

УДК 004.932

Г.Н. Востров, канд. техн. наук, доц.,
 А.В. Монастырский, магистр,
 Д.М. Полторац, магистр,
 Одес. нац. политехн. ун-т

АНАЛИЗ ПОТЕРЬ ПРИ КОМПРЕССИИ ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЕЙ И МЕТОДОВ ИХ УСТРАНЕНИЯ

Г.М. Востров, О.В. Монастырский, Д.М. Полторац. **Аналіз втрат при стиску відеопослідовностей та методів їх уникнення.** Розглядається проблема втрати якості зображень у процесі стиску при обробці відеопослідовностей. Представлено аналіз існуючих методів рішення даної проблеми та розглядаються основні передпосилки при проектуванні нового методу компенсації руху.

G.N. Vostrov, A.V. Monastyrskyy, D.M. Poltorak. **Analysis of quality losses in video image compression and methods for their elimination.** The quality losses problem in video compression applications is considered. The analysis of the existing techniques for resolving this problem is presented as well the main prerequisites in developing a new method for motion compensation are considered.

Сжатие видео является важной задачей, которая требует эффективного решения. Уже накоплено большое количество видеoinформации, и объем ее непрерывно растет. Это и записи конференций, и всевозможные учебные материалы, и просто художественное видео. Для хранения этой информации емкости современных носителей информации недостаточно, следовательно, необходимо устранять всякого рода ее избыточность, т.е. разрабатывать и развивать методы сжатия динамических изображений.

Проведен анализ причин потери качества при сжатии видеопоследовательностей. Предлагаются пути улучшения одного из самых эффективных методов повышения качества (устранения ошибки) — метода компенсации движения.

Укрупненно процесс сжатия видеопоследовательностей состоит из следующих этапов:

- подготовка изображений к дальнейшей обработке, например, преобразование цветовой модели или шумоподавляющая фильтрация;
- аппроксимация изображений;
- сохранение результатов.

Для оценки качества изображения использована метрика отношения сигнала к шуму (Peak-to-peak Signal-to-Noise Ratio — PSNR) [1]

$$\text{PSNR}(I, D) = 10 \log_{10} \frac{255^2 n^2}{\sum_{i=1, j=1}^{n, m} (I_{i,j} - D_{i,j})^2}, \quad (1)$$

где I — оригинальное изображение;

D — восстановленное изображение;

n — размер изображений по вертикали;

m — размер изображений по горизонтали.

Самым известным и широко применяемым методом аппроксимации видеопоследовательностей является метод оценки/расчета движения (Motion Estimation — ME), основанный на предположении о схожести соседних кадров видеопоследовательности [2]. Так как стандартная видеопоследовательность PAL — 25 кадров в секунду, это предположение для обычных фильмов оказывается достаточно правдоподобным. В соответствии с методом изображение разбивается на макроблоки стандартного размера 16×16 пикселей, хотя некоторые методы работают с различными размерами блоков. Из последовательности берется первый кадр и последующий, находящийся на расстоянии N кадров от него, и сжимаются одним из методов сжатия статических изображений [2]. Полученные кадры называют опорными. Все промежуточные кадры составляются из двух опорных кадров путем нахождения для каждого макроблока предсказываемого кадра наиболее подходящего ему макроблока из опорных кадров. Далее остается сохранить лишь вектор смещения для данного макроблока в результирующем

потоке. Координаты вектора смещения определяются следующим образом: $(x-v_x; y-v_y)$, где x, y — координаты верхнего левого угла обрабатываемого макроблока, а v_x, v_y — координаты верхнего левого угла аппроксимирующего макроблока.

Данная схема кодирования позволяет достичь довольно больших коэффициентов сжатия видеопоследовательностей. Количество векторов для изображения

$$V = 2(m/b_x)(n/b_y), \quad (2)$$

где b_x — размер макроблока по горизонтали;

b_y — размер макроблока по вертикали.

В свою очередь, размер изображения в байтах

$$S = mnCdb, \quad (3)$$

где C — количество байт для представления одного пикселя.

Например, для изображения с размерами $m=352$ и $n=240$ пикселей, глубиной цвета $C=3$ байта и при размере макроблока 16×16 пикселей ($b_x=b_y=16$), необходимо сохранить лишь $V=660$ координат векторов смещения. Даже если на хранение каждой компоненты вектора выделить 1 байт (это довольно много, т.к. обычно значения вектора лежат в диапазоне $-16 \dots 16$), то получится несжатый массив векторов размером в 660 байт. Так как размер изображения $S=253440$ байт, то коэффициент сжатия изображения составит 384 раза.

Качество сжатой последовательности сильно зависит от значения N . Чем оно больше, тем хуже будет качество, т.к. видеопоследовательностям в подавляющем большинстве случаев присуща динамичность. На практике для достижения высоких степеней сжатия значение N должно быть достаточно большим.

Аппроксимация является не единственным этапом в сжатии видеопоследовательности, на котором теряется качество. Рассмотрены и проанализированы остальные этапы, а также некоторые особенности методов, применяемых на данных этапах, которые приводят к потере качества.

Большинство алгоритмов используют преобразование из цветового пространства RGB в цветовое пространство YUV (YCbCr), где Y — компонента, отвечающая за яркость, а $U(Cb)$ и $V(Cr)$ — цветовые компоненты, для отделения яркости от цвета. Человеческое зрение более чувствительно к яркости, чем к цвету. Пользуясь этим, можно еще до аппроксимации уменьшить количество информации о цвете в 2 или 4 раза[1]. Упрощенно перевод из цветового пространства RGB в цветовое пространство YUV можно представить с помощью матрицы перехода

$$\begin{pmatrix} Y \\ Cb \\ Cr \end{pmatrix} = \begin{pmatrix} 0,299 & 0,587 & 0,114 \\ 0,5 & -0,4187 & -0,0813 \\ 0,1687 & -0,3313 & 0,5 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} + \begin{pmatrix} 0 \\ 128 \\ 128 \end{pmatrix}. \quad (4)$$

Обратное преобразование осуществляется умножением вектора YUV на обратную матрицу.

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1,402 \\ 1 & -0,34414 & -0,71414 \\ 1 & 1,772 & 0 \end{pmatrix} \begin{pmatrix} Y \\ Cr \\ Cb \end{pmatrix} - \begin{pmatrix} 0 \\ 128 \\ 128 \end{pmatrix}. \quad (5)$$

Очевидно, что при уменьшении информации о цвете снижается качество изображения.

Опорные кадры сжимаются алгоритмами с потерями и изначально вносят ошибку в результирующую видеопоследовательность.

Все промежуточные кадры, также как и опорные, перед кодированием переводятся в модель цветоопределения YUV.

Не всегда удается расположить опорные кадры достаточно близко друг к другу, это приводит к тому, что они могут находиться в различных сценах видеопоследовательности. Тогда аппроксимация будет происходить лишь по одному опорному кадру, что, естественно, приведет к положительным результатам лишь на слабодинамичных сценах.

Даже при близком расположении опорных кадров может возникнуть ситуация, когда некоторые детали промежуточного кадра не присутствуют ни в одном из опорных (в сцене наблюдается поворот объекта, неравномерное изменение освещения сцены, сильное движение камеры, и т.д.).

Разбиение изображения на блоки ведется без учета особенностей изображения. Стандартные размеры блока 16×16 довольно большие для точной работы МЕ. Влияние последнего фактора уменьшают путем разбиения макроблока на блоки меньших размеров, но все же это не решает проблему в глобальном плане.

Так как в качестве опорного кадра зачастую используют предыдущий аппроксимированный, то все артефакты из него автоматически переносятся в текущий предсказываемый.

Очень большая проблема связана с метрикой, по которой оценивается наилучший блок для аппроксимации. Можно найти наилучший по метрике, например, PSNR, блок, который в то же время будет совсем не лучшим с позиции зрительного восприятия.

Данные сравнительной таблицы демонстрируют падение качества изображения на различных этапах сжатия.

Падение качества на различных этапах сжатия видеопоследовательностей

Этапы сжатия			Качество изображения, PSNR		
			Стандартные видеопоследовательности [3]		
			Flower	Football	Mobile
Перевод из RGB в YUV			43,001831	44,729213	44,402490
Аппроксимация методом МЕ	Расстояние между опорными кадрами N	Качество опорных кадров			
	30	высокое	22,097891	23,444668	24,574162
	60	высокое	20,989432	21,350622	21,834961
	30	низкое	12,795367	14,217944	15,246159

Самыми распространенными стандартами для сжатия видеопоследовательностей являются MPEG2 и MPEG4, в которых для устранения ошибки применяется метод компенсации движения (Motion Compensation — MC) [4]. Суть метода заключается в следующем: если найденный лучший аппроксимирующий блок не дает хорошего качества, то вычисляется разница между оригинальным и найденным блоком, далее к массиву разницы применяется Дискретное Косинусное Преобразование (ДКП), затем этот массив квантуется и сжимается методом кодирования с переменной длиной (variable length coder) [4].

Метод хорошо проработан и успешно используется на протяжении последних лет, но все же имеет ряд недостатков:

- искажается информация об ошибке при обратном ДКП из-за квантования коэффициентов, полученных после ДКП; если присмотреться к краю объектов в фильмах, закодированных с помощью MPEG2 или MPEG4, легко можно увидеть четкое наличие муара (“эффект Гиббса”) [3];

- проявляется эффект “блочности” (на изображении отчетливо видны границы блоков) из-за того, что качество рядом стоящих блоков может сильно отличаться, т.к. метод работает на блоках 16×16 , и для каждого блока, независимо от соседних, принимается решение о сохранении ошибки;

- процедура ДКП довольно ресурсоемка (около 80% времени работы всего метода), что создает проблемы при реализации приложений для кодирования или декодирования видео в реальном времени.

По результатам анализа недостатков метода MC предлагается следующее:

- применять процедуру квантования самих значений ошибки, а не коэффициентов, полученных после декоррелирующего преобразования (таким преобразованием в MPEG2 и MPEG4 является ДКП), и до кодирования использовать преобразования с управляемыми потерями;

- работу вести не на блоках, а на целом изображении.

Данные модификации позволят:

- избежать проявления “эффекта Гиббса”, т.к. преобразования с управляемыми потерями, в отличие от ДКП, никогда не вносят дополнительную ошибку в изображение;

- уйти от поблочной обработки изображения и избавиться от эффекта “блочности” при работе на целом изображении;

- значительно повысить скорость работы метода из-за отсутствия необходимости использовать ДКП.

Предложенные модификации позволят повысить эффективность методов сжатия видеопоследовательностей, что откроет новые возможности для мультимедийных технологий.

Литература

1. Bhaskaran V., Konstantinos K. Image and Video Compression Standards: Algorithms and Architectures. — Boston: Kluwer Academic Publishers, 1995. — 432 p.
2. MPEG Video Compression Standard / Mitchell J.L., Pennebaker W.B., Fogg C.E., Le Gall D.J. Chapman and Hall, 1997. — 516 p.
3. ISO/IEC 14496-2 standart. Information tehnology. Coding of audio-visual objects. P. 2: Visual — 2001. — 517 p. — <http://www.iso.cn>. — 29.12.2003.
4. Kuhn P. Algorithms, Complexity Analysis and VLSI Architectures for MPEG-4 Motion Estimation. — Boston: Kluwer Academic Publishers, 1999.

Поступила в редакцию 3 октября 2003 г.