Image Segmentation by Graph Partitioning

Ana Sofia Torres and Fernando C. Monteiro

Polytechnic Institute of Bragança, Campus Santa Apolónia, Apartado 1134, 5301-857 Bragança, Portugal

Abstract. In this paper we propose an hybrid method for the image segmentation which combines the edge-based, region-based and the morphological techniques in conjunction through the spectral based clustering approach. An initial partitioning of the image into atomic regions is set by applying a watershed method to the image gradient magnitude.This initial partition is the input to a computationally efficient region segmentation process which produces the final segmentation. We have applied our approach on several images of the Berkeley Segmentation Dataset. The results reveal the accuracy of the propose method.

Keywords: Image Segmentation, Graph partitioning, Normalized cut, Watershed Transform.
PACS: 87.57nm

INTRODUCTION

Image segmentation is a fundamental problem in image processing and computer vision. In many methods, the problem of image segmentation have been considered as a clustering problem, in a pixel-by-pixel approach, with the association of a feature vector to each pixel, in some n-dimensional feature space, and using a clustering algorithm to assemble these feature vectors.

Spectral-based segmentation treats image segmentation as a graph partitioning problem. These methods use the eigenvectors of a matrix representation of a graph to partition image into disjoint regions with pixels in the same region having high similarity and pixels in different regions having low similarity. A common characteristic among these techniques is the idea of clustering/separating pixels or other image elements using the dominant eigenvectors of a $n \times n$ matrix derived from the pair-wise similarities, as measure by one or more cues, between pixels where n denotes the number of pixels in the image. It thus segments an image from a global point of view.

In [1], Shi and Malik, presented an image segmentation approach as a graph partitioning and proposed a global criterion, the normalized cut, for segmenting the graph. Many extensions of the normalized cuts or related have been proposed [2, 3, 4, 5].

One major issue for segmentation methods based on graph representations is the size of the corresponding similarity matrix. If the node set V contains the pixels of an image, the size of the similarity matrix is equal to the squared number of pixels, and therefore generally too large to fit into computer memory completely (e.g. for an image of 481×321 pixels - the size of the images from the Berkeley Segmentation Dataset [6] - the similarity matrix contains $\approx 24 \times 10^9$ cells).

In this paper we propose a method that immensely reduce the problem size by replacing the individual pixels with micro regions in order to reduce the number of nodes in the graph. However, it is very important that the atomic regions will already yield a meaningful segmentation, i.e. the atomic regions must be homogeneous and the edges contained in the image must correspond to segment boundaries. The basic idea of the method resembles the perceptual grouping task: abandoning pixels as the basic image elements, we instead use small image regions (atomic regions) of coherent structure to define the corresponding graph representation. By treating regions as the elementary unit for image processing, we can reduce the computational complexity without a corresponding loss of accuracy.

This paper is organized as follows: next section gives a short description of the method used for the image segmentation. Followed by the experimental results and discussion. The conclusions marks are given in the last section.

IMAGE SEGMENTATION

Image segmentation is the identification of the regions that are uniform relatively to some parameter, e.g. image intensity or texture. One of the objectives for developing new algorithms in this area is the increasing of accuracy and reducing of the computational cost.

Numerical Analysis and Applied Mathematics ICNAAM 2012
AIP Conf. Proc. 1479, 802-805 (2012); doi: 10.1063/1.4756259

We create the atomic regions with the watershed transform over the gradient image [7]. The watershed transform tends to produce an over-segmentation due to the number of local minima of the gradient. A way to overcome this problem is to filter the images and eliminate the weaker gradient through gradient minima suppression. The pre-segmentation stage with the watershed transform is shown in Fig. 1.

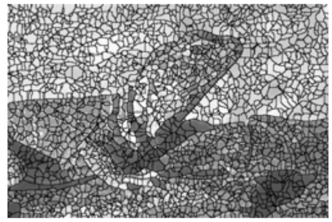


FIGURE 1. The atomic regions from watershed.

Normalized Cut

For a given image, we built the graph $G = (V,E,W)$, considering each pixel as a node (V) and defining the links between nodes (E) using a function of similarity. This weighted graph is used as a global descriptor of the image characteristics.

To reduce the number of nodes in the graph we replace the individual pixels by atomic regions which we obtain in the previous step. Thus, each atomic region is a node in a region similarity graph [8].

Shi and Malik [1] proposed the normalized cut segmentation for bipartitioning an image. Let V_A and V_B be two disjoint sets of the graph $V_A \cap V_B = \emptyset$. We define *links* (V_A,V_B) as the total weighted connections from V_A to V_B:

$$links(V_A,V_B) = \sum_{i \in V_A, j \in V_B} W_{i,j}. \tag{1}$$

Solving the *kNCut* problem is to finding a partition T that minimizes the function:

$$kNCut(T) = \frac{D_1 - W_{11}}{D_1} + \frac{D_2 - W_{22}}{D_2} + \ldots + \frac{D_k - W_{kk}}{D_k} \tag{2}$$

where $D_i = links(V_i,V)$, $D_i - W_i = (V_i, \overline{V}_i)$ and $\overline{V}_i$ represents the complement of V_i.

The multiclass partitioning problem of the graph G is formulated by an $n \times k$ indicator matrix $X = [x_1, \ldots x_k]$ shown in Eq. (3):

$$kNCut(T) = \frac{\mathbf{x}_1^T(D-W)\mathbf{x}_1}{\mathbf{x}_1^T D_{\mathbf{x}_1}} + \ldots + \frac{\mathbf{x}_k^T(D-W)\mathbf{x}_k}{\mathbf{x}_k^T D_{\mathbf{x}_k}} = k - \left(\frac{\mathbf{x}_1^T W\mathbf{x}_1}{\mathbf{x}_1^T D\mathbf{x}_1} + \ldots + \frac{\mathbf{x}_k^T W\mathbf{x}_k}{\mathbf{x}_k^T D\mathbf{x}_k}\right) \tag{3}$$

where $links(V_i, \overline{V}_i) = \mathbf{x}_i^T(D-W)\mathbf{x}_i$ and $links(V_i,V) = \mathbf{x}_i^T D\mathbf{x}_i$.

An approximate solution of this problem can be obtained by spectral methods, which implies the construction of the matrix W, calculating the eigenvalues and eigenvectors of the system:

$$(D-W)X = \Phi DX. \tag{4}$$

This is a NP-hard problem. If we considered the discrete values of x_i as continuous values y_i, using the method of Lagrange, the Eq. (4) can be expressed as a standard eigenvalue problem.

Making $\mathbf{y}_i = D^{1/2}\mathbf{x}_i$ and $Y = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k]$, the solution of the matrix W can be expressed by:

$$\widetilde{W}Y = Y\Lambda, \tag{5}$$

where $\widetilde{W} = D^{-1/2}WD^{-1/2}$ with $\Lambda = 1 = \lambda_1 \geq \ldots \geq \lambda_k$.

This sum is maximized by selecting the eigenvectors corresponding to the k largest eigenvalues of the continuous $\widetilde{W}$. Equation (3) becomes equivalent to:

$$\min_{X^T DX = I_k} kNCut(T) = k - \max_{Y^T Y = I_k} trace(Y^T \widetilde{W} Y) \tag{6}$$

in which $Y^T Y = I_k$ and $Y^T \widetilde{W} Y = Y^T Y \Lambda = I_k \Lambda = \Lambda$.

The LANCZOS algorithm provides an optimized solution for this problem with a time complexity of $O(n^{3/2}k)$, where n is the dimension of the matrix and K is the number of the eigenvectors.

Region similarity graph and pairwise spatial similarity

The quality of the segmentation depends on the pre-segmentation stage and on the affinity function between pixels. Using the centroid of each atomic region as a node of the graph, we measure the similarity of a pair of regions using the mean intensity of the regions and the gradient magnitude, or orientation energy, between them [9]. Let i and j be two atomic regions with a gradient magnitude OE^* between them, the intervening contours similarity contribution is given by:

$$w_i c(i,j) = exp\left[-\frac{max_i \in line(i,j) \left\|OE^*(\overline{x_i}, \overline{x_j})\right\|^2}{\sigma_{ic}^2}\right] \tag{7}$$

where $line(i,j)$ is the line between centroids $\overline{x_i}$ and $\overline{x_j}$ formed by t pixels.

The contribution of the mean intensity for the link weight according is given by:

$$w_I(i,j) = exp\left(-\frac{\left(I_{\overline{x_i}}, I_{\overline{x_j}}\right)^2}{\sigma_I^2}\right) \tag{8}$$

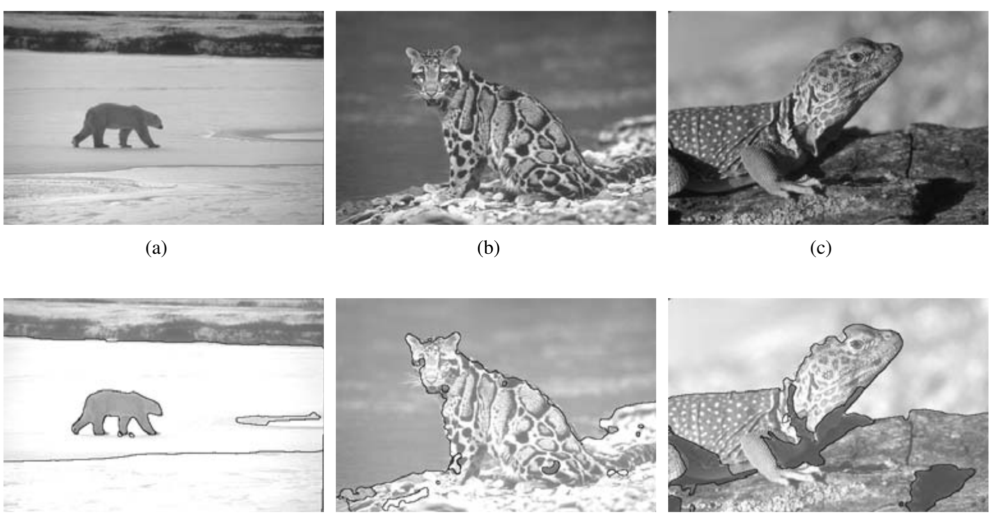
These cues are combined in a final link weight similarity function, with the values $\sigma_i c$ and σ_I selected in order to maximize the dynamic range of W:

$$W(i,j) = w_i c(i,j) \times w_I(i,j). \tag{9}$$

Additional features such as texture, could be added to the similarity criterion in the construction of the matrix.

RESULTS AND DISCUSSION

The images used in this paper were provided by the Berkeley Segmentation Database [6], a repository of natural images with corresponding human segmentations. The proposed method produces segmentation of high quality. The final segmentation results are shown in Fig. (2).



(a) (b) (c)

(d) (e) (f)

FIGURE 2. Example of image segmentation. First row: Original image. Second row: Segmentation result

Let S be the segmentation result and T be the ground truth, the measures of precision (P) and recall (R) can be expressed by:

$$P = \frac{matched\,(S,T)}{|S|} \qquad R = \frac{matched\,(S,T)}{|T|} \tag{10}$$

where $matched(S,T)$ is number of true positive pixels, $|S|$ and $|T|$ are the number of pixels in the segmentation result and the ground truth, respectively.

In probabilistic terms, precision is the probability that the result is valid, and recall is the probability that the ground truth data was detected. The two statistics may be distilled into a single figure of merit:

$$F = \frac{2RP}{R+P} \tag{11}$$

The results of the evaluation in terms of F-measure are presents in table 1.

TABLE 1. Quantitative results in term of Precision, Recall and F-measure.

Image	Precision	Recall	F
d)	0.5981	0.9379	0.7304
e)	0.5892	0.4521	0.5116
f)	0.7075	0.6252	0.6638

CONCLUSIONS

In this paper we have proposed an image segmentation method which combines edge-based and region-based information with spectral techniques through the morphological algorithm of watersheds. An initial partitioning of the image into primitive regions is set by applying a watershed simulation on the image gradient magnitude. This initial partition is the input to a computationally efficient graph partition process that produces the final segmentation. The latter process uses a region similarity graph representation of the image regions.

REFERENCES

1. J. Shi, and J. Malik, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**, 888–905 (2000).
2. C. Fowlkes, S. Belongie, F. Chung, and J. Malik, *IEEE Trans. Pattern Analysis and Machine Intelligence* **26**, 214–225 (2004).
3. E. Sharon, A. Brandt, and R. Basri, "Fast multiscale image segmentation," in *Proc. of IEEE International Conference on Computer Vision and Pattern Recognition*, 2000, vol. I, pp. 339–344.
4. S. Yu, "Segmentation using multiscale cues," in *Proc. of IEEE International Conference on Computer Vision and Pattern Recognition*, Washington DC, USA, 2004, vol. 1, pp. 247–254.
5. T. Cour, F. Benezit, and J. Shi, "Spectral segmentation with multiscale graph decomposition," in *Proc. of IEEE International Conference on Computer Vision and Pattern Recognition*, Software available at: *http://www.seas.upenn.edu/~timothee/software/ncut_multiscale/*, Washington, DC, USA, 2005, vol. II, pp. 1124–1131.
6. D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. of IEEE Int. Conference on Computer Vision*, 2001, vol. II, pp. 416–423, online at: *http://www.cs.berkeley.edu/projects/vision/grouping/segbench/*.
7. S. Beucher, and C. Lantuéjoul, "Use of watersheds in contour detection," in *International Workshop on Image Processing, Real-Time Edge and Motion Detection/Estimation*, Rennes, France, 1979.
8. F. C. Monteiro, *Region-Based Spatial and Temporal Image Segmentation*, Ph.D. thesis, Faculdade de Engenharia da Universidade do Porto, Porto (2008).
9. A. Torres, and F. Monteiro, *Proc. of the Int. Conference on Bio-inspired Systems and Signal Processing* pp. 434–437 (2012).