

ИССЛЕДОВАНИЕ ПОДХОДОВ К КОМПЕНСАЦИИ ДВИЖЕНИЯ ПРИ СЖАТИИ ДИНАМИЧЕСКИХ ИЗОБРАЖЕНИЙ

И.В. Рубан, Д.В. Сумцов, А.Ф. Тарасов
(Харьковский университет Воздушных Сил им. И. Кожедуба)

В статье рассматриваются современные подходы к сжатию динамических изображений. Предложены пути совершенствования алгоритмов компенсации движения с целью уменьшения вычислительной сложности при межкадровом кодировании.

сжатие динамических изображений, компенсация движения

Постановка проблемы. В настоящее время для сжатия динамических изображений используются алгоритмы обработки, предусмотренные такими общепринятыми стандартами сжатия видеоданных, как MPEG-1, MPEG-2, MPEG-4. Эти стандарты позволяют обеспечить обмен битовыми потоками между различными приложениями, поддерживают как постоянную, так и переменную скорость передачи, произвольный доступ, переключение каналов, масштабируемое декодирование, редактирование битового потока, а также такие специальные функции, как быстрое воспроизведение, быстрое обратное воспроизведение, обратное воспроизведение с нормальной скоростью, медленное движение, паузу и неподвижные изображения.

Возможность сжатия видеоданных обусловлена такими особенностями исходных данных и человеческого восприятия, как избыточность информации, плавность ее изменения и слабая чувствительность системы человеческого зрения к небольшим искажениям при восстановлении. Многие типы данных содержат в себе статистическую избыточность. Такие данные можно эффективно сжимать, используя компрессию без потерь. Однако, использование сжатия без потерь применительно к компрессии динамических изображений дает относительно небольшой выигрыш – объем данных по отношению к исходному уменьшается в 3-4 раза. Для достижения большего коэффициента сжатия приходится применять алгоритмы сжатия с потерями качества, которые не гарантируют точного восстановления исходных данных, но обеспечивают значительно лучшую степень сжатия, по сравнению с алгоритмами без потерь. При сжатии с потерями предполагается сжатие путем удаления несущей

ственной, релевантной информации. Такие методы основаны на удалении субъективной избыточности, то есть тех элементов изображения, которые можно удалить без заметного влияния на зрительское восприятие видео. Сжатие изображений основывается на корреляции пикселей, а компрессия видео может использовать не только корреляцию близких пикселей каждого кадра видеоряда (пространственную избыточность), но и корреляцию между соседними кадрами (временную избыточность) [1].

Анализ литературы. Подходы к сжатию динамических изображений и вопросы компенсации движения (КД) рассмотрены в литературе [1 – 3].

Основной материал. Для эффективного устранения пространственных и временных избыточностей, присутствующих в любом видеосигнале, современные стандартизованные подходы к видеокомпрессии предполагают использование как внутрикадровой, так и межкадровой компрессии. Типичный метод сжатия видео начинается с кодирования первого кадра с помощью некоторого алгоритма компрессии изображения. Затем осуществляется кодирование каждого последующего кадра. При этом высокие временные затраты стандартизованных видеокодеков обусловлены в основном низкой производительностью методов межкадрового кодирования, которые хотя и позволяют обеспечить межкадровую компрессию в десятки раз выше по сравнению с внутрикадровой, зато и занимают порядка 70-80% от общих временных затрат видеокодеков. Наиболее весомой по временным затратам является процедура оценки и компенсации движения (КД). Рассмотрим ее более подробно.

Модуль КД является составной частью практически всех современных видеокодеков. КД в видеопотоке – это процесс извлечения информации о характере и параметрах так называемого «оптического» двумерного движения объектов (то есть в плоскости кадра) по имеющейся видеоинформации. Отличия между кадрами могут появляться из-за движения объектов, не меняющих свою форму (например, автомобилей), перемещения видеокамеры (панорамная видеосъемка, наклоны, наезды, изменение фокуса, повороты и т.п.), изменения перекрытия объектов (открываются части сцены, которые были закрыты движущимся объектом), изменения освещенности и в силу других причин. За исключением изменения перекрытия и освещенности объектов съемки, эти отличия соответствуют перемещению групп пикселей по кадру. На рис. 1 приведена схема работы алгоритма КД на основе блоков.

При КД оценивается скорость перемещения движения объектов от кадров, и при определенных предсказаниях производится коррекция положения опорного изображения, по отношению к которому находится ошибка предсказания ϵ [2]. Набор чисел, оценивающих смещение, принято называть вектором движения.

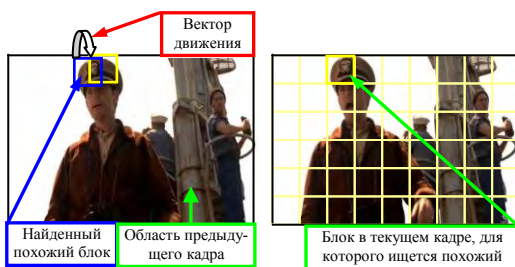


Рис. 1. Схема работы алгоритма КД

Особенности КД:

1. Степень сжатия видео выше, чем при кодировании разности кадров.
2. Стадия кодирования алгоритма КД в вычислительном отношении интенсивна.
3. Схема КД работает лучше всего для сцен с ограниченным движением.

КД основана на использовании ряда сложных алгоритмов. Сфера, где данная технология сжатия эффективна, как правило, сводится к видеоряду, в котором объект изменяет свое местоположение относительно неподвижного фона. Объекты, изменяющиеся по форме, приближающиеся или удаляющиеся (движущаяся камера), вращающиеся объекты, не подлежат эффективному сжатию посредством алгоритма КД. Схема алгоритма кодирования с КД приведена на рис. 2 [2].

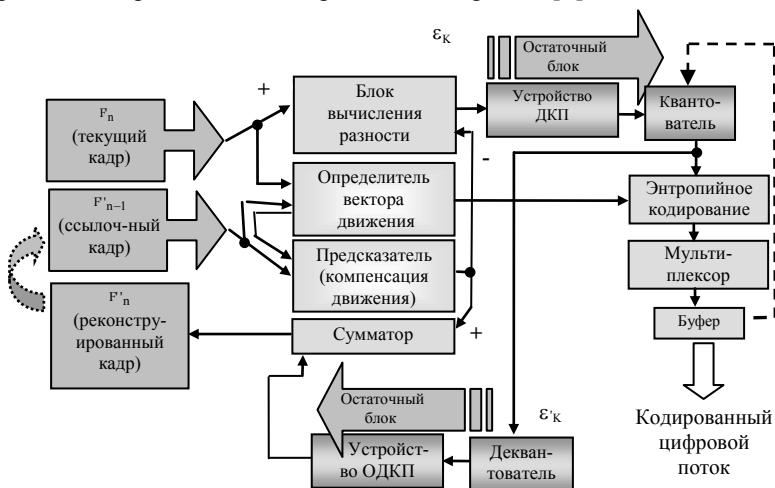


Рис. 2. Блок-схема видеокодера на основе КД и поблочного преобразования

Сигналы нового кадра F_n , поступающие с препроцессора, сравниваются с хранимым в памяти кодера кадром F'_{n-1} по блокам определенного размера. Когда определяется подходящая область на ссылочном кадре, вырабатывается вектор, описывающий направление и дистанцию движения этого участка (блок определения вектора движения). Вектор движения задает позицию наилучшего блока по отношению к текущему блоку. Он обозначает смещение на предсказанную ссылку ранее декодированного кадра, которая является минимальной в сравнении со всеми возможными ссылками. Задание вектора смещения элементов изображения уменьшает объем данных, необходимых для описания межкадровой разности (остаточного кадра). Векторы движения позволяют определить положение кодируемого отсчета в новом кадре (скомпенсировать его перемещение) и, таким образом, сохранить высокую точность предсказания. Информация о векторах и сведения из памяти поступают в предсказатель. В нем формируется видеосигнал предсказанной картинкой, которая сравнивается в блоке вычисления разности с оригинальным кадром для получения сигнала разностного изображения. Вычисляется ошибка предсказания ϵ , т.е. разность между фактическим и предсказанным с использованием вектора движения элементами изображения, что снижает временную избыточность.

Процесс уменьшения пространственной избыточности начинается с преобразования сигнала из временной в частотную область с целью декорреляции значений элементов в каждом блоке, или уплотнения как можно большего количества информации в наименьшее число коэффициентов преобразования [3]. При этом разложение может проводиться по различным ортогональным функциям. Наибольшее распространение получило преобразование на основе дискретного косинусного преобразования (ДКП). ДКП производится над блоками элементов разностной картинки (остаточными блоками). Результат ДКП – частотный спектр каждого блока. Чем меньше мелких деталей в блоке, тем меньше уровень высокочастотных компонент. С целью сжатия информации их отбрасывают.

С выхода преобразователя ДКП информация поступает на квантователь, где происходит ее округление и весовая обработка в соответствии с заложенной матрицей. После такой операции частотные составляющие с малой энергией исчезают.

После квантования сигналы попадают в этропийный кодер, где ошибка предсказания переупорядочивается и покоординатно кодируется вместе с векторами движения, мультиплексируется и передается вместе с основным потоком через буфер, от которого имеется обратная связь на квантователь.

Таким образом, формируется сжатый битовый поток или файл, который можно передавать по сетям или хранить на носителях цифровых данных.

Управление степенью компрессии и скоростью потока данных происходит через обратную связь с помощью изменения параметров матрицы квантования. Чем грубее квантование, тем больше нулевых значений в матрице коэффициентов и тем меньше объем данных, необходимых для передачи информации об изображении. Однако, с увеличением степени компрессии растут и необратимые искажения изображения из-за шумов квантования.

Для формирования сигнала точной предсказанной картинке необходимо в памяти кодера иметь копию передаваемого изображения с учетом всех искажений, возникающих при кодировании. Поэтому в кодер введена петля КД, состоящая из деквантователя, блока обратного ДКП (ОДКП) и сумматора. Фактически эта часть устройства представляет собой декодер.

Декодер использует вектор движения для нахождения блока прогноза. Выполняется инверсное квантование, инверсное ДКП, в результате чего формируется ошибка предсказания, которая складывается с прогнозом (декодированным изображением предыдущего кадра) для реконструкции версии исходного блока (декодированного изображения текущего кадра).

Наилучшее качество приближения, то есть минимальную ошибку предсказания, может гарантировать полный перебор всех возможных значений векторов смещения из допустимой области с подсчетом ошибки компенсации для них и выбор такого вектора, на котором достигается минимум ошибки.

Ошибка компенсации ϵ характеризуется энергией остатков остаточных блоков. Наиболее простой мерой энергии остатков, в вычислительном отношении, является сумма абсолютных отклонений (SAE, Sum of Absolute Error):

$$SAE = \sum_{i=0}^{N-M-1} \sum_{j=0}^{M-1} |C_{ij} - R_{ij}|, \quad (1)$$

где N , M – размер блока по горизонтали и вертикале; C_{ij} – значение яркости текущего блока с координатами i, j ; R_{ij} – значение яркости ссылочного блока с координатами i, j .

Когда найдено минимальное значение SAE вычисляется вектор движения, как разность между прошлым и настоящим положением блока

$$\vec{d} = (C_i - R_i, C_j - R_j). \quad (2)$$

Таким образом, целью компенсации движения является минимизация ошибки предсказания всех векторов движения для рассматриваемого блока по следующему критерию:

$$\varepsilon(\vec{d}, n) = \min_{d \in \{D\}} \varepsilon(\vec{d}_i, n), \quad (3)$$

где \vec{d} – искомый вектор движения; n – номер блока; $\vec{d}_i \in \{D\}$ – множество векторов в области поиска D .

Этот подход считается эталонным и является базой временных моделей во всех современных стандартах цифрового видеокодирования. Сравнение с КД на основе блоков является неотъемлемой частью любой работы, посвященной разработке нового алгоритма КД.

Однако, практическое его применение в системах сжатия не представляется возможным ввиду слишком большой вычислительной сложности. Например, если поиск вектора движения осуществляется в области $\Delta x = \Delta y = 8$ пикселей для каждого блока размером 16×16 , то область поиска состоит из $(16 + 2 \times 8)(16 + 2 \times 8) = 1024$ пикселей, по которым нужно вычислить SAE $(1 + 2 \times 8)(1 + 2 \times 8) = 289$ раз, в то время как каждое вычисление SAD требует 256 операций взятия модуля и 255 операций сложения (рис. 3).

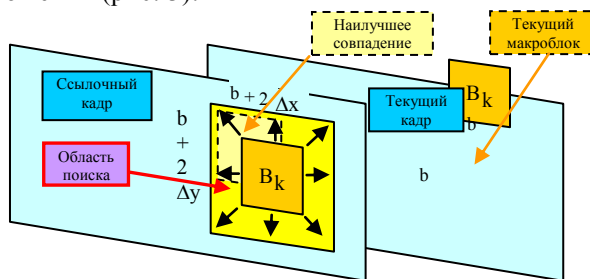


Рис. 3. Оценка движения

Для кадра размером 352×288 (стандарт CIF) получаем $22 \times 18 = 396$ макроблоков, для которых необходимо произвести $396 \times 289 \times (256 + 255) \approx 59$ млн. операций.

Полный перебор всех возможных вариантов требует значительных затрат времени, поэтому перспективной является разработка подоптимальных алгоритмов, которые осуществляют адаптивный поиск блоков – кандидатов. Такие методы позволяют существенно ускорить весь процесс сжатия, за счет небольшого снижения эффективности компрессии [3]. Кроме большой вычислительной сложности, данный метод об-

лаждает еще несколькими недостатками. Реальные объекты редко имеют четкие прямоугольные границы. Объекты часто перемещаются на нецелое число пиксельных позиций между кадрами. Многие типы перемещений трудно скомпенсировать, используя методы на основе блоков (деформацию объектов, их повороты, изменения масштаба, т.е. приближение или удаление объектов от видеокамеры, сложные движения, такие как клубы дыма).

Вывод. Оценка движения выполняется видеокодером и вносит значительную временную задержку в процессе сжатия. Правильный выбор прогнозной ссылки сокращает остаточную энергию после компенсации движения, которая в свою очередь максимизирует степень сжатия. Недостатком эталонного подхода к анализу и компенсации движения является необходимость использования значительной вычислительной мощности. Так, например, из требуемой вычислительной мощности порядка 20 млрд. операций/с при обработке динамических изображений примерно 75% операций требуется для реализации системы анализа и КД. В связи с этим, разработка новых алгоритмов КД, имеющих меньшую вычислительную сложность, является актуальной.

Перспективными являются задачи:

- выбора наименьшего числа точек, наилучшим образом передающих изменение рельефа детали объекта, и анализа движения этой детали с использованием только ее характерных точек;
- использование корреляционных свойств векторов движения соседних блоков для уменьшения области поиска вектора движения;
- упрощение процедуры сравнения блоков или снижение необходимого числа таких сравнений;
- адаптация размеров блока к характеристикам изображений.

ЛИТЕРАТУРА

1. Сэлмон Д. Сжатие данных, изображений и звука. – М.: Техносфера, 2004. – 368 с.
2. Ричардсон Я. Видеокодирование. H.264 и MPEG-4 – стандарты нового поколения. – М.: Техносфера, 2005. – 368 с.
3. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2005. – 1072 с.

Поступила 22.12.2005

Рецензент: доктор технических наук, профессор Ю.В. Стасев,
Харьковский университет Воздушных Сил им. И. Кожедуба.