ISSN 2395-1621

# Disparity Based Depth Image Acquisition Using SAD Algorithm

#1Arshiya Mulani, #2Ashika Pande, #3Priyanka Savat,
#4Prof.V.G.Puranik

2ashikapande11@gmail.com

#123B.E Student, E&TC Department,
#4Asst.Professor, E&TC Department

JSPM's BSIOTR, Pune, India.

**ABSTRACT**

**In this paper, a system proposed for obtaining depth information from stereo images. Using Sum of Absolute Difference (SAD) algorithm the depth information of image is calculated. Disparity mapping concept is used for estimating the depth of image. Disparity is calculated by using a mat lab software. After estimating the disparity the calculation part of depth is performed on FPGA (Field programming gate array) kit. The process of image acquisition, image filteration and object detection are also discussed in this paper.**

**Keywords-Stereo vision, SAD algorithm, Disparity, FPGA, Depth estimation.**

ARTICLE INFO

**Article History**

Received: 28th March 2017

Received in revised form :

28th March 2017

Accepted: 30th March 2017

**Published online :**

**30th March 2017**

## I. INTRODUCTION

The perception of 3-D visual is becomes more popular due to increase in stereoscopic television and 3-D cinema. If an image is defined with depth information along with hardware such technique becoming popular for practical application .The most common applications are games, movies, consumer electronics products[1].The most commonly method, the perception of 3D is often produced in stereoscopic vision. However, this type of technology has never been able to find out the actual distance of space and objects[5]. If this technology would be applied to autonomous robots, vehicles and pedestrian detection, or security systems, it requires being able to perceive the actual distance of space and objects. With the help of Microsoft Kinect,we see the some application. However, instead of hardware the depth data processing and the computation remains in Xbox 360 software.

The main aim of this paper is to implement a very efficient and cost-effective real-time video depth camera using two identical cameras entirely processed in hardware[3].For this project the limited resources are required, so that we aim to provide the highest resolution and depth image quality while satisfying the limits in FPGA.Normally, image processing techniques are commonly used for this purpose. The principle is to capture images with two cameras, and to verify by comparing the same object in two images through stereo matching algorithm[8]. Then, according to the triangular relationship that the imaging of the object has constituted in two cameras, the three dimensional information (3D coordinates, relative distance) can be estimated, thereby achieve depth perception. A stereo vision is the process of estimating the depth from their change in position between two or more images and such depth information is used to a home assistant robot for intelligent navigation. The stereo depth extraction calculates the disparity between right and left camera image[8].
The method in is for matching low-texture region. In this system Yang proposed an epipolar distance transform which helps estimate planar 3D structure at low texture areas in the form of distance along the epipolar line[2]. In they use SAD algorithm for matching stereo images in autonomous vehicle. Curve fitting tool(cftool) gives reference data to the autonomous vehicle to navigate and avoiding obstacle. Dong proposed a hardware based on depth extraction method for intelligent home assistant robot for calculating the sum of absolute differences and pipelined depth calculation for real-time operation[4].

## II. PROPOSED WORK

A. *Block Diagram*

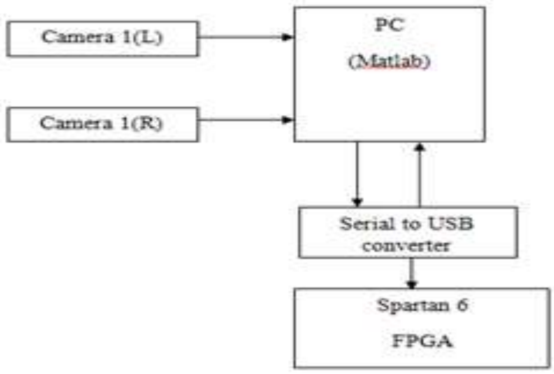



Fig 1. System Block Diagram

## III. PROPOSED SYSTEM

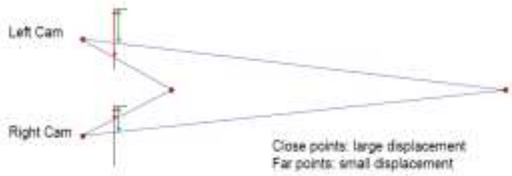
A) SAD

SAD is nothing but Summation of Absolute Difference, which we calculated by the following formula. It is the difference between the two images which we have captured by two cameras.

$$SAD = \sum_{x \in N} |L(x) - R(x)|$$

B) Disparity Calculation

The disparityis measure of displacement of a point between the two images.




1. Depth estimation from stereo cameras

When we look 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 parallax, and it can be exploited to extract geometrical information from a scene[6]. From that multiple captures of same scene from different viewpoints, it is possible to estimate the distance of the objects, i.e. the depthof the scene. By tracking the displacement of points between the alternate images, the distance of those points from the camera can be determined. Different disparities between the points in the two images of a stereo pair are the result of parallax. When a point in the scene is projected onto the image planes of two horizontally displaced cameras, the closer the point is to the camera baseline, the more difference is observed in its relative location on the image planes. Stereo matching aims to identify the corresponding points and retrieve their displacement to reconstruct the geometry of the scene as a depth map. Stereo matching has traditionally been used in machine vision e.g. to make machines aware of the surrounding environment in different applications. More recently it has been adapted to produce and transmit content for 3D televisions, especially for multi view displays where it can save significant bandwidth as opposed to sending all required views separately.

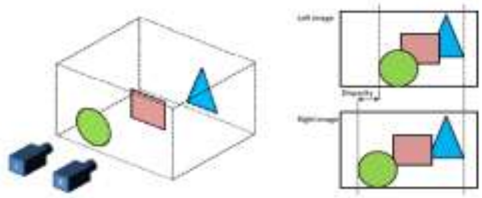


Fig 2. Depth estimation from stereo cameras

It also allows scaling of 3D content for different sizes and types of displays. As a passive method, stereo matching does not have to rely on explicitly transmitted and recorded signals such as infrared or lasers, which experience significant problems when dealing with outdoor scenes or moving objects.

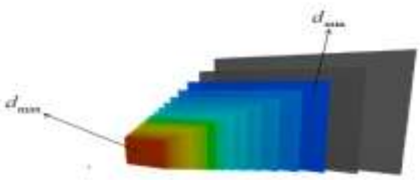

Fig. Steps in stereo matching

2. Disparity map quality

The quality of estimated disparity is commonly measured in bad pixels – the percentage of pixels in the estimated disparity that have different values than in the ground truth. Small errors can be allowed by requiring the estimate to be within a given threshold.

Close points are bright and far ones are dark[10].


Disparity plane



3. Disparity estimation

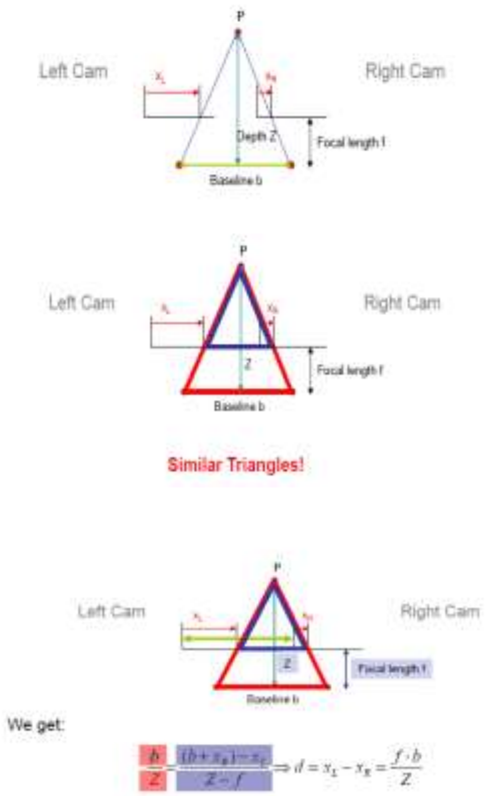
The whole concept of stereo matching is based on finding correspondences between the input images. In this paper correspondence between two points is determined by inspecting the pixel neighbourhood N around both points[7]. The pairing that has the lowest sum of absolute differences is selected as a corresponding point pair. In practice, a matching block is located for each pixel in an image. The relative difference in the location of the points on the image planes is the disparity of that point. Due to the assumption of being constrained into a 1-dimensional search space, these disparities can be represented as a 2D disparity map which is the same size as the image. Disparity of a point is closely related to the depth of the point. This is essentially a block matching scheme familiar from video compression (a.k.a. motion estimation), although in this application the search space can be constrained to be horizontal only (with certain assumptions). The matching block size is one of the most important parameters that affect the outcome of the estimation. Smaller blocks can match finer detail, but are more prone to errors, while large blocks are more robust, but destroy detail.

4. Depth

While the disparity map estimated from the stereo pair already distinguishes between objects at different distances, disparity is not the same as depth[9]. The relation depends on the camera configuration.

Relationship of disparity to depth

**We found the depth by comparing with disparity which is shown as below**




$$d = \frac{b \cdot f}{Z}$$

Using this relationship we can draw important conclusions. Disparity values are inverse proportional to the depth of a point *Z*. Far points have low disparity (for example the horizon has disparity of zero).

Close points have a high disparity

The disparity is proportional to the baseline *b*. The larger the baseline, the higher the disparity. The disparity resolution scales linearly with imager resolution. High resolutions allow accurate disparity measurements

## IV. APPLICATION

1.We can make Autonomous Robots using this technology.
2.Vehicles and pedestrian detection also can be done by this technique.
3.Optical scanner and image readerfor reading images and decoding optical information including one and two dimensional symbologies at variable depth of field.
4.Computer-Generated Holography as Generic Display Technology.for a wide range of display types,including 2D,stereoscopic,autostereoscopic,volumetric and true 3D imaging.
5.Real-time fusion of low-light CCD and uncooled IR imagery for color night vision.
6.Thefusion architectures that match opponent-sensor contrast to human opponent-color processing will yield fused image products of high image quality and utility.
7.Interactive video display system.A device allows easy and unencumbered interaction between a person and a computer display system using the person's (or another object's) movement and position as input to the computer.

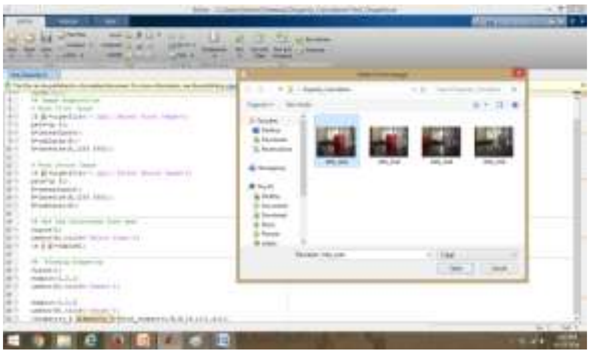
## V. PRIMARY RESULTS

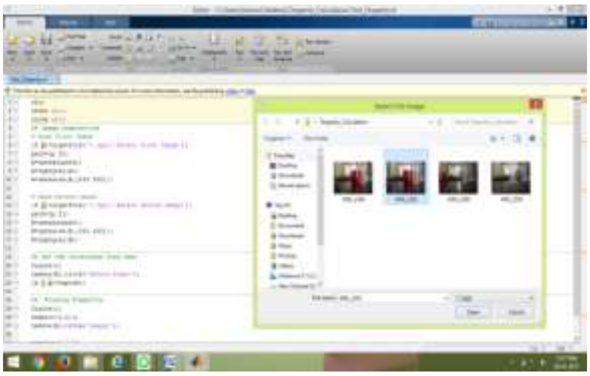
In this paper, we have presented a new method is proposed to capture image.The principle is to capture images with two cameras, and to verify by comparing the same object in two images through stereo matching algorithm. Then, according to the triangular relationship that the imaging of the object has constituted in two cameras, the three dimensional information (3D coordinates, relative distance) can be estimated, thereby achieve depth perception. A stereo vision is the process of estimating the depth from their change in position between two or more images and such depth information is used to a home assistant robot for intelligent navigation. The stereo depth extraction calculates the disparity between right and left camera image pair to estimate distance and position in image space. While stereo matching, the left and right images require computing through algorithm and the depth information would be available.
Now our Experiments on Real-Time images demonstrated on Mat-lab Simulation. Firstly we write a code on Mat-lab for finding a disparity between two images. In this code there are three major parts:

A. Image Acquisition

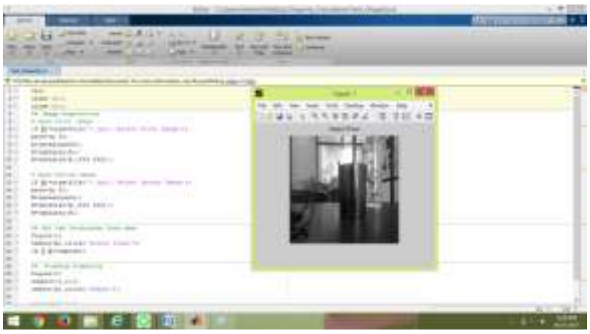
In image acquisition include firstly take two images and read that images is shown in below fig.

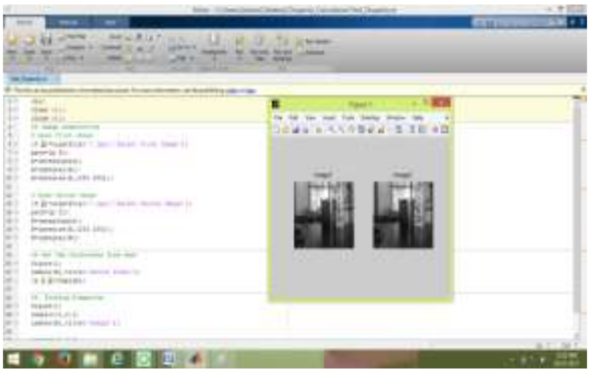




B. Get Co-Ordinate Form User

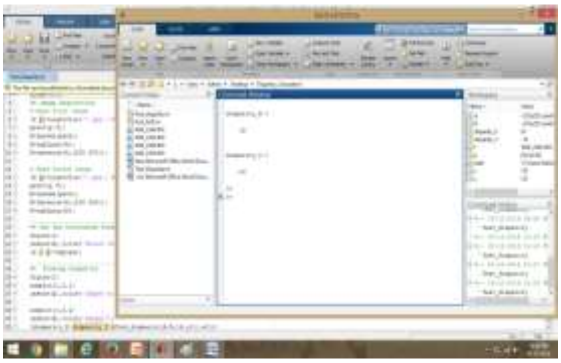
It includes the Co-ordinate between two images and form the new 3D images.





C. Finding Disparity

The disparity measures the displacement of a point between the two images and calculates the Disparity is as shown in below fig.



## VI. DISCUSSION

In this paper after analyzing advantages and drawbacks of some methods a new method is proposed to capture image. Here we use two cameras for capturing a depth image.Because the amount of computation is too large, disparity map is mostly computed from the existing image in a general computer processor. If upgrading image resolution, the processor would be unable to load such a large amount of computation. This hardware Intel MMX, Graph Process Units, and Digital Signal Processor. Whereas MMX and DSP platform is reduce the computation time but consume excessive power. So here we use FPGA to consume less power.

## REFERENCES

[1] Chien-Chung Wu,"The Development and Implementation of a Real-time Depth Image Capturing System using SoC FPGA",2016 30th International Conference on Advanced Information Networking and Applications Workshops.

[2] Q. Yang and N. Ahuja, "Stereo matching using epipolar distance transform," in IEEE Transactions on Image Processing, Vol. 21, No. 10,October 2012.

[3] R. A. Hamzah, R. A. Rahim, Z. M. Noh, "Sum of absolute differences algorithm in stereo correspondence problem for stereo matching in computer vision application," in Computer Science and Information Technology (ICCSIT), pp. 652-657,2010.

[4] S. K. Dong, S. L. Sang, H. C. Byeong, "A real-time stereo depth extraction hardware for intelligent home assistant robot," in IEEE Transactions on Consumer Electronics, Vol. 56, No. 3, August 2010,pp. 1782-1788.

[5] J. Zhao, J. Li Chai, G. Z. Men, "A fast quasi-dense matching method," in International Asia Symposium on Intelligent Interaction and Affective Computing,December 2009,pp. 100-103.

[6] X. T. Wang and X. Bo. Wang, "FPGA based parallel architectures for normalized cross-correlation," The 1st

International Conference on Information Science and Engineering (ICISE2009),December 2009, pp. 225 – 229.

[7] D. Scharstein and R. Szeliski, “A taxonomy and evaluation of dense two-frame stereo correspondence algorithms,” Int. J. Comput. Vision, vol. 47, nos. 1–3, Apr.–Jun. 2002, pp. 7–42.

[8] J. Sun, N. Zheng, and H. Y. Shum, “Stereo matching using belief propagation,” Pattern Anal. Mach. Intell., vol. 25, no. 7,2003, pp. 787–800.

[9] Y. Furukawa, B. Curless, S. Seitz, and R. Szeliski, “Manhattan-world stereo and Surface Reconstruction,” in Proc. IEEE Compute.Vision Pattern Recogn.Conf,Jun. 2009, pp.1422–1429

[10] T. Werner and A. Zisserman, “New techniques for automated architectural reconstruction from photographs,” in Proc. Eur. Conf. Comput. Vision,2002, pp. 541–555.