Object Detection and Extraction of Moving Human Silhouettes in Video Surveillance System

Lovely June Sanglyne[*], Gypsy Nandi
Department of Computer Science & Engineering and Information Technology, Don Bosco College of Engineering and Technology, Assam Don Bosco University, Guwahati, India

Abstract— Compared to the past decades, the growth in the use of video cameras have rapidly changed. Also, the use of video cameras specifically for surveillance has seen a rapid increase on a day-by-day basis. Video surveillance has become a budding approach for prevention of crimes as it facilitates the ability to handle a suspect when any crime occurs. Generally video surveillance consists of three phases including object detection, object classification and tracking. This paper presents a study on the three phases of video surveillance and is directed on object classification and estimating human pose(s) or action recognition by extracting the silhouette or shape of a human body. In this paper we use the Moore-Neighbor tracing algorithm and the adaptive background subtraction method for producing the foreground human blobs called human silhouettes. The silhouette obtained can be further used for human pose estimation or behavior analysis or action recognition.

Keywords— Video surveillance, segmentation, object detection, silhouette, action

I.   INTRODUCTION

The process of analysing video sequences is known as video surveillance. Video surveillance can be done in three ways - manual, semi-autonomous and fully-autonomous [1]. In manual video surveillance, video analysis is done by human. Semi-autonomous video surveillance includes some form of video processing as well as significant human involvement. In such a case, the video that is recorded in the presence of motion is sent to a human expert for further analysis. In a fully-autonomous system there is no involvement of human and the system does all the tasks required for video surveillance. During surveillance, video sequences are taken as the only input and then the system itself does both the low-level and the high-level decision making tasks. Motion detection and abnormal event detection are the low-level decision making task and the high-level decision making task respectively.

Video surveillance has become a dominant and central technique in monitoring the movements of a scene in any venue such as museums, airports and banks. This can be done for security-conscious, to detect suspicious activities or unlikely events. To do so, the video captured can be analysed by following a set of three major phases or steps as mentioned below:

- **Detection of Moving Object** – in this phase, recognition of the interested moving objects in the video is done by locating connected regions of pixels that represent the moving objects within the scene.
- **Tracking of Detected Object** – in this phase, the detected object such as a human is kept track of from frame to frame.
- **Action Analysis** – in this phase, identification and analysis of unusual behaviour or activity of an object is done. For instance, analyzing the movement of a human being to determine abnormal behaviour.

An object can be defined as anything that is of concern for further study. Objects are mainly characterized by their shapes and appearances such as primitive geometric shapes, points, skeletal models, object silhouette and contour. The likely feature selection includes colour, edges, texture and optical flow. Detection of moving object in video is a preliminary phase for video surveillance since it provides attention and also simplifies every tracking method and subsequent analysis steps. Object detection can be done either in each frame or when the object first appears in the video [2]. Detection handles segmentation of moving objects and separates them from motionless background objects. Segmentation is the process for partitioning the image into perceptually similar regions. Segmentation decreases computation time since it is used for extracting the human silhouette by partitioning image into parts. Two common approaches for detecting object are the use of information which is in a single frame, and the use of temporal information which is in the form of frame differencing. Temporal differencing highlights the region that changes in consecutive frames. To avoid or reduce false detection, this information is figured from a series of frames. After the object has been detected, the tracker will then produce the tracks from one frame to the next.

Some frequently used methods for detecting moving objects in video surveillance are background subtraction, statistical methods, temporal differencing and optical flow.

A. Background Subtraction

Background subtraction also known as foreground detection [3] is the process of detecting objects by determining the difference between the current image and background image. Changes occur from this comparison indicates a moving object. Results are used for further processing such as tracking of the detected object, object classification and behaviour analysis.

B. Statistical methods

Statistical methods overcome the disadvantage of the basic background methods by extracting change regions from background and dynamically update the background during processing. Pixels that belong to background image are recorded and updated dynamically, in each frame. Each pixel's statistics are compared with that of the background model to identify the foreground pixel. This approach is reliable in scenes that contain noise, illumination changes and shadow.

C. Temporal differencing methods

Temporal differencing methods extract moving regions by considering the pixel-wise difference between two or more consecutive frames in a video. This approach is adaptive to dynamic scene changes but fails to extract all relevant pixels of a foreground object texture when an object moves slowly, or when an object stops.

D. Optical flow

The optical flow describes the direction and time rate of pixels in a time sequence of two consequent images. A two dimensional velocity vector, carrying information on the direction and the velocity of motion is assigned to a given place of the picture.

Detected moving objects can be tracked using tracking approaches such as Point tracking, Kernel tracking or Silhouette tracking [2].Tracking is the process of determining or estimating the path of an interesting object such as human, in the plane of an image or across the frames from image sequences or video, as it moves around a scene. Motion tracking can be classified as shown in table I below.

Table I. Classification of Motion Tracking

Categories	
Based on positioning of imaging instrument	• Static camera tracking • Active motion tracking
Based on motion of objects	• with markers (used in 2D and 3D animations) • Marker-less (our interest)
Based on methodology or approach by which motion tracking can be set off	• motion based tracking • recognition based tracking
Based on the type of motion to be tracked in human motion tracking	• articulated motion • moving motion

We now briefly introduce main tracking approaches:

Point tracking: Point tracking methods are robust, reliable and precise. These are mainly used to track the vehicles. Objects are detected in consecutive frames represented by points (see figure 1(a)).

Kernel tracking: Kernel refers to the object shape and appearance such as a rectangular template or an elliptical shape with an associated histogram .Objects is tracked by computing the motion of the kernel in consecutive frames (Figure 1(b)). This motion is usually in the form of a parametric transformation such as translation, rotation, and affine.

Silhouette tracking: Silhouette tracking is one of the most appropriate methods for tracking human or complex object with shapes such as head, shoulder and hands (see fig 1(c)) because these objects cannot be well described by other representation such as simple geometric shapes.
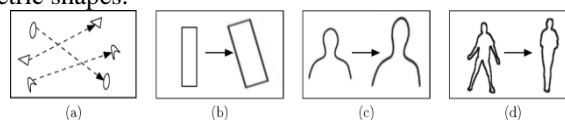

(a)   (b)   (c)   (d)

Fig. 1. Tracking approaches [2] (a) Multipoint correspondence point tracking, (b) parametric transformation of a rectangular patch,(c,d)Two examples of contour evolution(silhouette tracking).

Tracking methods are used for tracking the human motion or identifying the behaviour or pose estimation to detect the incident of the possible dangerous or abnormal cases. The advanced video surveillance system requires analysis to detect such suspicious events.

In this paper we use Silhouette tracking for extracting the human shape, therefore we discussed mainly on Silhouette tracking. Silhouette based methods provide an accurate shape description for an object by representing their edges or boundary (see figure 2(a)).Silhouette-based object tracker finds the object silhouette (see figure 2(b)) in each frame by

means of an object model which can be in the form of the object contour, edges or colour histogram. Silhouette trackers are further categorized into contour tracking and shape matching. Shape matching approaches track for the object silhouette by considering only the current frame .In contour tracking approaches, an initial contour to its new position in the current frame is considered.

The object silhouettes help in classifying the objects and analysing their behaviour. Before classifying moving objects, the object silhouette should be extracted from image sequences. After detecting an object, its actions can be recognized. Generally, the action recognition system can recognize different human actions such as walking, jumping, running, boxing, crawling and kicking.
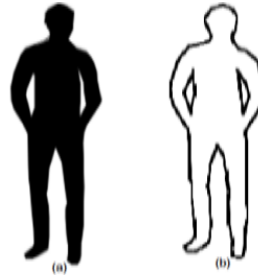

Fig. 2. (a) object silhouette, (b) control points on object contour.

II. RELATED WORK

This section discusses mainly about the related work on the three phases of video surveillance, that is, detection of moving object(s), tracking of detected object(s), and action analysis. Object classification methods used for classifying objects into human and non-human are also discussed.

E. Object Detection

Object detection is classified into point detectors, segmentation and background modelling. Several common object detection categories and their related work are tabulated in table II.

Table II. Categories of object detection

Categories	Representative Work
Point detectors	Moravec's detector [3], Harris detector [4], Scale Invariant Feature Transform [5]. Affine Invariant Point Detector [6].
Segmentation	Active contours [7]. Mean-shift [8], Graph-cut [9],
Background Modeling	Mixture of Gaussians [10], Wall flower [11],Eigenbackground[12], Dynamic texture background [13].

A. Point Detector

Images may have an expressive texture in their respective localities .Interest points in such images can be detected using Point detectors. Some of the commonly used interest point detectors are Moravec's interest operator [3], Harris interest point detector [4], SIFT detector[5], interest point detectors[6].

B. Segmentation

In contour evolution, object segmentation is attained when the contour tightly encloses the objects' boundary [7]. Comaniciu and Meer [8] proposed the mean-shift approach for finding clusters in the joint spatial+color. Mean-shift clustering approach is applicable in edge detection, image regularization [8], and tracking [9]. Shi and Malik[9] proposed the normalized cut for segmentation .

C. Background Modelling

Background model is the representation of the scene used for detecting moving objects. In case of background model, the process used for finding deviations from the model for each incoming frames is called background subtraction.

In the work of Stauffer and Grimson [10], a mixture of Gaussians was used for modelling the pixel color. Oliver et al. [11] proposed a holistic approach using the eigenspace decomposition, Toyama et al.[12]used a three-tiered algorithm to deal with the background subtraction problem. Monnet et al.[13] proposed methods which are capable for handling time-varying background such as the waves on the water, moving clouds, and so on.

D. Silhouette Tracking

Silhouette tracking is one of the appropriate methods for tracking moving objects such as pedestrians, vehicles, clutter, etc, in video surveillance. Silhouette tracking methods are categorised into shape matching and contour tracking.

Table III. Categories of Silhouette Tracking

Silhouette Tracking	Representative works
Matching shapes	Hausdorff [15], Histogram [16].
Contour formation	State space models [17,18], Variational methods [19], Heuristic methods [22, 23].

Shape matching

In 1993, Huttenlocher et al [15] used edge-based representation to perform shape matching. Hausdorff distance is used for measuring the most mismatched edges and for constructing a correlation surf by emphasizing much on those edges which is not drastically affected by object motion.

The minimum is selected as the new object position. For example, in the case of a walking human, the shape of the torso and the head do not change much, whereas the shape of the legs and arms are changed drastically by motion. The edges corresponding to arms and legs are removed to improve the tracking performance. In 2004 Kang et al.[16] used histograms of colour and edges as the object models by generating histograms from concentric circles.

Contour tracking

Contour tracking can be done either by using state space models or by using direct minimization methods such as gradient descent. In contour tracking using state space models, the object's state is defined by the shape and the motion parameters of the contour. The object state has been defined in terms of dynamics of the control points [17], affine motion parameters and spline shape parameters [18]. The measurements consist of image edges (see figure 3(a)).
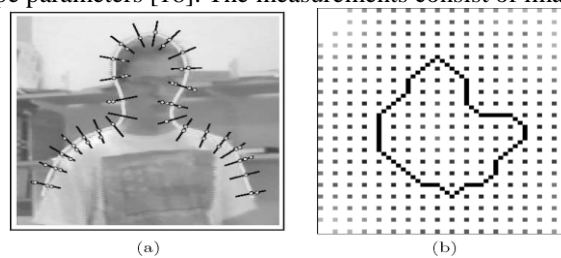

Fig. 3. (a) Edge observations along the contour normals [18]. (b) Level set contour representation[19]

The contour energy is defined in terms of temporal information [19, 20, 21] or appearance statistics generated from the object and the background regions [22, 23]. To develop the contour, Bertalmio et al. [19] use the optical flow constraint. Their main aim was to compute the flow of vector for each contour position using the level set representation (see Figure 3(b)). Yilmaz et al. [23] used the level set-based shape model that determines the object occlusions during the tracking process.

E. Object Classification

The approach presented in [24] makes use of the objects' silhouette information to classify detected objects. Detected objects are classified into human, vehicle and so on. The classification method developed by Collins et al. [25] recognizes four classes of objects. These are human, human group, vehicle and clutter.

F. Analysis

Behaviour analysis or action recognition methods in video surveillance are classified into three groups, according to the methods they used. These are general signal processing techniques [26], template matching [27] and state-space approaches [28].

III. PROPOSED WORK

Our proposed work for object detection and extraction of moving human silhouette consists of four modules. It includes background subtraction (for detecting the moving object), contour formation (shape representation for extracting the human silhouette), object classification (human or non-human) and behaviour analysis (silhouette tracking). Framework of our proposed work is shown in Figure 4.

At first the video is taken as input and is then converted into frames. Next, background subtraction is used for detecting human or the moving objects in a given input file, rejecting the background of the frame. Then, contour formation is used for forming the boundary pixel of the object. Contour formation of all objects is done before analysing the behaviour of the moving object. It is in fact the shape of the objects that help in classifying the objects to be human or non-human. Thus after object classification, if the object is human it can be further analysed.
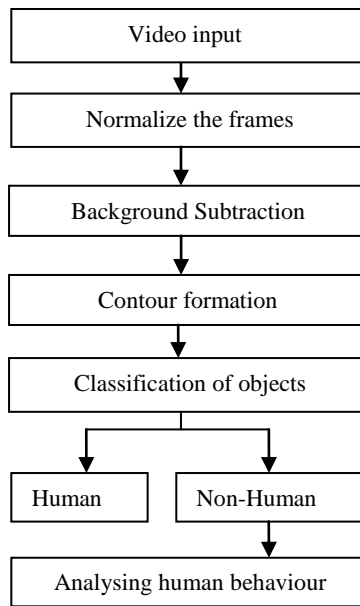
```
┌─────────────────────────┐
│      Video input        │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│   Normalize the frames  │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│  Background Subtraction  │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│    Contour formation    │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│  Classification of objects │
└─────────────────────────┘
        │         │
        ▼         ▼
   ┌────────┐ ┌────────────┐
   │ Human  │ │ Non-Human  │
   └────────┘ └────────────┘
            │
            ▼
┌─────────────────────────┐
│ Analysing human behaviour │
└─────────────────────────┘
```

Fig. 4. Framework of proposed work for object recognition

1. Background subtraction

In our proposed work for background subtraction, adaptive Statistical background model is used.The Process flow chart is shown in Figure 5.

```
┌─────────────────────────┐
│          Video          │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│    Convert to Frames    │
└─────────────────────────┘
        │         │
        ▼         ▼
┌────────────┐ ┌────────────┐
│ Background │ │  Current   │
│   Frame    │ │   Frame    │
└────────────┘ └────────────┘
        ▲         │
        │         ▼
┌──────────┐ ┌────────────┐
│ Morphing │ │ Difference │
└──────────┘ └────────────┘
        ▲         │
        │         ▼
        │   ┌──────────────────┐
        │   │ Threshold operation │
        │   └──────────────────┘
        │         │
        │         ▼
        │   ┌──────────────┐
        └───│ Moving Object │
            └──────────────┘
```
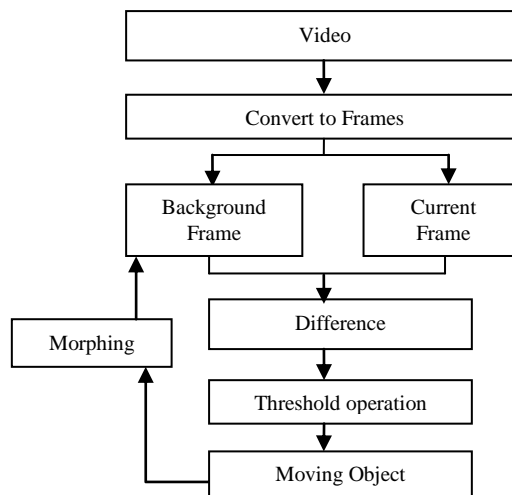
Fig. 5. Flowchart of Background Subtraction Method

Before classifying and tracking the human blob, the moving objects should be extracted from the background. The absolute value obtained is filtered with dynamic threshold per pixel [3]. The pixel where the difference is greater than the bounded threshold implies a foreground pixel. The success of motion detection depends on the threshold value. A lot of false change points will be produced if the threshold value is too small. The scope of changes in movement will be reduced if the threshold choice is too large [7].The background image is updated by morphing slightly toward the current frame. Incoming information are provided into the current background image. The faster new changes (moving objects) in the scene are updated to the background frame.

The idea behind the algorithm is as follows:
- Store the first frame image as a background image B.
- Perform subtraction of the current image C_k with B.

$$R_k(x,y) = B(x,y) - C_k(x,y) \qquad (1)$$

$$F_k(x,y) = \begin{cases} 1 & R_k(x,y) > T \\ 0 & R_k(x,y) \square\ T \end{cases}$$

- Update the background image B for every frame $_k$ by using a first order recursive filter. The equation is

$$B_{k+1} = \square\ \square\ B_k + (1 - \square\ \square\ ) C_k \qquad (2)$$

Where $\square\ \square$ is an adaptation coefficient and cannot be too large. If   is too large, it might form artificial "tails" behind the moving objects.

Algorithm for background subtraction

Input: Image of the video.

Output: Foreground Object F.

1. For each pixel(k) ,calculate the difference

R(k) -> | Backgroundframe(k) –Currentframe(k) |

2. If R(k)>threshold

   i. Set F(k) = True;

3. Else

   ii. Set F(k) = False;

4. Set the adaptation coefficient *a* and then update the Backgroundframe.

   i. B(k) -> *a**B(k)+(1-*a*)*C(k)

The advantage of this algorithm is that it is effective detection, can provide the complete feature data of the target. It overcomes the shortcomings of using Simple Background subtraction method to obtain the background image when there are frequent in the moves of the occasions such as changes in dynamic scene. This method reduces computation time significantly.

2. Contour Formation

Contour formation is used to find the boundary or edges of an object .The region inside the contour are called the silhouette of the object .In our proposed system we are using the Moore-Neighbor tracing algorithm to find the contour of the given binary image.

The idea behind this algorithm is as follows:

- Define **M(a)** to be the neighbour of pixel **a**.
- Let **p** denote the current boundary pixel.
- Let **c** denote the current pixel under consideration.
- Find a black pixel and declare it as a "start" pixel, **s.**
- By using this black pixel, extract the contour by going around the pattern in a clockwise direction.
- Backtrack when a black pixel, **P**, is hit. That is, move to the pixel from which **s** was entered.
- The algorithm terminates when the start pixel is visited for a second time.

***Algorithm for contour formation* [29]**

Input: Binary Image.

Output: The contour consisting of a sequence **B (b₁, b₂ ,b₃,b₄..., bₖ)** of boundary pixels .

Begin

1. Set **B** to be empty.
2. From bottom to top and left to right scan the cells until a black **start** pixel, **s,** of **p** is found
   i. Insert **s** in **B**.
   ii. Set p=s
3. Backtrack
4. Set **c** to be the next clockwise pixel in **M(p)**.
5. While **c** not equal to **s** do
   i. If **c** is black
      a. insert **c** in **B**
      b. set **p=c**
      c. backtrack
   ii. else
      a. advance the current pixel **c** to the next clockwise pixel in **M(p)**
6. end While

End

The algorithm terminates in two cases:

a) When it has rotated 3 times (each through 90 degrees clockwise), thus declare the pixel an isolated one.

b) When the current boundary pixel is the **start** pixel, thus declare that it has traced the contour of the pattern.
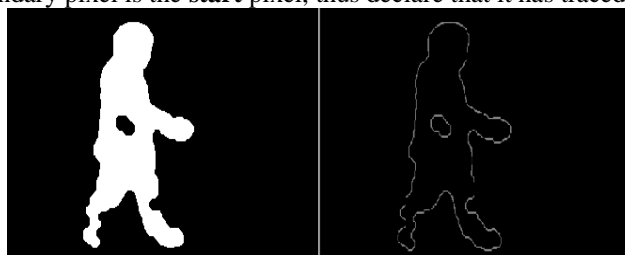

Fig. 6. Contour formation

The advantage of this algorithm is that it checks the whole neighbourhood of a boundary pixel in order to find the next boundary pixel. It will always be able to extract the outer boundary of any connected component, since this algorithm proceeds to check every pixel in the neighbourhood of the current boundary pixel, it is bound to detect the next boundary pixel.

3. Object Classification

In natural scene, different shapes may correspond to different moving objects such as people, vehicles, animals, natural phenomenon (such as rain, snow), plants and clutter. Under this assumption, object classification is a necessary process that can analyse moving regions to recognize human being from other moving objects. In our proposed system, object silhouettes are used to classify objects. Our object classification metric is based on the object shapes and their similarity.

The primary step of object classification is collecting the different poses of different objects. Then these samples are train with certain feature .Then find out the threshold value for human blob and check this threshold value with all objects blob while testing with real time video.
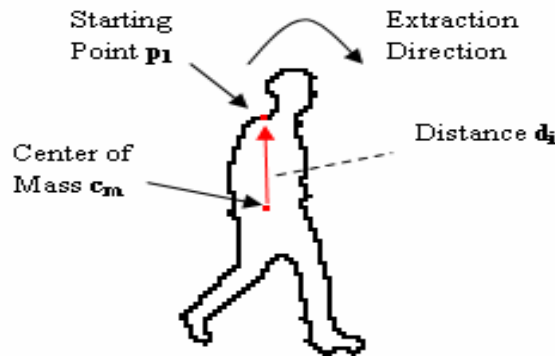

Fig.7. A sample silhouette for calculating the distance signal.[29]

Algorithm for object classification
Input: All Objects' blob
Output: Human Silhouette
1. Calculate the centre of mass of each object C_m.
   $C_m = (xC_m, yC_m)$, using the following equation

$$xC_m = \frac{\sum_i^n xi}{n}, \quad yC_m = \frac{\sum_i^n yi}{n} \qquad (1)$$

2. Find points S= {p1, p2,p3,....pn} of each object O.
3. Find DS={d1,d2,d3...dn} using the following equation
   $$di = Dist(C_m, p_i), \quad \square \quad i \in [1....n] \qquad (2)$$

   Where Dist function is the Euclidean distance
   $$di = \sqrt{(xi - xc)^2} + \sqrt{(yi - yc)^2}$$
4. Find the fix-sized distance signals

$$\overline{DS}[i] = DS\left[i * \frac{N}{C}\right]; \quad \square \quad i \in [1....c] \qquad (3)$$

5. Do normalization for the distance signal

$$\overline{DS}[i] = \frac{\overline{DS}[i]}{\sum_1^n \overline{DS}[i]} \qquad (4)$$

4. Analysing human behaviour

Action can be recognized once the object has been detected according to its type. If it is a human, its actions can be analysed.

Steps for analysing human behaviour:
1. Manually create the template pose database and store the extracted object silhouettes. The object silhouettes for all of the actions that can be recognized by the system contain key poses with integer IDs in the range [1,IDMAX].
2. Create action templates. Actions can be represented with a histogram of key poses (pose Ids) it matches. That is, create a histogram H_j for the action of the size of the total number of key silhouettes in the silhouette template database.
3. After creating action template database, test actions are recognized in real-time.In order to recognize an action we keep a circular list of the IDs of the matching silhouettes in the template pose database for the subject's silhouette.
   a. Let $A_T = \{S_{i-(w-1)}, S_{i-(w-2)}, ..., S_i\}$ be the fixed length list of the silhouettes of a test subject in the last w frames of video. For each S_i, a corresponding pose template match P_i is found in the silhouette pose template database by using the same distance metric used during training(section III.3)

b. Let $L_T = \{P_1, P_2,..., P_N\}$ represent the list of matched pose IDs, where $P_i \in [1, ID_{MAX}]$. A normalized histogram H_T of IDs is created by using the IDs in L_T.

c. The distance between H_T and each action template histogram H_j in the action database is calculated using Euclidian distance metric.

IV. CONCLUSION

The goal of this paper is to give an overview for real-time object classification and human action recognition using object detection and object silhouettes, which is applicable in outdoor environments as well as indoor environments. An adaptive statistical background model is discussed for background subtraction. Moore's algorithm has been discussed to trace the contour for all image objects. Silhouette based method is consider to separate human from other objects and therefore analyse their behaviour. Data from different environmental and physiological sensors will be used in future work.

REFERENCES

[1]   Su Kim, Hong Seok Choi, Kwang Moo Yi, Jin Young Choi, and Seong G. Kong, Intelligent Visual Surveillance - A Survey, in the International Journal of Control, Automation, and Systems 8(5): pp.926-939,2010

[2]   Yilmaz, A., Javed, O., & Shah, M, "Object tracking", *ACM Computing Surveys*, *38*(4), 13–es. doi:10.1145/1177352.1177355 ,2006

[3]   H. Moravec, "Visual mapping by a robot rover", In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*. pp. 598–600, 1979.

[4]   C. Harris, and M. Stephens, "A combined corner and edge detector," In *4th Alvey Vision Conference*, pp. 147–151, 1988.

[5]   D. Lowe, "Distinctive image features from scale-invariant keypoints," In *t. J. Comput.*, Vision 60, pp.2,91–110,2004.

[6]   K. Mikolajczyk, and C. Schmid, "An affine invariant interest point detector," In *European Conference on Computer Vision (ECCV)*., vol. 1,pp. 128–142, 2002.

[7]   V. Caselles, R. Kimmel,and G. Sapiro, "Geodesic active contours," In *IEEE International Conference on Computer Vision (ICCV)*., pp. 694–699,1995.

[8]   D. Comaniciu, and P. meer, "Mean shift analysis and applications," In *IEEE International Conference on Computer Vision (ICCV)*.,vol. 2, pp1197–1203, 1999.

[9]   J. Shi, and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Patt. Analy. Mach.Intell.*, 22, 8,pp. 888–905, 2000.

[10] C. Stauffer, and W. Grimson, "Learning patterns of activity using real time tracking," *IEEE Trans.Patt. Analy. Mach. Intell. 22*, 8, pp. 747–767, 2000.

[11] K. Toyama, B.B. J. Krumm, AND B. Meyers, "Wallflower: Principles and practices of background maintenance," In *IEEE International Conference on Computer Vision (ICCV)*, pp. 255–261, 1999.

[12] N. Oliver, B. Rosario, and A. Pentland, "A bayesian computer vision system for modeling human interactions," *IEEE Trans. Patt. Analy. Mach. Intell. 22*, 8,pp. 831–843, 2000.

[13] A. Monnet, A. Mittal, N. Paragios, and V. Ramesh, Background modeling and subtraction of dynamic scenes. In *IEEE International Conference on Computer Vision (ICCV)*., pp. 1305–1312. 2003.

[14] M. Kass, A. Witkin, and D. terzopoulos, "Snakes: active contour models," *Int. J. Comput. Vision 1*.,pp. 321–332, 1988.

[15] D. Huttenlocher, J. Noh, and W. Rucklidge, "Tracking nonrigid objects in complex scenes," In *IEEE International Conference on Computer Vision (ICCV)*., pp. 93–101, 1993.

[16] J. Kang, I. Cohen, and G. Medioni, "Object reacquisition using geometric invariant appearancemodel," In *International Conference on Pattern Recongnition (ICPR)*., pp.759–762, 2004.

[17] D. Terzopoulos, and R. Szeliski, "Tracking with kalman snakes," In *Active Vision*, A. Blake and A. Yuille, Eds. MIT Press., 1992.

[18] M. Isard, and A. Blake, "Condensation - conditional density propagation for visual tracking," *Int. J.Comput. Vision 29*, 1.,pp. 5–28, 1998.

[19] M. Bertalmio, G. Sapiro, AND G. RANDALL, "Morphing active contours," *IEEE Trans. Patt. Analy. Mach.Intell.*, 22, 7, pp. 733–737, 2000.

[20] A. Mansouri, "Region tracking via level set pdes without motion computation," *IEEE Trans. Patt.Analy. Mach. Intell.*, 24, 7, pp. 947–961, 2002.

[21] D. Cremers, and C. Schnorr, "Statistical shape knowledge in variational motion segmentation," *I.Srael Nent. Cap. J. 21*, pp. 77–86, 2003.

[22] R. Ronfard, "Region based strategies for active contour models," *Int. J. Comput. Vision 13*, 2, pp. 229–251, 1994.

[23] A. Yilmaz, X. Li, and M. Shah, " Contour based object tracking with occlusion handling in video acquired using mobile cameras," *IEEE Trans. Patt. Analy. Mach. Intell.* 26, 11, pp. 1531–1536, 2004.

[24] A. J. Lipton, H. Fujiyoshi, and R.S. Patil, " Moving target classification and tracking from real-time video," In *Proc. of Workshop Applications of Computer Vision*, pp.129–136, 1998.

[25] R. T. Collins et al., "A system for video surveillance and monitoring: VSAM final report". Technical report CMU-RI-TR-00-12, Robotics Institute, Carnegie Mellon University, May 2000.

[26] L. Wang, W. Hu, and T. Tan, "Recent developments in human motion analysis". Pattern Recognition, 36(3): pp. 585–601, March 2003.

[27] B. U. Toreyin, Y. Dedeoglu, A. E. Cetin, "HMM based falling person detection using both audio and video," IEEE International Workshop on Human-Computer Interaction, Beijing, China, Oct. 21, 2005 (in conjunction with ICCV 2005), Lecture Notes in Computer Science, vol. 3766, pages 211-220, Springer-Verlag GmbH, 2005.

[28] O. Chomat, J.L. Crowley, "Recognizing motion using local appearance, International Symposium on Intelligent Robotic Systems, University of Edinburgh," pp. 271-279, 1998.

[29] P.Rajashekar Reddy, V.Amarnadh, Mekala Bhaskar , "Evaluation of Stopping Criterion in Contour Tracing Algorithms," Vol. 3 (3) ,pp.3888-3894, 2012.