УДК 681.325

Ю.П. САЛЬНИК

Харьковский университет Воздушных Сил, Украина

МОДИФИЦИРОВАННЫЙ МЕТОД АРИФМЕТИЧЕСКОГО КОДИРОВАНИЯ С ФИКСИРОВАННОЙ МОДЕЛЬЮ

Проведен анализ существующих методов статистического кодирования, относящихся к классу оптимальных кодов. Предложен метод кодирования трансформанты преобразования Хаара с учетом известного распределения значений коэффициентов выделенных зональных последовательностей.

избыточность изображений, статистическое кодирование, зонально-пороговая селекция, статистика сглаженных гистограмм

Постановка проблемы и анализ литературы

Известно, что изображения, наряду с другими видами избыточности, обладают значительной статистической избыточностью [1, 2]. Снизить данную избыточность возможно за счет статистического кодирования [3].

Анализ наиболее распространенных методов сжатия, таких как JPEG, JPEG-2000, TIFF и др. показывает высокую эффективность применения кодов, относящихся к классу оптимальных [4, 5]. Так, использование кодирования Хаффмана, относящегося к этому классу, в алгоритме сжатия JPEG позволяет достичь коэффициентов сжатия до 20 раз.

В существующих методах сжатия визуальной информации используют методы статистического кодирования, которые выполняются по одной из двух схем:

1) двухпроходное кодирование, при котором при первом проходе моделируется источник, при втором источник кодируется в соответствии с сформированной моделью;

2) однопроходное (адаптивное) кодирование, при котором модель источника изменяется при поступлении очередного символа.

Такое кодирование обладает рядом недостатков: при первом подходе снижается эффективность кодирования из-за затрат времени на моделирование источника; при втором – из-за несоответствия получаемой модели кодируемому источнику.

Целью данной статьи является разработка адаптивного метода арифметического кодирования зональных последовательностей трансформанты ортогонального преобразования Хаара.

1. Определение количества информации

Один из подходов к определению количества информации, которая содержится в изображении, состоит в применении Шенноновской теории информации [1]. В ее основе лежит понятие энтропии $H(X)$ источника S с алфавитом X. При цифровом описании изображения энтропия цветовой составляющей определяется следующим выражением:

$$H(S) = -\sum_{i=1}^{N}\sum_{j=1}^{M} p(x_{i,j})\log(p(x_{i,j})), \qquad (1)$$

где $p(x_{i,j})$ – частота появления отсчета со значением $x_{i,j}$; N, M – размеры изображения по горизонтали и вертикали соответственно.

Поскольку для отсчетов $x_{i,j}$ характерна статистическая зависимость, то вероятности $p(x_{i,j})$ будут условными (1). Из-за чрезмерного объема необходимых вычислений оценить энтропию многоградационных изображений даже с использованием упрощенных схем (оценка энтропии Шрайбера) [1] не представляется возможным. Вместе с тем энтропию изображения можно получить по декоррелированным коэффициентам трансформанты его преобразования

$$H(S) = -\sum_{i=1}^{N}\sum_{j=1}^{M} p^*(y_{i,j})\log(p^*(y_{i,j})) , \qquad (2)$$

где $p^*(y_{i,j})$ – статистическая частота появления коэффициента со значением $y_{i,j}$:

$$p^*(y_{i,j}) = \frac{N_{y_{i,j}}}{N \cdot M} \$ \qquad (3)$$

$N_{y_{i,j}}$ – количество отсчетов со значением $y_{i,j}$.

2. Оценка энтропии трансформанты

Кодирование отсчетов оцифрованного изображения, представленного в цветовой модели RGB, выполняется с использованием распределения Бернулли. Такая модель не учитывает значительную статистическую избыточность, которой обладают фотографические изображения.

Использование зонально-пороговой селекции (ЗПС) коэффициентов преобразования Хаара, представленной в работах [6, 7] позволяет дополнительно снизить энтропию при сокращении психовизуальной избыточности изображения (наряду со снижением статистической избыточности при проведении ортогонального преобразования), что и показано в табл. 1.

Таблица 1

Значения энтропии H для изображений с различной степенью насыщенности при $Q_S = 3$, $Q_S = 5$, $Q_S = 7$ и $Q_S = 10$

Насыщенность изображений	По модели Бернулли	$Q_S = 10$	$Q_S = 7$	$Q_S = 5$	$Q_S = 3$
Высоко-насыщенное	8	2,85	2,62	2,1	1,45
Средне-насыщенное	8	1,82	1,53	1,31	1,047
Мало-насыщенное	8	0,755	0,708	0,674	0,641

В табл. 1 Q_S – параметр качества восстановленного изображения для s-й зональной последовательности (при $Q_S = 10$ порог селекции равен 0), а классификация изображений выполнена в соответствии с методом оценки насыщенности изображений, представленном в работе [7].

Значения энтропии трансформант до и после проведения ЗПС, представленные в таблице, показывают, что использование ортогонального преобразования позволяет более точно оценить энтропию изображения и снизить ее при использовании зонально-пороговой селекции в зависимости от выбранного параметра качества и класса насыщенности изображения от 6 до 12 раз.

Понятие энтропии источника является ключевым и в методах статистического кодирования. Рассмотрим его основные положения. Стоимостью кодирования f источника S с алфавитом Y называется величина $C(f,S) = \sum_{i=1}^{k} p^*(y_i)\left|f(y_i)\right|$, где k – мощность алфавита, $\left|f(y_i)\right|$ – длина кодового слова [3]. Тогда избыточность кодирования определяется выражением

$$R(f,S) = C(f,S) - H(S) . \qquad (4)$$

Для случая двоичного представления исходных зональных последовательностей Y выражение (4) примет следующий вид:

$$\begin{aligned} R(f,S) &= \sum_{i=1}^{k} p^*(y_i)\left|f(y_i)\right| + \sum_{i=1}^{k} p^*(y_i)\log_2(p^*(y_i)) = \\ &= \sum_{i=1}^{k} p^*(x_i)\left[\left|f(y_i)\right| + \log_2(p^*(y_i))\right], \end{aligned} \qquad (5)$$

где $\lceil\ \rceil$ – знак операции округления к ближайшему целому.

Анализ выражения (5) показывает, что величина избыточности кодирования тем меньше, чем меньше отличается длина кода исходного отсчета от логарифма его частоты. Наименьшей избыточностью обладают коды, относящиеся к классу оптимальных кодов. При оптимальном кодировании средняя длина передаваемых кодовых слов $\left|f(y_i)\right|$ приближается к энтропии изображения, т.е. $R(f,S) \to 0$. К классу оптимальных кодов относятся коды Хаффмана и арифметическое кодирование. Использование данных кодов позволяет обеспечить максимально возможные коэффициенты сжатия среди статистических кодов. Так, при использовании кодов Хаффмана длина кода символа равна

$$\left|f(y_i)\right| = \left\lceil -\log_2 p^*(y_i) \right\rceil .$$

При этом расчет суммарной кодовой последовательности приводит к накоплению ошибки округления, связанной с ограниченностью разрядной сетки вычислительных средств.

При использовании арифметических кодов длина кодовой последовательности равна

$$\left|f(y_{i_t},\dots,y_{i_{t+T}})\right| = \left\lceil -\log_2 \prod_{j=t}^{t+T} p^*(y_{i_j}) \right\rceil , \qquad (6)$$

где T – размер выборки из исходной последовательности.

Таким образом, главной отличительной чертой арифметического кода, по сравнению с кодом Хаффмана, является то, что в первом кодовая последовательность зависит не от всей исходной последовательности, а только от ее части $\left|f\left(y_{i_t},\dots,y_{i_{t+T}}\right)\right|$. Такой подход позволяет избежать ошибки кодирования при масштабировании исходной последовательности.

Если считать, что получение кода для каждого символа алфавита (при кодировании Хаффмана) или для выборки $(y_{i_t},\dots,y_{i_{t+T}})$ (при арифметическом кодировании) выполняются с абсолютной точностью, то суммарная длина кодовой последовательности $D_{теор}$ будет одинаковой при использовании любого из оптимальных кодов

$$D_{теор} = \sum_{j=1}^{N\cdot M} \left|f(y_{i_j})\right| = -\sum_{j=1}^{N\cdot M} \left\lceil \log_2 p(y_{i_j}) \right\rceil . \qquad (7)$$

Особенностью реализаций арифметических кодеков является то, что они не увеличивают размера исходных данных в худшем случае. Существующие реализации арифметического кодека требуют либо выполнения двух проходов по исходной последовательности (первый проход для подсчета частот символов, второй проход – собственно кодирование), либо однопроходного кодирования с накоплением частоты символов, что снижает эффективность сжатия, причем в первом случае также необходимо передавать на приемную сторону для декодирования статистическую модель исходной последовательности.

3. Сглаживание зональных последовательностей

Исследования показали, что гистограммы изображений (рис. 1) в отличие от их трансформант (рис. 2) не подчинены нормальному закону распределения (НЗР). Однако, рассмотрение зональных последовательностей (рис. 3) отдельно, кроме первой, обладающей глобальной чувствительностью, обеспечивает лучшее сглаживание НЗР.

Можно найти аналитическое выражение, выражающее лишь существенные черты полученных статистических распределений. Для этого необходимо получить численные параметры, входящие в выражение выравнивающей функции статистического распределения [8]. При выравнивании гистограмм с помощью нормального закона распределения такими параметрами является статистическое среднее m_y^* и статистическая дисперсия D_y^*:

$$p^*_{норм}(y) = \frac{1}{\sqrt{2\pi D_y^*}} \exp\left(-\frac{(y - m_y^*)^2}{2D_y^*} \right). \qquad (9)$$

Статистическое среднее и дисперсия рассчитываются по следующим выражениям [8]:

$$m_y^* = \sum_{i=1}^{K} y_i p^*(y_i)\,;\; D_y^* = \sum_{i=1}^{K} y_i^2 p^*(y_i) - m_y^{*2} \qquad (9, 10)$$

где K – количество коэффициентов ЗП.

Пример выполненного выравнивания статистического распределения модулей коэффициентов преобразования отсчетов средненасыщенного изображения представлен на рис. 4.

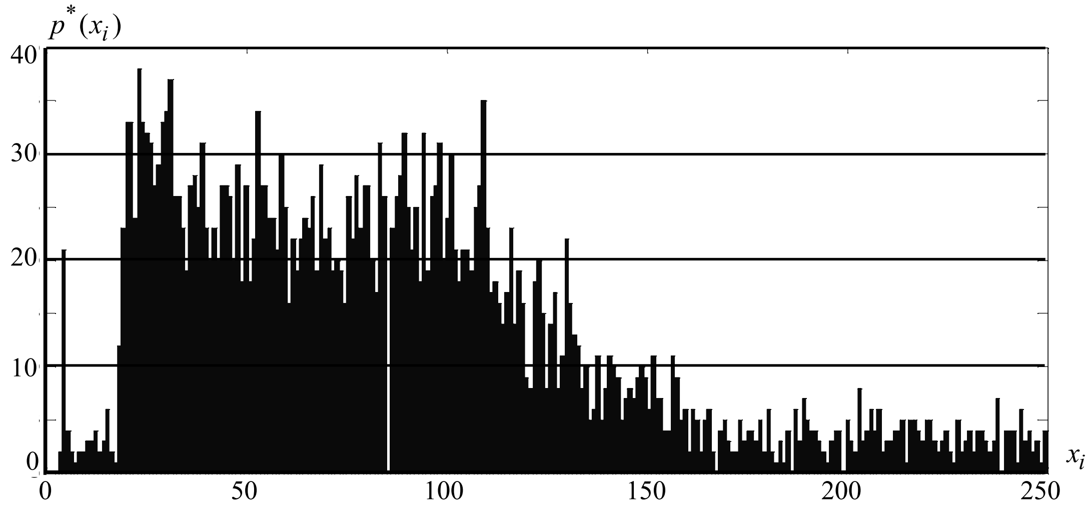


Рис. 1. Гистограмма распределения значений отсчетов изображения со средней степенью насыщенности

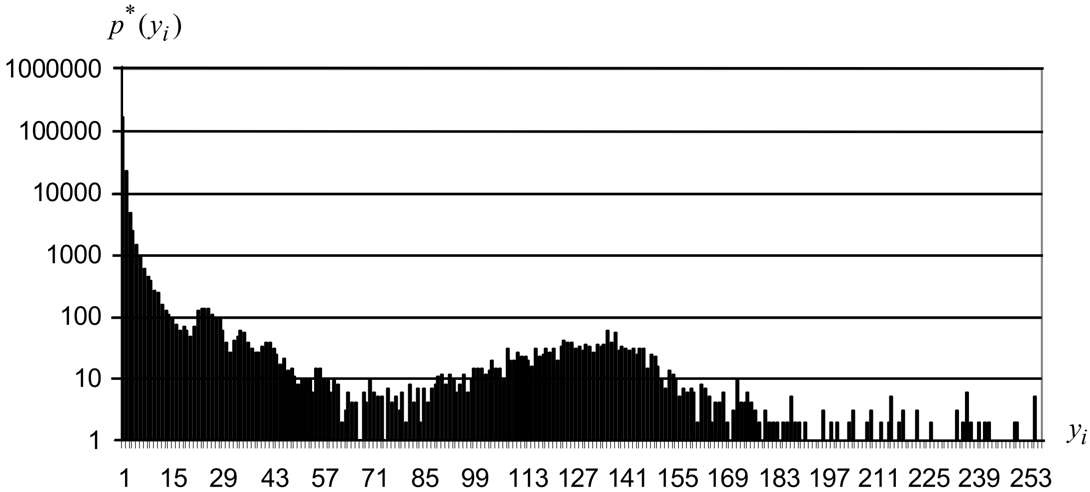
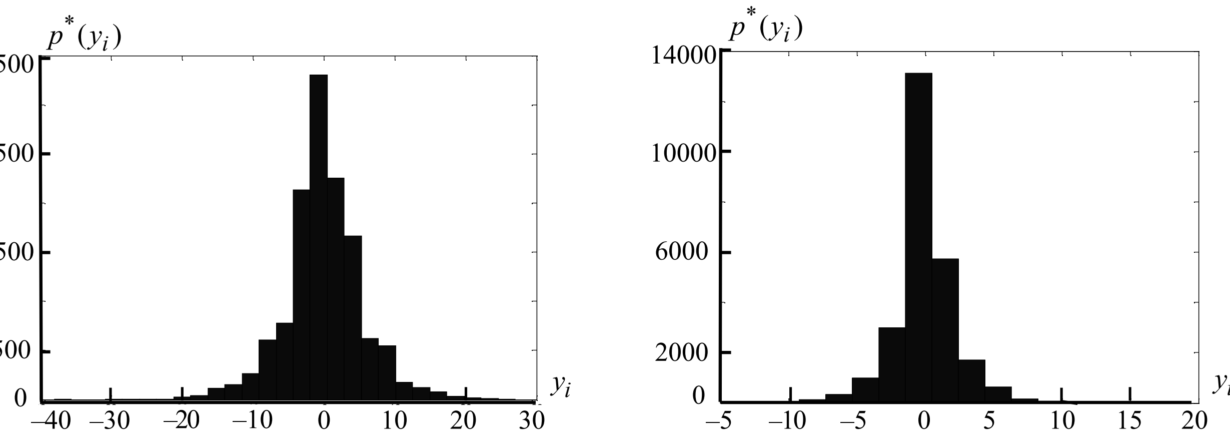


Рис. 2. Гистограмма распределения модулей значений y_i коэффициентов трансформанты преобразования Хаара изображения со средней степенью насыщенности



а б

Рис. 3. Гистограммы распределения значений y_i коэффициентов зональных последовательностей (ЗП) трансформанты изображения со средней степенью насыщенности: 3-я ЗП (а); 7-я ЗП (б)

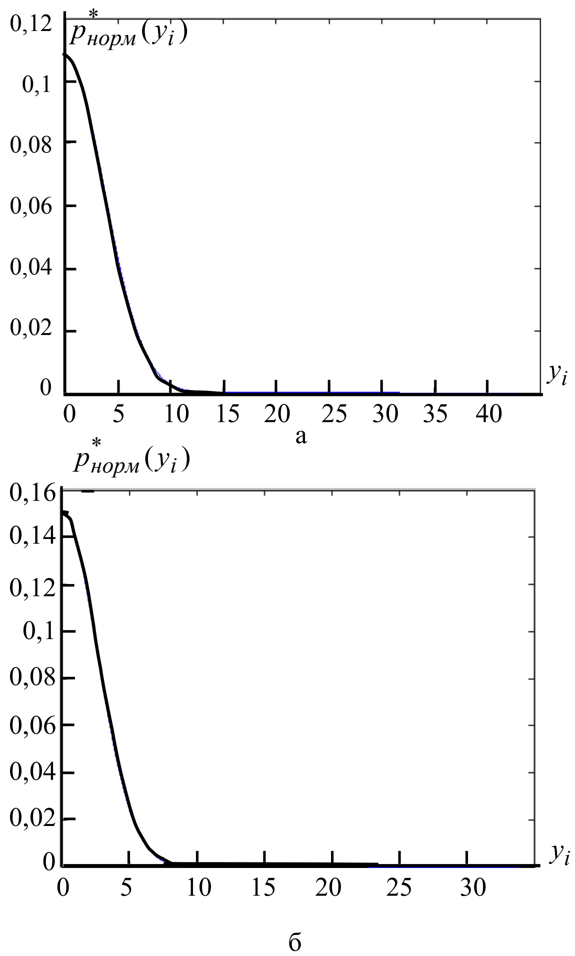


Рис. 4. Пример сглаживания ЗП НЗР:
а – 3-я ЗП; б – 7-я ЗП

Заключение

Модифицированный метод арифметического кодирования с фиксированной моделью позволяет:

– исключить необходимость двухпроходного кодирования при использовании фиксированной модели ЗП, полученной при присвоении обрабатываемому изображению класса насыщенности;

– избежать потери эффективности сжатия при однопроходном кодировании за счет известного статистического распределения значений коэффициентов ЗП;

– уменьшить объем данных, которые необходимо передать на приемную сторону, поскольку необходимо передавать только параметры моделей зональных последовательностей, таких как m_y^* и D_y^*.

Направлением дальнейших исследований является разработка алгоритмов и программных реализации предложенного метода кодирования.

Литература

1. Потапов В.Н. Теория информации. Кодирование вероятностных источников. – Новосибирск: МВССО РФ, 1999. – 271 с.
2. Ватолин Д., Ратушняк А. Смирнов М., Юкин В. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. – М.: ДИАЛОГ-МИФИ, 2002. – 384 с.
3. Миано Дж. Форматы сжатия изображений в действии. – М.: Триумф, 2003. – 354 с.
4. Бондарев В.Н, Трестер Г., Чернега В.С. Цифровая обработка сигналов: методы и средства. – Х.: Конус, 2001. – 398 с.
5. Прэтт У. Цифровая обработка изображений: Т. 1, 2. – М.: Мир, 1985. – 736 с.
6. Королев А.В., Бохан К.А., Сальник Ю.П. Порог зонально-пороговой селекции коэффициентов двумерного преобразования Хаара // ІКСЗТ. – 2004. – № 2. – С. 45 – 51.
7. Вентцель Е.С., Овчаров Л.А. Теория вероятностей и ее инженерные приложения. – М.: Наука, 1988. – 480 с.
8. Королев А.В., Бохан К.А., Сальник Ю.П. Оценка степени насыщенности изображений // Моделювання та інформаційні технології. – К.: ІПМЕ, НАНУ. – 2003. – Вип. 22. – С. 88 – 93.
9. Королев А.В., Бохан К.А., Сальник Ю.П. Зонально-пороговая селекция коэффициентов быстрого двумерного преобразования Хаара // Системи обробки інформації. – Х.: НАНУ, ПАНМ, ХВУ. – 2004. – Вип. 4. – С. 109 –113.

Поступила в редакцию 27.12.2004

Рецензент: д-р техн. наук, проф. О.Н. Фоменко, Харьковский университет Воздушных Сил, Харьков.