A Survey on Approaches of Object Detection

Sanjivani Shantaiya
Dept. of Computer Science & Engg.DIMAT
Raipur, India

Keshri Verma
Dept.of Computer Application.NIT
Raipur, India

Kamal Mehta
Dept.of Computer Science & Engg.Sagar Engg.
Bhopal, India

ABSTRACT

This research paper presents the extensive survey of the state-of-the-art detection of object from video and also classifies the detection approaches into different categories. Detection is generally performed in the context of higher-level applications that require the location and/or shape of the object in every frame. As the detection of objects from vide have progressed significantly in the past few years; human motion and behavior interpretation have naturally become the upcoming step. The research paper includes various approaches that have been used by different researchers for object detection.

General Terms

Object detection approaches, video frames.

Keywords

Object detection, motion recognition and tracking.

1. INTRODUCTION

The use of video is becoming prevalent in lot more applications such as monitoring of vehicular traffic, detection of pedestrians, identification of anomalous behavior in a parking lot or near an ATM, etc. The field of automated surveillance systems is nowadays of immense interest because to its implications in the prospects of security. Surveillance of vehicular traffic and human activities offers a context for the extraction of significant information such as scene motion and traffic statistics, object classification, human identification, anomaly detection, as well as the analysis of interactions between vehicles, between humans, or between vehicles and humans [1].

2. OBJECT DETECTION APPROACHES

The proliferation of high-powered computers, the availability of high quality and inexpensive video cameras, and the increasing need for automated video analysis has generated a great deal of interest in object detection and tracking of objects. A key task in tracking video data is the detection and tracking of moving objects, such as people and vehicles, through the video frames. Detection of moving objects in video streams is the first relevant step of information extraction in many computer vision applications, including video surveillance, people tracking, traffic monitoring, and semantic annotation of videos. There are various approaches for object detection shown in fig.3.

2.1 Feature-based object detection

In feature-based object detection, standardization of image features is important. One or more features are extracted and the objects of interest are modeled in terms of these features. Features may be shape, size or the color of objects.

2.1.1 Shape-based object detection

Shape-based object detection is one of the complex problems due to the difficulty of segmenting objects of interest in the images. The detection and shape characterization of the objects becomes more difficult for complex scenes where there are many objects with occlusions and shading.**Lu et al.** [11] proposes representation of athletes by the PCA-HOG descriptor as shown in fig.1, which can be computed by first transforming the athletes to the grids of Histograms of Oriented Gradient (HOG) descriptor and then apply Principal Component Analysis (PCA). PCA-HOG descriptor can be computed by the following procedures:

1. The image I is filtered by a symmetric low-pass Gaussian filter. Then, compute the image gradient along the x and y direction by a 1-D centered mask [− 1, 0, 1]

$$gx(x,y) = I(x+1,y) - I(x-1,y)\ \forall x,y$$
$$gy(x,y) = I(x,y+1) - I(x,y-1)\ \forall x,y$$

where gx(x, y) and gy(x, y) denotes the x and y components of the image gradient, respectively.

2. The magnitude m(x, y) and orientation θ(x, y) of the image gradient are computed by

$$\mathrm{m}(x,y) = \sqrt{gx(x,y)2 + gy(x,y)2}$$
$$\theta(x,y) = tan - 1(gy(x,y)/gx(x,y))$$

3. Then partition the image into $s_w \times s_h$ non-overlapping grids. For each grid, quantize the orientation ˜θ(x, y) for all pixels into s_b orientation bins weighted by its magnitude m(x, y).

4. Then normalize each feature by the sums of all features. The resulting feature vector, $H \in R^{nf}$,

$nf = sw \times sh \times sb$, is the HOG descriptor of the image I.

5. Let $\Gamma \in R^{np\times nf}$ denote the first np principal components learned from the HOG descriptors of training images. Then project the HOG descriptor H to the linear subspace spanned by the principal components Γ, i.e.

$$Y = \Gamma T\ (H - \bar{H})$$

where $\bar{H} \in R^{nf}$ is the mean HOG descriptor of all training images, and $Y \in R^{np}$ is PCA-HOG descriptor of the image I.

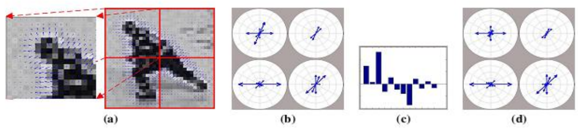

Fig 1: Examples of PCA-HOG descriptor a) Image Gradient b) HOG descriptor with a 2 x 2 grids and 8 orientation bins c) & d) PCA-HOG descriptors

One limitation with this method is; it might be difficult to target multiple objects.**Huabo Sun et al.** [27] propose a new method for moving object detection using variable resolution double-level contours. The proposed algorithm of moving object detection using variable resolution double-level contours is:-

Level 1: Contour detection

First of all, low-resolution a remotely sensed image is processed. Detect image edges with canny method and extract the contours of the object. Then calculate the minimum area matrix [m1:m2, n1:n2] that includes the contour.

Level 2: Contour segmentation with GAC texture gradient model

After this, the high-resolution remotely sensed image, with the same surface features, is processed. Then operate on the image with the geodesic active contour (GAC) texture gradient model. Initially set the contour in the range of [m1:m2, n1:n2] and the contour gradually reduces to detect the moving target. The contour is set and is based on the formula in equation below.

$$E = \iint g\delta(u)|\nabla u|dxdy + c\iint[1 - H(u)]gdxdy$$

where E stands for a closed curve that minimizes the energy function. The limitation of above method is although, compound eyes have large fields that are viewable, and the acute zone lies in just a small part of it. So, in most cases the target does not directly appear in the area of the acute zone. **Chiverton et al.** [30] proposed a new fully automatic object tracking and segmentation framework. The framework consists of a motion-based bootstrapping algorithm concurrent to a shape-based active contour. They propose two approaches of increasing computational intensity and accuracy that statistically estimate the foreground from the potential mixture of foreground and background enveloped by an alpha hull. The first approach, i.e., referred to here as single back-ground–foreground boosting (SB–FB), estimates the foreground using a single foreground–background model, which assumes that the tracked object possesses different photo-metric properties from any part of the background immediately surrounding the object. The second technique, i.e., referred to here as multiple background–foreground boosting (MB–FB), estimates the object shape using an assumption of local differential properties between the tracked object shape and smaller local regions surrounding the object.

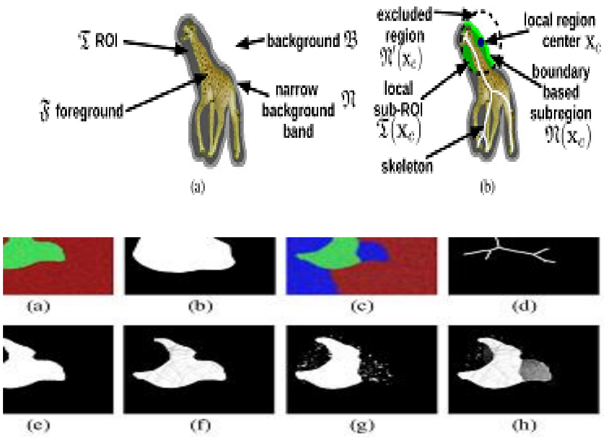

Fig 2: a. Globalized ROI b.Localized ROI Fig. 2 b-h. SB-FB vs. MB-FB [30]

2.1.2 Color-based object detection

Unlike many other image features (e.g. shape) color is relatively constant under viewpoint changes and it is easy to be acquired. Although color is not always appropriate as the sole means of detecting and tracking objects, but the low computational cost of the algorithms proposed makes color a desirable feature to exploit when appropriate. **Sebastien et al.** [7] dealt with object learning using color information. The GHOSP (Genetic Hybrid Optimization & Search of Parameters) algorithm is developed which uses multidimensional observations are taken from RGB color images which contain object to be learnt. They consider each color component separately and sample the intensities of

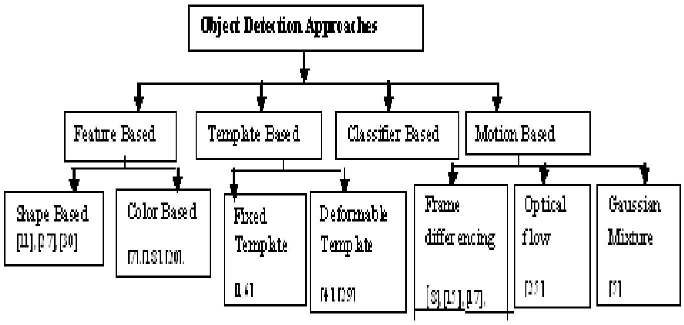

Fig 3: Taxonomy of approaches of object detection

pixels belonging to the three gray level images (one for each color component) in order to define a reasonable set of symbols. So the intensity or value v of every pixel is converted from [0, 255] into a symbol s using the following formula:

$$s = \frac{[v \cdot Mr - 1]}{255}$$

where M^r is the number of possible symbols. From this sampling they obtain a three component vector which will be used as input in the GHOSP algorithm. The problem with this is; the method is unable to perform correctly when the tracked object size is too small. In this case, the learning phase cannot be performed correctly due to the lack of learning data. **Zhenjun et al.** [18] used combined feature set which is built using color histogram (HC) bins and gradient orientation histogram (HOG) bins considering the color and contour representation of an object for object detection. The combined feature set is the evolvement of color, edge orientation histograms and SIFT descriptors.

Color histogram: They define a color histogram (HC) of 48 dimensions for both the object and its background. In each color component in RGB color space, 16 dimensions of histogram features are calculated.

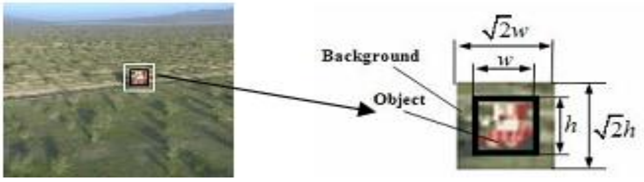

Fig 4: Object and background regions [18]

Gradient orientation histogram: They follow the idea to extract HOG features on gray value image windows. On each window, a histogram of 72 dimensions is extracted to descript the gradient orientation of an object.

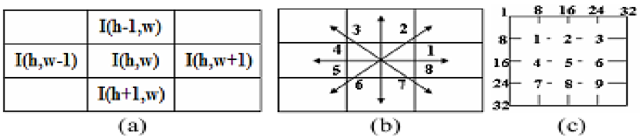

Fig 5: HOG Extraction. a. Mask for pixel gradient calculation b.Orientation bin for voting c. 9 blocks in the image window [18]

Limitation is that it will work if the background has the similar color. **Zhu et al.** [20], [23] employed the mean shift algorithm for segmentation which is followed by a manual refinement of the boundaries to generate smooth contours for creating the distance map. This algorithm would create a confidence map in the new image based on the color histogram of the object in the previous image, and use mean shift to find the peak of a confidence map near the object's old position. The mean-shift algorithm runs as follows:

1. Choose a search window:
   - Its initial location;
   - Its type (uniform, polynomial, exponential, or Gaussian);
   - Its shape(symmetric or skewed, possibly rotated, rounded);
   - Its size (extent at which it rolls off or is cut off).
2. Compute the window's (possibly weighted) center of mass.
3. Center the window at the center of mass.
4. Return to step 2 until the window stops moving (it always will).

The weakness of this algorithm is the tracking drift (or tracking failure) especially when the color distributions of target object and the background clutter (or other objects) become similar. **Saravanakumar et al.** [26] represented the objects using the properties of the HSV color space. Adaptive k-means clustering algorithm was applied to cluster objects centroids color values and co-ordinates were sent to next frame for clustering. A three dimensional representation of the HSV color space is a hexacone, with the central vertical axis representing intensity. Hue is defined as an angle in the range [0, 2∏] relative to the red axis with red at angle 0, green at 2∏/3, blue at 4∏/3 and red again at 2∏. Saturation is the purity of color and is measured as a radial distance from the central axis with values between 0 at the center to 1 at the outer surface.

2.2 Template-based object detection

If a template describing a specific object is available, object detection becomes a process of matching features between the template and the image sequence under analysis. There are two types of object template matching, fixed and deformable template matching.

2.2.1 Fixed template matching

Fixed templates are useful when object shapes do not change with respect to the viewing angle of the camera. Two major techniques have been used in fix template matching.

2.2.1.1 Image subtraction

In this technique, the template position is determined from minimizing the distance function between the template and various positions in the image. Although image subtraction techniques require less computation time than the following correlation techniques, they perform well in restricted environments where imaging conditions, such as image intensity and viewing angles between the template and images containing this template are the same.

2.2.1.2 Correlation

Matching by correlation utilizes the position of the normalized cross-correlation peak between a template and an image to locate the best match. This technique is generally immune to noise and illumination effects in the images, but suffers from high computational complexity caused by summations over the entire template. **Jiyan et al.** [16] proposes Content-Adaptive Progressive Occlusion Analysis (CAPOA) algorithm; which analyzes the occlusion situation within a given region of interest (ROI) and generates corresponding template mask. Detection of a reappearing target is somewhat difficult with this method.

2.2.2 Deformable template matching

Deformable template matching approaches are more suitable for cases where objects vary due to rigid and non-rigid deformations. Because of the deformable nature of objects in most video, deformable models are more appealing in tracking tasks. **Zhong et al.** [4] proposed a novel method for object detection using prototype-based deformable template models.. Deformed template is obtained by applying a parameterized deformation transform on the prototype. The prototype-based template combines both the global structure information and local image cues to derive an interpretation. It consists of three components: 1) a contour template which describes the prior knowledge about the object shape (a prototype), 2) a parameterized transformation which is applied to the prototype to deform it, and 3) a probabilistic distribution on the deformation parameters which controls the variation in the deformable template.

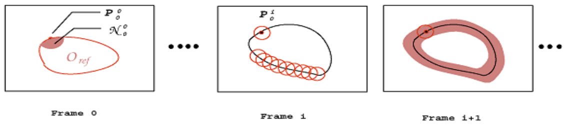

Fig 6: Computing color (gray scale) distance. The detected object in the first frame is used as the reference object [4]

Xiaobai Liu et al. [29] proposed hybrid online templates for object detection which uses different features such as flatness, texture, or edge/corner. The template consists of multiple types of features, including sketchs/edges, texture regions, and flatness regions. Sketch/edge regions usually consist of various links, ridges, such as corners and junctions. Texture regions are a large number of objects that are either too small or too distant to the camera. In contrast, flatness regions are always filled with homogeneous color or intensity. The limitation of this method is; as the discriminative power of features change along with the object movements, the hybrid template should be adaptively updated by either adjusting the feature confidences, or substituting the old features with the newly discovered ones from the currently observed frames.

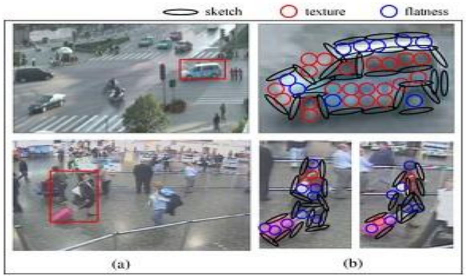

Fig 7: Hybrid object template consists of multiple types of features, including sketch, texture and flatness. a. Objects within images. B. Hybrid templates learned from the foreground regions. [29]

2.3 Classifier Based Object Detection

Liu et al. [9] presented a novel semiautomatic segmentation method for single video object extraction. Proposed method formulates the separation of the video objects from the background as a classification problem. Each frame was divided into small blocks of uniform size, which are called object blocks if the centering pixels belong to the object, or background blocks otherwise. After a manual segmentation of the first frame, the blocks of this frame were used as the training samples for the object-background classifier. A newly developed learning tool called ψ-learning was employed to train the classifier. Nevertheless, the object boundaries are not always perfectly located due to the classification error. **Yuhua et al.** [14] presented new face detection method from a video sequence. Firstly, a classifier with a set of parameters was built up based on the knowledge of the interest object. Then both positive and negative sample data were fed into the classifier to adjust those parameters. There was a mapping between the object and the classifier. For complex objects, multiple classifiers needed to be integrated, which was called cascade classifiers or boosted classifiers. The basic idea of these cascade classifiers are that several weak classifiers are used to cover different features of the object and combined to reach a better classification globally. The limitation with this method is; more object features need to be embedded to train the object model under different environment and light conditions.

Table 1. Analysis of object detection approaches

Approach	Methods/Algorithm Used	Authors	Single(S)/ Multiple(M)
Shape based (Feature based)	1. PCA-HOG descriptor, hybrid Hidden Markov Model 2.Gaussian filters, Geodesic Active Contour model 3. Bootstrapping algorithm, shape-based active contour	1.Wei-Lwun Lu et al. [11] 2.Huabo Sun et al. [27] 3. John et al. [30]	S S S
Color based (Feature based)	1.HMM model, GHOSP (Genetic Hybrid Optimization & Search of Parameters) algorithm 2. color histogram (HC) bins and gradient orientation histogram (HOG) bins, kalman filter 3.Scale Invariant Feature Transform (SIFT), Kalman filter 4. Particle filters and multi-mode anisotropic mean shift. 5. K-Means clustering,HSV color space, Connected component analysis	1.Sebastien et al. [7] 2. Zhenjun et al. [18] 3. Junda Zhu et al. [20] 4. Zulfiqar et al. [23] 5.Saravanakumar et al. [26]	S S S M S
Fixed Template based	1. Content-adaptive progressive occlusion analysis (CAPOA) algorithm, Local best match authentication (LBMA) algorithm, Kalman filter	1. Jiyan et al. [16]	S
Deformable Template based	1. prototype-based template 2.Adaptive tracking algorithm, Hybrid Template	1. Yu Zhong et al. [4] 2. Xiaobai Liu et al.[29]	S M
Classifier Based	1. Multilayer ψ-learning for object tracking, pyramid boundary refining algorithm 2. Bayes-based filter, Haar-like features, PSO-based searching algorithm 3. Annealed Gaussian-based PSO (AGPSO) algorithm.	1. Yi Liu et al. [9] 2. Yuhua et al. [14] 3. Xiaoqin et al. [22]	S S M
Frame Differencing (Motion based)	1.HMM model, Blob detection algorithm, Bayesian filtering 2.Background subtraction algorithm, particle filter tracker 3. Background elimination and background registration techniques. 4. Background subtraction algorithm, Kalman filter 5.Adaptive back-ground penalty with occlusion reasoning	1.Lan Wu[8] 2.Sourabh Khire et al. [15] 3.Vibha et al. [17] 4. Swantje et al. [19] 5. Zhang et al. [28]	M M S M M
Optical flow (Motion based)	1. SIFT / HOG descriptor, optical flow estimation, SVM Classifier	1. Navneet et al. [25]	M
Gaussian Mixture (Motion based)	1. Mixture of Gaussians, Kalman filter	1.Robert et al. [5]	S

2.4 Motion- based Object Detection

Viola, jones and snow use motion information for detection of pedestrian. They use different motion filters for effective detection of pedestrians [6]. A large variety of motion detection algorithms had been proposed as.

2.4.1 Thresholding technique over the interframe difference

These approaches rely on the detection of temporal changes either at pixel or block level. The first stationary frame is reference background frame. The foreground is detected by subtracting each frame with reference frame and using thresholding [8].This method assumes object moving continuously and foreground would not be detected if no motion between frames. **Lan Wu** [8] presented the task of detecting and tracking hockey players by automatically estimate the locations and sizes of the hockey players on the ice rink coordinate system as well as keeping the identities of the players given the two synchronized video sequences. However, the system sometimes fails to keep the identities of the players when the players are occluded in both views.

Input: Frame 226

Input: Frame 226

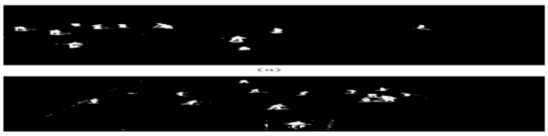

Fig 8: The Background subtraction results. a. The foreground blobs from the right view. b. The foreground blobs from the end view [8]

Khire et al. [15] focuses on detection of moving objects in a scene, for example pedestrians crossing a street as long as they stay in the scene. This technique also uses a median-based approach for initial background estimation and a novel two-pass approach for noise removal. Thus, for a training set of N frames, the background can be constructed simply as

$$BgR(x,y) = median\ t \in T\ \{IR(x,y,ti), \dots IR(x,y,ti - N - 1)\}\}$$

$$BgG(x,y) = median\ t \in T\ \{IG(x,y,ti), \dots IG(x,y,ti - N - 1)\}\}$$

$$BgB(x,y) = median\ t \in T\ \{IB(x,y,ti), \dots IB(x,y,ti - N - 1)\}\}$$

Limitation of the median-based approach for background estimation, and the two-pass approach for noise removal, though effective; is computationally expensive. **Hegde et al.** [17] proposed a technique for identifying a moving object in a video clip. Dynamic objects are identified using both background elimination and background registration techniques. The method may be affected by the presence of noise in image. **Johnsen et al.** [19] used Approximated median filter to perform background modeling. For the implementation, better results were obtained by scaling the increment and decrement by a step factor if the absolute difference between the current pixel and the median-modeled background pixel is bigger than a threshold.

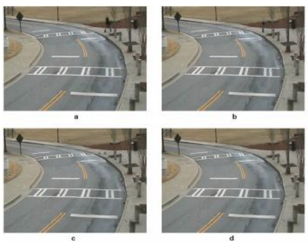

Fig 9: Background estimation using median fig. Based approach and F=a) 10 frames b) 20 frames c) 40 frames d) 60 frames [15]

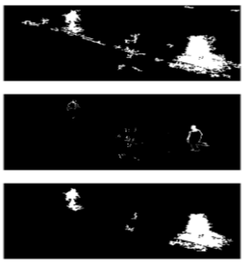

Fig10: illustrating two pass using approach with loose estimate, tight estimate and combination of two.

Foreground pixels are detected by calculating the Euclidean norm at time t:

$$||It(x,y) - Bt(x,y)|| > Te$$

where I_t is the pixel intensity value, B_t is the background intensity value at time t and T_e is the foreground threshold or by checking $|Ij,t - B\,j,t| > Ta$

for j = 1,..., c where Ta is the foreground threshold,

$$It = [I1,t \dots Ic,t]T\ , Bt = [B1,t \dots Bc,t]T$$

and c is the number of image channels. The foreground thresholds Te and Ta are determined experimentally. The foreground pixels were detected by determining the threshold Ta. **Qian Zhang et al.** [28] propose Adaptive background penalty with occlusion reasoning is proposed to separate the foreground regions from the background in the initial frame. First, adaptive background penalty with occlusion reasoning is developed to separate the foreground regions from the background. When the segmentation starts from the initial frame with overlapping objects, not all parts of the objects in the target view in Fig. 12(a) are also visible in other reference views, as shown in Fig. 11(a) and (b). The inter-view occlusions displayed in Fig. 11(c) contain not only the inter object occlusion but also the intra-object occlusion at the interior of the object. Since the inter-object occlusion is mainly located in the background regions, it deserves a larger value of αb_p to enforce its likelihood to be the background. It is achieved this in following equation to adaptively change the value of α_{bp} using the following motion statistics:

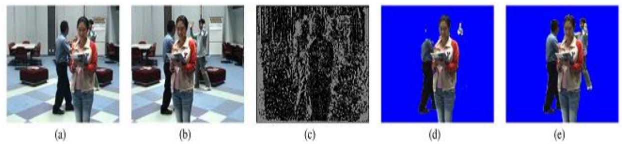

Fig: 11. Adaptive background penalty with occlusion reasoning.a) Left reference view b) Right reference view c) Combined occlusion d) result with constant αb_p e) Result with adaptive αb_p[28]

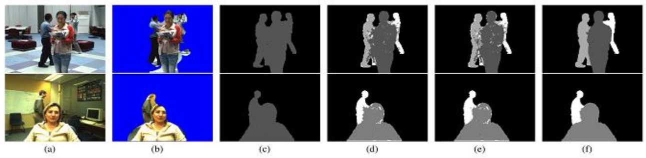

Fig 12: Segmentation of individual object. a) Target view of initial frame b) Extracted initial objects c) Foreground region.d)Initial labeling by depth clustering e) Improved classification using depth ordering.f) Objects segmentation results. [28]

$$\alpha bp = \frac{\log h(\mathrm{mp}|\mathrm{fp} = 0, \theta\mathrm{m})}{\log h(\mathrm{mp}|\mathrm{fp} = 1, \theta\mathrm{m}) + \eta}$$

where f_p=0 for the static background and f_p=1 for the moving object. m_p is the motion vector of p, and θ_m is the motion foreground and background distribution modeled in Fig. 12(b) using the histogram.η is a small value to avoid the division by zero. In Fig. 12 (c), multiple overlapping objects are first separated from the background as foreground regions. Segmentation of individual objects is equivalent to a k -class pixel labeling problem. By assuming that the human objects are in the different depth layers, a coarse labeling field shown in Fig. 12(d) can be obtained by the k-means clustering of the scene, some objects may suffer from complete occlusion.

2.4.2 Optical Flow

Optical flow is one of the widely used methods. It is a dense field of displacement vectors which defines the translation of each pixel in a region. In this method, the apparent velocity and direction of every pixel in the frame have to be computed. This method is very attractive in detecting and tracking objects in video with moving background or shot by a moving camera. Dalal et al. [25] developed a detector that could be used to analyze film and TV content, or to detect pedestrians from moving car applications in which the camera and the background often move as much as the people in the scene. It studies oriented histograms of various kinds of local differences or differentials of optical flow as motion features, evaluating these both independently and in combination with the Histogram of Oriented Gradient (HOG) appearance descriptors. Briefly, the HOG method tiles the detector window with a dense grid of cells, with each cell containing a local histogram over orientation bins. At each pixel, the image gradient vector is calculated and converted to an angle, voting into the corresponding orientation bin with a vote weighted by the gradient magnitude. Votes are accumulated over the pixels of each cell. The simplest approach is to treat the two flow components I^x, I^y as independent 'images', take their local gradients separately, find the corresponding gradient magnitudes and orientations, and use these as weighted votes into local orientation histograms in the same way as for the standard gray scale HOG. It can be called this family of schemes as Motion Boundary Histograms (MBH). One of the weaknesses with optical flow is that it works well if background is stationary only.

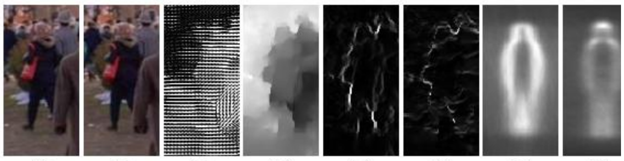

Fig16: Illustration of MBH descriptor.(a,b) Reference images at time t and t+1 (c,d) compound optical flow and flow magnitude showing motion boundaries.(e,f) Gradient magnitude of flow field (g,h) [25]

2.4.3 Gaussian Mixture

Bodor et al. [5] tried to develop automated intelligent vision-based monitoring systems. They detect objects appearing in a digitized video sequence with the use of a mixture of gaussians for background/foreground segmentation.

The algorithm [5] works as:-

1. The values of a pixel are modeled as a mixture of Gaussians.
2. At each iteration gaussians are evaluated using a simple heuristic to get which are likely to correspond to background.
3. Pixels that do not match with the "background Gaussians" are classified as foreground.
4. Foreground pixels are grouped using 2D connected component analysis

The limitation is rapid changes in the lighting of the outdoor scene such as those caused by the sun suddenly going behind/emerging from a cloud introduced some error in the tracking system.

3. DISCUSSION & CONCLUSION

This paper presents an extensive survey of object detection approaches and also gives a brief review of each approach. Various object detection approaches are discussed as feature based, template based, classifier based, motion based as per the reviewed papers. Feature based approaches may not provide effective detection for multiple objects and also affects if background and foreground object has got similar color. Detection of a reappearing target is somewhat difficult with template based object detection and also shape of fixed template is one of the limitations. With classifier approach, the object boundaries are not always perfectly located due to the classification error. Motion based approaches are most widely used and is found to be very effective approach for object detection, although it may be affected if video is noisy. One more important fact is observed that mostly researchers work on detection of single object. This survey on approaches of object detection with rich bibliography content can give valuable insight into this important research topic and encourage new research.

4. REFERENCES

[1] Alok K. Watve, Dr. Shamik Sural "*Object tracking in video scenes*", Technical seminar at IIT Kharagpur, 1998

[2] Yiwei Wang and John F. Doherty, Robert E. Van Dyck "*Moving Object Tracking in Video*," *1999*

[3] Dong Kwon Park, Ho Seok Yoon and Chee Sun Won, "*fast object tracking in digital video",* IEEE Transactions on Consumer Electronics, Vol. 46, No. 3, august 2000

[4] Yu Zhong, Anil K. Jain, M.-P. Dubuisson-Jolly," *Object Tracking Using Deformable Templates* ",IEEE transactions on pattern analysis and machine intelligence, vol. 22, no. 5, may 2000

[5] Robert Bodor, Bennett Jackson, Nikolaos Papanikolopoulos," Vision-*Based Human Tracking and Activity Recognition*", 2000

[6] Paul Viola, Michael Jones, Daniel Snow," *Detecting Pedestrians Using Patterns of Motion and Appearance* ", Proceedings of the International Conference on Computer Vision (ICCV), October 13, 2003, Nice, France.

[7] Sebastien Lef evre and Emmanuel Bouton and Thierry Brouard and Nicole Vincent*," A new way to use hidden markov models for object tracking in video sequences* "2003

[8] Lan Wu," *Multiview Hockey Tracking with Trajectory Smoothing and Camera Selection* ", 2005

[9] Yi Liu, Yuan F. Zheng, "*Video Object Segmentation and Tracking Using* ψ *-Learning Classification* ", IEEE transactions on circuits and systems for video technology, vol. 15, no. 7, July 2005

[10] Karthik Hariharakrishnan and Dan Schonfeld," *Fast Object Tracking Using Adaptive Block Matching* ",IEEE transactions on multimedia, vol. 7, no. 5, oct 2005

[11] Wei-Lwun Lu, James J. Little," *Simultaneous Tracking and Action Recognition using the PCA-HOG Descriptor* "Proceedings of the 3rd Canadian Conference on Computer and Robot Vision (CRV'06)

[12] Chieh-Chih Wang, Charles Thorpe, Martial Hebert, Sebastian Thrun, Hugh Durrant-Whyte," *Simultaneous Localization, Mapping and Moving Object Tracking*" Carnegie Mellon University Research Showcase, 2007

[13] Young Min Kim," *Object Tracking in a Video Sequence"*, 2007

[14] Yuhua Zheng1 and Yan Meng," *Object Detection And Tracking Using Bayes-Constrained Particle Swarm Optimization* ", ISBN 978-1-60021-992-4, 2007 Nova Science Publishers, Inc

[15] Sourabh Khire & Jochen Teizer," *Object Detection and Tracking* ", 2008

[16] Jiyan Pan, BoHu, and Jian Qiu Zhang," *Robust and Accurate Object Tracking Under Various Types of Occlusions* ", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 18, No. 2, February 2008

[17] Vibha L, Chetana Hegde, P Deepa Shenoy, Venugopal K R, L M Patnaik," Dynamic Object Detection, Tracking and Counting in Video Streams for Multimedia Mining ",IAENG International Journal of Computer Science,35:3,IJCS_35_3_16,21 august 2008

[18] Zhenjun Han, Qixiang Ye, Jianbin Jiao," *Online feature evaluation for object tracking Using kalman filter* "978-1-4244-2175-6/08/$25.00 ©2008 IEEE

[19] Swantje Johnsen and Ashley Tews," *Real-Time Object Tracking and Classification Using a Static Camera* "Proceedings of the IEEE ICRA 2009 Workshop on People Detection and Tracking Kobe, Japan, May 2009

[20] Junda Zhu, Yuanwei Lao, and Yuan F. Zheng, "*Object Tracking in Structured Environments for Video Surveillance Applications* ", IEEE Transactions On Circuits And Systems For Video Technology, Vol. 20, No. 2, February 2010

[21] Qiang Chen, Quan-Sen Sun, Pheng Ann Heng, and De-Shen Xia," *Two-Stage Object Tracking Method Based on Kernel and Active Contour* ", IEEE Transactions On Circuits And Systems For Video Technology, Vol. 20, No. 4, April 2010

[22] Xiaoqin Zhang, Weiming Hu, Wei Qu, and Steve Maybank," *Multiple Object Tracking Via Species-Based Particle Swarm Optimization* ",IEEE Transactions On Circuits And Systems For Video Technology, Vol. 20, No. 11, November 2010

[23] Zulfiqar Hasan Khan, Irene Yu-Hua Gu, and Andrew G. Backhouse," *Robust Visual Object Tracking Using Multi-Mode Anisotropic Mean Shift and Particle Filters*", IEEE Transactions On Circuits And Systems For Video Technology, Vol. 21, No. 1, January 2011

[24] Li Sun and Guizhong Liu," *Visual Object Tracking Based on combination of Local Description and Global Representatio*n ",IEEE Transactions On Circuits And Systems For Video Technology, Vol. 21, No. 4, April 2011

[25] Navneet Dalal, Bill Triggs and Cordelia Schmid," *Human Detection using Oriented Histograms of Flow and Appearance* "2011

[26] S. Saravanakumar, A. Vadivel and C.G. Saneem Ahmed," *Human object tracking in video sequences* "ICTACT Journal on Image and Video Processing, August 2011, Volume: 02, Issue: 01

[27] Huabo Sun, Lei Yan, Peter Mooney and Rui Liang*," A new method for moving object detection using variable resolution bionic compound eyes* " ,International Journal of the Physical Sciences Vol. 6(24), pp. 5618-5622, 16 October, 2011

[28] Qian Zhang and King Ngi Ngan," *Segmentation and Tracking Multiple Objects under Occlusion from Multiview Video* ", IEEE Transactions On Image Processing, Vol. 20, No. 11, November 2011

[29] Xiaobai Liu, Liang Lin, Shuicheng Yan, Hai Jin, and Wenbin Jiang," *Adaptive Object Tracking by Learning Hybrid Template Online* ",IEEE Transactions On Circuits And Systems For Video Technology, Vol. 21, No. 11, November 2011

[30] John Chiverton, Xianghua Xie and Majid Mirmehdi," *Automatic Bootstrapping and Tracking of Object Contours* ", IEEE Transactions on Image Processing, Vol. 21, No. 3, March 2012