Retinal Disparity Estimation and 2D Fusion in Stereo-Pairs

Priyanka.V.K[1], Prof.Vijayakumar.K, Assoc.Prof. Sudheesh Madhavan

PG Scholar, Dept.ECE, TocH Institute of Science and Technology, Ernakulam, India

ABSTRACT: Retinal disparity is the slight difference in the images formed in the retina due to the position and phase through which eyes view an object. In computer vision systems, stereo matching algorithms are very popular. Most of these techniques are used to implement a 3D vision system by analyzing the disparity map of the stereo pairs. But the case of 2D vision is slightly different from this. In two dimensional objects retinal or binocular disparity is encoded by receptive field position disparity. It is due to the slight horizontal shift between cameras, which resembles human eyes. This paper goes through one of the simple and accurate method for finding receptive field position disparity in stereo pairs and binocular fusion to provide a two dimensional vision to machine vision system

KEYWORDS: Binocular disparity; Binocular fusion; computer vision; Normalized cross-correlation; Receptive field position disparity; Stereo pairs; striate cortex

I. INTRODUCTION

Vision is the most dominant sense from which we can derive most of the information about the world. Human beings are blessed with the attributes of binocular vision, which is the coordinated use of eyes. Each eye has a monocular visual field of 150 degree. If the eyes are functioning normally and equally the images formed in the retinas will be same in size, illuminance and colour. But, it has slight difference due to the positioning and angle through which eyes view an object [8].This difference is usually termed as binocular disparity or retinal disparity. Striate cortex is the first site along the central visual pathways at which signals from left and right eyes combines to a single neuron [3].The process of encoding binocular disparity begins here. Receptive field (RF) position disparity and RF phase disparity are the major reasons for binocular disparity. In the case of two dimensional object RF position disparity is dominant. This position disparity can be experienced by just closing our eyes one by one. If observing an object far away from our eyes, this position disparity has the main contribution in binocular fusion, it is due to the distance to an object is inversely proportional to the disparity. Visual system utilizes binocular disparity to discriminate relative depth of object in space.

Binocular fusion is carried out in brain. The existence of corresponding retinal elements with their common relative subjective visual direction is the essence of binocular fusion. The term motor fusion refers to the ability align eye in such a manner that sensory fusion can be maintained. Due to the coordinated use of eyes we can view a wider visual field. There is a region, where both eyes field view overlap. This region is called as binocular field.

In computer vision, human vision processing is an important aspect, in which a model of human eye is created. In such systems, receptive field position disparity estimation is very essential because it can be used for depth perception and binocular fusion. The ultimate goal of this paper is to find a method for accurate disparity estimation and binocular fusion.

II. RELATED WORK

Andre R et al (1999) proposed a paper on correspondence estimation in image pairs [2]. Spatial image pairs can be obtained by recording a scene with two cameras at different position and same time. Correspondence estimation in spatial pair is called disparity estimation. Different methods were there to find the geometrical correspondence and photometric correspondence of image pairs. Block matching techniques, feature based matching, optical flow algorithm etc.were some among them. The latter provides one of the most powerful constraints in correspondence estimation. It

has been applied widely on parallel image pairs, and recently also on uncalibrated spatial pairs.

Zitova B et al (2003) developed a review of image registration methods [4]. Image registration is the process of overlaying two or more images of the same scene taken at different times, viewpoints, and sensors. The registrations geometrically align the reference and sensed images. The approaches are classified according to their nature as feature-based and area based techniques, and according to four basic steps of image registration procedure: mapping function design, feature detection, image transformation, and feature matching, Main advantages and limitations of the methods are also mentioned in the paper. The main objective of this paper is to provide a reference source for the researchers involved in image registration, regardless of application areas.

Banks J et al (1997) developed non parametric techniques for stereo matching [1]. This paper compares a number of stereo matching algorithms in terms of robustness and suitability to fast implementation. This includes area based algorithms and algorithms based on nonparametric transforms notably the rank and census transforms. The rank transform is is defined as the number of pixels in the window whose value is less than the centre pixel. The images will therefore be transformed into an array of integers whose value ranges from 0 to N-1 where N is the number of pixels in the window. Census Transform maps the window surrounding the centre pixel to a bit string. Results show that the rank and census transforms are robust with respect to radiometric distortion and introduce less computational complexity than conventional area based matching techniques.

Luca L et al (2013) proposed a method for autonomously learning representations of visual disparity between images from left and right eye, as well as appropriate saccadic movements to fixate objects with both eyes [6]. A sparse coding model encodes sensory information using binocular basis functions, while a reinforcement learner generates the eye movement, according to the sensed disparity. A multi-scale approach, which exploits binocular basis functions at different resolutions, is used to encode disparities in different ranges. The sparse disparity matching technique has the advantages of the small amount of calculation and the short matching time. This method just can gain some limited disparity information, so it is bad for the reconstruction of the scene.

Nuno R et al (2000) developed a comparative analysis of the performance and characteristics of a set of cross correlation algorithms[5]. The presented analysis was focused on the study of final matching error and computational load of the correlation functions. Some of the cross correlation measures are given below. R (u, v) denotes a reference window pixel, S(c, l) a search window pixel, $\bar{R}$ the local mean of reference window and $\overline{S(c,l)}$ the pixel mean in the block of search window being compared.

$$\bar{R} = \sum_{v=0}^{Rlength} \sum_{u=0}^{Rwidth} R(u, v) \tag{1}$$

$$\overline{S(c,l)} = \sum_{v=0}^{Rlength} \sum_{u=0}^{Rwidth} S(c + u, l + v) \tag{2}$$

Simple Cross correlation $SCC(c, l) =$

$$\sum_{v=0}^{Rlength} \sum_{u=0}^{Rwidth} R(u, v). S(c + u, l + v) \tag{3}$$

Normalized Cross-correlation $NCC(c, l) =$

$$\frac{\sum_{v=0}^{Rlength} \sum_{u=0}^{Rwidth} R(u,v).S(c+u,l+v)}{\sqrt{\sum_{v=0}^{Rlength}.\sum_{u=0}^{Rwidth} R^2(u,v).\sum_{v=0}^{Rlength}.\sum_{u=0}^{Rwidth} S^2(c+u,l+v)}} \tag{4}$$

Sum of Absolute Difference $SAD(c, l) =$

$$\sum_{v=0}^{Rlength} \sum_{u=0}^{Rwidth} |R(u, v) - S(c + u, l + v)| \tag{5}$$

Sum of Squared Difference $SSD(c, l) =$

$$\sum_{v=0}^{Rlength} \sum_{u=0}^{Rwidth}(R(u,v) - S(c+u, l+v))^2 \qquad\qquad (6)$$

III.   PROPOSED METHODOLOGY FOR DISPARITY ESTIMATION

Cameras should be placed in a single plane with slight horizontal shift. Image captured from cameras are the input to the proposed system. Before capturing the images camera calibrations has to be done properly .Because the image captured from both cameras should have same size, brightness and colour with slight position disparity for proper processing.
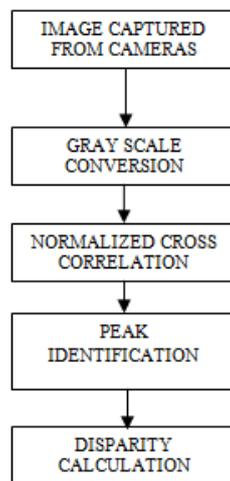

Figure 1 Disparity Estimation block diagram

A.   RGB to Gray Scale conversion

Image captured from cameras are in RGB format. Disparity estimation in RGB image complex because it has to deal with red, green and blue pixels. So before processing, RGB image is converted to Gray scale. Gray scale images are distinct from black and white images. This conversion will compress the image .So the complexity of processing can be reduced. RGB value are converted to intensity values by using equation,

I = (.2989 x R) + (.5870 x G) + (.1140 x B)   (7)

B.   Normalized Cross-Correlation

Normalized cross correlation is a similarity measure. For a similarity measure best match corresponds to maximum value of the function. Expression for normalized cross correlation is shown in equation (4).Due to changing intensity values and template size computational complexity increases. Normalized cross correlation requires more complex computations, but provides accurate result for disparity measurements. A comparative study conducted by analysing computation time and disparity [7] normalized correlation proved to be the best one.

C.   Peak Identification and Disparity Calculation

For a similarity measure, best match corresponds to maximum value of the computed output. So the next step is the peak or maximum value identification from computed results. The Horizontal and vertical coordinates of this peak value is the pixel shift in images required to get maximum correlation. Disparity value is computed by subtracting horizontal coordinates of maximum value from image size. Since disparity is the horizontal shift, vertical coordinates of peak value is neglected.

IV.   PROPOSED METHODOLOGY FOR IMAGE FUSION

Fusion of the left and right images can be done in two methods. Method I explains the technique for obtaining total visual field, which resembles combined vision of human eyes. Method II suggests a technique for obtaining common portion of stereo pairs, which is the binocular visual field of human vision.

A. Method I

If the computed disparity is equal to zero, either left or right image will be the fused image. If disparity value is greater than zero, disparity region in first (left) image is identified .This portion is fused in the left side of the second (right) image. An image representing total visual field is obtained as the final result.
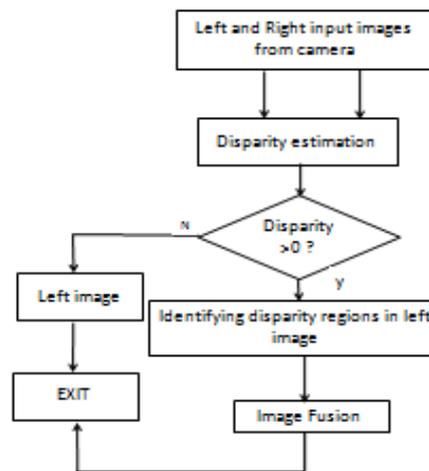

Figure 2 Proposed block diagram for method I

B. Method II

Using the computed disparity value disparity correction is performed in both images by removing left portion of the left image and right portion of the right image. This Disparity corrected images are fed to a pattern matching block. This block will give the common portion of the disparity corrected images.
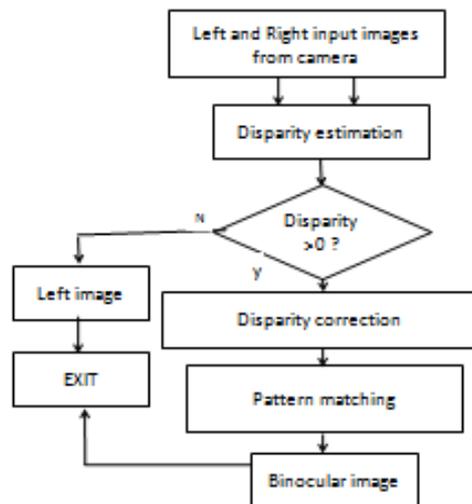

Figure 3 Proposed block diagram for method II

V. SIMULATION RESULTS

Simulation is done in matlab 2013a.Figure 4 shows the left and right input images with slight horizontal disparity.Normalized image cross correlation is carried out to find the disparity. Obtained disparity is 35 pixels for the given image pairs. Output obtained for image fusion is also shown in figure 5 and 6.

Figure 3 Left and Right input images

Figure 4 Image fusion by method I

Figure 4 Image fusion by method II

First method for image fusion would result in a large sized image corresponds to total visual field. Resultant image contains left portion of the left image fused with right image Second method resembles the binocular visual field of human vision, that is the common portion of the left and right image pairs.

VI. CONCLUSION
VII.

Disparity estimation is done by using Normalized cross correlation algorithm. Computed disparity can be used for calculating distance to object and also can be used for adjusting the focus of cameras. Receptive field position disparity estimation and image fusion was successfully completed by analyzing various left and right image pairs. Work can be extended to stereo pairs with receptive field position and receptive field phase disparity.

REFERENCES.

[1]   Jasmine banks, Mohammed Bennamoun, Peter Corke."Non_parametric techniques for fast and robust stereo matching ",Tencon '97. IEEE region 10 annual conference. Speech and image technologies for computing and telecommunications. Proceedings of IEEE (Volume:1) 1997, 365 - 368 vol.1

[2]   R.Andre ,E.Hendriks,J.Biemond,"Correspondence Estimation in Image Pairs",In proceedings of IEEE Signal processing Magzine,May1999.pp 29-45

[3]   Akiyuki Anzai, Izumi Ohzawa, and Ralph D. Freeman." Neural mechanisms underlying binocular fusion and stereopsis: Position vs. phase", Proceedings of the National Academy of Sciences of USA, May 13 1997, vol.94,no.10,pp5438-5443

[4]   Barbara Zitova,Jan Flusser,"Image registration methods: a survey",In proceedings of Image and Vision Computing 21 (2003) 977–1000 [2003]

[5]   Nuno.Roma,Jose Santos-Victor,and Jose Tome,"A Comparitive Analysis of cross correlation Matching Algorithms Using a Pyramidal Resolution approach,"2002

[6]   Luca.L et al,"Autonomous learning of active multi-scale binocular vision",In proceedings of Development and Learning and Epigenetic Robotics (ICDL), 2013 IEEE Third Joint International Conference, Publication Year: 2013 , Page(s): 1 – 6 ,IEEE 2013

[7]   Priyanka.V.K ,Vijayakmar.K,Sudheesh Madhavan,"Receptive field Position Disparity Estimation Using Cross correlation Algorithms", International Journal of Engineering Research and General Science Volume 3, Issue 1, January-February, 2015 ISSN 2091-2730

[8]   Gunter K. von Noorden, Emilio C. Campos "BINOCULAR VISION AND OCULAR MOTILITY: THEORY AND MANAGEMENT OF STRABISMUS".6th edition

BIOGRAPHY

Priyanka.V.K is a Post Graduate Student in TocH Institute of Science and Technology.She had completed her BTech in Electronics and communication Engineering from Cochin University College of Engineering Kuttanad. Currently pursuing her Masters in VLSI and Embedded systems. Area of interest includes Image processing ,artificial intelligence and robotics**.**.