International Journal of Multimedia and Ubiquitous Engineering
Vol.10, No.12 (2015), pp.289-296
http://dx.doi.org/10.14257/ijmue.2015.10.12.29

Camera Image Stitching using Feature Points and Geometrical Features

Chi-Woong Kim[1], Hyun Seong Sung[2] and Hee-Jun Yoon[3,*]

[1,3] Korea Institute of Science and Technology Information, South Korea
[2]Agency for Defense Development, South Korea, wigman98@gmail.com
[1]jwkim@kisti.re.kr, [2]wigman98@gmail.com, [3]k2@kisti.re.kr

Abstract

To obtain a panoramic image by using ordinary cameras instead of an expensive panoramic camera, various researches are being studied. We usually use Stitching of the images using a homography matrix that has been evaluated by the method based on feature points such as SURF to obtain a panoramic image. However, the above method has some problems that it need a lot of time to extract feature points and feature points are obtained differently depending on various external factors. Proposed method is stitching the images using spatial transformation from two cameras at the same time. We expect to be able to apply proposed method to various electronic devices to obtain a panoramic image easily.

Keywords: *Image Stitching, Panoramic image, Spatial transformation*

1. Introduction

Panoramic image usually represents wide-angle view of landscape. According to people's desire what is to have more and more high-resolution picture, various high-resolution panoramic cameras have been developed actively. Especially, panoramic cameras using multiple lenses are also developed to cover wide-angle view. Moreover, algorithms to acquire a panoramic image using non-panoramic camera such as an algorithm to take a panoramic picture by moving smartphone laterally is being studied. Since the use of smartphones increases exponentially, most people can take a panoramic picture using their smartphone easily even if they don't have any panoramic camera. But, when you want to take a panoramic video, your smartphone can not make it up to now at least, it's because that non-panoramic single camera cannot capture large area at once. Other non-panoramic single camera applications have problems as described above. In order to overcome the problem of a single camera application, we studied the method for using the two cameras to obtain panoramic images and videos.

In this paper, we propose a method to obtain panoramic images and videos using fixed non-panoramic two cameras. To make a panoramic image using two cameras, we use two image stitching methods - based on feature point and based on geometrical feature method. These different methods are complementary to each other, and can derive the best result. In order to obtain panoramic images and videos with the proposed method, only two fixed cameras are needed. As the technology advances, the size of camera is miniaturized and the camera resolution is increased, so, it is expected that the proposed method can be applied to other small devices such as laptops, smartphones as well as normal cameras and it also can be used in various fields.

[*] Corresponding Author

ISSN: 1975-0080 IJMUE

2. Image Stitching Based on Feature Point Detection

A representative method to make panoramic image using multiple cameras or images is image stitching. For several years, a lot of related works to obtain panoramic images from multiple images have been studied. To create a panoramic image by stitching two different images, it needs to transform each image appropriately. So, it is most important issue to create a panoramic image from two different images how to transform two different images appropriately.

Elan Dubrofsky proposed a homography estimation method for the appropriate image transformation [1]. According to Elan's study, a homography matrix is an invertible mapping of points and lines on the projective plane. Homography matrix is transformation matrix to convert one plane to the other plane. Figure 1 shows how world plane can transform to an image plane by homography matrix. H of the equation in Figure 1 is homography matrix to convert 2D image plane to 2D world plane.

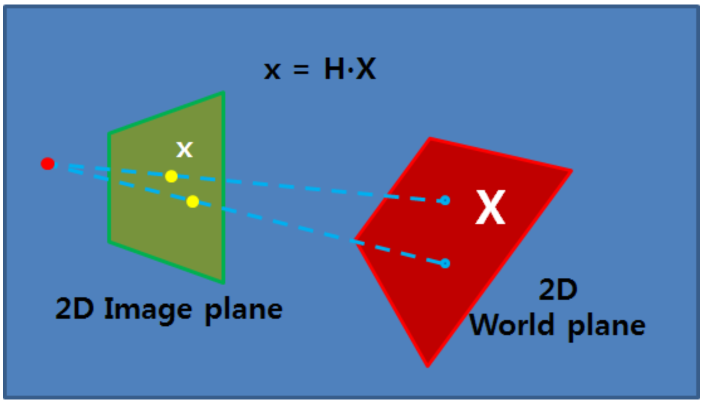


Figure 1. Plane-to-Plane Homography

In order to stitching two images located in separate image plane, it is necessary to convert from each image plane to world plane. So, the problem on how you can stitch two images accurately can be replaced by how you can find the homography matrix and how you can calculate the homography matrix more accurately. Most actively studied method to evaluate homography matrix is based on feature points matching. Coordinates of detected feature points from each image are compared and matched, and homography matrix is calculated from matched coordinates. So, various algorithms have been developed in order to find and to compare the feature points more accurately. SIFT [2], SURF [3], SFOP [4], IBR and EBR [5] are widely used feature detectors. Also in our study, we experimented the homography matrix evaluation using various algorithms for feature point detection and SURF algorithm shows the best performance in comparison with other algorithms. So, we use SURF algorithm for feature points extracting and getting the approximation of the nearest neighbor search [6] algorithm for matching. However, we found two problems when we use the homography evaluation algorithms based on feature point detection in our experiment.

First, feature points detecting and matching is largely depended on external factors such as brightness. Figure 2 shows feature point matching between two images of our office from two cameras. In Figure 2, a circle denote detected feature point, a line denote matching between feature points and the triangles denote the example of wrong feature points detecting and matching for a brightness of image. These wrong feature points matching induce the wrong Homography matrix extracting and it causes the very strange panoramic image obtaining. So, image stitching method based on fully featured point detecting and matching can be operated well in some specific environments or in some specific obtained images before. But it cannot be operated well in a situation that external factors are frequently changed. Second, feature point detecting and matching time is very long. Figure 3 shows that it needs very long time to calculate only one frame. In Figure 3, the processing time to make one panoramic image is more 2.8secs. Why it need a very long processing time using feature point algorithms in whole images is that the number of the pairs of feature points must be compared and increase exponentially. Our final goal is to make a panoramic image and video system using many normal cameras. So, if the number of cameras increases more and more, the processing time will increase exponentially and to obtain a panoramic image and video in a real-time will be impossible although the algorithm is optimized or the performance of hardware will be improved.

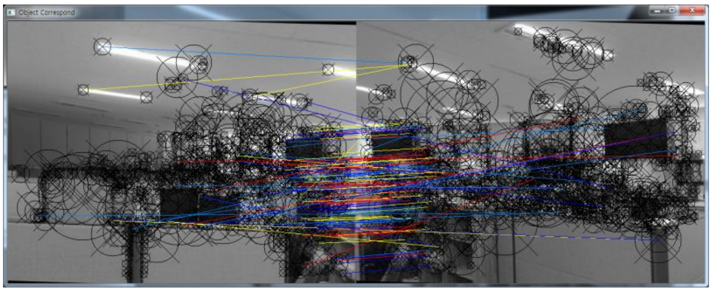


Figure 2. Example of Wrong Feature Points Matching

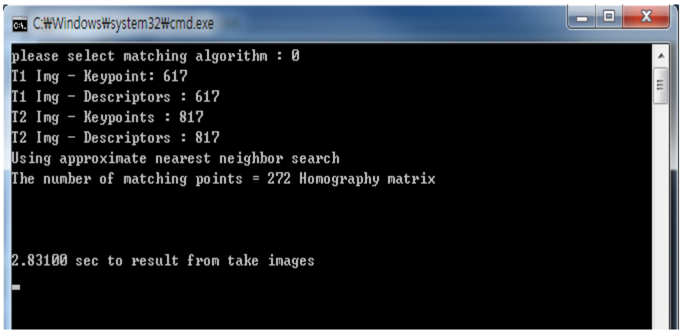


Figure 3. Processing Time for One Frame

In order to solve the problems of the method based feature point detection and to improve the performance of image stitching, we propose a new image stitching method. Proposed image stitching method is combination of the feature point based stitching algorithm and geometrical rotation of the images. Using the geometrical feature, we select the specific area in each image that the performance of the feature point matching can show the best. This method use the feature point detection based method and also use image rotation using a spatial relationship between the cameras. All printed material, including text, illustrations, and charts, must be kept within the parameters of the 8 15/16-inch (53.75 picas) column length and 5 15/16-inch (36 picas) column width. Please do not write or print outside of the column parameters. Margins are 3.3cm on the left side, 3.65cm on the right, 2.03cm on the top, and 3.05cm on the bottom. Paper orientation in all pages should be in portrait style.

3. Stitching Ased on Geometrical Feature

We have had various experiments to take a panoramic image from camera system configured as follows: we use two Logitech c920 webcams, each camera is rotated at an angle of 30 degrees to the left and right and two cameras are located at the same height. Feature points based stitching algorithms mentioned in session 2 mostly create a panoramic image from stored pictures but they do not use image which is taken directly from camera. After we study above algorithms, we found that these algorithms don't consider spatial feature of cameras. So, we focused on the relationship with the spatial position of cameras and acquired image from cameras. To improve the performance of stitching algorithm, we try out the new method combined of internal and external features properly. To solve the problem of feature points based algorithms which is greatly influenced by external factors, we use the geometrical external factors. And to solve the other problems that real-time processing is impossible because the number of pairs increase exponentially as the size of images increase larger, we reduce the feature points extracting area. Through these two points, we can reduce the processing time and the stitching accuracy is improved.

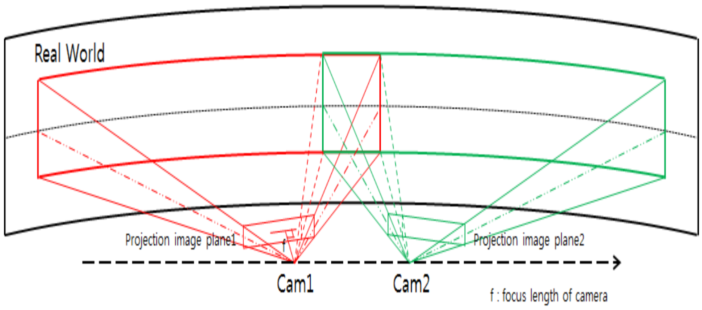


Figure 4. A Configuration Example of Two Cameras

When we suppose that two cameras are in the same coordinate space like Figure 4, real world is projected onto the virtual space the focal distance away from the cameras. Projection image plane 1 and projection image plane2 in Figure 4 are projected images from each camera. Projected images from each camera cannot be added up because each image plane is in a different coordinate space. For stitching two images in a different coordinate space, first, we rotate two planes to have the same normal vector in a real world space. Projection image plane 1 is rotated 30 degrees to the right, projection image plane 2 is rotated 30 degrees to the left as we install and then, it becomes that they have the same normal vector in a real world space. Second, in order to remove the overlapped area, we must translate the rotated planes to the x axis. However, translation value calculated from camera configuration can't be applied to the translation matrix because the units are different. So, we can use homography matrix described in section 2 to evaluate translation value. Generally, transformation of plane is performed by a combination of translation and rotation. Each (x,y,z) ∈ P is transformed to

$$\begin{bmatrix} r_{11}x + r_{12}y + r_{13}z + t_x \\ r_{21}x + r_{22}y + r_{23}z + t_y \\ r_{31}x + r_{32}y + r_{33}z + t_z \end{bmatrix} \tag{1}$$

Where (r11,r12….,r33) is the elements of rotation matrix R and tx, ty, tz is translation value to each axis. Because we set the rotation degree of each camera as 30 degree, rotation matrix R is:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\pm\frac{\pi}{6} & \sin\pm\frac{\pi}{6} \\ 0 & -\sin\pm\frac{\pi}{6} & \cos\pm\frac{\pi}{6} \end{bmatrix} \tag{2}$$

The following matrix multiplication yields the same result for the first three vector components:

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11}x + r_{12}y + r_{13}z + t_x \\ r_{21}x + r_{22}y + r_{23}z + t_y \\ r_{31}x + r_{32}y + r_{33}z + t_z \end{bmatrix} \tag{3}$$

This implies that the 4x4 matrix:

$$\mathrm{T} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4}$$

So, the following matrix multiplication yields transformation of plane to plane:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \sim \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p \\ q \\ 0 \\ 1 \end{bmatrix} \tag{5}$$

Where f is focal length of camera.
This can be simplified as follows:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \sim \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ r_{31} & r_{32} & t_z \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p \\ q \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ r_{31} & r_{32} & t_z \end{bmatrix} \begin{bmatrix} p \\ q \\ 1 \end{bmatrix} = \begin{bmatrix} fr_{11} & fr_{12} & ft_x \\ fr_{21} & fr_{22} & ft_y \\ r_{31} & r_{32} & t_z \end{bmatrix} \begin{bmatrix} p \\ q \\ 1 \end{bmatrix} = H \begin{bmatrix} p \\ q \\ 1 \end{bmatrix} \tag{6}$$

Where H is homography matrix.

Focal length f can be calculated by camera calibration method, typically z. zhang's calibration method [7]. Finally, the transformation matrix T for the plane to plane transformation is same to homography matrix. So, if we evaluate homography matrix, we can calculate translation value t_x. In order to evaluate homography matrix, we use SURF algorithm and nearest neighbor search algorithm. As you can be seen in Figure 4, when the objects are far enough away from cameras, overlapped areas are shown in each image. Feature point matching in the overlapped area has low probability of incorrect matching and has short processing time. We supposed the size of overlapped area is 20% of original image and we don't consider translation of y axis. Figure 5 shows the result of feature points matching in overlapped areas. Feature points matching in overlapped areas have improved accuracy because the probability of same feature points extracting is improved.

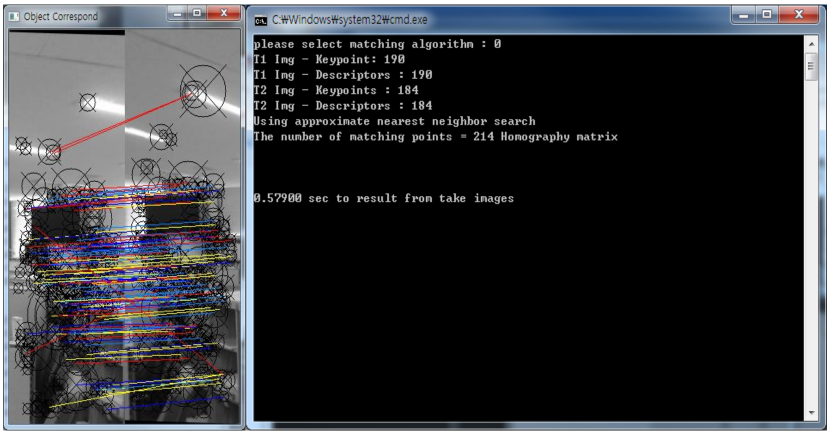


Figure 5. Result of Feature Point Matching in Overlapped Area

And as you can see in Figure 5, the processing time is more than five times faster. According to the research of Richard Szeliksi [8], 3D transformation matrix T of 2D image is $T=VRV^{-1}$. V and V^{-1} are viewing matrix for perspective transformation between 2D and 3D. R is rigid - rotation and translation - transformation matrix. It means that before and after we adopt the translation matrix and rotation matrix, it must be needed to the image projection to the 3D coordinate space and 2D coordinate space.

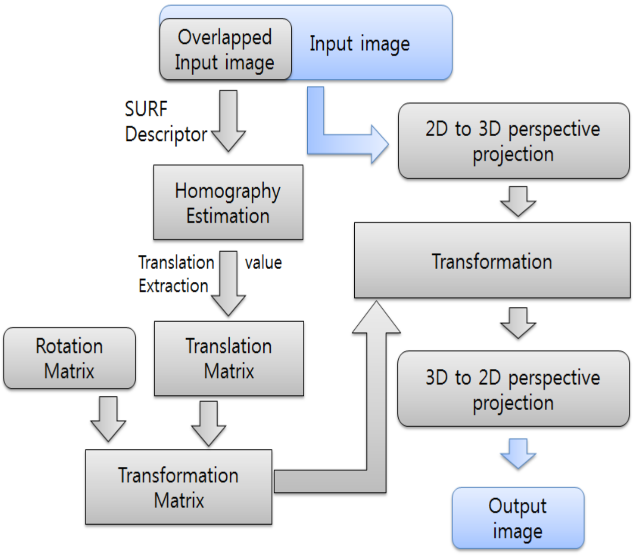


Figure 6. Diagram of Proposed Image Stitching

Proposed image stitching based on spatial transformation in this paper is shown in Figure 6 and result of image stitching is shown in Figure 7 Overlapped images are obtained from two images obtained from the two cameras. Then, each of the feature points from overlapped images are extracted using by SURF algorithm. Extracted feature points are using to create the homography matrix through approximate nearest neighbor search feature points matching algorithm. Final transformation matrix is calculated by homography matrix from overlapped images and geometrical rotation matrix.

Figure 7. Result Image of Image Stitching

4. Conclusion and Future Work

We proposed the image stitching method to create panoramic image using spatial transformation from normal two web camera. For the fewer feature points matching errors and faster processing time, we used two image stitching methods appropriately. Proposed method is less affected by external factors and processing time is shorter rather than use only image stitching using by homography matrix based on feature point matching. But, using homography matrix in the overlapped area improves the image stitching accuracy. The processing time to make one panoramic image is less than 1sec. To make the panoramic video using proposed method, it needs to be more advanced to reduce the processing time less than 0.3sec at least for the performance of more than 30fps. So, in the next study, we will exert effort to reduce the processing time to make one panoramic image and more accurate image stitching. Advantages of multi camera system is that you can use multi camera at the same time or separately as you want to. If the processing is possible more than 30fps in a normal system, the camera system can be applied to various fields like multi-camera smartphones. Also, as you can see in Figure 7, the boundary of the stitched area in the result image is very distinct. So it needs to adopt some proper de-blocking filter and we will implement de-blocking filter in the next study. And also, we will study very high resolution panoramic image system using a large number of cameras, more than 100 cameras. Although the number of camera is very large, we will study the real-time processing system using by distributed computing in a huge camera system.

Acknowledgment

This work was supported by the program of the Global hub for Experiment Data of Basic Science, 2015 funded by the NRF (N-15-NM-CR01-S01).

References

[1] E. Dubrofsky, “Homography estimation”, A Master’s thesis from Department of Computer Science, University of British Columbia, **(2009)**.

[2] G. Lowe, “Distinctive image features from scale-invariant key-points”, International journal of computer vision, vol. 60, no. 2, **(2004)**, pp. 91-110.

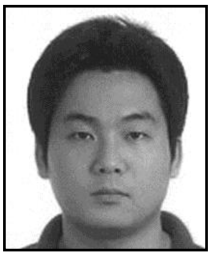
[3] H. Bay and T. Tuytelaars, “Surf: Speeded up robust features”, Computer Vision-ECCV 2006, **(2006)**, pp. 404-417.

[4] W. Forstner, T. Dickscheid and F. Schindler, “Detecting interpretable and accurate scale-invariant key-points”, Computer Vision-IEEE 12th International Conference on, **(2009)**, pp. 2256-2263.

[5] T. Tuytelaars and L. V. Gool, “A comparison of affine region detectors”, International journal of computer vision, vol. 65, no. 1-2, **(2005)**, pp. 43-72.

[6] S. Arya and D. M. Mount, "An optimal algorithm for approximate nearest neighbor searching in fixed dimensions", Proc. 5th ACM-SIAM Symposium discrete algorithms, **(1994)**, pp. 573-582
[7] Z. Zhang, "Camera calibration with one dimensional objects", Computer Vision-ECCV 2002, **(2002)**, pp. 161-174.
[8] R. Szeliski, "Video mosaics for virtual environments", Computer Graphics and Applications, IEEE, vol. 16, no. 2, **(1996)**, pp. 22-30.

Authors



Chi-Woong Kim, received the B.S., M.S degrees from the School of Computer Science at University of Seoul, in 2008 and 2010, respectively. He has been serving as a researcher of Global Science Experimental Data hub Center at Korea Institute of Science and Technology Information since 2010. His research interests include image processing, big data system and linux system.

Hyun Seong Sung, received the B.S., M.S and PH.D. degrees from the School of Computer Science at University of Seoul, in 2003, 2005 and 2010, respectively. He has been serving as a researcher at Agency for Defense Development, Korea since 2011. His research interests include image processing and information protection.

Hee-Jun Yoon, received the M.S from the Computer Engineering at Chungnam National University, KR in 1997. He has been serving as a Senior Researcher at Korea Institute of Science and Technology Information since 2000. His research interests include Large data processing, Parallel Computing, Cyber-Infra system, and Computer Education.

