Development of Monocular Vision based Targeting System for an Autonomous Defence Vehicle

Aakash Gupta*
Department of Mechanical Engineering
Indian Institute of Technology Madras
Chennai, India
aakashgupta18@gmail.com

P.V.Manivannan & M.Singaperumal
Department of Mechanical Engineering
Indian Institute of Technology Madras
Chennai, India
pvm@iitm.ac.in

***Abstract*—The present work deals with development of a Target tracking system based on Monocular Vision for an Autonomous Defence Vehicle without a global satellite view. The vehicle is also equipped with ultrasonic sensors to gauge distance from the target at any point of time. In this work, theoretical expressions have been arrived at for control of motors for accurate capture of target without image capture of the targeting pointer itself. This can greatly reduce the costs as the high resolution cameras required for capture of targeting pointer (generally a LASER pointer) are very costly. Theoretical expressions have also been determined for the approach distance in dependence only on camera, motor specifications and the target size. A prototype system was developed along with the necessary real time embedded controller and image processing algorithms for identifying the target and to control targeting system LASER pointer. Experiments were conducted to find out efficacy of developed system. The system is able to detect the target on different approach distances and camera cone angles. It should be noted that the target size will also be calculated on the fly by the robot itself. This makes it possible for deployment of this targeting module without much change to the existing system.**

Keywords*— *Autonomous Defence Vehicle, Monocular Vision, Approach Distance, Size Factor of the Target

I. INTRODUCTION

Autonomous defence vehicles have been a matter of great interest all over the world. They can be deployed in war zones to reduce human casualties. Generally, autonomous defence vehicles are capable of navigation, given the initial and final coordinates, obstacle avoidance as well as seek and destroy approach for target points. Multiple ways of positioning have been explored academically. Most important ones include Dead Reckoning, Global Vision, Global Positioning, Relative positioning, etc. Similarly, many methods have been suggested for obstacle avoidance and target capture. They involve proof of concept methods like colour / size-based differentiation to advanced theories like resultant forces and potential fields methods.

Although extensive work has been carried out in the field of navigation and obstacle avoidance [1][2][3], targeting still remains a relatively obscured field. Theoretical techniques and algorithms for image capture and processing have been widely explored e.g. Wang Kangling and Fu Xiaowei have explored feature based multi tracking algorithm [4], Canny J and Hildreth E. Maar have separately forwarded computational approaches to edge detection [5][6], while Arun Kulkarni has focused on fuzzy neural systems approach [7]. These techniques form the base of image processing in most of the targeting systems. However, the difference lies in the practical set-up and approach for using these algorithms.

Changhong Yu Zhiyong Wang and Tianding Chen have explored the concept of binocular stereovision for moving target detection [8]. Huadong Wang et al have explored an offline model for target detection wherein the robots collect and transmit the data, while it is processed separately by image queuing at a later point [9]. Toby P. Breckon et al have developed multi-modal automated target detection including thermal and infrared sensor inputs [10]. Another way of target detection via hyperspectral image processing [11] has been researched by Dimitris Manolakis, David Marden and Gary A Shaw, wherein the focus is on satellite capture of ground targets. Chein-I Chang has further refined this in his work on constrained sub-pixel target detection [12]

II. PRESENT WORK

Most of the targetting processes explored are inherently costly due to usage of high-resolution cameras, thermal imagery, use of satellites or costly computational power. The following work explores a single camera based solution for targetting, which is low cost and easy to mount on already existing bots. This becomes very important especially for Indian defence concerns.

This project aims to develop of a tracking system for an onboard vision (Monocular webcam) based defence vehicle without a global satellite view. The tracking system has both local and global intelligence systems

* Was a graduate student in the Department of Mechanical Engineering, Indian Institute of Technology Madras.

connected by a USART based communication module. Theoretical expressions have been arrived at for control of motors of targeting system for accurate capture of target without image capture of the targeting pointer itself. This can greatly reduce the costs as the high-resolution cameras required for capture of targeting pointer illumination.

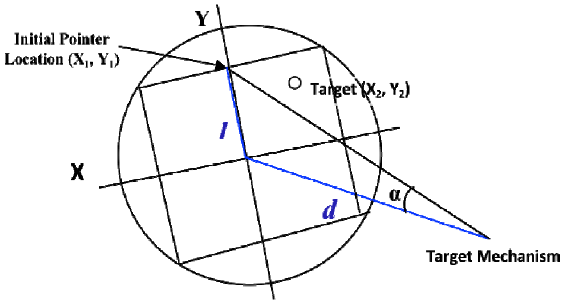
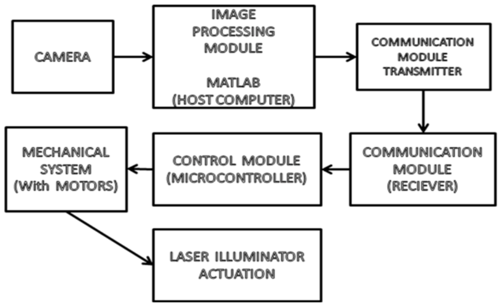
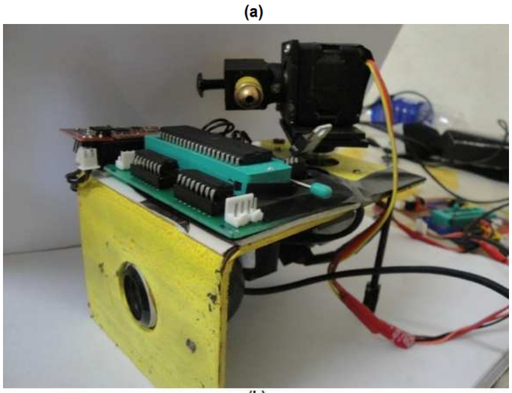


Fig.1 Various parameters used in mathematical model

Relations between displacement of the target pointer (y), rotation angle of motors (θ) and cone of vision angle of the camera (**α**) have been established (Refer Fig. 1). Furthermore, mathematical expressions have been determined for establishing the relation between the approach distance (d) and the size factor of the target (a) for a given system of motors and camera. Family of systems has been discussed to establish sensitivity to design elements by variation of one or two of these parameters keeping the others constant. The theoretical results thus obtained can be used for system design and implementation for various utilities, including but not limited to targeting mechanism of an autonomous defense vehicle.



(a)



(b)

Fig.2 (a) Functional block diagram of system and (b) Developed prototype targeting system

A prototype was developed to test the theoretical model (Fig.2). The model uses stepper motors (0.9° resolution) controlled via real time embedded controller using ATMega microcontroller. The image processing is done on a host machine using MATLAB based algorithm. A radio link has been provided between the host computer and the onboard microcontroller for fast transmission of instructions. Experiments were conducted with the prototype to test the accuracy of theoretical model as well as to establish the sensitivity of error in target capture with multiple design elements, primarily size of the object, features of camera used, motor characteristics and distance of approach. The results have been summarized which can work as look-up graphs while designing future systems based on various cost and efficiency trade-offs. Finally, comparison of the experimental and theoretical results has been provided. Possible reasons for errors have been explored and further work required in the field has been summarized

III. Development of a Mathematical model for the Targeting system

A mathematical model was developed to determine the exact correlation between the angle moved by the motor and the linear distance subtended by the laser pointer. The model is first solved for the y-axis for the pointer location at every instant. We write equation of the line in 3-d between motor and pointer at the previous instant and at the current instant. Knowing the least count of the motor, we get one equation by finding the angle between initial and final lines. Also, the initial and final lines make same angle with the y-axis, since the rotation is considered purely along y-axis first.

As shown in Fig. 1, let the initial position of the pointer be (x_1, y_1, 0) and the final position be (x_2, y_2, 0) as shown. So, the direction ratios of lines are (x_i, y_i, -d). Let the angle between the lines be θ. This will obviously depend on the least count of the motor. Now the angle made by the lines joining the initial position line and the final line can be given as

$$\cos\theta = \frac{(x_1 * x_2) + (y_1 * y_2) + d^2}{\sqrt{x_1^2 + y_1^2 + d^2} * \sqrt{x_2^2 + y_2^2 + d^2}} \quad (1)$$

Let us consider the movement about y-axis. Thus, the angle made by the final position line with the y-axis will be same as that made by the initial position line with the y-axis. So, we have

$$\frac{y_1}{\sqrt{x_1^2 + y_1^2 + d^2}} = \frac{y_2}{\sqrt{x_2^2 + y_2^2 + d^2}} \quad (2)$$

At any instant, x_1 and y_1 shall be known (Initial position). x_2 and y_2 shall be supplied by the image processing module through the global intelligence. With the knowledge of approach distance 'd', rotation required by the motor in radians (θ) can be found out. A special case can be considered with the initial point fixed as (0,l,0). Then, we have:

$$\cos\theta = \frac{l*y_2+d^2}{\sqrt{l^2+d^2}*\sqrt{x_2^2+y_2^2+d^2}} \quad (3)$$

$$\frac{l}{\sqrt{l^2+d^2}} = \frac{y_2}{\sqrt{x_2^2+y_2^2+d^2}} \quad (4)$$

Let us assume, $\frac{l^2}{l^2+d^2} = c$,

Thus, we have

$$c = \frac{y_2^2}{x_2^2+y_2^2+d^2} \quad (5)$$

$$\left(x_2^2+y_2^2\right)*\left(1-\frac{1}{c}\right)+d^2=0 \quad (6)$$

Solving (3) and (6), we have

$$\cos\theta = \frac{\left(l*d^2\right)+\left(y_2*l^2\right)}{y_2*\left(l^2+d^2\right)} \quad (7)$$

'd' and '*l*' are correlated by cone of vision angle of the given camera as, tanα = *(l/d)*. Substituting this value back in (7), we get

$$\cos\theta = \frac{\left(d+y*\tan\alpha\right)*\sin\left(2\alpha\right)}{2y} \quad (8)$$

With known values of d (distance at the time of targeting) and α (Camera property), we thus have a correlation between displacement of the target pointer and the motor rotation in a particular degree of freedom (y axis in our case).

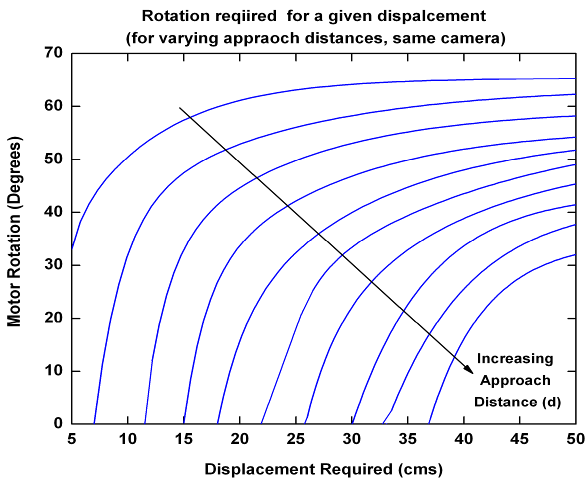


Fig.3 Relation between motor rotation and displacement required (Constant α, variable d)

Figure 3 depicts this relation for a family of approach distances 'd'. The camera cone of vision angle has been kept fixed at 50 degrees. It should be noted that as approach distance increases, i.e. the mechanism goes further away from the robot; lower values of displacement become impossible. Also, very few rotations are needed to cover a greater displacement on screen. Thus the accuracy suffers as we go farther from the target, given that the target size remains same. Similarly, for an increasing cone of vision, the angle needed to be rotated by the motor for the same displacement decreases as shown in Fig.4.

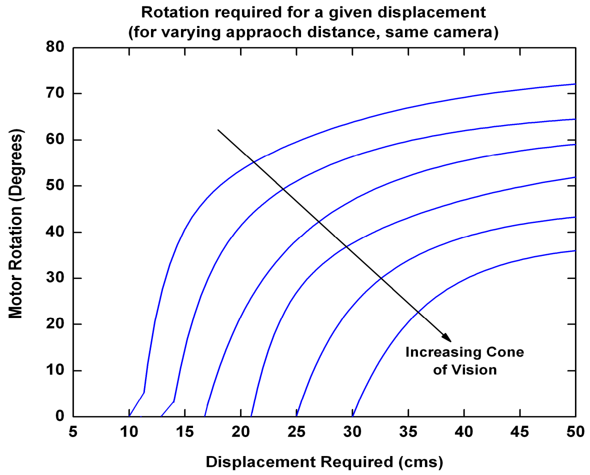


Fig.4 Relation between motor rotation and displacement required (Constant d, variable α)

IV. DETERMINATION OF TARGET APPROACH (D*)

We need to determine the optimum value of approach (d*) so as to minimize error. Farther we go from the target plane; more will be the displacement per unit motor rotation. However, nearer we go to the target, we lose more and more periphery data. Also, we could be limited in the approach due to safety reasons from the target. In this section, we find out an ideal value of approach distance 'd' in terms of target shape factor (a). For a target object with a shape factor of 'a' cms; the maximum possible error that can be tolerated will ±a. Thus, the least count of the motors should translate into this displacement in accordance to Equation.8.

$$\cos\theta = \frac{\left(\left(d+y\right)*\tan\alpha\right)*\sin\left(2\alpha\right)}{2y} \quad (9)$$

This can be restated as

$$d = \frac{2y*\left(\cos\theta+\sin^2\alpha\right)}{\sin\left(2\alpha\right)} \quad (10)$$

Substituting y = a, we have

$$d* = \frac{2a*\left(\cos\theta+\sin^2\alpha\right)}{\sin\left(2\alpha\right)} \quad (11)$$

Here, d* = Optimum approach distance. Thus, we can determine the approach distance for a given set of motors (θ), camera (α) and for a given target shape (a) which can be determined by initial image processing. Establishing the approach distance based on the size of the target is critical for precision positioning of the vehicle and forms a core input from targeting to locomotion in the overall system.

Figure 5(a) deals with this relationship for a given set of motors (i.e. fixed θ) but varying cameras (variable α). The graph stands to reason that as the cone of vision for the camera becomes narrow, it has to come closer to the target for the same level of accuracy, all other parameters remaining constant. Figure 5(b) deals with the relationship between the optimum approach distance and the size of the target, given that the camera is fixed but the motors are variable. The results indicate that there is no considerable difference for different motors. Thus, the camera specifications and target size affect the approach distance much more than motor selection.

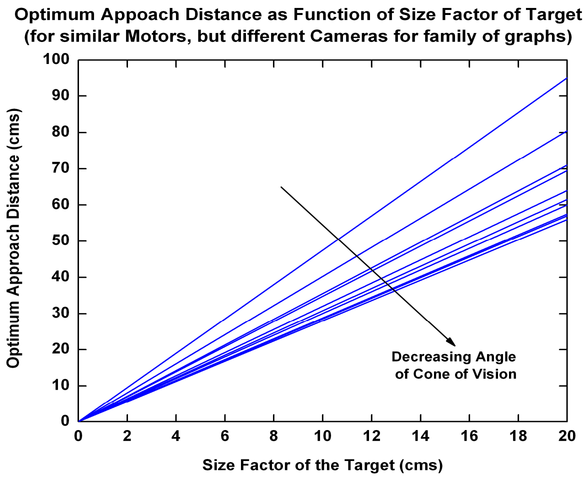


Fig.5(a) Relation between optimum approach distance and size of target object (variable α, constant θ)

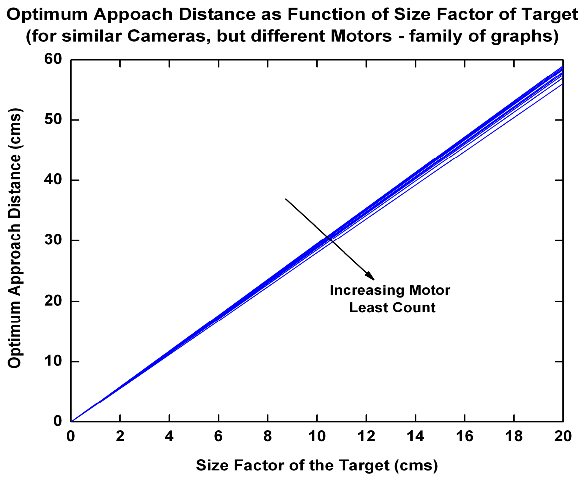


Fig.5(b) Relation between optimum approach distance and size of target object (constant α, variable θ)

V. PROTOTYPE SYSTEM DESIGN AND DEVELOPMENT

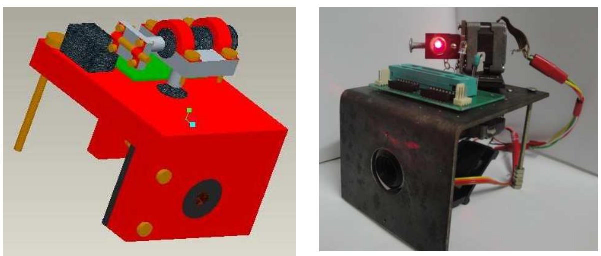


Fig.6 3D-Model and Prototype Targeting system

A targeting system prototype has been developed as shown Fig.6, to check the validity of the calculations. The targeting system has four modules: Mechanical module, Image processing module, RF communication link between the local targeting controller & image processing (global intelligence) system and Local target system controller module (microcontroller and motor driver circuit). The mechanical module is a 2-DOF system, which uses two stepper motors for rotation about two axis. A LASER pointer is used as a targeting pointer (representing Canon). A 3D model has been developed for the final system evolving from the initial prototype. Image processing module, uses a 2MP web camera for acquires image against a static background. A MATLAB based image processing algorithm has been developed used for capturing and processing static images for shape patterns, and calculations of their geometric parameters in MATALAB. The image processing process also includes noise cancellation, calculating the Euclidean distances for all the pixels in the image, comparison with a threshold for determination of different colours and arranging the matrix for determination of geometric parameters e.g. Centre of circle etc. The coordinates in pixels are converted to centimeters and then to pulses of the stepper motors as per equation 1. These calibrated values are sent via wireless communication to the micro controller, which runs the two stepper motors.

An air link based on radio communication is provided between the host computer and the microcontroller for fast and reliable transmission of instructions has been installed in the system. We used an integrated transceiver on both ends of the air link (msp430-cc2500 from TI ltd.). The transmission rate was fixed at 9600 baud. A basic schematic of the communication module is shown in the Fig.7. The module has a range of 10-15m, is bidirectional and operates on universal standards.

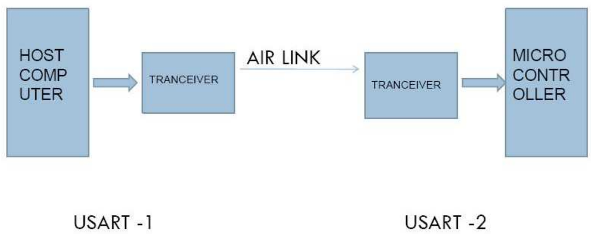


Fig.7 Schematic of the communication system

The motors are driven by quadruple half h-bridge motor drivers (L293-D). They are controlled using an 8-bit microcontroller (ATMEGA 16L). A set of embedded assembly instructions has been programmed into the microcontroller for driving the motors. The microcontroller uses USART (Universal Synchronous and Asynchronous Reception and Transmission) for real time interpretation of the signals received from the air link. It is enabled for interrupt-based operation to receive and interpret commands from the host computer.

VI. EXPERIMENTS

Experiments were conducted to validate the theoretical results. As specified earlier, these were conducted with a 2 DOF targeting mechanism with two stepper motors with minimum step being 0.9 degrees. The camera had a cone of vision angle of 27 degrees. A generic action flow sequence was followed to conduct the experiment. Each sectional experiment follows this action flow with some deviations particular to that section.

First step is to chose a target size, thereby fixing 'a'. Then, a camera is chosen with a certain cone vision angle value, α. Using these two parameters, we can look up figure 4, and find out the value for the optimum approach distance. The mechanism is then centered at that distance away from the target plane. The target is now randomly placed at any location against a static background, to reduce noise. The mechanism is now activated. The onboard camera scans the environment for target patterns and keeps updating the global intelligence module where image processing takes place, using MATLAB. Once the target is spotted, its size factor (a) is calculated using edge detection methods (Fig.8). The coordinates of the target are then calculated with respect to a floating reference point by the image-processing module. These coordinates are sent via a radio link to a microcontroller onboard the vehicle. The microcontroller uses interrupt based functioning to receive and interpret this message and accordingly activate the targeting canon motors, which make the system point at the centre of the target using algorithms based on Fig 3. The final position of the LASER pointer is noted along with the actual rotation of the motors. The theoretical equivalents of the same are calculated using Fig.3. Error in targeting is then calculated against different parameters.

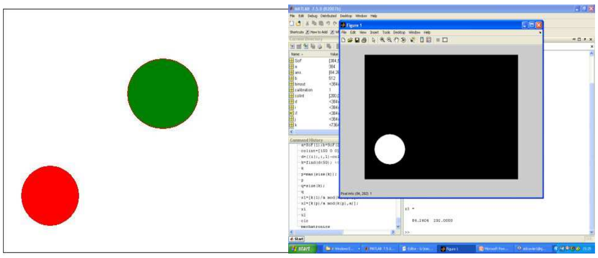


Fig.8 Target sample (Red circle) and Identified target After image processing

To test the system, the target was chosen as a red painted circle with a radius of 5cms. It was located at an arbitrary, but fixed position (1.5, -6, 0). As stated, the initial position of the pointer was kept at (0, l, 0) with *l* = 23cms here. Then, using the data from Fig.4, we have d* = 15cms, and y-displacement required = (y2-y1) = 21- (-6) = 27cms. Now, using data from the Fig.3, we calculate the number of rotations required for this displacement, θ (theoretical) = 64.5 degrees. Another set of experiments was conducted with d = 40cms. Again, θ (theoretical) = 28.3 degrees can be calculated using the data from the Fig. 3. The mechanism is now allowed to run and the final position of the LASER pointer is recorded. The distance of the pointer from the center of the target is noted as the error (Table 1). Similar experiments were conducted with variation of different design parameters. The results are summarized in the next section.

TABLE I: EXPERIMENTAL RESULTS (PILOT RUN)

α = 27 deg, y = 27cms				
Parameters	θ (Theory)	θ (Expe)	y (Expe)	Error
Units	(deg)	(deg)	(cms)	(%)
Experiments Set-1	64.5	64.8	27.3	1.11
(d = 15cms)	64.5	63	25.3	-6.30
Experiment Set-2	28.32	28.8	28.2	4.44
(d = 40cms)	28.32	25.2	25.6	-5.19

VII. RESULTS AND DISCUSSIONS

The experiments can be broadly classified in four groups. These explore the sensitivity of accuracy to: 1) Target size, 2) Use of different cameras (Variable cone of vision angle), 3) Use of different motors (Variable least count of motors) and 4) Approach distance. Each of these results, along with Fig.3 and Fig.4 can basically be used as look-up graphs while designing systems as per various requirements and cost trade-offs.

A. *Error variation with target size*

For this set of results, camera cone of vision, motor's degree per step, and distance of both from the target were kept constant. Only the target size factor was varied by varying the radius of the circle. This experiment is important to simulate the real life problem of various sizes of targets that the mechanism will have to face. For different sizes of the target, optimum approach distances were calculated and the mechanism was kept at those distances from the target. The distance of the final position of the LASER from the center of the target was noted and thus error dependence of target capture on size of the target was established.

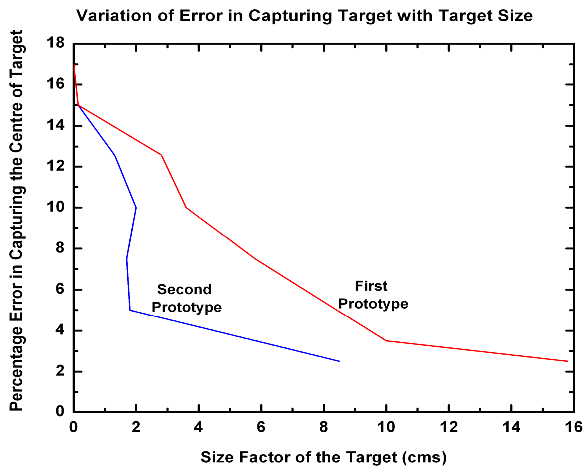


Fig.9 Variation in Error for different Target Size Factor

It was found that with the increase in target size, percentage error decreased uniformly. This may be due to the fact that with an increase in size the fractional error decreases. Also, in practical terms, deviation from center is more tolerable in a bigger target than a smaller one. Figure 9 shows the percentage of error variation with size of the object for two prototypes. Better manufacturing processes in the second prototype are also reflected in the graph as the percentage error decreases throughout the set as compared to the first prototype.

B. Variation in error with use of different cameras

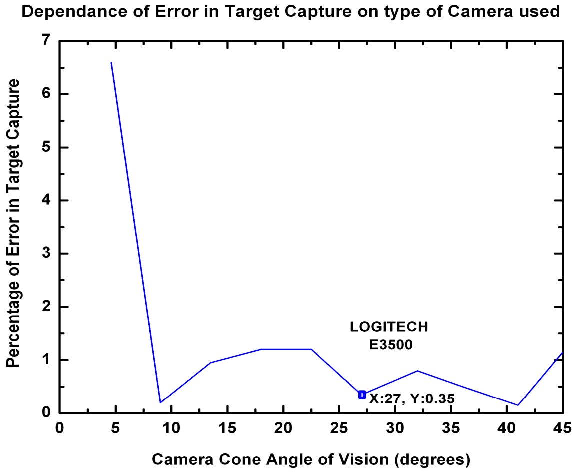


Fig.10 Variation in Error for different Camera Cone Angles

For this set of results, motor's degree per step, target size, and distance of robot from the target were kept constant. The camera cone of vision was changed. Herein, two different cameras Logitech Pro 9500, and Logitech e-3500 were used. The other results were simulated due to impracticality in changing cameras. Figure 10 shows the variation in error with cone angle of vision (α), which is unique for every camera. It is evident from the graph that the cameras with a very low value of α are prone to maximum error. This is probably due to the fact that the cameras with lower cone angle lack in capturing data by definition. Of course, the choice of cameras is a key cost versus efficiency trade-off and will change considerably for different environment (Constraints on distance of approach) and targets (Size factors, desired accuracy).

C. Variation in error with use of different motors

Camera cone of vision, approach distance, and distance of bot from the target were kept constant. Rotation amount by stepper motor with each step taken were varied. The results for least count less than 0.9 degrees were simulated. The use of stepper motor is very advantageous in these experiments because their least count can be set in software itself. Commercial stepper motors are available till 0.9 degrees per step.

The results for least count lesser than that had to be simulated. Figure 11 shows the variation in error with different least count for different motors. Commercially available stepper motors have a least count of 0.9 degrees per step. Below that, servo motors and some DC motors can be used depending upon the accuracy needed. The graph shows that the error decreases as a motor with better least count is used

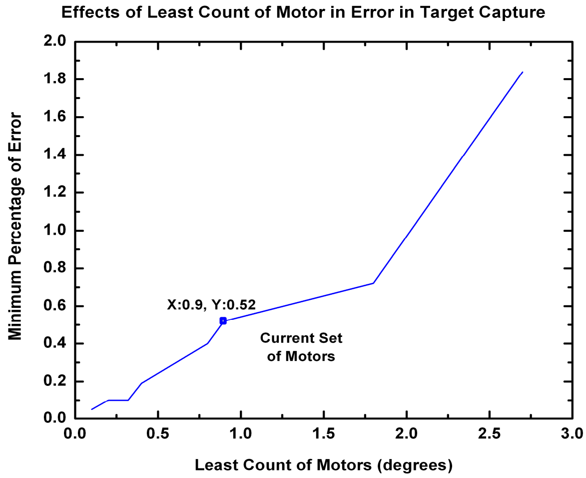


Fig.11 Variation of Error with different Motors

D. Variation of approach distance

Distance of approach is a critical parameter for controlling the error in the target capture. It depends upon the size of the target. More the distance of approach, more is the planar area available to the camera and more is the displacement of pointer per rotation. This may lead to increase in error. However, too close to the target will lead to problems like concavity of the motion of the pointer, and loss of data by the camera. Camera cone of vision, motor's degree per step, and the radius of the target were kept constant. Only the distance of mechanism from the target plane was varied. In real life application, this distance will indicate the distance at which the robot stops in front of its target to start the target capture process. This is perhaps the most important set of experiments, since it determines how close to the target the robot should reach for optimum targeting.

Two different subsets of experiments were conducted under this section to:

- Prove the existence of an optimum approach distance. Herein, the same target was captured at various approach distances and the trends were noted. Figure 12 shows the variation in error as the distance of approach is changed.
- Establish error dependence on optimum distances. Although this is a derivative from the size of the target, it becomes important since it will be the onboard knowledge available to the crew. Figure 12 proves the existence of an optimum thus validating the theory.

From the results, it may be noted that a major source of error is the quantized motion of the stepper motors due to which the least count of the mechanism is very high. This problem can be sorted out by using servo motors.

Other sources of error include manufacturing defect due to brazing operations on the bushes. Lack of calibration also adds to the error. It should also be noted that the errors are with respect to an idealized theoretical model as developed, which has inherent assumptions, which may not always comply.

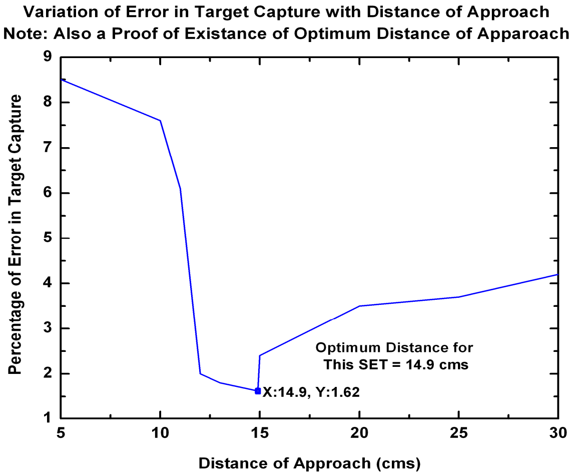


Fig.12 Variation in Error for different Distance of Approach

VIII. Conclusion

For a monocular vision based targeting system**,** theoretical expressions have been arrived at for control of motors for accurate capture of target without image capture of the targeting pointer itself. Theoretical expressions have also been determined for the approach distance in dependence only on camera, motor specifications and the target size. Experiments conducted with a prototype show error within acceptable limits establishing the accuracy of theoretical expressions. Sensitivity of error on various design factors like quality of camera, least count of motors, allowed distance of approach etc. has been explored in detail. Further work needs to be focused towards real time processing of moving targets and development of heuristic algorithms for better control.

IX. References

[1] Johann Borenstein, "Internal correction of dead-reckoning errors with a dual drive compliant linkage mobile robot," Journal of robotic systems, 12(4):257–273, 1995.

[2] E.Triendl D.J.Kriegman, "A mobile robot: sensing, planning, and locomotion," Springer-Verlag New York, Inc, New York, NY, USA, 1990.

[3] H. Arakawa et al. T. Hongo, "Autonomous robot vehicles," Springer-Verlag New York, Inc, New York, NY, USA, 1990.

[4] Wang kangling Fu Xiaowei, "A feature-based multi-target tracking algorithm," Computer Engineering and Applications, 2004.

[5] Canny J.A., "Computational approach to edge detection," IEEE Trans. Pattern Analysis and Machine Intelligence, 1986.

[6] Hildreth E. Maar, "Theory of edge detection," Proceedings Royal Soc. London, 1980.

[7] Arun D Kulkarni, Computer Vision and Fuzzy-Neural Systems, Prentice Hall, 2001.

[8] Changhong Yu Zhiyong Wang, Tianding Chen, "Binocular stereo vision of moving target detection," Proceedings of the 2009 International Symposium on Information Processing, 2009.

[9] Huadong Wang et al., "Scalable target detection for large robot teams," Proc. of the 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI-2011), Lausane, Switzerland, March, 2011, pp.363-370.

[10] Toby P. Breckon, Ji W. Han and Julia Richardson, Consistency in Multi-modal Automated Target Detection using Temporally Filtered Reporting, UK MoD Centre for Defence Enterprise (CDE) (Res. Contract number:11657).

[11] Manolakis, Marden and Shaw. Hyperspectral Image Processing for Automatic Target Detection Applications. Lincoln Laboratory Journal, Volume 14, Number 1, 2003

[12] Chein-I Chang, Daniel C. Heinz, "Constrained Subpixel Target Detection for Remotely Sensed Imagery," IEEE Transactions On Geoscience And Remote Sensing, VOL. 38, NO. 3, MAY 2000.