International Journal of Computer and Electrical Engineering, Vol. 2, No. 2, April, 2010
1793-8163

Region of Interest in Disparity Mapping for Navigation of Stereo Vision Autonomous Guided Vehicle

Anwar Hasni Abu Hasan, Rostam Affendi Hamzah, and Mohd Haffiz Johar

***Abstract*— Stereo vision system is a practical method for depth gathering of objects and features in an environment. This paper presents the region of interest (ROI) in disparity mapping for stereo vision autonomous guided vehicle (AGV) using block matching algorithm. This region is a reference sight of stereo camera. Stereo vision baseline is using horizontal configuration. The block matching technique is briefly described with the performance of its output. The disparity mapping is generated by the algorithm with reference to the left image coordinate. The algorithm is using Sum of Absolute Differences (SAD) which runs in Matlab software.The result is the AGV is able to navigate with using disparity mapping fron ROI.**

***Index Terms*— Block matching algorithm, disparity mapping, epipolar line, region of interest, 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. Autonomous Guided Vehicle (AGV)

An autonomous guided vehicle navigation 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 autonomous guided vehicle 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.

B. An Autonomous Guided Vehicle with Stereo Vision

Basically, the two stereo cameras on top of Autonomous Guided Vehicle AGV Fig 1 take 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 Fig. 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]:

$$\frac{b+x_l-x_r}{Z-f} = \frac{b}{Z} \qquad (1)$$

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

$$Z = f\frac{b}{d} \qquad (2)$$

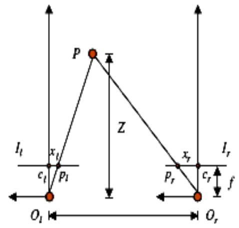
The corresponding two images representing the same point of the scene is called disparity matching. The set of displacements between matched pixels is usually indicated as disparity map [6].

Fig. 1 Stereo camera placement

Anwar Hasni Abu Hasan is with School of Electrical & Electronic, University Science of Malaysia, Penang, Malaysia. (Email: anwar@usm.edu.my)

Rostam Affendi Hamzah is with Faculty of Electronics & Computer Engineering, UTeM, Melaka, Melaka, Malaysia. (email: rostamaffendi@utem.edu.my)

Mohd Haffiz Johar is with School of Electrical & Electronic, University Science of Malaysia, Penang, Malaysia. (email : jojo@yahoo.com)



II. HARDWARE ARCHITECTURE

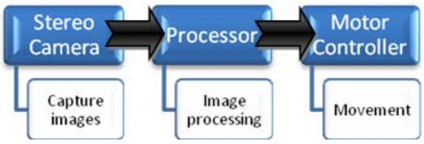


Fig. 3 Hardware implementation

Three important parts consist of stereo camera, processor Intel Core 2 Duo with 2GHz cpu clock and motor controller. The communication medium to each other is using the USB port Fig. 3. 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 Fig. 4 images are captured by using the calibrated stereo camera. Then it will be changed to grey level scale to enhance the images using histogram equalization method [9]. The equation is:

$$q = T(w) = Imax\sum_{r=0}^{w} Pw(r) \quad (3)$$

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 Fig. 5, the result of stereo image after rectification process using equation (4).

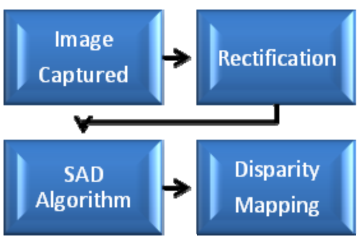
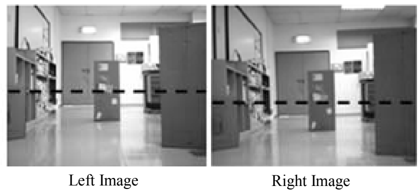
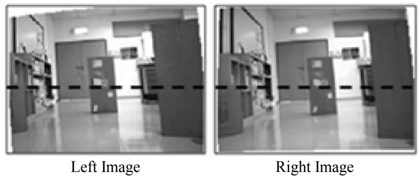


Fig. 4 Software implementation

$$I_{new}(x_0, y_0) = a_1 I_{old}(x_1, y_1) + a_2 I_{old}(x_2, y_2) + a_3 I_{old}(x_3, y_3) + a_4 I_{old}(x_4, y_4) \quad (4)$$



(a)



(b)

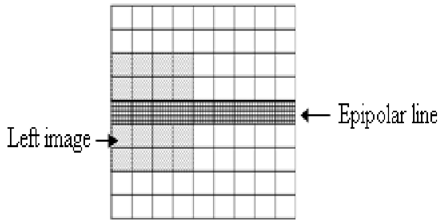
Fig. 5 Stereo image rectification process

With I_{new} and I_{old} as the original and the rectified image and the blending coefficients a_i separate for each camera With referring to dotted line in Fig. 5(a) shows stereo image before rectification process and Fig. 5(b) is the images with rectification process. 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].

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

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

Or in easy interpretation the reference point of the left image g_t minus with the right image g_{t-1} at the same epipolar plane. It sums up the intensities of all surrounding pixels in the neighbourhoods 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 Fig. 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

V. Filter

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.

$$\text{Distance} = ae^{bx} + ce^{dx} \tag{6}$$

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 autonomous vehicle.

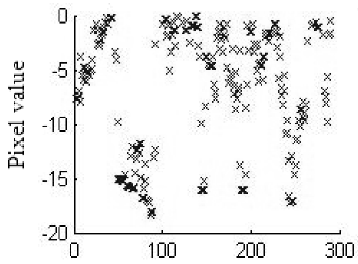
In the obstacle detection step, the points belonging to obstacles must be found. For data combination in image processing software, it is important to cluster or group the obstacle points in such a way that individual obstacles can be identified. If there are obstacles at closest range, the stereo program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

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

VI. Region of Interest in Disparity Mapping

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 AGV search for obstacles Fig. 7. This area is a reference box for stereo cameras to navigate. So, the AGV is designed to see the obstacles only in that particular area. The size of the region of interest is 320X20. From Fig. 8 shows the pixels spread in the region of interest in pixel value with the horizontal coordinate.

Fig. 7 Region of interest for AGV



VII. Stereo Vision Result

With some illustrations of basic stereo theory shown by Fig. 9, the obstacles detection for close object and far object can be concluded by Fig. 10. 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 Fig. 10. In disparity mapping the brighter pixels in the disparity map, the more close the corresponding pixel in the reference image. As a result Fig. 11, an obstacle could be identified as a group of pixels, which are generally brighter than their neighbourhood. 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.

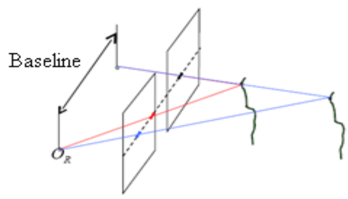


Fig. 9 Stereo vision of obstacles detection

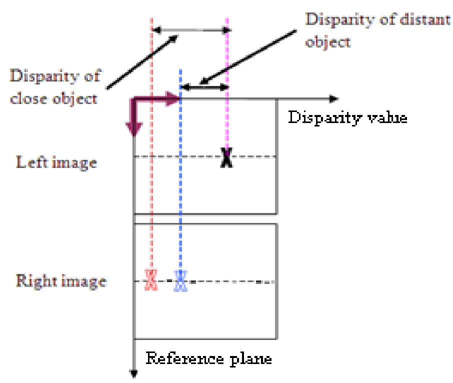


Fig. 10 Relation of disparity and distance

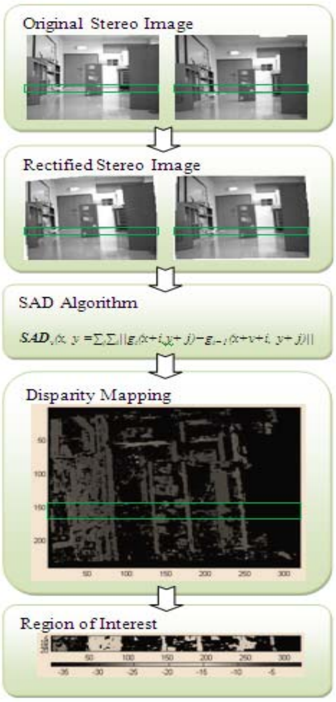


Fig. 11 Disparity mapping of stereo image

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

Therefore, the mapping of disparity values is undetected at that border. 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 AGV detection.

VIII. NAVIGATION OF AGV USING REGION OF INTEREST IN DISPARITY MAPPING

Fig. 12 shows the flowchart of controlling the AGV. The value of Ds represents the main distance of red line in Fig. 13. The θs represents the value of AGV 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 AGV reversed 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 AGV move forward to d distance and start to run the program again. The disparity mapping from Fig. 11 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 Fig. 14 (a)(b) as a reference. The result could be explained when the pixel values are big, the range is close to the AGV and if the pixel values are small, the range located faraway from AGV. So the pixel works contrary to the range shown by Fig. 15.

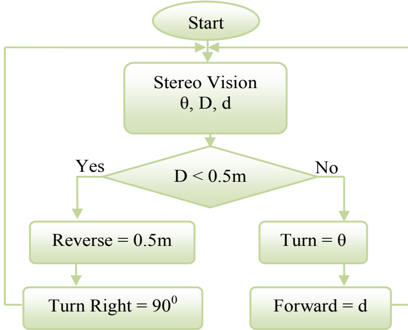
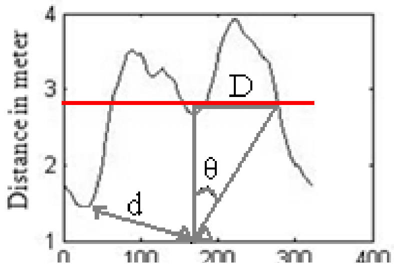
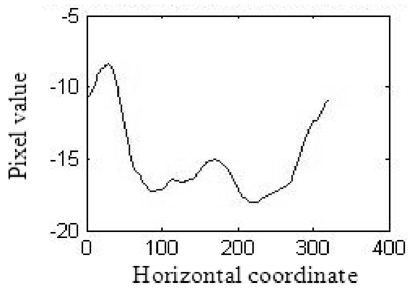
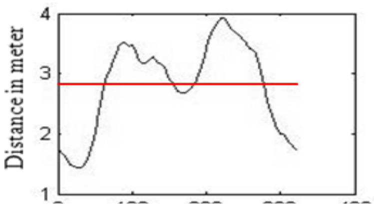


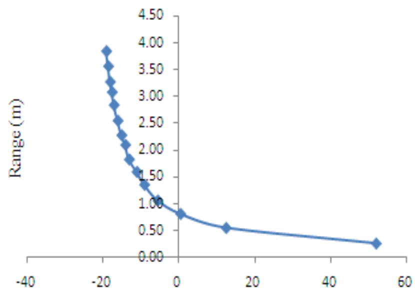
Fig. 12 Flowchart of autonomous navigation





(a)





The function of region of interest is to view a scene for navigation of AGV. The size of a new disparity mapping is adjustable to suitable environment view in Matlab programming. In this paper the size is explained above and the AGV is able to navigate efficiently with the selected size. For other applications the size should be increase or decrease depending on the surface of the scene or location

.

IX. CONCLUSION

A disparity map image is an efficient method for 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. The selection region of interest is a suitable method to AGV navigation. The AGV is able to avoid the obstacles with a reference from region of interest with several effective ranges. 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.

REFERENCES

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

[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, 2001. Steroscopic Camera for Autonomous Mini-Robots Applied in KheperaSot League. System and Circuit Technology, Heinz Nixdorf Institute, University of Paderborn.

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

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

[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.

Anwar Hasni Abu Hasan is a Senior Lecturer in the School of Electrical and Electronic Engineering at Universiti Sains Malaysia in Malaysia. His research interests include mechatronics, robotic and automation, quality control and instrumentation. He received his Ph.D. and M.Sc. from Hull University in UK.

Rostam Affendi Hamzah received the B.Eng. in Universiti Teknologi Malaysia and currently pursues his M.Sc. in electronic system design engineering at Universiti Sains Malaysia,Malaysia. His current research interests include stereo vision, image processing and embedded system.

Mohd Haffiz Johar received the B.Eng. in Universiti Sains Malaysia and currently pursues his M.Sc. by research at Universiti Sains Malaysia,Malaysia. . His current research interests include robotic and automation, stereo vision and image processing.