

Алгоритм автоматической сегментации растровых изображений, основанный на росте кластеров от максимумов R-величины.

Михаил Минченков
ЦОСиВТ, МФТИ
micker@cos.ru

1. АННОТАЦИЯ

Представлен новый алгоритм автоматической сегментации изображений, построенный на основе ранее предложенного релеевского детектора границ границ двумерных объектов [1]. В отличие от другого алгоритма – R-seg – основанного на том же детекторе, представленный обеспечивает более высокую точность в определении хода границ.

Это удалось достичь за счёт того, что начальные области для дальнейшего роста сегментов выбираются в максимумах R-величины, где согласно детектору, наиболее вероятно обнаружение границ. Гистограмма окрестности с центром в таких точках, как правило, имеет двухпиковую структуру. Разделение точек по гистограмме в таких окрестностях даёт очень точный ход границы. Однако, на нерезких границах её ход изменяется существенным образом, при изменении координат центра окрестности на небольшую величину. Что бы избежать этого предлагается делать сегментацию нескольких областей, центры которых лежат на некотором, достаточно маленьком расстоянии от выбранного максимума R-величины. Результаты каждой такой сегментации предлагается заносить в буфер, а окончательный решение – к какому сегменту отнести точку – предлагается делать исходя из значений буфера. Точка относится к тому, сегменту, к которому она была отнесена в большинстве случаев.

Дальнейший рост сегментов предлагается проводить по простейшей схеме: к сегментам добавляются те соседствующие точки, что удовлетворяют пороговым значениям, устанавливаемым для каждого из сегментов исходя средней яркости и её дисперсии по сегменту.

Ключевые слова: Сегментация, детектор границ.

2. ВВЕДЕНИЕ

Сегментация – процесс перехода рассмотрения изображения как совокупности пикселей, каждый из которых характеризуется своей яркостью и положением, к совокупности областей точек, характеризующихся большим числом характеристик: геометрических, текстурных, распределение яркости по сегменту и целым рядом иных.

Такой переход необходим для решения различных задач: классификации, (в т.ч. ландшафтно-тематической дешифровке аэрокосмических снимков) – увеличивается число признаков, по которым она производится; восстановлении 3-х мерных сцен и совмещении снимков сделанных с различных точек зрения – по выделенным опорным точкам в местах излома границ; в задачах слежения – уменьшая их вычислительную сложность, а так же целым рядом других.

Среди многочисленных требований к алгоритмам сегментации стоит отметить – точность выделяемых границ, 4-связанность выделяемых сегментов и ограниченность снизу их минимального размера.

Немаловажным требованием к алгоритмам сегментации в последнее время называется автономности работы, то есть без участия оператора. Это позволяет использовать их в потоковых обработках данных, таких как кодировании видеопотока.

Предлагаемый алгоритм, как отмечалось, построен на основе ранее предложенного [1] релеевского детектора границ площадных объектов. После построения R-изображения выбираются точки, в которых значения R-величины максимальны. Выбор только тех из них, что отстоят друг от друга на расстоянии не меньше линейного размера кластера, позволяет минимизировать взаимное наложение кластеров на следующем этапе.

На втором этапе создаются кластеры в окрестностях выбранных точек. На третьем этапе производится их рост до полной сегментации изображения. На последнем этапе происходит пост-обработка полученной сегментации: выявляются сегменты, которые могут быть объединены с соседями в силу их малой величины и близости пространственно-дисперсионных характеристик.

3. РЕЛЕЕВСКИЙ ДЕТЕКТОР ГРАНИЦ ПЛОЩАДНЫХ ОБЪЕКТОВ

Релеевский детектор границ площадных объектов на растровых изображениях был представлен и подробно описан в [1]. Его краткое описание в данной работе приводится для того, что бы облегчить чтение и понимание предлагаемого алгоритма сегментации.

Упоминания о гистограммных алгоритмах сегментации встречаются уже в [3] и имеют применение в наши дни в приложениях реального времени, где требования к ресурсам имеют больший приоритет, чем остальные. Недостатки этого подхода очевидны. Это и игнорирование пространственных связей между точками, и плохие результаты на снимках с большим числом различных объектов с неоднородной окраской. В последнем случае, в силу центральной предельной теоремы теории вероятности, гистограмма снимка будет стремиться к нормальному распределению с ростом числа объектов на снимке. В таком случае, выделение пиков не даст осмысленного результата.

Пример, когда такой подход даёт хороший результат – наличие на изображении 2-3 объектов с достаточно однородной окраской. Это реализуется для ограниченной части изображения, размерами порядка L – характерного линейного размера минимальной выделяемой области.

Рассмотрим область изображения, ограниченную круглой рамкой диаметра L . В силу выбора размера рамки, внутри этой области лежит 1 или 2 объекта. Ситуации 3 и более объектов маловероятны. Однозначного ответа на вопрос о количестве объектов внутри рамке дать практически не возможно.

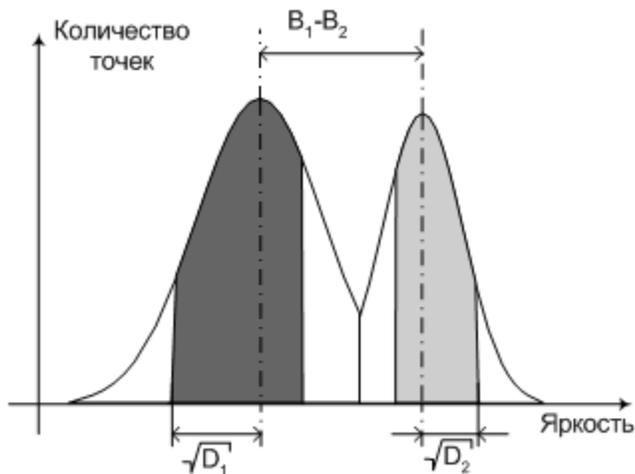


Рисунок 1 Типы деления рамок, используемые для построения R-изображения.

Введем численный критерий, характеризующий вероятность наличия границы двух объектов по различию статистических характеристик их распределений яркости в виде:

$$R_{\theta} = \frac{1}{1+Q_{\theta}}; \quad \text{где} \quad Q_{\theta} = \frac{\sigma_1 + \sigma_2 + \delta}{|B_1 - B_2|} \quad (1)$$

где B_i, σ_i – средняя яркость и корень из дисперсии яркости по объекту i (см рис 1); δ – величина, характеризующая зашумленность реального изображения. Модификация (2) позволяет повысить чувствительность функционала (1) к границам для объектов, с близкими величинами средних яркостей, но существенно различными дисперсиями:

$$R_{\theta} = \frac{1}{1+Q_{\theta}}, \quad \text{где} \quad Q_{\theta} = \frac{2 \cdot \max\{\min(\sigma_1, \sigma_2), \delta\}}{|B_1 - B_2|} \quad (2)$$

Критерий может быть обобщён на многоканальный случай (спектрозональные снимки, цветные фотографии, слои текстурных или иных признаков) если яркость B_i интерпретировать как вектор из соответствующих величин \vec{B}_i , а вместо σ_i использовать $|\vec{\sigma}_i \times \vec{e}|$, где $\vec{e} = (\vec{B}_1 - \vec{B}_2) / |\vec{B}_1 - \vec{B}_2|$ и $\vec{\sigma}_i$ так же вектора соответствующих величин.

Вероятность найти границу 3-х или более объектов в некоторой ограниченной окрестности точки значительно меньше, чем найти границу 2-х объектов или, тем более, не найти её вообще. Деление такой области на 2 примерно равные части и подсчёт величины (2) может дать оценку того, насколько велика вероятность нахождения границы в окрестности этой точки в направлении деления. Делая такую оценку по всем возможным направлениям делений и найдя такое, при котором достигается максимум, можно

утверждать, что если граница и есть, то наиболее точно она описывается этим делением.

Такая оценка делается для каждой точки входного изображения, с фиксированными размерами и формой окрестности, по ограниченному числу делений. Соответствующий пиксель (x,y) выходного изображения окрашивается яркостью, пропорциональной величине оценки в точке (x,y) . Коэффициент пропорциональности выбирается, исходя из максимальной яркости и величины дискретизации выходного изображения, которое получило название R-изображение.

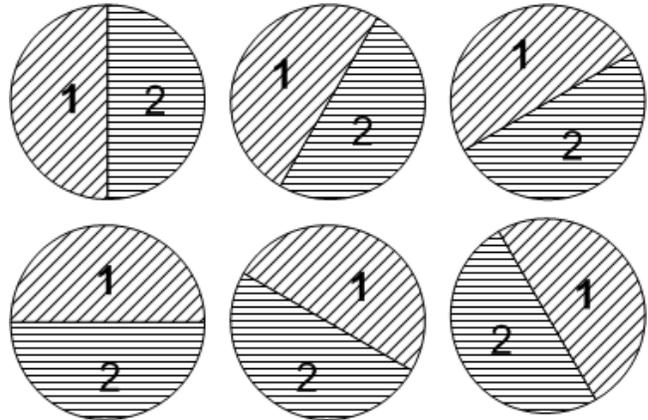


Рисунок 2 Типы деления рамок, используемые для построения R-изображения

Область выбирается круглой формы (см рис 2). Деления, в силу очевидных причин, вдоль диаметра для ряда углов. Диаметр области – величина, влияющая на точность нахождения границ: при меньших размерах выделяется больше границ, но они более точны, при больших размерах выделяется меньшее число границ, с меньшим качеством. Это связано с тем, что текстура проявляется различным образом на различных разрешениях.

Компромисс – разумное число границ с приемлемым качеством – предлагается достигать путем перемножения R-изображений полученных при различных значениях размера области L .

$$R_{multi} = \prod_{k=0}^m R(L_k) \quad (3)$$

Для того, что бы избежать потери границы там, где на одном из R-изображений она все же отсутствует, перед перемножением каждое из них подвергается линейному преобразованию, переводящему интервал яркостей $[0, \text{MAX_COLOR}]$ в интервал $[\epsilon, \text{MAX_COLOR}]$, а после перемножения результат подвергается гамма-коррекции с коэффициентом K_{γ} . (MAX_COLOR – максимальное значение яркости на принятом типе изображения). R-изображения перемножаются начиная с построенного при большей величине L , что соответствует постепенному уточнению границ. Такое изображение получило название мультиплексированного R-изображения.

4. АЛГОРИТМ ПОСТРОЕНИЯ КЛАСТЕРОВ

4.1 Выбор точек образования кластеров

Выбор точек, в которых создание кластеров, происходит в 2 этапа. Сначала ищутся все локальные максимумы R -изображения. Для этого изображение сканируется квадратной рамкой, по размеру примерно равной будущему кластеру (как будет видно из дальнейшего рассмотрения, характерная ширина и высота получаемых кластеров примерно $2L$). Если в рамке нет точки, с яркостью больше, чем в центре, то она выбирается, как возможный центр кластера.

На следующем этапе из этих точек выбираются только те, создание кластеров в которых не приведёт к их наложению на другие кластеры. То есть точки, отсеиваются так, чтобы каждые 2 из них не лежали на расстоянии ближе, чем характерный линейный размер кластера.



Рисунок 3 Исходное изображение «Самолёт», исходный размер 474×263 точек

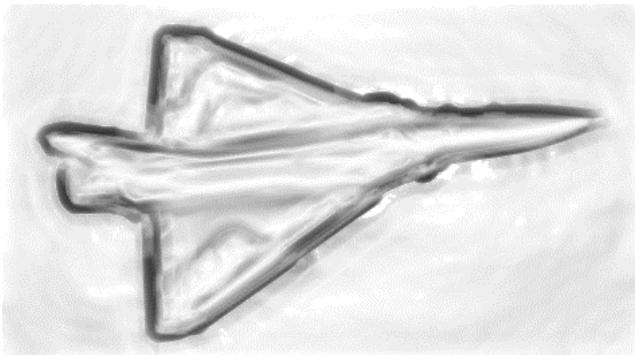


Рисунок 4 Мультиплексированное R -изображение (негатив по полиграфическим соображениям) снимка «Самолёт» ($L_i=64,32,16$)

4.2 Создание кластеров

Кластеры создаются поочерёдно в каждой точке (x_i, y_i) выбранной на предыдущем этапе.

4.2.1 Гистограммно-пространственное разделение точек окрестности.

Основой создания кластера служит гистограммно-пространственное разделение (ГПР) точек в окрестности на 2 группы. Рассматривается квадратная окрестность размером L с центром в (x_c, y_c) . По определению R -величины находится оптимальное, то есть дающее наибольшее

значение (1), деление (см. рис 2) рассматриваемой окрестности. Для каждой из двух полученных частей, находится её пик на гистограмме яркостей всей окрестности: его левая и правая границы определяется (см рис 1) как:

$$\begin{cases} Bleft_i = B_i - \sqrt{D_i}, \\ Bright_i = B_i + \sqrt{D_i}; \end{cases} \quad (4)$$

где i -номер пика ($i=1,2$); B_i и D_i - средняя яркость и дисперсия по части области; здесь и далее предполагается, что $B_1 < B_2$.

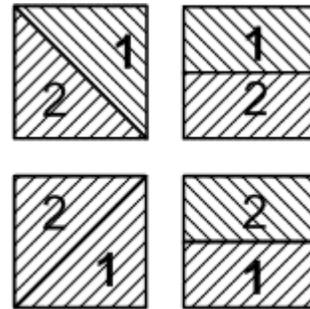


Рисунок 5 Типы деления рамок, использовавшиеся для гистограммно-пространственного разделения точек в окрестности.

В случае, если $Bleft_2 \leq Bright_1$, то есть пики перекрываются, значения $Bleft_2$ и $Bright_1$ пропорционально уменьшаются, так, чтобы между ними оставалось «нейтральное» расстояние $(B_2 - B_1)/3$. Цель этого изменения - уменьшить влияние точек, которые по яркостным характеристикам могут принадлежать как к одному сегменту, так и ко второму.

На последнем этапе ГПР области перебираются все её точки. Если пиксель пространственно лежит в полу-области i , а его яркость лежит в пределах от $Bleft_i$ до $Bright_i$, то пиксель относится к сегменту i .

4.2.2 Начальный этап создания кластера

Сегментация, получаемая непосредственным применением алгоритма ГПР не устойчива в том смысле, что при изменении координат центра области на 1-2 пикселя граница между получаемыми областями может измениться существенным образом. Неоднозначность, о том, какая же из границ более верна, удалось снять усреднением полученного хода границы.

Для этого, ГПР производится во всех областях центры которых отлежат друг от друга на незначительном расстоянии:

$$\begin{cases} x_i - L/2 \leq x_c \leq x_i + L/2, \\ y_i - L/2 \leq y_c \leq y_i + L/2; \end{cases} \quad (5)$$

здесь (x_i, y_i) - очередная точка максимума R -изображения, а (x_c, y_c) - точка центра квадратной окрестности, в которой производится ГПР.

Здесь можно оценить ранее упомянутый размер кластера. Он примерно равен расстоянию от левой границы рамки смещенной на $L/2$ влево, от центральной точки, до правой границы области смещенной на $L/2$ влево.

Результаты каждого разделения не наносятся на выходное изображение, а сохраняются в буферах. Для каждой точки создаётся три буфера: $B_0(i, j)$ инкрементируется, если соответствующая точка (i, j) (в координатах всего изображения, а не области) не была отнесена ни к одному из сегментов; $B_1(i, j)$ - если точка была отнесена к первому сегменту, $B_2(i, j)$ - второму.

После того, как ГПР произведено для всех (x_c, y_c) , осуществляется анализ буферов. Точка (i, j) отмечается цветом первого сегмента на выходном изображении если:

$$\left\{ \begin{array}{l} \sum_{k=0,1,2} B_k(i, j) \geq 3, \\ \forall B_1(i, j); (B_2(i, j) + B_0(i, j) = 0) \\ \frac{B_1(i, j)}{\max(B_2(i, j), B_0(i, j))} \geq \frac{2}{3}, \\ (B_2(i, j) + B_0(i, j) \neq 0) \end{array} \right. \quad (6)$$

Для второго сегмента аналогичные условия с учётом изменения индексов 1 и 2.

Стоит отметить, что коэффициент $2/3$ выбран, что бы обеспечить требование: точка должна быть отнесена не в простом большинстве случаев, а именно в преобладающем числе случаев. Требование, к сумме буферов позволяет не принимать решения относительно тех точек, для которых нет абсолютно никакой статистики.

4.2.3 Конечный этап создания кластера

На конечном этапе создания кластеров по всем точкам, для которых

$$\sum_{k=0,1,2} B_k(i, j) \geq 3 \quad (7)$$

проводится итерационный цикл, в процессе которого для каждой неокрашенной точки анализируется её 4-х связанная окрестность (4 точки прилегающие сверху, снизу, слева и справа):

- Если 3 из этих точек принадлежат одному сегменту, а другому принадлежит либо одна, либо ни одной, то центральная точка относится к первому сегменту.
- Если в окрестности точки нет ни одной окрашенной точки, то она остаётся пока неокрашенной.
- В остальных случаях, точка относится к тому из сегментов, для которого расстояние между яркостью точки и средней яркостью по пику меньше.

Цикл прекращается, когда количество окрашенных точек после очередного прохода не изменяется.

На рис 6 представлены кластеры, созданные в процессе сегментации снимка «Самолёт». Видно, что границы «Самолёт-небо» выявлены очень точно, особенно в районе плоскостей и двигателя. На границы областей на небе, (которое имеет градиентную окраску в силу объективных физических причин), кластеры созданы, но граница между ними не явна.

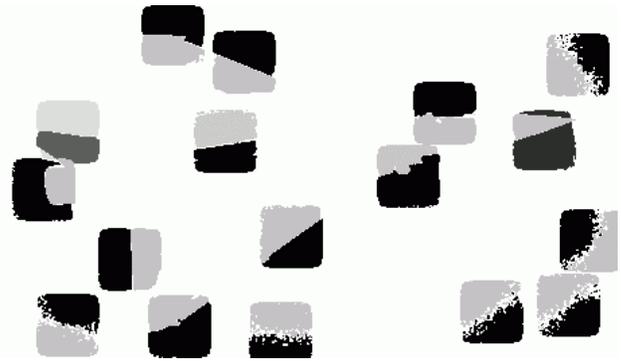


Рисунок 6 Кластеры, созданные в процессе сегментации снимка «Самолета», $L=16$

4.3 Рост кластеров

Кластеры наращиваются за счёт прилегающих к ним точек, чья яркость отличается не более чем на пороговое значение, которое постепенно увеличивалось в процессе их роста.

Перед тем, как начать рост кластеров они все перебираются: определяется их максимальная 4-связанная область, вычисляется средняя яркость пикселей и её дисперсия, удаляются все остальные области, принадлежащие этому кластеру. На основе полученных данных, для каждого кластера вычисляются величины:

$$\begin{aligned} B_{up}(i) &= B(i) - \sqrt{D(i)}, \\ B_{down}(i) &= B(i) + \sqrt{D(i)}; \end{aligned} \quad (8)$$

здесь $B(i)$ - средняя яркость; $D(i)$ - дисперсия яркости по кластеру, i - индекс сегмента.

Алгоритм рост кластеров - итерационный процесс, состоящий из нескольких этапов, продолжающийся до полной сегментации изображения.

На каждом шаге итерации кластеры перебираются в порядке возрастания значения дисперсии. Для каждого кластера:

1. Из какой либо точки сегмента запускается заливка, которая добавляет к сегменту помимо ранее выделенных точек, точки лежащие в 4-связанной окрестности кластера с яркостью в пределах от $B_{up}(i)$ до $B_{down}(i)$.
2. Производится поиск кластеров, касающихся рассматриваемой области. Найденные области объединяются, если выполняется условие:
$$B(i) - B(j) \leq \min \left\{ \sqrt{D(i)}, \sqrt{D(j)}, \left| \sqrt{D(i)} - \sqrt{D(j)} \right| \right\} \quad (9)$$
здесь $B(i)$ и $D(i)$ - текущее значение средней яркости и её дисперсии для i -го сегмента.
3. Производится поиск областей точек, целиком лежащих внутри рассматриваемой области. Если такие области находятся, то они выделяются как отдельный сегмент.
4. Определяются новые значения величин $B_{up}(i)$ и $B_{down}(i)$:

$$\begin{cases} B_{up}(i) = \max(B_{up}(i) + 1, \text{MAX_COLOR}), \\ B_{down}(i) = \min(B_{down}(i) - 1, 0); \end{cases} \quad (10)$$

здесь MAX_COLOR – максимальное значение яркости исходного изображения.

Как уже было отмечено, условием выхода из итерационного процесса является полная сегментация всего изображения.

После того, как полная сегментация изображения получена, производится её постобработка. Она состоит в переборе всех сегментов, выделенных на шаге 3 алгоритма роста кластеров и анализе их площади, средней яркости и дисперсии яркости. Каждая такая область присоединяется к внешнему сегменту, если либо её площадь меньше минимального размера выделяемого сегмента, либо для этих областей выполняется условие (9). В остальных случаях, такая область выделяется как отдельный сегмент.

Цель выделения отдельного сегмента на шаге 3 – ускорить процесс заливки. Так же, это позволяет во многих случаях выделять небольшие, но очень контрастные и значимые объекты, например, светлые облака над морем. Область не прибавляется ко внешней сразу после обнаружения (когда, например, её размер меньше минимального), что бы не испортить яркостно-дисперсионные характеристики внешнего сегмента.

Выполнение условия (9) для пары областей фактически означает, что их пики распределения яркости на гистограмме совпадают: расстояние между центрами пиков меньше, чем квадратный корень из дисперсии или разницы между дисперсиями. Шаг 2 позволяет объединять кластеры, представляющие один и тот же объект, но лежащие изначально в разных начальных областях.

Простые правила роста пороговых величин $B_{up}(i)$ и $B_{down}(i)$ были подобраны экспериментально. Сегментация не сильно отличалась при использовании иных, более продвинутых методов, методов. Более обещающим является выбор неких контролирующих рост кластера правил [5]: наложение ограничений на поведение, например, дисперсии или средней яркости по сегменту, с ростом числа точек.

5. ЭКСПЕРИМЕНТАЛЬНЫЕ РЕЗУЛЬТАТЫ

Алгоритм был опробован на большом числе изображений: реалистических сцен, аэрокосмических снимков и т.д. Было выявлено, что алгоритм способен точно выделять контрастные границы сложной формы (в отличие от R-seg который такие границы скругляет, как правило). Вместе с тем, алгоритм менее адаптирован для сегментации изображений с нечёткими, размытыми границами (которые очень хорошо выделяет R-seg).

Так же как и R-seg, представленный алгоритм сегментации выделяет большое число сегментов. Однако, это проблема может быть решена использованием алгоритма уменьшения детальности сегментации. Хорошие результаты могут быть получены, например, алгоритмом [2].

5.1 «Самолет»

На рисунке 7 представлен результат сегментации снимка «Самолёт» (рисунк 3), для которого в работе были приведены R-изображение (рисунк 4) и кластеры

(рисунк 6). Для наглядности, в работе были использованы относительно большой размер величины минимального выделяемого размера объекта. При этом, были получена высокая точность всех существенных границы.

Очень точно были определены контуры самолета, как в области плоскостей, так и в области фюзеляжа. Исключение составляет кабина пилота, которая, в силу своей малой величины и яркости окраски была отнесена к небу, а не фюзеляжу.



Рисунок 7 Результат сегментации снимка «Самолёт» (сосуществующие сегменты окрашены различными цветами).

Почти точно определены были камуфляжные линии, проблемы возникли там, где они пролегли по мелким деталям на плоскостях. Эти недостатки объясняются выбором размера минимальной выделяемой области ($L=32$).

Наличие границ на, как кажется, однородном небе объясняется тем, что на фотоснимках небо всегда имеет градиентную окраску. Поэтому, в результате сегментации были выделены область примерно однородных частей неба.

5.2 «Девушка»



Рисунок 8 Результат сегментации фотоснимка «Девушка», исходные размеры 500×318

На рисунке 8 представлен фотоснимок «Девушка» и границы выделенных сегментов, нанесенные белыми линиями. Видно, что были очень точно выделены все существенные границы. Так же точно были определены очертания текста в правом верхнем углу и складки одежды на девушке – изогнутые границы с мелкими деталями.

Обилие сегментов с прямоугольными изгибами на фоне стены объясняется тем, что когда-то изображение было сохранено в формате JPEG. Теперь, фон состоит и ряда

областей, однородных по окраске, с границами, проходящими по границам квадратов 8×8 , использованных при компрессии.

К недостаткам результатов сегментации стоит отнести слияние темных частей одежды с покрывалом (например, в районе левого локтя). Однако здесь разделение исключительно по цветовым признакам, по-видимому, не возможно. Человек проводит границу в таких случаях, используя априорные знания о форме предмета.

5.3 «Бурый Медведь»

На рисунке 9 представлен результат сегментации снимка «Бурый Медведь» из коллекции Corel. Видно, что были правильно определены границы «растительность-небо», «медведь-небо», «дерево-небо», а так же ряд менее значительных: между различными видами растительности. К недостаткам стоит отнести то, что не были обнаружены границы на верхней лапе медведя (слилось с очень темной частью дерева), а так же не была найдена граница между деревом (слева) и растительностью в нижней части снимка.

Стоит отметить особое внимание, на точность, с которой были очерчены контуры нижней лапы медведя – фактически различимы отдельные пальцы.

Так же как и в снимке «Самолёт» небо было разделено на большое число сегментов. Обилие, в сравнении с «Самолётом» количества сегментов, объясняется различием размеров изображений и минимального размера выделяемого сегмента.

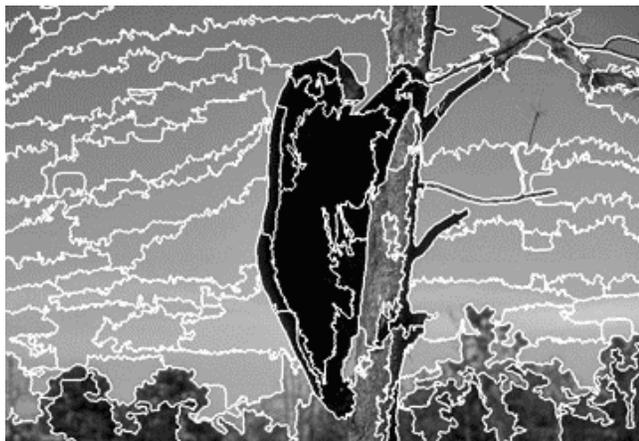


Рисунок 9 Результат сегментации снимка «Бурый Медведь» ($L_1=32$, $16,8$; $L=24$), исходный размер 720×494 точки

6. ЗАКЛЮЧЕНИЕ

Представлен новый алгоритм автоматической сегментации растровых изображений разработанные на основе ранее предложенного релеевского детектора площадных границ. Алгоритм обеспечивает очень точное выделение границ, в том числе с сильными изгибами.

7. ЛИТЕРАТУРА

[1] М.В. Минченков, Д.В. Юрин, А.В. Хельвас. «Алгоритм автоматической сегментации изображений на основе релеевского детектора границ двумерных объектов», труды

12-ой Международной Конференции по Графике и Компьютерному Зрению “GraphiCon-2002”.

[2] П.В. Непомнящий, Д.В. Юрин, «Построение иерархического дерева детальности изображения через поиск минимальных разрезов графа», труды 13-ой Международной Конференции по Графике и Компьютерному Зрению “GraphiCon-2003”.

[3] У. Прэтт, *Цифровая обработка изображений: пер. с англ.* – М: Мир, 1982, (в 2-х книгах)

[4] A. Sibert, “Segmentation based image retrieval”, труды конференции SPIE conference in storage and retrieval for Image and Video Databases VI, 1988.

Об авторе

Минченков Михаил Викторович, аспирант Московского Физико-Технического Института, контактный адрес электронной почты: micker@cos.ru