A Global Sampling Based Image Matting Using Non-Negative Matrix Factorization

NAVEED ALAM*, MUHAMMAD SARIM*, AND ABDUL BASIT SHAIKH*

ABSTRACT

Image matting is a technique in which a foreground is separated from the background of a given image along with the pixel wise opacity. This foreground can then be seamlessly composited in a different background to obtain a novel scene. This paper presents a global non-parametric sampling algorithm over image patches and utilizes a dimension reduction technique known as NMF (Non-Negative Matrix Factorization). Although some existing non-parametric approaches use large nearby foreground and background regions to sample patches but these approaches fail to take the whole image to sample patches. It is because of the high memory and computational requirements. The use of NMF in the proposed algorithm allows the dimension reduction which reduces the computational cost and memory requirement. The use of NMF also allow the proposed approach to use the whole foreground and background region in the image and reduces the patch complexity and help in efficient patch sampling. The use of patches not only allows the incorporation of the pixel colour but also the local image structure. The use of local structures in the image is important to estimate a high-quality alpha matte especially in the images which have regions containing high texture. The proposed algorithm is evaluated on the standard data set and obtained results are comparable to the state-of-the-art matting techniques.

Key Words: Alpha Matte, Non-Negative Matrix Factorization, Non-Parametric SamplingMethod.

1.   INTRODUCTION

In image matting, a foreground is separated from the background of the given image which can then seamlessly be composited in a new background to produce a novel scene. Porter and Duff in their seminal paper [1] introduced the composite equation for image matting.

$$I = \alpha F + (1-\alpha)B \qquad 0 < \alpha < 1 \qquad (1)$$

Equation (1) shows that an image I is the combination of two layers F, B and alpha, weighted by the blending proportion. The pixels which have $\alpha=0$ refer asdefinite background and $\alpha=1$ for definite foreground object while $0 << 1$ represents the semi-transparent foreground pixels. The problem is inherently ill posed because Equation (1) consists of more unknown terms F,B and than the known term I.

The problem is alleviated by manually aiding the information of the definite foreground, background and semitransparent regions in the form of a tri-level map as illustrated in **Fig. 1.** The information from the definite regions is used to solve Equation (1) for the semi-transparent foreground pixels.

Corresponding Author: (E-Mail: n.alam@fuuast.edu.pk)
* Department of Computer Science, Federal Urdu University of Arts, Science & Technology, Karachi.

Approaches like [2-4] have defined different affinities among the neighboring pixels because there exist a strong correlation among the local pixels that could be used for smoothness assumption. Sun et. al. [2] applied Poisson formulation to Equation (1), under the assumption that the intensity changes are locally smooth. Grady et. al. [3] proposed a method, similar to [2], that is based on Random Walks and define the affinity matrix in term of Gaussian function. Levin et. al. [4] formulated a novel quadratic cost function in α to obtained the alpha matte. Since matting information is propagated from the known regions to the unknown region pixels these approaches are prone to error propagation across the matte.

Some approaches like [5-7] fit statistical models on the pixels of the local known regions to estimate the colour and alpha value of each unknown region pixels. Recent methods [8-12] estimates alpha values by sampling local known regions pixels non-parametrically. These approaches fail where the local samples are not representative of the true colours of the unknown region pixels. In [8] authors have used Non-negative matrix factorization but it is limited to the foreground and background patch spatially closest to the unknown pixel and does not colour information in patch sampling. Like other approaches [8] is also limited to the local regions.

This work presents a global non-parametric sampling based approach to overcome the aforementioned problem present in the local non-parametric modeling. We sampled

FIG. 1. A NATURAL IMAGE ALONG WITH A MANUALLY DEFINED TRIMAP

image patches as in [11] instead of pixel to incorporate the local image structure along with colour information. NMF [13] is used to efficiently determine the similar patches of the known regions for each unknown region pixels. Only high confidence sample pairs are chosen using linear analysis to estimate the alpha matte.

2. RELATED WORK

The first reported work on non-parametric sampling was Corel Knockout [14]. Foreground colour is estimated by using weighted average of the local samples at the unknown foreground boundary. Background colour is also estimated first by taking weighted average and then refined.

Wang and Cohen [9] introduced the colour fitness criteria for the spatially close sample pixels. Higher confidence sample pairs are used to estimate alpha. Rhemann et. al. [10] used a similar approach [9] but the selection of the candidate samples are based on geodesic distance instead of Euclidean. Gastal and Oliveira [12] sample pixels that lies along rays shooting from the unknown pixels and samples are further shared among neighbouring pixels. Sarim et. al. [11] proposed a patch based image matting technique that model local image colour and structure in a local window to estimate colour and then the alpha matte. Considering only two spatially nearest known patches could produce high inconsistency in the resulted alpha matte. All of these techniques are inefficient in regions where the foreground and background have similar colour distributions or they are not representative of the true colour of unknown region pixels.

3. NFM FACTORIZATION

A dimension reduction technique NMF [13] is used to find the linear appropriate representation of non-negative data. If a M dimensional data with N samples is denoted by a matrix X_{MXN}, then the linear approximation is given

as:

$$X_{MXN} \approx W_{MXO}H_{OXN} = \sum_{i=1}^{0} w_i h_i^c \qquad (2)$$

where, W is a basic matrix with columns representing basic vector, H is a coefficient matrix with columns as coefficient vector and 0 is the order of decomposition. The optimal W and H for X can be obtained by iteratively minimizing the error function:

$$f(W,H) = \frac{1}{2}\left\|X - WH\right\|^2 \qquad (3)$$

In order to minimizing the cost function Equation (3), many algorithms area available [15]. In this paper, Alternate least square optimization method is used to minimize Equation (3) due to its fast convergence rate and consistency.

4. IMAGE MATTING USING NMF

Proposed algorithm follows as (1) collecting candidate samples, (2) selecting best candidate samples and (3) matte optimization.

4.1 Collecting Candidate Samples

Initially an image is divided into foreground, background and unknown regions. A global foreground patch space S^f is constructed by localising a patch of dimension *mxn* at each pixel in the known region. Since NMF algorithm requires matrix to be two dimensional so foreground model M^f is formed by arranging each $\psi_i \in S^f$, of size *mxn*, as a column of:

$$M^f = \begin{bmatrix} \psi_1^1 & \psi_2^1 & \psi_3^1 & \cdots & \psi_N^1 \\ \psi_1^2 & \psi_2^2 & \psi_3^2 & \cdots & \psi_N^2 \\ \psi_1^3 & \psi_2^3 & \psi_3^3 & \cdots & \psi_N^3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \psi_1^M & \psi_2^M & \psi_3^M & \cdots & \psi_N^M \end{bmatrix} \qquad (4)$$

Since we are using RGB (Red, Green and Blue) colour

space so the dimension of $M=mxnx3$ and N represents the total number of patches in foreground F. To find the similar samples for a patch ϕ_p in the unknown region, foreground model M^f can be factorized as:

$$\mathbf{M^f \approx W^f H^f} \qquad (5)$$

W^f is the $[mxnx3]xd$ basic matrix for RGB colour image, while H^f is the coefficient matrix. The constant d represents order of decomposition while N^f denote total number of patches in the foreground model M^f.

A patch ϕ_p is the unknown region can be approximated by projecting coefficient matrix H^p on the basis matrix W^f as:

$$\mathbf{\phi_p \approx W^f H^f} \qquad (6)$$

The most similar patch ψ_q, is the foreground model V^f, to ϕ_p is obtained by minimizing the L2 norm of H^p and H^i as:

$$\psi_q \text{ arg min}\left\|H^p - H^i\right\| i = 1,2,...N^f \qquad (7)$$

Colour of the center pixel of the patch ψ_q as taken as the foreground colour to the pixel p is the unknown region as shown in Fig. 2. If the foreground contains textures then using only the most similar patch to estimate foreground colour could lead to segmentation inaccuracies hence N most similar patches are used to find the optimum foreground colour. Similar approach is used to determine most similar patches in the background region.

4.2 Selecting Best Candidate Samples

After collecting multiple candidate samples, the best foreground and background pair of $[F_p, B_p]$, for a pixel p is obtained as:

$$\left\{\alpha_p, F_p, B_p\right\} = \underset{\alpha_i, F_j, B_k}{\text{arg min}} \left\|I - \left(\alpha_i, F_j + (1-\alpha_i)B_k\right)\right\| \qquad (8)$$

The alpha values used for optimization ranges in [0 1] with an interval of 0.01. The optimization is performed for all unknown region pixels. The resulted alpha matte might contain discontinuities where optimum pair is not

Mehran University Research Journal of Engineering & Technology, Volume 36, No. 4, October, 2017 [p-ISSN: 0254-7821, e-ISSN: 2413-7219]

765

representative of the true colour of unknown region pixels. Matte optimization is used to improve the initially estimated alpha matte as shown in Fig. 3.

4.3   Matte Optimization

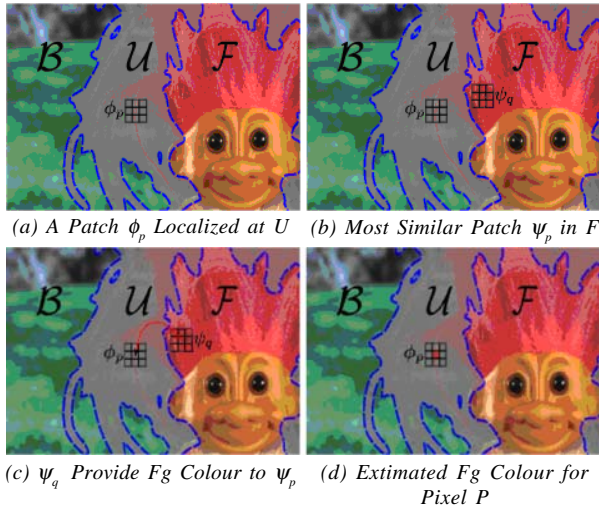A matte optimization approach similar to shared matting [12] and a cost function is defined as:

(a) A Patch ϕ_p Localized at U   *(b) Most Similar Patch ψ_p in F*

(c) ψ_q Provide Fg Colour to ψ_p   *(d) Extimated Fg Colour for Pixel P*

FIG. 2. ESTIMATION OF FOREGROUND COLOUR FOR THE UNKNOWN REGION

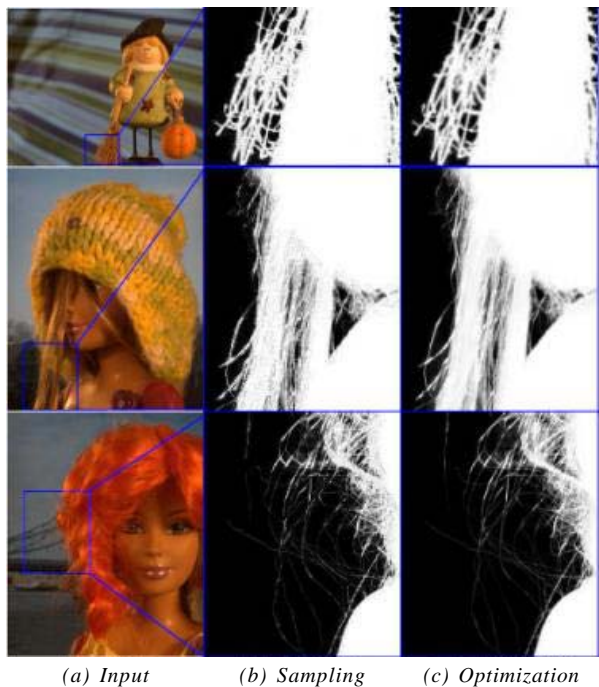*(a) Input   (b) Sampling   (c) Optimization*
FIG. 3. COMPARISON PROPOSED ALGORITHM RESULTS USING MATTE OPTIMISATION

$$\alpha = \arg\min \alpha^T L\alpha + \lambda(\alpha - \hat{\alpha})^T - D(\alpha - \hat{\alpha}) + \lambda(\alpha - \hat{\alpha})^T \hat{\Gamma}(\alpha - \hat{\alpha}) \quad (9)$$

Matting Laplacian L is taken as in [4] while λ and g are the weighting parameters. The matrix D whose diagonal elements are 1 for known and 0 for unknown pixels while matrix Γ having diagonal element 0 for known pixels and a confidence C for each unknown pixel which is computed as:

$$C = \exp\left(-\frac{1 - \{\alpha F + (1 - \alpha)B\}}{2\sigma^2}\right) \quad (10)$$

Where σ is a scaling parameter. This optimization process provides an alpha α propagation from pixels have high confidence value to the pixels with low confidence.

5.   RESULTS AND EVALUATION

The proposed algorithm is compared with four different matting techniques; (1) Closed form, (2) Robust, (3) Learning based and (4) Shared matting. To observe the effect of dimension reduction on the quality of estimated alpha matte an algorithm is implemented that compare the patches using simple SSD (Sum of Squared Difference).

5.1   Dataset

For Evaluation a benchmark image matting data set from [16] is used. The dataset contains different images having different level of complexity regarding matte estimation. The proposed work is evaluated on images containing long, thin and intricate boundaries to show the effectiveness of the method. The images also contain holes within the foreground object which pose a strong problem for matting algorithms. The dataset contain images having an approximate spatial resolution of 800x640.

5.2   Qualitative Evaluation

For qualitative evaluation two images with different level of difficulty are considered in Fig. 4. The first image has

sharp boundaries in the foreground region, other approaches have produced erroneous matte at the intricate boundaries while the proposed method has produced visually better alpha matte. Large transparent regions in the second image are considered very difficult to matte. Robust and Closed form and Learning based matting approaches have misclassified the boundary while Shared matting produced higher alpha values in the transparent regions. The proposed technique produced a better quality matte with proper foreground boundary and classification of the thread around the bag.

5.3   Quantitative Evaluation

Ground truth along with the alpha mattes obtained using different algorithms are shown in Fig. 5. The first image contains holes and sharp boundaries in the foreground region.

The proposed algorithm has produced better result while other techniques have failed to classify sharp boundaries and tend to fill foreground holes. The second image has complex texture in the background and detailed structure in the foreground. Closed-form and Learning based method performed poor at the intricate regions while Robust and Shared matting misclassify the pixels where colour overlapping is present. The proposed algorithm produces better results because it utilizes local image structure along with the colour information which allows it to produce better matte estimation for texture images. The third image contain intricate region like hair where other approaches tend to misclassify foreground and background pixels. The proposed technique generated better mattes as shown by the mean square error in Fig. 6, the images have approximate spatial dimensions of 800x640.
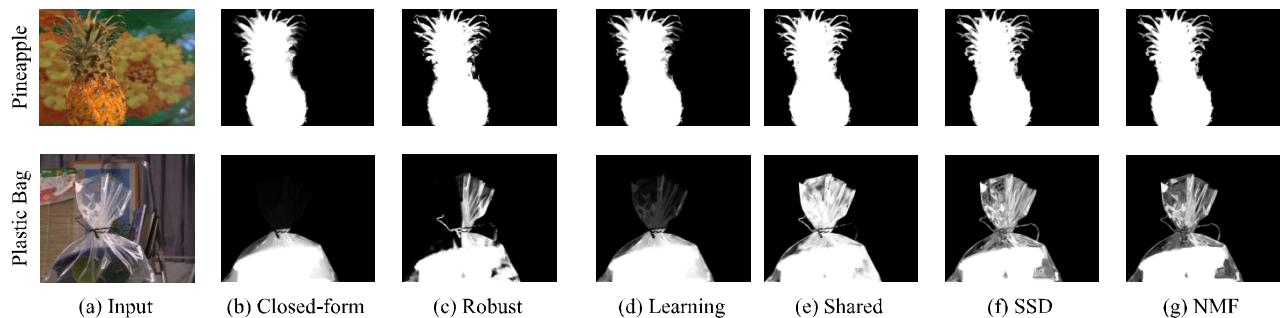

| (a) Input | (b) Closed-form | (c) Robust | (d) Learning | (e) Shared | (f) SSD | (g) NMF |

FIG. 4. VISUAL COMPARISONS OF ALPHA MATTES OBTAINED USING DIFFERENT APPROACHES [16]

| (a) Input | (b) Closed-form | (c) Robust | (d) Learning | (e) Shared | (f) SSD | (g) NMF | (h) Ground Truth |
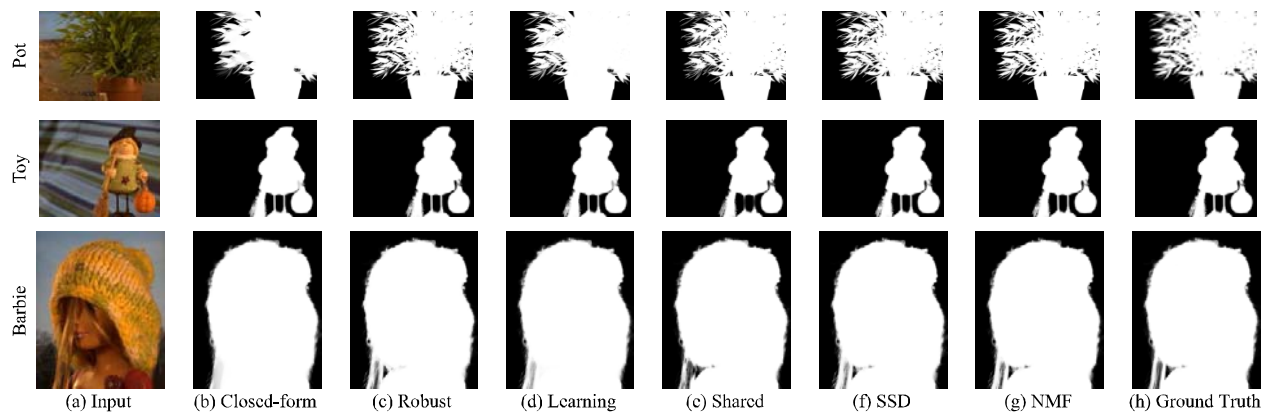
FIG. 5. COMPARISON ESTIMATED ALPHA MATTE USING DIFFERENT ALGORITHM [16]

Mehran University Research Journal of Engineering & Technology, Volume 36, No. 4, October, 2017 [p-ISSN: 0254-7821, e-ISSN: 2413-7219]
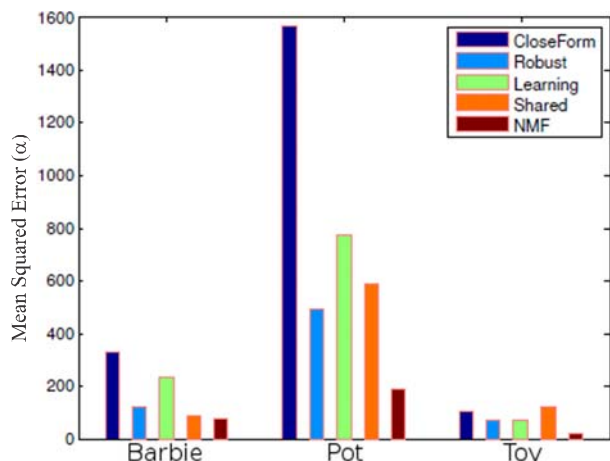
767

FIG. 6. MSE IN ALPHA OF DIFFERENT ALGORITHMS

6.   CONCLUSION

A global non-parametric sample based natural image matting using NMF is presented which not only consider the colour but also the local image structure to provide better alpha mattes. The comparison of proposed algorithm with simple SSD shows that the reduction in the dimension in search space using NMF results in efficient comparison with no visible lose to matte quality. The qualitative and quantitative comparison shows that the proposed method performs better than the state-of-the-art matting approaches especially for images that contain texture, intricate boundaries like hair and large transparent regions. The mean square error values suggest that the proposed approach estimates alpha matte with significant reduction in the overall error. In future, the use of weights in matrix factorization would be investigated to improve matte estimation. This would allow the algorithm to give different importance to different patches. The weights could be assigned based upon spatial distance and colour similarity of the patches.

ACKNOWLEDGEMENT

REFERENCES

[1]   Porter, T., and Duff, T., "Compositing Digital Images", ACM SIGGRAPH, pp. 253-259, New York, USA, 1984.

[2]   Sun, J., Jia, J., Tang, C.K., andShum, H.Y., "Poisson Matting", ACM Transactions on Graphics, Volume 23, No. 3, pp. 315-321, New York, USA, 2004.

[3]   Grady, L., Schiwietz, T., Aharon, S., and Westermann, R., "Random Walks for Interactive Alpha-Matting", VIIP, ACTA Press, pp. 423-429,Benidorm, Spain, 2005,

[4]   Levin, A., Lischinski, D., and Weiss, Y., "A Closed form Solution to Natural Image Matting", CVPR, pp. 61-68, New York, USA, 2006.

[5]   Ruzon, M.A.,and Tomasi, C.,"Alpha Estimation in Natural Images",CVPR, pp. 18-25, Hilton Head, SC, USA, 2000.

[6]   Chuang, Y.Y., Curless, B., Salesin, D.H., and Szeliski, R., "A Bayesian Approach to Digital Matting", CVPR, Volume 2, pp. 264-271,Kauai, HI, USA, 2001.

[7]   Hillman, P., and Hannah, J., "Natural Image Matting", Video, Vision and Graphics, pp. 211-18, Edinburgh, Scotland 2005.

[8]   Kun, W., Zheng, N.N., and Liu, W.X., "Natural Image Matting with Nonnegative Matrix Factorization", ICIP, pp. 1186-1189, Genoa, Italy, 2005.

[9]   Wang, J., Cohen, and M. F., "Optimized Color Sampling for Robust Matting", CVPR, pp. 1-8, Minneapolis, Minnesota, USA, 2007.

[10]   Rhemann, C., Rother, C., and Gelautz, M., "Improving Color Modeling for Alpha Matting", BMVC, Leeds, UK, 2008.

[11]   Sarim, M., Hilton, A., Guillemaut, J.-Y., and Kim, H., "Non-Parametric Natural Image Matting", International Conference on Image Processing ICIP, Cairo, Egypt, 2009.

[12]   Eduardo S.L.G., and Oliveira, M.M., "Shared Sampling for Real-Time Alpha Matting", Proceedings of Eurographics Computer Graphics Forum, Volume 29, No. 2, pp.575-584, Pisa, Italy, May, 2010.

[13]   Lee, D.D., andSeung, H.S., "Learning the Parts of Objects by Nonnegative Matrix Factorization", Nature, Volume 401, No. 6755, pp. 788-791, USA, 1999.

[14]   Berman, A.,Vlahos, P., and Dadourian, A., "Comprehensive Method for Removing from an Image the Background Surrounding a Selected Object", US Patent, Volume 6, pp. 134,345, USA, 2000.

[15]   Wang, Y., and Zhang, Y. "Non-Negative Matrix Factorization: AComprehensive Review", IEEE Transactions on Knowledge and Data Engineering, Volume 99, No. Pre-Prints, USA, 2012.

[16]   Rhemann,C., Rother, C.,Wang, J.,Gelautz,M., Kohli, P., and Rott, P., "A Perceptually Motivated Online Benchmark for Image Matting", CVPR, pp. 1826-1833, Miami, Florida, 2009.