

УДК 612.8

МЕТОДЫ УВЕЛИЧЕНИЯ СТЕПЕНИ СЖАТИЯ ИЗОБРАЖЕНИЙ ЭНТРОПИЙНЫМИ КОДЕРАМИ

Н. Н. Красильников,

д-р техн. наук, профессор

Санкт-Петербургский государственный университет аэрокосмического приборостроения

Рассмотрены перспективы применения предварительного разделения изображений на контексты при их сжатии энтропийными кодерами. Эксперименты показали, что применение такого подхода дает ощутимый результат, обеспечивая дополнительное сжатие изображений. Сформулированы направления дальнейших исследований.

In this paper some perspectives of application of preliminary contexts separation methods in images compression with use of entropy encoders are considered. Performed experiments demonstrated that such an approach leads to essential effect providing additional compression of images. The directions of further investigations are formulated.

Введение

В связи с широким применением цифровых методов в телевидении возникла проблема сжатия движущихся изображений как при съемке исходного материала, так и при его передаче или записи на компакт-диск. Несмотря на то, что в настоящее время разработаны весьма эффективные методы компрессии движущихся изображений, обеспечивающие сжатие цифрового потока, которым представляются движущиеся изображения, в десятки раз, проблему нельзя считать полностью решенной. Объясняется это тем, что при использовании этих методов (так называемых методов сжатия с потерей информации) в изображения вносятся искажения (шум преобразования), которые добавляются к изображениям при каждом новом цикле сжатия-восстановления. Поскольку при редактировании телевизионных программ сжатые изображения приходится неоднократно распаковывать из файла и снова записывать в файл со сжатием, результирующая помеха преобразования накапливается и качество изображения становится недопустимо



■ **Рис. 1.** Исходное изображение

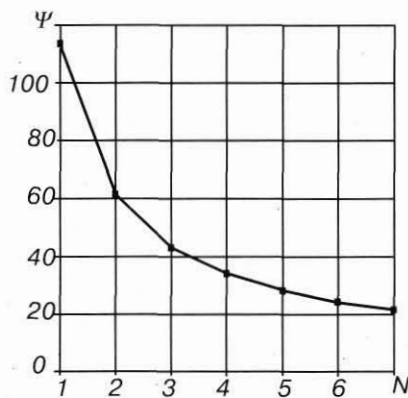
низким. Сказанное справедливо как для методов сжатия, в которых используется межкадровая избыточность движущихся изображений для увеличения степени сжатия, например, таких, как MPEG2, так и для методов, в которых для увеличения сжатия межкадровая избыточность не используется, как, например, в методе MJPEG. Аналогичная проблема встречается и в издательском деле при подготовке к печати неподвижных изображений.

Эффект накопления шума преобразования изображений при их многократном сжатии-восстановлении кодерами с потерями

Проиллюстрируем изложенное результатами эксперимента. Исходное изображение (рис. 1) было несколько раз сжато с записью в файл посредством архиватора JPEG, а затем восстановлено. При выполнении эксперимента величина сжатия была выбрана минимальной для используемого архиватора и составляла всего 1,624 раза, что соответствовало наивысшему качеству изображения, обеспечиваемого этим архиватором.

Результаты эксперимента представлены на рис. 2 в виде зависимости отношения максимального значения сигнала (перепад яркости от черного до белого) к среднеквадратичному значению шума преобразования Ψ от числа циклов сжатия-восстановления N . Из данного графика следует, что уже после первого цикла сжатия-восстановления изображения отношение сигнала к шуму становится равным $\Psi = 114$, а после второго цикла 61, что недопустимо мало.

Полученные результаты достаточно хорошо отражают существующее положение дела и в случае архивации движущихся изображений, поскольку характеристики JPEG при сжатии неподвижных изображений близки к характеристикам MJPEG. Помеха преобразования носит весьма специфический характер. Как правило, на



■ **Рис. 2.** Зависимость отношения сигнала к шуму преобразования от количества циклов сжатия-восстановления архиватором JPEG

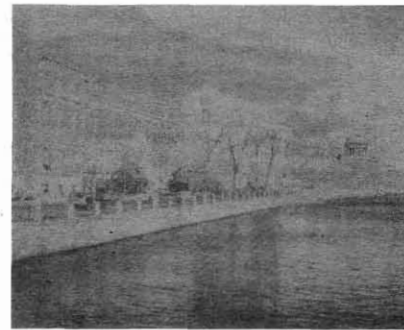
восстановленном изображении появляется блочная структура, контраст которой возрастает при увеличении количества циклов сжатия-восстановления. В некоторых архиваторах, например, в архиваторе JPEG последних версий графического редактора Adobe Photoshop, используются методы, маскирующие блочную структуру. В этом случае вид помехи несколько изменяется и становится таким, как показано на рис. 3, на котором можно видеть различие между исходным изображением и изображением, восстановленным после пяти циклов сжатия-восстановления.

По изложенным выше причинам в настоящее время методы компрессии изображений с потерями, обеспечивающие высокую степень сжатия, применяются только на конечной стадии производства, например, для записи на компакт-диск неподвижных изображений, работа над которыми уже завершена, или для передачи готовой телевизионной программы по цифровому каналу. В процессе же съемки исходного материала, а также в процессе его редактирования рекомендуется либо хранить изображения без сжатия, либо использовать для этой цели методы сжатия без потерь, обеспечиваемые энтропийными кодерами.

Сжатие изображений с разделением на контексты

Недостатком энтропийного кодирования является малая степень сжатия. Так, использование алгоритма LZW позволяет получить сжатие исходного изображения (см. рис. 1) лишь в 1,12 раза, что, конечно же, очень мало. В последнее время в связи со снятием патентных ограничений на арифметическое кодирование появилась реальная возможность использовать этот метод сжатия, применение которого к изображению на рис. 1 позволяет сжать его в 1,69 раза, т. е. даже чуть больше, чем при использовании JPEG в режиме максимального качества. Тем не менее, это не предел в отношении достижимых методов сжатия при использовании энтропийных кодеров и, в частности, арифметического кодирования.

Дело в том, что энтропийные кодеры ориентированы на работу со стационарными сигналами, статистика которых неизменна. Однако различные участки изображения, как правило, обладают различной статистикой. В качестве наглядного примера обратимся к рис. 4, на котором показано исходное изображение (см. рис. 1), «наклеенное» на паспарту. Из рисунка вид-

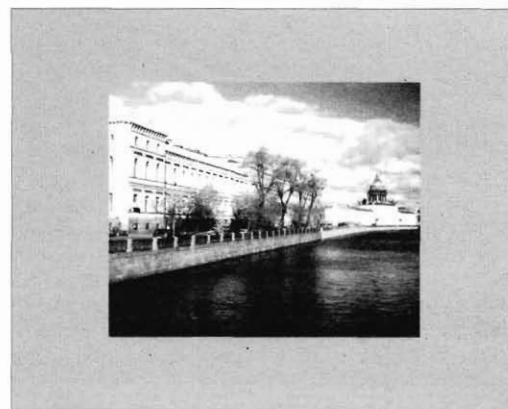


■ **Рис. 3.** Вид помехи преобразования исходного изображения после пяти циклов сжатия-восстановления (помеха преобразования наложена на равномерный серый фон)

но, что часть этого изображения содержит много мелких деталей, а часть представлена однородным серым фоном. При сжатии такого изображения кодер настраивается на среднюю статистику кодируемого изображения и, как результат, работает не оптимально, хотя вследствие наличия в кодируемом изображении больших однородных полей сжатие составляет 2,36 раза.

Ситуация может быть существенно улучшена, если изображение, показанное на рис. 4, разделить на две компоненты, или на два контекста: на однородную и детальную части, а затем сжимать их по отдельности. В этом случае дополнительно к информации детальной компоненты необходимо передать информацию о границах, разделяющих однородную и детальную части, а также интенсивность серого фона. Проведенный нами эксперимент показал, что при таком подходе величина сжатия увеличивается по сравнению со сжатием без разделения на контексты в 1,79 раза и становится равной 4,238.

Методы подстройки энтропийного кодера под изменяющуюся статистику сжимаемых текстовых сообщений (так называемые методы контекстного моделирования) достаточно хорошо известны в кругу разработчиков программ сжатия, которыми созданы так называемые алгоритмы PPM (Prediction by Partial Matching) [1]. Специалистам же, работающим в области сжатия изображений и, в частности, в области телевидения, такой подход менее известен и только сейчас он начинает ими использоваться [2].



■ **Рис. 4.** Исходное изображение, «наклеенное» на паспарту в виде равномерного серого поля

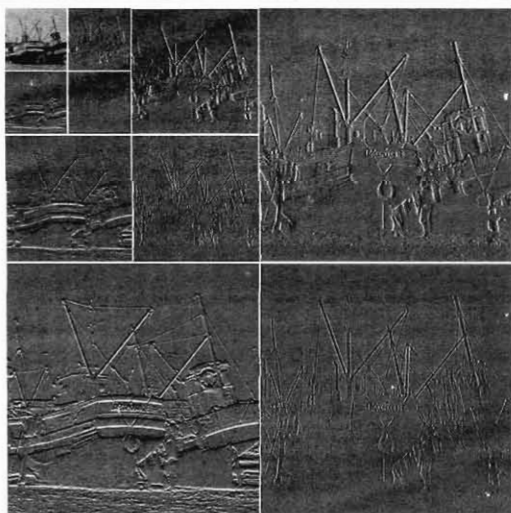


■ Рис. 5. Тестовое изображение «Яхта»

В качестве другого примера рассмотрим разделение на контексты графика, показанного на рис. 2. Это изображение можно разделить на отрезки прямых, представляющие собственно зависимость, экспериментальные точки, буквы и цифры, а также сетку. Мы провели эксперимент, в процессе которого изображение было разделено на указанные контексты. Эксперимент показал, что при сжатии посредством арифметического кодирования без разделения на контексты величина файла с заголовком составила 2,47 Кб, а при разделении – 1,98 Кб. Таким образом, разделение на контексты в данном случае увеличило общую степень сжатия в 1,246 раза, при этом сжатие составило 123 раза.

Разделение на контексты сжимаемого изображения может осуществляться как в пространстве изображений, так и в частотной области. Построение алгоритма разделения изображения на контексты составляет одну из важных проблем при разработке архиватора, использующего этот метод.

С позиций контекстного разделения изображений, перед их сжатием может быть рассмотрен ряд широко



■ Рис. 6. Компоненты, полученные в результате трехкратного вейвлет-преобразования тестового изображения «Яхта»

применяемых при обработке изображений преобразований, например, дискретное косинусное преобразование, а также вейвлет-преобразование, на базе которого разработан архиватор JPEG2000 [1, 3]. В этом архиваторе исходное изображение вначале рекурсивно подвергается трем вейвлет-преобразованиям, а затем кодированию полученных в результате этих преобразований компонент (матриц коэффициентов). При кодировании коэффициентов в зависимости от вида применяемого кодера и заданного режима его работы обеспечивается та или иная степень сжатия.

На рис. 6 в качестве примера показан результат трехкратного применения вейвлет-преобразования к текстовому изображению «Яхта» (рис. 5). Исходное изображение разделяется на ряд компонент. Компонента, расположенная в левой верхней части рисунка, носит название аппроксимации, а три других компонента – матриц вейвлет-коэффициентов (деталей). Это разделение исходного изображения на ряд компонент можно рассматривать как разделение его на ряд контекстов. Получающиеся при этом контексты имеют различные статистические характеристики. Так, аппроксимация имеет несимметричное распределение плотности вероятности величин коэффициентов, все коэффициенты этого контекста положительны, а их величины в среднем убывают с увеличением номеров индексов. Распределения других контекстов являются симметричными и хорошо аппроксимируются законом Лапласа. Кроме того, дисперсии вейвлет-коэффициентов других контекстов в среднем убывают по мере перехода к более высоким частотам. Используя для сжатия каждого из полученных контекстов свой энтропийный кодер, настроенный на статистические характеристики кодируемого контекста, можно получить большую величину сжатия всего изображения в целом по сравнению со сжатием изображения одним кодером без его предварительного вейвлет-преобразования. Собственно, описанный принцип и лежит в основе режима сжатия без потерь, реализуемого в JPEG2000. Величина сжатия без потерь, обеспечиваемая в этом случае, составляет около 4,8 раза при сжатии тестового изображения «Barbara» [4]. Сжатие изображений без потерь путем их предварительного разделения на контексты в результате вейвлет-преобразования с последующим применением энтропийных кодеров, как это следует из изложенного, может найти применение для записи исходного материала, получаемого при видеосъемке.

До сих пор разделение на контексты сжимаемых изображений рассматривалось нами с позиций обеспечения «хорошей» статистики отсчетов, кодируемых в дальнейшем энтропийными кодерами. Однако разделение сжимаемого изображения на контексты целесообразно и при использовании кодеров, вносящих потери. Примером этому может служить сжатие с потерями в рассмотренном выше архиваторе JPEG2000. В данном случае разделение сжимаемого изображения на контексты в результате вейвлет-преобразования преследует иные цели. Дело в том, что, как известно, в зрительной системе при передаче визуальной информации от сетчатки глаз в высшие отделы зрительной системы имеет место ее разделение на ряд частотных каналов, полосы пропускания которых находятся в октавных соотношениях, подобно тому, как это реализуется при рекурсивном применении вейвлет-преобразования. Другими словами, при этом имеет место субполосное разделение передаваемого сигнала [5, 6].

Примечательно, что зрительная система имеет различную чувствительность к сигналам, поступающим в ее высшие отделы по различным каналам. Например, она менее всего чувствительна к сигналам, спектр которых лежит в области верхних пространственных частот и, соответственно, поступает в зрительные центры по «высокочастотным» каналам. Это свойство зрительной системы широко используется при сжатии изображений, благодаря чему высокочастотные вейвлет-коэффициенты в JPEG2000, а также высокочастотные коэффициенты дискретного косинусного преобразования в JPEG, оказалось возможным квантовать на меньшее число уровней по сравнению с низкочастотными коэффициентами и за счет этого осуществлять сжатие изображения. Таким образом, в данном случае разделение изображения на контексты позволило согласовать уровень допустимых искажений, вносимых в изображение кодером, со свойствами зрительной системы человека.

Хорошей иллюстрацией к изложенному подходу может служить разработанный нами метод препроецирования изображения, направленный на увеличение его сжатия энтропийным кодером. В этом методе перед сжатием изображение разделяется на контексты, различающиеся между собой размерами и контрастом мелких деталей, которые в них записываются. После этого в зависимости от значимости деталей данного размера для зрительной системы человека эти контексты кодируются различными кодерами с различной степенью точности, при этом часть фрагментов может быть вообще устранена из изображений. В простейшем случае реализации этого алгоритма, описанном в работе [7], из изображения устраняются мелкие детали, контраст которых меньше порогового контраста зрения для деталей данного размера. При этом искажения, вносимые в изображение, для зрителя остаются незаметными, а вероятность скачков яркости в обработанном таким образом изображении заметно уменьшается, благодаря чему увеличивается степень его компрессии энтропийным кодером. В работе [7] показано, что применение описанного метода позволяет существенно увеличить степень сжатия изображений. Так, степень сжатия изображений «Яхта» (см. рис. 5) арифметическим кодером за счет разделения на контексты и устранения контекстов, которые не воспринимаются зрителем, была увеличена с 1,37 до 2,10, т. е. в 1,53 раза.

Важной особенностью описанного препроецирования является простота его реализации, а также то, что при его использовании совместно с энтропийным кодером не происходит накопления ошибки преобразования при повторных циклах сжатия-восстановления изображений.

Заключение

Основная проблема разделения изображения на контексты при его сжатии, как показали проведенные нами исследования, заключается в определении границ, разделяющих эти контексты, причем это разделение может быть как в спектральной области, так и в пространстве изображений. Иногда решение может быть простым, как, например, в рассмотренных выше случаях. Однако, как правило, форма границ, разделяющих различные контексты, носит более сложный характер, а сами переходы от одного контекста к другому являются размытыми. Более того, контексты могут формироваться по семантическому признаку, как это было показано выше, когда к одному контексту были

отнесены отрезки прямых линий, к другому – экспериментальные точки, к третьему – сетка и т. д.

Для достижения высокой степени сжатия целесообразно в качестве контекстов принять образы объектов, составляющих окружающий нас мир. При этом экземпляры этих объектов следует предварительно записать в память кодера-декодера. В рассматриваемом случае сжатие изображения будет включать в себя следующие шаги:

1) разделение сжимаемого изображения на контексты-образы;

2) совмещение контекстов-образов с экземплярами хранящихся в памяти кодера-декодера объектов путем применения к этим экземплярам аффинных преобразований (возможно, к их фрагментам), подобно тому как это делается при фрактальном методе сжатия;

3) собственно кодирование коэффициентов аффинного преобразования с указанием кодируемого образа.

Разделение сжимаемых изображений на контексты особенно эффективно для применения к цветным динамическим изображениям, например, в телевидении, при разделении изображений на неподвижный фон и движущиеся объекты, которые при высокой скорости движения могут передаваться с пониженным разрешением.

Большой интерес может также представлять разделение на контексты для увеличения сжатия телевизионных изображений при переходе от передачи одного плана к передаче другого плана, которые, в частности, могут сильно различаться цветовой тональностью. В этом случае первые кадры нового плана могут быть переданы также с пониженным разрешением и это, как показали исследования, будет незамечено зрителями. Степень допустимого понижения разрешения в обоих случаях может быть предварительно оценена исходя из известных зависимостей коэффициента эффективности зрительной системы, полученных в экспериментах с динамическими изображениями [8].

Работа финансируется из гранта РФФИ № 01-06-80297.

Литература

1. Ватолин Д., Ратушняк А., Смирнов М., Юкин В. Методы сжатия данных. – М.: Диалог – МИФИ, 2002. – 381 с.
2. Володин А. Б. Адаптивная контекстная компрессия изображений на основе нечеткой морфологии. – СПб.: Вестник молодых ученых, 2004.
3. Воробьев В. И., Грибунин В. Г. Теория и практика вейвлет-преобразования. – СПб.: Изд-во ВУС, 1999. – 204 с.
4. Marcellin M. W., Gormish M. J., Bilgin A., Boilek M. P. An Overview of JPEG2000 // Proceedings of IEEE. Data Compression Conference, 2000. – P. 523–541.
5. Глезер В. Д. Зрение и мышление. – Л.: Наука, 1985. – 246 с.
6. Бондарко В. М., Данилова М. В., Красильников Н. Н. и др. Пространственное зрение. – СПб.: Наука, 1999. – 218 с.
7. Красильников Н. Н., Волошина Н. В. Предварительная логическая фильтрация изображений для увеличения степени их сжатия энтропийными кодерами // Оптический журнал. 2003. – Т. 70. – № 5. – С. 39–43.
8. Красильников Н. Н., Шелепин Ю. Е., Красильникова О. И. Исследование эффективности зрительной системы человека при опознавании движущихся объектов // Физиология человека. – 2003. – Т. 29. – № 4. – С. 22–29.