An efficient Method for Synthesizing the Full View of the Large Scale Object on Successive Multiple Images

Thin Lai Lai Thein, Myint Myint Sein
University of Computer Studies, Yangon

tllthein@gmail.com; mlm@wwlmail.com

Abstract

This paper presents a robust method for panorama view reconstruction from multiple viewpoints of the large-scale object. The desired object must be acquired from multiple viewpoints. The information of the feature points of an object is obtained by establishing correspondence between views. We present a robust method based on image mosaicing method and panorama synthesizing. First, the feature of each object is extracted. Then the corresponding feature points of an object between the images are estimated. The different of minimum threshold value between two images is computed. After that, the full view of an object is reconstructed by integrating two images based on image mosaicing approach. This paper describes how to establish feature correspondences between images accurately and effectively. Image registration technique provides an initial estimation for establishing feature correspondences of point features. The linear solution with the reliable correspondences makes the computation of the geometric transformation between two images, called homography. Experiments on real images show the effectiveness of these techniques. The tested images can be taken from indoors and outdoors views. The resulted image is demonstrated as real- life situations.

Keywords: Image mosaicing, Synthesizing, Geometric transformation, Linear motion, Arbitrary view

1 Introduction

It has long been recognized that the large-scale object reconstruction is an important enabling technology for realistic, large-scale visualization of the world. Realism in simulation is enhanced by accurate modeling of the environment being simulated. Despite tremendous research efforts, automatic acquisition of 3D views from 2D imagery remains a challenging problem. The whole view of a large-scale object is impossible to take the whole scene at once. After that, in my proposed system find feature point of the image which difficult to detect the corresponding pair points and integrate the very small feature points of the large scale object. The advantage of this process is useful for showing irremovable objects from one place to another at museum.

A robust method for synthesizing the arbitrary view from multiple images of the large-scale object is presented. The desired object must be acquired from multiple viewpoints. Feature extraction algorithm is developed for extracting the feature points from each image. Histogram equalization and lighting adjusting are performed to improve the accuracy of feature extraction. An automatic panorama view modeling approach is presented. The corresponding feature points are extracted among images and find the matching automatically.

The major contributions of this paper are a large scale object can be generated by integrating the multiple images of an object. The large-scale object which can't take the whole view at once and it is impossible to take at remove. It can be solved insufficient distance between camera and object. It can be generated impossible to take the whole scene from obstacle object. It can be reduced the number of camera. In our system, we propose a method for synthesizing the arbitrary view. This approach can be applied to generate the intermediate view, the full view of large-scale object, 3D virtual views and so on. It is also possible to create the new scene by merging the synthesized views of multiple objects. In this approach, it is possible to produce a lot of intermediate views by using a few cameras. Then it will be provided to reduce the cost and number of camera in security system by using the synthesizing the arbitrary view of images.

Some reports have already been presented concerning the research work of the synthesizing the arbitrary view for large-scale virtual environment. S.E Chan and Williams [9] proposed the view interpolation method for image synthesizing by using image morphing technique. M.M.Sein et al., [10] presents an approach for reconstructing the arbitrary view of a large-scale object. This presents a new approach for synthesizing the arbitrary view based on the image morphing technique. T.Takahashi et al., [7] proposed a method for rendering views for large-scale scenes. T.L.L.Thein et al., [8] presented a method for reconstruction the entire view of multiple objects. This method is not only different from arbitrary synthesize view of multiple object but also it can be seen entire view of each object individuality. C.Doria et.al [3] describes a prototype system for automatically registration and integrating multiple views of objects from range data. Then the result can be used to construct geometric models of the objects. Y.Chen and G.Medioni [1] avoided the search in the view transformation space by assuming an initial approximate transformation for registration, which is improved with an iterative algorithm that minimizes the distance from points in a view to tangential planes at corresponding points in other views.

This paper organizes as following. In section 2, the proposed synthesizing method and its practical applications based on the image mosaicing method is presented. Section 3 represents the design of system and the section 4 shows the result of synthesizing the multiple images applying mosaicing techniques. Finally section 5 provides the conclusion for this paper.

2 Image Mosaicing Method

Many problems require finding the coordinate transformation between two images of the same scene. Image mosaicing is important to have a precise description of the coordinate transformation between a pair of images. The construction of mosaic images and the use of such images on several computer vision/graphics applications have been active areas of research in recent years. Feeling strongly interested in the field of immersive environments, we present the image mosaicing method, which creates larger images by collections of overlapping images from many related images. It is possible to construct a single image covering the entire visible area of the scene. Automatic image mosaicing has been an active research area in the fields of photogrammetry, computer vision, image processing, and computer graphics. At the early stage, the application was to composite images of satellites such as air planes. The geometric transformation was simple because the land can be assumed to be flat. The distance between the camera and the object was large. The main research topic was generated for large scale object which haven’t sufficient distance between camera and objects. Two images taken from the different viewpoints, the relationship between the images can be described by a linear projective transformation called a homography.

Once the homography between two images is obtained, we can construct a large-scale image by transferring one image to the other image with the homography matrix.

Figure 1.An Illustration of Image mosaicing

In our system, it is expected to get mosaic images and finally get one picture of panorama**.** Our method has many advantages such as high speediness, efficiency, quality and low costs. We presented image mosaicing approach and multiple views synthesizing. First, we extract features in the overlapped area in one image and then extract the corresponding features in the other image. After that, we integrate common parts of two images. In our implementation, image mosaicing is decomposed into 2 stages: we first find the perspective transformation between two overlapping images, and then compute the absolute coordinate transform for each image. And then find the best maps between two images.

2.1 Corresponding Relations Among the Images

Let $P(X,Y,Z)$ be a point in a 3D scene $P'(X',Y',Z')$ and $P''(X'',Y'',Z'')$ are observed from two different camera positions. P' and P'' are the projected image points relations to the different camera positions. Then the relation between the two camera systems can be expressed by the rotation and translation as

$$P'' = (R/T)P' \qquad (2.1)$$

where R be the 4×3 rotation matrix and T be the 4×1 translation vector, respectively. This above equation (2.1) can be expressed in term of the image point P' and P'' as follow:

$$Iq'' = S(R/T)q' \qquad (2.2)$$

where I= $(Z''/f''), S = (Z'/f')$.

The transformations accompanied with the above correspondences are applied to each view and all the views are registered in the coordinate basis of a reference view. After all the views are registered, correspondences are established between all sets of overlapping views on the basis of nearest neighbor that are within a distance. Views that have correspondences less than a threshold are considered as non overlapping views and their correspondences are rejected. The estimation of the feature points by matching two images at different resolutions. A feature based approach is used to match the images leading to the solution of a set of the linear equations derived from the feature-point correspondences. This linear system is solved by an iterative least square method [2].

2.2 Image Registration

Image registration is one of the fundamental tasks in image processing. It is used to match two or more images of the same scene taken at different times, from arbitrary view. It is the process of matching two images which are reference image and operated image. The corresponding coordinate points in the two images correspond to the same physical region of the scene being imaged. (See the Figure.2). Image registration is widely used in remote sensing, computer vision, video processing, and many others. In our system, the reference image could be the first frame of the image sequence. The large-scale image is created for the operated image and the coordinate system of the reference image is known. Image registration is the process of transforming one image into the coordinate system of another image. We find the coordinate transformations between all pairs of an image sequences. After that we cut off the image which is not overlapping region from operated image and then merged it to the reference image.

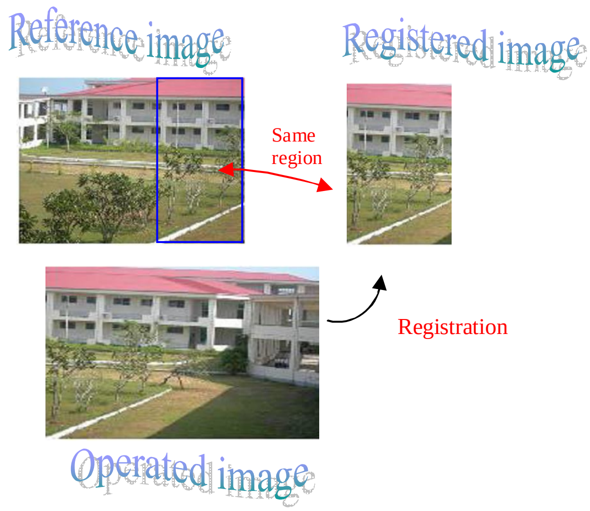


Figure2. An Illustration of Image Registration

2.3 Image Projection

After image registration, the feature points in each image can be transformed the coordinate frame. The set of all images and transformation parameters comprises the mosaic representation of the scene. The aim of image projection is to find the correspondences between two images. We computed the transformation parameters by using mean square distance between two images. The correspondences are not easy to find due to incorrect feature detection and non-robust feature matching. Therefore, reliable feature detection is needed. The feature detection approaches that canny edge detection usually requires the choice of a threshold value.

2.4 Feature-Based Image Mosaicing

The most difficult part in image mosaicing is to estimate the geometric transformation between images. The difference between images can be represented compactly by this transformation. The level of difficulty depends on the amount of the displacements between images. When the displacement is small, the images sequences can be solved easily with any approaches. However, when the displacement is large, a collection of images becomes more difficult. Planar projective transformation is the most accurate model whose number of parameters is larger than other transformation. To estimate this transformation from images is difficult because it is sensitive to noise. As the number of parameters increases, its computation requires reliable methods.

Image features are primitive elements which contain prominent pixel value changes. These features are described with image coordinates when we place a 2D coordinate system on the image plane. Once we establish feature correspondences between images, we can compute the geometric transformation from the image coordinates. We obtain the geometric relationships between images by using image features. Another difficult to obtain the geometric transformation between images is lack of texture in the images. We find the feature points between two images to solve this problem. The point features are overlapping regions in an image, which are common to use in image mosaicing. We extract the overlapping region between two images and we find a displacement maximizing in the overlapping region, and we determine the displacement by finding the best score. After obtaining the overlapping region of two images, we first auto detect the feature points in two images.

2.5 Image Synthesizing

The basic concept of using the planar projective transform is that it assumes scenes to be planar. By dividing a non-planar scene into multiple triangular patches, the planar projective transform can be applied for non-planar scene.

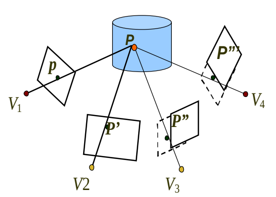


Figure 3 Multiple Views Synthesize

A 3D scenes point P projects onto the 2D points *p*, *p*', *p*'', *p*''' in four views. The relationship between the 3D and 2D spaces is represented by the camera projection matrices as follows:

$$p = [\mathrm{I}, 0]P,$$
$$p' = [A, v']P,$$
$$p'' = [B, v'']P,$$

where **I** is the identity matrix $A = [a_i^j], B = [b_i^j]$ and $C = [c_i^j]$ are the homography matrix. The vector *v*, *v*'', v''', are known as epipoles. The point, *p*'' in 3rd image can be obtained from the correspondence pair ***p*** and ***p'*** as,

$$p_l'' = p_k (p_i' T_{kjl} - p' T_{kil}), \qquad (2.3)$$

where $T_{ijk} = v_j' b_i^k - v_k'' a_i^j, (i, j, k = 1,2,3)$

We can virtually extend for multiple images by repeating this process.

3 Design Implementation

In the last few years the interest in mosaicing has grown in the vision community because of its many applications. The automatic construction of large and high resolution image mosaics is an active area of research in the fields of photogrammetry, computer vision, image processing, real rendering, robot -vision and computer graphics. Image mosaics involve aligning a sequence of image into a larger image and are an important issue in many virtual reality problems. Mosaicing is a common and popular method of effectively increasing the field of view of a camera, by allowing several views of a scene to be combined into single view. The traditional approach, which uses correlation intensity based [11] image registration, suffers from computation inefficiency and is sensitive to variations in image intensity. A feature-based approach is used to improve the efficiency of image mosaics.

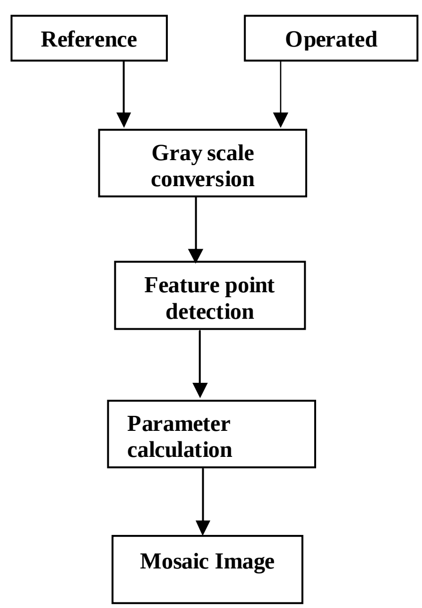


Figure4. Flowchart of Mosaic Creation System

Two planar views are grabbed by an affine transformation of the camera. One of the images is used as the reference image, and the second image is aligned with the reference image. To find the coordinate transformation between the two images, the feature points in each image are detected firstly. Next, a matching process is performed to estimate the corresponding feature points between these images. SSD (Sum of Squared Difference) method is used for feature points matching. System flowchart is shown in Figure 4.

3.1 Image Acquisition

NIKON COOLPIX 7600 digital camera is used for taking the images from arbitrary view positions. It gives a 640×480 resolution. Two images of a scene have been taken at different time, from different viewpoints. First, the original image has been converted to gray-scale image. The reference image and operated image of original images are shown in Figure.5.The proposed algorithm has been implemented with MATLAB R2006a version 7.2.0.232 and it has been tested using the image sequences. The test images were obtained in the outside and inside scenes by using hand held digital camera for minimum time interval. Each color image has 320×240 pixels and image format is JPEG format. Figure.6 shows the final result of image mosaicing method from two input images.

Reference Image Operated Image

Figure 5 Reference Image and Operated Image

Figure 6 Panorama View, Resulted by Using Two Images

4 Experimental Result

This section introduces the result of applying mosaicing techniques to synthesizing the multiple images. Figure 7 shows the overlapping region of two images. In this figure, the registered image is the overlapping region of two images. The registered image is obtained from operated image by using image registration. Image registration is the process of transforming one image into the coordinate system of another image. It is one of the fundamental tasks in image processing .It is used to match two or more images of the same scene taken at different times from arbitrary view. It is the process of matching which are reference image and operated image. The corresponding coordinate points in the two images correspond to the same physical region of the scene being imaged. In system, the reference image could be the first frame of the image sequences. The large scale image is created with operated image and the coordinate system of the reference image is known.

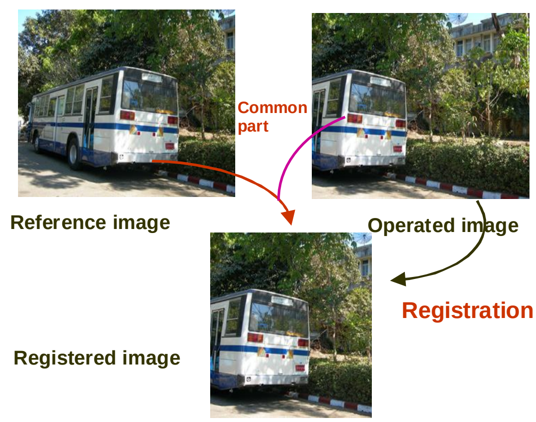


Figure 7 Overlapping Region of Two Images

The corresponding feature points on the two input images are detected and the distance in the overlapping region is calculated. The non overlapping region of the operated image is cut off (See the Figure 7.1) and then merged it to the reference image. Figure 7.2 shows the integrating both the reference image and non overlapping region of operated image. By synthesizing these images, the whole view of a scene is obtained.

Figure 7.1 Cut Off Non- Overlapping Region from Operated Image

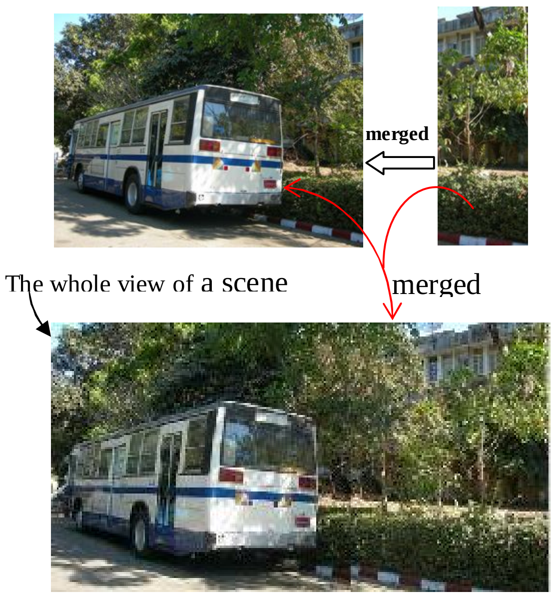
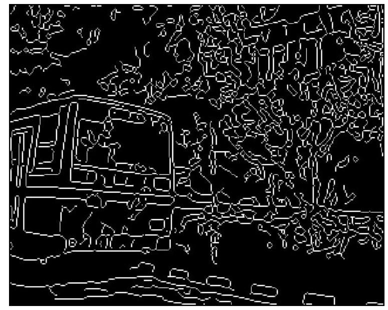


Figure 7.2 Synthesized Two Images

Before extraction the feature points from both images, edge detection is used for estimating the translational displacement between the images. Figure 7.3 shows the edge extraction from two input images.

(a)Edge Detection for Reference Image



(b)Edge Detection for Operated Image

Figure 7.3 Edge Extraction

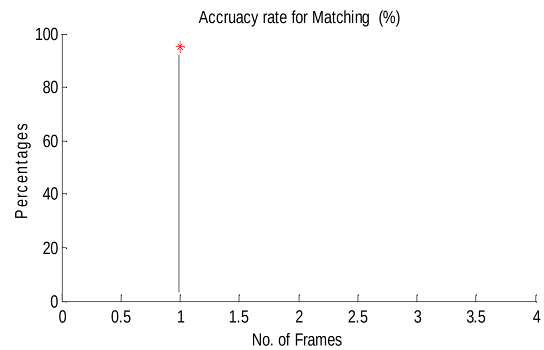
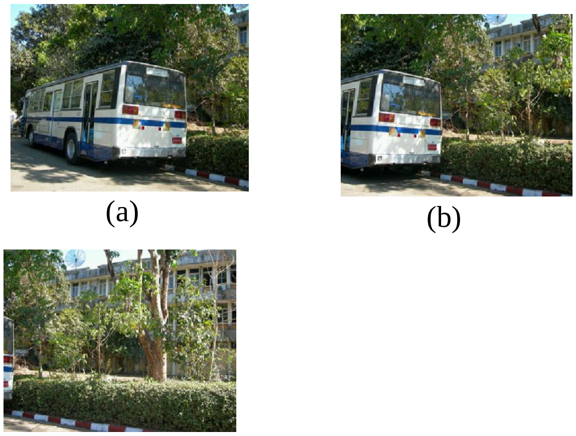


Figure 7.4 Accuracy Rates Matching Points for Figure 7.2

This figure shows the accuracy matching percentages between two images. The accuracy rate percentage is the ratio of the number of corresponding pairs and maximum number of feature points in images. The correct matches is (95.3630%) between two images. It took 49 seconds for the total processing time to synthesize the two images. The following equation is used for calculating the accuracy rates.

$$\text{Accuracy rates of matching} = \frac{(\text{no. of corresponding pairs}) \times 100\%}{(\text{Maximum no. of feature points})}$$



(a) (b)

(c) (d)

Figure 7.5(a), (b), (c), (d) Four Input Images

Figure 7.5 shows four input images in outside scene using digital camera for minimum time interval. Firstly, we perform a matching process to find the corresponding feature points between two images (figure (a), (b)) and create a sub-mosaic with these images. Then this sub-mosaic and next input images have been combined (See the Figure.7.9). By repeating this process, we get the whole view of outdoor scene (See the Figure.7.10). It takes 136 seconds for processing time. The correct matching points is (95.3630 %) for first image and second image. And it took (65.7330 %) for a correct matching points between submosaic and the third image. Then the correct match between submosaic and the fourth image is (72.9466 %). (See the Figure.7.7). Final result will get after sub-mosaic and the fourth image are synthesized.

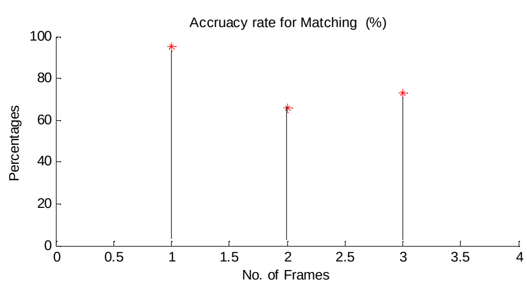


Figure7.7 Correct Matching Points for Four Input Images

Figure 7.8 A Sub-mosaic Image for input images (a),(b)

Figure 7.9 A Sub-mosaic Image for input images (a),(b)and (c)

Figure 7.10 The Result Image Using Four Input Images

5 Conclusion

The proposed system makes the full-view of large-scale object with multiple images at any position as real scene. We find the relative transform between each pair of overlapping images and then compute the absolute coordinate transform for each image by an error minimization technique. Image mosaicing is important to have a precise description of the coordinate transformation between a pair of images. Image mosaics are collection of overlapping images together with coordinate transformations that relate the different image coordinate systems. The transformation parameters between each successive image are calculated and then used in blending operation. For each image pair, we could get very good results but when we use these parameter in blending, it could cause some artifacts. We calculate the transformation parameters between any successive images then we use them to construct mosaicing. The perfect and good image has been received when blending the two, three, four images into one. But if we increase the number of images, the number of artifacts also increases. Therefore we proposed tree structure image mosaicing. As the transformation error between two image is less, instead of calculating transformation parameters once and use them construct a mosaic, we calculate transformation parameter between one-two, one two- three, one two three-four, etc.Transformation parameters for image1 and image 2 are used to construct a submosaic for each image pair. After constructing the submosaic of images, we use same method to take another submosaic by synthesizing two images. Then the transformation parameters are calculated between these submosaic and next input image. This process is continued until we get only one mosaic. The following figure shows tree structure mosaicing process.

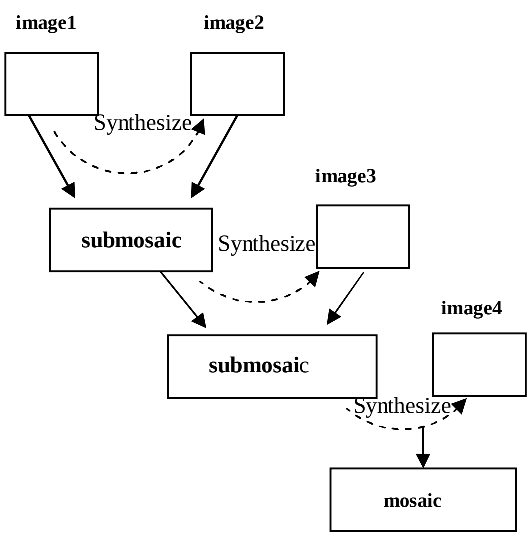


Figure. 8 Tree structure mosaic

References

[1] Y.Chen and G.Medioni, "Object modeling by registration of multiple range images", Image and Vision Computing, 10(3):145-155, Apr 1992.

[2] C.Chua and R.Jarvis, "3D free-from surface registration and object recognition", IJCV, vol.17, pp.77-99, 1996.

[3] C.Dorai et.al., "Registration and Integration of Multiple Object Views for 3D Model Construction", IEEE Tran. On pattern analysis and machine Intelligence, vol.20, no.1, January 1998.

[4] M.Irani, P.Anandan, and S.Hsu, "Mosaic based representations of video sequences and their applications", In fifth International Conference on Computer Vision (ICCV'95), pages 605-611, Cambridge, Massachusetts, June, 1995.

[5] B.Lucas and T.kanade, "An iterative image registration technique with an application to stere vision", Proc.of the 7th International Joint conference on Artificial Intelligence, pp.674-679, 1981.

[6] Ko Nishino et al.," Robust Simultaneous Registration of Multiple Range Images", ACCV2002: The 5th Asian Conference on Computer Vision, 23-25 January 2002, Melboume, Australia.

[7] T.Takahashi et al., "Arbitrary View Position and Direction Rendering for Large-Scale Scenes", IEEE International Conference on Computer Vision and Pattern Recognition, pp.1063-6919, vol2.2000.

[8] T.L.L.Thein, et al.,"An efficient Method for Reconstruction the Entire view of the Multiple Object by Pair-wise Registration Technique", SICE-ICASE International Joint Conference 2006 Oct. 18-21, 2006 in Bexco, Busan, Korea.

[9] S.E.Chan and Williams, "View Interpolation for Image Synthesis", Proc.SIGGARPH 93, pp. 279-299, 1993.

[10] M.M.Sein et.al, " Reconstruction the Arbitrary View of an Object Using the Multiple Camera System", IEEE International Symposiun on Micromechatronic and Human Science(MHS 2003), Nagoya, Japan,pp.83-88, Oct. 19-22,2003.

[11] R.Szeliski, H.Y.Shum, "Creating full view panoramic mosaic and environment maps", Proceeding of the SIGGRAPH'97, Los Angeles, CA August (1997) 251-258.