УДК 621.397

КОМПЕНСАЦИЯ ДВИЖЕНИЯ В ВИДЕОКОДЕКАХ НА ОСНОВЕ ОПТИМАЛЬНЫХ РАЗБИЕНИЙ КОДИРУЕМЫХ БЛОКОВ

д-р техн. наук, доц. С.В. ДВОРНИКОВ
(Военная академия связи им. Маршала Советского Союза С.М. Буденного, Санкт-Петербург);
В.В. ЦВЕТКОВ, А.А. УСТИНОВ
(ЗАО «НПФ «ТИРС», Санкт-Петербург)

Предлагается метод компенсации движения в видеокодеках на основе оптимальных разбиений кодируемых блоков. Сформулирована задача разбиения кодируемых трехмерных фрагментов подвижных изображений на непересекающиеся группы. Доказывается, что снижение межкадровых различий внутри каждой группы приводит к уменьшению числа значимых коэффициентов трехмерного преобразования. Обосновывается, что уменьшение числа значимых коэффициентов преобразования приводит к увеличению коэффициента сжатия.

Теоретические основы сжатия подвижных изображений на основе трехмерных ортогональных преобразований

Прямое трехмерное дискретное косинусное преобразование (ДКП-3) для куба размером $N{\times}N{\times}N$ задается следующим образом [1]:

$$F(i,j,k)=\sqrt{\frac{8}{N^3}}c(i)c(j)c(k)\sum_{z=0}^{N-1}\sum_{y=0}^{N-1}\sum_{x=0}^{N-1}f(x,y,z)\cos\left[\frac{(2x+1)\pi i}{2N}\right]\cos\left[\frac{(2y+1)\pi j}{2N}\right]\cos\left[\frac{(2z+1)\pi k}{2N}\right],$$

где $f(x, y, z)$ – значение яркостной или цветоразностной компонент пикселя с координатами $x,y,z\in[0,...,N-1]$; $F(i, j, k)$ – коэффициент преобразования с координатами $i,j,k\in[0,...,N-1]$, функция $c(k)$ определяется как

$$c(k)=\begin{cases}1/2, & k=0;\\ 1, & k\neq 0.\end{cases}$$

Обратное дискретное косинусное преобразование вычисляется следующим образом:

$$f(x,y,z)=\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\sum_{k=0}^{N-1}\sqrt{\frac{8}{N^3}}c(i)c(j)c(k)F(i,j,k)\cos\left[\frac{(2y+1)\pi j}{2N}\right]\cos\left[\frac{(2z+1)\pi k}{2N}\right].$$

Свойства трехмерного дискретного косинусного преобразования

К основным свойствам ДКП-3 следует отнести такие, как:

- локализация большой части энергии сигнала в небольшом числе коэффициентов преобразования. Данное свойство позволяет исключить наименее значимые коэффициенты из рассмотрения при кодировании с потерями [1];
- ДКП-3 представляет собой ортогональное преобразование, которое может быть реализовано путем последовательного выполнения одномерных ДКП по строкам, столбцам и оси времени [1].

Данные свойства определили широкое применение ДКП-3 при сжатии подвижных изображений. Типовая схема кодека на основе ДКП-3 предполагает выполнение следующих основных операций: входная обработка, ДКП-3, квантование коэффициентов преобразования, энтропийное кодирование. На приеме данные процедуры выполняются в обратной последовательности [1].

Однако в условиях кодирования высокодинамичных сцен ДКП-3 может быть не оптимальным в смысле концентрации энергии в незначительном числе коэффициентов, расположенных на временной оси (в различных кадрах кодируемого блока).

Для примера на рисунке показаны квантованные коэффициенты ДКП-3 для кодируемого блока размером 8 × 8 × 8. Данный блок был намеренно выбран из высокодинамичной области кадров тестового подвижного изображения «foreman».

Анализ представленных коэффициентов ДКП-3 показывает наличие ненулевых коэффициентов в каждой плоскости куба 8 × 8 × 8.

Наиболее распространенным подходом к устранению данного недостатка является изменение размеров кодируемых трехмерных блоков [2 – 8]. В основе данного подхода лежит оценка степени подвижности каждого трехмерного кодируемого куба по нескольким фиксированным градациям. Далее для каждой градации определяется оптимальный размер кодируемого куба, например, 16 × 16 × 1, 16 × 16 × 8 или 8 × 8 × 8. Недостатки такого подхода очевидны. Во-первых, градация на заданное число уровней подвижности не вполне обоснованна и носит эмпирический характер. Во-вторых, фиксированное число уровней градации не может описать всего многообразия степени изменения сцен в трехмерных кодируемых фрагментах.

а)

183	–3	0	–1	0	0	0	0
–12	–1	0	0	0	0	0	0
1	0	0	–1	0	0	0	0
–1	–1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

б)

3	5	0	1	0	0	0	0
–11	–1	0	–1	0	0	0	0
–1	0	0	0	0	0	0	0
–1	–1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
–1	0	0	0	0	0	0	0

в)

32	2	–1	0	0	0	0	0
–10	–2	1	0	0	0	0	0
–2	–1	0	0	0	0	0	0
–1	1	–1	0	0	0	0	0
0	0	0	0	0	0	0	0
–1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
–1	0	0	0	0	0	0	0

г)

17	2	0	0	0	0	0	0
–1	2	0	0	0	0	0	0
0	–2	–1	0	0	0	0	0
–1	1	0	0	0	0	0	0
–1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
–1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

д)

16	0	0	0	0	0	0	0
3	1	–1	0	0	0	0	0
–1	1	0	0	0	0	0	0
–1	–1	1	0	0	0	0	0
–1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
–1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

е)

2	–1	2	0	0	0	0	0
4	0	0	1	0	0	0	0
–3	1	1	0	0	0	0	0
–1	0	1	0	0	0	0	0
–1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

ж)

4	–2	–1	1	0	0	0	0
4	1	1	0	0	0	0	0
–1	1	1	0	0	0	0	0
–2	–1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

з)

1	–3	–1	0	0	0	0	0
2	1	0	1	0	0	0	0
–1	0	0	0	0	0	0	0
–1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Пример коэффициентов ДКП-3 для кодируемого блока размером 8×8×8:
а – 1-го слоя куба; *б* – 2-го слоя куба; *в* – 3-го слоя куба; *г*) 4-го слоя куба;
д) 5-го слоя куба; *е* – 6-го слоя куба; *ж* – 7-го слоя куба ; *з* – 8-го слоя куба

В следующем пункте рассмотрим развитие данной идеи на основе оптимального разбиения кодируемого блока на непересекающиеся наборы кадров.

Формализация задачи поиска оптимального разбиения трехмерных кубов на непересекающиеся области

Пусть на вход кодера поступают L кадров подвижного изображения размером $P_1 \times N_1$ пикселей. Разделим видеоданные на непересекающиеся области (кубы) размером $P \times N \times L$ пикселей в виде трехмерных матриц $A_{P\times N\times L}$. С целью упрощения дальнейших рассуждений трехмерную матрицу $A_{P\times N\times L}$ представим множеством из L одномерных векторов размерностью $P \times N$ элементов: $A_{P\times N\times L} \to \left\{\vec{A}_1, \vec{A}_2, ..., \vec{A}_L\right\}$. Отметим, что любой i-й вектор сформированного множества из L векторов получен путем последовательной N-кратной канкатенации P-мерных столбцов i-го слоя матрицы $A_{P\times N\times L}$ в единый вектор размерности $P \times N$. Используя терминологию линейной алгебры, множество $\left\{\vec{A}\right\} = \left\{\vec{A}_1, \vec{A}_2, ..., \vec{A}_L\right\}$ будем рассматривать как множество из L точек в евклидовом пространстве размерности $P \times N$ ($E^{P\times N}$).

Разделим исходное множество $\{\vec{A}\}$ на R подмножеств $\{\vec{A}\}_1$, $\{\vec{A}\}_2$, ..., $\{\vec{A}\}_R$ так, чтобы выполнялись условия:

$$\{\vec{A}\}_1 \cup \{\vec{A}\}_2 \cup, \ldots, \cup \{\vec{A}\}_R = \{\vec{A}\};$$

$$\{\vec{A}\}_1 \cap \{\vec{A}\}_2 \cap, \ldots, \cap \{\vec{A}\}_R = \{0\}.$$

Первое условие обеспечивает объединение R подмножеств в единое множество $\{\vec{A}\}$. Второе условие обеспечивает то, что R образованных подмножеств являются непересекающимися.

Исходя из необходимости достижения максимального сжатия, будем считать некоторое разбиение множества $\{\vec{A}\}$ оптимальным, если суммарное число ненулевых коэффициентов преобразования после выполнения ДКП-3 над каждым образованным подмножеством является оптимальным.

Введем функцию $f_{ДКП\text{-}3}\left(\{\vec{A}\}_i\right)$, которая определяет число ненулевых (значимых) коэффициентов преобразования, полученных после выполнения ДКП-3 над i-м подмножеством. Тогда формально задачу поиска оптимального разбиения можно записать следующим образом:

$$\sum_{k=1}^{R} f_{ДКП\text{-}3}\left(\{\vec{A}\}_k\right) \to \min_{\{\vec{A}\}_k \ \forall k=\overline{1,R}}. \quad (1)$$

Задача (1) предполагает поиск некоторого разбиения на конечном множестве возможных разбиений, для которого суммарное число ненулевых коэффициентов преобразования по каждому подмножеству разбиения минимально. В этом смысле данная задача является целочисленной. Кроме того, в ходе решения задачи необходимо для каждого разбиения выполнять ДКП-3 по всем сформированным подмножествам. В связи с этим метод решения данной задачи, основанный на переборе возможных разбиений с последующим выполнением ДКП-3 для найденных подмножеств, является сложным.

Поэтому рассмотрим постановку задачи поиска оптимального разбиения, которая не требует вычисления ДКП-3 для оценки количества значимых коэффициентов.

Известно, что ДКП-3 является наиболее эффективным, если корреляция между соседними элементами внутри обрабатываемого куба является относительно высокой [9]. Коэффициент корреляции между соседними элементами можно сопоставить со степенью их похожести или близости. Выберем в качестве меры близости средний квадрат ошибки (СКО). Тогда можно утверждать, что чем меньше СКО между элементами, тем выше коэффициент корреляции между ними. Следовательно, для решения задачи (1) без выполнения ДКП-3 необходимо найти такое разбиение исходного множества на R непересекающихся подмножеств, при котором суммарный СКО между элементами каждого из подмножеств был бы минимальным.

Для формальной постановки такой задачи необходимо математически описать следующее:

- способ разбиения исходного множества $\{\vec{A}\}$ на R непересекающихся подмножеств;

- способ вычисления СКО между элементами каждого из подмножеств.

Разбиение исходного множества $\{\vec{A}\}$ на R непересекающихся подмножеств зададим в виде матрицы X размером $R\times L$ элементов. На элементы матрицы X наложим ограничения:

$$X(i, j) \in \{1, 0\}; \quad (2)$$

$$\sum_{i=1}^{R} X(i, j) = 1, \quad \forall j = \overline{1, L}. \quad (3)$$

Ограничение (2) формально описывает способ разбиения в следующих случаях:

$$\text{если } X(i, j) = 1, \quad \text{то } \vec{A}_j \in \{\vec{A}\}_i;$$

$$\text{если } X(i, j) = 0, \quad \text{то } \vec{A}_j \notin \{\vec{A}\}_i.$$

Ограничение (3) обеспечивает выполнение условия непересекаемости подмножеств: каждый элемент $\vec{A}_j$ $j = 1, 2, ..., L$ может принадлежать только одному подмножеству.

Для вычисления СКО между элементами определим средний элемент i-го подмножества в виде [2]

$$\hat{\vec{A}}_i = \frac{[A]\vec{X}_i^T}{\vec{X}_i\,\vec{1}},$$

где $[A]$ – матрица размером $P\times L$ элементов; $\vec{X}_i$ – i-я строка матрицы X. Отметим, что каждый j-й столбец матрицы $[A]$ ($\vec{A}_j$), является j-м элементом исходного множества $\{\vec{A}\}$.

Если j-й элемент принадлежит i-му подмножеству, то данный элемент определим в виде $\vec{A}_{j\in\{\vec{A}\}}$. Тогда СКО в i-й группе определим следующим образом:

$$СКО_i = \sum_{j=1}^{n_i}\left(\vec{A}_{j\in\{\vec{A}\}_i} - \frac{[A]\vec{X}_i^T}{\vec{X}_i\,\vec{1}}\right)^T\left(\vec{A}_{j\in\{\vec{A}\}_i} - \frac{[A]\vec{X}_i^T}{\vec{X}_i\,\vec{1}}\right), \tag{4}$$

где n_i – число элементов множества $\{\vec{A}\}$, принадлежащих i-му подмножеству; $\vec{1}$ – единичный вектор-столбец размерности $L\times 1$.

Исходя из физического смысла матрицы X, принадлежность j-го элемента i-му множеству можно представить в виде произведения элемента на $X(i, j)$. В этом случае суммарное значение СКО по всем R подмножествам запишем в виде

$$СКО = \sum_{i=1}^{R}\sum_{j=1}^{L}\left(\vec{A}_j X(i,j) - X(i,j)\frac{[A]\vec{X}_i^T}{\vec{X}_i\,\vec{1}}\right)^T\left(\vec{A}_j X(i,j) - X(i,j)\frac{[A]\vec{X}_i^T}{\vec{X}_i\,\vec{1}}\right). \tag{5}$$

Раскрыв скобки в выражении (5), получим:

$$СКО = \sum_{i=1}^{R}\sum_{j=1}^{L}\left(\vec{A}_j^T\vec{A}_j X^2(i,j) - 2X^2(i,j)\frac{\vec{A}_j^T[A]\vec{X}_i^T}{\vec{X}_i\,\vec{1}} + X^2(i,j)\frac{\vec{X}_i\,[A]^T[A]\vec{X}_i^T}{(\vec{X}_i\,\vec{1})^2}\right). \tag{6}$$

Проанализируем каждое слагаемое в выражении (6) с учетом суммирования:

- первое слагаемое

$$\sum_{i=1}^{R}\sum_{j=1}^{L}\vec{A}_j^T\vec{A}_j X^2(i,j) = \sum_{j=1}^{L}\vec{A}_j^T\vec{A}_j\sum_{i=1}^{R}X^2(i,j) = \sum_{j=1}^{L}\vec{A}_j^T\vec{A}_j,$$

поскольку $\sum_{i=1}^{R}X^2(i,j) = \sum_{i=1}^{R}X(i,j) = 1$ в соответствии с ограничениями (2) и (3);

- второе слагаемое

$$2\sum_{i=1}^{R}\sum_{j=1}^{L}X^2(i,j)\frac{\vec{A}_j^T[A]\vec{X}_i^T}{\vec{X}_i\,\vec{1}} = 2\sum_{j=1}^{L}\vec{A}_j^T\sum_{i=1}^{R}\frac{X^2(i,j)[A]\vec{X}_i^T}{\vec{X}_i\,\vec{1}} = 2\sum_{j=1}^{L}\frac{\vec{X}_i[A]^T[A]\vec{X}_i^T}{\vec{X}_i\,\vec{1}};$$

- третье слагаемое

$$\sum_{i=1}^{R}\sum_{j=1}^{L}X^2(i,j)\frac{\vec{X}_i\,[A]^T[A]\vec{X}_i^T}{(\vec{X}_i\,\vec{1})^2} = \sum_{j=1}^{L}X^2(i,j)\sum_{i=1}^{R}\frac{\vec{X}_i\,[A]^T[A]\vec{X}_i^T}{(\vec{X}_i\,\vec{1})^2} = \sum_{i=1}^{R}\frac{\vec{X}_i\,[A]^T[A]\vec{X}_i^T}{(\vec{X}_i\,\vec{1})},$$

поскольку $\sum_{j=1}^{L}X^2(i,j) = \vec{X}_i\,\vec{1}$, при выполнении ограничения (2).

С учетом полученных выражений для 1 – 3 слагаемых выражение (6) окончательно запишем в виде

$$СКО = \sum_{j=1}^{L}\vec{A}_j^T\vec{A}_j - \sum_{i=1}^{R}\frac{\vec{X}_i\,[A]^T[A]\vec{X}_i^T}{(\vec{X}_i\,\vec{1})}. \tag{7}$$

Как было отмечено ранее, оптимальным разбиением является разбиение, при котором минимизируется величина СКО. Следовательно, задача поиска оптимального разбиения сводится к минимизации выражения (7) по всем возможным матрицам разбиений. Поскольку первое слагаемое в выражении (7) от искомой переменной не зависит, то окончательно задачу поиска оптимального разбиения можно записать как

$$\sum_{i=1}^{R}\frac{\vec{X}_i\,[A]^T[A]\vec{X}_i^T}{(\vec{X}_i\,\vec{1})}\to\max_{\vec{X}_i,\ \forall i=\overline{1,R}} \tag{8}$$

при ограничениях (2) и (3) на искомые значения элементов матрицы X.

Решением задачи (8) является матрица X, задающая оптимальное разбиение исходного трехмерного куба размером $P\times N\times L$ пикселей на R непересекающихся подмножеств.

Заключение. Для практической реализации алгоритма сжатия подвижных изображений на основе ДКП-3 и поиска оптимального разбиения трехмерных кубов на подмножества необходимо осуществить решение оптимизационной задачи (8). Данная задача относится к классу задач нелинейного целочисленного программирования. Её решение и анализ эффективности предложенного решения компенсации движения требует дальнейших исследований.

ЛИТЕРАТУРА

1. Беляев, Е.А. Сжатие видеоинформации на основе трехмерного дискретного псевдо-косинусного преобразования для энергоэффективных систем видеонаблюдения / Е.А. Беляев, Т.М. Сухов, Н.Н. Шостацкий // Компьютерная оптика. – Т. 34, № 2. – С. 260 – 272, 210.
2. Use of an Adaptive 3D-DCT Scheme for Coding Multiview Stereo Images / N.P. Sgouros [et al.] // IEEE Proceedings of ISSPIT. – 2005. – P. 180 – 185.
3. Chan, Y.L. Variable Temporal-length 3-D Discrete Cosine Transform Coding / Y.L. Chan, W.C. Siu // IEEE Transactions on Image Processing. – 1997. – Vol 6, №. 5. – P. 758 – 763.
4. Koivusaari, J.J. Simplified Three-Dimensional Discrete Cosine Transform Based Video Codec / J.J. Koivusaari, J.H. Takala // SPIE Proc. of Multimedia on Mobile Devices. – 2005. – P. 11 – 21.
5. Dugad, R. A Fast Scheme for Image Size Change in The Compressed Domain / R. Dugad, N. Ahuja // IEEE Transactions on Circuits and Systems for Video Technology. – 2001. – P. 461 – 474.
6. Koivusaari, J.J. Image Coding Using Adaptive Resizing in The Block-DCT Domain / J.J. Koivusaari, J.H. Takala, M. Gabbouj // SPIE Proc. of Multimedia on Mobile Devices II. – 2006. – P. 1 – 9.
7. Use of Adaptive Resizing in 3-D DCT Domain for Video Coding / J. Li [et al.] // Picture Coding Symp., Lisbon, Portugal. – 2007. – P. 7 – 9.
8. Dai, Q. Fast Algorithms for Multidimensional DCT-to-DCT Computation Between a Block and Its Associated Subblocks / Q. Dai, X. Chen, C. Lin // IEEE Trans. Circuits and Syst. for Video Tech. – 2003. – Vol. 13, № 7. – P. 717 – 725.
9. Ахмед, Н. Ортогональные преобразования при обработке цифровых сигналов / Н. Ахмед, К.Р. Рао. – Л.: Связь, 1980. – 248 с.

Поступила 16.05.2013

COMPENSATION OF MOTION IN VIDEO CODECS ON THE BASIS OF OPTIMAL PARTITION OF ENCODED BLOCKS

S. DVORNIKOV, V. TSVETKOV, A. USTINOV

A method for compensation of motion in video codecs on the basis of optimal partition of encoded blocks is proposed. The aim to divide encoded three-dimensional fragments of movable images into non-intersecting groups is formulated. It is proven, that reduction of interframe differences inside of each group leads to reduction of the number of significant coefficients of three-dimensional transformation. It is grounded, that reduction of the number of significant coefficients of transformation leads to increase of the coefficient of compression.