Visual Object Recognition, Tracking and Destruction for Surveillance System

Ms. Rajashree D. Gade
PG student, Department of Electronics
DPCOE, Wagholi
Pune, India
rajashree.gade25@gmail.com

***Abstract* — This paper is proposed for the purpose of detection, tracking and destruction of intruding object. The system will be mounted at suitable place from which complete and clear view of the area under surveillance can be captured with camera. It will detect intrusion and recognize the intruding object by comparing its features with features of the objects stored in database. If feature match is found, the intruding object will get tracked to find its velocity and get bombarded with bullets and bombs until object gets destroyed completely. Thus image will be captured, proceed and desired action will be performed on it. This system works as automatic detection, recognition and destruction of intrusion which is very helpful for military application. The very tight security and safety can be assured without endangering precious life of human solider. In the implementation point of view, we keep the system as simple as possible. Our system ensures low execution time. Simplicity also ensures low implementation cost.**

Keywords— Object tracking, Object recognition, and autonomous cannon

I. INTRODUCTION

This system is proposed for detection, tracking and destroying of intruding object. The system will be fixed at suitable location, from which complete and clear view of the surveillance area will cover and images of intrusion under this surveillance area can be captured with camera. The system is consist with a high resolution camera, image processing hardware, microcontroller, two servo motors and other supplementary hardware and mechanisms. The complete assembly set up is shown in Fig.1. Image Processing Hardware will acquire and stored images captured by camera. The images will be taken at some predefined interval of time. Then captured image will process for detecting intrusion. We have collected database for the objects those are to be destroyed. If captured image of intrusion is detected by Image Processing Hardware will extract the features of that intruding object and compare them with features of objects stored in database. If intruding object is matched with any one of the objects stored in database, object is said to be recognized. System will track that object to calculate its velocity of motion. This velocity information is needed to decide the angle and time instant at which projectile is to be launched at intruding object to destroy it. Position of the intruding object in the form of x-y co-ordinate is found and sent to microcontroller. Microcontroller will control the angle of rotation of two Servo Motors to position the cannon aiming at the intruding object. At last cannon will get fired.

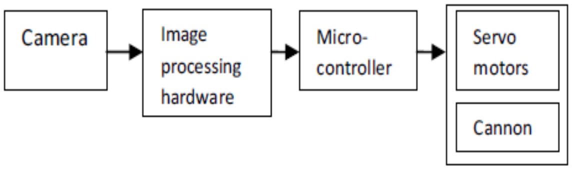


Fig1. Basic building blocks

II. ISSUES IN VIDEO SURVEILLANCE SYSTEMS

Motion detection and object tracking is a very rich research area in computer vision. The main issues that make this research area difficult are:

A. *Computational Expense*

If an algorithm for detecting motions and tracking objects is to be applied to real-time applications, then it needs to be computationally inexpensive so that a modern PC has enough power to run it. Yet many algorithms in this research area are very computationally expensive since they require computing values for each pixel in each image frame.

B. *Moving Background Rejection*

The algorithm also needs to be able to reject moving background such as a swaying tree branch and not mistakenly recognizes it as a moving object of interest. Misclassification can easily occur if the area of moving background is large compared to the objects of interest, or if the speed of moving objects is as slow as the background..

C. *Tracking Through Occlusion*

Many algorithms have devised ways of becoming robust against small occlusions of interested objects, but most algorithms still fail to track the object if it is occluded for a long period of time.

D. *Modeling Targets of Interest*

Many algorithms use a reasonably detailed model of the targets in objects detection and consequently require a large number of pixels on target in order to detect and track them properly. This is a problem for real-world applications where it is frequently impossible to obtain a large number of pixels on target.

E. Adapting to Illumination Variation

Real world applications will inevitably have variation in scene illumination that a motion detection algorithm needs to cope with. Yet if an algorithm is purely an intensity based method then it will fail under illumination variation.

F. Analyzing Object Motion

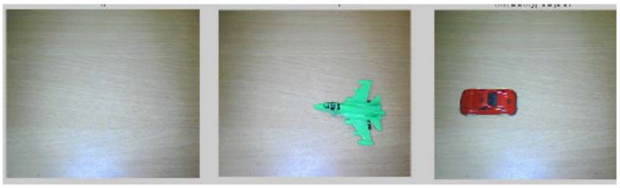
After objects have been correctly classified and tracked, an algorithm may want to analyze the object motions such as the gait of moving human, the speed of car (is it speeding or not?) etc. This could be difficult especially for non-rigid objects such as human if the object view is not in the right perspective for the algorithm.

G. Adapting to Camera Motion

Detecting moving entities from mobile camera video streams still remains a challenge in this research area

III. DATABASE COLLECTED

It shows captured images of background and two objects aero-plane and car. These are test objects taken for experimentation purpose.



a) Background Image b) Database object 1 c) Database object 2

Fig 2. Shows example of collected database.

IV. METHODOLOGY

Fig1. Shows basic building blocks of the complete system. Here is block-by-block explanation.

A. Camera

Camera used for experimentation purpose is "i-ballface2face, a USB webcam, with 640 by 480 resolution and64M color depth. A Night-Vision Camera and camera with different resolution and color depth can be used depending upon requirement of application. Camera is fixed on the system and should not move from its place once background image is captured, otherwise it will adversely affect accuracy of the system as subtraction is used to detect intruding object.

B. Image processing hardware

A computer with Intel 1.6 GHz processor and 512MBRAM was used as Image Processing hardware. Camera is connected to this PC via USB port. The image acquired by camera is processed by this hardware and result of processing is sent to Microcontroller. Microcontroller will control the angle of rotation of servo motor to position cannon accordingly.

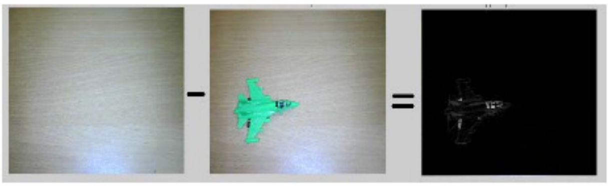


Fig 3. Subtraction of image captured by camera from background image to segment/detect intruding object.

C. Image processing software

When an object is observed form far distance, its overall shape and color is the prominent feature which distinguishes it from background and other objects. So while extracting features of the object for identification, its peripheral shape and color are considered, instead of other features like texture and other minute details. The complete algorithm for this prototypic system is implemented in Matlab7.2 software using Image Processing Toolbox.

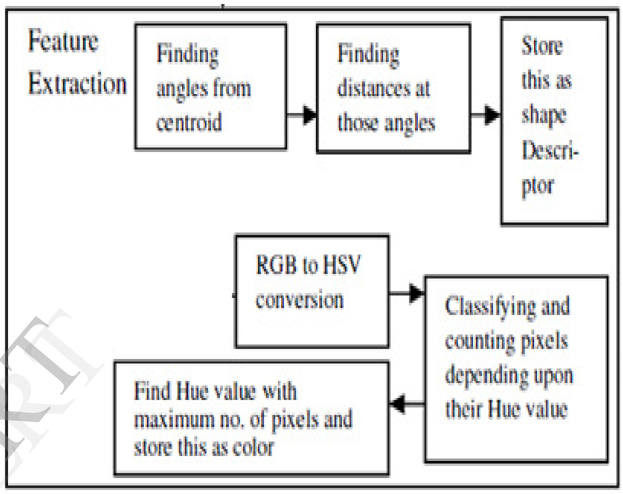

Feature Extraction
Finding angles from centroid
Finding distances at those angles
Store this as shape Descri-ptor
RGB to HSV conversion
Classifying and counting pixels depending upon their Hue value
Find Hue value with maximum no. of pixels and store this as color

Fig. 4 Block diagram of future extraction

D. Steps involved in image processing

1) Preprocessing

Background image has to be captured after installing camera at its place and care must be taken that camera shouldn't move once background is captured. Subtraction between background image and current image obtained from camera is used to detect intrusion. If there is no intrusion then previously captured background image and image taken from camera at any later time will have no difference and result of subtraction will be zero (a complete black image). But if some object has intruded in the scene then difference between those two images will be the object itself and is not zero, as in previous case. This subtraction process is shown in Fig.3. Result of subtraction is not suitable for feature extraction as outer edges of intruding object are not clearly visible. This problem can be depicted from Fig.5.

This is image obtained from inverting the result of subtraction which prominently indicates above mentioned problem. So this result of subtraction has to be preprocessed before feature extraction. Preprocessing involves following steps. As subtraction is obtained by subtracting color images, result of subtraction is also a color image. It is converted into binary image as shown in Fig.6.a. Our final aim of image preprocessing is to obtain image shown in Fig.6.f which is suitable for feature extraction. Canny edge detection is function is operated on Binary Image to detect the edges of the intruding

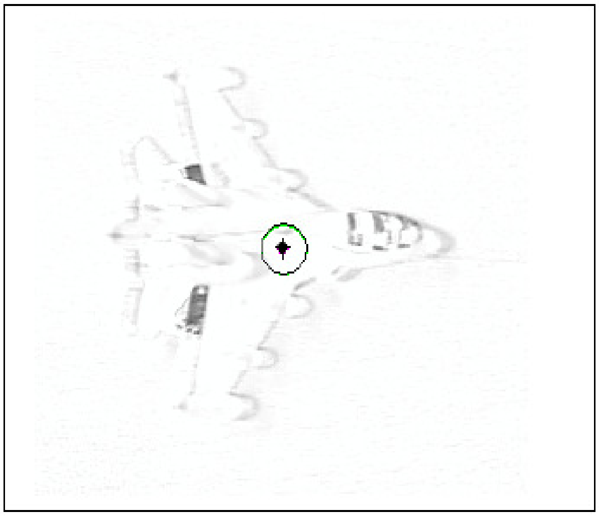


Fig. 5 complement of subtracted image

object depending upon threshold chosen adaptively. Correct choice of threshold leads to proper edge detection, which will increase overall accuracy. Problem of unclear and broken boundary as depicted in the Fig.5 is present till this stage. To remove this problem image is dilated with some suitable mask to join these broken edges. The output of dilation is shown in Fig.6.c. Holes in binary images are black portion of image surrounded by white boundary. These holes are filled with filling operation to get number of different unconnected white areas. These white areas are related to different objects in image. At this stage, we get various unwanted white regions other than white region related to the intruding object. These unwanted white regions are result of illumination variations, minor changes in the background at the time of capturing images, camera imperfections etc. These unwanted regions shave inherent property that their area is lesser than intruding object. So only white region with maximum area is kept which corresponds to intruding object as shown in Fig.6.e.Edge of this image is detected using canny edge detection method. This edge corresponds to boundary of intruding object with some error introduced due to dilation operation. Error that is less than 13% is proved to be acceptable. So mask for dilation operation must be chosen adaptively depending upon size of objects. This edge detected image is suitable for shape detection.

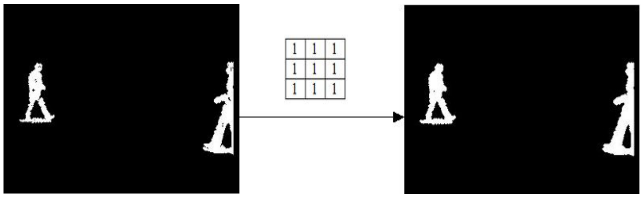
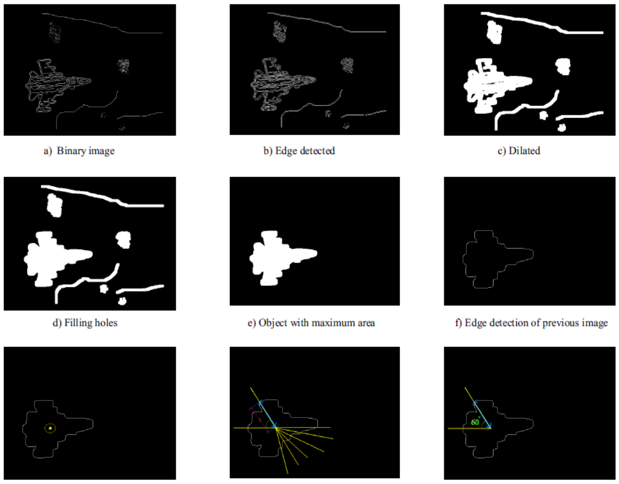


Fig 6g Example of dilation operation



a) Binary image

b) Edge detected

c) Dilated

d) Filling holes

e) Object with maximum area

f) Edge detection of previous image

g) centroid obtained and located

h) Finding distance between centroid and points of edge located at various angles

Fig 6 various steps involved in image processing and shape detection

2) *Feature extraction* : Followings steps are involved in feature extractions

a) *Shape description:* Shape of an object is nothing but distances of all the points on its boundary from some reference point. This reference point can be centroid (center of mass) of an object Centroid of object does not change though object is rotated. Center of circle is its centroid and distances of all points from centroid are equal. For square it will be different case. Similarly if we measure distances of some points on the boundary of an object as shown in Fig6.h then we can get shape descriptors. In our case we have considered those points on the boundary of the object which are separated by angle of 10 degrees. All the angles are measured from centroid of the object. Thus we have calculated 36 distances corresponding to 36different angles separated by 10 degrees. This angle separation can be reduced in order to increase accuracy. But along with reduction in angle separation, number of readings will increase and it will increase computation time. So there is tread off between ability of system to work in real time and its accuracy. Normalization of the data obtained above is done in order to enable scale invariance. Object viewed from various distances will not differ in their shape. But they will differ there in sizes. Normalization will enable comparison between objects those are present at various distances from camera.

Normalization is done by dividing all above 36 distances readings by largest distance reading. This results in get all 36 shape descriptors readings to range from 0 to1.Again shape descriptors obtained above can be made invariant to rotation of the object. To make it starting point or rotation invariant, circular shifting of readings (which issued in chain codes) can be implemented. This enables to compare objects having different rotational orientation. Thus we have made this system rotation invariant and scale invariant. Thus shape descriptor of the object is obtained.2.b Color detection: If we observe some object from far distance, we consider its gross features instead of fine details. If that object is having different colors on its different parts then the color which is occupying maximum area of that object is considered, and it is said that object that object is of that color. e.g. Fig7.a shows object is

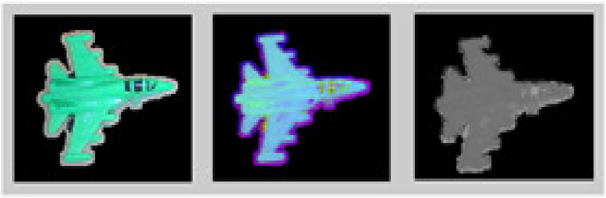


Fig 7 a) object b) HSV transformation c)only Hue plane

having green color covering most of its parts, and blue and black colors covering only some of its portion. So color of that object is considered to be green only. To find color of the object, color image obtained from camera is logically ANDed with preprocessed image shown in Fig.6.e. Result of ANDing is shown in Fig.7.a.Now this resulting image is converted into HSV image as shown in Fig7.b. Hue plane of HSV image contains only color information. All the values of Hue plan lies between 0 to 1. Depending upon their values, color is detected. E.g. red, green and blue colors can be distinguished as in HSV plane as red pixels have values >0.8 and <0.15, green pixels have values >0.15 but <0.48, blue pixels have values>0.48 but <0.8. Thus second gross feature of the objects that is its color is detected. Color and shape detection is also performed on database images. If match between any on the database object and intruding object is found then that object is said to be recognized. And it has to be destroyed. Fig.8 shows result of preprocessing and feature extraction on some sample database image.

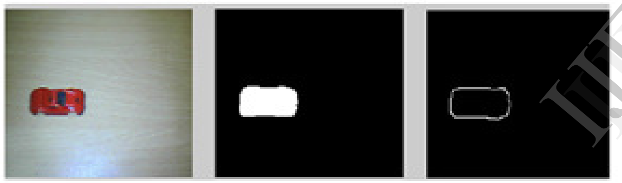


Fig 8 Preprocessing and feature extraction on database image

E. *Object tracking*

Once object is found to be matching with one of the database objects, it is tracked to find its velocity, before bullet or cannon-bomb is fired at it. Both reading, size of object and its distance from camera is obtained and stored in database beforehand. Perpendicular distance of intruding object from camera can be obtained by comparing its size in image, with size of the same object stored in database using formula given below,

$$\text{Distance} = \frac{\text{Distance of database object + size of intruding object}}{\text{Size of database objet}}$$

Velocity of intruding object can be found out by, finding relative distance covered by object in two consecutive frames captured by camera and time interval between two frames. This calculation of velocity and distance of object enables system to calculate angle of projectile (Bullet or Bomb) launch and time at which projectile has to be fired. This ensures accurate hitting of object by projectile and its destruction. The system will track and keep launching projectiles at intruding object until it gets destroyed and its motion stops.

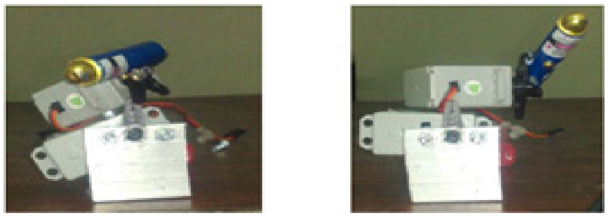


Fig 9 Servo motors mechanism for decoding angle of projectile launch

F. *Servo motor mechanism*

Two servo motors, one responsible for motion of cannon in X-direction and other in Y-direction together decides angle of projection. Angle of rotation of servomotor can be changed by changing width of PWM signal fed to motor. 89v51RD2 microcontroller with inbuilt PWM module is used in prototype system. The results obtained after Image Processing are fed to microcontroller which in turn controls firing angle and time for projectile launch.

V. RESULTS & CONCLUSION

15 different objects were stored in database. Out of total 80 trials, system correctly recognized, tracked and destroyed intruding objects in more than 74 trials. This brings 93% accuracy. For shape description, total 36 distances corresponding to 36 angles were calculated. If total 32distance readings of intruding object and database object are matched then object is said to be recognized. If number of readings used for shape description is increased, then it drastically improves accuracy. More readings are taken, more accurately shape of the object can be described. If number of readings taken is doubled it increases total execution time by 62%.A graph of number of readings verses execution time in milliseconds is shown in Fig. 9Above graph is plotted using number of readings for shape Figure 10. A graph of number of readings verses execution time inmilliseconds1488 2011

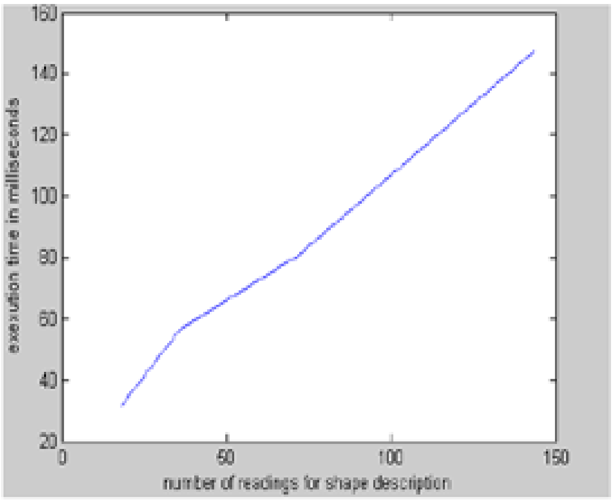


Fig 9a Graph of number of reading verses execution time in millisecond

Above graph is plotted using number of readings for shape description as 18, 36, 72, 144 and their corresponding execution time in milliseconds as 31.83, 56.49, 81.12, 147.64.Advanced algorithm such as SIFT (Scale Invariant Feature Transform) is used for object recognition purpose, its execution time found to be 950ms on the same system. So the algorithm used in this paper is at least 7 times faster than SIFT without much affecting overall required accuracy. Though this system is tested under constrained environment, subtraction between background image and camera captured image ensures that similar accuracy of the algorithm can be guaranteed when images captured are real-life images. Thus by implementing this automatic system one can ensure complete intrusion free area under surveillance. so, the precious life of human soldier can be save using this system. The same work can be done providing very tight security and safety.

REFERENCES

[1] He D.; Hujic D.; Mills J.K.; Benhabib B. "Moving object recognition using premarking and active vision"; Robotics and Automation, 1996. Proceedings. 1996 IEEE International Conference.

[2] Diplaros, A.; Gevers, T.; Patras, I.; "Combining color and shape information for illumination-viewpoint invariant object recognition", Image Processing, IEEE Transactions on Jan. 2006 Volume: 15,Issue:1

[3] Jae-Han Park; Seung-Ho Baeg; Jaehan Koh; Kyung-Wook Park; Moon-Hong Baeg; "A new object recognition system for service robots in the smart environment" Control, Automation and Systems,2007. ICCAS '07.

[4] Chyi-Yeu Lin; Setiawan, E. "Object orientation recognition based on SIFT and SVM by using stereo camera", Robotics and Biomimetics,2008. ROBIO 2008. IEEE International Conference.

[5] [Chensheng Wang; Fei Wang; "A Knowledge-Based Strategy for Object Recognition and Reconstruction"; Information Technology and Computer Science, 2009.ITCS 2009.

[6] Jen-Shiun Chiang; Chih-Hsien Hsia; Shih-Hung Chang; "An efficient object recognition and self-localization system for humanoid soccerrobot", SICE Annual Conference 2010.

[7] Yinghua Xue; Guohui Tian; Rongkuan Li; Haitao Jiang; "A new object search and recognition method based on artificial object markin complex indoor environment", Control Conference, 2008. CCC2008.

[8] Seung-Ho Baeg; Jae-Han Park; Jaehan Koh; Kyung-Wook Park; Moonhong Baeg; "An object recognition system for a smart home environment on the basis of color and texture descriptors", Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference.

[9] Johnson, A.E.; Hebert, M. "Using spin images for efficient object recognition in cluttered 3D scenes ", Pattern Analysis and Machine Intelligence, IEEE Transactions.

[10] Jianying Liu; Pengju Zhang; Fei Wang; Dept. of Electrical Eng., Inner Mongolia Vocational Coll. of Mech. & Electr. Technol.,Huhhot, China "Real-Time DC Servo Motor Position Control by PID Controller Using Labview", Intelligent Human-Machine Systems and Cybernetics, 2009. IHMSC '09.

[11] [11] Shimada, A.; Kishiwada, Y.; Arimura, M.; "AC servo motor position sensorless control using mechanical springs "; AdvancedMotion Control, 2006. 9th IEEE InternationalWorkshop.