

ОСОБЕННОСТИ МЕЖКАДРОВОГО СЖАТИЯ ВИДЕОИНФОРМАЦИИ В УСТРОЙСТВАХ ВИДЕОНАБЛЮДЕНИЯ И ВИДЕОРЕГИСТРАЦИИ

Введение

Вопросы совершенствования методов сжатия видеoinформации, не смотря на стремительный рост производительности вычислительной техники, увеличение пропускной способности интернет-каналов и информационной ёмкости устройств хранения данных, остаются очень актуальными. Развитию методов сжатия способствует внедрение систем телевидения высокой чёткости, систем объемного изображения, а также специализированных прикладных телевизионных систем и устройств.

К наиболее востребованным инструментам на рынке прикладного телевидения относятся системы видеонаблюдения и видеорегистрации [1]. Современные системы видеонаблюдения постоянно совершенствуются и развиваются, используя всё большее количество видеокамер, всё чаще применяются ip-камеры и технологии Интернет, широкое распространение получают специализированные системы видеонаблюдения с борта радиоуправляемых носителей, например – беспилотных летательных аппаратов. Общей тенденцией является увеличение количества передаваемой и записываемой видеoinформации и возрастающие требования к качеству изображения.

подавляющее большинство методов сжатия видеoinформации использует оценку и компенсацию движения для устранения временной избыточности видеoinформации, обусловленной схожестью соседних кадров. Это позволяет существенно сократить цифровой поток при прочих равных условиях. От точности компенсации движения непосредственно зависят степень сжатия и качество изображения. Поскольку сам метод оценки движения, как правило, не описывается в стандартах сжатия, к настоящему времени для этих целей разработано и применяется большое многообразие алгоритмов. Некоторые новые эффективные алгоритмы разрабатываются как комбинация уже существующих. В данной статье проводится анализ и экспериментальное исследование работы некоторых алгоритмов компенсации движения при сжатии видеоизображений в системах видеонаблюдения и видеорегистрации.

Анализ и экспериментальное исследование алгоритмов

Большинство видеокодеров построены на базе единой модели, включающей два основных пути прохождения данных: пространственное кодирование и предсказание с компенсацией движения [2]. На рис.1 показана структурная схема базового видеокодера.

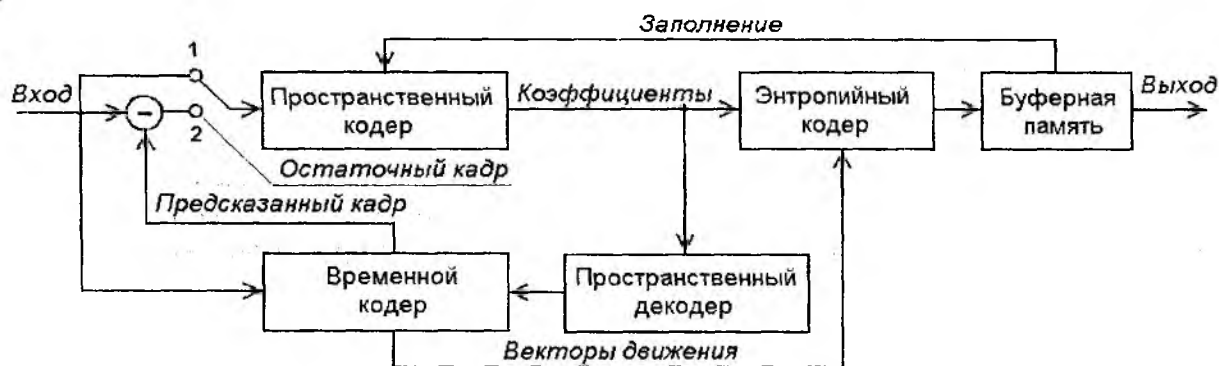


Рис. 1. Структурная схема базового кодера видеоданных

Кодер преобразовывает несжатую видеопоследовательность с помощью пространственного и временного кодеров, которые должны представить эту последовательность по возможно-

сти меньшим числом бит и с наиболее возможной точностью. Эти два требования чаще всего противоречат друг другу.

Процесс межкадрового кодирования состоит следующем. Несжатый видеокادر поступает на вход временного кодера, где сравнивается с опорным кадром. В качестве опорного чаще всего используется один из предшествующих кадров, реконструированных пространственным декодером. Реконструкция позволяет использовать в кодере точно такой же опорный кадр, что и в декодере, и тем самым избежать дрейфа ошибок при декодировании.

Функция временного кодера состоит в поиске в опорном кадре области, наиболее похожей на заданный блок текущего кадра. Смещение этой области относительно текущего блока фиксируется в виде вектора движения. Из найденных в опорном кадре наиболее похожих областей формируется предсказанный кадр, который вычитается из текущего кадра. Эта разность называется остаточным кадром и, по сути, является ошибкой предсказания. Коэффициенты пространственного кодера, а также вектора движения кодируются энтропийными кодерами, после чего поступают в буферную память. Большинство кодеров являются адаптивными, поскольку сигнал заполнения буферной памяти регулирует степень сжатия.

Главная задача межкадрового кодирования – уменьшение количества передаваемой информации за счёт снижения энергии остаточного кадра. Поэтому в большинстве случаев при поиске векторов движения применяются функции оценки подобия блоков, основанные на минимизации энергии разности текущего и предсказанного блоков [3]. Для блока размером $N \times N$ с координатами (x, y) оценочные функции вычисляются как

$$\begin{aligned} SAD(V_x, V_y) &= \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} |L_t(x+m, y+n) - L_{t-1}(x+V_x+m, y+V_y+n)|, \\ SSD(V_x, V_y) &= \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} (L_t(x+m, y+n) - L_{t-1}(x+V_x+m, y+V_y+n))^2, \end{aligned} \quad (1)$$

где (V_x, V_y) – значения вектора движения, L_t, L_{t-1} – значения яркости пикселей текущего и предыдущего кадров соответственно. Традиционно используется сумма абсолютных разностей (SAD – sum of absolute differences), поскольку вычисление данной функции не требует умножений.

Для большинства передаваемых сюжетов, снятых при большом отношении сигнал-шум, вектора соседних блоков мало отличаются, т.е. поле векторов движения достаточно гладкое. Поэтому для сокращения передаваемых данных дальнейшему кодированию и передаче подвергаются разностные вектора движения, имеющие компоненты

$$\Delta V_x^i = V_x^i - V_x^{i-1}; \quad \Delta V_y^i = V_y^i - V_y^{i-1}, \quad (2)$$

где i – номер кодируемого блока.

В системах видеонаблюдения и видеорегистрации не редки случаи, когда съёмка производится в условиях малой освещённости. При уменьшении отношения сигнал-шум в видеосигнале увеличивается вероятность ошибочного определения векторов движения.

С целью выявления условий, при которых возникают ошибочные вектора движения, автором проведен ряд экспериментов, включавших видеосъёмку при различных условиях и определение полей векторов движения для отснятых видеопоследовательностей.

Один из примеров ошибочного определения векторов движения показан на рис.2. В данном случае видеосъёмка проводилась в офисном помещении при боковом освещении светильником с люминесцентной лампой мощностью 40 Вт. Диагональ ПЗС матрицы видеокамеры (1/4)", разрешение 0,8 млн пикселей, диафрагменное число объектива 1,6. Для уменьшения влияния прочих факторов съёмка осуществлялась видеокамерой без межкадрового сжатия (формат Motion JPEG), скорость цифрового потока 25 Мбит/с, размер кадра видеоизображения 720x576 пикселей. Поле векторов движения получено программным способом

по двум соседним кадрам методом полного перебора при минимизации функции (1). Размер блока составил 8x8 пикселей, размер окна поиска – 32x32 пикселя, точность поиска вектора движения – 1 пиксель.



Рис. 2. Пример поля «ошибочных» векторов движения

В данном эксперименте видеокамера неподвижна и движение в кадре отсутствует, тем не менее, найденные вектора движения существенно отличны от нуля. Вектора соседних блоков заметно отличаются, как по абсолютной величине, так и по направлению, поле векторов перестаёт быть гладким. Причём большее число векторов движения найдено в слабо освещённой части кадра, т.е. алгоритм, по сути, отслеживает изменение распределения шума от кадра к кадру. Такая ситуация наблюдалась в большинстве проведенных экспериментов и особенно характерна для съёмок достаточно однородных по своей структуре объектов.

Нахождение «ложных» векторов движения не приводит к искажениям изображения напрямую, найденные вектора соответствуют минимальной энергии предсказанного кадра. Однако данный фактор может снижать эффективность сжатия видеoinформации кодером. Применение оценочных функций SAD и SSD не позволяет учесть цифровой поток, требуемый для передачи векторов движения. При сжатии с большими коэффициентами, что весьма характерно для систем видеонаблюдения и видеорегистрации, а также в случае однородных видеоизображений, хорошо поддающихся сжатию, цифровой поток для передачи векторов движения становится сравним с потоком, требуемым для передачи остаточных кадров. Так, при коэффициенте квантования 10 доля векторов движения составляет около 10 % от общего цифрового потока, при параметре 20 – около 20 %, при параметре 40 – около 35 % [4]. Поскольку сигнал заполнения буферной памяти регулирует поток данных, управляя коэффициентом сжатия, то при увеличении потока векторов движения следует ожидать увеличения потерь в пространственном кодере.

Описанной ситуации можно избежать, применив при поиске векторов движения оценочную функцию, учитывающую не только энергию остаточного кадра, но и битовые затраты на кодирование векторов движения. Одна из таких функций предложена в [5] в виде

$$Cost(V_x^i, V_y^i) = SAD(V_x, V_y) + \lambda \cdot r(V_x^i, V_y^i), \quad (3)$$

где $r(V_x^i, V_y^i)$ – количество бит, необходимое для передачи составляющих векторов движения текущего блока, λ – множитель, учитывающий степень компрессии.

Для минимизации выражения (3) нужно знать значение λ . Один из вариантов второго слагаемого и оценочной функции предложен в работе [6] в виде

$$Cost(V_x^i, V_y^i) = SAD(V_x, V_y) + k_l l \left(|V_x^i - V_x^{i-1}| + |V_y^i - V_y^{i-1}| \right), \quad (4)$$

где k_l – коэффициент, учитывающий степень компрессии текущего кадра. $l = 0,3$ – постоянный коэффициент, найденный в [6] экспериментально.

Второе слагаемое в (4) учитывает, что кодер передаёт разность составляющих векторов движения текущего и предыдущего блоков. При некоторых допущениях величина k_l может быть представлена в следующем виде:

$$k_l = \left(\sum_{i=1}^M S_{i-1}^i \right) / \left(\sum_{i=1}^M SAD_{i-1}^i \right), \quad (5)$$

где M – количество блоков в кадре; S_{i-1}^i – число бит для кодирования i -го блока в $(i-1)$ -м кадре; SAD_{i-1}^i – значения функции (1) для i -го блока в $(i-1)$ -м кадре.

Для оценки эффективности использования оценочной функции (4) при межкадровом сжатии в системах видеонаблюдения и видеорегистрации автором проведен ряд экспериментов по нахождению полей векторов движения для отснятых и описанных ранее видеопоследовательностей.

На рис.3 показано поле векторов движения, найденное для тех же условий, что и поле на рис.2, однако подобие блоков в сравниваемых кадрах оценивалось согласно выражениям (4) и (5).



Рис. 3. Поле векторов движения

Как видно из рисунка, найденное поле векторов движения состоит преимущественно из векторов малой величины и существенно более гладкое – вектора соседних блоков мало отличаются, что уменьшит битовые затраты на их передачу.

Проведена количественная оценка уменьшения скорости цифрового потока при сжатии с использованием оценочной функции (4). Эксперимент проведен для трёх видеопоследовательностей: первая и вторая получены при съёмке в офисном помещении и имеют отношение сигнал-шум 24 и 30 дБ соответственно, третья – при съёмке на улице пасмурным днём, отношение сигнал-шум 38 дБ. Значения отношения сигнал-шум рассчитаны по яркостной компоненте согласно выражению

$$PSNR = 10 \lg \frac{255^2 \cdot X \cdot Y}{\sum_{x=1}^X \sum_{y=1}^Y (I(x, y) - \bar{I}(x, y))^2}. \quad (6)$$

где $X = 720$, $Y = 576$ – размеры кадра в пикселях, $L(x, y)$ – отсчёты яркости произвольного кадра видеопоследовательности, $\bar{L}(x, y)$ – отсчёты яркости, усреднённые по ста кадрам.

В вычислении выражения (6) участвовали только те области изображения, в которых отсутствовало движение. Во всех трёх видеопоследовательностях движущиеся объекты занимали не более 20 % площади кадра.

Сжатие производилось по алгоритму MPEG-2 с различными коэффициентами квантования, поиск векторов движения осуществлялся методом полного перебора. Оценивалась относительная разница D в размерах видеопоследовательностей, сжатых с использованием оценочных функций $Cost$ (4) и SAD (1):

$$D = \frac{S_{Cost} - S_{SAD}}{S_{SAD}} \cdot 100\% \quad (7)$$

График зависимости величины D от коэффициента квантования f для трёх тестовых видеопоследовательностей показан на рис.4.

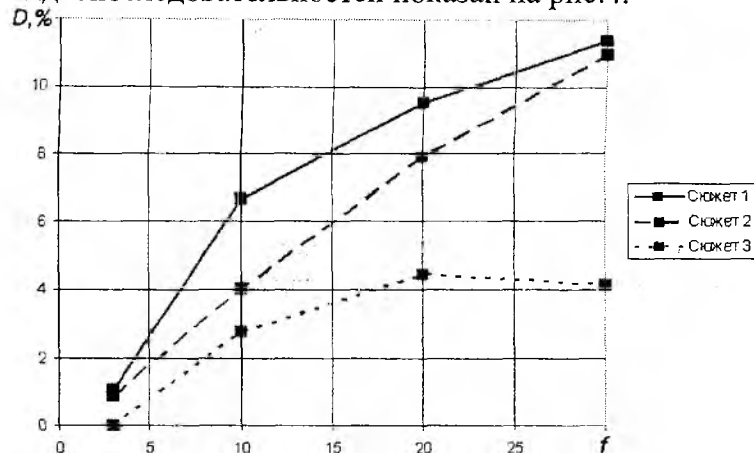


Рис. 4. Выигрыш в размере видеопоследовательностей, сжатых с использованием оценочных функций $Cost$ и SAD

Анализ полученных данных показывает, что по сравнению с функцией SAD использование оценочной функции $Cost$ даёт выигрыш при коэффициенте квантования 30 около 11 %. Наибольший выигрыш получается при сжатии видеопоследовательностей 1 и 2, имеющих меньшее отношение сигнал-шум, а также большее количество крупных однородных деталей, легче подвергающихся сжатию. Видеопоследовательность 3 имеет заметно лучшее отношение сигнал-шум, а также содержит достаточно

большое количество мелких деталей, что позволило получить выигрыш около 4 % при коэффициенте квантования 15 и более.

Также было уточнено значение коэффициента l в выражении (4). При $l = 0,34$ для тестируемых видеопоследовательностей получены наилучшие результаты.

В проведенном эксперименте использовался метод полного перебора как эталонный, поскольку в заданном окне поиска он позволяет найти абсолютный минимум оценочной функции. Однако такой метод поиска обладает высокой вычислительной сложностью, так как для определения наилучшего вектора движения при окне поиска размером $N \times N$ необходимо вычислить оценочную функцию $(N + 1)^2$ раз.

Для уменьшения количества операций предложено большое количество методов, основанных на ограниченном рассмотрении возможного числа кандидатов на вектора движения [3]. Важным свойством каждого такого алгоритма является его устойчивость к воздействию шума.

Рассмотрим свойства сигнала и шума, характерные для систем видеонаблюдения и видеорегистрации. На рис.5 показана частотно-контрастная характеристика (ЧКХ) использованной в экспериментах видеокамеры, снятая по штриховой мире в центре испытательной таблицы ТИГ-0249 для двух значений диафрагменного числа F при выключенном корректоре чёткости.

Завал ЧКХ видеокамеры на верхних частотах обусловлен рядом причин, главными среди которых являются конечная разрешающая способность объектива и матричного фотоприёмника [7]. Особенно заметна потеря чёткости при максимально открытой диафрагме

($F = 1,6$), что характерно для съёмки при малой освещённости. Учитывая также, что в пространственном спектре большинства изображений основная энергия сосредоточена в низкочастотной области, а энергетический спектр флуктуационного видеозума равномерный, можно сделать вывод, что отношение сигнал-шум в системах видеонаблюдения, работающих без специального освещения, на верхних частотах будет существенно меньше, чем на нижних. Следует полагать, что именно этими причинами обусловлено наличие «ложных» векторов движения в однородных областях без мелких деталей, наблюдаемое во всех проведенных экспериментах. Коррекция ЧКХ видеокамеры в области верхних частот не улучшает ситуацию, поскольку не увеличивает отношение сигнал-шум.

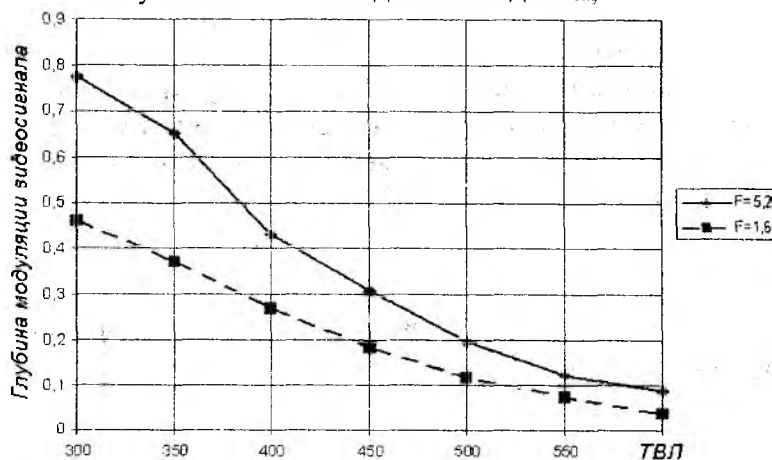


Рис. 5. Частотно-контрастная характеристика видеокамеры

Таким образом, для уменьшения ошибок в определении векторов движения целесообразно применять метод, в котором первоначально проводится поиск движения для низкочастотных компонентов изображения, имеющих большее отношение сигнал-шум, а затем произвести уточнение вектора с использованием высокочастотных компонентов спектра. Возникновение ошибок при таком методе поиска возможно лишь на этапе уточнения, что существенно уменьшает общую ошибку вычисления вектора движения.

Одним из вариантов иерархических алгоритмов является «средняя пирамида» [8]. В этом методе с помощью фильтра нижних частот из исходного кадра формируется кадр с вдвое меньшим разрешением по горизонтали и вертикали. Полученный кадр снова подвергается фильтрации, такая процедура повторяется P раз, где P – параметр алгоритма. Общее количество уровней иерархии составляет $P + 1$, исходный и опорный кадры соответствуют нулевому уровню. Передискретизирующий фильтр может быть описан выражением [8]:

$$L_p(x, y) = \sum_{u=0}^1 \sum_{v=0}^1 L_{p-1}(2x+u; 2y+v), \quad (8)$$

где $L_p(x, y)$ – яркость пикселя с координатами (x, y) кадра уровня p .

Для оценки эффективности использования иерархического метода поиска векторов движения в видеопоследовательностях с малым отношением сигнал-шум проведен эксперимент на описанных выше видеосюжетах. Поиск выполнялся при параметре алгоритма $P = 2$, формировались три уровня иерархии. Один пиксель уровня 2 соответствует блоку 4×4 на нулевом уровне и блоку 2×2 на первом уровне иерархии. На верхнем уровне пирамиды использован трёхшаговый поиск, на остальных уровнях – полный поиск в небольших окнах. Вектору движения, найденному на уровне p , присваивался вес 2^p .

Об эффективности иерархического метода можно судить по среднему модулю отклонения компонентов найденных векторов движения ($V_{x\hat{e}}, V_{y\hat{e}}$) относительно найденных методом полного перебора ($V_{x\hat{i}}, V_{y\hat{i}}$):

$$\Delta = \left(\frac{|V_{x\hat{i}} - V_{x\hat{e}}|}{V_{x\hat{i}}} + \frac{|V_{y\hat{i}} - V_{y\hat{e}}|}{V_{y\hat{i}}} \right) 100\%. \quad (9)$$

В проведенном эксперименте относительное отклонение Δ не превысило 4 %, что подтверждает целесообразность применения иерархического метода при поиске векторов движения в кодерах систем видеонаблюдения.

Выводы

Результаты проведенных экспериментов демонстрируют, что при сжатии видеoinформации в системах видеонаблюдения и видеорегистрации для поиска векторов движения эффективно использовать оценочные функции, учитывающие битовые затраты на передачу векторов движения. Это позволяет снизить скорость цифрового потока до 11 % при параметре квантования 30. Наибольший выигрыш наблюдается при сжатии видеоизображений с невысоким отношением сигнал-шум и с малым количеством мелких деталей. При сжатии с большими параметрами квантования выигрыш в скорости увеличивается, поскольку в этом режиме количество бит, требуемое для кодирования векторов движения, сравнимо с количеством бит для кодирования разностного блока.

Также показано, что для устойчивости алгоритма поиска векторов движения к воздействию шума в системах видеонаблюдения целесообразно применять иерархический метод поиска. Относительная ошибка поиска вектора движения при этом по сравнению с методом полного перебора не превышает 4 %.

Список литературы: 1. *Кругль, Г.* Профессиональное видеонаблюдение. Практика и технологии аналогового и цифрового CCTV / Г. Кругль ; пер. с англ. – М. : Секьюрити Фокус, 2010. – 640 с. 2. *Ричардсон, Я.* Видеокодирование. H264 и MPEG-4 – стандарты нового поколения / Я. Ричардсон ; пер. с англ. – М.: Техносфера, 2005. 368 с. 3. *Кубасов Д.* Обзор методов компенсации движения. [Электронный ресурс] / Кубасов Д., Ватолин Д. // Компьютерная графика и мультимедиа, Сетевой журнал. 2005. №3(2). URL: <http://cgm.computergraphics.ru/content/view/76>. 4. *Беляев, Е.А.* Алгоритмы оценки движения в задачах сжатия видеoinформации на низких битовых скоростях / Беляев Е.А., Тюрликов А.М. // Компьютерная оптика. – 2008. – Т. 32, №4. – С. 403–413. 5. *Kossentini, F.* Predictive RD Optimized Motion Estimation for Very Low Bit-Rate Video Coding / F. Kossentini, Y.-W. Lee, M.J.T. Smith and other // IEEE Journal On Selected Areas in Communications. – 1997. – Vol. 15, N.9. – P. 1752–1763. 6. *Потапов, П.В.* Оценочная функция для поиска векторов похожести при сжатии видеоданных / Потапов П.В., Корилов А.М. // Доклады ТУСУРа. – 2008. – № 2(18), ч.2. – С.87-90. 7. *Куликов, А.Н.* Реальная разрешающая способность телевизионной камеры. [Электронный ресурс] // Специальная техника. 2002, № 2. URL: http://www.ess.ru/sites/default/files/files/articles/2002/02/2002_02_04.pdf (дата обращения 15.06.2012). 8. *Kwon Moon Nam.* A Fast Hierarchical Motion Vector Estimation Algorithm Using Mean Pyramid / Kwon Moon Nam, Joon-Seek Kim, Rae-Hong Park // IEEE Transactions on Circuits and Systems for Video technology. – 1995. – Vol. 5, N.4. – P. 344–351.

Харьковский национальный
университет радиозлектроники

Поступила в редколлегию 08.09.2012