REALISTIC 3D VIEW GENERATION USING DEFORMAL STEREO MATCHING

Jae-Chul Kim, Jong-Hyun Park

LBS Research Team,Telematics Research Division, ETRI
161 Kajong-Dong, Yusong-Gu, Taejon, 305-350, Republic of Korea
kimjc@etri.re.kr

ABSTRACT:

In this paper, we propose an adaptive stereo matching algorithm to encompassing stereo matching problems in projective distortion region. Since the projective distortion region can not be estimated in terms of fixed-size block matching algorithm, we tried to use adaptive window warping method in hierarchical matching process to compensate the perspective distortions. In addition, probability theory was adopted to encompass uncertainty of disparity of points over the window in this study. The proposed stereo matching algorithm has tested on both disparity map and 3D model view. The experimental result shows that remarkable improvement is obtained in the projective distortion region.

1. INTRODUCTION

Stereo images are obtained from the different perspective position(Figure1), so that each image can have the effect of projective distortion as shown in Figure 2.
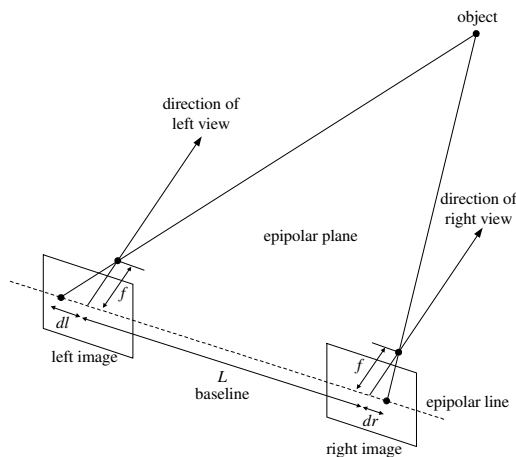

Figure 1. Stereo geometry with parallel axes.

Since the surface of a real object is projected on the right and left cameras, each projected-image has different viewing characteristics (Note Figure2 show that area of dr is larger than that of dl). A main problem of stereo matching heavily depends upon selecting an appropriate window size. However, same window size, both reference window and searching window, does not appropriate for projective distortion region. It is necessary to consider projective distortion in stereo matching process. Due to above result, many researcher (T. Kanade, 1994; Z. F. Wang, 1994; P. N. Belhumeur, 1992) have been interested in the effect of projective distortion in recent years.
Kanade (T. Kanade, 1994) presented an adaptive window method to reduce the effect of projective distortion. His method employs a statistical model of the disparity distribution within a window. By evaluating the local variation of the intensity and the disparity, the method can select an appropriate window size

and estimate disparity with the least uncertainty for each pixel of an image. Wang and Ohnishi (Z. F. Wang, 1994) presented a 3D-to-3D method. Instead of using a statistical model of the disparity distribution, the method uses a geometric model of the matching process (G. Medioni, 1985).
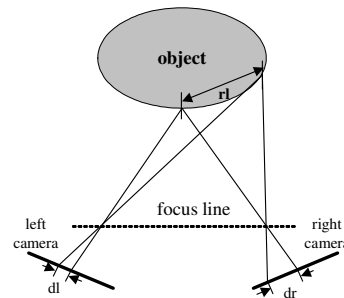

Figure 2. Projective distortion.

A transformational template-based matching method is used to recover the depth data from the projective distortion information. There has been no study that tried to use window warping technique (G. Wolberg, 1994) in the projective distortion.
The algorithm we propose is based on wondow warping technique. Furthermore Fant's algorithm (G. Wolberg, 1994) is used in order to warp a local mask in hierarchical block-matching process.

2. WINDOW WARPING ALGORITHM

2.1 Projective Distortion

Since surface normal is often very tilted with respect to the optical axes of camera, the projected stereo images has projective distortion. The evidence of the projective distortion can be found in the comparison of line profiles (Figure 3). In Figure 3, I_{xR} and I_{xL} (270th line profiles) are a line profile of right and left image. These are give a good account of projective distortion phenomenon in 'man' stereo images. Points m_1,and

m_2 represent matching end point between r_1 and l_1, on the other hand m_3 and m_4 are matching end point between r_2 and l_2.
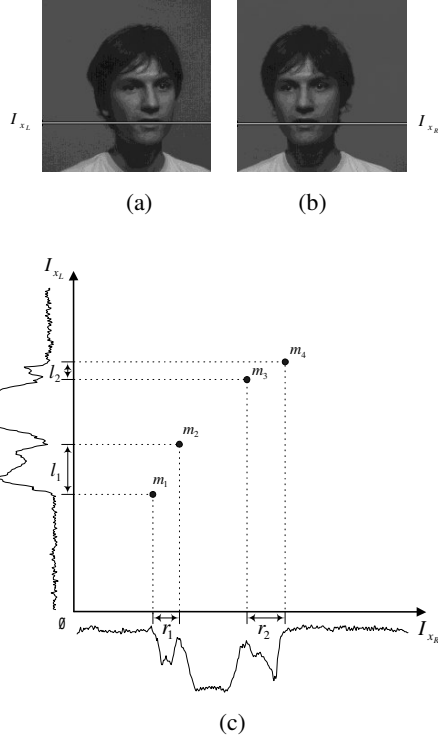

(a)　　　　　(b)

(c)

Figure 3. The original stereo images of 'man' and corresponding projective distortion result. (a) Line profile on left image. (b) Line profile on right image. (c) Correspondence of line profiles.

The length of l_1 is not equal to that of r_1. but l_1 corresponds to r_1 in real world coordination. Inevitably, other window-based stereo matching method is not fit to projective stereo image (T. Kanade, 1994). Since conventional window-based stereo matching methods use the same size in reference local window and searching local window, those methods cannot found correct matching point. In conclusion, analyses above show that wndow warping technique is necessary needed in matching process.
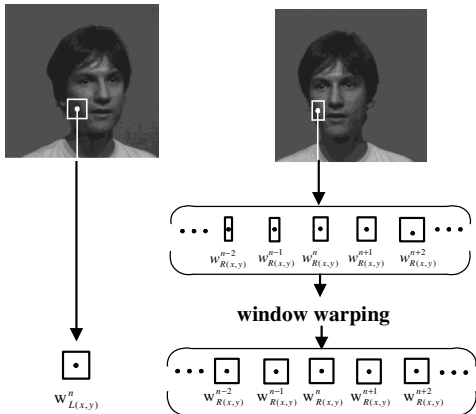
2.2 Window Warping using Fant's algoritm

Figure 4. The window warping using Fant's algorithm.

We use the Fant's algorithm so as to warp local windows in hierarchical BMA process. First, local window in low reliability region is warped to candidate windows. Second, we evaluate matching possibility between reference window and warped window.

The main benefit of this separable algorithm is the lower complexity in 1-D resampling. The fittest window among the candidate window can be represented as Equation (1)

$$w_\varphi = \arg\min_{k \in \Omega} \left| w_{L(x,y)}^n - w_{R(x,y)}^{n+k} \right|,$$

(1)

where Ω is window warping size, and w is a local window in a block matching. We warp a local mask to the horizontal direction under the assumption of epipolar geometry on an input image.

3. PROPOSED STEREO MATCHING WITH AN ADAPTIVE WINDOW WARPING ALGORITHM

Figure 5 shows a stereo matching method we proposed. The proposed matching algorithm consists of four steps. The first step is an object extraction. In the second step, hierarchical block matching is performed to avoid local minima. The multi-resolution images are used in hierarchical block matching because the upper disparity result has the disambiguate characteristic. In the third step, the reliability-map with two disparity-maps (left-to-right and right-to-left) is obtained. Reliability includes bi-directional consistency checking and local disparity variance. The last step is the block matching process such that the warping window is used in a low reliable region and the normal window is used in a high reliable region..
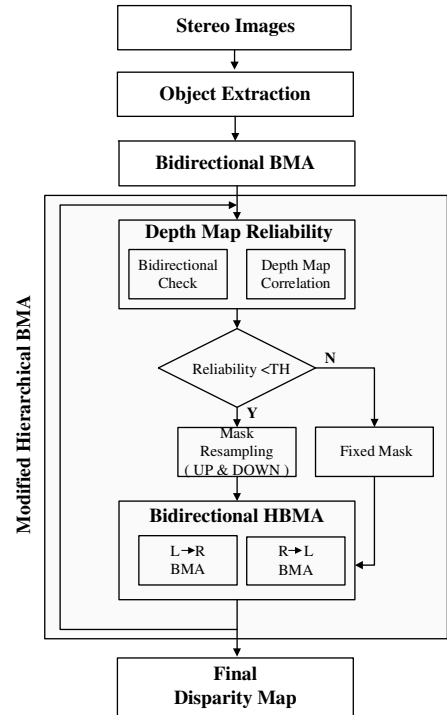

Figure 5. The block diagram of the proposed algorithm.

Since Fant's algorithm is faster than conventional re-sampling technique that uses low-pass filtering and sub-sampling, his algorithm is used as a warping procedure in this study. Also Fant's algorithm has one step process comparison with conventional re-sampling technique (two step). In particularly conventional method has a complicated process (low pass filtering or interpolation) in the re-sampling process

3.1 Object Extraction

Since indoor stereo images have a simple back ground, the object of stereo images can be separated from background. Processing time is reduced and accurate matching result is given. This object extraction algorithm consists of six steps : labeling, top-down merging, bottom-up merging, renumbering, small region merging, and object extraction.

3.2 Modified Hierarchical Block Matching Algorithm

In order to reduce noise sensitivity and reach higher efficiency simultaneously, both the left and right images are low-pass-filtered and sub-sampled. We propose a modified hierarchical block matching algorithm shown in Figure 6. Hierarchical block matching algorithm used in this paper consists of four steps as follows.
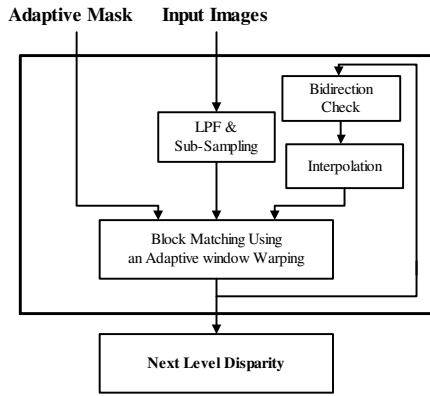

Figure 6. Modified hierarchical block matching algorithm.

Step 1 : Low pass filtering and sub-sampling
    Input images are passed a LPF(low pass fitering) and subsampled.
Step 2 : Block Matching Process
    Half-sized input images are used in stereo matching.
Step 3 : Bidirection checking
    Disparity map that found in previous process is checked using the bidirectional constraint.
Step 4 : Interpolation
    Remained disparity value is linear-interpolated.

3.2.1 Block Matching Algorithm: First process is a block matching process. The result of this process is used to generate reliability function. The matching process is performed in parallel sequence.
Distance-measure($D(x,y,d)$), we used, is Mean-Absolute-Distance (MAD). Each disparity map is obtained by calculating the MAD between a pair of stereo images for a particular window size as Equation (2).

$$D(x, y, d) = \arg\min_{d \in (x,y)} \left[\frac{1}{mn} \sum_{y'=-\frac{w}{2}}^{\frac{w}{2}} \sum_{x'=-\frac{w}{2}}^{\frac{w}{2}} |I_R(x + x', y + y') - I_L(x + d + x', y + y')| \right],$$

(2)

where d is the window shift in the right image, m and n are the mask sizes, w is the window size, and I_L and I_R are the gray-levels of the left and right images, respectively. Here we employ an epipolar constraint to reduce computational cost, assuming that an epipolar line can be determined first. The value of d that minimizes the MAD is considered to be the disparity at each pixel position(x,y).

3.3 Reliability Generation

In order to estimate reliability of matching points, we use matching constraints: uniqueness and smoothness. Other constrains are considered in matching process.

3.3.1 Uniqueness Constraint: A given pixel from one image can match no more than one pixel from the other images. However, bidirectional check shows that each disparity value has small difference in Figure 7. And the bidirectional check distance is defined as

$$\delta = |D_L(x, y) - D_L(x + D_R(x, y))|.$$

(3)

The uniqueness condition can be tested for each sampling position (x,y). The deviation δ can be seen as measure of the perturbation.
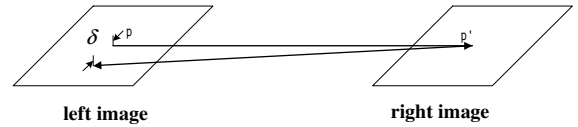

Figure 7. The bidirectional check.

Where δ describes the difference given by the relation (3). In order to consider how much input image satisfy the uniqueness constraints, the following reliability function is employed:

$$f_{R1} = \frac{T_{bc} - \delta}{T_{bc}}$$

(4)

Where T_{bc} : threshold
    δ : bidirectional check distance

where f_{R1} represents the uniqueness reliability, δ is bi-directional check error, and T_{bc} is minimum threshold.

3.3.2 Smoothness Constraint:

The matching pixels must have similar intensity values(i.e. differ lower than a specified threshold) or the matching windows must be highly correlated. Thus, the next probability is proposed by using this characteristic. We calculate a variance in a disparity map that satisfies uniqueness constraint, note (Equation 4). Each value is extracted from weight-function:

$$w_s(x,y) = \begin{cases} 0 & , \delta > T_{bc} \\ 1 & , \delta < T_{bc} \end{cases} \quad . \tag{5}$$

Consequently, we define local variance in the extracted disparity map:

$$\sigma_D^2(x,y) = \left[\frac{1}{N_V} \sum_{i=-\frac{w}{2}}^{\frac{w}{2}} \sum_{j=-\frac{w}{2}}^{\frac{w}{2}} \{w_s(i+x,j+y) \cdot I(i+x,j+y)\}^2 \right]$$
$$- \left[\frac{1}{N_V} \sum_{i=-\frac{w}{2}}^{\frac{w}{2}} \sum_{j=-\frac{w}{2}}^{\frac{w}{2}} w_s(i+x,j+y) \cdot I(i+x,j+y) \right]^2, \tag{6}$$

$$N_V = \sum_{i=-\frac{w}{2}}^{\frac{w}{2}} \sum_{j=-\frac{w}{2}}^{\frac{w}{2}} w_s(i+x,j+y).$$

where

Using these conventions, we define the smoothness reliability:

$$f_{R2} = e^{-\xi \sigma_D^2}, \tag{7}$$

where ξ is an damping parameter to control the decreasing velocity of f_{R2}. The function value of f_{R2} is large if σ_D^2 is small, whereas f_{R2} will take small value in opposite case.

3.3.3 Total Reliability

The final reliability function is defined as a linear combination of f_{R1} and f_{R2} :

$$f_{total} = \lambda_1 f_{R1} + \lambda_2 f_{R2}, \tag{8}$$

where λ_1 and λ_2 are weight coefficients satisfying the relation $\lambda_1 + \lambda_2 = 1$.

Here, we consider local intensity and smoothness constraint. The number of upper disparity vectors is 9 and that of current state vectors is 4. Each disparity map is obtained by calculating the mean absolute difference (MAD) and smoothness of upper disparity vectors between a pair of stereo. In the modified hierarchical BMA, we define cost functions S_i and D_i as:

$$S^i(x,y,d) = \frac{1}{9} \sum_{y^i=-1}^{1} \sum_{x^i=-1}^{1} D^{i-1}(x^i,y^i) - D^i, \tag{9}$$

$$D^i(x,y,d) = \arg\min_{d \in (x,y)} \left[\frac{1}{mn} \sum_{y'=-\frac{w_\varphi}{2}}^{\frac{w_\varphi}{2}} \sum_{x'=-\frac{w_\varphi}{2}}^{\frac{w_\varphi}{2}} \left| I^i_R(x+x',y+y') \right. \right.$$
$$\left. \left. - I^i_L(x+d+x',y+y') \right| \right] \tag{10}$$

and we can compose our cost function as

$$E^i = \gamma_1 D^i + \gamma_2 S^i, \tag{11}$$

where Di denotes the cost function at level of the ith pyramid, the other parameters are the same as Equation(2), γ_1 and γ_2 are weight coefficients satisfying the relation $\gamma_1 + \gamma_2 = 1$.

4. EXPERIMENTS

The simulation result indicates that proposed algorithm shows outstanding ability in low texture and projective distorted region such as the nose and cheek in Figure 9 and Figure 13.

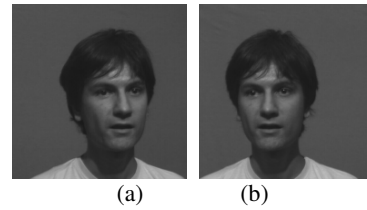| Image \ Items	"Man" image	"Claude" image
Image size	384x384	720x288
Actual disparity	About –100~100	About –30~30
Searching range	-120~120	-45~45
Block size	7x5	7x5
T_{bc}	6	5
Weights (λ_1 / λ_2)	0.5/0.5	0.5/0.5
Weights (γ_1 / γ_2)	0.4/0.6	0.4/0.6

Table 1. Parameters for correspondence estimation

(a)   (b)

Figure 8. The original stereo images of 'man', (a) Left image, (b) Right image.

(a)   (b)

Figure 9. "claude " image pair. (a) Left image, (b) Right image.
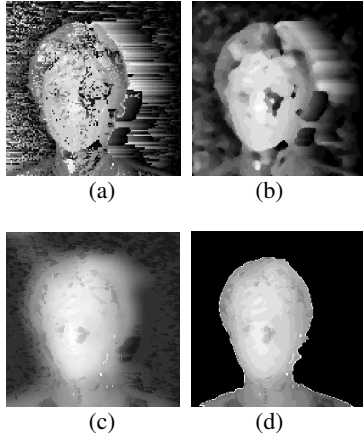
Figure 10. Disparity map. (a) Block-Matching algorithm, (b) Population-Based Incremental Learning algorithm, (c) Proposed algorithm without object extraction, (d) Disparity map by using proposed algorithm.
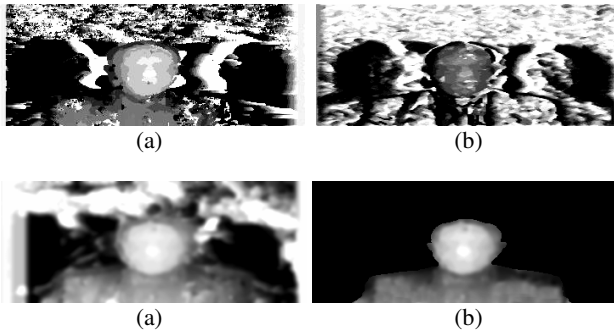

Figure 11. Disparity map of "claude" stereo image. (a) Block-Matching algorithm with fixed-size window(7x5), (b) Population-Based Incremental Learning algorithm, (c) Proposed algorithm without object extraction, (d) Disparity map by using proposed algorithm.
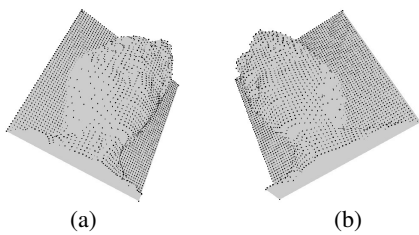

Figure 12. Isometric plot of the disparity maps computed by proposed method. (a) Down left view, (b) Down right view.
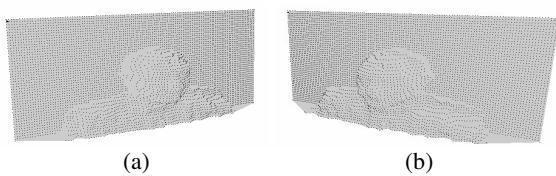

Figure 13. Isometric plot of the disparity maps computed by the proposed method. (a) Down left view, (b) Down right view.
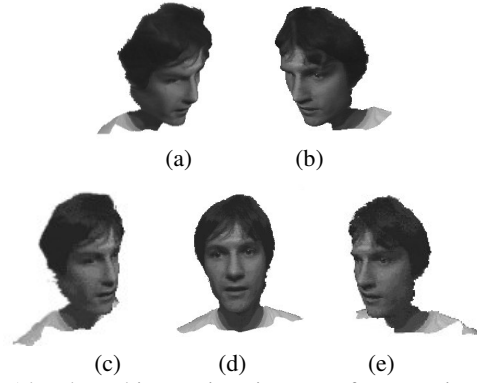

Figure 14. The arbitrary-view images of man using texture mapping and 3D model rotating scheme. (a) Upper left view image, (b) Upper right view image, (c) Left view image, (d) Center view image, (e) Right view image.
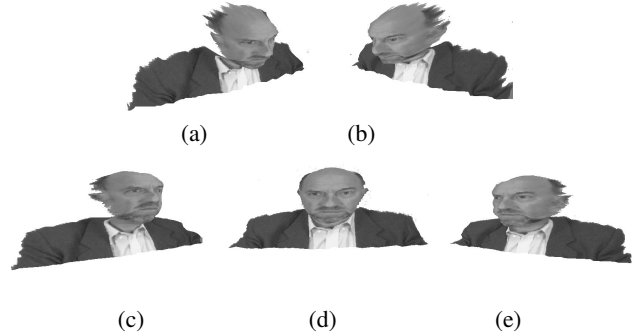

Figure 15. The arbitrary-view images of man using texture mapping and 3D model rotating scheme. (a) Upper left view image, (b)Upper right view image, (c) Left view image, (d) Center view image, (e) Right view image.

As the Figure 10(a) and Figure 11(a) indicate, since matching information doesn't exist in background, mismatching (noise) phenomenon is dominant. Also mismatching result occurs in projective distortion region and occluded region. The PBIL(Population Based Incremental Learning) (6) contains neighborhood considering characteristics, noise phenomenon is reduced. But mismatching phenomenon is much increased in the projective distortion region and occluded region(Figure 10(b) and Figure 11(b)). Figure 10(d) and Figure11(d) are the final result of proposed stereo matching algorithm. The result shows that considerable improvements are obtained especially in projective distortion region. And other mismatching problems, i.e. noise, luminance, boundary problem, is reduced.

In order to demonstrate performance of algorithm suggested, we generate the texture mapping images (Figure 14 and 15). These different view images are generated by 3D model rotating scheme in the different viewing angle. Figure 12 and Figure 13 show that proposed algorithm is superior to conventional algorithm, i.e. reduced noise, increase of matching reliability in the face (especially nose, cheek, and hair area).

5. CONCLUSION

A new stereo matching approach using probability-based window warping algorithm was presented to improve conventional stereo matching method. Since the projective

distortion regions can't be estimated with any fixed-size block-matching algorithm, the window warping technique is used in block matching process. The position of a reference window to warping is adaptively obtained according to the degree of reliability calculated by estimates disparities in the previously. In order to reduce noise sensitivity and reach higher efficiency simultaneously, we used a modified hierarchical block-matching algorithm.

The experimental results show that considerable improvements are obtained especially in projective distortion region, and texture-mapping result demonstrates the enhanced performance of proposed algorithm.

References

G. Wolberg, Digital Image Warping, IEEE, pp. 153-160, 1994.

G. Medioni and R. Nevatia, Segment-Based Stereo Matching, Computer Vision, Graphics, and Image Processing, vol. 31, no. 1, pp. 2-18, July, 1985.

K. P. Han, T. M. Bae, and Y. H. Ha, A Compact Stereo Matching Algorithm Using Modified Population-Based Incremental Learning, The Journal of The Institute of Electronics Engineers of Korea, vol. 36, no. 10, pp. 103-112, Oct. 1999.

P. N. Belhumeur and D. Mumford, A Bayesian Treatment of the Stereo Correspondence Problem Using Half-occlusion Regions, IEEE Conf. on Computer Vision and Pattern Recognition, pp.506-512, 1992.

T. Kanade and M. Okutomi, A Stereo Matching Algorithm with an Adaptive Window, IEEE Trans. on Patt. Anal. Machine Intell., vol. 16, no. 9, pp. 920-932, Sep. 1994.

Z. F. Wang and N. Ohnish, Intensity-Based Stereo Vision: from 3-D to 3-D, SPIE, vol. 2354, pp. 434-443, Nov. 1994.