

Кластеризация пикселей для сегментации цветового изображения

М. Харинов

Федеральное государственное бюджетное учреждение науки Санкт-Петербургский институт информатики
И автоматизации Российской академии наук (СПИИРАН)
khar@iias.spb

Аннотация

В докладе рассматривается проблема сегментации изображения посредством иерархической последовательности кусочно-постоянных приближений, минимально отличающихся от изображения по суммарной квадратичной ошибке. Предлагается получать приближения комбинированным методом кластеризации/сегментации на основе метода Уорда кластеризации пикселей изображения, в котором предусматривается снижение числа сегментов в бинарной иерархии кластеров за счет реклассификации пикселей из одного кластера в другой. В случае цветowych изображений, предлагаемый метод применяется сначала для выравнивания и снижения числа кластеров пикселей в каждой из RGB-компонент, а затем для кластеризации, в частности, сегментации, цветового изображения в целом. В результате достигается заметное улучшение качества приближений изображения, а их кодирование в инвариантном виде обеспечивает разметку изображения для обнаружения объектов.

Ключевые слова: изображение, приближение, бинарная иерархия, минимизация, суммарная квадратичная ошибка, метод Уорда, метод Оцу, модель Мамфорда–Шаха.

1. ВВЕДЕНИЕ

Доклад относится к области распознавания цифровых изображений на первоначальной стадии «приведения данных к удобному для распознавания виду» [6], «локализации объектов» [5], извлечения и упорядочения «глобально-локальной информации» [1]. Более конкретно, речь идет о *проблеме сегментации*, трактуемой нами как стадия автоматического выделения на изображении, доступных для анализа и распознавания вложенных изображений, именуемых *объектами*, которые обнаруживает на исходном изображении компьютер. Допускается, что в общем случае, пиксели объекта составляют на изображении *кластер* из того или иного числа несмежных связанных сегментов. Таким образом, в нашем понимании, термины «изображение», «объект», «кластер» употребляются в сходном смысле.

Если изображение состоит из изображений, то последние, в свою очередь, также разделяются на изображения, и сегментация выполняется посредством итеративного разделения изображения до тех пор, пока оно не распадется на некоторое множество *тривиальных* кластеров из одинаковых пикселей, которые полагаются неделимыми. Таким образом, мы вычисляем не единственное разбиение изображения, как при традиционной сегментации, а иерархию разбиений изображения на объекты, которая отражает неоднозначность изображения. Другим отличием нашей обобщенной интерпретации сегментации является то, что для каждого объекта автоматически генерируется целочисленная *инвариантная* метка, обозначение объекта, которое не зависит от предусмотренных преобразований или искажений исходного изображения. При этом числовые значения инвариантных меток кодируют иерархию сопоставляемых им объектов.

2. ПОСТАНОВКА ЗАДАЧИ

Для нахождения иерархии объектов решается задача оптимизации качества требуемых разбиений [2], которое при каждом значении числа кластеров k оценивается по величине суммарной квадратичной ошибки $E^{(k)}$ изображения относительно своего *кусочно-постоянного приближения* с усредненными по кластерам значениями пикселей. Так как последовательность оптимальных приближений изображения, описываемая минимально возможными значениями суммарной квадратичной ошибки $E_{\min}^{(k)}$, в общем случае, не является иерархической, возникает задача аппроксимации оптимальных приближений иерархией *квазиоптимальных* приближений, описываемых значениями $E^{(k)}$:

$$E^{(k)} \approx E_{\min}^{(k)}, \quad k = 1, 2, 3, \dots, g, \quad (1)$$

где g – общее число *тривиальных* кластеров.

При этом полагается, что для последовательности $E^{(k)}$ сохраняется свойство выпуклости:

$$E^{(k)} \leq \frac{E^{(k-1)} + E^{(k+1)}}{2}, \quad k = 2, 3, \dots, g-1, \quad (2)$$

которое характерно для мажорируемой последовательности оптимальных значений $E_{\min}^{(k)}$.

Если для кластеров пикселей устанавливается условие, что они не должны состоять из нескольких несмежных сегментов, то задача кластеризации сводится к задаче иерархической сегментации изображения, т.е. получению последовательности приближений изображений разбиениями с $k = 1, 2, 3, \dots, g$ связными сегментами, и соответствующей последовательности $\tilde{E}^{(1)}, \tilde{E}^{(2)}, \tilde{E}^{(3)}, \dots, \tilde{E}^{(g)}$ значений суммарной квадратичной ошибки. Если условие не минимизирует, а только ограничивает число связанных сегментов в составе кластера, то ставится задача получения приближений с промежуточными значениями $\tilde{E}^{(1)}, \tilde{E}^{(2)}, \tilde{E}^{(3)}, \dots, \tilde{E}^{(g)}$ суммарной квадратичной ошибки:

$$\tilde{E}^{(k)} \geq \tilde{E}^{(k)} \geq E^{(k)} \geq E_{\min}^{(k)}, \quad k = 1, 2, 3, \dots, g. \quad (3)$$

По существу, требуется алгоритм комбинированной кластеризации/сегментации, который в одном предельном случае обеспечивает иерархическую квазиоптимальную кластеризацию пикселей, а в другом – иерархическую *сегментацию* изображения, при которой число кластеров совпадает с числом связанных сегментов.

3. АНАЛИТИЧЕСКОЕ ОБОСНОВАНИЕ

Суммарная квадратичная ошибка $E^{(k)}$ для приближения из k кластеров вычисляется суммированием суммарных квадратичных отклонений E_i значений пикселей каждого кластера от их среднего значения:

$$E^{(k)} = \sum_{i=1}^k E_i, \quad (4)$$

Если i -й кластер рассматривать как результат слияния j вложенных кластеров $i = 1 \cup 2 \cup \dots \cup j$, то суммарное квадратичное отклонение E_i для i -го кластера получается суммированием квадратичных отклонений E_p для вложенных кластеров и некоторой неотрицательной добавки, характеризующей средний взвешенный разброс квадратов евклидовых попарных расстояний $\|I_p - I_q\|^2$ между центрами I_p, I_q кластеров p и q из n_p и n_q пикселей в цветовом RGB пространстве [12]:

$$E_i = \sum_{p=1}^j E_p + \frac{1}{2} \cdot \frac{\sum_{p=1}^j \sum_{q=1}^j n_p n_q \|I_p - I_q\|^2}{n_i}, \quad (5)$$

где $n_i = \sum_{p=1}^j n_p$ – число пикселей в i -м кластере.

Если в качестве j вложенных кластеров иметь в виду тривиальные кластеры, на которые распадается обсуждаемый i -й кластер, то в (5) сумма E_p , очевидно, обращается в 0.

Если в качестве тривиальных кластеров рассматриваются отдельные пиксели, то в (5) следует также положить $n_p = n_q = 1$ и $j = n_i$. Тогда (5), с точностью до множителя, совпадает с выражением для дисперсии значений пикселей i -го кластера, которое активно применяется в [8 – 10] для теоретической и практической разработки методов кластеризации. Таким образом, (5) обобщает аналогичную формулу из [8 – 10] на случай тривиальных кластеров из нескольких или многих пикселей.

Для случая $j = 2$ пары вложенных кластеров, из (4), (5) несложно получить [12] выражения для приращений суммарной квадратичной ошибки ΔE_{merge} и $\Delta E_{correct}$ при слиянии пары кластеров 1 и 2, а также их коррекции посредством реклассификации пикселей кластера $3 \subset 1$ из кластера 1 в кластер 2:

$$\Delta E_{merge} = \frac{\|I_1 - I_2\|^2}{\frac{1}{n_1} + \frac{1}{n_2}}, \quad (6)$$

$$\Delta E_{correct} = \frac{\|I_3 - I_2\|^2}{\frac{1}{n_3} + \frac{1}{n_2}} - \frac{\|I_3 - I_1\|^2}{\frac{1}{n_3} - \frac{1}{n_1}}, \quad (7)$$

где n_1, n_2, n_3 – число пикселей, в кластерах 1, 2, 3, а I_1, I_2, I_3 – трехмерные центры кластеров 1, 2, 3.

Именно выражение (6) минимизируется при слиянии сегментов изображения в версиях [4, 12] модели сегментации Мамфорда-Шаха [14], и либо с дополнительным множителем, либо с аддитивной добавкой используется в ее версиях [7, 13]. Формула (7) интересна тем, что из нее выводится классический метод K -means [2] и обобщенный метод K -meanless [10], который в нашей версии [12] отличается тем, что предусматривает операции с наборами пикселей.

4. ИНВАРИАНТНОЕ ПРЕДСТАВЛЕНИЕ

Для автоматической разметки поля изображения целочисленными метками, обозначениями объектов используется *инвариантное представление* изображения, разработанное и запатентованное в задачах стеганографии [11]. Представление $H(u, v, w)$ цветового изображения (u, v, w) из трех RGB -компонент называется *инвариантным*, так как не зависит от умножения значений пикселей изображения на положительное число $\alpha > 0$ и сдвига значений пикселей любой компоненты на одинаковую величину:

$$H(u, v, w) = H(\alpha \cdot u, \alpha \cdot v, \alpha \cdot w) = H(u + c_R, v + c_G, w + c_B), \quad (8)$$

где c_R, c_G, c_B – *тривиальные* компоненты изображения, состоящие из одинаковых пикселей.

Преобразование H коммутирует с преобразованием P :

$$HP = PH, \quad (9)$$

где под преобразованием P можно понимать: а) произвольную перестановку цветовых компонент; б) преобразование цветовой компоненты в негатив; в) масштабирование изображения посредством дублирования пикселей. Перестановочность аргументов в $H(u, v, w)$ означает, что функция преобразования цветовой компоненты зависит только от двух других компонент и не зависит от номера компоненты:

$$H_{u,v} : H(u, v, w) \equiv (H_{v,w}u, H_{u,w}v, H_{u,v}w), \quad (10)$$

причем, не меняется при перестановке индексов: $H_{u,v} = H_{v,u}$. При этом преобразование цветовой компоненты серого изображения из одинаковых компонент совпадает с преобразованием данной цветовой компоненты при двух других тривиальных:

$$H_{c,c}u = H_{u,u}u, \quad (11)$$

где c – *тривиальная* матрица из одинаковых пикселей, например, из одних 0. Таким образом, программная реализация преобразования H цветового изображения поддерживает одно и то же преобразование одномерной компоненты в составе серого изображения из трех одинаковых цветовых компонент или монохромного изображения при обнулении остальных.

При заданном алгоритме разделения нетривиального кластера на пару вложенных инвариантное представление $(H_{v,w}u, H_{u,w}v, H_{u,v}w)$, по существу, является результатом целочисленной «оцифровки» средних яркостей вложенных кластеров. Оно строится посредством кодирования знаков разностей средних яркостей вложенных кластеров в итеративном алгоритме, подобном алгоритму получения двоичного кода числа, в котором для учета при кодировании неделимых тривиальных кластеров используется не двоичное, а троичное значение коэффициентов разложения по степеням 2. При этом по результирующему инвариантному представлению вычисляются предыдущие – посредством арифметических операций [11].

5. КЛАСТЕРИЗАЦИЯ И СЕГМЕНТАЦИЯ

Комбинированный метод сегментации изображения посредством кластеризации пикселей поясняется на примере стандартного изображения «Лена» из 256×256 пикселей и трех одинаковых RGB -компонент (рис. 1).

На рис. 1 показаны кусочно-постоянные приближения изображения, в которых усредненные по кластерам значения

пикселей на первом шаге кодирования [11] просто заменены значениями 0 и 2 числовых кодов инвариантного представления яркости и нормализованы на рабочий диапазон.



Рис. 1. Кластеризация пикселей и сегментация изображения в инвариантном представлении

Суть обсуждаемого метода состоит в том, что сначала для изображения методом Уорда [16] строится квазиоптимальная бинарная иерархия кластеров, которая получается итеративным слиянием кластеров начального разбиения изображения на пиксели из условия $\Delta E_{merge} = \min$, где ΔE_{merge} раскрывается в (6). Затем берется разбиение множества пикселей изображения на пару вложенных кластеров (рис. 1, слева), и число сегментов итеративно редуцируется за счет реклассификации пикселей неизолированных сегментов из одного вложенного кластера в другой. Очередность реклассификации устанавливается из условия $\Delta E_{correct} = \min$, где $\Delta E_{correct}$ вычисляется согласно (7). Коррекция пары вложенных кластеров по числу сегментов продолжается до выполнения установленного правила останова, например, условия неповторяющихся площадей сегментов (рис. 1, в центре), или, при безусловной коррекции, продолжается до тех пор, пока во вложенных кластерах не останется по одному сегменту (рис. 1, справа).

По завершению формирования вложенных кластеров, каждый из них рассматривается как самостоятельное изображение, для которого обновляется иерархия вложенных кластеров, и вся процедура повторяется до тех пор, пока множество пикселей изображения не распадется на *тривиальные* кластеры из одинаковых пикселей. По мере формирования иерархии кластеров пикселей продолжает формироваться и инвариантное представление, которое визуально выглядит как приближение изображения на очередной итерации, так как связано с ним *изотонным* преобразованием, сохраняющим яркостный порядок [11].

В отличие от модели Мамфорда-Шаха [4, 7, 13, 14], в методе Уорда [16] при минимизации суммарной квадратичной ошибки $E^{(k)}$ принимаются во внимание все пары кластеров. Однако в частном случае серого изображения эквивалентный результат получается при анализе только пар кластеров с последовательными значениями средней яркости. При этом квазиоптимальные приближения (рис. 1, слева) можно получать в простейшем гистограммном алгоритме [3], если вместо эвристического критерия слияния [3] воспользоваться критерием $\Delta E_{merge} = \min$.

С другой стороны, в случае серого изображения для разделения очередного кластера надвое можно воспользоваться классическим методом Оцу [15], что представляет интерес для сравнения (рис. 2).

Рис. 2 иллюстрирует поведение суммарного числа сегментов s в кластерах и среднее квадратичное отклонение

$$\sigma^{(k)} = \sqrt{\frac{E^{(k)}}{3 \cdot N}}$$

приближения от изображения в зависимости от числа k кластеров в изображении из N пикселей и трех одинаковых *RGB* компонент. Сплошные серые кривые описывают наш метод при выполнении разделения кластера

надвое посредством обращения последнего слияния в методе Уорда, а черные пунктирные кривые описывают разделение кластера надвое методом Оцу.

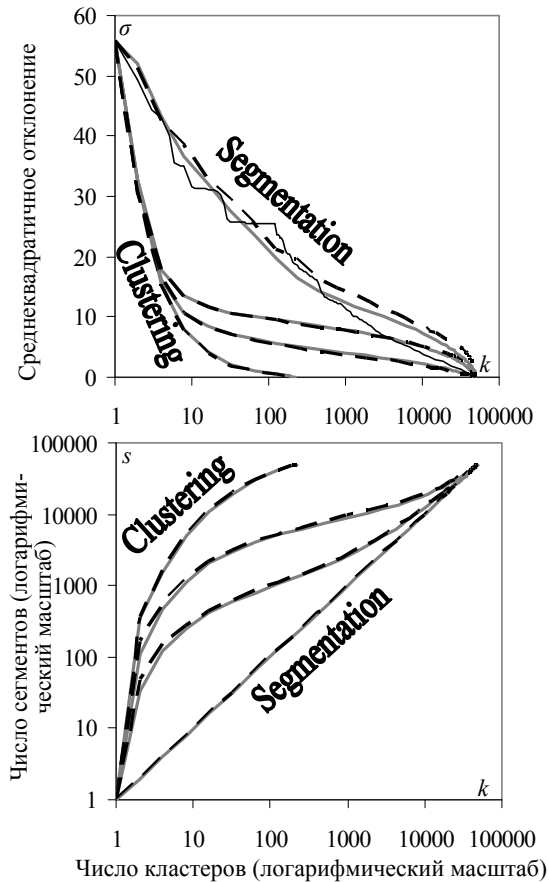


Рис. 2. Сопоставление методов по качеству приближений σ (вверху) и числу сегментов s (внизу) в зависимости от числа кластеров k

На верхних графиках рис. 2 тонкая сплошная черная кривая описывает результаты сегментации в модели Мамфорда-Шаха по версии FLSA [7]. Переплетающиеся с ней серая и пунктирная кривые описывают сегментацию изображения посредством кластеризации с безусловной коррекцией кластеров. Нижняя пара кривых описывает квазиоптимальную кластеризацию пикселей методами Уорда и иерархическим методом Оцу без коррекции кластеров по числу сегментов. Промежуточные пары кривых описывают кластеризацию пикселей по очереди с итеративной коррекцией, до останова при отсутствии повторений сегментов по площади и средней яркости (третья и вторая снизу пара кривых).

На нижних графиках рис. 2 соответствующие кривые следуют в обратном порядке. Верхняя пара кривых отражает лавинообразное возрастание числа сегментов с ростом числа кластеров при кластеризации без коррекции. Совпадение числа сегментов с числом кластеров при сегментации изображения описывается нижней прямой. Промежуточные кривые описывают подавление возрастания числа сегментов за счет коррекции разделения кластеров.

В целом, наш метод позволяет получать приближения изображения с заранее установленным качеством, так как область между кривыми на рис. 2 оказывается доступной при варьировании правил останова коррекции. Для серого изображения варианты комбинированной иерархической кластеризации/сегментации на основе метода Уорда и иерархического метода Оцу являются взаимно-

дополнительными, причем коррекция кластеров не нарушает согласованности результатов, как по качеству разбиений, так и по количеству сегментов (рис. 2).

Так как метод Оцу [15] не имеет очевидного обобщения для цветового изображения, то в этом случае из двух вариантов нашего метода остается применимым только метод кластеризации/сегментации на основе метода Уорда, который поясняется на примере обработки цветового изображения «Лена» из 512×512 пикселей (рис. 3).

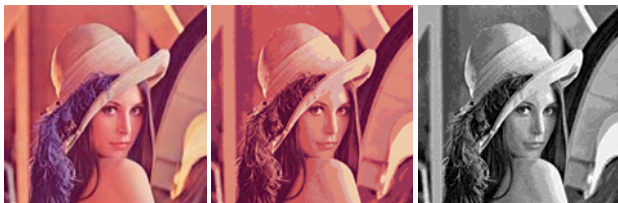


Рис. 3. Цветовое изображение (слева) и его квазиоптимальное приближение в исходном (в центре) и в инвариантном (справа) представлениях

Демонстрируемые на рис. 3 представления получены методом Уорда сначала снижением числа цветов в изображении с 148279 до 3434 за счет снижения числа тонов в каждой цветовой компоненте исходного изображения до 25. Затем, повторным применением метода Уорда для кластеризации трехмерных пикселей рассчитаны 3434 квазиоптимальных приближения изображения и столько же инвариантных представлений. Из них на рис. 3 показано приближение изображения в 8 цветах вместе с соответствующим инвариантным представлением, которое является серым:

$$H_{v,w}^u = H_{u,v}^w = H_{u,v}^w, \quad k = 1, 2, \dots, 8, \quad (12)$$

что не является тривиальным свойством, так как оно нарушается, например, при преобразовании одной из исходных цветовых компонент в негатив. Тем не менее, как показывает эксперимент, совпадение компонент инвариантного представления характерно для начального числа кластеров и согласуется с тем, что образы одних и тех же объектов в различных цветовых компонентах помечаются одинаковыми метками. Обнаруженное свойство планируется проиллюстрировать в докладе на примере выделения объектов на дистанционном изображении из 15 млн. пикселей.

6. ЗАКЛЮЧЕНИЕ

Скоростные вычисления с миллионами приближений изображения выполняются в ограниченном объеме оперативной памяти благодаря интенсивному применению структуры данных динамических деревьев Слэйтора-Тарьяна, которые не имеют альтернатив при программировании операций с произвольными множествами пикселей [11, 12], но, оптимизируя вычисления, никак не влияют на прозрачный смысл алгоритмов. Для содержательной разработки, развития и обоснования программно-алгоритмического обеспечения гораздо более важна интерпретация инвариантного представления как массива кодов информации, содержащихся в пикселях изображения [11]. Упомянутая интерпретация позволяет предложить целочисленную оценку количества информации в пикселе изображения и сопоставить ее с классическими оценками, что планируется обсудить в дальнейших работах.

7. ССЫЛКИ

[1] Абламейко С.В., Недзьведь А.М., Белоцерковский А.М., Руцкая Е.А. Сегментация трехмерных изображений компьютерной томографии на основе гло-

бально-локальной информации // Вестник Белорусского государственного университета. Сер. 1. Физика. Математика. Информатика, 2009. – № 1. – С. 58 - 64.

- [2] Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: классификация и снижение размерности. – М.: Финансы и статистика, 1989. – 607 с.
- [3] Arifin A.Z., Asano A. Image segmentation by histogram thresholding using hierarchical cluster analysis // Pattern Recogn. Letters 2006. Vol. 27. №. 13. – P. 1515 - 1521.
- [4] Бугаев А.С., Хельвас А.В. Поисковые исследования и разработка методов и средств анализа и автоматического распознавания потоковой информации в глобальных информационных системах. Шифр «Лацкан» // Отчет по НИР – М.: МФТИ, 2001. – Т 1. – 140 с.
- [5] Визильтер Ю.В., Желтов С.Ю. Проблемы технического зрения в современных авиационных системах // Механика, управление и информатика, 2011. – № 6. – С. 11 - 44.
- [6] Gurevich I.B. and Yashina V.V. Descriptive approach to image analysis: image formalization space // Pattern Recogn. Image Anal.: Adv. Math. Theory Appl. 2012. Vol. 22, №.4. – P. 495 - 518.
- [7] Crisp D.J., Tao T.C. Fast Region Merging Algorithms for Image Segmentation // The 5th Asian Conf. on Computer Vision (ACCV2002), Melbourne, Australia, 23 – 25 January 2002. – P. 1 - 6.
- [8] Dvoenko S.D. Clustering of a set of objects// Pattern Recognition and Information Processing (PRIP-2007) / Proc. of the 9th Int. Conf., Minsk, 2007. Vol. 1. – P. 93 - 97.
- [9] Dvoenko S.D. Meanless k-means as k-meanless clustering with the bi-partial approach // Pattern Recognition and Information Processing (PRIP'2014) / Proc. of the 12th Int. Conf., Minsk, 2014. – P. 50 - 54.
- [10] Двоенко С.Д. Неиерархический дивизимный алгоритм кластеризации // Автоматика и телемеханика, 1999. №4, — С. 117–124.
- [11] Харинов М.В. Запоминание и адаптивная обработка информации цифровых изображений; под ред. Р.М. Юсупова. – СПб: Изд-во С.-Петербург. ун-та, 2006. – 138 с.
- [12] Харинов М.В. Обобщение трех подходов к оптимальной сегментации цифрового изображения // Труды СПИИРАН, 2013. – Вып. 2 (25). – С. 294 - 316.
- [13] Koepller G., Lopez C., Morel J. A Multiscale Algorithm for Image Segmentation by Variational Method // SIAM J. on Numerical Analysis, Vol. 31, 1994. – № 1. – P. 282 - 299.
- [14] Mumford D., Shah J. Boundary detection by minimizing functionals, I // Proc. IEEE Comput. Vision Patt. Recogn. Conf., San Francisco, 1985. – P. 22 - 26.
- [15] Otsu N.A. Threshold Selection Method from Gray-Level Histograms // IEEE Trans. on systems, MAN, and CYBER-NETICS, January 1979. Vol. SMC-9, № 1. – P. 62 - 66.
- [16] Ward J.H., Jr. Hierarchical grouping to optimize an objective function. // J. Am. Stat. Assoc. 1963. Vol. 58, Issue 301. – P. 236 - 244.

Об авторах

Харинов Михаил – старший научный сотрудник СПИИРАН. E-mail: khar@iias.spb.su.