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

KEY WORDS: Vision Sciences, Matching, Pattern, Recognition, Detection

ABSTRACT:
In this paper, we propose an adaptive stereo matching algorithm to encompass stereo matching problems in projective distortion region. Since the projective distortion region cannot 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 (Figure 1), 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 Figure 2 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).

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, \( I_R \) and \( I_L \) (270th line profiles) are a line profile of right and left image. These give a good account of projective distortion phenomenon in ‘man’ stereo images. Points \( m_1 \), and
$m_2$ represent matching end point between $r_2$ and $l_2$, on the other hand $m_3$ and $m_4$ are matching end point between $r_2$ and $l_2$.

![Figure 3](image1.png)

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 window warping technique is necessary needed in matching process.

2.2 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)

$$w_\theta = \arg\min_{w \in \Omega} |w_{L(x,y)} - w_{R(x,y)}^w|,$$

where $\Omega$ 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.
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 particular 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.

![Diagram of modified hierarchical block matching algorithm]

**Figure 6. Modified hierarchical block matching algorithm.**

**Step 1 : Low pass filtering and sub-sampling**
Input images are passed a LPF (low pass filtering) 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).

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

(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

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

(3)

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

![Diagram of the bidirectional check]

**Figure 7. The bidirectional check.**

Where \( \delta \) 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:

\[
f_M = \frac{T_{bc} - \delta}{T_{bc}}.
\]  

(4)

Where \( T_{bc} \) : threshold
\( \delta \) : bidirectional check distance

where \( f_M \) represents the uniqueness reliability, \( \delta \) is bidirectional 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:

\[
 w_i(x, y) = \begin{cases} 
 0 & , \delta > T_w \\
 1 & , \delta < T_w 
\end{cases} .
\]

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

\[
 \sigma'_{yi}(x, y) = \left[ \frac{1}{N_i} \sum_{i=1}^{N_i} \sum_{j=1}^{N_i} w_i(i+x, j+y) \cdot I(i+x, j+y) \right]^2 \\
 - \left[ \frac{1}{N_i} \sum_{i=1}^{N_i} \sum_{j=1}^{N_i} w_i(i+x, j+y) \cdot I(i+x, j+y) \right]^2 ,
\]

where

\[
 N_i = \sum_{i=1}^{N_i} \sum_{j=1}^{N_i} w_i(i+x, j+y) .
\]

Using these conventions, we define the smoothness reliability:

\[
 f_{R2} = e^{-\xi \sigma'_{yi}} ,
\]

where \( \xi \) is an damping parameter to control the decreasing velocity of \( f_{R2} \). The function value of \( f_{R2} \) is large if \( \sigma'_{yi} \) 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 \( f_{R2} \):

\[
 f_{total} = \lambda_1 f_{R1} + \lambda_2 f_{R2} .
\]

where \( \lambda_1 \) and \( \lambda_2 \) 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:

\[
 S'(x, y, d) = \frac{1}{9} \sum_{y'=1}^{9} \sum_{x'=1}^{9} D^{-1}(x', y') - D' ,
\]

and we can compose our cost function as

\[
 E' = g_1 D' + g_2 S' ,
\]

where \( D_i \) denotes the cost function at level of the ith pyramid, the other parameters are the same as Equation(2), \( g_1 \) and \( g_2 \) are weight coefficients satisfying the relation \( g_1 + g_2 = 1 \).

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.
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 Figure 11(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


