Stereo Matching Using an Adaptive Window Warping

Jae-Chul Kim, Ho-Keun Lee, Sun-Kyu Kwon, and Yeong-Ho Ha
School of Electronic and Electrical Eng., Kyungpook National University
1370 Sankyuk-Dong, Buk-Gu, Taegu 702-701, Korea
Phone: +82-53-950-5535, Fax: +82-53-957-1194
E-mail address: yha@ee.knu.ac.kr

Abstract: In this paper, we propose a window warping method to solve stereo matching problems in projective distortion region. Because the projective distortion region can't be estimated with fixed-size block matching algorithm, we use the window warping technique in block matching process. The position of a reference window to resample is obtained adaptively according to the degree of reliability in disparity estimated previously. The initial disparity and reliability are obtained by applying a well known hierarchical strategy. The experimental result shows that considerable improvement is obtained in the projective distortion region.

Keywords: *Stereo matching, hierarchical BMA, projective distortion, matching constraints, window warping.*

1. Introduction

Stereo vision is useful for obtaining 3D depth information with matching two images taken from different view points[1,2]. The three-dimensional structure of object can be recovered from image pairs. The depth is calculated by the parameters of camera and the disparity value of images[2]. If the two-dimensional search space is reduced to an one-dimensional one by epipolar constraint, image matching can be tremendously simplified. The process to find the corresponding point is called as stereo matching. In Fig. 1, *L*, *f*, *dl*, and *dr* are stereo baseline, the focal length of camera, and distances from the center point of left and right images, respectively.

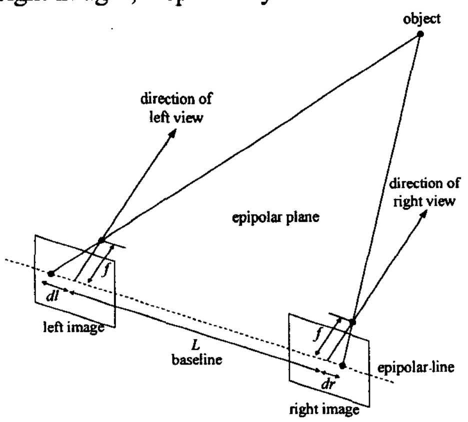


Fig. 1. Stereo geometry with parallel axes.

Stereo images are obtained from the different perspective position. So each image has the effect of projective distortion in Fig. 2. If the surface of real object(rl) is projected on the right and left cameras, each projected-image has different view characteristic(**dr>dl**).

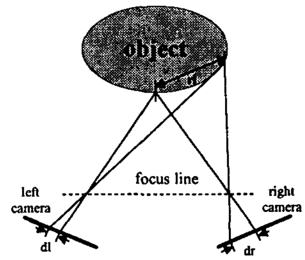


Fig. 2. Projective distortion.

Recently, the effect of projective distortion has received much attention and has been studied by some researchers[3-5]. Kanade presented an adaptive window method to reduce the effect of projective distortion[3]. His method employs a statistical model of the disparity distribution within an 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 Ohishi presented a 3D-to-3D method to slove the stereo matching problem. Instead of using a statistical model of the disparity distribution, the method uses a geometric model of the imaging process[4]. A deformable template-based matching method is used to recover the depth information from the projective distortion information. Neither of above two studies considered window warping about projective distortion. Difference of projected-images causes a very serious problem in stereo matching. This problem can not be overcome by using fixed-size block matching[5]. The proposed algorithm solves the problem based on window warping technique.

The paper is organized as follows: Section 2 presents the basic conception of window warping. Section 3 describes the proposed disparity estimation using an adaptive window warping. In section 4, we show the example of experimental result. In section 5, we conclude the paper and give some remarks

2. Window Warping Algorithm

2.1 Projective Distortion

Even in the points that are visible in all the views, the surface normal is often very tilted with respect to the cameras' optical axes. This causes stereo images to have serious *perspective distortion* between different views of the same point, as Fig.3 shows, resulting in a significant difference in the luminance profile of stereo homologous patches.

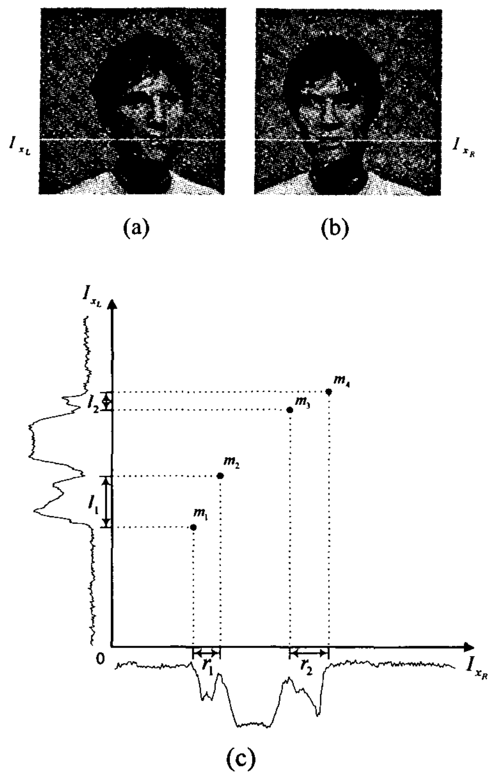


Fig. 3. The original stereo images of 'man' and corresponding projective distortion result. (a) Line profile on left image. (b) Line profile on right image. (c) Correspondence of line profiles.

Fig. 3 (c) shows that the length of l_1 is not equal to that of r_1. This result is the same as l_2 and r_2. By a fixed-size *BMA*, therefore, we cannot search for correct disparity values. Here, we need to resample the local mask in a matching process.

2.2 Fant's Resampling Algorithm

We use the Fant's algorithm so as to resample a local mask in *hierarchical BMA process.*

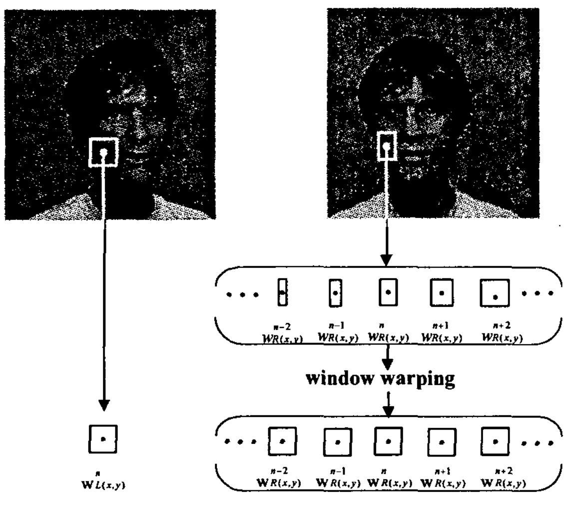


Fig. 4. The window warping using Fant's algorithm.

The algorithm treats the input and output as streams of pixels that are consumed and generated at rates determined by the spatial mapping. It is assumed that the input is mapped onto the output along a single direction, i.e, with no folds. As each input pixel arrives, it is weighted by its partial contribution to the current output pixel and integrated into an accumulator. The central benefit of this separable algorithm is the reduction in complexity of 1-D resampling algorithm. We define the fittest window($\mathbf{w}_\varphi$) in the candidate windows:

$$\mathrm{w}_\varphi = \arg\min_{k\in\Omega}\left|\mathrm{w}^{n}_{L(x,y)} - \mathrm{w}^{n+k}_{R(x,y)}\right|, \tag{1}$$

where Ω is window warping size, w is a local window in a block matching. We resample a local mask to the horizontal direction because we assumed epipolar geometry on input image.

3. Proposed Stereo Matching with an Adaptive Window Warping Algorithm

We propose a window re-sampling method in stereo matching in Fig. 5. The proposed matching algorithm consists of four steps. In the first step, hierarchical block matching[6] 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 second step, we obtain the reliability-map with two disparity-maps (left-to-right and right-to-left). Reliability is composed of bi-directional consistency check and local disparity variance. The third step is the block matching process. The re-sampled mask is used in a low reliable region and the normal mask is used in a high reliable region. Here, Fant's algorithm is used as a re-sampling procedure. It is faster than conventional re-sampling technique that uses low pass filter and subsampling.

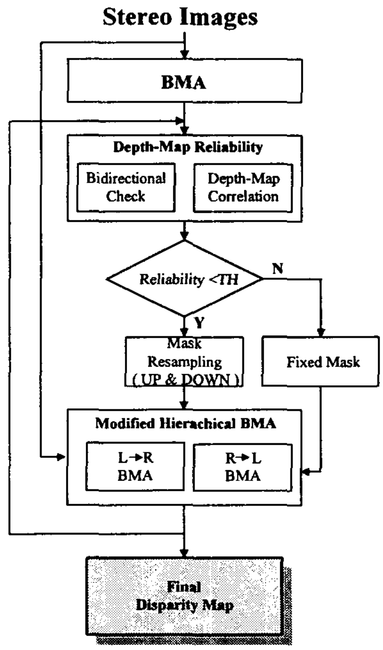


Fig. 5. The block diagram of the proposed algorithm.

3.1 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, i.e.,

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

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).

3.2 Reliability Generation

In order to estimate reliability of matching points, we use matching constraints: *uniqueness, smoothness*.

Though other constrains are not considered, only by using these constraints, the characteristic of stereovision can be reflected. Thus a pixel which doesn't satisfy these constraints has low reliability. Contrarily, a pixel which satisfies these constraints, has high reliability.

3.2.1 Uniqueness Constraint

A given pixel from one image can match no more than one pixel from the other image. However, bidirectional check shows that each disparity value has small difference in Fig. 6. And the bidirectional check distance is defined as

$$\delta = \left| D_L(x,y) - D_L(x + D_R(x,y)) \right|. \tag{3}$$

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

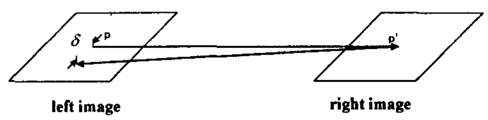


Fig. 6. The bidirectional check.

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:

$$f_{R1} = \frac{T_{bc} - \delta}{T_{bc}}, \quad \begin{cases} T_{bc} : \text{threshold} \\ \delta : \text{bidirectional check distance} \end{cases} \tag{4}$$

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

3.2.2 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 (eq. 3). Each value is extracted from *weight-function*:

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

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

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

where $$N_V = \sum_{i=-\frac{w}{2}}^{\frac{w}{2}} \sum_{j=-\frac{w}{2}}^{\frac{w}{2}} w_s(i+x, j+y). \tag{6}$$

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 $\sigma_D{}^2$ is small, whereas f_{R2} will take small value in opposite case.

3.2.3 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 $\lambda_1 + \lambda_2 = 1$.

3.3 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 in Fig. 7. Hierarchical block matching algorithm used in this paper is composed four steps as follows.

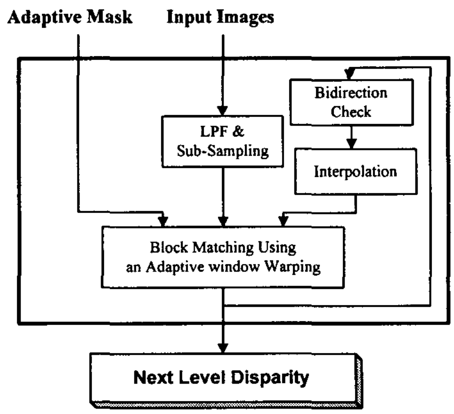


Fig.7. Modified hierarchical block matching algorithm.

Step 1 : *Low pass filtering and sub-sampling*

- Input images pass a LPF and are subsampled.

Step 2 : *Block Matching Process*

- Half-size 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.

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 D^i and S^i as:

$$D^i(x,y,d) = \arg\min_{d\in(x,y)} \left[\frac{1}{mn} \sum_{y'=-\frac{w_\varphi}{2}}^{\frac{w_\varphi}{2}} \sum_{x'=-\frac{w_\varphi}{2}}^{\frac{w_\varphi}{2}} \left| I^i{}_R(x+x', y+y') - I^i{}_L(x+d+x', y+y') \right| \right] \tag{9}$$

$$S^i(x,y,d) = \frac{1}{9} \sum_{y^i=-1}^{1} \sum_{x^i=-1}^{1} D^{i-1}(x^i, y^i), \tag{10}$$

and we can compose our cost function as

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

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

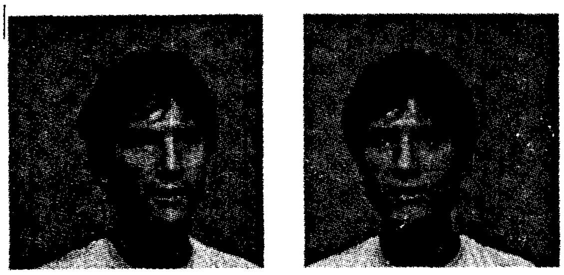
4. Experiment

In experiment we used the 'man' stereo images and the parameters were shown in table 1.

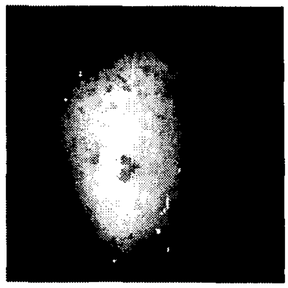
Table 1. Parameters for correspondence estimation

Window Size(w)	T_{bc}	λ_1	λ_2	γ_1	γ_2
7x5	3	0.5	0.5	0.4	0.6

The simulation result indicates that proposed algorithm shows outstanding ability in low texture and projective distorted region such as the nose and cheek[Fig. 8].



(a) (b)



(c)

Fig. 8. The original stereo images of 'man' and disparity map obtained by using proposed algorithm. (a) Left image. (b) Right image. (c) Disparity map.

5. Conclusion

In this paper, we have presented a stereo matching algorithm using an adaptive window warping. The algorithm selects a window adaptively for each pixel so that it produces the disparity which has the least uncertainty. The first key idea for the algorithm is that it employs a probability model that represents uncertainty of disparity of points over the window. And the second key idea is that it employs the window warping technique in matching process. The experimental result has demonstrated a clear advantage of this algorithm.

References

[1] M. W Matlin, *Sensation and perception*, pp. 184-229, Allyn and Bacon, 1988.

[2] D. Marr and T. Poggio, "Computational Theory of Human Stereo Vision," *Proc. of Roy. Soc. London B*, vol. 204, pp. 301-308, 1979.

[3] Kanade and 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.

[4] Z. F. Wang and N. Ohnish, "Intensity-based Stereo Vision: from 3-D to 3-D," *SPIE*, vol. 2354, pp. 434-443, Nov. 1994.

[5] 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.