제 18 회 한국정보처리학회 추계학술발표대회 논문집 제 9 권 제 2 호 (2002.11)

Novel Scene Generation Along The Viewing Positions

Dae-Hyun Kim*, Tae-Eun Kim**, Jong-Soo Choi*
*Dept. of Image Eng., Graduate School of AIM, Chung-Ang University
**Dept. of Multimedia, Nam-Seoul University

ABSTRACT

In this paper, we propose an algorithm to automatically generate a novel scene along the altemated viewing positions using the bidirectional disparity morphing. To compute the disparity between two reference images, we use the 2-step block based matching algorithm and define three occluding pattems so as to smooth the computed disparities, especially for occluded regions. We discuss the advantages of this algorithm compared to the commonly used schemes and show some experimental results with real data.

1. Introduction

There has been increased interest both for computer vision and graphics in image-based rendering (IBR) methods, which deal with how to produce a novel scene from an arbitrary viewpoint, given a set of reference images. So far, the following two methods are prevalent to realize the view interpolation [1,2,5,6].

A) ***3D model based method*** developed in computer vision is a natural solution to IBR by transferring image textures onto it. It can present an arbitrary 2D view of an object as long as a complete 3D model exists. However, the task to construct the 3D model of a real scene automatically based on its 2D images is very difficult.

B) ***2D image based method*** has recently been developed for generating novel scenes from multiple images without recovering 3D structure. Because IBR is essentially based on 2D image processing, the errors in 3D shape reconstruction do not affect the quality of the generated new images as much as for model-based rendering. This implies that the quality of the input images can be well preserved in the generated novel scenes.

In this paper, we propose an algorithm to automatically generate a novel scene along the altemated viewing positions using the bidirectional disparity morphing (BDM).

As illustrated in figure 1, it needs a dense disparity map to generate a precise novel scene. However, it is expensive for computing a disparity map and false disparity caused by occlusion makes low the quality of the novel scene.

In order to speed up the computation of the disparity, we propose 2-step block based matching algorithm (BBMA). This method, at first step, roughly determines the corresponding candidate and then closely searches the corresponding block within a restricted region. In this procedure, an occlusion is detectable because the estimation error of the occluded block is larger than that of the visible block.

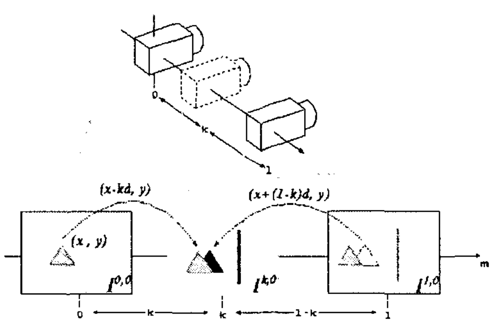


Figure 1. Parallel camera configuration [14]

From the properties that the disparity of the occlusion in the left image is similar to the disparity of the left layer of the occluding boundary and the right image is vice versa, we can define three occluding patterns. These patterns remarkably refines the occluding boundary.

Synthesis is performed separately from both left and right images, and then the individual novel scenes are dissolved into the final synthesized scene to minimize the effect of disparity estimation errors[1].

The paper is organized as follows. Section 2 describes the BDM for computing an intermediate coordinate and proposes how to fill the holes. Section 3 defines 2-step block based matching algorithm (BBMA) and computes a disparity map while detecting the occluding regions. Then we define three occluding patterns used to compensate the false disparities. Section 4 shows the experimental results and evaluates the proposed algorithm in terms of the performance. Finally section 5 makes the conclusions.

2. Bidirectional Disparity Morphing(BDM)

We assume that the cameras used to capture a given stereopair were separated by a horizontal distance and arranged in the parallel camera configuration as depicted in figure 1. A point in the scene generates corresponding points p_l and p_r in the left and right images, respectively. If the scene point is visible in both images, the disparity is defined as the distance, in pixels, between the corresponding points. Due to the parallel camera setup, the disparity is only in the horizontal direction and the relationship is given by

$$p_r = \begin{bmatrix} x_r \\ y_r \end{bmatrix} = \begin{bmatrix} x_l + d_{L\to R}(x_l, y_l) \\ y_l \end{bmatrix} = p_l + \begin{bmatrix} d_{L\to R}(x_l, y_l) \\ 0 \end{bmatrix} \tag{1}$$

where $d_{L\to R}(x_l, y_l)$ is the disparity from the left pixel to the corresponding right pixel.

The distance between the left and right images is normalized to one. The desired novel scene is then parameterized by morphing coefficient(s) where $0 < s < 1$. The mapping from the given left image to the position in the novel scene is a function of disparity as shown by the linear relationship,

$$p_s = \begin{bmatrix} x_s \\ y_s \end{bmatrix} = \begin{bmatrix} x_l + s \cdot d_{L\to R}(x_l, y_l) \\ y_l \end{bmatrix} \tag{2}$$

Equation (2), however, does not model the effects that changes in visibility have on image content[8]. From the standpoint of morphing, change in visibility results in two types of conditions : ***folds*** and ***holes*** [7,8,12].

Folds can be easily resolved by using Z-buffering algorithm[7,8]. Unlike ***folds***, ***holes*** cannot always be eliminated by using image information alone. Even though it is prevalent in existing image interpolation[8,4], the neighborhood interpolation approach has a limit for larger disparity scenes. In this paper, therefore, we first generate the novel scenes of both left and right images using BDM, and then these two novel scenes are used to fill the opposite image's holes because the holes in one's novel scene correspond to the visible regions in the opposite image. As a result, we can improve the visible realism of the novel scenes.

Synthesis is performed separately from both left and right images, and then the each novel scene is dissolved into the final synthesized scene to minimize the effect of disparity estimation errors[1]. The relationship is given by

$$IP_s(x,y) = \begin{cases} w_{L\to R} \cdot IP_{s,L\to R}(x,y) + w_{R\to L} \cdot IP_{s,R\to L}(x,y) & if\, IP_{s,L\to R} \neq 0, IP_{s,R\to L} \neq 0 \\ IP_{s,L\to R}(x,y) & if\, IP_{s,L\to R} \neq 0, IP_{s,R\to L} = 0 \\ IP_{s,R\to L}(x,y) & if\, IP_{s,L\to R} = 0, IP_{s,R\to L} \neq 0 \\ Nearest\ neighbor\ pixel & if\, IP_{s,L\to R} = 0, IP_{s,R\to L} = 0 \end{cases} \tag{3}$$

where $IP_s(x,y)$ is a dissolved intensity value at (x,y), and $IP_{s,L\to R}(x,y)$ and $IP_{s,R\to L}(x,y)$ are the interpolated intensity values from both left and right image, respectively. If $IP_{s,L\to R}(x,y)$ and $IP_{s,R\to L}(x,y)$ are not zero, the intensity value is determined by the weighted average of the intensity values at the corresponding points. The weighted average gives a priority to the information from the nearer viewing position[2].

3. Computation of Disparity Map

3.1 2-step block based matching algorithm

A) Restriction of searching range

Because we assume the parallel camera setup, we have $y = y_L = y_R$ and $x_L \geq x_R$ [3]. These constraints can restrict the searching range half

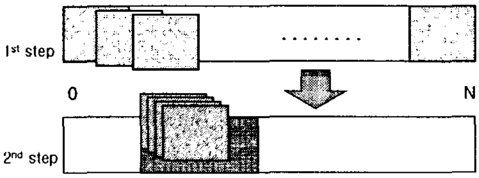


Figure 2. 2-step block-based matching algorithm

B) 2-step search procedure

We propose a 2-step search procedure. As shown in figure 2, we first search roughly the corresponding candidate block within all searching range and then closely investigate the corresponding block within the restricted narrow searching range.

C) Detecting an occluding region

A block which different layers are mutually overlapped, in general, presents larger estimation error compared to other blocks and it has a higher probability that the occlusion occurs. Therefore we deem the 8×8 block to be an occlusion if its estimation error is larger than a threshold.

3.2 Three occluding patterns

The boundary of the occlusion intends to be vertical because of the parallel camera setup. The occlusion in the left image is occurred when either the foreground moves to left or the camera moves to right. And the boundary of the occlusion is seen in the left side of the foreground. That is, the disparity of the occlusion is similar to the disparity of the left layer of the occluding boundary. The occlusion in the right image represents the opposite properties. From these properties, we can define three occluding patterns. Figure 3(a) presents the median mask for both visible blocks. Figure 3(b) and (c) present the left and right occluding patterns, respectively. The grayed m_0 presents the central pixel of the median mask and it is located on the boundary of the occlusion.

With these patterns, we perform the 3×3 median filter that is replaced by the median of the pixels within the mask so as to smooth the computed disparity map.

m_1	m_2	m_3
m_4	$\mathbf{m_0}$	m_5
m_6	m_7	m_8

(a)

m_1	m_2	m_3
m_4	m_5	$\mathbf{m_0}$
m_6	m_7	m_8

(b)

m_1	m_2	m_3
$\mathbf{m_0}$	m_4	m_5
m_6	m_7	m_8

(c)

Figure 3. Three occluding patterns. (a) pattern of the both visible blocks, (b) pattern of the left occluded blocks, (c) pattern of the right occluded blocks

4. Experimental results

We make an experiment on the proposed algorithm with two image sequences shown in figure 4. The experimented images are 640×480 RGB color images, which are acquired from the 1×8 camera arrays in the university of Tsukuba in Japan.

8×8 block is used for computing a disparity map and the overall searching range is $0 \le d \le 127$, and local searching range is $-7 \le d \le 7$. In the first step, a block is moved by 4 pixels to determine the candidate block and then closely investigated.

(a) Kid sequence (b) City sequence

Figure 4. Experimented image sequences

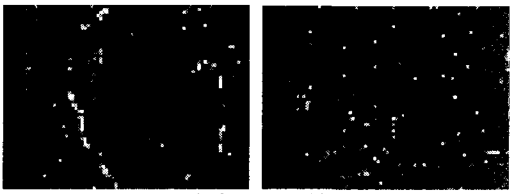
4.1 Detecting an occluded block

We use the mean absolute difference (MAD) of Eq.(4) as a matching criteria and determine the corresponding block which MAD is minimum. However, it is not reasonable to simply select the minimum block as a correspondence since the occluded block doesn't have any correspondence.

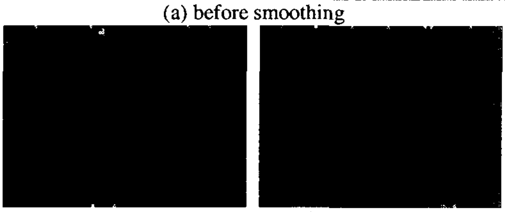
$$E(d_x) = \sum_{(x,y)\in B} |I_m(x,y) - I_{m+1}(x+d_x, y)| \qquad (4)$$
$$[\hat{d}_x] = \arg\min_{(d_x)}[E(d_x)]$$

We set the threshold up 10% of the maximum estimation error through the repeated experiments and decide whether it is a correspondence or not.

4.2 Smoothing with three occluding patterns



(a) before smoothing



(b) after smoothing

Figure 5. Disparity maps before and after smoothing

We smooth the disparity maps with three occluding patterns defined in section 3.2 in order to improve the reliability of the computed disparity maps. Figure 5 presents the disparity maps of figure 4 before and after the smoothing is performed.

White blocks in figure 5(a) present the occluded blocks and false disparities. After being smoothed with three occluding patterns, the boundary between the foreground and background becomes sharper and the false disparities are well corrected.

4.3 Synthesis

Once we compute the disparity maps from both left and right images, we can interpolate the novel scenes by using BDM. Then two interpolated scenes are dissolved to minimize the effect of the disparity estimation errors. Figure 6 presents some examples of the novel scenes for various morphing coefficients(s) and it is also visibly satisfactory.

4.4 Performance evaluation

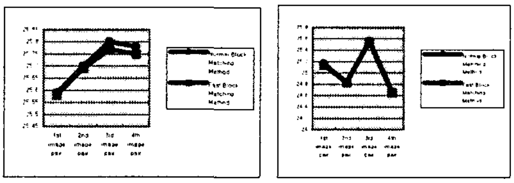
We use the peak signal noise ratio (PSNR) so as to evaluate the proposed algorithm in terms of the performance.

Table 1 compares the amount of the computation via the PC profile, and figure 7 illustrates the individual performance before and after applying the 2-step BBMA. As a result, the proposed 2-step BBMA cannot only be even cheaper about 53%, but also preserve its performance.

Figure 8 compares the results of whether we apply the median filter with three occluding patterns or not. This also presents that the proposed algorithm is averagely 0.05dB higher. Even though they don't improve the overall quality of the images, three occluding patterns can remarkably refine the novel images around the occluding boundary and make sharper the occluding boundary.

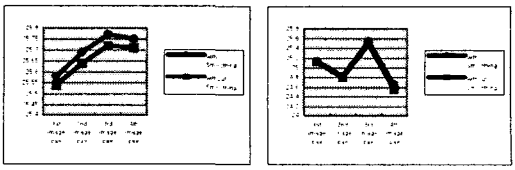
Table 1. Profile for computation of disparity map

	normal block matching	2-step fast block mathcing
Required time(msec)	955.574	511.634



(a) Kid (b) City

Figure 7. Performance for various morphing coefficients



(a) Kid (b) City

Figure 8. Performance for three occluding patterns

5. Conclusion

In this paper, we proposed the algorithm to automatically generate a novel scene using the bidirectional disparity morphing (BDM) and three occluding patterns from the parallel stereopair.

2-step BBMA is able to restrict the searching range and accelerate the speed of the computation of the disparity, thanks to the parallel camera setup.

We defined three occluding patterns to smooth the computed disparity map. These patterns vary the location of the central pixel of the median mask according to the visibility.

The result of the evaluation of the proposed algorithm presents that 2-step BBMA lowers the cost with keeping up the previous performance, and the smoothing with three occluding patterns slightly increases overall PSNR of the novel images as it makes the boundary between the foreground and the background sharper.

References

[1] J. S. McVeigh, M. W. Siegel, and A. G. Jordan, "Intermediate view synthesis considering occluded and ambiguously referenced image regions," *Signal Processing: Image Communication*, pp. 21 - 28 , Vol. 9, 1996.

[2] K. Satoh, I. Kitahara, Y. Ohta, "3D Image Display with Motion Parallax by Camera Matrix Stereo," *Proc. of 3rd Int. Conf. on Multimedia Computing and Systems (ICMCS'96)*, 1996.

[3] Redert, E. Hendriks, J. Biemond, "3D Scene Reconstruction with Viewpoint Adaptation on Stereo Displays," *IEEE Trans. On Circuits and Systems for Video Technology*, pp.550-562, Vol. 10, No.4, June 2000.

[4] K. Satoh, Y. Ohta, "Occlusion Detectable Stereo Using A Camera Matrix," *Proc. 2nd Asian Conf. on Computer Vision (ACCV'95)*, pp.331-335, Vol.2, 1995.

[5] Takeo Kanade, P. J. Narayanan, P. W. Rander, "Virtualized reality: concepts and early resu-lts," *Proc. of IEEE Workshop on Representation of Visual Scenes*, pp.69-76, 1995.

[6] Hideo Saito, S. Baba, M. Kimura, S. Vedula, T. Kanade, "Appearance-based virtual view gener-ation of temporally-varying events from multi-camera images in the 3D room," *Proc. of 2nd Int. Conf. on 3-D Digital Imaging and Modeling*, pp.516–525, 1999.

[7] S. M. Seitz, C. R. Dyer, "View Morphing," *Proc. of SIGGRAPH'96*, pp.21-30, 1996.

[8] S. E. Chen, L. Williams, "View Interpolation for Image Synthesis," *Proc. of SIGGRAPH 93*, pp.279-288, 1993.

[9] C. L. Zitnicj, T. Kanade, "A Cooperative Algorithm for Stereo Matching and Occlusion Detect-ion," *IEEE Trans. On Pattern Analysis and Mac-hine Intelligence*, pp.675-684, Vol.22, No.7, July 2000.

[10] J. R. Ohm, E. Izquierdo M., "An Object-Based System for Stereoscopic Viewpoint Synthesis," *IEEE Trans. On Circuits and Systems for Video Technology*, pp.801-811, Vol.5, No.5, Oct. 2000.

[11] Y. Sugaya, Y. Ohta, "Stereo by integration of two algorithms with/without occlusion handling," *Proc. 15th International Conference on Pattern Recognition(ICPR2000)*, Vol.1, pp.109-113, Sep. 2000.

[12] S. M. Seitz, C. R. Dyer, "Physically-valid view synthesis by image interpolation," *Proc. of IEEE Workshop on Representation of Visual Scenes*, pp.18-25, 1995.

[13] M. Lhuillier, L. Quan, "Image Interpolation by Joint View Triangulation," *Proc. of Computer Vision and Pattern Recognition(CVPR'95)*, 1995.

[14] http://www.esys.tsukuba.ac.jp/research/3DTV/3dtv.html

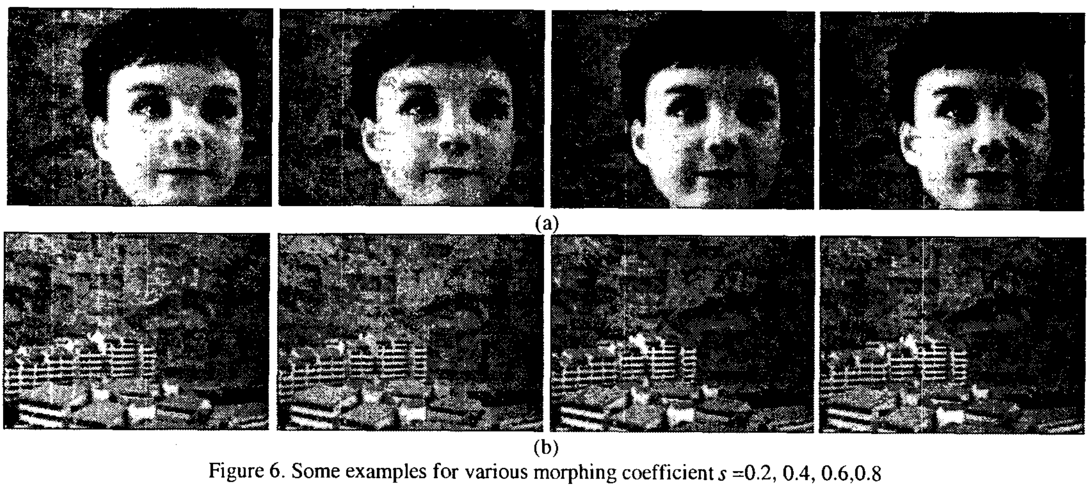


Figure 6. Some examples for various morphing coefficient s =0.2, 0.4, 0.6,0.8