We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Using Optical Flow as an Additional Constraint for Solving the Correspondence Problem in Binocular Stereopsis

Yeon-Ho Kim[1] and Soo-Yeong Yi[2]
[1]Purdue University
[2]Seoul National University of Technology
[1]USA
[2]Korea

1. Introduction

A stereo matching algorithm for 3D structure reconstruction relies on the correlation of image features in a pair of images. Usually, two calibrated cameras are used to capture scenes, and the left and right images are rectified to reduce the search space from 2D image region to 1D line. Then, the disparity is defined by the horizontal difference between the corresponding points on the search line. To solve the correspondence problem, various image features such as image intensity, edge, color, infrared light, pixel motion, etc. are employed either separately or integrated together. Among these image features, pixel motion or optical flow has been rarely used as one of the matching features (A. Scheuing & H. Niemann, 1986, G. Sudhir et al, 1995, A. M. Waxman & J. H. Duncan, 1986). In this chapter, we focus on the use of optical flow as an additional constraint in solving the correspondence problem.

The stereo matching algorithm for 3D structure reconstruction can be divided into two groups based on the matching primitives: intensity based matching (also referred to as area based matching) and feature based matching. The intensity based matching method searches the best matching points using only intensity values of pixels. The method can further be divided into two groups depending on the smoothness constraints in minimization of matching cost function: local (window-based) method and global method. More details on the intensity based method can be found in a survey by Scharstein (D. Scharstein & R. Szeliski. 2002, M. Z. Brown et al, 2003). The intensity based matching method produces a dense disparity map without any additional post-processing, but usually needs an exhaustive search. Since the intensity based method uses the intensity value at each pixel directly, this method may suffer from the varying illumination problem.

The feature based matching method searches the best matching points using some special symbolic feature points, such as line, contour, corner, etc. Since this method calculates disparity values only on the pixels corresponding to the feature point, this method does not require an exhaustive search. Also, symbolic features are less sensitive to illumination changes than pixel intensity, so the calculated disparity is more reliable than that of the intensity based matching method in the case of varying illumination. However, this method

requires an additional process for extracting image features, and the produced disparity map is not as dense compared to that produced by the intensity based matching method unless a surface-fitting step is applied. More details on the feature based matching method can be found in a survey by Dhond (U. R. Dhond & J. K. Aggarwal, 1989).

In this chapter, we propose to use optical flow as an additional constraint in solving stereo correspondence problem in both intensity and feature based stereo matching methods. For the intensity based matching method, optical flow is used to reduce mismatching, which is described in Section 2.1. For the feature based matching method, optical flow simplifies the matching procedure as described in Section 2.2. In Section 3, we present some preliminary results on the surface reconstruction of a human hand in stereo image sequences using our proposed methods.

2. Modification of two matching techniques

2.1 Modification of an intensity based matching technique

In this section, we restrict the intensity based method to the local (window-based) method. The local intensity based matching technique calculates the disparity based on two geometric constraints. The first is the epipolar constraint that reduces two-dimensional search space to one-dimensional one. The second constraint is the assumption that the disparity is constant in small image regions (O. Faugeras et al, 1993). Based on these two constraints, small image regions are matched between two images along the epipolar line. If left and right images are rectified, then the vertical position of the search line in the right image is the same as that of the corresponding point in the left image. Then the disparity is defined as the horizontal distance between the corresponding points. To find the best matching pixel, various matching criteria are used. One of the most basic criteria is the sum of squared distances (SSD) defined as:

$$SSD_{Intensity}(x,y,d) = \sum_{i,j}[I_L(x+i,y+j) - I_R(x+i+d,y+j)]^2, \qquad (1)$$

where I_L and I_R are left and right images, (x,y) is the pixel coordinates where the disparity d is calculated, and the indices i and j are used to cover the entire pixels within a rectangular window. The best matching pixel has the minimum SSD value. This method is very simple but requires an assumption that the illumination conditions are same for the left and right images. But in real stereo images, due to the different setting of each camera or relative distance from each camera to the light sources, this assumption is not always satisfied. To ameliorate this problem, we use another matching criterion that is more robust to the variance of the illumination. This criterion is referred to as the normalized sum of squared distance (NSSD)(O. Faugeras et al, 1993) and defined as follows:

$$NSSD_{Intensity}(x,y,d) =$$

$$\frac{\sum_{i,j}[(I_L(x+i,y+j) - \overline{I_L(x,y)}) - (I_R(x+i+d,y+j) - \overline{I_R(x+d,y)})]^2}{\sqrt{\sum_{i,j}(I_L(x+i,y+j) - \overline{I_L(x,y)})^2} \times \sqrt{\sum_{i,j}(I_R(x+i+d,u+j) - \overline{I_R(x+d,y)})^2}}, \qquad (2)$$

where $\overline{I_L(x,y)}$ is the mean of $I_L(x+i,y+j)$ within the window in the left image and $\overline{I_R(x+d,y)}$ is the mean of $I_R(x+i+d,y+j)$ within the window in the right image.

In addition to the intensity based matching criterion, we use another matching criterion that uses optical flow. For each left and right image sequence, the optical flow is calculated using our robust motion estimation technique presented in (Y.-H. Kim et al, 2005). The optical flow corresponding to a pixel in the left image is matched to that corresponding to a pixel in the right image along the epipolar line using the following motion correspondence matching criterion in a similar way of the intensity based matching technique described above:

$$NSSD_{Motion}(x,y,d) =$$
$$\frac{\sum_{i,j}[(I_L(x+i,y+j)-\overline{I_L(x,y)})-(I_R(x+i+d,y+j)-\overline{I_R(x+d,y)})]^2}{\sqrt{\sum_{i,j}(M_L(x+i,y+j)-\overline{M_L(x,y)})^2} \times \sqrt{\sum_{i,j}(M_R(x+i+d,u+j)-\overline{M_R(x+d,y)})^2}},$$
(3)

where $M(x,y)=u^2(x,y)+v^2(x,y)$ is the squared magnitude of optical flow vector which consists of the horizontal (u) and vertical (v) components and other notations are same as defined in Eqs. (1) and (2).

Finally, we integrate the intensity based matching criterion and the motion based matching criterion in a single matching criterion as follows:

$$NSSD_{Integrated}(x,y,d) = NSSD_{Intensity}(x,y,d) \times (1 + NSSD_{Motion}(x,y,d)),$$
(4)

2.2 Modification of a feature based matching algorithm

In this section, we first describe the Marr-Poggio-Grimson (MPG) algorithm that is one of the most popular feature based matching methods. Then explain how we modify this algorithm using optical flow.

The MPG algorithm, implemented by Grimson (W. E. L. Grimson, 1981), solves the stereo correspondence problem based on the Marr and Poggio's computational model of human stereopsis (D. Marr & T. Poggio. 1979). The MPG algorithm consists of following three main steps: feature extraction, feature matching, and matching analysis.

The first step of the MPG algorithm is to extract features in images. For this purpose, the Laplacian of Gaussian $(\nabla^2 G)$ operator (referred to as the primal sketch operator) is first applied to the image where the operator is formed by:

$$\nabla^2 G(x,y) = \frac{r^2 - 2\sigma^2}{\sigma^4} \exp(-r^2/(2\sigma^2)),$$
(5)

where (x,y) is the pixel coordinate and $r^2 = x^2 + y^2$. The width of the function is represented by the distance between the first zeros on either side of the origin and denoted by $w_{2D} = 2\sqrt{2}\sigma$. From the image filtered by this operator, zero crossing points are then detected as final image features.

The second step of the MPG algorithm is feature matching. Based on the Marr and Poggio's hierarchical model of human stereopsis, the MPG algorithm employs a coarse-to-fine approach for feature matching. The disparity values obtained at a coarse level of images are used for the disparity calculation at the next finer level of images. At each level, for each zero-crossing point in the left image, its matching candidates in the right image are searched within the scan line where the center of the line is the coordinates of the zero-crossing point

in the left image and the width of the line is $2w_{2D}$. Within this search line, the points with similar local orientations are selected as its matching candidates. Instead of using the local orientation at zero-crossings as originally proposed by Grimson (W. E. L. Grimson, 1981), we use binary neighborhood comparison -- implemented by Tanaka and Kak (S. Tanaka & A. C. Kak, 1990) - as the candidate selection criterion in our implementation.

The third step of the MPG algorithm is matching analysis. If just only one matching is found in the scan line, then the distance between the corresponding points will be the final disparity. However, within the search line, more than one matching candidates can be found and, in that case, further consideration is required to determine the final disparity. For this purpose, Marr and Poggio used two constraints. One is referred as to *uniqueness*: each zero-crossing point in the left image must have only one disparity value. The other one is referred as to *continuity*: disparity value must vary smoothly. To implement these two constraints of Marr and Poggio's theory, Grimson grouped matching candidates by three types depending on the position of the matching candidates in the search line: convergent, divergent, or zero. If more than one matching candidates are found, then the one having the same type of matching as the dominant matching type in the neighbourhood is accepted.

In our modification of the MPG algorithm, we simplify the final matching analysis step by employing optical flow. First, we calculate squared magnitude of the optical flow vectors in the left image and the right image separately. Then in the multiple matching candidates, we choose the one having the minimum difference of the squared magnitude of the optical flow vectors between the corresponding points.

3. Results

In this section, we present some preliminary results of the 3D structure reconstruction using our modified matching algorithms described in the previous section. Fig. 1 shows left and right image sequences we used here. These image sequences include a human hand that is moving at the front of the left and right cameras. To enrich the texture of the hand, we used a glove covered by random-dot pattern.

For each image sequence, optical flow is obtained frame by frame using our robust optical flow estimation method (Y.-H. Kim et al, 2005). Generally, differential optical flow techniques based on the image derivatives are able to estimate only the small motion. The small optical flow does not provide significant differences between neighbour pixels and is not adequate for the matching criterion. To ameliorate this problem, we accumulate optical flow vectors calculated from several image frames to get larger optical flow. Fig. 2 shows the squared magnitude of the accumulated optical flow vectors in three-dimensional space. The width and depth of the three-dimensional space represent the axes of the horizontal and vertical direction of the image plane, and the height of the three dimensional space represents the axis of the squared magnitude of the optical flow vectors. Figs. 2(a, b) show the squared magnitude of the optical flow vectors from the 1st image frames to the 2nd image frames for each left and right image sequence respectively. Figs. 2(c, d) show the squared magnitude of the accumulated optical flow vectors from the 1st image frames to the 3rd image frames for each left and right image sequence respectively. Figs. 2(e, f) show the squared magnitude of accumulated optical flow vectors from the 1st image frames to the 5th image frames for each left and right image sequence respectively. As shown in this figure, the squared magnitude of optical flow and the difference of these values between neighbour pixels increase as the number of image frames used for the optical flow calculation increases.
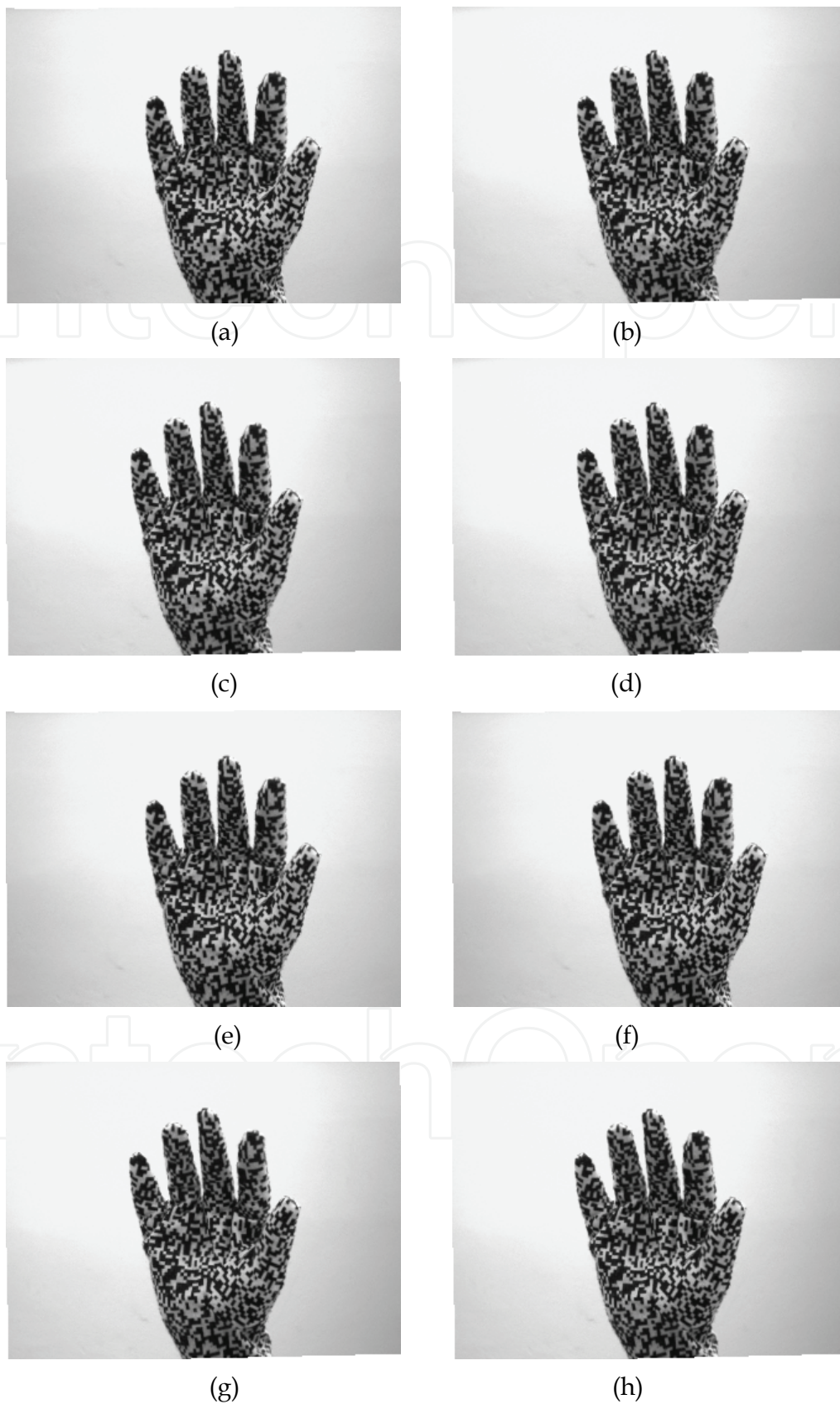
Fig. 1. Left and right image sequences: (a, b) The 1st image frames, (c, d) The 2nd image frames, (e, f) The 3rd image frames, and (g, h) The 5th image frames.
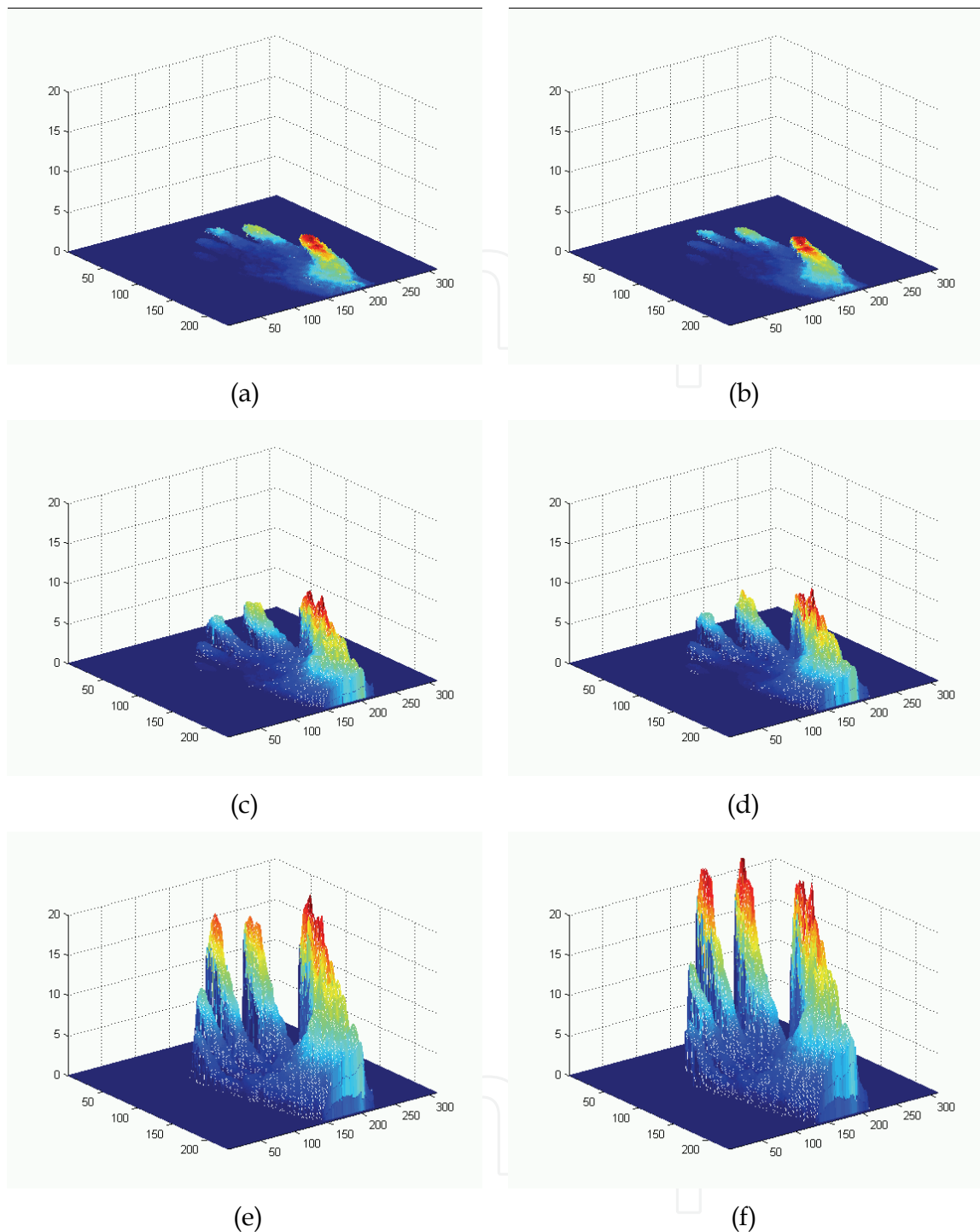
(a)   (b)

(c)   (d)

(e)   (f)

Fig. 2. Squared magnitude of (a, b) optical flow from the 1st image frames to the 2nd image frames,   (c, d) accumulated optical flow from the 1st image frames to the 3rd image frames, (e, f) accumulated optical flow from the 1st image frames to the 5th image frames for the left and the right image sequences respectively.

3.1 Results of a modified intensity based matching technique

Fig. 3(a) shows the disparities obtained from the original intensity based matching algorithm using the first image frames of the left and right image sequences. No motion

Using Optical Flow as an Additional Constraint for Solving
the Correspondence Problem in Binocular Stereopsis
341

information is used for this result. The width and the depth of the three dimensional space represent the axes of the vertical and horizontal direction of the image plane and the height of the three dimensional space represents the axis of the disparity. Fig. 3(b) shows the disparities obtained from the modified intensity based matching algorithm using the intensity matching criterion and additional motion matching criterion. The intensity matching criterion uses pixel intensities in the first image frames of the left and right image sequences. The motion matching criterion uses the optical flow from the 1st image frames to the 2nd image frames for each left and right image sequence respectively. For the integrated matching criterion, Eq. (3) is used. Fig. 3(c) shows the disparities obtained from the modified matching algorithm in which the intensity based matching criterion uses the first image frames of the left and right image sequences and the motion based matching criterion uses accumulated optical flow from the 1st image frames to the 5th image frames for each left and right image sequence respectively. We can see the number of mismatches decreases by using the motion matching criterion and also by increasing the number of image frames to calculate optical flow. The mismatches in Fig. 3(c) are separated as shown in Fig. 4(a) using our de-nosing technique (that counts the pixels in a certain cubic area around a pixel in interest and if the number of count is less than a threshold then the pixel in interest is discarded). Fig. 4(b) shows the finally reconstructed surface of hand using a surface fitting function (spaps) in MATLAB (C. de Boor, 2004).

## 3.2 Results of a modified feature based matching algorithm

Fig. 2 shows different levels of zero-crossing points for the test image sequences filtered by the Laplacian of Gaussian operator with different w_{2D} values. In Fig. 2, the value of w_{2D} for (a, b) is 32, for (c, d) 16, for (e, f) 8, and for (g, h) 4. Fig. 6(a) shows disparities obtained using the original feature based matching algorithm from the first image frames of the left and right image sequences. No motion information is used for this result. The width and the depth of the three dimensional space represent the vertical and horizontal axes of the image plane and the height of the three dimensional space represents the axis of the disparity. Fig. 6(b) shows the disparities obtained using our modified feature based matching algorithm. This algorithm uses zero-crossing features obtained from the first image frames of the left and right image sequences and optical flow obtained from the 1st image frame and the 2nd image frame for each left and right image sequence respectively. Although our modified feature based matching algorithm uses simpler matching procedure than the original feature based matching algorithm as described in Section 2.2, we can see no significant difference between the disparity results in Fig. 6(a) and those in (b). Fig. 6(c) shows the disparities obtained using our modified feature based matching algorithm using the accumulated optical flow from the 1st image frame to the 5th image frame for each left and right image sequence respectively. There is no significant difference between the result in Figs. 6(b) and that of (c). Fig. 7(a) shows the disparities after separating noisy disparities using the same de-nosing algorithm used in Section 3.1. The final reconstructed hand surface result is shown in Fig. 7(b).
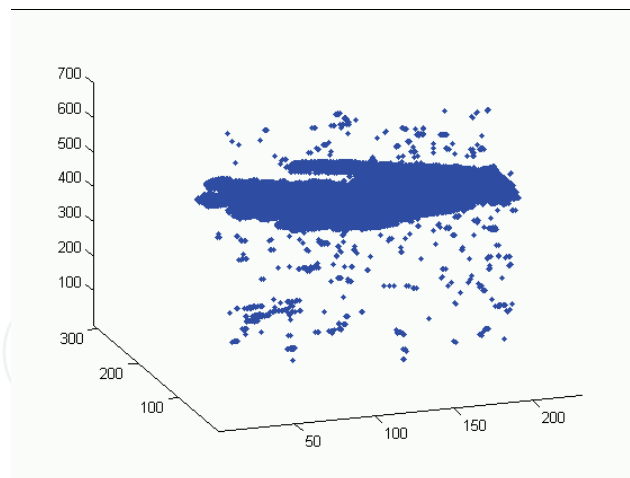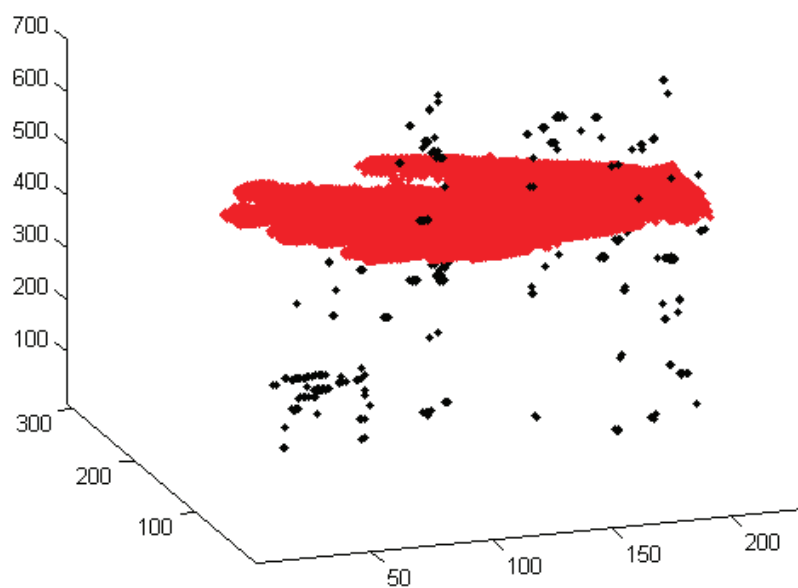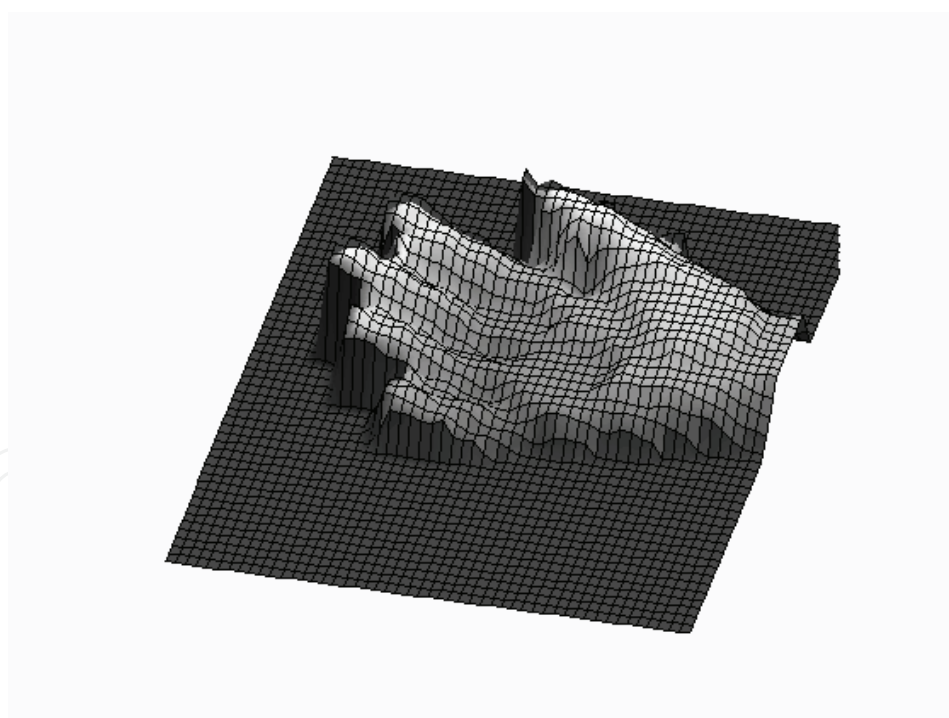
(a)
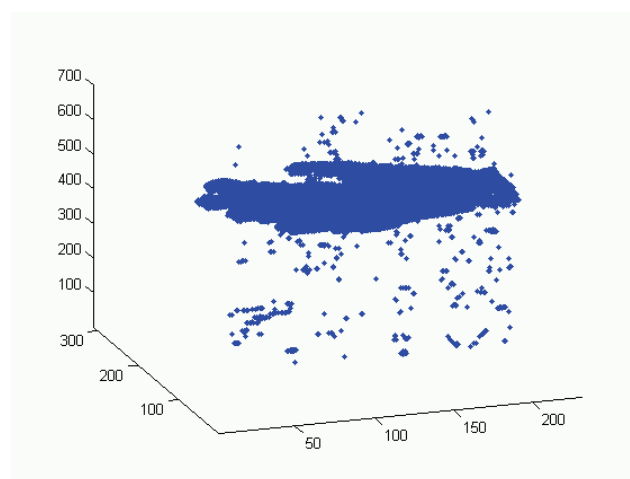
(b)
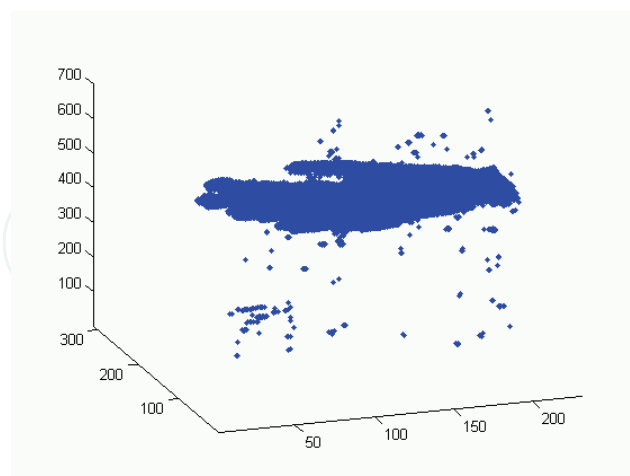
(c)

Fig. 3. Calculated disparities: (a) using the original intensity based stereo matching technique, (b) using the modified intensity based stereo matching technique with optical flow from the 1st to the 2nd image frames, and (c) using the modified intensity based stereo matching technique with accumulated optical flow from the 1st to the 5th image frames.

(a)

(b)

Fig. 4. (a) Noise separated disparities, (b) A human hand surface reconstruction result using our modified intensity based stereo matching method.
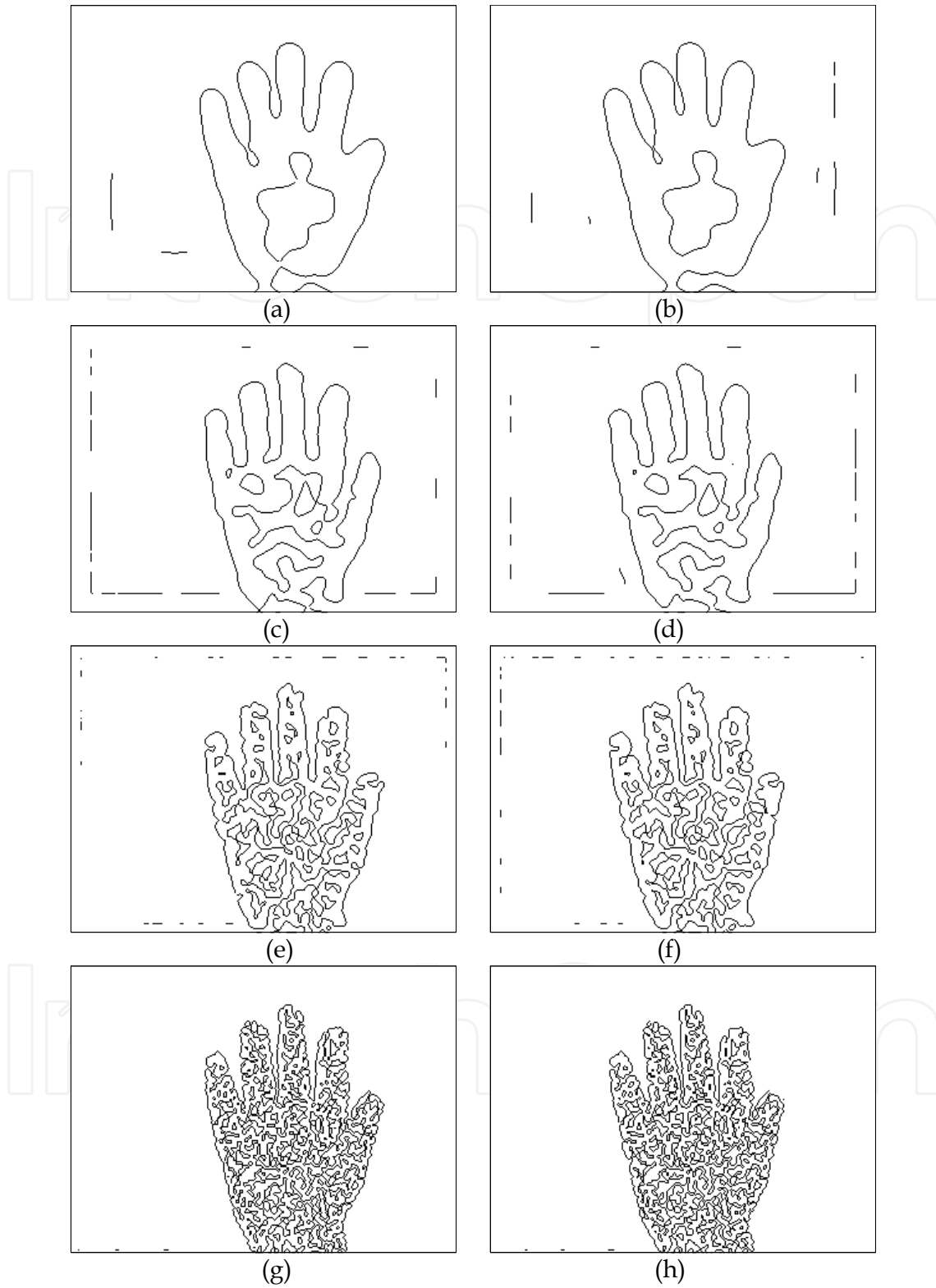
Fig. 5. (a) Zero-crossing feature points in the left and right images which are filtered by LoG operator with different w_{2D}: The value of w_{2D} is (a, b) 32, (c, d) 16, (e, f) 8, and (g, h) 4, respectively.
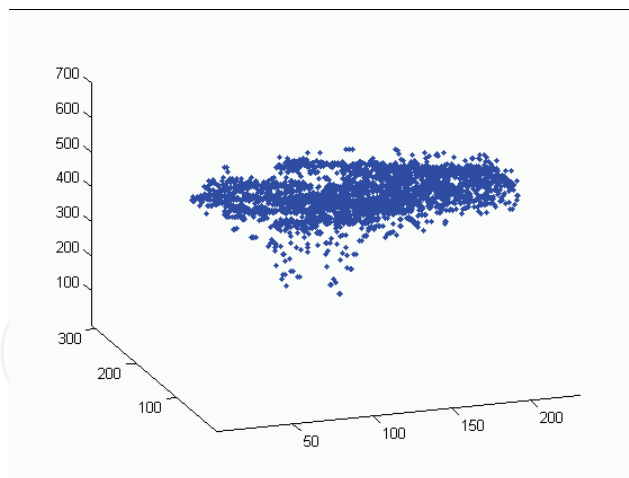
(a)
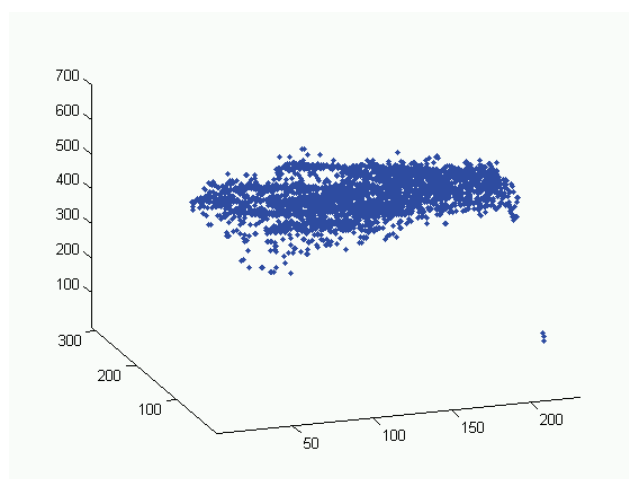
(b)
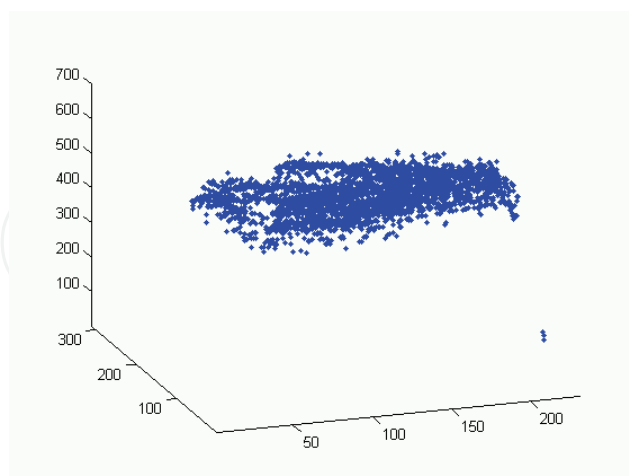
(c)

Fig. 6. (a) Calculated disparities: (a) using the original MPG algorithm, (b) using our modified MPG algorithm with optical flow from the 1st to the 2nd image frame, and (c) using our modified MPG algorithm with accumulated optical flow from the 1st to the 5th image frames.
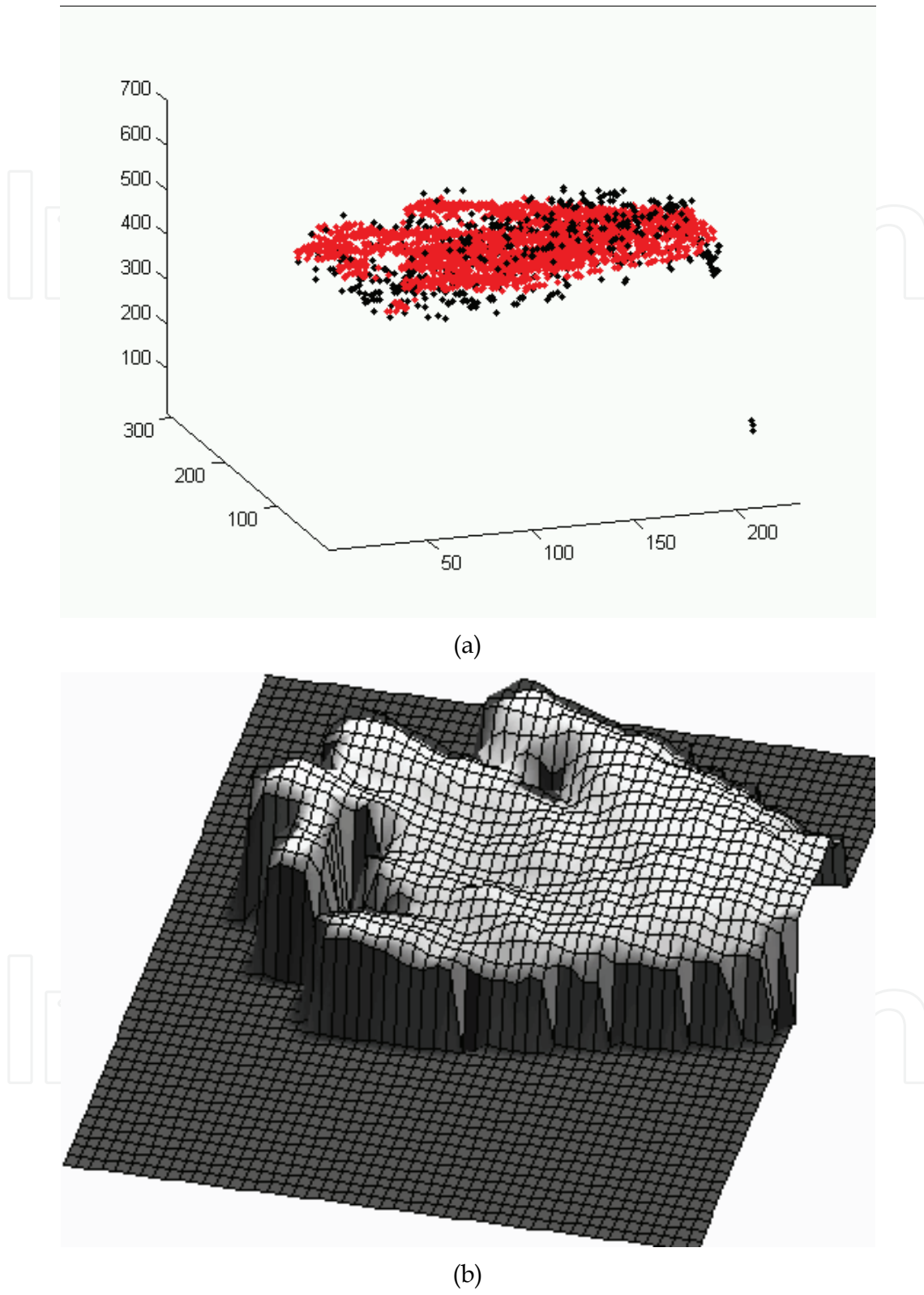
(a)

(b)

Fig. 7. (a) Noise separated disparities, (b) A hand surface reconstruction result using our modified MPG algorithm.

3. Conclusion

In this chapter, we proposed two modified stereo matching algorithms. One is the modification of an intensity based matching algorithm and the other one is the modification of a feature based matching algorithm. For the modification of an intensity based matching algorithm, we employed an additional matching criterion using optical flow to reduce the number of mismatching disparities. For the modification of a feature based matching algorithm, we simplified the matching procedure in the original MPG algorithm using optical flow. We presented some preliminary experimental results of the 3d structure reconstruction obtained by each modified matching algorithm for a pair of image sequence including a human hand.

8. References

A. M. Waxman & J. H. Duncan. (1986). Steps toward stereo-motion fusion. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 8, No. 6, November, (715-729)

A. Scheuing & H. Niemann. (1986). Computing depth form stereo images by using optical flow. *Pattern Recognition Letters*, 4, (205-212)

C. de Boor. (2004). *Spline Toolbox for use with MATLAB*. The Math Works, Inc.

D. Marr & T. Poggio. (1979). A computational theory of human stereo vision. *Proc. Royal Soc. London*, B204

D. Scharstein & R. Szeliski. (2002). A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision*, 47, (7-42)

G. Sudhir; S. Banerjee; K. K. Biswas & R. Bahl. (1995). A cooperative integration of stereopsis and optic flow computation. *Journal of the Optical Society of America A*, Vol. 12, No. 12, December, (2564-2572)

M. Z. Brown; D. Burschka & G. D. Hager. (2003). Advances in computational stereo. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 8, (993-1008)

O. Faugeras; B. Hotz; H. Mathieu; T. Viéville; Z. Zhang; P. Fua; E. Théron; L. Moll; G. Berry; J. Vuillemin; P. Bertin & C. Proy. (1993). Real time correlation-based stereo: algorithm, implementations and applications. *INRIA Technical Report*, (2013)

S. Tanaka & A. C. Kak. (1990). *Analysis and Interpretation of Range Images: Chapter 2. A Rule-Based Approach to Binocular Steropsis*. Springer-Verlag, New York

U. R. Dhond & J. K. Aggarwal. (1989). Structure from stereo – a review. *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 19, No. 6, (November-December), (1489-1510)

W. E. L. Grimson. (1981). A computational implementation of a theory of human stereo vision. *Phil. Trans. Royal Soc. London*, B292

Y.-H. Kim, A. M. Martínez, & A. C. Kak. (2005). Robust motion estimation under varying illumination. *Image Vision Computing*, Vol. 23, No. 4, (365-375)

Stereo Vision

Edited by Asim Bhatti

The book comprehensively covers almost all aspects of stereo vision. In addition reader can find topics from defining knowledge gaps to the state of the art algorithms as well as current application trends of stereo vision to the development of intelligent hardware modules and smart cameras. It would not be an exaggeration if this book is considered to be one of the most comprehensive books published in reference to the current research in the field of stereo vision. Research topics covered in this book makes it equally essential and important for students and early career researchers as well as senior academics linked with computer vision.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Yeon-Ho Kim and Soo-Yeong Yi (2008). Usinig Optical Flow as an Additional Constraint for Solving the Correspondence Problem in Binocular Stereopsis, Stereo Vision, Asim Bhatti (Ed.), ISBN: 978-953-7619-22-0, InTech, Available from:
http://www.intechopen.com/books/stereo_vision/usinig_optical_flow_as_an_additional_constraint_for_solving_the_correspondence_problem_in_binocular_

INTECH
open science | open minds