2013 13th International Conference on Computer-Aided Design and Computer Graphics

Automatic Matting of Identification Photos

Wenshuang Tan[1], Tiantian Fan[1], Xudong Chen[1], Yaobin Ouyang[1] , Dong Wang[2], Guiqing Li[1]

[1]South China University of Technology, [2]South China Agricultural University

Guangzhou, China
Corresponding author: ligq@scut.edu.cn

***Abstract*—this paper proposes a framework for matting identification photos automatically. We firstly adopt image processing operations such as skin detection, k-means clustering, Grabcut segmentation and canny operator to generate a trimap with three regions, i.e., background, foreground and unknown regions. Then, we improve the Bayesian matting by introducing an alpha regularization term. Experiments demonstrate that our system can achieve about 86% acceptance rate for real photo datasets.**

Keywords: image matting; automatic trimap; Bayesian matting; identification photos.

I. INTRODUCTION

Image matting was initially developed for image composition in computer animation [5], however, most of matting algorithms such as Poisson matting [3], Bayesian matting [2], KNN matting [1] and so on, require a trimap image (consists of background, foreground and unknown regions) as input. To generate trimap images, users have to interactively draw strokes for indicating partial or whole foreground and background regions. For specific applications, it is impossible or at least inefficient. For example, in online identification photo processing systems, there are thousands of images submitted within a work day. Furthermore, a higher quality requirement on human faces makes the matting of identification photos more challenge.

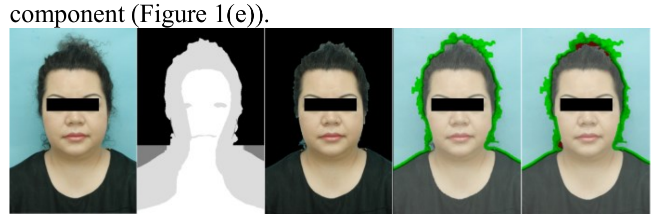
We propose a framework to address automatic matting of identification photos based on two specific requirements of such kind of images: (1) The background of original images is near monochrome, and the foreground only contains a human bust and occupies a major region of the image; (2) The positions of background and foreground are relatively fixed. Our main contributions are summarized as follows:

- An automatic matting framework is presented for identification photo processing;
- A smooth term is introduced to improve the Bayesian matting [2] such that α can smoothly change across the boundary between the background and foreground regions; furthermore, it greatly alleviates the blur phenomenon near the boundary;
- For identification photo matting, our framework can achieve about 86% correct rate which is an acceptable percent for automatic batch processing.

II. METHODS

A. Automatic generation of trimaps

We first employ the eye detector of OpenCV to locate two eyes of the person and estimate the horizontal line of his/her chin. This line partitions the image into two: the upper part above the chin and the bottom part below it. We then perform k-means clustering in the upper area with three centers. This partitions the region into three components: background, hair and human face. In addition to these inputs, we create a minimal foreground template which is the common foreground of a set of segmented images. Finally, the image (Figure 1(a)) is partitioned into four parts (Figure 1(b)). We then employ Grabcut to segment the whole image into two semantic parts: the foreground and the background (Figure 1 (c)). Along the boundary, a trimap will be automatically generated. To achieve this, we firstly detect hairline edges in the background using canny operator (Figure 1(d)). It can find most of salient hair lines spreading in this region. We then scan the hair region (only the upper part) to search for the area including background color component (Figure 1(e)).



(a) (b) (c) (d) (e)

Figure 1. Pipeline illustration of the proposed approach: (a) The source image; (b) the source image is partitioned into 4 areas: background (black), possible background (deep gray), possible foreground (light gray) and foreground (white); (c) the image is segmented into two components using Grabcut; (d) detection of hairlines (green pixels) in the background side; (e) detection of unknown areas inside the hair of the foreground side (deep red pixels).

B. Smoothness constraints for alpha maps in Bayesian matting

Bayesian Matting [2] maximizes the following condition probability $\mathrm{P}(\mathrm{F}_i, \mathrm{B}_i, \alpha_i|\mathrm{C}_i)$, equivalently maximizes

$$\mathcal{E}_{ORG}(i) = \ln\mathrm{P}(\mathrm{F}_i, \mathrm{B}_i, \alpha_i|\mathrm{C}_i) \approx \ln\mathrm{P}(\mathrm{C}_i|\mathrm{F}_i, \mathrm{B}_i, \alpha_i) + \ln\mathrm{P}(\mathrm{F}_i) + \ln\mathrm{P}(\mathrm{B}_i), \tag{1}$$

assuming a pixel C_i is blended by $\mathrm{C}_i = (1-\alpha_i)\mathrm{F}_i + \alpha_i\mathrm{B}_i$,

The transition from background to foreground regions is usually smooth in a natural image. To simulate this phenomenon, we introduce a smooth term to force α smoothly varying:

$$\mathcal{E}_{ALP}(i) = \ln\mathrm{P}(\alpha_i) = -\mathrm{w}_i(\Delta\alpha_i)^2 \tag{2}$$

where

$$\Delta\alpha_i = \alpha_i - \frac{1}{8}\sum_{j=0}^{7}\alpha_{i_j} \tag{3}$$

is the Laplacian operator with i_j being the eight-connected pixels, and w_i are weights used to control the shape of the

978-1-4799-2576-6/13 $31.00 © 2013 IEEE
DOI 10.1109/CADGraphics.2013.60

cross profile of the alpha map, which we will be detailed later. It finally results in the following global energy

$$\mathcal{E}_{IMP} = \mathcal{E}_{ORG} + \lambda\mathcal{E}_{ALP}, \tag{4}$$

where $\mathcal{E}_{ORG} = \sum_{i=0}^{N-1} \mathcal{E}_{ORG}(i)$, $\mathcal{E}_{ALP} = \sum_{i=0}^{N+M-1} \mathcal{E}_{ALP}(i)$ with N, the number of unknown pixels, and M, the number of known pixels enclosing the unknown pixels, and $\lambda = 2000$ in our experiment. We introduce following weights to constrain the value of the Laplacian operator of the current pixel $\mathrm{w}_i = e^{-d_i/\sigma_w^2}$, where d_i is the distance from pixel i to the known regions and $\sigma_w = 2$ in our experiments.

Notice that $\mathcal{E}_{ALP}$ in (2) does not include parameters B_i and F_i, we can maximize it using the block-decent coordinate strategy. Therefore, we only need to solve a global linear system to maximize (3) to obtain all unknown α_i. Figure 2 shows that forcing the shape of the cross profile of α map as harmonic as possible can only guarantee its smoothness and can't eliminate blur phenomenon. The blue curve indicates that Laplacian operator of α changes in the pattern.

(a) (b) (c) (d)

Figure 2. Transition of α map across boundaries: (a) varying curves of alpha from background to foreground; (b) a composite example with the red curve as the shape of α; (c) zoom-in of local region; (d) result of the same local region with the blue curve as the shape of α. (d) shows a better transition across the boundary while the purple curve will result in the worst result among three curves.

III. EXPERIMENTAL RESULTS

Our framework was implemented by C++ of Visual Studio 2010 and it directly called OpenCV library for most of the basic image operations such as grabcut, k-means, canny operator, and detection operators of face, nose, eyes, and mouth. We test our method with two datasets which are real data uploaded by individual users on a day (645 blue and 432 white/gray background photos), and achieve 89.612% and 84.027% acceptable rates, respectively. Figure 3 shows automatic composition with different backgrounds. Figure 5 shows that the color transition across face boundary regions of our method looks the most pleasing. As comparison, we also present the closed-form matting [7] and KNN matting [1] in Figure 4 which depicts that our matting can capture more details.

Figure 3. From left to right, the first two are respectively original image and original image with trimap. The last three are composite results with white, red and blue background.

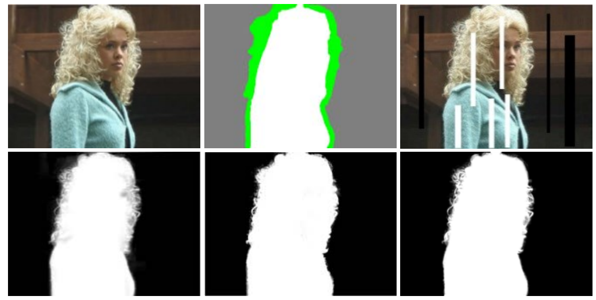


Figure 4. Natural image matting: top from left to right are respectively original image, input trimap for KNN matting and our approach, and scribble input for closed-form matting; bottom from left to right are generated α maps of the closed-form, KNN and our mattings.

ACKNOWLEDGMENT

This paper is partially supported by the Fundamental Research Funds for the Central Universities (No. 2013ZM087, 2012zm0062), the Open Project Program of the State Key Lab of CAD&CG (No.A1309, Zhejiang Univ.), and NSFC (No. 61202294). Closed-form matting and KNN matting in Figure 4 are generated by using codes from their authors.

REFERENCES

[1] Chen, Q, Li, D., and Tang, C.. KNN matting. CVPR 2012.

[2] Chuang, Y.-Y., Curless, B., Salesin, D.H., and Szeliski, R. . A Bayesian approach to digital matting". CVPR 2001, II, 264-271.

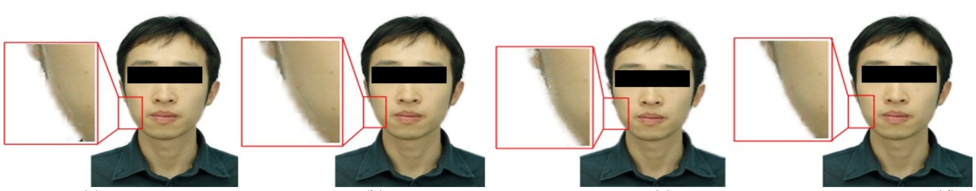
[3] Sun, J., Jia, J., Tang, C.-K., and Shum, H.-Y.. Poisson matting. ACM Trans. Graph. (SIGGRAPH 2004), 23(3): 315–321.

[4] Sindeyev, M., Konushin, V., Vezhnevets, V.. Improvements of Bayesian Matting. GraphiCon 2007, 88-95.

[5] Wang, J. and Cohen, M. F.. Image and video matting: a survey. Foundations and Trends in Computer Graphics and Vision, 3(2): 97-175, 2007.

[6] Wang, J. and Cohen, M.F.. Optimized color sampling for robust matting. CVPR 2007

[7] Levin, A., Lischinski, D., and Weiss, Y.. A closed form solution to natural image matting. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(2):228-242, 2008



(a) (b) (c) (d)

Figure 5. Two examples for (a) Bayesian matting [2], (b) Improved Bayesian matting [4], (c) Robust Matting [6], (d) Our improved Bayesian matting.