнего левого соседнего

пикселя, шестой режим, в котором разность между верхним и верхним левым соседними пикселями пикселя вычитают из его левого соседнего пикселя, и седьмой режим, в котором значение пикселя вычитают из среднего для левого и верхнего соседних пикселей пикселя.

8. Медиасистема, обеспечивающая предсказательное кодирование без потерь медиакодекта изображения или видео, система содержит процесс макроблочного разбиения для разбиения входных данных изображения на макроблоки,

многорежимный процесс дифференциальной импульсно-кодовой модуляции (ДИКМ), действующий на отдельном макроблоке входных данных изображения для выбора одного из множественных режимов предсказания ДИКМ, который создает остаточное распределение для макроблока, которое более точно совпадает с оптимальным двусторонним несмещенным распределением энтропийного кодирования Голомба-Райса по длине серии (RLGR), и применяет выбранный режим предсказания ДИКМ к макроблоку, и

процесс энтропийного кодирования для осуществления кодирования Голомба-Райса по длине серии для остатков ДИКМ макроблока.

9. Медиасистема по п.8, дополнительно содержащая процесс преобразования цветового пространства для преобразования входных данных изображения в формат цветового пространства YCoCg до кодирования.

10. Медиасистема по п.8, в которой режимы предсказания ДИКМ содержат режимы, способные создавать распределения, близкие к оптимальному распределению энтропийного кодирования RLGR для макроблоков, содержащее изображения которых, в основном, является горизонтальным главным краем, вертикальным главным краем, диагональными краями рампы, полосами и полосчатыми горизонтальными рампами.

11. Медиасистема по п.8, в которой режимы предсказания ДИКМ содержат: первый режим, в котором значение пикселя вычитают из его левого соседнего пикселя, второй режим, в котором значение пикселя вычитают из его верхнего соседнего пикселя, третий режим, в котором значение пикселя вычитают из минимума или максимума его левого и верхнего соседних пикселей, четвертый режим, в котором значение пикселя вычитают из среднего для его верхнего и верхнего правого соседних пикселей, пятый режим, в котором значение пикселя вычитают из его верхнего левого соседнего пикселя, шестой режим, в котором разность между верхним и верхним левым соседними пикселями пикселя вычитают из его левого соседнего пикселя, и седьмой режим, в котором значение пикселя вычитают из среднего для левого и верхнего соседних пикселей пикселя.

12. Компьютерно-считываемый носитель данных, на котором хранятся компьютерно-выполняемые программные команды, предназначенные, при выполнении в компьютерной медиасистеме обработки для осуществления способа кодирования данных изображения или видео, способ содержит этапы, на которых преобразуют данные изображения в формат цветового пространства YCoCg, разбивают данные изображения на макроблоки, для макроблока данных изображения определяют, какой из группы доступных режимов предсказания ДИКМ создает остатки, ближайшие к оптимальному двустороннему несмещенному распределению для кодирования RLGR, если этот определенный режим предсказания ДИКМ создает остатки, распределение которых достаточно близко к двустороннему несмещенному

оптимальному распределению, применяют режим предсказания ДИКМ к макроблоку, и

подвергают энтропийному кодированию RLGR остатки макроблока.

13. Компьютерно-считываемый носитель данных по п.12, в котором способ дополнительно содержит этапы, на которых определяют, создает ли применение определенного режима предсказания ДИКМ к макроблоку плоские остатки, и если да, то кодируют макроблок как указание плоского режима макроблока без энтропийного кодирования RLGR остатков такого плоского макроблока.

14. Компьютерно-считываемый носитель данных по п.13, в котором способ дополнительно содержит этап, на котором подвергают энтропийному кодированию RLGR указание режима макроблока с использованием отдельного контекста кодирования RLGR вместо того, чтобы подвергать остатки энтропийному кодированию RLGR.

15. Компьютерно-считываемый носитель данных по п.13, в котором способ дополнительно содержит этапы, на которых определяют, создает ли режим предсказания ДИКМ, создающий остаточное распределение, ближайшее к оптимальному двустороннему несмещенному распределению, остаточное распределение, достаточно близкое к оптимальному двустороннему несмещенному распределению, и если недостаточно близкое, подвергают энтропийному кодированию RLGR макроблок без применения режима предсказания ДИКМ к макроблоку.

16. Способ декодирования данных изображения или видео, подвергнутых предсказательному кодированию без потерь, содержащий этапы, на которых подвергают энтропийному декодированию RLGR режим макроблока, режим предсказания ДИКМ и закодированные двусторонним несмещенным энтропийным кодированием остатки ДИКМ для каждого из совокупности макроблоков с использованием отдельных контекстов кодирования RLGR,

если режим макроблока для макроблока является плоским режимом макроблока, декодируют пиксели макроблока с использованием демодуляции ДИКМ, которая является обратной к режиму предсказания ДИКМ для всех нулевых остатков, декодированному по RLGR,

в противном случае, если режим предсказания ДИКМ для макроблока является отсутствием режима предсказания ДИКМ, декодируют пиксели макроблока без демодуляции ДИКМ,

в противном случае демодулируют декодированные по двустороннему несмещенному RLGR остатки ДИКМ с использованием демодуляции ДИКМ, которая является обратной к режиму предсказания ДИКМ, декодированному по RLGR, и компонуют макроблоки для формирования декодированных данных изображения.

17. Способ по п.16, содержащий этап, на котором преобразуют декодированные данные изображения из формата цветового пространства YCoCg в отображаемый формат цветового пространства.

18. Декодер изображения или видео, подвергнутого предсказательному кодированию без потерь, содержащий энтропийный декодер по длине серии Голомба-Райса (RLGR), предназначенный для декодирования остатков ДИКМ и режима предсказания ДИКМ макроблока, закодированных по двустороннему несмещенному RLGR,

демодулятор ДИКМ для применения обратного режима предсказания ДИКМ к остаткам ДИКМ, закодированным по двустороннему несмещенному RLGR, и

перекомпоновщик макроблоков для компоновки макроблока с другими декодированными макроблоками для формирования данных восстановленного изображения.

5 19. Декодер изображения или видео, подвергнутого предсказательному кодированию без потерь, по п.18, содержащий обратный преобразователь YCoCg для преобразования восстановленного изображения из цветового пространства YCoCg в цветное пространство, пригодное для отображения изображения.

10

15

20

25

30

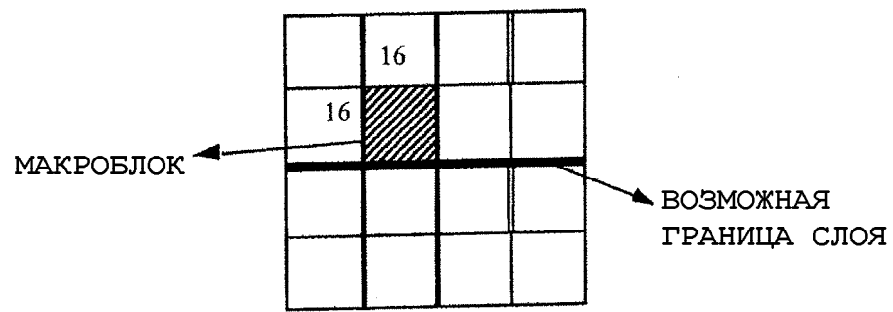
35

40

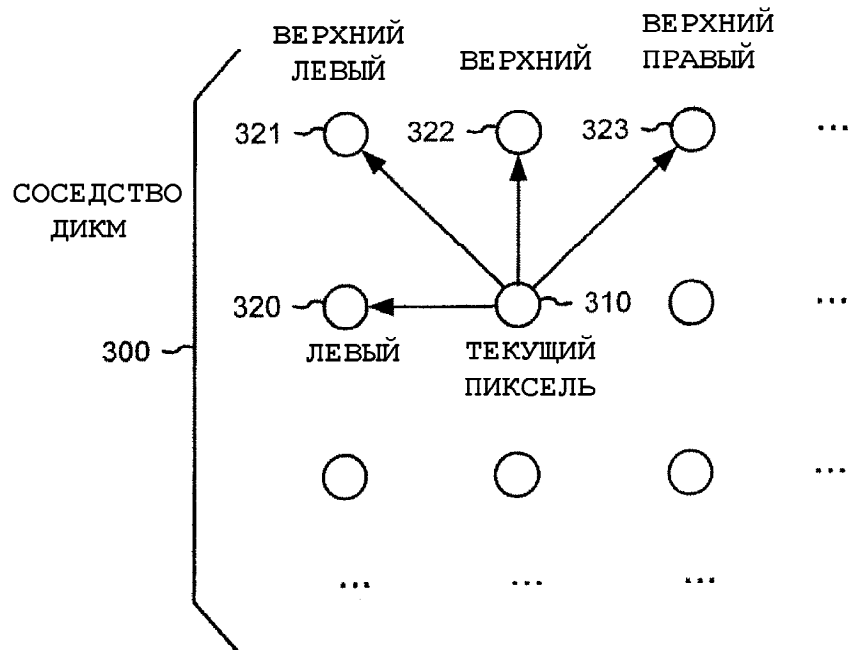
45

50

КАДР ИЗОБРАЖЕНИЯ ИЛИ ВИДЕО

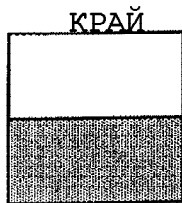
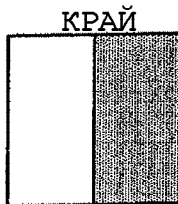
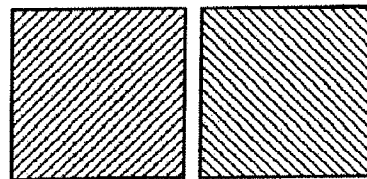
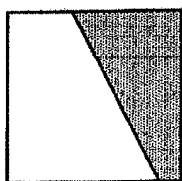
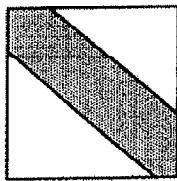
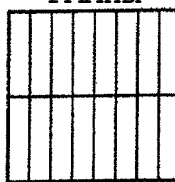
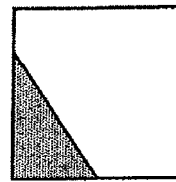


ФИГ. 2



ФИГ. 3

РЕЖИМЫ ДИКМ

РЕЖИМ 1: ГОРИЗОН-
ТАЛЬНЫЙ ГЛАВНЫЙРЕЖИМ 2: ВЕРТИ-
КАЛЬНЫЙ ГЛАВНЫЙРЕЖИМ 3: ДИАГОНАЛЬНЫЕ
КРАЯ РАМПЫРЕЖИМ 4:
ДИАГОНАЛЬНЫЙ
КРАЙ РАМПЫРЕЖИМ 5:
ДИАГОНАЛЬНАЯ
ПОЛОСАРЕЖИМ 6:
ПОЛОСЧАТЫЕ
ГОРИЗОНТАЛЬНЫЕ
РАМПЫРЕЖИМ 7:
КРАЙ

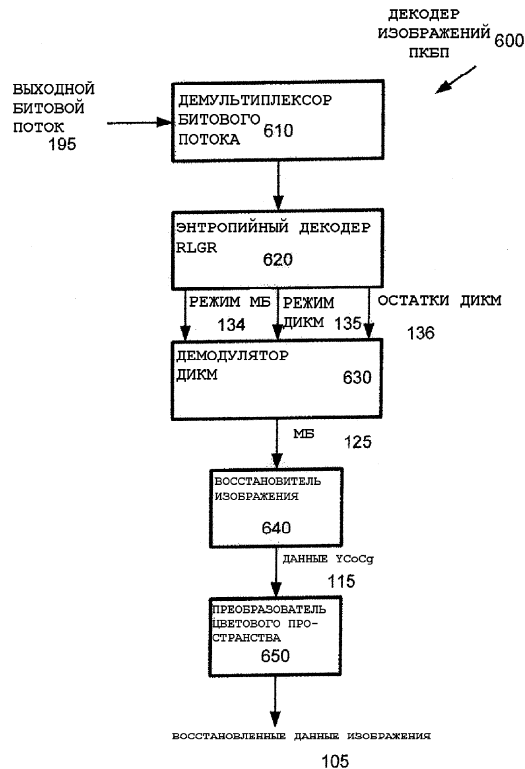
ФИГ. 4

500

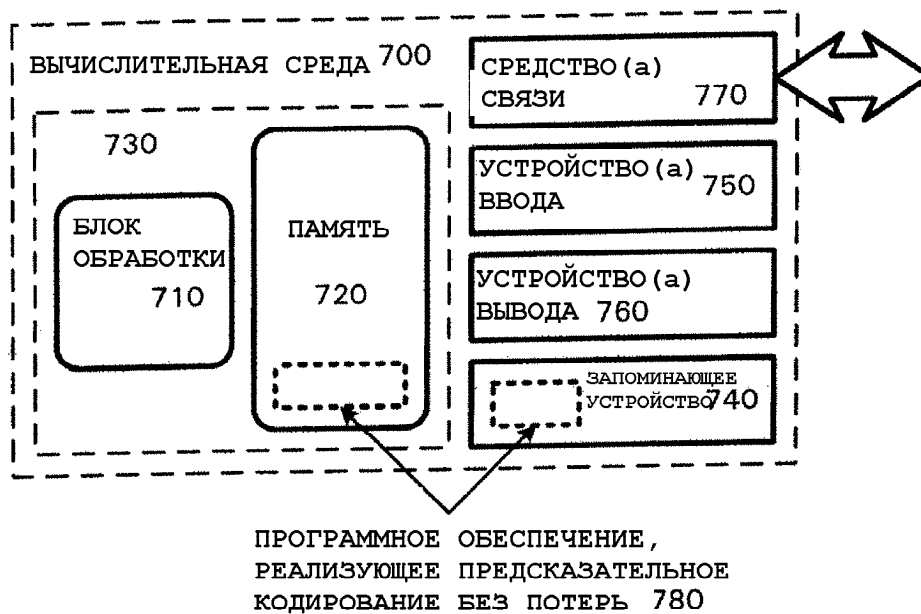
```

PLC(IN ImageDand i, OUT BitStream b)
для всех макроблоков делать
    найти и применить лучший режим ДИКМ
    если МБ плоский, то продолжать //пропустить остальное
//для блока
    кодировать по RLGR остатки ДИКМ
конец цикла
конец PLC
  
```

ФИГ. 5



ФИГ. 6



ФИГ. 7