

В.Г. ИВАНОВ, д-р. техн. наук,
Ю.В. ЛОМОНОСОВ, канд. техн. наук,
М.Г. ЛЮБАРСКИЙ, д-р физ.-мат. наук (г. Харьков)

АНАЛИЗ И КЛАССИФИКАЦИЯ МЕТОДОВ СЖАТИЯ ИЗОБРАЖЕНИЙ

Построена классификация перспективных методов сжатия данных по критериям эффективности и качества сжатия. Отмечается, что одним из наиболее перспективных методов кодирования изображений является агрегативный подход, объединяющий различные методы сокращения избыточности и методы распознавания образов.

Ключевые слова: сжатие изображений, классификация, перспективные методы.

Постановка проблемы. Профессионально-техническая и социально-культурная среда современного человека становится все более электронной, а главной характеристикой этого процесса является огромный объем цифровых данных, который создается, хранится и циркулирует в этой среде [1]. Поскольку значительная часть этих данных является графической, аудио или видеоинформацией, требования к техническим параметрам средств связи и системам хранения становятся чрезвычайно высокими [2, 3]. Поэтому эффективное функционирование и развитие коммуникативно-компьютерных систем хранения, обработки, передачи и поиска мультимедийной информации невозможно без использования методов сжатия, многообразие которых требует их корректного сравнения и классификации [4 – 6].

Анализ литературы. Цитируемая литература в основном состоит из работ, в которых наиболее полно и ясно изложено то или иное направление в области сжатия данных. К сожалению, законченной классификации методов и алгоритмов сжатия данных, отвечающих современному состоянию этой области, нет ни в отечественной, ни в зарубежной литературе [1 – 17].

Цель статьи. Провести анализ и сопоставление методов сжатия данных с целью получения их классификации, взаимосвязи и направлений дальнейшего развития.

Классификация методов сжатия. Избыточность данных является центральным понятием цифрового сжатия данных [7, 8].

В общем случае (см. рис. 1) причины возникновения избыточности данных могут быть следующие [8]:

1. Непреднамеренное завышение размерности вектора сведений $W(t)$, предоставляемых получателю. Оно может быть вызвано неполным учетом свойств получателя, которому, вообще говоря, для правильного восприятия состояний объекта достаточно располагать вектором меньшей размерности. Таким образом, первая причина возникновения избыточности данных связана

с физическими и психофизическими свойствами получателя, семантическими и прагматическими аспектами восприятия информации. Процесс разработки методов сокращения такого вида избыточности трудно формализуем и носит в значительной степени эвристический характер.



Рис. 1. Обобщенная модель информационных технологий со сжатием данных

2. Непреднамеренное завышение числа компонент $\{X_m(t)\}_{m=1, \overline{M}}$ процесса $X(t)$, порождаемого объектом, что приводит, как правило, к повышению размерности сообщения $Y(t)$ и оценок $\bar{Y}(t)$ и $\bar{X}(t)$. Это может быть связано с неоптимальностью методов получения сведений $W(t)$. Следовательно, вторая причина возникновения избыточности вызвана тем, что для решения задач классификации, распознавания образов, идентификации объектов и других, связанных с выработкой решений, объект рассматривается в пространстве состояний завышенной размерности.

3. Проектирование системы ведется на основании осредненных по множеству реализаций процесса $X(t)$ или предельных значений характеристик процесса $X(t)$, например, из расчета максимально возможной ширины его спектра. В этом случае говорят о возникновении естественной (статистической) избыточности сообщений, которая характеризуется наличием корреляционных связей между отдельными компонентами $\{X_m(t)\}_{m=1, \overline{M}}$ процесса $X(t)$, а также тем, что распределение вероятностей передачи этих компонент и

символов сигнала, которыми кодируется каждая компонента, отличаются от равномерного.

С учетом изложенного все методы сжатия данных по классификационному признаку – виду устраняемой избыточности – могут быть разделены предварительно на классы методов, реализующих сжатие данных с учетом позиций 1 – 3. При этом методы уменьшения размерности $W(t)$ и $X(t)$ должны рассматриваться в тесной связи с конкретными задачами. Например такими, как классификация, идентификация, обнаружение и выделение объектов и др. Поскольку эти методы существенно уменьшают объем обрабатываемой в системе информации, то их также необходимо отнести к методам сжатия данных.

В традиционной постановке сжатие данных обычно связывается с устранением третьего вида избыточности, то есть с использованием методов, основанных на теории информации и обобщенной теории дискретных сигналов.

Известно, что изображения наряду с большой информационной емкостью 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. Мы определим следующие виды избыточности изображений, на основании которых и будет осуществлена классификация методов сжатия: кодовая избыточность, межэлементная избыточность (статистическая избыточность), психовизуальная избыточность, структурная избыточность, которая включает в себя контурно-текстурную и содержательную избыточность. Сжатие данных достигается в том случае, когда сокращается или устраняется избыточность одного или нескольких из вышеуказанных видов.

Возникновение кодовой избыточности можно пояснить на основе теоретико-информационного понятия энтропии источника, которое ввел Клод Шеннон в созданной им теории [6]. Если дискретная случайная переменная r_K , распределенная в интервале $[0, 1]$, представляет значение яркости изображения, и каждое значение r_K появляется с вероятностью

$$P_r(r_K) = \frac{n_K}{n}, \quad K = 0, 1, 2, \dots, L - 1,$$

где L – общее число уровней яркостей, n_K – число пикселей, имеющих значение яркостей K ; n – общее число элементов в изображении. Если число битов, используемых для представления каждого из значений r_K , равно $\ell(r_K)$, то среднее число битов (энтропия), требуемых для представления значения одного элемента, равно:

$$L_{cp} = \sum_{K=0}^{L-1} \ell(r_K) P_r(r_K). \quad (1)$$

Если присваивать кодовые слова с меньшим числом бит более вероятным значениям, то можно достичь сжатия данных. Такой подход называют неравномерным кодированием. Когда значения яркости изображения кодируют некоторым способом, требующим большего числа символов, чем это строго необходимо (т.е. код не минимизирует выражение (1)), то говорят, что изображение имеет кодовую избыточность. 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 не обеспечивается минимизация выражения (1) 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) [6, 7]. Можно сказать, что компрессия без потерь является экстремальным случаем сжатия, при котором энтропия данных остается неизменной.

Здесь мы приходим к другой важной проблеме: каково наиболее компактное представление информации, если допускается неточное восстановление сжатых данных. Такое сжатие называется сжатием с частичной потерей информации. Сжатие с потерями по существу предполагает уменьшение энтропии исходной информации на основе удаления статистической (межэлементной), психовизуальной и структурной избыточности.

Как показали результаты измерения некоторых статистических характеристик второго порядка для различных изображений, функция автокорреляции зависит от наличия деталей в изображении, т.е. функция автокорреляция качественно связана со структурой изображения и является убывающей функцией. Примерно так же ведет себя и огибающая энергетического спектра. Отсюда следует, что большая часть энергии видеосигнала сосредоточена на низких частотах [9], иначе говоря, соседние элементы сильно коррелированы. Это отражает другую важную форму избыточности данных, которая напрямую связана с межэлементными связями внутри изображения – межэлементную избыточность.

Причиной возникновения межэлементной избыточности изображений является высокая разрешающая способность дискретного поля изображения, которая реализуется только вдоль контуров; на всех гладких участках

изображения она расходуется впустую, т.е. возникает межэлементная избыточность и увеличивается объем информации.

При сжатии видеoinформации кроме вероятностно-статистических свойств изображения (кодовая и межэлементная избыточность) весьма важно учитывать и особенности получателя изображений. Зрение человека обладает ограниченными возможностями и характеризуется некоторыми известными (отчасти) особенностями [6, 10]. 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.

Установлено, что 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 [11, 12].

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) [12]. 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. Однако необходимо сказать, что несмотря на очевидность такого вывода, большинство систем кодирования и передачи изобразительной информации основаны на устранении избыточности именно в яркости и цвете, а не в форме и положении объектов [6 – 9]. Поэтому есть все основания ожидать, что перспективные более емкие алгоритмы кодирования будут основываться главным образом на структурах изображения [10, 13].

Способы сжатия изображений на основе такого подхода (контурно-текстурной избыточности) заключаются в его сканировании и обнаружении повторяющихся однородных областей. Полученное таким образом описание изображения можно эффективно использовать для решения задач распознавания и классификации, а методы выделения признаков в этих задачах – для сжатия изображений. Таким образом, структурное описание изображений является общей частью решения проблем распознавания, классификации и сжатия изображений.

Существует и другой вид избыточности структурного типа, которая определяется семантической природой изображения на основании цели обработки и анализа изображений – содержательная избыточность [12, 14].

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 [6, 9].

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 одним 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 [15, 16, 17].

На основании вышеизложенного на рис. 2. представлена схема классификации, взаимосвязи и развития методов сжатия, отвечающая современным тенденциям исследования в области кодирования и обработки изображений.

На основании анализа методов сжатия и схемы их классификации, на рис. 3. представлена обобщенная модель технологической среды сжатия данных, которая отражает концептуальный подход к кодированию изображений на основе объединения независимых методов сокращения избыточности и методов распознавания образов.

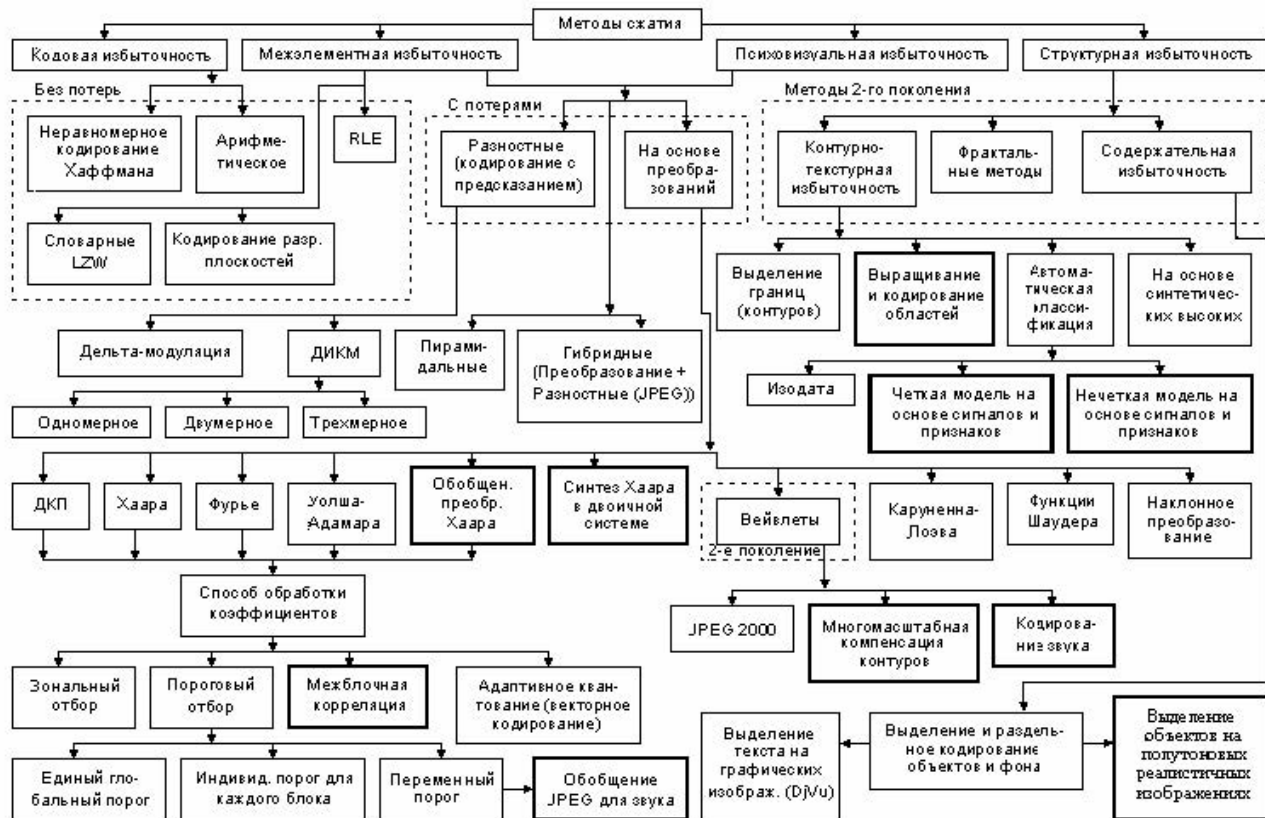


Рис. 2. Схема классификации, взаимосвязи и развития методов сжатия данных

Обобщенная модель технологической среды сжатия данных включает в себя декомпозиционные составляющие моделей разложения сигналов в ряды Фурье, вейвлет-анализа и контурно-текстурной модели на принципах группирования данных (автоматической классификации), позволяющая оставаться при этом в рамках классических методов теории и практики JPEG-сжатия изображений.

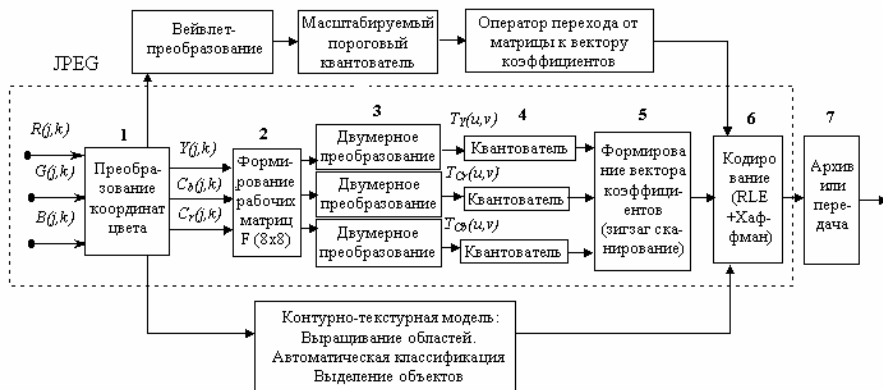


Рис. 3. Обобщенная модель технологической среды сжатия данных

Модель позволяет ослабить ограничения методов теории информации и теории кодирования, которые в недостаточной степени учитывают свойства зрительной системы человека, и в общей схеме кодирования изображений, состоящей из этапа формирования последовательности сообщений, а затем этапа формирования кодовых слов, основной упор делают на втором этапе, состоящем, в основном, из методов устранения межэлементной и кодовой избыточности сообщений. Результаты из области физиологии зрения и последних достижений технологий сжатия [1] дают веские аргументы в пользу применения общей контурно-текстурной модели для обработки и кодирования изображений, которая акцентирована, в первую очередь, на отбор сообщений для кодирования, используя методы распознавания образов, а само кодирование производится классическими методами теории информации [11, 15 – 17].

Выводы. Таким образом проведенный анализ и приведенная схема классификации методов сжатия данных вместе с предложенной обобщенной моделью технологической среды сжатия данных позволяет определить перспективные направления исследований в данной предметной области и выявить резервы как действующих JPEG-форматов сжатия изображений, так и указать новые подходы на основе методов автоматической классификации и

выделения объектов с целью дальнейшего их кодирования с различной степенью качества.

Список литературы: 1. Головных А. Цифровая среда обитания // СНИР. Компьютеры и коммуникации. – К.: Издат. дом "Софт Пресс". – 2003. – № 1. – С. 68-70. 2. Миронов С. Электронные архивы для промышленности // Открытые системы. – 2005. – № 2. – С. 56-60. 3. Новиков С. Передача данных видеонаблюдения по IP-сетям // Открытые системы. – 2003. – № 9. – С. 57-59. 4. Молодчик П. Видеокомпрессия: настоящее и будущее // Компьютерное обозрение. – К.: "Издательский Дом ИТС", 2002. – № 33. – С. 49-51. 5. Д. Мюррей, Уван Райнер. Энциклопедия форматов графических файлов. – К.: Издат. группа ВНУ, 1997. – 672 с. 6. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2005. – 1072 с. 7. Мановцев А.П. Основы теории радиотелеметрии. – М.: Энергия, 1973. – 592 с. 8. Орищенко В.И., Санников В.Г., Свириденко В.А. Сжатие данных в системах сбора и передачи информации. – М.: Радио и связь, 1985. – 184 с. 9. Претт У. Цифровая обработка изображений. В 2-х кн. – Кн. 2. – М.: Мир, 1982. – 480 с. 10. Кунт М., Икономопулос А., Кошер М. Методы кодирования изображений второго поколения // ТИИЭР. – М.: Мир, 1985. – Т. 73. – № 4. – С. 59-86. 11. Завалишин Н.В., Мучник И.Б. Модели зрительного восприятия и алгоритмы анализа изображений. – М.: Наука, 1974. – 344 с. 12. Александров В.В., Горский Н.Д. Представление и обработка изображений: Рекурсивный подход. – М.: Наука, 1985. – 192 с. 13. Сэлмон Д. Сжатие данных, изображений и звука. – М.: Техносфера, 2004. – 368 с. 14. Павлидис Т. Алгоритмы машинной графики и обработки изображений. – М.: Радио и связь, 1986. – 400 с. 15. <http://www.dejavu.research.att.com/djvu/sci/djvuspec>. 16. Ричардсон Ян. Видеокодирование. Н. 264 и MPEG-4 – стандарты нового поколения. – М.: Техносфера, 2005. – 368 с. 17. Иванов В.Г., Любарский М.Г., Ломоносов Ю.В. Сокращение содержательной избыточности изображений на основе классификации объектов и фона // Проблемы управления и информатики. – К., 2007. – № 3. – С. 93-102.

УДК 004.627

Аналіз і класифікація методів стиснення зображень / Иванов В.Г., Ломоносов Ю.В., Любарский М.Г. // Вісник НТУ "ХПІ". Тематичний випуск: Інформатика і моделювання. – Харків: НТУ "ХПІ". – 2008. – № 49. – С. 78 – 86.

Побудована класифікація перспективних методів стиснення даних за критеріями ефективності і якості стиснення. Наголошується, що одним з найбільш перспективних методів кодування зображень є агрегативний підхід, об'єднуючий різні методи скорочення надмірності і методи розпізнавання образів. Лл.: 3. Бібліогр.: 17 назв.

Ключові слова: стиск зображень, класифікація, перспективні методи.

UDC 004.627

Analysis and classification of methods of compression of images / Ivanov V.G., Lomonosov U.V, Lyubarsky M.G. // Herald of the National Technical University "KhPI". Subject issue: Information Science and Modelling. – Kharkov: NTU "KhPI". – 2008. – № 49. – P. 78 – 86.

Classification is built perspective methods of compression of data on the criteria of efficiency and quality of compression. It is marked that one of the most perspective methods of encoding of images is aggregate approach, uniting the different methods of reduction of surplus and methods of recognition of patterns. Figs: 3. Refs: 17 titles.

Key words: compression of images, classification, perspective methods.

Поступила в редакцію 10.10.2008