Nilesh G. Nayi / International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com
Vol. 2, Issue 3, May-Jun 2012, pp.2182-2185

Image Matting for Natural Image

Nilesh G. Nayi
Electronics and Communication Department,
Vishwakarma Government Engineering College,
Chandkheda, India

ABSTRACT

The image matting problem is to extract a foreground object from a still natural image with the limited help from user. The foreground layer consists of three color layers and opacity layer. In this paper we propose a user assisted image matting workflow. Starting from a manual marking trimap, we perform hard segmentation and allow the user to quickly turn the segmentation result into a trimap. The algorithm used to apply soft matting process and allow the user to refine the result by fine-tuning the matting parameters in distinct image regions.

Keywords – **alpha, matting**

I. INTRODUCTION

Separating a foreground object in an image from its background is an important operation in image as well as video editing, with many applications in the entertainment industry. For instance, once an object has been separated from its background, it may be blended with another background scene. It is a technique in many image editing applications and has been extensively studied for more than two decades.

In this paper we address the problem of foreground object extraction. It is also known as the image matting problem. The first formal introduction of the matting problem was given by Porter and Duff in 1984[1]. The original purpose of their work is to introduce the alpha channel as the way to blend the foreground and background images. A given image I is considered to be a linear combination of a background image B and a foreground image F using the compositing equation:

$$I = \alpha F + (1-\alpha)B \qquad (1)$$

Where I, F and B are 3D vectors of RGB values. The alpha matte α takes values in the range [0, 1]. The pixel with corresponding α=1 is said to be a pure foreground and α=0 is said to be a pure definite background pixel. Otherwise, it is a mixed pixel. The problem is to reconstruct the alpha matte α, F and sometimes B images from the source image I using some additional user input.

Image is usually accompanied by a trimap – ternary image which divides the image area into three regions: foreground, background and unknown. The former two act as hard constraints while the latter denotes the area matting, where the matting algorithm is applied.

II. PREVIOUS WORK

From the human intervention point of view, the single image based matting methods can be classified into three types: trimap based [2, 3, 4, 5, 6], scribble based [7, 8], and automatic [9].

Several approaches were used to take the advantage of typical alpha matte structure: large areas with α = 0 or 1 with rather thin soft edge between them [10,11,6].

GrabCut algorithm [11] uses interactive hard segmentation and border matting by constraining the shape of soft object edge.

Coherence matting algorithm [12] uses hard segmentation and a smoothness constraint which attracts alpha to a precomputed value obtained from edge feathering. However, this works well only for blurry but distinct edges, i.e. with focusing or anti-aliasing effects. Large fuzzy areas cannot be processed in this way.

Knockout algorithm requires a precise trimap, ideally with unknown region containing only pixels with 0<α<1. When processing an unknown region pixel, F and B values are estimated by averaging color along the foreground/background region border in the neighborhood of the pixel being processed. α value is then calculated for each color component independently and weighted average is used as final α value. While being very fast, this algorithm produces poor results when F and/or B values in the pixel are inconsistent with the color along the corresponding region boundary. This happens in many images and the algorithm produces incorrect and noisy results.

Bayesian matting [2] also uses color statistics, but performs per pixel color distribution estimation. Pixels are processed starting from foreground and background region borders contracting unknown region step by step. Pixels processed on earlier steps provide new foreground and background samples in addition to pixels from known regions. Used color model is a set of oriented Gaussians.

Algorithm involves Bayesian framework to maximize the likelihood of F, B and α values. Conditional probability for F, B and α given observed color I can be written using Bayes's rule as:

$$P(F,B,\alpha \mid I) = P(I \mid F,B,\alpha)\, P(F)P(B)P(\alpha)/ P(I) \qquad (2)$$

where P(I|F,B,α) is estimated using the distance between I and the mix of F and B (i.e. by the norm of the difference of the left hand side and right hand side of equation (1)), P(F) and P(B) are estimated via probability density of foreground and background Gaussians, P(α) is ignored (assuming all α values to be equiprobable), P(I) is constant relatively to maximization parameters.

III. TRIMAP

As mentioned in previous section, alpha matting is an ill-posed problem. Therefore, it is needed that additional information about the image before to proceed with alpha estimation. The user manually segments the input image into three regions, called trimap. A trimap is composed of three regions: a known foreground F (white), a known background B (black), and an unknown region U (gray). The input image is shown in fig. 1(a) and their trimap as shown in fig. 1(b). The foreground and background region provides the additional information that is needed to estimate α in the unknown region. One of the important factor affecting the performance of a matting algorithm is how accurate the trimap is. Ideally, the unknown region in the trimap should only cover truly mixed pixels. In other words, the unknown region around the foreground boundary should be as thin as possible to achieve the best possible matting results.

(a) (b)

Fig. 1: (a) Input Image (b) Trimap

IV. ALPHA MATTE ESTIMATION

Matting consists of two main tasks: alpha matte's estimation, and foreground (and background) color's computation. Given an image I for which the complete set of pixels is denoted by $\Omega = \{1,\cdots,n\}$ where n is the total number of pixels, and given a set of labeled pixels $\Omega_l \subset \Omega$ for which we know the α values, alpha matte estimation is defined as computing the α values of the set of unlabeled pixels $\Omega_u = \Omega - \Omega_l$. Alpha matte estimation is to estimate the alpha value of the unlabeled pixels (unknown region) with a global alpha-color model trained from some chosen labeled pixels (known region). The chosen pixels closer to the unlabeled pixel can help more in the matte's estimation.

Therefore it suits particularly to the case when a trimap is provided and the unknown region is slim. We choose for each unknown pixel two subsets from its nearby labeled foreground and background pixels respectively.

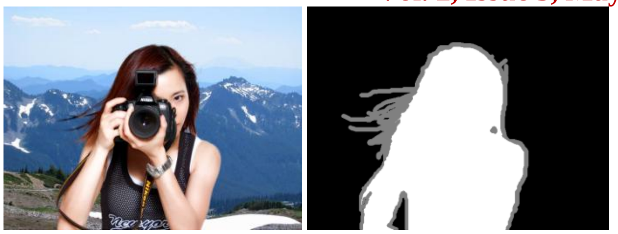
To choose the subsets from the labeled pixels, we first select two subsets $Q_l^f \subseteq \Omega_l^f$ and $Q_l^b \subseteq \Omega_l^b$, in which for any pixel j we have $D_j < D_{th}$, where D_j means the shortest Euclidean distance of j to the pixels in Ω_u on the regular lattice, and D_{th} is a distance threshold. For each unknown pixel i, we then select two subsets $Q_l^{f'} \subset Q_l^f$ and $Q_l^{b'} \subset Q_l^b$, in which the pixels have the shortest distance to i. We set the two subsets having the same number (e.g. 80) of pixels. To compute D_j, we use the algorithm in that can finish in a linear time. D_{th} is determined by $D_{th} = (\gamma_d|\Omega_u|)/|\Omega| + \sqrt{2}$, where γ_d is a constant and is empirically set to γ_d= 1.2. It is designed with the strategy that, when the unknown region is thicker, it is larger and more labeled pixels near the pixel being estimated can be selected, and otherwise, it is smaller and more distant labeled pixels can be chosen. Note that D_{th} is set with the distance to the unknown region instead of the pixel being estimated.

For each pixel j in the subsets $Q_l^{f'}$ and $Q_l^{b'}$, we set a weight $w_j = 1/(D_j)^{\gamma w}$. Where, γ_w = 0.25 is an empirically determined constant. We further create a diagonal matrix W_{Qi} with the w values of the pixels in $Q_l^{f'} \cup Q_l^{b'}$, whose size is t × t where t =$|(Q_l^{f'} \cup Q_l^{b'})|$. Now introducing α_{Qi} to denote the vector composed of the alpha values of the pixels in $Q_l^{f'} \cup Q_l^{b'}$ and X_{Qi} to represent the matrix constructed but with the data values of the pixels in $Q_l^{f'} \cup Q_l^{b'}$. For any pixel $i \in \Omega_u$, we have its alpha value's estimation with a linear model as below:

$$\alpha_i = x_i^{'T} X_{Qi}^T W_{Qi} \left(W_{Qi} X_{Qi} X_{Qi}^T W_{Qi} + \lambda_r I_{(t)} \right)^{-1} \alpha_{Qi} \qquad (2)$$

IV. RESULT

Here we compare the MSE of matting results with Bayesian matting [2], Closed form [8], Knockout using the test set of Wang-Cohen [5] in Table 1. From the computed α and given groundtruth α*, the MSE is calculated as MSE=$\sum_i (\alpha_i - \alpha_i^*)$.



(a) (b)

(c) (d)

Figure 2: (a) Input Image (T5) (b) Trimap (c) Groundtruth (d) Alpha matte

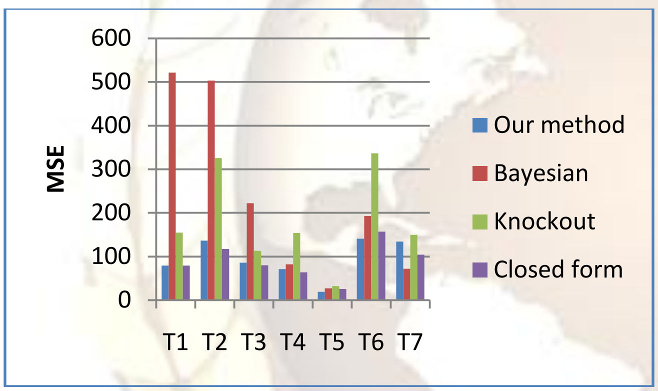


Table 1: MSE of alpha matte of test image sets

(a) (b) (c)

(d) (e)

Fig. 3: (a) Input image (b) Trimap (c) Alpha matte (d) Extracted foreground (e) Composite image

V. CONCLUSION

Trimap is one of the important parameter for the based image matting, as trimap is accurate we get accurate alpha matte, which is necessary for object extraction. Using the image matting we can also extract the multiple object, provided that trimap of image. Still, applying this process gives some error in complex fuzzy boundaries, like hair, fur.

REFERENCES

[1] T. Porter and T. Duff, Compositing digital images, *in Proceedings of ACM SIGGRAPH*, pp. 253–259, July 1984.
[2] Y. Chuang, B. Curless, D. Salesin, and R. Szeliski, A Bayesian approach to digital matting, *In CVPR01*, 2001.
[3] J. Sun, J. Jia, C.-K. Tang, and H.-Y. Shum, Poisson matting, *In ACM Trans. Graph*, volume 23, pages 315–321, 2004.
[4] L. Grady, T. Schiwietz, S. Aharon, and R. Westerman, Random walks for interactive alpha-matting, *In VIIP05*, 2005.
[5] J.Wang and M. Cohen, Optimized color sampling for robust matting, *In CVPR07*, 2007.
[6] J.Wang, M. Agrawala, and M. F. Cohen, Soft scissors: An interactive tool for realtime high quality matting, *SIGGRAPH'07*, 2007.
[7] J. Wang and M. Cohen, An iterative optimization approach for unified image segmentation and matting, *In ICCV05*, 2005.
[8] A. Levin, D. Lischinski, and Y. Weiss, A closed-form solution to natural image matting, *In IEEE TPAMI*, number 3, pages 228–242, 2008.
[9] A. Levin, A. R. Acha, and D. Lischinski, Spectral matting, *In CVPR07*, 2007.
[10] Apostoloff, N. and Fitzgibbon, A. Bayesian video matting using learnt image priors, *Proc. Of IEEE CVPR*, pp. 407–414, 2004.
[11] Rother, C., Kolmogorov, V., Blake, A. GrabCut: interactive foreground extraction using iterated graph cuts, *ACM Transactions on Graphics*, 23(3):309–314, 2004
[12] Shum, H.-Y., Sun, J., Yamazaki, S., Li, Y. and Tang, C.-K. , Pop-up light field: An interactive image-based modeling and rendering system, *ACM Transactions on Graphics*, 23(2):143–162, 2004.