



УДК 681.3.397

© Ш. С. Фахми, 2010

## РАЗВИТИЕ ТРИАНГУЛЯЦИОННОГО ПОДХОДА КОДИРОВАНИЯ И ДЕКОДИРОВАНИЯ НЕСТАЦИОНАРНЫХ ИЗОБРАЖЕНИЙ

Фахми Ш. С. – канд. техн. наук, ведущий науч. сотр. БЦСП НИИТ, тел.: (812) 234-36-75, e-mail: Shakeebf@mail.ru

Рассмотрен метод кодирования и декодирования нестационарных сигналов, основанный на пирамидально-рекурсивном разбиении исходного изображения. Сжатие выполняется путем поиска опорных точек, а восстановление – применением триангуляции. Рассмотрены структурные схемы подсистем кодирования и декодирования по опорным точкам, а также приведены количественные оценки результатов моделирования алгоритмов.

The method of encoding and decoding non-stationary signals based on pyramidal-recursive splitting of the initial image is considered. Compression is carried out by search of reference points, and restoration by triangulation. Skeleton subsystems diagrams of encoding and decoding on reference points are considered, and also the results of algorithm simulation are qualitatively estimated.

*Ключевые слова:* пирамидально-рекурсивный метод, опорные точки, полигон.

### Введение

Немаловажным и, пожалуй, наиболее перспективным направлением в области сжатия и восстановления полутоновых изображений являются триангуляционные методы, применение которых известно из задач распознавания образов [1, 2] и из задач восстановления двумерных изображений по нерегулярно расположенным опорным точкам (ОТ). Основные преимущества триангуляции заключаются в следующем. Во-первых, триангуляция ОТ естественным образом подстраивается под данные – там, где ОТ разрежены, треугольники крупнее. А там, где есть сгущение – мельче. Число треугольников не превышает удвоенного числа опорных точек. Во-вторых, у прямоугольной сетки для адекватного отображения достаточно изменчивых поверхностей требуется сильно измельчить сетку, что требует больших вычислительных мощностей и ведет к образованию неустойчивости, что отсутствует в случае треугольной сетки.

Развитие и систематизация триангуляционных алгоритмов кодирования и декодирования изображений, предложенные в [3], тесно связаны с блочным

методом проектирования систем на кристалле, когда проектируется комбинация новых сложно-функциональных блоков и блоков многократного использования из библиотек (от разных поставщиков и от собственных проектных групп). Многочисленные стандарты и отсутствие универсального инструмента применения этих блоков делает эту задачу нетривиальной [4]. При проектировании видеосистем, основанных на пирамидально-рекурсивном подходе (ПРП), сложность аппаратной реализации возрастает по мере увеличения степени сжатия и сохранения точности кодирования. Однако быстрое развитие технологии систем на кристалле и повышение плотности размещения вентилях на кристалле делает некоторое возрастание сложности преодолимой проблемой, а развитие триангуляционных методов и разработка алгоритмов кодирования и декодирования по опорным точкам, на основе ПРП более актуальными [3].

Системы распознавания и обработки реалистических объектов изображений должны обеспечивать передачу всех свойств моделируемого объекта: объемность, расположение, передача полутонов, тени, освещение, текстуры поверхности. Чем выше степень реалистичности изображения, тем больше требуется вычислений для его формирования [5].

В качестве подхода для развития триангуляционных методов в данной статье предлагается пирамидально-рекурсивный метод, существенно снижающий вычислительные затраты при поиске опорных точек в процессе кодирования и обеспечивающий эффективный способ построения регулярной триангуляционной сетки при декодировании.

В [6] показано, что наилучшее приближение к энтальпии сигнала изображения обеспечивается разбиением изображений на треугольники.

Это объясняется следующими причинами:

- треугольник является простейшим полигоном, вершины которого однозначно задают грань;
- любую область можно гарантировано разбить на треугольники;
- вычислительная сложность алгоритмов разбиения на треугольники существенно меньше, чем при использовании других полигонов;
- для треугольника легко определить три его ближайших соседа, имеющих с ним общие грани.

#### **Пирамидально-рекурсивный подход**

В отличие от триангуляционных методов регулярность пирамидальных структур предопределяет их удобную реализацию и эффективное использование в системах обработки нестационарных изображений за счет параллельной обработки. Иерархия описаний различной степени общности способствует контекстно-независимой структуризации изображений. Эта структуризация может служить основой для моделирования индуктивных (при вводе данных) и дедуктивных (при анализе пирамиды «сверху вниз») процессов человеческого восприятия.

Основные преимущества пирамидально-рекурсивного подхода заключается в следующем [7, 8]:



– простота и скорость алгоритмов декодирования (для задач эффективного сжатия, анализа и синтеза изображений);

– удобство отображения закодированной информации (задачи синтеза, хранения и поиска изображений);

– наиболее важным достоинством используемых в системах кодирования алгоритмов пирамидально-рекурсивного разбиения является возможность обеспечения поэтапной обработки изображения, при которой выполнение операций каждого последующего обрабатывающего модуля возможно до полного окончания работы предыдущего модуля, и по мере выполнения операций обработки, в зависимости от класса задач, обеспечивается нужное приближение к эpsilon-энтропии сигнала изображения;

– возможность обработки закодированных данных без восстановления исходного изображения для большей части операций обработки.

Суть пирамидально-рекурсивного подхода заключается в следующем. Предполагается наличие множества опорных отчетов на исходном изображении, в результате триангуляции которых получаем регулярную двумерную полигональную сетку. Аппроксимация вершин сетки по яркости приводит к восстановлению результирующего изображения (рис. 1).

Точность, скорость и сложность алгоритмов кодирования и декодирования изображений напрямую зависят от числа и местоположения опорных точек.

Для нахождения опорных точек в [9–14] предложены эмпирические варианты, основная идея которых заключается в следующем:

– методом наименьших квадратов находится уравнение плоскости (S), проходящей через точки полигона с минимальным отклонением от заданного порога;

– разбивается полигон на одинаковые по площади полигоны, если отчеты превышают по яркости заданного порога (первая степень детализации);

– если яркости отчетов не превышают заданного порога, то в пределах полигона опорная точка определяется наименьшим расстоянием от плоскости (вторая степень детализации).

Общность подхода к решению поставленной задачи нашла отражение в предложенной их систематизации [3].

#### **Алгоритмы кодирования и декодирования по опорным точкам**

Процесс кодирования и декодирования нестационарных сигналов изображений заключается в следующем: все изображение рассматривается как единая область (полигон) и на ней проверяется критерий однородности (наличие или отсутствие каких-либо объектов или опорных точек) (см. рис. 1). На этапе кодирования, если критерий выполняется, то изображение считается содержащим лишь один фон (далее обозначается как пустой полигон). В противном случае проводится разбиение этого полигона на одинаковые равные полигоны и на каждом из них вновь вычисляется критерий однородности. Затем выполняется проверка, и дальнейшему

разбиению подвергаются лишь те полигоны, для которых критерий не соблюдается. Далее в пределах каждого из полученных полигонов осуществляется поиск оптимальной точки, максимально характеризующей данный полигон. Итак, в результате анализа всего изображения получаем полигональную двумерную сетку, содержащую как пустые полигоны, так и полигоны, содержащие ОТ [14].

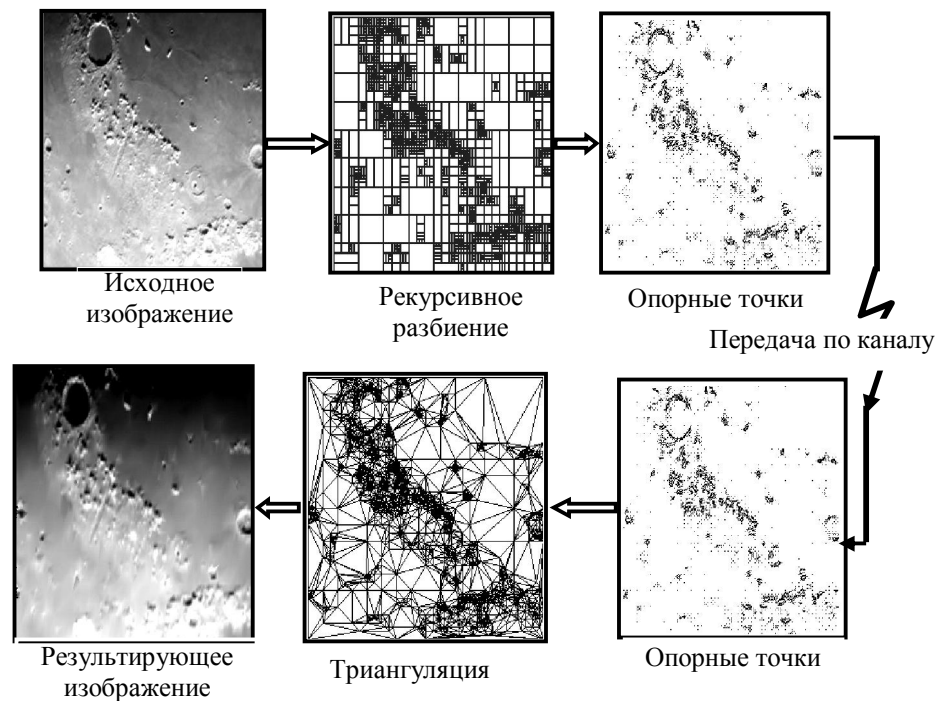


Рис. 1. Процесс кодирования и декодирования по опорным точкам

Каждая опорная точка характеризуется своими абсолютными координатами и яркостью в пределах полигона. Методы и алгоритмы разбиения, алгоритмы поиска объектов и критерии однородности в рамках данной статьи не рассматриваются, а с некоторыми из них можно познакомиться в работах [9, 11–13]. Далее на этапе декодирования каждая опорная точка соединяется с ближайшими соседними точками, и в итоге все точки будут соединены сеткой, называемой регулярной триангуляцией. Полученная треугольная сетка путем аппроксимации двумерной поверхности по яркости восстанавливает результирующее изображение. В данной статье приводятся обобщенные алгоритмы кодирования и декодирования нестационарных сигналов изображений, основанные на рекурсивном разбиении исходного изображения, структурная схема системы и результаты моделирования процессов сжатия и восстановления для различных изображений.



Существует хорошо известный, но мало изученный метод триангуляции Делоне [1]. Этот метод эффективен с точки зрения достижения высокого коэффициента сжатия ( $K_{сж}$ ). К сожалению, за это приходится платить сложностью реализации. Если на прямоугольной сетке можно за несколько операций определить ячейку, куда попала точка, в которой мы хотим узнать значение яркости, то в случае триангуляции поиск треугольника и анализ формы аппроксимирующих треугольников превращаются в нетривиальную задачу. Кроме того, сам процесс построения сетки непересекающихся треугольников с вершинами в опорных точках достаточно сложен.

### **Структурная схема системы кодирования по опорным точкам**

В настоящий момент, с появлением новой элементной базы в виде систем на кристалле и современных САПР на ее основе, открылись новые возможности создания специализированных систем обработки нестационарных сигналов и почти сняты многие ограничения по скорости и сложности для проектирования систем обработки видеoinформации, что, казалось бы, должно привести к решению основных задач восприятия видеоданных техническими системами [15]. Однако проблема не исчезла, а лишь возросло понимание ее значимости и специфических сложностей с учетом наличия фотоприемника в составе вычислителя [3]. Сложившаяся ситуация вынуждает исследователей делать все больше упор на развитие новых подходов и способов не только в обработке изображений, но и в эффективном представлении сжатых видеоданных [9], чтобы максимально сокращать разрыв между «машинным» и «человеческим» восприятием изображений [7, 8].

Пирамидально-рекурсивные методы разбиения основываются на одном из двух базовых свойств сигнала яркости: разрывности (нестационарности) и однородности. В первом случае подход состоит в разбиении изображений на основании резких изменений сигнала, таких как перепады яркости на изображении. Вторая категория методов используют разбиение изображения на области, однородные в смысле заранее выбранных критериев, например пороговая обработка, выращивание областей, слияние и разбиение областей [13, 14].

Регулярность пирамидальных структур позволяет создавать эффективные видеосистемы обработки изображений из-за возможности распараллеливания алгоритмов (рис. 2) [13, 14]. На этапе кодирования распараллеливанию подлежит процесс разбиения и поиска опорных точек на исходном изображении (рис. 2, а), а на этапе декодирования – процесс триангуляции (аппроксимации полигонов) опорных точек (рис. 2, б). Поскольку основное процессорное время занимают именно эти два процесса, следовательно, увеличение числа элементарных процессоров для выполнения этих функции и применение модульного проектирования приводит к увеличению производительности системы кодирования и декодирования в системах реального времени, обеспечивающую скорость не меньшую 30 кадр/с [3].

А для эффективного использования памяти необходимо ее разделить на три части: память заявок на разбиения полигонов, память опорных точек и память изображений (рис. 2). Однако развитие технологии систем на кристалле и

появление современных САПР на их основе позволили перейти от теоретического моделирования к реализации быстрых пирамидально-рекурсивных алгоритмов обработки видеoinформации и в перспективе решать следующие проблемы:

– реализацию алгоритмов пирамидальных представлений в виде параметризованных библиотечных СФ-блоков на базе технологии «система на кристалле» как параллельных, так и последовательных типов;

– создания мультимедийных систем на основе четырехмерной пространственной обработки видеоданных.

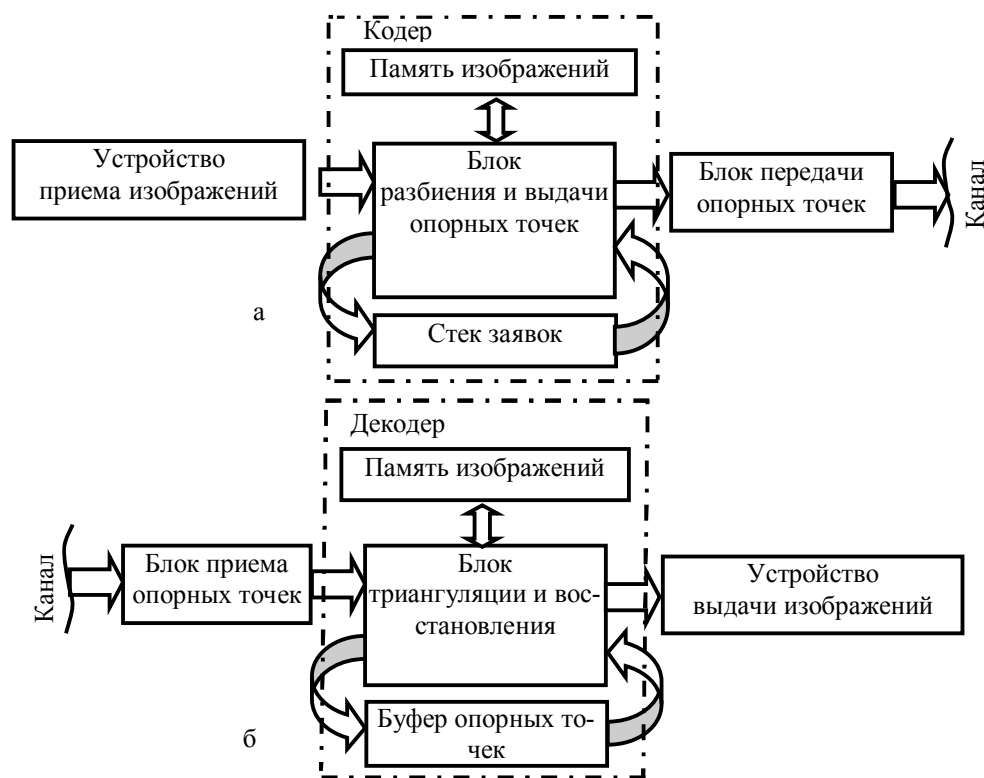


Рис. 2. Структурная схема системы кодирования (а) и декодирования (б) изображений

Синтез алгоритмов сжатия видеoinформации и определение параметров устройств для реализации алгоритмов ПРМ во многом опирается на максимальном приближении к энтропии нестационарного сигнала изображений [3]. На современном этапе твердотельной революции на кристалле размещается всё большее количество вентилях, и, как следствие, проектировщики назначают всё меньшую цену сложности алгоритмов кодирования по сравнению с ценой пропускной способности канала при сохранении качества восстановленных изображений.



### Результаты моделирования

Несмотря на очевидную целесообразность применения предлагаемого метода для сжатия и восстановления нестационарных изображений, при этом, считая полезным оценить коэффициенты сжатия на примерах изображений, предлагаемых в разработанной классификации по двум признакам: мере нестационарности и широкополосности сигналов изображений [16]. При этом сравнительный анализ коэффициентов сжатия для различных методов (дискретное вейвлет-преобразование (ДВП), дискретное косинусное преобразование (ДКП) и ПРМ [17, 19]) проводился при одинаковой субъективной оценке качества восстановленного изображения и для заданной пропускной способности канала передачи (рис. 3).

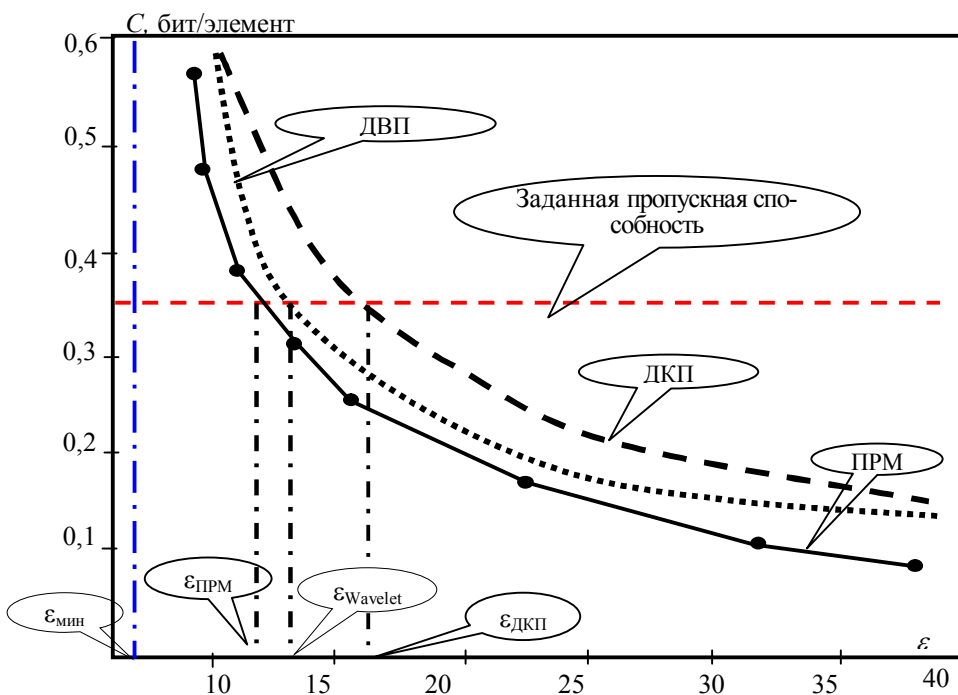


Рис. 3. Зависимость требуемой пропускной способности от допустимой СКО

Важно отметить оценку влияния геометрических и амплитудных отклонений, учитываемых в процессе сжатия, на качество восстановленного изображения. При этом наиболее существенным оказывается нахождение оптимального сочетания геометрических и амплитудных искажений, которое характеризуется их равным влиянием на качество восстановленного изображения.

Под амплитудным отклонением подразумевается тот диапазон допустимых искажений яркости точек исходного изображения при сжатии для получения полигональной сетки. А геометрические отклонения – это допустимые отклонения в координатной системе (т. е. в пределах полигона). Однако следует отметить, что



совокупность этих отклонений образует своеобразную апертуру в трехмерном пространстве, за пределы которого в процессе сжатия не допускаются отклонения яркостей и координат точек изображения. Сама апертура может быть вытянута вдоль координатных осей или оси яркости и имеет различные формы и пространства [10, 14].

В результате моделирования алгоритмов ПРМ определены основные характеристики системы для аппаратной реализации алгоритмов, полагаемых в основу оценки требуемой производительности СБИС кодера и декодера на базе технологии «система на кристалле»:

- объем рабочей памяти =  $(61, 1642)^1$ ;
- число обращений в память =  $(90399, 143037)$ ;
- длина сжатого описания исходного изображения в битах =  $(2412, 61306)$ ;
- плотность описания опорных точек в битах =  $(1,39, 1,25)$ ;
- количество опорных точек =  $(326, 8451)$ .

А также найдены оптимальное сочетание числа полигонов, получаемых при разбиении исходного полигона, числа уровней разбиения, среднеквадратическая ошибка кодирования  $\varepsilon$  и требуемой пропускной способности  $C$  при многократном проведении экспериментов в алгоритмах с фиксированным расположением ОТ с целью достижения высокой степени сжатия и приемлемом качестве [12, 18].

ПРМ алгоритмы кодирования и декодирования реализованы на языке высокого уровня C/C++, используемом на начальной стадии системного проектирования. Апробация ПРМ проведена при использовании разбиения на каждом шаге кодирования априорного интервала на 2, 3 и 4 части на изображениях двух типов контрастности (гладкие и контрастные). Легко заметить, что при хорошем субъективном качестве восстановленных изображений коэффициент сжатия с использованием ПРМ на 2 % больше, чем ДВП, и на 1,5 % больше, чем ДКП для гладких снимков, а для контрастных изображений на 1,4 % больше, чем ДВП и ДКП (рис. 4).

Исследования проводились над изображениями размером  $256 \times 256$  точек (полигон нулевого уровня) и яркостным разрешением 8 градаций. Ёмкость памяти для хранения цифрового изображения  $N = 256^2 \cdot C$ .

#### **Заключение**

Кратко можно сформулировать основные выводы, которые являются основными свойствами предлагаемого подхода при проектировании систем кодирования и декодирования нестационарных изображений.

1. Структура данных представляется в виде некоторой регулярной, иерархической, не зависящей от содержания данных структуры. Эта структура содержит множества опорных точек и способы их поиска, а также компактно расположенную в памяти ЭВМ, причем структура и специальная нумерация позволяют абстрагироваться от исходного многомерного скалярного или векторного информационного поля и сохраняющие в то время все свойства исходного изображения.

<sup>1</sup> Первое число для гладких изображений, а второе для контрастных.





2. Разработка алгоритмов решения конкретных задач производится таким образом, чтобы обеспечить возможность поэтапного приближения к энтропии исходного сигнала. Для этого на каждом уровне разбиения исходного изображения анализируется результат восстановления по величине ошибки кодирования и при необходимости производится дальнейшее разбиение и уточнения результирующего изображения.

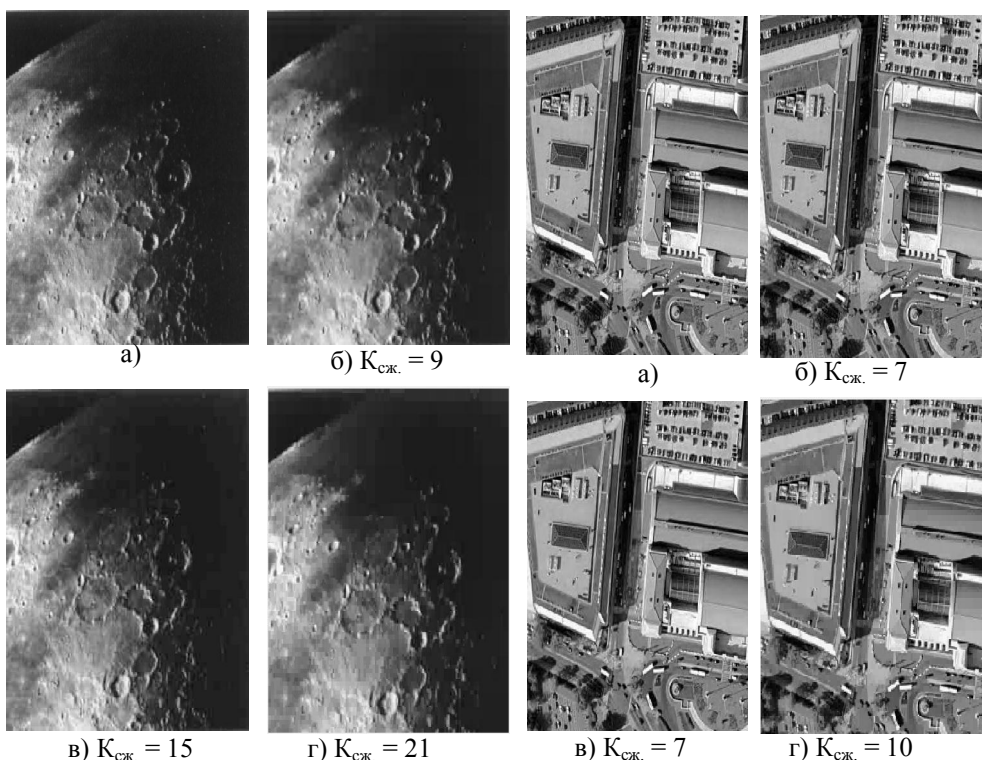


Рис. 4. Результаты сжатия двух классов изображений:  
а – исходное изображение; б – ДВП; в – ДКП; г – ПРМ

3. Элементы сжатого описания представляют собой не элементы изображения или информационного поля, а опорные точки, имеющие похожие или одинаковые свойства. Поэтому сложность ПРМ определяется не разрешающей способностью аппаратуры (количеством отчетов изображения), а количеством опорных точек, т. е. числом элементов самой структуры, которое для конкретного класса может быть значительно меньше числа отчетов исходного изображения.

#### Библиографические ссылки

1. *Эксперименты на дисплее. Первые шаги вычислительной физики.* М., 1989.



2. *Техническое зрение роботов* / под ред. проф. Л. Пью. М., 1987.
3. *Твердотельная революция в телевидении* / В. В. Березин, А. А. Умбиталиев, Ш. С. Фахми, А. К. Цыцулин, Н. Н. Шипилов. М., 2007.
4. *Немудров В., Мартин Г.* Системы на кристалле. Проблемы проектирования и развития. М., 2004.
5. *Эксперименты с машинным зрением* / В. П. Андреев, Д. А. Белов и др.; под ред. И. А. Овсеевича. М., 1987.
6. *Фахми Ш. С., Шах В. В., Шмидт В. К.* Алгоритм сжатия и восстановления методом деления на треугольники // Изв. СПб. ГЭТУ. 1993. Вып. 448.
7. *Александров В. В., Горский И. Д.* Представление и обработка изображений. Рекурсивный подход. Л., 1985.
8. *Александров В. В., Горский И. Д.* Алгоритмы и программы структурного метода обработки данных. Л., 1983.
9. *Фахми Ш. С., Шах В. В., Шмидт В. К.* Алгоритм сжатия и восстановления изображений с пирамидально-рекурсивной структурой данных / СПбГЭТУ. Деп. в ВИНТИ 11.02.93, № 350-В93.
10. *Фахми Ш. С.* Оценка степени сжатия изображений с использованием рекурсивного метода разбиения // Изв. ГЭТУ. 1993.
11. *Фахми Ш. С.* Аналитическая модель оценки эффективности кодирования по опорным точкам // Изв. ЛЭТИ. 2006.
12. *Фахми Ш. С., Шах В. В., Шмидт В. К.* Аналитическая модель для оценки степени сжатия информации по опорным точкам // Методы и аппаратно-программные средства цифровой обработки сигналов. 1993. № 1.
13. *Фахми Ш. С.* Кодирование и декодирование видеоинформации // Вопросы радиоэлектроники, сер. Техника телевидения. 2007. Вып. 2.
14. *Фахми Ш. С.* Полигональная рекурсивная обработка видеоинформации // Вопросы радиоэлектроники, сер. Техника телевидения. 2008. Вып. 1.
15. *Зубарев Ю. Б., Кривошеев М. И., Красносельский И. Н.* Цифровое телевизионное вещание. Основы, методы, системы. М., 2001.
16. *Фахми Ш. С., Зубакин И. А., Цыцулин А. К.* Классификация нестационарных изображений и разработка методики оценки алгоритмов кодирования источника // Материалы 63-й научно-технической конференции профессорско-преподавательского состава университета. СПб., 2010.
17. *Прэнтл У.* Цифровая обработка изображений. М., 1982.
18. *Фахми Ш. С.* Аналитическая модель оценки степени приближения к эпсилон-энтропии на основе пирамидально-рекурсивного метода кодирования изображений // Вестник Тихоокеанского государственного университета. 2010. № 1.
19. *Фахми Ш. С., Зубакин И. А.* Адаптивный алгоритм кодирования видеоинформации на основе трехмерного дискретного косинусного преобразования // Изв. вузов России. Радиоэлектроника. 2010. Вып. 1.