Implementation of Object-based Multiview 3D Display Using Adaptive Disparity-based Segmentation

Jae-Sung Park, Seung-Cheol Kim, Kyung-Hoon Bae, Eun-Soo Kim
3D Display Research Center/ National Research Lab. of 3D Media, School of Electronic Eng., Kwangwoon University, 447-1Wolge-Dong, Nowon-Gu, Seoul 139-701, Korea

Abstract
In this paper, implementation of object-based multiview 3D display using object segmentation and adaptive disparity estimation is proposed and its performance is analyzed by comparison to that of the conventional disparity estimation algorithms. In the proposed algorithm, firstly we can get segmented objects by region growing from input stereoscopic image pair and then, in order to effectively synthesize the intermediate view the matching window size is selected according to the extracted feature value of the input stereo image pair. Also, the matching window size for the intermediate view reconstruction (IVR) is adaptively selected in accordance with the magnitude of the extracted feature value from the input stereo image pair. In addition, some experimental results on the IVR using the proposed algorithm is also discussed and compared with that of the conventional algorithms.

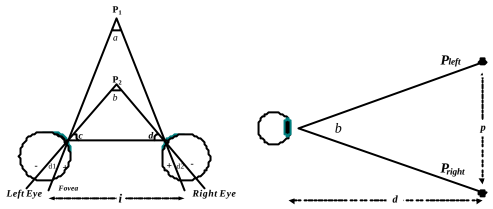
1. Introduction
The conventional stereoscopic display system has been implemented by imitating the human visual system (HVS). Among them, the binocular stereoscopic display system that is based-on stereo camera has been mostly researched [1]. But, in these stereoscopic display systems, the 3D feeling of depth is suddenly disappeared when the viewing point is just moved away from the point taken the stereo views through two cameras and as a result, the viewing region is very limited [2]. Therefore, recently, a lot of research on the multiview 3D display system has been done to solve the

shortcoming of the conventional binocular stereoscopic display method [3]. That is, the multiview IVR technique is suggested as a new approach to extend the limited viewing region of the binocular display system. As a result, the intermediate view reconstruction technique makes it possible to display more natural 3D image in some range of views angle by increasing the number of viewpoints. Another approach is the multiview 3D display system in which multiple cameras are used to get the multiview images. But this system has some problems. That is to say, the excessive amount of data and the discontinuity between viewpoints occurs as a number of viewpoints are increased [4]. Accordingly, the easy and best way to solve this kind of problems without increase of the number of cameras and image data to be processed is the IVR technique in which multiview intermediate view images can be digitally synthesized by using the given stereo image pair taken by a stereo cameras.

In this paper, implementation of object-based stereoscopic display system using adaptive disparity estimation is proposed. Firstly input stereo image pair are segmented by using region-growing algorithm. This algorithm is a procedure that groups pixels or sub-region into larger regions and give very good segmentations that correspond well to the observed edges. Secondly intermediate view of various viewpoints images are synthesized by using adaptive matching window size. Therefore, the proposed intermediate view reconstruction techniques makes it possible to display more natural 3D image in some range of views angle by increasing the number of viewpoints.

2. Stereovision and Occlusion Information

2.1 Depth information according the camera configuration
The stereo image pairs are acquired by the binocular camera system arranged in parallel-axis configuration. This camera configuration is called the stereovision, corresponding to the human visual system (HVS). Stereovision is the ability to use binocular disparity to determine whether objects are closer or further than the fixation point of the eyes [1][3]. Fig.1 shows the disparity and depth perception in HVS.



(a) (b)

Fig.1 (a) Disparity and depth perception in HVS. (b) Screen disparity and viewing angle.

If two eyes are focused to main point P_1 in space, it occurred main angle a . If focused to the other point P_2, occurred angle b . Point P_1 and P_2 perceive different depth and locate in different distance each other. In the Fig.1 (a), the summation of a , $d1$, c, $d2$, d is 180 degree and b , c, d is 180. Therefore, this modeled by Equation (1).

$$a - b = d1 + d2 \tag{1}$$

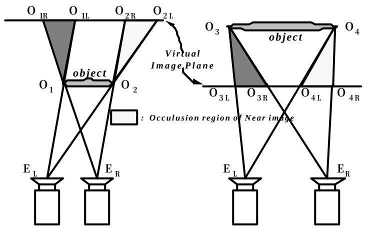
It means that differences of two angle equal to $d1+d2$, and it called retinal disparity. Let's examine stereovision from camera configuration. Camera configuration has to correspond following some cues; distance between axis, direction of axis, distance between object and camera. In the Fig.1 (b), screen disparity p and viewing distance d are given, then viewing angle b is modeled by Equation (2).

$$b = 2\arctan\frac{p}{2d} \tag{2}$$

This geometric relationship is called the epipolar constraint. Many stereo matching algorithms make the assumption that conjugate epipolar lines are collinear (or near collinear), which enables them to restrict the search for homologous image points to one dimension.

2.2 Depth information and Occlusion region

In this paper, it is used to estimate the disparity and to identify the occluded regions in the extreme left and right images. Disparity map can be obtained from mapping the disparity information to the reference image. Disparity map includes the occlusion region, and it depends on the positioning relation of virtual image plane and objects. Fig. 2 shows this.



(a) (b)

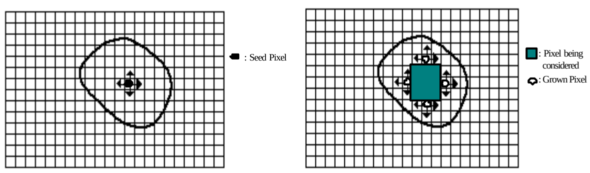
Fig.2 (a) Occlusion region in front of virtual image plane (b) Occlusion region in rear of virtual image plane

Occlusion regions are mainly appeared on the outline of objects within the stereo image pair. In case of having an occlusion region in front of virtual image plane, the occlusion region based on the left view is appeared on the right side of the object. Inversely, occlusion region based on the right view is appeared on the left side of the object. Principles of horizontal parallax applied the segmented objects.

3. Adaptive Disparity-Based Segmentation

3.1 Segmentation by using region growing

Firstly input stereo image pair are segmented by using region growing algorithm. Region growing is a procedure that groups pixels or sub-region into larger regions. It starts with a set of "seed" points and from these grows regions by appending to each seed point those neighboring pixels that have similar properties. Region growing methods often give very good segmentations that correspond well to the observed edges. Fig.3 is procedure of segmentation by using region growing



(a))

Fig.3 Procedure of segmentation using region growing algorithm (a) Start of growing a region (b) Growing process after a few iterations

3.2 Extraction of feature value by using bi-directional feature extraction algorithm

In this paper, a feature-based disparity estimation scheme is used to extract the feature points right after detecting the edge of object by the stereo image. Here, the feature point is a discontinuous point displayed by the pixel brightness depending on the characteristics of the stereo image. Generally, one directional matching method that is mostly used for disparity estimation is hard to find the correct matching point and as a result, wrong assignment of the disparity can be happened. Therefore, in this paper, bi-directional matching method is used to extract the feature points which are the same corresponding points between image of the right and left directions through the feature extraction algorithm. Canny mask operator is also used to extract the feature value for feature-based disparity estimation. This operator removes noises by using Gaussian mask and then it executes the edge detection operation such as Sobel mask. This operator has an advantage that is not sensitive to noises than other mask operators.

As a cost function, MSE (mean square error) is used to search for the corresponding pixel or block having smallest error correct disparity. Equation (3) expresses MSE function to extract the disparity of the stereo image in which N_x , N_y mean the block size and I_L and I_R represent the block of left image and the block of right image corresponding I_L .

$$MSE = \frac{1}{NxNy}\left[\sum_{i=1}^{Nx}\cdot\sum_{j=1}^{Ny}\left\{I_L(i, j) - I_R(i+k, j)\right\}^2\right] \qquad (3)$$

3.3 Adaptive disparity estimation by adaptive matching window size

The matching window size is determined by comparing the extracted feature point values and the predetermined threshold values. The threshold value is set from 1 to 5 and the initial threshold value and the matching window size are established, respectively. And then, the size value of feature point is determined by the threshold value. For the region having small feature value such as similar region, large matching window is selected, while small matching window is selected for the region having large feature value such as the edge of object. This method can not only reduce miss-estimation of the feature vector mostly happened in the conventional fine disparity estimation with small matching window size, but also reduce the blocking effect in the disparity estimation with big matching window size although representation of the fine movement of the object is difficult.

That is, in this paper, as a method to extract the disparity from the stereo image the block and pixel-based disparity estimation algorithm is selectively used depending on the matching window determined by feature-based disparity estimation.

3.4 Application of intermediate view reconstruction

By using the IVR technique, more natural stereopsis can be acquired through synthesizing the multiview stereo images from the limited given stereo image. As a result, offering of stereopsis to many observers is possible by implementation of available stereoscopic 3D display system. In this paper, more natural intermediate views can be reconstructed by using interpolation scheme with a weighted average. Equation (4) shows the case of interpolation with a weighted average by position of viewpoint.

$$I_P(i,j) = (1-a)\cdot I_r(i-\hat{d}_{ij}(i,j),j) + a\cdot I_l(i-\hat{d}_{ij}(i,j),j) \qquad (4)$$

Also, the mismatched disparity in the synthesized intermediate views by the proposed adaptive disparity estimation method is compensated through the process of disparity stability and misallocated disparity is replaced with the nearby disparity value. Through this process of intermediate views reconstruction, multiview 3D image can be created and displayed as more natural 3D image.

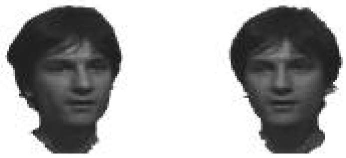
4. Simulation Result & Analysis

In the computer simulation, 'man' stereo image pairs of 256×256 pixel size are used as shown in Fig. 4. And, simulation results of the proposed method are compared with the conventional pixel-based matching and the pixel and feature-based matching methods.

(a) Left image (b) Right image

Fig. 4 Original stereo image pair of ‘man’

Fig. 5 is segmented objects from the ‘man’ image. Using the input stereo image pair, we obtain the disparity map including the occlusion region information. In the stereo image pair, position of the occlusion region corresponds to the left (reference) image and Right image.



(a) Left object (b) Right object

Fig. 5 Segmented objects of stereo image pair 'man'

Fig. 6 shows the disparity maps extracted from each of matching methods. Considering the disparity map of the Fig. 6, it is found that the disparity map of the proposed method is more stable than that of the other methods.

(a) (b) (c)

Fig 6. Disparity maps by synthesized 'man' image
(a) Pixel-based method (b) Pixel and feature-based method (c) Proposed method

Fig. 7 is intermediate view image by segmented objects from stereo image pair 'man'. Fig. 7 shows the intermediate view synthesized by each algorithm. In the pixel-based matching algorithm, it is found that mismatching is still occurred and some disparities are improperly assigned around the edge region of an object as shown in Fig. 7(a). Also, the feature-based matching algorithm can synthesize the intermediate view image in detail by accurate matching, but it is found that some mismatching. On the other hand, the image synthesized by the proposed method is well reconstructed even in the edge of the object and comparatively correctly matched by comparison to the other methods.

(a) (b) (c)

Fig 7. Synthesized 'man' image from each method
(a) Pixel-based method (b) Pixel and feature-based method (c) Proposed method

5. Conclusion

In this paper, a new 3D intermediate views reconstruction technique using an adaptive disparity estimation algorithm is proposed and its performance is analyzed by comparison to that of the conventional disparity estimation algorithm. In the proposed algorithm, in order to effectively synthesize the intermediate view the matching window size is selected according to the feature value of the input stereo image. By doing this, the mismatching probability can be reduced through coarse matching in the similar area and fine matching in the area having large feature values such as the edge part of object. From some experimental results on the IVR using the proposed algorithm is also discussed and compared with that of the conventional algorithms.

5. Acknowledgements

This research has been conducted by the ITRC project of the Korea Ministry of Information and Communication and the NRL project of the Korea Ministry of Science and Technology.

6. References

[1] Grinberg, V. "Geometry of binocular imaging," Proc. of SPIE, Vol. 2177, pp. 56 - 65, February 1994.

[2] M. Ebroul Izquierdo, "Stereo image analysis for multi-viewpoint teleprsence applications", Signal Processing : Image Comm. 11, 231-254, 1998.

[3] D.V. Papadimitriou and T.J. Dennis, "Epipolar line estimation and rectification for stereo image pairs", IEEE Trans. on image processing, vol. 5, no. 4, April 1996.

[4] Jens-Rainer Ohm and Karsten Muler, "Incomplete 3-D multiview representation of video objects," IEEE Trans. on Circuits and Systems for Video Tech., vol.9, no.2, March 1999.