Sparse Codes as Alpha Matte

Jubin Johnson
JUBIN001@e.ntu.edu.sg
Deepu Rajan
ASDRAJAN@ntu.edu.sg
Hisham Cholakkal
HISHAM002@ntu.edu.sg

School of Computer Engineering
Nanyang Technological University
Singapore

Matting is a useful tool for image and video editing where foreground objects need to be extracted and pasted onto a different background. A matte is represented by α which defines the opacity of a pixel and is a value in [0, 1], with 0 for background (B) pixels and 1 for foreground (F) pixels. There are three main approaches for image matting: In sampling-based approaches, a foreground-background sample pair is picked from few candidate samples taken from F and B regions by optimizing an objective function. This (F,B) pair is then used to estimate α at a pixel with color I by

$$\alpha_z = \frac{(I-B)(F-B)}{\|(F-B)\|^2}. \tag{1}$$

α-propagation based methods assume correlation between the neighboring pixels under some image statistics and use their affinities to propagate alpha values from known regions to unknown ones. The third category is a combination of the two in which the matting problem is cast as an optimization problem.

The method proposed in this paper is based on sampling. However, there is one important difference between our method and other sampling-based approaches. Matting is cast as a sparse coding problem wherein the sparse codes directly give the estimate of the alpha matte. Hence, there is no need to use the matting equation that restricts the estimate of α from a single pair of foreground and background samples. This allows the matting framework to determine α based on more relevant F and B samples than with only one of each.

A dictionary of color values of F and B pixels is employed to determine the sparse codes for a pixel in an unknown region. The sum of the sparse codes for F pixels directly provides the α. Initially, the pixels in the trimap are classified into high-confidence and low-confidence based on probabilistic segmentation. Since the feature used for coding is color and the complexity of a region for matting is dependent on the overlap of foreground and background colors, we use probabilistic segmentation [2] as a cue to determine the confidence of a pixel as follows:

$$p(I_i) = \frac{p_f(I_i)}{p_f(I_i)+p_b(I_i)}, \tag{2}$$

where $p_f(I_i)$ is the foreground color probability value given by

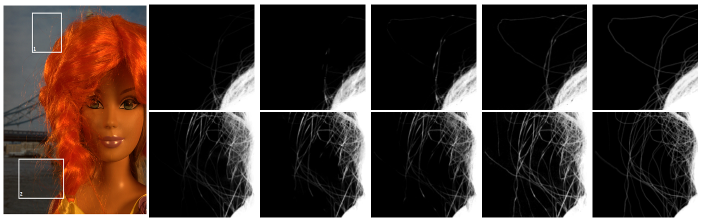
$$p_f(I_i) = exp\left(-\frac{\sum_{k=1}^{m}\|c(I_i)-c(f_k)\|^2}{m\cdot\delta}\right), \tag{3}$$

where $c(\cdot)$ is the RGB color value, m is the number of spatially close foreground samples. A similar formulation exists for background color probability $p_b(I_i)$.

(a) (b) (c)

Figure 1: Alpha matte extracted using our proposed sparse coding method. (a) Input image, (b) Extracted matte and (c) Ground truth.

The size of the dictionary for high-confidence pixels is smaller than that for low-confidence pixels. A universal sample set is generated using a superpixel-based sampling strategy, which is detailed in the paper. For a given unknown pixel of low-confidence, the final dictionary is a larger subset of the universal sample set than that of high-confidence pixels.



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

Figure 2: Visual comparison of alpha matte generated by our proposed method using sparse coding with other state-of-the-art methods. Top and bottom rows show zoomed in regions of windows 1 and 2 respectively. (a) Input image, (b) Closed form, (c) Weighted color and texture, (d) Comprehensive sampling, (e) Proposed method and (f) Ground truth.

Given the final dictionary $\mathbf{D}$ for an unknown pixel i, its alpha matte is determined by sparse coding as

$$\begin{aligned} \beta = argmin\|v_i - \mathbf{D}\beta_i\|_2^2 \quad s.t \;\; \|\beta_i\|_1 \leq 1\;;\; \beta_i \geq 0, \\ \alpha = \sum_{p\in F}\beta^{(p)}, \end{aligned} \tag{4}$$

where v_i is the feature vector at i composed of (R,G,B,L,a,b). The sparse codes β_i are generated using a modified version of the Lasso algorithm [3]. The sparse coding procedure is presented with an appropriate set of F and B samples and the sparse coefficients sum up to less than or equal to 1. In order to avoid negative sparse coefficients, the second constraint forces all coefficients to be positive. The sparse codes corresponding to atoms in the dictionary that belong to foreground are added to form the α for the unknown pixel.

The alpha matte obtained by sparse coding is further refined to obtain a smooth matte by considering the correlation between neighboring pixels' matte. We adopt the post-processing approach [4] where a cost function consisting of the data term and a confidence value together with a smoothness term is minimized with respect to α.

Implementation of cost function optimization is described in the paper. The contribution of each part of our proposed method is analyzed with quantitative and qualitative experiments conducted on a benchmark database [1] used universally for image matting evaluation. Our conclusion is that the simplicity of the sparse coding model, coupled with its ability to break away from the $F-B$ pair assumption in matting, makes it a useful tool for future insight into understanding the matting process.

[1] http://www.alphamatting.com.

[2] J. Ju, J. Wang, Y. Liu, H. Wang, and Q. Dai. A progressive tri-level segmentation approach for topology-change-aware video matting. In *Computer Graphics Forum*, volume 32, pages 245–253, 2013.

[3] J. Mairal, F. Bach, J. Ponce, and G. Sapiro. Online learning for matrix factorization and sparse coding. *The Journal of Machine Learning Research*, 11:19–60, 2010.

[4] E. Shahrian, D. Rajan, B. Price, and S. Cohen. Improving image matting using comprehensive sampling sets. In *CVPR*, 2013.