Improving Image Matting using Comprehensive Sampling Sets

Ehsan Shahrian[1], Deepu Rajan[1], Brian Price[2] and Scott Cohen[2]
[1]Nanyang Technological University, [2]Adobe Research
ehsa0004@e.ntu.edu.sg, asdrajan@ntu.edu.sg, bprice@adobe.com, scohen@adobe.com

Abstract

In this paper, we present a new image matting algorithm that achieves state-of-the-art performance on a benchmark dataset of images. This is achieved by solving two major problems encountered by current sampling based algorithms. The first is that the range in which the foreground and background are sampled is often limited to such an extent that the true foreground and background colors are not present. Here, we describe a method by which a more comprehensive and representative set of samples is collected so as not to miss out on the true samples. This is accomplished by expanding the sampling range for pixels farther from the foreground or background boundary and ensuring that samples from each color distribution are included. The second problem is the overlap in color distributions of foreground and background regions. This causes sampling based methods to fail to pick the correct samples for foreground and background. Our design of an objective function forces those foreground and background samples to be picked that are generated from well-separated distributions. Comparison on the dataset at and evaluation by www.alphamatting.com shows that the proposed method ranks first in terms of error measures used in the website.

1. Introduction

Accurate extraction of a foreground object from an image is known as alpha or digital matting. It has a fundamental role in image and video editing operations. The process is mathematically modeled by considering the observed color of a pixel as a combination of foreground and background colors using the compositing equation given by

$$I_z = \alpha_z F_z + (1 - \alpha_z) B_z, \tag{1}$$

where F_z and B_z are the foreground and background colors of pixel z that are linearly combined using α_z to represent its observed color I_z. The opacity parameter α takes values in the range $[0, 1]$ with pixels having $\alpha = 1$ belonging to the foreground and those having $\alpha = 0$ belonging to the background.

The estimation of seven unknowns for each pixel from three compositing equations - one for each color channel - is highly ill-posed. Typically, matting approaches rely on constraints such as assumption on image statistics [10, 9] or the availability of a trimap to reduce the solution space. Trimaps partition the image into three regions - known foreground, known background and unknown regions that consist of a mixture of foreground (F) and background (B) colors. The trimaps could be drawn by the user or generated automatically [16] or semi-automatically[8].

Current alpha matting approaches can be categorized into alpha propagation based and color sampling based methods. Alpha propagation based matting methods [10, 15, 6, 8, 2] assume that neighboring pixels are correlated under some image statistics and use their affinities to propagate alpha values of known regions toward unknown ones. A closed form solution for alpha matting is proposed in [10] by minimizing a quadratic cost function based on α. The assumptions of large kernels by [8] and local color line of [10] are relaxed in KNN matting [2] using nonlocal principles and K nearest neighbors.

Color sampling based methods collect a set of known foreground and background samples to estimate alpha values of unknown pixels. Different combinations of spatial, photometric and probabilistic characteristics of an image are used [5, 18] to find the known samples that best represent the true foreground and background colors of unknown pixels. Once the best known foreground and background samples are selected for pixel z, its alpha value is computed as

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

This approach can be further sub-divided into parametric and non-parametric methods. Parametric sampling methods like [3, 13, 16] usually fit parametric statistical models to the known foreground and background samples and then estimate alpha by considering the distance of unknown pixels to known foreground and background distributions. Non-parametric methods including [11, 1, 18, 5, 7, 14] simply collect set of known F and B samples to estimate alpha values of unknown pixels. However, the quality of the ex-

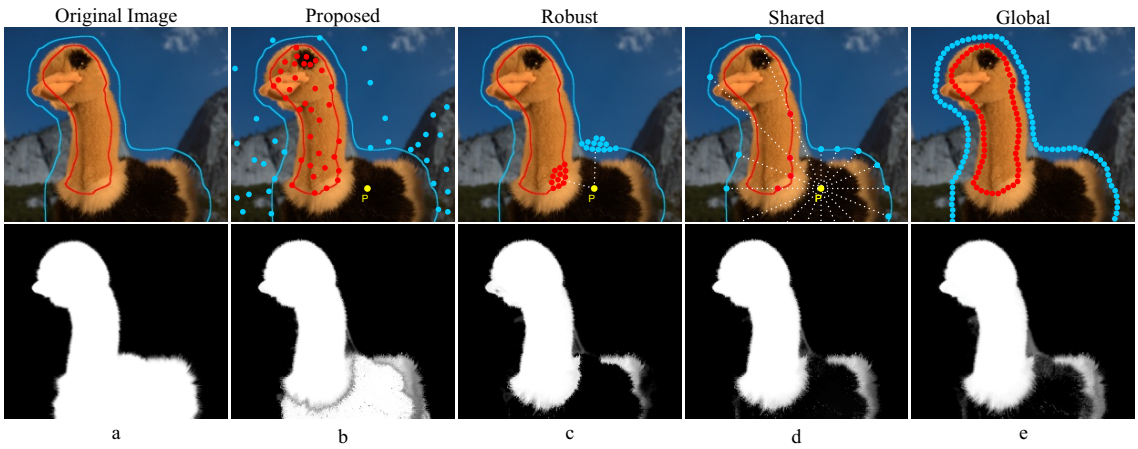

Figure 1. (a) Original Image and Sampling strategies of Proposed, Robust, Shared and Global matting methods. (b) Groundtruth matte and estimated mattes by Proposed, Robust [18], Shared [5] and Global matting [7] methods.

tracted matte is highly dependent on the selected samples. It degrades when the true foreground and background colors of unknown pixels are not in the sample sets. We call this the *missing true samples* problem. Hence, the challenge is to select a comprehensive set of known samples that encompass the different F and B colors in the image. A brief review of sampling strategies is presented next. A comprehensive review on image matting methods can be found in [17].

Mishima's [11] well-known blue screen matting method extracts a foreground object placed in front of a monochrome background (painted blue) by forming two clusters - one corresponding to the background and the other to the foreground and estimates alpha using the relative position of an unknown pixel to the clusters.

In the knockout system [1], known regions are extrapolated into unknown region and a weighted sum of known samples are used to estimate true foreground and background of unknown samples. The weights of known samples are proportional to their spatial distances to the unknown samples.

Robust matting [18] collects a few samples that are spatially close to the unknown pixel as shown in the first row of Fig. 1(c) in which the unknown pixel is shown in yellow and the foreground and background samples are shown in red and blue, respectively. The selection of best known background and foreground samples from the candidate set is done with respect to a color fitness parameter. It works better than the knockout system because only good samples that linearly explain the observed color of unknown pixels are used for matting. However, the quality of estimated mattes degrades when the true samples are not in the sets of known samples.

Shared matting [5] divides the image plane into disjoint sectors containing equal planar angles and collect samples that lie along rays that are emanated from unknown pixels as shown in the first row of Fig. 1.(d). It collects these samples from the boundaries of foreground and background regions as specified by a trimap. Different combinations of spatial, photometric and probabilistic characteristics of the image are used to find best samples for unknown pixels. The weighted color and texture matting [14] uses the same sampling approach. Here, too, if the desired sample for unknown pixels does not lie on the boundary, the extracted matte will not be accurate.

In order to avoid missing true samples, the largest set of known samples, among all other sampling based approaches, are built by collecting all known boundary samples in Global matting [7] as shown in the first row of Fig. 1(e). A simple cost function and an efficient random search are used to find the best samples among a huge number of known samples for every unknown pixel. Finally, the estimated alpha matte is refined by solving a global optimization problem. Once again, the true samples may still be missed if they are not on the boundary of the trimap from where the samples are collected.

The drawback of collecting samples only around the boundaries of known regions is illustrated in Fig. 1. The original image of doll with trimap boundaries of known F and B regions is shown in the first row of Fig. 1.(a). The doll has black and light brown colors but only light brown color samples are on or near the foreground boundaries. Thus, sets of collected known foreground samples by robust and shared matting methods do not contain black colors and therefore, the black region of the doll is wrongly estimated as background as shown in Fig. 1(c) and (d) in second row. This problem is further compounded due to the overlapping color distributions of foreground and background. More importantly, when all samples along the boundary are selected as shown in second row of Fig. 1(e) for global sampling, the matte is still inaccurate because foreground black color samples are inside the region and are excluded from the set

of candidate samples. Hence, it is important that the set of candidate samples should be comprehensive enough to represent all color variations in foreground and background regions.

In this paper a sampling method is presented that takes advantage of highly correlated boundary samples as well as samples inside of the F and B regions to build a set of comprehensive samples that covers a large range of diverse color distributions in the image, thus avoiding the missing true samples problem. A new objective function is proposed that contains measures of chromatic distortion, spatial and color statistics in the image. The color statistics helps especially in the case when foreground and background color distributions overlap leading to erroneous samples for F and B. Finally the estimated mattes are refined using conventional Laplacian approach. The selected samples according to the proposed method are shown in Fig. 1(b) in row 1 and the estimated matte is shown in second row of Fig. 1(b). The improvement in performance by selecting a comprehensive set of samples compared to other sampling strategies can be seen clearly. Later, we show in the experimental results that the proposed method obtains state-of-the-art performance on a benchmark dataset.

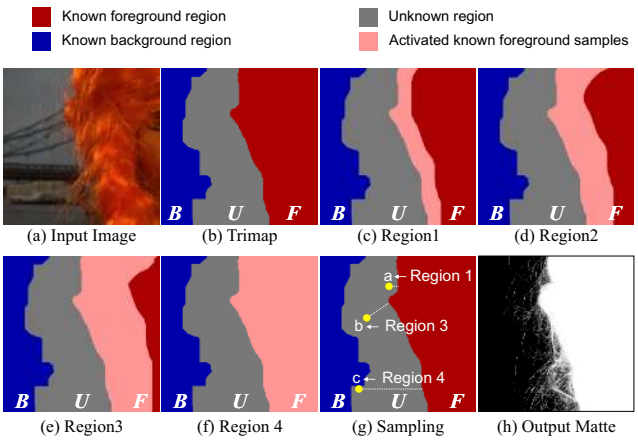

Figure 2. Illustration of sampling approach in foreground region. (a) Original Image, (b) Colored trimap showing foreground, background and unknown region, (c) - (f) Region 1, Region 2, Region 3 and Region 4 in foreground from where potential samples are obtained, (g) Unknown samples at varying distance from foreground. Each of them will receive known foreground samples from different regions based on its distance to the foreground, e.g., sample a receives from Region 1 while b receives from Region 3. (h) The generated alpha matte.

2. Proposed Method

In this section, we first describe how a comprehensive set of samples are generated followed by the process by which candidate samples are selected. Next, we illustrate the problem when the F and B color distributions overlap and finally formulate an objective function whose optimization leads to the true foreground-background (F, B) pair for an unknown sample.

2.1. Gathering comprehensive sample set

The goal of the sampling process is to find the best foreground/background combination to represent the color at a given pixel. We achieve this in two ways. First, the range over which samples are gathered is varied according to the distance of a given pixel to the known foreground and background. The motivation for this is that the closer an unknown sample is to known regions, the higher is the likelihood of a high correlation with known samples and thus known samples can estimate true samples robustly. This is intuitively obvious and indeed, previous sampling based methods have employed this criterion. However, they were constrained to use the known F and B samples that lie around boundaries of known region. By removing this restriction and instead adjusting the sampling range, we collect samples near boundaries as well as from inside F and B regions generating a more comprehensive sample set. Second, we ensure that every color distribution in a pixel's sampling range is represented in the sample set. Together, these two ideas allow our method to avoid missing true samples and to generate more accurate foreground/background samples.

The trimap is divided into regions to obtain a set of known F and B samples which form foreground-background pairs for an unknown pixel. The process is best explained with the help of Fig. 2 which shows region partitioning for foreground regions only. A similar strategy is used for background regions. Fig. 2(a) shows part of an original image whose trimap consisting of background, foreground and unknown regions labeled as B, F and U is shown in Fig. 2(b). The foreground region is divided into four regions (for illustration only), labeled as activated known foreground samples, as shown in Fig. 2(c)-(f). The width of the regions increases going from (c) to (f) as each region subsumes the previous regions. For instance, the activated region in Fig. 2(e) includes the regions in Fig. 2(d) and Fig. 2(c). Two questions arise: (i) how many such regions are needed? and (ii) what is the width of each region? The number of regions is determined by the size of the foreground region; however, we need as many regions as to cover the entire foreground as seen in Fig. 2(f).

The widths of the regions follow an incremental sequence starting from the region closest to the boundary. This is because for an unknown pixel that is close to the boundary, it is usually true that the correlation would be highest with pixels in a narrow region close to the boundary. However, the farther an unknown pixel is from the boundary, it is not possible to establish correlation with any narrow band of pixels. Hence, we would like to cover a larger

area to capture as much representative samples as possible. Thus, the widths of the regions is narrowest at the boundary, but increases as the regions cover more and more of the foreground/background region.

For each region, a two-level hierarchical clustering is applied. In the first level, the samples are clustered with respect to color through Gaussian mixture models (GMM) in which the number of components of the GMM is same as number of peaks in the color histogram of samples in the region. In the second level, the same clustering process is applied on samples of each cluster but with respect to spatial index of pixels. The mean value of the color in each cluster at the second level constitutes the set of candidate samples in each region. Thus, we obtain a comprehensive sample set that includes samples from all color distributions thereby handling the missing samples problem.

Our approach of parametrically determining the sample set differs from most current methods that sample non-parametrically. We observe that constructing a comprehensive sampling set covering all possible foreground and background colors is more important than whether that set was constructed parametrically or non-parametrically.

2.2. Choosing candidate samples

Each pixel in the unknown region collects a set of candidate samples that are in the form of a foreground-background pair. Once again, in the following discussion, we consider only foreground region to illustrate the idea. Consider the yellow pixel labeled a in Fig. 2(g). Since it is close to the foreground region, its candidate pixels will come from the region closest to the boundary of the foreground, viz., the region in Fig. 2(c) marked in pink. This is because the color correlation of the unknown pixel is likely to be the highest with pixels in this region. Thus, the foreground candidate samples for pixel a are the means of the clusters generated in the corresponding region, as discussed in the previous subsection. However, the pixel marked b is further away from the boundary and hence, would need a larger collection of foreground candidate samples. This is obtained as the means of clusters of pixels in the pink region of Fig. 2(e). Finally, the pixel marked c is very far from the boundary and the entire foreground region is utilized to generate foreground samples for this pixel.

The above discussion about Fig. 2 considered only the foreground region. Using exactly the same method, background candidate samples can be obtained for the unknown pixels. The candidate samples for each unknown pixel is in the form of a foreground-background pair (F, B). For example, if an unknown pixel obtains 4 candidate foreground samples and 3 candidate background samples, then the total number of candidate (F, B) pairs is 12. The next task is to choose the best (F, B) from among the candidate samples.

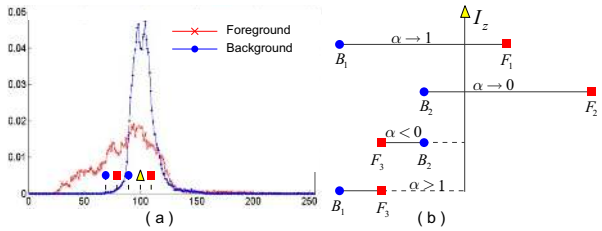

Figure 3. Illustration of the problem of sample selection when foreground and background have similar color. (a) Color distribution of foreground and background regions, (b) Effect of overlapped distribution on α.

2.3. Handling overlapping color distributions

In addition to selecting a representative set of candidate samples, the proposed method also addresses the problem encountered in current sampling based matting methods that involves overlapped color distributions of foreground and background regions. This problem was addressed in [14] by considering texture feature in addition to color. Here, we do not introduce any new feature; instead, we show that the overlapping problem can be alleviated to a large extent by considering only color. The problem is illustrated in Fig. 3(a) where the overlapped color distributions of known foreground and background regions which generated (F, B) pairs are shown in red and blue. Four pairs (F_1, B_1), (F_2, B_2), (F_3, B_2) and (F_3, B_1) are collected to estimate α of the unknown pixel z as shown in Fig. 3(b). The estimated alpha values show that the pixel is considered as foreground by (F_1, B_1) and as background by (F_2, B_2). It is hard to find the best pair because the pixel is close to F_1 and B_2 in color space and the alphas estimated by other pairs are out of the range $[0, 1]$. The question, therefore, is which pair can accurately estimate α? We propose that the (F, B) pair that is generated from the least overlapping color distributions of a foreground and background cluster should be selected. The objective function that we describe in the next section contains a term specifically to address the overlapped color distributions.

2.4. Selection of best (F, B) pair

Once the set of candidate (F, B) pairs is determined for unknown pixels, the task is to select the best pair that can represent the true foreground and background colors and estimate its α using eq. (2). The selection is done through a brute-force optimization of an objective function based on photometric and spatial image statistics. It consists of three parts as follows:

$$O_z(F_i, B_i) = K_z(F_i, B_i) \times S_z(F_i, B_i) \times C_z(F_i, B_i) \quad (3)$$

where K indicates chromatic distortion and S and C contain spatial and color statistics of the image. Next, we discuss each of these parameters.

K: The linear model of compositing equation successfully explains the color of a pixel as a convex combination when the estimated color is close to its observed color. Hence, K accounts for chromatic distortion. For a certain (F, B) pair, the estimated color of the unknown pixel is obtained by the compositing equation. The distortion between estimated color and observed color is given by

$$K_z(F_i, B_i) = \exp\left(-\left\|I_z - (\alpha F_i + (1-\alpha)B_i)\right\|\right) \quad (4)$$

where I_z is observed color of pixel z. It has a high value for (F, B) pairs whose estimated color is close to the observed color.

S: This term involves the distance between an (F, B) pair and unknown pixel in the spatial domain and favors spatially close pairs.

$$\begin{aligned} S_z(F_i, B_i) = &\exp\left(\frac{-\left\|z - F_i^s\right\|}{\frac{1}{|S_z^F|}\sum_{F_k \in S_z^F}\left\|z - F_k^s\right\|}\right) \\ &\times \exp\left(\frac{-\left\|z - B_i^s\right\|}{\frac{1}{|S_z^B|}\sum_{B_k \in S_z^B}\left\|z - B_k^s\right\|}\right) \end{aligned} \quad (5)$$

where S_z^F is set of known foreground samples for pixel z and $\left|S_z^F\right|$ is its cardinality. The spatial coordinates of foreground sample F_i is shown by F_i^s. The Euclidean spatial distance between the pixel and foreground sample F_i is indicated by $\|z - F_i^s\|$ that shows spatial cost required for F_i to reach the unknown pixel.

C: This term is biased towards those (F, B) pairs that come from well separated distributions and is formulated as

$$C_z(F_i, B_i) = \frac{d(F_i, B_i)}{M_z} \quad (6)$$

where $d(F_i, B_i)$ is Cohen's d value of color distributions [4] that generated F_i and B_i, M_z is a scaling factor which is the maximum Cohen's d value among the set of (F, B) pairs for pixel z. To reiterate, the color distributions are of the pixels within a cluster in the foreground/background region. Cohen's d value is inversely proportional to the overlap of distributions; it is high for distributions that are well separated and is computed as

$$d(F_i, B_i) = \frac{\mu_{F_i} - \mu_{B_i}}{\sqrt{\frac{(N_{B_i}-1)\sigma_{B_i}^2 + (N_{F_i}-1)\sigma_{F_i}^2}{N_{B_i}+N_{F_i}-2}}} \quad (7)$$

where μ_{F_i}, $\sigma_{F_i}^2$ and N_{F_i} are the mean, variance and population size of the distribution that generated sample F_i.

2.5. Pre and Post-processing

The proposed method uses a pre-processing step to expand known regions to unknown regions according to the following condition: An unknown pixel z is considered as foreground if, for a pixel $q \in F$,

$$(D(z,q) < E_{thr}) \wedge (\|I_z - I_q\| \leq (C_{thr} - D(z,q)) \quad (8)$$

where *D(z,q)* is the Euclidean distance between pixels z and q in spatial domain and E_{thr} and C_{thr} are threshold in spatial and color spaces which are empirically set as 9 in our experiments. A similar formulation is applied to compare the unknown pixel with a background pixel.

The alpha matte obtained by estimating α for each pixel using the best (F, B) pair in eq. (2) is further refined to obtain a smooth matte by considering correlation between neighboring pixels. In particular, we adopt the post-processing method of [5] where a cost function consisting of the data term $\hat{\alpha}$ and a confidence value f together with a smoothness term consisting of the matting Laplacian [10] is minimized with respect to α. The confidence value is the value of the objective function in eq.(3) for the selected (F, B) pair. The cost function is given by [5]

$$\begin{aligned} \alpha = \arg\min\ &\alpha^T L \alpha + \lambda(\alpha - \hat{\alpha})^T \Sigma(\alpha - \hat{\alpha}) \\ &+ \gamma(\alpha - \hat{\alpha})^T \hat{\Gamma}(\alpha - \hat{\alpha}) \end{aligned} \quad (9)$$

where λ is a large weighting parameter compared to the estimated alpha $\hat{\alpha}$ and its associated confidence f while γ is a constant (10^{-1}) that indicates the relative importance of data and smoothness terms. Σ is a diagonal matrix with values 1 for known foreground and background pixels and 0 for unknown ones, while diagonal matrix $\hat{\Gamma}$ has values 0 for known foreground and background pixels and f for unknown pixels.

3. Experimental Results

In the first experiment, the performance of the proposed matting method is evaluated on a benchmark dataset [12]. It consists of 8 images with three types of trimaps -small, large (coarse) and user defined and is available at *www.alphamatting.com*. The ground truth alpha mattes are hidden from the public and an independent quantitative evaluation is provided in terms of the mean squared error (MSE), the gradient error and the sum of absolute difference (SAD). Next, we illustrate the effectiveness of the proposed sampling method in alleviating the problem of missing true samples. Finally, we evaluate the performance on the dataset described in [14], which contains images with significant overlap in color distributions of foreground and background; we show that the proposed method outperforms other color sampling based methods.

3.1. Evaluation on benchmark dataset

Table 1 shows the quantitative evaluation of the proposed matting method when compared to current matting methods

Table 1. Evaluation of Matting methods by alpha matting website over set of bench mark images with three trimaps with respect to SAD , MSE and Gradient errors .

	Sum of Absolute Differences				Mean Square Error				Gradient Error			
	overall rank	avg. small rank	avg. large rank	avg. user rank	overall rank	avg. small rank	avg. large rank	avg. user rank	overall rank	avg. small rank	avg. large rank	avg. user rank
1. Proposed Method	**4.8**	4	4.8	5.8	**4.8**	4.3	4.6	5.4	**4.7**	4.5	4.5	5.1
2. SVR Matting	**5.7**	7	5.3	4.9	**5.4**	7.3	4.3	4.8	**5.5**	6.3	6.1	4.1
3. Weighted Color and Texture Matting	**6**	5	7	6	**6.7**	5.6	7.5	6.9	**9.3**	8.8	8.9	10.4
4. Shared Matting	**6.8**	6.6	8.3	5.4	**7.6**	7.4	9	6.4	**6.2**	6	6.4	6.3
5. Global Sampling Matting	**8**	6.1	9.5	8.5	**6.8**	4.1	8.9	7.4	**6.3**	6.3	6.4	6.3
6. Segmentation-based matting	**8.5**	9	7.9	8.6	**8.7**	9.1	8.3	8.6	**6.4**	7.9	5.6	5.8
7. Fast Automatic Matting	**8.5**	7.9	9	8.8	**9.8**	9.9	10.1	9.5	**11.1**	11.6	12	9.8
8. Improved color matting	**8.9**	8.8	8.4	9.6	**8.6**	8.9	9	8	**6.7**	7.4	6.5	6.1
9. Local Spline Regression (LSR)	**9.7**	11.3	7.5	10.4	**10.7**	11.5	9.5	11	**11.7**	12.5	11.3	11.3
10. Global Sampling Matting (filter version)	**10**	9.3	10.6	10	**11.8**	10.8	12.3	12.4	**9.8**	9.4	9.4	10.8

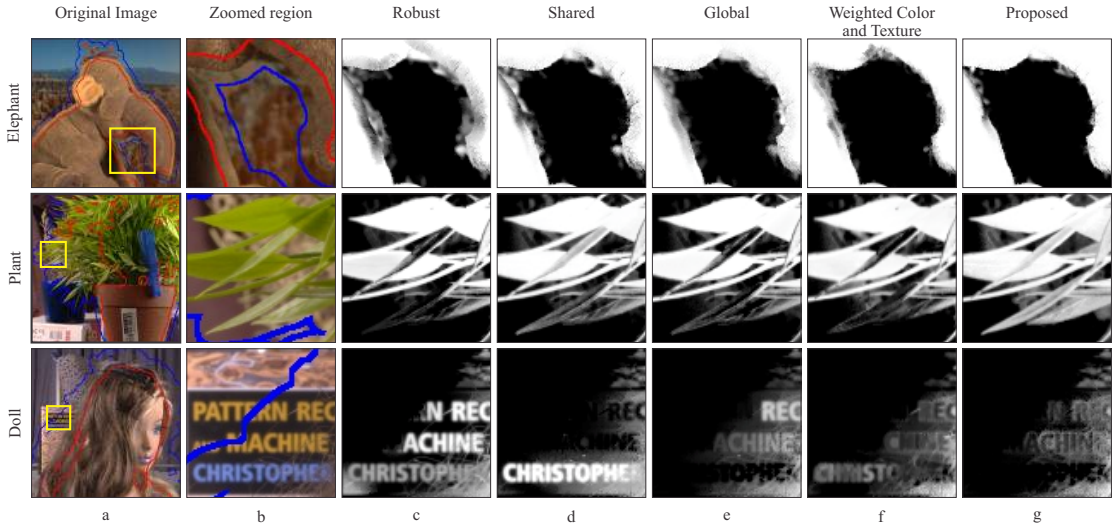

Figure 4. Visual Comparison of matting methods. (a) Original Image, (b) Zoomed area. Estimated mattes by (c) Robust [18], (d) Shared [5], (e) Global sampling [7], (f) Weighted color and texture [14] and (g) proposed method.

as evaluated by the alpha matting website. Only the top 10 methods are shown in the table. "Average small/large/user ranks" refers to the average rank over images for each of the three types of trimaps. The overall rank is the average over all the images and for all the trimaps. The proposed method performs best with overall ranks of 4.8, 4.8 and 4.7 with respect to SAD, MSE and gradient errors, respectively. SVR matting [19] , Weighted Color and Texture [14], Shared [5] and Global [7] matting methods have SAD ranks of 5.7, 6, 6.8 and 8 respectively. Visual comparison of the proposed method with some other sampling based matting methods are shown in Fig. 4. Original images and zoomed areas are shown in Fig. 4(a) and (b), respectively. The estimated mattes for zoomed areas by sampling based matting methods of Robust [18], Shared [5], Global sampling [7] and Weighted Color and Texture [14] are shown in Fig. 4(c-f). The Elephant (first row) has similar color as background, which makes it hard for color sampling based methods (Robust, Shared and Global) to discriminate between foreground and background as shown Fig. 4(c,d,e). Using only boundary samples makes it hard for Robust, Shared and Global Matting method to estimate true foreground colors for plant's leaves in unknown region. The problem is same for Doll image in last row in which some characters on the book are considered as foreground as shown in Fig. 4(c,d,e). The problem of overlapped color distribution is compensated in [14] using texture. However, it uses a sampling strategy similar to shared matting and hence, it still suffers from missing true samples problem as seen in Fig. 4(f) for Plant and Doll images.

The proposed methods takes advantage of comprehensive sampling to cover all true samples and also selects the best foreground and background pairs that are generated from well-separated distributions. These two characteristics of the algorithm help to generate a more accurate matte as shown in Fig. 4(g). Moreover, the standard deviation of matting methods over three types of trimaps on the set of benchmark images with respect to SAD is computed. The proposed method achieves the lowest overall standard deviation rank, 7.625, among more than 25 matting methods on the site. In comparison, the overall standard deviation ranks of Shared [5], Global sampling [7], Robust [18], Closed form [10] and weighted color and texture [14] methods are 8.375, 12.25, 17.5, 15 and 8.75 respectively which indicates that the quality of estimated mattes by the proposed method has less dependency on the trimap.

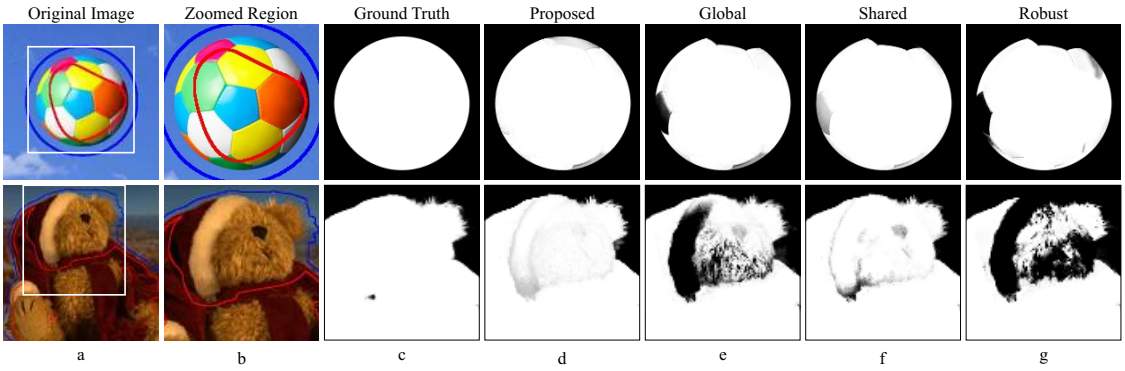

Figure 5. Illustration of missing true samples. (a) Original Images.(b) Zoomed region, (c) Ground truth matte. Estimated mattes by (d) Proposed, (e) Global [7], (f) Shared [5] and (g) Robust [18] matting methods.

3.2. Missing true samples

Current sampling based matting methods fail to estimate the true foreground and background colors of pixels when the set of collected samples do not contain the true colors. Fig. 5(a) shows two original images with their corresponding boundaries of known F and B regions. Zoomed regions and ground truth alpha mattes of zoomed regions are shown in Fig. 5(b) and (c), respectively.

The boundary of the foreground in the ball image does not contain blue colors of the ball. Because of this, the set of samples collected by the global, shared and robust matting methods cannot sample from the blue color distribution and are unable to estimate the true foreground colors of these parts of the ball. They are wrongly estimated as background as shown in Fig. 5(e),(f) and (g). A similar situation arises in the second image whereby global, shared and robust matting miss the true black and white colors in the foreground as shown in Figs. 5(e), (f), and(g), respectively. The proposed method uses samples inside of known regions to complement the set of highly correlated boundary samples to solve the problem of missing true samples by sampling from all color distributions. The visual comparison between ground truth mattes and estimated ones by the proposed method in Fig. 5(c) and (d) shows that the collected set of samples by proposed method is comprehensive enough to solve the problem of missing true samples.

3.3. Effect of overlapping distributions

The effectiveness of the proposed objective function containing *only* color information in reducing the effect of overlapping distributions is shown by comparing proposed method with Shared [5], Robust [18], Closed form [10] and Weighted Color and Texture [14] matting methods on the dataset described in [14]. Three of ten images with zoomed parts are shown in Fig. 6(a) and (b). The color similarity between foreground and background in the images is illustrated in Fig. 6(c) in which histograms are computed from the red channel of the images.

In the first row, the background texture contains colors similar to the leaves making it hard for sampling based methods to find the true samples as shown in Fig. 6(e) and (f) for Robust and Shared Matting. False correlations are increased due to color similarity for closed form matting in Fig. 6(d). The strong edges intensify the problem by blocking the propagation of alpha. Weighted color and texture matting, which was developed precisely to address this problem, uses 3×3 window to obtain texture information to complement color in matting. However, when texture information is not captured properly as shown in Fig. 6(g), the performance degrades considerably. Similar comments can be made about the Doll image. For the flower image, the performance of weighted color and texture matting is better, probably because the texture is not as strong as in the other two images. The estimated mattes by our proposed method are shown in Fig. 6(h), which are comparable to the ground truth shown in Fig. 6(i). It takes advantage of an effective objective function to find the best known samples even when they come from overlapped color distributions. Of course, if the overlap is to such an extent that the colors are not distinguishable, then color sampling based methods fail and we would need to resort to additional features such as texture. Even so, the feature should be able to capture the texture variation accurately. This may require additional parameters such as the size of the window over which the texture feature is captured. In this paper, we show the effectiveness of using only color information.

4. Conclusion

A new sampling based image matting method is proposed that uses a new sampling strategy to build a comprehensive set of known samples by sampling from all color distributions in known regions. This set includes highly correlated boundary samples as well as samples inside the F and B regions to capture all color variations and solve the problem of missing true samples. Moreover the problem of overlapping color distributions of foreground and back-

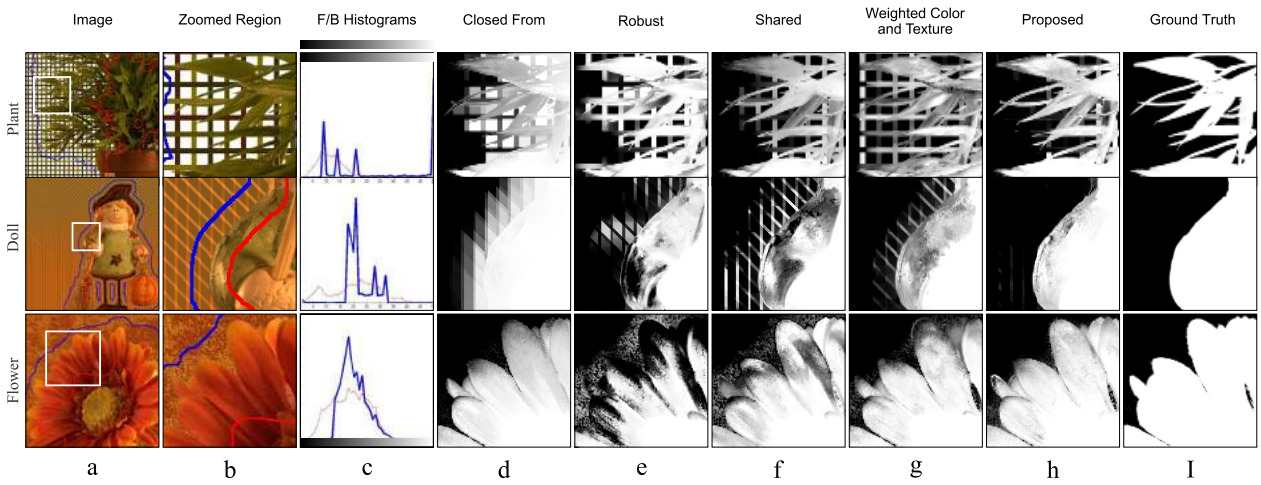

Figure 6. Visual comparison of matting methods on the dataset of [14] to illustrate cases when foreground and background color distributions overlap. (a) Original image, (b) Zoomed area, (c) Foreground and background color distributions on red channel, (d) Closed form [10], (e) Robust [18], (f) Shared [5], (g) Weighted color and texture [14], (h) Proposed method, (i) Ground truth.

ground are handled by using an effective objective function that chooses the best (F, B) pair by choosing those samples that were generated from well-separated distributions. Finally, the quality of the estimated mattes is improved using conventional Laplacian refinement. Experimental results shows that the proposed method achieves state-of-the-art performance in terms of standard error measures on a benchmark dataset.

References

[1] A.Berman, A.Dadourian, and P.Vlahos. Method for removing from an image the background surrounding a selected object, 2000. U.S Patent 6,134,346,.

[2] Q. Chen, D. Li, and C. Tang. Knn matting. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 869–876, 2012.

[3] Y. Chuang, B. Curless, D. Salesin, and R. Szeliski. A bayesian approach to digital matting. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 7–15, 2001.

[4] J. Cohen. *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum, 1988.

[5] E. Gastal and M. Oliveira. Shared sampling for real time alpha matting. In *Proc. Eurographics , 2010*, volume 29, pages 575–584, 2010.

[6] L. Grady. Random walks for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(11):1768–1783, 2006.

[7] K. He, C. Rhemann, C. Rother, X. Tang, and J. Sun. A global sampling method for alpha matting. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2049–2056, 2011.

[8] P. Lee and Y. Wu. Nonlocal matting. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2193–2200, 2011.

[9] A. Levin, R. Acha, and D. Lischinski. Spectral matting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(10):1699–1712, 2008.

[10] A. Levin, D. Lischinski, and Y. Weiss. A closed-form solution to natural image matting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(1):228–242, 2007.

[11] Y. Mishima. Soft edge chroma-key generation based upon hexoctahedral color space, 1994. U.S. Patent 5,355,174.

[12] C. Rhemann, C. Rother, J. Wang, M. Gelautz, P. Kohli, and P. Rott. A perceptually motivated online benchmark for image matting. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1826–1833, 2009.

[13] M. Ruzon and C. Tomasi. Alpha estimation in natural images. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 18–25, 2000.

[14] E. Shahrian and D. Rajan. Weighted color and texture sample selection for image matting. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 718–726, 2012.

[15] J. Sun, J. Jia, C. Tang, and H. Shum. Poisson matting. *ACM Transactions on Graphics (TOG)*, 23(3):315–321, 2004.

[16] J. Wang and M. Cohen. An iterative optimization approach for unified image segmentation and matting. In *Proc. International Conference on Computer Vision*, volume 2, pages 936–943, 2005.

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

[18] J. Wang and M. Cohen. Optimized color sampling for robust matting. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, 2007.

[19] Z. Zhanpeng, Z. Qingsong, and X. Yaoqin. Learning based alpha matting using support vector regression. In *IEEE Conference on Image Processing (ICIP)*, 2012.