A HYBRID STEREO ALGORITHM WITH 3D DISPARITY SPACE FOR RECONSTRUCTION OF SURFACE MORPHOLOGY WITH DEPTH DISCONTINUITY

Gyung-Bum Kim*

*Applied Opto-Mechatronics Lab., Department of Mechanical Design Engineering
Chungju National University, Chungju, Chungbuk, 380-702, Korea

Abstract: A critical issue in area-based stereo matching lies in selecting a fixed rectangular window size. Previous stereo methods doesn't deal effectively with occluding boundary due to inevitable window-based problems, and so give inaccurate and noisy matching results in areas with steep disparity variations. In this paper, a hybrid stereo vision approach is presented to estimate rapid, accurate, detailed and smooth disparities in depth discontinuity. It makes the smoothing of depth discontinuity reduced by evaluating corresponding correlation values and intensity gradient-based similarity in the three-dimensional disparity space. In addition, it investigates maximum connected match candidate points and then devises the novel variable window to treat with general surface morphologies. We show how our results improve on those of closely related techniques for accuracy, robustness, matching density and computing speed

Keywords: Correlation values, Disparity, Intensity gradient-based similarity, Match candidate points, Projective distortion, Three-dimensional(3D) disparity space

1. Introduction

For machine stereo vision available for 3D reconstruction, the appropriate window size selection is critical to obtain smooth, detailed and accurate disparity map in area-based stereo methods. These stereo methods have used fixed windows experimentally selected, and then the contradiction of disparity maps are followed according to selecting window size.

In this paper, a hybrid stereo algorithm is tried to estimate a rapid, detailed and accurate disparity in depth discontinuity. For match candidate selection, the incorporation of both a new correlation function and intensity-gradient based similarity makes smoothing around depth discontinuity decreased. The connectivity of match candidate points in 3D disparity space is investigated to improve robustness to noise. A new variable window[1] is devised to cope with general surfaces as well. Sub-pixel disparity is accurately estimated through formulating the error function in the variable window.

The proposed hybrid stereo method has obtained a detailed, accurate and robust disparity, and also has relatively quick disparity estimation speed. Experiments show that the proposed algorithm easy to effectively implement outperforms previous ones.

2. Disparity distribution models

All stereo algorithms have been so far developed to examine how much support they receive from their local neighborhood. Prazdny[2] incorporates the function to specify the support amount between neighboring points based on disparities into stereo algorithm. Three requirements for the function are set. Kanade's model[2] is exactly the same as that of Pradzny in that the disparity difference between two pixels follows a zero-mean gaussian distribution whose variance increases with their distance apart. The only difference between Pradzny's model and Kanade's one is that the variance is proportional to the square of the distance between the pixels, instead of the distance itself.

Analyzing above disparity distribution models, physical meanings that small match areas are straight planar and parallel to baseline, are implied. However, the larger matching areas, more distorted planes or surfaces are expected. In addition, match error is seriously generated in depth discontinuity, due to projective distortion resulting from different viewpoints. Thus, Arbitrary match support areas to deal with various surface morphologies are needed. In order to incorporate these ideas into our algorithm, we have devised a novel variable window Ψ and then proposed the disparity model shown by Eq. (1).

$$d(u,v) - d(0,0) = d_c \quad (u,v) \in \Psi \tag{1}$$

3. Hybrid stereo approach with 3D disparity space

3.1 Match candidate selections in 3D disparity space

The selection of match candidates is performed in the 3D disparity space composed of image row, column and disparity, as shown in Fig. 1. The 3D space voxels $\Gamma(u,v,d)$ have correlation values which mean the probabilities to be matched disparity d for the left image pixel.

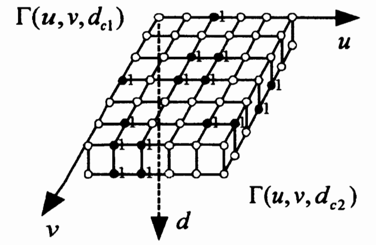


Fig. 1 Descriptions of 3D disparity space

This correlation values use the sum of absolute difference ε_{SAD} within fixed windows Ω, and are defined as Eq. (2). These correlation values have to have high match probabilities for correct matches. The lower ε_{SAD}, the correlation is high and then correlation values are close to 1.

$$\begin{aligned}
&\varepsilon_{SAD(u,v,d)} = \sum_{(u',v')\in\Omega} \left| f_L(u+u',v+v') - f_R(u+u'+d,v+v') \right| \\
&\varepsilon_{\max} = \max_d \varepsilon_{SAD(u,v,d)} \\
&\varepsilon_{\min} = \min_d \varepsilon_{SAD(u,v,d)} \\
&\text{where} \quad d_{\min} \le d \le d_{\max} \\
&\rho(u,v,d) = \frac{1}{1+\beta\dfrac{\varepsilon_{SAD(u,v,d)} - \varepsilon_{\min}}{\varepsilon_{\max} - \varepsilon_{\min}}}
\end{aligned} \tag{2}$$

In opposite case, the values are close to 0. In order to improve the consistency of computed correlation values, the voxels of the 3D space are updated by averaging the correlation values in certain support areas.

In this paper, the 3D support to cope with sloping and more general surfaces is devised in the 3D disparity space, and the criterion function for the choice of match candidate voxels is suggested in Eq. (3) using Eq. (2) and intensity-gradient based similarity S_G.

$$\begin{aligned}
&\Gamma(u,v,d) = \frac{1}{N_v} \sum_{(u',v',d')\in\Phi} \rho(u+u',v+v',d+d') \\
&S_G = \frac{|\nabla f_L(u,v)| + |\nabla f_R(u+d,v)|}{2} - |\nabla f_L(u,v) - \nabla f_R(u+d,v)| \\
&\mathit{if}\ \Gamma(u,v,d) > \Gamma_{Th}\ \mathit{and}\ S_G > 0 \\
&\quad \Gamma(u,v,d) = 1 \\
&\mathit{else} \\
&\quad \Gamma(u,v,d) = 0
\end{aligned} \tag{3}$$

This intensity gradient-based similarity represents the measure of the intensity variation and is computed with the sobel operator. Here, for simplicity, the 3D support area Φ is defined as a cube with a fixed row, column and disparity(depth), and N_v is the voxel number of 3D support. The positive S_G means the intensity changes of corresponding point are consistent. Here, we have tried to improve the reliability of the selection of match candidates by incorporating intensity changes into the criterion function, as well.

3.2 Connectivity of match candidates

Fig. 1 shows match candidate points are described with both black color and '1' in the 3D space, and they mean voxels corresponding to a specific disparity d_{c1}, d_{c2}. In the case of projecting the arbitrary point of 3D surface morphology to stereo images, we have surveyed that a same disparity near to an interested pixel is very likely to exist through the assumption of several stereo methods. To include this physical meaning, the connectivity of a voxel is investigated with labeling techniques, and then the 3D disparity space is renewed at every disparity plane. Here, connected components C of match candidates can be found at a voxel. The connected components mean the set of a voxel with the same disparity, and also the connected number of the match candidates in each component can be computed. The integer disparity in the updated 3D space is estimated by determining the plane with maximum connected candidates in the connected components. This can be expressed as follows.

$$\tilde{d} = \arg\max_{d}\ N_c(L) \text{ of } C(\Gamma(u,v,d)) \quad \text{where } N_c\text{:Number of connected label}(L) \tag{5}$$

3.3 Sub-Pixel Disparity Estimation

In the case of projecting an arbitrary point of 3D surface morphology to stereo images, there is ambiguity of ±1/2 pixel in an image coordinate. These problems make stereo matching get more difficult. Sub-pixel disparity estimation for identifying the ambiguity is performed with only the intensity of the variable window and then the intensity changes resulting from viewpoints is assumed not to be. The possible solution to the smoothing of depth discontinuity is to use non-uniform and variable support regions, that is the variable window, since the boundary blurring is caused by support regions that span object boundary.

This variable window is defined as Eq. (4) where Ω_v is to limit the variable window range.

$$\Psi = \left\{ w \mid w \in C\left(\Gamma(u,v,\tilde{d})\right) \cap \Omega_v \right\} \quad \text{where } \Omega_v\text{:Variable window mask} \tag{4}$$

The intensities of corresponding points in the variable window are given by Eq. (5) where $\delta\tilde{d}(w_i)$ is a correction value for sub-pixel disparity.

$$f_L(w_i) = f_R\left(w_i - \tilde{d}(w_i)\right) + n(w_i),\ w_i \in \Psi \quad \text{Where } \tilde{d}(w_i) = d_0(w_i) + \delta\tilde{d}(w_i) \tag{5}$$

The intensity function f_R is linearized at $w_i - d_0(w_i)$ and then high order terms are assumed to be negligible. Thus, as shown in Eq. (6), the correction value can be determined to minimize the error function E, which is defined as the sum of squared noise in the variable window.

The proposed hybrid method with the variable window makes match error reduced rather in a general surface morphology with both depth discontinuity and smooth areas. The rapid and accurate stereo method has been verified through experiments in next section.

$$E = \sum_{w_i \in \Psi} n^2(w_i),\quad \frac{\partial E}{\partial \delta\tilde{d}} = 0$$
$$\delta\tilde{d} = -\frac{\sum_{w_i\in\Psi} f_R'\left(w_i - d_0(w_i)\right)\left(f_L(w_i) - f_R\left(w_i - d_0(w_i)\right)\right)}{\sum_{w_i\in\Psi} f_R'^2\left(w_i - d_0(w_i)\right)} \tag{6}$$

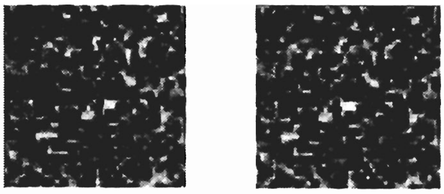
4. Experiments

Fig. 2(a) is synthetic stereo images which are constructed with the disparity range of 0 to 24, including depth discontinuity. An actual disparity map of the stereo images is described in Fig. 2(b). The larger the disparity, the brighter the disparity map is represented. Fig. 3 shows disparity maps estimated by area-based(SSD, NCC), pixel-based[3] and the proposed method. A window Ω was applied as 7 in area-based methods, and several parameters $\Omega, M_{Th}, \beta, \Phi$ were inputed as 3, 0.7, 9, 3 in the proposed method, respectively. Large matching errors occur due to the smoothing of depth discontinuity resulting from rectangular window based problems in area-based methods. Meanwhile, no the proposed method gives more detailed discontinuity than pixel-based one, but it represents more accurate disparity than area-based techniques.

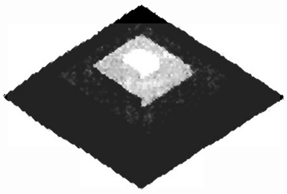
Table 1 is to evaluate estimated disparity maps for RMS(root-mean-squared) error, BMR(bad match ratio), NMR(no match ratio), computation time. Equations of these items are in Eq. (7).

SSD makes RMS error minimized according as the window size is larger, but we confirm via BMR that match errors instead get bigger, due to projective distortion. Also, computation time gets rapidly increased. The results of NCC are similar to those of SSD as well. MLMHV gives excellent results of 34% compared with the RMS error of the proposed method.

We are aware that the pixel-based method, MLMHV gives detailed and sharp disparities in depth discontinuity through RMS and BMR values. However, one of the reasons that these values are low is not to include non-matched pixels, and so MLMHV has very large NMR. It has been expected that the bigger NMR can be generated in stereo images with noise and severe intensity variations and it takes long time to reconstruct a surface morphology with the stereo images.

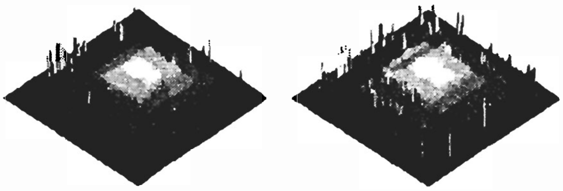


(a) Stereo images

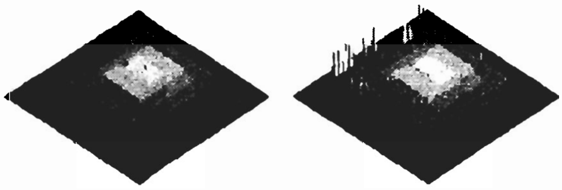


(b) A true disparity map with depth discontinuity

Fig. 2 Synthetic stereo images



(a) SSD (b) NCC



(c) MLMHV (d) Proposed

Fig. 3 Computed disparity maps

$$\mathrm{RMS} = \left(\frac{1}{N_m} \sum_{(u,v)} \left| d_T(u,v) - \tilde{d}(u,v) \right|^2 \right)^{\frac{1}{2}}$$

$$\mathrm{BMR} = \frac{100}{N_m} \sum_{(u,v)} \left(\left| d_T(u,v) - \tilde{d}(u,v) \right| > \Delta_d \right)$$

$$\mathrm{NMR} = \frac{100(N_p - N_m)}{N_p} \tag{7}$$

where d_T : True disparity

N_m : Total number of matched pixels

Δ_d : Disparity error tolerance

N_p : Total number of pixels

The proposed method reduces 12%, 13%, 72% in terms of RMS error, BMR and computation time, compared with area-based methods, respectively and 82% for NMR compared with MLMHV as well. Finally, we confirm that the proposed method makes up for the weak points of two stereo techniques.

Table 1 Evaluations of stereo methods for Fig. 3

Method (Ω)	RMS (pixel)	BMR (%)	NMR (%)	Time (sec)
SSD(3)	2.11	23.88	0.31	3.05
SSD(11)	1.87	28.11	0.36	40.25
SSD(15)	2.01	30.12	0.24	74.67
NCC(3)	2.32	24.35	0.60	3.58
NCC(11)	2.02	28.32	0.45	42.09
NCC(15)	2.11	30.31	0.14	77.65
MLMHV	1.21	16.95	6.93	1.05
Proposed	1.86	20.74	1.21	0.83

5. Conclusions

A hybrid stereo algorithm with 3D disparity space has been proposed to give a detailed, sharp and smooth disparity in a surface morphology including depth discontinuity. In 3D disparity space, a correlation value and intensity-gradient based similarity are defined, and then smoothing in depth discontinuity is reduced with them. Maximum connected match candidates in the 3D disparity space are investigated to deal effectively with general surface morphology.

The performance of the proposed method has been evaluated and verified with several items. The proposed algorithm gives more excellent data for each item and is easy to implement. In addition, it has been regarded the hybrid stereo algorithm could be widely applied to machine vision system required three-dimensional information

Acknowledgement

The work was supported by research funding 2003 provided by Chungju National University and Korea Science and Engineering Foundation Grant.(R05-2004-000-10512-0)

References

[1] Kim, G. B. and Chung, S. C, 2004, An Accurate and Robust Stereo Matching Algorithm with Variable Windows for 3D Measurements, *Mechatronics*, 14, 715-735.

[2] Kanade, T. and Okutomi, M., 1994, A Stereo Matching Algorithm with an Adaptive Window:Theory and Experiment, *IEEE Trans. Pattern Anal. Machine Intell.* 16(9), 920-932.

[3] Cox, I. J., Hingornni, S. L. and Rao, S. B., 1996, Maximum Likelihood Stereo Algorithm, *Computer Vision and Image Understanding* 63(3), 542-567.