Multiple Moving Object Detection and Tracking in a Video Sequence

Amit S. Kakad
Dept of Electronics and Telecommunication,
Mauli College of Engineering and Technology,
Shegaon, India
ameetkakad@gmail.com

Dr C. M. Jadhao
Principal and Professor,
Dept of Electronics and Telecommunication,
Mauli College of Engineering and Technology,
Shegaon, India
cmjadhao@gmail.com

Prof. S. S. Mhaske
Dept of Electronics and Telecommunication,
Mauli College of Engineering and Technology,
Shegaon, India
mhaskesantosh12@gmail.com

Prof. N. B. Bhawarkar
Dept of Electronics and Telecommunication,
Mauli College of Engineering and Technology,
Shegaon, India
nb.bhawarkar@gmail.com

Abstract— After the development of computers and later on of fast processing software's, video processing has found a vital place in the area of research. Also due to its presence in security related applications, many countries in the world has made it of utmost importance for them. This paper explains an algorithm which takes the video sequences as an input and starts extracting the frames from these sequences. After extraction, equalization of each and every frame is done to have a well-defined input for processing it further. Through detection tracking algorithm we have detected and tracked the moving objects until the vision of the camera. Further these detected objects were classified according to shape based criterion. Detection of objects took place in two different approaches; one was feature as a color based and second was multiple objects detection.

Keywords-*MMODT, Object detection, object tracking, object classification*

*****

I. Introduction

In imaging science, image processing is processing of images using mathematical operations by using any form of signal processing for which the input is an image, a series of images, or a video, such as a photograph or video frame; the output of image processing may be either an image or a set of characteristics or parameters related to the image. Most image-processing techniques involve treating the image as a two-dimensional signal and applying standard signal-processing techniques to it. Images are also processed as three-dimensional signals where the third-dimension being time or the z-axis.

Video Analysis has several advantages and has tremendously grown in last decade or so. Video processing has several ranges of techniques which include frame enhancement, frame restoration, frame reconstruction, feature Extraction and recognition and compression. Video based moving object detection is one of the applications of video analysis which has found its existence in modern Video Surveillance Systems. Human being recognizes a multitude of object in images with little effort, despite of the fact that image of an object may vary somewhat in different viewpoints, in many different sizes and scales or even when they are translated or rotated.

Object detection and tracking refers to finding the location of an object when it moves from frame to frame. In case of tracking, object could be anything like vehicles on road, plane in sky or person walking on road. No of tracking algorithms have been developed [1][2]. Obviously to detect something in a video sequence we need a clear and enhance video for processing for which equalization of input sequence needs to be done [3] [4]. Particle filter is also a good option for these algorithms [5].

II. Frame Extraction and Histogram Equalization

The quality of video service gets degraded by several technical limitations such as poor lightening conditions, bad exposure level, and unpleasant skin color tone. Focus is on overall intensity level of the captured frames from the camera. Measuring the lowest intensity and highest intensity level value and to distribute evenly over the region in order to get the balanced frame intensity level [3]. After this evenly distribution of intensity level striking difference is noted between the original video sequence and the equalized video sequence.

Let f be a given image represented by pixel intensities ranging from 0 to L − 1. L is the number of possible intensity values, often 256. Let p denote the normalized histogram of f with a bin for each possible intensity.

Histogram for an image f is given as

$$p_n = \frac{no\ of\ pixels\ with\ intensiy\ n}{total\ no\ of\ pixels}$$

The histogram equalized image g will be defined by

$$g_{i,j} = \text{floor}((L-1)\sum_{n=0}^{f_{i,j}} p_n)$$

Where L is no of possible intensities values, floor() rounds down to the nearest integer.

III. OBJECT DETECTION

A. Feature Based Detection

Image as in frame for video is a different kind of data which includes a huge amount of information, such as color information, objects, edges, pixel definition, dimensions and others [6]. Commonly used well-known color spaces include RGB& CMY. To represent the RGB space, a cube can be defined on the R, G, and B axes. White is produced when all three primary colors are at M, where M is the maximum light intensity, say M=255. The main diagonal axis connecting the black and white corners defines the Intensity.

$$I(R,G,B)=R+G+B$$

From these intensity levels each of the specific intensity level is subtracted and the result we get is the specific color detected object.

B. Motion Based Detection

Importance of object detection and tracking is due to its enhanced automation in public security surveillance as well as in traffic control and pedestrian flow analysis [7]. Background subtraction, also can be said as Foreground Detection, is a technique wherein an image's foreground is extracted for further processing. Generally an image's regions of interest are objects (humans, cars, text etc.) in its foreground. After the stage of image preprocessing object localization is required which may make use of this technique. Background subtraction is an approach for detecting moving objects in videos from static cameras. The rationale in the approach is that of detecting the moving objects from the difference between the current frame and a reference frame, often called “background image”, or “background model”.

Background Subtraction algorithm follows the logic of Exclusive OR Gate, Truth table for which is as given below in Table No.1

Background Frame	Next Frame	Detected Objects
0	0	0
1	0	1
0	1	1
1	1	0

Table No.1 Truth Table for Background Subtraction

Now what the above logic describes is that if there is any change or deviation with respect to background frame than that change is subtracted from next coming frame and result is the intruder i.e. the objects are detected. Flow Graph for Background Subtraction is as shown in following Fig.1

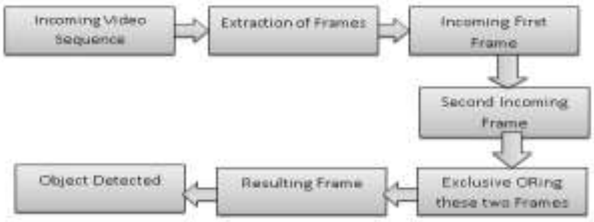


Fig.1 Flow Graph for background subtraction

Blob detection is aimed at detecting regions in a digital image that differ in properties, such as brightness or color, compared to surrounding regions. Informally, a blob is a region of an image in which some properties are constant or approximately constant; all the points in a blob can be considered in some sense to be similar to each other.

IV. OBJECT TRACKING

Once the object is detected, next task is to do the motion analysis.It finds its application in tracking the objects in bad weather conditions. So a adaptive algorithm is required to track the required objects. Here two tracking filters are compared i.e kalman and particle filter [8]. Particle Filter is more adequate for multitracking tasks in complex situations, and a Kalman Filter is to be chosen as solution in simpler ones, such as surveillance applications in low populated areas. Such kind of recursive algorithm is provided by a Kalman filter [9].

Kalman filter is a recursive predictive filter that is based on the use of state space techniques and recursive algorithms. It estimates the state of a dynamic system. This dynamic system can behave abnormally due to some noise which is generally white noise. To improve estimated state Kalman filter uses measurements that are related to state but are disturbed as well. Kalman filter consists of two steps

1. Prediction
2. Correction.

Logical flow for kalman filter is as shown in Fig. 2

Fig.2 Logical flow for kalman filter

In the case of kalman filters each iteration begins with predicting the process’s state using a linear dynamics model.

A) State Prediction: For each time step T, a Kalman filter first makes a prediction Xt and it is given by

$$X_t = AX_{t-1} + Bu_t$$

X_{k-1} is process state at time t-1, A is process transition matrix and B is control vector. Ut which converts control vector into state space

B) Error Covariance Prediction: The Kalman filter concludes the time update steps by estimating the error covariance forward by one time step

$$P_t = AP_{t-1}A^T + Q$$

P_{t-1} is a matrix representing error covariance in the state prediction at time t-1and Q is the process noise covariance. Lower the value of prediction error covariance Pt we can trust more on prediction of the state X_t. If the process is precisely modeled, then the prediction error covariance will be low.

By using the time update step kalman filter will predict the state xt -1 and find the error covariance at time k. then after during the measurement update steps kalman filter uses measurements to correct its prediction

A). Kalman Gain: Kalman filter computes a Kalman gain K_t, which is used for correcting the state estimate X_t.

$$K_k = P_k H^T (HP_k H^T + R_k)^\wedge - 1$$

Where H is a matrix used for converting into measurement space from state space and R is measurement noise covariance.

B) State Update: Using Kalman gain K_t and measurements Z_t from time step t, where we can update the state estimate:

$$X_t = X_{t-1} + X_t(Z_t - HX_{t-1})$$

C). Error Covariance Update: The final step of the Kalman filter's iteration is to update the error covariance P_t

$$P_t = (1 - K_t H)P_{t-1}$$

If the measurements are accurate then the updated error covariance will be decreased.

V. OBJECT CLASSIFICATION

Before thinking about classifying the objects in a video sequence, we need to take into consideration overall conditions for classifier [10]. Following conditions needs to be understood

A. Performance under real time constraints
It talks about the generic constraints put up by a system while it is being operated in real time.
B. Being robust to several conditions
It focuses its attention towards parameters like changes in lighting and atmospheric conditions, shakes of camera due to strong winds, color variation, spurious object landing in camera etc.
C. Resolving a multi class problem
It talks about the classes that are going to be considered for the classification. It discusses about various groups on which classification can be based.

Further different classification techniques were studied [11]. Object Classification methods are summarized in following Fig.3

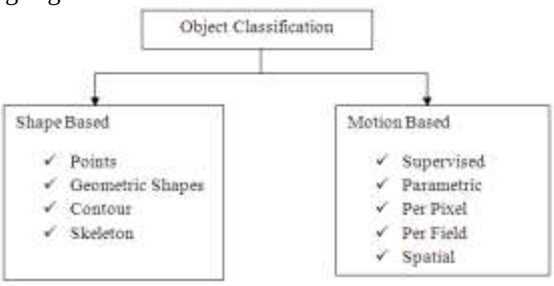


Fig.3 Object Classification Methods

As seen above the object classification techniques are broadly classified in two categories

A. Shape Based
B. MotionBased

In shape based category it makes use of the objects' 2D spatial information. Common features used in shape-based classification schemes are the points (centroid, set of points),primitive geometric shapes(rectangle or ellipse), skeleton, silhouette and contour.
Motion based category uses temporal tracked features of objects for the classification. They are further divided into following types

1. Training sample used
Here in supervised classification the process of using samples of known informational classes (training sets) to classify pixels of unknown identity, whereas in unsupervised classification it examines a large number of unknown pixels and divides it into number of classes based on natural groupings present in the image values.

2. Assumption on Parameter on Data
In Parametric classifiers parameters like mean vector and covariance matrix are used. There is an assumption of Gaussian distribution whereas in Non Parametric classifier there is no such assumption and no usage of statistical parameters.

3. Pixel Information Used
In per pixel classifier methodologies like Maximum Likelihood, Support Vector Machine and ANN are used whereas Sub pixel classifiers has the capability to handle the mixed pixel problem, suitable for medium and coarse spatial resolution images.

4. Object Oriented Classifiers
Pixels of the image are united into objects and then classification is performed on the basis of objects. It involves 2 stages: image segmentation and image classification Image segmentation unites pixels into objects, and a classification is then implemented on the basis of objects.

5. Spatial Information Classifiers
This classification uses both spectral and spatial information initial classification images are generated using

parametric or non-parametric classifiers and then contextual classifiers are implemented in the classified images.

Combination of these techniques led to development of our logic of classification i.e. classifying the objects on basis of size and structure. Here area of each object is calculated and a threshold is defined for each kind of object detection.

VI. RESULTS

Here we have dealt with frame extraction, histogram equalization of these extracted frames, detection of object on basis of colors, Generalized object detection and tracking and then classifying them on certain parameters. Fig.4 shows the result for frame extraction for an video sequence

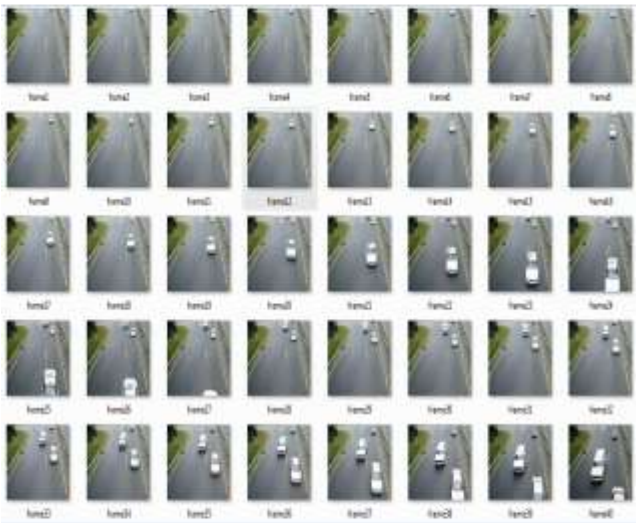
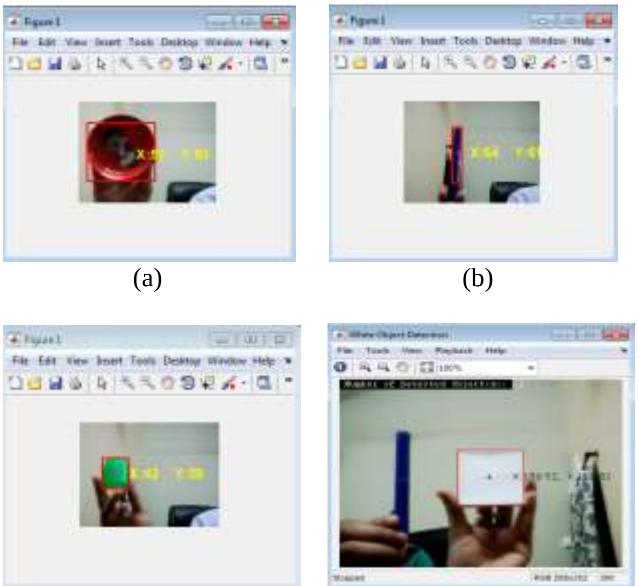


Fig. 4 Extracted frames for video sequence

Following Fig.5 shows detection of red, blue, green and white colored object.



(a) (b)

(c) (d)

Fig.5 Individual Red, green, Blue and White Color object detection

Following Fig. 6 shows red, grren and blue colored object in one sequence

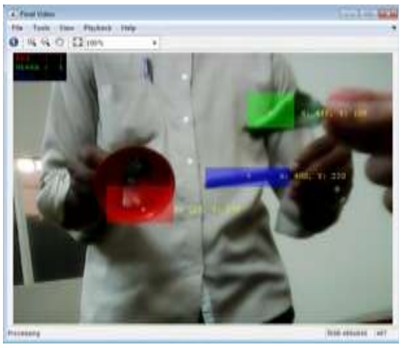


Fig.6 Red, Green and Blue colored object detection

Following Fig.7 shows object and blob detection through background subtraction

(a) (b)

Fig.7 Object and Blob detection

Following Fig.8 shows multiple object detection and tracking in a video sequence

(a) (b)

Fig.8 Multiple Object detection and tracking

Following Fig.9 shows object classification in video sequence

Fig.9 Object classification in video sequence

VII. CONCLUSION

Main motive of the paper was of Object Detection and then to track its location. Two different approaches were followed for this motive, one was to detect and track using specific features like color and other was to detect using background subtraction and then track it using an Adaptive recursive filter. Many experimental results have been presented which started off with extracting the frames from input video sequence or live sequences and then performing the equalization on these frames so as to have a clear and well defined video for further processing. Results for equalized video too have been demonstrated. Striking difference is noted in these two videos.

Many industries are going for complete automation. Robotics is a field which comes to their rescue; here a robot with Multiple Moving Object Detection and Tracking (MMODT) can be used in an industry which can differentiate the required objects. Red, Green, Blue, White Colored Objects can be classified as in case on production line in industry. Particular classified object in real time can be very useful in emergency times and effective action can be taken before any bigger damage is done as in case of terrorist activity. Also in a system as of bank lockers, where any unauthorized entry in particular time zone can be easily warned by this system

REFERENCES

[1] A.Yilmaz, O.Javed, and M.Shah, “Object tracking: A survey,” ACM Computing Surveys, vol. 38, no. 4, 2006.

[2] Sherin Cherian, C.Senthil Singh, “Real Time Implementation of Object Tracking through webcam” International Journal of Research in Engineering and Technology eISSN:2319-1163 Pissn:2321-7308.

[3] Ms K.Lavanya, Prof J.R. Balakrishnan, “Super-Imposing Improve Quality Of Inter andIntra Frame and Extract Histogram” International Journal of Innovative Research in Science, Engineering and Technology,Volume 3, Special Issue 3, March 2014.

[4] S. Bileschi and L. Wolf, “Image representations beyond histograms of gradients: The role of Gestalt descriptors,” in Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2007, pp. 1-8.

[5] Yu Huang, Joan Llach, Chao Zhang, “A method of small object detecion and tracking based on Particle filters” Thomson Corporate Research Princeton, NJ08540, US.

[6] Gaurav Mandloi , “A Survey on Feature Extraction Techniques for Color Images”, International Journal of Computer Science and Information Technologies, Vol. 5 (3) , 2014, 4615-4620.

[7] Swantje Johnsen and Ashley Tews, “Real-Time Object Tracking and Classification Using a Static Camera” Proceedings of the IEEE ICRA 2009Workshop on People Detection and TrackingKobe, Japan, May 2009.

[8] M. Marrón, J.C. García, M.A. Sotelo, M. Cabello, D. Pizarro, F. Huerta, J. Cerro, “Comparing a Kalman Filter and a Particle Filter in a Multiple Objects Tracking Application”, Electronics Department, University of Alcala, Alcalá de Henares, SPAIN

[9] Rachel Kleinbauer, “Study Report in the field of Study Geodesy and Geoinformatics”.

[10] Insaf Setitra, “Object Classification in Videos—An Overview”, Journal of Automation and Control Engineering, Vol. 1, No. 2, June 2013

[11] Pooja Kamavisdar , Sonam Saluja , Sonu Agrawal, “A Survey on Image Classification Approaches and Techniques”, ISSN (Print) : 2319-5940 ISSN (Online) : 2278-1021 International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 1, January 2013

[12] C. Stauffer and W. E. L. Grimson, “Adaptive Background MixtureModels for Real-time Tracking,” in IEEE Computer Society Conferenceon Computer Vision and Pattern Recognition, 1999.

[13] Theo Gevers, Arnold W.M. Smeulders, “Color-based object recognition”, Pattern Recognition 32 (1999) 453—464, Journal of Pattern Recognition K. Elissa, “Title of paper if known,” unpublished.