

Design and Implementation of Dynamic Target Tracking with Stereo Vision System

Hsiao-Yu Wang , Hung-Yuan Chung*

Abstract—The purpose of this study is an intelligent control system on two mobile robots applied to moving target detection and tracking with improved diamond search algorithm using non-static binocular stereo vision. Binocular stereo visions are mounted on the top of one robot to track another robot. The method of stereo vision is precisely based on this principle, using two cameras. Imaging a single object from various angles, to find the same points in pictures through the proposed algorithm, the same point has different position in different pictures, we can use the triangular positioning to restore the actual positions coordinates. Based on the surrounding image data, the target detection using improved diamond search algorithm of computer vision system and the robot behavior motion depended on the fuzzy control strategy and the circuit design on chip. With proposed scheme could take less time to achieve target tracking effectively.

Keywords-Fuzzy control, Mobile robot, Moving target tracking, Stereo vision, Improved Diamond Search

INTRODUCTION

Mobile Robots are concerned to be safety and correctly moving in real-world environment. The tasks in real-world environment those are unconstructed, so that there exists a need to investigate different sensing method to improve the automatic tasks. Most robots that navigate in uncertain environment using infrared sensors or laser range sensors as their primary spatial sensor. However these methods could limit in the angle problem. In order to overcome this problem some researchers using vision systems.

In recent years, motion detection has been widely adopted in vision systems, since motion detection is very effective to exploit temporal redundancy of static CCD signals. There is still a lot of need for the methods that can find out motion vectors more accurately and faster. Of motion detection algorithms, full search algorithm (FSA) yields the optimal motion vectors but requires much computation.

To relieve the computational problem, there have been many algorithms. Block-matching algorithm (BMA) technique has been applied to reduce the computation time to remove the temporal redundancy between two or more successive images. Images are divided into equally sized rectangular blocks [1] and used to determine the displacement of the best matched block from the previous image as the motion vector to the block in the current images within the search window [2]. Based on a block distortion measurement, the displacement of the best-matched block

will be described as the motion vector (MV) to the block in the current image [3]. Full search (FS) [4], three-step search (TSS), new three-step, and diamond search (DS) [7] are methods developed search (NTSS) [5], hexagon-based search (HEXBS) [6] based on the block-matching technique.

During the last few years, fuzzy logic control has been suggested as an alternative approach of mobile robots [16][17] for target tracking. Since the first paper on fuzzy sets [8] was published, fuzzy logic control has attracted great attention from both the academic and industrial communities to use in smooth and real-time control.

In this paper, the hardware architecture of mobile robot details is presented in Section 2. A non-static stereo vision system is proposed as sensing method proposed in Section 3. Since non-static cameras let the background not stay in static situation, moving object and dynamic background will increase difficulty for moving object detection. However, we can detect moving target under complicated dynamic background by proposing in Section 4 to use an improved diamond search algorithm in adapting background model. At last we use fuzzy control of mobile robots to track another mobile robot which is shown in Section 5. The rule of the fuzzy set we use is offered by Zadh, and the mechanism of fuzzy is Min-Max offered by Mandani, which combined with image tracking system to achieve the real-time tracking control system [18]. The experimental results show that the proposed scheme can detect the target correctly and track it smoothly in real time. The experimental results are shown in Section 6. Finally, conclusions are given in Section 7.

HARDWARE ARCHITECTURE OF MOBILE ROBOT

Fig. 1 shows the experimental setup of two wheeled mobile robot (WMR) in an intelligent space. The overall system includes the following three parts: two mobile robots (one is including two DC motors, one microprocessor, one expansion circuit card, wireless device, and mechanism), two Webcams, and one personal notebook (including image processing card and wireless device). The right picture in Fig. 1 shows the model of another wheeled mobile robot considered as target in the visual tracking control problem.

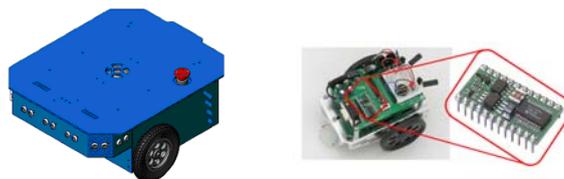


Fig. 1. (a)WMR Tracking Robot Fig. 1. (b) Target Robot

Manuscript received January 18, 2011.

Hsiao-Yu Wang is with the No.300, Jhongda Rd., Jhongli City, Taoyuan County 32001, Taiwan (R.O.C.) (phone: 0972709165; e-mail: 985201091@cc.ncu.edu.tw).

Hung-Yuan Chung is with the No.300, Jhongda Rd., Jhongli City, Taoyuan County 32001, Taiwan (R.O.C.) (phone: (03)422-7151-34575; e-mail: hychung@ee.ncu.edu.tw).

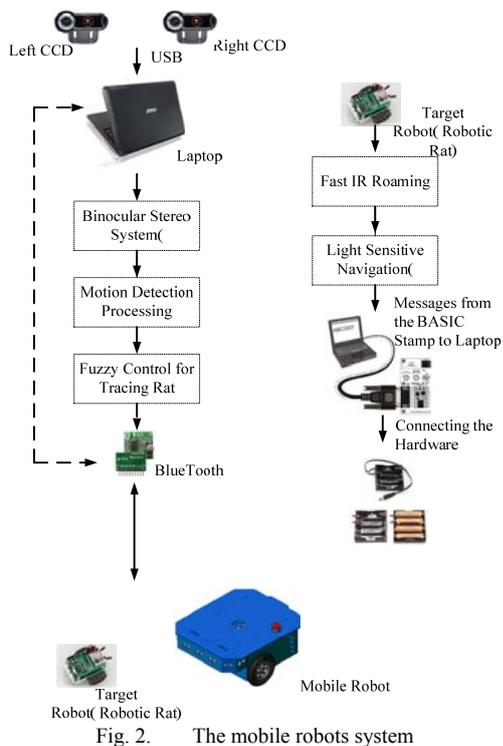


Fig. 2. The mobile robots system

The mobile robots system is shown in Fig. 2. The mobile robots equips with two Logitech cameras to track a dynamic moving target, which is supposed to be well recognizable object with appropriate dimensions in the image plane. The cameras are mounted on top of the mobile robot and its optical-axis faces the target of the robot.

STEREO VISION SYSTEM

In our stereo system, we construct the three dimensional coordinates by stereo visions mounted on the top of one robot as shown in Fig. 5. It comprises two cameras mounted on a bar; two cameras are all synchronized by the command of the signals. We define the X-axis as the horizontal direction and along the view direction of lens, the front and back direction defined as Z-axis and the vertical direction defined as Y-axis. In stereo system, it is most important to find the corresponding points. As defining the original points, the original point of X-axis is the middle of the two cameras, the center of lens is defined as the original point of Z-axis, and the same height of cameras is defined as the original point of Y-axis.



Fig. 3. (a) The vertical view



Fig.3. (b) The front view

After defining the coordinates, we measure the focal length f of the camera which is similar to the guidebooks of the camera and then solve the dense depth (the 3D coordinates Z) out by:

$$Z = \frac{L \cdot f}{dL + dR} \quad (1)$$

Where dL is the image width of target to left CCD optic axis and dR is the image width of target to right CCD optic axis. The distance between two cameras is L , which shown in Fig. 4:

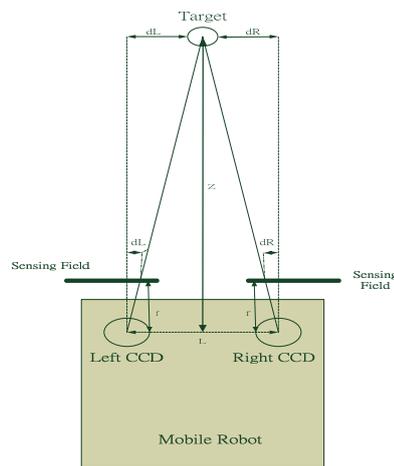


Fig. 4. The 2 CCD image projection and triangular relation
Now calculating the other two 3D coordinates X, Y by using the two dimensional coordinates and triangular relation. In the horizontal direction, as depicted in Fig. 5,

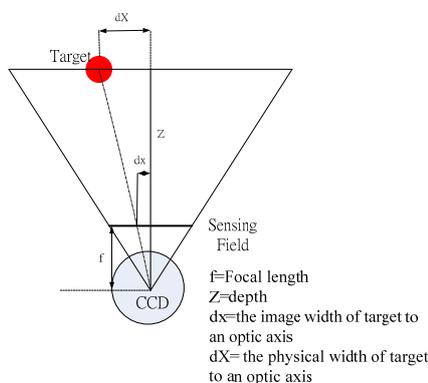


Fig. 5. The horizontal of 2D image projection and triangular relation

We obtained the relation between depth, focus of length, dX , and dY . And find out the physical width of target to an optic axis. This is expressed as:

$$\frac{dX}{dx} = \frac{Z}{f} \quad (2)$$

$$dX = \frac{Z \times dx}{f} \quad (3)$$

After transformation of coordinates, we obtained the X in three dimensional coordinates. In the vertical direction, as depicted in Fig.6, finding out the physical height of target to an optic axis. This is expressed as:

$$\frac{dY}{dy} = \frac{Z}{f} \quad (4)$$

$$dY = \frac{Z \times dy}{f} \quad (5)$$

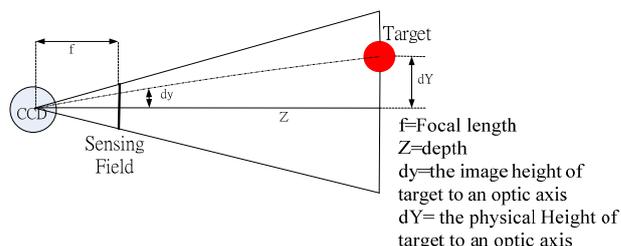


Fig. 6. The vertical of 2D image projection and triangular relation

COMPUTER VISION ALGORITHM

Improved Diamond Searching Algorithm

Motion vector of a block is highly correlated to the motion vectors of the adjacent blocks and also the motion vector of the temporal neighboring blocks. In this thesis, we propose a motion estimation algorithm that exploits new diamond search pattern which adaptively selects different size of shape in a search pattern. Besides that, based on the predicted motion vector, the accuracy of the prediction is improved by changing initial search center position.

It is widely known that the block motion field of natural video is usually gentle and smooth. Nie and Ma [9] suggested a zero-motion prejudgment (ZMP) to speed up the search. Without exception, we will further utilize both zero and center biased motion vector characteristics, however in an improved manner. The ZMP [11] uses a fixed threshold of SAD 512 for ZMV decision. We use an adaptive threshold to decide zero motion vector because we observed many zero motion vectors have SAD values larger than 512.

Based on the rapid block-matching motion estimation using adaptive small or medium diamond search pattern [10][12][13]. Our proposed improved algorithm also has two stages which are described as follows:

Dynamic Threshold (noted as THRESHOLD)

To illustrate our algorithm, we let the search window be Ω and the displacement with respect to the current block located at (x, y) be (u, v) . The SAD (Sum Absolute Difference), for the current block q between current frame n and reference frame $n-\tau$, is used for the block distortion measure, which is obtained through

$$SAD_{n,q;u,v} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |f_n(i, j) - f_{n-\tau}(i+u, j+v)| \quad (6)$$

Where $-\Omega \leq u, v \leq \Omega$ and N is the block size, $f(\cdot)$ is the pixel intensity.

In our algorithm, these redundant search points are avoided via a dynamic block difference threshold. Ideally, the dynamic threshold for the current q th block (*THRESHOLD*) should be very close to the final minimum SAD when using full search.

With this goal in mind, *THRESHOLD* is chosen adaptively to the content of block motion through a linear model. Let $SAD_{n,q;u,v}$ be the SAD of the current q th block in the n th frame. We use the variable of reference frame to obtain *THRESHOLD*, average minimum SAD ($ASAD_{n,m;0,0}$).

Since the minimum SAD of the q th block is not known yet, $ASAD_{n,q;\min}$ is calculated using all the blocks prior to the q th block as:

$$ASAD_{n,q;\min} = \frac{1}{q-1} \sum_{z=1}^{q-1} SAD_{n,z;\min} \quad (SAD \neq 0) \quad (7)$$

the computation time will increase and *THRESHOLD* is less accurate due to the value of SAD equals to zero. We calculate the average minimum SAD without $SAD=0$.

The dynamic threshold (*THRESHOLD*) is computed using a linear model with $ASAD_{n,q;\min}$.

$$THRESHOLD = ASAD_{n,q;\min} + \varepsilon \quad (8)$$

Where ε is a constant factor and is set as 256. Those parameters are carefully selected so that no or minimum premature endings happen.

Motion vector prediction:

Let $b_{t,i}$ is the current block in the t^{th} frame. The neighboring blocks are shown in Fig 7. To get the prediction,

$$Let \ B = \begin{bmatrix} b_{t-1,i} & b_{t-1,i0} & b_{t,i1} & b_{t,i2} & b_{t-1,i3} \end{bmatrix} \quad (9)$$

$$= \begin{bmatrix} b_0 & b_1 & b_2 & b_3 & b_4 \end{bmatrix}$$

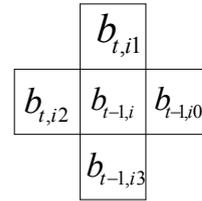


Fig. 7. The neighboring blocks used to predict the motion vector

Denote the blocks used to predict the current block's motion vector (every element of B is also a 2D vector), and $\lambda = [a_0 \ a_1 \ a_2 \ a_3 \ a_4]$ denote its weighting coefficients,

$$which \ satisfies \ \sum_0^4 a_i = 1 \quad (10)$$

The prediction types of predicting motion vector, $b_{t,i}^p$, of the current block $b_{t,i}$ are classified into two classes:

Case 1: regular motion (The motion is slow motion), that is,

If $\|b_i - b_0\| \leq THRESHOLD, \ \forall i \in [0, 4]$, then

$$b_{t,i}^p = sign(\cdot) \times I[abs(b_{t-1,i} + b_{t,i1} + b_{t,i2})/3] \quad (11)$$

Where $sign(\cdot)$ is a sine function, depending on the sine of $(b_{t-1,i} + b_{t,i1} + b_{t,i2})$, and $I[\]$ denotes rounding towards zero.

When we get the predicted motion vector, we DON'T use any block searching method to find the final solution. Because of the dynamic motion CCDs which will cause the stable background seems like in regular moving in the image, so that the blocks of background are regular moving vectors.

Case 2: irregular motion (The motion is quick motion), that is,

If $\|b_i - b_0\| > THRESHOLD, \ \exists i \in [0, 4]$, then

$$b_{t,i}^p = I[\lambda B^T], \quad (12)$$

Where $I[\]$ denotes rounding towards zero. When we get this predicted motion vector, we can use MDSP [9] to find the final solution. Go to Step1.

Step1

An MDSP is formed by repositioning the minimum SAD Eq. (13)

$$SAD = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |f_n(i, j) - f_{n-1}(i+u, j+v)| \quad (13)$$

Found in previous step as the center of MDSP. The MDSP search points are checked one-by-one. If the $SAD < THRESHOLD$, the search stops immediately; If the new minimum SAD search point is at the center of the newly formed MDSP, then go to next step for converging to the final solution; otherwise, this step is repeated.

Step2

With the minimum SAD search point in the previous step as the center, an SDSP is formed. The SDSP search points are checked one-by-one. If the $SAD < THRESHOLD$, the search

stops immediately (Final Sub-Step Stop); Otherwise, identify the new minimum SAD search point for the SDSP, which is the final solution for the motion vector

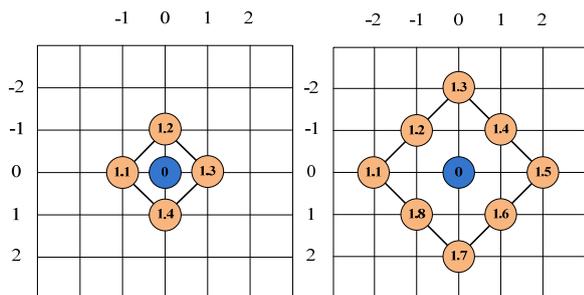


Fig. 8. (a) SDSP

Fig. 8. (b) MDSP

FUZZY CONTROL SYSTEM

In order to acquire a tracking system, we use the block diagram below to control our direction of mobile robot.

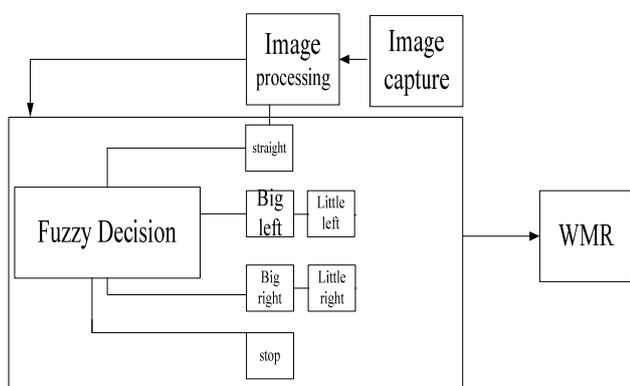


Fig. 9. Block diagram of fuzzy control selection system

Parameters utilize fuzzy logic to choose the most suitable strategy for the robot. It has two inputs for the fuzzy decision unit. The angle “ θ ” and the angular velocity “ $\dot{\theta}$ ” are input variables; each variable is ranged by membership function of five triangles PB (Positive Big), PS (Positive Small) and NB (Negative Big), NS (Negative Small) and ZO (zero) as shown in Fig. 10 and Fig. 11.

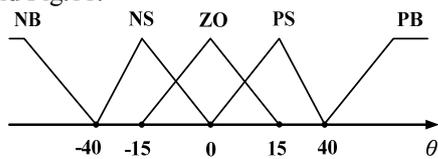


Fig. 10. Membership function of “ θ ”

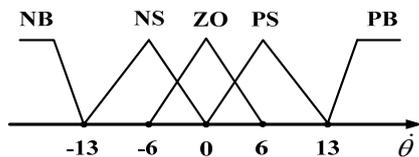


Fig. 11. Membership function of “ $\dot{\theta}$ ”

Membership functions of output u are the navigation angle of WMR, which is ranged by membership function of five triangles PB (Positive Big), PS (Positive Small) and NB (Negative Big), NS (Negative Small) and ZO (zero). Two inputs and five fuzzy partitions will construct twenty five rules as the fuzzy rule. The corresponding rule table is shown in Table I. The defuzzification interface is a mapping from

the action of fuzzy control space which is defined over an output u .

TABLE I.
FUZZY RULE BASE

$\theta \backslash \dot{\theta}$	PB	PS	ZO	NS	NB
PB	NB	NB	NB	NS	ZO
PS	NB	NS	NS	ZO	PS
ZO	NB	NS	ZO	PS	PB
NS	NS	ZO	PS	PS	PB
NB	ZO	PS	PB	PB	PB

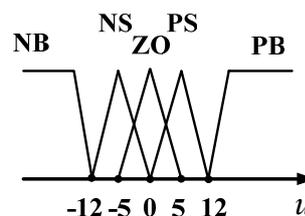


Fig. 12. Membership function of “output u ”

Consider a mobile robot driven by two differential wheels. The center of motion, denoted by C , is located at the midpoint between the left and right driving wheels. Assuming that the robot moves on the planar surface without slipping, the tangential velocity v_c and angular velocity w_c at the center C can be written as

$$v_c = \frac{r_w}{2}(w_r + w_l) \quad (14)$$

$$w_c = \frac{r_w}{d_w}(w_r - w_l) \quad (15)$$

Where w_r and w_l denote the rotational velocities of the right and left driving wheels, respectively, r_w is the radius of the wheels, and d_w is the azimuth length between the wheels.

The kinematic equation of the mobile robot is given by:

$$\dot{x}_c = v_c \cos(\theta_p) \quad (16)$$

$$\dot{y}_c = v_c \sin(\theta_p) \quad (17)$$

$$\dot{\theta}_c = \omega_c \quad (18)$$

Where coordinates (x_c, y_c) indicate the position of the robot with respect to the world coordinate system and θ_c is the heading angle of the robot. The triplet (x_c, y_c, θ_c) is used for defining the robot posture and represented by vector.

EXPERIMENT RESULTS

A. Performance Results of Stereo System

In this section, Table II shows the overall comparison of different depth implementation result in our stereo system. In Table II shows the error rate of different depth implementation result. The conclusion we observed from Table 2 that the far motion objects we detect, the higher error rate we get; however all the error rates are not over 5% so we assume that the tracking system is not influenced by our stereo system. The distance between 2 CCDs is 13cm.

TABLE II.
THE ERROR RATE OF DIFFERENT DEPTH IMPLEMENTATION RESULT

The actual depth(cm)	Z's error	Y's error	X's error
100	1	0.23	0.17
200	5	0.05	0.25
300	9	0.21	0.48
400	15	0.52	0.57
500	22	1.13	0.69
Error rate	Z	Y	X
100	1.00%	0.23%	0.17%
200	2.50%	0.03%	0.13%
300	3.00%	0.07%	0.16%
400	3.75%	0.13%	0.14%
500	4.60%	0.23%	0.14%

B. Experiments Results of Tracking System

In this section, we use two background methods to detect moving pixels. One is the traditional background subtraction method and another is proposed improved diamond search with frame difference subtraction method. Table III shows the performance between proposed method and other two traditional methods.

TABLE III.
THE PERFORMANCE COMPARISON

Average time form 3different video <u>with dynamic back ground</u>	Proposed method [without control system]	Diamond search method [without control system]	Traditional frame subtraction [without control system]
Detect SAD time	1.74ms	1.13 ms	No
Define the motion object	32.28 ms	37.79 ms	97.28 ms
Error rate	0.37%	53.76%	6.86%
Total time	51.84 ms	54.32 ms	147.76 ms

According to the result from Table III, we choose the lower error rate in the tracking system with fuzzy control so that we compare the results of two methods with control tracking system. Fig.13 to Fig.16 is the results of tracking simulation in different direction. In all figures, (a) is the final image in the tracking system, (b) is the route image of the method with tracking robot and target robot, (c) showing the iteration time comparison. ((X1, Y1) is the center of gravity in tracking robot; (X2, Y2) is the center of gravity in target robot).

Proposed Method tracking result with left target:



Fig. 13. (a) The final image

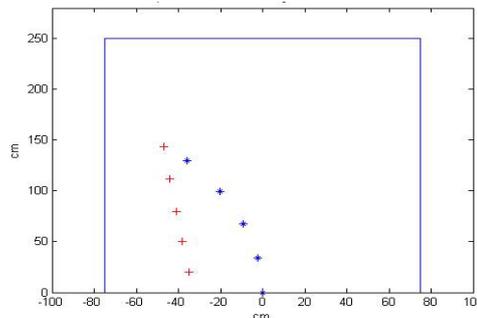


Fig. 13. (b) Route image (Red: Target, Blue: Tracking)

(X1, Y1)	(X2, Y2)	Proposed method(s)
(0, 0)	(-35, 20)	1.67
(-2.28, 34.05)	(-38, 50)	1.63
(-9/08, 67.48)	(-41, 80)	1.61
(-20.29, 99.71)	(-44, 112)	1.59
(-35.70, 130.16)	(-47, 144)	1.49
Average Iteration Times		1.60

Fig. 13. (c) Iteration time comparison

Traditional Method tracking result with left target:



Fig. 14. (a) The final image

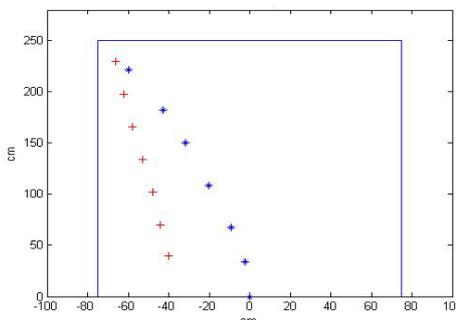


Fig. 14. (b) Route image (Red: Target, Blue: Tracking)

(X1, Y1)	(X2, Y2)	Traditional frame subtraction(s)
(0, 0)	(-40, 40)	1.99
(-2, 34)	(-44, 70)	2.06
(-9, 67)	(-48, 102)	1.95
(-20, 108)	(-53, 134)	1.93
(-31, 150)	(-58, 166)	1.96
(-43, 182)	(-62, 198)	2.04
(-60, 221)	(-66, 230)	2.11
Average Iteration Times		2.03

Fig. 14. (c) Iteration time comparison

Proposed Method tracking result with right target:



Fig. 15. (a) The final image

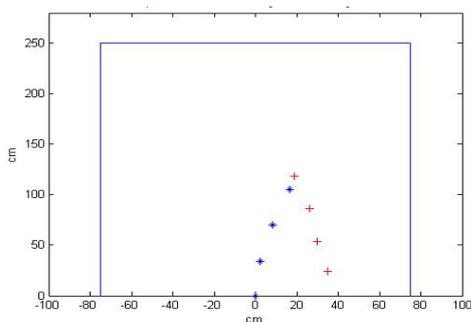


Fig. 15. (b) Route image (Red: Target, Blue: Tracking)

(X1, Y1)	(X2, Y2)	Proposed method(s)
(0, 0)	(35, 24)	1.63
(2, 34)	(30, 54)	1.56
(8, 69)	(26, 86)	1.75
(16, 105)	(19, 118)	1.61
Average Iteration Times		1.63

Fig. 15. (c) Iteration time comparison

Traditional Method tracking result with right target:



Fig. 16. (a) The final image

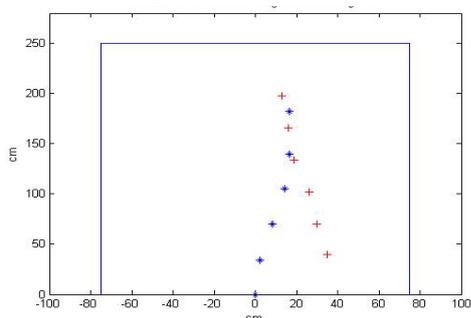


Fig. 16. (b) Route image (Red: Target, Blue: Tracking)

(X1, Y1)	(X2, Y2)	Traditional frame subtraction(s)
(0, 0)	(35, 40)	2.03
(2, 34)	(30, 70)	1.96
(8, 69)	(26, 102)	2.18
(14, 106)	(19, 134)	2.12
(16, 182)	(16, 166)	2.04
Average Iteration Times		2.066

Fig. 16. (c) Iteration time comparison

CONCLUSIONS

This paper proposed improved diamond search approach has improved motion detection accuracy and provides researcher with a new insight. The special experiments are that the CCDs and the background are all active. The better testing accuracy and the shorter time reaction of the wheeled mobile robot (WMR) are presented by the discovery using the improved diamond search with frame difference method and also compared with traditional

method results. A definitive conclusion is that the improved diamond search with background subtraction is adopted to suitable for the active condition.

At present, the majority of the processing time is spent on the control part. In order to reduce the control processing time we are currently looking for FPGA instead of Parallax Stamp 2px to realize the real time dynamic tracking system.

REFERENCES

- [1] Libo Yang, Keman Yu, Jiang Li, and Shipeng Li, "an effective variable block-size early termination algorithm for h.264 video coding", *IEEE Trans. on Circuits And Systems For Video Technology*, Vol. 15, No. 6, June 2005.
- [2] S. Zhu and K.-K. Ma, "A new diamond search algorithm for fast block matching–motion estimation," *IEEE Trans. Image Processing*, vol. 9, pp. 287-290, Feb. 2000.
- [3] J. Kim and R. Park, "A fast feature-based block matching algorithm using integral projections," *IEEE Jour. Sel. Areas in Comm.*, vol.10, no.5, pp. 968-971, Jun. 1992.
- [4] Andy Chang, Oscar C. Au, Y. M. Yeung, "A Novel Approach to Fast Multi-Block Motion Estimation" *Multimedia and Expo, 2003. ICME '03. Proceedings. 2003 International Conference on*, 105–108 2003.
- [5] R. Li, B. Zeng, and M. L. Liou, "A new three-step search algorithm for block motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 4, pp. 438–442, Aug. 1994.
- [6] Ce, Z., Xiao, L., Chau, L.P., Lim, K.P., Ang, H.A., Ong, C.Y.: Anovel hexagon-based search algorithm for fast block motion estimation. *IEEE. Trans. Circuits. Syst. Video. Technol.* 12, 349–355 2002
- [7] Shan, Z., Ma, K.K.: A new diamond search algorithm for fast block-matching motion estimation. *IEEE. Trans. Image. Process.* 9, 287–290, 2000
- [8] Zadeh, L.A., "Fuzzy Sets," *Inform. Control*, Vol. 8, pp. 338-353, 1965.
- [9] Y. Nie and K.-K. Ma, "Adaptive rood pattern search for fast block-matching motion estimation," *IEEE Trans. Image Processing*, vol. 11, pp. 1442-1449, Dec. 2002.
- [10] A. Tourapis, O. C. Au, and M. L. Liou, "Highly efficient predictive zonal algorithm for fast block-matching motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, pp. 934-947, Oct. 2002.
- [11] Yi, X., Nam, L.: Rapid block-matching motion estimation using modified diamond search algorithm. *IEEE. Trans. Circuits. Syst.* 6, 5489–5492, 2005
- [12] G.R. Bradski, "Computer vision face tracking for use in a perceptual user interface," in *IEEE Workshop on Applications of Computer Vision*, 1998, pp. 214–219.
- [13] Lam, C.W., Po, L.M., Cheung, C.H.: A novel cross-diamond search algorithm for fast block motion. *IEEE. Trans. Circuits. Syst. Video. Technol.* 12, 1168–1177 (2002)
- [14] Kim, B.G., Song, S.K., Mah, P.S.: Enhanced block motion estimation based on distortion-directional search patterns. *Pattern. Recogn. Lett.* 27, 1325–1335 (2006)
- [15] E. Menegatti, A. Pretto, A. Scarpa, and E. Pagello, "Omnidirectional Vision Scan Matching for Robot Localization in Dynamic Environments," *IEEE Trans. on Robotics and Automation*, vol. 22, pp. 523–535, 2006.
- [16] C. C. Wong, H. Y. Wang and S. A. Li, "PSO-based Motion Fuzzy Controller Design for Mobile Robots," *International Journal of Fuzzy Systems*, vol. 10, no. 1, pp. 284-292, Mar. 2008.
- [17] R. C. Luo, and T. M. Chean, "Autonomous mobile target tracking system based on gray-fuzzy control algorithm," *IEEE Trans. on Ind. Electron.* vol. 47, no. 4, pp. 920–931, 2000.
- [18] S. F. Lin, S. S. Liao and J. Y. Chen, "Design of fuzzy predictor in the moving target system," *International Journal of Fuzzy Systems*, vol. 3, no. 1, pp. 323–330, 2001.