

УДК 519.687

*М.В. Минченков, А.В. Хельвас, Д.В. Юрин*  
Московский Физико-Технический институт  
(государственный университет)

## **АЛГОРИТМ АВТОМАТИЧЕСКОЙ СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ ДЛЯ СИСТЕМ ТЕХНИЧЕСКОГО ЗРЕНИЯ**

Предложен эвристический алгоритм сегментации реалистичных изображений, основанный на использовании функционала обнаружения границ.

Сегментация цифровых изображений позволяет перейти от представления данных в виде набора точек с различной яркостью к описанию сцены, как композиции объектов, каждый из которых характеризуется формой, площадью, взаимным расположением, яркостными, текстурными и другими признаками. Она является важной составляющей систем технического зрения и искусственного интеллекта. Сегментация нашла широкое применение не только в робототехнике [1,2], в задачах обработки аэрокосмических снимков [3,4,5] (ландшафтно-тематическая дешифровка, поиск в базе данных), медицинских диагностике [6], распознавании текстов [7], задачах обнаружения и распознавания. В процессе восстановления трехмерных сцен сегментация позволяет снизить вычислительную сложность задачи в части траекторного анализа и совмещения изображений, снятых с подвижной камеры, за счет фильтрации, основанной на принадлежности отметок на разных кадрах к сегментам имеющим похожие статистики.

Множество работ было посвящено различным методам сегментации: эвристические подходы [6,7,8], методы глобальной оптимизации [3,9], подходы, основанные на выделении границ [4,10,11] и др. Однако все разработанные на сегодняшний день алгоритмы имеют ограниченную область применения.

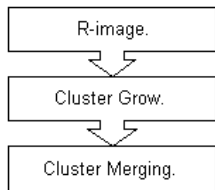


Рис. 1. Схема алгоритма сегментации

Предлагаемый эвристический подход основан на поиске границ площадных объектов и дальнейшего роста кластеров от наиболее вероятных центров объектов (рис. 1). Так как в большинстве случаев получаемая сегментация имеет чрезвычайно высокую детальность, то при необходимости она может быть понижена слиянием соседствующих сегментов, обладающих наиболее близкими дисперсионно-яркостными характеристиками.

Предлагаемый алгоритм сегментации может быть использован без вмешательства оператора и сохраняет работоспособность на изображениях с нечёткими границами. Существенна также возможность обобщения алгоритма на многослойные изображения для сегментации не только по интенсивности, но также текстуре, цвету и иным признакам [3].

### **Вероятностная характеристика обнаружения границы двух объектов в окрестности данной точки**

По-видимому, один из первых методов сегментации был основан на пороговом ограничении [1]: на гистограмме изображения находятся и нумеруются все пики, каждому из них сопоставляется сегмент. Каждый пиксель относится к тому сегменту, в пределах пика которого оказывается его яркость. Достоинства этого метода: простота, высокое быстродействие и минимальные требования к вычислительным ре-

сурсам. Недостаток – полное игнорирование пространственных взаимосвязей между точками изображений.

Наличие на изображении исключительно чётких границ и отсутствие сильно размытых объектов ещё не гарантирует успеха подобного рода сегментации. В силу центральной предельной теоремы теории вероятностей, гистограмма изображения будет близка к нормальному распределению при большом количестве различных объектов, присутствующих на нём. Вместе с тем, такой метод применим, если на изображении присутствуют лишь два-три достаточно однородных объекта. В силу своей однородности и малого количества пики на гистограмме отстоят друг от друга на заметном расстоянии, выделяются практически однозначно и соответствуют различным объектам. Такая ситуация реализуется для маленьких частей изображения – величиной порядка размера минимальной выделяемой области (далее характерный линейный размер минимальной области обозначен  $L$ ).

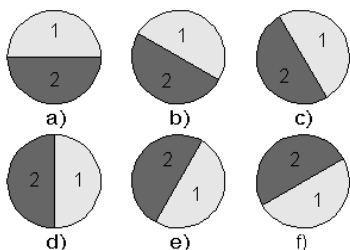


Рис. 2. Различные типы деления круглой рамки

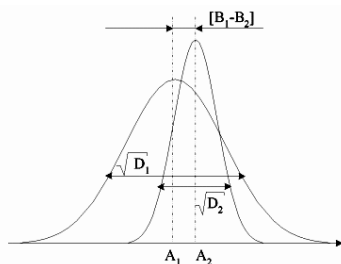


Рис. 3. Распределение яркостей, при котором критерий (2) указывает на различимость, а мера (1) нет

Рассмотрим фрагмент изображения, ограниченный круглой рамкой диаметра  $L$ . Гистограмма такого фрагмента может содержать один пик или два, по количеству лежащих внутри рамки объектов. Ситуации трёх и более пиков, в силу выбора размеров рамки, маловероятны. Представляет ли гис-

тограмма рамки двухпиковую структуру однозначно ответить практически не возможно. Численный критерий вероятности нахождения границы объектов вблизи данной точки изображения поможет упростить задачу.

Будем делить изображение в рамке на две части равной площади как показано на рис 2. Пики гистограммы, соответствующие различным частям рамки, будут практически неразличимы, если внутри рамки изображен один объект. Введем количественную меру различия распределений яркостей в двух половинках рамки по аналогии с критерием Рэлея различимости спектральных линий:

$$Q_{\theta} = \frac{\sigma_{1\theta} + \sigma_{2\theta} + \delta}{|B_{1\theta} - B_{2\theta}|}, \quad B_{1\theta} \neq B_{2\theta}, \quad \sigma_i = \sqrt{D_i} \quad (1)$$

здесь  $\theta$  обозначает направление деления рамки (рис. 2),  $B_i$  и  $D_i$  – средняя яркость и её дисперсия в соответствующей половинке рамки,  $\delta$  - константная величина, учитывающая погрешности и шумы реального изображения. Выражение (1) аналогично обратному значению линейного дискриминанта Фишера.

Однако, человеческий глаз способен воспринимать не только различия по яркости, но и по её разбросу (рис. 3), например, белый зверь на фоне чистого снега. Что бы учесть это, преобразуем выражение (1), а так же приведём численный критерий к диапазону  $[0,1]$ :

$$R = \begin{cases} \max_{\theta} \{1/(1 + Q_{\theta})\}, & \text{при } B_{1\theta} \neq B_{2\theta}, \\ 0, & \text{при } B_{1\theta} = B_{2\theta}; \end{cases} \quad (2)$$

$$\text{где } Q_{\theta} = \frac{\max(\min(\sigma_{1\theta}, \sigma_{2\theta}), \delta)}{|B_{1\theta} - B_{2\theta}|};$$

Определённая таким образом величина  $R$  (2) может быть интерпретирована как вероятность нахождения границы двух объектов в точке изображения, соответствующей центру рамки. Действительно, когда  $R$  близко к нулю, сумма полу-

ширин распределений яркости значительно больше расстояние между их центрами (распределения существенно перекрываются). При  $R$  близком к единице  $Q_\theta$  мало, расстояние между максимумами распределений яркости велико по сравнению с их полуширинами.

Заметим, что в отличие от поиска границ на основе анализа пиков гистограммы, величина (2) не только упрощает процедуру, но и в ряде случаев позволяет разделить текстурированные объекты (если различаются их средние яркости), что в методе порогового ограничения по гистограмме не возможно.

Изображение, значения пикселей которого вычисляются по формуле (2) будем называть  $R$ -изображением (рис 5-б). Его анализ позволяет указать точки, в окрестности которых проходит граница объектов и области, которые с большой долей вероятности представляют отдельный объект.

Критерий (2) легко обобщить на многоканальный случай (спектральные каналы или слои текстурных и иных признаков). Достаточно представить  $\sigma_i$  и  $B_i$  в виде векторов, каждая компонента которых есть соответствующая величина по  $j$ -му каналу. Вместо  $\sigma_1$  и  $\sigma_2$  в (2) подставляются абсолютные значения скалярных произведений соответствующих векторов корней из дисперсии на единичный вектор в направлении  $\vec{B}_1 - \vec{B}_2$ . То есть,

$$R = \begin{cases} \max_{\theta} \{1/(1+Q_\theta)\}, & \text{при } \vec{B}_{1\theta} \neq \vec{B}_{2\theta}, \\ 0, & \text{при } \vec{B}_{1\theta} = \vec{B}_{2\theta}; \end{cases} \quad (3)$$

$$\text{где } Q_\theta = \frac{\max(\min(|\sigma_{1\theta} \times \vec{e}|, |\sigma_{2\theta} \times \vec{e}|), \delta)}{|\vec{B}_{1\theta} - \vec{B}_{2\theta}|} \text{ и } \vec{e} = \frac{\vec{B}_{1\theta} - \vec{B}_{2\theta}}{|\vec{B}_{1\theta} - \vec{B}_{2\theta}|};$$

На различных разрешениях текстура проявляется по-разному, и использование информации, полученной при различных масштабах мелкости, позволяет более точно определять границы текстурированных объектов, не отвлекаясь при

этом на их содержание. В предлагаемом алгоритме это реализуется мультиплексированием – пиксельным перемножением  $R$ -изображений, построенных при различных размерах рамки.

$$R_{mult} = \prod_{k=0}^m R(L_k); \quad (4)$$

Так как  $R \in [0,1]$ , то и  $R_{mult} \in [0,1]$ .  $R_{mult}$  увеличивается там, где граница обнаруживается на всех масштабах и, наоборот, понижается в отсутствие границы на некоторых  $L$ .



Рис. 4. Исходное изображение «Самолёт»

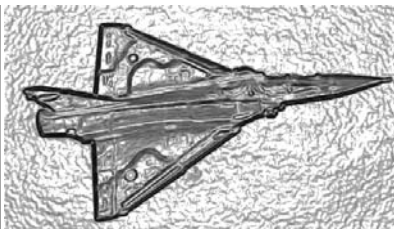


Рис. 5. Негатив  $R$ -изображения рис. 4, построенного по формуле (3).  $L=8$ ,  $\delta=1$

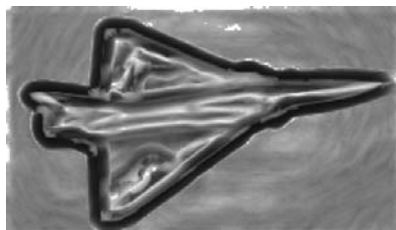


Рис. 6. Негатив  $R$ -изображения рис. 4, построенного по формуле (3).  $L=32$ ,  $\delta=1$

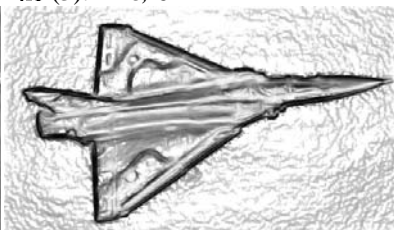


Рис. 7. Негатив мультиплексированного  $R$ -изображения (4) рис. 4,  $L_0=8$ ,  $L_1=16$ ,  $L_2=32$ ,  $\delta=1$ ,  $\varepsilon=0.06$

Безвозвратную потерю границы в точках, где  $R$ -величина приняла нулевое значение на одном из масштабов,

предлагается избегать, линейно преобразуя  $R(L_k)$  в  $[\varepsilon, 1]$ , где  $0 < \varepsilon < 0.5$ .

Пример R-изображения, посчитанного по формуле (4), приведён на рис. 7. Границы существенных объектов (контуры самолёта) стали более отчётливыми, чем на рис. 5-6, а малоинформативных стали менее заметны.

### Заливка R-изображения

Идея сегментации изображения способом заливки вспомогательного серого изображения, на котором точкам вблизи границ объектов соответствуют большие значения была предложена в [10]. Закраска осуществлялась методом порогового ограничения по последовательно возрастающему  $\sigma$ , зависящего от дисперсии всего изображения. Недостатком такого метода является игнорирование нечётких границ, (в силу большой дискретности порога) которым соответствует небольшой перепад яркости.

Использование в предлагаемом алгоритме заливки Брезенхэма [11], а так же малая дискретность набора пороговых ограничений лишают его указанных недостатков. Предлагаемый алгоритм (рис 8) строится как циклический по всем градациям яркости R-изображения. *SelectLevel()* на выходном изображении закрашивает цветом  $C_1$  все несегментированные точки, в которых R-величина не превышает текущего значения порога  $T$ . На этапе выполнения *OperateSelected()* такие точки заливкой Брезенхэма группируются в области, для каждой из которых подсчитываются значение площади ( $S_{region}$ ) и количество ( $N_{neig}$ ) ранее выделенных сегментов, лежащих в её 4-связанной окрестности. Для каждой области вызывается функция *DetClustColor()* (рис 8), которая определяет принадлежность каждой из областей к тому или иному сегменту.

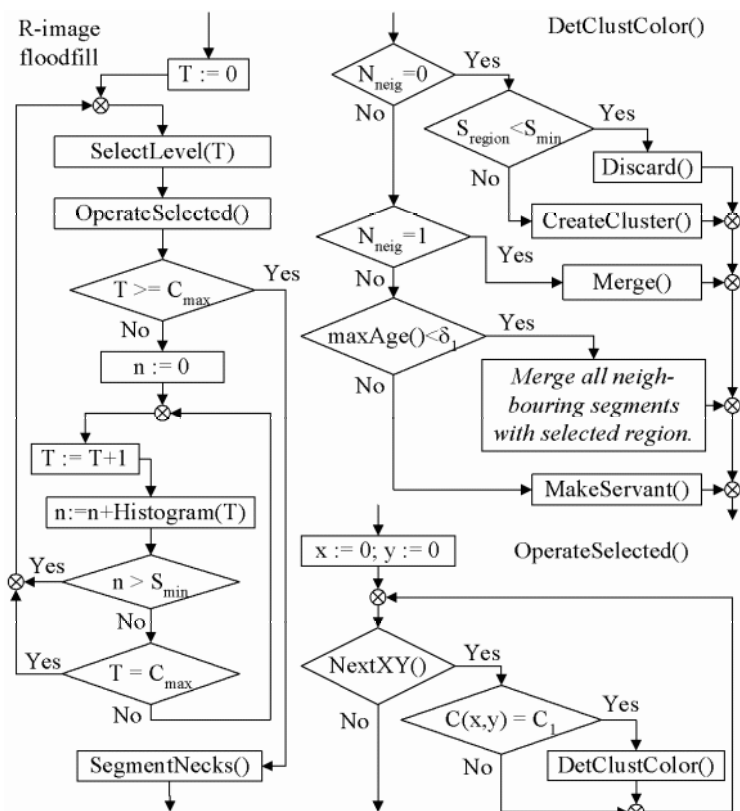


Рис. 8. Общая схема алгоритма заливки R-изображения и детализация функций OperateSelected() и DetClustColor()

Область, в зависимости от  $N_{neig}$  и «возраста» (возвращаемого функцией  $maxAge()$  самого старого, из прилежащих сегментов, на котором были образованы касающиеся её сегменты, может:

1. Образовать новый отдельный сегмент ( $CreateCluster()$ ).
2. Увеличить один существующих сегментов ( $Merge()$ ).
3. Объединить всех соседствующих с ней сегментов.
4. Отнесена к служебному сегмента  $C_2$  ( $MakeServant()$ ).



5. Перекрасится в цвет фона  $C_0$  ( $Dicard()$ ).

Наличие служебного сегмента  $C_2$  ускоряет работу алгоритма в целом за счет предотвращения излишних вызовов  $Dicard()$  и заливки Брезенхэма узких областей, соответствующих границам сегментов, которые покрывают, как правило, единой паутиной всю область изображения. Точки с цветовым индексом  $C_2$  не могут быть добавлены к какому-либо сегменту на этапе заливки и могут не являться 4-связанной областью.

Для перехода к следующей итерации и функции  $SelectLevel()$  с параметром  $T_{new}$  точек, лежащих в интервале от  $T$  до  $T_{new}$  должно быть более чем  $S_{min}$ . Для этого функция  $Histogram(T)$  возвращает количество «незалитых» точек  $R$ -изображения, яркость которых равна  $T$ .

Определение принадлежности к тому или иному сегменту точек, окрашенных в процессе заливки в цвет  $C_2$ , происходит во время работы функции  $SegmentNecks()$ . Сначала все они перекрашиваются в цвет фона. Затем ширина полученных проливов уменьшает по правилам алгоритма утоньшения линии, а оставшиеся окрашиваются исходя из количества точек лежащих в их 4-связанной окрестности.

В экспериментальной работе бралось  $\delta_1=3$ .

**Уменьшение детальности сегментации слиянием кластеров.**

Проблеме оптимального объединения группы сегментов на изображении посвящено немало работ. В [3] описан подход, использующий функционал Мамфорда-Шаха.

$$E(u, K) = \int_{\Omega} \|u - g\|^2 + \lambda \ell(K); \quad (5)$$

здесь  $g$  – функция исходного изображения,  $u$  – его аппроксимация в результате сегментации,  $K$  – текущий набор

границ сегментов суммарной длины  $\ell(K)$ ,  $\lambda$  - коэффициент, учёта границ.

Как видно, значение этого функционала зависит от среднеквадратичного отклонения аппроксимации от исходного изображения и длины границ сегментов. С ростом числа сегментов значение функционала сначала уменьшается за счёт первого слагаемого, а в последствии растёт, за счёт второго. В зависимости  $\lambda$ , возможны так же и предельные случаи – функционал либо убывает ( $\lambda=0$ , сегментация аппроксимирует исходное изображение всё более и более точно), либо возрастает (слишком большое  $\lambda$ , значение функционала от точности аппроксимации не зависит, а длина границ постоянно возрастает).

Наибольшее уменьшение (5) достигается слиянием пары сегментов, для которых минимальна величина:

$$\Delta E = \frac{|O_i||O_j|}{|O_i| + |O_j|} \|u_i - u_j\| - \lambda \ell(K_{ij}); \quad (6)$$

здесь  $|O_i|$  - площадь  $i$ -го сегмента,  $\|u\|$  - норма вектора  $u$ , в случае одноканального изображения, когда  $u$  – скаляр, просто его абсолютное значение. Нетрудно заметить, если значение (6) положительно, то происходит уменьшение (5), иначе его увеличение.

Этап слияния кластеров построен по жадному алгоритму и представляет собой циклический процесс минимизации функционала (5):

1. Просматриваем всевозможные пары соседствующих сегментов. Пара, для которой значение (6) минимально сливается.

2. Если знак (6) для выбранной пары сегментов не изменяется, то переходим к п.1, иначе прекращаем слияние сегментов.

## Экспериментальные результаты

Программная реализация алгоритма была протестирована на большом числе снимков различного происхождения. По полиграфическим соображениям в данной статье не представлены результаты сегментации сложных по составу изображений, набор представленных результатов ограничен простыми для анализа сценами.



Рис. 10. Исходное изображение  
«Самолёт»



Рис. 11. Результат заливки R-изображения, представленного на рис. 5 ( $S_{min}=32$ )

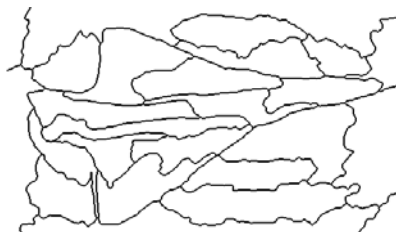


Рис. 12. Результат заливки R-изображения, представленного на рис. 6 ( $S_{min}=512$ )



Рис. 13. Результат заливки R-изображения, представленного на рис. 7 ( $S_{min}=128$ )

На рис. 11-13 приведены результаты заливки R-изображений представленных на рис. 5-7. Здесь черной линией отмечены границы сегментов. Сопоставление результатов сегментации при различных размерах рамки  $L=8$  и  $L=32$  (см рис. 5-6 и рис. 11-12) показывает, что чем больше размер рамки, тем меньше количество сег-

ментов и тем меньше точность определения хода границ (например, передняя часть фюзеляжа самолета). Границы сегментов в местах сильного излома (например, задняя часть фюзеляжа: сопло и хвостовое оперение) представляются в виде огибающей низкой кривизны. Это связано с тем, что  $R$ -изображение по построению содержит в себе яркостную информацию исходного снимка, усреднённую по некоторой окрестности точки.

Комплексирование информации полученной при различных размерах рамки (см рис. 7 и рис. 13) позволяет добиться компромисса: с одной стороны – небольшое число сегментов, как при больших размерах рамки, с другой – точность границ, как при малых размерах. Анализ хода границ, представленных рис 13 показывает, что контурные линии практически всего самолёта (левой и правой плоскостей, носовой части фюзеляжа, а так же части сопла) выделены достаточно точно. Наибольшая погрешность хода границы - в области хвостового оперения, где имеется сильный изгиб границы «самолёт-небо».

Необходимо также обратить особое внимание, на сегменты вне самолёта – на небе. Их наличие объясняется неоднородностью окраски неба, которая отчётливо заметна при преобразовании гистограммы (например, повышение контраста) исходного изображения. В данном случае это малоинформативные сегменты, однако, в иных случаях эта особенность разработанного метода сегментации может быть применена для обнаружения замаскированных объектов.

Снимок «Орёл» взят из коллекции “*Corel professional photos CD*”. Негатив мультиплексированного  $R$ -изображения ( $L_1=8$ ,  $L_2=12$ ,  $L_3=24$ ,  $\delta=4$ ,  $\varepsilon=0.06$ ) построенного по цветному оригиналу приведён на рис. 14. Отчётливо видны контуры орла, а также очертания волн моря в нижней части изображения. Как и в предыдущем примере, в процессе сегментации было выделено большое количество сегментов (рис. 16). По-

сле уменьшения детальности сегментации (рис. 17) при  $\lambda=25000$  число сегментов свелось к 21.



Рис. 14. Изображение «Орёл» (“Corel professional photos CD”) приведённое к оттенкам серого

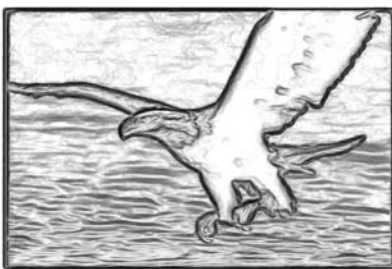


Рис. 15. Негатив R-изображения (4) рис. 14 ( $L_1=8, L_2=12, L_3=24, \delta=4, \varepsilon=0.06$ )



Рис. 16. Границы всех сегментов, полученных заливкой R-изображения рис. 15. ( $S_{min}=72$ )



Рис. 17. То же что и рис. 15 после уменьшения детальности сегментации ( $\lambda=25000$ )

Анализ границ сегментов, приведённых на рис. 17, показывает, что контур орла выделился достаточно точно. Выделилось всё хвостовое оперение. Разделение на различные сегменты хвостового оперения представляется оправданным, в силу того, что они достаточно различимы по яркости. Из-за сложной формы, не совсем точно выделились когти орла. Более того, не были выделены выступающие когти, величина которых значительно меньше, чем размер рамки. Голова орла представлена набором объектов, характеризующихся различными значениями средней яркости. Объединение существен-

но различных по яркости глаза и клюва птицы в один сегмент произошло на этапе уменьшения детальности сегментации и обусловлено мелкостью этих объектов. Форма клюва птицы определена не точно из-за мелких пятен различной освещенности, хорошо заметных на цветном изображении.

## Заключение

Предложен вероятностный критерий нахождения границы объектов в окрестности точки изображения. Представлен алгоритм автоматической сегментации изображений. Качество работы алгоритма было продемонстрировано на результатах сегментации некоторых снимков, типичных для систем технического зрения; из-за полиграфических ограничений для демонстрации выбраны наиболее простые изображения. Алгоритм был протестирован на большом количестве снимков различного происхождения (“*Corel professional photos CD*”, аэрокосмические изображения, изображения полученные с цифровых и пленочных фотоаппаратов, бортовой камеры мобильного робота “*Pioneer-2AT*”) и показал удовлетворительное качество сегментации.

## СПИСОК ЛИТЕРАТУРЫ

1. Мошкин В.И., Петров А.А., Титов В.С., Якушиенков Ю.Г. Техническое зрение роботов, М., Машиностроение, 1990. -272 с..
2. James Bruce, Tucker Balch, Manuela Veloso Fast and inexpensive color image segmentation for interactive robots // Proceedings of the 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '00) — 2000 — V.3. — P.2061-2066.
3. Koepfler G., Lopez C., Morel J.M. A Multiscale Algorithm for Image Segmentation by Variational Method // SIAM Journal on Numerical Analysis — V.31 — N 1. — P. 282-299.

4. *Newsam S., Bhagavathy S., Fonseca L., Kenney C., Manjunath B. S.* Object based representations of spatial images // 51st International Aeronautical Congress — 2000.
5. *Siebert A.* Segmentation based Image Retrieval // SPIE Conference on Storage and Retrieval for Image and Video Databases — 1998 — V.1.
6. *Huseyin Tek, Benjamin B. Kimia* Volumetric Segmentation of Medical Images By Three-Dimensional Bubbles // CVIU — 1997.
7. *Williams P.S., Alder M.D.*, Generic texture analysis applied to newspaper segmentation // Proc. ICNN96 — 1996 — V.3 — P. 1664-1669.
8. *Maxwell B.A.* Segmentation and Interpretation Using Multiple Physical Hypotheses of Image Formation // Tech.Rep. Carnegie Mellon University, 1996.
9. *Marroquin J.L.* Probabilistic Solution of Inverse Problems // Tech. Rep. Massachusetts Institute of Technology, 1985.
10. *Deng Y., Manjunath B.S., Shin H.* Color Image Segmentation // IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR '99 — 1999 — V.2. — P.446-451.
11. *Ma W.Y., Manjunath B.S.* Edge flow: a framework for boundary detection and image segmentation // IEEE International Conference on Computer Vision and Pattern Recognition, CVPR'97 — 1997.
12. *Уилтон Р.* Видеосистемы персональных компьютеров IBM PC и PS/2. Руководство по программированию/ Пер. с англ. К.Г.Смирнова; Под ред. В.Л.Григорьева. — М.: Радио и Связь, 1994.

Получено 29.05.2002