Obstacle Detection and Avoidance using Stereo Vision System with Region of Interest (ROI) on FPGA

Mr. Rohit P. Sadolikar[1], Prof. P. C. Bhaskar[2]

[1,2]Department of Technology, Shivaji University, Kolhapur-416004, Maharashtra, India.

Abstract— **Stereo vision is an area of study in the field of Machine vision that recreate the human vision system by using two or more 2- dimensional views of the same scene to obtain 3D distance information about a scene. The ability of a machine to capture 3D information from the real world in a similar fashion to a human being is called Stereo Vision System. A proposed system can detect an object and avoid it by using stereo vision system by implementing the Sum of Absolute Difference Algorithm (SAD). Sum of Absolute Difference Algorithm (SAD) can create the Disparity Image and from the amount of difference in disparity image, the object is near or far is decided. The system contains two cameras positioned on front side of autonomous system at the same height capturing the respective images from slightly different angle of the same view or scene. In application as obstacle detection fast localisation of object is important for that a new approach is proposed. SAD algorithm with Region of interest (ROI) is implemented on Spartan 3E FPGA system.**

Keywords— *Stereo Vision, Region of interest (ROI), Spartan 3E FPGA, SAD, Disparity Image.*

I. INTRODUCTION

Recent development in unmanned space travel, agricultural automation, Archaeological exploration and evolution of military devices shows that the demand for Unmanned Ground Vehicles is strongly increasing. Still most of the implementations appear more like enhanced remote-controlled cars [20] than autonomously acting and decision-taking, tasks "intelligent", and robots. As tele-operation is difficult and expensive, costs and operation times shall be minimised by the introduction of autonomous vehicles. The application of Obstacle detection is carried out with stereo vision system and SAD Algorithm. Stereo vision is resembles with human vision system. Like human vision system it has two cameras position side-by-side on front of autonomous system. Each camera captures respective images and these two images are compared for differences with the help of template matching algorithm that is Sum of Difference Algorithm (SAD). The template matching algorithm is used to get the disparity image which gives difference by comparing both left and right camera capture views. If difference is more the object is near and if difference is less, object is far away from system.

Generally the 3D information is reconstructed in stereo vision system using template matching algorithm which is slower and highly computational. The highly computational reconstruction of full 3D scene is unnecessary in application such as 'Obstacle Detection'. A fast localisation of Object is important for obstacle detection in autonomous navigation framework. A new approach is proposed in this system by deciding Region of Interest from the two images for localisation of obstacle. The system has two cameras situated on front side and as the autonomous vehicle moving forward direction, the centre part of both the images should taken into account for localization of object.
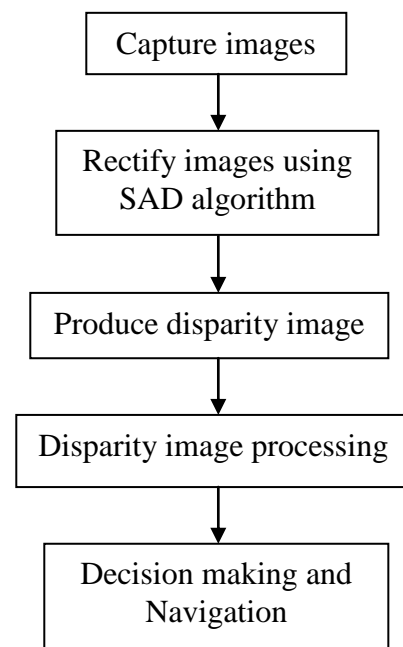

Figure 1: Flow of Proposed system

The flow chart gives the information of steps executing in system: firstly two cameras situated on front side of the system are triggered using MATLAB installed on laptop. As per requirements of the stereo vision system the system must moving and taking pictures. Before transferring these pictures to FPGA for further processing, the Region of Interest (ROI) of these pictures are consider. These modified pictures are simultaneously transferring from laptop to FPGA. On the FPGA side, difference of the left and right modified images is determined with the Sum of Absolute Difference Algorithm

(SAD). The SAD Algorithm is implemented using VHDL which is Hardware description language. The Total SAD gives a Disparity image from the difference of both the images. Depending on the difference from disparity image, the depth of object is determined. An object is presented or not and also the distance of object from system that is object is near or far from the system is determined. If the difference in disparity image if less then, the object is far away from system and if difference is more then, object is near to the system.

II.   BACKGROUND & RELATED WORK:

Stereo vision refers to the issue of determining the 3-D structure of a scene from two or more images taken from distinct viewpoints [2].In paper [3] and [4] Some interesting details about the developed sensor systems and proposed detection and avoidance algorithms can be found. The obstacle avoidance sensor systems found in literature can generally be divided into two main categories. The first category involves the use of ultrasonic sensors. They are simple to implement and can detect obstacles reliably. On the other hand, the second category involves vision-based sensor systems. This category can be further divided into the stereo vision system (which is applied to the detection of objects in 3D) and the laser range sensors which can be applied to the detection of obstacles both in 2D and 3D, but can barely used for real-time detection [5]. A number of researchers over the years have used sonars [8] and lasers [6] for obstacle avoidance. Departing from the traditional sonar ring, Nourbakhsh et al. [7], [8] showed that a non-conventional arrangement of sonars including angled sensors leads to more robust behaviour and, in particular, prevents decapitation (i.e., collision with an object at a height above the ring of sensors).Other research, with Stereo Vision systems mounted in fixed positions, exist to extend capabilities of security monitoring systems, and improve human face-recognition and tracking algorithms [9].

Stereo matching has computational complexity and the matching technique directly affects the accuracy of disparity map, various matching algorithms have been devised. They can be classified as area-based [12], feature-based [13] and phase-based [14]. In the area-based approach, the correspondence problem is solved by matching image intensity patterns. [15] Employed the area-based approaches such as dynamic programming, sum of absolute differences (SAD) algorithm and Census transform respectively on FPGAs. An interesting but very computationally demanding local method is presented in [16]. It uses varying weights for the pixels in a given support window, based on their colour similarity and geometric proximity. An interesting method is proposed in [17], an adaptive window size is used in SAD algorithm for template matching for fast processing. The algorithm reported in [19] achieves almost real-time performance. It is once more based on SAD but the correlation window size is adaptively chosen for each region of the picture. Apart from that, a left to right consistency check and a median filter are utilized.

III.   STEREO VISION

Stereo Vision is a well-known ranging method as it resembles the basic mechanism of Human eye. Human vision system has two eyes located side-by-side in the front of their head. Thanks to this positioning, each eye takes a view of the same area from a slightly different angle. Each eye captures its own view and the two separate images are sent on to the brain for processing. The small differences between the two images add up to a big difference in the final picture. It is a three-dimensional *stereo* picture. Stereo vision is an area of study in the field of Machine vision that attempts to recreate the human vision system by using two or more 2D views of the same scene to derive depth information about the scene by using disparity image.

Stereo vision for mobile robots has some specific requirements. The first requirement is that the algorithm to compare the two images and get the resulting disparity image. Second, mobile robots tend to move around and take pictures. This means the stereo algorithm needs to handle image sequences. This provides the algorithm a better chance to get the correct disparity images or refine it. Third, mobile robots are typically moving on a plane ground. To avoid obstacles on the ground, the disparity image can be calculated based on the plane ground. Fourth, stereo is not the only sensor on a mobile robot but Fusion of multiple sensors.

IV.   SUM OF ABSOLUTE DIFFERENCE ALGORITHM

SAD is used as a similarity measure in case of block matching in several image-processing applications. SAD takes the absolute value of the difference between each pixel in the reference image block and the corresponding pixel in the target image block. The differences are accumulated to create a metric of block similarity. In order to find the similarity of a reference image (generally smaller in size) in a bigger image, known as the target image, the reference is placed on the image and the SAD is calculated; it is then moved to the next location on the image and the SAD is again calculated. This is repeated until the entire image SAD is calculated. The greater the value of the SAD operation the lesser will be the similarity between the reference and the specific part of the image for which the SAD is specified. Because of the simplicity of the SAD it is considered to be a very fast and effective metric for similarity measure calculation in images. It takes into account all the pixels that are present in a window without attaching any specific bias to any values. Since the calculation of SAD for one block in an image is done without affecting the calculation for other blocks in its vicinity, it can calculate the SAD in parallel.

The Sums of Absolute Differences (SAD) algorithm takes the pixel data from a stereo pair and iterates through its contents in a linear fashion to find the disparity of the two images. It determines the disparity of a single pixel by searching through a scan window around each pixel within its disparity range and finding the scan window with the lowest absolute difference between each corresponding pixel for its red, green, and blue values.

V.   PROPOSED SYSTEM

The proposed system of obstacle detection and avoidance has following approach:

The system consist two cameras situated on front side, at same height and with slight different angle. Both the cameras are operated by MATLAB on laptop. Then through laptop this two images send on Spartan 3E FPGA Board where template matching SAD algorithm is implemented. According to figure the left part of right image and right part of left image is unwanted so the center part of image is considered for obstacle detection. As the system moves in forward direction, the centre part of both images consider as a Region of Interest (ROI). Now this Region of Interest of images is compare by the SAD Algorithm and getting the 3D disparity image. This two captured images are transferred to FPGA for image processing where the Sum of Absolute Algorithm (SAD) is implemented.

$$SAD_{total}(x,y,r,c) = \sum_{j=0}^{9} SAD9_j(x,y,r,c)$$

where $SAD9_j(x,y,r,c)$ defined as

$$SAD9_j(x,y,r,c) = \sum_{j=0}^{9} \left| L_{(x+i,y+j)} - R_{((x+r)+i,(y+c)+j)} \right|$$

i)   With $0 \leq x; y <$ frame size

ii)   With (r,c) being the motion vector

iii)   With L(x,y) being a reference frame   pixel at (x, y) of Left image

iv)   With R(x , y) being a target frame pixel at (x, y) of Right image

All data units L_i and R_i are considered to be unsigned 8 bits numbers.

Subtraction of two unsigned numbers (e.g., $A - B$) is performed by adding A with a bit inverted   B i.e. ($\bar{B} = 2n-1-B$) and adding a one:
$$A + (2n-1-B) + 1 = 2n+A-B.$$

Assuming that $B \leq A$, the resulting carry (2^n) of the addition can be ignored. The SAD_9 operation can be performed in three steps:

i)   Compute $(A_i - B_i)$ for all 9×9 pixel locations.

ii)   Determine which $(A_i - B_i)$'s are negative, i.e., when no carry was generated and compute $(B_i - A_i)$.

iii)   Add all 9 absolute values together.

The slight modification is done with the algorithm as follows:

Algorithm:

1)   Decide the Region of interest (ROI) of both reference image and target image and modified so.

2)   Iterate through the pixels of the left modify image that will be processed with the pixels of right modify image.
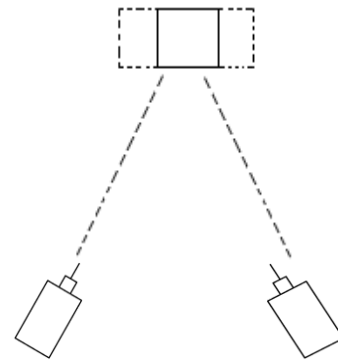

Figure 2: Stereo Vision with ROI

3) Take the absolute difference between pixels from both the images. Save the difference and again repeat it for next pixels.

4) Ensure that the compared pixel is not outside of the bounds of the right image or area of interest.

5) At the end iterate through all of the pixels within the image and get the sum of all absolute differences which gives the resulting disparity image.

This figure set indicates the left and right camera images and the third image is difference of two images taken by using Sum of Absolute Difference Algorithm.

Why Region of Interest:

In stereo vision system, a template matching algorithm is required for getting 3D disparity image which is useful for getting the surrounding information. For obstacle detection using stereo vision system, a fast processing and accurate algorithm is required. In previous papers different algorithms are mentioned for getting accurate information of objects. A template matching algorithm is comparing the two images and produces a 3D disparity image. But for obstacle detection, it is unnecessary to construct a full 3D disparity image which a time consuming process. For obstacle detection a fast localisation of object is important and for that the Region of Interest from the two images is required. Now if region is determine then only that much part of images are forwarded for template matching so that fast localisation of object is carried out.

An obstacle detection and avoidance is dependent on difference in disparity image. If the difference is exceed the threshold difference value then the object is detected. As soon as the object is detected the systems start changing direction to avoid it. A system takes around 40-50ms to detect and avoid the obstacle. The system detects obstacle within 30-40cm.

Hardware Requirement:

Device Utilization Summary (estimated values)			[-]
Logic Utilization	Used	Available	Utilization
Number of Slices	129	2448	5%
Number of Slice Flip Flops	82	4896	1%
Number of 4 input LUTs	250	4896	5%
Number of bonded IOBs	29	66	43%
Number of GCLKs	2	24	8%

Figure 3: FPGA Resource Utilization

The SAD algorithm used in proposed system is implemented on Spartan 3E FPGA XC3S250E. It contains 172 I/O's, 2k total slices. As mentioned in paper [18], a number of 4input LUTs utilization is 6% and number of Slices used is 7%. As shown in figure a number of 4-input LUTs utilization is 5% and number of slices used is 5%. So there is minimum utilization of resources of FPGA.

VI.   RESULT :

From the resulting images, a first image indicate Right camera image (Figure 4) and second image is Left camera image (figure 5).

Figure 4: Right Camera Image.

Figure 5: Left Camera Image.

Figure 6 and figure 7 indicate gray images with Region of Interest (ROI) of Left and Right cameras image. Figure 8 indicate the disparity image depending on disparity object in path is detected.
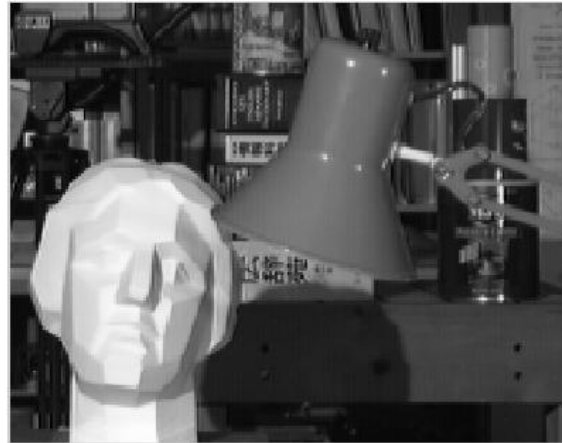

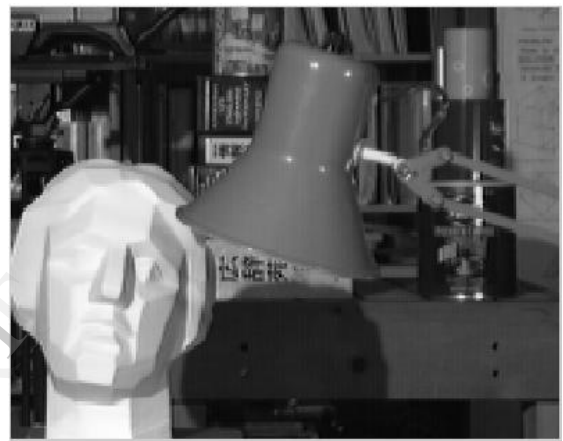Figure 6: Right image with ROI.

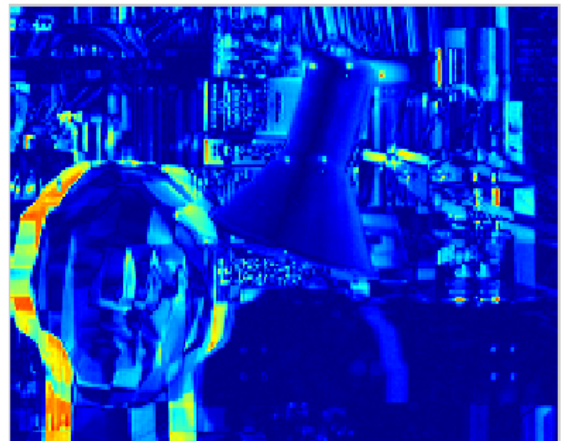Figure 7: Left image with ROI.

Figure 8: Disparity Image.

All the images are from MATLAB program of SAD Algorithm with ROI.

VII. CONCLUSION

The SAD algorithm is a sub method of an area based method and with the ROI; it can be much faster than other methods for determining the disparity from both the images. In this case, SAD algorithm has implemented on FPGA-based hardware architectures to process the image and determine object's depth information. For Obstacle detection application fast localization of obstacle is important and by using Region of Interest (ROI) with SAD algorithm in the obstacle detection application, fast localisation of object is carried out and image processing overhead is minimised. The proposed system also minimises resource utilization of FPGA.

REFERENCE

1) Borenstein, J., Koren, Y.: Real-time obstacle avoidance for fast mobile robots in cluttered environments. IEEE Transactions on Systems, Man, and Cybernetics 19(5), 1179–1187 (1990)

2) Ohya, A., Kosaka, A., Kak, A.: Vision-based navigation of mobile robot with obstacle avoidance by single camera vision and ultrasonic sensing. IEEE Transactions on Robotics and Automation 14(6), 969–978 (1998)

3) Vandorpe, J., Van Brussel, H., Xu, H.: Exact dynamic map building for a mobile robot using geometrical primitives produced by a 2d range finder. In: IEEE International Conference on Robotics and Automation, Minneapolis, USA, pp. 901–908 (1996)

4) Y., and Zelinsky, A. "An Algorithm for Real-time Stereo Vision Implementation of Head Pose and Gaze Direction Measurement", Nara Institute of Science and Technology 8916-5 Takayamacho, Ikoma-city, Nara, Japan, and The Australian National University.

5) Molton N., Se S., Brady J.M., Lee D., and Probert P. "A stereo vision-based aid for the visually impaired", Department of Engineering Science, University of Oxford, Oxford, OX1 3PJ, U.K, appears in *Image and Vision Computing* 16 (1998) 251–263.

6) Labayrade, R., Aubert, D., Tarel, J.P.: Real time obstacle detection in stereovision on non flat road geometry through "v-disparity" representation. In: IEEE Intelligent Vehicle Symposium, Versailles, France, vol. 2, pp. 646–651 (2002)

7) O. Brock and O. Khatib. High-speed navigation using the global dynamic window approach. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 1999.

8) I. Nourbakhsh. The sonars of Dervish. *The Robot Practitioner*, 1(4),1995.

9) I. Nourbakhsh, R. Powers, and S. Birchfield. Dervish: An officenavigating robot. *AI Magazine*, 16(2):53–60, 1995.

10) D. Marr and T. Poggio, "A Computational theory of Human Stereo Vision", Proceedings of the Royal Society of London – Series B: Biological Sciences, Vol. 204, pp. 301-328, 1979.

11) Matsumoto Y., and Zelinsky, A. "An Algorithm for Real-time Stereo Vision Implementation of Head Pose and Gaze Direction Measurement", Nara Institute of Science and Technology 8916-5 Takayamacho, Ikoma-city, Nara, Japan, and The Australian National University.

12) G. C. DeAngelis, I. Ohzawa, R. D. Freeman, "Depth is encoded in the visual cortex by a specialized receptive field structure*", Nature*, 11, 352(6331) pp. 156-159, 1991.

13) Sungchan Park, Hong Jeong, "High-speed parallel very large scale integration architecture for global stereo matching", Journal of Electron Imaging, Vol. 17, 2008

14) K. Ambrosch, M. Humenberger, W. Kubinger, A. Steininger, "Hardware implementation of an SAD based stereo vision algorithm", IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-6, 2007.

15) J. Woodfill and B. V. Herzen, "Real-Time Stereo Vision on the PARTS Reconfigurable Computer", IEEE Symposium on FPGA-Based Custom Computing Machines, pp. 201, 1997.

16) K.J. Yoon, I.S. Kweon, "Adaptive Support-Weight Approach for Correspondence Search", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 4, April 2006.

17) Chuen-Horng Lin "A Stereo Matching Algorithm Based on Adaptive Windows" International Journal of Electronic Commerce Studies Vol.3, No.1 , pp.21-34, 2012.

18) Saad Rehman, Rupert Young "An FPGA Based Generic Framework for High Speed Sum of Absolute Difference Implementation" European Journal of Scientific Research ISSN 1450-216X Vol.33 No.1 (2009), pp.6-29,2009.

19) Yoon, S., Park, S.K., Kang, S., Kwak, Y.K.: Fast correlation-based stereo matching with the reduction of systematic errors. Pattern Recognition Letters 26(14), 2221–2231 (2005).