

Пирамидальный алгоритм сегментации изображений

П. А. Чочиа

Институт проблем передачи информации им. А. А. Харкевича РАН, Москва, Россия

Поступила в редколлегию 24.03.2010

Аннотация—Исследуется задача сегментации изображения (разбиения на однородные области) на основе использования цветовых и текстурных различий областей. Для ее решения предлагается двухуровневый иерархический пирамидальный алгоритм сегментации. Критерием однородности является оценка близости элементов и областей изображения в объединенном текстурно-цветовом пространстве признаков. Вводится и исследуется метрика в этом пространстве. Результаты проверяются на множестве тестовых изображений различных типов.

1. ВВЕДЕНИЕ

Одним из фундаментальных этапов анализа видеоинформации является сегментация изображения, означающая разбиение поступающего изображения на множество областей, однородных в соответствии с теми или иными критериями, и тем самым ассоциируемых с объектами наблюдаемой сцены или их частями. Результатом сегментации является получение карты областей (сегментов) изображения. Мы рассматриваем задачу текстурно-цветовой сегментации, предполагая при этом, что исходные данные представлены в виде обычного цветного RGB изображения, а однородность областей будет определяться на основе оценок их яркостных, цветовых и текстурных характеристик.

Исследованию вопроса сегментации изображений посвящено большое количество работ, тем не менее, процедура сегментации остается одной из фундаментальных проблем анализа поступающей видеоинформации. В большинстве случаев алгоритмы сегментации строят либо на использовании базовых свойств изображений, таких как однородность на связных областях и разрывность на их границах [1–3], либо на преобразовании сигнала в пространство признаков и классификации в этом пространстве [4]. Среди подходов к построению алгоритмов сегментации можно выделить следующие четыре класса, из которых два первых в основном ориентированы на однокомпонентные (монохромные, черно-белые) изображения, тогда как третий и четвертый классы допускают анализ многокомпонентных данных.

1. Разбиение изображения проведением контуров. Заключается в обнаружении контуров с помощью тех или иных операторов, а также в их прослеживании, связывании, и составлении из них границ областей. Известно много различных алгоритмов обнаружения контурных границ. Большинство из них основано на том или ином способе фильтрации, преобразующем изображение в некоторый двумерный массив, значения которого соответствуют вероятности нахождения контура в соответствующей точке. Большинство контурных операторов основаны на вычислении первой производной (операторы Робертса, Собеля [5, 6] и др. [1–3, 7–9]) или второй производной (оператор Лапласа [1–3]). Результаты их применения выглядят неплохо с точки зрения визуальной оценки, однако не удовлетворяют ряду формальных критериев, в частности требованиям непрерывности и минимальной толщины контурных линий. Другой известный подход основан на обнаружении так называемых точек перехода через ноль второй производной (“zero crossing”) [9]. Он обеспечивает выполнение указанных требований, однако

в результате возникает много несвязанных между собой пятен, недостаточно хорошо соответствующих реальным контурам. Имеются также некоторые другие методы и алгоритмы [1–3, 7, 10]. Альтернативой алгоритмам фильтрации является прослеживание контуров [5, 8]. Для решения задачи сегментации такой подход имеет ряд преимуществ, среди которых: возможность получения карты с непрерывными контурными линиями; прослеживание линии в случае небольших нарушений контурных перепадов (разрывов); проведение контурных линий минимальной толщины. Из недостатков можно отметить: сильная зависимость от шумов на изображении, необходимость предварительного выбора начальных точек прослеживания контуров, определенные сложности в точках пересечения или ветвления контуров, необходимость дополнительного анализа и фильтрации полученной карты контуров.

2. Морфологический подход. Из подобных методов наиболее известна сегментация по водоразделам [13, 14]. Она заключается в интерпретации гладких областей как локальных бассейнов, а контуров между ними как водоразделов. Такой подход достаточно эффективен в случае монохромных изображений, но слабо применим к цветным и текстурным изображениям.

3. Разбиение изображения на однородные области. Здесь в первую очередь необходимо упомянуть пороговые методы, использующие как глобальные, так и адаптивные пороги [1, 11, 12]. Подобные методы просты в реализации и вычислительно эффективны, но не универсальны и применимы лишь в случае монохромного сигнала. Известен ряд методов, основанных на процедуре выращивания областей [15, 16], которая состоит в группировке элементов и мелких областей изображения в более крупные, начиная из так называемых “центров кристаллизации”. Проблемы данного метода — выбор критериев близости и подходящего алгоритма остановки процесса выращивания областей [17]. Альтернативой выращиванию является методы, основанные на первичном разбиении изображения на множество малых непересекающихся областей и осуществлении последующего их слияния и разделения [18].

4. Классификация в пространстве признаков. Она заключается в выборе отображения набора входных данных в некоторое многомерное пространство признаков и последующее решение классической задачи кластеризации — разбиении выбранного пространства на классы, базируясь на плотности распределения в нем [4, 19]. Данные методы особенно удобны, когда заранее известно количество классов объектов, содержащихся в изображении. Также как правило, предполагается, что обнаруженные в пространстве признаков кластеры отображают связанные области [20]. Классическая процедура сегментации может дать неверные результаты, если отображения различных областей исходного изображения в пространстве признаков заметно перекрываются. Для преодоления данной проблемы предлагается расширение пространства признаков, например, путем добавления значений координат изображения [4].

Возможны два пути улучшения результатов сегментации. Первый состоит в нахождении более чувствительных и соответствующих задаче признаков, а второй — в отыскании иного подхода к разработке алгоритма. В настоящей работе используются оба пути. В части выбора признаков исследуется комбинация яркостно-цветовых и текстурных признаков с построением многокомпонентного пространства и выбором метрики в данном пространстве. Для алгоритма сегментации выбран иерархический подход, основанный на преобразовании изображения в граф первичных кластеров (для этого используется пирамидальный алгоритм) и последующем упрощении карты первичных кластеров.

Характеристики “цвет” и “яркость” являются измеримыми и удобными признаками для классификации [22, 23], в то время как понятие “текстура” является достаточно расплывчатым и нечетко формулируемым [21]. Обобщенно говоря, можно сказать, что под текстурой понимают смесь большого числа в той или иной степени упорядоченных и похожих элементов или образов (примитивов), когда ни один из этих конкретных образов не имеет принципиального значения, и когда наблюдателя интересует только общий вид и эффект, производимый

достаточно большими областями. При описании текстур обычно используются следующие характеристики: регулярность, направленность, крупность, повторяемость (эффективный период или размер), форма, цвет и контраст примитивов, и некоторые другие [23, 24]. Одни из них могут быть четко сформулированы, тогда как другие нет.

Проблема текстурного анализа исследуется много лет, и основная трудность заключается в отыскании подходящих характеристик, описывающих текстуру. В анализе текстур имеются три основных подхода: структурный, спектральный и статистический. Структурный подход основан на предварительном выделении примитивов и последующей оценке их размещения, тем самым он лучше подходит для анализа регулярных текстур. Спектральные методы используют свойства Фурье-спектра и наиболее подходят для отыскания периодичности в изображении. Более универсальным и адаптируемым является статистический подход. При этом в качестве статистических характеристик чаще всего используются: матрица смежности, автокорреляция, разности уровней серого, длины серий по направлениям, локальная частота флуктуаций яркости [23–27]. Иногда используется иерархический подход, позволяющий оценивать характеристики на разных масштабах. После 1980-х годов возник интерес к многоканальному и многомасштабному [28] текстурному анализу, а еще позже — к использованию вейвлетов [29, 30] и Габоровских фильтров [30–33], тесно связанных с пирамидальным представлением изображения [30].

2. СТРУКТУРА АЛГОРИТМА СЕГМЕНТАЦИИ

Задача ставится как текстурно-цветовая сегментация изображения. Исходные данные являются многокомпонентными; поэтому подход к построению алгоритма сегментации был выбран из множества двух последних классов методов: разбиение изображения на однородные области путем анализа сходства в пространстве признаков. Главным критерием при этом является расстояние между проекциями элементов в пространство признаков, на основе чего осуществляется объединение соседних элементов, или же проведение границы между ними. Хотя в данном алгоритме сегментации существенно и используется пространство признаков, однако в отличие от классического решения задачи кластеризации анализируется не плотность распределения, а оценивается близость точек в этом пространстве. Тем самым одним из важнейших вопросов становится выбор метрики в нем.

Рассматриваемый алгоритм содержит два уровня (стадии). На первом изображение разбивается на множество мелких кластеров с помощью пирамидального алгоритма, а на втором происходит их окончательное слияние. Принципиальная структура алгоритма следующая.

Первоначально строится объединенное N -мерное пространство признаков, содержащее преобразованные цветовые и выделенные текстурные характеристики, и осуществляется переход к так называемому векторному изображению, которое является двумерным массивом N -мерных векторов в этом пространстве. Полученное векторное изображение обрабатывается пирамидальным алгоритмом, преобразующим его в первичный граф кластеров, в котором каждый элемент изображения оказывается соотношенным тому или иному кластеру — соответствующему узлу графа. Тем самым, каждый узел графа отображает соответствующий кластер, а ребра графа — возможные связи между соседствующими кластерами. Необходимые характеристики кластеров и их связей с соседними кластерами объединяются в список кластеров, используемый на второй стадии. Полученный на данной стадии граф кластеров представляет лишь предварительный результат сегментации, поскольку размеры первичных кластеров оказываются значительно меньше размеров объектов на изображении, а их число, соответственно, значительно больше числа интересующих объектов.

На второй стадии попарно анализируются соседние кластеры, имеющие общие границы, осуществляется слияние тех кластеров, параметры которых достаточно близки, и происхо-

дит упрощение построенного графа. Поскольку слияние кластеров приводит к изменению их средних параметров, то этот процесс выполняется по итеративной схеме. Критерием окончания процедуры слияния служит отсутствие пар кластеров с расстоянием, меньше некоторого заданного порога. Результатом данной операции и является сегментированное изображение.

3. ТЕКСТУРНО-ЦВЕТОВОЕ ПРОСТРАНСТВО ПРИЗНАКОВ

Полное текстурно-цветовое пространство признаков является объединением двух более простых подпространств — цветовых и текстурных признаков.

Подпространство цветовых признаков. В качестве цветовых признаков используются следующие характеристики: яркость, цветность и насыщенность (BHS). Данное пространство цветовых признаков совпадает с обычным RGB цветовым пространством с точностью до координатного преобразования. Существуют различные способы преобразования значений координат из пространства RGB в пространство BHS (см., например, [2]); мы воспользовались следующими формулами. Пусть r, g, b — координаты точки в исходном RGB пространстве, и $u = \min(r, g, b)$. Тогда

$$\begin{aligned} B &= \sqrt{(r^2 + g^2 + b^2)/3} \\ H &= \begin{cases} (2\pi/3)(b-u)/(g+b-2u) + \pi/3, & \text{если } r = u, \\ (2\pi/3)(r-u)/(b+r-2u) + \pi, & \text{если } g = u, \\ (2\pi/3)(g-u)/(r+g-2u) + (5\pi/3), & \text{если } b = u, \end{cases} \\ S &= 1 - u/(r+g+b). \end{aligned} \quad (1)$$

Подпространство текстурных признаков. Как показывает анализ литературы, в качестве текстурных характеристик используются самые различные оценки локальной изменчивости изображения. В настоящей работе использован многомасштабный набор оценок изменчивости по множеству направлений, которые вычисляются для каждой точки изображения на основе значений локальной вариации v и локального изменения (осцилляции) градиента g вдоль выбранных направлений. Такой подход близок известному методу разностей яркостей, обсуждаемому в обзорах [27, 34]. В качестве анализируемой информации используются значения яркости B получаемые согласно уравнению (1). Признаки вычисляются следующим образом.

Рассмотрим отрезок длиной $2L$, центрированный в точке (m, n) , и проходящий под углом $\alpha_k = k\pi/K$, где $(k = 0, \dots, K-1)$. Равномерно расположим на нем $2L+1$ точек (отсчетов), и каждой точке i отрезка $(-L \leq i \leq L)$ присвоим значение x_i , равное значению ближайшего элемента изображения. Пусть $d_i = x_{i+1} - x_i$ есть разность между соседними точками на отрезке (т.е. градиент в точке i вдоль направления отрезка). В качестве весовой функции выберем следующую: $q_i = \cos(i\pi/2L)/\sum \cos(i\pi/2L)$. Тогда взвешенные суммы подъемов и суммы спусков вдоль отрезка будут равны

$$v^+ = \sum_{i=-L}^{L-1} q_i d_i, \text{ если } d_i \geq 0, \quad \text{и} \quad v^- = \sum_{i=-L}^{L-1} q_i (-d_i), \text{ если } d_i < 0. \quad (2)$$

Дабы уменьшить влияние контурных перепадов на текстурные отклики, значение *локальной вариации по направлению* выбирается как *наименьшее* из v^+ и v^- :

$$v_k = \min(v^+, v^-). \quad (3)$$

Значение *осцилляции локального градиента* g определяется как число перемен знака градиента d_i на отрезке $(-L \leq i \leq L)$, при условии, что $d_i > \delta$, где δ — выбранный порог чувствительности. Произведение локальной вариации и осцилляции локального градиента по направлению

k дает предварительное значение текстурного признака: $\tilde{t}_k = v_k g_k$. Для выравнивания признаков высоко- и малоконтрастных текстур используется, как это предлагается в [31], функция гиперболического тангенса:

$$t_k = \tanh(\alpha \tilde{t}_k). \quad (4)$$

Полученное значение t_k считается *текстурным признаком* по направлению k в соответствующей точке (m, n) изображения. Значения t_k являются достаточно зашумленными; для повышения однородности можно рекомендовать использование фильтрации какого-нибудь вида, например, усреднения по локальному двумерному фрагменту размерами $h \times h$ точек.

Текстура, как очевидно, является масштабно-зависимой характеристикой. Для различения текстур разной периодичности требуется выделение признаков, соответствующих различным пространственным частотам. Чтобы воспользоваться для этого формулой (4), необходимо взять отсчеты с другим шагом, т.е., при сохранении числа точек равным $2L + 1$, соответствующим образом изменить размер самого отрезка. Так, увеличив в 2^n раз размер отрезка, диапазон частот увеличится до в 2^n раз более низких частот; влияние высокочастотной составляющей спектра можно подавить с помощью фильтрации (например, путем сглаживания). Другим путем изменения диапазона частот может являться предварительное масштабирование самого изображения в те же 2^n раз при сохранении размеров отрезков неизменными; такой подход представляется предпочтительным.

Таким образом, текстурные признаки вычисляются по формулам (2)–(4) для каждой точки изображения по всем направлениям ($k = 0, \dots, K - 1$) и масштабам $1, 1/2, 1/4, \dots, 2^{S-1}$, где K и S — общее число направлений и масштабных уровней. Таким образом, мы получаем набор v_k^s и g_k^s масштабно-зависимых значений локальной вариации и осцилляции локального градиента, а текстурные признаки будут равны

$$t_k = \tanh(\alpha v_k^s g_k^s). \quad (5)$$

Можно также оценить среднюю величину текстуры в точке следующим образом:

$$T = \sum_{s=1}^S w^s \sum_{k=0}^{K-1} t_k^s; \quad (6)$$

здесь w^s есть весовой коэффициент масштабного уровня s .

Суммарное число текстурных признаков для каждой точки изображения равно KS . В реальности достаточными представляются $K = 4$ и $S = 4$, т.е. всего формируется и анализируется 16 текстурных признаков. Объединенный набор текстурных и цветовых признаков для точки изображения составляет вектор признаков, а все возможное многообразие таких векторов — пространство признаков.

4. МЕТРИКА В ПРОСТРАНСТВЕ ПРИЗНАКОВ

Одним из важнейших аспектов при решении задачи сегментации изображения как выделения однородных областей, является вопрос “похожести” элементов и/или кластеров изображения, т.е. близости их признаков. Поскольку как элементы, так и кластеры могут быть отображены в одно и то же пространство признаков, то вопрос измерения близости может быть переформулирован как введение метрики в пространстве признаков. Важно заметить, что универсальной метрики не существует, ее нет даже для цветового подпространства. Метрика, предлагаемая в настоящей работе, разработана специально для задачи сегментации. Между цветовыми и текстурными признаками имеется существенное различие, поэтому было

принято решение использовать две отдельные метрики: чисто цветовую в цветовом подпространстве и чисто текстурную в текстурном подпространстве, а затем объединить их в общую текстурно-цветовую метрику.

Задача построения цветовой метрики известна достаточно давно, и обычно она решается поиском равноконтрастной цветовой метрики [2, 35, 36]; предлагается много различных метрик в цветовом пространстве, среди них хорошо известны $\mathbf{U}^*\mathbf{V}^*\mathbf{W}^*$, $\mathbf{S}\Theta\mathbf{W}^*$, \mathbf{Lab} , и т.д. Мы тестировали некоторые из них и пришли к выводу, что они дают близкие и во многих случаях удовлетворительные результаты. Тем не менее, было принято решение использовать новую метрику, близкую $\mathbf{S}\Theta\mathbf{W}^*$, но которая лучше соответствует задаче сегментации и позволяет достичь более точных результатов в темных областях изображения¹.

Цветовая метрика, используемая в алгоритме сегментации, задается следующей формулой:

$$D_c(x, y) = \left[w_B(B_x - B_y)^2 + w_H F\left(\frac{B_x + B_y}{2B_0}\right) F\left(\frac{S_x + S_y}{2S_0}\right) (\min(|H_x - H_y|, 2\pi - |H_x - H_y|))^2 + w_S F\left(\frac{B_x + B_y}{2B_0}\right) (S_x - S_y)^2 \right]^{1/2}. \quad (7)$$

По сути, эта формула ничто иное, как модифицированный вариант корня из суммы квадратов разностей значений яркости, цветности и насыщенности двух точек x и y в цветовом подпространстве признаков. B_i , H_i , и S_i ($i = x, y$) суть, соответственно, компоненты яркости, цветности и насыщенности, w_B , w_H , w_S , — весовые коэффициенты, B_0 and S_0 — пороги подавления, а $F(\cdot)$ — функция подавления, имеющая вид:

$$F(x) = \begin{cases} x, & \text{если } x < 1, \\ 1, & \text{если } x \geq 1. \end{cases} \quad (8)$$

Целью введения функции $F(\cdot)$ является подавление излишнего влияния разностей значений цветности и насыщенности в области малых значений яркости, а также значений цветности в области малой насыщенности.

Текстурная метрика задается формулой:

$$D_t(x, y) = \sum_{s=1}^S w^s \sum_{k=0}^{K-1} |t_k^s(x) - t_k^s(y)|, \quad (9)$$

где x и y — две точки текстурного подпространства, $t_k^s(x)$ и $t_k^s(y)$ — значения текстурных откликов уровня s и направления k в этих точках, а w^s — весовой коэффициент уровня s .

Объединенная текстурно-цветовая метрика определяется как взвешенная сумма расстояний в цветовом (7) и текстурном (9) подпространствах:

$$D(x, y) = w_c D_c(x, y) + w_t D_t(x, y). \quad (10)$$

Здесь x и y — две точки объединенного подпространства, являющиеся отображениями точек или кластеров изображения, а w_c и w_t — цветовой и текстурный весовые коэффициенты.

5. АЛГОРИТМ ПИРАМИДАЛЬНОГО ПРЕОБРАЗОВАНИЯ (ПЕРВИЧНАЯ СЕГМЕНТАЦИЯ)

Задачей стадии первичной сегментации является преобразование исходного векторного изображения в первоначальный набор кластеров небольшого размера, в котором каждый сформированный кластер характеризуется как собственными параметрами, так и параметрами связи

¹ Исследования по оптимизации выбранной метрики и алгоритма вторичной сегментации проводились при участии Д.В. Сушко

с соседними кластерами. Эту задачу удобно решать с помощью пирамидально-рекурсивного алгоритма [37, 38].

Пирамидальный алгоритм обработки изображения требует задания способа вычисления уменьшенного изображения уровня $n + 1$ на основе имеющегося изображения уровня n , и применения этой процедуры рекурсивно до предельного уменьшения размера изображения. Часто для этого выбирают элементарный домен некоторой формы, позволяющей плотно покрыть всю площадь изображения, и задают способ нахождения значения элемента следующего уровня по значениям элементов домена предыдущего уровня. Рекурсивное применение данной процедуры позволяет построить дерево (граф), в котором каждый элемент изображения некоторого уровня (кроме самого нижнего) является узлом, связанным с элементами домена предыдущего уровня, а число нисходящих связей определяется формой и размерами выбранного домена. В случае, когда в качестве домена выбран квадрат размерами 2×2 элемента, мы получаем так называемое *квадродерево* (Рис. 1,а).

Алгоритм пирамидальной сегментации осуществляет прямой (вверх) и обратный (вниз) проходы по квадродереву.

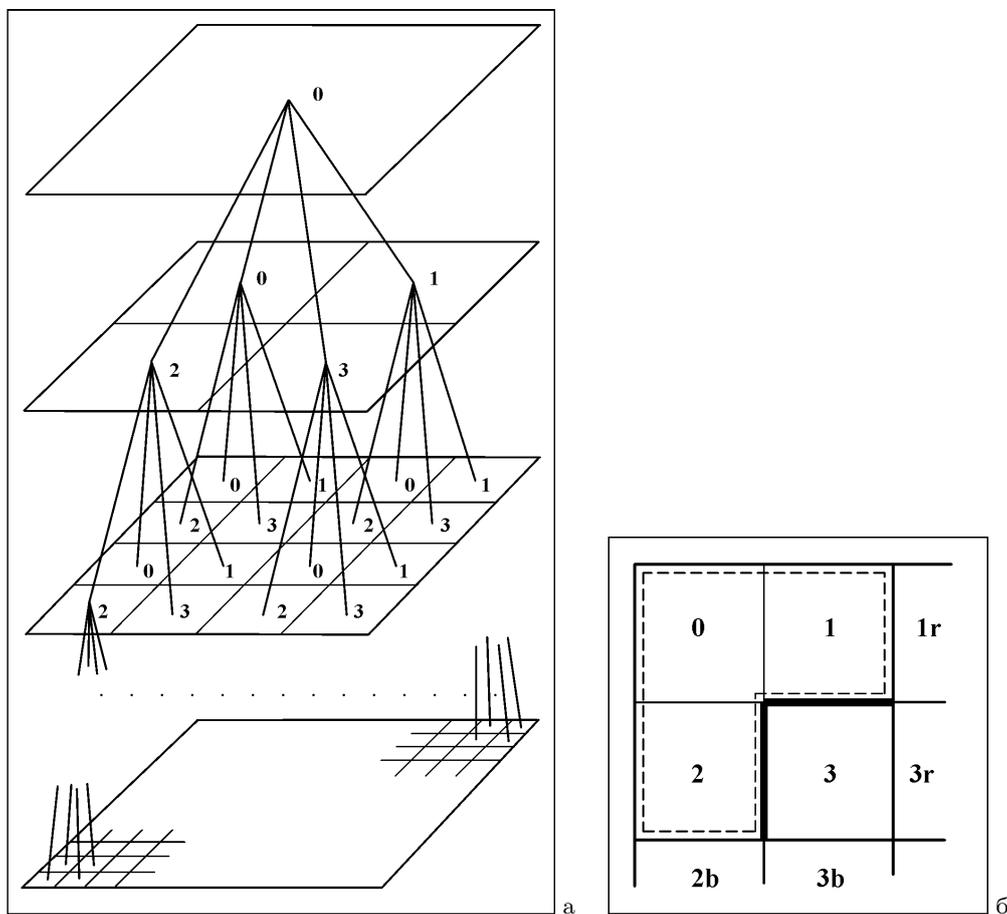


Рис. 1. Пирамидальная структура представления данных — квадродерево (а); домен из четырех элементов и его соседние элементы справа и снизу (б).

При **прямом проходе** вверх по квадродереву происходит рекурсивный анализ всех уровней пирамиды, начиная с самого нижнего (исходного изображения) и заканчивая верхним уровнем, состоящим из одного узла; одновременно с этим строится само квадродерево. На каждом шаге

на основе анализа четырех нижних узлов уровня n создается узел уровня $n + 1$ (см. Рис. 1,а), а в соответствующей структуре нового узла запоминается информация об узлах предыдущего уровня, соединенных с данным узлом, их средней яркости, наличии контуров. Таким образом, каждый узел является вершиной некоторого квадродерева, охватывающего расположенные под ним элементы изображения, и содержит информацию о поддеревьях предыдущего уровня.

Основной задачей пирамидального этапа сегментации является объединение соседних элементов, имеющих близкие признаки и не разделенных контуром. Эта процедура требует отслеживания контурных линий на всех уровнях пирамиды. Считается, что два вертикально или горизонтально соседствующих элемента разделены контуром в том случае, если расстояние между их отображениями в пространстве признаков (в смысле задаваемой формулой (10) объединенной текстурно-цветовой метрики) превышает некоторый заданный порог. Для второго и последующих уровней процедура обнаружения контуров также учитывает наличие контура на предыдущем уровне.

Анализ и объединение элементов домена размерами 2×2 (Рис. 1,б) состоит из следующих шагов.

1. Анализ наличия контурных перепадов между парами соседних элементов: **0-1**, **2-3**, **0-2** и **1-3**. Пары элементов (узлов графа), не разделенные контуром, считаются принадлежащими одному и тому же кластеру. На Рис. 1,б показан контур, проходящий между элементами **1** и **3**, а также между элементами **2** и **3**. Тем самым, элементы **0**, **1**, и **2** соответствуют некоторому одному кластеру, а элемент **3** — другому кластеру.

2. Выбор наиболее представительного набора элементов, принадлежащих одному кластеру нижнего уровня (в данном случае в такой набор входят элементы **0**, **1**, и **2**), и передача усредненной информации соответствующему узлу следующего уровня. Оставшийся элемент (номер **3** на рисунке) считается “подвешенным”, информация о нем наверх не передается и сохраняется только в самом описании текущего узла.

3. Создание карты контуров для следующего уровня. Это осуществляется с помощью анализа внутренних элементов самой четверки, плюс четверки соседних с ней элементов в последовательной развертке, т.е. справа и снизу; согласно примеру на Рис. 1,б будет проверяться наличие контурных перепадов между элементами **1-1r** и **2-2b**.

После окончания прямого прохода по всему квадродереву, каждый узел любого уровня (кроме корневого узла) будет либо принадлежать поддереву, начинающемуся на одном из верхних уровней, либо сам являться корнем некоторого поддерева.

При **обратном проходе** вниз по квадродереву происходит синтез первичных кластеров. На каждом из уровней вершинам новых поддеревьев (т.е. тем, которые оказались “подвешены” на шаге 2 прямого прохода) присваиваются новые номера кластеров; данные номера распространяются вниз на все элементы, связанные с данным поддеревом. Эта процедура рекурсивно распространяется на все дерево с верхнего уровня (вершины дерева) до нижнего (самого изображения); элементы изображения, получившие при этом одинаковые номера, образуют первичные кластеры (первичные сегменты) обрабатываемого изображения.

Процедура синтеза на каждом из узлов дерева состоит в следующем.

1. Если текущий узел не имеет присвоенного номера, он получает следующий свободный номер кластера.

2. Номер кластера, присвоенный текущему узлу, присваивается всем узлам следующего, более низкого, уровня, связанным с текущим узлом при прямом проходе.

Результатом данной процедуры является первичная карта кластеров сегментируемого изображения. Для дальнейшего анализа формируется граф соседства кластеров (каждому кластеру соответствует узел графа) и собирается нужная информация, описывающая свойства

каждого кластера (число элементов, объединенных в кластере, длина границы, средний цвет и т.д.). Ребра графа соединяют пары соседствующих кластеров, имеющих общую пространственную границу, и отображают характеристики их границы, например, длину общей части. Следует заметить, что одним из преимуществ использования предлагаемого пирамидального подхода к сегментации является тот факт, что он во многих случаях позволяет решить проблему разделения областей, имеющих границу с разрывами. То, как выглядят изображения после операции пирамидальной (первичной) сегментации, можно видеть на Рис. 2(б)–6(б).

Согласно структуре алгоритма сегментации, разбиение образовавшихся кластеров на более мелкие части в дальнейшем невозможно. Следовательно, результаты первой стадии должны удовлетворять двум следующим требованиям (качественному и количественному). Качественное требование состоит в том, что первичный набор кластеров должен быть как можно более детализированным, чтобы обеспечить разделение всех областей изображения, которые должны быть отдельными на финальной карте сегментов. Количественное требование заключается в том, что число кластеров первичного набора не должно быть излишне большим, чтобы предотвратить перегруженность процедуры анализа на второй стадии.

6. УПРОЩЕНИЕ ГРАФА КЛАСТЕРОВ (ВТОРИЧНАЯ СЕГМЕНТАЦИЯ)

Если целью первой стадии сегментации является максимальная детализация изображения, то задача второй стадии — максимальное сокращение числа сегментов при условии минимальной потери (т.е. слияния разнородных) объектов. Таким образом, вторичная сегментация сводится к упрощению графа кластеров, построенного во время первичной сегментации. Это достигается сравнением признаков пар соседних узлов и слиянием тех из них, признаки которых близки. Для сравнения признаков кластеров используется *мера расстояния* между ними. Если расстояние между парой соседних узлов меньше некоторого задаваемого порога, то эти узлы сливаются, а их признаки усредняются.

Признаки новых кластеров, получаемых в результате слияния, формируются как усреднение признаков объединяемых кластеров или элементов. Тем самым они могут быть отображены в то же пространство признаков, а значит, допустимо использовать ту же самую метрику (10) и для измерения межкластерного расстояния. Однако такой подход никак не использует доступную пространственную информацию о кластере и его соседях, такую как число элементов в кластере, длина границы и т.д. Можно модифицировать меру (10) путем добавления новых характеристик в пространство признаков.

В нашем исследовании мы использовали минимальное расширение пространства признаков; в качестве дополнительных характеристик были добавлены: размер кластера (число элементов в кластере) $P(u)$ и интегральный текстурный уровень (6). Метрика также была соответствующим образом модифицирована. В расширенном пространстве признаков расстояние между кластерами u и v задается следующей формулой:

$$\begin{aligned} \tilde{D}(u, v) = & [\tilde{w}_c + \tilde{w}_t(1 - t(u, v)p(u, v))] s(u, v)D_c(u, v) + \\ & + [(1 - s(u, v))(\tilde{w}_c + \tilde{w}_t) + \tilde{w}_t s(u, v)t(u, v)p(u, v)] D_t(u, v). \end{aligned} \quad (11)$$

Здесь $D_c(u, v)$ и $D_t(u, v)$ — цветовая и текстурная части в пространстве признаков, задаваемые формулами (7) и (9), с той лишь разницей, что вместо признаков отдельных элементов подставляются соответствующие признаки кластеров. Весовые параметры \tilde{w}_c и \tilde{w}_t в (11) аналогичны константам w_c и w_t в (10), но значения их могут быть другими. Функции p , t и s зависят от размеров, уровня текстуры и цветовой насыщенности кластеров u и v , и задаются

следующими формулами:

$$\begin{aligned} p(u, v) &= F(\min(P(u), P(v))/P_0); \\ t(u, v) &= (T(u) + T(v))/2T_{\max}; \\ s(u, v) &= 0.5 + 0.5F(\max(S(u), S(v))/S_0). \end{aligned}$$

Здесь T_{\max} — нормализующая константа, задающая максимальное значение уровня текстуры; S_0 и P_0 — пороговые константы, имеющие размерность цветовой насыщенности и размеров кластера соответственно. Пороговая функция F задается выражением (8).

Формула (11) заменяет в расширенном пространстве признаков формулу (10). Коэффициенты перед $D_c(u, v)$ и $D_t(u, v)$ должны рассматриваться как модифицированные веса цветовой и текстурной составляющих соответственно. Сумма этих коэффициентов — величина постоянная, равная $(\tilde{w}_c + \tilde{w}_t)$. Введение функций p , t и s в формулу (11) преследует цель сбалансировать метрику в вырожденных случаях, когда один или несколько значений признаков, такие как уровень текстуры, цветовая насыщенность или размеры кластера становятся малы. Так, например, если размеры кластера u или v становятся меньше, чем P_0 , модифицированный вес цветовой части возрастает, а модифицированный вес текстурной части одновременно уменьшается. Если цветовая насыщенность обоих кластеров u и v становится меньше S_0 , или же увеличивается уровень текстуры, то модифицированный вес цветовой части уменьшается, а текстурной — увеличивается.

Процедура вторичной сегментации построена по следующей итеративной схеме. Набор узлов графа (кластеров) анализируется в порядке убывания размеров кластеров. Для каждого очередного узла, который в данном случае выступает в роли базового узла (базового кластера), просматривается набор соединенных с ним узлов (соседних кластеров). Если согласно (11) расстояние между базовым узлом и очередным соседним не превышает заданного порога Θ , то этот (соседний) узел сливается с базовым. Это означает, что все связи соседнего узла переориентируются на базовый узел, а также корректируются все связи и характеристики объединенного узла (кластера). Затем операция сравнения повторяется со следующим соседним узлом. По окончании списка соседних узлов осуществляется переход к следующему базовому узлу. Процедура заканчивается, когда обработаны все узлы графа.

Наши исследования показали, что не удастся подобрать единых универсальных значений весовых коэффициентов w в (11) и порога Θ таких, чтобы получить приемлемые результаты за один проход. Для преодоления этой проблемы процедура вторичной сегментации выполняется в несколько этапов: основное слияние кластеров (состоящее из первоначального слияния маленьких кластеров и последующего слияния больших кластеров), слияние темных кластеров, и, наконец, слияние текстурных кластеров. На каждом из шагов используются различные значения коэффициентов w и порога Θ . Динамический порог Θ зависит от шага итерации и размеров кластеров:

$$\Theta = \alpha(1/(P + P_0) + \beta). \quad (12)$$

Здесь P — размер наименьшего кластера в паре; P_0 задает размер кластера, считающегося “маленьким”; β ограничивает уменьшение порога для “больших” кластеров ($P \gg P_0$), а коэффициент α задает общую регулировку порога. Этап основного слияния кластеров выполняется по вышеописанной схеме итеративно, с постепенным увеличением динамического порога Θ в (12).

Процедура слияния темных кластеров обрабатывает только кластеры с низким значением яркости B . Вводятся также некоторые дополнительные изменения: текстурный весовой коэффициент \tilde{w}_t в (11) принимается равным нулю, и меняются весовые коэффициенты в (7), принимая значения $w_H \ll w_B$ и $w_S \ll w_B$.

Процедура слияния текстурных кластеров применяется только к множеству кластеров с высоким значением уровня текстуры $T(u)$ в (6). При этом на весовые коэффициенты в (11) налагается условие $w_t \ll w_c$.

Окончательные результаты сегментации представлены на Рис. 2(в)–5(в), 6(г), 7(б), а также на Рис. 8, где границы между обнаруженными сегментами проведены в виде белых линий поверх исходного изображения.

7. ЭКСПЕРИМЕНТЫ И РЕЗУЛЬТАТЫ

Алгоритм сегментации был реализован на персональном компьютере, и эксперименты проводились со многими различными типами реальных изображений без существенных ограничений. Тем самым параметры (значения коэффициентов и порогов) подбирались таким образом, чтобы получить наилучшие результаты “в среднем” по сравнительно большому набору тестовых изображений различного вида.

Разработанный алгоритм продемонстрировал удовлетворительные результаты на большинстве из доступных изображений, среди которых: тестовые изображения сцен ограниченной сложности, реальные сцены и портреты, изображения из различных баз данных и библиотек, изображения, составленные из текстур альбома [39], а также целый ряд других изображений. Важно заметить, что поскольку алгоритм ориентирован на анализ изображений в обобщенном текстурно-цветовом пространстве, то его возможности сокращаются, если исходное изображение отображается лишь в его подпространство, что происходит, например, при сегментации черно-белого изображения. Тем не менее, даже в таком случае алгоритм демонстрирует приемлемое качество (см. Рис. 8(b)).

Несколько примеров обработки различных изображений показано на Рис. 2–8. Границы сегментированных областей проведены белыми линиями, а внутренние части сегментов заполнены средним цветом (за исключением Рис. 8, где границы нанесены поверх исходного изображения).

Иерархический пирамидальный подход демонстрирует свою эффективность для сегментации изображений. Он допускает модификации путем использования различных цветовых, текстурных или обобщенных метрик. Следует отметить, что предложенный алгоритм может быть применен не только к обычным цветным RGB изображениям, но также и к любым другим многокомпонентным (например, многозональным) изображениям. Необходимо только модифицировать рассмотренную цветовую метрику (7). Расширение пространства признаков в некоторых случаях также может дать улучшение результатов сегментации.

СПИСОК ЛИТЕРАТУРЫ

1. Гонсалес Р. Вудс Р. *Цифровая обработка изображений*. М.: Техносфера, 2005.
2. Прэтт У. *Цифровая обработка изображений. т. 1 и 2*. М.: Мир, 1982.
3. Rosenfeld A., Kak A. C. *Digital Picture Processing. vol. 1 and 2*. New York: Academic Press, 1982.
4. Дуда Р., Харт П. *Распознавание образов и анализ сцен*. М.: Мир, 1976.
5. Робертс Л. Автоматическое восприятие трехмерных объектов. В кн.: *Интегральные роботы. т.1*. М.: Мир, 1973, стр. 162–208.
6. Sobel I. E. *Camera Models and Machine Perception*. Ph.D. thesis. Palo Alto, Calif.: Stanford University, 1970.
7. Prewitt J. M. S. Object Enhancement and Extraction. *Picture Processing and Psychopictorics*. New York: Academic Press, 1970, pp. 75–150.

8. Фурман Я. А., Кревецкий А. В., Передреев А. К., и др. *Введение в контурный анализ и его приложения к обработке изображений и сигналов*. М.: Физматлит, 2003.
9. Clark J. J. Authenticating Edges Produced by Zero-Crossing Algorithms. *IEEE Trans. Pattern Analysis Machine Intelligence*, 1989, vol. 12, no. 8, pp. 830–831.
10. Canny J. A Computational Approach for Edge Detection. *IEEE Trans. Pattern Analysis Machine Intelligence*, 1986, vol. 8, no. 6, pp. 679–698.
11. Sahoo P.K., Soltani S., Wong A.K.C., Chan Y.C. A Survey of Thresholding Techniques. *Computer Vision, Graphics, Image Processing*, 1988, vol. 4, pp. 233–260.
12. Jain R., Kasturi R., Schunk B. *Computer Vision*, McGraw-Hill, New York, 1995.
13. *Image Analysis and Mathematical Morphology. vol. 2*. Serra J. (ed.). New York, Academic Press, 1988.
14. *Special issue on Mathematical Morphology and Nonlinear Image Processing*. Pattern Recognition, 2000, vol. 33, no. 6, pp. 875–1117.
15. Fu K.S., Mui J.K. A Survey of Image Segmentation. *Pattern Recognition*. 1981, vol. 13, no. 1, pp. 3–16.
16. Haralick R.M., Shapiro L.G. Image Segmentation Techniques. *Computer Vision, Graphics, Image Processing*. 1985, vol. 29, no. 2, pp. 100–132.
17. Haralick R.M. Shapiro L.G. *Computer and Robot Vision. vol. 2*. Reading, MA, Addison-Wesley, 1993.
18. Shapiro L.G., Stockman G.C. *Computer Vision*. N.J., Prentice Hall, Upper Saddle River, 2001.
19. Jain A.K., Dubes R.C. *Algorithms for Clustering Data*. Prentice Hall, 1988.
20. Matas J. Kittler J. Spatial and Feature Space Clustering: Applications in Image Analysis. *6th Int. Conf. on Computer Analysis and Patterns*. Prague, Czech Republic, 1995.
21. Яне Б., *Цифровая обработка изображений*. М.: Техносфера, 2007.
22. Ohta Y., Kanade T., Sakai T. Color Information for Region Segmentation. *Computer Graphics, Image Processing*. 1980, vol. 13, pp. 224–241.
23. Pal N.K., Pal S.K. *A Review on Image Segmentation Techniques*. Pattern Recognition, 1993, vol. 26, no. 9, pp. 1277–1293.
24. Haralick R.M. Image Texture Survey. *Fundamentals in Computer Vision*. Cambridge, 1983, pp.145-172.
25. Haralick R.M. Statistical and Structural Approaches to Textures. *Proc. IEEE*. 1979, vol. 67, no. 5, pp. 786–804.
26. Bovik A.C., Clark M., Geisler W.S. Multichannel Texture Analysis Using Localized Spatial Filters. *IEEE Trans. PAMI*. 1990, vol. 12, no. 1, pp. 55–73.
27. Van Gool L., Dewaele P., Oosterlinck A. Texture Analysis Anno 1983. *Computer Vision, Graphics, Image Processing*. 1985, vol. 29, pp. 336–357.
28. Roan S.J., Aggarwal J.K. Multiple Resolution Imagery and Texture Analysis. *Pattern Recognition*. 1987, vol. 20, no. 1, pp. 17–31.
29. Chang T., Kuo C.J. Texture Analysis and Classification with Three-Structured Wavelet Transform. *IEEE Trans. Image Processing*. 1993, vol. 2, no. 4, pp. 429–441.
30. Pichler O., Teuner A., Hosticka B.J. A Comparison Of Texture Feature Extraction Using Adaptive Gabor Filtering Pyramidal And Tree Structured Wavelet Transforms. *Pattern Recognition*. 1996, vol. 29, pp. 733–742.
31. Jain A.K., Farrokhnia F., Unsupervised Texture Segmentation Using Gabor Filters. *Pattern Recognition*. 1991, vol. 24, no. 12, pp. 1167–1186.
32. Dunn D., Higgins W.E. Optimal Gabor Filters For Texture Segmentation. *IEEE Trans. Image Processing*. 1995, vol. 4, pp. 947–964.

33. Weldon T.P., Higgins W.E., Dunn D.F., Efficient Gabor Filter Design For Texture Segmentation. *Pattern Recognition*. 1996, vol. 29, pp. 2005–2015.
34. Carton E.J., Weszka J.S., Rosenfeld A. Some Basic Texture Analysis Techniques. *TR-288. Computer vision laboratory, Computer Science Center. Univ. of Maryland*, 1974.
35. Jain A.K. Color Distance and Geodesics in Color 3 Space. *JOSA*. 1972, vol. 62, no. 11, pp. 1287–1291.
36. MacAdam D.L. Projective Transformations of the ICI Color Specifications. *JOSA*. 1935, vol. 27, no. 9, pp. 294–299.
37. Hunter G.M., Steiglitz K. Operation of Images Using Quad Trees. *IEEE Trans., PAMI-1*. 1979, no. 2, pp. 145–153.
38. Rosenfeld A., Quadrees and Pyramids for Pattern Recognition and Image Analysis. *Proc. 5th Int. Conference on Pattern Recognition*. Miami Beach, 1980, pp. 802–811.
39. Brodatz P. *Textures: A Photographic Album for Artists and Designers*. New York, Dover Publications, 1966.

Статью представил к публикации член редколлегии Н. А. Кузнецов

8. ИЛЛЮСТРАЦИИ



Рис. 2. Сегментация изображения “Шары”: (а) исходное изображение; (б) первичная сегментация — 2875 сегментов; (в) окончательная сегментация — 35 сегментов.



Рис. 3. Сегментация изображения “Тукал”: (а) исходное изображение; (б) первичная сегментация — 2671 сегмент; (в) окончательная сегментация — 22 сегмента.

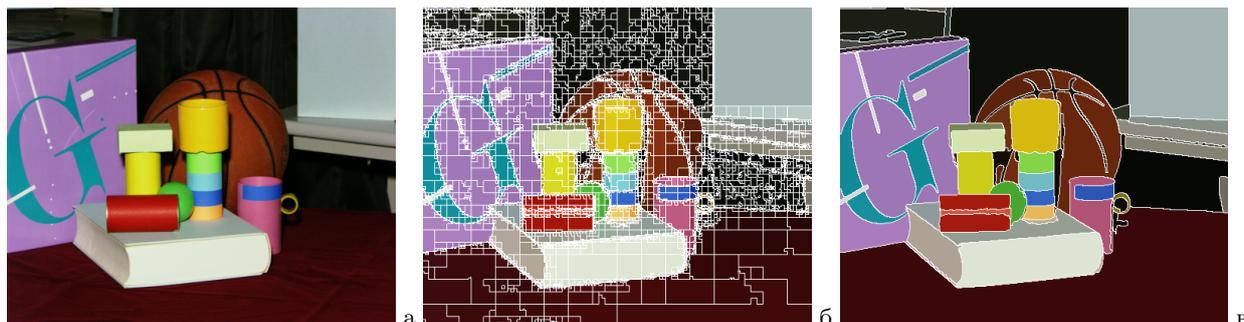


Рис. 4. Сегментация изображения “Сцена”: (а) исходное изображение; (б) первичная сегментация — 7269 сегментов; (в) окончательная сегментация — 75 сегментов.



Рис. 5. Сегментация изображения “Диктор”: (а) исходное изображение; (б) первичная сегментация — 4720 сегментов; (в) окончательная сегментация — 50 сегментов.

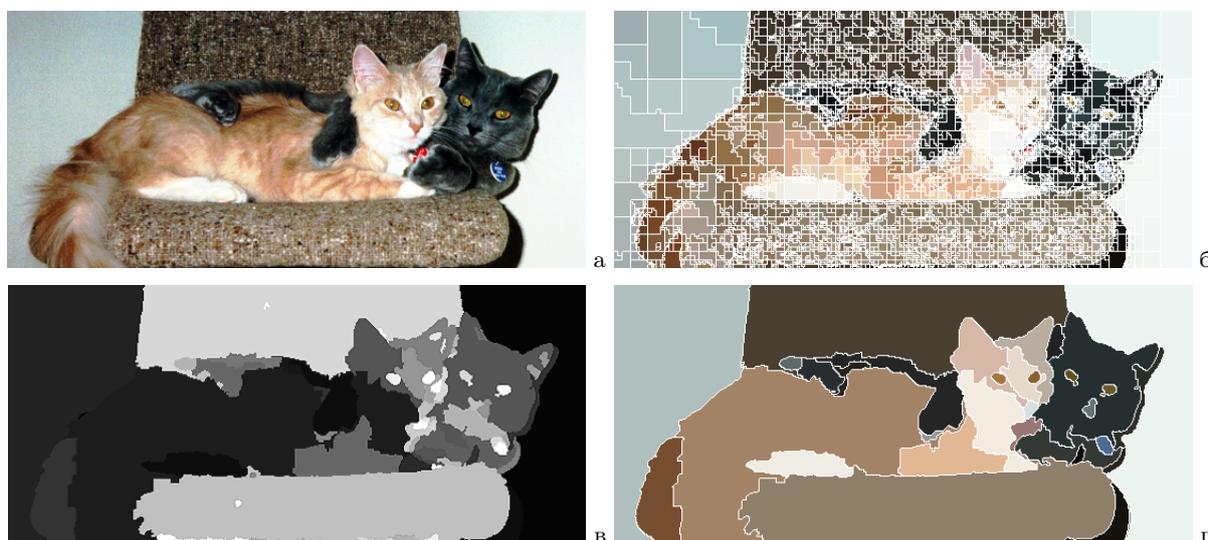


Рис. 6. Сегментация изображения “Кошки”: (а) исходное изображение; (б) первичная сегментация — 7308 сегментов; (в) средняя интенсивность текстурного отклика в градациях полутонов; (г) окончательная сегментация — 33 сегмента.



Рис. 7. Сегментация изображения “Замок”: (а) исходное изображение; (б) результат сегментации — 39 сегментов.



Рис. 8. Сегментация текстурных изображений: (а) “Зерна” — 13 сегментов; (б) “Текстуры” (скомпиллировано из альбома [39]) — 5 сегментов.