



Volume: 2, Issue: 6, 163-166  
June 2015  
www.allsubjectjournal.com  
e-ISSN: 2349-4182  
p-ISSN: 2349-5979  
Impact Factor: 3.762

**Ashwini Abhyankar**  
Department of Electronics &  
Telecommunication,  
RMDSSOE, SavitribaiPhule  
Pune University, Warje,  
Pune, 411058, India.

**V. D. Shinde**  
Department of Electronics &  
Telecommunication,  
RMDSSOE, SavitribaiPhule  
Pune University, Warje,  
Pune, 411058, India.

## Advance Trends in Object Detection

**Ashwini Abhyankar, V. D. Shinde**

### Abstract

An object recognition system finds objects in the real world from an image of the world, using object models which are known a priori. This task is surprisingly difficult. Humans perform object recognition effortlessly and instantaneously. Algorithmic description of this task for implementation on machines has been very difficult. Here we will discuss different steps in object recognition and introduce some techniques that have been used for object recognition in many applications. We will discuss the different types of recognition tasks that a vision system may need to perform. We will analyze the complexity of these tasks and present approaches useful in different phases of the recognition task. The object recognition problem can be defined as a labeling problem based on models of known objects. Formally, given an image containing one or more objects of interest (and background) and a set of labels corresponding to a set of models known to the system, the system should assign correct labels to regions, or a set of regions, in the image. The object recognition problem is closely tied to the segmentation problem: without at least a partial recognition of objects, segmentation cannot be done, and without segmentation, object recognition is not possible.

**Keywords:** Object recognition, Feature extraction, Video surveillance.

### 1. Introduction

An object recognition system finds objects in the real world from an image of the world, using object models which are known a priori. This task is surprisingly difficult. Humans perform object recognition effortlessly and instantaneously. Algorithmic description of this task for implementation on machines has been very difficult. In this chapter we will discuss different steps in object recognition and introduce some techniques that have been used for object recognition in many applications. We will discuss the different types of recognition tasks that a vision system may need to perform. We will analyze the complexity of these tasks and present approaches useful in different phases of the recognition task. The object recognition problem can be defined as a labeling problem based on models of known objects. Formally, given an image containing one or more objects of interest (and background) and a set of labels corresponding to a set of models known to the system, the system should assign correct labels to regions, or a set of regions, in the image. The object recognition problem is closely tied to the segmentation problem: without at least a partial recognition of objects, segmentation cannot be done, and without segmentation, object recognition is not possible. In this report, we discuss basic aspects of object recognition. We present the architecture and main components of object recognition and discuss their role in object recognition systems of varying complexity.

### 2. System Components

An object recognition system must have the following components to perform the task:

1. Model database (also called model base)
2. Feature detector
3. Hypothesizer
4. Hypothesis verifier

The feature detector applies operators to images and identifies locations of features that help in forming object hypotheses. The features used by a system depend on the types of objects to be recognized and the organization of the model database. Using the detected features in the image, the hypothesizer assigns likelihoods to objects present in the scene. This step is used to reduce the search space for the recognizer using certain features. The model base is organized using some type of indexing scheme to facilitate elimination of unlikely object candidates from possible consideration. The verifier then uses object models to verify the hypotheses and refines the likelihood of objects. The system then selects the object with the highest likelihood, based on all the evidence, as the correct object.

### Correspondence:

**Ashwini Abhyankar**  
Department of Electronics &  
Telecommunication,  
RMDSSOE, SavitribaiPhule  
Pune University, Warje,  
Pune, 411058, India.

All object recognition systems use models either explicitly or implicitly and employ feature detectors based on these object models. The hypothesis formation and verification components vary in their importance in different approaches to object recognition. Some systems use only hypothesis formation and then select the object with highest likelihood as the correct object. Pattern classification approaches are a good example of this approach. Many artificial intelligence systems, on the other hand, rely little on the hypothesis formation and do more work in the verification phases. In fact, one of the classical approaches, template matching, bypasses the hypothesis formation stage entirely.

An object recognition system must select appropriate tools and techniques for the steps discussed above. Many factors must be considered in the selection of appropriate methods for a particular application. The central issues that should be considered in designing an object recognition system are:

#### **A. Object or model representation**

How should objects represented in the model database? What are the important attributes or features of objects that must be captured in these models? For some objects, geometric descriptions may be available and may also be efficient, while for another class one may have to rely on generic or functional features. The representation of an object should capture all relevant information without any redundancies and should organize this information in a form that allows easy access by different components of the object recognition system.

#### **B. Feature extraction**

Which features should be detected, and how can they be detected reliably? Most features can be computed in two dimensional images but they are related to three-dimensional characteristics of objects. Due to the nature of the image formation process, some features are easy to compute reliably while others are very difficult.

#### **C. Feature-model matching**

How can features in images be matched to models in the database? In most object recognition tasks, there are many features and numerous objects. An exhaustive matching approach will solve the recognition problem but may be too slow to be useful. Effectiveness of features and efficiency of a matching technique must be considered in developing a matching approach.

#### **D. Hypotheses formation**

How can a set of likely objects based on the feature matching be selected, and how can probabilities be assigned to each possible object? The hypothesis formation step is basically a heuristic to reduce the size of the search space. This step uses knowledge of the application domain to assign some kind of probability or confidence measure to different objects in the domain. This measure reflects the likelihood of the presence of objects based on the detected features.

#### **E. Object verification**

How can object models be used to select the most likely object from the set of probable objects in a given image? The presence of each likely object can be verified by using their models. One must examine each plausible hypothesis to verify the presence of the object or ignore it. If the models are geometric, it is easy to precisely verify objects using camera

location and other scene parameters. In other cases, it may not be possible to verify a hypothesis

### **3. Complexity of Object Detection**

Images of scenes depend on illumination, camera parameters, and camera location. Since an object must be recognized from images of a scene containing multiple entities, the complexity of object recognition depends on several factors. A qualitative way to consider the complexity of the object recognition task would consider the following factors:

#### **A. Scene constancy**

The scene complexity will depend on whether the images are acquired in similar conditions (illumination, background, camera parameters, and viewpoint) as the models. As seen in earlier chapters, scene conditions affect images of the same object dramatically. Under different scene conditions, the performance of different feature detectors will be significantly different. The nature of the background, other objects, and illumination must be considered to determine what kind of features can be efficiently and reliably detected.

#### **B. Image-models spaces**

In some applications, images may be obtained such that three-dimensional objects can be considered two-dimensional. The models in such cases can be represented using two-dimensional characteristics. If models are three-dimensional and perspective effects cannot be ignored, then the situation becomes more complex. In this case, the features are detected in two-dimensional image space, while the models of objects may be in three dimensional spaces. Thus, the same three-dimensional feature may appear as a different feature in an image. This may also happen in dynamic images due to the motion of objects.

#### **C. Number of objects in the model database**

If the number of objects is very small, one may not need the hypothesis formation stage. A sequential exhaustive matching may be acceptable. Hypothesis formation becomes important for a large number of objects. The amount of effort spent in selecting appropriate features for object recognition also increases rapidly with an increase in the number of objects.

#### **D. Number of objects in an image and possibility of occlusion**

If there is only one object in an image, it may be completely visible. With an increase in the number of objects in the image, the probability of occlusion increases. Occlusion is a serious problem in many basic image computations. Occlusion results in the absence of expected features and the generation of unexpected features. Occlusion should also be considered in the hypothesis verification stage. Generally, the difficulty in the recognition task increases with the number of objects in an image. Difficulties in image segmentation are due to the presence of multiple occluding objects in images.

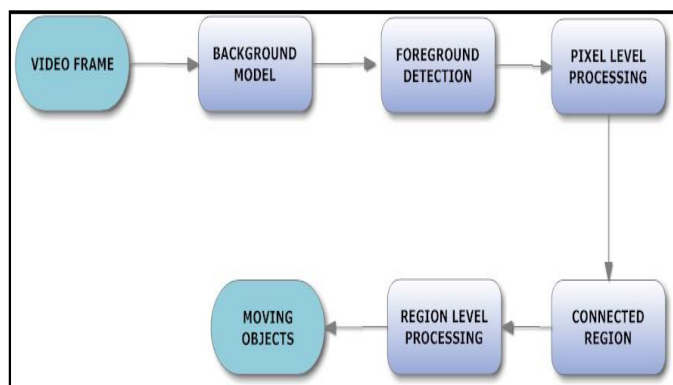
### **4. Video Surveillance**

Video surveillance is a process of analyzing video sequences. It is an active area in computer vision. It gives huge amount of data storage and display. There are three types of Video surveillance activities. Video surveillance activities can be manual, semi-autonomous or fully-autonomous. Manual video surveillance involves analysis of the video content by a human. Such systems are currently widely used. Semi-autonomous video surveillance involves some form of video

processing but with significant human intervention. Typical examples are systems that perform simple motion detection. Only in the presence of significant motion the video is recorded and sent for analysis by a human expert. By a fully-autonomous system, only input is the video sequence taken at the scene where surveillance is performed. In such a system there is no human intervention and the system does both the low-level tasks, like motion detection and tracking, and also high-level decision making tasks like abnormal event detection and gesture recognition. Video surveillance system that supports automated objects classification and object tracking. Monitoring of video for long duration by human operator is impractical and infeasible. Automatic motion detection which can provide better human attention. There are varieties of applications in video surveillance like access Control, person identification, and anomaly detection. Intelligent visual surveillance (IVS) refers to an automated visual monitoring process that involves analysis and interpretation of object behaviors, as well as object detection and tracking, to understand the visual events of the scene. Main tasks of IVS include scene interpretation and wide area surveillance control. Scene interpretation detects and track moving objects in an image sequence. It is used to understand their behaviors.

### 5. Moving Object Detection

Moving object detection is the basic step for further analysis of video. Every tracking method requires an object detection mechanism either in every frame or when the object first appears in the video. It handles segmentation of moving objects from stationary background objects. This focuses on higher level processing. It also decreases computation time. Due to environmental conditions like illumination changes, shadow object segmentation becomes difficult and significant problem. A common approach for object detection is to use information in a single frame. However, some object detection methods make use of the temporal information computed from a sequence of frames to reduce the number of false detections. This temporal information is usually in the form of frame differencing, which highlights regions that changes dynamically in consecutive frames. Given the object regions in the image, it is then the tracker's task to perform object correspondence from one frame to the next to generate the tracks.



**Fig 2:** Framework of Moving Object Detection System

The first step is to distinguish foreground objects from stationary background. To achieve this, we can use a combination of various techniques along with low-level image post-processing methods to create a foreground pixel map at every frame. We then group the connected regions in the

foreground map to extract individual object features such as bounding box, area, perimeter etc.

#### A. Foreground Detection

The main purpose of foreground detection is to distinguishing foreground objects from the stationary background. Almost, each of the video surveillance systems uses the first step is detecting foreground objects. This creates a focus of attention for higher processing levels such as tracking, classification and behavior understanding and reduces computation time considerably since only pixels belonging to foreground objects need to be dealt with. The first step is the background scene initialization. There are various techniques used to model the background scene. The background scene related parts of the system is isolated and its coupling with other modules is kept minimum to let the whole detection system to work flexibly with any one of the background models. Next step in the detection method is detecting the foreground pixels by using the background model and the current image from video. This pixel-level detection process is dependent on the background model in use and it is used to update the background model to adapt to dynamic scene changes. Also, due to camera noise or environmental effects the detected foreground pixel map contains noise. Pixel-level post-processing operations are performed to remove noise in the foreground pixels. Once we get the filtered foreground pixels, in the next step, connected regions are found by using a connected component labeling algorithm and objects' bounding rectangles are calculated. The labeled regions may contain near but disjoint regions due to defects in foreground segmentation process. Hence, some relatively small regions caused by environmental noise are eliminated in the region-level post-processing step. In the final step of the detection process, a number of object features like area, bounding box, perimeter of the regions corresponding to objects are extracted from current image by using the foreground pixel map.

#### B. Pixel Level Post-Processing

The output of foreground detection contains noise. Generally, it affects by various noise factors. To overcome this dilemma of noise, it requires further pixel level processing. There are various factors that cause the noise in foreground detection such as: Camera Noise: Camera noise presents due to camera's image acquisition components. This is the noise caused by the camera's image acquisition components. This noise is produce because of the intensity of a pixel that corresponds to an edge between two different colored objects in the scene may be set to one of the object's color in one frame and to other's color in the next frame. Background Colored Object Noise: The color of the object may have the same color as the reference background Difficult to detect foreground pixels with the help of reference background. Reflectance Noise: Reflectance noise is caused by light source. When a light source moves from one position to another, some parts in the background scene reflect light. We can use low pass filter and morphological operations, erosion and dilation, to the foreground pixel map to remove noise that is caused by the items listed above. Our aim in applying these operations is removing noisy foreground pixels that do not correspond to actual foreground regions, and to remove the noisy background pixels near and inside object regions that are actually foreground pixels. Low pass filters are used for blurring and for noise reduction. Blurring is used in pre-processing tasks, such as removal of small details from an image prior to large object extraction, and bridging of small

gaps in lines or curves. Gaussian low pass filter is used for pixel level post processing. A Gaussian filter smooths an image by calculating weighted averages in a filter coefficient. Gaussian filter modifies the input signal by convolution with a Gaussian function.

### C. Detecting Connected Regions

After detecting foreground regions and applying post-processing operations to remove noisy regions, the filtered foreground pixels are grouped into connected regions. After finding individual regions that correspond to objects, the bounding boxes of these regions are calculated.

### D. Region Level Post-Processing

As pixel-level noise is removed, still some artificial small regions remain just because of the bad segmentation. To remove this type of regions, regions that have smaller sizes than a pre-defined threshold are deleted from the foreground pixel map. Once segmenting regions we can extract features of the corresponding objects from the current image. These features are size, center-of-mass or just centroid and Bounded Area of the connected component. These features are used for object tracking and classification for the further processing in event detection.

## 6. Motion Detection Methods

The motion detection methods are classified according to the method of finding moving object. Different motion detection methods are described as follows:

### A. Temporal differencing

The Frame differencing method uses the two or three adjacent frames based on time series image to subtract and gets difference images, its working is very similar to background subtraction after the subtraction of image it gives moving target information through the threshold value. This method is simple and easy to implement, and also it is similar to the background subtraction. But this method is highly adaptive to dynamic scene change; however, it generally fails in detecting whole relevant pixels of some types of moving objects. Additional methods need to be adopted in order to detect stopped objects for the success of higher level are computationally complex and cannot be used real-time without specialized hardware.

### B. Background subtraction

It is particularly a commonly used technique for motion segmentation in static images. It will detect moving regions by subtracting the current image pixel-by-pixel from a reference background image that is created by averaging images over time in an initialization period. The basic idea of background subtraction method is to initialize a background firstly, and then by subtracting current frame in which the moving object is present that current frame is subtracted with background frame to detect moving object. This method is simple and easy to realize, and accurately extracts the characteristics of target data, but it is sensitive to the change of external environment, so it is applicable to the condition that the background is known.

### C. Optical flow

The optical flow method uses the motion target of the vector characteristics which changed with time to detect motion area in image sequences. It gives better performance under the

moving camera, but this algorithm is very complex and complicated computation and also it needs special hardware support, so it is difficult to meet the requirements of real-time video processing.

## 7. Conclusion

Automatic motion detection which can provide better human attention. There are varieties of applications in video surveillance like access Control, person identification, and anomaly detection. Intelligent visual surveillance (IVS) refers to an automated visual monitoring process that involves analysis and interpretation of object behaviors, as well as object detection and tracking, to understand the visual events of the scene. Main tasks of IVS include scene interpretation and wide area surveillance control. Scene interpretation detects and tracks moving objects in an image sequence. It is used to understand their behaviors.

## 8. References

1. Maiti A, Sivanesan A. "Cloud Controlled Intrusion Detection and Burglary Prevention Stratagems" published in Future Internet Communications, 2012, Date of Conference: 25-27 April, 2012.
2. Sindhulakshmi K, Soundarya J. International Journal of Engineering Science and Innovative Technology (IJESIT) March on 'Cloud Controlled Intrusion Detection and Burglary Prevention Stratagems in Home Automation Systems'. 2014, 3(2).
3. Turaga P. IEEE Paper on "Integrating sensors and cameras in real time monitoring systems.
4. Feng-Li Lian, Yi-Chun-Li. IEEE paper on "Voting-Based Motion Estimation for Real-Time Video Transmission in Networked Mobile Camera Systems" in, 2013.
5. Bodor R. IEEE paper on "Vision-Based Human Tracking and Activity Recognition".