Gradient Color Tensor based Approach for Spectral Matting

Adam Ghorbel, Marwen Nouri and Emmanuel Marilly
Alcatel-Lucent Bell Labs France, Multimedia Technologies Domain, Centre de Villarceaux, Route de Villejust, 91620, Nozay, France

Keywords: Image Matting, Spectral Matting, Gradient Color Tensor, Matting Laplacian, Affinity Matrix, Alpha Matte.

Abstract: Image matting aims to extract foreground objects from a given image in a fuzzy mode. One of the major state-of-the-art methods in this field is spectral matting. It automatically computes fuzzy matting components by using the smallest eigenvectors of a defined Laplacian matrix that is generated from affinities computation between adjacent pixels in an image. Results obtained by such approach are coarsely related to the ability of defining an affinity matrix that it should be able to well separate between different pixels' clusters. To accomplish better matting and get better results, we propose a new spectral matting approach. We use a color tensor gradient of color images in order to enhance the affinity computation process.

1 INTRODUCTION

Since it was first mathematically established by Porter and Duff (Porter, 1984), image matting has been a high-value key in image editing, film production and interactive entertainment applications. This technique is used to extract the foreground defined by a specific object from an arbitrary scene which is essential for image composition tasks.

Starting from the assumption that an image is a combination of two distinct layers which are called Foreground (F) and background (B), image matting allows the extraction of the foreground layer out of the background layer of an input image. This process is done by estimating the opacity value at each pixel, typically referred as alpha value, in order to generate a matting image, typically known as alpha matte, with alpha is a real value varying between 0 and 1. Hence, image matting can be seen as a generalization of binary segmentation.

The image matting problem is an ill posed problem. In the case of a color image, we have to estimate seven unknowns from the three color component measures. Frequently, much additional information can be provided by the user whether in the form of Trimap which divide the image into three regions (a define foreground, a define background and unknown region that is considered as a mixture of foreground and background colors) or in the form of Scribbles such brush strokes to solve the image matting process.

In this paper, we propose an enhancement of spectral algorithm (Levin, 2008) based on color tensor gradient constraints to compute pixels' affinities.

2 LITTERATURE SURVEY

Usually, image matting can be classified into 2 types: Semi-Automatic approaches and Automatic approaches.

2.1 Semi-automatic Image Matting

Many semi-automatic image matting approaches have been reviewed in (Wang, 2007). These approaches can be classified into two major categories which are: sampling based methods and affinity based methods.

2.1.1 Sampling based Image Matting

Sampling based methods are based on two assumptions. First, adjacent pixels which have similar colors in a given image have often local correlation. Second, the foreground and background color of the unknown pixel are estimated by taking into account the foreground and background samples

Ghorbel A., Nouri M. and Marilly E..
Gradient Color Tensor based Approach for Spectral Matting.
DOI: 10.5220/0004215404260430
In *Proceedings of the International Conference on Computer Vision Theory and Applications* (VISAPP-2013), pages 426-430
ISBN: 978-989-8565-47-1

which are the nearby known pixels specified by the user. In 2000, Ruzon and Tomasi (Ruzon, 2000) have modelled the foreground and background as a mixture of un-oriented Gaussians. Furthermore, a Bayesian matting method has been proposed by Y Chang et al. (Chuang, 2001). In this framework, the foreground and background distributions are modelled using a mixture of oriented Gaussians and a maximum-likelihood criterion is used to estimate the final matte.

2.1.2 Affinity based Image Matting

In contrast of sampling-based methods, affinity based matting methods solve alpha values by defining various affinities between nearby pixels. The Poisson matting method proposed by (Sun, 2004) solves the matte from its relative gradient field estimated from a given image. This work is based on the fact that an image can be modified by treating the gradient interactively or automatically and that the intensity changes in both the foreground and background are locally smooth. In 2005, Gragy (Grady, 2005) proposes the Random walks method which computes the alpha values based on the affinity between neighboring pixels. This affinity is calculated with the measurement of the color distances in the RGB channels by using a Local preserving projection (LPP). In 2007, Bai and Sapiro (Bai, 2007) have used weighted geodesic distances to estimate alpha values with a random walker technique. A recent affinity based method which offers both trimap and scribble based matting is Closed-Form matting has been proposed by Levin (Levin, 2006). To overcome the over smoothing problem in the closed form matting, Zhu (Zhu, 2010) has introduced a new matting cost function by including a gradient constraint into the cost function and incorporating pixels from some special windows.

However, semi-automatic matting methods are a time and memory consuming, so it is interesting to solve the alpha matte automatically without any user input.

2.2 Automatic Image Matting

The first automatic image matting method, known as Spectral Matting, has been proposed by Levin (Levin, 2008). This approach combines both spectral segmentation methods (Yu, 2003) with soft image matting. Based on analyzing the smallest eigenvectors of a defined Matting Laplacian matrix (Levin, 2006), a set of fuzzy matting components are automatically extracted. These components which are obtained by a linear transformation of the eigenvectors are then combined to form the final alpha matte based on minimizing a matte cost function. In this framework, a new mode of user control over the generated matte is proposed. The user can directly control the final result by presenting to him a various matting components to choose from.

However, the spectral matting has some limitations and does not generate a high quality final alpha matte. Recently, modified spectral matting methods have been proposed to overcome some limitations and to obtain automatic and accurate alpha matte. Hu et al. (Hu, 2010) have introduced a spectral matting method based on the palette-based component classification. This type of classification allows classifying components as foreground regions, background regions or unknown regions.

In 2012, Hu et al. (Hu, 2012) have proposed a spectral matting based on components-Hue-Difference. Moreover, an improved spectral matting method by iterative K-means clustering and the modularity measure is proposed recently by Yu (Yu, 2012).

3 PROPOSED APPROACH

Spectral matting (Levin, 2008) is based on some hypothesis which are sometimes not true in the case of natural images. Below, we identify some of these violations that lead to distort the quality of the final alpha matte:

- Violation of the color-line model: in the case of complex intensity variation, the image data does not satisfy the color line model which is defined in next sub-section.
- In real images, the affinity matrix that is used to define the matting Laplacian is rarely able to perfectly separate different pixel clusters or components.

The proposed enhanced spectral matting solution based on the improvement of the matting Laplacian computation will be described below.

3.1 Improved Spectral Matting

Spectral matting relies on the evaluation of an affinity function between each pair of adjacent image pixels within small windows. Hence, pixels are implicitly described by their similarity to every other pixel in the input image. This similarity measurement is done by considering the covariance of the data and their intensities. Indeed, the affinity matrix in spectral matting is the following:

$$A_q(i,j) = \begin{cases} \delta_{ij} - \frac{1}{|w_q|}\left(1 + (I_i - \mu_q)^T \left(\sum_q + \frac{\varepsilon}{|w_q|} I_{3\times 3}\right)^{-1} (I_j - \mu_q)\right) \\ 0, otherwise \end{cases}, (i,j) \in w_q \quad (1)$$

Accordingly, the affinity between two pixels of the same color increases while the affinity between pixels of different color is zero. In other way, nearby pixels with similar colors have high affinity while nearby pixels with different colors have small affinity. In consequence, this affinity measurement captures the fact that an image is composed by connected components or different clusters.

However, in the case of natural image with complex intensity variations, the affinity matrix fails sometimes to separate between different clusters of the image data.

To overcome this limitation, we have incorporated gradient color tensor information in the affinity computation process. Based on the color tensor, we estimate the gradient which is a measure of the rate of change of the image intensity between neighbouring pixels.

The flowchart of the proposed solution is presented in Figure1:

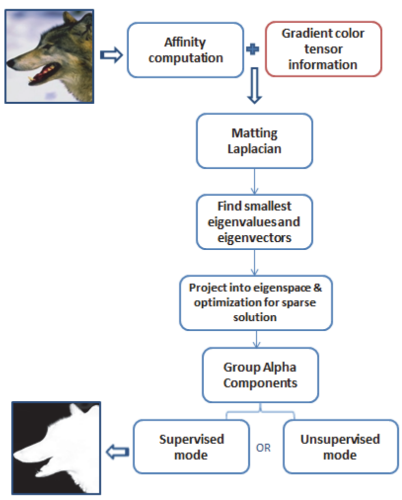

Figure 1: The flowchart of the proposed method.

Based on the assumption that a RGB image is a vector valued function defined over a manifold of the x and y plane, hence, mathematically, its gradient can be presented by a tensor (Di Zenzo, 1986).

The gradient vector for the space coordinate x (respectively for y) can be defined as following:

$$\nabla x = \left(\frac{\partial R}{\partial x}, \frac{\partial G}{\partial x}, \frac{\partial B}{\partial x}\right)$$
$$\nabla y = \left(\frac{\partial R}{\partial y}, \frac{\partial G}{\partial y}, \frac{\partial B}{\partial y}\right) \quad (2)$$

Then, the gradient is evaluated as the vector sum of the square gradient of the individual components or separate bands of the input RGB image. Regardless that an RGB image is naturally represented by a 3-order tensor, typically the 1-mode is the height, the 2-mode is the width and the 3-mode is the color space, hence, we can define the color tensor coefficients as:

$$f_x = \left|\frac{\partial R}{\partial x}\right|^2 + \left|\frac{\partial G}{\partial x}\right|^2 + \left|\frac{\partial B}{\partial x}\right|^2$$
$$f_y = \left|\frac{\partial R}{\partial y}\right|^2 + \left|\frac{\partial G}{\partial y}\right|^2 + \left|\frac{\partial B}{\partial y}\right|^2 \quad (3)$$
$$f_{xy} = \frac{\partial R}{\partial x} \times \frac{\partial R}{\partial y} + \frac{\partial G}{\partial x} \times \frac{\partial G}{\partial y} + \frac{\partial B}{\partial x} \times \frac{\partial B}{\partial y}$$

And the color tensor may be defined as:

$$T = \begin{bmatrix} f_x & f_{xy} \\ f_{xy} & f_y \end{bmatrix} \quad (4)$$

Once defined the structure of the color tensor, we calculate tensor coefficients sequentially for each small window w_j. For a scalar image, the tensor T has a unique egeinvalue λ_1, different from zero and which is equal to the maximum value of the quadratic form. This value is the square of the modulus of the gradient, denoted $Gamp$:

$$\|\nabla I\| = \sqrt{\lambda_1} \quad (5)$$

This data gives more knowledge about the discontinuity between pixel clusters and allows to adapt or improve the estimation of the affinity between neighboring pixels. In a windows w_j, the more the value of $\|\nabla I\|$ is high, the more its pixels will tend to be separated. In our proposed solution, we empirically set a threshold δ, which provides the best possible results based on images used in (Levin, 2008). We defined that if the amplitude of the tensor is greater than $\gamma = 0.09$, the pixels of the corresponding window does not come from the same cluster. We have translated as follows:

$$A_q(i,j) = \quad (6)$$
$$\begin{cases} \delta_{ij} - \frac{1}{w_q}\left(1 + (I_i - \mu_q)^t \times Gamp_q{}^2\right)\left(\sum_q \frac{\varepsilon}{|w_q|} I_{3x3}\right)^{-1} \left((I_j - \mu_q) \times Gamp_q{}^2\right); if\ \gamma > 0.09 \\ \delta_{ij} - \frac{1}{w_q}\left(1 + (I_i - \mu_q)^t\right)\left(\sum_q \frac{\varepsilon}{|w_q|} I_{3x3}\right)^{-1} (I_j - \mu_q); if\ \gamma > 0.09 \\ 0;\ otherwise \end{cases}$$

4 EXPERIMENTAL RESULTS

In order to evaluate our improved spectral matting approach, we carried out our study on a set of natural images referred in image matting researches. The algorithm is implemented in Matlab R2011b on a 2.40 GHZ CPU and its running time was nearby the same as (Levin, 2008). The number of clusters using K-means algorithm is set as 10 which gives us better results. The number of the smallest eigenvectors in finding matting components is set as 50.

To quantitatively evaluate and compare our proposed approach with (Levin, 2008), we use the following three metrics: Sum of Absolute Differences, Mean Absolute Error and Root Mean Squared Error.

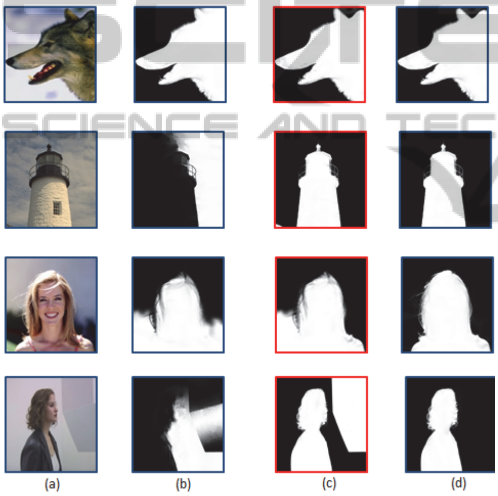

Figure 2: Image matting results with the two spectral matting algorithms.

Figure2 shows the image matting results of tested images with our proposed method and the original one (Levin, 2008). Figure 2(a) is a wolf image, a tower image, a wind image and Amira image, respectively. Figure 2(B) is the alpha matte results using the spectral matting algorithm (Levin, 2008). Figure 2(C) is the alpha matte results using our proposed approach that are framed in red. Figure 2(D) is the ground truth images.

Table 1: SAD errors for the estimation of alpha mattes.

SAD	Spectral Matting (Levin 2008)	Proposed method
Wolf	518089	93001
Tower	6016516	54189
Wind	2005183	1324222
Amira	30548255	19202333

Table 2: MAE errors for the estimation of alpha mattes.

MAE	Spectral Matting (Levin 2008)	Proposed method
Wolf	4,7444	0,8517
Tower	91,8047	0,8269
Wind	31,0804	20,5255
Amira	107,8681	67,8048

Table 3: RMSE errors for the estimation of alpha mattes.

RMSE	Spectral Matting (Levin 2008)	Proposed method
Wolf	2,8823	2,3506
Tower	9,9809	1,4805
Wind	6,6760	5,4085
Amira	10,7515	8,5210

Furthermore, Table1, Table2 and Table3 show the errors results of the relative metrics for those two spectral matting algorithms.

It can be seen from those three tables that our proposed method outperforms the spectral matting method using SAD, MAE and RMSE.

Moreover, in order to highlight the performance of the proposed method, we infer the visual quality for image matting from human observers. This is due the fact that the three metrics mentioned above may not be representative of errors noticed by a human. Thus, the figure 3 shows some results of generated final alpha mattes. Figure 3(a) is the real images. Figure 3(b) is alpha mattes generated by spectral matting (Levin, 2008). Figure 3(c) is alpha mattes generated by our spectral matting.

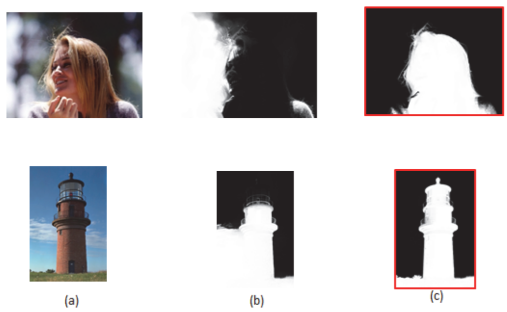

Figure 3: Results of unsupervised image matting.

Both quantitative measurements and visual results show that the proposed spectral matting method can generate better alpha mattes for natural images without user intervention compared to the original spectral matting method.

5 CONCLUSIONS

In this paper, we presented an improved spectral matting approach by using the gradient color tensor of color images. Gradient information are integrated into the affinity matrix computation process. Experiments through both quantitative measurements and visual results show that our proposed method has better performance in the foreground extraction process than the original spectral matting. In future, we intend to incorporate a GPU version of the spectral matting with gradient color tensor information in order to speed up the processing time. We also are looking for applying our proposed method to a specific use case of immersive communication that we call "Presentation at Distance" in order to improve video matting and human actions classification.

REFERENCES

Porter, T, Duff, T, 1984.Compositing Digital Images. In: *Proceedings of ACM SIGGRAPH.*

Wang, J, Cohen, M. F., 2007. Image and Video Matting: A survey. *Foundations and trends in computer Graphics and vision*.

Ruzon M, Tomasi C, 2000. Alpha estimation in natural images. In *CVPR.*

Y.-Y Chuang, B.Curless, D.H.salesin, szeliski, 2001. A Bayesian approach to digital matting. IEEE CVPR.

J. Sun, J. Jia, C.-K. Tang, H.-Y. Shum,2004. Poisson matting. In *Proceedings of ACM SIGGRAPH*.

L. Grady, T. Schiwietz, S. Aharon, R. Westermann, 2005. Random walks for interactive alpha-matting. In *Proceedings of VIIP 2005*.

X. Bai and G. Sapiro, 2007. A geodesic framework for fast interactive image and video segmentation and matting. In *Proceedings of IEEE ICCV*.

A. Levin, D. Lischinski, and Y. Weiss, 2006. A closed form solution to natural image matting. In *Proceedings of IEEE CVPR*.

A. Levin, A. Rav-Acha A., D. Lischinski, 2008. Spectral matting. *IEEE transactions on pattern analysis and machine intelligence*, 30(10), pp.1699–712.

Jun Zhu, Dengsheng Zhang, Guojun Lu, 2010. An Enhancement to close form method for natural image matting. *Digital image computing: Techniques and Applications*.

Yu, S. X, Shi, J, 2003. Multiclass spectral clustering. In *Proceedings of international Conference on Computer Vision*.

Wu-Chih Hu, Deng-Yuan Huang, Ching-Yu Yang, Jia-Jie Jhu, Cheng-Pin, 2010. Automatic and accurate Image matting. In: *Proceedings of the 2nd international Conference on Computational Collective Intelligence-Technology and Applications*.

Wu-Chih Hu, and Jung-Fu-Hsu, 2012. Automatic image matting using component-Hue-Difference-Based Spectral Matting. In: *Proceedings of ACIIDS'12 of the 4th Asian Conference on Intelligent Information and Database Systems*, Berlin, Heidelburg.

Tung-Yu Wu, hung-hui juan, Horng-Shing Lu, 2012. Improved spectral matting by iterative k-means clustering and the modularity measure. *IEEE International Conference an Acoustics, Speech, and Signal Processing*, Kyoto international conference center, Japan.

Silvano Di Zenzo, 1986: A note on the gradient of a multi-image. *Computer Vision, Graphics and Image processing*.