HUMAN MOTION DETECTION FOR REAL TIME GESTURE RECOGNITION: A SURVEY

Lovendra Solanki, Research Scholar, Singhania University, Jhunjhunu(Rajasthan) INDIA;
J.L. Raheja, Senior Principal Scientist, CSIR-CEERI,Pilani (Rajasthan) INDIA ;
P.S. Bhatnagar, Director, BKBIET, Pilani (Rajasthan) INDIA

Abstract

The goal of this paper is to review of various moving human object detection methods. This paper focuses on detection of moving human objects in video surveillance system, before tracking the objects in the scene. Detection of person is first low level important task for any video surveillance application. In this survey, we described background subtraction with alpha, statistical method, temporal frame differencing and optical flow methods to detect moving object. After moving objects are detected, to further track people and analyze their activities, it is very necessary to correctly distinguish them from other moving objects, therefore shape based and motion based classification is studied under human body classification.

Keywords: Object detection, background subtraction, statistical methods, temporal frame differencing, human body classification

1 INTRODUCTION

When the umpire on the cricket ground raises his finger towards the sky everybody watching him knows that the batsman is out, similarly when we see a known person we raise our hand to say Hello to other person, and the other person replies back by either raising his hand or just smiling at the other person in acknowledgement. These are the two most common examples of human gestures to which we are very familiar with. But how do we make computer understand the gesture is complex and challenging task to understand, a lot of research have been done in that area to develop algorithms or methods by which computer can understand the human gesture in order to bridge the gap between man and the machine[1].

1.1 Application of Gesture Recognition

Gesture recognition finds its application in many areas, such as[6] : smart surveillance, advanced user interface, motion-based diagnosis etc. An overview of the some of the application that employ gesture interactions is presented here:1)The strong need of smart surveillance systems [7,8] stems from security-sensitive areas such as banks, department stores, cinema halls, sports stadiums, parking lots and borders. Surveillance cameras are already prevalent in commercial establishments, while camera outputs are usually recorded in tapes or stored in video archives and these video data are used as a forensic tool, losing its primary benefit as an active real-time media but what is needed is the real-time analysis of surveillance data to alert security officers to a burglary in progress, or to a suspicious individual wandering around in the parking lot. Nowadays, the tracking and recognition techniques of face [9],[10],[11],[12] and gait [13],[14],[15],[16] have been strongly motivated for the purpose of access control. As well as the obvious security applications.2) With the growing amount of pollution caused by the traffic it is hazardous for traffic signalling person to sustain in that environment for longer duration of time, human gesture can play a great amount of role in helping out by replacing human being with smart display system which can monitor the traffic in real time and can generate appropriate signal for traffic control. 3) Communication among people is mainly realized by speech, therefore speech understanding has already been widely used in early human-machine interfaces. Sign language is an important case of communicative gestures for the deaf people and American, Chinese and Taiwan sign languages are common example of human gesture [3],[4],[5],[6] and it is a good way to help the disabled to interact with computers. Other applications in the user interface domain include sign-language translation, controls, and signalling in high-noise environment such as factories, airports, games and sports [8]. 4) Another important application domain is advanced user interfaces in which human motion analysis is usually used to provide control and command. Arm action of bowlers in cricket, different body movements of athletes and dance steps of dancers are monitored and analysed for further study. 5)Interactive games is the most emerging application of human body gesture where Freeman et al. [8] player’s hand or body position is used to control movement and orientation of interactive game objects such as cars, very recently interactive boxing playing game is introduced where the monitor recognises the motion of the player and adjusts its position to save from punch . Play Station 2 has introduced the Eye Toy, a camera that tracks hand movements for interactive games [10].

1.2 Purpose of the Survey

The primary purpose of this paper is as follows:

1] This paper aims to provide a complete survey of the most recent developments in vision-based human motion analysis so that the beginners who are new to this field can have hand on knowledge of this emerging field. It covers the latest
research ranging mainly from 1980 to 2012. It thus contains many new references not found in previous surveys.

2) This study highlights progresses made in the development of detection of human body for vision-based gestural interfaces as well as challenges that still remain.

1.3 Classification of Human Gesture

From the application area mentioned in the previous section it is clear that gestures made by human being are very expressive and meaningful involving physical movements of the fingers, hands, arms, head, face, or body for conveying meaningful information or interacting with the external world. Different body parts can be used for representing various classes of gestures they are :

a) Hand and Arm Gestures: Recognition of hand poses[1,2,18,19] , sign languages, and entertainment applications (allowing children to play and interact in virtual environments).

b) Head and Face Gestures: Some examples of head and face gesture are: 1) nodding or shaking of head 2) direction of eye gaze 3) raising the eyebrows 4) opening the mouth to speak 5) winking 6) flaring the nostrils and 7) looks of surprise, happiness, disgust, fear, anger, sadness, contempt, etc.

c) Body Gestures: Involvement of full body motion, as in: 1) tracking movements of two people interacting outdoors 2) analyzing movements of a dancer for generating matching music and graphics and 3) recognizing human gait for medical rehabilitation and athletic training.

1.4 Gesture Recognitions Approach

The human gesture recognition approach can be broadly divided into contact based approach[13] and vision based approach[11,12].

A comparison of these two methods are summarised in the Table I

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Contact-based Approach</th>
<th>Vision-based Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>User cooperation Required</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Interference to user</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Accuracy of system</td>
<td>High</td>
<td>Very (Depending on method used)</td>
</tr>
<tr>
<td>Flexible to configure</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Flexible to use</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Occlusion problem</td>
<td>No (Yes in case of metallic obstacle)</td>
<td>Yes</td>
</tr>
<tr>
<td>Health issues of user</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Costly</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

From the above comparison it can be observed that both the approaches have its advantages and disadvantages, where contact based approach is expensive, requires the user cooperation and can be uncomfortable to wear for long time but they are precise. On the other side vision-based devices do not require user cooperation but they are more difficult to configure and suffer from occlusion problems. Contact based devices are more precise except the ultrasonic’s. Also, they generally have no occlusion problems except the magnetic (metal obstacles) and ultrasonic sensors (mechanical obstacles). Concerning health issues, it is observed that some contact-devices can cause allergy because of the mechanical sensor material and cancer risk from magnetic devices etc. Due to health related issues and the cost factor of the contact based devices vision based technology has gained upper hand over the contact based approach that’s why researches are putting in lot of efforts to find out the most efficient vision based method for the study if human gesture and recognition.

1.5 Vision Based Approach for Human Gesture Recognition

The taxonomy which is generally followed for gesture recognition [11, 20, 21] is shown in Fig. 2. Here three major tasks are done in the process of human motion analysis: human detection, human tracking and human behaviour understanding. Although they do have some overlap (e.g., the use of motion detection during tracking), this general classification provides a good framework for discussion throughout this survey. In our survey we will discuss only human motion detection as it is the main steps in human gesture recognition in computer vision.

Figure 1: Block diagram of Human Gesture Classification

1) Contact Based Approach for Gesture Recognition: This method uses either mechanical or optical sensor which are used for digitizing hand, finger(e.g data gloves) and leg motions(e.g. body suit) into multi-parametric data. Extensive description of various tracking devices have been discussed by Sturm and Zeltzer [13].

2) Vision Based Approach for Gesture Recognition: Vision Based methods require only a camera, without the use of any extra devices an interaction between humans and computers is possible and can be implemented using software and/or hardware. This approach also has few drawbacks as these systems need to be background invariant, lighting insensitive, person and camera independent to achieve real time performance.

Table I: Comparison of Contact based and Vision based approach

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Contact-based Approach</th>
<th>Vision-based Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>User cooperation</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Interference to user</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Accuracy of system</td>
<td>High</td>
<td>Very (Depending on method used)</td>
</tr>
<tr>
<td>Flexible to configure</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Flexible to use</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Occlusion problem</td>
<td>No (Yes in case of metallic obstacle)</td>
<td>Yes</td>
</tr>
<tr>
<td>Health issues of user</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Costly</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>
2. HUMAN DETECTION

Moving object detection is the first step of nearly every system of vision-based human motion analysis. It starts with human detection in which foreground (moving human objects) is separated from background. Detection is described as low level processing.

2.1 Moving Object Detection

Moving object detection[32],[34] in video sequences is one of the most essential as well as critical tasks in automatic image processing which aims at detecting regions corresponding to moving objects such as vehicles and people in natural scenes (see Fig 3). Detecting moving spots provides a focus of attention for later processes such as tracking and activity analysis because only those changing pixels needed to be considered. However weather changes, illumination fluctuations, shadow, repetitive motions (e.g. waving tree leaves), camera noise make motion segmentation poses difficulty to process image data[23,24]. At present, most segmentation methods use either temporal or spatial information of the images.

Now we will study different methods for detecting moving objects from background.

i]Background Subtraction:

Background subtraction [69] is most basic and popular method for motion segmentation in this method moving region is detected by differencing the current and reference background frame in pixel by pixel manner. Lipton et al. [22] presented a two-frame differencing scheme where the pixels that satisfy the following equation are marked as foreground.

\[
|I_{(x,y)} - B_{(x,y)}| > \text{Threshold} \tag{1}
\]

Here  denotes the current frame pixel and  is background image pixel. The pixels where the difference is above a threshold are classified as foreground.

After creating a foreground pixel map, some morphological post processing operations such as erosion, dilation and closing are performed to reduce the effects of noise and enhance the detected regions. The reference background is updated with new images over time to adapt to dynamic scene changes. This technique is suitable where the background is static (e.g. motion detection inside a room where there is no illumination or background variation).

In the approached proposed by Yang & Levine [31] could handle some inconsistencies due to illumination variation in which the background is modelled by taking median value of pixel color over a series of images. The median as well threshold value is determined using histogram procedure was used to create the difference image because of the algorithm.

Heikkila and Silven[25] uses the simple version of this scheme where a pixel at location  in the current image  is marked as foreground if equation (1) is satisfied. The background image  is updated by the use of an Infinite Impulse Response (IIR) filter as follows:

\[
B_{i+1} = \alpha I_t + (1-\alpha)B_i
\]

This method fails to detect stopped objects in the scene or if the moving object is quite big (e.g. an elephant) therefore additional methods are required to be adopted in order to detect these types of objects for the success of higher level processing. In order to overcome shortcomings of two frame differencing in some cases, three frame differencing can be used [20]. Generally most researchers have shown their interest in building different adaptive background models in order to reduce the influence of dynamic scene changes on motion segmentation. K.P. Karmann et al [29] and M. Kilger [30] proposed adaptive background model based on Kalman filter, this model adapted to the temporal changes of weather and lighting to some extent.

In the paper by Pandey et al[33], based on background subtraction for detecting moving human body in real-time application method. They establish reliable back ground model, used dynamic threshold method to detect moving object and update the background in real time, they combined contour projection analysis with shape analysis to remove the shadow effect [70], the proposed algorithm overcomes the shortcomings and deficiencies of the traditional method of object detection and this algorithm is fast and simple, able to detect moving human body better and it has a broad applicability.
ii] Temporal Differencing or Frame Differencing

The approach of temporal differencing [35],[36],[37],[38],[39],[40] makes use of pixel-wise difference between two or three consecutive frames in an image sequence to extract moving regions. The advantage of temporal differencing is that it is very adaptive to dynamic environments but the disadvantage of this method is that it does a poor job of extracting the entire relevant feature pixels, e.g., possibly generating holes inside moving entities. In order to overcome shortcomings of two frame difference in some cases, three frame differencing can be used [35].

Collins et al. developed a hybrid method that combines three-frame differencing with an adaptive background subtraction model for their VSAM project [35]. The hybrid algorithm of Collins et al. successfully segments moving regions in video without the defects of temporal differencing. The background subtraction in this case was very fast and surprisingly effective for detecting moving objects in image sequences.

iii] Statistical Method

Recently, some statistical methods [70] to extract change regions from the background are inspired by the basic background subtraction methods described above. The statistical approaches use the characteristics of individual pixels or groups of pixels to construct more advanced background models, and the statistics of the backgrounds can be updated dynamically during processing. Each pixel in the current image can be classified into foreground or background by comparing the statistics of the current background model. We will discuss this approach in some detail as it is becoming increasingly popular due to its robustness to noise, shadow, change of lighting conditions.

In the concept given by C.Stauffer and W.E.L. Grimson.[26],[27],[28] here each pixel is modelled separately by a mixture of K Gaussian

$$P(I_i) = \sum_{i=1}^{K} \omega_{i,j} \eta(I_i; \mu_{i,j}, \Sigma_{i,j})$$

Where K=3-5 in [10] and K=4 in [27]. In [26],[28] it is assumed that $\Sigma_{i,j} = \sigma_{i,j}^2 I$

The background is updated, before the foreground is detected, as follows:

1. If $I_i$ matches component i, i.e., $I_i$ is within $\lambda$ standard deviations of $\mu_{i,j}$ (where $\lambda$ is 2 in [26] and 2.5 in [27]), then the ith component is updated as follows:

   $$\omega_{i,j} = \omega_{i,j-1}$$
   $$\mu_{i,j} = (1-\rho)\mu_{i,j-1} + \rho I_i$$
   $$\sigma_{i,j}^2 = (1-\rho)\sigma_{i,j-1}^2 + \rho(I_i - \mu_{i,j})^T(I_i - \mu_{i,j})$$

   Where $\rho = \alpha P_i(I_i;\mu_{i,j-1},\Sigma_{i,j-1})$

2. Components which $I_i$ don’t match are updated by

   $$\omega_{i,j} = (1-\alpha)\omega_{i,j-1}$$
   $$\mu_{i,j} = \mu_{i,j-1}$$
   $$\sigma_{i,j}^2 = \sigma_{i,j-1}^2$$

3. If it does not match any component, then the least likely component is replaced with a new one which has $\mu_{i,j} = I_i, \Sigma_{i,j}$ large, and $\omega_{i,j}$ low.

After the updates, the weights $\omega_{i,j}$ are renormalized.

The foreground is detected as follows. All components in the mixture are sorted into the order of decreasing $\omega_{i,j}^T \Sigma_{i,j}$. So higher importance gets placed on components with the most evidence and lowest variance, which are assumed to be the background.

Let $B = \arg \min_b \left( \frac{\sum_{i=1}^{b} \omega_{i,j}^T \Sigma_{i,j}}{\sum_{i=1}^{K} \omega_{i,j}^T} > T \right)$

for some threshold T. Then components B are assumed to be background. So if $I_i$ does not match one of these components, the pixel is marked as foreground. Foreground pixels are then segmented into regions using connected component labelling. Detected regions are represented by their centroid [11]. This algorithm gave the results which were reliable and real-time outdoor tracker which can deal with lighting changes and clutter but it suffered from one main disadvantage that it processing rate was slow.

Pflinder [17] used a simple scheme, where background pixels are modelled by a single value, updated by

$$B_i = (1-\alpha)B_{i-1} + c_i$$

and foreground pixels are explicitly modelled by a mean and covariance, which are updated recursively. It requires an empty scene at start-up.

Haritaoglu et al. [41] built a statistical model by representing each pixel with three values: its minimum and maximum intensity values, and the maximum intensity difference between consecutive frames observed during the training period. The model parameters were updated periodically.

McKenna et al. [42] used an adaptive background model combining color and gradient information, in which each pixel’s chromaticity was modelled using means and variances, and its gradient in the x and y directions was modelled using gradient means and magnitude variances. Background subtraction was then performed to cope with shadows and unreliable color cues effectively.

In variable adaptation rate as proposed by Porikli and Tuzel [48] the main goal is to update only when it needed. Though Stauffer and Grimson [28] made necessary arrangements for the maintenance of every frame but in that case no significant pixels changes was needed to be updated at every frame. Porikli et al. propose to adapt the time period of the maintenance mechanism following the illumination change score. The idea is that no maintenance is needed if no
illumination change is detected otherwise quick maintenance is necessary.

Magee [44] used a variable adaptation frame rate following the pixel’s activity, which improves temporal history storage for slow changing pixels while running at high adaptation rates for less stable pixels.

To reduce computation time, Atrey et al. [43] proposed to use an exponential sampling technique. The goal was to apply only the MOG on the region for interest (ROI) and thus reduce the computational efforts in smaller part region. Results showed significant gain in processing speed with a minor loss in accuracy.

Zuo et al. [45] propose a switching based background modeling approach called MSBM. In this approach regions are classified as regions with high or low complexity using an entropy measure. For background regions with high complexity of pixels value distribution, the MOG (Mixture of Gaussian) model is used to guarantee the accuracy of moving object detection. Otherwise a running average is applied to reduce the computational load. Results [45,47] show that this approach possesses almost the same detection accuracy and much higher image processing frame rate than MOG model.

Another approach proposed by Liang et al. [46] use a mean shift algorithm which classifies the background pixels as single mode or multiple mode pixels so that the single mode pixel values are updated with a IIR filter, while the multi-mode pixel values are modelled by the MOG. This approach has a faster speed than the original MOG.

iv) Optical Flow

Optical flow methods make use of the flow vectors of moving objects over time to detect moving regions in an image. [49]. It is generally used to describe coherent motion of points or features between image frames. Motion segmentation based on optical flow [49],[50],[51],[52],[53],[54],[55] uses characteristics of flow vectors of moving objects over time to detect change regions in an image sequence.

Meyer et al. [49] performed monofonic operation which computed the displacement vector field to initialize a contour-based tracking algorithm, called active rays, for the extraction of articulated objects which would be used for gait analysis.

The work by Rowley and Rehg [53] also focused on the segmentation of optical flow fields of articulated objects. Its major contributions were to add kinematic motion constraints to each pixel, and to combine motion segmentation with estimation in expectation maximization (EM) computation. However the addressed motion was restricted to 2-D affine transforms. Also, in Bregler’s work [55], each pixel was represented by its optical flow. These flow vectors were grouped into blobs having coherent motion and characterized by a mixture of multivariate Gaussians.

The main advantage of this method is that it can detect motion in video sequences even from a moving camera. But optical flow method is computationally complex and very sensitive to noise, and cannot be applied to video streams in real-time without specialized hardware.

2.2 Human Body Classification

The image sequences captured by cameras mounted for specific task probably captures images which may include people, vehicles, and other moving objects such as flying birds, flowing clouds, flickering of light etc. To further track people and analyze their activities, it is very necessary to correctly distinguish them from other moving objects is what is defined by human body classification. This is required [22],[35],[55],[57],[58],[59],[60],[61] to extract the region corresponding to people from all moving blobs obtained by the moving object detection as discussed previously. In some cases this step may not be required under some situations where the moving objects are known to be human.

In this section we will describe two categories towards moving object classification they are Shape based and Motion based classification.

i) Shape based classification

The features used in shape-based classification schemes are the bounding rectangle, area, silhouette and gradient of detected object regions. There are different descriptions of shape information of motion regions such as representations of point, box, silhouette and blob are available for classifying moving objects.

The approach presented by Lipton et al. [22] makes use of the objects’ silhouette contour length and area information to classify detected objects into three groups: human, vehicle and other. The method depends on the assumption that humans are, in general, smaller than vehicles and have complex shapes. Dispersedness is used as the classification metric and it is defined in terms of object’s area and contour length (perimeter) as follows:

\[
\text{Dispersedness} = \frac{\text{Perimeter}^2}{\text{Area}}
\]

Classification is performed at each frame and tracking results are used to improve temporal classification consistency.

Object classification by Kuno and Watanabe [56] described silhouette-based shape representation as a reliable method of human detection for visual surveillance systems. They used simple shape parameters of human silhouette patterns to separate humans from other moving objects. The advantage of this method was to use simple shape parameters of human silhouette patterns to classify humans from other moving objects such as birds, vehicles etc and these shape parameters were the mean and the standard deviation of silhouette projection histograms and the aspect ratio of the circumscribing rectangle of moving regions.

Saptharishi et al.[37] propose a classification scheme which uses a logistic linear neural network trained with differential learning to recognize two classes: vehicle and people.

Papageorgiou et al.[62] presented a method that makes use of the Support Vector Machine (SVM) classification trained by wavelet transformed object features (edges) in video images from a sample pedestrian database. This method is used to recognize moving regions that correspond to humans.

The classification method developed by Collins et al. [35] uses view dependent visual features of detected objects to
train a neural network classifier to recognize four classes: human, human group, vehicle and clutter. The inputs to the neural network are the dispersedness, area and aspect ratio of the object region and the camera zoom magnification. Like the previous method, classification is performed at each frame and results are kept in a histogram to improve temporal consistency of classification.

The classification method proposed by Brodsky et al. [63] uses a Radial Basis Function (RBF) classifier which has a similar architecture like a three-layer back-propagation network. The input to the classifier is the normalized gradient image of the detected object regions.

ii) Motion based classification

In classifying non rigid moving objects the meaningful human motion is used as strong indicator for motion classification because of its periodic property. Cutler and Davis [57] describe a similarity-based technique to detect and analyze periodic motion. By tracking an interesting moving object, its self-similarity is computed as it evolves over time. As we know, for periodic motion, its self-similarity measure is also periodic. Therefore time-frequency analysis is applied to detect and characterize the periodic motion, and tracking and classification of moving objects are implemented using periodicity.

Lipton’s proposed [22],[58], residual flow to analyze rigidity and periodicity of moving objects. It is expected that rigid objects present little residual flow, where as a non rigid moving object such as a human being has a higher average residual flow and even display a periodic component. Based on this useful cue, human motion is distinguished from motion of other objects, such as vehicles. Some methods in the literature use only temporal motion features of objects in order to recognize their classes [57],[58],[59]. In general, they are used to distinguish non-rigid objects (e.g. human) from rigid objects (e.g. vehicles). The method proposed in [57] is based on the temporal self-similarity of a moving object. As an object that exhibits periodic motion evolves, its self-similarity measure also shows a periodic motion. The method exploits this clue to categorize moving objects using periodicity as residual flow generated by human motion will have a periodicity. By using this indicator, human motion, can be distinguished from other objects such as vehicles.

Cutler and Davis [57] described a similarity-based technique to detect and analyze periodic motion. By tracking moving object of interest, they computed its self-similarity as it evolved over time. As we know, for periodic motion, its self-similarity measure was also periodic. Therefore they applied time–frequency analysis to detect and characterize the periodic motion and implemented tracking and classification of moving objects using periodicity.

Two common approaches mentioned above, namely shape-based classification and motion-based classification can also be effectively combined for moving object classification [41].

Stauffer [61] proposed a novel method based on time co-occurrence matrix to hierarchically classify both objects and behaviours. It is expected that more precise classification results can be obtained by using extra features such as color and velocity. In a word, finding people [49],[64],[65] in images is a particularly difficult object recognition problem. Generally, human detection follows the processing described above. However several of the latest papers provide an improved version in which the combination of component-based or segment-based method and geometric configuration constraints of human body is used [66],[67],[68].

In the work proposed by Mohan et al [66], the system was structured with four distinct example-based detectors that were trained to separately find four components of the human body: the head, legs, left arm, and right arm. After ensuring that these components were present in the proper geometric configuration, a second example-based classifier was used to classify a pattern as either a person or a non-person. Although this method was relatively complex, it might provide more robust results than full-body person detection methods in that it was capable of locating partially occluded views of people and people whose body parts had little contrast with the background.

Conclusion

Over the last two decades, a large amount of research has been conducted on human motion detection and tracking. This review paper gives valuable insight into this important research topic and encourages new researchers in the area of moving object detection. Human motion analysis is a challenging problem due to large variations in human motion and appearance, camera viewpoint and environment settings.

We have extensively studied human detection involving motion segmentation and object classification. Four types of techniques for motion segmentation are addressed in the paper: background subtraction, temporal differenting or frame differencing, statistical methods and optical flow. The statistical methods may be a better choice in outdoor or where illumination variation occurs because of which it is described in some detail here in the paper. To segregate human from other moving objects shape based and motion based classification is studied under detection.

We believe that no perfect system exists one has to find the most suitable algorithm depending upon kind of object to be detected , place of tracking , illumination variation, motion etc. then one can get best results for the purpose of detection and tracking of human in real time.

References


[14] Moeslund T.B. ; Computer Vision Based Human Motion Capture: A Survey


[21] Robers Bodor, Bennett Jackson and Nikolaos Papanikolopoulos; Vision Based Human Tracking and Activity Recognition, AIRVL Department of Computer Science & Engineering, University of Minnesota.


ISSN No: 2250-3536 Volume 3, Issue 3, May 2013


Biographies

Lovendra Solanki has received his M.E degree from M.B.M Engineering College, Jodhpur. He has 20 years of teaching experience in the field of Electronics Engineering. He is the author of 3 books on Digital Electronics. Currently he is Associate Professor at B K Birla Institute Of Engineering & Technology, Pilani and pursuing Ph.D from Singhania University, Junjhuunu. He may be reached at dean.ece@bkbit.ac.in
Dr. P. S. Bhatnagar is currently the Director of B K Birla Institute of Engineering & Technology, Pilani. Previously, he was Scientist at CEERI, Pilani, Rajasthan, for 24 years. He has published 106 technical papers and is the author of 3 books on Electronics and Antenna. He has 2 patents on antenna. He is recipient of more than twelve technical awards and is a member of many technical societies. Dr. P. S. Bhatnagar may be reached at director@bkbit.ac.in

Dr. J. L. Raheja has received his M.Tech from IIT Kharagpur and PhD degree from Technical University Munich, Germany. At present, he is Senior Principle Scientist, Digital Systems Group, in Council of Scientific and Industrial Research- Central Electronics Engineering Research Institute (CSIR-CEERI), Pilani, Rajasthan. He has published 48 papers in international journals. His area of interest is Cartographic Generalisation, Digital Image Processing and Human Computer Interface. He is frequently visiting Germany and UK as guest scientist. Dr. J. L. Raheja may be reached at jagdish@ceeri.ernet.in