

# GENERATING PANORAMA IMAGE BY SYNTHESIZING MULTIPLE HOMOGRAPHY

*Seongdo Kim, YoungJung Uh, Hyeran Byun*

Department of Computer Science, Yonsei University, Seoul, Korea  
{seongdo,sheldon,hrbyun}@yonsei.ac.kr

## ABSTRACT

This paper presents a method to generate image mosaics of a panoramic scene. In general, the relation between images which is required for mosaicing cannot be expressed by a single homography due to geometrical condition of the scene, even if the images are taken at the same position. Many existing methods are using only one homography to make panorama image while ignoring the geometrical variations. Therefore, they experience a lot of distortions and misalignments from input images which contain several planes which cannot be handled by one homography. In this paper, we present a novel method that utilizes synthesis of multiple homography to warp the images. Moreover, our method determines the number of homography automatically, without user's input. By our method, various distortions of shapes and mismatches can be reduced.

*Index Terms*—Image Mosaicing, Image Stitching, Multiple Homography Synthesis, Panorama

## 1. INTRODUCTION

Recent years, image mosaicing has been under extensive study and various approaches exist in the literature. Current status of image mosaicing enabled many commercial products to be intimate to end-users. The center of previous researches on image mosaicing lies in feature-based alignment and stitching [1][2], but there still remain challenging problems especially to resolve various constraints. In more detail, they usually work well on single-homography cases such as distant scenes which are explainable by one plane or a cylindrical surface [3], but not on intricate cases. Consequently, degree of distortion in the result image will grow while the formation of planes gets more complex. It is also true for Photoshop CS5 which is well-known commercial software. The state-of-arts algorithm introduced Dual-homography method and applied unidentical warping to each pixel to deal with this problem [4]. They adopt Dual-homography warping to represent two planes which are ground plane and distant plane to synthesize a new homography on each pixel. We extend this idea to cover more large range of situations by adopting multiple number of homography.

In this paper, we present a novel method to find the multiple homography, generate the synthesis of multiple homography, and construct a panorama image with better quality than existing methods. To do this, we firstly extract the feature points and match them to make matching point pairs. Next, we classify the matching point pair into proper number of clusters which have corresponding homographies for the input image pair. After that, we synthesize the multiple homography introducing the flood-distant factor. To generate natural boundary between the warped images, we find the optimal cut on the overlapped area using  $\alpha$ - $\beta$  swap algorithm of graph-cut [2], and stitch them. Color and gradient [4] were used for energy formulation of graph-cut.

The organization of this paper is as follows. The proposed method is described in Section 2 including algorithm flow and details. Section 3 is about experiments of our method and comparison with a recently proposed method and a commercial software. The conclusion is given in section 4.

## 2. IMAGE STITCHING WITH SYNTHESIZED HOMOGRAPHY

To compute the multiple homography and combine them, first we extract feature points from the overlapped input image pair, and match them to make matching point pairs. The matching point pairs will be bundled into the optimal number of clusters to find corresponding multiple homography. After that, we synthesize the multiple homography using the flood-distance factor, which are introduced by our proposed method. The synthesized homography will be used to align the images. Figure 1 shows the entire procedure of our method. In this section, we will describe our procedure in detail.

### 2.1. Estimating Optimal Homography Set

For the input image pair, we extract features using scale invariant feature transform (SIFT) [5], and match them to make matching point pairs. We bundle these matching point pairs into several clusters. J. Gao et al. used K-means clustering [4]. However, they used Euclidean distance on 2D image space for K-means clustering which is not reasonable enough to classify the matching point pairs into clusters while considering planes in the scene. Moreover, it is easy

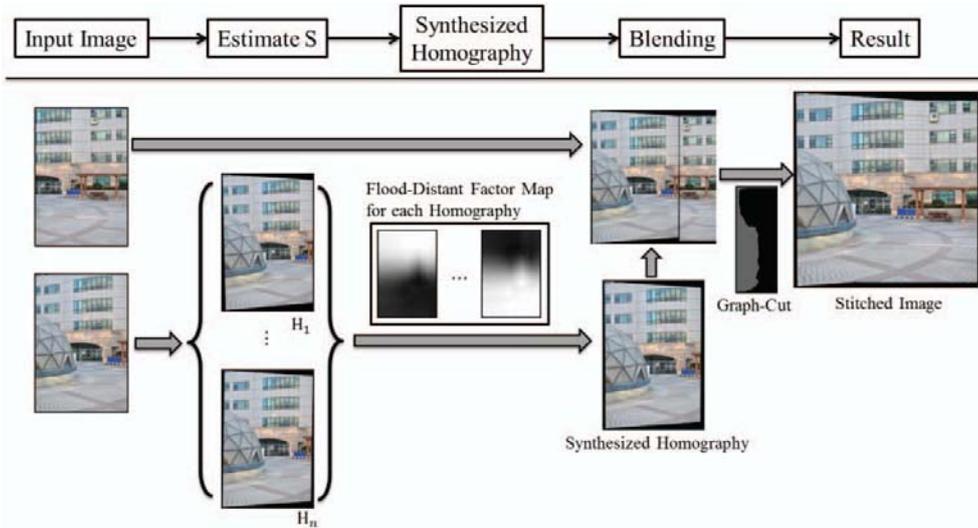


Figure 1: The workflow of our panorama construction system. Proper number of homography will be automatically determined via our method.

to wrongly decide whether the matching point pairs belong to the ground plane or the distant plane, because the result of K-mean clustering depends on the scattering of the matching point pairs. To cluster the matching point pairs and compute corresponding homography in a reasonable way, we propose a novel method as follows.

The input image pair can be expressed by more than two homography, but the number of proper homography is hard to predict. To solve this problem, we define the arbitrary set  $S = (n, \mathcal{H})$  where  $n$  is the number of homographies, which is same as the number of clusters of the matching point pairs, and  $\mathcal{H} = \{H_i\}$  is a set of homography matrices for the corresponding clusters respectively.  $H_i$  is calculated using cluster  $\mathcal{P}_i$  of the matching point pairs using RANSAC-based robust method implemented in open computer vision library (openCV).  $\hat{S}$  is the optimal set which has the smallest cost. To get this optimal set, we define the objective function as

$$\hat{S} = \arg \min_{S_i} \mathcal{C}(S_i) \quad (1)$$

where  $S_i = (n_i, \mathcal{H}_i)$  is an randomly generated set by arbitrarily dividing the image sector, and  $\hat{S} = (\hat{n}, \hat{\mathcal{H}})$  is the optimal set. We generate a set  $S_i$  with the randomly divided clusters of the matching point pairs. The function  $\mathcal{C}(S_i)$  means cost of  $S_i$ , which is defined as

$$\mathcal{C}(S_i) = \left( \frac{\sum_{j=0}^{n_i} \mathcal{J}(H_j, \mathcal{P}_j)}{\sum_{j=0}^{n_i} |\mathcal{P}_j|} \right)^{-1} \quad (2)$$

where  $\mathcal{J}(H, \mathcal{P})$  is the number of inliers of the matching point pairs in cluster  $\mathcal{P}$ .  $|\mathcal{P}|$  is the number of matching point pairs in  $\mathcal{P}$ .

In our experiment, generally 2~3 homography were found in one image pair.

## 2.2. Weight Estimation for Multiple Homography

Multiple homography were acquired in prior section by optimizing the set  $\hat{S}$ , however, we cannot directly use these homography to align the images. For each pixel of the image, first, we should decide how to use multiple homography to warp it. Depending on pixel's location, effect of each homography should differ, and we can determine the weight of each homography by considering the distance to each cluster. The Dual-Homography method calculates the weight of each homography for each pixel using distance ratio of the distant plane's cluster and the ground plane's cluster by Euclidean distance from the pixel to the nearest feature point in the cluster. However, calculating weight costs more time as the number of feature points grows because they use  $\|\cdot\|_2$ . Moreover, if a pixel is far from points in both clusters, then homography for the pixel can be disturbed by the opposite cluster and cause misalignment. It is due to the distance ratio converges to 1:1 because distances to both clusters grow together (Figure 2). To eliminate the distortion caused by the unsatisfactory combination of homographies, we consider the farther cluster much less than the closer one. In this section, we propose a novel method to calculate weight introducing flood-distance factor to solve this problem.

As Dual-Homography method represented, distance to the cluster is the most important factor for deciding the weight to each homography. However, the method has a problem that direct computation of the distance will cost large amount of time as the number of the feature points increases. Hence we adopt flooding method with Manhattan distance instead of naïve Euclidean distance to resolve the problem. Every matched feature point is used as seed point.

We calculate the distance from every seed point to each pixel. In order to disregard farther cluster, we introduce an

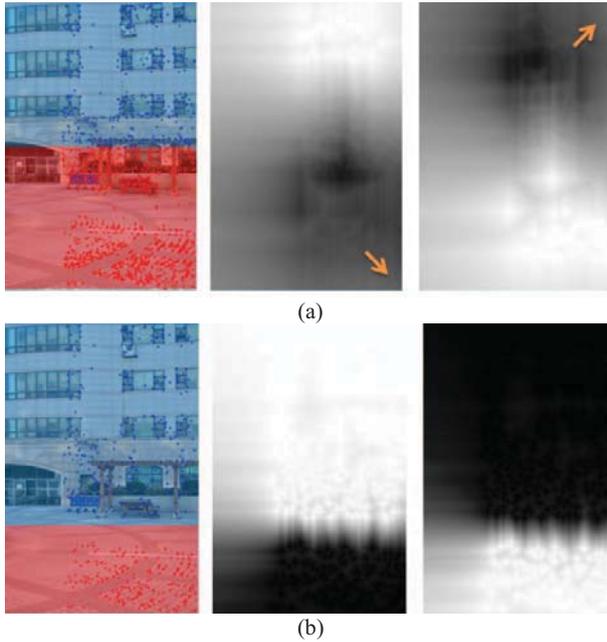


Figure 2. Weight map comparison with 2-sector case. Left column shows division of planes. (a) Weigh map of Dual-Homography method. Orange arrows indicate the region with wrongly calculated weight caused by opposite homography, (b) Weight map calculated by our method.

increasing factor  $r$  to define a distance measure from pixel  $p'$  to the cluster  $\mathcal{P}_k$  as

$$f_k^{p'} = (f_k^p + 1) \times r \text{ for } p' \in \mathcal{N}_p \quad (3)$$

where  $f_k^p$  is flood-distance factor of pixel  $p$  from the closest seed point in the cluster  $k$ ,  $\mathcal{N}_p$  is 4-neighboring pixel of  $p$ . We fix increasing factor  $r$  as 1.004 experimentally. The increasing factor is to reduce the interference of distant homography which exists in Dual-Homography method (Figure 2(a)). Then we can calculate the weight of each homography for a pixel  $p$  as

$$\omega_p^k = 1 - \frac{f_k^p}{\sum_{l=0}^{\hat{n}} f_l^p} \quad (4)$$

where divider is sum of all flood-distance factors. Using this weight, we can compute the synthesized homography  $H_p$  for  $p$  as

$$H_p = \sum_{k=1}^{\hat{n}} \omega_p^k H_k \quad (5)$$

where  $\hat{n}$  is number of cluster,  $H_k$  is homography for cluster  $\mathcal{P}_k$ .

### 2.3. Stitching based on Seam-Cut

Traditional researches adopt the blending method which works well only when the images coincide exactly each other. However, even if image was properly warped, not all region of the image exactly coincides. Therefore, we adopt cut based stitching on overlapped region of the images

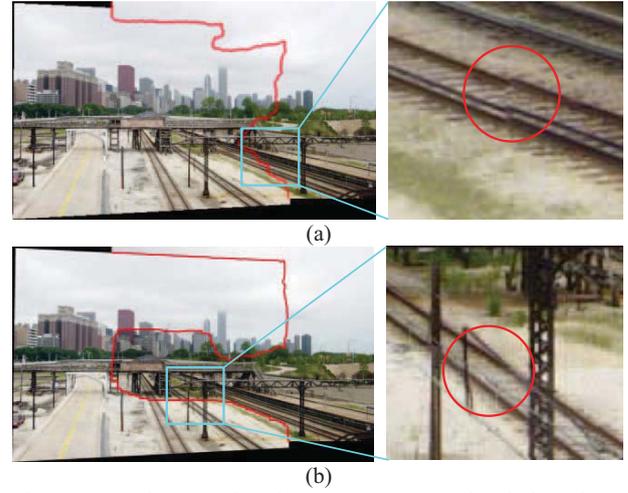


Figure 3. Result comparison between our proposal and Photoshop PS5. (a) Result of Photoshop CS5 contains misalignment of railway on the blending boundary. (b) Result our method have no misalignment on the blending boundary.

instead of blending method. We find the optimal boundary by graph-cut to generate the most naturally stitched image.

An image can be expressed as a discrete grid which can also be treated as Markov random field. We choose a label for each pixel which is a node in the grid to denote the image which the pixel value comes from. To find the optimal label for each pixel, we use the graph-cut based  $\alpha$ - $\beta$  swap algorithm [6]. As beginning of the procedure, we first construct a graph with the overlapped region of the images. In addition, outer borderline of the overlapped region also participates to make graph consider surrounding condition. The energy of label  $f_i$  for pixel  $p_i$  is defined as

$$E(f_i) = E_d(f_i) + E_s(f_i, f_j) \quad (6)$$

where  $E_d$  is data cost,  $E_s$  is smoothness cost, and  $f_i, f_j$  are label assigned to pixel  $p_i, p_j$  respectively. The data cost is unary energy when  $p_i$  has label  $f_i$ , and the smoothness cost is pairwise energy when  $p_i$  has label  $f_i$ , and its neighbor pixel  $p_j$  has label  $f_j$ . We adopt data cost equation from [7], which is defined as

$$E_d(f_i) = \|\nabla I_{f_i}(p_i)\| \quad (7)$$

where  $I_{f_i}$  is color intensity of image which is assigned with label  $f_i$ . The smoothness cost is defined in [7] as

$$E_s(f_i, f_j) = \left( \|I_{f_i}(p_i) - I_{f_j}(p_i)\| + \|I_{f_i}(p_j) - I_{f_j}(p_j)\| \right) + \left( \|\nabla I_{f_i}(p_i) - \nabla I_{f_j}(p_i)\| + \|\nabla I_{f_i}(p_j) - \nabla I_{f_j}(p_j)\| \right) \quad (8)$$

By equations (7) and (8), the cut of the graph tend to occur in the pixels which have least difference in color and gradient.

## 3. RESULT

In our experiments, we compare our result with state-of-the-art software, Photoshop CS5's Photomerge function [8], and

recently proposed algorithm, Dual-homography method [4]. Photoshop uses a single cylindrical homography and Dual-homography uses two homography to express the distant plane and the ground plane. We use test images which include more than two homography, and do not have large translation of camera center.

Figure 3(a) shows misalignment in the image which is stitched by Photoshop. The upper side of seam-cut border is not well-matched. It implies that making panorama with a single homography is only suitable for stitching images with simple geometrical relation. However, our result in figure 3(b) does not have misaligned area because we consider various spatial relations among input images via multiple homography.

Figure 4 is comparison of Photoshop, and Dual-homography method and our method. As shown in figure 4(b), it solves the misalignment problem which is caused by using single homography as figure 4(a). However, distortions and windiness appear on the far region from the matching point pair clusters. It is due to interference of the opposite side homography. As explained above, our method solves this problem by introducing the flood-distance with increasing factor (figure 4(c)).

The computation time of our approach with 1200×960 input image is as follows. As first phase, optimal homography set estimation runs in around 2~3 seconds while dual-homography need 5~10 seconds. The second phase takes 5~6 seconds for the homography synthesis and warping. As the last phase, computation time of the stitching is 5~10 seconds. It means that our method not only generates qualified stitched result, but also is faster than Dual-homography by introducing flood-distance factor.

#### 4. CONCLUSION

In this paper, we propose an advanced feature-based panorama construction method, and our method can generate naturally stitched image using the multiple homography even if input images have various geometric shapes. This cannot be achieved via existing single homography method or Dual-homography method without post processing. Our proposed method finds the proper number of homography to express the relation among the matching points of an image pair. The synthesized homography of multiple homography for each pixel can be estimated fast by using flood-distance to generate panorama image. Experimental results show that our proposed method generates the most naturally stitched image rather than state-of-the-art methods.

#### 5. REFERENCES

[1] M. Brown and D.G. Lowe, Recognizing Panoramas. In Proceedings of the 9th International Conference on Computer Vision (ICCV2003), pp. 1218-1225, 2003.

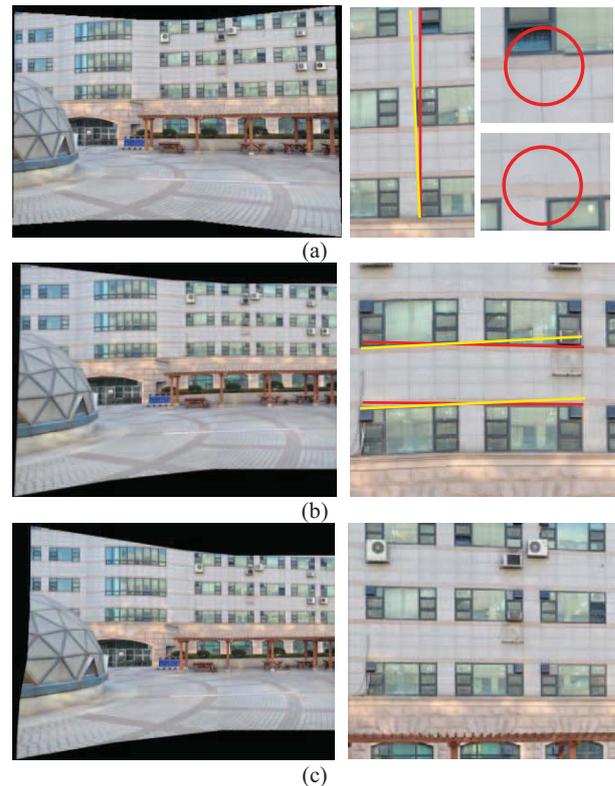


Figure 4. Result comparison (a) Photoshop PS5, (b) Dual-Homography, (c) Our result.

[2] Jiang Gao, Texture Dependent Feature Selection for Panorama. 16th IEEE International Conference on Image Processing (ICIP2009), pp. 1353-1356, 2009.

[3] H.-Y. Shum and R. Szeliski, Construction of Panoramic Image Mosaics with Global and Local Alignment. International Journal of Computer Vision, vol. 36(2), p.p. 101-130, 2000.

[4] J. Gao, S. J. Kim, and M. S. Brown, Constructing Image Panoramas using Dual-Homography Warping. Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), p.p. 49-56, 2011.

[5] D. G., Lowe, Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision (IJCV), vol. 60(2), p.p. 91-110, 2004.

[6] Y. Boykov, O. Veksler, and R. Zabih, Fast Approximate Energy Minimization via Graph Cuts. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 23, no. 11, p.p. 1222-1239, 2001.

[7] A. Agarwala, M. Dontcheva, M. Agrawala, S. Druncker, A. Colburn, B. Curless, D. Salesin, and M. Cohen., Interactive Digital Photomontage. ACM Transactions on Graphics, vol. 23(3), p.p. 294-302, 2004.

[8] Adobe Photoshop CS5, <http://www.adobe.com/products/photoshop>