

УДК 621.391

ЭФФЕКТИВНОЕ КОДИРОВАНИЕ ИЗОБРАЖЕНИЙ ОТ РАСПРЕДЕЛЕННЫХ В ПРОСТРАНСТВЕ ВИДЕОКАМЕР С УСТРАНЕНИЕМ МЕЖРАКУРСНОЙ ИЗБЫТОЧНОСТИ

О.ДЖ. АЛЬ-ФУРАЙДЖИ, В.К. КОНОПЕЛЬКО, В.Ю. ЦВЕТКОВ

Белорусский государственный университет информатики и радиоэлектроники
П. Бровка, 6, Минск, 220013, Беларусь

Поступила в редакцию 10 декабря 2011

Предложен метод распределенного эффективного кодирования многокурсовых изображений, учитывающий межкадровую избыточность их перекрывающихся фрагментов, формируемых несколькими видеокамерами. Метод основан на определении границ областей соответствия фрагментов многокурсового изображения с использованием реперных точек. Метод позволяет повысить коэффициент сжатия многокурсовых изображений за счет исключения повторной передачи информации об областях соответствия их перекрывающихся фрагментов.

Ключевые слова: распределенное эффективное кодирование, многокурсовые изображения, межкадровая избыточность изображений.

Введение

В настоящее время постоянно расширяется круг задач, связанных с применением сетевого видеомониторинга, предполагающего использование множества видеокамер, объединенных через телекоммуникационную сеть с серверами обработки и распределения видеоданных. Это охрана территорий, оптическая локация и навигация, виртуальный туризм и т.д. [1]. Одной из центральных проблем в данных системах является сжатие видеoinформации в целях сокращения ее объема при передаче, уменьшения нагрузки на каналы связи и сетевые узлы, а также экономии памяти при записи и хранении. Для сжатия видеоданных используются методы эффективного кодирования, основанные на устранении статистической (кодирование Хаффмана [2], арифметическое кодирование [3]), пространственной (JPEG [4], JPEG2000 [5], EZW [6], SPIHT [7], SPECK [8]) и временной (Cinepak [9], MotionJPEG [10], MPEG [11]) избыточности видеоданных. Данные методы позволяют достичь высоких коэффициентов сжатия при кодировании изображений от одного источника (т.е. однокурсовых). Однако они не эффективны при кодировании подвижных и неподвижных изображений в условиях сетевого видеомониторинга, т.к. не учитывают корреляцию между фрагментами многокурсового изображения, формируемыми несколькими видеокамерами с перекрывающимися секторами обзора. Данный недостаток приводит к избыточной передаче многокурсовых изображений, увеличению загруженности сети и нагрузки на серверы обработки видеоданных. Он может быть устранен в результате определения границ соответствующих областей фрагментов многокурсового изображения и исключения повторной передачи информации об этих областях. Идеи распределенного кодирования впервые сформулированы в работах [12, 13], однако они носят общий характер и не могут непосредственно использоваться для сжатия видеоданных в условиях сетевого видеомониторинга.

Целью работы является разработка метода распределенного эффективного кодирования многокурсовых изображений с устранением межкадровой избыточности их фрагментов на основе определения соответствующих областей этих фрагментов.

Методы определения соответствующих областей и сшивки изображений

Для определения границ соответствующих областей перекрывающихся фрагментов многокурсовых изображений могут использоваться различные методы локализации и параметризации реперов и поиска соответствия.

Наиболее распространенными методами локализации и параметризации реперов являются SIFT (Scale Invariant Feature Transform) [14] и SURF (Speeded-Up Robust Features) [15].

Метод SIFT [14] использует многоуровневое кратномасштабное представление изображения, формируемое с помощью обычного и дифференциального преобразований Гаусса [16], для локализации реперных областей изображения, устойчивых к изменению яркости, контраста и действию шума. Как правило, используются 4 уровня кратномасштабного представления, каждый из которых объединяет группу из 5 аппроксимированных образов изображения одного масштаба, полученных в результате свертки с ядром Гаусс-фильтра размером 11×11 , 15×15 , 21×21 , 29×29 пикселей и 4 дифференциальных гауссовских образов, полученных в результате пиксельной разности аппроксимированных образов одного масштаба с близкими значениями ядра Гаусс-фильтра. Аппроксимированные образы верхнего уровня в 4 раза меньше по сравнению с аппроксимированными образами соседнего нижнего уровня и формируются в результате децимации.

В методе SURF реперные области локализуются в пространстве Хессиян-матрицы [15], формируемой в результате свертки интегрального изображения и бокс-фильтров (аппроксимированных Гаусс-фильтров). Использование интегрального изображения позволяет существенно сократить вычислительную сложность локализации реперных областей [16]. Метод SURF оперирует пятью уровнями кратномасштабного представления изображения, на каждом из которых формируются по 4 аппроксимированных образа. При этом используются 30 различных бокс-фильтров размером 9×9 , 15×15 , 21×21 , 27×27 , 39×39 , 51×51 , 75×75 , 99×99 , 147×147 , 195×195 , 291×291 и 387×387 пикселей. Поиск реперных точек осуществляется в трехмерном пространстве с использованием методов, предложенных в работах [17, 18].

Для пары перекрывающихся фрагментов (рис. 1,а) многокурсового изображения локализация и параметризация реперов включает следующие шаги: выделение реперных областей (рис. 1,б), локализация реперов и формирование идентификационных параметров (дескрипторов) реперов (рис. 1,в). Эти шаги выполняются для каждого фрагмента независимо.

Следующим этапом является поиск соответствий фрагментов, который включает два основных этапа: поиск уникальных реперов на каждом фрагменте (рис. 2,а) и поиск соответствий между уникальными реперами двух фрагментов (рис. 2,б).

Наиболее эффективным методом поиска соответствующих областей является RANSAC (RANdom SAmple Consensus) [19]. Он основан на вычислении матрицы соответствия, которая устанавливает связь между четырьмя случайно выбранными уникальными реперами на одном из фрагментов и уникальными реперами другого фрагмента. Поиск продолжается до тех пор, пока ошибка совмещения фрагментов, преобразованных согласно матрицы соответствия, не станет меньше заданного порога.

После определения соответствующих областей фрагментов многокурсового изображения может осуществляться их сшивка, которая предполагает геометрическое преобразование фрагментов (рис. 3,а) на основе информации об относительном положении соответствующих реперов на этих фрагментах, обрезку изображений в области перекрытия (рис. 3,б) и совмещение обрезанных изображений со сглаживанием границы (рис. 3,в) [20].

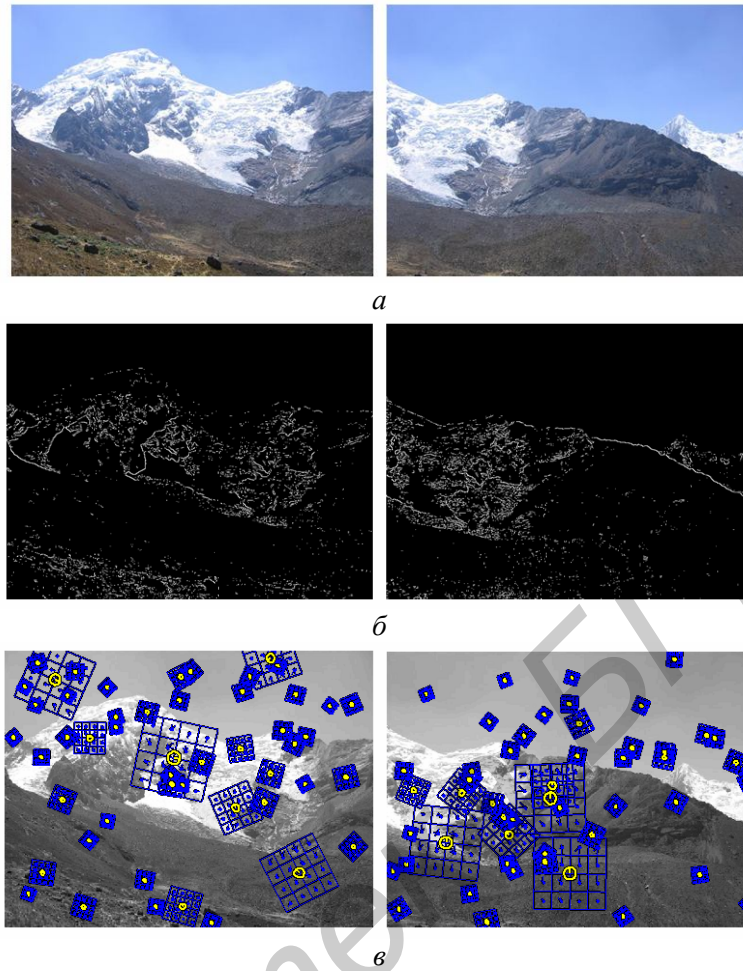


Рис. 1. Операции локализации и параметризации реперов на изображениях: *a* – исходные изображения; *б* – выделение реперных областей; *в* – локализация реперов и формирование их идентификационных параметров

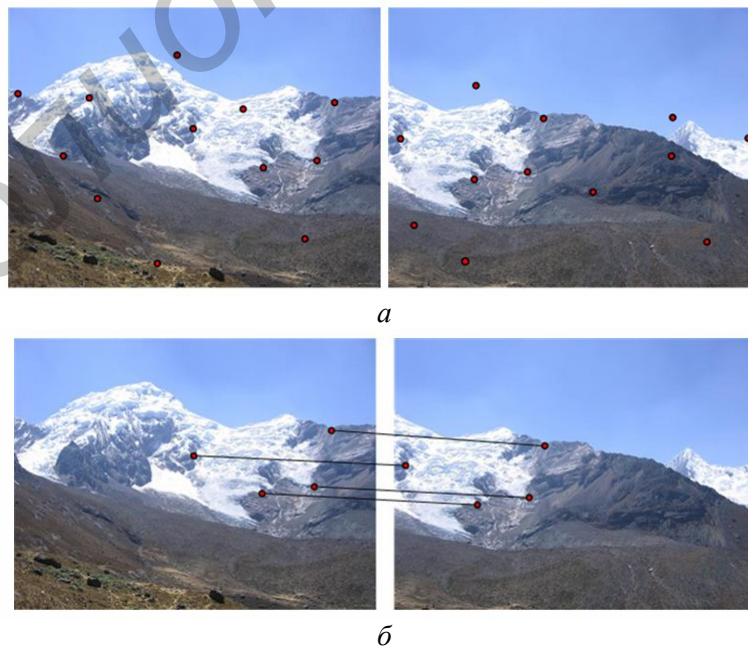


Рис. 2. Операции поиска соответствующих областей изображений: *a* – поиск уникальных реперов; *б* – поиск соответствующих реперов

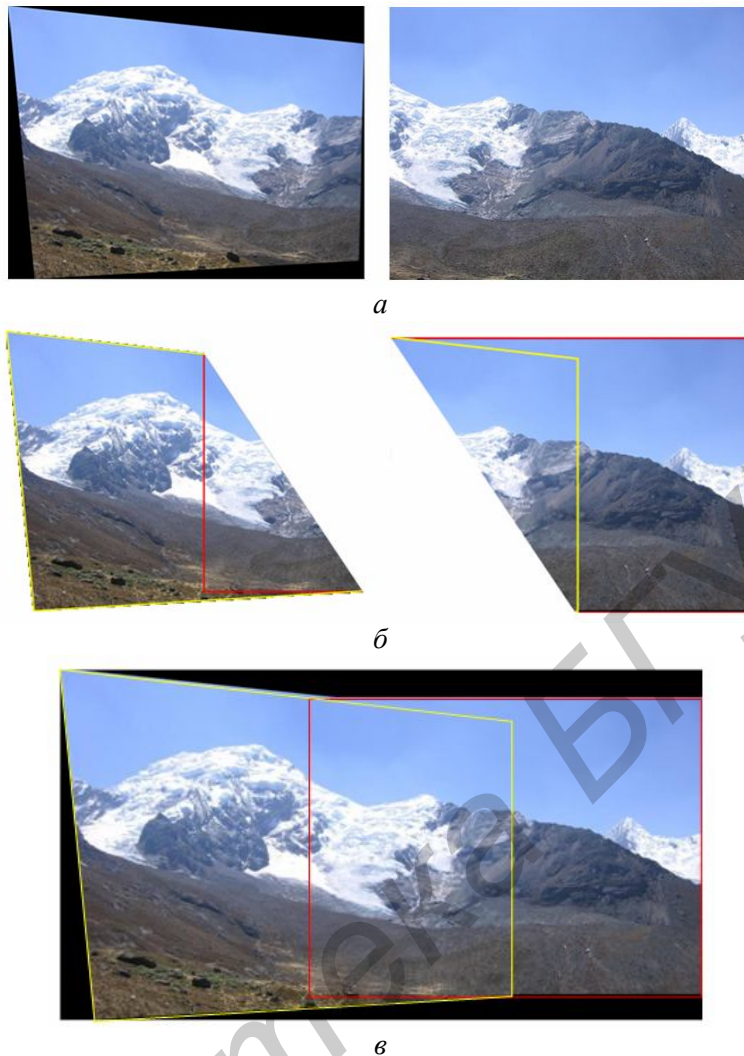


Рис. 3. Операции шивки изображений: *а* – геометрическое преобразование фрагментов; *б* – обрезка фрагментов в области перекрытия; *в* – совмещение обрезанных фрагментов

Метод распределенного эффективного кодирования многокадровых изображений с устранением межкадровой избыточности фрагментов

Для уменьшения объема передаваемых видеоданных и снижения нагрузки на каналы передачи, узлы коммутации и маршрутизации в условиях сетевого видеомониторинга предлагается метод распределенного эффективного кодирования многокадровых изображений, учитывающий межкадровую избыточность их перекрывающихся фрагментов, формируемых несколькими видеокameraми. Метод основан на определении границ областей соответствия фрагментов многокадрового изображения с использованием реперных точек. Суть метода заключается в использовании информации о соответствующих областях фрагментов многокадрового изображения для исключения повторного кодирования и передачи информации о пикселях этих областей. Метод позволяет повысить коэффициент сжатия многокадровых изображений за счет устранения межкадровой избыточности их фрагментов, уменьшить нагрузку на каналы и узлы в сети за счет исключения повторной передачи информации о пикселях соответствующих областей и снизить нагрузку на серверы обработки многокадровых изображений за счет распределенной обработки перекрывающихся фрагментов на уровне видеокодеров.

Метод состоит из следующих шагов.

1. Выделение реперных областей на фрагментах многокадрового изображения. Для каждого i -го фрагмента $M_F(i) = \|m_F(i, y, x)\|_{(y=0, Y-1, x=0, X-1)}$ многокадрового изображения фор-

мируются матрицы $\tilde{M}_Y(i) = \|\tilde{m}_Y(i, n)\|_{(n=0, \tilde{N}(i)-1)}$ и $\tilde{M}_X(i) = \|\tilde{m}_X(i, n)\|_{(n=0, \tilde{N}(i)-1)}$ координат контурных пикселей, устойчивых к изменению масштаба, яркости и контраста фрагментов, где $i = \overline{0, I-1}$; Y, X – размеры фрагментов по вертикали и горизонтали; I – число фрагментов; n – номер контурного пикселя; $\tilde{N}(i)$ – число контурных пикселей, устойчивых к изменению масштаба, яркости и контраста i -го фрагмента.

2. Локализация реперов на изображении. Для каждого i -го фрагмента $M_F(i)$ многоакурсного изображения формируются матрицы $\hat{M}_Y(i) = \|\hat{m}_Y(i, n)\|_{(n=0, \hat{N}(i)-1)}$ и $\hat{M}_X(i) = \|\hat{m}_X(i, n)\|_{(n=0, \hat{N}(i)-1)}$ координат реперных пикселей, где $\hat{N}(i)$ – число реперов i -го фрагмента, удовлетворяющее условию $\hat{N}(i) \leq \tilde{N}(i)$.

3. Параметризация реперов. Для каждого i -го фрагмента $M_F(i)$ многоакурсного изображения формируется набор $\hat{M}_D(i) = \|\hat{M}_D(i, n)\|_{(n=0, \hat{N}(i)-1)}$ идентификационных параметров реперов, где $\hat{M}_D(i, n) = \|\hat{m}_D(i, n, p)\|_{(p=0, P-1)}$ – идентификатор n -го репера i -го фрагмента; $\hat{m}_D(i, n, p)$ – p -й элемент идентификатора n -го репера; P – число элементов в идентификаторах.

4. Поиск уникальных реперов в пределах каждого фрагмента. Для каждого i -го фрагмента $M_F(i)$ многоакурсного изображения формируются матрицы $M_Y(i) = \|m_Y(i, n)\|_{(n=0, N(i)-1)}$ и $M_X(i) = \|m_X(i, n)\|_{(n=0, N(i)-1)}$ координат и набор $M_D(i) = \|M_D(i, n)\|_{(n=0, N(i)-1)}$ идентификационных параметров уникальных реперных пикселей, где $M_D(i, n) = \|m_D(i, n, p)\|_{(p=0, P-1)}$ – идентификатор n -го уникального репера i -го фрагмента; $m_D(i, n, p)$ – p -й элемент идентификатора n -го уникального репера; $N(i)$ – число уникальных реперов i -го фрагмента, удовлетворяющее условию $N(i) \leq \hat{N}(i)$. Идентификаторы уникальных реперов i -го фрагмента должны удовлетворять условию

$$\neg \exists M_D(i, n_1) \neg \exists M_D(i, n_2) (\forall m_D(i, n_1, p_1) \exists m_D(i, n_2, p_2) (m_D(i, n_1, p_1) = m_D(i, n_2, p_2))) \quad (1)$$

при $n_1 \in [0, N(i)]$, $n_2 \in [0, N(i)]$, $p_1 = \overline{0, P-1}$, $p_2 \in [0, P]$.

5. Обмен идентификационными параметрами реперов между видеокодерами соседних узлов. Если фрагменты $M_F(i)$ и $M_F(j)$ многоакурсного изображения сформированы видеокамерами i и j с перекрывающимися секторами обзора, то между соответствующими видеокодерами осуществляется обмен координатными и идентификационными параметрами уникальных реперов

$\{m_Y(i, n_1), m_X(i, n_1), M_D(i, n_1)\}_{(n_1=0, N(i)-1)}$ и $\{m_Y(j, n_2), m_X(j, n_2), M_D(j, n_2)\}_{(n_2=0, N(j)-1)}$ фрагментов $M_F(i)$ и $M_F(j)$. Обмен координатными и идентификационными параметрами уникальных реперов может осуществляться через концентратор или непосредственно (рис. 4).

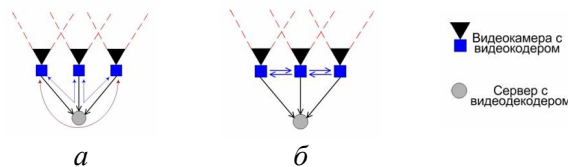


Рис. 4. Обмен идентификационными параметрами реперов между видеокодерами: a – через концентратор; b – напрямую

6. Поиск соответствий фрагментов. Для каждого i -го фрагмента $M_F(i)$ многокурсового изображения в i -м видеокодеке формируются наборы $\{\{\ddot{m}_Y(i, n_k), \ddot{m}_Y(k, n_k)\}, \{\ddot{m}_X(i, n_k), \ddot{m}_X(k, n_k)\}\}_{(k=0, \overline{j(i)-1}, n_k=0, \overline{N_k(i)-1})}$ координат соответствующих уникальных реперов, где $\ddot{I}(i)$ – число фрагментов, перекрывающихся с рассматриваемым i -м фрагментом; $\ddot{N}_k(i)$ – число соответствующих уникальных реперов для i -го и k -го фрагментов. Уникальный репер i -го фрагмента с номером n_i ($n_i = \overline{0, N(i)-1}$), координатами и идентификатором $\{m_Y(i, n_i), m_X(i, n_i), M_D(i, n_i)\}$ и уникальный репер k -го фрагмента с номером n_k ($n_k = \overline{0, N(k)-1}$), координатами и идентификатором $\{m_Y(k, n_k), m_X(k, n_k), M_D(k, n_k)\}$ считаются соответствующими, если выполняется условие

$$\forall m_D(i, n_i, p_i) \exists m_D(k, n_k, p_k) (m_D(i, n_i, p_i) = m_D(k, n_k, p_k)), \quad (2)$$

где $p_i = \overline{0, P-1}$, $p_k \in [0, P]$.

При выполнении условия (2) фиксируются координаты соответствующих реперов: $\ddot{m}_Y(i, n_k) = m_Y(i, n_i)$, $\ddot{m}_Y(k, n_k) = m_Y(k, n_k)$, $\ddot{m}_X(i, n_k) = m_X(i, n_i)$, $\ddot{m}_X(k, n_k) = m_X(k, n_k)$.

7. Обрезка фрагментов. В каждом i -м видеокодеке формируется обрезанный фрагмент $\bar{M}_F(i) = \|\bar{m}_F(i, y, x)\|_{(y=0, \overline{Y-1}, x=0, \overline{X-1})}$ многокурсового изображения в результате удаления половинной области перекрытия, где \bar{Y} , \bar{X} – размеры обрезанного фрагмента, $\bar{Y} \leq Y$, $\bar{X} \leq X$.

8. Эффективное кодирование обрезанных фрагментов. Каждый i -й видеокодер осуществляет эффективное кодирование i -го обрезанного фрагмента $\bar{M}_F(i)$ с коэффициентом $CR_C(i)$ сжатия, в результате чего формируется сжатый фрагмент $M_C(i) = \|m_C(i, l)\|_{(l=0, \overline{L_D(i)-1})}$, где $L_D(i)$ – длина (число бит) сжатого i -го фрагмента. Значения коэффициента $CR_C(i)$ сжатия и длины $L_D(i)$ сжатого i -го фрагмента связаны между собой выражением

$$L_D(i) = \frac{BD_P \cdot \left(Y \cdot X - \frac{S_o}{2} \right)}{CR_C(i)}, \quad (3)$$

где $S_o = K_o \cdot Y \cdot X$ – площадь перекрытия фрагментов; K_o – коэффициент перекрытия фрагментов.

Полный коэффициент $CR_D(i)$ сжатия i -го фрагмента многокурсового изображения с учетом координатных и идентификационных параметров реперов вычисляется с помощью выражения

$$CR_D(i) = \frac{BD_P \cdot Y \cdot X}{L_D(i) + BD_D \cdot P \cdot N(i)}, \quad (4)$$

где BD_P – битовая глубина фрагментов многокурсового изображения; BD_D – битовая глубина идентификаторов.

При равномерном распределенном эффективном кодировании фрагментов многокурсового изображения (когда площади перекрытия и коэффициенты сжатия фрагментов примерно одинаковы) выражение (4) определяет коэффициент сжатия многокурсового изображения в целом.

9. Передача сжатых фрагментов на сервер. Каждый i -й видеокодер осуществляет передачу i -го сжатого фрагмента $M_C(i)$ многокурсового изображения на сервер.

В результате выполнения данных шагов на декодер поступают сжатые обрезанные фрагменты многокурсного изображения, которые после декодирования подвергаются геометрическому преобразованию и совмещению с другими фрагментами.

Оценка эффективности распределенного кодирования многокурсных изображений с устранением межкурсной избыточности фрагментов

Оценка эффективности разработанного метода распределенного кодирования многокурсных изображений с устранением межкурсной избыточности произведена с использованием коэффициента сжатия. На рис. 5 представлены зависимости полного коэффициента $CR_D(i)$ сжатия обрезанного фрагмента многокурсного изображения от числа $N(i)$ уникальных реперов (рис. 5,а) и коэффициента K_O перекрытия (рис. 5,б) при фиксированных значениях $BD_p = 8$, $Y = 480$, $X = 640$, $BD_D = 8$, $P = 64$. Предполагается, что обрезанный фрагмент при распределенном эффективном кодировании и необрезанный фрагмент при обычном эффективном кодировании сжимаются с одинаковым коэффициентом $CR_C(i)$.

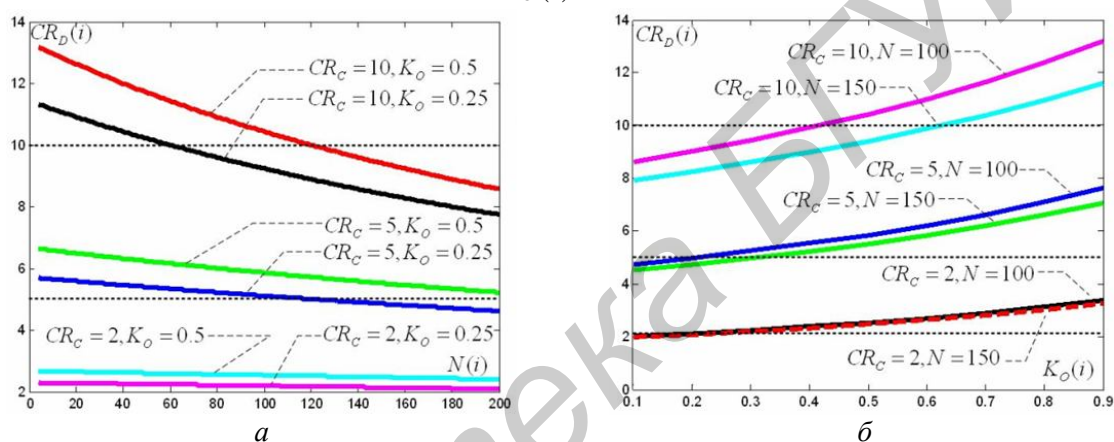


Рис. 5. Зависимости полного коэффициента $CR_D(i)$ сжатия фрагмента многокурсного изображения: а – от числа $N(i)$ уникальных реперов; б – от коэффициента K_O перекрытия

Из рис. 5 видно, что выигрыш от использования распределенного эффективного кодирования многокурсных изображений достигает 30%. С уменьшением коэффициента K_O перекрытия, увеличением коэффициента $CR_C(i)$ сжатия и числа $N(i)$ уникальных реперов эффективность распределенного кодирования снижается. Это справедливо для распределенного эффективного кодирования неподвижных и подвижных многокурсных изображений без устранения временной избыточности (например, с использованием метода MotionJPEG).

Заключение

Предложен метод распределенного эффективного кодирования многокурсных изображений, учитывающий межкурсную избыточность их перекрывающихся фрагментов, формируемых несколькими видеокameraми. Метод основан на определении границ областей соответствия фрагментов многокурсного изображения с использованием реперных точек. Метод позволяет повысить коэффициент сжатия многокурсных изображений за счет устранения межкурсной избыточности их фрагментов, уменьшить нагрузку на каналы и узлы в сети за счет исключения повторной передачи информации о пикселях соответствующих областей и снизить нагрузку на серверы обработки многокурсных изображений за счет распределенной обработки перекрывающихся изображений на уровне видеокodеров. Показано, что предложенный метод позволяет повысить коэффициент сжатия на величину до 30% по сравнению с обычным эффективным кодированием при площади перекрытия фрагментов около 50% и использовании для поиска соответствий не более 20 уникальных реперов.

EFFECTIVE CODING OF IMAGES FROM DISTRIBUTED VIDEO CAMERAS IN THE SPACE WITH ELIMINATING FRAGMENTS REDUNDANCY BETWEEN THE MULTI-VIEWPOINTS

O.J. AL-FURAJI, V.K. KANAPELKA, V.Yu. TSVIATKOU

Abstract

A method of distributed effective coding for the multi-viewpoint images is proposed. This method takes into account the redundancy between viewpoints of their overlapping fragments which are generated by multiple video-cameras. It is based on determining the region boundaries of corresponding fragments in the multi-viewpoint image using feature points. The method improves the compression ratio of the multi-viewpoint images by eliminating the re-transmission of information about corresponding regions of their overlapping fragments.

Литература

1. *Kopf J., Uyttendaele M., Deussen O. et al.* // ACM Transactions on Graphics. 2007. Vol. 26, №3. P. 1–10.
2. *D.A. Huffman* // Proceedings of the I.R.E. 1952. Vol. 40. P. 1098–1101.
3. *Witten I.H., Neal R.M., Cleary J.G.* // Communications of the ACM. 1987. Vol. 30, №6. P. 520–540.
4. *Wallace G.K.* // Communications of the ACM. 1991. Vol. 34, №4. P. 31–44.
5. *Christopoulos C., Skodras A., Ebrahimi T.* // IEEE Transactions on Consumer Electronics. 2000. Vol. 46, №4. P. 1103–1127.
6. *Shapiro J.M.* // IEEE Transactions on Signal Processing. 1993. Vol. 41, №12. P. 3445–3462.
7. *Said A., Pearlman W.A.* // IEEE Transactions on Circuits and Systems for Video Technology. 1996. Vol. 6, №3. P. 243–250.
8. *Islam A., Pearlman W.A.* // Visual Communications and Image Processing, Proceedings of SPIE. 1999. Vol. 3653. P. 294–305.
9. *Wootton C.* Focal Press: First edition. USA, 2005.
10. *Pons J., Malumbers M.P., Garcia R.* // IASTED, ACTA Press (AI'2000). 2000. P. 302–308.
11. *Le Gall D.* // Communications of the ACM. 1991. Vol. 34, №4. P. 46–58.
12. *Slepian D., Wolf J.K.* // IEEE Transactions on Information Theory. 1973. Vol. 19, №4. P. 471–480.
13. *Wyner A.D., Ziv J.* // IEEE Transactions on Information Theory. 1976. Vol. 22, №1. P. 1–10.
14. *Lowe D.G.* // International Journal of Computer Vision. 2004. Vol. 60, №2. P. 91–110.
15. *Bay H., Ess A., Tuytelaars T. et al.* // Computer Vision and Image Understanding. 2008. Vol. 110, №3. P. 346–359.
16. *Gonzales R.C., Woods R.E.* Prentice-Hall: Second edition, 2002.
17. *Neubeck A., Gool L.V.* // ICPR '06 Proceedings of the 18th International Conference on Pattern Recognition. 2006. Vol. 3. P. 850–855.
18. *Brown M., Lowe D.* // In British Machine Vision Conference, Cardiff, Wales, 2002. P. 656–665.
19. *Fischler M.A., Bolles R.C.* // Communications of the ACM. 1981. Vol. 24, №6. P. 381–395.
20. *Brown M., Lowe D.G.* // IJCV. 2007. Vol. 74, №1. P. 59–73.