ISSN NO. 2320-5407

Journal homepage: http://www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH

RESEARCH ARTICLE

A SURVEY OF VARIOUS TECHNIQUES FOR HUMAN MOTION DETECTION FROM VIDEO

R.SHANMUGA PRIYA[1], S.KARTHICK[2], C.THULASIYAMMAL[2]

1. Doing II year ME(Embedded system) in The Kavery Engineering College, Mecheri.

2. Assistant Professor of Electrical and Electronic Engineering,The Kavery Engineering College, Mecheri.

Manuscript Info

Manuscript History:

Received: 12 February 2014
Final Accepted: 14 March 2014
Published Online: April 2014

Key words:

Corresponding Author

R.SHANMUGA PRIYA

Abstract

This paper presents a survey of various techniques related to human motion detection system .The system captures video and identifies those moving objects in the video which are characteristically human or non-human. In order to function, they require some methods for detecting people from a given input video. In this paper, we discuss a sample of techniques for finding people using optical input. These techniques are classified with respect to the human identification using background subtraction and direct detection methods used to describe human appearance (shape, color, motion). The results with the system indicate the ability to minimize the false detections and missed detections.

.

1. Introduction:

The field of computer vision is alarmed with problems that involve interfacing computers with their surrounding environment through visual resources. One such problem is object recognition. It involves detecting the presence of a known entity in an image or a video sequence, given some idea about what that entity should look like. As humans, we take this capability for established, as our brains are extraordinarily talented at both learning new objects and identifying them shortly. In order of computer vision, this same crisis has proven to be one of the most difficult in this field.

Human activity detection has applications in various fields, the most important is surveillance system. Before the complexity of human motion can be understood, we primarily need some automatic systems for finding humans in an image or a video sequence. Once the human action is detected, depending on the application, the system can do further process. This paper selects a representative sample of papers from the broad literature on human body detection and presents another study of the various techniques

2. Techniques Overview:

Our primary aim is to find human in a given video or an image. The related literature can be divided into techniques which require background Subtraction, motion segmentation and other techniques which can detect humans directly from the input.

Background subtraction techniques usually split the foreground and background objects. Subsequently takes the foreground object from the video and then classify it into like human, non-human, animal etc., based on shape, colour, activity or other descriptions.

Direct techniques operate on features extracted from image or video patches and classify them as human or non-human. We can also categorize techniques based on the descriptions which are used to identify a given input as human or non-human. These descriptions include shape, colour (skin colour detection), motion, or combinations of these.

3. Techniques using background subtraction

3.1 Beleznai et al. [2004]

This thesis treats the intensity difference between an input frame and a reference image as a multi-modal probability distribution, and the mode detection is performed by using mean shift computation. The mean shift computation is performed in a fast way using integral images or summed area tables, which gives the method real-time performance in a manner which is independent of the size of the window used. The mode detection procedure is able to locate secluded humans, but for separating partially occluded and grouped humans, a model-based validation process is used. The human body model is very simple and consists of three rectangular regions. Within each cluster of humans, a maximum similar configuration of humans is identified.

3.2 Haga et al. [2004]

In this thesis, a moving object is classified as human based on the spatial uniqueness of the image movement (called criterion A1 by the authors), temporal uniqueness of the human motion (A2), and the temporal motion continuity (A3). First, the moving object is detected by background subtraction, and then A1 A2, and A3 are evaluated .The spatial uniqueness of image activity is a measure of uniformity of local motion within a region. Temporal uniqueness is correspondingly defined in the time direction. A linear classifier human and non-human data in the A1-A2-A3 space, and is used to classify new input data.

3.3 Lee et al. [2004]

In this paper, a shape-based approach for categorization of the objects is using the following background subtraction based on the frame differencing. The aim is to detect the humans for hazard assessment. The objective is to classify the intruder as human or animal or non-human like vehicle based on the shape of its boundary contour. The system classifies the contour of the entity into different categories using the following procedure. The data points on the contour are reduced by a curve evolution Method which uses a relevance measure to remove vertices from the contour. By this technique, the contour is reduced to 60 data points, which basically amounts to a polygon estimation expressed as bend angle vs. normalized length. The likeness between contours is measured using the L2 norm. For this, a new quick matching algorithm is developed, which can be used to categorize the object as human or animal or vehicle.

3.4 Zhou and Hoang [2005]

This thesis presents a technique to detect and track a human body in a video or image. At first, the background subtraction is performed to detect the foreground object, which involves temporal differencing of the consecutive frames. After this step, the classification of the object is based on two approaches: the first approach is a codebook, and the second approach involves tracking of the object and if the object can be tracked effectively, it is considered to be a human. For the first approach, the size of foreground blob is normalized to 20x20, and then a shape feature vector of the foreground object is created. However, to create the shape vector of object, the mask figure and boundary of human are created. The distance from the boundary of human body to the left side of bounding box is used as feature vector. This is compared against the feature vectors of the human images in the codebook. The minimum of all distortions for the all the features vectors in the codebook is found, and that if that is below the threshold, then the object is classify as human. Tracking is based on colour histograms, motion and size of the foreground blob. Fake alarms due to static oscillatory motions are also identified and removed, to handle objects like shaking trees in air. Other features of the technique include shadow removal.

3.5 Jiang et al. [2004]

This paper is based on fusion of infrared (IR) images with Images from a normal camera. Humans display a feature signature in IR images due to their skin temperature, but these images usually have low contrast. They can be fused with images from a model camera to obtain superior detection output. The proposed technique first computes pixel saliencies in the two images (IR and visible) at numerous scales, and fusion is performed based on relative saliencies in the two images (called the perceptual contrast difference in the paper).

4 Direct detection

4.1 Cutler and Davis [2000]

In this paper it focus on detecting periodic motions and is applicable to the detection of typical periodic biological motion patterns such as walking, running etc. The video from a moving camera is first stabilized then frame differencing and thresholding is performed to detect independently moving regions. The operations are then used to attain a set of tracked objects. Every segmented object is aligned along the time axis (to remove transformation and its size is also made constant across time. The object's temporal self-likeliness matrix is computed via similarity measures (such as Correlation) which are periodic for periodic motions. Time-frequency study based on the short-time Fourier transform (STFT) is applied and autocorrelation is used for robust periodicity detection and analysis. A lattice fitting method is used to classify human or non-human, animal and vehicle, and the experiments demonstrate that the technique can distinguish the motion of a human from a cat. Not only is the system capable of detecting periodic human activity but it also has understanding of the period which is useful for extracting more information about gait such as stride length.

4.2 Utsumi and Tetsutani [2002]

This paper uses the aspect that the relative positions (geometric distances) of various body parts are common to all people; even though the pixel values may get variations because of the clothes and the illumination. This method uses a structure known as the distance map which is built by taking an image of a individual human and dividing it into MxN blocks. A distance matrix of size MN x MN is then computed in which every element expresses the distance between colour distributions present in a pair of blocks. Then, using those distance maps for a large database of human and the non-human images, a statistical model is built for distance maps of every type, which consists of the average and covariance matrix for every block. The two distributions are compared using the Mahalanobis distance and are found to be very similar excluding for some elements. These few elements show a data projection matrix which is the important model used for detection. Given a new input image, image patches at multiple locations and scales are compared to the model and a threshold is used to classify a patch as human or non-human.

4.3 Dalal and Triggs [2005]

The best part of this paper is that it uses a histogram of gradients as the feature space for building a classifier. It utilizes the fact that the outline of an object can be well represented by a distribution of local intensity gradients or edge directions. This is done by isolating the image in small spatial parts as cells and finding the histograms of edge orientations over all the pixels of the cell. The joined histogram entries form the characteristic representation after local contrast normalization in overlapping descriptor blocks. The authors experiment with several orientation and spatial interval resolutions and normalization schemes to obtain the maximum performance. For further categorization, a Dataset of human and non-human examples is created, and a linear classifier is trained using Support Vector Machine (SVM) on the gradient histogram features from the two classes. This classifier can then be applied to a new input image at various scales for detecting humans.

5 Summaries

We have discussed several techniques in the recent literature for human motion detection from video. We have organized them according to techniques which use background subtraction and those which operate directly on the input. In the first category, we have ordered the techniques based on type of background subtraction used and the model used to represent a human motion. In the second category, we have ordered the method based on the human model and the classifier model used. Overall, there seems to be an increasing trend in the recent literature towards robust methods which operate directly on the image rather than those which require background subtraction as a first.

Bibliography

1 Neeti A. Ogale, "A Survey of Techniques for Human Detection from Video,", unpublished.

2 C. Beleznai, B. Fruhstuck, and H. Bischof. Human detection in groups using a fast mean shift procedure.International Conference on Image Processing, 1:349–352, 2004.

3 R. Cutler and L. S. Davis. Robust real-time periodic motion detection, analysis and applications. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(8):781–796, 2000.

4 Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 1063–6919, 2005.

5 H. Eng, J. Wang, A. Kam, and W. Yau. A bayesian framework for robust human detection and occlusion handling using a human shape model. International Conference on Pattern Recognition, 2004.

6 D. M. Gavrila and J. Giebel. Shape-based pedestrian detection and tracking. IEEE Intelligent Vehicle Symposium, 1:8–14, 2002.

7 T. Haga, K. Sumi, and Y. Yagi. Human detection in outdoor scene using spatio-temporal motion analysis. International Conference on Pattern Recognition.

8 Ju Han and B. Bhanu. Detecting moving humans using color and infrared video. IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, 30:228–233, 2003.

9 Lijun Jiang, Feng Tian, Lim Ee Shen, Shiqian Wu, Susu Yao, Zhongkang Lu, and Lijun Xu. Perceptual-based fusion of ir and visual images for human detection. International Symposium on Intelligent Multimedia, Video and Speech Processing, pages 514.

R.Shanmuga priya, Doing II year ME(Embedded system) in The Kavery Engineering College, Mecheri.
S.Karthick: Assistant Professor of Electrical and Electronic Engineering,The Kavery Engineering College, Mecheri.
C.Thulasiyammal: Assistant Professor of Electrical and Electronic Engineering,The Kavery Engineering College, Mecheri.