# Receptive Field Position Disparity Estimation Using Cross-Correlation Algorithms

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

PG Scholar[1],Dept. of ECE, Toc H Institute of Science and Technology, priyankavk2011@gmail.com

**Abstract**—In human vision, retinal disparity is an important attribute which is the main reason for stereopsis. Mechanisms used for encoding binocular disparity through simple cells in striate cortex are difference in receptive field position and difference in receptive field phase. In the case of two dimensional objects position disparity has the major contribution in binocular fusion. Since human brains are more similar to correlators, cross correlation algorithms are the most simplest and natural method for disparity estimation. This paper is a comparative study of different cross correlation measures that can be used for the estimation of receptive field position disparity.

**Keywords**— Binocular disparity, computer vision, Correlation, Disparity map , Image registration, Normalized cross correlation ,Receptive field position disparity, Striate cortex.

## INTRODUCTION

Human beings are blessed with the attributes of binocular vision. 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 [12].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 two major reasons for binocular disparity [9]. In the case of two dimensional object RF position disparity is the major concern. 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.

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 important due to the horizontal shift between cameras. Disparity estimation is important, because it can be used for adjusting the focus of cameras, so that cameras can be converged to a single object. Disparity estimation is also useful for binocular fusion in computer vision systems and distance to object calculation. Most of the earlier works in this area were used for plotting disparity map, which gives the depth information of the object, and it is used to provide three dimensional visions to the system. This paper analyses the receptive field position disparity in two dimensional images. Different methods are there to compute receptive field position disparity. But correlation measures are the most simplest and natural method for a system, that emulate human vision processing. The ultimate goal of this paper is to compare different cross- correlation measures that can be used to estimate receptive field position disparity.

## RELATED WORKS

Andre R et al (1999) developed a paper on correspondence estimation in image pairs [2].Image pairs can be classified as spatial and temporal. Spatial images 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. Proposed paper analyses various techniques used for finding geometric correspondence and photometric correspondence of image pair. Additional constraints can be introduced to enhance the quality of the estimation results. The improvements discussed in this paper include the estimation of all pseudo correspondences, the incorporation of image restoration models, modeling of specular reflectivity of scene surfaces, the use of image sequences instead of pairs. 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 [5]. 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 were 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 drawbacks of the methods are also mentioned in the paper. Scopes in future researches are also discussed 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 non parametric 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.

Yan H et al (2008) designs a robust phase correlation based sub-pixel feature matching technique and its application in pixel-to-pixel image registration, DEM generation and motion estimation [7]. In particular, a median shift propagation (MSP) technique has been introduced to refine the unreliable motion estimation in image areas either featureless or subject to significant spectral changes. This paper proposes to use a phase fringe filter and a highly robust technique in the direct Fourier-based phase correlation algorithm for translational shift estimation. The local phase correlation based feature matching may fail in areas either featureless or with significant spectral differences between an image pair, a frequency based motion estimation technique and a motion flow refinement techniques are designed to improve the unreliable local motion estimates around these areas. Using the robust phase correlation based local matching algorithm, we can derive pixel-to-pixel image registration and disparity mapping for DEM generation in most synthetic and real images from different sensor.

Fengjun H et al (2013) developed a detailed study on a local matching algorithm and a semi-global matching algorithm which are representatives in stereo matching [11]. For the local matching they choose the SAD algorithm, while semi-global matching algorithm takes the SGBM algorithm. Experiments show that when the SAD window is small or large, there will be larger matching errors and low precision in results. In SGBM algorithm, BT(Birchfield and Carlo Tomasi) algorithm is used in matching cost calculation and smooth restraints are added in the energy formula and the experiments show that SGBM stereo matching algorithm is much better than SAD algorithm since it has authority in real time. Matching algorithm aims to get a disparity map with denseness and high precision. But as a result of factors like shelter, the light change and lack of texture, there will be wrong matching points in disparity map. Matching algorithm can handle the problem of shelter, but requires large amount of calculation, further research is required to obtain precise stereo matching information and meets the real time requirement.

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 [10]. 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.

Wang D et al (2008) formulate a new energy function followed by the use of graph cuts to refine the disparity map which takes segment as node [8]. Firstly, the robust disparity plane fitting is modeled and the method of Singular Value Decomposition (SVD) is used to solve least square. To ensure reliable pixel sets for the segment, they filter out outliers through three main rules, namely; judging reliable area, cross-checking and measuring the distance between previous disparities to the computed disparity plane. Secondly, improved hierarchical clustering algorithm is applied to merge neighbour.

Martin S et al(2005) developed local cross correlation model for stereo correspondence[6].As the disparity gradient of a stimulus increases, observers' ability to solve the correspondence problem and thereby estimate the disparities becomes poorer. It finally fails altogether when the critical gradient reaches the disparity gradient limit. They investigated the cause of the disparity-gradient limit as a part of this work, and developed a local cross-correlator similar to ones proposed in the computer vision literature and similar to the disparity-energy model of neurons in cortex V1. The cross-correlator exhibits poorer performance as the disparity gradient increases. To establish this they conducted a psychophysical experiment in which observers were presented sawtooth waveforms defined by disparity. Thus, human observers and a cross-correlators exhibit similar behavior, which insist humans to use such an algorithm to estimate disparity. As a result, disparity estimation is done with local estimates of constant disparity, which places a restriction on the highest possible stereo resolution.

Nuno R et al (2000) developed a comparative analysis of the performance and characteristics of a set of similarity measure algorithms proposed in the past few years [4]. The presented analysis was focused on the study of final matching error and computational load of the considered correlation functions. The studied similarity function constitute asset of different cross correlation based matching algorithms. 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. The advantages of using pyramidal resolution approach was also considered, to obtain fast running times and a few arithmetic operations with significant loss in final matching error.

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

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

$$\text{Simple Cross correlation } \mathrm{SCC(c,l)} = \sum_{v=0}^{Rlength} \sum_{u=0}^{Rwidth} \mathrm{R(u,v).S(c+u, l+v)} \qquad (3)$$

$$\text{Normalized Cross-correlation} \mathrm{NCC(c,l)} = \frac{\sum_{v=0}^{Rlength} \sum_{u=0}^{Rwidth} \mathrm{R(u,v).S(c+u,l+v)}}{\sqrt{\sum_{v=0}^{Rlength} . \sum_{u=0}^{Rwidth} \mathrm{R^2(u,v)}. \sum_{v=0}^{Rlength} . \sum_{u=0}^{Rwidth} \mathrm{S^2(c+u,l+v)}}} \qquad (4)$$

$$\text{Sum of Squared Difference } \mathrm{SSD(c,l)} = \sum_{v=0}^{Rlength} \sum_{u=0}^{Rwidth} \mathrm{(R(u,v) - S(c+u, l+v))^2} \qquad (5)$$

$$\text{Sum of Absolute Difference } \mathrm{SAD(c,l)} = \sum_{v=0}^{Rlength} \sum_{u=0}^{Rwidth} \mathrm{|R(u,v) - S(c+u, l+v)|} \qquad (6)$$

According to literature survey, it is found that human observers and a cross-correlators exhibit similar behavior. In a system that emulate human vision processing ,most simplest and accurate method is to use cross correlation algorithm as a disparity measure. Another challenging task was to choose a better cross correlation algorithm.

## OBJECTIVE OF THE PROJECT

Main objective of the paper is to compare different correlation measures found during literature survey and choose a best disparity measure that can be used for estimating receptive field position disparity in the computer vision systems. Comparison is carried out by computing the disparity value and computation time.

## PROPOSED METHODOGY

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. For analyzing various cross correlation algorithms Cameras are arranged in a setup to get a disparity of 100 pixels. Proposed methodology is shown in Figure 1.

## RGB TO GRAY SCALE CONVERSION

Image captured from cameras are in RGB format. Disparity estimation in RGB image is 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 one-bit bidirectional black and white images. Gray scale images have many shades of gray in between. This conversion will compress the image .So the complexity of processing can be reduced. RGB value are converted to intensity values by using equation,

$$I = (.2989 \text{ x } R) + (.5870 \text{ x } G) + (.1140 \text{ x } B) \qquad (7)$$

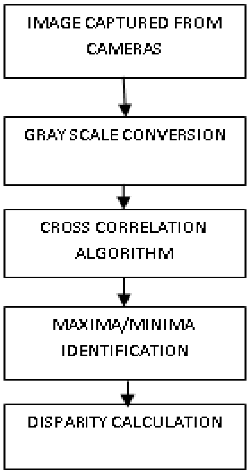


Figure 1 Methodology for the system

## CROSS CORRELATION ALGORITHMS

Disparity can be computed by using similarity measures and disparity measures. Mainly different correlation functions can be used for this. Functions like SCC, NCC are some of the similarity measures used. SSD and SAD are some of the dissimilarity measures. To find a best cross correlation algorithm that meets the requirement of proposed system an analysis is performed by testing correlation measures like SCC, NCC, SSD and SAD. Computation time and disparity value are the various parameters used for analysis.

## MAXIMA/MINIMA IDENTIFICATION

For a similarity measures the best match is obtained when the returned value is maximum and for dissimilarity measure best match corresponds to minimum value. Horizontal and vertical coordinates of this maxima/minima value is the pixel shift in images required to get maximum correlation. Disparity value is computed by subtracting horizontal coordinates of maximum value of the disparity estimation algorithm from the size of the image. Since disparity is the horizontal shift, vertical coordinates of peak value is neglected.

## SIMULATION RESULTS

Simulation is done in matlab 2013a. Figure 2 shows the input images with a disparity of 100 pixels. Both images are analyzed using correlation measures like SCC, NCC, SAD, and SSD .The figure 3 shows the graph plotted by analyzing different cross correlation measures.

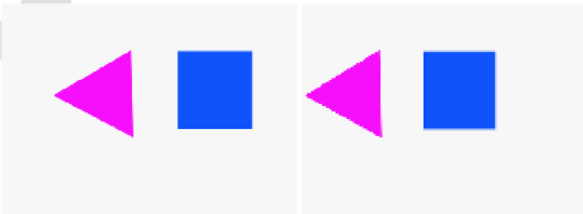

Figure 2. Input images with disparity of 100 pixels

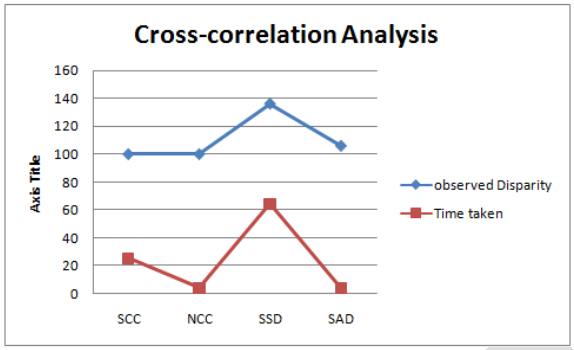


Figure 3.Cross correlation Analysis

ACKNOWLEDGMENT

First of all, I am grateful to God Almighty, for showering His blessings upon me for making me capable of doing this project. I am deeply indebted to the Management of Toc H Institute of Science and Technology and Department of Electronics and Communication for their whole hearted support.

## CONCLUSION

According to the simulation results, normalized correlation gave accurate and best result comparing to other measures. Sum of absolute difference correlation require less computation time. Sum of squared difference correlation is more sensitive to variations. Simple cross correlation also provides exact results, but computation time is comparatively large. Results can be modified by analyzing other parameters like error in computation, variations with colour, brightness, contrast etc. Other correlation measures like ZNCC(Zero Mean Normalized Cross Correlation), LSAD (Locally Scaled Sum of Absolute Difference), ZSAD (Zero Mean Sum of Absolute Difference) etc can be added in the comparative study to choose a best cross correlation measures. Also combination of different correlation measures can be used for improving results.

## REFERENCES:

[1] Banks.J et al,"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.94no.10,pp5438-5443

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

[5] Zitova.B et al,"Image registration methods: a survey",In proceedings of Image and Vision Computing 21 (2003) 977–1000 [2003]

[6] Martin S. Banks et al," Local cross-correlation model of stereo correspondence", Proc. SPIE 5666, Human Vision and Electronic Imaging X, 53 (March 18, 2005)

[7] Hongshi yan, jian guo liu,"Robust phase correlation based feature matchinig for image co- registration and dem generation",In proceedings of The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B7. Beijing 2008

[8] Daolei wang & kah bin lim:,"A new segment-based stereo matching using graph cuts",In proceedings of National university of Singapore.,Department of mechanical engineering. SETN 2008, LANI 5138, pp. 365-370, 2008

[9] Randolph Blake, Hugh Wilson,"Review Binocular vision", ELSEVIER,Vision Research 51 (2011) 754–770

[10] 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

[11] Fengjun HU & yanwei Zhao,"comparative research of matching algorithms for stereo vision,"In proceedings of Journal of computational information systems , 9: 13 (2013) 5457–5465

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