

Improving Image Segmentation Quality Via Graph Theory

Xiangxiang Li, Songhao Zhu

*School of Automatic, Nanjing University of Post and Telecommunications,
Nanjing, 210046, China*

zhush@njupt.edu.cn

Abstract

Image segmentation is a fundamental process in many image, video, and computer vision applications. It is very essential and critical to image processing and pattern recognition, and determines the quality of final result of analysis and recognition. This paper presents a semi-supervised strategy to deal with the issue of image segmentation. Each image is first segmented coarsely, and represented as a graph model. Then, a semi-supervised algorithm is utilized to estimate the relevance between labeled nodes and unlabeled nodes to construct a relevance matrix. Finally, a normalized cut criterion is utilized to segment images into meaningful units. The experimental results conducted on Berkeley image databases and MSRC image databases demonstrate the effectiveness of the proposed strategy.

Keywords: Semi-Supervised, Graph Theory, Over-segmented

Introduction

As a significant branch of image processing and computer vision, the segmentation and extraction of target objects attracts researchers' attention. Moreover, image segmentation and target extraction has been widely used in the area of pattern recognition, computer vision, artificial intelligence, etc. Therefore, further research on this matter can not only provide the perfect solution to image segmentation and target extraction, but also give impetus to the development of pattern recognition, computer vision, artificial intelligence, etc. That is the initial motivation for us to study image segmentation.

Tradition image segmentation methods include Mean Shift Algorithm^[1], Normalized Cut Algorithm^[2] and K-Means Algorithm^[3]. As insensitive to smooth area and texture area, the mean shift algorithm not only has very good adaptability and robustness, but also achieves good segmentation results. With the aid of graph theory, the normalized cut algorithm segments each image into meaningful units. However, this algorithm is sensitive to parameters and involves large amount of computation. Due to the limitation of clustering, segmentation results obtained by the k-means algorithm are often not so satisfactory.

To achieve satisfactory results of image segmentation, this paper presents a novel image segmentation scheme based on the graph theory. The basic idea of the

proposed scheme is described as follows. To access semantic description of image features, each image is first segmented coarsely and represented as a graph model. Here, the nodes in the constructed graph are not from pixels but from blobs, and the edge information denote the relationship between blobs correspondingly. Then, a semi-supervised learning algorithm, a manifold algorithm is utilized to construct a relevance matrix estimate the relevance between labeled nodes and unlabeled nodes. Here, the constructed relevance matrix is a modified form of traditional relevance matrix. Finally, a normalized cut criterion is utilized to segment images into meaningful units in human perception.

Relevance Model

Implementing Image Preprocessing

With the increasing of image resolution, each area in a given image will become more detailed as more image pixels are utilized to represent it. That is, the number of image pixels in each perceptive area is becoming more and more, which makes it difficult to implement image processing efficiently and effectively based on image pixels. Therefore, in this paper, mean shift algorithm is utilized to deal with the issue of image preprocessing. Specifically, with appropriate parameters, each image is segmented into several regions coarsely.

Constructing Relevance Graph

It is known to all that the image segmentation quality using graph-theory techniques is largely based on the relevance between segmented regions. That is, it is crucial to construct an appropriate associated graph. Different from existing methods [4-6] constructing an associated graph based on the features in segmented regions, such as color information, edge information, or the combination of these two types of information, this paper proposes an improved feature extraction to achieve good segmentation results.

The values of neighboring pixels from the same region are generally similar, therefore each pixel is considered as a node in an associated graph where the weight between two nodes is formulated as:

$$W_{ij} = \exp(-\theta_p \|p_i - p_j\|^2 - \theta_v \|v_i - v_j\|) \quad (1)$$

where p_i and p_j are the positions of the i^{th} pixel and the j^{th} pixel respectively, v_i and v_j are the values of the i^{th} pixel and the j^{th} pixel respectively, and θ_p and θ_v are two constants controlling the weight of pixel position and pixel value respectively.

Different from the idea in literate [4], this paper constructs an associated graph where each pre-segmented region is considered as a node and weight between two nodes is formulated as:

$$W_{ij} = \exp(-\theta_r \|r_i - r_j\|^2 - \theta_m \|m_i - m_j\|)$$

(2)

where r_i and r_j are the positions of the i^{th} region and the j^{th} region respectively, m_i and m_j are the mean values of the i^{th} region and the j^{th} region respectively, and θ_r and θ_m are two constants controlling the weight of region position and mean value respectively.

It can be seen from formula (2) that the relevance between pixels within the same region is higher than the relevance between pixels from different regions. Therefore, the formula (2) describes the information of image edge.

Constructing Relevance Matrix

As an algorithm combining the characteristic of both supervised learning and unsupervised learning, semi-supervised learning technology has become a hot area of pattern recognition and machine learning research in recent years. It mainly deals with the following issue: learn to classify a mass of unlabeled samples into one of the pre-defined categories with a handful of labeled training samples of each class. There are two critical assumptions of semi-supervised learning technology^[7]: (1) Neighboring nodes are very likely to have the same label; (2) Nodes with the same structure are very likely to have the same label.

For unsupervised learning technology, the process of label propagation from labeled nodes to unlabeled nodes is generally performed in an iteration plan, which needs a large amount of computational cost. To reduce the computational cost and improve the propagation performance, the associated matrix between pre-segmented regions is constructed using the similar semi-supervised learning technology as described in literature [8].

Specifically, the weights between the k th node and the other nodes are formulated as follows:

$$\Gamma_k = \alpha [I - (1 - \alpha)P]^{-1} \Lambda_k \quad (3)$$

where α denotes the ratio of initial label information of each node and label information propagated from its neighboring nodes, I is an identity matrix. Λ is an $N \times 1$ vector which specifies the initial labels of each of the N nodes in the constructed relevance graph, where the value of each labeled node is 1 and the value of each unlabeled node is 0. P is a $N \times N$ matrix formulated as follows:

$$P = D^{-1}W \quad (4)$$

where W is a $N \times N$ affinity matrix as defined by formula (2), and D is a $N \times N$ diagonal matrix with its (i, i) -element equal to the sum of the i th row of W :

$$d_{ii} = \sum_{j=1}^n w_{ij} \quad (5)$$

The nature of the algorithm described above equals to the energy function of minimized images[9]:

$$Q = \sum_{i,j=1}^N W_{ij} \left| \frac{\Gamma_{ik}}{\sqrt{d_{ii}}} - \frac{\Gamma_{jk}}{\sqrt{d_{jj}}} \right|^2 - \mu \sum_{i=1}^N |\Gamma_{ik} - \Lambda_{ik}|^2 \quad (6)$$

and the above formula contains two constraints: (1) smooth constraint and equipped constraint, where the first constraint requires that labels of neighboring nodes are similar and the second constraint requires that the difference between initial labels of a node and propagated labels from neighboring nodes is as small as possible.

The local smooth between nodes can be estimated reasonably well using the above formula (3), the relevance matrix constructed in this paper is formulated as follows:

$$\Gamma = \alpha [I - (1 - \alpha)P]^{-1} \quad (7)$$

where the meanings of α , I and P are exactly the same as those in the above formula (3).

The Proposed Segmentation Algorithm

From the perspective of pattern recognition, the process of image segmentation can be considered as a label propagation process. Specifically, suppose there are K labels and each node in the constructed relevance graph is labeled with a label from these K labels. The segmentation algorithm used here is the algorithm proposed by J. Shi and J. Malik[2]:

$$S(Y) = \frac{1}{K} \sum_{k=1}^K \frac{y_k^T \Gamma y_k}{y_k^T D y_k} \quad (8)$$

where y_k is a $N \times 1$ segmentation vector with $y_{ik}=1$ if the i th region is contained in the k th region and $y_{ik}=0$ otherwise, and $Y=[y_1, y_2, \dots, y_K]$ is a segmentation matrix with $YY^T=I$. Moreover, the meanings of D and Γ are the exactly the same as that in formula (4) and formula (7) respectively.

The optimum solution of the above formula (8) is the subspace formed by the least K eigenvalues of the matrix:

$$D^{\frac{1}{2}} (I - (1 - \alpha)P) D^{\frac{1}{2}} \quad (9)$$

Experimental Results

The experimental results on Berkeley image database are illustrated in Table 1.

Table 1: Experimental comparisons on Berkeley image database.

Algorithm	Probabilistic Rand Index	Variation of Information	Global Consistency Error	Boundary Displacement Error
MShift ^[1]	0.7958	1.9725	0.1888	14.41
NCut ^[2]	0.7242	2.9061	0.2232	17.15
MNCut ^[10]	0.7559	2.4701	0.1925	15.10
Ours	0.8082	1.8759	0.1880	13.05

It can be seen from the above table that the proposed algorithm outperforms all the other algorithms, which confirms the algorithm from the following two aspects: (1) by considering the relationship among all the image data containing labeled and unlabeled images, the proposed algorithm can better approximate the relevance between the un-segmented image and each segmented image in the database than traditional algorithms; (2) by constructing an improved relevance matrix, the proposed algorithm can segment each image into perceptive regions more accurately than classical algorithms.

Conclusions

To improve the performance of image segmentation, this paper presents an image segmentation algorithm based on the graph theory. An image is first segmented roughly into some regions, and map these roughly segmented regions into a graph structure; then, a relevance matrix is constructed by semi-supervised learning; finally, roughly into some regions are further refined into perceptive regions by normalized segmentation rules. Experimental results conducted on Berkeley image database and MSRC image dataset demonstrate the most competitive performance of the proposed image theory segmentation method compared with traditional algorithms.

Acknowledgments

This work is supported by Postdoctoral Foundation of China under No. 2014M550297, Postdoctoral Foundation of Jiangsu under No. 1302087B, Education Reform Research and Practice Program of Jiangsu Province under No. JGZZ13_041, and Graduate Bilingual Teaching-Learning Pilot Program of Pattern Recognition and Intelligent Systems of Jiangsu Province.

References

- [1] D. Comaniciu and P. Meer. Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis And Machine Intelligence*, 24(5): 603-619, 2002.

- [2] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis And Machine Intelligence*, 22(8): 888-905, 2000.
- [3] S. Sakaida, Y. Shishikui, Y. Tanaka, and I. Yuyama. Image segmentation by region integration using initial dependence of the K-means algorithm. *Systems and Computers*, 29(14): 68-80, 1998.
- [4] Y. Boykov and G. Lea. Graph cuts and efficient n-d image segmentation. *International Journal on Computer Version*, 70(2): 109-131, 2006.
- [5] L. Mansilla and P. Miranda. Image Segmentation by Oriented Image Foresting Transform with Geodesic Star Convexity. *Proceeding of IEEE Conference on Computer Analysis of Images and Patterns*, 2013: 572-579.
- [6] M. Donoser and D. Schmalstieg. Discrete-Continuous Gradient Orientation Estimation for Faster Image Segmentation. *Proceeding of IEEE Conference on Computer Vision and Pattern Recognition*, 2014: 3158-3165.
- [7] O. Chapelle, J. Weston, and B. Scholkopf. Cluster kernels for semi-supervised learning. *Proceeding of IEEE Conference on Advances in Neural Information Processing Systems*, 2002: 585-592.
- [8] D. Zhou, O. Bousquet, T. Lal, J. Weston, and B. Scholkopf. Learning with local and global consistency. *Proceeding of IEEE Conference on Advances in Neural Information Processing Systems*, 2003: 1-8.
- [9] T. Kim and K. Lee. Learning full pairwise affinities for spectral segmentation. *Proceeding of IEEE Conference on Computer Vision and Pattern Recognition*, 2010: 2101-2108.
- [10] T. Cour, F. Benezit, and J. Shi. Spectral segmentation with multiscale graph decomposition. *Proceeding of IEEE Conference on Computer Vision and Pattern Recognition*, 2005: 1124-1131.