



УДК 621.327

ОЦЕНКА КАЧЕСТВА ДИНАМИЧЕСКИХ ИЗОБРАЖЕНИЙ ПРИ ИХ СЖАТИИ

Алатар А.И., mih01@mail.ru, Михайлов А.А., mih01@mail.ru

Южно-Российский государственный политехнический университет (НПИ)
имени М.И. Платова, г. Новочеркасск

В статье обоснован выбор критерия качества сжатого динамического изображения, что позволило осуществить целенаправленное управление сжатием динамического изображения. При этом сравнены сжатие посредством квантования и дискретизации, кодирования с предсказанием и *wavelet* сжатием. Исследованы алгоритмы *wavelet* сжатия путем разложения сигнала на два компонента равного размера: среднего (аппроксимация (*approximation*)) подсигнала (*subsignal*) и различие (*difference*) (деталь (*detail*)).

Ключевые слова: критерии качества сжатого изображения, сжатие посредством квантования и дискретизации, кодирование с предсказанием, алгоритмы *wavelet* сжатия, среднее подсигнала, различие детали.

ASSESSMENT OF THE QUALITY FOR COMPRESSED DYNAMIC IMAGES

Alatar A. I., Mikhaylov A.A

Platov South-Russian State Polytechnic University (NPI), Novocherkassk

This article explains the choice of criterion of quality of compressed dynamic image, which allowed to implement a targeted control to compress the dynamic image. At the same time, compression by means of quantization and discretization, coding with prediction and wavelet compression are compared. Investigated algorithms of wavelet compression decomposition of a signal into two components of equal size: approximation subsignal and the difference (detail).

Keywords: criteria of quality of compressed image, compression by quantization and sampling, coding with prediction, algorithms for wavelet compression, approximation subsignal, difference detail.

Если файл изображения занимает много места на диске, то его размеры можно уменьшить, не изменяя размер изображения. Наилучшее сжатие достигается форматом *JPG*, но даже если изображение уже в этом формате, его можно уменьшить, так как у формата *JPG* есть адаптивная схема сжатия, позволяющая сохранять изображение с разной степенью сжатия. Однако чем меньше места занимает изображение, тем больше деталей теряется при сжатии и повторные сохранения в формате *JPG* также приводят к ухудшению качества изображения.

Идея, лежащая в основе всех алгоритмов сжатия с потерями, довольно проста: на первом этапе удалить несущественную информацию, а на втором этапе к оставшимся данным применить наиболее подходящий алгоритм сжатия без потерь. Основные сложности заключаются в выделении этой несущественной информации. Подходы здесь существенно различаются в зависимости от типа сжимаемых данных. Для видеоданных кодируют только движущиеся объекты, а незначительные изменения на неподвижных объектах просто отбрасывают.

Существуют два основных класса алгоритмов:

1. Сжатие без потерь *A* (*lossless compression*), если существует такой обратный алгоритм A^{-1} , что для любого h -изображения $A[h]=h_1$ имеем $A^{-1}[h_1]=h$. Сжатие без потерь применяется в таких графических форматах представления изображений, как: *GIF*, *PCX*, *PNG*, *TGA*, *TIFF*, и применяется при обработке особо ценной первичной информации, когда даже малейшие искажения нежелательны



2. Сжатие с потерями (*lossy compression*), если оно не обеспечивает возможность точного восстановления исходного изображения. Парный к A алгоритм примерного восстановления изображения будем обозначать как A^* . Пара (A, A^*) подбирается так, чтобы обеспечить большие коэффициенты сжатия при сохранении визуального качества. Сжатие с потерями применяется в графических форматах: *JPEG*, *JPEG2000* и т.д.

Все алгоритмы и утверждения относятся как к изображениям, так и к произвольным последовательностям, элементы которых могут принимать конечное количество значений. При этом следует учитывать, что не существует идеальных алгоритмов, сжимающих без потерь любой набор данных.

Целью статьи является сравнение сжатия динамического изображения посредством квантования и дискретизации, кодирования с предсказанием и *wavelet* сжатием.

Критерии качества сжатого изображения. Главным критерием качества сжатого изображения является визуальная оценка изображения для количественной оценки изменения в сжатом изображении. Самый простой способ оценки – это вычисление непосредственной разности сжатого и исходного изображений. Примем $p_{i,j}$ как пиксел, находящийся на пересечении i -ой строки и j -ого столбца исходного изображения, а под $\bar{p}_{i,j}$ – соответствующий пиксел сжатого изображения. То для любого пиксела ошибка кодирования $e(i, j) = \bar{p}_{i,j} - p_{i,j}$, а для всего изображения, состоящего из N строк и M столбцов, суммарная ошибка

$$\sum_{i=1}^N \sum_{j=1}^M e(i, j) = \sum_{i=1}^N \sum_{j=1}^M \bar{p}_{i,j} - p_{i,j}.$$

Чем больше суммарная ошибка, тем сильнее искажения на сжатом изображении. Однако чаще применяется оценка с использованием среднеквадратичного отклонения

$$\frac{1}{N \times M} \sqrt{\sum_{i=1}^N \sum_{j=1}^M (\bar{p}_{i,j} - p_{i,j})^2}.$$

Другой подход заключается в следующем: пиксели итогового изображения рассматриваются как сумма пикселей исходного изображения и шума. Критерием качества при таком подходе называют величину отношения сигнал-шум:

$$SNR = \sqrt{\frac{\sum_{i=1}^N \sum_{j=1}^M \bar{p}_{i,j}^2}{\sum_{i=1}^N \sum_{j=1}^M (\bar{p}_{i,j} - p_{i,j})^2}}.$$

Обе эти оценки являются объективными критериями верности воспроизведения, т.к. зависят исключительно от исходного и сжатого изображения. Тем не менее, эти критерии не всегда соответствуют субъективным оценкам. Если изображения предназначены для восприятия человеком, единственное, что можно утверждать: плохие показатели объективных критериев чаще всего соответствуют низким субъективным оценкам, в то же время хорошие показатели объективных критериев вовсе не гарантируют высоких субъективных оценок.

Сжатие посредством квантования и дискретизации. С процессом квантования и дискретизации связано понятие визуальной избыточности. Значительная часть



информации на изображении не может быть воспринята человеком: например, человек способен замечать незначительные перепады яркости, но гораздо менее чувствителен к цветности. Также, начиная с определённого момента, повышение точности дискретизации не влияет на визуальное восприятие изображения. Таким образом, некоторая часть информации может быть удалена без ухудшения визуального качества. Такую информацию называют визуально избыточной [1].

Самым простым способом удаления визуальной избыточности является уменьшение точности дискретизации, но на практике этот способ можно применять только для изображений с простой структурой, т.к. искажения, возникающие на сложных изображениях, слишком заметны.

Для удаления избыточной информации чаще уменьшают точность квантования, но нельзя уменьшать её бездумно, т.к. это приводит к резкому ухудшению качества изображения.

Наиболее простой способ удаления избыточной информации является равномерное квантование, но, если важно более точно сохранить цветопередачу, можно использовать один из следующих подходов: либо использовать равномерное квантование, но значением закодированной яркости выбирать не середину отрезка, а математическое ожидание яркости на этом отрезке, либо использовать неравномерное разбиение всего диапазона яркостей.

При этом на сжатых изображениях возникают отчётливые ложные контуры, которые значительно ухудшают визуальное восприятие. Существуют методы, основанные на переносе ошибки квантования в следующий пиксел, позволяющие значительно уменьшить или даже совсем удалить эти контуры, но они приводят к зашумлению изображения и появлению зернистости. Перечисленные недостатки сильно ограничивают прямое применение квантования для сжатия изображений.

Большинство современных методов удаления визуальной избыточной информации используют сведения об различной чувствительности человеческого глаза к информации о цветности и яркости изображения.

Несмотря на простоту описанных методов, в чистом виде они применяются редко, чаще всего они служат одним из шагов более эффективных алгоритмов.

Кодирование с предсказанием. Кодирование с предсказанием весьма эффективный метод сжатия информации без потерь. Если совместить кодирование с предсказанием и сжатие посредством квантования, получится весьма простой и эффективный алгоритм сжатия изображения с потерями [2]. В рассматриваемом методе будет квантоваться ошибка предсказания. Именно точность её квантования будет определять как степень сжатия, так и искажения, вносимые в сжатое изображение. Выбор оптимальных алгоритмов для предсказания и для квантования - довольно сложная задача. На практике широко используют следующие универсальные (обеспечивающие приемлемое качество на большинстве изображений) предсказатели:

$$\begin{aligned}\hat{p}_{i,j} &= 0,97 \cdot p_{i,j-1}, \\ \hat{p}_{i,j} &= 0,5 \cdot p_{i-1,j} + 0,5 \cdot p_{i,j-1}, \\ \hat{p}_{i,j} &= 0,75 \cdot p_{i-1,j} + 0,75 \cdot p_{i,j-1} - 0,5 \cdot p_{i-1,j-1}.\end{aligned}$$

Самым простым и в тоже время весьма популярным способом кодирования с потерями является дельта-модуляцию. Этот алгоритм использует предсказание



на основе одного предыдущего пиксела, т.е. $\hat{p}_k = \alpha \cdot p_{k-1}$. Полученная после этапа предсказания ошибка квантуется следующим образом:

$$e_k = \begin{cases} \xi, & e_k > 0, \\ -\xi, & e_k \leq 0. \end{cases}$$

Данный способ квантования имеет неоспоримое преимущество: результат предсказания может быть закодирован единственным битом.

Наиболее заметны два вида искажений – размытие контуров и некая зернистость изображения. Это наиболее типичные искажения, возникающие из-за перегрузки по крутизне и из-за так называемого шума гранулярности. Такие искажения характерны для всех вариантов кодирования с потерями с предсказанием, но именно на примере дельта-модуляции они видны лучше всего.

Шум гранулярности возникает в основном на однотонных областях, когда значение ξ слишком велико для корректного отображения малых колебаний яркости (или их отсутствия).

Ситуация перегрузки по крутизне принципиально отличается от ситуации с шумом гранулярности. В этом случае величина ξ оказывается слишком малой для передачи резкого перепада яркости. Из-за того, что яркость закодированного изображения не может расти также быстро, как яркость исходного изображения, возникает заметная размытость контуров.

Легко заметить, что шум гранулярности уменьшается вместе с уменьшением ξ , но вместе с этим растут искажения из-за перегрузки по крутизне и наоборот. Это приводит к необходимости оптимизации величины ξ . Для большинства изображений рекомендуется выбирать $\xi \in [5; 7]$.

Вейвлет сжатие. Вейвлеты – математические функции, предназначенные для анализа частотных компонент данных [3 – 5]. В задачах сжатия информации вейвлеты используются сравнительно недавно, тем не менее исследователям удалось достичь впечатляющих результатов. Вейвлеты не требуют предварительного разбиения исходного изображения на блоки, а могут применяться к изображению в целом. Наиболее простым и базисным вейвлет – преобразованием является преобразование Хаара [6]. Пусть имеется одномерный дискретный сигнал $f(f_1, f_2, \dots, f_n)$. Преобразование Хаара разлагает каждый сигнал на два компонента равного размера. Первый из компонентов называется средним или аппроксимацией (*approximation*), а второй известен как различие (*difference*) или деталь (*detail*). Точная формула для среднего значения подсигнала (*subsignal*), $a^1 = (a_1, a_2, \dots, a_{N/2})$, на первом уровне для одного сигнала длины N , т. е. $f(f_1, f_2, \dots, f_n)$ имеет вид

$$a_n = \frac{f_{2n-1} + f_{2n}}{\sqrt{2}}, \quad n = 1, 2, 3, \dots, N/2,$$

и детализирующий подсигнал, $d^1 = (d_1, d_2, \dots, d_{N/2})$, на этом же уровне представляется как

$$d_n = \frac{f_{2n-1} - f_{2n}}{\sqrt{2}}, \quad n = 1, 2, 3, \dots, N/2.$$

Сигнал распадается на две составляющие: приближенное значение исходного (с уменьшенным в два раза разрешением) и уточняющую информацию.



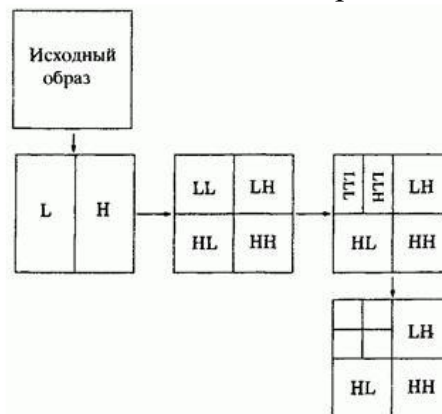
Это преобразование обратимо: т.е. из пары чисел можно легко восстановить исходную пару

$$\begin{aligned}f_{2n-1} &= a_n + d_n \\ f_{2n} &= a_n - d_n.\end{aligned}$$

Двумерное *wavelet* – преобразование Хаара (2D) – простая композиция одномерных преобразований [6]. 2D получается в результате применения одномерного *wavelet* – преобразования (1D) последовательно к строкам и столбцам изображения. Если исходные данные представлены в виде матрицы, то сначала выполняется преобразование для каждой строки, а затем для полученных матриц выполняется преобразование для каждого столбца.

Эти значения формируют два новых сигнала $a\{a_n\} \ n \in \mathbb{Z}$ и $d\{d_n\} \ n \in \mathbb{Z}$, один из которых является огрубленной версией исходного сигнала (каждой паре элементов f соответствует их среднее арифметическое), а другой содержит детализирующую информацию, необходимую для восстановления исходного сигнала. К сигналу a можно применить аналогичную операцию и также получить два сигнала, один из которых является огрубленной версией a , а другой содержит детализирующую информацию, необходимую для восстановления a .

При этом формируются области A , H , V и D с размерностью (число строк/2) (число столбцов/2). A (область приближения) включает информацию о глобальных свойствах проанализированного изображения. Удаление спектральных коэффициентов от этой области приводит к самому большому искажению в исходном изображении. H (горизонтальная область), включает информацию о вертикальных строках скрытых в изображении. Удаление спектральных коэффициентов от этой области исключает горизонтальные детали из исходного изображения. V (вертикальная область) содержит информацию о горизонтальных строках скрытых в изображении. Удаление спектральных коэффициентов от этой области устраняет вертикальные детали из исходного изображения. D (диагональная область) охватывает информацию о диагональных деталях скрытых в изображении. Удаление спектральных коэффициентов от этой области приводит к минимальному иска-



жению в исходном изображении. ПХ подходит для приложения, когда матрица изображения имеет число строк и столбцов, кратное числу 2.

Пирамидальное разложение (рис. 1) вычисляет *wavelet* – преобразование, применяя итерации поочередно к строкам и столбцам. На первом шаге вычисляются полусуммы и полуразности для всех строк (только одна итерация, а не все *wavelet* – преобразование). Это действие производит средние в левой половине



матрицы и полуразности – в правой половине. На втором шаге вычисляются полусуммы и полуразности для всех столбцов получившейся матрицы. В итоге в левом верхнем квадранте будут стоять средние четырех квадрантов исходного образа, а в остальных квадрантах будут находиться соответствующие полуразности.

Рис. 1. Пирамидальное разложение образа

Шаги 3 и 4 оперируют со строками и столбцами, в результате чего средние величины будут сконцентрированы в левой верхней подматрице (одной шестнадцатой всей исходной таблицы). Эти пары шагов применяются к все более и более маленьким подматрицам, до тех пор пока в верхнем левом углу не будет стоять среднее всей исходной матрицы, а все остальные пиксели преобразуются в разности в соответствии с ходом алгоритма.

Wavelet – преобразования Хаара являются поддиапазонными. Они разбивают образ на подобласти, из которых одна область содержит большие числа (средние значения в случае преобразования Хаара), а другие области состоят из малых чисел (разностей в нашем случае). Однако эти области, называемые поддиапазонами, не просто являются семействами больших и малых чисел. Они отражают различные геометрические свойства трансформируемого образа.

Нижний правый поддиапазон обозначим *HH*, на котором отражаются диагональные особенности образа (в нашем случае отсутствующие). Самым интересным остается верхний левый поддиапазон, целиком состоящий из средних величин (он обозначается *LL*). Этот квадрант, являющийся уменьшенной копией исходного образа с пониженным качеством, содержит следы от обеих линий.

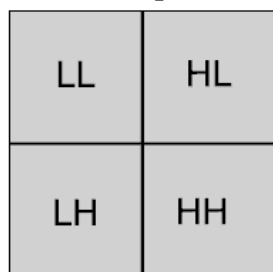


Рис. 2. Однократное применение двумерного *wavelet* – преобразования пирамидального разложения

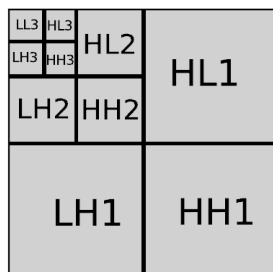


Рис. 3. Трёхкратное применение двумерного *wavelet* – преобразования пирамидального разложения

Двумерное дискретное *wavelet* – преобразование состоит из поочередного одномерного *wavelet* – преобразования строк и столбцов этой матрицы.

Сначала выполняются одномерные *wavelet* – преобразования каждой строки в отдельности, преобразованная строка записывается на прежнее место.

Далее *wavelet* – преобразования применяются ко всем столбцам. В результате изображение разбивается на четыре равные части (рис. 2) стандарт-

ные квадранты преобразованного изображения: *LL*, *LH*, *HL*, *HH*. Квадрант *LL* соответствует низкочастотным *wavelet* – коэффициентам, *HH* – высокочастотным *wavelet* – коэффициентам (*L*–Low, *H*–High) [7].

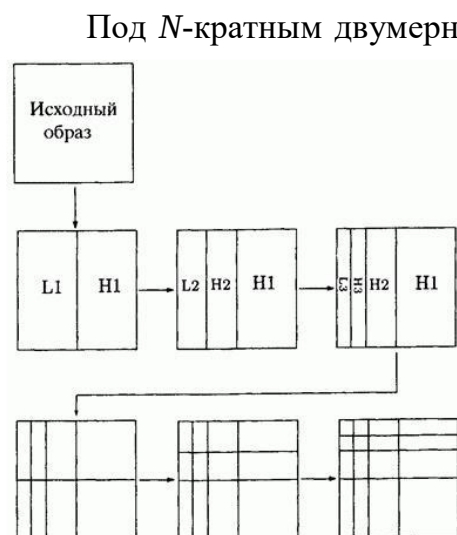


Рис. 4. Стандартное вейвлетное разложение

Под N -кратным двумерным *wavelet* -преобразованием понимается применение N раз двумерного *wavelet* -преобразования, причём очередное двумерное *wavelet* - преобразование применяется к младшей четверти матрицы (квадрант *LL* на рис. 2). В итоге N -кратное преобразование получаем стандартные квадранты изображения (рис. 3 при $N=3$). Квадранты N -кратного двумерного *wavelet* – преобразования имеют аналогичное обозначение.

Стандартное разложение (рис. 4) начинается вычислением *wavelet* – преобразований всех строк изображения. К каждой строке применяются все итерации процесса, до тех пор, пока самый левый элемент каждой строки не станет равен среднему значению чисел этой строки, а все остальные элементы будут равны

взвешенным разностям. Получится образ, в первом столбце которого стоит среднее столбцов исходного образа. После этого стандартный алгоритм производит *wavelet* – преобразование каждого столбца. В результате получится двумерный массив, в котором самый левый верхний угловой элемент равен среднему всего исходного массива. Остальные элементы верхней строки будут равны средним взвешенным разностям, ниже стоят разности средних, а все остальные пиксели преобразуются в соответствующие разности.

Представление изображения с помощью вейвлетов позволяет в процессе восстановления и само восстановленное изображение после удаления из уточняющих матриц малых по модулю коэффициентов добиваться эффективного сжатия, сохраняя при этом визуальное качество изображения. Несмотря на простоту, преобразование Хаара сравнительно редко используется на практике, т.к. существуют другие вейвлеты, обладающие рядом преимуществ (например, вейвлеты Добеши или биортогональные вейвлеты).

Список цитируемой литературы

1. Gonzalez R.C., Woods R.E. Digital image processing. – Boston: Addison – Wesley, 2001. – 813 p.
2. Pratt W.K. Digital Image Processing. – N.Y.: Wiley Interscience, 2001. – 738 p.
3. Чуи Ч. Введение в вейвлеты. – М.: Мир, 2001. – 412 с.
4. Яковлев А.Н. Введение в вейвлет-преобразования. – Новосибирск: НГТУ, 2003. – 104 с.
5. Mallat S.A Wavelet Tour of Signal Processing.–N.Y.: Academic Press, 1999.–851 p.
6. Anuj B., Rashid A. Image compression using modified fast Haar wavelet transform // World Applied Sciences Journal. – 2009. – V. 7. – № 5. – P. 647–653.
7. Шокуров А.В., Михалёв А.В. Оптимальное использование вейвлет – компонент // Успехи математических наук.–2007.–Т. 62.– № 4.– С. 171–172.