Target Distance Estimation Using Monocular Vision System for Mobile Robot

M. N. A. Wahab, N. Sivadev, K. Sundaraj

School of Mechatronics Engineering,
Universiti Malaysia Perlis,
Perlis, Malaysia.
mohdnadhir@yahoo.com

Abstract— **Mobile robot with vision could be useful for many applications and purposes. However, the vision system needs to be robust, effective, robust and fast to achieve on its goal. Somehow it needs stereo vision system to estimate the depth of the object. In this paper, a monocular vision system is introduced to the mobile robot to enhance their capabilities for calculating the distance or depth approximately. The main topic explains method used in monocular vision system, which is Hough transforms and how this method operates. This paper also shows the basic mathematical morphology which is implemented in Hough transform for image processing.**

Keywords-Mobile robot; monocular vision system; Hough transform; mathematical morphology; image processing

I. INTRODUCTION

Mobile Robot equipped with visual capabilities may give a great challenge for vision systems since it requires skills for developing the solution on complex image to understand the problem. It could be said that vision system is the most crucial part in mobile robot applications. It can enhance the function and mobility of the mobile robot drastically. For example, to serve human or play soccer, the robot must be able to find and maintain the fix interest point on a moving object while the robot itself is moving. This capability has been proven to be difficult for artificial systems [1].

Although many researches have been done in mobile robot vision system, most of the existing algorithms are not designed with real-time color based segmentation [2]. Robust adaptive vision against variable lighting and color also had been studied [3]. While in the other aspect such as accuracy, a simple heuristic approach in shifting pixels to lie within a rectangle is being introduced but the transformation seems arbitrary [4].

There are also some researches works on calibration and distortion correction by using variable camera models [5, 6]. However, the accuracy of localization affected by color pattern and it starts to have attention from the researchers. Different patterns are compared and the error of each object detected is analyzed [7]. Despite the fact that the simulation result is good, but error still occur in the model for actual robot. There is also a research that related to general or common problems in designing a robotic vision [8]. The vision system can give the best outcome depends on the system architecture and environment.

The development of a vision system which requires robustness and efficiency is still very difficult. For a practical vision system, minimum requirement on the speed of image processing is 25fps (frames per second) and also robustness for low level processing tasks to maximize the time available for high level decision making process.

In this paper, our approach is using a wireless camera as monocular vision system that placed on the mobile robot to identify and distinguish between the obstacle and target. We also will be familiarized with basic mathematical morphologies and the Hough transform approach for our monocular vision system. The main novelty for this paper is the implementation distance estimation by using monocular vision system.

II. MATHEMATICAL MORPHOLOGY

Mathematical morphology is a useful tool for image processing. Basic operators are used when we transform color image into binary image by applying sub-sampling techniques and threshold-based segmentation [9]. The morphologies used are dilation, erosion, closing and opening.

A. Dilation Operation

$$A \oplus B = \{z \in E \mid (B^S)_{\approx} \cap A = \emptyset \qquad (1)$$

where A is the image that is to be dilated, B is a set of coordinate points known as the structuring element.

The basic effect of the dilation operator on a binary image is to enlarge the boundaries around the regions of foreground pixels gradually. Thus, areas of foreground pixels grow in size while holes within those regions become smaller.

B. Erosion Operation

$$A \ominus B = \{z \in E \mid B_{\approx} \subseteq A\} \qquad (2)$$

The basic effect of the erosion operator is to erode away the boundaries around the regions of foreground pixels. Thus area of foreground pixels shrink in size and the holes within those regions become larger. Figure 1 shows the result after dilation and erosion operation compare with original image.

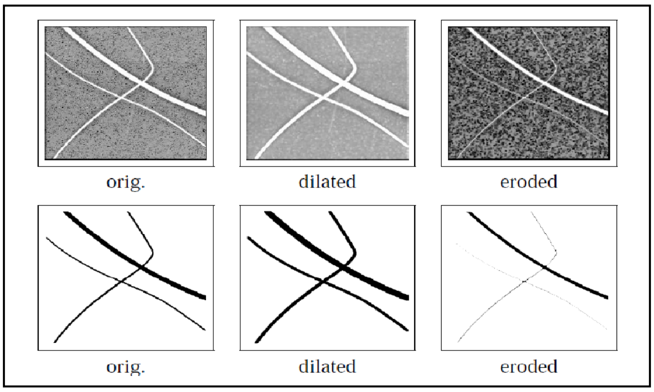

Figure 1: Effect of Dilation and Erosion Operation [9].

C. Opening Operation

$$A \circ B = (A \ominus B) \oplus B \tag{3}$$

The opening operation is a combination of erosion operation followed by a dilation operation.

D. Closing Operation

$$A \bullet B = (A \oplus B) \ominus B \tag{4}$$

The closing operation is a dilation operation then followed by an erosion operation using the same structuring element for both operations. Figure 2 shows the result after opening and closing operation.

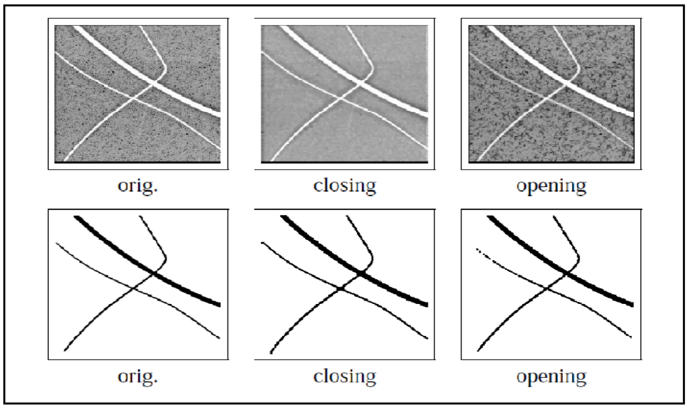

Figure 2: Effect of Opening and Closing Operation [9].

III. Hough Transform

The Hough transform (HT) is named after Paul Hough who discovered the method in 1962. It is a powerful global method to detect edges. This formula can be applied in circle detection or line detection [10].

A. Circle Hough Transform

The Hough transform can be used to determine the parameters of a circle when a number of points that fall on the perimeter are known. A circle with radius R and center (a, b) can be described with the parametric equations

$$x = a + R\cos(\square) \tag{5}$$

$$y = b + R\sin(\square) \tag{6}$$

When the angle □ sweeps through the full 360 degree range, the points (x, y) trace the perimeter of a circle. If an image contains many points with some fall on perimeters of circles then the job of the search program is to find parameter triplets (a, b, R) to describe each circle. Since the fact that the parameter space is 3D, it causes the direct implementation of the Hough technique becomes more expensive in computer memory and time.

The first step in this method is to convert the acquired image into a binary image with a threshold operation. Then it is followed by the opening morphology to remove background noise. The closing morphology is used afterward to remove small holes. The edges are identified by using the boundary operation morphology [11].

B. Line Hough Transform

When we have straight lines in an image, first we note for every point (x_i, y_i) in the image with all the straight lines passing through the point that satisfied Equation (7) for variables such as line slope and interception point (m, c) in Figure 3 [12].

$$y_i = mx_i + c \tag{7}$$

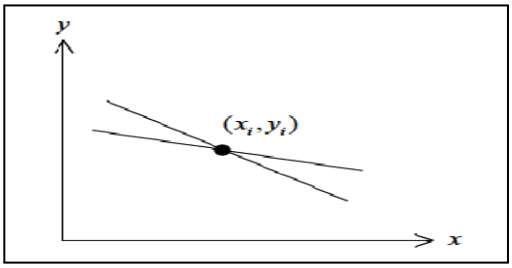

Figure 3: Lines through a Point in the Cartesian Domain [12].

If we set straight our variables (m, c) as a function of the image point coordinates (x_i, y_i), then Equation (7) becomes:

$$c = y_i - mx_i \tag{8}$$

Equation (8) describes a straight line on a graph of c against m as shown in Figure 4. At this point, it is easy to find that each different line through the point (x_i, y_i) is corresponds to one of the points on the line in the (m, c) space.

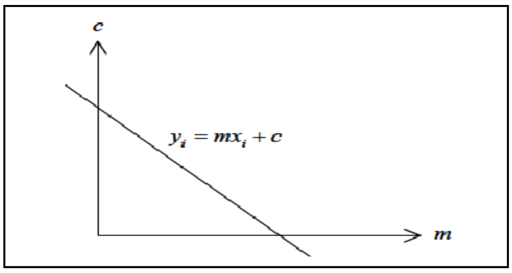

Figure 4: The (m, c) Domain [12].

Consider two pixels P1 and P2 lay on the same line in the (x, y) space, for each pixel, we can represent all the possible lines through it by a single line in the (m, c) space. Hence, a line in the (x, y) space that passes through both pixels must lay on the intersection of the two line in the (m, c) space which is represent by the two pixels. This means that all pixels which lay on the same line in the (x, y) space are represented by line which all pass through a single point in the (m, c) space as seen in Figure 5 and Figure 6.

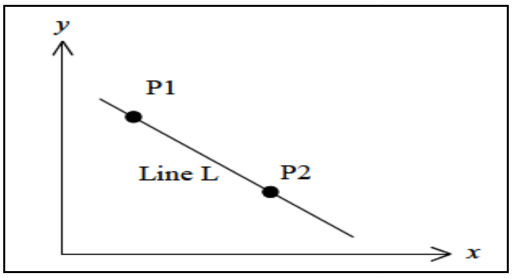

Figure 5: Points on the Same Line [12].

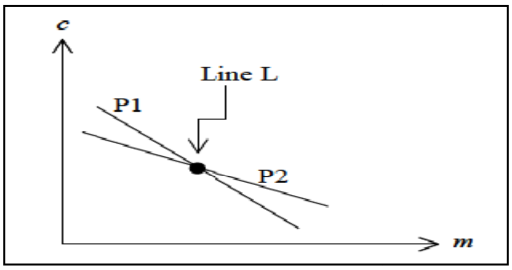

Figure 6: The Mapping of P1 and P2 from Cartesian Space to the (m, c) Space [12].

These are the steps of algorithm that used to detect lines in the images with simple explanation [12]:

1) By using any suitable edge detection scheme, we find all the edge points in the image.
2) Quantize the (m, c) space into a two-dimensional matrix H with appropriate quantization levels.
3) Initialize the matrix H to zero.
4) Each element of H matrix, H (m_i, c_i), which is found to correspond to edge point in incremented by 1. The result is a histogram or a vote matrix showing the frequency of edge points corresponding to certain (m, c) values.
5) The histogram H is threshold where only the large valued elements are taken. These elements correspond to lines in the original image.

From the explanation of Hough Transform algorithm, we can see the advantages and disadvantages between circle detection and line detection. We decided to use Circle Hough Transform in our system because it performs faster in detecting the orange golf ball which is circle shape.

IV. EXPERIMENTAL RESULTS

We are using Miabot (Merlin Intelligent Autonomous Robot) as our mobile robot and a wireless camera is placed in the centre of the robot as shown in Figure 7. The camera also had tested on other positions and the result shows the best position to capture the image and the field view as seen in Figure 7.

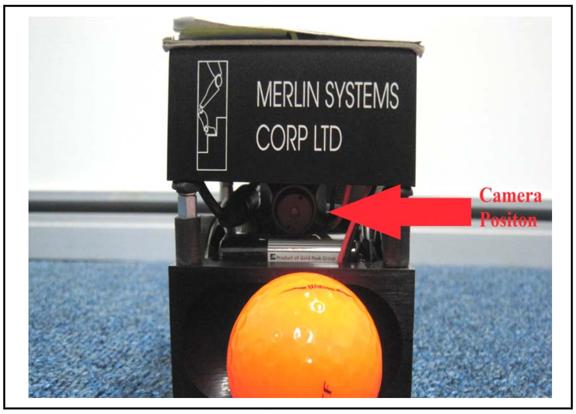

Figure 7: The Camera Position.

In our experiment, an orange golf ball is assigned as the target and other objects as obstacles. The overall system architecture is shown in Figure 8. There are two sub systems in the main system which is image processing part and mobile robot part. The system starts with image acquisition from wireless camera. Afterwards, the image will go through standard pre-processing such as filtering and smoothing the image. Hough Transform algorithm will be applied to detect the circle object before the distance estimation of the object proceeds. The information about estimated distance to the object will be passed to mobility part. After the target had been identified, few possible paths will be created before the best path being selected. The mobile robot moves and tried to avoid any obstacle that may be on the path until reached the target.

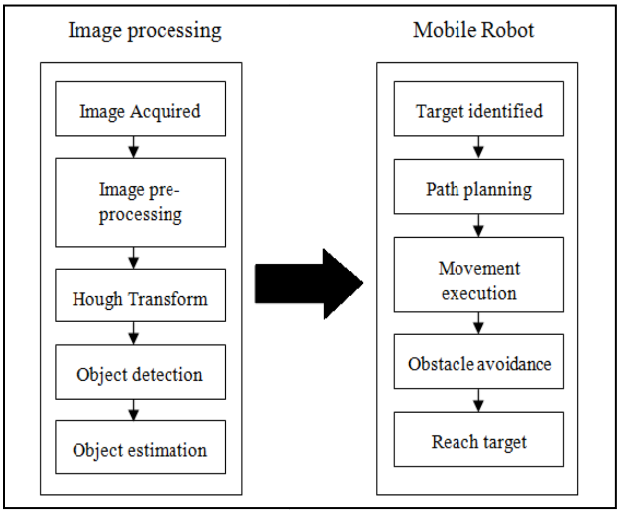

Figure 8: System Architecture.

The orange golf ball's position had been tested with different distance from the mobile robot. From the closest position to the mobile robot until to 100cm. Figure 9 shows the ball's position from the view of wireless camera on the mobile robot for the distance of 3cm, 10cm, 20cm, 30cm, 40cm and 50 cm. From Figure 10, the images of ball's position in distance of 60cm, 70cm, 80cm, 90cm and 100cm are shown respectively.

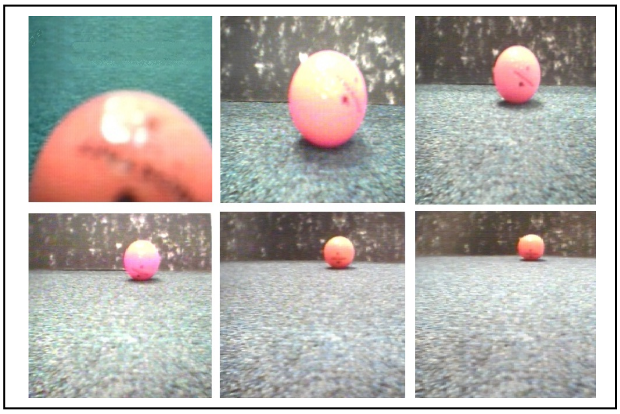

Figure 9: View from the Wireless Camera on Mobile Robot for Different Ball's Position

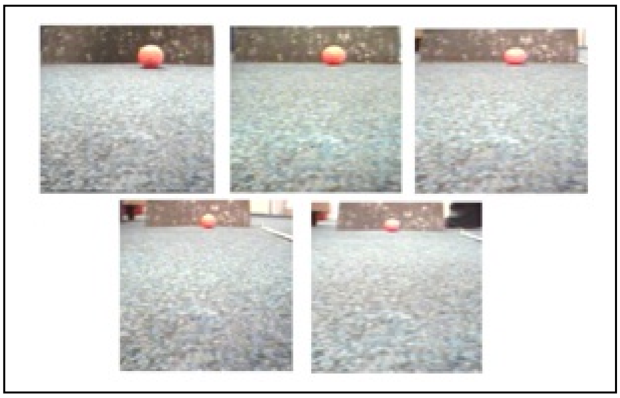

Figure 10: Different Ball's Position from the View of Mobile Robot.

As we can see from the figure, color of the orange golf ball is different for every ball's position depends on the lighting and environment. These variables may give different contrast and effect to the object's brightness. It may also affect the target tracking outcome.

From the images that obtained from the camera, the comparision between distance and diameter is done. The diameter is obtained by counting the pixels in the image. Table 1 shows the comparision between distance and diameter. The counting of pixel can been done by using various types of software with the sourcecode available online.

TABLE I. RELATIONSHIP BETWEEN DISTANCE AND DIAMETER

DISTANCE (CM)	DIAMETER (PIXEL)
10	139
20	83
30	60
40	47
50	38
60	36
70	29
80	25
90	23
100	21

The graph of diameter against distance is constructed from the table as shown in Figure 11. It clearly shows that diameter becomes smaller once the distance becomes larger. However after some points of distance, the pixels become approximately constant. The different of the pixel may be barely seen. Hence, we can say that the monocular vision system is only effective for the distance of 100cm and below. After this point, the system only can be used for detection purpose but not for depth estimation. Accuracy of the system is also reducing after 70cm.

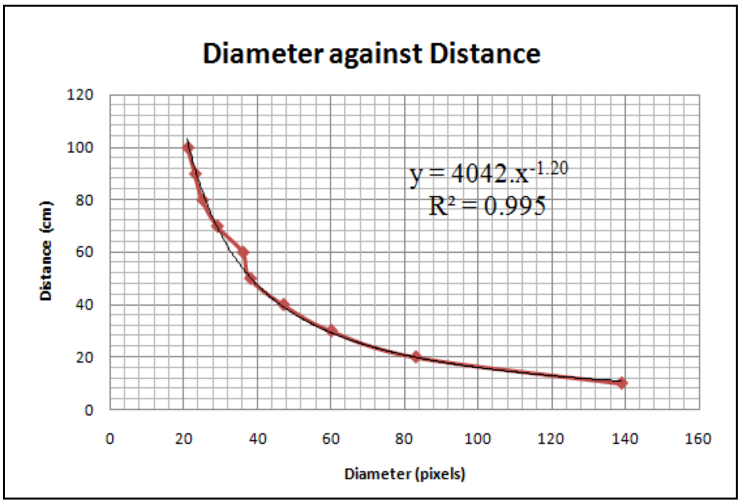

Figure 11: Graph Diameter against Distance.

As we can see from Figure 9, there is some extreme condition when the ball's position is less than 10cm. The object of interest is partially visible to the camera. This occasion is called occluded ball. For this situation, the method from [13] is used to obtain the diameter of the ball. The steps are listed below:

1) Obtain the height (H) of the occluded ball.
2) Obtain the width (W) of the occluded ball.
3) By using Equation (9), calculate the diameter (D).

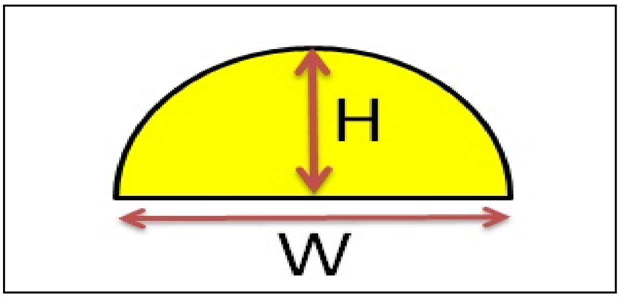

Figure 12: Calculation of circle's diameter from an occluded image.

$$D = 2\left(\frac{H}{2} + \frac{W^2}{8H}\right) \quad (9)$$

For an example, assume the occluded ball had the height of 3mm and with the width of 15mm as shown in Figure 13, then the diameter is obtained as calculated from Equation (10).

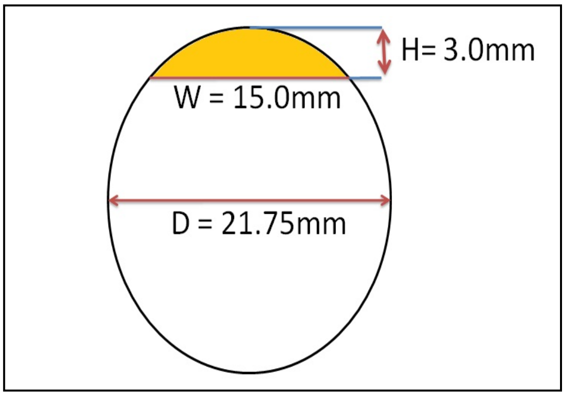

Figure 13: Example of calculation.

$$D = 2\left(\frac{H}{2} + \frac{W^2}{8H}\right)$$

$$= 2\left(\frac{3.0}{2} + \frac{15^2}{8 \times 3.0}\right)$$

$$= 2\left(1.5 + \frac{225}{24}\right)$$

$$= 2(1.5 + 9.375)$$

$$= 21.75mm \quad (10)$$

V. CONCLUSION

Research on monocular vision is unpopular due to the number of constraints and complexities. However, this vision system is possible to be developed and could be help for future works. This system will be revised from time to time to improve the overall performance of the system. For the time being, we are only looking into several algorithms to develop this system. Even though this work might seem small but it might be the stepping stone in enhancing the monocular vision system.

ACKNOWLEDGMENT

The authors would like to acknowledge Politeknik Tuanku Syed Sirajudin for their support and permitted us to use their equipments and facilities in making this research successful.

REFERENCES

[1] Micheal R. Blackburn and Hoa G. Nguyen, “Autonomous Visual Control of a Mobile Robot,” in ARPA Image Understanding Workshop, 1994, pp.1143-1150.

[2] J. Bruce, T. Balch and M. Veloso, “Fast and Inexpensive Color Image Segmentation for Interactive Robot,” IEEE/RSJ International Conference on Intelligent Robot and Systems, 2000.

[3] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[4] Ilario Dichiera, “Robot Soccer with Global Vision”, Hobors Dissertation, 1999.

[5] Chunmiao Wang, Hui Wang, William YC Soh, and Han Wang, “A Real Time Vision System for Robotiv Soccer”, 4^{th} Asian Conference on Robotics and its Applications, Singapore , 2001.

[6] Klancar Gregor, Kristan Matej, and Karba Rihard, “Wide Angle Camera Distortions and Non-Uniform Ilumination,” Mobile Tracking,Robotics and Autonomous Systems, 2004, pp.125-133.

[7] J. Bruce and M. Veloso, “Fast and Accurate Vision Based on Pattern Detection and identification”, IEEE International Conference on Robotics and Automation, 2003, pp. 1142-1147.

[8] B. S. Farroha, M. E. Valdez and R.G. Deshmukh, “Problems in Robotic Vision”, IEEE Proceedings of Southeastcon ’91, vol. 1, 1991, pp. 341-345.

[9] M. R. Haralick, R. S. Sternberg and X. Zhuang, “Image Analysis Using Mahtematical Morphology”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 4, July 1987, pp. 532-550.

[10] Peter E. Hart, “How the Hough Transform was Invented,” IEEE Signal Processing Magazine, November 2009, pp.18-22.

[11] H. Rhody, “Lecture 10: Hough Circle Transform,” Rochester Institute of Technology, October 2005.

[12] G. Hamarneh, K. Althoff, and A. G. Rafeef, “Line Detection using Hough Transform,” Project Report for the Computer Vision Course, September 1999.

[13] J. Page, Math Open Reference : Radius of an arc or segment, 2009.