ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 11, May 2015

A Review on Depth Estimation for Computer Vision Applications

Dineesh Mohan, Dr. A. Ranjith Ram
Department of Electronics & Communication Engineering Government College of Engineering Kannur

Abstract*—in this review paper, depth estimation techniques using cues from two images are discussed. In this paper, different approaches for depth estimation like Vergence, Stereo Disparity, Stereo Matching, Familiar Size, Defocus Cue, Convex Optimization, Sum of Absolute Differences Algorithm are reviewed and finally disparity based depth estimation using Sum of Absolute Difference(SAD) matching algorithm is proposed. Depth is determined by using maximum and minimum disparity.

***Index Terms*—Correspondence problem, Sum of Absolute Difference, epipolar constraint, disparity, fundamental matrix, depth.**

I. INTRODUCTION

Many scientists have been working to create robots that perform manual work for years. It is important to estimate the position of objects in the robot's visual field. Depth information is used to achieve this. When looking out of the side window of a moving car, the distant scenery seems to move slowly while the lamp posts flash by at a high speed. This effect is called parallax, and it is used to extract the geometrical information from a scene. The distance of the objects is obtained by using multiple captures of the same scene at different viewpoints. i.e depth of the scene can be determined. The distance of the points from the camera can be determined by tracking the displacement of those points between the captured images. Disparity means pixel displacement between corresponding points in the multi view images or stereo images. Stereo vision systems reconstruct 3D scenes by matching two or more images taken at slightly different viewpoints. Stereo vision systems are based on two forward-facing cameras, where each camera gives a 2D projection of a scene. The matching of image points is obtained by comparing a region in one image with matching regions in the other image and selecting the most likely match based on some similarity measure. These correctly matching points can be used to calculate the depth of the point. Depth information is used for mobile robot navigation and obstacle detection in real time applications. The rest of the paper is organized as follows. In section II, we present the current state of the art related to depth estimation using multiple images. In section III, we discuss the various methods of depth estimation. In Section IV we propose the work that is intended as the first authors M. Tech. project and in Section V we discuss its applications. Finally Section VI summarizes this paper with some concluding remarks.

II. CURRENT STATE OF THE ART

The human visual system is prepared for the depth perception. This perception is possible by a combination of different and complementary physiological and psychological structures and functions. They are explained below as Two eyes: the most important source of depth perception is the two eyes, sharing an important area of vision. However, the fronto-parallel hypothesis is only respected when looking at something placed in the infinity. The angle of obliqueness (parallax) also provides information about the distance of the object. Focus: the crystalline is an elastic tissue which allows changing the focal distance of the eye and, hence, focusing in a wide range of distances. This information helps the brain computing the distance of the focused plane. Differences in brightness: For constant illumination, the depth can be perceived in terms of the brightness. This method has been applied to compute the distance to stars (however, the hypothesis of constant brightness was not true), and works in daily live to help the brain knowing the distance, as perceived in figure 18. Several researches have been conducted in the recent decades for the depth estimation. The main difficulty encountered in the stereo vision systems is the stereo matching which determines the spatial displacement between each two corresponding pix- els in a stereo pair. This process is called the correspondence problem. Intensive research has been conducted recently to solve the problem of finding the correspondences between the right and the left images. i.e the correspondence problem. Generally Stereo matching algorithms can be categorized into two major classes : local and global methods. Global methods formulate the problem in terms of an energy function. Then all disparities can be determined simultaneously. This is done by applying energy minimization techniques. Most of these methods are computationally expensive. Local methods have higher computational efficiency and they are more suitable. For real-time applications. The problem of reconstructing a three dimensional scene from several viewpoints was first investigated in the fields of human stereopsis and aerial photography. Recently, the scene reconstruction problem was treated as a matching problem where the objective was to match points or features between two or more images. After match is obtained, the three dimensional position of the point is

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 11, May 2015

determined by triangulation method assuming the camera positions were known. In 1999, Mudenagudi. U et. al. [1] proposed an algorithm that estimates depth using defocused stereo image pairs. This algorithm provides simultaneous recovery of depth and image restoration. Rajagopalan. A.N et. al. [2] a paper in 2004, on depth estimation. This method uses the fusing of defocus and stereo cues. It use Stereo based constraints to obtain improved depth estimates. In 2008, depth estimation using vertical edge is proposed by Yiming Nie et. al. [3]. This paper presents an image processing technique that can fast estimate vertical edges depth from binocular vision images. This method method is a compromise between local methods and global optimization. In 2010, Cem Karaoguz et. al. [4] presented on depth estimation that is based on the active vision(stereo disparity, vergence and familiar size) methods. In the same year Changyin Zhou et. al. [5] proposed an algorithm for the depth estimation. This algorithm is based on the diffusion. In 2011, Sidhu et. al. [6] presented a paper based on robust disparity estimation even in the presence of occlusions. The algorithm employs the Sum of Absolute Difference (SAD) approach for the disparity estimation. This paper explains the removal of occlusions by using Left-Right consistency constraint. In 2012, Patrik Kamencay et. al. [7] proposed a method on depth map estimation based on hybrid Method. This method is based on the combination of the Belief Propagation and Mean Shift algorithms. This hybid algorithm use image filtering and modified SAD (Sum of Absolute Differences) stereo matching method. In 2013, Ashraf Anwar Fahmy [8] proposed a method on depth estimation based on a well known matching algorithm, SURF, and epipolar constraints. In the same year Stefanoski. N et. al. [9] proposed a method based on the diffusion of depth features. Linear- runtime feature diffusion algorithm was presented in this paper which aims for the fast and accurate processing of such high resolution data. In 2014, Chirag. S et. al. [10] presented a paper on depth estimation using Sum of Absolute Difference (SAD) Algorithm is based on correlation techniques used for the disparity estimation.

III. METHODS OF DEPTH ESTIMATION

There are various methods for depth estimation found in literature, of which a few are discussed in this section.

A. Vergence

When both eyes are positioned such a way that the optical axes intersect on the surface of an object and it allows the projection of the object to fall on the foveae of both retinae. Thus stereo fixation on the object is obtained. This type of eye movement is call edvergence. It is an important source of information about depth in the human visual system [11]. The depth estimation by vergence triangulation foranartificial systemisshowninFig.1.Herecorrelation basedvergence control algorithm [12] isusedtoachieve fixationontheobject surface.

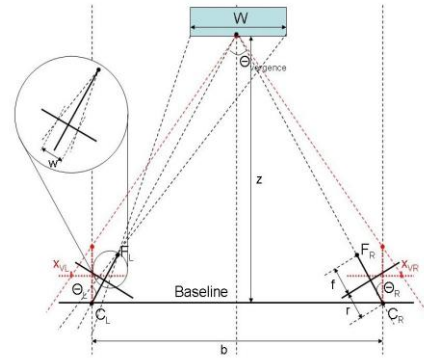

Fig. 1. Analytical model of the active vision system and depth estimation methods

B. Stereo Disparity (SD)

An object which is not at the stereo fixation and it projects to different locations on the left and right retinae according to scene depth and the horizontal baseline separating the eyes [11]. The difference between these two locations is called the stereo disparity (SD). It is a common depth cue used in artificial vision systems [13], [14]. Here depth is computed from the stereo disparity. An active rectification process [15] is used to obtain absolute depth information (i.e. the distance from the baseline to the stereo fixated object). Disparity maps are computed using the block matching algorithm [16] which was used in the active vision case. Using this algorithm, disparities are refined via post processing (sub-pixel interpolation and post-filtering). Color based segmentation methods [17], [18] are used here to obtain the disparities of the object in the disparity maps and the average of these disparities was taken for depth estimation. The depth z is computed as

$$z = (bf/d) + r + f \qquad (1)$$

Where d is the disparity, $d = x_{VL} - x_{VR}$(where x_{VL} and x_{VR} are the projections of the object on the virtual left and right image planes).f is the focal length of the cameras and r is the distance from the center of rotation of the cameras to the image planes.

C. Stereo matching

When different viewpoints from the same scene are compared, a problem a rises that is associated with the mutual identification of images. The solution to this problem is commonly referred to as matching. The matching process consists of identifying each physical point within different images. However, matching techniques are not only used in stereo or multivision procedures but widely used for image retrieval or fingerprint identification where it is important to allow rotational and scalar distortions. There are various constraints that are generally satisfied by true matches thus simplifying the depth estimation algorithm, such as similarity, smoothness, ordering and uniqueness. The matching process is a conceptual approach to identify similar characteristics in

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 11, May 2015

different images. It is, then, subjected to errors. The matching is, hence, implemented by means of comparators allowing different identification strategies such as minimum square errors (MSE), sum of absolute differences (SAD) or sum of squared differences (SSD). The characteristic compared through the matching process can be anything quantifiable. Thus, we will see algorithms matching points, edges, regions or other image cues.

D. Familiar Size (FS)

The depth of an object can be estimated from the size of its projection on the camera images if the real size of the object is known. Various approaches [19] exist for this operation. The depth z can be derived as

$$Z= ((fW/w)+r+f)\cos\theta \quad (2)$$

Where θ is the camera angle and cosθ ≈ 1. W is the size of the object. w is the retinal size.

E. Combination of Methods

The combination of methods use Bayesian cue integration [20] to obtain the depth estimation methods.

F. Using defocus cue

This method [21] regularization based approach is used for the simultaneous depth estimation and image restoration from defocused observations. It uses two defocused observations of a scene that are captured with different camera parameters for the depth estimation. This method consists of two steps. In the first step, depth estimate is obtained for the focused image. In the second step, fast optimization is used for refining the solution.

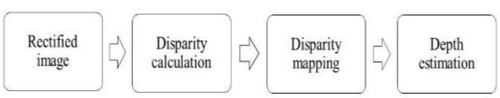

Fig. 2. Depth estimation process using disparity measure obtained from two images

G. Convex Optimization Approach

This method [22] is used for the robust depth estimation from a stereo pair under varying illumination conditions. A spatially varying multiplicative model is developed to account for brightness changes induced between left and right views. The depth estimation is formulated as a constrained opti- mization problem in which an appropriate convex objective function is minimized.

H. Using object placement relation

Object placement [23] is one of vision cues usually used to identify 3-d position efficiently. Extraction of such information is not so trivial. This method presents an adaptive algorithm which defines placement information as a constraint and it is used to estimate depth from a single scene image having many arbitrary objects.

I. Using Sum of Absolute Differences Algorithm

In this method, the rectified images have to run through several processes starting from stereo correspondence until the disparity mapping. Then depth is obtained from the mapping of the disparity values by using intensity of pixel value for each matching point. Sum of Absolute Differences (SAD) algorithm is used to solve the correspondence problem. The depth estimation process is shown in Fig. 2.

IV. PROPOSED M.TECH.PROJECT

In this section, we propose a disparity based depth esti- mation technique. The differences between the in liers points resulting from epipolar constraint to get the maximum and the minimum disparity. The SAD algorithm is used for finding the matched points in the stereo images. The depth is calculated from the minimum and the maximum disparity. Here the depth information is calculated directly from the inliers resulting from the fundamental matrix. Mean depth is obtained from the mean value between the maximum disparity and the minimum disparity. The main steps in the proposed M.Tech project is shown in Fig. 3.

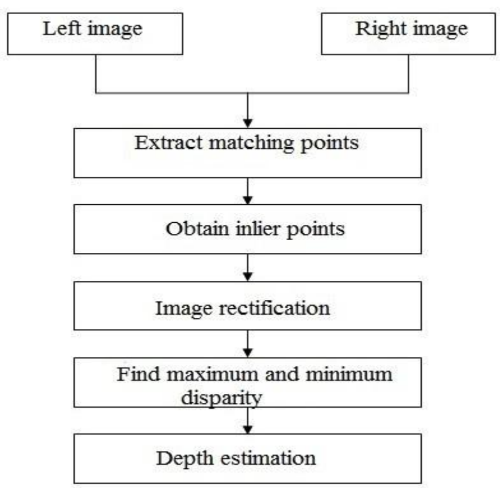

Fig. 3. Steps involved in the proposed method of depth estimation using disparity

V. APPLICATIONS

This research is focused on the study of developing stereo vision depth estimation algorithm for the mobile robot applications. Depth estimation is suitable for the mobile robot navigation and obstacle detection in real time applications. Mobile robots are able to navigate both indoor and outdoor environments. Depth information is used to navigate themselves and respond to the environment. In robotics, depth is used to extract information about the relative positionof3D objects in the vicinity of autonomy in scientific applications for digital stereo vision includes the extraction of information from aerial surveys or calculation of 3D heliographical information such as obtained by the NASA STEREO project us system. It is also used in scientific applications for digital stereo vision includes the extraction of information from aerial surveys or calculation of3D heliographical information such as obtained by the NASA STEREO projects.

VI. CONCLUSION

The depth is an important cue of a scene, which is lost in standard image acquisition systems. For that reason, and given that many applications need this information, several strategies have been proposed to extract the depth. We have seen active methods, which project some energy onto the scene to process the reflections, and passive methods, only dealing with the natural received energy from the scene. Among this last option, we found monocular systems, working with a single perspective, and stereo or multiview systems, which work with more than one single perspective. We have shown why these last algorithms have to solve the matching problem, or finding the same physical points in two or more images. Several strategies, again, are available in this category. The analysis has revealed advantages and disadvantages in every system, regarding energy needs, computational load and, hence, speed, complexity, accuracy, range, hardware implementation or price, among others. Thus, there is not a concluding winner among all the analyzed solutions. Instead of that, we will have to think about the final application of our algorithm, to make the correct choice. We have presented an efficient and robust algorithm in depth estimation which is suitable for mobile robot navigation and obstacle detection in real time applications.. This term paper proposes to develop an algorithm for depth estimation as an M.Tech project.

REFERENCES

[1] U.Mudenagudi and S.Ghaudhuri,“ Depth estimation using defocused stereo image pairs,” The Proceedings of the Seventh IEEE International Conference on Computer Vision, vol.1,pp.483-488,1999.

[2] A.N.Rajagopalan, S.Chaudhuri and U.Mudenagudi, “Depth estimation and image restoration using defocused stereo pairs,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.26, pp.1521-1525, November2004.

[3] Yiming Nie, TaoWu, Xiangjing An and Hangen He,“ Vertical Edge Depth Estimation,” International Conference on Computer Science and Software Engineering, vol.1,pp.129-132, December2008.

[4] Cem Karaoguz, Andrew Dankers, Tobias Rodemann and MarkDunn, “An Analysis of Depth Estimation within Interaction Range,”International Conference on Intelligent Robots and Systems, October2010.

[5] Changyin Zhou, Columbia, O.Cossairt and S.Nayar, “Depth from Diffusion,”IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.1110-1117, June 2010.

[6] H.S.Sidhu, S.Kumar and A.Das, “A robust area based disparity estimation technique for stereo vision applications,”International Conference on Image Information Processing (ICIIP), pp.1-4, 2011.

[7] Patrik Kamencay, Martin Breznan, Roman Jarina, Peter Lukac and Martina Zacharia sova, “Improved Depth Map Estimation from Stereo Images Based on Hybrid Method, ”Proceedings of the IEEE International Conference on Radio Engineering, vol.1April2012.

[8] T.Syeda-Mahmood and D.Ponceleon,“ Learning video browsing behavior and its application in the generation of video previews,” ACM Multimedia, pp.119–128,October2001.

[9] N.Stefanoski, C.Bal, M.Lang and O.Wang, “Depth estimation and depth enhancement by diffusion of depth features, ”IEEE International Conference on Image Processing(ICIP), pp.1247-1251,September 2013.

[10] Chirag S.Panchal, Abhay B.Upadhyay, “Depth Estimation Analysis Using Sum of Absolute Difference Algorithm, ”International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, vol.1,January2014.

[11] S.E.Palmer, “Vision Science: Photons to Phenomenology,MITpress” 1edition,1999.

[12] A.Bernardino and J.Santos Victor, “Correlation based vergence control usinglog-polarimages,”InIntelligentRoboticSystems, 4thInternational Symposium, 1996.

[13] M.Z.Brown, D.Burschka, and G.D.Hager, “Advances in computational stereo,”IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.1, pp. 9931008, September 2003.

[14] D.Scharstein, R.Szeliski, and R.Zabih, “A taxonomy and evaluation of dense two-frame stereo correspondence algorithms, ”Proc. IEEE International Conference on Image Processing, pp.477–480, September 2001.

[15] A.Dankers, N.Barnes, and A.Zelinsky, “Active vision-rectification and depth mapping,” In Australasian Conference on Robotics and Automation, 2004.

[16] K.Konolige, “Small vision systems: hardware and implementation,”In Eighth International Symposium on Robotics Research, pp.111-116, 1997.

[17] A.Dankers, N.Barnes, and A.Zelinsky, “Mapzdf segmentation and tracking using active stereo vision: Hand tracking case study, “Computer Vision Image Underst., vol.2, pp.74-86,2007.

[18] A.Denecke, I.Ayllon-Clemente, H.Wersing, J.Eggert, and J.Steil, “Figure-ground segmentation using metrics adaptation in level set methods, ”In European Symposium on Artificial Neural Networks(ESANN), 2010.

[19] C.Zhang, V.Willert, and J.Eggert, “Tracking with depth-from-size,” International conference on Neural Information Processing, pp.274- 283, 2009.

[20] M.Ernst and M.Banks, “Humans integer at visual and haptic information in a statistically optimal fashion,” pp.429433, January2002.

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 11, May 2015

[21] K.Ranipa and M.Joshi, "A practical approach for depth estimation and image restoration using de focus cue ,"IEEE International Workshop on Machine Learning for Signal Processing(MLSP) pp.1-6,2011.

[22] W.Miled, J.Pesquet and M.Parent, "A Convex Optimization Approach for Depth Estimation Under Illumination Variation," IEEE Transactions on Image Processing vol.18,pp. 813-830,January 2009.

[23] N.Futragoon, P.Kanongchaiyos, "Enhanced depth estimation by using object placement relation, "IEEE International Conference on Robotics and Biomimetics pp.1899-1904,2009.

AUTHOR BIOGRAPHY

Dineesh Mohan received the B-Tech degree in electronics and communication engineering in 2012 from Cochin university of Science and Technology kerala,. Currently doing MTech degree in signal processing and embedded system at kannur university kerala

Dr. A. Ranjith Ram received the B-Tech degree in electronics and communication engineering from university of Calicut. Received M-Tech degree from College of Engineering Thiruvananthapuram. Recently got Doctorate Degree from IIT Bombay and currently work as Assistant Processor on ECE Dept at GCE Kannur.

.