International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-3 Issue-1, October 2013

Dense Stereo Correspondence Algorithm for Robotic Applications

Deepambika V.A, Arunlal S.L.

Abstract— ***Stereo vision, the passive sensing technique for inferring the three dimensional position of objects of a scene under study is having great applications in the field of machine vision, robotics, image analysis and image reconstruction. Robotics require computationally fast and easy to implement stereo vision algorithms that will provide reliable and accurate results under real time constraints. By using some similarity measure, the stereo correspondence, tries to find out the matching pixels or objects between left and right views of the scene. Since the focus is on real time application the local winner-take-all optimization in the disparity computation process is done in this study. The correspondence is done by using fast block matching Sum of Absolute Differences (SAD) algorithm .With the help of camera parameters and the disparity map obtained from this algorithm, the depth map of the scene under study is extracted by using the principle of triangulation. To simplify the correspondence search, rectified stereo image pairs are used as inputs.***

Index Terms— ***Stereo correspondence, Sum of Absolute Differences (SAD), Disparity, Depth.***

I. INTRODUCTION

The use of visual information to control a robot has been used extensively in many sectors for various applications. When humans grasp objects, they usually do so with the aid of vision. Visual information is used to locate and identify things, and to decide how they should be grasped. In a similar manner, machine vision can be used to coordinate a robotic arm [1].One method of obtaining depth information is by using a stereo vision system**.**

A. STEREO VISION

In computer stereo vision, two cameras are placed horizontally and the difference between the views is used to map the depth of the environment. Stereo vision is a widespread passive sensing technique for inferring the three dimensional position of objects from two or more simultaneous views of a scene. This system provide high resolution depth map and there is no interference with other sensor devices when multiple robots are present in the same environment. In this system the depth of information is not directly measured thought the system but it has to be

Manuscript published on 30 October 2013.
***** Correspondence Author (s)
Deepambika V. A*, Department of Electronics and Communication, LBS Institute of Science and Technology for Women, Trivandrum, (Kerala). India.
Arun Lal S.L Scientist, ISRO, Thiruvananthapuram, (Kerala). India.

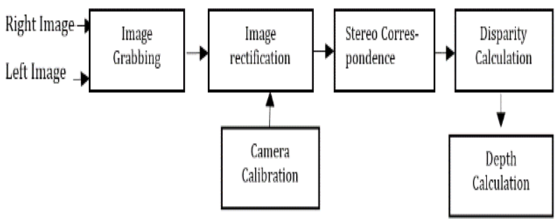


Fig.1 Stereo vision system

Explored from the binocular images with the help of disparity. Fig.1 shows the system overview.

Stereo vision systems do not need complex hardware; two coupled cameras are the minimal requirement. Cameras capture left and right view of the scene under study. *Rectification* transforms each image plane so that pairs of conjugate epipolar lines become collinear and parallel to one of the image axes. The rectified images can be thought of as acquired by a new stereo camera obtained by rotating the original cameras around the optical centre [2].By rectification a 2-D correspondence search is reduced to a 1-D search, typically y axis will be same for two rectified images. Therefore both cameras need to be calibrated first to get the camera parameters [3]. The knowledge of the camera parameters is used to rectify both images.

After rectification stereo correspondence algorithm will find the matching pixels (conjugate pair) of two given input images based on Lambertian surface assumption .The correspondence algorithm will provide disparity map of the scene. From this disparity values the depth of objects can be calculated by using camera parameters and triangulation principle.

Stereo correspondence, based on some similarity measure aims to find the best matching pixels of two input images. Based on accuracy and efficiency, the matching algorithms can be grouped two main categories-local methods and global methods [4]. Global methods are accurate but computational cost is higher due to their iterative nature. Therefore global approaches are not suitable for real time applications. They are used for applications such as precise 3D surface modelling, especially when dealing with object surfaces with complex reflectance behaviour and poor texture. Local methods can be of type area based or feature based approach. Feature based algorithms gives faster results which involves the extraction of edges and corners. These algorithms find the correspondence between some feature points of the stereo image pair, and usually give only sparse disparity maps. Area based approaches involve subdividing the whole view into sub regions and applying a photometric similarity measure to all regions. The disparity map obtained will be dense.

The most common pixel-based matching costs include sum of squared intensity differences (SSD) [4], and sum of absolute intensity differences (SAD) [5].Key to the real-time aspect of stereo is the use of algorithms based on SAD, correlation or basic variants. SAD is easier to compute than the other measures such as sum of square difference (SSD) and Normalized cross correlation (NCC) [6,7]. Of these the Sum of absolute Differences (SAD) function is the most commonly used cost matching function for real time applications. It has decent accuracy and computational speed.

II. SUM OF ABSOLUTE DIFFERENCES ALGORITHM

Sum of Absolute Differences (SAD) is a simple block matching algorithm. It is faster and suitable for real time applications and it gives descent disparity map. The correspondence is done by comparing blocks of pixels in each image. The dissimilarity is the sum of differences in pixel intensities and hence the name Sum of Absolute Differences. The match with the highest similarity, the SAD minimum wins. Disparity map will be dense. Computing the sum of differences between the pixel intensity values for a block around the pixel(x, y) is simple. Let f (i, j) be the intensity of the pixel at coordinates (i, j) in the reference image. For the second image, intensity g (i, j) is used, N defines the extend of block in either direction around the center pixel, the block has therefore width and height (2N+1). Here the rectified images are used as inputs the search is done along the x direction only as there is no y shift for the same pixels in the left and right images.

$$y = y_{LEFT} = y_{RIGHT} \qquad (1)$$

The variable‘d’isthe disparity, the displacement of corresponding matching points from one image to the other.

$$SAD(x, y, d) = \sum_{i=x-N}^{x+N} \sum_{j=y-N}^{y+N} |f_{i,j} - g_{i+d,j}| \qquad (2)$$

Each pixel in the left image is compared with every pixel on same epipolar line in right image. SAD values for all pixels are computed and the disparity is computed using winner-take-all (WTA) principle. The winner is simply the disparity associated with the SAD minimum, in ideal case SAD minimum will be zero.

$$disparity(x, y) = \min_{d \in D} \,(sad(x + d, y)) \qquad (3)$$

The Eqn. (3) describes the disparity for the block around the pixel(x, y), which varies from zero to *dmax,* where *dmax* represents the highest disparity value of the stereoscopic images called disparity range. Each block from the left image is matched to a block in the right image by shifting the left block over the search area of right image [2]. Difference in index values for the SAD minimum corresponds to disparity for that area. This is continued for each block until the disparity map is completely filled and the resulting image is known as Disparity Space Image (DSI).

A. Performance Improvements

By applying a simple sliding window scheme, the algorithm will speed up. This works in the horizontal as well as vertical direction.

Horizontal sliding window

As shown in Fig.2, a window is moved from one column to next from left to the right across a horizontal scan line. The computed rows are still the same, except for the pixel left and right.Fig.5.1 shows a window that is moved one column to the right across a horizontal scan line

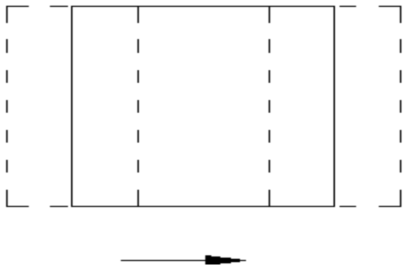


Fig. 2: Horizontal Sliding Window: window slides across horizontal scan line.

Vertical sliding window

The sliding window scheme is also being applied when moving to the next horizontal line as shown in vertical sliding window Fig.3 After completing a pass on one horizontal line, the window moves to the line below. The computed columns are still the same, except for the pixel above and below.

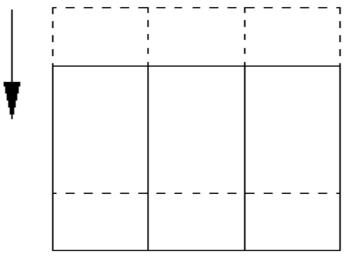


Fig.3: Vertical Sliding Window- window slides across vertical scan line

The results of the matching cost computation (SAD) comprise the disparity space image (DSI).

B. Depth Computation

From the disparity values obtained depth (Z) can be estimated by using triangulation and camera parameters. The depth Z is given by

$$Z = \frac{f.b}{x_l - x_r} = \frac{f.b}{d} \qquad (4)$$

Where d is the disparity f, focal length of the camera, b the baseline distance.

C. Steps of SAD Algorithm

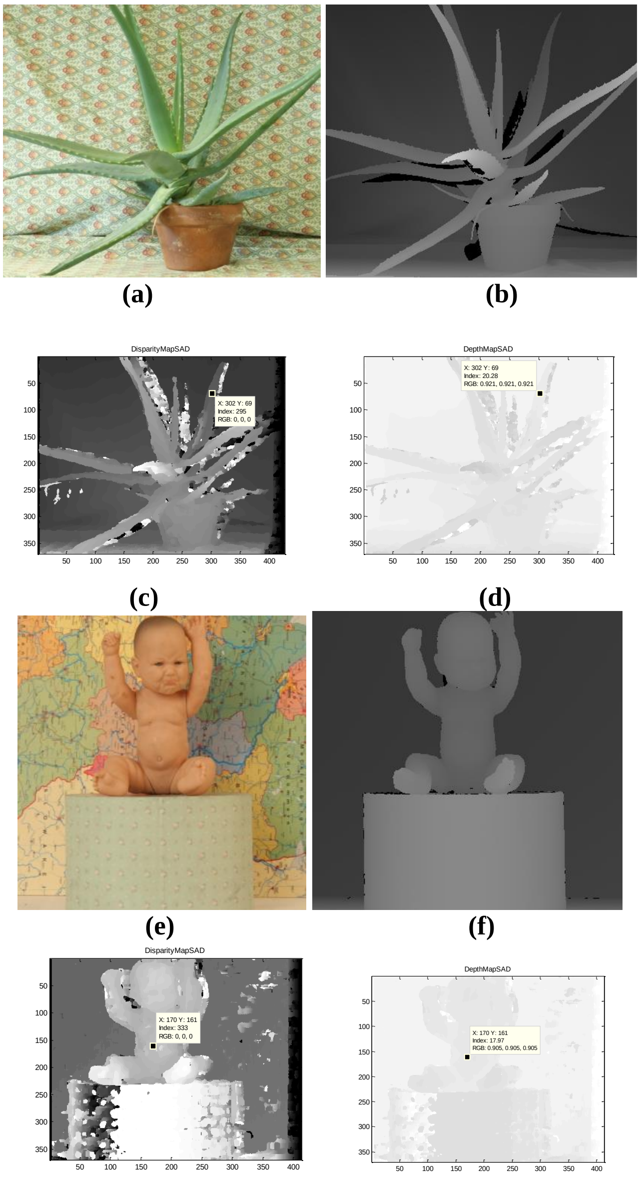
- Read the Rectified input images.
- Select a square window from left image
- Slide this window over right image and compute SAD minimum
- Obtain the difference in index values corresponding to the minimum value
- Aggregate this index values (disparity) for the entire image
- Obtain depth value using triangulation.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-3 Issue-1, October 2013

III. RESULTS AND ANALYSIS

To verify the SAD algorithm, the rectified stereo image pairs from Middlebury stereo data set [8] are used. Each dataset of this database is made up of a pair of stereo images and the corresponding disparity map. The experiments for this study used the Matlab . For the study and validation of our algorithms Aloe, Baby, Monopoly image pairs and their ground truth data are used. Fig.2 below shows left view of input images and its disparity map and the depth maps. Fixed window size of 9X9 pixels is used for obtaining the depth map and the corresponding depth value. The disparity values of SAD algorithm are mapped to disparity mapping. To visualize the depth value for any point in an input image, a user interface is developed. A mouse click on any point on the image will show the corresponding depth value.



(a) (b) (c) (d) (e) (f) (g) (h)

Fig.4 Dense disparity map & depth maps for Middlebury images and their corresponding Ground truth disparity: Input images-(a) Aloe-Right view (e) Baby-Right view. (b) & (f) Ground truth disparities. (c)& (g) Disparity map from SAD (d) & (h) corresponding depth map Depth map.

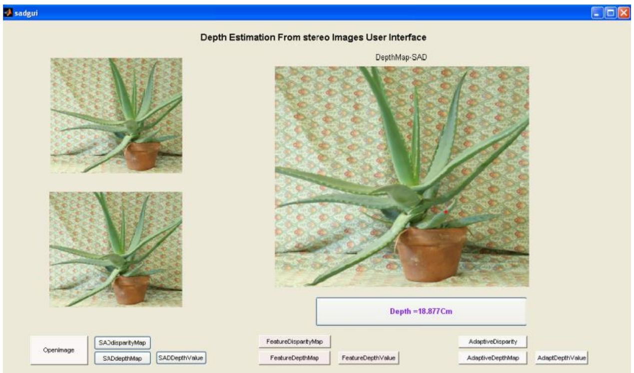


Fig.5 User Interface – Depth estimation from Stereo vision

A. Evaluation of results

To evaluate the performance of a stereo algorithm or the effects of varying some of its parameters, we need a quantitative way to estimate the quality of the computed correspondences [4].Compute error statistics with respect to some ground truth data compute the following quality measure based on known ground truth data:

1. RMS (root-mean-squared) error (measured in disparity units) between the computed depth map $d_C(x, y)$ and the ground truth map $d_T(x, y)$, i.e.,

$$E = \left(\frac{1}{N}\sum_{(x,y)} |d_C(x,y) - d_T(x,y)|^2\right)^{\frac{1}{2}} \qquad (6)$$

where N is the total number of pixels.

Table 1. RMS Error of PNG images.

widow size	% RMS Error PNG Images		
	Aloe	Baby	Bowl
3	1.987368	2.225251	4.948374
5	1.792845	1.765533	4.403559
7	1.785006	1.576174	4.114675
9	1.797745	1.49025	3.910918
11	1.812445	1.448657	3.783502
13	1.837924	1.445373	3.700022
15	1.865363	1.455771	3.65993
17	1.892802	1.460697	3.663164
19	1.92661	1.467812	3.615444
21	1.899736	1.484777	3.606107
23	1.938109	1.50229	3.606107
25	1.971272	1.526918	3.600615

Fig.6 Evaluation of SAD Algorithm : Window size Vs % Error for PNG&PPM images

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

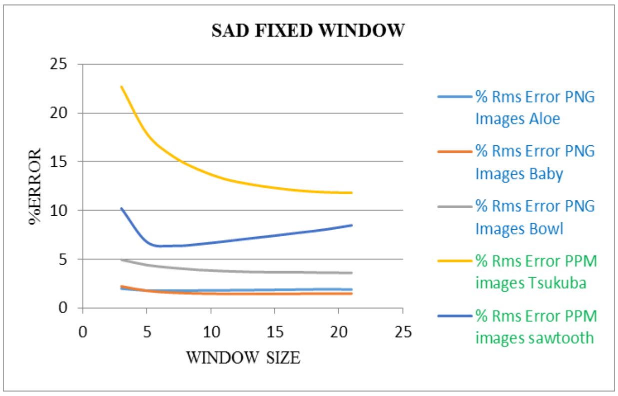


Fig.6 Evaluation of SAD Algorithm: Window size vs % Error for PNG and PPM images

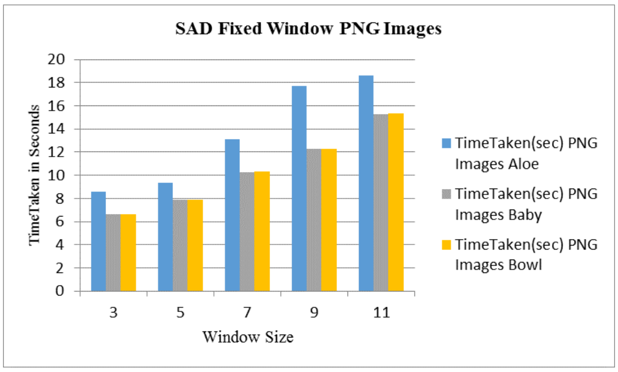


Fig.7 Evaluation of SAD Algorithm: Window size vs. Time Taken for PNG images

Table .2 RMS Error of PPM images.

widow size	% RMS Error :PPM images	
	Tsukuba	Sawtooth
3	22.6901836	10.2048411
5	17.8734223	6.765293862
7	15.5858594	6.364324525
9	14.210132	6.536930584
11	13.2600947	6.815589668
13	12.7019965	7.129811442
15	12.3149718	7.420282419
17	12.0263258	7.736663356
19	11.8665645	8.057743649
21	11.8146398	8.478018447

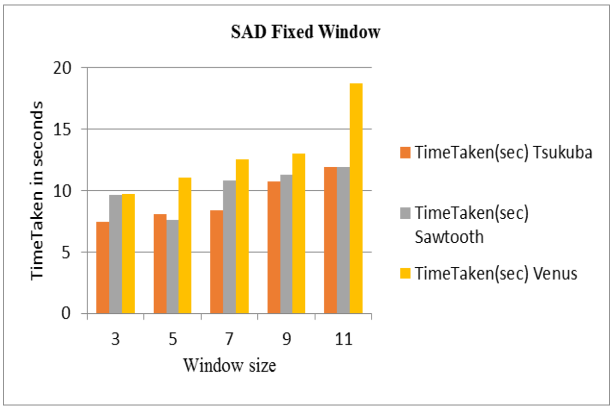


Fig.8 Evaluation of SAD Algorithm: Window size vs. Time Taken for PPM images.

IV. CONCLUSION

From the above evaluation it is seen that error decreases as the window size increases, but the time taken for the depth estimation will also increase. For PNG images - images are taken under good illumination condition ,when the window size increases beyond 11X11 for Aloe,13X13 for Baby and 17X17 for Bowl error increases .From the disparity map it can be seen that edges of the disparity space image obtained by the SAD algorithm fattens and this will leads to increase in error. In general this algorithm will give a better performance - less than 5% error-for images taken under good illumination condition. Also the computational cost is less for this algorithm ie, gives fast result for small window sizes.

For the case of PPM (Portable Pixel Map) images- taken under extreme bad illumination condition-the algorithm shows more error for Tsukuba max error is 22.69%, and for Sawtooth it is 10.20% .For this condition illumination invariant similarity measure is to be considered. In summary this SAD algorithm is better for real time application due to its simplicity and fast performance, but it requires images with good texture.

REFERENCES

1. Radhakrishnamurthy, H. C., Murugesapandian, P., Ramachandran, N., & Yaacob, S. (2007). Stereo vision system for a bin picking adept robot.W.-K. Chen, Linear Networks and Systems (Book style).Belmont, CA: Wadsworth, 1993, pp. 123–135.
2. Fusiello, Andrea, Emanuele Trucco, and Alessandro Verri. "A compact algorithm for rectification of stereo pairs." Machine Vision and Applications 12.1 (2000): 16-22.
3. Tsai, R. Y. (1986). An efficient and accurate camera calibration technique for 3D machine vision. In Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 1986.
4. Scharstein, Daniel, and Richard Szeliski. "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms." International journal of computer vision 47.1-3 (2002): 7-42.
5. Matthies, Larry, Takeo Kanade, and Richard Szeliski. "Kalman filter-based algorithms for estimating depth from image sequences." International Journal of Computer Vision 3.3 (1989): 209-238.
6. Matthies, Larry. "Stereo vision for planetary rovers: Stochastic modeling to near real-time implementation." International Journal of Computer Vision 8.1 (1992): 71-91
7. Faugeras, Olivier, et al. "Real-time correlation-based stereo: algorithm, implementations and applications." (1993).
8. http://vision.middlebury.edu/stereo/data/
9. Lane, R. A., and N. A. Thacker. "Tutorial: overview of stereo matching research." Imaging Science and Biomedical Engineering Division, Medical School, University of Manchester (1998).
10. Sonka, Milan, Vaclav Hlavac, and Roger Boyle. "Image processing analysis and machine vision." (1999).
11. Wang, Liang, et al. "How far can we go with local optimization in real-time stereo matching." 3D Data Processing, Visualization, and Transmission, Third International Symposium on. IEEE, 2006.
12. Robert, L., et al. "Applications of nonmetric vision to some visually guided tasks." Visual Navigation (1996): 89-135.

Deepambika V A, currently working as Assistant Professor, in department of Electronics and Communication Engineering at LBS Institute of Technology for Women, Thiruvananthapuram, Kerala, India. She has obtained B.Tech and M.Tech degrees from University of Kerala.

Arunlal S L currently working as Engineer in. ISRO, Thiruvananthapuram, Kerala, India. He has obtained M.Tech degree from M S University Tirunelveli.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)