

Автоматическое определение размера суперпикселей при сегментации изображений методом Quick Shift

КОЛОСОВСКИЙ МАКСИМ АЛЕКСАНДРОВИЧ

Алтайский государственный технический университет (Барнаул), Россия
e-mail: maxim.astu@gmail.com

КРЮЧКОВА ЕЛЕНА НИКОЛАЕВНА

Алтайский государственный технический университет (Барнаул), Россия

Сегментация изображения означает разбиение изображения на множество покрывающих его областей. Цель многих задач анализа изображений заключается в сегментации на области, с которыми связана существенная для данной задачи информация. Например, при обработке спутниковой фотографии областями интереса могут быть области городских построек, лесных массивов и сельскохозяйственных полей [1], в более сложной задаче нужно будет не просто понимать тип местности, но и выделять конкретные объекты: здания, автомобили, людей и т.д. Такая сегментация, разделяющая объекты на изображении, называется семантической.

Для выполнения семантической сегментации удобно перейти от пиксельного представления к более высокоуровневым структурам – суперпикселям. Суперпикселями называются области изображения, объединяющие пиксели с похожими признаками (интенсивность, цвет, текстура и т.д.). Таким образом, мы переходим к более компактному и удобному представлению изображения. Представление будет тем компактней, чем более крупные суперпиксели мы будем делать, и тем меньше будет их общее количество. Однако качественное разделение на суперпиксели подразумевает то, что в один суперпиксель не попадутся части сразу нескольких объектов, что ограничивает размеры суперпикселей сверху.

Для разных изображений максимальный возможный размер суперпикселей различен. Например, при сегментации аэроснимка, на котором мы разделяем поля пшеницы, лес и озеро, суперпиксели могут быть очень больших размеров, т.к. искомые области будут занимать большие непрерывные участки. Однако если на изображении искомые объекты гораздо мельче, например, автомобили и отдельные постройки, то и суперпиксели должны быть меньшего размера. Заметим также, что размер суперпикселя зависит не только от изображения, но и от задач семантической сегментации, например, в одних случаях будет достаточно выделить область построек, а в других придется выделять каждую постройку как отдельный объект. Цель данной работы найти механизм определения оптимального размера суперпикселей по данному изображению.

На данный момент уже изобретено много алгоритмов выделения суперпикселей, в данной работе мы остановились на алгоритме QuickShift [2], т.к. он имеет параметр, регулирующий, насколько близкие пиксели будут попадать в один суперпиксель, а значит и косвенно определяющий размер суперпикселей. Также отметим, что этот алгоритм достаточно широко используется, т.к. он гораздо быстрее своего точного и медленно прародителя MeanShift, однако не сильно уступает ему в качестве. Мы используем конкретную реализацию QuickShift, описанную в [3] и распространяемую в открытой библиотеке VLFeat [4].

Общая идея данной реализации этого алгоритма в том, что мы будем кластеризовать пиксели изображения, взяв в качестве признаков компоненты цвета пикселя и его координаты на изображении. Для каждого пикселя оценим, как много похожих пикселей вокруг него (назовем это плотностью), для каждого пикселя найдем ближайших к нему пикселей с большей плотностью, построив этими связями ориентированное дерево. Если задать порог максимального расстояния для объединения в кластеры, то дерево распадется в лес, в котором деревья будут соответствовать суперпикселям. При помощи этого порога мы можем регулировать размер суперпикселей: чем больше порог, тем более крупные суперпиксели получим.

Проделав ряд экспериментов, мы определили, что подход, при котором дерево разбивается ограничением максимальной длины ребра, требует различного порога в зависимости от плавности перехода: при резком переходе можно использовать большой порог, однако если эта же граница будет размыта, то переход может не найтись, т.к. переходы между соседними пикселями не такие уже длинные. Для улучшения метода был предложен другой подход в разбиении дерева в лес. Граница будет определяться не порогом перехода между пикселями, а максимальным разбросом параметров внутри суперпикселя, что улучшает определение границ суперпикселей.

Литература

1. Л. Шапиро, Дж. Стокман. Компьютерное зрение. М.: Бином. Лаборатория знаний, 2006.
2. Quick Shift and Kernel Methods for Mode Seeking, A. Vedaldi and S. Soatto, in Proceedings of the European Conference on Computer Vision (ECCV), 2008.
3. Really quick shift: Image segmentation on a GPU, B. Fulkerson and S. Soatto. In Proceedings of the Workshop on Computer Vision using GPUs. September 2010.
4. A. Vedaldi and B. Fulkerson. VLFeat: An open and portable library of computer vision algorithms. <http://www.vlfeat.org/>, 2008.