

УДК 004.627

СЕМАНТИЧЕСКОЕ СЖАТИЕ ВИДЕОИНФОРМАЦИИ В СИСТЕМАХ ВИДЕОНАБЛЮДЕНИЯ

С. А. Кузьмин

Санкт-Петербургский государственный университет аэрокосмического приборостроения, факультет радиотехники, электроники и связи, кафедра электронных и телевизионных систем

Получена 9 апреля 2012 г.

Аннотация. Показано, что нецелесообразно сжатие видеопотока путем значительного снижения его параметров. Поэтому далее рассматривается сжатие на основе использования особенностей сигнала. Описаны алгоритмы выделения объектов и семантического сжатия видеопотока. Особенностью алгоритма сжатия является дополнение видеопотока информацией о параметрах объекта. В ходе эксперимента проведено исследование коэффициента сжатия предложенного алгоритма.

Ключевые слова: выделение объекта, параметры объекта, сжатие видеоинформации, JPEG, метайнформация.

Abstract. In the first part of article the possibilities of video compression by downscaling of parameters (bit depth, resolution) are studied and rejected. In the second part of the article the approach to compression, which utilizes saliencies of signal, is studied. Algorithms of object detection and semantic compression are described. The key feature of compression algorithm is supplement of video stream by information about parameters of objects. Final part consists of experimental study of compression ratio of proposed algorithm.

Keywords: object detection, parameters of object, video compression, JPEG, metadata.

Введение

Рассматривается задача сжатия видеопотоков, передаваемых по каналу связи от камер наблюдения к конечному потребителю. При этом сжатие не должно сильно ухудшать качество видеoinформации, чтобы наблюдатель мог принимать верные решения.

Пример расчета скорости ввода информации для случая передачи несжатых видеопоследовательностей кадров размером 576x704 пикселей (формат 4 CIF) с частотой 25 кадров в секунду при условии, что используются 3 канала изображения, квантованные на 256 градаций яркости каждый, от одной камеры:

$$D=K*X*Y*C*R*f=1*576*704*3*1*25=29700 \text{ КБ/с} = 29 \text{ Мб/с},$$

где K – количество камер, X – количество пикселей в строке, Y – количество строк, C – количество каналов цвета, R – количество уровней квантования ($256=2^8=1$ байт) на канал цвета, f – кадров/с.

Время передачи такого потока данных с помощью сетевой карты Ethernet (пропускная способность $V=10$ Мбит/с= $1,25$ Мб/с) $T = D/V = 23,2$ сек, т.е. для приема и просмотра видеопотоков от одной видеокамеры в режиме реального времени требуется сжатие примерно в 23,2 раза. Таким образом, требуется сжатие для согласования пропускной способности канала связи с объемом информации источника сообщения в случае необходимости получения видеопотоков в реальном времени.

Самый простой способ сжатия – уменьшать какие-то из перемножаемых показателей (X , Y , C , R , f). Была исследована возможность уменьшения разрешения (X , Y) и количества уровней квантования (R). Критерием качества переданного изображения выступало значение пикового отношения сигнал-шум PSNR. Методика эксперимента: уменьшить разрешение децимацией строк, увеличить изображение до исходных размеров билинейной интерполяцией, уменьшить количество уровней квантования, вычислить PSNR путем сравнения с исходным изображением. Отсортированные по убыванию результаты приведены в таблице 1. Обычным шрифтом отмечены случаи, когда ухудшение

качества не приводило к невозможности различать контуры объектов. Курсивом, отмечены случаи, когда ухудшение качества искажало форму объекта. Потребитель должен иметь хорошее качество изображения, позволяющее идентифицировать объекты, что достигается при $PSNR \geq 31$ дБ [1]. Эти случаи отмечены жирным шрифтом. Формат записи (1:4) в столбце «Уменьшение площади» показывает знаменатели по осям в процедуре децимации – количество сохраняемых столбцов и строк (X/1, Y/4).

Таблица 1. Изменение показателя качества изображения при уменьшении разрешения и снижении количества уровней квантования.

Уменьшение площади, раз / Количество бит на канал	PSNR при R=8	PSNR при R=7	PSNR при R=6	PSNR при R=5	PSNR при R=4	PSNR при R=3	PSNR при R=2	PSNR при R=1
1 (исходное)	∞	∞	43,36	35,83	29,32	23,13	17,00	10,63
2 (1:2)	34,32	34,32	33,66	31,90	28,00	22,73	16,88	10,58
4 (2:2)	31,41	31,31	30,97	29,97	27,02	22,39	16,76	10,55
3 (1:3)	30,97	30,89	30,65	29,55	26,89	22,39	16,79	10,64
6 (2:3)	29,38	29,32	29,10	28,35	26,17	22,08	16,68	10,54
4 (1:4)	28,73	28,68	28,49	27,83	25,87	21,99	16,65	10,53
9 (3:3)	27,71	27,67	27,52	26,99	25,27	21,71	16,57	10,52
8 (2:4)	27,71	27,67	27,52	26,99	25,29	21,72	16,55	10,50
12 (3:4)	26,54	26,51	26,39	25,98	24,57	21,39	16,45	10,48
16 (4:4)	25,85	25,82	25,72	25,36	24,13	21,17	16,41	10,46

Таким образом, при требовании хорошего качества изображения максимальный коэффициент сжатия k при применении снижения этих параметров видеопотока, составляет $k(2:2, R=7)=4*8/7=4,57$ раз. Исследования уменьшения количества каналов (децимации цветоразностных сигналов по форматам 4:1:1 и т.п.) не проводилось, т.к. в этом случае потребовалось бы для каждого случая рассчитывать таблицы, аналогичные 1. Даже упомянутая схема 4:1:1 (8 пикселей яркостного канала и по 1 пикселю от двух цветоразностных каналов вместо $24=2(\text{строки}) \times 4(\text{пикселя}) \times 3(\text{канала})$) дает дополнительное

сжатие потока лишь в 2,4 раза, то есть коэффициент сжатия не превысил бы $k(2:2, R=7, 4:1:1)=4,57*24/10=10,97$ раз. Таким образом, снижение параметров видеопотока не дает требуемого коэффициента сжатия, так как постоянное снижение частоты кадров нежелательно, а способы организации переменной частоты кадров требуют усложнения передачи видеoinформации (реализации буфера кадров).

Поэтому необходимо исследовать возможность сжатия кадров, связанную с характером видеопоследовательности.

Семантический анализ видеопоследовательностей

Передачу информации по каналам связи следует также рассматривать не только как способ документирования наблюдаемой сцены, но и как способ организации реакции на события в ней. Чем быстрее дойдет информация до наблюдателя или автоматического устройства принятия решения, тем быстрее будет и реакция.

Для удобства принятия решения и поиска в архивах важно сформировать семантическое описание сцены и дополнить видеоряд метаинформацией об объектах и событиях. Оператор в пункте наблюдения должен иметь возможность нажать на интересующий объект на экране и сразу же получить измеренные характеристики и все кадры, относящиеся к объекту, с помощью сформированного инвертированного индекса. Инвертированным индексом назовем множество номеров кадров, в которых появлялся интересующий объект. Эта термин взят из алгоритмов работы поисковых систем в Интернет. Инвертированный индекс строится на приемной стороне на основе метаинформации (идентификаторов объектов и других параметров), переданной вместе в кадрами. На основании видеопотока и метаинформации можно создать изображение дополненной реальности – показ векторов скоростей перед объектами, показ траекторий движения сзади объектов и т.п. Таким образом, необходимо разработать способ семантического анализа и способ сжатия видеoinформации и метаинформации в один поток данных.

Существующие алгоритмы сжатия видеоинформации либо не обеспечивают метаданные для объектов (Motion JPEG, Dirac, VP3-Theora, VP8-WebM, RealVideo), либо являются кодеками с низким количеством ключевых кадров (серия стандартов MPEG), что ведет к задержке при кодировании в реальном времени (чем больше длина группы кадров, тем больше задержка между моментом получения первого и последнего кадров в группе $T_{\text{CONDENSATION}}$). Например, камера установлена таким образом, чтобы транспортное средство на очень высокой скорости проходило через её поле зрения не менее, чем за 10 кадров. При количестве кадров в группе больше 10, автомобиль при определенных обстоятельствах может пройти за время $T_{\text{TRANSPASSING}}$ через все поле зрения раньше, чем будет завершено накопление группы кадров $T_{\text{CONDENSATION}}$, сжатие T_{ENCODING} , передача группы кадров $T_{\text{TRANSMISSING}}$, декодирование T_{DECODING} и организация реакции T_{REACTION} , что нежелательно в системах видеонаблюдения:

$$T_{\text{TRANSPASSING}} < T_{\text{CONDENSATION}} + T_{\text{ENCODING}} + T_{\text{TRANSMISSING}} + T_{\text{DECODING}} + T_{\text{REACTION}}.$$

Рассмотрим случай наблюдения потока транспортных средств, движущихся по шоссе в поле зрения камеры. Информацию от видеокамер наблюдения и извлеченную из видеопотоков семантику необходимо передать в районный центр видеомониторинга.

Извлечение семантики подразумевает анализ видеок кадров с целью выделения областей разных классов. В случае разнообразия внешнего вида и разброса значений признаков объектов интереса, не позволяющего создать однозначные модели вида искомым транспортным средствам, мы сталкиваемся с ситуацией априорной параметрической неопределенности. В условиях априорной параметрической неопределенности относительно яркости искомым объектам нельзя по одному кадру выделить все объекты при приемлемом уровне ошибок классификации, т.к. области одинаковой семантики находятся в разных яркостных диапазонах (светлые и темные объекты, светлый и темный фон, тени транспортных средств и расположенных около улицы деревьев, зданий, оград, столбов).

Отделить объекты от фона возможно, если известны модели объектов или модель фона, но в рассматриваемом случае априорной параметрической неопределенности относительно яркости искомым объектам нельзя задать их яркостные эталонные описания, в то время как фоновая составляющая сцены может быть оценена [2]. На основании кадров видеосъемки наблюдаемого участка сцены строится попиксельная модель фона. Исследования зависимости качества сжатого видеопотока изображения от алгоритма оценки фона приведены в третьей части статьи. Оценка фона возможна, т.к. камера является неподвижной относительно наблюдаемой сцены, что верно для других систем мониторинга и охраны. Если бы камера двигалась, то её движение можно алгоритмически скомпенсировать в случае небольшой скорости перемещения объектива [3].

На основании сравнения кадра с моделью фона (статической составляющей видеопотока) возможно выделение динамической составляющей – движущихся объектов и теней. Тени транспортных средств могут быть обнаружены по уменьшению яркости участков текущего кадра относительно яркости оценки фона. Обнаруженные участки теней могут быть использованы для коррекции границ (уменьшения размеров) выделенных областей с объектами на изображении динамической составляющей. Это приводит как к повышению коэффициента сжатия, так и улучшает визуально воспринимаемое качество изображения, а также повышает точность классификации объектов.

При применении метода вычитания оценки фона из кадра возникает проблема раздробленности частей «движущихся областей» объектов. Для решения этой проблемы предложено дополнять выделенные области информацией об их внутренних связях (переходах от одних частей объекта к другим) и границах (внешних контурах). Эта дополнительная информация – разница контурных препаратов кадра и оценки фона – названа «движущиеся контуры». Блок-схема алгоритма показана на рис.1

После объединения препаратов «движущихся областей» и «движущихся контуров» с помощью логической операции «ИЛИ» следует этап

постобработки, на котором повышается связность частей объектов и понижается уровень шума. Результирующее бинарное изображение представляет собой динамическую составляющую сцены в текущем кадре – движущиеся объекты и тени[4].

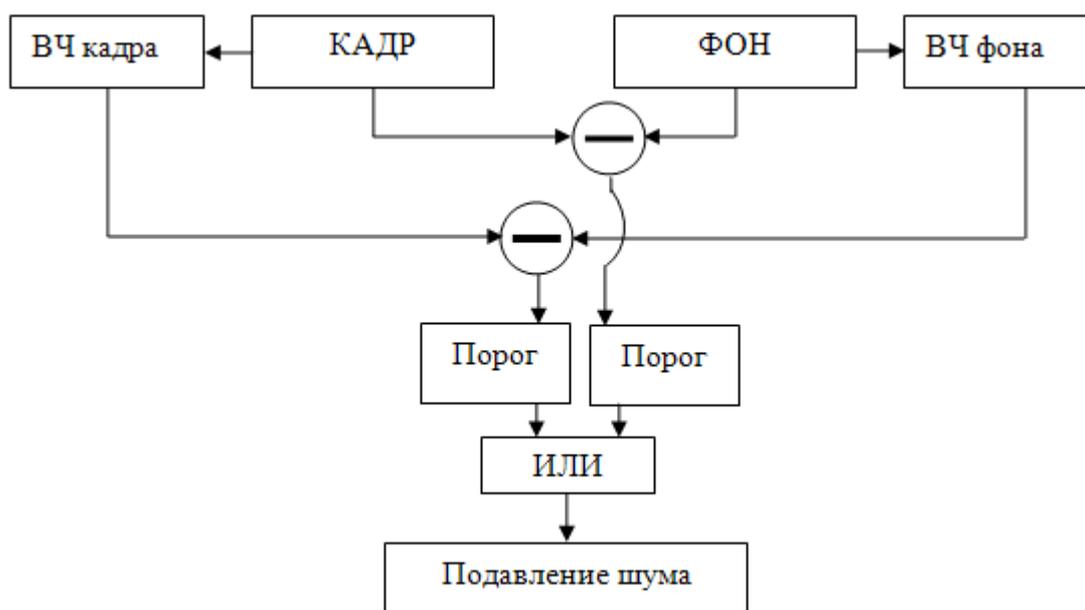


Рис. 1. Обнаружение объектов на основе информации о движущихся областях и движущихся контурах.

Затем бинарное изображение сегментируется (присвоение номеров пикселям по их принадлежности к областям), области описываются с помощью набора признаков движения, позиции и формы. Вычисленные значения параметров областей сравниваются с допустимыми значениями для интересующих объектов в сцене (превышение минимальной площади, нахождение объекта на дороге, отношение длины и ширины объекта превышает минимальную величину). Прошедшие отбор области считаются сопровождаемыми объектами. На основании сопоставления наборов описаний объектов в соседних кадрах осуществляется обнаружение прихода новых и ухода старых объектов, сопровождение известных объектов. В результате формируется *семантическое описание сцены* в виде массива, описывающего результаты сегментации, и таблицы, описывающей множество обнаруженных

объектов и связанных с ними событий. Таблица будет передана построчно вместе с областями, а массив будет передан в виде изображения альфа-канала.

Семантическое сжатие видеoinформации

Возможность сжатия связана с характером видеопоследовательности: есть быстро меняющиеся положение автомобили и есть практически неизменная на коротком временном отрезке фоновая составляющая сцены, представляемая оценкой фона. Оценка фона передается на приемную сторону и поскольку после этого часть информации кадра уже находится на приемной стороне, то становится возможным сжатие путем передачи только динамической составляющей видеопотока – движущихся объектов (рис. 2). Для каждого передаваемого кадра происходит вкрапление движущихся объектов в изображение оценки фона. Информация о движущихся объектах должна передаваться очень часто, в то время как оценку фона можно обновлять время от времени (например, раз в 10 минут).

Этот принцип описан в работах [2, 4]. Сжатие называется семантическим, т.к. используется тот факт, что нас интересуют лишь движущиеся области изображения и необходимо передавать только их. При этом возможен синтез изображения, близкого к исходному изображению от видеокамеры, при использовании матирования изображения с движущимися областями из кадра, поступающего от камеры, и изображения оценки фона. Под матированием [5] понимается способ объединения информации из нескольких изображений с использованием альфа-канала. При этом битовая плоскость, представляющая альфа-канал, является своеобразным триггером, пропускающим яркость пикселей одного из объединяемых изображений в итоговое изображение и непропускающим яркость второго изображения. В качестве альфа-канала выступает изображение с движущимися объектами.

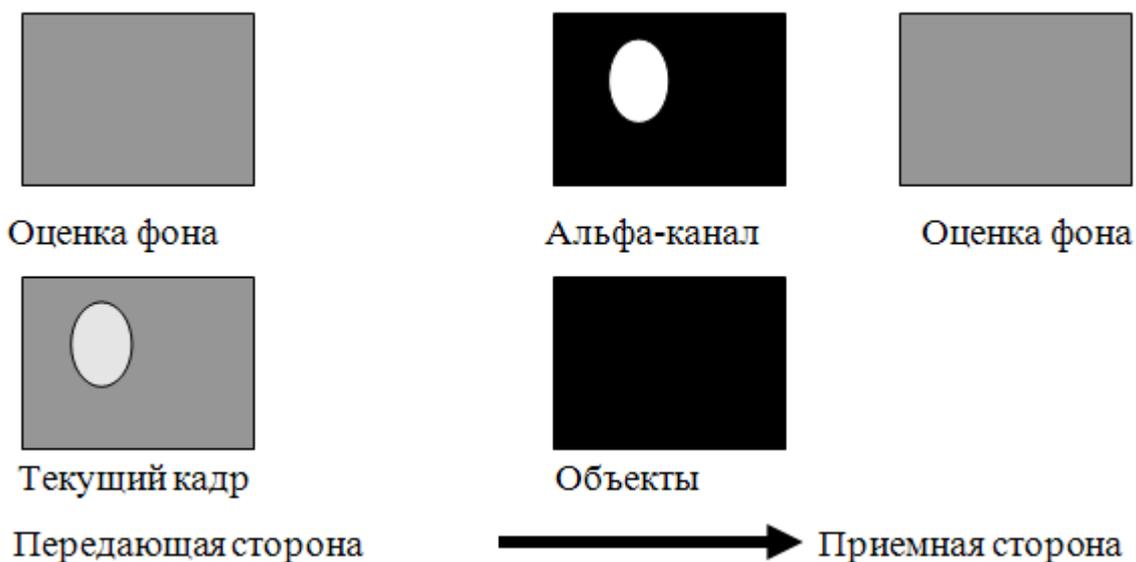


Рис. 2. Типичный цикл передачи информации при семантическом сжатии состоит в передаче информации об областях с движущимися объектами.

Пример операции матирования показан на рис. 3. Формула для матирования:

$$I'[x,y] = I[x,y] * \alpha[x,y] + B[x,y] * (1 - \alpha[x,y]),$$

где $I'[x,y]$ – яркость пикселя с координатами $[x,y]$ многобитного изображения, синтезированного на основе двух кадров и альфа-канала;

$\alpha[x,y]$ – яркость пикселя с координатами $[x,y]$ однобитного (бинарного) изображения альфа-канала,

$I[x,y]$ – яркость пикселя с координатами $[x,y]$ многобитного изображения, поступающего от камеры;

$B[x,y]$ – яркость пикселя с координатами $[x,y]$ многобитного изображения оценки фона.

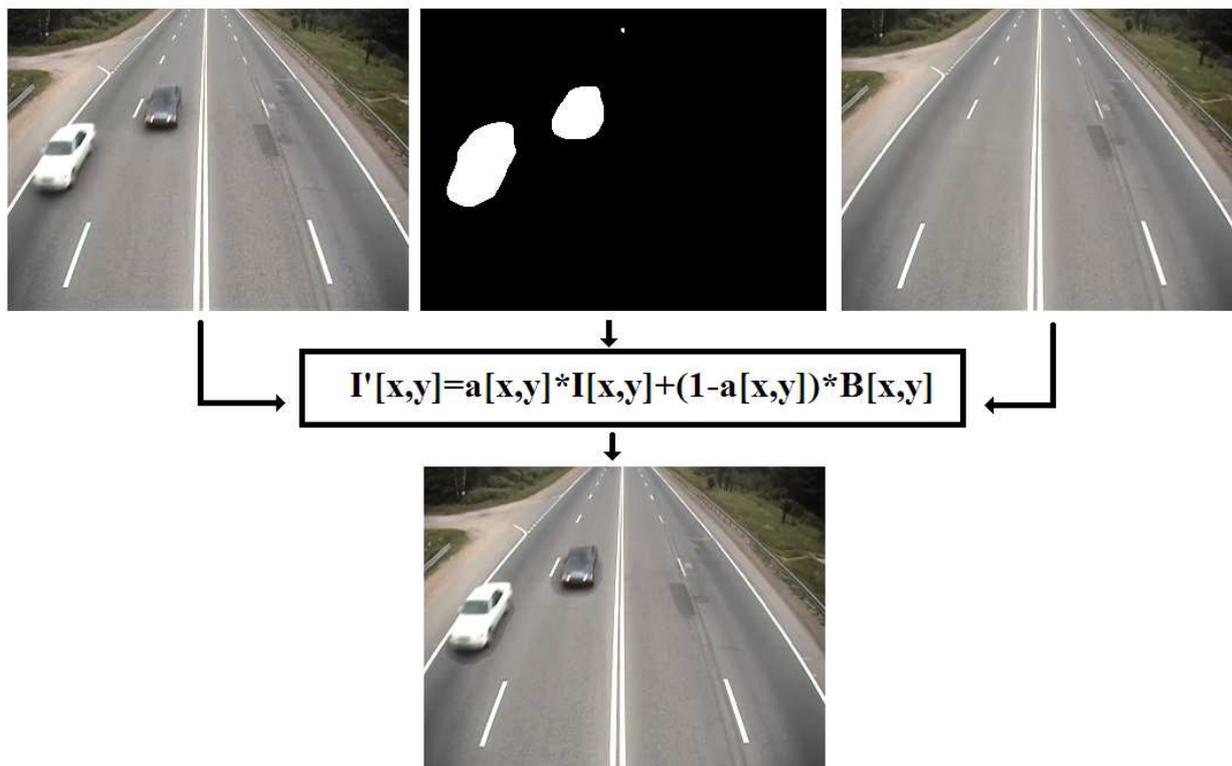


Рис. 3. Пример операции матирования.

На качество синтезированного изображения влияют: 1) динамика изменения освещенности во входном кадре; 2) ошибки оценки фона; 3) ошибки выделения в сформированном изображении альфа-канала; 4) используемые алгоритмы сжатия передаваемых областей.

Для уменьшения потока данных, передаваемых в канал связи, можно передавать изображение только тех участков текущего кадра, которые в результате операции матирования, были бы включены в синтезированное изображение (рис. 4). Для показанного кадра 95% его площади принадлежат фону и не представляют для нас никакого интереса. Таким образом, передачей только динамической составляющей сцены можно получить сжатие порядка $k=20$ раз.

Т.к. бинарное изображение с выделенными областями уже сегментировано, т.е. пиксели изображения пронумерованы номерами областей, то по каналу связи можно передавать только содержимое прямоугольников, охватывающих области, и бинарное изображение альфа-канала. Этот подход напоминает MPEG 1-2 с тем отличием, что там передаются блоки и связанные с

ними векторы движения, а здесь передаются области и связанные с ними характеристики областей.

Движущиеся области требуют дополнительного сжатия, чтобы гарантировано сократить скорость ввода информации в канал связи до требуемого коэффициента сжатия. Собственные исследования алгоритмов в двух разных направлениях сжатия областей – вейвлетном и обмен разрешения на количество бит на пиксель (дизеринг–обратный дизеринг) – привели к сравнительно высоким коэффициентам сжатия, но низкому качеству декодированных (матированных) изображений ($PSNR < 30$ дБ). Поэтому альфа-канал и каждая область сжимаются с помощью JPEG, а метаинформация об областях записывается в тэги файла (формат Exif предусматривает возможность записи тестовых комментариев и до 64 Кб бинарных данных). Изображение оценки фона передается отдельным кадром. Размер сжатого файла кадра оценки фона в 3-5 раза больше, чем сумма размеров файлов альфа-канала и областей.



Рис. 4. Динамическая составляющая сцены – движущиеся объекты.

Номер и тип кадра (оценка фона, альфа-канал или область) детектируется по значению тэгов принятых файлов. На основании одного номера кадров для всех файлов-областей одного кадра собирается динамическая составляющая. Метаданные в полях Exif представляют из себя значения признаков области, разделенные символом «;» (формат таблиц comma separated values). После

декодирования альфа-канала его изображение бинаризуется, чтобы избежать ошибок при матировании. Блок-схемы процедур кодирования и декодирования одного кадра представлены на рис. 5-6.

Таким образом, формат потока данных имеет следующий вид:

Передача JPEG-файла с изображением оценки фона; *Начало кадра 001:* передача JPEG-файла с альфа-каналом; передача JPEG-файла с областью 1; ... передача JPEG-файла с областью L; *Начало кадра 002:* ... *Начало кадра 250:* передача JPEG-файла с альфа-каналом; передача JPEG-файла с областью 1; ... передача JPEG-файла с областью M; Передача JPEG-файла с изображением оценки фона; *Начало кадра 251:* передача JPEG-файла с альфа-каналом; передача JPEG-файла с областью 1; ... передача JPEG-файла с областью N; *Начало кадра 252:* ...



Рис. 5. Блок-схема алгоритма кодирования одного кадра.

Было проведено исследование качества восстановленных изображений по показателю PSNR на четырех клипах (рис.7) в зависимости от метода оценки фона при разных значениях порогового показателя Quality в алгоритме сжатия JPEG.



Рис. 6. Блок-схема алгоритма декодирования одного кадра.

Методика: 1) кодирование: выделить области – создать альфа-канал, создать изображение динамической составляющей, сегментировать альфа-канал, сжать изображение оценки фона в JPEG, сжать изображение альфа-канала в JPEG, сжать изображение динамической составляющей в JPEG; 2) декодирование: бинаризовать сжатое изображение альфа-канала, на основе трех изображений (альфа-канал, динамическая составляющая, фон) вставить динамическую составляющую в декодированное после сжатия изображение фона с помощью матирования, вычислить PSNR и коэффициент сжатия.

Измеряемые характеристики:

1) коэффициент сжатия $k=D/D'$, где D – размер файла несжатого видеопотока, D' – размер файла сжатого видеопотока. Также широко используется показатель «бит на пиксель»:

у исходных изображений $bpp = D/X \cdot Y = X \cdot Y \cdot C \cdot R / X \cdot Y = C \cdot R = 24$,

у сжатых изображений $bpp' = D'/X \cdot Y = D/X \cdot Y \cdot k = bpp/k$. Отсюда следует иная формулировка коэффициента сжатия $k = bpp/bpp'$. В данном эксперименте для каждого кадра суммируются размеры файлов изображения альфа-канала и областей $D'[T] = D'(\alpha[T]) + \sum D'(O_i[T])$;

2) показатель качества изображения рассчитывается по формуле

$$PSNR = 20 \log_{10} \frac{255}{\sqrt{MSE}}, \text{ где } MSE = \frac{1}{X \cdot Y} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} |I(x, y) - I'(x, y)|^2, I(x, y) - \text{ кадр до}$$

сжатия, $I'(x, y)$ – кадр после сжатия.



Рис. 7. Клипы из эксперимента: верхний ряд слева направо – England, car flow, нижний ряд слева направо – highway, Module.

Результаты исследований для значений параметра Quality в алгоритме JPEG 80 и 95 сведены в таблицу 2.

Таблица 2. Зависимость качества изображения PSNR от коэффициента сжатия k при использовании разных алгоритмов оценки фона

PSNR, дБ/ k	Клип England	Клип car flow	Клип highway	Клип Module
28		Медиана 24,3 Медиана 16,2 Мода 14,6		
29		Мода 11,1	Среднее 25,3 Калман 25,1 Калман 16,7 Среднее 16,7	
30		Калман 30,0 Среднее 29,2 Калман 19,6 Среднее 19,2		
31	Среднее 59			
32	Калман 58,7 Мода 58,5 Медиана 57,2 Среднее 41		Медиана 27,8 Мода 27,6	
33	Калман 40,5 Мода 40,4 Медиана 39,5		Мода 18,3	Среднее 45,4 Калман 45,1
34			Медиана 18,4	Медиана 48,1 Мода 47,9 Среднее 32,0 Калман 31,6
35				Медиана 34,1 Мода 33,8

Вывод по неприведенным в таблице значениям PSNR для промежуточных значений Quality – в среднем изменение величины Quality на 5 единиц приводит к изменению PSNR на 0,2-0,5 дБ. Жирным шрифтом выделены случаи, когда превышен требуемый коэффициент сжатия $k=23,2$. При этом нас интересуют случаи, когда $PSNR \geq 31$ дБ.

Для оценки фона можно использовать различные алгоритмы, но желательно соблюсти компромисс между качеством изображения и удобством (т.к. размеры сжатого видео будут примерно одинаковы для всех алгоритмов оценки). Исследование показало, что на видеопотоках плохого качества (car flow) результаты лучше у алгоритма вычисления среднего значения во времени и у фильтра Калмана, так как в таких клипах часто изменяются значения яркости пикселей и вычисленные медиана или мода получаются не настолько выраженными относительно искаженных значений яркости. На видеопотоках высокого качества результаты лучше у алгоритмов попиксельного вычисления медианы и моды во времени. Соответственно, в зависимости от погодных условий и иных факторов ухудшения качества снимаемого видеопотока, целесообразно переключаться между алгоритмами оценки фона: в плохих условиях использовать усреднение во времени, в хороших – медиану во времени. С точки зрения удобства использования алгоритмов – мода и медиана требуют буфер на десятки кадров, что эквивалентно десяткам Мб памяти, в то время как среднее требует памяти на 2 кадра (кадр скользящего среднего и текущий кадр) и фильтр Калмана требует памяти на 21 массив коэффициентов площадью с кадр.

Сравнение предложенного алгоритма с другими кодеками видеосжатия показало, что по коэффициенту сжатия он лучше, чем Indeo Video 5.10 в 4 раза, но хуже Xvid 1.2.2 также в 4 раза.

Выводы:

– сжатие путем уменьшения параметров видеопотока, таких как разрешение и количество уровней квантования, нецелесообразно из-за получающегося небольшого коэффициента сжатия;

– построение инвертированного индекса для сопровождаемых объектов позволяет быстро формировать отчеты об интересующих объектах;

– описанный способ семантического анализа, включающий в себя выделение областей, вычисление параметров областей, фильтрацию областей по несовпадению с заданными целевыми показателями, сопоставление областей текущего кадра с описаниями сопровождаемых объектов, позволяет создать семантическое описание кадра;

– предложенный способ сжатия видеопотока и внедрения в него семантического описания сцены имеет следующие достоинства: 1) гораздо более протяженный временной отрезок между передачей ключевых кадров (изображений оценки фона), чем в традиционных алгоритмах; 2) увеличение временного промежутка не приводит к задержке в кодировании и декодировании, т.к. текущий обрабатываемый кадр не зависит от соседних кадров; 3) приход и уход объектов в ряде других алгоритмов сжатия требует создания ключевых кадров, а в данном алгоритме это не требуется; 4) расширенная функциональность – передача метаинформации о сопровождаемых объектах;

– в зависимости от погодных условий и иных факторов ухудшения качества снимаемого видеопотока, целесообразно переключаться между алгоритмами оценки фона: в плохих условиях использовать усреднение во времени, в хороших – медиану во времени.

– возможно использование других форматов сжатия (JPEG2000 и т.п.) и записи метаинформации (IPTC, XPM).

Литература

1. *Шелухин О.И., Иванов Ю.А.* Влияние помехоустойчивости широкополосных систем беспроводного доступа на качество передачи потокового видео// «Журнал Радиоэлектроники» (электронный журнал). – №9 – 2010. URL: <http://jre.cplire.ru/alt/sep10/5/text.pdf> (дата обращения 05.04.2012)

2. *Кузьмин С.А.* Семантическое сжатие видеопоследовательностей с движущимися средствами// Доклады 11-й Международной конференции и выставки «Цифровая обработка сигналов и её применение». – М., 2009. Т. 2. С. 420-423.
3. *Кузьмин С.А.* Обновление оценки фона с учетом глобального движения// Материалы 6-ой международной конференции «Телевидение: передача и обработка изображений». – СПб, 2008. С. 20-23.
4. *Кузьмин С. А.* Анализ и сжатие видеoinформации в условиях наблюдения объектов на статическом фоне// Труды 7-ой международной конференции «Телевидение: передача и обработка изображений», СПб, 2009. С. 146-150.
5. Синдеев М., Конушин В., Вежнев В. Матирование изображений // Компьютерная графика и мультимедиа (электронный журнал). – № 5(1) – 2007. URL: <http://cgm.computergraphics.ru/content/view/167> (дата обращения 05.04.2012)