IMPROVED SPECTRAL MATTING BY ITERATIVE K-MEANS CLUSTERING AND THE MODULARITY MEASURE

Tung-Yu Wu, Hung-Hui Juan and Henry Horng-Shing Lu

Institute of Statistics, National Chiao Tung University, Hsinchu 30010, Taiwan

ABSTRACT

Spectral matting is a useful technique for image matting problem. A crucial issue of spectral matting is to determine the number of matting components which has large impacts on the matting performance. In this paper, we propose an improved framework based on spectral matting in order to solve this limitation. Iterative K-means clustering with the assistance of the modularity measure is adopted to obtain the hard segmentation that can be used as the initial guess of soft matting components. The number of matting components can be determined automatically because the improved framework will search possible image components by iteratively dividing image subgraphs.

Index Terms**—** Image matting, Spectral matting, Modularity

1. INTRODUCTION

Image matting is a technique which estimates the foreground alpha matte and extracts the foreground object from an image. In image matting, the image color of a pixel is assumed to be a linear combination of the corresponding foreground and background colors, according to the so-called compositing equation (1).

$$I_i = \alpha_i F_i + (1-\alpha_i) B_i \tag{1}$$

In (1), I_i denotes the color of the i-th pixel, F_i and B_i denote the foreground and background colors respectively, and α_i is the foreground alpha matte at the i-th pixel. In the image matting problem, input image I is given, foreground alpha matte α jointly with the foreground image F and background image B need to be solved at each pixel. Since the unknown variables are much more than equations, it is an under-constrained problem and it is challenging to extract a precise α with or without limited user input.

Spectral matting proposed by Levin et al. [1-2] is one of the popular image matting techniques [3-5]. It can extract mattes not only two layers but also multiple layers [6]. They use the eigenvectors of matting Laplacian to obtain matting components. To determine the proper number of matting components is a major challenge of spectral matting [2]. This fundamental difficulty also exists in all spectral segmentation methods. To overcome this challenge, we propose an iterative dividing process to find enough matting components automatically. Given an initial number of matting components, it will automatically find the matting components. A stopping criterion, modularity measure [7], is adopted in this paper to terminate the dividing process.

The rest of the paper is organized as follows. Spectral matting is briefly introduced in section 2. In section 3 we re-formulate the problem and propose the improved framework to solve the major limitation of spectral matting. Results of visual and quantitative comparisons are provided in section 4. Finally, conclusions are given in section 5.

2. OVERVIEW OF SPECTRAL MATTING

Spectral matting, which combines spectral segmentation techniques with soft image matting, is a closed form solution to the image matting problem. A cost function is derived and the image matting problem is linked to an optimization problem. Compositing equation (1) can be rewritten and approximated by (2) in a small window w. By using (2) as the approximation of alpha matte, a cost function is defined by (3).

$$\alpha_i \cong \sum_c a^c I_i^c + b, \quad \forall i \in w \tag{2}$$

$$J(\alpha, a, b) = \sum_{j \in I} \left(\sum_{i \in w_j} \left(\alpha_i - a_j I_i - b_j\right)^2 + \varepsilon a_j^2 \right) \tag{3}$$

Now the image matting problem is transformed to an optimization problem. To minimize the cost function (3), they minimize $J(\alpha,a,b)$ by choosing the best a and b first. As a result, the objection function becomes a quadratic form in α, as shown in (4), where L matrix is called matting Laplacian, whose (i,j)-th element is defined by (5), where μ_k is a 3x1 mean vector and Σ_k is a 3x3 covariance matrix of pixel's color information I_i in a window w_k with the window size $|w_k|$, ε is a parameter used for the regularization term in (3).

$$J(\alpha) = \alpha^T L \alpha \tag{4}$$

$$\sum_{k \left| \substack{i \in w_k \\ j \in w_k} \right.} \delta_{ij} - \frac{1}{|w_k|} \left(1 + (I_i - \mu_k)^T \left(\Sigma_k + \frac{\varepsilon}{|w_k|} I_{3X3} \right)^{-1} (I_j - \mu_k) \right) \tag{5}$$

Now the optimization problem is to minimize the cost function defined by (6) (for more user input information as supervised matting, there are more constraint equations).

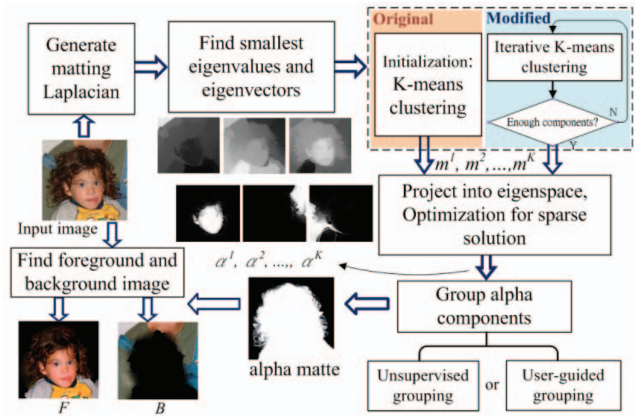


Fig. 1. Flowchart of the spectral matting framework [1-2] and the improved framework, the example image is from [1].

$$\min_{\alpha} J(\alpha) = \min_{\alpha} \alpha^T L \alpha \qquad \text{s.t.} \quad 0 \le \alpha_i \le 1 \ \forall i \tag{6}$$

It is still not easy to directly solve the optimization problem (6). Consequently, the spectral method, which analyzing the smallest eigenvectors of matting Laplacian, is adopted to segment the input image into multiple components first and then the hard segmentation is used to produce soft matting components. Finally, the grouping mechanism which can be unsupervised or supervised is used to obtain the foreground matte. Correspondingly, the image compositing equation (1) is extended to K-layer compositing (7) and a layer image operated by its corresponding alpha component, α^k, can be regarded as a soft image component in the input image.

$$I_i = \sum_{k=1}^{K} \alpha_i^k L_i^k \qquad \text{s.t.} \sum_{k=1}^{K} \alpha_i^k = 1 \ \forall i \in I \tag{7}$$

As a result, Fig. 1 summarizes the flowchart of the spectral matting framework presented in [1-2].

3. IMPROVED FRAMEWORK FOR SPECTRAL MATTING

The optimization of minimizing the cost function (4) can be approximated as a vector partitioning problem which try to segment vectors formed from the smallest eigenvectors of the matting Laplacian L.

Cluster indication vectors m^1, m^2, …, m^K, which are the output of the dash-line block in Fig. 1, are defined by (8), where n_k is the number of pixels in the k-th component.

$$m_i^k = \begin{cases} 1 & \text{if } i\text{-}th \text{ pixel} \in \text{the } k\text{-}th \text{ component } G_k \\ 0 & \text{otherwise} \end{cases} \qquad {m^k}^T m^k = n_k \tag{8}$$

If we consider the K-layer matting, the costs of matting components belonging to the foreground at the output of hard segmentation can be written as following

$$J = \sum_{k \in F} {m^k}^T L m^k = \sum_{k \in F} \left({m^k}^T U (\Lambda - \beta I) U^T m^k + \beta {m^k}^T m^k \right) \tag{9}$$

where $U \Lambda U^T$ is the eigendecomposition of L with eigenvalues $\lambda_1 \leqq \lambda_2 \leqq \ldots \leqq \lambda_n$, and β is a threshold of eigenvalues. We can represent (9) as

$$J = \sum_{k \in F} {m^k}^T U (\Lambda - \beta I) U^T m^k + \beta \sum_{k \in F} n_k \tag{10}$$

To minimize (10) is equivalent to (11) as follows,

$$\begin{aligned} \max_{m^1 \ldots m^K} -J &= \max_{m^1 \ldots m^K} \sum_{k \in F} {m^k}^T U (\beta I - \Lambda) U^T m^k - \beta \sum_{k \in F} n_k \\ &= \max_{m^1 \ldots m^K} \sum_{k \in F} \sum_{j=1}^{n} (\beta - \lambda_j) \left(\sum_{i=1}^{n} U_{ij} m_i^k \right)^2 - \beta n_F \\ &\cong \sum_{k \in F} \sum_{j=1}^{p} \left(\sum_{i \in G_k} \sqrt{\beta - \lambda_j}\, U_{ij} \right)^2 - \beta n_F \end{aligned} \tag{11}$$

where n_F is the number of pixels in the foreground. Let us define the p-dimensional vertex vector v_i and cluster vector V_k as (12), then (11) becomes (13).

$$(v_i)_j = \sqrt{\beta - \lambda_j}\, U_{ij}, \quad V_k = \sum_{i \in G_k} v_i \tag{12}$$

$$\max_{m^1 \ldots m^K} -J \cong \max_{m^1 \ldots m^K} \sum_{k \in F} \|V_k\|^2 - \beta n_F \tag{13}$$

If we neglect the second term in (13), the optimization problem looks like the vector partition problem which tries to find a division of vectors such that the sum of square norm of cluster vectors is maximal. However, if we consider the cost of not only the foreground matte but also the background matte, (13) should be modified to (14), where n is the number of pixels in the image. Since n is a constant, now the optimization problem is equivalent to the vector partitioning problem.

$$\max_{m^1 \ldots m^K} -J^{All} \cong \max_{m^1 \ldots m^K} \sum_{k=1}^{K} \|V_k\|^2 - \beta n \tag{14}$$

To solve the vector partitioning problem in a feasible way, K-means clustering can be regarded as a heuristic algorithm. When K-means clustering is finished, each image pixel is classified to an image segment correspondingly.

To obtain an appropriate hard segmentation of an image can be regarded as a graph partitioning problem where the matting Laplacian L describes the graph structure. To find possible matting components, the number of matting components is a critical parameter. Hence, this study aims to find a proper graph partition pattern from the test image automatically. Modularity measure [7-8] is a powerful measure used to quantify the fitness of a division pattern for a given graph. It is defined by the number of edges within subgraphs minus the expected number in an equivalent graph with edges placed at random. Suppose that there are K clusters in a graph of n vertices. The modularity of the graph, Q, is defined as

$$Q = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{K}\left(A_{ij} - \frac{d_i d_j}{2E}\right)m_i^k m_j^k \quad (15)$$

where A_{ij} is the number of edges between vertices i and j; E is the total edges in the graph; d_i and d_j are the degrees of the vertices i and j respectively. Since L describes the graph structure, A_{ij} and d_i are defined.

As mentioned, the major limitation of spectral matting is that the number of matting components is a critical parameter which influences the matting performance but lacks of an algorithm to understand the appropriate setting. When the number of matting components is smaller than the true image components, it may result in bad matting components which consist of dissimilar image objects. To solve this problem, we do not apply only once K-means clustering to the smallest eigenvectors of matting Laplacian. Instead, we apply iterative K-means clustering for the purpose of further dividing a cluster if it contains dissimilar image objects. Correspondingly, the graph is divided into subgraphs and every subgraph may be divided again. To stop the dividing process, the modularity (15) is adopted to judge whether the dividing pattern can increase the modularity of a subgraph or not. Besides, with the modularity we can search a range of small numbers to find the best number of clusters to divide the subgraph. Since the modularity is used to find the community structure in networks, it is similar to find image components in an image. Consequently, the dividing process is listed in the following steps.

1) Pick up p smallest eigenvectors of L to form the data matrix, which represents n data in p-dimensional vector space. Apply K-means clustering using number of clusters C_0.
2) For each cluster of 1), form the data matrix, which represents n_k data in p-dimensional vector space. Apply K-means clustering using number of clusters 2, 3,…, C_1, and calculate the corresponding modularity using (15).
3) If the maximum total modularity of sub-clusters divided from the cluster is more than the original modularity of the cluster, then execute the dividing pattern of K-means result. Otherwise, the cluster will not be further divided any more. When a cluster cannot be further divided, it enters the "*indivisible state*".
4) For each cluster which does not enter the indivisible state, recursively perform 2) to 3) until all clusters enter the indivisible state.
5) After 4), cluster indication vectors m^1, m^2, …, m^K, are used to generate the soft results of matting components.

As a result, the iterative K-means clustering divides the image into multiple image components automatically. Compared with the original method (only once K-means), it can find more image components potentially with the assistance of the modularity measure. The flowchart of the improved spectral matting is also shown in Fig. 1. The first advantage of the improved spectral matting is that the parameter of the number of clusters is not critical now because the dividing process is automatically performed to find appropriate image components. The second advantage of the improved spectral matting is that finding sufficient image components is critical to obtain a pure foreground alpha matte in some cases. As shown in Fig. 1, matting components are grouped to form the foreground alpha matte. If one matting component contains one region belonging to the foreground and the other region belonging to the background, the foreground matte may be imperfect. To avoid this case, it is necessary to find sufficient image components.

4. RESULTS AND DISCUSSIONS

Fig. 2 shows the increment of modularity measure in our iterative dividing approach, and a reference value of the modularity in original approach is also presented with 40 matting components. The modularity of our iterative dividing approach increases and surpasses the reference value. Iterative dividing process stops when there is no more increment of modularity. The number of final matting components is also automatically determined.

The hard decision results of the related example are shown in Fig. 3. There are still some components which can be extracted in the original method with fixed matting components number, and our iterative approach can divide them further. If some image components are not separated in the hard decision, it leads that some matting components which should be extracted are mixed. The performance of spectral matting is affected without enough components. Note that our iterative dividing approach may result in more components in a visible object. However, they can be further combined by the grouping mechanism show in Fig. 1.

There is an example where the original approach fails to split possible image components in Fig. 4. The iterative dividing approach can find more possible components than the original approach can because the better hard decision is used to produce the soft matting components.

Although a lager initial number of matting components can be set in original spectral matting algorithm, it is also hard to determine how many components is enough and to measure whether the segmentation of matting components is proper or not. Our approach can find components iteratively and find whether there are still possible components in an existing component using modularity measure. Table 1 lists the modularity of the original approach in spectral matting with different initial number of matting components and that of our iterative dividing approach. Note that the modularity of the original approach may not increase when the number of matting components increases because it may not find an appropriate dividing pattern of the image.

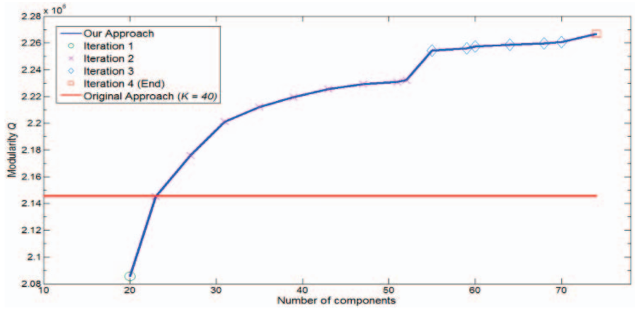


Fig. 2. Increment of the modularity in the proposed iterative dividing graph approach. The test image is *Peppers* (from [9]).

(a) (b) (c)

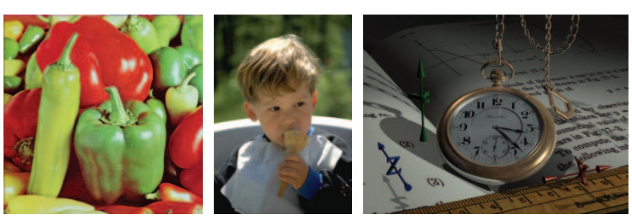
Fig. 3. (a) The input image is *Peppers*. (b) Hard segmentation result of the original approach. (c) Hard segmentation result of our iterative dividing approach.

(a) (b) (c)

Fig. 4. Image components of *Peppers* that present the benefit of the iterative dividing approach. (a) One component of the original approach. (b) and (c) The corresponding components of our iterative dividing approach after appropriate grouping.

Table 1. The modularity of each method. K denotes for the number of matting components. The test images are from [9].

Method / Image	Original approach			Our approach with C_0=20
	K=20	K=40	K=80	Q (Final K)
*Peppers 512*512*	2.1292E+06	2.1456E+06	2.1452E+06	2.26689E+06 (74)
*Kid 703*487*	2.8301E+06	2.8252E+06	2.9416E+06	2.97863E+06 (74)
*Watch 640*480*	2.4509E+06	2.4448E+06	2.5476E+06	2.626182E+06 (57)



(a) (b) (c)

Fig. 5. The test images in Table 1. (a) *Peppers*. (b) *Kid*. (c) *Watch*. All test images are from [9].

5. CONCLUSIONS

Spectral matting combines the spectral segmentation with image matting that leads to a useful solution of image matting in image and video editing. In this paper, we propose an improved framework to solve the technique limitation of spectral matting, that is, to determine the appropriate number of matting components. Iterative K-means clustering with the assistance of the modularity measure can divide the input image to sufficient image components. As a result, an appropriate hard segmentation result can be generated for obtaining appropriate soft results of matting components. The number of matting components can be determined automatically because the improved framework will search sufficient image components by iteratively dividing image subgraphs. Furthermore, it is possible to prevent the situation that a matting component consists of both foreground and background objects. Hence, the proposed framework can improve the performance of spectral matting for a wide spectrum of applications.

6. ACKNOWLEDGMENTS

The authors would like to acknowledge the supports from National Science Council, Thematic Research Program in Academia Sinica, National Center for Theoretical Sciences, Center of Mathematical Modeling and Scientific Computing at National Chiao Tung University in Taiwan. The authors would also like to acknowledge the National Center for High-performance Computing at Taiwan in providing huge computing resources to facilitate this research.

7. REFERENCES

[1] A. Levin D. Lischinski and Y. Weiss, "A Closed Form Solution to Natural Image Matting," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Feb 2008.

[2] A. Levin, A. Rav-Acha, D. Lischinski, "Spectral Matting," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Oct 2008.

[3] L. Grady, T. Schiwietz, S. Aharon, and R. Westermann, "Random Walks for Interactive Alpha-Matting," *Proc. Int'l Conf. Visualization, Imaging and Image Processing*, Sept 2005.

[4] J. Wang and M. Cohen, "Simultaneous Matting and Compositing," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2007.

[5] C. Rhemann, C. Rother, A. Rav-Acha, and T. Sharp, "High Resolution Matting via Interactive Trimap Segmentation," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2008.

[6] D. Singaraju and R. Vidal, "Estimation of Alpha Mattes for Multiple Image Layers," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Jul 2011.

[7] M. E. J. Newman, "Modularity and community structure in networks," *Proc. Natl. Acad. Sci. USA (PNAS)*, June, 2006.

[8] M.E.J. Newman, "Finding community structure in networks using the eigenvectors of matrices," *Phys. Rev. E*, 2006.

[9] Image database available at the web site of http://www.petitcolas.net/fabien/watermarking/image_database/