

СЕКЦИЯ 2 ОБРАБОТКА И ПЕРЕДАЧА ИЗОБРАЖЕНИЙ

УДК 681.3.06

КОДИРОВАНИЕ И СЖАТИЕ СЕГМЕНТИРОВАННОГО ИЗОБРАЖЕНИЯ

А.И. МИТЮХИН

*Институт информационных технологий Белорусского государственного университета
информатики и радиоэлектроники, Республика Беларусь*

Поступила в редакцию 18 марта 2019

Аннотация. В статье рассматривается алгоритм эффективного представления изображения, полученного после процесса сегментации. Предлагается комбинированный метод описания сегментированного объекта, включающий в себя выполнение двух последовательных этапов цифровой обработки исходных данных. Приведена сравнительная оценка эффективности сжатия и информационного содержания исходных данных с применением цепного кода Фримена, энтропийного кодирования и рассматриваемого комбинированного алгоритма. Представлен пример сокращения размерности обработки изображения сегментированного объекта.

Ключевые слова: изображение, граница объекта, сегментация, энтропия, форма, последовательность, окрестность, источник.

Введение

Операции распознавания некоторой изолированной области, присутствующей на изображении (снимке), предшествует этап сегментации. Желательно, чтобы число признаков, описывающих процесс сегментации, было минимально возможным. При этом, для повышения вероятности правильного последующего процесса распознавания, описание сегментированного объекта должно выполняться с определенной степенью точности. В статье рассматривается математический подход, основанный на комбинировании уже существующих алгоритмов обработки изображений для достижения уменьшения вычислительных затрат на представление и хранение данных, являющих объектом интереса

Теоретические принципы

Пусть имеется снимок $f(x, y)$ отдельного объекта определенной формы. В результате сегментации на дискретной сетке в пространственной области получено бинарное изображение $g(x, y)$. Изображение можно записать в виде 2D-последовательности (вектора):

$$g = g(x, y) = ((x_0, y_0), \dots, (x_n, y_n)), \quad (1)$$

где (x, y) – пространственные целочисленные переменные, n – длина границы, $n \in \mathbb{Z}^+$. Пиксели с координатами (x_i, y_i) и (x_{i-1}, y_{i-1}) удовлетворяют свойству m -компонентной окрестности, $0 \leq i \leq m$. При значительной величине n последовательности g , а также при необходимости выполнения требования точной сегментации и высокого разрешения вычислительные затраты на представление g могут оказаться недопустимыми.

Для определенных приложений, например, хранения отдельных деталей изображаемых объектов интереса, 2D-последовательность g можно представить в виде случайной последовательности $g(x, y)$ символов дискретного источника информации без памяти:

$$G = \{g_0, g_1, \dots, g_m\}, \tag{2}$$

где m определяет вид окрестности (на практике используется 4-компонентная или 8-компонентная окрестность). Источник, формирующий g , полностью описывается множеством $P = \{p(g_0), p(g_1), \dots, p(g_m)\}$ вероятностей символов. Знание распределения множества вероятностей P позволяет рассчитать энтропию источника G .

$$H = -\sum_{i=0}^m p(g_i) \log p(g_i). \tag{3}$$

В свою очередь, значение H определяет верхнюю и нижнюю границы средней длины l эффективного q -ичного префиксного кода, с помощью которого можно выполнить сжатие данных. В результате появляется возможность получения точной оценки энтропийного кодирования источника, эффективности сжатия. Недостатком энтропийного алгоритма кодирования является сравнительно низкий коэффициент сжатия:

$$L = \frac{N}{M} = \frac{N}{nl} \leq 3, \tag{4}$$

где N – затраты в виде числа носителей информации (например, двоичных чипов) на представление, хранение исходных данных без сжатия, M – затраты в виде количества чипов необходимых на представление, хранение исходных данных со сжатием.

Известно, что определенная эффективность описания контуров, замкнутых границ, линий, достигается с использованием цепного кода Фримена [1]. Кодирование состоит в присваивании отрезку контура (границы, линии) одноразрядного десятичного числа в соответствии с направлениями, показанными на рис. 1. Следует отметить как достоинство метода применение цепного кода, это то, что для пространственной «синхронизации» процесса декодирования (восстановления изображения) достаточно знания координат (x_i, y_i) только одного пиксела. Этот пиксел задает начальную пространственную фазу кода.

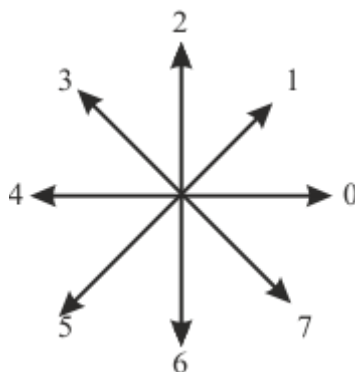


Рис. 1. Кодирование векторов направлений в 8-компонентной окрестности

Уменьшение вычислений при обработке изображений с применением цепного кода происходит за счет кодирования исходных данных равномерными словами длиной два или три двоичных символа (чипа). Таким образом, начальный этап обработки изображения включает в себя сегментацию изображения $f(x, y)$ для получения последовательности $g = g(x, y)$. Далее последовательность кодируется цепным кодом $c = c_0, c_1, \dots, c_n$. В статье рассматривается применимость к цепному коду линейного ортогонального преобразования, позволяющего повысить эффективность сжатия исходных данных. Известно, среди всех линейных преобразований оптимальным по отношению к критерию наименьшего значения

среднеквадратической ошибки при неполной размерности (неточном задании координат) является декоррелирующее преобразование в базисе дискретных функций [1, 2]. Рассматривается подход, основанный на преобразовании цепного кода $c = c_0, c_1, \dots, c_n$ в систему координат, базовым векторным пространством которого служат собственные векторы ковариационной матрицы $\text{cov}(c)$ цепного кода c . Вычисление вектора собственных значений ковариационной матрицы $\text{cov}(c)$ позволяет произвести оценку возможной эффективности сжатия обрабатываемых данных:

$$\Lambda = (\lambda_0, \lambda_1, \dots, \lambda_n) \quad (5)$$

При векторном описании сегментированного изображения, соответственно, кода c как выхода источника (2), можно утверждать, что среднее количество информации этого источника, приходящееся на символ, равно энтропии (3). Так как значение энтропии не зависит от того, в каком порядке следуют символы источника, кодовую последовательность c запишем в виде матрицы \mathbf{C} размером $(n \times 2)$. Из практических и вычислительных соображений временные затраты на обработку входного процесса всегда уменьшаются, если заранее исключить составляющую входа отражающую математическое ожидание. Тогда преобразованию будут подлежать входные данные:

$$\mathbf{B} = \mathbf{C} - E(\mathbf{C}), \quad (6)$$

где $E(\mathbf{C})$ – среднее значение матрицы \mathbf{C} .

Обозначим матрицу собственных векторов через \mathbf{A} . Прямое преобразование матрицы \mathbf{B} можно записать следующим образом:

$$\hat{\mathbf{B}} = \mathbf{A}^T \mathbf{B}, \quad (7)$$

где \mathbf{A}^T – ядро преобразования в базисе собственных векторов.

Имея распределение значений дисперсий (5) коэффициентов преобразования (7), можно решить задачу уменьшения размера входа обработки. Матрицы \mathbf{A} и \mathbf{B} представляются меньшим числом коэффициентов. К этим укороченным матрицам применяется обратное преобразование:

$$\mathbf{B}'' = \mathbf{A}'' \hat{\mathbf{B}}'', \quad (8)$$

где \mathbf{A}'' – укороченная матрица ядра обратного преобразования \mathbf{A} , $\hat{\mathbf{B}}''$ – укороченная матрица, полученная из матрицы коэффициентов преобразования $\hat{\mathbf{B}}$. В результате достигается уменьшение размера входа обработки данных.

Пример. Кодирование изображения цепным кодом и сжатие.

На рис. 2. показано изображение кафедрального собора. Требуется компактно представить форму собора, используя его граничные пиксели. После выполнения сегментации объекта и кодирования с помощью 8-компонентной окрестности получен двумерный цепной код c длиной $n = 42$. Элементы кода показаны в таблице. Кодовой последовательности c соответствует матрица \mathbf{C} размером (42×2) .

Цепной код сегментированного объекта

№ пиксела	0	1	2	3	4	5	6	7	8	9	10	11	12	13
Элементы цепного кода	0	7	6	7	6	6	6	6	6	6	6	7	0	7
№ пиксела	14	15	16	17	18	19	20	21	22	23	24	25	26	27
Элементы цепного кода	7	6	6	4	4	4	4	3	4	4	4	4	4	2
№ пиксела	28	29	30	31	32	33	34	35	36	37	38	39	40	41
Элементы цепного кода	2	2	1	1	2	2	2	2	2	1	2	2	2	2
Элементы цепного кода	2	1	1	2	2	2	2	2	2	2	2	1	1	2



Рис. 2. Изображение кафедрального собора

Вычислив среднее значение $E(\mathbf{C})$ и используя выражение (6), находим матрицу входа с нулевым математическим ожиданием \mathbf{B} . Далее вычисляется ковариационная матрица $\text{cov}(\mathbf{B})$ размером 42×42 . По матрице $\text{cov}(\mathbf{B})$ вычисляется распределение Λ собственных значений λ_i
 $\Lambda = (\lambda_0, \lambda_1, \dots, \lambda_{41}) = (0, 0, \dots, 75)$.

Как видно, имеется только одно ненулевое собственное значение равное 75. Таким образом, возможна полная декорреляция исходных данных и тем самым компактное их представление и описание для определенных практических применений. Вычисление матрицы собственных векторов \mathbf{A} ковариационной матрицы $\text{cov}(\mathbf{B})$ позволяет определить по выражению (7) матрицу коэффициентов преобразования:

$$\hat{\mathbf{B}} = \begin{pmatrix} \hat{b}_{00} & \hat{b}_{10} & \dots & \hat{b}_{410} \\ \hat{b}_{10} & \hat{b}_{11} & \dots & \hat{b}_{141} \end{pmatrix}^T = \begin{pmatrix} 0 & 0 & \dots & 6,1237 \\ 0 & 0 & \dots & -6,1237 \end{pmatrix}^T. \quad (9)$$

Из (9) следует, что для описания данных в области преобразований, достаточно сохранить значения коэффициентов $\hat{b}_{4,10}$ и $\hat{b}_{1,41}$. В этом случае затраты на хранение изображения (без информации о координатах начального пиксела) составят величину $M = 2$. Промежуточное обратное преобразование вычисляется по формуле (8). Точное восстановление сегментированной границы реализуется с использованием вектора:

$$\hat{\mathbf{b}}'' = (6,1237 - 6,1237)^T, \quad (10)$$

среднего вектора $E(\mathbf{C})$ в области исходных данных (всегда можно вычислить заранее) и только одной строки \mathbf{A}'' матрицы ядра обратного преобразования \mathbf{A} .

Проведем сравнительную оценку информационного содержания исходных данных и эффективность сжатия с применением цепного кода Фримена, энтропийного кодирования и рассматриваемого комбинированного алгоритма.

1. Кодирование граничных пикселей цепным равномерным кодом на основе 8-компонентной окрестности потребует $N = 2nl = 252$ двоичных символов (чипов).

2. Вычисление энтропии дискретного изображения (см. таблицу и (3)), когда учитываются вероятностные характеристики источника (2)

$$P = \{p(g_0), \dots, p(g_7)\} = \left\{ \frac{1}{42}, \frac{7}{84}, \frac{2}{7}, \frac{1}{84}, \frac{17}{84}, \frac{1}{84}, \frac{11}{42}, \frac{7}{84} \right\}, \quad \text{даёт следующее значение:}$$

$$H = 2,1789 \frac{\text{бит}}{\text{символ}}.$$

Полученное значение соответствует предельному минимальному значению средней длины \bar{l} кода. На практике средняя длина \bar{l} слова энтропийного кодирования находится в пределах $H + 1 > \bar{l} > H$. В наилучшем случае это означает, что исходные данные могут быть представлены с помощью $M \approx 183$ двоичных чипов. В сравнении с равномерным кодированием цепным кодом, коэффициент сжатия (выигрыш в записи данных) $L_{entropy}$ (4) в описании составляет величину

$$L_{entropy} = \frac{N}{M} = 1,377.$$

3. Комбинированный метод описывается с помощью 8-битового двоичного кода. На описание элементов вектора $\hat{\mathbf{b}}''$ (9) и двух координат начального пиксела границы требуется $M' = 32$ чипов. Коэффициент сжатия равен 7,785.

Как видно, предлагаемый метод обеспечивает сравнительно высокую степень сжатия с нулевой ошибкой восстановления изображения после процедуры декодирования.

Заключение

Подход на основе объединения метода кодирования сегментированных изображений объектов посредством цепного кода и метода ортогонального преобразования в базе дискретных собственных функций позволяет осуществлять эффективное описание данных. Применение комбинированного алгоритма для приложений, связанных с распознаванием, хранением значительных информационных массивов, позволяет снижать технические затраты на хранение данных, более экономно использовать вычислительный ресурс, уменьшать энергетические затраты.

ENCODING AND COMPRESSING THE IMAGE SEGMENTED

A.I. MITSUKHIN

Abstract. An algorithm for effective representation of the image obtained after segmentation is described. The combined method for describing segmented object including two successive phases of digital processing of source data is proposed. The comparative evaluation of the effectiveness of the compression and information content of the raw data using chain code Freeman, entropy coding and reporting combined algorithm is given. An example of reducing dimension processing segmented object is shown.

Keywords: image of the object boundary, segmentation, entropy, form, consistency, a neighborhood source.

Список литературы

1. Jahne B. Digital Image Processing. Concepts, Algorithms, and Scientific Applications. E-bok, 2013.
2. Mitsukhin A. // Proceedings 59th IWK. Ilmenau Scientific Colloquium, Technische Universität Ilmenau, Band 59, 2017, Heft 2.2.02, 6 Seiten.