

УДК 621.391

## ОЦЕНКА ЭФФЕКТИВНОСТИ МЕТОДОВ СЖАТИЯ ДЛЯ КОДИРОВАНИЯ МНОГОРАКУРСНЫХ ИЗОБРАЖЕНИЙ С ПОДВИЖНЫХ ОБЪЕКТОВ

В.Ю. ЦВЕТКОВ

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

Поступила в редакцию 1 апреля 2014

Произведена оценка эффективности методов сжатия для кодирования многокурсовых изображений, формируемых камерами, установленными на подвижных объектах.

*Ключевые слова:* сжатие изображений, кодирование многокурсовых изображений.

### Введение

Развитие систем дистанционного зондирования Земли, пассивной оптической локации и мобильного видеонаблюдения привело к формированию проблемы эффективной передачи видеоданных с подвижных объектов. Для таких систем характерно использование радиоканала с ограниченной пропускной способностью, передача многокурсовых изображений от одной или нескольких перемещающихся камер и снижение временной избыточности видеоданных из-за параллакса, проявляющегося на изображениях при перемещении камеры. Таким образом, актуальной задачей является оценка эффективности методов сжатия в рассмотренных условиях. Для этого необходимо сопоставить характеристики многокурсовых изображений, радиоканалов и кодеков сжатия в условиях передачи многокурсовых изображений с подвижных объектов.

### Характеристики многокурсовых изображений с подвижных объектов

Возможны следующие способы формирования многокурсовых изображений с подвижных объектов (рис. 1): перекрывающиеся кадры одной видеопоследовательности, сформированные в различные моменты времени одной перемещающейся камерой; перекрывающиеся кадры нескольких видеопоследовательностей, сформированные в различные моменты времени одной или несколькими циклически перемещающимися камерами; перекрывающиеся кадры нескольких видеопоследовательностей, сформированные в один момент времени несколькими камерами (неподвижными или перемещающимися, установленными на одном или нескольких носителях).

Видеопоследовательность, формируемая перемещающейся камерой, может быть разделена на перекрывающиеся фрагменты, кадры которых представляют собой разнокурсовые изображения одних и тех же объектов (3D-изображения). В результате обработки многокурсового изображения формируются плоские, круговые, сферические, секторные панорамы, включающие преобразованные разнокурсовые фрагменты многокурсового изображения. По способу формирования панорамные изображения могут быть отнесены к классу многокурсовых изображений, хотя и не являются таковыми по способу представления (хранения) и воспроизведения.

Видеокамеры, используемые в системах видеонаблюдения, формируют цветные или полутоновые кадры с размерами от  $320 \times 180$  до  $2048 \times 1536$  пикселей с частотой  $F_K = [1, 120]$  кадр/с. Тепловизоры формируют полутоновые кадры с размерами  $640 \times 480$ ,  $320 \times 240$ ,  $160 \times 120$  пикселей с частотой  $F_K = [1, 120]$  кадр/с.

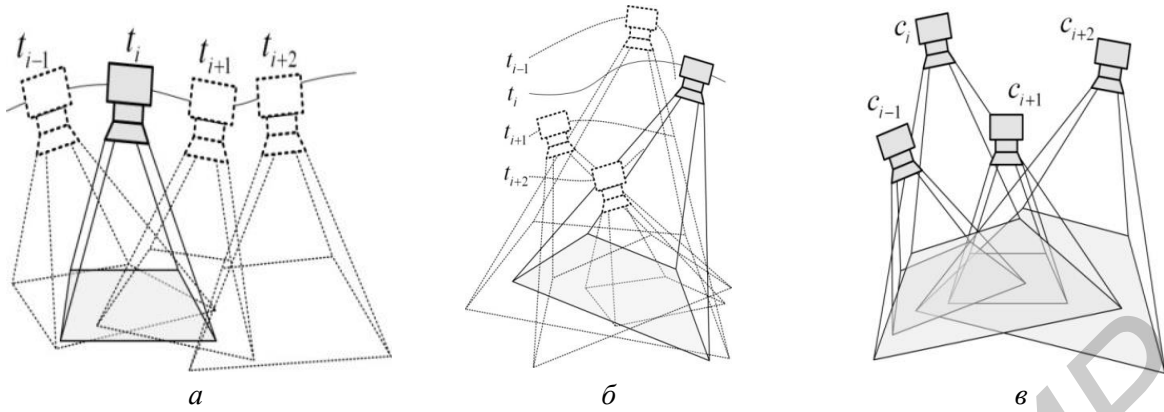


Рис. 1. Способы формирования многокурсных изображений:

*a* – от одной перемещающейся камеры; *б* – от одной или нескольких циклически перемещающихся камер; *в* – в один момент времени от нескольких камер

Скорость  $B_K$  формирования 3D изображений (в бит/с) от одной перемещающейся видеокамеры определяется с помощью выражения

$$B_K = V_K F_K D_B, \quad (1)$$

где  $V_K = Y_K X_K N_C$  – объем видеоданных (в байтах), приходящихся на один кадр размером  $Y_K \times X_K$  пикселей;  $N_C$  – число каналов ( $N_C = 1$  для полутонового кадра,  $N_C = 3$  для цветного кадра);  $D_B$  – битовая глубина (число бит на пиксель) кадра.

Для определения скорости  $B_K$  формирования стереоизображений необходимо удвоить соответствующее значение для одной камеры.

Скорость  $B_{MVP}$  формирования плоской панорамы определяется с помощью выражения

$$B_{MVP} = V_{MVP} F_K D_B, \quad (2)$$

где  $V_{MVP} = Y_K X_{MVP} N_C$  – объем видеоданных, приходящейся на один фрагмент плоской панорамы размером  $Y_K \times X_{MVP}$  пикселей;  $X_{MVP} = X_K v_X / (S_X F_K)$  – размер фрагмента плоской панорамы по горизонтали;  $S_X = 2L_X \operatorname{tg}(\gamma_X/2)$  – размер (в метрах) области видимости на расстоянии  $L_X$  от камеры;  $v_X$  – скорость (в м/с) носителя камеры;  $\gamma_X$  – угол (в градусах) обзора камеры по горизонтали. Скорость  $B_{CP}$  формирования круговой панорамы с помощью одной камеры определяется с помощью выражения

$$B_{CP} = V_{CP} F_K D_B, \quad (3)$$

где  $V_{CP} = Y_K X_{CP} N_C$  – объем видеоданных, приходящихся на одну круговую панораму размером  $Y_K \times X_{CP}$  пикселей;  $X_{CP} = 360X_K/\gamma_X$  – размер круговой панорамы по горизонтали;  $F_{CP} = \omega_K/360$  – частота (панорам/с) формирования кругового панорамного видеоизображения;  $\omega_K$  – угловая скорость (град/с) поворота камеры.

При использовании двух однотипных камер, установленных на одном или двух объектах-носителях, для формирования стереоскопического видеоизображения, круговой стереопанорамы, плоской стереопанорамы для расчета скоростей формирования изображений принимаются соответствующие удвоенные значения, установленные для одной камеры. Необходимым условием формирования панорам является перекрытие изображений на 25 % и более. Поэтому для вычисления скоростей формирования панорамного изображения от двух однотипных камер принимаются соответствующие значения, установленные для одной камеры, умноженные на 1,75. Для двух разнотипных камер (видеокамера и тепловизор), установленных на одном объекте-носителе, скорости формирования многокурсных изображений вычисляются с учетом того, что размер и частота кадров определяются минимальными значениями для двух камер.

При отдельной передаче видео от  $N$  однотипных камер, установленных на одном или нескольких объектах-носителях, формируется 3D-видеоизображение. Для определения его размеров  $Y_{NK} \times X_{NK}$ , объема  $V_{NK}$  и скорости  $B_{NK}$  передачи принимаются соответствующие значения, установленные для одной камеры, умноженные на  $N$ . В случае использования одного носителя для полного кругового обзора периметра необходимое число  $N$  устанавливаемых на него однотипных камер вычисляется с помощью выражения  $N = 360 / (0,75\gamma_X)$ , где  $0,75$  – коэффициент, учитывающий перекрытие изображений по горизонтали. Для сферического обзора пространства необходимое число  $N$  устанавливаемых на одном носителе однотипных камер вычисляется с помощью выражения  $N = 4 \cdot 180 / (180 \cdot 0,56 \sin^2(\gamma_Y/2)) = 4 / (0,56 \sin^2(\gamma_Y/2))$ , где  $\gamma_Y$  – угол обзора камеры по вертикали;  $0,56$  – коэффициент, учитывающий перекрытие изображений по вертикали.

Скорость  $B_{NK}$  передачи круговой видеопанорамы вычисляется с помощью выражения

$$B_{NK} = V_{NK} F_K D_B, \quad (4)$$

где  $V_{NK} = Y_{NK} X_{NK} N_C$ ;  $X_{NK} = 360 X_K / \gamma_X$ ;  $Y_{NK} = Y_F$ .

Для сферической видеопанорамы  $N = 4 \cdot 180 / (180 \sin^2(\gamma_Y/2)) = 4 / \sin^2(\gamma_Y/2)$ ;  $Y_{NK} = N Y_F$ ;  $X_{NK} = N X_K$ .

Расчеты по выражениям (1)–(4) показывают, что скорость формирования многокамерных изображений от нескольких десятков видеокамер или тепловизоров при минимальном качестве составляет около 100 Мбит/с (рис. 2). С увеличением числа камер и качества многокамерных изображений скорость их формирования быстро растет.

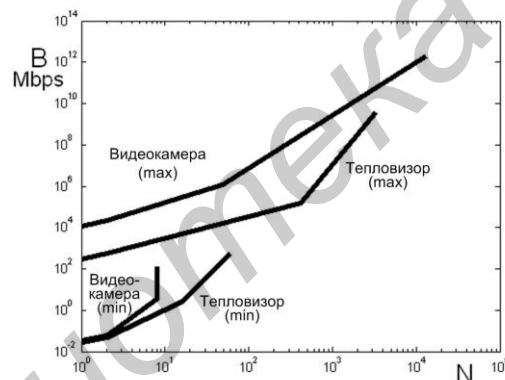


Рис. 2. Зависимости минимальной (min) и максимальной (max) скоростей формирования многокамерных изображений от числа камер

Требования к качеству формирования и передачи многокамерных изображений с подвижных объектов определяются задачами их обработки, в число которых входят обнаружение, идентификация, сопровождение, определение координат местоположения цели, трехмерная реконструкция, панорамирование.

### Характеристики радиоканалов передачи видеоданных с подвижных объектов

Единственным в своем роде стандартом, устанавливающим требования к радиоканалам передачи видеоданных с подвижных объектов, является STANAG 4609, разработанный НАТО в 2005 году и постоянно модифицируемый [1]. Стандартом предусмотрена передача видеоданных по радиоканалам с подвижных объектов на скоростях от 32 Кбит/с до 1485 Мбит/с. С учетом используемого частотного диапазона максимальная скорость передачи, доступная для носимых и возимых систем, составляет до 0,5 Мбит/с; для самолетов – до 360 Мбит/с; для космических аппаратов до 1500 Мбит/с.

### Характеристики кодеков сжатия видеоданных

Радиоканал имеет жесткие ограничения по скорости передачи и дальности, обусловленные массогабаритными и энергетическими характеристиками радиоаппаратуры.

Для передачи многоракурсных изображений с подвижных объектов необходимо согласовать скорость их формирования с пропускной способностью радиоканала. Это осуществляется с помощью сжатия.

Коэффициент  $CR$  сжатия рассчитывается с помощью выражения

$$CR = B_f / BR, \quad (5)$$

где  $BR$  – пропускная способность радиоканала (с учетом помехоустойчивого кодирования);  $B_f$  – скорость формирования многоракурсных изображений.

Требуемый коэффициент сжатия для передачи многоракурсных изображений по радиоканалам может достигать  $20 \times 10^{10}$  раз. Однако реально возможный коэффициент сжатия значительно ниже. Во-первых, максимально возможный коэффициент сжатия определяется используемым кодеком. Во-вторых, метод и максимально допустимый коэффициент сжатия ограничены задачами обработки многоракурсных изображений, восстановленных после сжатия. Кроме того, пропускная способность радиоканала ограничена условиями его применения.

Для сжатия видеоданных с подвижных объектов используются, как правило, два кодека: MPEG-2 и H.264. Для них установлены максимальные коэффициенты сжатия 165 и 430 раз соответственно при частоте 25 кадров/с и кодировании RGB 8 бит/пиксель [1]. Эти значения близки к максимально возможным коэффициентам сжатия для данных кодеков и соответствуют значению  $PSNR$  пикового отношения сигнал-шум на уровне 20–25 дБ [2].

Значение  $PSNR$  определяется с помощью выражения

$$PSNR = 10 \log_{10} \left( (2^{B_p} - 1)^2 / MSE \right), \quad (6)$$

где  $MSE = \frac{1}{n} \sum_{i=0}^{n-1} (X_i - \tilde{X}_i)^2$  – среднеквадратическая ошибка;  $X_i, \tilde{X}_i$  – исходное и восстановленное значения  $i$ -го пикселя;  $n$  – число пикселей в изображении.

В системах стационарного видеонаблюдения используются также MJPEG, MJPEG2000 и MPEG-4 [3]. Максимально возможные коэффициенты сжатия для этих кодеков при  $PSNR$  на уровне 20 дБ составляют около 50, 150 и 300 раз соответственно [2]. При  $PSNR$  на уровне 25 дБ коэффициенты сжатия для данных кодеков составляют около 30, 100 и 140 раз.

Коэффициенты сжатия при  $PSNR$  на уровне 25 дБ являются предельно допустимыми для задач обработки многоракурсных изображений, не критичных к потере мелких деталей. В некоторых условиях к таким задачам могут быть отнесены обнаружение, сопровождение и определение координат местоположения цели. Задача обнаружения цели допускает использование кодеков любого типа (табл.). Другие задачи обработки многоракурсных изображений требуют более высокого качества и, соответственно, значительно меньших коэффициентов сжатия.

**Скорости передачи и максимально допустимые коэффициенты сжатия изображений**

Задачи обработки многоракурсных изображений	PSNR, дБ	Коэффициент сжатия для различных кодеков				
		MJPEG	MJPEG 2000	MPEG-2	MPEG-4	H.264
Обнаружение цели	25	30	100	85	140	200
Сопровождение цели						
Определение координат цели						
Идентификация	30	10	25	–	–	–
Сопровождение цели						
Определение координат цели						
3D реконструкция	30	10	25	27	35	55
Панорамирование						
Сопровождение цели						
Определение координат цели						

Для задач идентификации цели в системах стационарного видеонаблюдения (распознавание лиц, номеров автомобилей) используется несжатое видео или видео с поккадровым сжатием в форматах MJPEG и MJPEG2000 при  $PSNR$  около 30 дБ [4], что соответствует

максимально допустимым коэффициентам сжатия до 10 раз для MJPEG и до 25 раз для MJPEG2000.

Для задач 3D-реконструкции и панорамирования могут использоваться изображения, сжатые с помощью кодека JPEG до 10 раз. В случае использования JPEG2000 максимально допустимый коэффициент сжатия при PSNR = 30 дБ возрастает до 25 раз, в случае MPEG-4 – до 35 раз, а в случае H.264 – до 55 раз [2]. Видеокодек MPEG-2 обеспечивает примерно в 2 раза меньший коэффициент сжатия по сравнению с H.264, т.е. примерно 27 раз для PSNR на уровне 30 дБ. Аналогичные максимально допустимые коэффициенты сжатия могут быть приняты для задач сопровождения и определения координат местоположения цели. Задача сопровождения при размерах цели порядка 50 пикселей сводится к задаче идентификации (допускает сжатие не более чем в 7 или 20 раз в зависимости от кодека MJPEG или MJPEG2000), а при размерах цели порядка 4×4 пикселей сводится к задаче обнаружения (допускает сжатие в десятки – сотни раз).

### Сопоставление характеристик многокадровых изображений, радиоканалов и кодеков

На рис. 3 представлены зависимости требуемой пропускной способности радиоканала от коэффициента сжатия для некоторых типов многокадровых изображений, сформированных видеокамерой в различных условиях с учетом рассмотренных ограничений (см. табл.), где S – системы космического базирования; A – системы авиационного базирования; H/M – носимые и возимые системы.

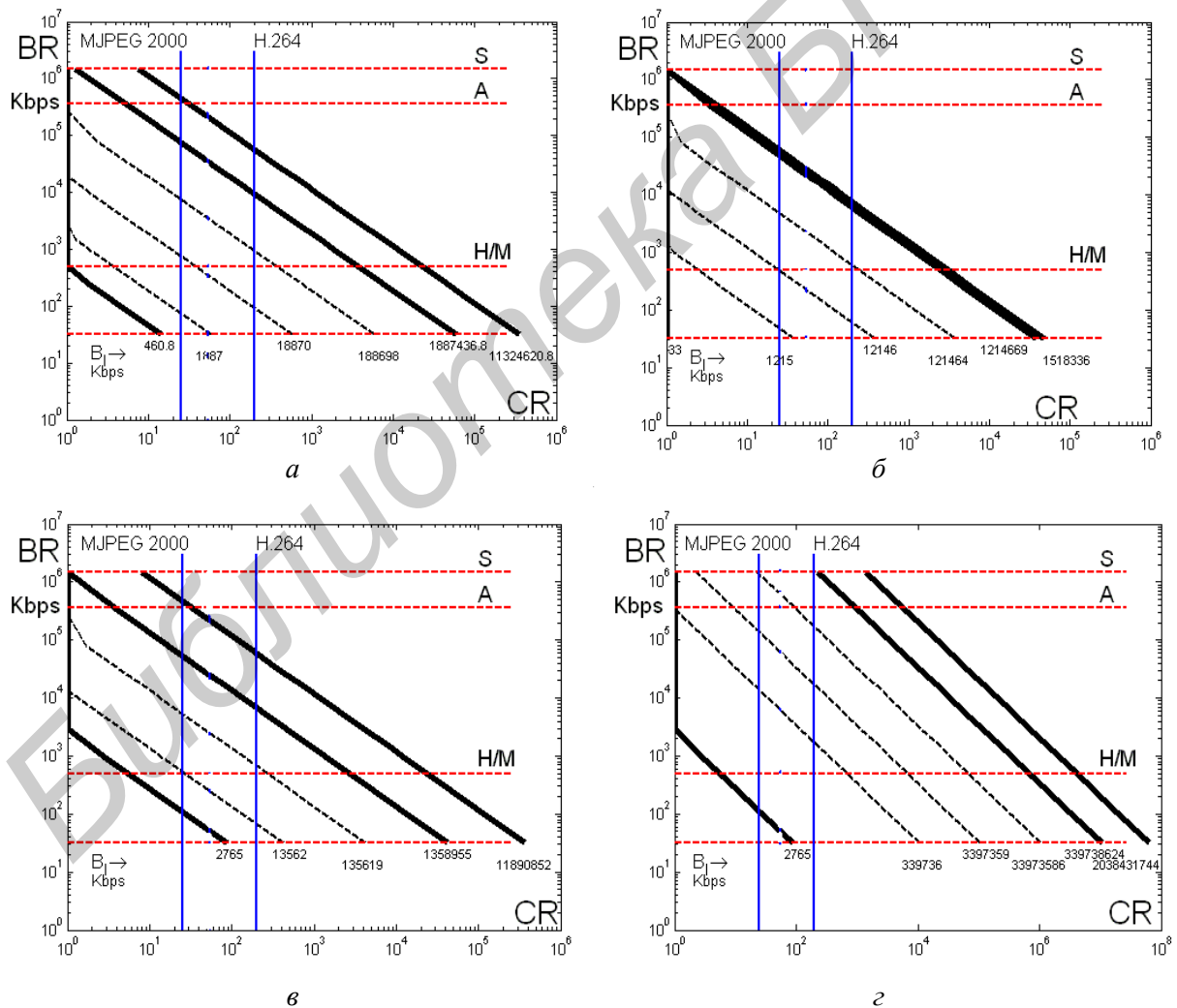


Рис. 3. Зависимости требуемой пропускной способности радиоканала от коэффициента сжатия для различных многокадровых изображений: а – 3D изображения; б – плоская панорамы; в – круговая панорамы; г – круговая видео-панорамы

Анализ данных зависимостей показывает следующее. В носимых и возимых системах формирования и передачи многокурсных изображений возможна передача 3D-изображений и плоских панорам в среднем разрешении с использованием кодеков любого типа. Круговые панорамы и видеопанорамы могут передаваться только в низком разрешении при высоких коэффициентах сжатия (5–25 раз для круговых панорам и кодека MJPEG2000; 8–100 раз для круговых видеопанорам и кодека H.264). В системах формирования и передачи многокурсных изображений авиационного и космического базирования передача 3D-изображений, плоских и круговых панорам в высоком разрешении возможна с использованием кодеков любого типа. Круговые видеопанорамы могут передаваться в среднем разрешении при использовании кодека H.264 и сжатии в 25–200 раз или в низком разрешении при использовании кодека MJPEG2000 и сжатии до 25 раз.

### **Заключение**

Ограничения на пропускную способность радиоканалов и максимально допустимый коэффициент сжатия в системах формирования и передачи многокурсных изображений с подвижных объектов являются основными причинами, препятствующими использованию камер высокого разрешения и решению задач обработки многокурсных изображений, связанных с 3D реконструкцией, панорамированием, идентификацией, сопровождением и определением координат местоположения цели. Необходимым условием расширения области применения камер высокого разрешения и задач обработки многокурсных изображений является увеличение максимально допустимого коэффициента сжатия, что требует разработки новых методов кодирования многокурсных изображений, учитывающих специфику их формирования и передачи.

## **PERFORMANCE EVALUATION OF COMPRESSION METHODS FOR MULTI-ANGLE IMAGES FROM MOVING OBJECTS**

V.Yu. TSVIATKOU

### **Abstract**

Evaluating the effectiveness of compression methods for encoding multi-angle images that are formed cameras mounted on moving objects is produced.

### **Список литературы**

1. STANAG 4609 NATO Digital Motion Imagery Standard // NATO Standardization Agency, Brussels, Belgium: 1<sup>st</sup> edition, 2005.
2. MSU Video Codecs Comparison. – [Электронный ресурс] – Режим доступа: [http://compression.ru/video/codec\\_comparison/codec\\_comparison.html](http://compression.ru/video/codec_comparison/codec_comparison.html). – Дата доступа: 16.03.2013.
3. Системы безопасности. 2010. № 4.
4. Аппаратная видеоаналитика для охраны стратегических объектов. – [Электронный ресурс] – Режим доступа: <http://synesis.ru/blog/article/ptitsyn-apparatnaya-videoanalitika-dlya-oxranyi-strategicheskikh-obektov>. – Дата доступа: 16.03.2013.