Intermediate View Synthesis for Multi-view 3D Displays Using Belief Propagation-Based Stereo Matching

Conference Paper · December 2008
DOI: 10.1109/ICCIT.2008.212 · Source: IEEE Xplore

Hong Jeong
Pohang University of Science and Technology

Intermediate View Synthesis for Multi-view 3D Displays using Belief Propagation-based Stereo Matching

Changming Jin, Hong Jeong
Department of E.E.E, Pohang University of Science and Technology, Korea
bigboy@postech.ac.kr, hjeong@postech.ac.kr

Abstract

Intermediate view synthesis is an important step for Multi-view 3D displays. In this paper we have proposed an intermediate view synthesis approach, which is based on stereo matching process. This makes that multi-view images can be created by only the two-view images, which are captured by two-view camera (stereo camera).In order to obtain accurate disparity map for intermediate view synthesis, Belief Propagation-based stereo matching was be used in this paper. Experimental results show that the proposed approach can provide the intermediate view with high quality.

1. Introduction

Human beings see different viewpoint images of an object through two eyes, the right and left. Then, the human brain recognizes the three-dimensional (3D) information of the object by synthesizing them with the binocular disparity of a stereo input image pair[2].

Most conventional 3D display systems have been implemented by imitating this human visual system, so that these systems normally need a pair of cameras for capturing the left and right images of an object, and a pair of projectors for projecting the captured left and right images to the screen, and some special optical devices for separately inputting the projected left and right images to the corresponding eyes between the left and the right image lead to the experience of depth.

Now the most popular 3D display technology is multi-view 3D display systems, which is glasses-free 3D technology. We focus on the lenticular multi-view 3D display. Examples of lenticular multi-view 3D display are Philips multi-view 3D display and Newsight multi-view 3D display[11].

The principle of lenticular multi-view 3D displays is depicted in Figure 1. In front of a conventional LCD, cylindrical lenses are attached. The lenses, which can direct light, accommodate that for a specific viewing angle, the viewer only sees a subset of the pixels of the underlying LCD. In the two-view display, in order to give the viewer the experience of 3D, the left eye perspective image and right eye perspective image are presented on an LCD panel utilizing only the associated pixel groups of columns. Since each eye sees a different perspective image, if images are a stereoscopic pair, the viewer can sees image in 3D. In multi-view display, multiple perspective image more than two are used[10,11].

Two-view displays have only one pair of images and must be viewed from a certain location. Multi-view displays have a larger number of stereo pairs and can be viewed from multiple locations. Moreover, multi-view displays can be watched by multiple viewers at the same time.

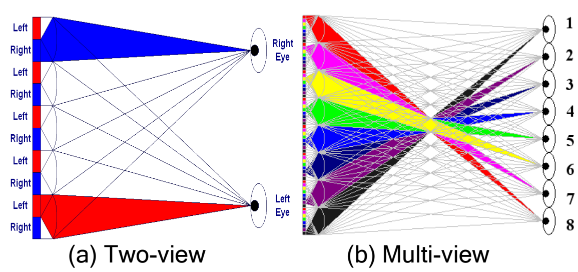


Figure 1. The principle of the lenticular screen

Figure 2. shows the lenticular 3DTV display system, which is composed of the multi-view cameras, the multi-view signal synthesis, and the lenticular multi-view screen. In this system, first the cameras capture the multi-view images, then multi-view images are woven to one 3D image, finally 3D image is displayed

on the lenticular multi-view screen. Through lenticular multi-view screen, the viewer can get the experience of 3D.

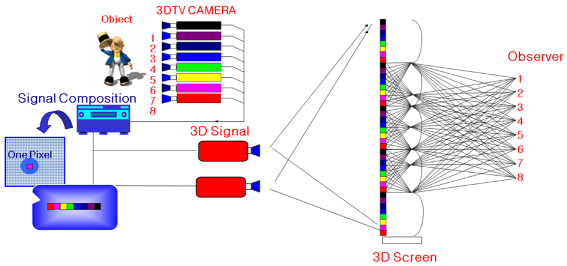


Figure 2. Multi-view 3D displays system

However, for practical applications, multi-view 3D displays suffer some problems which are high price of multi-view cameras, large memory size of multi-view images data and low speed of large data transmission.

To solve the problems, we propose an intermediate view synthesis approach, which is based on stereo matching process. This makes that multi-view images can be created by only the two-view images, which are captured by two-view camera (stereo camera).

The proposed method basically runs using a pair of stereo images captured by stereo cameras to generate the intermediate view. Using stereo matching, a disparity map yields a simple and fast way of generating the intermediate view based on local image warping.

The synthesized intermediate view can be used as parts of multi-view images.

In Section 2, we introduce the stereo matching and disparity. In Section 3, the method for intermediate view synthesis is described. Section 4 shows some experimental results with the stereoscopic image pair. Conclusions are included in Section 5.

2. Related Works: Stereo Matching and Disparity

Stereo matching process is to estimate the disparity map, which represent the depth information of objects in 3D real world.

Since the stereoscopic cameras are set to make cameras have parallel optical and vertical axes and coplanar plans, corresponding point search in stereo images is constrained to horizontal scan lines. Thus in a pair of stereo images, disparity can be defined as the difference in the horizontal coordinates of corresponding point.

Figure 3. shows a 3D object point P project on both the left and right image plane. f is the focal length, B is the base length, and z is depth of object point P [1,5,8,9].

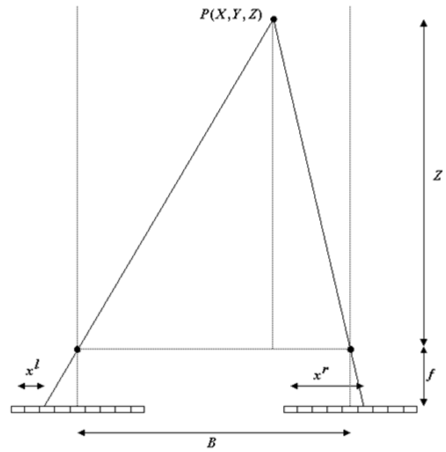


Figure 3. The geometric structure of stereo matching

$$d^{LR} = X_R - X_L = -\frac{Bf}{z} \qquad (1)$$

Where d^{LR} is the disparity of corresponding point from the left image to the right image, disparity presents the geometric relation of coordinates of corresponding point between the left and right image.

3. Proposed Method

Figure 4. shows the geometric structure of intermediate view between a stereo pair. A 3D object point P also project on intermediate view plane. By finding coordinates of corresponding point, intermediate view will be synthesized by mapping pixels of the left and right image.

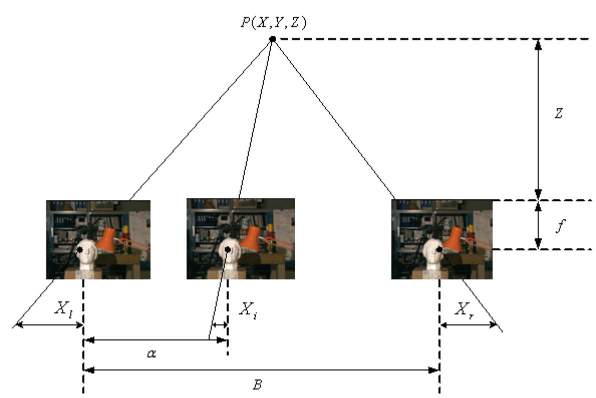


Figure 4. The geometric structure of intermediate view between a stereo pair

From formula (1), we can get similar geometric relation of coordinates of corresponding point between the left and intermediate image.

$$d^{LI} = X_I - X_L = -\frac{\alpha Bf}{z} = \alpha d^{LR} \qquad (2)$$

$$(0 \le \alpha \le 1)$$

Where d^{LI} is the disparity of corresponding point from the left image to the intermediate image. α is the location parameter , which stands for the relative normalized distances of the desired intermediate image between the stereo image pair.

$$X_I = X_L + \alpha d^{LR} \qquad (3)$$

$$(0 \le \alpha \le 1)$$

Where d^{LI} is used as a mapping function between corresponding point in the left image and the intermediate image.

Figure 5. shows a brief summary about the proposed method. It is composed of the following four steps. First, disparity map of the input stereo image is estimated. Next, in order to render the intensity of the intermediate image, pixel mapping from input image to intermediate image is carried out using disparity. Then, occlusions in the intermediate image are filled by compensation. Finally, the noises in the intermediate image are removed by interpolation.

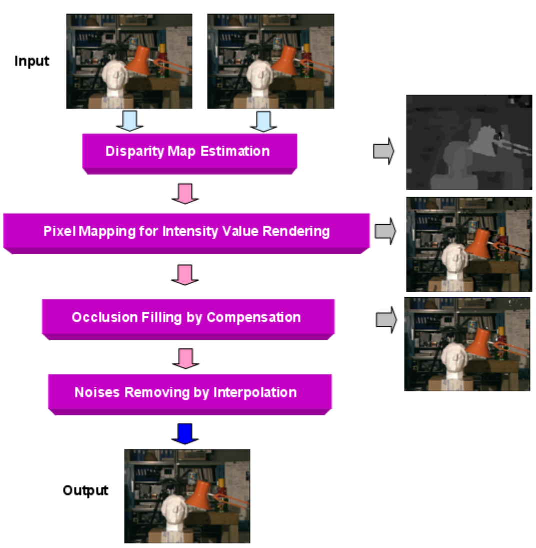


Figure 5. Overall configuration of the proposed system

3.1 Disparity Map Estimation

To estimate disparity map, we use stereo matching based on belief propagation (BP). BP is an iterative algorithm for computing posterior probabilities on a graphical model most commonly used in artificial intelligence and information theory, using global information[3,4,6,7]

In Figure 6, stereo matching using BP can get relative accurate disparity map.

Figure 6. Ground truth of disparity map and disparity map obtained by Belief Propagation

3.2 Pixel Mapping for Intensity Value Rendering

To render the unknown intensity value of the pixel in the intermediate image, we should find the coordinate of correspond image pixel(from right image to intermediate image). Thus for finding the coordinate of correspond image pixel, we map the pixel of corresponding image pixel (from right image to intermediate image) by using disparity. Then through the coordinate of corresponding image pixel, we provide the pixel in the intermediate image, with known intensity value of the pixel in the right image.[10,12]

$$I_{IR}(X_I, Y) = I_{IR}(X_R + (1-\alpha) * d^{RL}, Y) = I_R(X_R, Y) \quad (4)$$

$$(0 \le \alpha \le 1)$$

Where I_R is the intensity value of the pixel in the right image, I_{IR} is the intensity value of the pixel in projected right image, d^{RL} is the disparity of corresponding point from the right image to the left image.

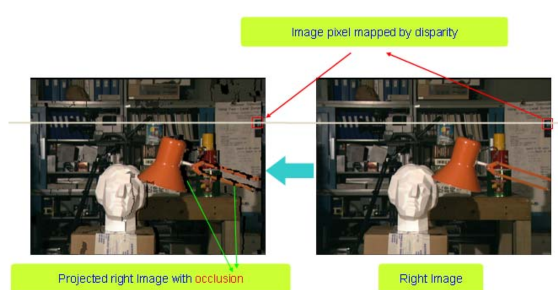


Figure 7. Projected right image with occlusion and original right image

As shown in Figure 7, occlusions occur after mapping pixel from the right image to the intermediate image. For the sake of filling occlusion, occlusion filling process will be introduced in following part 3.3.

3.3 Occlusion Filling by Compensation

Occlusion, which occur after mapping pixel from the right image to the intermediate image, can be filled by using left image, since the left image have the intensity information about occlusion. First, for finding the pixel coordinate of occlusion region in the intermediate image, we map the pixel of corresponding pixel (from left image to intermediate image) by using disparity. Then through the coordinate of corresponding pixel in occlusion region, we provide the pixel in the intermediate image, with known intensity value of the pixel in the left image.

$$I_{IL}(X_I, Y) = I_{IL}(X_L + \alpha * d^{LR}, Y) = I_L(X_L, Y) \quad (5)$$
$$(0 \le \alpha \le 1)$$
$$I_I(X_I, Y) = I_{IL}(X_I, Y) + I_{IR}(X_I, Y) \qquad (6)$$

Where I_L is the intensity value of the pixel in the left image, I_{IL} is the intensity value of the pixel in projected left image, d^{LR} is the disparity of corresponding point from the left. image to the right image.

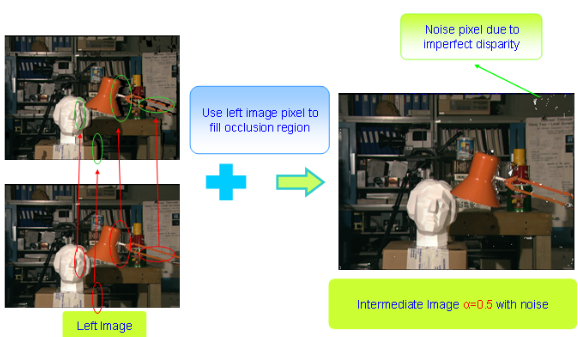


Figure 8. Projected right image with occlusion, original left image and intermediate image with noise

In Figure8, the marked black (Intensity=0) of the projected right image are occlusion regions. The marked part in the left image is to compensate the occlusion regions. After compensation, the intermediate image is generated with no occlusion. It shows that occlusions have been filled.

Although occlusion have been removed, noises, which are marked white part (Intensity=255), occur in the intermediate image. For sake of removing noises, interpolation process will be introduced in following part 3.4.

3.4 Noises Removing by Interpolation

The noises in intermediate image due to imperfect disparity mapping, can be removed by interpolation method. We use the nearest interpolation method since the number of noises is small,

$$I_I(X_I, Y) = I_I(X_I \pm 1, Y) \text{ if } I_I(X_I, Y) = 255 \quad (7)$$

In Figure 9, noises have been removed by the nearest interpolation method.

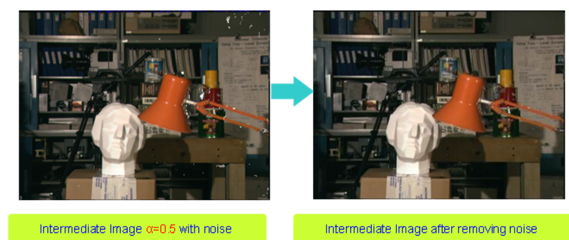


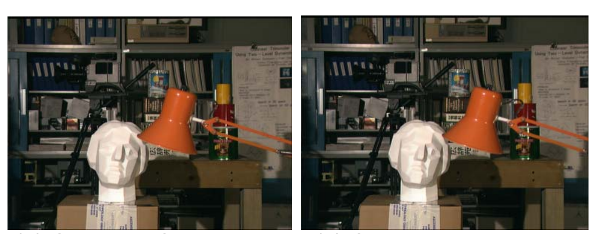
Figure 9. Intermediate image with noises and intermediate image after remove noises

4. Experiment Result

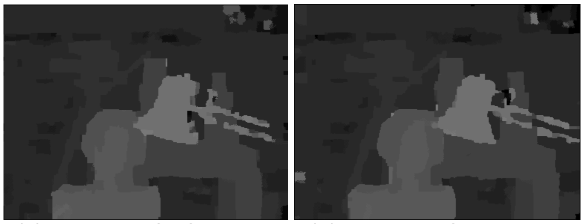
We tested our method on a pair of stereo image ‘Tsukuba’. Though there is no standard evaluation method, the intermediate view image synthesized (a=0.5) appears to be accurate and natural by eyes observation.

Using intermediate view image from stereo images, multi-view images are also implemented. It is synthesized by the 8-view image.

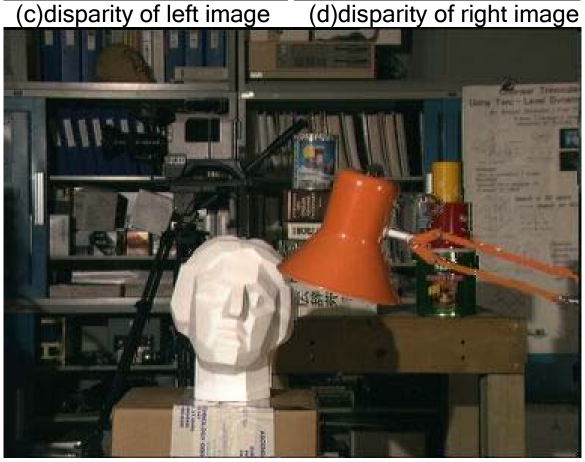
Experiment result proves that multi-view images can be created by only the two-view images by our method.



(a) Original left image (b) Original right image



(c)disparity of left image (d)disparity of right image



(e) Synthesized intermediate view $\alpha = 0.5$

Figure 10. Projected right image with occlusion and original right image

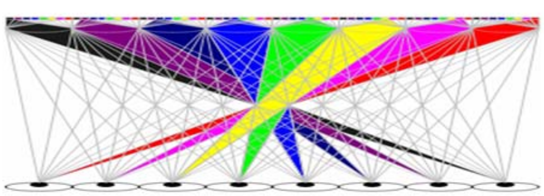


Figure 11. 8-view synthesis for 3D display

5. Conclusion

In this paper we have proposed an intermediate view synthesis approach, which can be applied in multi-view 3D displays. This approach is based on stereo matching process. This makes that multi-view images can be created by only the two-view images, which are captured by two-view camera (stereo camera).

We proposed possible way of deal with some problems of multi-view 3D displays, which are high price of multi-view cameras, large memory size of multi-view images data and low speed of large data transmission.

In order to obtain accurate disparities for intermediate view synthesis, BP based stereo matching was be used in this paper.

The disparity map yields a simple and fast way of generating the intermediate view in 3D displays based on local image warping.

We also proposed possible way of removing occlusion and noise.

Experimental results show that the proposed approach can provide the synthesized intermediate view with high quality.

6. References

[1] Hong Jeong, Yuns Oh, "A parallel real time implementation of stereo matching", *Parallel and Distributed Processing Symposium., Proceedings 15th International*, April 2001, pp. 6.

[2] Umesh R. Dhond and J. K. Aggarawal. "Structure from stereo a review", *IEEE Transactions on Systems, Man, and Cybernetics*, nov/dec 1989, pp. 553-572.

[3] Q. Yang et al. "Real-time global stereo matching using hierarchical belief propagation", *The British Machine Vision Conference*, 2006.

[4] P.F. Felzenszwalb and D.R. Huttenlocher. "Efficient belief propagation for early vision", *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2004, pp. I261 – I268.

[5] S. Kimura et al. "A convolver-based real-time stereo machine", *Proc. Computer Vision and Pattern Recognition*, 1999, pp. 457–463.

[6] H. Hirschmuller. "Improvements in real-time correlation-based stereo vision", *IEEE Workshop on Stereo and Multi-Baseline Vision*, Dec. 2001, pp. 141 – 148.

[7] N. N. Zheng J. Sun and H. Y. Shum. "Stereo matching using belief propagation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(7): pp. 787 – 800.

[8] Darius Burschka Myron Z. Brown and Gregory D. Hager. "Advances in computational stereo", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, , Jul. 2003, pp. 993–1008.

[9]O.Veksler. "Fast variable window for stereo correspondence using integral images", *Computer Vision and Pattern Recognition*, 2003, pp. 556–564.

[10] Chaohui, L., Ping, A., and Zhaoyang, Z., "Efficient stereo disparity estimation for intermediate view synthesis", *Circuits and Systems, The 47th IEEE Midwest Symposium*, July 2004, pp. iii- 483-6.

[11]Berretty, R.-P. M.,Peters, F. J.,Volleberg, G. T. G., “Real-time rendering for multiview autostereoscopic displays”, *Proceedings of the SPIE*, 2006, pp. 208-219.
[12] Potter, M.A., and Jong, K.A.D, “Stereo Vision for View Synthesis”, *1996 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'96)*, 1996, pp. 852.