

SHAPE-BASED IMAGE SEGMENTATION USING NORMALIZED CUTS

Wenchao Cai^{1,3}, Jue Wu^{2,3}, Albert C. S. Chung^{1,3}

¹Department of Computer Science and Engineering and ²Bioengineering Programs,
³Lo Kwee-Seong Medical Image Analysis Laboratory,
The Hong Kong University of Science and Technology, Hong Kong.

ABSTRACT

To segment a whole object from an image is an essential and challenging task in image processing. In this paper, we propose a hybrid segmentation algorithm which combines prior shape information with normalized cut. With the help of shape information, we can utilize normalized cut to correctly segment the target whose boundary may be corrupted by noise or outliers. At the same time, we introduce the use of segmentation results of the normalized cut to guide the shape model, and thus avoid searching the shape space. The proposed method was demonstrated to be effective by our experiments on both synthetic and real data.

Index Terms— Image shape analysis, Image segmentation

1. INTRODUCTION

As a low level processing, image segmentation is an essential task in the fields of image processing and computer vision. In recent years, a lot of researchers have contributed in this field. For example, region growing techniques [1], active contour models [2] and the methods based on the level set framework [3] or MRF models [4] have made significant progress in image segmentation. In these methods, various low level features are employed, such as color, intensity and edges. However, in some situations, for example, in medical images, due to the low contrast of images or large variation of intensity inside the object, the targeted objects may be corrupted and the above methods may cause leakage of the boundary. It is difficult to segment the whole object. To solve these problems, Shi and Malik [5] proposed the normalized cut based on the graph theory and spectral clustering [6]. Normalized cut tries to ignore the details of the image to obtain a "big picture" of the structure in an image. Although this method can get a better result than other related methods in terms of segmenting the whole object, there are still problems in accurate segmentation due to its bias of equipartition and the graph structure [7].

It is natural to introduce prior information to get a better segmentation. Shape information of objects is widely used by researchers. Introducing shape information not only makes the computer "understand" the object, but also helps the computer to "infer" the blurred or occluded parts of the object. Wang *et al.* [8] put serial landmarks on the boundary of object to represent the shape. However, this approach as a shape representation suffers from problems such as inability to accurately capture high curvature locations, difficulty in handling topological changes, and finding point correspondences [9]. To overcome these problems, based on the level set methods [3], Leventon *et al.* [10] and Tsai *et al.* [9] utilized the signed distance map to represent the shape and greatly improved the results of segmentation. Because of its good properties, we also employ this representation in this paper.

In this work, our main contribution is to improve the performance of the normalized cut by introducing the shape information. Like most shape-based methods, the proposed method can correctly segment the object, even though a part of the boundary is missing or many noisy regions accompany the object. Besides, another advantage of the proposed method is that the eigenvectors from the normalized cut can generate the parameters of the shape model directly. In [10] and [9], considering the shape variation, they had to optimize an objective function with respect to the shape model to determine the parameters of the shape model, which increased the cost of computation. However, in our method, we can obtain the parameters easily through projecting the segmentation of the normalized cut onto the shape space.

The overview of our method is as follows. First, we extract a shape space consisting of a mean shape and principal components by using PCA (Principal Component Analysis) from a training set. Second, we obtain the eigenvectors of Laplacian matrix [11] derived from an affine matrix. Third, we project the segmentation generated from the eigenvectors onto the shape space to obtain a parametric shape model; and then, based on the constraint of the shape model, we modify the similarities of pairwise pixels, which are elements of the affine matrix. Finally, we solve the modified Laplacian matrix again and update the shape model. After several iterations of modifying the affine matrix and updating shape model, the final convergent result will segment the object from the image successfully.

2. PROPOSED METHOD

2.1. Normalized Cuts

In graph partition, the image is converted into an undirected weighted graph. Every pixel in the image corresponds to a vertex of a graph. And the weight on one edge is assigned according to the similarity between two corresponding pixels. The criteria of similarity are different in different applications. In general, the similarity can be defined by the distance, color, gray level, textures and so on. Considering the computation of the algorithm, we usually restrict the relationship of pixels in a neighborhood.

Unlike the edge cut algorithms [6, 12, 13] which only minimize the cut cost between classes, the normalized cut measures the total similarity within classes as well as the total dissimilarity between different classes.

Given a connected graph $G = (V, E)$ with a vertex set V and edge set E , the goal is to find an edge cut which separates the graph into k disjoint subsets such that $V = \bigcup_{i=1}^k V_i$ and $\forall(i, j), V_i \cap V_j = \emptyset$. We define that an edge cut C is a set of edges whose removal makes the graph disconnected. Generally, we want to minimize the cost of a cut:

$$Cut(C) = \sum_{e \in C} w_e. \quad (1)$$

In this paper, we try to segment an object from the background region. We, therefore, focus on discussing *two-way cut* of a graph. For convenience, denote the two parts as A and B . Then we can reform the cost of a cut:

$$Cut(A, B) = \sum_{p \in A, q \in B} w_{pq}. \quad (2)$$

Shi and Malik [5] pointed out that this criterion tended to cut a small segment from the graph, and based on the normalized cut, they proposed a new criterion which is a minimal cut penalizing small segments. The measure of normalized cut can be written as:

$$NCut(A, B) = \sum_{p \in A, q \in B} w_{pq} \left(\frac{1}{\sum_{p \in A} D_p} + \frac{1}{\sum_{p \in B} D_p} \right), \quad (3)$$

where $D_p = \sum_{q \in V} w_{pq}$ is defined as the degree of the vertex p . From Eq. 3, the normalized cut measures both the total dissimilarity between the different groups like Eq. 2, as well as the total similarity within groups.

To minimize the normalized cut, they deduced an efficient computational technique based on a generalized eigenvalue problem. First, we define W as the affine matrix of the graph, and $W(i, j) = w_{ij}$; D is a diagonal matrix, in which each element along the diagonal represents the degree of correspondent vertex i , D_i ; x is a label vector, if vertex i is in sub-graph A , $x(i) = 1$; otherwise $x(i) = -1$. Based on these notations, minimizing normalized cut results in the following objective function:

$$\min NCut(A, B) = \min \frac{y^T (D - W) y}{y^T D y}, \quad (4)$$

with the conditions $y_i \in \{1, -b\}$ and $y^T D 1 = 0$, where b is a constant less than 1.

Although this optimal problem is NP hard [5], if we ignore the discrete constraint, the eigenvector of $D^{-1}(D - W)$ corresponding to the smallest non-zero eigenvalue is an approximation. Then corresponding eigenvector is split into two parts by a threshold [13]. Yu and Shi [14] extended the normalized cut into k -way partition and tried to find a closer approximation through orthonormal transform.

Chung [11] proposed a "normalized" definition of the Laplacian matrix $L = D - W$ as $D^{-1/2} L D^{-1/2}$. The eigenvectors for this "normalized" Laplacian, when multiplied by $D^{-1/2}$, are exactly the generalized eigenvectors we used to compute the normalized cut. Because of many good properties of the "normalized" Laplacian matrix [11], the normalized cut method decomposes this matrix instead of $D^{-1}(D - W)$.

2.2. Parametric Shape Model

There are several methods to represent the shape of an object, e.g. landmarks, binary images and signed distance maps. Although it is a natural and popular way to put landmarks along the boundary, it is difficult to manage the landmarks, including removing and inserting landmarks according to different features. Binary images are sensitive to the boundary marked by hand in training samples. To avoid these problems, we employ the signed distance maps to represent shape. First of all, we have to segment the whole object and align all the training samples manually. Given the boundary of an object, a signed distance map is generated by computing the nearest distance between the point and the boundary. Pixels in the object region are assigned negative distance, pixels outside are positive, and the zero level corresponds to the boundary. Fig. 1 shows some distance maps obtained from a training set. Given a training sample,

we compute the distance map and reform the map into a vector V . From a training set of N samples, we obtain N corresponding vectors $V_i (i = 1, \dots, N)$, and the mean of the set $\bar{V} = 1/N \sum_{i=1}^N V_i$. Here we utilize PCA to extract the variation of the shapes. First, we subtract the mean from every sample, and then we obtain the covariance matrix:

$$M = \sum_{i=1}^N (V_i - \bar{V})(V_i - \bar{V})^T. \quad (5)$$

The principle components correspond to the eigenvectors with largest eigenvalues of matrix M . Then we obtain a shape space consisting of (1) mean shape \bar{S} , that is \bar{V} , and (2) m ($m < N$) eigenvectors $S_i \{1, \dots, m\}$ corresponding to m largest eigenvalues $\lambda_i^2 \{1, \dots, m\}$ (See Fig. 1). In this space, we can represent every shape by m coefficients x_i combined with \bar{S} and S_i as follows:

$$S = \bar{S} + \sum_{i=1}^m x_i S_i. \quad (6)$$

The shape S generated by Eq. 6 preserves the principal shape of the object. In this paper, we only consider the translation of the object. However, small variations of the shape size and orientations are included in the training samples.

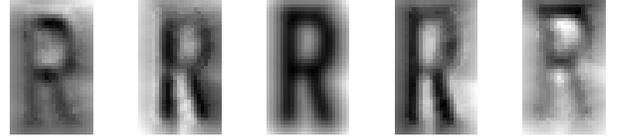


Fig. 1. Mean and variability of the shape of character R . From left to right: $\bar{S} - \lambda_2 S_2$; $\bar{S} - \lambda_1 S_1$; \bar{S} ; $\bar{S} + \lambda_1 S_1$; $\bar{S} + \lambda_2 S_2$.

2.3. Update Shape Model and Affine Matrix

Although the normalized cut can effectively extract the salient components in an image, it was shown that it could obtain bad results even in very simple situation due to its equipartition bias and the sinusoidal structure in the eigenspace [7]. Yu and Shi [15] tried to introduce some prior information to improve the performance by clustering some pixels manually in advance. Different from them, in this paper we incorporate the shape constraint into the normalized cut, which is inspired from the method presented in [7]. Tolliver *et al.* [7] modified the diagonal elements of the normalized Laplacian matrix using shape to maximize the lower bound of the second smallest eigenvalue of the matrix. In our method, we use the shape information to modify the whole affine matrix instead of the diagonal elements of the Laplacian matrix. Because the affine matrix represents the similarity of all pairwise pixels, in the ideal case, the similarity of two pixels belonging to the same group is much larger than the similarity of those belonging to different groups. Under the constraint of shape information, we can increase the similarity of pairwise pixels both in the same region (inside or outside of the object); and decrease the similarity, if they are in different regions. These adjustments make the object easier to be segmented.

Details of our algorithm are described below:

1. Segment the object from the training set, and compute the distance maps of all the samples;
2. Calculate the mean shape \bar{S} , and use PCA to extract m principal components $S_i \{i = 1, \dots, m\}$;
3. Construct the affine matrix $W^{(0)}$ from the input image based on distance, intensity and edge information;

4. For the k^{th} iteration, compute l smallest eigenvectors $Y_i \{i = 1, \dots, l\}$ of matrix $D^{(k)(-1)}(D^{(k)} - W^{(k)})$ using the normalized Laplacian matrix;
5. For every Y_i , search for a best threshold for the cut, and compare the corresponding segmentation result Y'_i with current shape model S (initialized by \bar{S}), $Eval_i = Y'_i \cdot S$, where \cdot represents dot product;
6. Find the maximum $Eval_{max}$ and the corresponding Y'_{max} , if the difference between $Eval_{max}$ and the last value $BestMatch$ (initialized by 0) is greater than the given threshold T_0 , then project Y'_{max} onto shape space to obtain the coefficients $x_i \{i = 1, \dots, m\}$ of PCA and the updated shape model $S = \bar{S} + \sum_{i=1}^m x_i S_i$, and $BestMatch = Eval_{max}$; otherwise, converge and exit with the shape model;
7. Use current shape model to construct a new affine matrix W_s :

$$W_s(i, j) = \begin{cases} 1 & \text{if } S(i) * S(j) > T_1; \\ W(i, j) & \text{if } abs(S(i) * S(j)) \leq T_1; \\ 0 & \text{otherwise;} \end{cases} \quad (7)$$

where $S(i)$ is the value of pixel i in the distance map of shape S , T_1 is a positive threshold;

8. Update the affine matrix $W^{(k+1)} = (1 - \lambda)W^{(k)} + \lambda W_s$, where λ is a parameter controlling the weight of shape constraint. Goto Step 4.

As mentioned above, the normalized cut tends to partition the image into two parts of similar size. Therefore, if the target takes up only a small part of the image or the contrast between the target and surroundings is low, the second smallest eigenvector may not segment the target. In general, different eigenvectors contain different components of the image. Therefore we calculate several smallest eigenvectors in Step 4. In Step 5, because the eigenvectors are continuous, we have to search a threshold to minimize the normalized cut by Eq. 4; then we use the shape model as a template to match every segmentation. Obviously, $Eval$ reaches the maximum when shape and segmentation are aligned exactly. Thus, we can find a coarse segmentation of the target, which may lose some boundary details. Through projecting the segmentation onto shape space in Step 6, we use the shape information to constraint the segmentation, since reconstruction from the shape space will subtract the unwanted parts and restore the missing parts of the object. In Step 7, according to the recovered shape, we construct W_s to adjust the similarities of pairwise pixels by combining it with the affine matrix W . We set most of the elements of the new affine matrix to 0 or 1, the minimum and maximum degrees of similarity respectively, to decrease or increase the similarity of pairwise pixels, while preserving the relationship around the boundary, where the shape information is not reliable.

3. EXPERIMENTAL RESULTS

To verify the performance of the proposed algorithm, we have performed the experiments on both synthetic and real data. Some parameters are set as follows. $T_0 = 50.0$, $T_1 = 1.0$, $\lambda = 0.3$; the radius of the neighborhood $R = 10$ and the affine matrix is:

$$W(i, j) = \begin{cases} \alpha e^{-\frac{\Delta I(i, j)^2}{\sigma_1^2}} e^{-\frac{D(i, j)^2}{\sigma_2^2}} & D(i, j) < R \\ 0 & \text{otherwise;} \end{cases}, \quad (8)$$

where $\Delta I(i, j)$ is the difference of intensity; $D(i, j)$ is the distance between i and j ; σ_1, σ_2 are regularized parameters, α is a term for

edge information, and $\alpha = 0.3$ if pixels i and j lie on two sides of an edge, otherwise $\alpha = 1.0$. Due to the limitation of pages, we mainly show binary images of the output.

3.1. Synthetic Data

We first tested our method on a data set including 26 letters ($A - Z$) and 10 numbers ($0 - 9$). For every letter and number, there were 50 samples with size 24×40 pixels. The training set consisted of 40 images. We added various levels of noise and outliers to the 10 remaining images and segmented them by the proposed algorithm. All of them could be segmented correctly if σ of Gaussian noise was less than 0.3 or no more than three 1-pixel width lines were added into the image. R is one of the most difficult letters to segment, because there is a hole in the letter and a narrow gap between two legs of R which is easy to be corrupted by noise. Its segmentation results with noise and outliers are shown in Fig. 2 and Fig. 3. In Fig. 2b, we

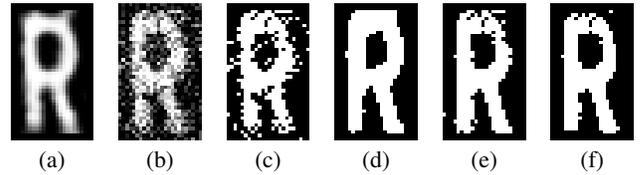


Fig. 2. Segmentation of character R . (a) Original image; (b) Input image corrupted by Gaussian noise ($\sigma = 0.3$); (c) Segmentation result by original NCut; (d) One of projective shape models; (e) Result of first iteration; (f) Final segmentation of our method.

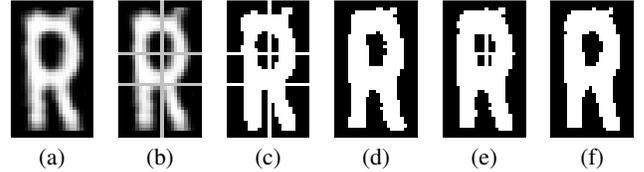


Fig. 3. Segmentation of character R . (a) Original image; (b) Input image corrupted by three lines; (c) Segmentation result by original NCut; (d) One of projective shape models; (e) Result of first iteration; (f) Final segmentation of our method.

added the Gaussian noise with $\sigma = 0.3$ to the original image, which eroded many pixels in R and cluttered the background. The segmentation of normalized cut (Fig. 2c) cannot eliminate the noisy pixels effectively. Using this noisy result (Fig. 2c), the proposed method can recover the shape (see Fig. 2d) of the target. The gradual effect of modification of the affine matrix can be seen in Fig. 2e and Fig. 2f. Another similar situation was shown in Fig. 3. It is more difficult to separate the lines from R than to eliminate the noise in Fig. 2 without using the shape constraint. Due to the effect of the noise and outliers, we still can see some corruption from the final results. In Fig. 2f, the boundary of final segmentation is not very smooth, and in Fig. 3f, the hole of R becomes smaller. If more noise is added, which can greatly corrupt the letter R , the proposed method may not work when $\lambda = 0.3$. In this case, we can increase the weight of shape information to put more emphasis on the shape constraint, e.g. $\lambda = 0.5$. However, if the value of λ is too large, we cannot make full use of the available image information. Thus the segmentation results cannot be very accurate.

3.2. Real Data

We also tested our method on the real data set of corpora callosa in human brain MR images. The data set contained 30 images of two

persons. Due to the difference of image size (49×35 pixels and 46×54 pixels), we obtained the shape space from the training sets of two persons (12 samples for each), and segmented the remaining images. Some results in difficult situations (there is an irrelevant part inside the arc of corpus callosum) are shown in Fig. 4 and Fig. 5. From the images, we find that the corpus callosum cannot be segmented by using the second smallest eigenvector of the normalized cut directly. This is because the corpus callosum occupies only a small part of the image. As such, we have to search several segmentation results for the best match. However, the initial segmentations (see Fig. 4d and Fig. 5e) are not satisfactory, since they include the irrelevant parts. Combining with the shape constraint, we can finally obtain much better results through several iterations (see Fig. 4g and Fig. 5g)).

To compare with the normalized cut, we also ran the program [16] of Yu and Shi [14] to segment the corpora callosa. To get the best result, we adjusted the number of grouping classes (8 for the image in Fig. 4a and 4 for the image in Fig. 5a) and separated the class corresponding to corpora callosa by hand. The results are shown in Fig. 4h and Fig. 5h.

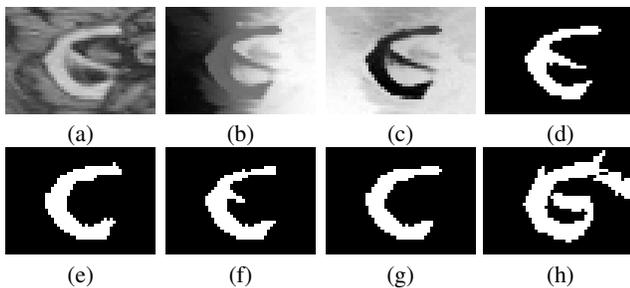


Fig. 4. Segmentation of corpora callosa. (a) Input image; (b) and (c) are 2^{nd} and 3^{rd} smallest eigenvector; (d) Segmentation result from (c); (e) Projective shape model; (f) Intermediate segmentation result; (g) Final segmentation of our method; (h) Segmentation result by Ncut of Shi's version.

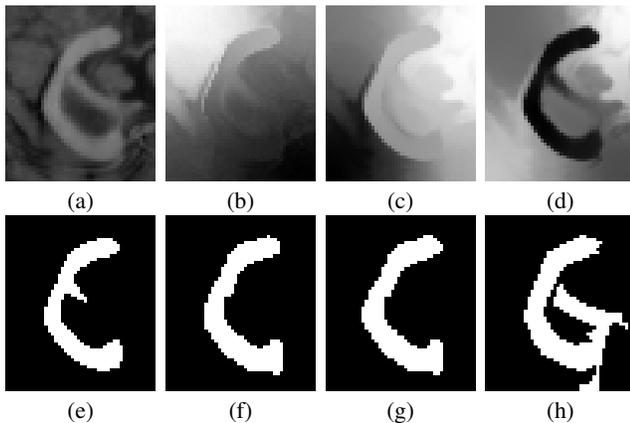


Fig. 5. Segmentation of corpora callosa. (a) Input image; (b)-(d) are 2^{nd} - 4^{th} smallest eigenvector; (e) Segmentation result from (d); (f) Projective shape model; (g) Final segmentation of our method; (h) Segmentation result by Ncut of Shi's version.

In the experiments on real data, our method outputs better results than the original normalized cut and our method can successfully separate corpus callosum from the extra component, which is connected to corpus callosum with similar intensity.

The computational efficiency of the normalized cut depends on the size of the affine matrix (The analysis can be found in [5]). According to our experiments, even if the initial normalized cut only

segment part of the boundary of the object, our method can converge after only three to five iterations. It means we need to decompose a sparse matrix five times at most.

4. CONCLUSION

This paper presents a new algorithm which uses the shape constraint to improve the performance of the normalized cut. Using the shape model, we can find the corresponding eigenvector to segment the target, even if the target is blurred or it is only a small part of the entire image which causes trouble to the conventional normalized cut method. Through projecting the coarse segmentation result onto PCA shape space, we can discard the unwanted parts and restore the missing parts of the object. With the modification of the affine matrix based on the shape constraint, we make the normalized cut more stable and robust. Our experiments have shown the robustness and accuracy of this extension.

5. REFERENCES

- [1] C.R. Brice and C.L. Fennema, "Scene analysis using regions," *AI*, vol. 1, no. 3-4, pp. 205-226, 1970.
- [2] M. Kass and A. Witkin *et al.*, "Snakes - active contour models," *IJCV*, vol. 1, no. 4, pp. 321-331, 1987.
- [3] S. Osher and J. Sethian, "Fronts propagation with curvature dependent speed: Algorithms based on hamilton-jacobi formulations," *J. of Comp. Phy.*, vol. 79, pp. 12-49, 1988.
- [4] F.R. Hansen and H. Elliott, "Image segmentation using simple markov field models," *CGIP*, vol. 20, no. 2, pp. 101-132, October 1982.
- [5] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. PAMI*, vol. 22, no. 8, pp. 888-905, 2000.
- [6] W.E. Donath and A.J. Hoffman, "Lower bounds for the partitioning of graphs," *IBM J. Research and Development*, pp. 420-425, 1973.
- [7] D. Tolliver and G.L. Miller *et al.*, "Corrected laplacians: Closer cuts and segmentation with shape priors," in *CVPR*, 2005, pp. II: 92-98.
- [8] S. Wang and T. Kubota *et al.*, "Shape correspondence through landmark sliding," in *CVPR*, 2004, pp. I: 143-150.
- [9] A. Tsai and A. J. Yezzi *et al.*, "A shape-based approach to the segmentation of medical imagery using level sets," *IEEE Trans. MI*, vol. 22, no. 2, pp. 137-154, 2003.
- [10] M.E. Leventon and W.E.L. Grimson *et al.*, "Statistical shape influence in geodesic active contours," in *CVPR*, 2000, pp. I: 316-323.
- [11] F.R.K. Chung, *Spectral graph theory*, American Mathematical Society, 1997.
- [12] J. Cheeger, "A lower bound for the smallest eigenvalue of the laplacian," in *Problems in Analysis*, 1970, pp. 195-199, Princeton University Press.
- [13] S. Guattery and G. L. Miller, "On the quality of spectral separators," *SIAM Journal on Matrix Analysis and Applications*, vol. 19, no. 3, pp. 701-719, 1998.
- [14] S. X. Yu and J. Shi, "Multiclass spectral clustering," in *ICCV*, 2003, pp. 313-319.
- [15] S. X. Yu and J. Shi, "Grouping with bias," in *NIPS*, 2001, pp. 1327-1334.
- [16] "<http://www.cis.upenn.edu/~jshi/software/>," .