Image Segmentation by Directed Region Subdivision

T. Gevers and V. K. Kajcovski
Faculty of Mathematics & Computer Science
University of Amsterdam
Kruislaan 403, 1098 SJ Amsterdam, The Netherlands

Abstract

In this paper, an image segmentation method based on directed image region partitioning is proposed.

The method consists of two separate stages: a splitting phase followed by a merging phase. The splitting phase starts with an initial coarse triangulation and employs the incremental Delaunay triangulation as a directed image region splitting technique. The triangulation process is accomplished by adding points as vertices one by one into the triangulation. A top-down point selection strategy is proposed for selecting these points in the image domain of grey-value and color images. The merging phase coalesces the oversegmentation, generated by the splitting phase, into homogeneous image regions.

Because images might be negatively affected by changes in intensity due to shading or surface orientation change, we propose homogeneity criteria which are robust to intensity changes caused by these phenomena for both grey-value and color images.

Performance of the image segmentation method has been evaluated by experiments on test images.

1 Introduction

The basic idea common to split-and-merge based image segmentation methods is to split or merge image regions until they all satisfy a specific homogeneity criterion. Because of the extensive literature on split-and-merge image segmentation, we leave the reader to the references, [3], [6], for example. The most common used data structure in split-and-merge algorithms is the quadtree data structure because of its simplicity and computational efficiency [8]. However, a major drawback of quadtree-structured split-and-merge algorithms is their inability to adjust their tessellation to the underlying structure of the image data because of the rigid rectilinear nature of the quadtree structure.

To optimize this drawback, an image segmentation method is proposed in this paper that employs the incremental Delaunay triangulation as a directed region partitioning technique which adjusts the image tessellation to the semantics of the image data.

The paper is organized as follows. In Section 2, the incremental Delaunay triangulation is incorporated into the split-and-merge concept. Appropriate homogeneity criteria, to determine whether an image region is homogeneous, are proposed in Section 3 for grey-value and color images. Because the incremental Delaunay triangulation adds points one by one into the triangulation, point selection strategies are proposed in Section 4. Experimental results are discussed in Section 5 and finally several conclusions will be drawn.

2 Split-and-merge Concept based on Directed Image Region Partitioning

The quadtree structure is a very popular geometric data structure in split-and-merge segmentation algorithms. Although the quadtree is simple and computational efficient, it has an inherent drawback which is its insensitivity to image semantics. The rigid rectilinear image tessellation process causes errors in segmentation (edges of segments consist of only horizontal and vertical orientations) and affects negatively the number of split and merge operations.

To reduce these side effects, we perform the splitting phase by an incremental Delaunay triangulation. A Delaunay triangulation of a set of points is a triangulation in which the circumcircle of any of its triangles does not contain any other point in its interior, [7], for example.

The image segmentation method is then as follows.

The initial triangulation consists of two disjoint triangles from which the vertices correspond to three arbitrary image corners of the image domain.

The top-down adaptive image splitting phase starts with the initial coarse triangulation and successively examines triangles T by computing homogeneity function $P(T)$ where $P : T \rightarrow$ {TRUE,FALSE}. If T is not homogeneous then a point in T is obtained by point selection function $L(T)$ where $L : T \rightarrow \vec{x} \in T$. The point selected by $L()$ is entered as a vertex into the triangulation process to locally refine the image tessellation defined by the triangulation. The splitting phase continues until all triangles satisfy $P()$. The merging phase starts with the triangulation produced by the splitting phase. It merges adjacent image regions satisfying $P()$.

The image segmentation method depends on the two important functions: $P()$ and $L()$. In Section 4, we consider $L()$. In the next section, $P()$ is discussed for both grey-value and color images.

3 Homogeneity Criteria

3.1 Grey-value Images

The intensity of light reflected by a point on a matte surface is given by [2]:

$$I = k_d I_a + \frac{k_d I_s}{d + d_0}(\overline{N} \cdot \overline{L}) \qquad (1)$$

I_a is the amount of ambient light falling on the point, I_s is the intensity of the source and k_d is a reflection coefficient. Parameter d is the distance from the point to the point source and d_0 is adjusted if light sources are larger than point sources. Unit vector $\overline{N}$ is the normal vector describing the orientation of the point on the surface and the direction to the light source is described by unit vector $\overline{L}$. A major cause of intensity changes are due to surface orientation change $cos\theta = \overline{N} \cdot \overline{L}$. If neighbour pixels have intensity differences caused by variations of $cos\theta = \overline{N} \cdot \overline{L}$, homogeneity criterion $P()$ should be able to manage these intensity changes adequately. Therefore, $P()$ makes use of low-order bivariate polynomial functions as approximating functions. If their order is not too high, these functions approximate well the changes in intensity caused by surface orientation change. These low-order polynomials can be written as:

$$f_m(\vec{x}) = \sum_{i+j \leq m} a_{ij} x^i y^j,\ m \leq 3 \qquad (2)$$

A least-square solver calculates vector $\vec{a}$ and mean fitting error ϵ for image region T:

$$\epsilon = \sqrt{\sum_{\vec{x} \in T} \frac{(g(\vec{x}) - f_m(\vec{x}))^2}{N}} \qquad (3)$$

where $g(\vec{x})$ is the grey-value image and N is the total of surface interior pixels contributing to the sum.

If the mean noise variance σ in T is estimated, the mean fitting error ϵ can be compared to this. If the noise is additive, stationary and has zero mean (i.e. white noise), an estimate of σ is obtained by a local least-square 3x3 linear fit to the image, and not including pixels with high slopes to keep edges from contributing to the sum:

$$\sigma = \sqrt{\sum_{\vec{x} \in Q} \frac{(g(\vec{x}) - f_1(\vec{x}))^2}{9}} \qquad (4)$$

where Q is a 3x3 neighborhood.

Then, the functional approximating criterion defined over grey-value image region T is:

$$P(T) = \begin{cases} \textbf{TRUE}, & \text{if } \epsilon \leq \sigma \\ \textbf{FALSE}, & \text{otherwise} \end{cases} \qquad (5)$$

where ϵ is the mean fitting error of T and σ the mean noise variance of the image.

3.2 Color Images

Changes in intensity in grey-value images caused by surface orientation change has motivated the need for the functional approximating criterion.

However, for color images, the shading problem can be bypassed by using shadowing and surface orientation invariant color features for image segmentation. A comparison of different color features for image segmentation is discussed in [4]. Shadowing and shading invariant color features are obtained as follows [4]. Again, the diffuse reflection model is taken:

$$I_C = k_C I_a + \frac{k_C I_s}{d + d_0}(\overline{N} \cdot \overline{L}) \qquad (6)$$

for $C = (R, G, B)$ and k_R, k_G and k_B are coefficients of reflection for respectively the red, green and blue component.

Now, consider the normalized colors $r = I_R/(I_R + I_G + I_B)$, $g = I_G/(I_R + I_G + I_B)$ and $b = I_B/(I_R + I_G + I_B)$. These normalized colors discount surface orientation change:

$$\frac{I_R}{I_R + I_G + I_B} =$$

$$\frac{k_R(I_a + \frac{I_s}{d+d_0}(\overline{N} \cdot \overline{L}))}{(k_R + k_G + k_B)(I_a + \frac{I_s}{d+d_0}(\overline{N} \cdot \overline{L}))} = \qquad (7)$$

$$\frac{k_R}{k_R + k_G + k_B}$$

Equal argument holds for green and blue.

Normalized colors also discount shadowing. Points on surfaces not illuminated by a point source have only ambient light intensity:

$$\frac{I_R}{I_R + I_G + I_B} = \frac{k_R I_a}{(k_R + k_G + k_B) I_a} = \frac{k_R}{k_R + k_G + k_B} \tag{8}$$

Equal argument holds for green and blue.

Hence, image pixels measuring the intensity of diffuse reflected light from points on the same physical homogeneous colored surface will produce similar normalized colors discounting shadows and surface orientation change.

However, a problem arises when using rgb for image segmentation, because normalized colors become unstable and meaningless due to noise when intensity $(I_R + I_G + I_B)$ is small. Also rgb is redundant because $b = 1 - r - g$. Therefore, b is discarded and L^* is incorporated in the color system yielding the L^*rg color system, where L^* is based on the scaling of intensity:

$$L^* = \begin{cases} 116(\frac{Y}{Y_0})^{\frac{1}{3}} - 16, & \text{if } \frac{Y}{Y_0} > 0.008856 \\ \\ 903.3(\frac{Y}{Y_0}), & \text{if } \frac{Y}{Y_0} \leq 0.008856 \end{cases} \tag{9}$$

$$r = \frac{R}{R + G + B} \tag{10}$$

$$g = \frac{G}{R + G + B} \tag{11}$$

R, G and B are the red, green and blue images recorded by a color camera and $Y = 0.299R + 0.587G + 0.114B$. Y_0 is a nominally white object-color stimulus.

Because normalized colors are shading invariant for the diffuse reflection model, $P(T)$ is considered homogeneous if the values of r and g are constant, possibly contaminated by noise, and equals the functional approximating criterion for $m = 0$, $f_0(\vec{x}) = a_{00}$. A least-square solver computes a_{00} and fitting errors ϵ_r and ϵ_g for image region T:

$$\epsilon_c = \sqrt{\sum_{\vec{x} \in T} \frac{(g_c(\vec{x}) - a_{00})^2}{N}} \tag{12}$$

where $g_c(\vec{x})$ is the normalized color image for $c = (r, g)$.

Because r and g become unstable due to noise when L^* is small, they are not considered in regions with an average value of L^* below threshold t_{L^*} (i.e. dark regions). Instead of r and g, L^* is used. The value of t_{L^*} is directly proportional to the amount of noise in the original R, G and B images.

To measure the upper bounds of σ_c, Q is a 3x3 neighborhood with average L^* value equal to t_{L^*} to account for the worst case of the propagation of noise error:

$$\sigma_c = \sqrt{\sum_{\vec{x} \in Q} \frac{(g_c(\vec{x}) - f_1(\vec{x}))^2}{9}} \tag{13}$$

for $c = (r, g)$.

Homogeneity criterion $P()$ for color images is given by:

$$P(T) = \begin{cases} \textbf{TRUE}, & \text{if } \epsilon_r \leq \sigma_r \text{ and } \epsilon_g \leq \sigma_g \\ \\ \textbf{FALSE}, & \text{otherwise} \end{cases} \tag{14}$$

if average value of L^* in T exceeds t_{L^*}, otherwise $P()$ is defined by:

$$P(T) = \begin{cases} \textbf{TRUE}, & \text{if } \epsilon_{L^*} \leq \sigma_{L^*} \\ \\ \textbf{FALSE}, & \text{otherwise} \end{cases} \tag{15}$$

where $\epsilon_{L^*} = \sqrt{\sum_{\vec{x} \in T} \frac{(g_{l^*}(\vec{x}) - f_m(\vec{x}))^2}{N}}$ and $\sigma_{L^*} = \sqrt{\sum_{\vec{x} \in Q} \frac{(g_{L^*}(\vec{x}) - f_1(\vec{x}))^2}{9}}$.

4 Point Selection Strategies

In this section, we propose point selection strategies which guide the image subdivision process by adding points as vertices one by one into the triangulation. Because the purpose of the image segmentation method is to extract geometrically homogeneous image regions, meaningful points to be added into the triangulation are defined to be on locations where the image surface exhibits geometrically distinctive discontinuities.

Maximum Difference: The maximum difference strategy is to select the point at the location where its pixel value has maximum difference value with respect to the approximating function:

$$L(T) = \max_{\vec{x} \in T} \{(g(\vec{x}) - f_m(\vec{x}))^2\} \tag{16}$$

Edges: The Canny's edge detector [1] is used here to estimate surface normal discontinuities.

Corners: Corners are defined on locations where the image surface exhibits high values of isophote curvature. The point selection strategy uses the corner detector of [5] as a measurement of cornerness.

Lines: Lines are calculated by applying the Canny edge detector [1] followed by a line fitting technique [9]. The two end-points of a straight-line are inserted as vertices in the triangulation in such a way that the straight line will be translated into an edge of the triangulation.

For color images, the point selection strategies combine the results obtained from the r and g images on a max-basis to select significant points in T if average value of L^* in T exceeds t_{L^*}. Otherwise, the point selection strategy is based on the L^* image.

5 Experiments

In this section, the results are presented of the experiments that have been carried out with the proposed image segmentation method. The goal of the experiments is to determine to what extent the triangulation point selection strategies affect the number of split and merge operations and the homogeneity criteria the quality of the segmentation results. *md* denotes the maximum difference point selection strategy. *ed* and *ld* are point selection strategies based on edge and line information respectively and *cd* is based on corner information.

The image segmentation method has been written in the C programming language on a SPARC-LX workstation with the UNIX operating system. Images have 128∗128 spatial and 8 bits intensity resolution.

5.1 The CUBE Image

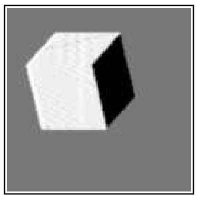


Figure 1: Original image

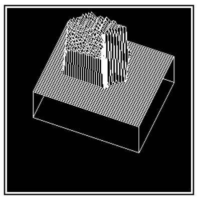
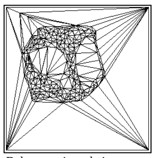
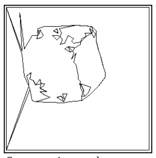
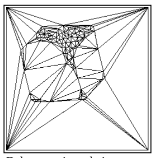


Figure 2: 2-D characteristics

The original grey-value image is shown in Figure 1. The two-dimensional grey-value characteristics are given in Figure 2.







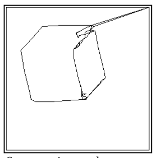
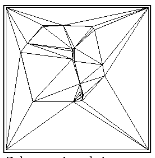
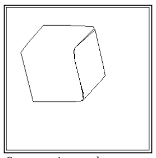
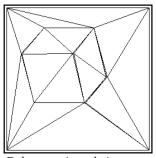


Figure 3: *md* Figure 4: *md* Figure 5: *ed* Figure 6: *ed*







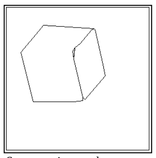


Figure 7: *cd* Figure 8: *cd* Figure 9: *ld* Figure 10: *ld*

In this subsection, we investigate to what extent the triangulation point selection strategies affect the number of split operations defined by the number of iterations. Therefore, we applied the point selection strategies on the same test image with the same homogeneity criterion. In this way, differences in the number of iterations and execution time is caused by the different triangulation point selection strategies, see table 11, where *iterations* is the number of iterations the splitting phase consists of, *triangles* the number of triangles after the splitting phase and *regions* the number of regions as output of the merging phase. According to table 11, we propose a top-down point selection strategy which starts with *ld*. If a line is present in T, the end-point are added as vertices into the triangulation. In case no meaningful lines can be detected, point selection strategy *cd* takes over. Then, in case of no meaningful corners in T, *cd* is replaced by *ed*. Finally, *ed* is substituted by *md* if necessary.

CUBE	iterations	triangles	regions	user time
md	262	526	26	7.738
ed	107	216	18	0.970
cd	26	54	6	0.553
ld	6	16	3	0.286

Figure 11: Table of measurements

5.2 The Cone Image

The grey-value version of a synthetically generated color image is shown in Fig. 12 consisting of three differently homogeneous colored objects (from left to right: a yellow cone, a blue square and a pink cone) on a background with diffuse reflection. The image is clearly contaminated by intensity changes due to noise, shadows and surface curvature. The results after the splitting and merging stage, superimposed on the image, based on color are shown in 13 and 15 respectively, those based on grey-value with $m = 2$ are shown Fig. 14 and 16. The color image is excellent segmented discounting shadowing and surface orientation change. The objects of the grey-value segmentation result are well segmented discounting surface curvature according to the definition of $P()$ for grey-value images.

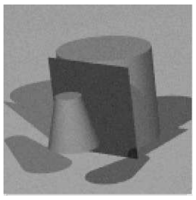


Figure 12:

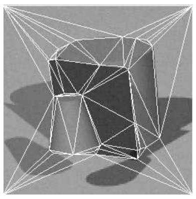


Figure 13: Color

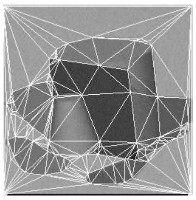


Figure 14: Grey

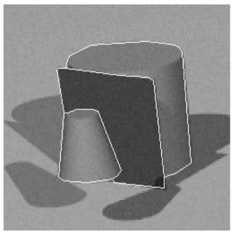


Figure 15: Color

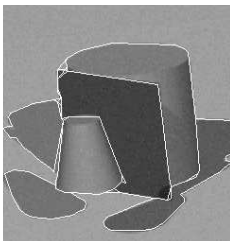


Figure 16: Grey

6 Conclusions

In this paper, an image segmentation method based on directed region partitioning has been proposed.

The image segmentation method consists of two separate stages: a splitting followed by a merging phase. Splitting starts with an initial coarse triangulation and employs the incremental Delaunay triangulation as a directed image region splitting technique. The merging phase coalesces the oversegmentation generated by the splitting phase into homogeneous image regions. Because the incremental Delaunay triangulation is accomplished by adding points as vertices one by one into the triangulation process, a top-down point selection strategy has been proposed for selecting these points. Appropriate homogeneity criteria have been proposed for grey-value and color images which are robust to changes in intensity due to shading or surface orientation change.

Excellent performance is shown, even in noisy images, by the experiments carried out on test images.

References

[1] Canny, J., *A Computational Approach to Edge Detection*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 8, No. 6, pp. 679-698, 1986.

[2] Foley, J. D. and van Dam, A., *Fundamentals of Interactive Computer Graphics* , Addison-Wesley Publishing Company, USA, 1982.

[3] Fu, K. S, and Mui, J. K, *A Survey on Image Segmentation* , Pattern Recognition, vol. 13, pp. 3-16, 1981.

[4] Gevers, T., *A New Approach to Segmentation of Color Images* , M.Sc. Thesis, Faculty of Mathematics and Computer Science, University of Amsterdam, The Netherlands, 1990.

[5] ter Haar Romeny, B. M., Florack, Lm, M., J., Koenderink, J. J., Viergever, M. A., *Scale-space: its Natural Operators and Differential Invariants* , Proc. Image Processing in Medical Imaging, 1991.

[6] Haralick, R. M. and Shapiro, L. G., *Survey: Image Segmentation Techniques* , Computer Vision Graphics and Image Processing, vol. 29, pp. 100-132, 1985.

[7] Lee, D. L. and Schachter, B. J., *Two Algorithms for Constructing a Delaunay Triangulation* , International Journal of Comput. and Inf. Sciences, vol. 9, no. 3, pp. 219-424, 1980.

[8] Samet, H., *The Quadtree and Related Hierarchical Data Structures*, Comput. Surveys, vol. 16, No. 2, pp. 187-230, 1984.

[9] Sharaiha, Y. M. and Christofides, N., *An Optimal Algorithm for the Straight Segment Approximation of Digital Arcs* , CVGIP, Vol. 55, No. 5, pp. 397-407, 1993.