

УДК 004.932.2

СЕГМЕНТАЦИЯ ИЗОБРАЖЕНИЙ НА ОСНОВЕ АНАЛИЗА РАССТОЯНИЙ В ПРОСТРАНСТВЕ ПРИЗНАКОВ

П. А. Чочиа

*Институт проблем передачи информации РАН,
127994, Москва, ГСП-4, Большой Каретный переулок, 19
E-mail: chochia@iitp.ru*

Рассматривается задача сегментации (разбиения на однородные области) изображения, использующая яркостные цветовые и текстурные различия. Критерием однородности служит оценка близости точек в объединённом пространстве признаков. Предлагается метрика в этом пространстве. Реализация осуществляется иерархическим алгоритмом на основе анализа расстояний в пространстве признаков.

Ключевые слова: обработка изображений, распознавание, сегментация.

Введение. Задача сегментации может быть сформулирована как разбиение исходного изображения на множество непересекающихся связных областей, ассоциируемых с объектами наблюдаемой сцены или их частями в соответствии с некоторыми выбранными критериями. Результатом является получение карты областей (сегментов) изображения, т. е. отображение множества точек изображения в конечный и сравнительно небольшой набор значений $1, \dots, Q$, где Q есть финальное число сегментов. Такое разбиение может служить как предварительным шагом для последующего морфологического или семантического анализа и интерпретации наблюдаемой сцены, так и самостоятельным результатом, например, в задачах обнаружения.

Подходов к решению задачи сегментации изображений много и исследованию данного вопроса посвящено большое количество работ, тем не менее сегментация по-прежнему остаётся одной из фундаментальных проблем анализа поступающей видеoinформации. В большинстве случаев алгоритмы сегментации строятся 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 четыре класса. Первые два в основном ориентированы на однокомпонентные (монохромные или чёрно-белые) изображения, третий и четвёртый классы допускают анализ многокомпонентных данных.

1. Разбиение изображения проведением границ. Подход заключается в обнаружении границ контурными операторами, в их прослеживании, связывании и составлении из них замкнутых границ областей. Большинство алгоритмов обнаружения контурных границ основано на фильтрации, преобразующей изображение в двумерный массив, значения которого соответствуют вероятности нахождения контура в соответствующей точке. Наиболее известные используют вычисление первой производной (операторы Робертса, Собела [5, 6] и др. [1–3, 7, 8]) 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. Альтернативой алгоритмам фильтрации является прослеживание контуров [5, 8]. Хорошо известен так называемый метод "zero crossing", основанный на обнаружении точек перехода через нуль второй производной [9]: он создаёт контурные линии минимальной толщины, однако при этом появляется много ложных линий, слабо

соответствующих реальным контурам. Интересные результаты достигаются алгоритмом Кэнни [10]. Существуют также и другие методы и алгоритмы [1–3, 7]. Сегментация на основе проведения границ используется редко, в частности для некоторых специфических классов анализируемых изображений.

2. Морфологический подход. Из подобных методов наиболее известна сегментация по водоразделам [11, 12], которая состоит в интерпретации гладких областей как локальных бассейнов, а контуров между ними как водоразделов. Подход достаточно эффективен для монохромных изображений, но слабо применим к цветным и текстурным.

3. Разбиение изображения на однородные по значению области. Здесь в первую очередь необходимо упомянуть пороговые методы, использующие глобальные и адаптивные пороги [3, 13]. Такие методы просты в реализации, но подобрать удовлетворительные пороги не всегда возможно. Известны методы выращивания областей [14, 15], заключающиеся в группировке элементов и мелких областей изображения в более крупные, начиная из так называемых «центров кристаллизации». Проблемы данного метода — выбор подходящих мер близости элементов и критериев останковки процесса выращивания областей [16]. Альтернативой выращиванию служат методы, основанные на первичном разбиении изображения на множество малых областей и последующем их слиянии или разделении [17].

4. Классификация в пространстве признаков. Состоит в выборе отображения набора входных данных в некоторое пространство признаков и последующем решении задачи кластеризации — разбиении выбранного пространства на классы на базе плотности распределения в нём [4, 18]. Данные методы удобны, если количество классов объектов в изображении заранее известно. Если отображения различных областей исходного изображения в пространстве признаков перекрываются, классическая процедура сегментации может оказаться неверной. Преодолеть проблему предлагается расширением пространства признаков, например добавлением значений координат изображения [4].

Возможны два пути улучшения результатов сегментации. Первый состоит в нахождении более соответствующих задаче признаков, а второй — в отыскании иного подхода к разработке алгоритма. Нами использованы оба. В части выбора признаков исследуется комбинация пространств яркостных, цветовых и текстурных признаков с выбором модифицированной метрики в данном пространстве. Для алгоритма выбран иерархический подход, основанный на анализе расстояния в пространстве признаков.

Признаки яркость и цвет являются измеримыми и удобными для классификации [19, 20], в то время как понятие «текстура» достаточно расплывчато и нечётко формулируется [21]. Обобщённо говоря, 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 интерес представляет 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, контраст примитивов и некоторые другие [20, 22]. Одни из упомянутых характеристик могут быть чётко сформулированы, тогда как другие — нет.

Главная трудность в анализе текстур заключается в отыскании подходящих характеристик для их описания; основными являются три подхода: структурный, спектральный и статистический. В структурном предварительно выделяются примитивы и затем оценивается их размещение, что эффективно при анализе регулярных текстур. Спектральные методы используют свойства фурье-спектра и подходят для отыскания периодичности в изображении. Более универсален статистический подход. В качестве статистических характеристик чаще всего выделяют матрицу смежности, автокорреляцию, разности уровней серого, длины серий по направлениям, локальную частоту флуктуаций яркости [22–25]. Иногда применяется иерархический подход, позволяющий оценивать характеристики

на разных масштабах. После 1980-х годов возник интерес к многоканальному и многомасштабному [26] текстурному анализу, а позже — к использованию вейвлетов [27, 28] и габоровских фильтров [28–31], тесно связанных с пирамидальным представлением изображения [28].

Особенности задачи сегментации изображений. Сегментацию изображения можно переформулировать как задачу разделения двумерного сигнала на связные области (гладкие в смысле анализируемых характеристик), отличающиеся от соседствующих областей по своим яркостным, цветовым и/или текстурным характеристикам. Согласно модели входного сигнала [32] следует отметить следующие принципиальные особенности сегментации изображения.

1. Изображения являются двумерным сигналом и отличаются высокой пространственной корреляцией характеристик соседствующих элементов.

2. Области на изображениях во многих случаях имеют достаточно большие размеры, причём на протяжении областей значения характеристик могут заметно изменяться.

3. Возможное число классов объектов на изображениях, как правило, заранее неизвестно.

4. При сегментации изображения обычно не ставится задача объединения в один класс близких по характеристикам, но не соседствующих между собой областей.

Из-за указанных особенностей элементы одного и того же объекта изображения в пространстве признаков могут отображаться в достаточно протяжённую область сложной формы. В свою очередь, области, соответствующие различным объектам, могут оказаться весьма близкими или даже пересекаться. К тому же при поэлементном отображении теряется информация о пространственном соседстве точек изображения. Всё это свидетельствует о том, что сегментация изображения путём решения классической задачи распознавания имеет ограниченное применение.

Удовлетворить указанным особенностям должен подход, использующий критерии близости как выбранных признаков, так и пространственного расположения элементов изображения. В качестве такового был принят иерархический подход, основанный на пирамидальном алгоритме [33], позволяющем сравнивать признаки пространственно соседствующих пар элементов и объединять те, которые в этом смысле оказываются близки. Согласно данному алгоритму изображение представляется в виде графа, на нижнем уровне которого расположены элементы изображения, а ветви соответствуют отдельным объектам. Ключевым здесь является выбор оценки близости соседних вершин графа, что в применении к рассматриваемой задаче означает выбор пространства признаков и способа оценки близости точек в этом пространстве. Сопоставлять вершины графа, соответствующие соседним элементам и/или сегментам изображения, удобно с помощью расстояния в пространстве признаков, которое и будет критерием при принятии решения об их слиянии или же проведении между ними границы.

Выбор пространства признаков. Требуется выбрать такое пространство признаков, которое бы адекватно отображало необходимые признаки и в котором было бы возможно найти удобную и не слишком сложную меру расстояния. Для этого принято решение строить полное пространство признаков как объединение нескольких более простых пространств (подпространств): яркостного, цветового и текстурного.

Яркостное пространство является наиболее простым. Оно одномерно, и во всём диапазоне яркостей можно предполагать постоянство контрастной чувствительности [1, 3]. В такой ситуации удачной метрикой будет модуль разности яркостей $|B_x - B_y|$.

С выбором цветового пространства ситуация существенно сложнее. Предлагается много различных цветовых представлений, таких как XYZ, $U^*V^*W^*$, $S\Theta W^*$, $L^*a^*b^*$, $L^*u^*v^*$ и др. [1, 34]. Тестирование метрик в ряде указанных пространств показало, что при их ис-

пользовании достигаются близкие результаты сегментации. Было решено остановиться на пространстве **BHS**, в котором хорошо разделяются яркостная и цветовая составляющие.

В качестве текстурного пространства в данной работе используется набор многомасштабных оценок изменчивости яркости по множеству направлений, которые вычисляются для каждой точки изображения на основе значений локальной вариации и локальной осцилляции (изменения знака) градиента вдоль выбранных направлений. Такой подход существенно развивает метод разностей яркостей, обсуждаемый в обзорах [25, 35], и близок идее оценки функций двумерными вариациями [36].

Пространство **BHS** отличается от обычного цветового пространства **RGB** координатным преобразованием. По сути, координаты B , H и S являются полярными координатами пространства **RGB**. Пусть r , g , b — координаты точки в исходном пространстве **RGB** и $u = \min(r, g, b)$. Тогда

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

Подпространства **B** и **HS** составляют яркостное и цветовое пространства признаков.

Текстурные характеристики оцениваются по значениям яркости B . Проводится анализ профиля на локальном участке (число подъёмов или спусков). С одной стороны, это подобие частотного анализа, с другой — аналог первой вариации Кронрода [36], предложенной для оценки двумерных сигналов и характеризующей число «выступов». Применимость такой характеристики к реальным изображениям проанализирована в работе [37].

Значения текстурных признаков вычисляются следующим образом. Рассмотрим отрезок длиной $2L$, центрированный в анализируемой точке и проходящий под углом $\alpha_k = k\pi/K$, где $k = 0, \dots, K - 1$. Расположим на нём равномерно $2L + 1$ точек (отсчётов) 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 (другими словами, градиент в точке i вдоль направления отрезка). Выберем также спадающую к краям весовую функцию $q_i = \cos(i\pi/2L)/\sum \cos(i\pi/2L)$. Тогда взвешенные суммы положительных v^+ и отрицательных v^- перепадов вдоль отрезка будут равны

$$v^+ = \sum_{i=-L}^{L-1} q_i d_i \quad \text{по всем } d_i > 0; \quad v^- = - \sum_{i=-L}^{L-1} q_i d_i \quad \text{по всем } d_i < 0. \quad (2)$$

Анализируемый отрезок может попасть на контурную границу объектов или же участок с заметным яркостным трендом, что внесёт существенные изменения в значения одной из сумм (2). В предположении, что на коротком отрезке длиной $2L + 1$ имеется не более одного такого яркостного перепада, снизить его влияние на значения текстурных откликов можно выбором наименьшего из значений v^+ и v^- . Таким образом, значение локальной вариации v_k по направлению k определяется как

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

Значение локальной осцилляции градиента g определяется как число перемен знака градиента d_i на отрезке $(-L \leq i \leq L)$ таких, что $|d_i| > \delta$, где δ — порог чувствительности. Для уменьшения влияния шума можно рекомендовать в качестве x_i брать значение среднего вдоль небольших отрезков, перпендикулярных основному отрезку анализа текстуры и центрированных в каждой из i $(-L \leq i \leq L)$ точек отрезка.

Произведение локальных вариации v_k и осцилляции g_k градиента $\tilde{t}_k = v_k g_k$ даёт предварительное значение текстурного признака по направлению k . Для выравнивания признаков высоко- и малоконтрастных текстур используется, как это предлагается в [30], функция гиперболического тангенса:

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

Полученное значение t_k считается текстурным признаком по направлению k в определённой точке изображения. Как правило, значения t_k являются достаточно зашумлёнными; для повышения однородности можно рекомендовать применение низкочастотной фильтрации какого-нибудь вида, например вычисления медианы по небольшому локальному двумерному фрагменту размерами $h \times h$ точек.

Текстура является масштабно-зависимой характеристикой. Для различения текстур разных масштабов необходимо выделение признаков, соответствующих различным пространственным частотам. Можно воспользоваться для этого формулами (2)–(4), сохранив число точек на отрезке $(2L + 1)$, но взяв отсчёты с другим шагом и пропорционально изменив размеры отрезка. Так, увеличив шаг в n раз, получим сдвиг диапазона текстурной чувствительности в сторону более низких частот. При этом следует подавить влияние высоких частот путём низкочастотной фильтрации (например, сглаживанием). Другим способом изменить диапазон частот можно с помощью предварительного масштабирования (сжатия) самого изображения в те же n раз при сохранении размеров отрезков анализа неизменными. С точки зрения количества вычислений такой подход представляется предпочтительным.

Таким образом, текстурные признаки вычисляются по формулам (2)–(4) для каждой точки изображения по всем направлениям $0, \dots, M - 1$ и масштабам $1, 1/2, 1/4, \dots, 2^{-(N-1)}$, где M и N — общее число направлений и масштабных уровней. Тем самым мы получаем набор масштабно-зависимых значений локальной вариации v_m^n и осцилляции g_m^n локального градиента. Текстурные признаки при этом вычисляются как

$$t_m^n = \tanh(\alpha v_m^n g_m^n). \quad (5)$$

Можно также оценить общий уровень текстуры в точке следующим образом:

$$T = \sum_{n=1}^N w^n \sum_{m=0}^{M-1} t_m^n, \quad (6)$$

здесь w^n есть весовой коэффициент масштабного уровня n .

Суммарное число текстурных признаков для каждой точки изображения равно $M \times N$. В реальности достаточными представляются $M = 4$ и $N = 4$, т. е. всего 16 формируемых и анализируемых текстурных признаков. Объединённый набор текстурных, яркостных и цветовых признаков для точки изображения составляет полный вектор признаков, а всё возможное многообразие таких векторов — пространство яркостно-цвето-текстурных признаков. Комбинацию трёх указанных пространств можно назвать объединённым многокомпонентным пространством признаков.

Метрика в пространстве признаков. Важнейшей частью иерархического алгоритма сегментации является принятие решения о слиянии соседних кластеров (вершин графа), что осуществляется сравнением близости их отображений в пространстве признаков. Уже отмечалось, что данная задача может быть сформулирована как определение расстояния (выбор метрики) в пространстве признаков. Универсальной метрики в объединённом пространстве признаков, к сожалению, не существует, таковой нет даже для цветового подпространства. Предлагаемая здесь метрика разработана специально для задачи сегментации.

Поскольку между яркостными, цветовыми и текстурными признаками имеется существенное различие, решено задать отдельные метрики в каждом из подпространств, а затем объединить их в общую метрику. Заметим, что мы используем здесь термин «метрика» в редуцированном смысле — исключительно как некоторый способ вычисления расстояния между парой точек в пространстве признаков. Отметим также, что выбор способа вычисления расстояния (метрики) производился в применении к одному и тому же пирамидальному алгоритму сегментации [33].

Задача различения цветов и построения цветовой метрики известна достаточно давно и обычно решается поиском равноконтрастной цветовой метрики [1, 34, 38, 39]. Как уже отмечалось выше, тестирование метрик в различных пространствах показало, что при их использовании в задаче сегментации достигаются близкие и во многих случаях приемлемые результаты. Нами выбрано пространство **BHS**, поскольку в нём удаётся хорошо разделить яркостную и цветовую составляющие, а также построить модифицированную метрику, которая лучше соответствует задаче сегментации и позволяет достичь более точных результатов в тёмных областях изображения. Метрики в каждом из подпространств задаются следующим образом.

Яркостная метрика $D_B(x, y)$ определяется просто как модуль разности яркостей B_x и B_y в двух точках x и y одномерного подпространства **B**:

$$D_B(x, y) = |B_x - B_y|. \quad (7)$$

Цветовая метрика вычисляется по следующей формуле:

$$D_C(x, y) = [w_H \varphi_B(x, y) \varphi_S(x, y) (\min(|H_x - H_y|, 2\pi - |H_x - H_y|))^2 + w_S \varphi_B(x, y) (S_x - S_y)^2]^{1/2}. \quad (8)$$

Здесь H_i и S_i ($i = x, y$) суть значения компонент цветности и насыщенности; w_H и w_S — весовые коэффициенты; $\varphi_B(x, y) = F((B_x + B_y)/B_0)$, $\varphi_S(x, y) = F((S_x + S_y)/S_0)$, где B_0 и S_0 — значения уровней нечувствительности, а

$$F(z) = (2/\pi) \arctg(z) \quad (9)$$

является функцией подавления: $0 \leq F(z) < 1$. В данном случае $F(z)$ служит для подавления избыточного влияния цветности в области малых значений яркости и насыщенности, а также подавления насыщенности в области малых яркостей.

Текстурная метрика рассчитывается по формуле

$$D_T(x, y) = \sum_{n=1}^N w^n \sum_{m=1}^M |t_m^n(x) - t_m^n(y)|, \quad (10)$$

где $t_m^n(x)$ и $t_m^n(y)$ — значения текстурных откликов уровня n и направления m в точках x и y ; w^n — весовой коэффициент уровня n .

Объединённая яркостно-цвето-текстурная метрика определяется как взвешенная сумма расстояний в яркостном (7), цветовом (8) и текстурном (10) подпространствах:

$$D(x, y) = w_B D_B(x, y) + w_C D_C(x, y) + w_T D_T(x, y). \quad (11)$$

Здесь x и y — две точки объединённого подпространства, являющиеся отображениями точек или кластеров изображения; w_B , w_C и w_T — весовые коэффициенты отдельных метрик. Если входное изображение не полностью задействует пространство признаков, можно использовать лишь часть подпространств: например, яркостное и текстурное — в случае исходного монохромного изображения или яркостное и цветное, если изображение заведомо не содержит текстуры.

Первичная сегментация — пирамидальный алгоритм. Стадия первичной сегментации служит для преобразования исходного изображения в набор кластеров (подсегментов) небольшого размера. На данном этапе каждый объект изображения может ещё отображаться не одним кластером, а набором из нескольких кластеров. Каждый сформированный кластер при этом должен характеризоваться как собственными параметрами, так и параметрами связи с соседними кластерами. Границы между объектами должны проходить по границам между кластерами.

Такую задачу удобно решать с помощью пирамидально-рекурсивного алгоритма [33, 40, 41]. Пирамидальный алгоритм обработки изображения требует задания способа вычисления уменьшенного изображения уровня $n + 1$ на основе имеющегося изображения уровня n и применения данной процедуры рекурсивно до предельного уменьшения размера изображения. Часто для этого выбирают элементарный домен некоторой формы, обеспечивающей плотное покрытие всей площади изображения, и задают способ нахождения значения элемента следующего уровня по значениям элементов домена предыдущего уровня. Рекурсивное применение такой 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 размерами выбранного домена. В случае если в качестве домена выбирается квадрат размерами 2×2 элемента, мы получаем так называемое квадродерево. Оно и будет использоваться в дальнейшем.

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

При прямом проходе 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 четырёх нижних узлов уровня n , составляющих один домен, создаётся узел уровня $n + 1$. В новом узле запоминается информация об узлах, входящих в домен предыдущего уровня, а также о прохождении контуров между ними. Таким образом, каждый узел является вершиной некоторого квадродерева, охватывающего расположенные под ним элементы изображения, и содержит информацию о поддеревьях предыдущего уровня.

Основная задача пирамидального этапа сегментации — объединение соседних элементов, имеющих близкие признаки и не разделённых контуром. Эта процедура требует прослеживания 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 выбранной метрики (11) 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 состоит из следующих шагов.

1. Анализируется наличие контурных перепадов между парами соседних элементов внутри домена. Пары элементов (узлов графа), не разделённые контуром, считаются принадлежащими одному и тому же кластеру.

2. Производится выбор наиболее представительного набора элементов, принадлежащих домену нижнего уровня (n). Они связываются с узлом следующего уровня ($n + 1$) и их усреднённые значения записываются в этот узел. Оставшиеся элементы (узлы) считаются «подвешенными»; информация о них на следующий уровень не передаётся и сохраняется только в описании текущего узла на уровне n .

3. Строится карта контуров домена с помощью анализа множества элементов, состоящего из внутренних элементов домена и четвёрки соседних с ним элементов, следующих в последовательной развёртке, т. е. ближайших справа и снизу. Карта контуров каждого домена записывается в узел следующего уровня, расположенный над данным доменом.

4. Производится усреднение яркостных, цветовых и текстурных параметров элементов, составляющих домен, которые также записываются в узел следующего уровня.

По окончании прямого прохода по всему квадродереву каждый узел любого уровня 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 (тем, которые оказались «подвешены» на шаге 2 прямого прохода) присваиваются новые номера кластеров, которые распространяются вниз на все элементы, связанные с данным поддеревом. Эта процедура рекурсивно распространяется на все дерево, начиная с верхнего уровня, т. е. вершины дерева, до нижнего — самого изображения. Элементы изображения, получившие при этом одинаковые номера, образуют первичные кластеры (первичные сегменты) обрабатываемого изображения.

Таким образом, процедура синтеза на каждом из узлов уровня n дерева сводится к следующему. Просматриваются все элементы очередного домена, т. е. узлы уровня $n - 1$. Если рассматриваемый узел относится к числу связанных с узлом уровня n , ему присваивается номер узла n . Если же узел оказывается подвешенным, он получает очередной свободный номер.

Результатом данной процедуры является первичная карта кластеров сегментируемого изображения. Согласно структуре алгоритма сегментации разбиение образовавшихся кластеров на более мелкие части в дальнейшем невозможно. Следовательно, результаты первой стадии должны удовлетворять качественному и количественному требованиям. Качественное требование состоит в том, что первичный набор кластеров должен быть достаточно детализированным, чтобы обеспечить разделение всех областей изображения, которые должны оказаться отдельными на финальной карте сегментов. Количественное требование заключается в том, что число первичных кластеров не должно быть излишне большим, чтобы предотвратить перегруженность процедуры анализа на второй стадии. Отметим также, 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.

Вторичная сегментация — упрощение графа кластеров. Целью второй стадии является максимальное сокращение числа сегментов при условии минимальной потери (т. е. ошибочного слияния) объектов. Для дальнейшего анализа строится связный граф соседства кластеров, сформировавшихся на нижнем уровне по окончании первичной сегментации. Каждому кластеру соответствует узел графа. Также вычисляется информация, описывающая свойства каждого кластера. В неё входят число элементов, объединённых в кластере, и средние значения признаков: яркости, цвета и текстуры. Рёбра графа соединяют пары соседствующих кластеров, имеющих общую пространственную границу.

Вторичная сегментация заключается в упрощении графа кластеров, построенного во время первичной сегментации. Это осуществляется сравнением признаков соседних узлов

и слиянием тех узлов, признаки которых близки. Аналогично рассмотренной выше метрике в пространстве признаков для сравнения близости кластеров вводится модифицированная мера расстояния между ними. Если расстояние между парой соседних узлов меньше некоторого задаваемого порога, то эти узлы сливаются, а их признаки усредняются.

Предложенная метрика (11) не охватывает всех возможностей сравнения близости кластеров; на данном этапе доступна дополнительная информация. Способ измерения расстояния (11) был модифицирован добавлением характеристик: размеров сравниваемых кластеров (число элементов в кластере) $P(u)$ и $P(v)$, а также интегральных текстурных уровней (6) $T(u)$ и $T(v)$. Модифицированное расстояние между кластерами u и v в расширенном пространстве признаков в результате задаётся формулой

$$\begin{aligned} \tilde{D}(u, v) = & [w_C + w_T(1 - t(u, v)p(u, v))]s(u, v)(w_B D_B(u, v) + w_C D_C(u, v)) + \\ & + [(1 - s(u, v))(w_C + w_T) + w_T s(u, v)t(u, v)p(u, v)]D_T(u, v), \end{aligned} \quad (12)$$

где $D_B(u, v)$, $D_C(u, v)$ и $D_T(u, v)$ — яркостная, цветовая и текстурная метрики в соответствующих подпространствах признаков, задаваемые формулами (7)–(10), с той лишь разницей, что вместо признаков отдельных элементов или узлов подставляются соответствующие усреднённые значения признаков кластеров. Функции p , t и s зависят от размеров, текстуры, цветовой насыщенности кластеров u и v и задаются выражениями

$$p(u, v) = F(\min(P(u), P(v))/P_0), \quad t(u, v) = (T(u) + T(v))/2T_{\max},$$

$$s(u, v) = 0,5 + 0,5F(\max(S(u), S(v))/S_0).$$

Здесь T_{\max} — константа, определяющая максимальное значение уровня текстуры, функция $F(z)$ задаётся выражением (9), а S_0 и P_0 — уровни нечувствительности, имеющие размерности цветовой насыщенности и размеров кластера.

Формула (12) заменяет в расширенном пространстве признаков формулу (11). Множители перед $(D_B(u, v) + D_C(u, v))$ и $D_T(u, v)$ суть модифицированные веса яркостно-цветовой и текстурной составляющих соответственно. Сумма этих коэффициентов постоянна и равна $(w_C + w_T)$. Введение функций p , t и s в формулу (12) преследует цель сбалансировать метрику вблизи вырожденных ситуаций, когда один или несколько признаков, таких как уровень текстуры, цветовая насыщенность или размеры кластера, становятся малы. Так, например, если размеры кластера u или v становятся меньше значения P_0 , вес цветовой части возрастает, а вес текстурной части одновременно уменьшается. Если цветовая насыщенность обоих кластеров u и v уменьшается или же увеличивается уровень текстуры, то соответственно снижается модифицированный вес цветовой части и повышается вес текстурной.

Процедура вторичной сегментации построена по следующей итеративной схеме. Набор узлов графа (кластеров) анализируется в порядке убывания размеров кластеров. Для каждого очередного узла u , который в данном случае выступает в роли базового узла (базового кластера), просматривается набор соединённых с ним узлов v_i (соседних кластеров). Если согласно (12) расстояние между базовым узлом и очередным соседним узлом v_i не превышает заданного порога θ (т. е. $\tilde{D}(u, v_i) < \theta$), то узел v_i сливается с базовым узлом u . Это означает, что все связи узла v_i переориентируются на узел u , а также корректируются все связи и характеристики объединённого узла (кластера). Затем операция сравнения повторяется со следующим соседним узлом v_{i+1} . По окончании списка соседних узлов осуществляется переход к следующему базовому узлу. После циклической обработки всех узлов графа процедура заканчивается.

Подобрать такое значение порога θ , чтобы получить приемлемые результаты за один проход, не удаётся. Поэтому процедура вторичной сегментации выполняется в несколько итераций, каждая из которых в свою очередь включает этапы слияния маленьких кластеров, больших, тёмных и текстурных. Значение порога θ задаётся уровнем итерации и размерами кластеров:

$$\theta = \alpha(1/(P_{\min} + P_0) + \beta). \quad (13)$$

Здесь P_{\min} — размер наименьшего кластера в паре; P_0 — порог чувствительности размеров кластера; β ограничивает снижение порога для «больших» кластеров ($P_{\min} \gg P_0$); коэффициент α задаёт общую регулировку порога. Слияние кластеров выполняется итеративно с постепенным увеличением P_0 и соответственно уменьшением θ .

Процедура слияния тёмных кластеров обрабатывает только кластеры с низким значением яркости B . Вводятся также некоторые дополнительные изменения: текстурный весовой коэффициент w_T в (12) принимается равным нулю и вчетверо уменьшаются весовые коэффициенты w_H и w_S в (8).

Процедура слияния текстурных кластеров применяется только к множеству кластеров с высоким значением уровня текстуры $T(u)$ согласно (6). При этом значение коэффициента w_T в (12) увеличивается вчетверо, а значения w_B и w_C уменьшаются вдвое.

Эксперименты по выбору параметров преобразований. Сегментация изображений общего вида является трудноформализуемой задачей с точки зрения выбора критериев оптимизации, необходимых для автоматического поиска параметров, предполагающего количественную оценку результатов преобразования. Для этого при сегментации изображений потребовалась бы априорная информация о числе и форме сегментов, которая может быть точной лишь при использовании синтезированных или же искусственно скомпилированных изображений-мозаик известных форм. Оценка качества сегментации произвольных изображений неизбежно требует морфологического анализа и интерпретации наблюдаемой сцены — задача ещё весьма далёкая от автоматического решения в общем виде. Поэтому вопрос о качестве сегментации таких изображений решается путём экспертного визуального оценивания.

Проверка алгоритма и эксперименты по выбору параметров рассмотренных метрик проводились как на искусственных, так и на реальных изображениях без существенных ограничений. Для этого использовалось более 250 изображений различных классов: *а*) синтезированные с помощью компьютерного моделирования изображения трёхмерных сцен с геометрическими фигурами, освещёнными несколькими цветными источниками; *б*) текстурные мозаики с границами заданной формы, составленные из классических текстур альбома Бродатца [42]; *в*) фотоснимки сцен ограниченной сложности с цветовыми и текстурными объектами известной простой формы; *г*) реальные изображения произвольного характера и сложности, включающие пейзажи, портреты, текстуры, изображения животных и т. д. из различных баз данных и библиотек. Изображения классов *а*, *б*, *в* и частично *г* — около 120 изображений — составляли тестовую группу настройки параметров.

По классам тестовых изображений ограниченной сложности *а*, *б* и *в* качество сегментации оценивалось сравнением количества обнаруженных сегментов и точности проведения их границ с известным разбиением сцены, а по классу *г* — экспертным сравнением. Критерием являлась минимизация расхождения разбиений по всему ансамблю изображений группы настройки параметров. Параметры подбирались так, чтобы получить наилучшие результаты в среднем; выбранные значения фиксировались и применялись далее для произвольных изображений. Приемлемость выбранных параметров проверялась на оставшихся изображениях из класса *г*.

Выбор коэффициентов первоначально осуществлялся для яркостно-цветового и текстурного подпространств признаков по отдельности на основе изображений из классов *а* и *б*,

после чего добавлялись изображения классов v и z , при сегментации которых уже необходимо задействовать полное пространство признаков. В результате экспериментов коэффициенты цветовой метрики (8) составили значения $w_H = 6,7$ и $w_S = 1,0$; текстурные коэффициенты в (6) и (10) — значения $w^n = 1 - 0,12n$, где n — номер масштабного уровня. Коэффициенты объединённой метрики в (11) были выбраны $w_B = 3,3$, $w_C = 1,0$, $w_T = 0,65$. Значения α и P_0 в (13) зависят от номера итерации: $\alpha = 1 - 0,7i/I$ и $P_0 = S^{(0,2+0,8i/I)}$, где i — номер итерации, I — общее число итераций, а S — общее число элементов в изображении.

Примеры сегментации тестовых цветного и текстурного изображений показаны на рис. 1 и 2, а сегментации реальных цветных изображений — на рис. 3–5. На рисунках a представлены исходные варианты, а на рисунках b — результаты сегментации. Внут-



Рис. 1

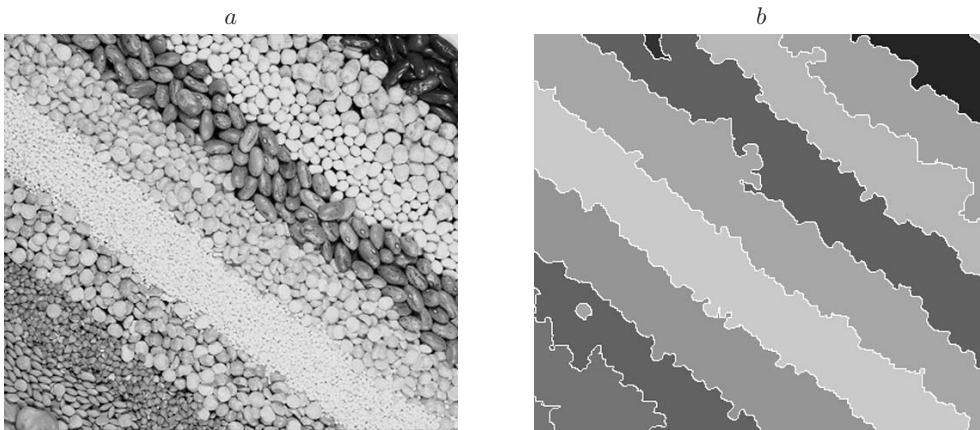


Рис. 2



Рис. 3



Рис. 4

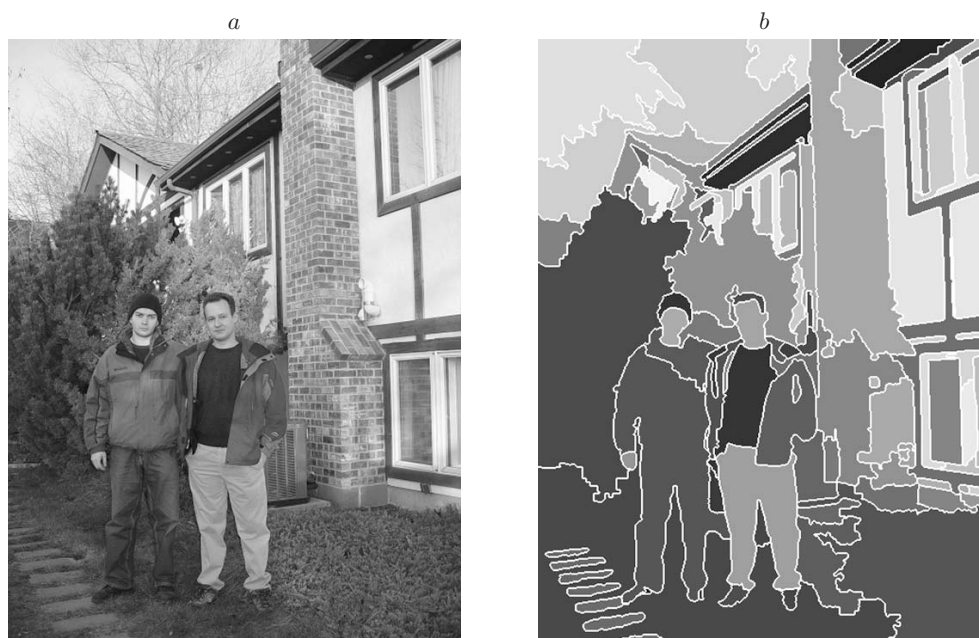


Рис. 5

рение части каждой из сегментированных областей заполнены средним значением по сегменту, а границы между ними проведены белыми линиями.

Заключение. В работе показано, что расстояние в пространстве яркостно-цвето-текстурных признаков может являться базой для построения алгоритма сегментации изображений. Критерием однородности при этом служит оценка близости точек в объединённом пространстве признаков. Предложена метрика в этом пространстве, использующая, в частности, новый способ оценки текстурных признаков на изображении, основанный на многомасштабной ориентированной локальной вариации и локальной осцилляции градиента.

С учётом анализа расстояний в пространстве признаков разработан иерархический алгоритм сегментации изображений, показавший эффективность и приемлемые результаты на подавляющем большинстве доступных изображений, среди которых: тестовые изображения сцен ограниченной сложности, изображения из различных баз данных и библиотек, а также целый ряд изображений реальных сцен, портретов, пейзажей и т. д. без существенных ограничений.

Алгоритм сегментации исследовался для объединённого яркостно-цветового и текстурного пространств признаков, однако возможна его модификация с привлечением иных признаков и метрик. Он может быть применён не только к обычным изображениям, но и к другим одно- или многокомпонентным массивам данных, удовлетворяющих свойствам изображений [1, 32], для чего требуется только выбрать набор признаков и метрику, а также подобрать соответствующие параметры.

СПИСОК ЛИТЕРАТУРЫ

1. **Прэтт У.** Цифровая обработка изображений: Пер. с англ. М.: Мир, 1982. Т. 1. 312 с.; Т. 2. 480 с.
2. **Rosenfeld A., Kak A. C.** Digital Picture Processing. N. Y.: Academic Press, 1982. Vol. 1. 435 p.; Vol. 2. 349 p.
3. **Гонсалес Р., Вудс Р.** Цифровая обработка изображений: Пер. с англ. М.: Техносфера, 2012. 1104 с.
4. **Дуда Р., Харт П.** Распознавание образов и анализ сцен: Пер. с англ. М.: Мир, 1976. 512 с.
5. **Робертс Л.** Автоматическое восприятие трехмерных объектов // Интегральные роботы: Пер. с англ. М.: Мир, 1973. Т. 1. С. 162–208.
6. **Sobel I. E.** Camera models and machine perception: Ph. D. Thesis. Palo Alto, USA: Stanford University, 1970. 89 p.
7. **Prewitt J. M. S.** Object enhancement and extraction // Picture Processing and Psychopictorics. N. Y.: Academic Press, 1970. P. 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. 592 с.
9. **Clark J. J.** Authenticating edges produced by zero-crossing algorithms // IEEE Trans. Pattern Anal. Mach. Intel. 1989. **12**, N 8. P. 830–831.
10. **Canny J.** A computational approach for edge detection // IEEE Trans. Pattern Anal. Mach. Intel. 1986. **8**, N 6. P. 679–698.
11. **Image Analysis and Mathematical Morphology.** Vol. 2: Theoretical Advances /Ed. J. Serra. N. Y.: Academic Press, 1988. 411 p.
12. **Pattern Recogn.** Special Issue on Mathematical Morphology and Nonlinear Image Processing. 2000. **33**, N 6. P. 875–1117.
13. **Jain R., Kasturi R., Schunk B.** Machine Vision. N. Y.: McGraw-Hill, 1995. 549 p.
14. **Fu K. S., Mui J. K.** A survey of image segmentation // Pattern Recogn. 1981. **13**, N 1. P. 3–16.
15. **Haralick R. M., Shapiro L. G.** Image segmentation techniques // Comput. Vis., Graph. and Image Process. 1985. **29**, N 2. P. 100–132.
16. **Haralick R. M., Shapiro L. G.** Computer and Robot Vision. Reading: Addison-Wesley, 1993. Vol. 2. 630 p.
17. **Shapiro L. G., Stockman G. C.** Computer Vision. Upper Saddle River: Prentice Hall, 2001. 580 p.
18. **Jain A. K., Dubes R. C.** Algorithms for Clustering Data. Upper Saddle River: Prentice Hall, 1988. 320 p.
19. **Ohta Y., Kanade T., Sakai T.** Color information for region segmentation // Comput. Graph. and Image Process. 1980. **13**, N 3. P. 224–241.
20. **Pal N. K., Pal S. K.** A review on image segmentation techniques // Pattern Recogn. 1993. **26**, N 9. P. 1277–1293.
21. **Яне Б.** Цифровая обработка изображений: Пер. с англ. М.: Техносфера, 2007. 584 с.

22. **Haralick R. M.** Image texture survey // Fundamentals in Computer Vision. Cambridge: CUP, 1983. P. 145–172.
23. **Haralick R. M.** Statistical and structural approaches to textures // Proc. IEEE. 1979. **67**, N 5. P. 786–804.
24. **Bovik A. C., Clark M., Geisler W. S.** Multichannel texture analysis using localized spatial filters // IEEE Trans. PAMI. 1990. **12**, N 1. P. 55–73.
25. **Van Gool L., Dewaele P., Oosterlinck A.** Texture analysis Anno 1983 // Comput. Vis., Graph. and Image Process. 1985. **29**, N 3. P. 336–357.
26. **Roan S. J., Aggarwal J. K.** Multiple resolution imagery and texture analysis // Pattern Recogn. 1987. **20**, N 1. P. 17–31.
27. **Chang T., Kuo C. J.** Texture analysis and classification with three-structured wavelet transform // IEEE Trans. Image Process. 1993. **2**, N 4. P. 429–441.
28. **Pichler O., Teuner A., Hosticka B. J.** A comparison of texture feature extraction using adaptive Gabor filtering pyramidal and tree structured wavelet transforms // Pattern Recogn. 1996. **29**, N 5. P. 733–742.
29. **Jain A. K., Farrokhnia F.** Unsupervised texture segmentation using Gabor filters // Pattern Recog. 1991. **24**, N 12. P. 1167–1186.
30. **Dunn D., Higgins W. E.** Optimal Gabor filters for texture segmentation // IEEE Trans. Image Process. 1995. **4**, N 7. P. 947–964.
31. **Weldon T. P., Higgins W. E., Dunn D. F.** Efficient Gabor filter design for texture segmentation // Pattern Recogn. 1996. **29**, N 12. P. 2005–2015.
32. **Чочиа П. А.** Двухмасштабная модель изображения // Кодирование и обработка изображений. М.: Наука, 1988. С. 69–87.
33. **Chochia P. A.** A pyramidal image segmentation algorithm // Journ. Commun. Technol. Electron. 2010. **55**, N 12. P. 1550–1560.
34. **Plataniotis K. N., Venetsanopoulos A. N.** Color Image Processing and Applications. Berlin — Heidelberg: Springer, 2000. 356 p.
35. **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.
36. **Кронрод А. С.** О функциях двух переменных // Успехи математических наук. 1955. **5**, вып. 1(35). С. 24–134.
37. **Милюкова О. П., Чочиа П. А.** Об оценке сложности изображений с помощью двумерных вариаций // Информационные процессы. 2012. **12**, № 4. С. 362–371.
38. **Jain A. K.** Color distance and geodesics in color 3 space // JOSA. 1972. **62**, N 11. P. 1287–1291.
39. **MacAdam D. L.** Projective transformations of the ICI color specifications // JOSA. 1935. **27**, N 9. P. 294–299.
40. **Hunter G. M., Steiglitz K.** Operation of images using quad trees // IEEE Trans. PAMI-1. 1979. N 2. P. 145–153.
41. **Rosenfeld A.** Quadrees and pyramids for pattern recognition and image analysis // Proc. of the 5th Intern. Conf. Pattern Recognition. Miami Beach, USA, 1980. P. 802–811.
42. **Brodatz P.** Textures: A Photographic Album for Artists and Designers. N. Y.: Dover Publications, 1966. 126 p.

Поступила в редакцию 19 декабря 2013 г.