International Journal of Hybrid Information Technology
Vol. 10, No. 9 (2017), pp.1-18
http//dx.doi.org/10.14257/ijhit.2017.10.9.01

Robust Multi-homography Method for Image Stitching under Large Viewpoint Changes

Wuxia Yan[1], Chuancai Liu[2*] and Furong Peng[3]

[1,2,3] *School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China*
[1]*wxyan127@outlook.com*, [2]*chcailiu@163.com*, [3]*pengfr@njust.edu.cn*

Abstract

Image stitching technique has many applications in computer vision. Traditional approaches have achieved much success under the assumption that the input images must have little or no parallax. However, this assumption cannot be guaranteed in real-world problems, for example, building images with ground under large viewpoint changes. The traditional methods only use a whole homography for alignment, which will result in severe distortions. To deal with the problem of stitching building images with ground under large viewpoint changes, we propose a robust multi-homography image composition method. It mixes multiple homographies for stitching building images with ground under large viewpoint changes. By calculating different homographies from different types of features, multiple homographies are then be blended with Gaussian weights to construct panorama. The chosen images under various conditions are used to evaluate our proposed method with other methods. The experimental results demonstrate that our proposed method performs higher alignment accuracy and more natural panorama than the traditional projective transformation and some representative state-of-the-art image stitching methods.

Keywords: *image stitching, multi-homography, Hough transformation, projective transformation*

1. Introduction

Image stitching method is the process of integrating images with overlapping regions into an entire and seamless panorama with high resolution. It has been widely used in the field of computer vision. Furthermore, lots of commercial products of image stitching for computer, digital cameras and mobile devices are popular. However, these methods and products can only work well with the constraint that input images must be taken under little or no parallax. This constraint implies that images have to be taken by cameras with little rotation or under the condition that the camera is sufficiently far away from the scene. Otherwise, artifacts such as ghosting, seam, distortion or broken image structure will appear on the final panoramic images.

To the best of our knowledge, most image stitching algorithms share the same pipeline: image matching, image alignment and image blending. Moreover, the most important step is image alignment, from which we can get the transformation model. Image alignment includes linear (rigid) methods and non-linear (non-rigid) methods [1]. Here, both the two methods have pros and cons of themselves. Such as, the linear transformation is often used in global warping methods. It can preserve global image structure but cannot handle projective deformation which introduced by large viewpoint changes. In contrast, non-

Received (August 2, 2016), Review Result (November 21, 2016), Accepted (December 5, 2016)
* Corresponding Author

ISSN: 1738-9968 IJHIT

linear transformation was often used in local warping models. It can deal with large parallax better than linear transformation, but cannot preserve global image structure.

In general, existing methods have two ways to deal with the misalignment and distortions caused by the large parallax:

(1) Improving alignment function. An intuitive approach to get more accurate stitching results is to construct a better alignment function. By combining the linear and non-linear transformation approaches, we can reduce the locally geometric distortion over the entire panorama [2] to get a more natural result. These non-linear transformations include mesh warping, local warping and image morphing. Several pieces of work have shown the validity of combining with global warping and local warping. For example, Rueckert et al. [3] divided the deformation function into affine projective transformation and B-spline transformation. Leventon et al. [4] added a pre-alignment process to elastic alignment. However, these methods have the same assumption that input images must be taken under little or no viewpoint changes. In addition, the union of warping methods introduce high computational complexity and discontinuity to image alignment.

(2) Post-processing methods. Such as, seam cutting method [5], shape-preserving image resizing method [6] and other methods. They are used to optimize image stitching results. The seam cutting methods find the optimal seam line from overlapping regions to align images. A weighted blending strategy for pixels around the seam line is applied to smooth the stitching transition. However, it requires additional computing resource to find the optimal seam line. On the other way, image resizing methods [6] are attempted to ensure that the local regions of the input image undergo a geometric similarity transformation while preserving the edge structure.

However, in practical use, it is hard to get convincing panoramas by these two ways of handling distortion on building image with ground under large parallax, as shown in Figure 1. It is also difficult to get convincing stitching results by traditional stitching method. Because the building parts have richer textures than the smoothed ground parts, the traditional transformation matrix can only be got from the matched features of the building parts. This results in distortion on the ground parts of the panoramas. Therefore, in this paper, we focus on building images with ground under large viewpoint changes.

Figure 1. Building Image with Ground

Here, for this type of images, a straight forward strategy is to improve alignment function. In this paper, we discuss how to deal with this case using the multi-homography method. The main contributions of this paper are summarized as follows:

(1) We present a robust image stitching method, which is based on non-linear transformation for image stitching. It can handle the case when the viewpoint of building image with ground changes dramatically.

(2) The proposed method with multi-homography preserves more image information than the traditional projective transformations, which have advantage in stitching images like building images with ground. It can also be easily extended to input images which containing more than two predominant parts.

(3) Different feature descriptors are used to describe image features on different image parts in this paper. More specifically, some feature descriptors are failed in extracting feature from smooth ground parts or failed in feature matching. By comparing the performances of several feature extraction algorithms, the one with the best performance is chosen as the feature descriptor for the ground part in this paper.

This proposed method focuses on the multi-homograghy stitching method. It includes the following steps: content-based image segmentation, affine-invariant feature extraction and weighted warping with multi-homography. The rest of this paper is organized as follows: the related image stitching methods are reviewed in Section 2. Stitching process based on multi-homography is systematically introduced, then stitching on building images with ground are explained in details in Section 3. In Section 4 the proposed method is verified through three groups of experiments, and the final conclusion is drawn in Section 5.

2. Related Work

On the observation that arbitrarily non-linear transformation can be got by the non-linear stitching methods, moreover, linearly projective transformations which used the global warping introduce severe projective distortion in size and shape. There is a strong motivation to combine these two alignment approaches to get a more robust image stitching method. Following this motivation, many local warping models have been proposed recently to reduce the image distortion caused by globally linear image stitching methods. For example, smoothly varying affine (SVA) approach [7] estimated a globally affine transformation and relaxed by EM formulation to get a smoothly varying affine model. As-projective-as-possible (APAP) warping method [8] was a locally projective transformation. It allowed local deviations for better alignment accuracy. Shape preserving half projective (SPHP) warping method [9] spatially employed projective transformation for overlapping regions and similarity transformation for non-overlapping regions. Chang et al. [10] combined the pixel-based method and feature-based method for image stitching, and followed by a shape preserving aggregation. Mesh warping was used to parameterize image stitching to alleviate distortion problem. Zhang *et al.*, [11] proposed method for stitching stereo images by stitching the left input stereo images, the disparity maps, and the right input stereo images step by step. Zhang *et al.*, [12] proposed parallax tolerant image stitching method by locally formulating image warping problem as a mesh warping problem. Lin *et al.*, [13] used local projective transformation and a global similarity transformation on non-overlapping regions. All these models attempted to use locally non-linear warping to get more natural panoramas.

For images containing two or more predominant parts, feature-based image stitching methods take no account of image content and feature distribution. Thus, Gao *et al.*, [14] proposed the dual-homography warping to handle the special scene images. It contained two predominant planes. Different weights were provided by the moving direct linear transformation (DLT) method [15] for the pixels that close or far from the seam line. However, it is very difficult to alleviate projective distortion. In order to separate the feature points from the two planes (distant plane and ground plane [14]), Gao et al. took two seed points from the top and bottom of the image. After separating the two planes, the

corresponding feature points were got from the two planes separately and the outliers were removed by RANSAC (Random Sample Consensus) [16]. The whole procedure was complicated. It included the acquiring of spatially varying weights, post-processed seam line blending and image resizing. In addition, the way of getting the weights by scanning all the refined features resulted in low efficiency. These processes also resulted in high complexity.

Therefore, a robust multi-homography method to stitch images containing two or more predominant parts is proposed in this paper. We take the building images with ground part under large viewpoint changes as an example and extend this method to the general case. For buildings with rich texture on the top of image and ground part with less or no texture on the bottom, it is reasonable to get different homographies, respectively. Hence, we exploit multi-homography to get more accurate alignment than projective transformation with single homography. The main difference between the dual-homography method [14] and our method is that we segment the image based on image contents before getting the corresponding homographies. Then, blend the entire images with a different strategy.

Let $p=[x,y]^T$ and $q=[x',y']^T$ be the matched points from image pair $\{I_0,I_1\}$. I_0 and I_1 can be written as the following equations:

$$\begin{aligned} f_0(\cdot) &= f_0(p) \\ f_1(\cdot) &= f_1(q) \end{aligned} \tag{1}$$

where $f \in R^{2\times 2}$.

In homogeneous coordinates $\overset{\sqcup}{p}=[p^T,1]^T$, $\overset{\sqcup}{q}=[q^T,1]^T$ and $f=[f,1]$. The alignment process is to transform I_0 to I_1 with geometric transformation $T(\cdot)$:

$$f_1(\cdot) : T(f_0(\cdot)) \tag{2}$$

where $:$ indicates equality up to scale and the transformation function $T \in R^{3\times 3}$. The cost function M is to measure difference of image and the transformed image with varies strategies as follow:

$$M = \min(f_1, T(f_0)) \tag{3}$$

or

$$M = \max(f_1, T(f_0)) \tag{4}$$

Generally, there are two independent means about function M: (1) M is function of measuring diversity. Then, the parameters of geometric space can be got by searching for the globally minimized value. Hence, M is distance function, sum of squared difference (SSD) function or sum of absolute differences (SAD) function. (2) M is the function measures the similarity. Then, the parameters of geometric space can be got by searching for the globally maximized value. Hence, M also present by cross-correlation function or mutual information. Hence, in this paper, we use SAD function as shown in Equation 5 to get the minimized value for measuring the alignment accuracy. SAD method is to minimize the sum of absolute differences between the overlapping pixels.

$$SAD_{x,y} = \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} \left| f_{x+i,y+j} - T(f)_{i,j} \right| \tag{5}$$

Furthermore, the outline of the proposed method is shown in Figure 1. It consists of content-based image segmentation, invariant feature extraction and weighted warping with two homographies. The details are as follows.

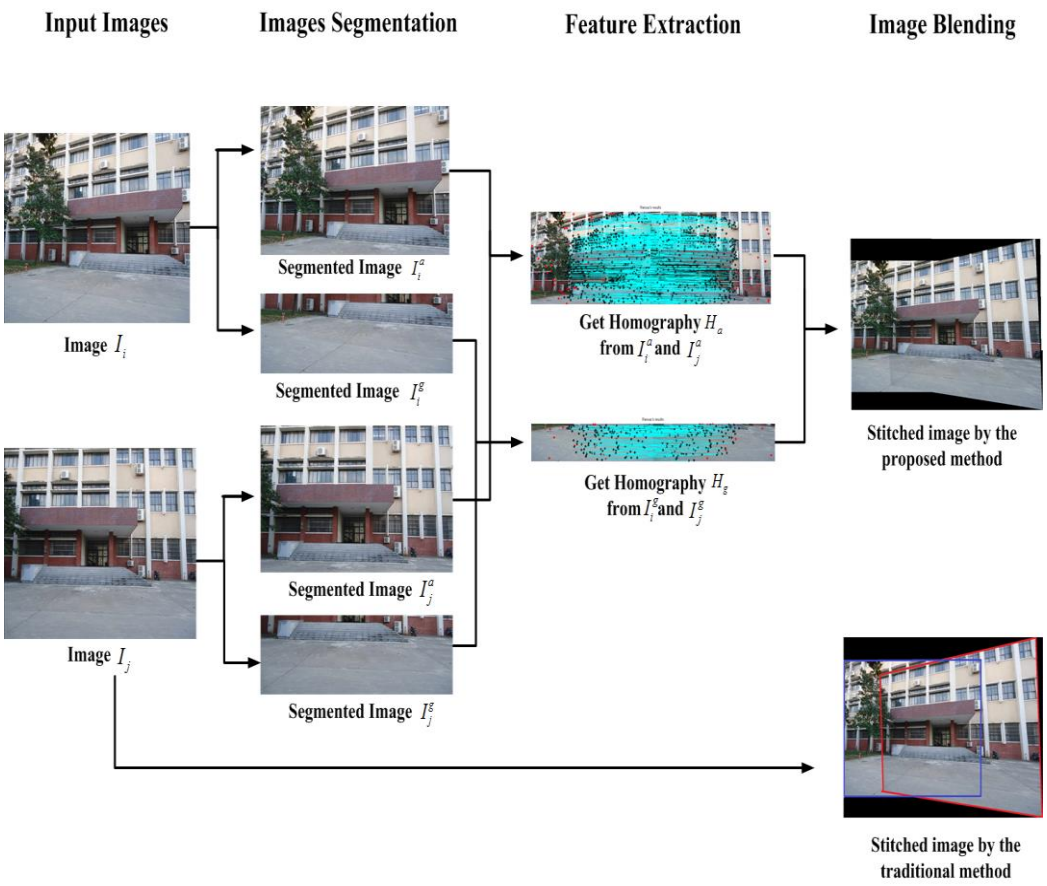


Figure 2. The Framework of the Proposed Method

3. Content-based Image Segmentation

To solve the problem of how to detect the segmented line on the images, Gao et al. [14] adopted K-means clustering with two seed points based on spatial location to get the homographies. However, it took too much time to get homographies for projective transformation. Therefore, the initial idea of this paper is to segment the image into building part and ground part directly. Then, we get the homography, respectively. In this section, content-based image segmentation method for detecting the line to segment the building and the ground is presented. We only deal with images containing building on the top and the ground on the bottom.

We begin with segmenting a pair of overlapping images $\{I_0, I_1\}$ into different parts $\{I_0^b, I_0^g\}$ and $\{I_1^b, I_1^g\}$ (the building part and the ground part, respectively) according to their contents. We project I_0^b (I_1^b) and I_0^g (I_1^g) from image I_0 (I_1) using forward image mapping method with bilinear transformation. The problem in Equation 2 can be rewritten in the following projective form

$$\overset{\text{도}}{q} = H\ p \tag{6}$$

where $H \in R^{3\times3}$. Homography H_0 ($H_0 \in R^{3\times3}$) can be got by matching the corresponding invariant feature transformation (SIFT) [18] from the building part. H_1 ($H_1 \in R^{3\times3}$) can be computed by matching the corresponding Affine-SIFT (ASIFT) [18] features from the ground part. The process is simplified by segmenting the image into two different parts, rather than using a clustering method to group the feature points.

On the other way, we use the plentiful relationships in man-made structures for line detection. The geometric characteristics of architectural scenes, such as, the parallelism and orthogonality of lines [19-20] are exploited in this paper. Furthermore, the detection of line segment is an important and foundational issue in computer vision. The most well known feature extraction technique is Hough transformation (HT) [21-23]. It is concerned with the identification of lines in the image. It replaces line detection in Cartesian coordinate space with point detected in Hough space. We get the parallel line segments of the building by Hough transformation. Then, use the line l which located at the bottom of the building as the segmented line. The experimental results in Section 6.1 have validated our supposition.

4. Affine-invariant Feature Extraction

Different feature descriptors are applied to detect features on the two predominate image parts. After we failed in testing many other affine-invariant feature descriptors, such as SIFT, MSER (maximally stable extremal regions) [24], SURF (speeded-up robust features) [25], SUSAN (smallest univalue segment assimilating nucleus) [26] and Harris [27], we chose SIFT and ASIFT as the feature descriptor of the building part and the ground part, respectively. The specific details are in Section 6.2.

5. Weighted Warping with Multi-homography

To prevent the unstable distribution of the homographies in blending images, the weighting coefficient serves to regularize the process. The projective distortion from different transformations was greatly reduced.

We model the homography H for the pair of overlapping images I_0 and I_1 from the two homographies H_b and H_d as

$$H = (1-w_0)H_b + w_0 H_g \tag{7}$$

The weighting factor w_0 is based on various distance from corresponding pixels of the two parts. It assigns low weights to pixels that far away from image segmented boundaries and high weights otherwise. The weight w_0 is given by the Gaussian weights as:

$$w_0 = \exp(-\|x_0 - x_1\|^2 / \gamma^2) \tag{8}$$

where x_0 and x_1 is the corresponding pixel coordinates come from I_0 and I_1. γ is a scale parameter which chosen based on the experimental conditions. We set γ as 10 in the experimental section.

We further extend this two homographies example to general image stitching case, where the images containing two or more predominant parts, with multi-homography under a weights function as follow:

$$H = w_0 H_0 + w_1 H_1 + ... + w_n H_n \tag{9}$$

where $n \geq 0$ and $w_i = w_i / (w_0 + .. + w_i + .. + w_n)$. This equation is suitable for images containing $n+1$ predominant parts. The solution is the same as the process in this paper. Firstly, we segment image based on contents. Then, different features are extracted and multiple homographies can be obtained. Finally, seamless panoramas are got with weights function in Equation 9.

6. Experimental Results and Analysis

Three experiments are designed to validate our proposed method from three parts of the entire stitching pipeline as follows: feature extraction, Hough transformation, comparing the results of our method with traditional projective transformation method and some representative state-of-the-art image stitching methods.

In our experiments, we select input images with large viewpoint changes that due to large rotation and translation of camera. An extensive range of input images under large viewpoint changes are shown in Figure 3, which including different types of images as follows:

(1) The ground planes with different types. Some images with smoothed ground plane (Figure 3(a), 3(b), 3(m), 3(n), 3(o) and 3(p)), some ground plane are covered with shadow (Figure 3(c), 3(d), 3(i) and 3(j)) and the others are lawn (Figure 3(e) and Figure 3(f)) or floor tiles (Figure 3(g) and Figure 3(h), Figure 3(k) and Figure 3(l)).

(2) The buildings are different too. Some buildings have rich texture. The others are the smoothed wall (Figure 3(m), 3(n), 3(o) and 3(p)).

The proposed method typically takes from 8 to 13 seconds on PC with 2.93 GHz CPU and 2G RAM to stitch images with resolution under Windows 7 OS and Matlab 2014a system platform.

6.1. Experiment for Hough Transformation

In this paper, we suppose that the images with building occupied the top position and ground plane located at the bottom. The building part contains rich texture in architectural scenes, but the smoothed ground part contains less or no texture. This reminds us to use the parallel line on the building as the segmented line. We employ constraints available from geometric relationships of the buildings, especially the parallelism. The parallel line which lies at the lowest location of the building is chosen as the segmented line of the image.

To detect the lines effectively we employ Hough transformation. The detection results are shown in Figure 3. From all the detected lines, the green lines on each image (some of them are the stairs of buildings, some are the segmented line of ground and building) are selected as the separating line for image segmentation. The results have validated our supposition and meet the requirements of this proposed method for content-based image segmentation.

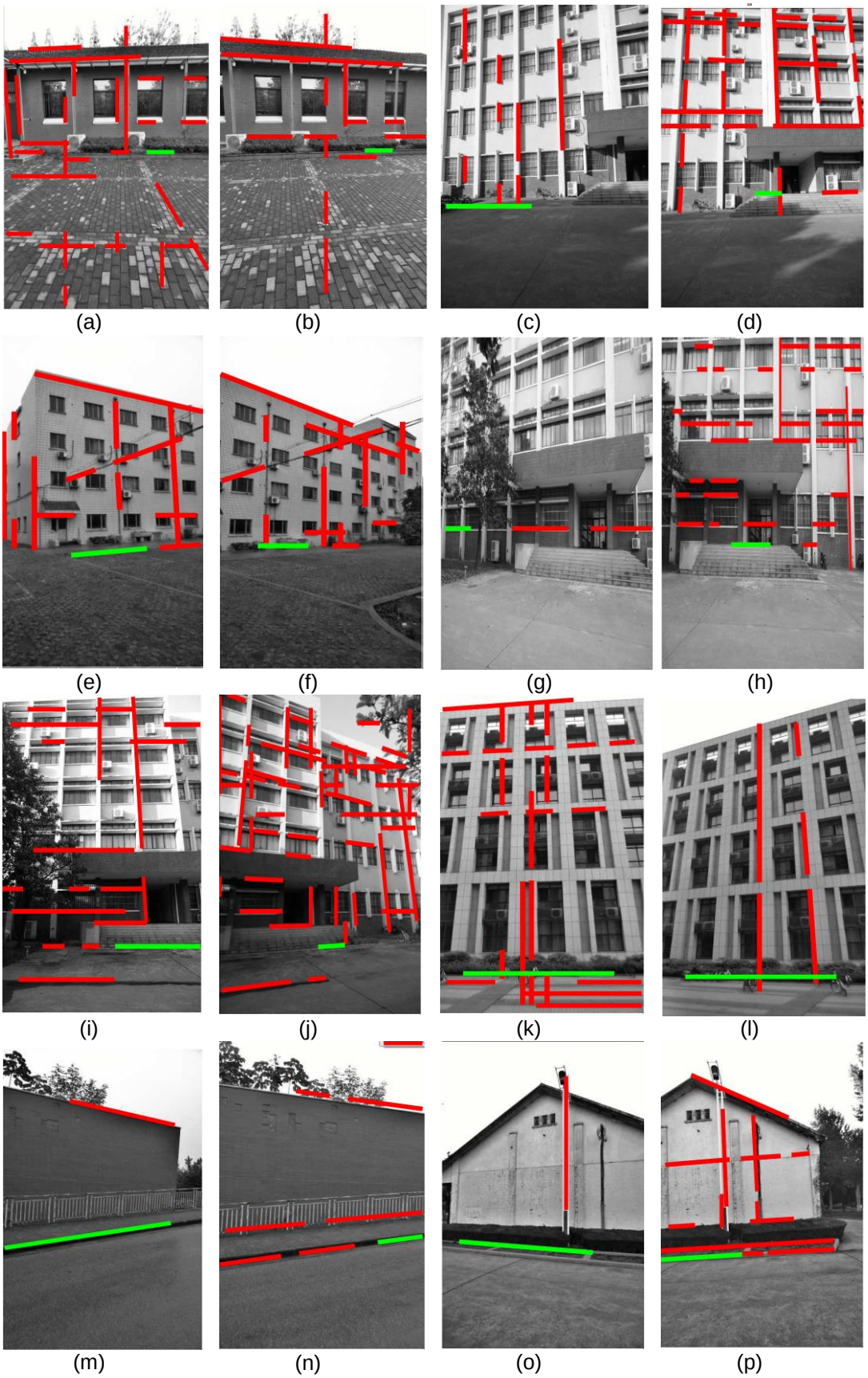


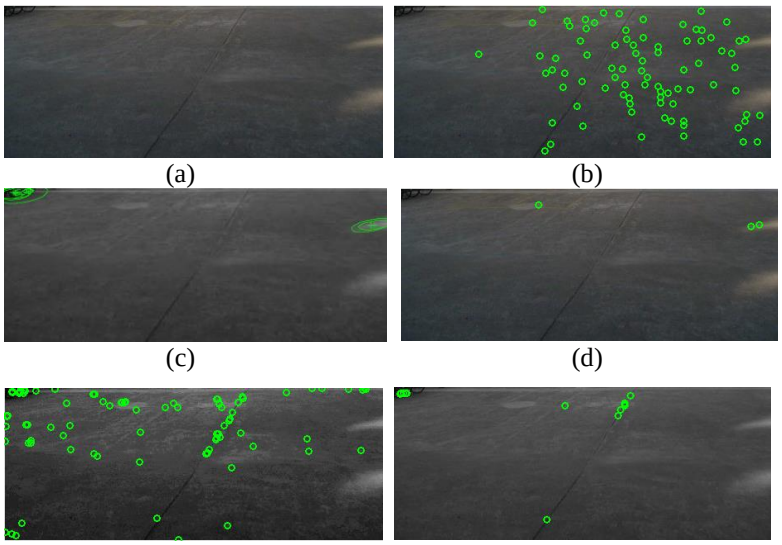
Figure 3. Hough Detection Results on Different Image Pairs

6.2. Experiment for Feature Extraction on Ground Parts

The input images under various conditions are evaluated in these experiments. After segmenting the input images with the line segment detected by Hough transformation, we focus on testing the ground parts in this section, as shown in Figure 4. The experimental results on the left ground part (Figure 4(a)) with different algorithms are shown in Figure 5.

(a) (b)

Figure 4. A Pair of Corresponding Ground Part. (a) Comes from the Left Image and (b) Comes from the Right Image



(a) (b)

(c) (d)

(e) (f)

Figure 5. Different Detection Results on Ground Plane by Algorithms. (a) SIFT, (b) ASIFT, (c) MSER, (d) SURF, (e) SUSAN and (f) Harris

We tested different feature descriptor algorithms to extract features on the smoothed ground parts for image stitching in this section. The performances of all the six algorithms on a pair of input images (Figure 4(a)) are provided in a noticeable way in Figure 6. The results on all the eight pairs of ground parts are shown in Figure 8. Three fields are evaluated: (1) the number of extracted features on the ground part of the left images (Figure 8(a)) and (2) the corresponding right image (Figure 8(b)); (3) the number of matched features after removing the outliers (Figure 8(c)). The experiment results indicated that it is easy to extract features on the rich texture images, such as, the ground with floor tiles (the first and the fourth pairs of images in Figure 8) and the grass filed (the seventh pair of images in Figure 8). The performances on slippery terrain are much lower.

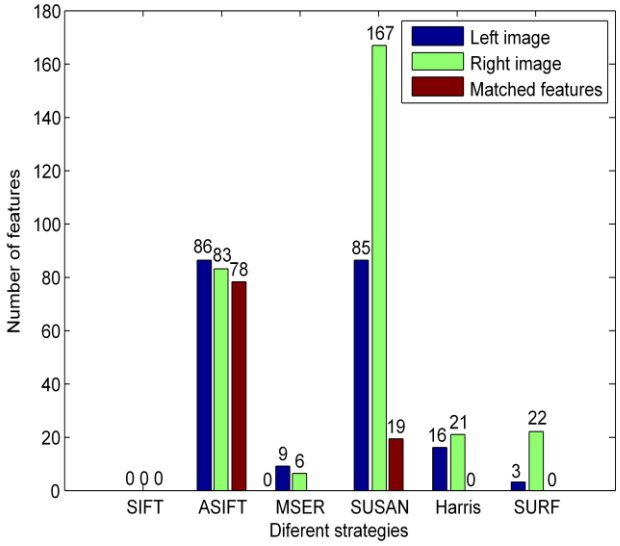


Figure 6. The Detected Features of Ground Plane

Unexpectedly the performance of SIFT is the lowest. SIFT descriptor which is proposed by Lowe [17] is robust to the changes of rotation, scaling, illumination and local affine distortion. However, it cannot suitable for the applications on smooth regions in this paper. It is difficult to detect the main direction of the gradient in each pyramid level in such smooth images. ASIFT improves SIFT with the two left over parameters (the two camera axis orientation parameters) to become more robust to images with large viewpoint changes. We can see that the performances of ASIFT on feature extracting and feature matching achieve the highest value from all the six algorithms. Take Figure 4 as an example, the performance of ASIFT for feature extracting is shown in Figure 5(b) and the feature matching result is shown in Figure 7. SUSAN and ASIFT methods outperform the rest of other methods. SUSAN which is based on non-linear filtering has been used as edges and corners descriptor. It is not sensitive to images noise, but it may get false corner features if boundaries are blurry. SUSAN performs well in this test on the smoothed ground parts, but the performance of matched features is not as steady as ASIFT.

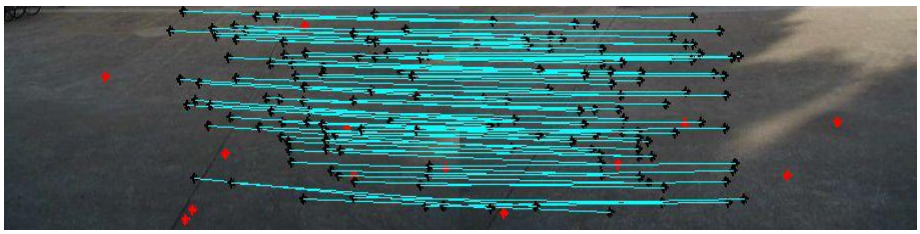
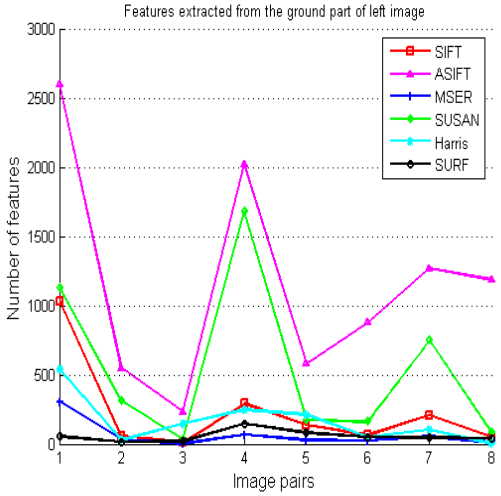
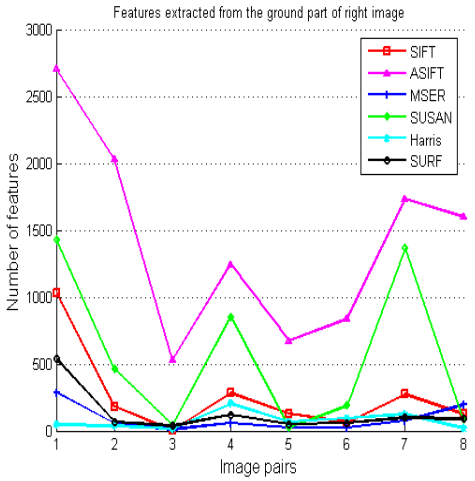


Figure 7. The Matched Result after Using RANSAC

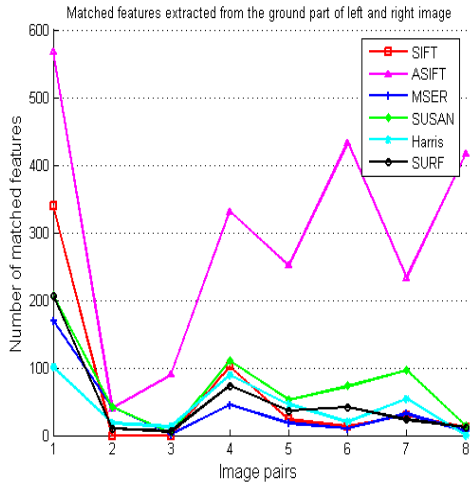
From Figure 8, we can see that the rest of other methods have almost the same low performances. The reasons are as follows: MSER based on watershed segmentation is to find blob-like structures as the distinctive property, but it performs as low as SIFT on smooth images. Harris represents the gradient information with a structure tensor; therefore, it fails to get features in many pairs of input images with fewer gradients from the smoothed ground parts. It fails in the feature matching either. SURF descriptor is robust to common image deformations, but it fails in viewpoint changes and illumination invariant.



(a)



(b)



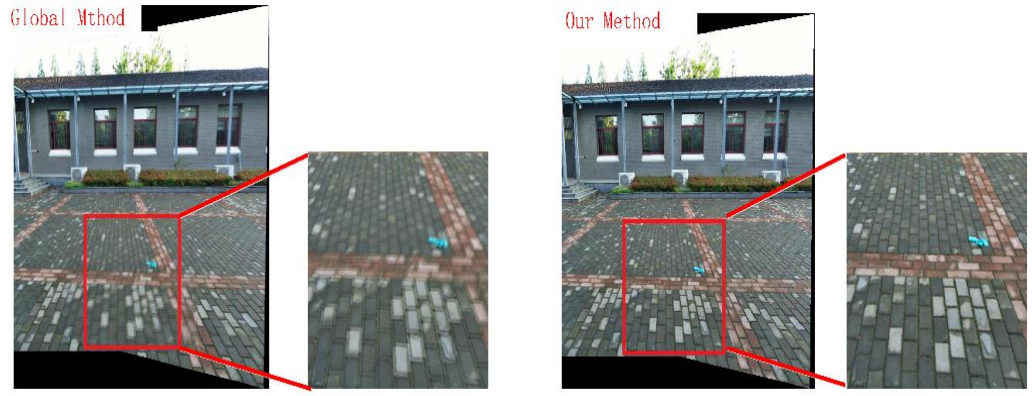
(c)

Figure 8. Comparing Numbers of Extracted Features and Matched Features. (a) Extracted Features on the Ground Part of the Left Image, (b) Extracted Features on the Ground Plane of Right Image and (c) the Matched Features between Different Methods

The experimental results demonstrate that not all the feature descriptors are suitable for describing features in images with less or no texture. We employ SIFT as feature descriptor for the building parts and ASIFT as feature descriptor for the ground parts in this paper.

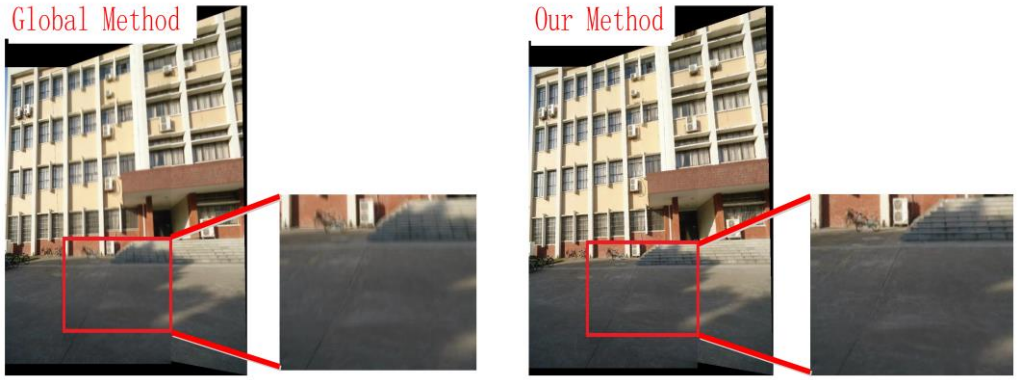
6.3. Comparing of Different Image Stitching Methods

SIFT and ASIFT feature descriptors are used to extract features on building and ground parts. Hence, after removing the outliers by RANSAC, the exact features matched by SAD algorithm. The homography is calculated through projective transformation stitching pipeline. Then, instead of stitching the segmented image parts respectively, we get the entire stitched image with the multi-homography which scored by Gaussian weights as described in Section 5. The results are shown in Figure 9.



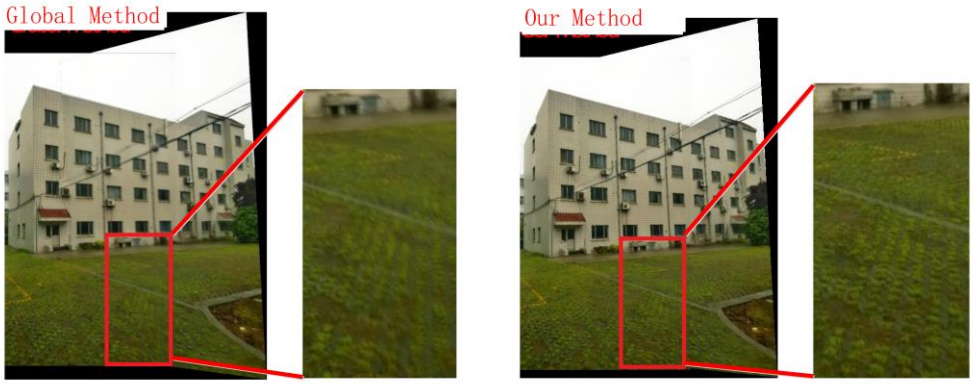
(a)

(b)



(c)

(d)

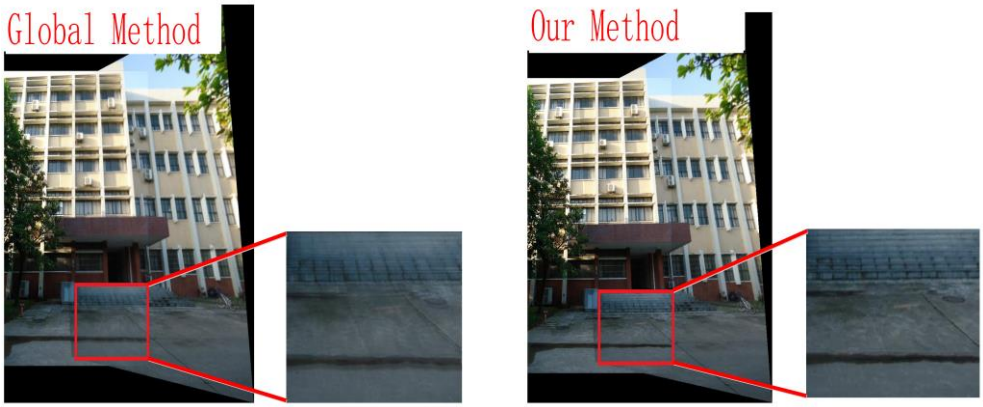


(e)

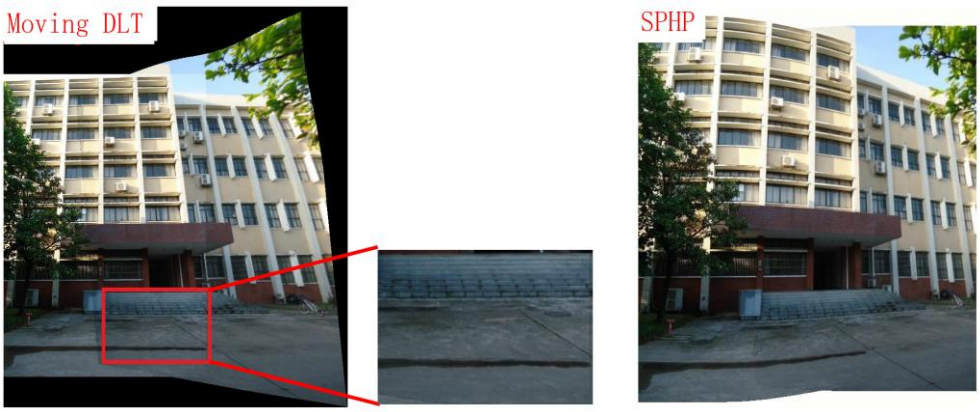
(f)

(g)

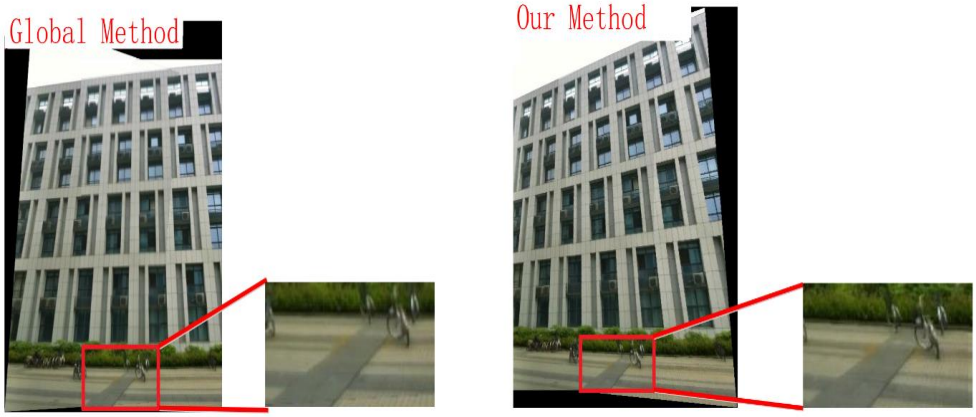
(h)



(i)

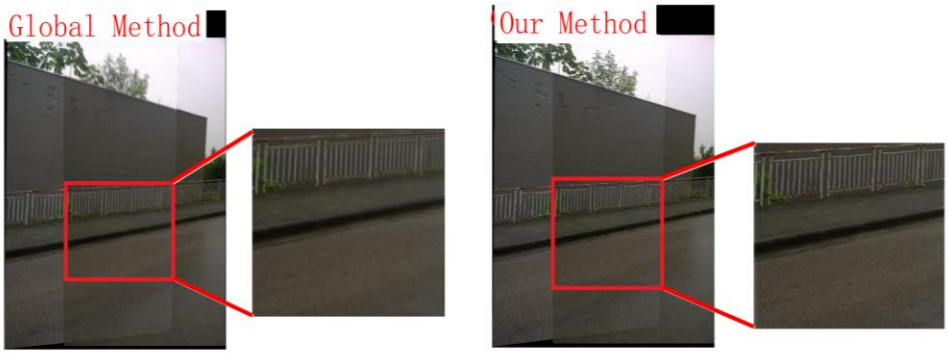


(j)



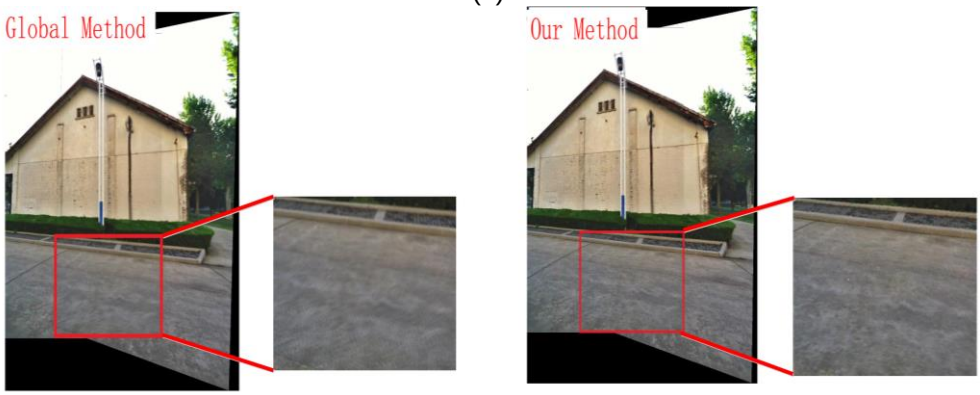
(k)

(l)

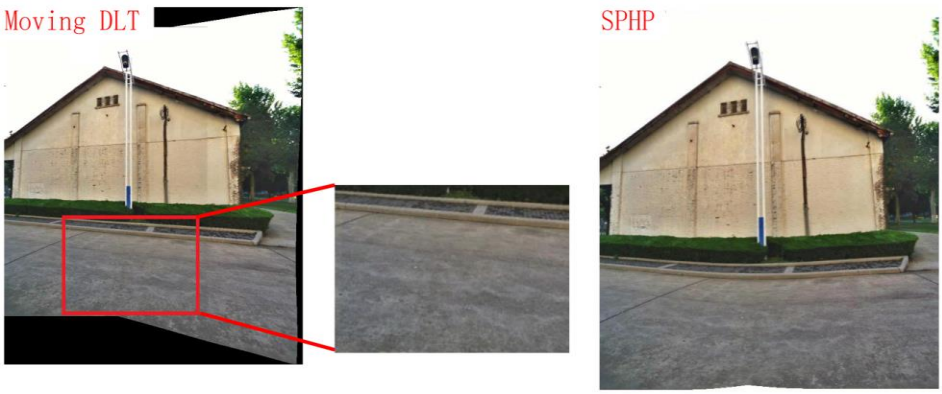


(m)

(n)



(o)



(p)

Figure 9. Comparisons of Various Stitching Methods

In Figure 9, we compared the proposed method to the traditional global projective transformation image stitching method, which include the state-of-the-art methods, *i.e.*, moving DLT method and SPHP method, on a range of challenging images with large viewpoint changes. All the results appear good at first glance. However, we highlight the noticeable regions where result from the global transformation, moving DLT and SPHP exhibit misalignment.

Here, it should be note that most of the misalignment happen in the bottom part of the panoramic image, such as, the stairs and the ground part. This implies that the single-homography based approach (*i.e.*, the global method) is got with the matching features on the building part of image, resulting in misalignments in the ground parts. Moreover, because moving DLT method cannot align the stairs well, the stairs suffer from ghosting, as indicated by the red square in Figure 9(b), 9(d) and 9(f). We can see from Figure 9(b), 9(d), 9(f), 9(h), 9(j), 9(l), 9(n), 9(p), the SPHP method with three steps in the stitching process can preserve the image contents well. However, it failed to preserve image structure. Therefore, comparing with these methods our method maintains high accuracy alignment and shows robustness to large viewpoint changes.

7. Conclusion

This paper has presented a multi-homography image stitching method under large viewpoint changes. Multiple homographies are employed for different image parts. Then, they blended with Gaussian weights to lessen projective distortion. We further extend the proposed method to stitch images with largely different blocks. Experimental comparisons to the traditional projective transformation and some of the state-of-the-art stitching methods show that the proposed method achieves higher accuracy alignment under large viewpoint changes.

References

[1] B. Zitova and J. Flusser, “Image Registration Methods: A Survey”, Image and Vision Computing, vol. 21, no. 11, **(2003)**, pp. 977-1000.
[2] J. Jia and C. K. Tang, “Image Stitching Using Structure Deformation”, Pattern Analysis and Machine Intelligence, IEEE Transactions, vol. 30, no. 4, **(2009)**, pp. 617-631.
[3] D. Rueckert, C. Hayes, C. Studholme, P. Summers, M. Leach and D. J. Hawkes, “Non-rigid Registration of Breast MR Images using Mutual Information”, Medical Image Computing and Computer-Assisted Interpenetration MICCA, **(1999)**.
[4] M. E. Leventon and W. E. L. Grimson, “Multi-modal Volume Registration Using Joint Intensity Distributions”, IN MICCAI, **(1999)**, pp. 1057-1066.
[5] A. Agarwala, M. Dontcheva, M. Agrawala, S. Drucker, A. Colburn, B. Curless, D. Salesin and M. Cohen, “Interactive digital photomontage”, ACM Transactions on Graphics, vol. 23, no. 3, **(2004)**, pp. 292-300.

[6] G. X. Zhang, M. M. Cheng, S. M. Hu and R. R. Martin. "A Shape-preserving Approach to Image Resizing", Computer Graphics Forum, vol. 29, no. 7, **(2009)**, pp. 1997–1906.
[7] W. Y. Lin, S. Liu, Y. Matsushita, T. T. Ng and L. F. Cheong, "Smoothly Varying Affine Stitching", Computer Vision and Pattern Recognition (CVPR), IEEE Conference, IEEE, **(2011)**.
[8] J. Zaragoza, T. J. Chin, Q. H. Tran, M. S. Brown and D. Suter, "As-projective-as-possible Image Stitching with Moving DLT," Pattern Analysis and Machine Intelligence IEEE Transactions, vol. 36, no. 7, **(2013)**, pp. 339-2346.
[9] C. H. Chang, Y. Sato and Y. Y. Chuang, "Shape-preserving Half Projective Warps for Image Stitching", Computer Vision and Pattern Recognition (CVPR), IEEE Conference, **(2014)**.
[10] C. H. Chang, C. J. Chen and Y. Y. Chuang, "Spatially-varying Image Warps for Scene Alignment", Pattern Recognition (ICPR), 22nd International Conference, IEEE, **(2014)**.
[11] F. Zhang and F. Liu, "Casual Stereoscopic Panorama Stitching", Computer Vision and Pattern Recognition (CVPR), IEEE Conference, **(2015)**.
[12] F. Zhang and F. Liu, "Parallax-tolerant Image Stitching", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), **(2014)**.
[13] C. C. Lin, S. U. Pankanti, K. N. Ramamurthy and A. Y. Aravkin, "A Daptive As-natural-as-possible Image Stitching", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, **(2015)**.
[14] J. Gao, S. J. Kim and M. S. Brown, "Constructing Image Panoramas Using Dual-homography Warping", IEEE Conference on Computer Vision and Pattern Recognition, **(2011)**.
[15] S. Schaefer, T. Mcphail and J. Warren, "Image Deformation Using Moving Least Squares", ACM Transactions on Graphics, vol. 25, no. 3, **(2006)**, pp. 533–540.
[16] M. A. Fisher and R. C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography", Communications of the ACM, vol. 24, no. 6, **(1991)**, pp. 726–740.
[17] D. G. Lowe, "Distinctive Image Features from Scales-invariant Keypoints", IJCV, vol. 60, no. 2, **(2004)**, pp. 91–110.
[18] J. M. Morel and G. Yu, "Asift: A New Framework for Fully Affine Invariant Image Comparison", SIAM Journal on Imaging Sciences, vol. 2, no. 2, **(2009)**, pp. 439–469.
[19] D. Liebowitz, A. Criminisi and A. Zisserman, "Creating Architectural Models from Images", Computer Graphics Forum, vol. 19, no. 3, **(1999)**, pp. 39–50(12).
[20] P. Debevec, "Modeling and Rendering Architecture from Photographs: Modeling and Rendering Architecture from Photographs", **(1999)**.
[21] P. C. Hough, "Method and Means for Recognizing Complex Patterns", Ph.D. dissertation, US, **(1960)**.
[22] R. Duda and P. Hart, "Use of the Hough Transformation to Detect Lines and Curves in Pictures", Communications of the ACM, **(1972)**.
[23] N. Bennett, R. Burridge and N. Saito, "A Method to Detect and Characterize Ellipses Using the Hough Transformation", Pattern Analysis and Machine Intelligence IEEE Transactions, vol. 21, no. 7, **(1999)**, pp. 652–657.
[24] J. Matas, O. Chum and Urban, "Robust Wide-baseline Stereo from Maximally Stable Extremal Regions", BMVC, vol. 22, no. 10, **(2002)**, pp. 761–767.
[25] H. Bay, A. Ess, T. Tuytelaars and L. Van Gool, "Speeded-up Robust Features (surf)", Computer Vision and Image Understanding, vol. 110, no. 3, **(2009)**, pp. 346–359.
[26] S. M. Smith, "Susan-a New Approach to Low Level Image Processing", International Journal of Computer Vision, vol. 23, no. 1, **(1995)**, pp. 45–79.
[27] C. Harris and M. Stephens, "A Combined Corner and Edge Detector", Alvey Conference, vol. 53, no. 3, **(1999)**, pp. 147–151.

Authors

Wuxia Yan, she received her BS and MS degrees in 2009 and 2011 respectively in the school of Computer Science of Qufu Normal University, and she is currently pursuing her Ph.D. degree in the school of Computer Science and Engineering, Nanjing University of Science and Technology (NUST). Her current research interests include image processing and computer vision.

Chuancai Liu, he received the BS degrees in 1997, MS degrees in 1991, Ph.D. degrees in 1997 respectively, and he is currently professor in the school of Computer Science and Engineering, Nanjing University of Science and Technology (NUST). His current research interests include image processing and computer vision.

Furong Peng, he received the BS degrees in 2009 in the school of Computer Science of Nanjing University of Science and Technology, China. He is currently a Ph.D. candidate in the Computer Science and Engineering Department of Nanjing University of Science and Technology, China. His research interests include data mining, machine learning, and recommendation system.

