Consistent Matting for Light Field Images

Donghyeon Cho, Sunyeong Kim, and Yu-Wing Tai

Korea Advanced Institute of Science and Technology (KAIST)

Abstract. We present a new image matting algorithm to extract consistent alpha mattes across sub-images of a light field image. Instead of matting each sub-image individually, our approach utilizes the epipolar plane image (EPI) to construct comprehensive foreground and background sample sets across the sub-images without missing a true sample. The sample sets represent all color variation of foreground and background in a light field image, and the optimal alpha matte is obtained by choosing the best combination of foreground and background samples that minimizes the linear composite error subject to the EPI correspondence constraint. To further preserve consistency of the estimated alpha mattes across different sub-images, we impose a smoothness constraint along the EPI of alpha mattes. In experimental evaluations, we have created a dataset where the ground truth alpha mattes of light field images were obtained by using the blue screen technique. A variety of experiments show that our proposed algorithm produces both visually and quantitatively high-quality matting results for light field images.

Keywords: Image Matting, Light field image, EPI.

1 Introduction

Image Matting aims to extract a soft and accurate alpha matte of foreground given a trimap of an image. Generally, colors of an image can be expressed as a linear combination of foreground and background colors as follows:

$$I = \alpha F + (1 - \alpha)B, \tag{1}$$

where F, B and α represent the foreground, the background, and the mixing coefficients, respectively. Since most matting algorithms were developed for matting a single image, it is less effective when facing multiple input images, e.g. multiple sub-images of a light field image, where consistent alpha mattes across the multiple images are necessary. In this paper, we introduce a new image matting algorithm targeting for a light field image.

A light field image consists of $m \times n$ sub-images where each sub-image was captured from slightly different perspectives. The correlation among the sub-images are encoded in the epipolar plane image (EPI), and the estimated alpha mattes across sub-images also need to follow the EPI constraint. Otherwise flickering artifacts will appear when moving an interpolated view point from one sub-image to another sub-image.

D. Fleet et al. (Eds.): ECCV 2014, Part IV, LNCS 8692, pp. 90–104, 2014.

By Equation (1), image matting is an ill-posed problem because the number of unknowns is more than the number of equations that can be derived from a single image. State-of-the-art matting algorithms can be categorized into two groups: color sampling based, and alpha propagation based methods. The color sampling based methods [7,28,21,11,13,24,23] sample foreground and background colors from the known regions, *i.e.*the definite foreground and the definite background regions, to estimate alpha mattes within the unknown region. The alpha propagation based approaches [26,17,32,14,18,16,3,4] assumes local/nonlocal smoothness of alpha values and propagate alpha values from the known regions to the unknown regions.

In the light field image matting problem, although the number of input images have increased, the number of unknown have also increased which makes it also an ill-posed problem. However, because of the EPI correlation among the sub-images, we can sample foreground and background colors across sub-images. Even if a true color sample in a sub-image is missing, we can still reliably estimate the true color sample from another sub-image. This allows us to achieve better performance than existing color sampling matting techniques. In addition, using the EPI constraint, we can propagate alpha values not only from the known regions to the unknown regions within a sub-image, but also along EPI of alpha mattes across sub-images. This provides an accurate and consistent alpha estimation across sub-images. As demonstrated in our experimental results, our algorithm reduces weaknesses and maximize strengths of both kinds of image matting techniques.

We evaluate and compare performance of our proposed algorithm and state-of-the-art image matting algorithms. To quantitatively compare the performance, we created a new light field matting dataset where the ground truth alpha mattes are obtained by using the blue screen matting procedures introduced in [22]. Our evaluations show that our algorithm produces both visually and quantitatively high-quality matting results for light field images, and have outperformed existing matting algorithms in term of both accuracy and consistency.

2 Related Works

We review previous works that are the most relevant to our work. In particular, we discuss the works related to the two categories of image matting and the works related to light field image processing.

As aforementioned, most image matting techniques can be categorized into color sampling based and alpha propagation based methods. The color sampling based methods [7,28,21,11,13,24,23] solve the matting problem by finding color samples from the known foreground and background pixels to estimate alpha mattes in unknown regions. In [7], the Bayesian matting by Chuang *et al.* analyzes unknown pixels using local color distribution by statistical methods. Robust Matting [28] collects color samples with respect to the color composite equation and are spatially close to the unknown pixels. Shared matting [11] and weighted color and texture matting [24] find the best samples by combining spatial, photometric, and probabilistic information measured by color and texture,

respectively. In [13], He *et al.* proposed the global sampling matting which uses all color samples in the known regions to find the best combination of foreground and background samples for matte estimation. Recently, Shahrian *et al.*[23] proposed the comprehensive sampling matting which uses Gaussian Mixture Model (GMM) to cover all color variations in the foreground and background regions of an image for accurate alpha matte estimation.

The alpha propagation based approaches [26,17,32,14,18,16,3,4] analyze statistical correlation among pixels to propagate alpha values from the known regions to the unknown regions. The Poisson Matting by Sun *et al.* [26] estimates an alpha matte by solving a Possion equation to reconstruct an alpha matte from gradients subject to the boundary condition of alpha matte in the known regions. Levin *et al.* [17] introduced the color line model and propose the matting Laplacian to solve the matte estimation problem in a closed form. This work is later extended by He *et al.* [14] who proposed the large kernel matting for matting high resolution images. Lin *et al.* [18] introduced motion regularization for matting motion blurred objects. Based on the nonlocal principle, Lee and Wu [16] introduced the nonlocal matting which propagate alpha values across nonlocal neighbor of a pixel. This work is later extended by Chen *et al.* [3] who proposed the KNN matting to propagate alpha across the k nearest nonlocal neighbors, and by Chen *et al.* [4] who combined local and nonlocal smoothness prior for alpha propagation.

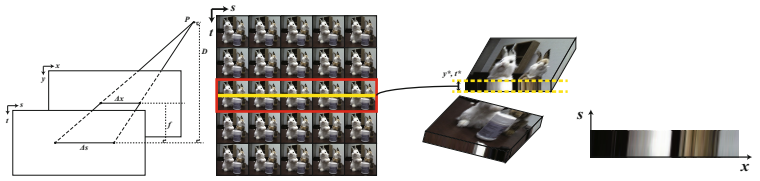
In light field image processing, since Ng *et al.* [20] introduce the prototype of micro lens array light field camera, a lot of follow up works have been proposed. The work by Dansereau *et al.* [9] estimates a depth map of the corresponding elements in a scene using gradient vector. Bishop and Favaro [1] estimate depth map by evaluating aliasing across multiple views. Wanner *et al.* [29,31] use epipolar plane images to estimate depth map with consideration of global and local consistent. The work by Goldluecke and Wanner [12] computes depth maps by using the local derivative constraint with a convex prior derived from a 4d light-field image. Recently, Wanner and Goldluecke [30] suggested a depth map reconstruction method for reflective and transparent surfaces through 4D light-field image analysis.

Comparing our work with previous works, as far as we are aware, this is the first work to seriously address the matting problem in light-field images. Utilizing the additional EPI information, we introduce a method to construct color sample sets across multiple sub-image, and to encode the EPI smoothness constraint in the matting Laplacian by introducing nonlocal smoothness term across multiple sub-images. To facilitate future research in light field matting, we have also created a dataset to quantitatively evaluate performance of matting algorithms.

3 EPI in Light Field Images

3.1 The EPI Constraint

A light field image is typically represented as a 4D function, $L(x, y, s, t)$, which records the intensity of a light ray passing through two parallel planes, $x - y$



(a) (b) (c)

Fig. 1. (a) Parallel plane representation of a light field image. (b) The multiple sub-images of a light field image after decoding. (c) After stacking the images within the yellow line region in (b), we have an epipolar plane image in $x-s$ plane with fixed y,t.

and $s-t$ planes, in a 3D space as illustrated in Figure 1(a). To capture a light field image, one can use a camera array or a consumer level micro lens light field camera, e.g. Lytro [19]. After decoding [10,5], we can obtain multiple sub-images where each sub-image represents image captured from slightly different perspective as illustrated in Figure 1(b).

Since a light field image captures light rays in a 3D space, a light ray from an object at different distance from a camera would pass through the two parallel image planes at different angle. This relationship is captured in the EPI of a light field image. For instance, if we fixed the index of y and t in $L(x, y, s, t)$, we can plot the EPI of $x-s$ plane as illustrated in Figure 1(c). Mathematically, we can derive [2,8,29]:

$$\Delta x = \frac{f}{D}\Delta s, \tag{2}$$

where f is the distance between the two parallel image planes and D is the distance of a object from a camera as illustrated in Figure 1(a). Using Equation (2), we can obtain pixel correspondences across sub-images in x-direction by measuring the image gradients in the EPI of $x-s$ plane. Similarly, we can obtain pixel correspondences in y direction through measuring the image gradients in the EPI of $y-t$ plane. Since colors of correspondent pixels across sub-images come from the same light ray in 3D, the estimated foreground/background colors as well as the alpha values are expected to be the same across the same EPI correspondents. This defines the EPI constraint across the multiple sub-images of a light field image.

3.2 Color Sample Correspondences in EPI

Using the EPI constraint, we can define pixel correspondences across the multiple sub-images and expect the color values as well as the alpha mattes along the EPI correspondences to be identical. In practice, because of the mixing effect in matting areas, the EPI constraint may not hold along the matting boundary since the measured intensity is the result of alpha blending of two light rays from different direction as illustrated in Figure 2(a).

In order to utilize the EPI constraint, we assume the EPI of foreground and background are spatially smooth in the unknown region of a trimap. Intuitively,

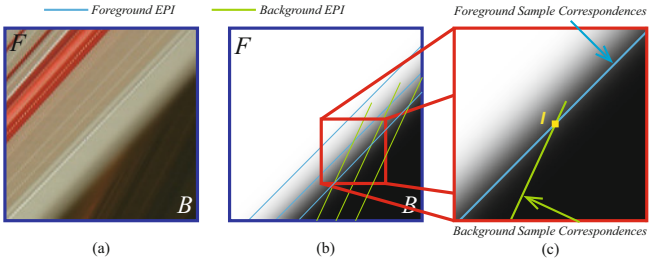


Fig. 2. (a) EPI of matting boundary. (b) Alpha matte of the EPI image. (c) Foreground and background color sample correspondences of a pixel in the matting area.

we assume the depth of foreground and background within the unknown region are similar to the depth of foreground and background in the known region. Thus, we can propagate the EPI constraint from the known region to the unknown region through extrapolation (details will be described in Section 4.2). Figure 2(c) illustrates the color sample correspondences of a pixel in a matting area using the propagated EPI constraint. Note that the foreground and the background color sample correspondences are defined differently since the foreground and background EPI have different light ray direction. In the next section, we will discuss how to use the EPI color sample correspondences to select color samples for better alpha matte estimation.

4 Consistent Matting for Light Field Images

4.1 Pre-processing: Color Samples Collection

In order to efficiently process color samples from the known regions, we follow the steps in [23] to construct a comprehensive color sample sets of foreground and background in a light field image. Specifically, we follow the two-level hierarchical clustering process by first using colors to estimate the global color distribution of foreground and background described by the Gaussian mixture models (GMM). In the second level, we include spatial index of pixels to estimate local color distribution of foreground and background. This provides us the comprehensive sample set which covering all possible foreground and background colors in a light field image. We have also followed the steps to expand trimap regions in color samples collection. This gives us a more accurate estimation of color distribution. We refer readers to [23] for further details of these steps in constructing the comprehensive sample set of foreground and background.

4.2 EPI Estimation and Propagation

Although EPI estimation is not the focus of this paper, the performance of our algorithm depends on the accuracy of estimated EPI since the EPI defines the pixel correspondences between sub-images of a light field image. In our implementation, we use the depth estimation method by [27] to estimate the EPI

in the definite foreground and the definite background regions. A median filter is applied to reduce effects of noise in the estimated EPI. In the unknown regions, we propagate the EPI from definite foreground and definite background regions by assuming the EPI in the unknown regions are spatially smooth. The propagation is achieved by solving the Poisson equation by setting zero gradient of EPI in the unknown regions, subject to the boundary constraint of the estimated EPI along the trimap, to extrapolate the EPI in the unknown regions. The extrapolation of foreground and background EPI are performed independently. Thus, two different set of EPI correspondences will be defined within the unknown regions as illustrated in Figure 2. We note that there are better EPI estimation algorithms such as the works by [31,12,30]. In experiments, we found that our EPI estimation and propagation algorithm provides sufficient accuracy for our examples. In implementation, the center view trimap is provided by user manually, and we use it to estimate EPIs in definite foreground and definite background regions, and propagate the EPI propagation to unknown regions in the center view. After that, the center view trimap is propagated to other sub-images automatically using the estimated foreground EPI.

4.3 Color Sample Selection

In previous sampling based matting algorithms, color samples are selected to minimize the linear composite error defined by the matting equation in Equation (1). A major challenge in this process is that there are multiple pairs of foreground and background samples that can minimize the error but the estimated alpha values can be totally different. Researches in sampling based matting algorithm have extensively focused on how to resolve this ambiguity by using different cues or making different assumptions about the true color samples. In this section, we describe how to resolve this ambiguity by using the EPI correspondence defined in the previous section.

Assumption. *We assume foreground and background are located at different depth from camera such that the foreground and the background color sample correspondences are misaligned as illustrated in Figure 2(c). If foreground and background are located at the same (or very closed) depth, the observed intensity across multiple sub-images will be identical. In such case, we will apply the method in [23] to select the optimal color samples based on color similarity, spatial distance, and distribution of sample sets. We will also assume the EPI correspondences are accurate.*

In a practical scenario, we have two different cases as illustrated in Figure 3 in handling the color sample selection problem:

Case 1: Background Samples along EPI of Foreground Are Known
This case happens when a background sample is partially occluded in one sub-image, but is disoccluded in another sub-image. The disoccluded background samples can be detected using the background color sample correspondence defined by the background EPI where pixels along background EPI are not entirely included in the unknown regions.

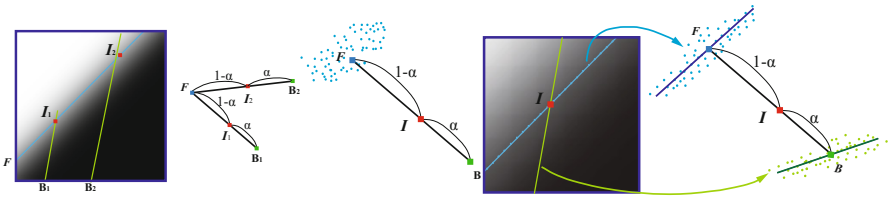


Fig. 3. (a) Case 1: If $B_1 \neq B_2$, we solve the alpha in a closed form. If $B_1 = B_2$, we solve the alpha using the comprehensive sample set. (b) Case 2: Foreground and Background samples are solved individually along its EPI, and the median F and B are selected to solve the alpha.

To estimate the color of a foreground sample with known background colors, we can derive multiple equations along the foreground color sample correspondence defined by the foreground EPI:

$$\begin{aligned} I_1 &= \alpha F + (1-\alpha)B_1, \\ &\vdots \\ I_n &= \alpha F + (1-\alpha)B_n, \end{aligned} \tag{3}$$

where $\{I_1, \ldots, I_n\}$ are the observed intensity along the foreground color sample correspondence, and $\{B_1, \ldots, B_n\}$ are the known background colors. Thus, we have number of equations more than or equal to the number of unknown when $n \geq 2$. When $n = 2$, we can obtain the alpha in a closed form:

$$\hat{\alpha} = 1 - \frac{I_1 - I_2}{B_1 - B_2}. \tag{4}$$

When $n > 2$, we solve the alpha using the least square error method by computing the weighted average α across the solution of all pairs of pixels in the foreground color sample correspondences with known background color. Weighting factors for each α are determined by the distance between two background color samples. Larger weight coefficient is given to α with longer distance between background color samples because it is more reliable to estimate α value than the inverse case when denominator is close to zero. With the estimated alpha, $\hat{\alpha}$ and the known background colors, B_i, the foreground color, F, can be computed accordingly. In order to avoid errors caused by image noise, we apply this method only when the number of disoccluded pixels is more than 4 along the EPI correspondence.

In a degenerated case when colors of all background samples are identical, e.g. homogeneous background color, we use the comprehensive sample set collected from the known foreground region to estimate the alpha. Specifically, the observed intensity I and the known background color B form a line constraint where the true foreground color F must be located along the line extrapolated from $I - B$. This line constraint guides the searching of true foreground color

sample from the comprehensive sample set. In the case when there are multiple foreground color samples that satisfied the line constraint, we choose the solution which produces the minimum differences of alpha value around the solution within the local neighborhood of a pixel.

Case 2: Background Samples along EPI of Foreground Are Not Known. When alpha matte area is large, background pixels closed to foreground region will be occluded/partially occluded in all sub-images. In this case, we apply the following method to estimate foreground and background samples. Again, we assume the foreground EPI and the background EPI are misaligned and they are accurately estimated.

For a pixel with different foreground and background color sample correspondence, we first compute the foreground and background sample pairs for each pixels along the foreground EPI and the background EPI independently. This process is done by using the method proposed in [23] to select the optimal color samples from the comprehensive sample set collected from the known foreground and background regions. Next, assuming that majority of the estimated color samples are correct, we apply a simple linear regression to fit a line to the estimated foreground and background samples. Note that when computing the foreground color, only the estimated foreground samples along the foreground EPI were used. This is the same case for the background color estimation. After fitting the color line, we sort the color samples along the estimated color line and choose the median foreground color and the median background color as the true foreground color and the true background color respectively. Once the true foreground and background colors are estimated, the alpha value can be computed as:

$$\hat{\alpha} = \frac{(I - B) \cdot (F - B)}{||F - B||^2}. \tag{5}$$

While this method is simple, we find that the method robust and reliable. In a degenerated case where all foreground and background colors along the two different EPIs are identical, this reduces the problem to the conventional setting of color sample estimation since the two EPIs do not provide additional information to assist the color sample selection.

Implementation: Solving Alpha in the Order to Reduce Ambiguity Since pixels are interconnected by the foreground and the background EPI, and each pixel in the unknown region have two different set of correspondences, once the foreground color of a pixel with known background samples (Case 1) is estimated, the estimated foreground color can be used to estimate background color of a pixel where its background samples are unknown (Case 2). Using this strategy, we can significantly reduces the ambiguity in color sample selection by reducing the Case 2 scenario to the Case 1 scenario using the estimated color as the known color samples. In our implementation, we use a priority queue to rank the pixels in the unknown region according to the number of known color samples along the foreground or background EPI. The priority of a pixel in the priority queue will be updated once color of a pixel along the EPI correspondence is solved. Intuitively, using this strategy, alpha value of pixels

are solved progressively from the boundary of background region towards the boundary of foreground region.

Due to the small angular resolution of the light-field images, propagation method on the EPI image has ambiguity. In order to reduce this ambiguity, we take multiple directional EPI images, 0°, 45°, 90° and 135° and average the estimated alpha using following confident weighting factor:

$$W_i = \frac{w_i}{\sum w_i}, \quad w_i = exp(-\frac{\|I - \alpha_i F_i - (1-\alpha_i)B_i\|_2^2)}{2\sigma}), \\ i = (0°, 45°, 90°, 135°) \tag{6}$$

where α_i, F_i, B_i and W_i are estimated α, foreground, background samples and confidence weighting at each direction respectively. By using this weighting strategy, more reliable and accurate alpha and color samples can be estimated.

4.4 Consistent Matting with the EPI Smoothness Term

The previous color sampling step estimates alpha value of each pixel independently although the selected color samples is guided by the EPI constraint. In this section, we describe the process to further improve the alpha matte by considering smoothness among neighboring pixels. This is also a common post-processing step in many previous matting algorithms.

Using the results, $\hat{\alpha}$, from the Equation (4) or Equation (5) as a data term, and the smoothness term defined by the matting Laplacian matrix L [17], we can obtain the final alpha by:

$$\alpha = \arg\min \alpha^T L \alpha + \lambda(\alpha - \hat{\alpha})^T D(\alpha - \hat{\alpha}) \tag{7}$$

where λ is a weighting parameter, and D is a diagonal matrix. Its diagonal element is a large constant for the known pixel, and a confidence $c = \exp(||I - (\hat{\alpha}F + (1-\hat{\alpha})B||^2/\sigma^2)$ for the unknown pixel.

In order to further consider the smoothness constraint along the EPI of the extracted foreground, we extend Equation (7) to include an additional nonlocal smoothness term in L, and solve the alpha matte of multiple sub-images simultaneously. In particular, we extend Equation (7) as follows:

$$\alpha = \arg\min \alpha^T \mathbf{L} \alpha + \lambda(\alpha - \hat{\alpha})^T \mathbf{D}(\alpha - \hat{\alpha}), \tag{8}$$

and

$$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}, \mathbf{L} = \begin{bmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{bmatrix}, \mathbf{D} = \begin{bmatrix} D_1 & 0 & 0 \\ 0 & D_2 & 0 \\ 0 & 0 & D_3 \end{bmatrix} \tag{9}$$

where L_{ii}, $i = \{1, 2, 3\}$, are the matting Laplacian of the sub-image I_i, and L_{ij}, $i \neq j$, are the cross sub-images smoothness term with each entry defined as:

$$L_{ij}(x, x') = \exp(-\frac{||I_i(x) - I_j(x')||^2}{2\sigma_c^2}),$$

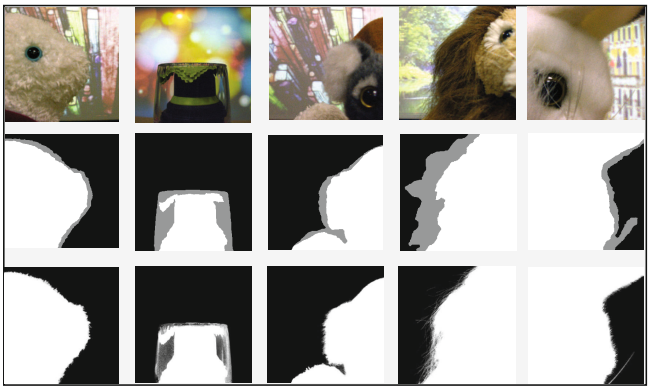


Fig. 4. Our dataset for testing

if x and x' are the foreground EPI correspondence between I_i and I_j, and $L_{ij}(x, x') = 0$ if otherwise. For a sub-image I_1, I_2 is the sub-image next to I_1 in horizontal direction, and I_3 is the sub-image next to I_1 in vertical direction. Thus, we can solve the alpha matte of three sub-images simultaneously with consideration of the EPI smoothness of alpha matte across the sub-images. Although we can solve the alpha matte of sub-images altogether by further extending Equation (9) to include more sub-images, the computation cost increases dramatically deal to the large linear system. In experiments, we found that using more adjacent sub-images does not improve much in accuracy. Thus, we only solve the alpha matte of three sub-images simultaneously.

5 Experimental Results

5.1 Light Field Matting Dataset

We follow the steps in [22] to create a new dataset to evaluate the performance of matting algorithms applied on light field images. In order to derive a high-quality ground truth alpha matte, we placed the matting objects in front of a monitor, and we displayed four different single-colored background (i.e. black, red, green, blue). The images are captured with a Lytro camera mounted on a tripod, and we use the method in [5] to decode the captured light field images from its RAW image data. This gives us 45 sub-images (7×7 without four corners) of a light field image. Each of the sub-image is of resolution 300×300 with 12-bits per color channel. With the different monochrome color background, we apply the blue screen matting [25] to get the ground truth alpha matte by triangulation. To capture the images for testing, we change the background on the monitor with natural background images. Finally, the images were cropped at a bounding box that was casually drawn around the foreground objects, resulting in the test scenes. Our dataset consists of 5 testing images, and the foreground objects were chosen to cover different properties with hard and soft boundaries, as well

Table 1. Quantitative Comparisons in term of RMSE and Consistency

	Data01		Data02		Data03		Data04		Data05	
	RMSE	CONS	RMSE	CONS	RMSE	CONS	RMSE	CONS	RMSE	CONS
GSM	8.259	5.267	15.997	8.567	13.967	5.900	10.341	6.057	13.167	7.498
KNN	10.508	7.681	20.085	5.575	12.822	8.315	13.605	8.597	18.197	11.855
COM	**7.776**	5.501	16.015	5.375	11.627	6.330	10.347	6.740	13.731	7.761
VIDEO	10.356	5.532	29.253	5.480	19.020	6.383	10.209	5.389	15.314	8.184
OUR	7.870	**5.178**	**14.491**	**4.595**	**10.277**	**5.792**	**9.897**	**5.277**	**11.194**	**6.929**

as translucency. Figure 4 shows our testing images, trimaps and the ground truth alpha mattes. The center view trimap is provided by user manually, and it is propagated to other sub-images automatically using the foreground EPI.

5.2 Evaluations

We evaluate the performance of our algorithm in term of RMSE and consistency. The RMSE is computed as follows:

$$RMSE(\alpha) = \sqrt{\frac{1}{N}\sum_i (\alpha_i^* - \alpha_i)^2}, \tag{10}$$

where α^* is the ground truth alpha and N is the total number of pixels. The consistency is evaluated as follows:

$$CONS(\alpha) = \sqrt{\frac{1}{N}\sum_i (\frac{1}{N_{EPI_i}} \sum_{j \in EPI_i} ||\alpha_i - \alpha_j||^2)}, \tag{11}$$

where α_i and α_j are the EPI correspondences defined by the foreground EPI in $x-s$ plane and $y-t$ plane respectively. The estimated alpha matte of the center view is used for evaluation.

5.3 Comparisons

We compare the performance of our algorithm with the state-of-the-art matting algorithms: global sampling matting [13], KNN matting [3], comprehensive matting [23], and video matting [6]. Table 1 summarizes the comparisons. In Figure 5, we also show the qualitative comparisons on the results of our dataset. As presented in the Table 1 and Figure 5, our algorithm achieves the minimum RMSE in most cases. Also, our results are more consistent across sub-images deal to the EPI smoothness term in our consistent matting algorithm.

Additionally, we also apply our algorithm to a real world example from UCSD in [15]. Figure 6 shows the comparisons using UCSD light field data [15] which are captured by cameras in a row. We use 20 sequential sub-images which include a pink gorilla as the foreground with the background. Since the sub-images are arranged horizontally, we have only the $x-s$ plane in EPI. As shown in Figure 6, our matting result is similar to the results from previous method, but the alpha matte in EPI is smoother which shows that our result is more consistent.

Fig. 5. Qualitative comparisons on dataset. We compare our estimated alpha mattes with results from previous methods. (a) Inputs, (b) Results from global sampling matting [13], (c) Results from KNN matting [3], (d) Results from comprehensive matting [23], (e) Results from video matting [6], (f) Our results, (g) Ground truth.

6 Limitation and Discussion

In this paper, we assume foreground and background have sufficient distance such that the directions of foreground EPI and background EPI are very different from each other. If this assumption is violated, *i.e.*the foreground EPI and background EPI are parallel to each other, our approach is less effective since the EPI constraint cannot be used to resolve the ambiguity in color sample selection.

Our another assumption is that the EPIs in unknown regions can be smoothly extrapolated from known regions. If there is significant depth changes in matting area, the errors in extrapolated EPIs will be propagated to our matting results since the EPI correspondences are incorrect. Similar to previous matting

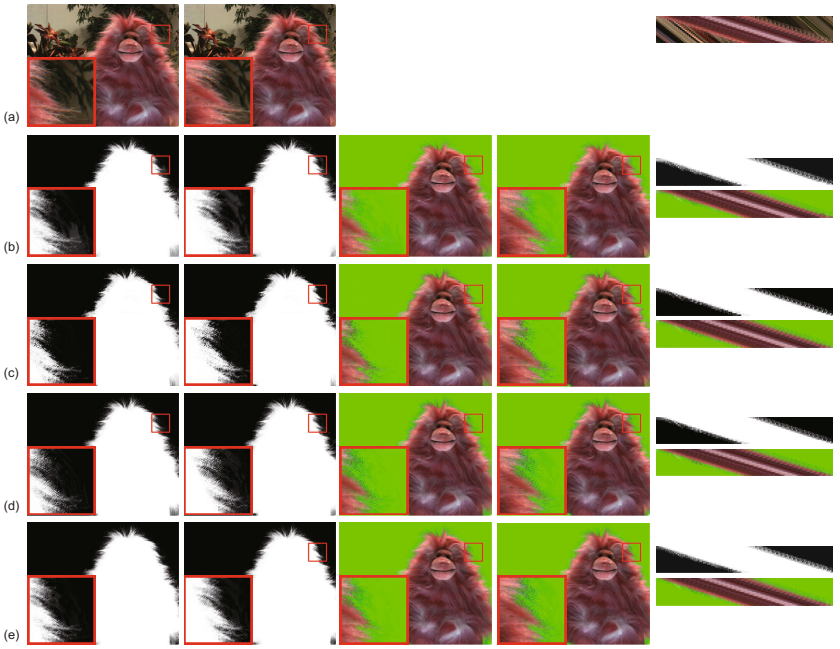


Fig. 6. A horizontal light field sub-images form UCSD data [15]. 20 sub-images are used for consistency. In the example, from the top, (a) input images and (b)-(e) alpha mattes by Global sampling matting [13], KNN matting [3], comprehensive matting [23], and our proposed matting. The first two columns are estimated alpha mattes and the second two columns are the combined image with the green background. In the righthand column, we provide the EPIs to compare the slopes for the consistency.

algorithms, we also assume noise level is low in the input images. If there is significant amount of noise, our EPI estimation method may be broken. In such case, better EPI estimation algorithm should be utilized.

7 Conclusion

In this paper, we have presented a method for light field image matting which estimates consistent alpha mattes of foreground across multiple sub-images in a light field image. By using the EPI constraint, we can define different sets of pixel correspondences for foreground and background. In the color sample selection, we have presented a method to estimate foreground samples alpha with known and unknown background samples. We have also introduced a method to include the EPI smoothness constraint and proposed to solve alpha matte of multiple sub-images simultaneously. In the experimental evaluations, we have created a new dataset with ground truth alpha mattes to quantitatively compare the performance of our algorithm with the performance of state-of-the-art image matting algorithms. Our algorithm outperforms previous work in term of both

RMSE and consistency. As for future work, we are interested in extending our work in other applications that utilize light field image data.

Acknowledgement. We thank the anonymous reviewers for their valuable comments. This research is supported by the MSIP, Korea, under the IT/SW Creative research program supervised by the NIPA (NIPA-2013-H0503-13-1011), and the Multi-ray camera System funded by the Samsung Electronics Co., Ltd (DMC R&D center) (IO130806-00717-01).

References

1. Bishop, T.E., Zanetti, S., Favaro, P.: Plenoptic depth estimation from multiple aliased views. In: IEEE ICCV Workshops (2009)
2. Bolles, R.C., Baker, H.H., Marimont, D.H.: Epipolar-plane image analysis: An approach to determining structure from motion. IJCV 1(1), 7–55 (1987)
3. Chen, Q., Li, D., Tang, C.K.: Knn matting. In: IEEE CVPR (2012)
4. Chen, X., Zou, D., Zhou, S.Z., Zhao, Q., Tan, P.: Image matting with local and nonlocal smooth priors. In: IEEE CVPR (2013)
5. Cho, D., Lee, M., Kim, S., Tai, Y.W.: Modeling the calibration pipeline of the lytro camera for high quality light-field image reconstruction. In: IEEE ICCV (2013)
6. Choi, I., Lee, M., Tai, Y.-W.: Video matting using multi-frame nonlocal matting laplacian. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part VI. LNCS, vol. 7577, pp. 540–553. Springer, Heidelberg (2012)
7. Chuang, Y.Y., Curless, B., Salesin, D.H., Szeliski, R.: A bayesian approach to digital matting. In: IEEE CVPR (2001)
8. Criminisi, A., Kang, S.B., Swaminathan, R., Szeliski, R., Anandan, P.: Extracting layers and analyzing their specular properties using epipolar-plane-image analysis. Computer Vision and Image Understanding (CVIU) 97(1), 51–85 (2005)
9. Dansereau, D.G., Bruton, L.T.: Gradient-based depth estimation from 4d light fields. In: IEEE International Symposium on Circuits and Systems (ISCAS) (2004)
10. Dansereau, D.G., Pizarro, O., Williams, S.B.: Decoding, calibration and rectification for lenselet-based plenoptic cameras. In: IEEE CVPR (2013)
11. Gastal, E.S.L., Oliveira, M.M.: Shared sampling for real-time alpha matting. In: Eurographics (2010)
12. Goldluecke, B., Wanner, S.: The variational structure of disparity and regularization of 4d light fields. In: IEEE CVPR (2013)
13. He, K., Rhemann, C., Rother, C., Tang, X., Sun, J.: A global sampling method for alpha matting. In: IEEE CVPR (2011)
14. He, K., Sun, J., Tang, X.: Fast matting using large kernel matting laplacian matrices. In: IEEE CVPR (2010)
15. Joshi, N., Matusik, W., Avidan, S.: Natural video matting using camera arrays. In: ACM SIGGRAPH (2006)
16. Lee, P., Wu, Y.: Nonlocal matting. In: IEEE CVPR (2011)
17. Levin, A., Lischinski, D., Weiss, Y.: A closed-form solution to natural image matting. IEEE Trans. on PAMI 30(2), 162–8828 (2008)
18. Lin, H., Tai, Y.W., Brown, M.S.: Motion regularization for matting motion blurred objects. IEEE Trans. on PAMI 33(11), 2329–2336 (2011)
19. Lytro: The lytro camera, `https://www.lytro.com`

20. Ng, R., Levoy, M., Brédif, M., Duval, G., Horowitz, M., Hanrahan, P.: Light field photography with a hand-held plenoptic camera. Tech. rep. (2005)
21. Rhemann, C., Rother, C., Gelautz, M.: Improving color modeling for alpha matting. In: British Machine Vision Conference (BMVC) (2008)
22. Rhemann, C., Rother, C., Wang, J., Gelautz, M., Kohli, P., Rott, P.: A perceptually motivated online benchmark for image matting. In: IEEE CVPR (2009)
23. Shahrian, E., Rajan, D., Price, B., Cohen, S.: Improving image matting using comprehensive sampling sets. In: IEEE CVPR (2013)
24. Shahrian, E., Rajan, D.: Weighted color and texture sample selection for image matting. In: IEEE CVPR (2012)
25. Smith, A.R., Blinn, J.F.: Blue screen matting. In: ACM SIGGRAPH (1996)
26. Sun, J., Jia, J., Tang, C.K., Shum, H.Y.: Poisson matting. ACM TOG 23(3), 315–321 (2004)
27. Tao, M.W., Hadap, S., Malik, J., Ramamoorthi, R.: Depth from combining defocus and correspondence using light-field cameras. In: IEEE ICCV (2013)
28. Wang, J., Cohen, M.F.: Optimized color sampling for robust matting. In: IEEE CVPR (2007)
29. Wanner, S., Goldluecke, B.: Globally consistent depth labeling of 4D lightfields. In: IEEE CVPR (2012)
30. Wanner, S., Goldluecke, B.: Reconstructing reflective and transparent surfaces from epipolar plane images. In: Weickert, J., Hein, M., Schiele, B. (eds.) GCPR 2013. LNCS, vol. 8142, pp. 1–10. Springer, Heidelberg (2013)
31. Wanner, S., Straehle, C., Goldluecke, B.: Globally consistent multi-label assignment on the ray space of 4d light fields. In: IEEE CVPR (2013)
32. Zheng, Y., Kambhamettu, C.: Learning based digital matting. In: IEEE ICCV (2009)