

Сжатие данных портретных изображений

(Розділ 5)

Аннотация. Предложены и проанализированы методы и средства сжатия портретных изображений на основе выделения объектов и фона и кодирования их с разной степенью качества. Предложенный метод разделяет элементы изображения на два класса при помощи вейвлет-преобразования изображения и выделения контуров в пространстве коэффициентов высокочастотных областей многомасштабного анализа. Обработка выделенных областей осуществляется при помощи JPEG-технологий на базе дискретного косинусного преобразования, а также с помощью многомасштабного вейвлет-анализа.

Ключевые слова: сжатие изображений, выделение объектов и фона, вейвлет-преобразования, битовая плоскость, JPEG-технологии.

Abstract. Proposed and analyzed the methods and means of the portrait image compression based on the selection of objects and background, and encoding them with varying degrees of quality. The proposed method separates the pixels into two classes using wavelet transform and image edge enhancement in areas of high coefficient space multiresolution analysis. Processing selections made using JPEG-based technologies discrete cosine transform and using multiresolution wavelet analysis.

Keywords: image compression, the allocation of objects and the background, the wavelet transformation, the bit plane, JPEG-technology.

Введение и постановка задачи. Термин “сжатие данных” означает уменьшение объема данных, используемого для представления определенного количества информации. При этом между понятиями данные и информация должны быть проведены четкие различия. Они не являются синонимами. Данные фактически являются тем средством, с помощью которых информация передается или хранится, и для представления одного и того же количества информации может быть использовано различное количество данных. Если данные несут несущественную либо повторяющуюся информацию, то говорят, что эти данные избыточны. Избыточность данных является центральным понятием цифрового сжатия изображений, которые используются практически во всех технологических и культурных областях человеческой деятельности.

Методы цифровой обработки изображений решают две основные задачи. Это повышение качества изображения для улучшения его визуального восприятия человеком и обработка изображений для их эффективного хранения, передачи, анализа и поиска в глобальной сети интернет, а также в системах компьютерного зрения, в том числе биометрических системах идентификации личности по изображениям лица и другим параметрам.

Изображения наряду с большой информационной емкостью (информативностью) обладают и большой информационной избыточностью. Поэтому одна из основных проблем обработки заключается в извлечении из исходных изображений лишь необходимой информации, т.е. в ее сжатии путем устранения избыточности. Определим следующие виды избыточности изображений: кодовая избыточность, межэлементная избыточность (статистическая избыточность), психовизуальная избыточность, структурная избыточность. Сжатие данных достигается в том случае, когда сокращается или устраняется избыточность одного или нескольких из вышеуказанных видов. Данные, в которых нет избыточности-сжать нельзя.

Возникновение кодовой избыточности можно пояснить на основе теоретико-информационного понятия энтропии источника, которое ввел Клод Шеннон в созданной им теории. Согласно этой теории среднее число бит(энтропия), необходимое для кодирования символов источника сообщений, должно быть согласовано с вероятностями появления этих символов. Более вероятным значениям символов присваиваются кодовые слова с меньшим числом бит. Такой подход называют неравномерным кодированием. Физической природой возникновения кодовой избыточности в этом случае является то, что изображения, как правило, состоят из объектов, имеющих регулярную, в некотором смысле предсказываемую морфологию (форму) и отражательные свойства поверхности. Прямым следствием этого является тот факт, что на большинстве изображений определенные значения яркости оказываются более вероятными, чем другие (т.е. гистограммы большинства изображений не являются равномерными). Обычное двоичное кодирование значений яркости таких изображений присваивает кодовые слова одинаковой длины как более вероятным, так и менее вероятным значениям. В результате и появляется кодовая избыточность. Важным следствием теории информации является тот факт, что энтропия сжатых данных совпадает с энтропией исходного источника. При этом предполагается, что по сжатым данным можно полностью восстановить исходную информацию. Такой подход принято называть сжатием без потерь (энтропийным сжатием).

Причиной возникновения межэлементной избыточности изображений является высокая разрешающая способность дискретного поля изображения], которая реализуется только вдоль контуров (границы резкого изменения цвета); на всех гладких участках изображения она расходуется впустую, т.е. возникает межэлементная избыточность (соседние пиксели окрашены в близкие цвета и сильно коррелированы), что приводит к увеличению объема информации.

При сжатии видеоинформации кроме вероятностно-статистических свойств изображения (кодированная и межэлементная избыточность) весьма важно учитывать и особенности получателя изображений. Зрение человека обладает ограниченными возможностями и характеризуется некоторыми известными (отчасти) отличительными особенностями. Использование конкретных особенностей зрения для сокращения избыточности изображений называется психофизической обработкой, заключающейся в том, что часть информации, переносимой изображением, может быть исключена (потеряна) без соответствующих негативных последствий для восприятия этого изображения человеком. Так высокочастотные и низкочастотные шумы и искажения воспринимаются слабее среднечастотных. Также искажения проявляются в меньшей степени вблизи перепадов яркости на границах объектов (пространственное маскирование).

Процесс зрительного восприятия той или иной сцены меньше всего напоминает количественный анализ яркости или координат цвета отдельных элементов изображения. Наблюдатель скорее стремится отыскать в изображении наиболее важные отличительные характеристики такого типа, как контуры или текстурные области, и образовать из них комбинации, поддающиеся распознаванию. Наличие таких областей определяет структурную избыточность изображений, которая является следствием упорядоченности реальных сцен, состоящих из более или менее однородных областей, которые в конечном счете образуют предметы той или иной сцены. Опыт человеческого восприятия одной и той же сцены в различных графических, черно-белых и цветных вариантах показал, что для сохранения семантики изображения более важным оказывается не тип поля, а форма и взаимное расположение в нем целостных структурных единиц (объектов).

Следовательно, при структурированном (контурно-текстурном) представлении цвет и яркость элементов играют по всей видимости вспомогательную роль, а успех восприятия определяется организацией экономного описания таких элементов, как контур или область. Поэтому есть все основания ожидать, что перспективные более емкие алгоритмы

кодирования будут основываться главным образом на структурах изображения [1,2].

Существует и другой вид избыточности структурного типа, которая определяется семантической природой изображения на основании цели обработки и анализа изображений – содержательная избыточность.

В задачах сжатия изображений устранение содержательной избыточности заключается в выделении объектов и фона и кодировании их с различным визуальным качеством. Это самый сложный и эффективный подход к решению задачи сжатия изображений, где в максимальной степени объединяются и используются методы распознавания образов и эффективного кодирования данных. Сложность такого подхода обусловлена недостаточной формализацией такого понятия как “объект” и отсутствием возможности иметь обязательный список обрабатываемых объектов, каждый из которых задан детерминировано, например, уравнением границ и распределением отсчетов сигнала внутри границ, что является характерным для задач распознавания и классификации.

Содержательная (семантическая) избыточность тесно связана с конкретной решаемой задачей. Для одного и того же изображения она может быть различной в зависимости от целей обработки или получателя информации, поэтому формальное ее измерение затруднительно. В частности, два изображения можно считать семантически (содержательно) эквивалентными, если они позволяют выработать идентичные решения после их анализа человеком или ЭВМ.

Тогда один из возможных подходов к передаче данных с устранением содержательной избыточности может быть такая организация последовательности передаваемых данных, при которой у получателя информации обеспечивается наиболее быстрое принятие решения. Когда решение принято, передачу (или выборку данных из памяти) можно прекратить, что эквивалентно сокращению объема передачи или сжатию данных. Следовательно, первой должна выбираться, передаваться, анализироваться и качественно кодироваться семантически наиболее насыщенная часть данных. Такой подход к кодированию на основе сокращения содержательной избыточности применяется во всех новейших информационных технологиях сжатия изображений [2,3,4].

В течение многих лет доминирующей тенденцией в исследованиях проблемы сжатия данных и изображений являются методы обработки сигналов и методы теории информации, которые используют декоррелирующие свойства

различных линейных преобразований и возможности квантования и энтропийного кодирования коэффициентов этих преобразований.

Развитие этих методов привело к созданию известных JPEG-форматов, которые используют субоптимальные косинусные преобразования Фурье и методы вейвлет-преобразований, позволяющие в 20 и более раз сократить объем графических данных, сохранив при этом хорошее визуальное качество [2]. Повысить количественные и качественные характеристики этих форматов можно в результате сокращения содержательной избыточности изображений. Суть такого подхода заключается в том, что в памяти с высоким качеством хранится только часть изображения, представляющая интерес в данном конкретном применении.

Основной проблемой данного подхода к сжатию изображений остается выделение значимых (информативных) участков изображений (объектов) и участков менее информативных (фона). И хотя решению этой проблемы посвящено большое количество работ [1,2], в которых получены интересные результаты, в целом задача кодирования изображений на основе этих принципов еще далека от своего окончательного решения. Такой же вывод можно сделать и в отношении систем видеоналитики и портретной идентификации, в которых первым и основным этапом обработки является поиск лица на изображениях в условиях изменения освещенности, ракурса или мимики лица [5].

В целях получения приемлемой сложности решения задачи автоматического выделения объектов часто используют поиск на изображении окон (шаблонов), имеющих простую форму и стилизованное содержание. Такой подход использован в системе для нахождения лиц [5]. Модель системы предполагает, что все лица имеют определенный размер и занимают область известной формы, а для подходящей маркировки используется алгоритм кластеризации и большое число изображений-образцов, который требует одинаковые условия освещенности на исходном изображении.

В той же работе [5] описано средство поиска лиц с помощью нейронных сетей, преобразующих окно шаблона во фронтальную проекцию. Однако если внешний вид объекта изменяется, то сравнение с шаблоном становится более сложным, поскольку требуется применить гораздо большее их количество. Системы, которые используют методы корреляционного поиска на мультимасштабной пирамиде сжатых изображений, требуют анализа очень большого объема данных, что приводит к чрезвычайно большим вычислительным затратам.

В цифровых библиотеках изображений средства поиска пытаются оценить семантику на уровне объектов. Такие системы вначале сегментируют изображение, а затем фокусируются на некоторых сегментах. Сегментация здесь является ключевым вопросом, поскольку она определяет какую информацию при сравнении изображений можно отбросить как несущественную. Система BlobWorld построена на этих принципах. Изображения в ней представляются как наборы областей когерентного цвета и текстуры. Однако форма этих областей представлена грубо и дальнейший выбор объекта осуществляется в диалоговом режиме с пользователем в полуавтоматическом режиме. Наиболее продвинутое и эффективное решение в области распознавания лиц и биометрической идентификации опубликовано Национальным институтом стандартов и технологий США (<http://www.frvt.org/FRVT2006/>)

Кодирование изображений с различной степенью качества реализовано в графическом формате DjVu [6]. Основная идея этого формата заключается в отделении на изображении текста (передний план) от фона (задний план) и использовании для каждого из них наиболее подходящего метода сжатия .

Однако применяемый в формате метод мультимасштабной кластеризации не всегда дает ожидаемые результаты при определении, что есть передний план (объект), а что есть фон, а при кодировании полутоновых изображений типа портрет метод DjVu вообще перестает работать и сводится к традиционному алгоритму вейвлет-компрессии, как в JPEG 2000 .

Основная часть. Предлагаемый ниже метод устраняет эти недостатки и разделяет элементы изображения на два класса (объект и фон) при помощи вейвлет-преобразования изображения и выделения контуров в пространстве коэффициентов высокочастотных областей многомасштабного анализа. Конечным результатом выделения объектов на изображении является построение битовой плоскости, которая содержит информацию о расположении объекта и фона. Дальнейшая обработка выделенных областей осуществляется при помощи JPEG-технологий на базе дискретного косинусного преобразования, а также с помощью многомасштабного вейвлет-анализа.

Для анализа алгоритма из библиотеки стандартных изображений <http://www.icsl.ucla.edu>, был выбран файл zelda.bmp, который представлен на рис.1. Параметры исходного изображения: размер 256x256, глубина цвета в градации серого 8 бит. Вейвлет коэффициенты исходного изображения, после одного шага преобразования, представлены на рис. 2.

На практике вейвлет-преобразование проводится с помощью биортогональных вейвлет-базисов по формулам разложения (1) и синтеза (2). Если $\{a_i\}$ – коэффициенты разложения остаточного члена, то $\{b_i\}$ и $\{c_i\}$ – коэффициенты, соответственно нового остаточного члена в следующем масштабе представления

$$b_i = \sum_s h_s a_{2i+s} \quad \text{и} \quad c_i = \sum_s g_s a_{2i+s}. \quad (1)$$

Аналогичным образом, если известны коэффициенты $\{b_i\}$ и $\{c_i\}$, то коэффициенты $\{a_i\}$ можно восстановить в соответствии с выражением

$$a_i = \sum_s h_s b_{i-s} + \sum_s g_s c_{i-s}. \quad (2)$$

Для разложения и восстановления сигнала используют две разные пары квадратурных зеркальных фильтров: пара $\{\tilde{h}, \tilde{g}\}$ при разложении и пара $\{h, g\}$ при восстановлении. Основное уравнение, в данном случае имеет вид

$$\overline{h(\omega)}\tilde{h}(\omega) + \overline{h(\omega + \pi)}\tilde{h}(\omega + \pi) = 2.$$

Для этого равенства существует семейство решений:

$$\tilde{h}_{n,m}(\omega) = \sqrt{2} \left(\frac{1 + e^{i\omega}}{2} \right)^n; \quad h_{n,m}(\omega) = \sqrt{2} \left(\frac{1 + e^{i\omega}}{2} \right)^2 P_m \left(\sin^2 \frac{\omega}{2} \right) e^{-im\omega}, \quad (3)$$

где $P_m(x) = \sum_{s=0}^{m-1} C_{m-1+s}^s x^s$ – многочлен Дебеша.

Для вейвлет преобразования исходного изображения, рис. 1, использовалась пара квадратурных зеркальных фильтров h и \tilde{h} с параметрами $n=1$ и $m=2$, выражение (3), общий вид, которых представлен на рис. 3.



Рис. 1. Исходное изображение.

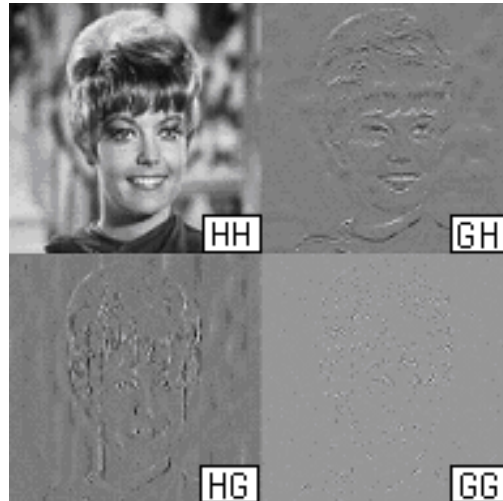


Рис. 2. Вейвлет коэффициенты после одного шага преобразования.

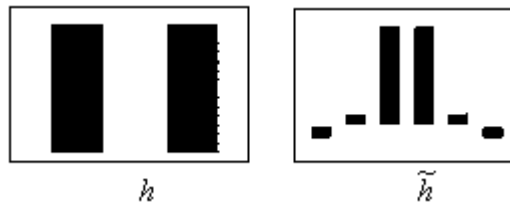


Рис. 3. Пара квадратурных зеркальных фильтров с параметрами $n=1$ и $m=2$.

При построении разделяющей битовой плоскости, использовались вейвлет коэффициенты HG и GH областей (рис. 2), которые были получены в соответствии с выражением (1). После усредненного суммирования выбранных вейвлет-коэффициентов $(HG+GH)/2$ получим изображение, которое представлено на рис. 4.



Рис. 4. Изображение полученное в результате усредненного суммирования вейвлет коэффициентов HG и GH областей.

Полученное изображение имеет более равномерное распределение яркости по сравнению с оригиналом (рис. 1), или остаточным членом разложения NN (рис. 2). Это качество достигается за счет применения

соответствующей пары фильтров h и \tilde{h} при вейвлет преобразовании исследуемого изображения. Исходное изображение раскладывается на “грубую” (НН) и ”тонкую”(GG) структуру, (рис. 2). Области НG и GН имеют более гладкий рельеф яркости на всей плоскости, но при этом сохраняют наиболее существенные ее перепады. Выбор именно этих составляющих (НG и GН) обусловлен тем, что они получены в результате обработки исходного изображения одной парой фильтров h и \tilde{h} , но только в обратной последовательности их применения. При этом, область НG содержит наиболее выраженные перепады яркости в вертикальном направлении, что обусловлено последовательностью операций свертки исходного изображения с соответствующими фильтрами h и \tilde{h} . Область GН наоборот, имеет резкие скачки яркости в горизонтальной плоскости.

Усредненное суммирование НG и GН плоскостей позволяет сохранить резкие изменения яркости, которые обязательно присутствуют в обоих направлениях. Усредненное суммирование НG и GН областей формирует плоскость изображения для пороговой обработки (рис. 4) по формуле

$$X_{i,j} = (HG_{i,j} + GN_{i,j})/2, \quad (4)$$

где $X_{i,j}$ -точка усредненного изображения, $i,j=1..n$, n -количество строк и столбцов в матрице изображения размерностью $n \times n$.

Применение пороговой обработки к изображению представленного на рис. 4, позволяет получить совокупность точек, между которыми значения разности яркости превышает пороговое значение ($|X_i - X_{i+1}| \geq KX_{cp}$), (рис. 5, а). Формирование плоскости $Z_{i,j}$, которая содержит точки превышения порогового значения, (рис. 5, а) согласно правилу

$$Z_{i,j} = \begin{cases} 1, \text{ если } |X_{i,j} - X_{i,j+1}| \geq KX_{cp}, \text{ или } |X_{i,j} - X_{i+1,j}| \geq KX_{cp}; \\ 0, \text{ если } |X_{i,j} - X_{i,j+1}| < KX_{cp}, \text{ или } |X_{i,j} - X_{i+1,j}| < KX_{cp}. \end{cases} \quad (5)$$

где $i,j=1 \dots n-1$, n -количество строк и столбцов в матрице изображения

Пороговое значение привязывалось к среднему значению яркости X_{cp} на всей плоскости изображения (рис. 4), и выбиралось таким образом, чтобы исключить наличие случайных точек, т.е. тех точек, которые не имеют в своем ближайшем окружении себе подобных. Таким образом, производится фильтрация одиночных точек преодолевших пороговое значение. На рисунке (5 а) наглядно видны область с высокой плотностью перепадов яркости и область,

которая практически не содержит смежных точек превышения порога. Пороговая обработка изображения представленного, на рис. 4 проводилась по строкам и столбцам матрицы изображения.

Затем, для отделения объекта от фона, производится последовательное сканирование изображения по строкам и столбцам до выявления первой и последней не нулевой точки. После разделения изображения на области объекта и фона, необходимо создать непрерывную и замкнутую границу между этими областями. На рис. 5 б, представлены граничные точки, которые разделяют области объекта и фона. Эта граница раздела имеет разрывы и не является замкнутой линией. Для формирования непрерывного и замкнутого контура требуется соединить точки разрыва границы между связными фрагментами контура (см. рис. 6). Точкой разрыва линии границы считается любая точка имеющая только одну смежную точку, т. е. одного "соседа". Для соединения точек разрыва линии границы X_j необходимо найти минимальное расстояния между точками разрыва принадлежащим разным подмножествам $\{R_i\}$, т.е. концом связного фрагмента границы и началом другого. После соединения граничных точек между собой получим битовую плоскость разделяющую объект и фон, рис. 5, в. Построение битовой плоскости является завершающим этапом выделения объекта на плоскости изображения. Точки внутри контура считаются принадлежащими объекту и им присваивается значение 1, точки лежащие вне контура считаются фоном и им присваивается значение 0, рис. 5 в.

В случае, когда компрессия изображений осуществляется при помощи JPEG-технологии на основе дискретного косинусного преобразования, то предварительно осуществляется разбиение плоскости изображения на фрагменты размером 8×8 [JPEG]. После одного шага вейвлет преобразования плоскость изображения представленная на рис. 4 уменьшилась в 4 раза по сравнению с плоскостью исходного изображения, рис. 1. Следовательно, размер фрагмента, для битовой плоскости, должен быть 4×4 , что полностью согласуется с разбиением исходного изображения (рис. 1) на фрагменты 8×8 . Каждый фрагмент может принадлежать трем областям: область объекта, область фона, область границы битовой плоскости. Таким образом, фрагментами области объекта можно считать те фрагменты, которые принадлежат, как области самого объекта, так и области границы битовой плоскости. Все остальные сегменты плоскости изображения принадлежат области фона.

На рис. 5, г представлена битовая плоскость разделения изображения после сегментного разбиения на фрагменты 4x4.

В том случае, когда сжатие исходного изображения (рис. 1) производится с использованием многомасштабного вейвлет преобразования, то сегментация битовой плоскости (рис. 5, в) не производится.

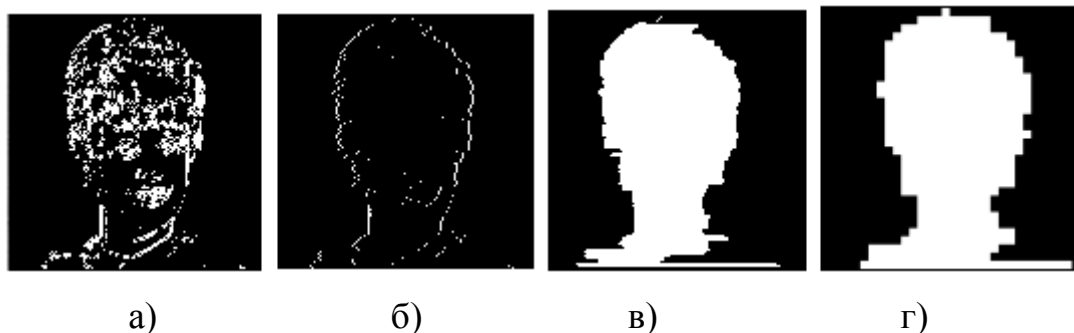


Рис.5 а) Совокупность точек превысивших разностный порог; б) граничные точки разделения объекта и фона; в) битовая плоскость разделения; г) битовая плоскость после сегментации.

Для кодирования изображений с различной степенью точности при использовании JPEG-технологий необходимо исходное изображение (рис. 1) разбить на сегменты размером 8x8 и осуществить кодирование объекта и фона изображения по методу JPEG-технологии .

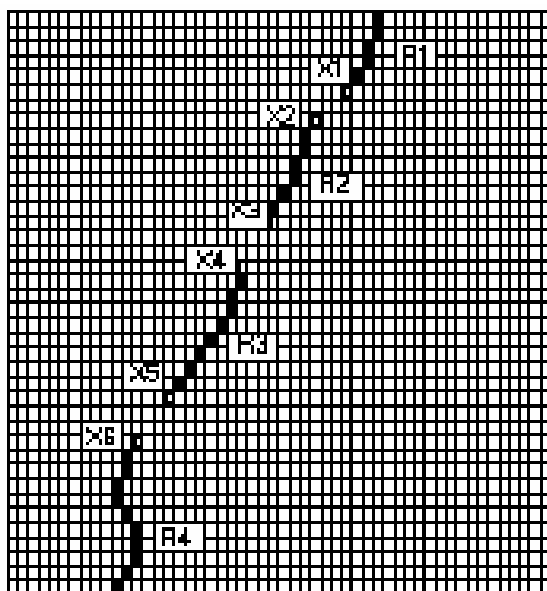


Рис. 6. Связные фрагменты R_i и точки разрыва линии границы X_j между областями объекта и фона.

Согласно метода JPEG-компрессии ядро одномерного косинусного преобразования имеет вид:

$$L_x(0) = \frac{1}{\sqrt{N}} \sum X(m); \quad L_x(k) = \sqrt{\frac{2}{N}} \sum_{m=0}^{N-1} \left(X(m) \cos \frac{(2m+1)k\pi}{2N} \right). \quad (7)$$

В этом выражении $X(m)$ - значение отчетов в строках соответствующих рабочих матриц, полученных на втором шаге JPEG-технологий, $L_x(k)$ – коэффициенты дискретного косинусного преобразования, а значения m и k изменяются соответственно от 0 до $N-1$ и от 1 до $N-1$ с шагом один.

Различная степень точности кодирования объекта и фона определяется количеством косинусных коэффициентов, которые используются в рабочих матрицах объекта и фона. Для фона выберем минимальное значение -1 (имеется ввиду первый коэффициент рабочей матрицы 8×8 в каждом сегменте изображения), а для объекта максимальное -64 коэффициента. Таким образом, сжатие изображения будет осуществляться только за счет сегментов, которые входят в область объекта на битовой плоскости. Область фона будет представлена только одним коэффициентом косинусного преобразования в каждом сегменте области фона. Такой подход позволяет обеспечить максимальное сжатие области фона изображения, а в области объекта обеспечить такую степень компрессии и качества изображения, которая являлась бы приемлемой для решения поставленной задачи.

На рис. 7 представлена зависимость общего коэффициента сжатия всей плоскости изображения ($K_{сж}$) от среднеквадратической ошибки (СКО) на плоскости объекта.



Рис. 7. Зависимость $K_{сж}$ всей плоскости изображения от СКО объекта.

На рис. 8, представлено исходное изображение после кодирования методом JPEG компрессии. Рисунок (8, а) показывает восстановленное изображение после отдельного кодирования объекта и фона с различной

степенью качества, а рис. (8, б) после кодирования всей плоскости изображения с одинаковой степенью качества. СКО на плоскости объекта в обоих вариантах составляет 12%, а увеличение коэффициента сжатия с 21 раза до 26,4 обусловлено большей степенью компрессии области фона (в данном случае К сжатия фона составляет 64 раза, так как только один косинусный коэффициент рабочей матрицы 8x8 представляет каждый сегмент области фона). При JPEG кодировании изображений битовую плоскость разделения необходимо кодировать энтропийными методами и хранить или передавать вместе с самим изображением в сжатом виде, что в незначительной мере влияет на степень компрессии самого изображения.

Кодирование изображений на основе вейвлет преобразований с предварительным выделением объекта имеет некоторые отличия от JPEG кодирования области объекта и фона.

Выделение объекта на плоскости изображения производится таким же образом, как и в случае преобразования в JPEG формат. Однако, после формирования битовой плоскости разделения изображения на объект и фон, (рис. 5, г) применить её необходимо только к вейвлет-коэффициентам представленных в областях HG, GH и GG на каждом шаге разложения.



а) К сжатия =26,4 СКО объекта=12%. б) К сжатия 21, СКО объекта=12%.

Рис. 8. Изображение после отдельного кодирования методом JPEG-технологий.

Сама битовая плоскость, при каждом шаге разложения изображения, подвергается соответствующему масштабированию, которое должно быть согласовано с каждым шагом вейвлет преобразования. Таким образом, битовая плоскость в соответствующем масштабе представления применяется к областям HG, GH и GG на каждом шаге вейвлет преобразования, что позволяет обеспечить сохранность вейвлет-коэффициентов области объекта на участках HG, GH и GG, которые в свою очередь отображают высокочастотную

составляющую исходного изображения. В результате, все вейвлет-коэффициенты, которые принадлежат области объекта на всех этапах разложения сохраняются, а коэффициенты, которые принадлежат области фона в соответствующих масштабах игнорируются. Необходимо заметить, что использование битовой плоскости на остаточном члене разложения (область НН) приведет к потере низкочастотной составляющей области фона, что в свою очередь полностью удалит из исходного изображения текстуру фона. При таком подходе, общий коэффициент сжатия будет более высоким, но на восстановленном изображении текстура фона будет полностью отсутствовать или заменена на абстрактную текстуру, которая не имеет никакого отношения к обрабатываемому изображению. Использование вейвлет преобразования при кодировании изображений с разной степенью точности не требует отдельного кодирования и хранения битовой плоскости вместе с изображением. Битовая плоскость используется только на этапе разложения изображения, при восстановлении вейвлет-коэффициенты будут распределены в обратном порядке Z-сканирования плоскости разложения. Этот факт можно отнести к преимуществу вейвлет кодирования изображений по сравнению с методом JPEG.

На рис.9 представлена зависимость коэффициента сжатия от СКО при кодировании изображения с использованием вейвлет преобразования.



Рис. 9. Зависимость коэффициента сжатия (К сж) от среднеквадратической ошибки (СКО) при вейвлет анализе изображения.

Для отметки СКО =12 % степень сжатия при использовании вейвлет преобразования соответствует 24,5 раза (кривая – Вейвлет). При использовании тех же фильтров разложения и восстановления, но с разделением изображения на объект и фон степень сжатия увеличилась до 28,5 раза при том же значении СКО=12 % (кривая – Вейвлет-объект). Следует заметить, что при аналогичном

значении ошибки восстановления ($СКО=12\%$) степень сжатия при JPEG кодировании изображения с выделением объекта равна 26,4 раза (рис. 7), что превышает соответствующий показатель при обычном вейвлет анализе на 8% (рис. 9, кривая – Вейвлет).

На рис. 10 (а, б) представлены восстановленные изображения полученные после вейвлет кодирования всей плоскости изображения, (рис. 10 а), а так же после раздельного кодирования области фона и объекта, (рис. 10 б), при одинаковом значении ошибки восстановления, $СКО=12\%$.



а) K сжатия =24,5, $СКО=12\%$
(без выделения объекта)

б) K сжатия =28,5, $СКО=12\%$,
(с выделением объекта)

Рис. 10. Восстановленные изображения после выделения объекта и вейвлет-кодирования.

Заключение. Предложенный метод разделения изображения на объект и фон, позволяет кодировать их отдельно, как с различной степенью качества, так и с использованием различных методов. Это дает возможность представлять наиболее информативные области изображения с более высоким качеством, чем остальные его составляющие. Благодаря такому подходу можно сохранить высокое качество информативных областей (объектов), менее информативные области (фон) представить с большей компрессией, но с худшим качеством и таким образом повысить общий коэффициент сжатия для всего изображения.

Предложенный подход позволил выявить резервы классической схемы JPEG-кодирования и получить лучшие количественные характеристики по сравнению с вейвлет-кодированием на основе формата JPEG 2000.

Использование метода выделения объекта в схеме классического JPEG-формата дает более высокое повышение эффективности кодирования, чем выделение объекта в схеме многомасштабного вейвлет преобразования изображения (JPEG 2000). Это объясняется естественной локальной природой

многомасштабного анализа, схема преобразований которого уже априори приспособлена для разделения изображений. Метод выделения объекта(лица) на изображении устойчив к изменениям внешности, степени освещенности, мимики лица и может с успехом использоваться в криминалистических системах портретной идентификации.

Литература

1. Кунт М., Методы кодирования изображений второго поколения / М. Кунт, А. Икономопулос, М. Кошер // ТИИЭР. – 1985. – Т. 73, – № 4. – С. 59-86.
2. Гонсалес Р. Цифровая обработка изображений./ Р. Гонсалес, Р. Вудс. – М.: Техносфера, 2005. – 1072 с.
3. Иванов В.Г. Сокращение содержательной избыточности изображений на основе классификации объектов и фона / В.Г. Иванов, М.Г. Любарский, Ю.В. Ломоносов // Проблемы управления и информатики. – Київ, 2007. – № 3. – С. 93–102.
4. Ричардсон Ян. Видеокодирование. H. 264 и MPEG-4 – стандарты нового поколения / Пер. с англ. – М.: Техносфера, 2005. – 368 с.
5. Компьютерное зрение. Современный подход / Д. Форсайт, Д. Понс; Пер. с англ. – М.: Вильямс, 2004. – 928 с.
6. Technical Papers from AT&T Labs: Электронный ресурс. – Режим доступа: <http://djavuzone.org/techpapers/index.html>.