

МЕТОДЫ ОБРАБОТКИ ИЗОБРАЖЕНИЙ, ОСНОВАННЫЕ НА СЕГМЕНТАЦИИ

Маркин А.Е.

*Маркин Андрей Евгеньевич - студент магистратуры, бакалавр,
кафедра робототехники и мехатроники,
отдел магистратуры,
Донской государственной технической университет, г. Ростов-на-Дону*

1. Сегментация

Мозг человека, получая изображения от глаз, предоставляет его в качестве однородного объекта, не разбивая его на точки в матрице. И уже далее обрабатывает полученный цельный объект. Данный процесс (объединение точек (слияние)) получил название – сегментация. Сложностей при обработке у человека не происходит, все действия происходят на уровне подсознания. А при компьютерной обработке изображения реализация происходит программно. На начальных этапах анализа изображения происходит сегментация. Основное воздействие на точность, скорость, а также последующую релевантность обработки изображения несет качество сегментации.

Интенсивные исследования в области сегментации изображения производились и ведутся отечественными и зарубежными учеными. В рамках заданной области исследования хотелось бы выделить работы таких авторов, как Баяковский Ю.М., Богуславский А.А., Б. Рассел, Дж. Малик, М. Андретто, Юй-Ли, Андреев Ю.С., Юй-Цзинь Чжан, Дж. Ши, П. Виола, М. Джонс, Вежневцев В.П., Казанов М.Д., Сергеев В.В.

Уже было разработано большое количество подходов к сегментации изображения, но до сих пор нет ни одного универсального способа. Применимость методов сегментации зависит от текстуры, качества, размера, происхождения, предмета исследования, а также множества различных параметров. Исходя из запросов, применяются разные способы. Зачастую выбранный метод сегментации применим только для конкретного типа исследования, как правило, методы узкоспециализированы и не могут быть использованы для изображений, несоответствующих «типичным случаям».

2. Методы, основанные на сегментации

Важный этап в задачах распознавания образов является формирование входного вектора, то есть набора важных признаков (параметров), на основе которого возможно осуществить обработку информации и определить наличие тех или иных образов в массиве данных. В задачах анализа изображений такими параметрами считают наборы контуров и однородных областей, а также данные о них, периметр, площадь, центры масс и оси симметрии выделенных областей и контуров. Из этого следует, что важнейший этап к распознаванию образов – сегментация, то есть деление изображения на однородные области по критерию [2].

2.1. Иерархическая кластеризация

Методы, основанные на поочередном объединении (или разъединении) кластеров согласно принципу близости друг к другу, относят к иерархическим. Иерархические методы разделяют на две стратегии, на агломеративные и дивизионные. Агломеративная стратегия подразумевает объединение исходных данных в кластеры, а дивизионная предполагает разделение.

При агломеративной стратегии кластеризации растрового изображения каждая точка представляет собой отдельный кластер, затем друг с другом объединяются кластеры, имеющие похожие пространственно-цветовые характеристики. В итоге, все пиксели могут объединиться в один кластер, но, как правило, этого не требуется, поэтому на определённом этапе необходимо прекратить процесс. Для остановки включается порог «похожести» кластеров, по достижению которого объединение кластеров прекращается [4]. Другим способом остановки кластеризации считается оценка уровня различия между кластерами, как только этот параметр начинает уменьшаться, кластеризация останавливается на данном этапе.

Дивизионный подход подобен агломеративному только с различием, что кластеризация идёт в обратном порядке: первоначально все точки объединены в один кластер, а потом происходит их поочередное разделение. Данный способ кластеризации используется реже.

Для иллюстрации иерархических способов кластеризации применяют дендрограммы. Дендрограмма является специализированным видом диаграмм. Объекты на дендрограмме располагаются по иерархическим уровням, чтобы подчеркнуть их взаимное сродство.

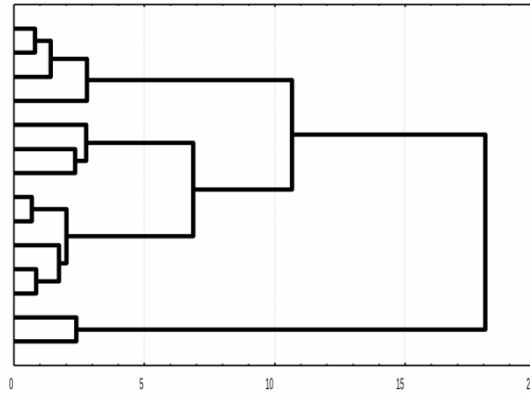


Рис. 1. Дендрограмма

2.2. Нейросетевые методы

Задачу сегментации можно решить при помощи самоорганизующихся карт Кохонена. Самоорганизующаяся карта Кохонена - это нейронная сеть с самообучением, выполняющая задачу визуализации и кластеризации. Данную идею сети предложил финский учёный в области машинного обучения, Т. Кохоненом. Этот подход является способом проецирования многомерного пространства в пространство с более низкой размерностью (чаще всего, двумерное), используется при решении задач моделирования, прогнозирования и др. Самоорганизующаяся карта состоит из узлов (нейронов). Любой из узлов описывается двумя векторами. Первый - вектор веса m , имеет размерность, что и входные данные. Второй – вектор координат узла на карте r . Как правило, узлы располагают в вершинах постоянной решётки с квадратными или шестиугольными ячейками.

Размерность входных данных - известна, по ней строится начальный вариант карты. В процессе обучения векторы веса узлов приближаются к входным данным. Для каждого исследования выбирается узел подобный вектору веса, и значение его вектора веса близится к наблюдению. Также к наблюдению приближаются векторы веса нескольких узлов, расположенных рядом, из этого следует, если во множестве входных данных два наблюдения были похожи, на карте им будут соответствовать близкие узлы. Повторяющийся процесс обучения, перебирающий входные данные, останавливается, когда достигается заданная аналитиком погрешность или по совершении заданного количества итераций. Процесс состоит из этапов: инициализации карты, то есть первоначального задания векторов веса для узлов и цикла; выбора следующего наблюдения; нахождения для него лучшей единицы соответствия (best matching unit, BMU, или Winner) – узла на карте, вектор веса которого меньше всего отличается от наблюдения; определения количества соседей BMU и обучения — изменения векторов веса BMU и его соседей с целью их приближения к наблюдению; определения ошибки карты.

При сегментации изображения входными параметрами будут служить координаты точек на изображении и положение их цвета в цветовой модели [1].

2.3. Выделение границ областей

Иной подход к сегментации основан на выделении контуров из изображения. Границы объектов определяются перепадами яркости или изменением цвета. Чтобы определить границы, вычисляется градиент яркости для каждой точки. Вычисление значений градиента происходит по матрице Собела.

Матрица

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

- для вычисления производной dC_x по оси x и

$$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

- для вычисления производной dC_y по оси y .

Также существуют другие матрицы, Робертса:

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

и

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

и Привита:

$$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

и

$$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

При вычислении цвета в RGB для каждого пикселя исходного изображения составляется матрица A. Размер матрицы A 3x3 из значений соответствующих цветов всех смежных пикселей. После этого происходит умножение каждого элемента на значение каждого цвета из матрицы A на соответствующую матрицу Собела для получения значений производных по направлениям x и y.

По направлению X

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

$$A = \begin{matrix} 124 & 191 & 242 \\ 168 & 205 & 249 \\ 149 & 242 & 255 \end{matrix}$$

По направлению Y

$$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

$$\begin{aligned} |grad|_x &= -1 * 124 + 0 * 191 + 1 * 242 - \\ &- 2 * 168 + 0 * 205 + 2 * 249 - \\ &- 1 * 149 + 0 * 242 + 1 * 255 \\ |grad|_y &= -1 * 124 - 2 * 191 - 1 * 242 + \\ &0 * 168 + 0 * 205 + 0 * 249 + \\ &1 * 149 + 2 * 242 + 1 * 255 \end{aligned}$$

Пример вычисления градиента цвета.

Модуль градиента в каждой точке рассчитывается по:

$$|grad| = \sqrt{dC_x^2 + dC_y^2} = \sqrt{dr_x^2 + dg_x^2 + db_x^2 + dr_y^2 + dg_y^2 + db_y^2} \quad (1)$$

Чтобы уменьшить влияния помех задаётся пороговое значение, ниже которого градиента приравниваются к нулю.

Изображение модулей градиентов считают картой границ (рис. 2).



Рис. 2. Изображение, модули градиентов его яркости и выделенные контуры

Затем для каждой полученной границы находится контур максимума градиента, а остальные области отсекаются. Замкнутые контуры можно считать выделенными сегментами изображения [6].

Близким к алгоритму по матрице Собела, является алгоритм водоразделов. Этот подход подразумевает вычисление градиента яркости изображения заранее. Если изобразить в виде рельефа, то низменности будут соответствовать однородным участкам изображения, а хребты - областям сильного изменения яркости, то есть границам. Такая аналогия позволяет рассматривать однородные области как водосборные бассейны, а границы - как водоразделы (рис. 3).

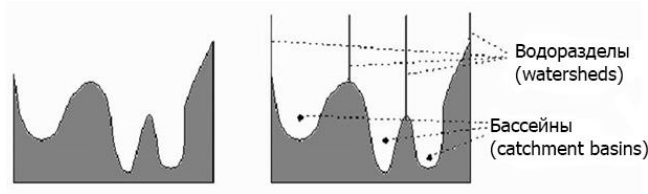


Рис. 3. Алгоритм водораздела

Причём водоразделы, обязательно замкнутые линии, так как они ограничивают водосборные бассейны.

Данный алгоритм чувствителен к шуму, поэтому для корректной сегментации важно задавать пороговое значение, снижение ниже которого должны отбрасываться [1].

2.4. Графовые методы

В случае с графовым методом, изображение представляется как неориентированный граф. Пиксели – вершины графа, ребра – соединение между пикселями, а вес ребер определяется схожестью пикселей. Когда вес ребер обратно пропорционален расстоянию между вершинами, и выполняется разбивка графа по выбранному признаку, в таком виде представляется граф. Либо разрез графа происходит так, чтобы суммарный вес в подграфе был минимальным. Полученные подграфы объединяют пиксели одного сегмента [5].

2.5. Метод mean-shift

Метод среднего сдвига (mean-shift) был предложен в 1975 году учеными Фукунагой и Хостетлером. Преимущество заключается в отсутствие необходимости в изначальных знаниях о количестве кластеров или о каких-либо их признаках. Задается функция оценки плотности (kernel density estimation). Затем для множества кластеризуемых точек строится функция градиента ядра оценки плотности и задаётся вектор среднего сдвига как вектор наибольшего увеличения градиента. Затем выполняются последовательные сдвиги точек на получаемые на каждом этапе векторы среднего сдвига до получения сходимости. Соответственно, точки, пришедшие в одну и ту же точку сходимости, считают элементами одного и того же кластера.

Для n точек x_1, x_2, \dots, x_n в d -размерном пространстве R^d задается функция оценки плотности :

$$\hat{f}_n(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) \quad (2)$$

где h - размер окна, $K(x)$ – функция ядра. Функция ядра подчиняется следующим условиям:

$$\int K(x)dx = 1; \int xK(x)dx = 0; \lim_{\|x\| \rightarrow \infty} \|x\|^d k(x) = 0. \quad (3)$$

Для широко распространенного класса симметрических ядер

$$k(x) = c_{k,d} k(\|x\|^2), \quad (4)$$

где $k(x)$ - функция профиля, $c_{k,d}$ - нормировочная константа.

Два наиболее распространённых вида ядер – это ядро Епашникова, где

$$k_E = \begin{cases} 1-x, & \text{если } 0 \leq x \leq 1 \\ 0, & \text{если } x > 1 \end{cases} \quad (5)$$

и, соответственно,

$$K_E(x) = \begin{cases} \frac{1}{2} C_d^{-1} (d+2)(1-\|x\|^2), & \text{если } \|x\| \leq 1 \\ 0, & \text{в противном случае} \end{cases} \quad (6)$$

Ядро Гаусса, где

$$k_N(x) = \exp\left(-\frac{1}{2}x\right), x \geq 0 \quad (7)$$

и, соответственно,

$$k_N(x) = (2\pi)^{-d/2} \exp\left(-\frac{1}{2}\|x\|^2\right). \quad (8)$$

Теперь функции ядра оценки можно представить по формуле:

$$\hat{f}_{h,k}(x) = \frac{C_{k,d}}{nh^d} \sum_{i=1}^n K\left(\left\|\frac{x-x_i}{h}\right\|^2\right). \quad (9)$$

Определяем функцию градиента:

$$\nabla \hat{f}_{h,k}(x) = \frac{2C_{k,d}}{nh^{d+2}} \sum_{i=1}^n (x-x_i)g\left(\left\|\frac{x-x_i}{h}\right\|^2\right), \quad (10)$$

где $g = -k'(x)$.

Для вычисления среднего сдвига воспользуемся формулой:

$$m(x) = \frac{\sum_{i=0}^n x_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=0}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)} - x, \text{ где } g = -k'(x). \quad (11)$$

Точки последовательно сдвигаются (рис. 4) на получаемые вектора до тех пор, пока не приходят в центры сходимости. Все точки, пришедшие в один и тот же центр, объединяются в кластер. Единственным параметром алгоритма является размер окна h [1].

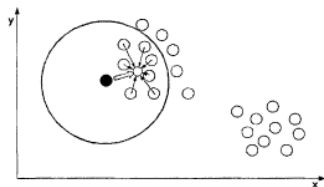


Рис. 4. Средний сдвиг

При сегментации изображения встает задача вычисления расстояния между отдельными пикселями. Соответственно, эту задачу удобно разбить на два этапа: определения пространственного расстояния и определение цветового расстояния. Таким образом, изображение можно представить в виде двухмерной сетки из p -мерных векторов. Где $p=1$ для полутоновых изображений и $p=3$ для цветных. Двухмерная сетка будет описывать пространственную область (*spatial domain*), то есть взаимное положение пикселей, а p -мерный вектор - область значений (*range domain*), то есть цветовое пространство. Соответственно параметр h можно разбить на ширину окна для пространственной области h_s (*spatial bandwidth*) и на ширину окна области значений h_r (*feature (range) bandwidth*) [3].

Список литературы

1. Zhang Yu-Jin. Advances in Image and Video Segmentation/ IRM Press, 2006.
2. Luo Y., Zhang Y.J., Gao Y.Y. et al. Extracting meaningful region for content-based retrieval of image and video. // SPIE, 2001.
3. Ridder D., Kittler J., Lemmers O., Duin R. The Adaptive Subspace Map for Texture Segmentation. // Proc. 2000 Int'l Conf. Pattern Recognition. Sept., 2000. Vol. Pp. 216-220.
4. Roman-Roldan R., Gomez-Lopera J.F., Atae-allah C., Martinez-Aroza J.. & Luque Escamilla PL. A measure of quality for evaluating methods of segmentation and edge detection. // Pattern Recognition, 2001.
5. Cour T., Benzeit F., Shi J. Spectral Segmentation with Multiscale Graph Decomposition // Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005. Vol. 2. Pp. 1124-1131.
6. Wilson D.L., Baddeley A.J., Owens R.A. A new metric for gray-scale image comparison. // International Journal Computer Vision, 1997. Vol. 24. Pp. 5-17.