IEEK Transactions on Smart Processing and Computing

Detection of View Reversal in a Stereo Video

Ji Deok Son and Byung Cheol Song

School of Electronic Engineering, Inha University / Incheon, South Korea bcsong@inha.ac.kr

***** Corresponding Author: Byung Cheol Song

Received July 14, 2013; Revised July 29, 2013; Accepted August 12, 2013; Published October 31, 2013

** Extended from a Conference: Preliminary results of this paper were presented at the ITC-CSCC 2013. This present paper has been accepted by the editorial board through the regular reviewing process that confirms the original contribution.*

** Short Paper*

***Abstract*:** This paper proposes a detection algorithm for view reversal in a stereoscopic video using a disparity map and motion vector field. We obtain the disparity map of a stereo image was obtained using a specific stereo matching algorithm and classify the image into the foreground and background. Next, the motion vector field of the image on a block basis was produced using a full search algorithm. Finally, the stereo image was considered to be reversed when the foreground moved toward the background and the covered region was in the foreground. The proposed algorithm achieved a good detection rate when the background was covered sufficiently by its moving foreground.

***Keywords*:** 3D image, Stereoscopic video, View reversal, Stereo matching, Motion estimation, Gradient

1. Introduction

Recently, the 3D video contents for TV, movies, and games has increased. The two views of a frame in a stereoscopic video, where one is referred to as the left view and the other is referred to as the right view, are intended for presentation to the left eye and right eye of a human visual system [1]. These 2D images are combined in the brain to give the perception of 3D depth.

On the other hand, the left view is shown in the right eye and the right view is shown in the left eye when the left view and right view are switched due to editing errors or software flaws. Fig. 1 shows a stereo image of a cylinder in front of a triangle. As a normal case, the triangle in Fig. 1(a) is focused behind the cylinder. Therefore, the stereoscopic image can be perceived. When view reversal occurs in this same scene, the triangle is focused in front of the cylinder, as shown in Fig. 1(b). That is, the object that is actually located in the foreground, is focused at the back, and the object, which is actually located in background, is focused at the front. This makes it difficult to recognize a real depth because the covered region of the triangle caused by cylinder provides a clue that the cylinder is in front of the triangle. As a result, the 3D effect cannot be perceived correctly when view reversal occurs. In addition, fatigue and dizziness can be experienced due to flickering at the covered region. Therefore, a detection algorithm is needed for view reversal even though this phenomenon does not occur frequently. Fig. 2 is a stereo image represented by an anaglyph. As an image separation technique, the left and right views are filtered with near-complementary colors and are viewed through respective color-filter glasses to direct both views to the appropriate eye [2]. Fig. 2(a) is a normal state and this image is perceived as 3D objects. On the other hand, in the reversal case (Fig. 2(b)), the 3D depth cannot be perceived correctly. Such a phenomenon becomes worse in the FPR (Film Patterned Retarder) type.

This paper proposes an effective algorithm to detect such a view reversal using disparity and motion information. First, a stereo image is divided into a foreground and background using disparity information that can be obtained by stereo matching. Next, the motion information caused by motion estimation is obtained, and eliminated some outliers. Finally, reversal can be detected if the covered region that occurs when the foreground moves to background or the background moves to the foreground. The simulation results show that the maximum detection rate of the proposed algorithm is 100% when a moving foreground sufficiently covers the static background, and 62.5% when the moving foreground

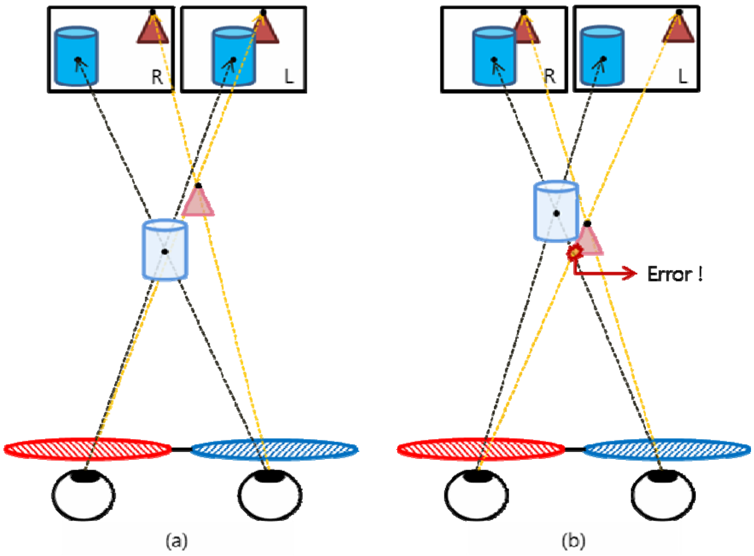

Fig. 1. Comparison of the normal case with the reversal case (a) Normal case, (b) Reversal case.

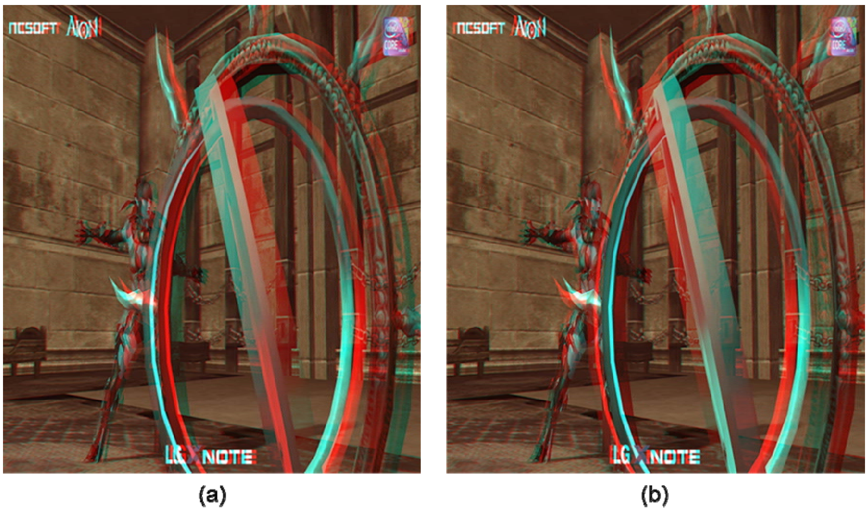

Fig. 2. Stereo image represented by an anaglyph (a) Normal case, (b) Reversed case.

covers the moving background.

The remainder of this paper is organized as follows. Section 2 describes the related work and in the proposed scheme section and a detection algorithm for view reversal is proposed. The performance evaluation section describes the simulation environment and results. Finally, the last section concludes the paper.

2. Related Work

A previous work used disparity and motion vector information to detect view reversal [3]. Stereo matching was performed to obtain disparity data and divide the image into the foreground and background. If the foreground covers the background region due to the movement of the foreground, view reversal can be detected by observing whether the covered region is in the foreground or background. Therefore, disparity information, such as accuracy, and motion information, such as motion vector size, was used to obtain an occurrence probability of view reversal. The lower SAD (Sum of Absolute Difference) produced by stereo matching induces the more accurate disparity. Therefore, higher scores were assigned to the disparity vectors with lower SADs. Similarly, higher weights were given to the motion vectors with lower SADs, which were obtained by a motion estimation. Higher scores were assigned to the larger motion vectors because the larger motion can have a wider covered region. Finally, all the mentioned information was used for scoring and the score ranges from 0% to 100%. A higher score means that the probability of view reversal is high.

In this paper, disparity and motion were used to detect view reversal, but a decision was made as to whether reversal occurs or not instead of scoring.

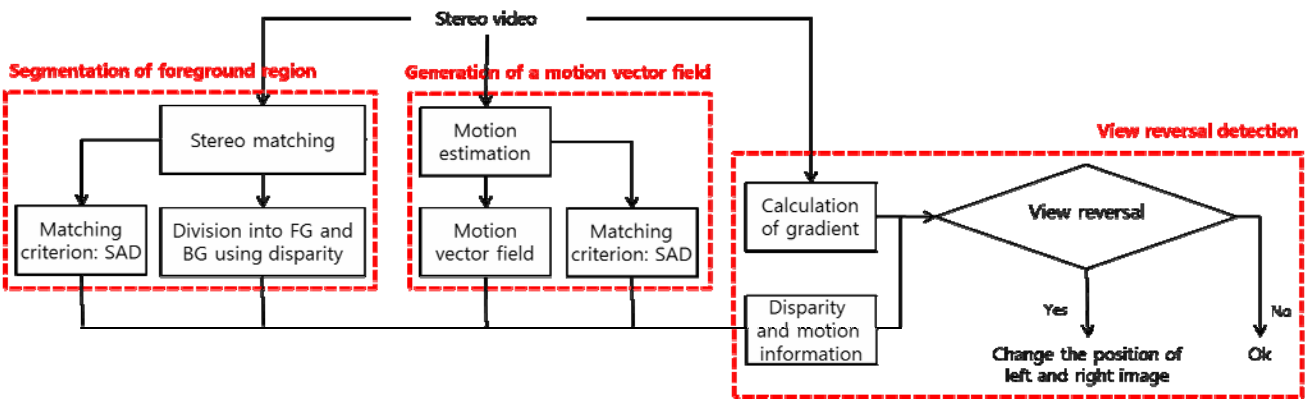

Fig. 3. Block diagram of the proposed algorithm.

3. The Proposed Scheme

This paper proposed a detection algorithm for view reversal in stereo videos. First, a disparity map employing stereo matching was obtained, and an input frame was classified into the foreground and background. Next, a motion vector field was produced using a specific motion estimation algorithm, and the gradient values on a block basis were calculated to exclude the flat regions from the candidates. Finally, view reversal cases were detected by observing whether the covered region was in the foreground or background.

Fig. 3 shows an overall block diagram for the proposed algorithm, and the details are described as follows.

3.1 Segmentation of the foreground region

Stereo matching is the process of taking two or more images and estimating a 3D model of the scene by finding the matching pixels in the images and converting their 2D positions into 3D depths [4]. To perform stereo matching, it was assumed that two pairs of an input stereo frame were already rectified. Therefore, block matching for stereo matching was performed only in the horizontal direction. A typical three-step hierarchical search was used to reduce the computational complexity. The disparity $\hat{d}$ which has the minimum SAD in the specific search range, was searched and a bi-directional check was performed to examine the accuracy of the block matching [5]. The final disparity was obtained by applying a closing operation and median filtering [6].

Next, a disparity histogram was derived, and the target frame was classified into a foreground and background using the disparity histogram. Here, the property that the disparity of the foreground is generally located on the right side of the disparity histogram and that of the background is located on the left side was used [7].

3.2 Generation of a motion vector field

A full search algorithm was used for motion estimation. The matching criterion between the current block B_{cur} and candidate blocks in all positions belonging to the specific search range were calculated. Among these candidates, the block with a minimum SAD was found, and a motion vector was assigned between the selected block B_{selec} and the current block. In addition, a bi-directional check was performed. For example, if B_{cur} in the current frame is matched with B_{selec} in the previous frame, B_{cur} was checked to determine if it was the best match of B_{selec} in the same search fashion. Through this procedure, the reliable blocks and their motion vectors could be found.

3.3 View reversal detection

In the flat region, the obtained disparity map and motion vector field might not be reliable. Therefore, gradient values were calculated to exclude the blocks in the flat region from the possible candidates.

$$G_x(x,y) = I(x-1,y) - I(x+1,y) \quad (1)$$

$$G_y(x,y) = I(x,y-1) - I(x,y+1) \quad (2)$$

Eqs. (1) and (2) were used to calculate the gradient values on a block basis in horizontally and vertically. Among the gradient values in each block, the number of gradient values that are higher than a particular threshold τ_1 were counted. If the counted value is greater than a specific threshold, τ_2, the current block was checked to determine if it satisfied the gradient condition.

Both the SAD values calculated during stereo matching and the motion estimation were exploited. The blocks with smaller SADs than a specific threshold were chosen to increase the detection accuracy.

On the other hand, an overlapped region occurs if the foreground moves toward the adjacent background or background moves to the adjacent foreground. The overlapped region should be the background. For example, object A of Fig. 4 moves toward object B in the current frame. The location of object B becomes object C in the next frame. If foreground A moves toward background B, object B will be occluded. On the other hand, object A will be occluded if view reversal occurs due to the reversed disparity information. In this case, an incorrect decision that object B is the foreground can be made. Fortunately, if

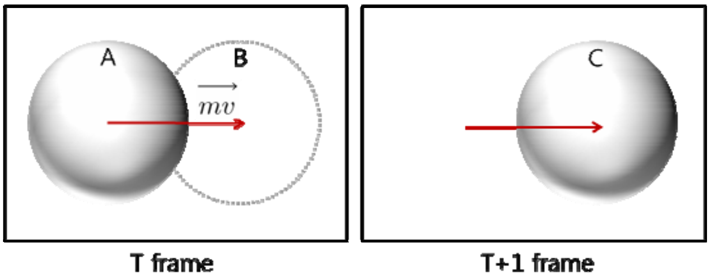

Fig. 4. Basic concept of a view reversal detection.

Flag1 =Flag2=Flag3=Flag4=0;
If (the number of the pixels whose $G_x(x,y) > \tau_1$ and $G_y(x,y) > \tau_1$ is larger than τ_2)
    Flag1 = 1;
If ($SAD_A^d < Th_d$ && $SAD_B^d < Th_d$ && $SAD_C^d < Th_d$ && $SAD_A^m < Th_m$)
    Flag2 = 1;
If(the direction of MV_A is toward object B, and $\|MV_B\| < \delta$)
    Flag3 = 1;

If(Flag1==Flag2==Flag3==1)
    If(object C is foreground)
        Normal;
    Else if(object C is background)
        Reversal;

Fig. 5. Pseudo code for detecting view reversal.

object A moves toward object B, as shown in Fig. 4(a), object B will be occluded by object A, which is already decided to be the background. Therefore, this view reversal can be detected by observing such a phenomenon.

Fig. 5 shows the pseudo code for detecting view reversal. First, objects A, B and C need to satisfy the afore-mentioned gradient condition. Next, the SADs derived from stereo matching, i.e. SAD^d values, should be smaller than a threshold Th_d , and the SAD obtained by motion estimation for object A, i.e. SAD_A^m, must be lower than a threshold Th_m. Next, if object B has a sufficiently small motion vector, and object A moves toward object B, a part of object B is occluded in the next frame. Finally, view reversal can be detected by determining if object C is the background or foreground.

4. Performance Evaluation

Twenty six stereoscopic video sequences extracted from various movies of 24fps and game videos of 30fps were used to evaluate the performance of the proposed algorithm. The video sequences were categorized into three classes. In Class 1, the background is sufficiently occluded by the moving foreground and the background is static. In Class 2, the background and foreground has motion. In Class 3, the foreground has motion, such as rotation or zoom. The video resolution is 1920×1080 except for two sequences of class 1, whose resolutions are 3840×1080 and 1440×576, respectively.

The false positive rate (FPR), which means the number of incorrectly detected sequences among the total number of sequences, was used to evaluate the performance of the proposed algorithm. A false negative rate (FNR) indicates that the number of missed sequences among the total number of sequences.

Fig. 6 shows the FPRs and FNRs according to the threshold, Th_d / n, i.e. an average matching error per pixel during a disparity estimation. Note that the search block size for stereo matching n was set to 32×32, and the detection rates for the Th_d / n values of 8 to 20 were calculated. In the case of class 1, when Th_d / n is small, the FPR is also small. In addition, the false detection disappears when Th_d / n reaches a specific value. In the case of class 2, the FNR increased with decreasing Th_d / n. This is because the number of candidates of class 2 is normally less than that of class 1. Note that only the candidates with a still background or a small motion were chosen. In the case of class 3, the FPR is zero, but the FNR is not negligible. This is because the proper candidates are rarely in the case of class 3.

Table 1 shows the maximum detection rates when there is no false detection. Note that class 1 where the background is still and the foreground sufficiently occludes the background is the best case. Class 2 shows a lower detection rate because the background has the same direction as the foreground. The sequences of class 3 also have some motion, but the foregrounds in those sequences

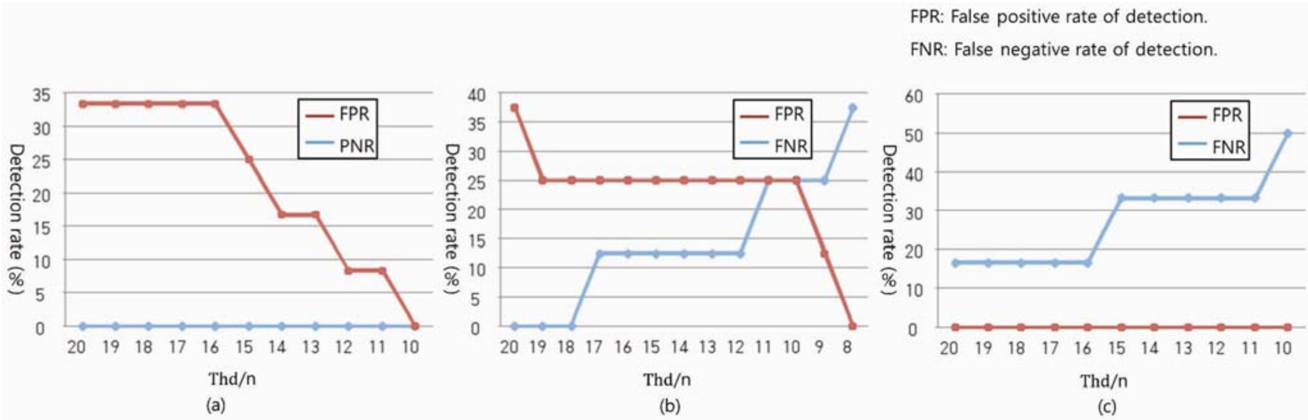

Fig. 1. Detection rate (a) Class 1, (b) Class 2, (c) Class 3.

Table 1. Maximum detection rate.

	Detection	Detection rate
Class 1	12/12	100%
Class 2	5/8	62.50%
Class 3	5/6	83.33%

sufficiently occlude their adjacent backgrounds. Therefore, class 3 also has a lower detection rate than class 1.

5. Conclusion

This paper proposed a detection algorithm for view reversal to prevent visual fatigue. View reversal was detected by examining whether a covered region is in the foreground or background because the foreground and background are reversed when view reversal occurs. The experimental results showed that the maximum detection rate of class 1 was 100% and that of class 2 and 3 was 62.5% and 82.33%, respectively.

Acknowledgement

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MEST) (No. 2012000446)

References

[1] A. Puri, R. V. Kollarits and B. G. Haskell, "Basics of stereoscopic video, new compression results with MPEG-2 and a proposal for MPEG-4," *Signal Processing: Image Communication*, vol. 10, pp. 201-234, Jul. 1997. Article (CrossRef Link)

[2] L. M. J. Meesters, W. A. IJsselsteijn and P. J. H. Seuntiens, "A survey of perceptual evaluations and requirements of three-dimensional TV," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 14, no. 3, pp. 381-391, Mar. 2004. Article (CrossRef Link)

[3] J. D. Son and B. C. Song, "Left-right inversion detection in side-by-side format stereo videos," *IEEK Conf.*, Nov. 2012.

[4] R. Szeliski, Computer Vision: Algorithms and Applications, Springer, 2010. Article (CrossRef Link)

[5] M. E. Izquierdo, "Stereo image analysis for multi-viewpoint teleprecense applications," *Signal Processing: Image Communications*, vol. 11, no. 3, pp. 231-254, Jan. 1998. Article (CrossRef Link)

[6] R. Gonzalez and R. Woods, Digital Image Processing, Prentice-Hall, 2007.

[7] H. G. Kim, J. K. Kang and B. C. Song, "Three-dimensional perception improvement using sharpness adjustment and hardware implementation," *in IEEE International Conference on Multimedia and Expo Workshop*, Jul. 2013. Article (CrossRef Link)

Ji Deok Son received her B.S. degree in electronic engineering from Inha University, Incheon, Korea in 2012. Currently, she is pursuing a M.S. degree in electronic engineering from Inha University. Her research interests include image processing, super-resolution, and video coding.

Byung Cheol Song received his B.S., M.S., and Ph.D. degrees in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 1994, 1996, and 2001, respectively. From 2001 to 2008, he was a senior engineer at Digital Media R&D Center, Samsung Electronics Co., Ltd., Suwon, Korea. In March 2008, he joined the School of Electronic Engineering, Inha University, Incheon, Korea, and is currently an associate professor. His research interests are in the general areas of image/video processing, super-resolution, 3D vision, multimedia system design, and content-based multimedia retrieval.