An Image Segmentation Method for Chinese Paintings by Combining Deformable Models with Graph Cuts

Ning He[1] and Ke Lu[2]

[1] School of Information, Beijing Union University,
Beijing, 100101, China
xxthening@buu.edu.cn
[2] College of Computing & Communication Engineering,
Graduate University of Chinese Academy of Sciences,
Beijing, 100049, China
luk@gucas.ac.cn

Abstract. In recent years researchers have developed many graph theory based algorithms for image setmentation. However, previous approaches usually require trimaps as input, or consume intolerably long time to get the final results, and most of them just consider the color information. In this paper we proposed a fast object extraction method. First it combines deformable models information with explicit edge information in a graph cuts optimization framework. we segment the input image roughly into two regions: foreground and background. After that, we estimate the opacity values for the pixels nearby the foreground/background border using belief propagation (BP). Third, we introduce the texture information by building TCP images' co-occurrence matrices. Experiments show that our method is efficient especially for TCP images.

Keywords: Graph cuts, Deformable Model, Traditional Chinese Painting (TCP), Belief Propagation (BP), Po-Occurrence Matrix.

1 Introduction

Segmentation of Traditional Chinese Painting (TCP) images is a good first step to separate them from generalized images. With the steady growth of computer power, rapidly declining cost of storage, and ever-increasing access to the Internet, digital acquisition of information has become increasingly popular in recent years[1]. Many organizations have a large digital images content available for on-line access. Various museums are constructing digital archives of art paintings and preserve the original artifacts. More and more artists attempt to exhibit and sell their productions on the Internet. Effective indexing, browsing and retrieving art images is an important and imperative problem need to be addressed, while image segmentation is a crucial step.

Deformable models and Graph cuts are two of the most important frameworks that have been used in the last decade to solve the image segmentation problem. Recent studies such as [2-4] are in the favor of combining deformable model approaches and

J.A. Jacko (Ed.): Human-Computer Interaction, Part I, HCII 2011, LNCS 6761, pp. 571–579, 2011.

graph cuts optimization to benefit from the advantages of deformable models in describing a wide variety of segmentation techniques and meanwhile take advantage of the fast global optimization associated with the use of graph cuts. The energy formulation of the model relaxes the global piecewise constant constraint in [3] and [5]. The energy will be optimized using graph cuts. In this paper, we introduces a new method for interactive segmentation that providing better boundary placement than deformable models segmentation and stronger region connectivity and less short-cutting than graph cuts methods. The experiment result shows that the technology is an effective method for the segmentation of TCP images.

2 The Features of TCP Images

Throughout its long history, TCP images have carried their own particular style. they look different from general image. Color and texture features are used here to characterisze the particularity of TCP images.

The color feature representation the corresponding percentage of occurrence of each color within a certain neighbourhood. Howe et al. have proposed in [6] a color-shape based method in which a quantized color image I' is obtained from the original image I by quantizing pixel colors in the original image. The authors tested on RGB, YUV, YIQ color space using SVM classifier. The result is in Table 1.

Autocorrelation [7] measures the coarseness of an image by evaluating the linear spatial relationships between texture primitives. Large primitives give rise to coarse texture and small primitives give rise to fine texture. If the primitives are large, the autocorrelation function decreases slowly with increasing distance whereas it decreases rapidly if texture consists of small primitives. Typically, TCP images have larger feature values compared to non-TCP images.

Table 1. Test result on color space

Classification rate on TCP test images			False classification rate on non-TCP test images		
RGB	YUV	YIQ	RGB	YUV	YIQ
0.902	0.942	0.940	0.117	0.09	0.076

3 The Proposed Method

Graph cuts algorithm[8,9] has established the relationship between the optimize of the energy function and the maximum flow/minimum cut algorithm in graph theory. In many cases can get the local optimal solution and in some cases get the global optimal solution. However, for a large amount of data processing, the graph cuts algorithm has the limitation in the memory requirements and computation velocity.

Different to reference [10], we consider each layer by accelerating the convergence of graph segmentation algorithm to reduce computation time. In this paper we use Gaussian Mixture Model (GMM), mean value and covariance matrix as the initial conditions, without having to use narrow-band. Although the number of nodes in every layer has not decreased, the parameters of images be obtained easily, this method can converge quickly and get the satisfied results.

Similar to reference [11], we use Gaussian Mixture Model (GMM) as the image color model. z is an array of image pixels, α and k are the non-transparent and the Gaussian component index arrays, respectively. Here α can only take 0 (background) or 1 (foreground). The parameters of each Gaussian component is defined as $\theta = \{\pi(\alpha,k), \mu(\alpha,k), \sum(\alpha,k)\}$, they are the weight of Gaussian component, mean vector and covariance matrix, respectively. The problem of image segmentation can be transformed into an deformable model optimization problem, and then use the graph guts method to solve. Here, the definition of the Gibbs energy function as:

$$\mathrm{E}(\alpha, \mathrm{k}) = \sum_{i} \mathrm{D}(\alpha_i, k) + \sum_{m,n} \mathrm{V}(\alpha_m, \alpha_n) \cdot \tag{1}$$

here, the optimal solution of α and k get the minimum of the energy function. Where $D(\alpha_{i,k}) = -\log P(\alpha_{i,k})$, $V(\alpha_m, \alpha_n) = [\alpha_m \neq \alpha_n] \exp(-\beta \|C_m - C_n\|^2)$, $[\alpha_m \neq \alpha_n]$ is the indicator function. The value is 0 when they are equal, otherwise is 1. $P(\alpha_{i,k})$ is Gaussian probability, $V(\alpha_m, \alpha_n)$ is smooth term.

4 Boundary Extraction

Belief propagation (BP) proposed by Pearl [12], are widely used in the encoding and decoding, stereo vision, image restoration and other fields. For acyclic graphs, belief propagation can get the global optimal solution, and for cyclic graph ,can only get the approximate solution. In practice, even for Markov Random Field (MRF) situations with the complex acyclic images can be obtained the optimal solution.

Consider using the MRF framework and the belief propagation algorithm to solving the boundary extraction problem. Each pixel in the image area R is as a MRF node. And the 4-connected neighboring pixels in R of boundary are included in the MRF. The aim of the image segmentation for the initial results in the foreground/background boundary width w to estimate the opacity. By minimizing the following defined energy function to get boundary of images:

$$E = E_{\mathrm{Data}}(\alpha_i) + E_{\mathrm{Smooth}}(\alpha_i, \alpha_j) + E_{\mathrm{Texture}}(\alpha_{\mathrm{i}}) \quad (i, j \in R) \tag{2}$$

The equation include in the data, smoothness and texture terms. Their significance are introduced as follow.

4.1 Data Term and Smoothness Term

The values $\alpha \in [0,1]$ is continuous, we use belief propagation algorithm to discrete it into K layers (e.g. $K = 10$), makes the each node in MRF can only be one of the K states.

To obtain estimates of the image pixel value α is equivalent to giving a label of each node.

The boundary of the foreground and background are smoothness, the adjacent pixels should have similar labels. If the value α of two nodes corresponding to the MRF in a clique has large differences, then the smooth item has greater penalty item, based on the above considerations, smooth item is defined as follows:

$$\text{Smooth}(\alpha_i, \alpha_j) = 1 - \exp(-\beta \frac{(\alpha_i - \alpha_j)^2}{2\delta_s^{\ 2}}) \tag{3}$$

where δ_ε is the experience value, i.e. it can be set to $K/2$, $\beta = \left(2\left\langle (x_p - x_q)^2 \right\rangle\right)^{-1}$ is parameter like a regularized factor .

Smoothness term defined reflects the degree of similarity between adjacent pixels, and data item of the model is the fitting item of foreground and background. If a pixel is assigned a label closer to the real value, obviously the data will correspond to smaller penalty value, so that the total energy function is reduced. We define the following data item:

$$\text{Data}\,(\alpha_i) = 1 - \frac{\mathrm{L_i(p)}}{\sum_{k=1}^{K} \mathrm{L_k(p)}} \tag{4}$$

To calculate the similarity $L_k(p)$ of pixels p and labels k. First, find N pixels, which is the smallest Euclidean distance with pixel p in foreground or background pixels were recorded as a collection of s_{obj} and s_{bkg}. Get pixel $p_i \in s_{\text{obj}}$, note RGB color vector C_i; get pixel $p_j \in s_{\text{bkg}}$, note the RGB color vector C_j. When mix C_i and C_j to α_k, the distance between pixel p and C is:

$$\left\| \alpha_k C_i + (1-\alpha_k) C_j - C \right\|^2 . \tag{5}$$

where $\|\cdot\|$ is the 2- norm. Thus, the similarity of pixel p and labels $L_k(p)$ can be defined as follows:

$$L_{\mathrm{k}}(p) = \text{Min} \exp(-\frac{D_{ij}(\alpha_k)}{\delta_d^{\ 2}}) \cdot \tag{6}$$

The value of δ_d set by the user manually, generally it sets 15.

4.2 Texture Term

Co-occurrence matrix is the description of texture information in image areas , it has the advantages such as computation simply, can reflect the texture direction and distance. Some images have rich texture information relatively, contributed to the pixel to find the optimal label.

Co-occurrence matrix determined by the offset distance and offset angle. For convenience, we construct the 8-neighborhood of four co-occurrence matrix, corresponding to the 0°, 45°, 90°, 135° four cases, shown in Fig.1.

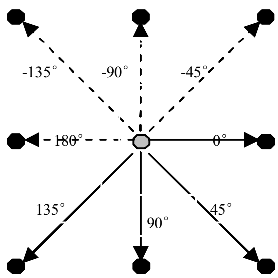

Fig. 1. Symbiotic matrix

For the pixels p of 8-connected must be considered the adjacent pair-pixels in 8 directions $\{p, p_i\}(i=1...8)$. Let p and p_i are quantized a and b, corresponding to the direction of the foreground and background of the co-occurrence matrix were M_{obj} and M_{bkg}. If the angle formed by p and p_i is one of the solid arrow lines in Fig.1, the corresponding textures penalty value of foreground and background can be defined as follows, respectively:

$$\begin{aligned} \text{Penalty}_{\text{obj}} &= \frac{\text{M}_{\text{bkg}}(a,b)}{\text{M}_{\text{bkg}}(a,b)+\text{M}_{\text{obj}}(a,b)} \\ \text{Penalty}_{\text{bkg}} &= \frac{\text{M}_{\text{obj}}(a,b)}{\text{M}_{\text{bkg}}(a,b)+\text{M}_{\text{obj}}(a,b)}. \end{aligned} \tag{7}$$

If the angle between p and p_i shown in Fig.2 is the dotted arrow line, obtained the value by symmetry. When the labels of p and p_i is α_i and α_j, respectively. To introduce the increment item in energy function.

$$\Delta\text{E} = (\alpha_i + \alpha_j) * \text{Penalty}_{\text{obj}} + [(1-\alpha_i) + (1-\alpha_j)] * \text{Penalty}_{\text{bkg}} \tag{8}$$

As MRF is 4-connected, co-occurrence matrix calculated using the eight different directions. In order to resolve inconsistencies between them, consider the following skills:

$$\begin{aligned}\Delta\mathrm{E} = \Delta\mathrm{E}(\alpha_{\mathrm{i}}) + \Delta\mathrm{E}(\alpha_{\mathrm{j}}) &= [\alpha_i * \mathrm{Penalty}_{\mathrm{obj}} + (1-\alpha_i) * \mathrm{Penalty}_{\mathrm{bkg}}] + \\ &[\alpha_j * \mathrm{Penalty}_{\mathrm{obj}} + (1-\alpha_j) * \mathrm{Penalty}_{\mathrm{bkg}}]\end{aligned} \tag{9}$$

Therefore, $\Delta\mathrm{E}(\alpha_{\mathrm{i}})$ and $\Delta\mathrm{E}(\alpha_{\mathrm{j}})$ same as the data item, only consider the current pixel label information without considering the 8-connected neighboring pixel labels. so that the co-occurrence matrix texture information can be propagated in the 4 -connected MRF. In summary, texture items in energy function can be defined as follows:

$$\mathrm{Texture}(\alpha_i) = \sum_{k=1}^{8} [\alpha_i * \mathrm{Penalty}_{\mathrm{obj}}^{\mathrm{k}} + (1-\alpha_i) * \mathrm{Penalty}_{\mathrm{bkg}}^{\mathrm{k}}] \cdot \tag{10}$$

Where k is the eight different directions.

5 Experimental Results

5.1 Running Time Analysis

We selected four representative TCP images from Internet. We implement the max flow-min cut to solve graph cuts method. Without using sophisticated data structures. The algorithm achieves the running time of $O(nm + n^2 \log U)$, where n is the number of nodes, m is the number of edges, and U is the largest edge weight. However, the simple topology of the graph used in this work makes the algorithm run much faster in practice. By obtaining the best polynomial fit to observed data, the minimum cut algorithm used in this work has running time of $O(n^{1.2})$. In Fig.2, the running time is shown as a function of the number n of nodes.

5.2 Comparison of Segmentation Accuracy

Fig.3 shows a sample comparison of performing the segmentation of graph cuts and the proposed method. Our approach clearly extracts the objects, while the graph cuts method failed to correctly identify some of the foreground pixels. Some background pixels wrongly identified as objects. The algorithm was implemented in Matlab 7.1 on a 2GHZ CPU with 1GB RAM. For quantitative assessment, we compared our segmentation results with the graph cuts by calculating the error rate. The results show in Table.2. Our method achieves better result. Fig.4 is in the same experimental conditions, we segmented the flower and birds respectively. The results are robust.

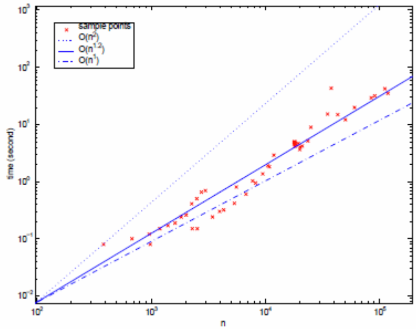

Fig. 2. The observed running time of the minimum cut algorithm used in our approach is $O(n^{1.2})$. The performance levels of $O(n)$ and $O(n^2)$ as shown, respectively.

(a) Original images (b) Results of our method (c) Results of graph cuts

Fig. 3. Compare results of proposed method and graph cuts method

(a) Original images (b) object is the flower (c) object is the birds

Fig. 4. The different segmentation results according to the different foreground

Table 2. Error Rate Comparation

Images	Proposed Method	Graph Cut Method
Person1	0.56%	1.82%
Tiger	0.35%	3.56%
Flower1	0.89%	1.29%

6 Conclusion

We have presented a graph cuts based deformation model approach to TCP images segmentation. First, we transform a multi-labels cut problem into a singe s-t minimum cut problem. Then, the method integrate the color and texture features of TCP images to extract the boundary. In the actual operation received good results.

Acknowledgments. This work was supported by the NSFC (China) Grant No.s 61070120, 60982145; Beijing national science fundation Grant No.4112021; Beijing educational commission science fundation; the opening project of Shanghai key laboratory of integrate administration technologies for nformation security, No.AGK2010005; the opening project of key laboratory of communication and information system (Beijing jiaotong university).

References

1. Jiang, S., Gao, W., Wang, W.: Classifying Traditional Chinese Painting Images. ICICS-PCM 2003, Singapore, December 15-18 (2003)
2. Criminisi, A., Sharp, T., Rother, C., Perez, P.: Geodesic Image And Video Editing. ACM Transactions on Graphics (2010)
3. Ding, L., Yilmaz, A.: Interactive Image Segmentation Using Probabilistic Hypergraphs. Pattern Recognition 43(5), 1863–11873 (2010)
4. Xu, N., Bansal, R., Ahuja, N.: Object Segmentation Using Graph Cuts Based Active Contours. In: CVPR, Los Alamitos, CA, USA, vol. 2, pp. 46–53 (2003)

5. Sun, J., Li, Y., Kang, S.B., Shum, H.-Y.: Flash Matting. ACM Trans. Graph 25(3), 772–778 (2006)
6. Liu, P., Jia, K., Wang, Z., Lv, Z.: A New and Effective Image Retrieval Method Based on Combined Features. In: Proc. IEEE Int. Conf. on Image and Graphics, pp. 786–790 (2007)
7. Oraintara, S., Nguyen, T.T.: Using Phase and Magnitude Information of the Complex directional Filter Bank for Texture Image Retrieval. In: Proc. IEEE Int. Conf. on Image Processing, vol. 4, pp. 61–64 (2007)
8. Boykov, Y., Jolly, M.P.: Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D Images. In: International Conference on Computer Vision (ICCV), vol. I, pp. 105–112 (2001)
9. Kolmogorov, V., Zabih, R.: What Energy Functions Can Be Minimizedvia Graph Cuts? IEEE TPAMI 26, 147–159 (2004)
10. Lombaert, H., Sun, Y., Grady, L., Xu, C.: A Multilevel Banded Graph Cuts Method for Fast Image Segmentation. In: Proceedings of International Conference on Computer Vision, Beijing, China, pp. 1259–265 (2005)
11. McGuinness, K., O'Connor, N.E.: A comparative evaluation of interactive segmentation algorithms. Pattern Recognition 43, 434–444 (2010)
12. Felzenszwalb, P.F., Huttenlocher, D.R.: Efficient Belief Propagation for Early Vision. In: CVPR, IEEE Int. Conf. on Computer society, vol. 1, pp. 261–268 (2004)