Stitching of Heterogeneous Images Using Depth Information

Jun-Tae Lee, Jae-Kyun Ahn, and Chang-Su Kim
School of Electrical Engineering, Korea University, Seoul, Korea
E-mails: {jtlee, jaekyunahn, cskim}@mcl.korea.ac.kr

***Abstract*—We propose a novel heterogeneous image stitching algorithm, which employs disparity information as well as color information. It is challenging to stitch heterogeneous images that have different background colors and diverse foreground objects. To overcome this difficulty, we set the criterion that objects should preserve their shapes in the stitched image. To satisfy this criterion, we derive an energy function using color and disparity gradients. As the gradients are highly correlated with object boundaries, we can find the optimal seam from the energy function, along which two images are pasted. Moreover, we develop a retargeting scheme to reduce the size of the stitched image further. Experimental results demonstrate that the proposed algorithm is a promising tool for stitching heterogeneous images.**

I. Introduction

Image synthesis is a technique of creating a new image from single or multiple images. The interest in image synthesis has increased significantly due to the rapid growth of smart phone applications and digital cameras. Two important categories of image synthesis are image stitching and image composition. Image stitching generates a panoramic image from multiple images, which are registered to one another. Although advanced lens have been developed, the fields of view of cameras are constrained by physical limitations. Hence, many image stitching methods have been proposed to combine multiple images into a single wide view image. On the other hand, image composition pastes a region from a source image to a target image. It is employed in various applications, such as texture interpolation and image melding. Image composition focuses on the natural transition of regions to create a faithful image. In both categories, the challenging issue is to synthesize seamless images without visible artifacts.

In image stitching, input images often have luminance differences, and a stitched image usually yields noticeable artifacts, referred to as seams. Therefore, the quality of a stitched image is measured by how seamless the resulting mosaic is. Levin *et al.* [1] proposed an image stitching algorithm, which combines two pre-aligned images in the gradient domain. Since their algorithm performs the image stitching in the gradient domain, artifacts from luminance differences are reduced efficiently in the spatial domain. However, it considers only pre-aligned images as the input. Lin *et al.* [2] proposed an alignment algorithm for image stitching, which uses local affine transforms between two input images. Their algorithm determines affine parameters of every pixel in input images from pre-computed features, and uses the parameters to overcome inconsistent misalignments. In [3], [4], image stitching techniques have been proposed to handle the ordering, orientation, scale or illumination differences of input images automatically. Jia *et al.* [5] investigated both intensity and structural inconsistencies. They determined initial stitching using gradients and detected structural features along the stitched line. Then, they solved a Poisson equation to generate a seamless result. Notice that most of these image stitching algorithms have the implicit assumption that input images should be parts of the same scene.

Image composition, which uses images with monotonous but different background, has been researched as well. There is a pre-defined transition region in a source image, and it is smoothly transformed to a target image. Since the source image contains background as well as objects in the transition region, direct composition causes inconsistency in most cases. Also, the photometric differences between the source and the target, such as textures, illuminations and tones, cause artifacts. Perez *et al.* [6] proposed an algorithm, which solves a Poisson equation subject to a Dirichlet boundary condition. Jia *et al.* [7] further optimized the boundary condition of Perez *et al.*'s algorithm. Their algorithm results in more natural image composition. It prevents the boundary of the source from invading the structural background of the target, suppressing visible distortions. Darabi *et al.* [8] proposed an image melding method, which gradually synthesizes the transition region. Their algorithm applies the patch-matching scheme [9] to the image synthesis and fuses patches in a robust manner against scale, rotation and illumination differences. It is effective in the transition region with monotonous background. These image composition algorithms can achieve natural results in transition regions. However, these algorithms are effective only if images contain monotonous background.

In this paper, we introduce a novel stitching algorithm to combine heterogeneous images. In the heterogeneous image stitching, objects should be preserved. In other words, it is desired that a resultant image maintains as many objects as possible. To achieve this goal, it is necessary that objects overlap one another. In the proposed algorithm, we introduce an energy function to find an optimal seam for the overlapping.

This work was supported partly by the National Research Foundation of Korea(NRF) grant funded by the Ministry of Science, ICT & Future Planning (MSIP) (No. 2009-0083495), and partly by the National Research Foundation of Korea(NRF) grant funded by the Korea government (MSIP) (No. 2012-011031).

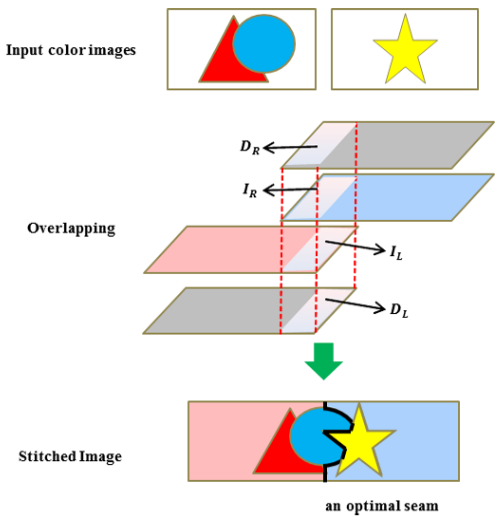


Fig. 1. An example of the panoramic stitching: To obtain a stitched image, we overlap two input images with disparity data and estimate the optimal seam within the overlapping region. I_L and D_L denote the color and disparity maps for the left image, and I_R and D_R for the right image, respectively.

The energy function is based on the color and disparity information of input images. Along the optimal seam, we stitch the input images. Moreover, after the stitching, we can retarget the stitched image further by rearranging the disparity data.

The rest of this paper is organized as follows: Section II describes the proposed heterogeneous image stitching algorithm in detail. Section III presents experimental results and analysis. Conclusions and future works are in Section IV.

II. Proposed Algorithm

The proposed image stitching algorithm is composed of two steps: panoramic stitching and image retargeting.

First, in the panoramic stitching step, two heterogeneous input images with depth information are pasted along a seam. To find the optimal seam, we extend the Avidan *et al.*'s seam carving technique [10]. We exploit the depth information, as well as the color information, to find the optimal seam. Fig. 1 illustrates the panoramic stitching. The optimal seam is determined to lie on the boundaries between the circle and the star, which yields the natural transition between the left and the right images.

Second, in the image retargeting step, we resize the stitched image to provide a compact image. To preserve object silhouettes, we compare the disparities of overlapping pixels and rearrange the objects in the overlapping region. We also control disparity ranges to reduce visible artifacts.

A. Seam Estimation

Let us briefly describe the seam carving technique [10], on which the proposed algorithm is based. The seam carving technique finds optimal seams repeatedly, which avoid salient features of an image, and carves them out. During the carving, the image maintains rectangular shapes, since each row contributes one pixel to the seam. The least noticeable pixels are selected as the seam to preserve visual coherence of the image. To quantify the noticeability, the energy function at each pixel p is defined as

$$e_1(p) = \|\nabla I(p)\|_1, \tag{1}$$

where the operator $\|\cdot\|_1$ denotes the l_1 norm, and $\nabla I(p)$ is the image gradient at pixel p. To keep important pixels, the optimal seam s^* should be selected to minimize the seam cost, given by

$$s^* = \min_s E(s) = \min_s \sum_i e(s_i), \tag{2}$$

where $s_i = (x(i), i)$ is the pixel in the ith row on the seam s. Since the seam should be connected, a neighbor constraint $|x(i) - x(i+1)| \le k_0$ is imposed. To find the optimal seam satisfying the constraint, all paths are considered based on dynamic programming, by accumulating the pixel energy from top to bottom. The cumulative energy M at pixel $p = (x, y)$ is given by

$$M(x, y) = e_1(p) + \min(\{M(x \pm k, y-1) | k \le k_0\}). \tag{3}$$

Then, starting from the smallest $M(i, j)$ of the bottom row, s^* is determined by the backtracking.

B. Panoramic Stitching

For the image stitching, we use four input images: two color images and two disparity maps. All inputs have the same width and height. We denote the color and disparity of the left image in the overlapping region by I_L and D_L. Similarly, I_R and D_R indicate those of the right image, as shown in Fig. 1. Within the overlapping region, we find the optimal seam to stitch the two color images and make a panoramic image. We extend the seam carving technique [10] to find the optimal seam for the stitching. Since the proposed algorithm uses disparity maps as well as color images, we incorporate the disparity information into the energy function, which is given by

$$e(p) = e_c(p) + e_d(p) + e_s(p), \tag{4}$$

where $e_d(p)$ and $e_c(p)$ are the color and disparity terms, respectively, and $e_s(p)$ is the similarity term. To achieve the natural transition between the left and right images, it is desirable that the seam lies along object silhouettes and avoid passing through object interiors. Since the color and disparity gradients are highly correlated to the object silhouettes, we define the color and depth terms using the gradients. First, the color term is defined as

$$e_c(p) = \frac{1}{\alpha}\left(\|\nabla I_L(p)\|_1 + \|\nabla I_R(p)\|_1\right), \tag{5}$$

where α is a normalization factor for each row r, *i.e.* $\alpha = \max_{q \in r}(\|\nabla I_L(q)\|_1 + \|\nabla I_R(q)\|_1)$. The color term indicates that, if pixel p has high gradient magnitudes in the left or right images, it is likely to be included in the seam. Similarly, the disparity term is defined as

$$e_d(p) = \frac{1}{\beta}(w_L(p)\|\nabla D_L(p)\|_1 + w_R(p)\|\nabla D_R(p)\|_1), \tag{6}$$

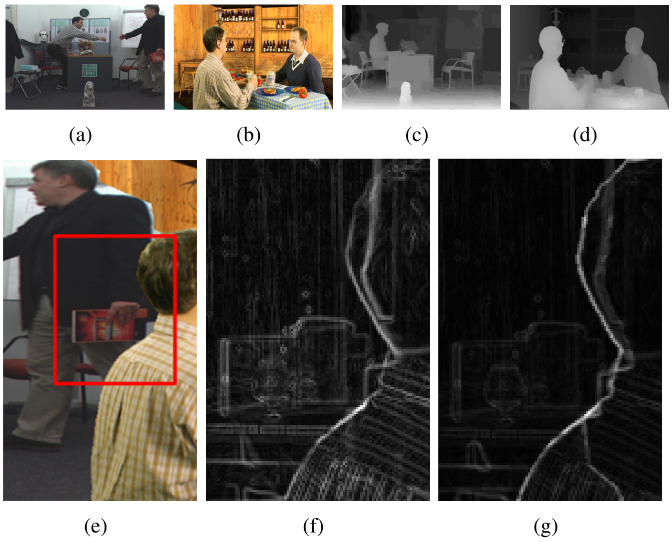


Fig. 2. Seam estimation: (a) left color image, (b) right color image, (c) left disparity image, (d) right disparity image, (e) overlapping region after stitching, (f) energy function without the disparity term, and (g) energy function with the disparity term.

where β is another normalization factor. Also, $w_L(p)$ or $w_R(p)$ is a weight function, given by

$$w(p) = \begin{cases} 1 & \text{if } D(p) > \delta \\ 0.5 & \text{otherwise,} \end{cases} \tag{7}$$

where δ is the most commonly found disparity in the input image. Note that the disparity term improves the quality of the panoramic image significantly. While the original seam carving minimizes the energy function, the proposed algorithm chooses a pixel at each row to maximize the energy. Thus, the optimal seam tends to adhere to object boundaries based on the color and disparity terms. However, candidate pixels in the current row occasionally have almost identical $e_d(p)$ values, which causes the seam to be found in wrong places. To prevent this, we assign a high energy to the candidate pixel, which has the similar disparity gradient to the seam pixel in the upper row. Thus, the similarity term $e_s(p)$ is defined as

$$\begin{aligned} e_s(p) = \frac{1}{\gamma} \sum_{q \in p(i-1, j \pm b)} &\{\langle \nabla D_L(p), \nabla D_L(q) \rangle \\ &+ \langle \nabla D_R(p), \nabla D_R(q) \rangle\}, \end{aligned} \tag{8}$$

where γ is a normalization factor and the operator $\langle \cdot, \cdot \rangle$ indicates the inner product.

Then, we determine the optimal seam s^* that maximizes the seam cost,

$$s^* = \max_s E(s) = \max_s \sum_{i=1}^{H} e(s_i), \tag{9}$$

where $E(s)$ denotes the seam cost function. To examine all possible seams with the neighbor constraint, we compute the cumulative function M as follows,

$$M(x, y) = e(x, y) + \max(\{M(x \pm k, y-1) | k \leq k_0\}). \tag{10}$$

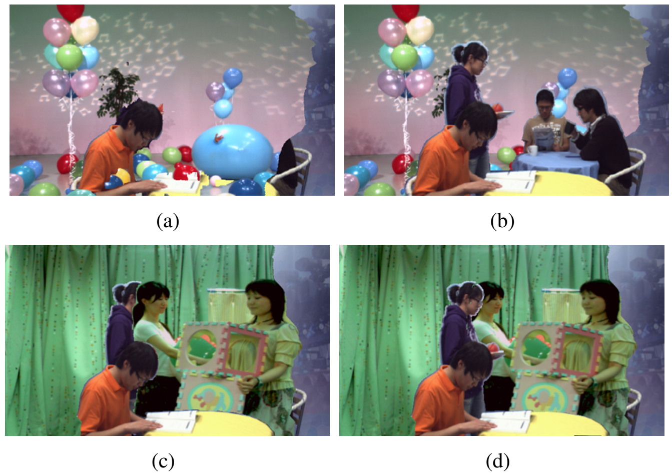


Fig. 3. Retargeted images: (a) initial retargetted image using the original disparities and (b), (c), (d) three retargeted images after different disparity adjustment.

Then, as in [10], we backtrack the optimal seam.

To illustrate the effectiveness of the disparity term, Fig. 2 visualizes the energy functions without and with the disparity term. We see that the disparity term clarifies object boundaries.

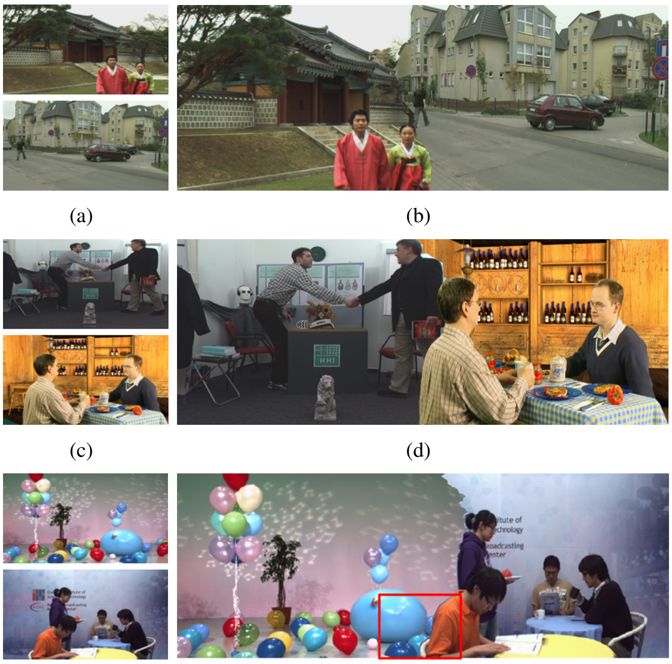
C. Image Retargeting and Disparity Adjustment

We develop an image retargeting method, which resizes the stitched image more compactly. After the image stitching, the overlapping region often covers a dominant part of the stitched image. Thus, the further cropping from a new seam causes the removal of several important objects. Hence, instead of finding a new seam, we maintain the overlapping region and indicate front objects to be included in the retargeting result. Since foreground objects are often more important than the background, we choose the pixel with the larger disparity between two overlapping pixels. However, if two disparities are almost identical, objects appear irregularly after this pixel selection. For example, if a floor has partly larger disparities than objects, parts of the floor might become visible through objects. To overcome this, we adjust the disparity range by adding or subtracting a certain value. In this way, a user can select desirable objects in the resized image easily through interactions. Fig. 3 shows examples of retargeted images. Fig. 3(a) is an initially retargeted image. Figs. 3(b), (c), and (d) shows results of the disparity adjustment. Note that the disparity adjustment enables a user to synthesize various retargeted images easily.

III. Experimental Results

We evaluate the performance of the proposed algorithm on various multi-view images with disparity maps. The test images are selected from “Akko and Kayo,” “Lovebird,” “Beer,” “Street,” “Book Arrival,” “Balloons,” and “Cafe” sequences. Since these sequences have different spatial resolutions, they are resized to 640×360 before the stitching.

Fig. 4 shows the panoramic images, which are stitched by the proposed algorithm. Although the input images are



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

Fig. 4. Panoramic stitching: the left column shows two heterogeneous input pairs, while the right column shows the stitched result for each pair.

heterogeneous and uncorrelated to each other, the stitched images exhibit natural transition in the overlapping regions. Note that object silhouettes are preserved faithfully. Especially, the two images in Fig. 4(a) have similar tones, and thus they are combined to synthesize a faithful image in Fig. 4(b). On the other hand, the images in Figs. 4(c) and (d) have quite different background colors. Thus, after the stitching, the background shows inevitable distortions along the seam. We are currently developing a technique to fuse these different background data smoothly.

Fig. 5 compares the stitched images, which are obtained without and with the disparity terms. In Fig. 5(a), balloons in the floor occlude the chair, and the shoulder silhouette is also broken. In contrast, in Fig. 5(b), the seam lies along the edges of the chair and the shoulder. This is because the disparity gradients have high magnitudes at the object boundaries.

Fig. 6 shows the retargeted results from the stitched images in Figs. 4(b) and (d). We see that the resultant images include most of salient objects and they are rearranged properly according to the depths. Simulation results on other test images also confirmed that the proposed algorithm is an effective stitching tool for heterogeneous images.

IV. Conclusions

In this work, we proposed a novel image stitching algorithm, which can paste heterogeneous input images faithfully. The stitching of heterogeneous images is difficult, since they often have different background colors and diverse foreground objects. To overcome this difficulty, the proposed algorithm uses disparity maps to estimate the optimal seam, along which the images are pasted. Consequently, the stitched image maintains most objects without destroying their shapes. Also, we developed a retargeting scheme for stitched images. By employing rearranged disparities, the proposed algorithm determines relative depths of the objects in the left and right images and render those objects naturally in the retargeted image. Further research issues include the extension of the proposed algorithm for the stitching of videos and the smooth melding of background regions.

(a) (b)

Fig. 5. Comparison of the panoramic stitching without the disparity term (a) and with the disparity term (b).

(a) (b)

Fig. 6. Retargeting of the stitched images in Figs. 4(b) and (d).

References

[1] A. Levin, A. Zomet, S. Peleg, and Y. Weiss, "Seamless image stitching in the gradient domain," *in Proc. Eur. Conf. Comput. Vis.,* pp. 377-389, May 2004.

[2] W. -Y. Lin, S. Liu, Y. Matsushita, T. -T. Ng, and L. -F. Cheong, "Smoothly varying affine stitching," *in Proc. IEEE Conf. Comput. Vis. Pattern Recog.,* pp. 345–352, Jun. 2011.

[3] M. Brown, and D. -G. Lowe, "Automatic panoramic image stitching using invariant features," *Int. J. Comput. Vis.,* vol. 74, no. 1, pp. 59–73, Aug. 2007.

[4] Y. Li, Y. Wang, W. Huang, and Z. Zhang, "Automatic image stitching using sift," *in Proc. Audio, Lang. Image Process.,* pp. 568–571, Jul. 2008

[5] J. Jia, and C. K. Tang, "Eliminating structure and intensity misalignment in image stitching," *in Proc. Int. Conf. Comput. Vis.,* vol. 2, pp. 1651–1658, Oct. 2005

[6] P. Perez, M. Gangnet, and and A. Blake, "Poisson image editing," *ACM Trans. Graphics,* vol. 22, no. 3, pp. 313–318, Jul. 2003.

[7] J. Jia, J. Sun, C. -K. Tang, H. -Y. Shum, "Drag–and–drop pasting," *ACM Trans. Graphics,* vol. 25, no. 3, pp. 631–637, Jul. 2006

[8] S. Darabi, E. Shechtman, C. Barnes, D. -B. Goldman, and P. Sen, "Image melding: combining inconsistent images using patch-based synthesis," *ACM Trans. Graphics,* vol.31, no.3, pp.82, Jul. 2012

[9] C. Barnes, E. Shechtman, D. -B. Goldman, and A. Finkelstein, "The generalized patchmatch correspondence algorithm," *in Proc. Eur. Conf. Comput. Vis.,* pp. 29-43, Sept. 2010

[10] S. Avidan, and A. Shamir, "Seam carving for content-aware image resizing," *ACM Trans. Graphics,* vol. 26, no. 3, pp. 10, Aug. 2007