Automatic Recognition of Image Processing to Detect and Extract Moving Objects Using MATLAB

M. Sasikala¹
Research Scholar
H.H.The Rajah’s College (Autonomous), Pudukkottai – India

R. Malathi²
Research Guide
H.H.The Rajah’s College (Autonomous), Pudukkottai – India

Abstract: In an image there are primarily two sources of information that can be used for detection and tracking of objects: visual features (e.g. color, texture and shape) and motion information. Robust approaches have been suggested by combining the statistical analysis of visual features and temporal analysis of motion information. A typical strategy may first segment a frame into a number of regions based on visual features like color and texture, subsequently merging of regions with similar motion vectors can be performed subject to certain constraints such as spatial neighborhood of the pixels. This research work presents an object class detection approach which fully integrates the complementary strengths offered by shape matchers. Like an object detector, it can learn class models directly from images, and can localize novel instances in the presence of intra-class variations, clutter, and scale changes. Like a shape matcher, it finds the boundaries of objects, rather than just their bounding-boxes. This is achieved by a novel technique for learning a shape model of an object class given images of example instances. Furthermore, also integrate Hough-style voting with a non-rigid point matching algorithm to localize the model in cluttered images. As demonstrated by an extensive evaluation, our method can localize object boundaries accurately and does not need segmented examples for training (only bounding boxes).

Keywords: Active shape model; image processing; image feature extraction; matlab; pixel.

I. INTRODUCTION

Automatic recognition systems for still and moving objects can be invalid in security applications, such as monitoring border areas, buffer zones and restricted areas. A simple recognition system would comprise a camera fixed high above the monitored zone, where images of the zone are captured and consequently processed. Processing the captured images can be in three phases, namely, detection of a moving object, extraction of the object and finally recognition of the object. Optical flow and background subtraction have been used for detecting moving objects in image sequences.

The making of video surveillance system best requires fast, reliable and robust algorithm for moving object detection and tracking system. Identifying moving objects from a video sequence is a fundamental and critical task in many computer-vision applications. A common approach is to perform background subtraction, which identifies moving objects from the portion of a video frame that differs significantly from a background model. There are many challenges in developing a good object detection algorithm. First, it must be robust against changes in illumination Second, it should avoid detecting nonstationary background objects such as swinging leaves, rain, snow, and shadow cast by moving objects.

Finally, its internal background model should react quickly to changes in background such as starting and stopping of vehicles. After identifying moving object in a given scene, the next step in video analysis is tracking. Tracking can be defined as the creation of temporal correspondence among detected objects from frame to frame. This procedure provides temporal identification of the segmented regions and generates cohesive information about the objects in the monitored area such as
trajectory, speed or direction. In everyday life, humans visually keep detecting and tracking a multitude of objects with certain objectives in mind.

II. LITERATURE SURVEY

S. Sravani et al Automated motion detection has become an increasingly important subject in traffic surveillance systems. Automatic video analysis from traffic surveillance cameras is a fast-emerging field based on computer vision techniques. It is a key technology to public safety, intelligent transport system (ITS) and for efficient management of traffic. In recent years, there has been an increased scope for automatic analysis of traffic activity.

Guiming Shi et al In this paper, digital image processing technology used MATLAB programming system developed by each module design and system integration. Proposed passenger identification system is designed to identify the real-time passenger number, density and velocity inside the station, and to determine the level of traffic safety situation.

Atena Khodarrahmi et al an robust algorithm is proposed for automatic detection moving vehicles at night or in environments with low level of light which has quality problems. In this algorithm, first preprocessing steps were conducted. Then all of vehicles in frame identify and classify according their type. Finally, the moving vehicles detected.

A.Anitha et al This introduces an automatic recognition of object, which uses image processing to detect and extract moving objects within a restricted area, and a neural network to recognize the extracted object. Experimental results provides a simple, efficient and fast solution to the problem of detecting, extracting and recognizing moving objects within one system.

Gottipati et al in this foreground detection based moving object detection and vehicle tracking algorithm is implemented targeting a wide class of applications. An AVI file is read and it is decomposed into R, G and B components. Various operations are carried out and the moving objects are detected. Thresholds at various phases are decide the possibility of identifying the moving object of certain sizes. Moving objects also tracked in it.

III. METHODOLOGY

A. Active shape model

Active Shape Model (ASM) is a model-based methods, which makes use of a prior model of what is expected in the image, and typically attempt to find the best match position between the model and the data in a new image. One can make measurements whether the target is actually presented after matching the model. In face recognition application, we collect training images, represent all shapes with a set of landmarks, to form a Point Distribution Model (PDM) respectively.

An image can be recognized by combining the texture and the shape which is extracted from the already trained model. In this work an active shape model is trained with a set of training images. The block diagram of the proposed object recognition system is shown in Fig.1.

The appearance model has parameter \( c \), controlling the shape and texture according to

\[ x = \bar{x} + Q_x c \]
Active Shape Models allow rapid location of the boundary of objects with similar shapes to those in a training set, assuming we know roughly where the object is in the image. They are particularly useful for:

- Objects with well defined shape (e.g., bones, organs, faces etc).
- Cases where we wish to classify objects by shape/appearance.
- Cases where a representative set of examples is available.
- Cases where we have a good guess as to where the target is in the image. However, they are not necessarily appropriate for objects with widely varying shapes (e.g., amorphous things, trees, long wiggly worms etc).
- Problems involving counting large numbers of small things.

B. Moving Object Detection

Distinguishing foreground objects from the stationary background is both an important and difficult research topic. The first step in almost all visual surveillance system is to detect objects in the foreground. This change not only create a greater focus processing levels such as tracking and greatly reduces the computation time, because only the pixels belonging to the foreground object to be processed. Scene changed short and long term trends, such as repetitive movements (such as the resignation of the leaves), light reflections, shadows, camera noise and lights make reliable detection of objects and fast to sudden changes difficult.

The image has been divided into separate objects from the background out. Object recognition and segmentation problem in most cases are closely linked. In some applications, it can be easily segmented object. In the segmented object has not been the case, identify the issues are closely related to the segmentation problem. As the main chain or the cylinder axis, the shape of the cross section of two dimensions, and defining how scan section is scanned along a curve in space rules. The cross-section along the axis can be altered smoothly. Axis plotted relative to the central axis of the cylinder, and in cross-section at each point is perpendicular to the central axis and the cost of the missing can take many different forms. The exact form of these functional requirements will be determined.

C. Image Database

Meaningful training of a neural network is vital if we are to generalize it successfully. Two aspects in training are noted here: firstly the number of images used and, secondly, the sufficiency of the extracted patterns from these images. This section will briefly describe the objects considered in this work; bearing in mind the assumed scenario for this application which is monitoring a secured area with a surveillance camera being positioned higher than the ground level. The assumed secured area is a border controlled area or a buffer zone between two countries, thus the choice of objects in this application. The considered objects in this work are classified into three groups: humans, animals, and vehicles. Images of objects from each group will be used for training and later on generalizing (or testing) the neural network. The objects in each group were as follows: human (Female and Male), animal (Gout), vehicle (Car, Jeep, Motorbike, and Loader).

IV. ALGORITHM USE

A. Active Shape Model Algorithm

1. Examine a region of the image around each point $X_i$ to find the best nearby match for the point $X'$
2. Update the parameters ($X_t$, $Y_t$, $s$, $θ$, $b$) to best fit the new found points $X$
3. Apply constraints to the parameters, $b$, to ensure plausible shapes (e.g., limit so $|bi| < 3√λi$)
4. Repeat until convergence.
B. Shape model

An object is described by points, referred to as landmark points. The landmark points are (manually) determined in a set of training images. From these collections of landmark points, a point distribution model is constructed as follows. The landmark points \((x_1, y_1), \ldots, (x_n, y_n)\) are stacked in shape vectors:

\[
X = (x_1, y_1, \ldots, x_n, y_n)^T.
\]

Principal component analysis (PCA) is applied to the shape vectors \(X\) by computing the mean shape:

\[
x = \frac{1}{s} \sum_{i=1}^{s} x_i
\]

The covariance:

\[
S = \frac{1}{s-1} \sum_{i=1}^{s} (x_i - x)(x_i - x)^T
\]

and the eigensystem of the covariance matrix. The eigenvectors corresponding to the largest eigenvalues are retained in a matrix \(\Phi = (\phi_1 | \phi_2 | \cdots | \phi_k)\). A shape can now be approximated by

\[
x \approx \overline{x} + \Phi b
\]

where \(b\) is a vector of \(k\) elements containing the model parameters, computed by

\[
b = \Phi^T (x - x)
\]

C. Active shape model matching

The Active Shape Model algorithm is a fast and robust method of matching a set of points, on a shape model with new images. The shape parameters, \(b\) for the model, along with parameters defining the global pose (the position, orientation and scale) define the position of the model points in an image, \(X\). Each step is an iterative approach to improve the fit of the points, to an image \(X\), involves first examining the region of the image around each current model point \(X_i, Y_i\) to find the best nearby match \((X_i', Y_i')\) and then updating the parameters \((t_x, t_y, s, \theta, b)\) to best fit the model to the new found points of \(X\). This is repeated until convergence.

In this work, a gray-level appearance model is described that is an alternative to the construction of normalized first-derivative profiles and the Mahalanobis distance-cost-function of the original ASM.

D. Image Features

A Taylor expansion approximates a function \(f\) around a point of interest \(x_0\) by a polynomial of (some) order \(N\). The coefficients in front of each term are given by the derivatives \(f^n\) at \(x_0\):

\[
f(x) = \sum_{n=0}^{N} \frac{1}{n!} f^{(n)}(x_0)(x - x_0)^n
\]
Derivatives of images are computed by convolution with derivatives of Gaussians at a particular scale. This motivates the use of a filter bank of multistate Gaussian derivatives to describe local image structure. Given a set of filtered images, features are extracted for each location by taking the first few moments of the local distribution of image intensities (the histogram) around each location. The most suitable choice for a window function to compute this histogram is a Gaussian, since every other choice induces spurious resolution.

E. Training the shape model

1. Construct shape model. Train the gray-level appearance model.
2. Compute the 60 feature images for each training image.
3. For each landmark, at each resolution, construct a set of training samples with 60 features as input and output zero or one depending on whether the sample is in or outside the object. Samples are taken from a grid around the landmark for each training image. For each training set, a KNN classifier is constructed with selected optimal features.

V. PROBLEM STATEMENT

Object tracking fundamentally entails estimating the location of a particular region in successive frames in a video sequence. Properly detecting objects can be a particularly challenging task, especially since objects can have rather complicated structures and may change in shape, size, location and orientation over subsequent video frames. Various algorithms and schemes have been introduced in the few decades, that can track objects in a particular video sequence, and each algorithm has their own advantages and drawbacks.

To overcome the different challenges issue as discussed in previous section there are following main component of object detection and tracking

In this thesis our aim is to improve the performance of object detection and tracking by contributing originally to two components

a) Motion segmentation
b) Object tracking.

Automated tracking of objects can be used by many interesting applications. An accurate and efficient tracking capability at the heart of such a system is essential for building higher level vision-based intelligence. Tracking is not a trivial task given the non-deterministic nature of the subjects, their motion, and the image capture process itself. The objective of video tracking is to associate target objects in consecutive video frames. We have to detect and track the object moving independently to the background. In this there are four situations to be considered in the account:

Single camera and single object,
Single camera and multiple objects,
Multiple cameras and single object,

Multiple cameras and multiple objects.

VI. IMPLEMENTATION

The recognition of Object in image sequences is an important and challenging problem that enables a host of human-computer interaction applications. In this work a new method for recognizing object has been proposed. Most studies on continuous object recognition have been done with frames obtained by processing the videos with regular/equal intervals.

A. Preprocessing Steps

The images are resized to 640*480 pixels to reduce memory space and execution time. For giving the edge input while training the model a binary mask of the input training images are needed. In binary mask creation, the image is read first and the dimensions are captured.

Fig.1. Preprocessing

B. Shape-freetexture

In the second step, other training data are marked by contour points. The location of the contour points of the image is the same in length as the coordinates.

Fig.2. Results obtained for the shape-freetexture
C. Testing image

The next step marks for the area contours as shown in fig.5 and it is displayed with the help of the testing image and matched to other training sets. After approximating model points that may be related to the training sets, they are displayed as final output.

![Testing image](image)

**Fig.3. Mean Contour Region obtained**

D. Plot the points for images

The gray world algorithm makes use of the PCA, which transforms the hand region into a data space of Eigenvectors that results from different reference hand poses. This data space is called Pose Eigen Space (PES). The reference poses are generated by means of a rotating virtual hand and are distributed equally. Before projection into PES, the monochrome images used are normalized first by subtracting the average intensity from individual pixels and then dividing the result by the standard deviation. Variations resulting from different illuminations are averaged.

VII. CONCLUSION AND FURTHER WORK

A. Conclusion

A technique has been presented for recognizing and tracking a moving non-rigid object or person in a video sequence. The objective function for active shape models has been extended to color images. We have evaluated several different approaches for defining an objective function considering the information from the single components of the color image vectors. This tracking technique does not require a static camera (except to initialize the landmark points for the object to be recognized).
Thus, it can be applied when using a pan–tilt–zoom (PTZ) camera for video tracking. However, the profile length has to be adapted to the pan, tilt, and zoom parameters of the PTZ camera.

In both our indoor and outdoor experiments, the median computation of the minima in the energy functions proved favourable. In general, the error in fitting an ASM to the real contour of an object was lower when using color information than when just using intensity information.

B. Future work

The presented approach can be extended in several ways. First, the training stage models only positive examples. This could be extended by learning a classifier to distinguish between positive and negative examples, which might reduce false positives. One possibility could be to train both our shape models and the discriminative models. At detection time, we could then use the bounding-box delivered to initialize shape matching based on our models. Moreover, the discriminative power of the representation could be improved by using appearance features in addition to image contours. Finally, in this paper we have assumed that all observed differences in the shape of the training examples originate from intra-class variation, and not from viewpoint changes. It would be interesting to add a stage to automatically group objects by viewpoint, and learn separate shape models.

References

2. Guiming Shi, Chang Liu and Shuyuan Zhao. Research on Passenger Recognition System Based on Video Image Processing in Railway Passenger Station.