

УДК 621.327

С.В. Карпенко

Національний авіаційний університет, Київ

СОЗДАНИЕ ПОДХОДА ДЛЯ ИСКЛЮЧЕНИЯ ТРЕХМЕРНОЙ ИЗБЫТОЧНОСТИ ИЗОБРАЖЕНИЙ

Разрабатывается подход к сжатию данных, основанный на исключении трехмерной структурной избыточности. Обосновывается, что трехмерная избыточность устраняется на основе учета закономерностей в динамических диапазонах одновременно по трем координатам. Проводится оценка степени сжатия видеоданных на основе сокращения трехмерной избыточности.

избыточность видеоданных, динамический диапазон

Введение

Развитие распределенных информационных систем (ИС) неразрывно связано с ростом цифровых объемов обрабатываемых данных. За последнее десятилетие наибольший удельный вес занимают видеоданные [1]. Такая особенность сопровождается противоречием между требованиями с одной стороны относительно достоверности информации, а с другой стороны относительно своевременности ее получения.

Препятствием для решения данного противоречия являются ограниченные возможности информационных систем по скорости обработки и передачи данных в телекоммуникационных системах. В связи с этим создаются подсистемы сжатия данных [1 – 6].

Сжатие данных достигается в результате сокращения избыточности. Наибольшее сокращение объемов видеоданных происходит на базе устранения психовизуальной избыточности. Но такое направление связано с потерями части информации. Значит, для решения противоречия необходимо основываться на снижении таких видов избыточности, которые не связаны с внесением погрешности. Недостаток такого подхода проявляется в недостаточной степени компрессии [1, 2].

Отсюда **цель статьи** состоит в сокращении времени на обработку и передачу данных в информационных системах при сохранении заданной степени достоверности.

Проведенный анализ различных подходов относительно процессов устранения избыточности выявил, что дополнительное увеличение степени сжатия возможно за счет учета структурных закономерностей одновременно по трем координатам [4, 5].

Таким образом, **научная задача** заключается в создании подхода для сжатия данных без внесения погрешности на основе устранения структурной избыточности в трехмерном пространстве.

Обоснование подхода для исключения трехмерной избыточности изображений

Для организации трехмерной обработки необходимо сформировать трехмерные структуры данных A (ТСД), имеющие в общем случае параллелепипедную топологию. При обработке изображений трехмерные структуры могут образовываться на следующих уровнях укрупнения:

1) последовательность кадров изображений $U = \{U_1, \dots, U_\theta, \dots, U_\Theta\}$, где U_θ – θ -й кадр; Θ – количество кадров в видеопоследовательности. В этом случае $(i; j)$ -й элемент u_{ij} принадлежащий θ -у кадру будет $(\theta; i; j)$ -м элементом $u_{\theta i j}$ трехмерной структуры данных A . Размер вертикальных плоскостей равен размеру кадра изображения $Z_{стр} \times Z_{стб}$, а длина ТСД будет равна количеству вертикальных плоскостей Θ :

$$A = \{u_{\theta i j}\}, i = \overline{1, Z_{стр}}, j = \overline{1, Z_{стб}}, \theta = \overline{1, \Theta};$$

2) плоскости в трехцветной цветовой модели оцифрованного полноцветного изображения U_θ . В частном случае для модели RGB-представления $U_\theta = \{U_R^{(\theta)}, U_G^{(\theta)}, U_B^{(\theta)}\}$, где $U_R^{(\theta)}, U_G^{(\theta)}, U_B^{(\theta)}$ – плоскости цветовых R, G и B составляющих. Для такой трехмерной структуры размер горизонтальных плоскостей будет равен размеру кадра изображения, а высота ТСД будет равна количеству цветковых плоскостей. Компоненты позиционируются в ТСД как $u_{\vartheta i j}$, где ϑ – номер, определяющий цветовую плоскость: 1 – R-плоскость, 2 – G-плоскость, 3 – B-плоскость:

$$A = \{u_{\vartheta i j}\}, i = \overline{1, Z_{стр}}, j = \overline{1, Z_{стб}}, \vartheta = \overline{1, 3};$$

3) отдельные фрагменты $U_{\theta, f}$ изображения, где f – индекс фрагмента изображения. Для этого варианта формирования ТСД размеры ее будут равны $Z_{стр} \times Z_{стб} \times Z_B$, где Z_B – количество гори-

зонтальных плоскостей. Тогда трехмерная структура задается как $A = \{u_{fij}\}$, $i = \overline{1, Z_{стр}}$, $j = \overline{1, Z_{стб}}$, $f = \overline{1, Z_B}$, где u_{fij} – $(f; i; j)$ -й элемент, расположенный в f -м фрагменте на позиции $(i; j)$.

Избыточность на первом уровне обусловлена в основном наличием одинаковых по содержанию областей в соседних кадрах. Данный вид избыточности сокращается в основном за счет кодирования с пополнением кадров [4]. Области, меняющие свой масштаб и положение в кадровой последовательности, обрабатываются на основе методов компенсации движения [1, 4]. На втором уровне представления изображений в виде ТСД избыточность вызвана главным образом структурным подобием между цветовыми плоскостями. Полноцветное изображение в цифровом виде рассматривается в виде цветоразностных моделей. Это обусловлено психовизуальными особенностями зрительной системы человека, а именно яркостная составляющая имеет большее значение для восприятия изображений, чем цветовые составляющие [1, 4]. Поэтому цветовые плоскости обрабатываются отдельно. Для цветовых плоскостей допускаются потери качества при восстановлении, что создает возможность для увеличения степени сжатия. Сокращение избыточности на нижнем уровне – внутрикадровой обработки – достигается в результате учета закономерностей только по одному или по двум направлениям.

Общие характеристики процессов исключения избыточности на разных уровнях обработки видеоданных состоят в том, что: наибольшие степени сжатия достигаются за счет учета психовизуальных характеристик зрительной системы; учитываются в основном две закономерности; сжатие видеоданных на разных уровнях содержит в себе внутрикадровую обработку.

Ограниченность внутрикадровой обработки двумя координатами вызвано двумерной природой изображений. При этом существуют следующие особенности:

1. Отсутствует методика оптимального разбиения изображений на отдельные фрагменты правильных размеров. Это приводит к таким недостаткам: происходит дробление когерентных областей и областей, окрашенных одним цветом на более мелкие части; появление эффекта блочности при восстановлении изображений с заданной степенью сжатия; не учитывается наличие закономерностей между смежными фрагментами изображения.

2. Обработка всего кадра изображения как одного блока не позволяет выявить важных особенностей, свойственных локальным фрагментам, т.е. снижается адаптация процесса компрессии к нестационарности отдельных частей кадра.

Поэтому для повышения степени сжатия *предлагается* из отдельных фрагментов изображения формировать трехмерные структуры данных. Трехмерная внутрикадровая обработка объясняется: возможностью для организации обработки одновременно по трем координатам. Это позволяет устранить большее количество избыточности по сравнению с количеством, исключаемой избыточности в двумерных областях; возможностями по сравнению с двумерной обработкой для сокращения количества операций. Такое преимущество обусловлено дополнительным сокращением количества операций при распараллеливании процесса обработки; снижением количества разрядов, отводимых на представление служебной информации, приходящейся на один элемент видеоданных; возможностью выявить закономерности в динамических диапазонах отдельных частей кадра, имеющих нестационарные свойства; упрощением процедуры формирования одного кода-номера для различных фрагментов изображения; дополнительным увеличением коэффициента сжатия для случаев, когда количество разрядов, отводимое на представление кода-номера значительно меньше, чем длина машинного слова. Тогда за счет дополнительного присоединения элементов из других уровней трехмерной структуры достигается повышение степени заполнения машинного слова.

Для осуществления трехмерной обработки видеоданных метод кодирования должен обладать следующими свойствами:

- 1) учитывать характеристики изображений одновременно по трем координатам;
- 2) адаптироваться к содержанию ТСД;
- 3) обрабатывать видеоданные в реальном времени без внесения погрешности;
- 4) обеспечивать дополнительное сокращение избыточности за счет перехода от двумерной к трехмерной обработке, так как не все коды позволяют учесть третью дополнительную координату.

В соответствии с требованиями к методу кодирования необходимо в первую очередь определить тип закономерностей, существующий в ТСД, построенных на основе локальных частей изображения. Наиболее универсальными структурными характеристиками, отличающимися незначительной зависимостью от нестационарности вероятностно-статистических свойств и свойств когерентности изображений, являются ограничения на динамический диапазон видеоданных.

Если рассматривать организацию видеоданных в виде плоскости $A_v = \{a_{ji}\}$; $1 \leq j \leq n_{стб}$, $1 \leq i \leq n_{стр}$, то под динамическим диапазоном элемента понимается минимальное значение ψ_{ij} из двух максимумов λ_i – по строкам ($i = \overline{1, n_{стр}}$) и χ_j –

по столбцам ($j=1, \overline{n_{\text{стб}}}$) [5]:

$$\psi_{ij} = \min(\lambda_i, \chi_j), \quad (1)$$

где ψ_{ij} – минимальное значение ($i; j$)-го элемента; λ_i, χ_j – максимальные значения соответственно в i -й строке и j -м столбце $\max_{1 \leq i \leq n_{\text{стр}}} \{a_{ji}\} = \chi_j - 1$;

$n_{\text{стб}}, n_{\text{стр}}$ – соответственно количество столбцов и количество строк в двумерном массиве видеоданных.

Наличие ограничений в динамическом диапазоне фрагментов обусловлено: разной цветовой насыщенностью изображений. Максимальное значение динамического диапазона соответствует областям, окрашенным в белый цвет. Минимальное значение динамического диапазона имеют области черного цвета; когерентностью областей кадра (диапазон будет однородным для всей области); существованием фрагментов сильнонасыщенных мелкими деталями. Таким фрагментам соответствует неравномерность распределения динамического диапазона по элементам. Количество $V^{(2)}$ фрагментов A_v , элементы которых a_{ji} удовлетворяют ограничениям

$$a_{ji} \leq \psi_{ji} - 1, \quad 1 \leq j \leq n_{\text{стб}}, \quad 1 \leq i \leq n_{\text{стр}} \quad (2)$$

равно [3, 5]:

$$V^{(2)} = \prod_{j=1}^{n_{\text{стб}}} \prod_{i=1}^{n_{\text{стр}}} \psi_{ji}. \quad (3)$$

Количество $H_v^{(2)}$ информации в A_v с учетом ограничений (2) находится по формуле

$$H_v^{(2)} = \log_2 \left(\prod_{j=1}^{n_{\text{стб}}} \prod_{i=1}^{n_{\text{стр}}} \psi_{ji} \right) = \sum_{j=1}^{n_{\text{стб}}} \sum_{i=1}^{n_{\text{стр}}} \log_2 \psi_{ji}. \quad (4)$$

Понятно, что величина $V^{(2)}$ будет уменьшаться при понижении значений ψ_{ij} . В этом случае количество информации будет сокращаться. Из анализа соотношения (1) следует, что одним из направлений уменьшения величин ψ_{ij} является введение дополнительной величины h_{ij} , значение которой будет меньше величин λ_i и χ_j :

$$h_{ij} < \lambda_i \quad \text{и} \quad h_{ij} < \chi_j. \quad (5)$$

Величина h_{ij} может быть введена за счет перехода от двумерного представления фрагментов изображений к трехмерному представлению. Для варианта трехмерной организации видеоданных к двум координатам добавляется третья (по вертикали). Динамическим диапазоном элемента трехмерной структуры данных (ТСД) будет величина ψ_{ijz} , равная

$$\psi_{ijz} = \min \{ \chi_{j,z}; \lambda_{i,z}; h_{j,i} \}; \quad (6)$$

$$1 \leq j \leq n_{\text{стб}}, \quad 1 \leq i \leq n_{\text{стр}}, \quad 1 \leq z \leq n_c,$$

где $\chi_{j,z}, \lambda_{i,z}$ – динамический диапазон соответственно для j -го столбца и i -й строки z -го сечения ТСД:

$$\max_{1 \leq i \leq n_{\text{стр}}} \{ a_{jiz} \} = \chi_{jz} - 1; \quad 1 \leq j \leq n_{\text{стб}}, \quad 1 \leq z \leq n_c; \quad (7)$$

$$\max_{1 \leq j \leq n_{\text{стб}}} \{ a_{jiz} \} = \lambda_{iz} - 1; \quad 1 \leq i \leq n_{\text{стр}}; \quad 1 \leq z \leq n_c, \quad (8)$$

а $h_{j,i}$ – динамический диапазон для ($j; i$)-й вертикали ТСД

$$\max_{1 \leq z \leq n_c} \{ a_{jiz} \} = h_{ji} - 1; \quad 1 \leq j \leq n_{\text{стб}}; \quad 1 \leq i \leq n_{\text{стр}}; \quad (9)$$

n_c – количество сечений (высота ТСД).

Характеристики динамических диапазонов могут учитываться одновременно по трем координатам ТСД. Наличие дополнительных закономерностей по третьей координате (выполнение условия (3)) вызвано: наличием структурных особенностей, состоящих в том, что динамический диапазон по вертикали будет меньше, чем динамический диапазон горизонтального сечения ТСД; нестационарностью распределения цветовой насыщенности для различных фрагментов изображения. Количество $V^{(3)}$ трехмерных структур $A_v = \{ a_{jiz} \}; \quad 1 \leq j \leq n_{\text{стб}}, \quad 1 \leq i \leq n_{\text{стр}}, \quad 1 \leq z \leq n_c$, составленных из элементов a_{jiz} , удовлетворяющих ограничению $a_{jiz} \leq \psi_{jiz} - 1$ равно

$$V^{(3)} = \prod_{j=1}^{n_{\text{стб}}} \prod_{i=1}^{n_{\text{стр}}} \prod_{z=1}^{n_c} \psi_{jiz}. \quad (10)$$

Выражение (11) позволяет оценить информативность $H_v^{(3)}$ трехмерных структур данных в зависимости от системы ограничений $\Psi = \{ \psi_{jiz} \}, \quad 1 \leq j \leq n_{\text{стб}}, \quad 1 \leq i \leq n_{\text{стр}}, \quad 1 \leq z \leq n_c$:

$$H_v^{(3)} = \sum_{j=1}^{n_{\text{стб}}} \sum_{i=1}^{n_{\text{стр}}} \sum_{z=1}^{n_c} \log_2 \psi_{jiz}. \quad (11)$$

Сравним количество информации в фрагменте изображения для случаев его двумерной и трехмерной организации.

Количество информации $H_v^{(2)}$ в двумерном фрагменте определяется по выражению (4). Если данный фрагмент занимает z -ю позицию в трехмерной структуре, то по соотношению (11) количество информации $H_{v,z}^{(3)}$ в двумерном фрагменте как части ТСД равно

$$H_{v,z}^{(3)} = \sum_{j=1}^{n_{\text{стб}}} \sum_{i=1}^{n_{\text{стр}}} \log_2 \psi_{jiz}. \quad (12)$$

Поскольку размер двумерного фрагмента в обоих случаях одинаков, то сравнение между величинами $H_v^{(2)}$ и $H_{v,z}^{(3)}$ сводится к сравнению величин ψ_{ij} и ψ_{ijz} . Анализ выражений (1) и (6) показывает, что между динамическими диапазонами ψ_{ij} и ψ_{ijz} z -го фрагмента изображения существует следующая взаимосвязь:

$$\psi_{jiz} = \min\{\psi_{ji}; h_{ji}\}; 1 \leq j \leq n_{\text{стб}}, 1 \leq i \leq n_{\text{стр}}. \quad (13)$$

Тогда выполняется неравенство

$$\psi_{jiz} \leq \psi_{ji}, \quad 1 \leq j \leq n_{\text{стб}}, \quad 1 \leq i \leq n_{\text{стр}}. \quad (14)$$

С учетом неравенств (14), получим $H_{v,z}^{(3)} \leq H_v^{(2)}$.

При этом чем, больше позиций (ijz) , которым соответствует основание ψ_{jiz} равно h_{ji} , тем меньше количество информации для трехмерной организации относительно двумерной.

Для формирования компактного представления трехмерных структур данных необходимо разработать кодирование, учитывающее закономерности для динамических диапазонов одновременно по трем координатам.

Оценка степени сжатия видеоданных на основе трехмерного кодирования. Для варианта, когда кодограммы равномерной длины формируются для трехмерных полиадических чисел, содержащих переменное количество элементов, минимальное значение степени сжатия k_{\min} равно

$$k_{\min} = n_{\text{стб}} n_{\text{стр}} n_c b / H^{(3)} \leq k.$$

Результаты экспериментальной обработки фрагментов сильнонасыщенных изображений представлены в табл. 1.

Таблица 1

Зависимость k_{\min} от $n_{\text{стб}}$ $n_{\text{стр}}$ n_c

$n_{\text{стб}}$ $n_{\text{стр}}$ n_c	512	2048	4096
k_{\min}	1,4	1,9	1,98

Из анализа табл. 1 вытекает, что: минимальное значение степени сжатия достигает 2 раз; степень сжатия увеличивается при повышении размеров обрабатываемых трехмерных структур в среднем на 60%.

Выводы

1. Обоснованы требования для метода трехмерного кодирования, который должен: учитывать характеристики изображений одновременно по трем координатам; адаптироваться к конкретному содержанию ТСД; обрабатывать видеоданные в реальном времени без внесения погрешности; обеспечивать дополнительное сокращение избыточности за счет перехода от двумерной к трехмерной обработки.

2. Доказано, что наиболее универсальными структурными характеристиками, отличающимися не значительной зависимостью от нестационарности вероятностно-статистических свойств и свойств когерентности изображений, являются ограничения на динамический диапазон видеоданных.

3. Обосновано, что дополнительное увеличение количества сокращаемой избыточности по третьей координате трехмерных структур видеоданных вызвано: наличием структурных особенностей, состоящих в том, что динамический диапазон по вертикали будет меньше, чем динамический диапазон горизонтального сечения ТСД; нестационарностью распределения цветовой насыщенности для различных фрагментов изображения.

4. Построено выражение, позволяющее оценить информативность ТСД в зависимости от ограничений на динамический диапазон. Это позволило оценить минимальное значение степени сжатия изображений, достигающее 2 раз. Степень сжатия увеличивается при повышении размеров обрабатываемых трехмерных структур в среднем на 60%.

Список литературы

1. Ватолин В.И., Ратушняк А., Смирнов М., Юкин В. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. – М.: ДИАЛОГ – МИФИ, 2002. – 384 с.
2. Акушский И.Я., Заболоцкий В.Н. О комбинаторном подходе к идее сжатия информации // Цифровая вычислительная техника и программирование. – 1971. – № 6. – С. 5-17.
3. Риордан Дж. Введение в комбинаторный анализ. – М.: Изд-во иностр. литр., 1963. – 287 с.
4. Прэнтт У. Цифровая обработка изображений: Том 1, 2. – М.: Мир, 1985. – 736 с.
5. Баранник В.В. Рельефное представление изображений пирамидальным кодированием // Інформаційно-керуючі системи на залізничному транспорті. – 2002. – № 1. – С. 17-25.
6. Баранник В.В. Метод трехмерного кодирования данных Системы обработки інформації. – Х.: ХВУ. – 2003. – Вип. 1. – С. 42-46.

Поступила в редколлегию 7.09.2007

Рецензент: д-р техн. наук, проф. В.И. Хаханов, Харьковский национальный университет радиоэлектроники, Харьков.