Bilgisayar Bilimleri Dergisi
Vol:1, No:1, 2016
© Anatolian Science

A Note on Depth Estimation from Stereo Imaging Systems

Baris Baykant Alagoz
Department of Computer Engineering, Inonu University, Malatya, Turkey

Abstract: The depth extraction from visual information is one of underpinning research area of robotics and there is a growing trend in development of autonomous and intelligent systems for real-life applications. These unmanned systems need reliable depth estimations in order to move in three-dimensional space, autonomously. Inspiring from biological vision systems, stereo imaging systems promise a solution for the depth estimation from binocular image pairs provided by stereo cameras. One of the major problems in the depth estimation from the stereo image pair is low depth resolution. This paper discusses the depth resolution problem and presents a depth resolution analysis for stereo imaging systems.

Keywords: Stereovision, disparity, focal length, baseline, depth estimation.

1. INTRODUCTION

Depth maps generation from stereo images has been a key problem of machine vision for two decades [1, 2]. Developments in three dimensional (3D) image reconstruction and autonomous systems applications [3] has stimulated the researches on depth estimation by using stereo imaging systems. Nowadays, the field of stereo vision became a promising research area and its application extending from measurement science to entertainment sector. Driving assistants for cars [4] and unmanned aerial vehicles (UAV) [5] are some robotic applications requiring a reliable depth estimation from visual information captured from cameras. However, artificial depth perception from camera images, which is as efficient as biological vision systems, is still an unsolved problem [1]. Many works addressed for improvement of stereo matching algorithms [6-19].

Depth map estimation from stereo correspondence has been taken into account as the the problem of accurately estimation of pixel disparities between binocular image pairs that is indeed the two different spatial views of the same scene. The other problem is the projection limitations of a stereo system, which severely reduces the resolution of depth maps. Even though the stereo matching algorithms can perform very efficient stereo correspondence between stereo images and yield accurate disparity maps, the depth resolution of the estimated depth maps is limited due to image resolution accordingly the pixel disparity resolution in stereo image pair. So, it is important to tuning stereo imaging system parameters to obtain a satisfactory depth resolution. This paper addresses analysis of stereo vision system parameters to enhance the depth resolution of stereo imaging systems.

Stereo matching is defined as the process of finding the best correspondence points between a stereo image pair [20]. In addition to hardware limitations, the accuracy of the disparity maps is limited depending on a number of methodological complications such as occlusion, ambiguity, illumination variation and projective distortion [20-22]. Stereo image rectification and camera calibration to enhance performance of stereo matching algorithms were addressed in detail [22-24].

Depth estimation from stereo images contains two main tasks: The first task is the estimation of the disparity map from the stereo image pair by means of a stereo matching algorithm. Disparity map contains pixel disparities between the projection of the same object in the left and right images. The second is the transforming the disparity map into a depth map as illustrated in Figure 1. The depth map is composed of the depth estimations of each pixel of projected objects in the stereo image and therefore it provides useful visual information for the depth perception, three dimensional reconstruction [25] and autonomous motion of the robotic system [3].

Accurate construction of disparity map is an important task for stereo vision system, which determines accuracy of the depth map estimation.

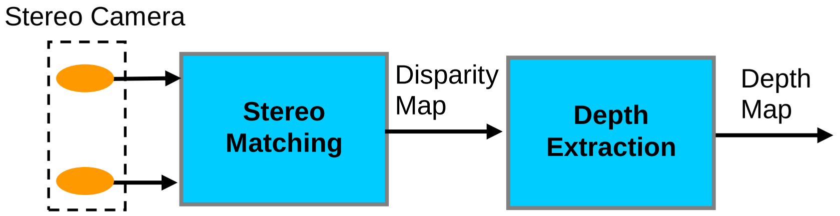

Figure 1. Fundamental blocks of a stereo dept estimation system

Depth resolution for canonical stereoscopic system was previously briefly discussed by Cyganek et al [22]. Importance of depth resolution in real stereo imaging applications were emphasized by Hsia et al. [5]. A detailed discussion and a systematic analysis are still needed for successful and reliable depth estimation by stereo imaging applications. This study presents a theoretical investigation on factors affecting depth estimation performance of stereo vision systems. Effect of stereo camera system parameters such as focal length, vertical pixel resolution of cameras, the length of base line between the camera pairs on the depth resolution of stereo images are analyzed.

2. METHODOLOGY

The widely used method for the extraction of the dept map from the disparity map is based on triangular transformation between real and projected points in two fronto-parallel camera system as illustrated in Figure 2(a). The parameters f and b represent focal length and base line of the fronto-parallel stereo imaging system as illustrated in Figure 2(a). This standard system (Canonical Stereoscopic System) is composed of two parallel cameras with focal lengths (f), displaced by a base distance (b) [22, 26]. The figure represents the projection of a surface located in $P(x, y, z)$ of the real world coordinate system. This surface is assumed to be projected by a pixel in the left projection screen (P_L) and the right projection screen (P_R). Stereo correspondence for $P(x, y, z)$ surface is valid under the following condition of the projections:

$$|L(i, j) - R(i+d, j)| = 0, \tag{1}$$

where, $L(i, j)$ and $R(i+d, j)$ are pixel location of the $P(x, y, z)$ surface projection on P_L and P_R, respectively. The pixel disparity from the left image matrix (L) to the right image matrix (R) is denoted by d . Disparity map (D) of the left image is composed of all pixel disparities from L to R image matrixes and defined as:

$$D(i, j) = \min_d \{|L(i, j) - R(i+d, j)|\}, \tag{2}$$

Estimation of disparity map was performed by stereo matching algorithms.
Real disparity on the projection screen in Figure 2(a) is defined as,

$$d_r = R_r - L_r, \tag{3}$$

where, R_r is real location of projection on the projection screen P_R and expressed according the vertical resolution of pixel, α_i, as $R_r = (i+d)/\alpha_i$. The L_r distance is the real location of projection on the left screen P_L and expressed as $L_r = i/\alpha_i$. One can express the real disparity matrix D_r to express relation between the real disparity and the pixel disparity as the flows:

$$D_r = D/\alpha_i, \tag{4}$$

Considering the triangular in Figure 2(a), the following equation is written,

$$\frac{Z-f}{Z} = \frac{b}{b + (w/2 - L_r) + (R_r - w/2)}, \tag{5}$$

Considering equation (5), the depth of the (i, j) pixel of the left image (L) is written as,

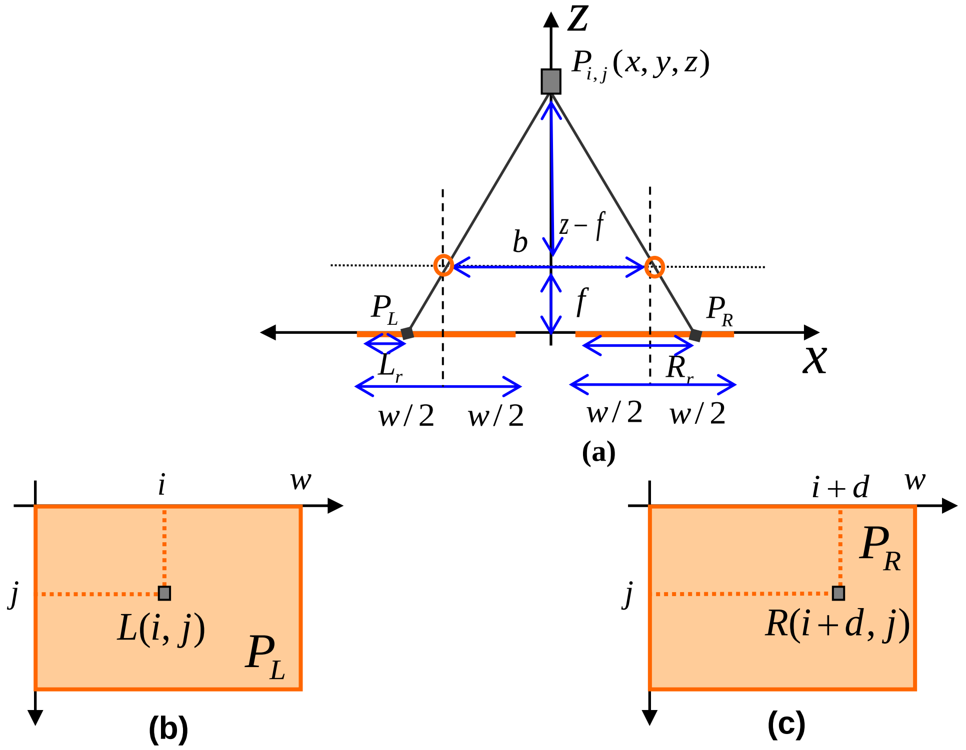

Figure 2. (a) Projection model used for fronto-parallel stereo vision systems, (b) Left image and (c) Right image

$$Z(i,j) = \frac{f\,b}{D_r(i,j)} + f\,, \tag{6}$$

Considering equation (4), depth map matrix can be formed depending on the pixel disparity matrix as the following,

$$Z(i,j) = \frac{f\,b\,\alpha_i}{D(i,j)} + f\,, \tag{7}$$

Parameters f ,b and α_i are structural parameters of stereo vision systems. The term of $f\,b\,\alpha_i$ is called as depth gain (G_z) of stereo system. In this case, the depth map can be expressed as,

$$Z(i,j) = \frac{G_z}{D(i,j)} + f\,. \tag{8}$$

Here, the depth gain, $G_z = f\,b\,\alpha_i$, is important structural parameter of the stereo imaging system and it means the depth per disparity. Tuning of G_z provides a desired depth resolution for stereo imaging system. The depth resolution of the system was defined as the number of pixel disparity per unit depth and it can be expressed as,

$$\alpha_z = D(i,j)/Z(i,j) = \frac{G_z}{Z^2(i,j)} + \frac{f}{Z(i,j)}\,. \tag{9}$$

Process of tuning the parameter G_z, while capturing stereo images, is referred to as stereo focusing and this process is very useful for obtaining a desired depth resolution for an object at the distance z from the stereo camera system. Stereo focusing aims to improve details in depth maps for the remote objects, in a similar manner; the lens focusing of camera system improves the visual details of a far object in the photography.

Figure 3(a) shows the depth versus disparity characteristic of the conventional stereo imaging system (Figure 2(a)) for 3 different depth gain configuration. The beginning of characteristics, which are marked by circles, shows the maximum measurable range, which is obtained for the disparity value of one. This figure also reveals that depth gain increases the measurable range of the stereo imaging systems. Figure 3(b) shows corresponding depth resolution versus disparity characteristics.

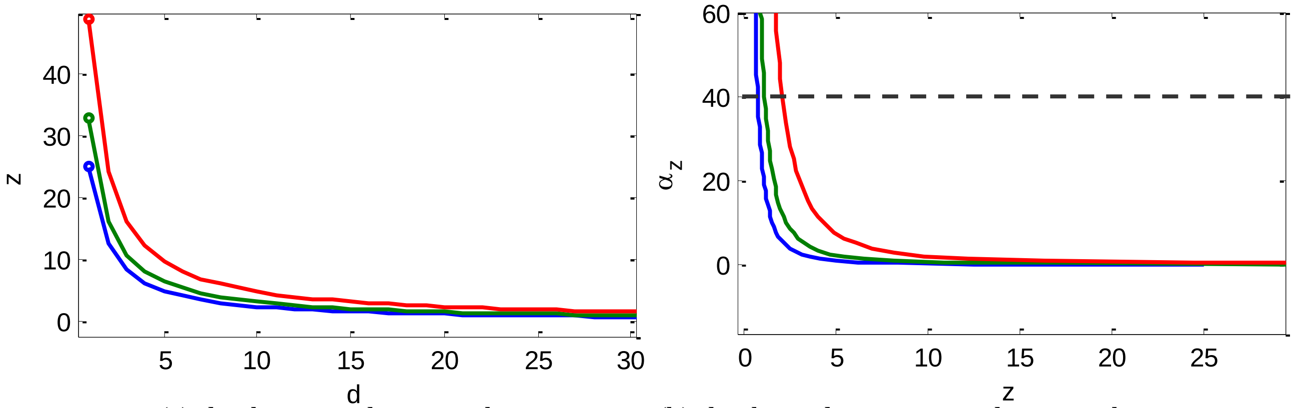

Figure 3(a) depth versus disparity characteristics; (b) depth resolution versus disparity characteristics for different G_z

3. STEREO FOCUSING STRATEGIES FOR FRONTO-PARALLEL STEREO IMAGING APPLICATIONS

The stereo focusing should be carries out to obtain a desired depth resolution depending on requirements of applications. Different stereo focusing strategies can be used depending on the freedom degrees of the depth gain parameters (f , b , α_i). The stereo imaging systems exhibit different freedom degrees of f , b , α_i corresponding to the nature of the application. The range of parameters f and α_i are limited by camera features used in stereo imaging. The parameter b depends on camera localization in stereo imaging. Some possible stereo vision application and depth resolution adjustment strategies are summarized as,

- *Moving Single-Camera Stereo Applications*: In this type application, stereo images is obtained by consecutive snap shot of scene by moving a camera. In a long-range application, a camera mounted on a plane or satellite used for the stereo imaging of scene for 3D reconstruction of the landscape [25]. In very short range application, microscopes or electron microscopes can be used for 3D reconstruction of microscopic structures by consecutive image capturing from different location [27]. In aerial application, the parameter b determined by the speed and capturing period of the moving camera, and therefore it provides very large degree of freedom for depth resolution adjustment process. For this type of application, by considering equation (9), the value of parameter b for a desired depth resolution in a target distance can be calculated for a desired depth resolution as,

$$b = \frac{Z^2}{\alpha_i}\left(\frac{\alpha_z}{f} - \frac{1}{Z}\right) \tag{10}$$

-*Fixed Baseline Stereo Camera System Applications*:
In this type application, stereo image pair is obtained by two camera system with a constant base line distances. This type of stereo imaging generally used in robotic applications such as humanoids, unmanned ground or aerial vehicles. For these systems, parameters f can be adjustable by the lens of cameras. The value of parameter f for a desired depth resolution at target distance can be found by

$$f = \frac{Z^2 \alpha_z}{b\alpha_i + Z} \quad (11)$$

-*Variable Baseline Stereo Camera System Applications*:
In some applications, stereo image pairs are obtained by two camera system and base line can be adjustable. So, the both parameters f and b can be used to adjust for a desired dept resolution according to equation (10) and (11). This type stereo imaging systems allows fine tuning of depth resolution for robotics, 3D reconstruction and measurement applications.

CONCLUSION
As summary, depth resolution of stereo vision system are analyzed on the bases of structural parameters of stereo vision hardware systems. This analyses allows the adjustment of stereo vision system to obtain improved depth resolution for applications utilizing stereo imaging systems. Hence, stereo focusing concept based on depth resolution adjustment is suggested. A future work should be carried out for the experimental study of stereo focusing.

REFERENCES
[1] Tomas Montserrat, Jaume Civit, Oscar Divorra Escoda, Jose-Luis Landabaso, Depth estimation based on multiview matching with depth/color segmentation and memory efficient belief propagation. In 2009 16th IEEE International Conference on Image Processing (ICIP),2009, pp. 2353-2356.
[2] Hoon Yoo, Joon-Jae Lee, and Byung-Gook Lee, "Depth extraction of three-dimensional objects using block matching for slice images in synthetic aperture integral imaging." Applied optics 50.29 (2011): 5624-5629.
[3] M. Bertozzi, A. Broggi, and A. Fascioli, "Vision-based intelligent vehicles: State of the art and perspectives," Robotics and Autonomous Systems, vol. 32, pp. 1–16, 2000.
[4] Mohamed El Ansari, St´ephane Mousset, Abdelaziz Bensrhair, George Bebis, Temporal Consistent Fast Stereo Matching for Advanced Driver Assistance Systems (ADAS), 2010 IEEE Intelligent Vehicles Symposium University of California, San Diego, CA, 2010,pp.825-831
[5] Kuo-Hsien Hsia, Shao-Fan Lien and Juhng-Perng Su (2012). Stereo Matching Method and Height Estimation for Unmanned Helicopter, Machine Vision - Applications and Systems, Dr. Fabio Solari (Ed.), ISBN: 978-953-51-0373-8, InTech, Available from: http://www.intechopen.com/books/machine-vision-applications-andsystems/
stereo-matching-method-and-height-estimation-for-unmanned-helicopter
[6] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo sorrespondence algorithms," International Journal of Computer Vision, vol. 47, no. 1-3, pp. 7–42, 2002.
[7] G.P. Otto and T.K. Chau. A region-growing algorithm for matching of terrain images. Image and Vision Computing, 7(2):83-94, 1989.
[8] Y. Boykov, O. Veksler and R. Zabih, Fast Approximate Energy Minimization via Graph Cuts, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 11, 2001.
[9] V. Kolmogorov and R. Zabih, Computing Visual Correspondence with Occlusions using Graph Cuts, IEEE International Conference on Computer Vision, Vol. I:508-515 2001.
[10] J. Sun, N.-N. Zheng and H.-Y. Shum, StereoMatching Using Belief Propagation, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, No. 7, July 2003.
[11] J. Sun, Y. Li, S. B. Kang and H.-Y. Shum, Symmetric Stereo Matching for Occlusion Handling, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. II:399-406, 2005.
[12] T. Kanade and M. Okutomi, "AStereo Matching Algorithm with an Adaptive Window: Theory and Experiment," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 16, no. 9, pp. 920-932, Sept.1994.

[13] Li Tang *, Chengke Wu, Zezhi Chen, Image dense matching based on region growth with adaptive window,, Pattern Recognition Letters 23 (2002) 1169–1178.
[14] S. Kang, R. Szeliski, and J. Chai, “Handling Occlusions in Dense Multi-View Stereo,” Proc. IEEE CS Conf. Computer Vision and Pattern Recognition, vol. 1, pp. 102-110, 2001.
[15] Yong Seok Heo, Kyoung Mu Lee, Sang Uk Lee, Robust Stereo Matching Using Adaptive Normalized Cross-Correlation, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 33, NO. 4, APRIL 2011
[16] S.B. Marapane, M.M. Trived, Region-based stereo analysis for robotic applications, IEEE Trans. Systems Man Cybernet 19 (1989) 1447-1464.
[17] A. Knotanzad, A. Bokil, Y.W. Lee, Stereopsis byconstraint learning feed-forward neural networks, IEEE Trans. Neural Networks 4 (1993) 332-342.
[18] Elisabetta Binaghi , Ignazio Gallo, Giuseppe Marino, Mario Raspanti, Neural adaptive stereo matching, Pattern Recognition Letters 25 (2004) 1743–1758.
[19] Kyu-Phil Han , Kun-Woen Song, Eui-Yoon Chung, Seok-Je Cho,
Yeong-Ho Ha,Stereo matching using genetic algorithm with adaptive Chromosomes, Pattern Recognition 34 (2001) 1729}1740.
[20] Pooneh Bagheri Zadeh, Cristian V. Serdean, Stereo Correspondence Matching using Multiwavelets, 2010 Fifth International Conference on Digital Telecommunications IEEE, 2010, 153-157.
[21] D. Scharstein and R. Szeliski, “A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms,” International Journal of Computer Vision, vol. 47, pp. 7-42, April 2002.
[22] Bogusław Cyganek, J. Paul Siebert, An Introduction to 3D Computer Vision Techniques and Algorithms, 2009 John Wiley & Sons, Ltd. ISBN: 978-0-470-01704-3.
[23] Ayache, N. and Hansen, C. (1988) Rectification of Images for Binocular and Trinocular Stereovision. INRIA Technical Report 860.
[24] Fusiello, A., Trucco, E. and Verri, A. (2000) A compact algorithm for rectification of stereo pairs. Machine vision and Applications, 12, 16–22.
[25] Gaurav Gupta, M.S. Rawat, R. Balasubramanian and Rama Bhargava, B.G. Krishna, Region growing stereo matching method for 3D building reconstruction, Int. J. Computational Vision and Robotics, Vol. 2, No. 1, 2011.
[26] S. T. Barnard, M. A. Fischler, “Depth map edge and intensity based stereo”, in Proc of the 7th Int. Joint Conference on Artificial Intelligence, , Vancouver, 1981, pp.631-636.
[27] Boyde, A. Colour-coded stereo images from the tandem scanning reflected light microscope (TSRLM). Journal of microscopy, 146(2), (1987),137-142.