

Human Motion Analysis with the Help of Video Surveillance: A Review

Kavita V. Bhaltilak, Harleen Kaur, Cherry Khosla

*Department of Computer Science and Engineering,
Lovely Professional University, Phagwara, Punjab-144401*

Abstract - Recent research in computer has persuaded more studies in human motion observation as well as analysis. Visual analysis of human motion is more active component in computer vision. With the help of video surveillance we are able to work in this area conveniently. Human motion analysis concerns the detection, tracking, recognition of human activities and also human behaviours. This paper helps to comprehend multiple techniques of human motion detection and behaviour understanding.

Index Terms - Human motion detection, Tracking, Human behaviour understanding, Background Subtraction.

1. INTRODUCTION

The analysis of human actions by a computer is gaining more and more interest. A significant part of this task is to register the motion, a process known as human motion capture. Human motion capture as the process of capturing the large scale body movements of a subject at some point [1]. There are various types of human activities as gestures, actions, interactions, and group activities. Fig1. Shows the basic levels of human motion analysis. Basic model of human motion analysis comprised of three tier system initiating with human detection as a lower level followed by human body tracking, and behavior understanding as a higher level [18].

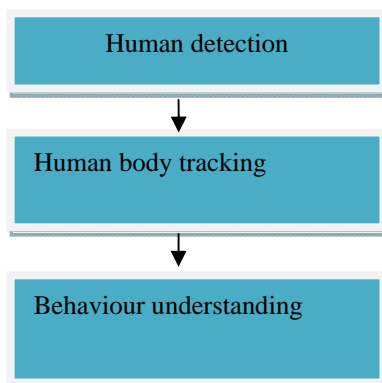


Fig1. Basic model of human motion analysis [18]

1.1 Applications of Human Motion Analysis

Human motion detection is considered to be a key requirement in many applications such as security purpose, motion detection, psychology study, image processing, etc. Therefore, reliable human motion detection is required for the success of these applications.

Human motion analysis has been investigated under several research projects worldwide. Defence Advanced Research Projects Agency (DARPA) project on Video Surveillance

and Monitoring (VSAM) [3], whose purpose was to develop an automatic video understanding technology that enabled a single human operator to monitor activities over complex areas such as battlefields and civilian scenes. The real-time visual surveillance system W4 [4] employed a combination of shape analysis and tracking, and constructed the models of people's appearances to make itself capable of detecting and tracking multiple people as well as monitoring their activities even in the presence of occlusions in an outdoor environment. Researchers in the UK have also done much research on the tracking of vehicles and people and the recognition of their interactions [5]. In addition, companies such as IBM and Microsoft are also investing on research on human motion analysis [6] [7].

1.1.1 Video Surveillance System

Human motion analysis is applicable in many areas such as banks, department stores, parking lots, sports stadiums, office buildings, courts, police stations, museums, shopping malls and borders [8]. Surveillance cameras are already prevalent in commercial establishments, while camera outputs are usually recorded in tapes or stored in video archives. The real-time analysis of surveillance data is needed to alert security officers to a burglary in progress. Video information is integrated with other situation awareness management software (SAMS). For example, when an alert occurs, different sets of software determine what happened. Say an access door alarm goes off that information will be sent to SAMS and an operations centre. The information from the appropriate cameras will be retrieved to understand that incident here is what happened and how. If appropriate, information can be sent to the police or similar agencies.

1.1.2 Human Computer Interaction (HCI)

There has been growing interest in the development of new approaches and technologies for bridging the human computer barrier. The use of hand gestures as a natural interface serves as a motivating force for research, its representations and recognition techniques [9]. Use of hand gestures provides an attractive interface device for human computer interaction. Using hands as a device can help people communicate with computers in a more intuitive way. "Human computer interaction can be defined as the discipline concerned with the design, evaluation, and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them"[25]. Now-a-days it is an important

application domain in advanced user interface in which human motion analysis is usually used to provide control and command. The user activity has three different levels: physical [22], cognitive [23], and affective [24]. The physical aspect determines the mechanics of interaction between human and computer while the cognitive aspect deals with ways that users can understand the system. The affective aspect deals with user's emotions and other stimuli.

1.2 Purpose of the Survey

The main purpose of this survey is to provide a complete survey of the most recent developments in this area, so that the new people who are interested in this field to gain knowledge. The importance and popularity of human motion analysis has led to several previous surveys. Highlights progresses made in the development of detection of human body for vision-based gestural interfaces as well as challenges that still remain. This paper selects a taxonomy based on functionalities including detection, tracking and behaviour understanding within human motion analysis systems. It aims to provide concrete knowledge to discuss different techniques for human detection as well as human behaviour understanding.

2. HUMAN BODY MOTION DETECTION

Almost every system of vision-based human motion analysis starts with human body detection. Fig 2. Show the procedure for detecting moving object, this are basic steps, Collect video frames first then detect moving objects such as human, car, birds etc. Next step is to identify human body parts from rest of moving objects. And the last step is to detect human body. The subsequent processes such as tracking and action recognition are greatly dependent on it. This paper present selective sample of papers on human body detection and its various methods[19]. Human body detection can be well illustrate by several techniques as presented in the following section.

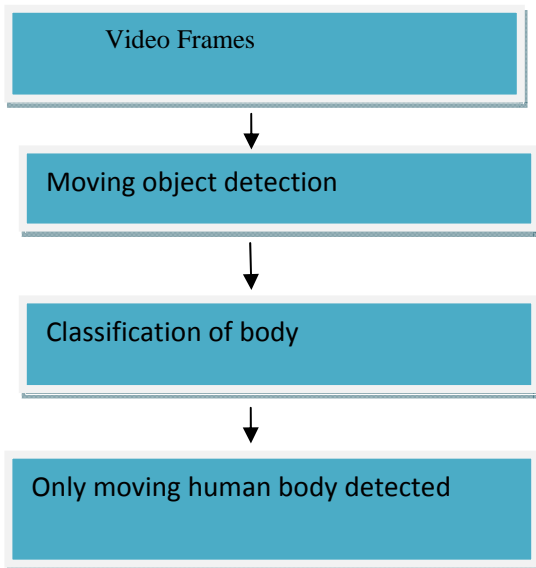


Fig2. Human body detection process [18]

2.1 Background Subtraction Method

Detection of moving human in videos from static camera is widely performed by background subtraction method. The origin of this approach is that of detecting the moving objects from the difference between the existing frame and a reference frame, frequently called the “background copy”, or “background replica” [12]. The background image must be a representation of the scene with no moving objects and must be kept regularly updated so as to adapt to the varying luminance conditions and geometry settings. More difficult models have extended the concept of “background subtraction” beyond its literal meaning. Fig3 (a)(b).The example of background subtraction. The key of this method lies in the initialization and update of the background image. The effectiveness of both will affect the accuracy of test results [16]. Background subtraction is most in this method moving region is detected by differencing the current and reference background frame in pixel by pixel manner [17].

The process of algorithm is described as follow:

- Sequences of Video Frames
- Frame Separation
- Image Sequence
- Separation of Image Sequence in current Frame
- Image and Background Frame Image
- Perform Background Subtraction
- Detection of Moving Object
- Perform Background Updating
- Noise Removal
- Shape Analysis



Fig3(a). Example of background subtraction



Fig3(b). Example of background subtraction

2.2 Statistical Method

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. Some statistical methods [2] to extract change regions from the background are inspired by the basic background subtraction methods. The statistics of the backgrounds can be updated dynamically during processing.

2.3 Temporal Differencing

The approach of temporal differencing 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 on other hand it does a poor job of extracting the entire relevant feature pixels [3]. Within the context of hand gesture recognition, spatiotemporal gesture segmentation is the task of determining, in a video sequence, where the gesturing hand is located, and when the gesture starts and ends. Existing gesture recognition methods typically assume either known spatial segmentation or known temporal segmentation, or both. This paper introduces a united framework for simultaneously performing spatial segmentation, temporal segmentation and recognition. A gesture can be recognized even when the hand location is highly ambiguous with scanty information about the gesture. Thus, the method can be applied to continuous image streams where gestures are performed in front of moving cluttered backgrounds.

3. HUMAN BODY TRACKING

Object tracking in video streams has been a popular topic in the computer streamline. Tracking is a particularly important issue in human motion analysis since it serves as a means to prepare data for pose estimation and action recognition. In contrast to human detection, human tracking belongs to a higher-level computer vision problem [15]. However, the tracking algorithms within human motion analysis usually have considerable intersection with motion segmentation during processing. Tracking over time typically involves matching objects in consecutive frames using features such as points, lines or blobs [13]. Tracking may be considered to be equivalent for establishing coherent relations of image features between frames with respect to position, velocity, shape, texture, color, etc. There are many human motion extraction and detection methods reported by researchers of face recognition. To combine image information and knowledge of the face is the main trend. It's very difficult in reality to track a human when walking. We use Gait Energy Image (GEI) to characterized human walking properties.

Human body motion tracking can be evaluated either by facial features extraction or by motion of body parts of human .The detail discussion of these tracking methodologies is presented in following segments .

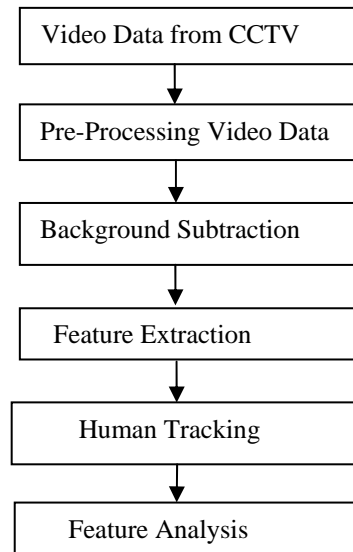


Fig4. General frame work of intelligent video surveillance system (VISS)[15]

3.1 Human Recognition by Integrating Face and Gait

Human recognition is very difficult task in reality, especially when a person doing some motion. This paper proposed an innovative system, which combine cues of face profile and gait silhouette from single camera video sequences [14]. For optimal face profile recognition need to collect low resolutions video frames to construct high resolution face profile. A fusion system, which combine face and gait cues from low resolution video sequences, is a practical approach to accomplish the task of human recognition.

3.2 Human Body Points Tracking

For tracking a human body, important body features are extracted based on the anatomical knowledge. These essential human body points are called as motion data. In biomechanics, the human body information is used to study and observe human movement [15]. The motion data can be used to identify the posture, pose, action and movement of an actor. Therefore, understanding the motion of body part is essentially needed for person identification. Fig4. Represent Intelligent Video Surveillance System (IVSS) is the most recent smart surveillance system which is capable to identify automated human activity and behaviour from the video. Motion detection of human has a very wide area of applications. Thomas and Erik [1] classified them into three major areas which are surveillance, control and analysis. Milan, Vaclav and Roger [10] separated three types of motion-related problems into motion detection which is usually used for security purposes, motion object detection and location. To obtain the object that causes the motion and using an recognition engine instead of just image-matching as suggested in Milan, Vaclav and Roger classify the object to check whether if it is a human or not. Therefore, this project is not fixed to only the first application area [1]. It can also be applied to applications in control areas or analysis areas where human motion is to be classified to distinguish them from other objects that cause motion. However, in this paper, we would mainly focus on the application of the algorithm in surveillance systems.

We have incorporated some of the methods proposed in this article. Several rules have been formulated for the task of locating the contour of the human body. The background of the targeted area is assumed to be non-moving and considerations of sudden change in lightings are ignored as well. However, the considerations of other factors are taken into consideration. Basically, the initial plan was to use a technique called image segmentation to abstract the foreground image from the source image obtained and later processed to filter out noises or small images disturbance. They focus on every single human figure which comes in our CCTV footage.

4. BEHAVIOUR UNDERSTANDING

Behaviour understanding involves action recognition and description. Human behaviour understanding can guide the development of many human motion analysis systems. It is an important area of future research in human motion analysis. It is obvious that the basic problem of human behaviour understanding is how to learn the reference action sequences from training samples, how to enable both training and matching methods.

Behaviour understandings mainly done by three methods are as follows.

4.1 Dynamic Time Warping

Dynamic time warping (DTW), used widely for speech recognition in the early days, is a template-based dynamic programming matching technique. It has the advantage of conceptual simplicity and robust performance, and has been used in the matching of human movement patterns recently [20].

4.2 Hidden Markov Models

The use of HMMs touches on two stages: training and classification. In the training stage, the number of states of an HMM must be specified, and the corresponding state transformation and output probabilities are optimized in order that the generated symbols can correspond to the observed image [21].

4.3 Psychology Study

Cognitive behavioural approaches with this reference we are trying to study the behaviour of the criminals related to bank robberies. Also studied the term behaviourism which was classified into

I) Subjectivism

II) Mentalism

Human motions which merely reflected person own mode of thinking. The concept of self efficacy comes under this topic, which helps to identify the self controlling responses towards their environments. Human’s activity has three modalities represent in Fig5. These three modalities are inseparable. If any side of the figure is removed, it ceases to be a triangle and no longer exists.

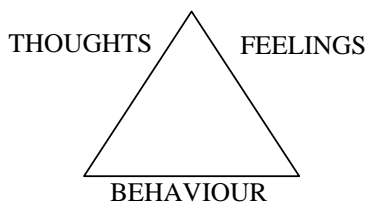


Fig5. Inter-dependence of thoughts, feelings and behaviour [11]

Each of the three modalities forming the triangle can be described as having three dimensions as follows:

Intensity: its experienced strength.

Frequency: how often a type of event occurs.

Duration: the time lapsed since its first occurrence [11].

Thoughts, feelings and behaviour these three things are interdependent to each other (Fig 5). If thought comes first, subsequently a rest component occurs suddenly and vice - versa.

Objective of “looking at people” is to analyze and interpret human action and the interactions between people and other objects; better understanding of human behaviours is the most interesting analysis. For instance, the W4 system [2] can recognize some simple events between people and objects such as carrying an object, depositing an object, and exchanging bags. Human motion understanding still emphasizing on tracking and recognition of some standard posture and simple action analysis e.g., the definition and classification of a group of typical actions running, standing, jumping, climbing, pointing, etc [5].

5. CONCLUSION

Human motion analysis is currently one of the most active topics in computer science. Human motion analysis concerns mainly detection, tracking and recognition of people behaviour. Bearing in mind a general processing framework of human motion analysis systems, we have presented an overview of recent developments in human motion analysis. Human detection involves motion segmentation and object classification. Four types of techniques for motion segmentation are addressed: background subtraction, statistical methods, temporal differencing and optical flow the statistical methods may be a better choice in more unconstrained situations. This survey include two different methods of behaviour understanding DTW and HMM.

REFERENCES

[1] Thomas B. Moeslund and Erik Granum Laboratory of Computer Vision and Media Technology, Aalborg University, Aalborg, Denmark,2000

[2] Lovendra Solanki, Research Scholar, Singhania University, Jhunjhunu (Rajasthan) INDIA 2013

[3] R.T. Collins, et al., A system for video surveillance and monitoring: VSAM Anal report,CMU-RI-TR-00-12, Technical Report, Carnegie Mellon University, 2000.

[4] I. Haritaoglu, D. Harwood, L.S. Davis, W4: real-time surveillance of people and their activities, IEEE Trans. Pattern Anal. Mach. Intell. 22 (8) (2000)

[5] P. Remagnino, T. Tan, K. Baker, Multi-agent visual surveillance of dynamic scenes. Image Vision Comput. 16 (8) (1998) 529–532.

[6] C. Maggioni, B. Kammerer, Gesture computer: history, design, and applications, in: R. Cipolla, A. Purtiland (Eds), Computer Vision for Human–Machine Interaction, Cambridge University Press, Cambridge, MA, 1998.

[7] W. Freeman, C. Weissman, Television control by hand gestures, Proceedings of the International Conference on Automatic Face and Gesture Recognition, 1995, pp. 179–183. 09–830.

[8] M. Valera and S.A. Velastin, School of Computing & Information Systems, Kingston University, UK, IEE, 2005

[9] Prashan Premaratne et al, "Hand grsture teacking and recognition of system" 2012

[10] Milan Sonka et al, "Image Processing, Analysis, and Machine Vision", PWS Publishing, 1999.

[11] James McGuire, University of Liverpool, Dept of Clinical Psychology, United Kingdom

- [12] Neeti A. Ogale, Department of Computer Science University of Maryland, College Park, MD
- [13] Prithviraj Banerjee and Somnath Sengupta, Department of electronic and ECE, Indian Institute of Technology, Kharagpur, India
- [14] Xiaoli Zhou, Bir Bhanu and Ju han, Center for Research in Intelligent System, University of California, USA, 2005
- [15] Win Kong et al Department of Electrical, University Kebangsaan, Proceedings of the World Congress on Engineering and Computer Science 2013 Vol I WCECS 2013, 23-25 October, 2013,
- [16] Rupali S. Rakibe, Bharati D, M.E (Electronics and Telecommunication), GHRCEM Wagholi, Pune. 2013
- [17] L. Zhao, C. Thorpe, Recursive context reasoning for human detection and parts identification, Proceedings of the IEEE Workshop on Human Modeling, Analysis and Synthesis, June 2000
- [18] Jagdish Lal Raheja et al, Electronic Computer Technology (ICECT), 2011 3rd International Conference , volume 2, page 199-203, 2011
- [19] Liang Wang et al, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing , 2003
- [20] C. Myers, L. Rabinier, A. Rosenberg, Performance tradeoffs in dynamic time warping algorithms for isolated word recognition, IEEE Trans. Acoust. Speech Signal Process. 28 (6) (1980) 623-635.
- [21] A.B. Poritz, Hidden Markov models: a guided tour, Proceedings of the International Conference on Acoustic Speech and Signal Processing, 1988, pp. 7-13.
- [22] A. Chapanis, Man Machine Engineering, Wadsworth, Belmont (1965).
- [23] D. Norman, "Cognitive Engineering", in D. Norman and S. Draper (eds), User Centered Design: New Perspective on Human-Computer Interaction, Lawrence Erlbaum, Hillsdale (1986).
- [24] R.W. Picard, Affective Computing, MIT Press, Cambridge (1997).
- [25] Alan Dix, Human-computer interaction, 2nd. University of Michigan, Europe, 1998