The Solution of Stereo Correspondence Problem Using Block Matching Algorithm in Stereo Vision Mobile Robot

Rostam Affendi Hamzah
Faculty of Electronics & Computer Engineering, UTeM, Malaysia
rostamaffendi@utem.edu.my

Afifah Maheran Abdul Hamid
Faculty of Electronics & Computer Engineering, UTeM, Malaysia
afifah@utem.edu.my

Sani Irwan Md Salim
Faculty of Electronics & Computer Engineering, UTeM, Malaysia.
sanirwan@utem.edu.my

***Abstract* - This paper presents the solution of stereo correspondence problem occur in comparing stereo images on stereo vision mobile robot using block matching algorithm. The algorithm is using Sum of Absolute Differences (SAD). Left image works as a reference block to the right image and the output is disparity mapping with the left coordinate system. The stereo vision baseline is based on horizontal configuration. The block matching technique is briefly described with the performance of its output. The curve fitting tool will determine the range of each obstacle detected in disparity mapping. The programming activities are using Matlab software starting from capturing images until navigation of mobile robot.**

Keywords - Block matching, disparity mapping, epipolar line, obstacle detection, stereo vision

I. INTRODUCTION

The origin of the word "stereo" is the Greek word "stereos" which means firm or solid, with stereo vision, the objects are seen solid in three dimensions with range [1]. In stereo vision, the same seen is captured using two sensors from two different angles. The captured two images have a lot of similarities and smaller number of differences. In human sensitivity, the brain combines the captured to images together by matching the similarities and integrating the differences to get a three dimension model for the seen objects. In machine vision, the three dimension model for the captured objects is obtained finding the similarities between the stereo images and using projective geometry to process these matches. The difficulties of reconstruction using stereo is finding matching correspondences between the stereo pair.

A. Stereo Vision Mobile Robot

A stereo vision mobile robot requires a number of various capabilities, including the ability to execute uncomplicated goal-achieving actions, like reaching a given location; to react in real time to unexpected events, like the sudden appearance of an obstacle; to build, use and maintain a map of the environment; to determine the robot's position with respect to this map; to form plans that pursue specific goals or avoid undesired situations; and to adapt to changes in the environment [2]. In a navigation of a stereo vision mobile robot research, the goal is to build up a vehicle that can navigate at a certain speeds using stereo camera whether in outdoor or indoor environments such as fields or building. These vehicles require massive computational power and suitable algorithm in order to adapt their sensitivity and control capabilities to the certain speed of motion and in avoidance of obstacles.

Essentially, a stereo camera Figure 1 on top of mobile robot takes pictures of a same object but with different view. The two dimension images on the plane of projection represent the object from camera view. These two images contain some encrypted information, the image-depth of each other. This information is the third dimension of two dimension images. Therefore, the object distance and its depth can be determined by using the stereo cameras [3].

With referring to Figure 2 the distance between two points of view is called the baseline [4]. The baseline's distance end to end affects on the range resolution, which establish the range of depth that can be calculated. A difference between scenes of the same object on the two images is called disparity [5]. The stereo system uses horizontal baseline. The cameras are placed at the same elevation. The process of stereo vision is then typically defined as finding a match between features in left and right images [6].

The stereo baseline b is the distance between the centres of the projection O_l and O_r. The x_l and x_r are the coordinates of p_l and p_r with respect to the principal points cl and cr. Where f is a focal length of the camera. The depth of P is uniquely determined by the following equation [4]:

$$\frac{b+x_l-x_r}{Z-f} = \frac{b}{Z} \tag{1}$$

And let the disparity d = $x_r - x_l$, then the depth of point P is:

$$Z = f\frac{b}{d} \tag{2}$$

The corresponding two images representing the same point of the scene is called disparity matching. The set of

Figure 1. Stereo camera placement

978-0-7695-4043-6/10 $26.00 © 2010 IEEE
DOI 10.1109/ICCRD.2010.167

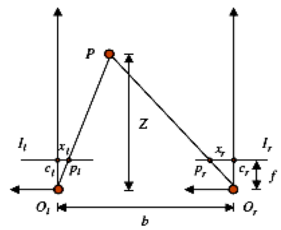


Figure 2. Basic principle of stereo vision

displacements between matched pixels is usually indicated as disparity map [6].

II. HARDWARE ARCHITECTURE

The hardware implementation Figure 3 has three important parts consist of stereo camera, processor Intel Core 2 Duo with 2GHz central processing unit clock and motor controller. The communication medium to each other is using the USB port. The cameras are CMOS type Logitech 1.3M pixel will capture the stereo images and send to processor using the USB port. Before that after the placement of stereo camera needs to be calibrated [3]. The calibration process is using Tsai's method. This method produces the values of extrinsic parameters such as the rotation and translation of images. Another result from Tsai's method is intrinsic parameters that generate focal length, principal point, skew, distortion coefficient and pixel size [7][8]. The processor processes the images and changes it to the digital values. Then the algorithm uses the values to map the disparity mapping. The motor controller receives the signals from processor to move or turn to a certain value according to the disparity mapping.

III. SOFTWARE ARCHITECTURE

The Matlab software is used to implement the software part. From Figure 4 images are captured by using the

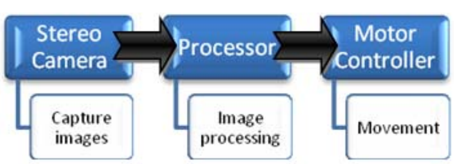


Figure 3. Hardware implementation

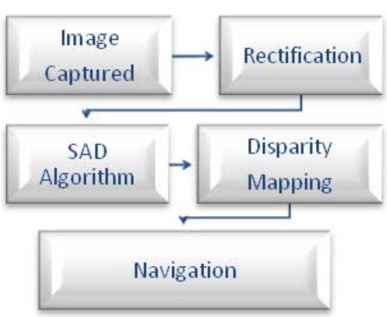


Figure 4. Software implementation

calibrated stereo camera. Then it will be changed to grey level scale to enhance the images using histogram equalization method [9]. The equation is:

$$s = T(r) = \int_0^r p_r(w)dw \qquad (3)$$

P_r is the original probability density function and $T(r)$ is final result of histogram equalization is independent of the density inside the integral. The data from camera calibration using Tsai's method will be used to rectify the stereo images to obtain corresponding epipolar lines that parallel to the horizontal scan-lines for both images. See Figure 5, the result of stereo image after rectification process using the equation (4). Where I_{old}, I_{new} are the original and the rectified image with the blending coefficients a_i separate for each camera.

$$I_{new}(x_0, y_0)=a_1I_{old}(x_1, y_1) + a_2I_{old}(x_2, y_2) + a_3I_{old}(x_3, y_3) + a_4I_{old}(x_4, y_4) \qquad (4)$$

With referring to dotted line in Figure 5(a) shows stereo image before rectification process and Figure 5(b) is the images with rectification process. Figure 5(c) is a perspective transformation is applied to an image the effect on the sampling is modeled by the change in local area. Depending on the location in the image the local area may shrink causing a loss of pixels, grow causing the creation of pixels, or remain constant. The output rectification process is look like a distortion image. But in fact that the image has been improved to be aligned each other with reference to epipolar line. The Sum of Absolute Difference (SAD) is applied to solve the stereo correspondence problem between two images. The SAD algorithm from the equation (4) shows an area-based correspondence algorithm [10][11][12]. The reference point of the left image g_t is minus with the right image g_{t-1} at the same epipolar plane.

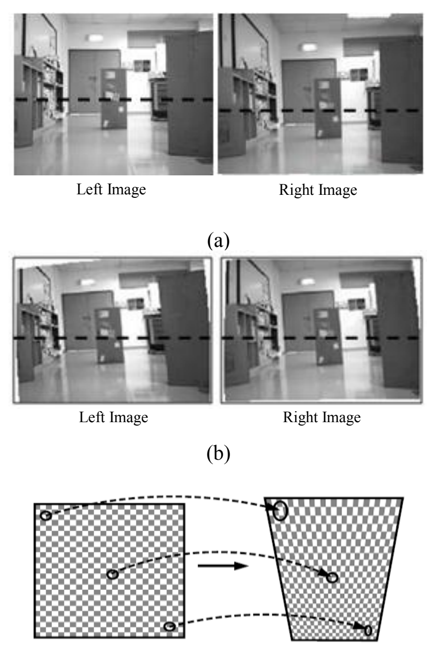


Figure5.. Stereo image rectification process

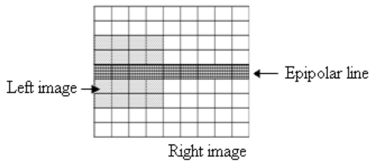


Figure6.. SAD block matching algorithm

It computes the intensity differences for each center pixel of (i, j) as follows:

$$SAD_v(x, y) = \sum_j \sum_i \| g_t(x+i, y+j) - g_{t-1}(x+v+i, y+j) \| \quad (5)$$

The equation above can be explained as reference point of the left image minus with the right image at the same epipolar plane. It sums up the intensities of all surrounding pixels in the neighborhoods for each pixel in the left image. To calculate stereo correspondence of stereo images, this paper is using block matching technique. Each block from the left image is matched into a block in the right image by shifting the left block over the searching area of pixels in right image as shown in Figure 6.

In order to find corresponding pairs of stereo points, they first have to be compared for different disparities, after which the best matching pairs can be determined. The maximum range at which the stereo vision can be used for detecting obstacles depends on the image and depth resolution. Absolute differences of pixel intensities are used in the algorithm to compute stereo similarities between points. By computing the sum of the absolute differences SAD for pixels in a window surrounding the points, the difference between similarity values for stereo points can be calculated. The disparity associated with the smallest SAD value is selected as best match [13].

IV. Range Estimation And Obstacle Avoidance

Together with the stereo camera parameters from software calibration and the disparity between corresponding stereo points, real-world distances can be retrieved. In order to find corresponding pairs of stereo points, they first have to be compared for different disparities, after which the best matching pairs can be determined. The maximum range at which the stereo vision can be used for detecting obstacles depends on the image and depth resolution [13]. The equation below shows the distance or depth of obstacles that has been used in this paper.

$$Distance = ae^{bx} + ce^{bx} \quad (6)$$

The value of a, b, c and d is a constant value from curve fitting tool in Matlab. The value of *x* represents a pixel value in disparity mapping. The calculated distance has an effective range of detection for the stereo vision mobile robot.

In the obstacle detection step, the points belonging to obstacles must be found. For data combination in image processing software, it is important to cluster or group the obstacle points in such a way that individual obstacles can be identified. If there are obstacles at closest range, the stereo program will filter using Gaussian low pass filter [9](7) to determine the object and cluster the pixels location.

$$H(u, v) = e^{-D^2(u,v)/2\sigma^2} \quad (7)$$

Figure7. Region of interest for mobile robot navigation

The disparity values from the SAD algorithm are mapped to disparity mapping with the reference of left image coordinate. The region of interest is the area within the green line box that the mobile robot searching for obstacles Figure 7. This area is a reference box for stereo cameras to navigate. So, the mobile robot is designed to see the obstacles only in that particular area.

V. Stereo Vision Result

With some illustration of basic stereo theory shown by Figure 8, the obstacles detection for close object and far object can be concluded by Figure 9. Which is the baseline range are permanently setup and the left plane works as a reference plane. The result is the disparity values at a close object bigger than the disparity of a distant object shown by Figure 9. In disparity mapping the brighter pixels in the disparity map, the more close the corresponding pixel in the reference image. As a result Figure 10, an obstacle could be identified as a group of pixels, which are generally brighter than their neighborhood. The original images are gray level scale but not aligned each other (left and right). The rectification process produces aligned images with same image size about 320x240. At the border of both images the rectification process unable to allocate the pixel values because of the original images has been reconstructed. So, the SAD algorithm cannot identify the disparity values at the border of rectified images. Therefore, the mapping of disparity values is undetected at that border.

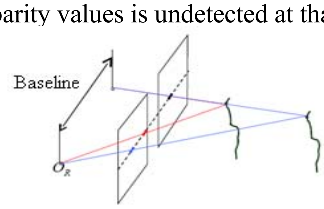


Figure8. Stereo vision of obstacles detection

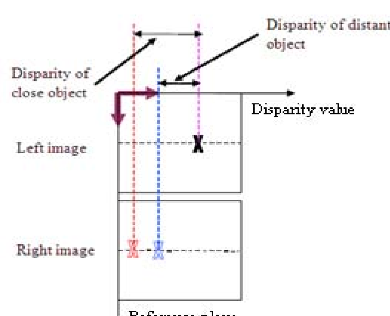


Figure9. Relation of disparity and distance

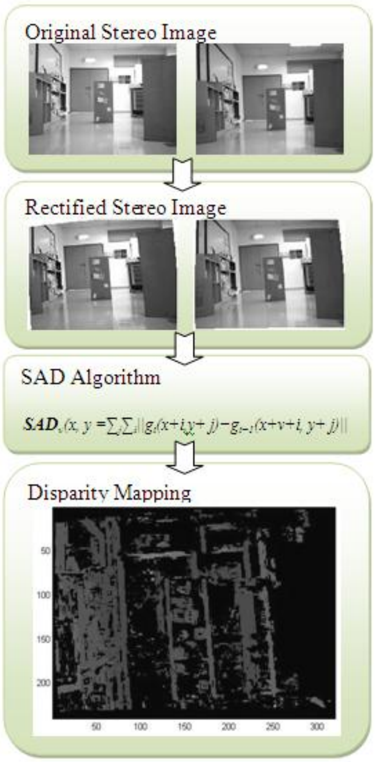


Figure10. Disparity mapping of stereo image

TABLE I. RESULT OF RANGE AND DISPARITY VALUE

Pixel Value	52	12.5	0.5	-5.5	-9	-11	-13	-14	-15	-16	-17	-17.5	-18	-18.5	-19
Range (m)	0.3	0.6	0.9	1.2	1.5	1.8	2.1	2.4	2.7	3	3.3	3.6	3.9	4.2	4.5

The result of the disparity values is shown by TABLE 1. The maximum detected range is 4.5 meter and -19 disparity value. And the minimum detected range is about 0.3 meter with 52 disparity value. This region refers to an effective range of mobile robot detection.

VI. STEREO VISION MOBILE ROBOT NAVIGATION USING DISPARITY MAPPING

Figure 11 shows the flowchart of controlling the mobile robot navigation. This flowchart is only for the experiment purpose. It's changeable according to the environment of navigation. The value of Ds represents the main distance of red line in Figure 12. The θs represents the value of mobile robot turning in degrees and the value of ds is minimum detected range of obstacle. After the main program activated, the process makes a decision by comparing the range of obstacle (D) if the value below than 0.5 meter then the mobile robot reverse and turn 90 degrees to the right. And start to run the program again. If the distance of the obstacle more than 0.5 meter, it turns to a certain degrees (θ) to the left or right depends on which area is the faraway object detected. After that the mobile robot move forward to d distance and start to run the program again.

The disparity mapping from Figure 10 can be plotted to the digital values. Therefore it can be used as a relationship with the distance estimation. The result is a graph with the horizontal coordinate of left image Figure 13 (a)(b) as a reference. The result could be explained when the pixel values are big, the range is close to the mobile robot and if the pixel values are small, the range located faraway from mobile robot. So the pixel works contrary to the range shown by Figure 14.

VII. CONCLUSION

The process of stereo correspondence or stereo matching using block matching algorithm is able to establish a correspondence by matching image pixel intensities. The output is a disparity mapping which storing the depth or distance of each pixel in an image. Each pixel in the map corresponds to the same pixel in an image, but the grey level corresponds to the depth at that point rather than the gray-shade or color. Viewing a gray level histogram equalized disparity map image, objects that are lighter are closer, and darker object are farther away. Stereo vision is also a type of passive sensor, meaning that it uses the radiation available from its environment. It is non-intrusive as it does not need to transmit anything for its readings.

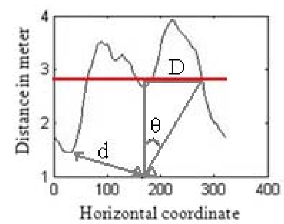
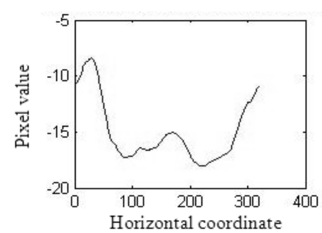
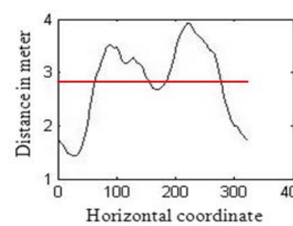


Figure12. Stereo vision parameters



(a)



(b)

Figure13. Distance estimation work opposite to the pixel value

REFERENCES

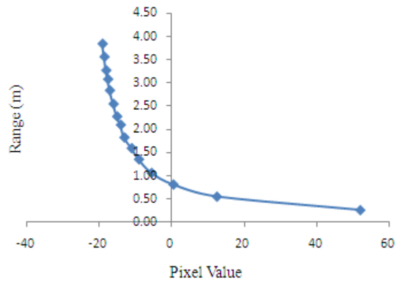


Figure14. Relation of pixel value and range

[1] Rovira-Más, F., Q. Zhang, and J. F. Reid., Automated agricultural equipment navigation using stereo disparity images. Trans. ASAE 47(4). 2004.

[2] Jones, E., J. Radford, D. Kumar, B. Fulkerson, R. Walters, and R. Mason, 2006. Autonomous off-road driving in the DARPA Grand Challenge. In Proc. PLANS Conference. IEEE

[3] Teerapat Chinapirom, U.W., and Ulrich Rückert, Steroscopic Camera for Autonomous Mini-Robots Applied in KheperaSot League. System and Circuit Technology, Heinz Nixdorf Institute, University of Paderborn. 2001.

[4] Konolige, K., Stereo Geometry. SRI International. 1999.

[5] B. Klaus, P.H., Robot Vision (MIT Electrical Engineering and Computer Science Series). MIT Press, McGraw-Hill Book Company, Cambridge, MA, 1986.

[6] Digney, S.S.a.B., Autonomous Cross-Country Navigation Using Stereo Vision. CMU-RI-TR-99-03, Carnegie Mellon University Technical report. 1999.

[7] Tsai, R.Y., An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Miami Beach FL, 1986, pp. 364–374.

[8] J.Y. Bouguet, "Camera Calibration Toolbox for Matlab", *MRL-Intel Corporation*, USA, 2003.

[9] Gonzalez, R.C., Digital Image Processing using Matlab. Pearson, Prentice Hall, 2002, pp 65-104

[10] Fusiello, E.T., and A. Verri, A compact algorithm for rectification of stereo pairs", Machine Vision and Applications. 12(1) 2000, pp.16-22.

[11] Mattoccia, L.D.S.a.S., Fast stereo matching for the videt system using a general purpose processor with multimedia extensions. Fifth IEEE International Workshop on Computer Architecture for Machine Perception, Padova, Italy, 2000.

[12] Kuhl, A., Comparison of Stereo Matching Algorithms for Mobile Robots. The University of Western Australia Faculty of Engineering, Computing and Mathematics, 2005.

[13] Johan C. van den Heuvel, J.C.M.K., Obstacle Detectection For People Movers Using Vision And Radar. TNO Physics and Electronics Laboratory Oude Waalsdorperweg 63, The Netherlands, 2003.