MOTION TRACKING TECHNIQUES: A REVIEW

Subreena Gulzar[1] Dr. Rakesh Joon[2]

[1]M.Tech Student, ECE Department, Ganga Institute Of Technology And Management, Kablana, Jhajjar, Haryana,India

[2]HOD, ECE Department, Ganga Institute Of Technology And Management, Kablana, Jhajjar, Haryana, India

***Abstract* -- Individual detection approach based on the video and visual images in a major concern for monitoring a location. Tracking application majorly uses the video analytics as they can be able to analyze and track each individual for identification. This problem of feature selection, object representation, dynamic shape, and motion estimation are the active areas of research and the new solutions have been proposed. Behavioral detection plays a vital role in several recognition and identification process. This paper deals with the various techniques used for surveillance and the respective pitfalls in the current surveillance system.**

Keywords—object detection, body detection, motion tracking, gesture recognition, machine learning

I. INTRODUCTION

Numerous researches have been done in a detection system for processing, manipulation and interpretation of visual images. Video analytics plays an important role in a wide variety of defense-activity, monitoring and surveillance-related systems for air and ground environments. Many robust tracking systems which have been developed in the recent years have a capability to detect the object in real time scenarios. Body motion tracking and its various associated problems such as feature extraction, analyzing object shape and tracking object motion were the active areas of research.[1] The researchers are continuously proposing new algorithms to enhance the system capability for tracking the object in motion and extracting the features to analyze the gestures. The terrorist attack in France on 23rd March 2018 and the 2017 Barcelona attack where the group was held with 120 gas canisters being ready to bomb a major target have revealed that even with such high-security systems there's still a loophole that needs to be mended. Prevention of terrorist attack and the detection of such threats is a major challenge to ensure public safety. In order to improve the security and ensure the safety of people at places with high commotion, smart surveillance cameras and security systems need to be deployed. The cameras need to perform the following operations: a) detect a human body and b) decode and analyze the suspicious gesture of an individual and. Detection of the human body or any object in itself is a cumbersome task. A robust object detection requires to detect in all kind of lighting conditions, at any angle or rotation plane and even if the image blend with the background. [2] After the detection of the human body, to study the suspicious behavior of human, gestures of the particular individual need to be examined. This motion of body can be a gesture, an action or a form of interaction with an object or another human. [3] Gestures are the motion of the body which is used to study the intention of a person. [4] A person may show a gesture to communicate, to respond to someone or to perform any malicious activity. Based on this the gestures are categorized as dynamic and static. [5] A dynamic gesture changes over a certain period of time while the static gesture is observed at the squirt of time. These gestures highlights suspicious behavior of an individual who might pose threat to the security. The surveillance systems or any video frame based system aims to improve the accuracy in the perception about the features extracted to analyze the behavior and prediction of the behavior of a suspect based on the features extracted. Many researches has been done in this field and these researches pertain to the installation of high-security surveillance systems at airports, banks and railway stations. [6] These systems generate an alarm to warn the security agencies working in this field. Sometimes there are many false alarms with the security systems which causes a lot of havoc. This requires feature extraction and classification techniques which detect any kind of threat to the security and reducing the rate of any false alarms. The system for detection works on input video frames, apply various pre-processing techniques on the video frames and then the feature is extracted from the video frames. This gives the output images with detected suspicious body motion as illustrated in Figure 1.

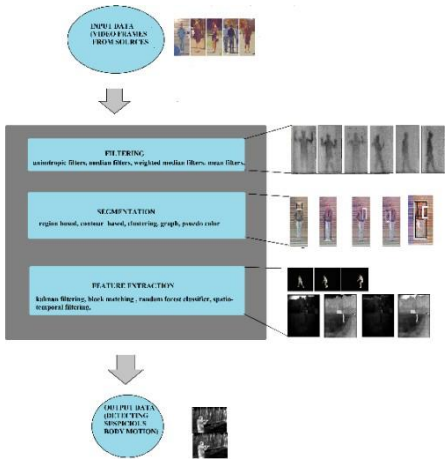

Fig.1 : The Working Model

II. Literature Review

Various researchers have studied the gestures in body motion in order to detect the suspicious behavior of any suspect. It starts with already applied techniques and recent trends in the detection of the suspect with an abnormal behavior. The review is then directed to various approaches for the studying the abnormal human behavior detection based on suspicious gestures exhibited by the suspect. The detection of a suspect in motion requires analysis of the video frames for detection of the object, tracking the body motion and studying any kind of suspicious gestures.

A. Detection and Tracking Of Object

The author [2] proposes that object detection deals with detection of an object by detecting the constituent components and then combining the components using a classifier. The method deals with processing of the model to distinguish an object class such that the object class has a high value for inter-class variations and a low value for intra class variations. In order to detect the human body, various properties need to be studied. The properties include the shape of the human body, positon of the body limbs, dresses worn by the human body, etc. Various approaches are used to detect the human body which includes model-based approach, example-based approach and component-based approach. The system employs the component based people detection. Various components of a human body are compared and the image frames with a high matching have a high candidate score. The highest component score for each component is given as input into the support vector machine. The final output determines if the given pattern is a person. The results exhibit that how the change in appearance affects the view point and this change affects geometric properties.

The various other techniques [5] were described to enhance the efficiency and effectiveness of object detection and tracking. It was used in various surveillance system applications such as Traffic Monitoring, understanding human activity, observation of people and vehicles within a busy environment, security in shopping malls or offices, etc. The researcher gives a review of various techniques for object detection and tracking. Background subtraction can be done using support vector machine and adaptive boosting. Frame difference is the most effective technique for object tracking. Brightness constraint is also used for computing the optical flow for detecting motion based object segmentation and tracking. Kalman filters, Contour tracking, Multi-object Data Association & State Estimation, Multiple Hypothesis Tracking and Mean Shift Method are other techniques for tracking the object in video frames.

Another approach for the object detection by [7] presented a technique for the detection and tracking of the boundary of an object in Wireless sensor networks. The scheme proposed is a collaborative one that focused on the detection of a node whether the node is failure prone or not and then tracks the respective node for the motion in the network. According to the algorithm proposed, in order to conserve the restricted video sources and frames available, the discernible change in the shape can be neglected This helps to detect and monitor the boundary of the object in a controlled environment. The scheme restricts the conditions for analyzing the nodes based on one or two hop neighbor involved in the tracking of the object motion by assigning weights to each tracking node and this helps to achieve optimization in tracking the body motion. Another scheme was proposed to discard the failure nodes for recovery which significantly reduces the report data size. Hence, the proposed approach detects failure nodes to eliminate inconsistencies in boundary data and improves the efficiency of the system.

B. Tracking of Body Motion

The researcher [8] aims at detecting and understanding the human activity. The technique involves two-level analysis of human behavior. The first level focusses on the extraction of the region of interest by detecting the human body motion and the second level detects the temporal motion patterns. During the process, features are extracted in three levels: individual object features (position, velocity, shape, and color), object features corresponding to the environment (relative position, velocity) and the relationship between the two. The technique used is template matching where the normal template of the video frame is compared to abnormal activity frame and analyzed for the tracking of body motion. The technique employs the advanced study of the Motion-history image (MHI) and Motion Energy Image (MEI) using background subtraction and binarization. It studies the various generations of surveillance systems for human activity recognition. Then the comparative analysis of various technique is done and performance is evaluated for various techniques.

The researcher [9] described that significant progress has been made in object detection and tracking during the last few years. In order to develop a real time tracking system, it should be capable enough to track the object in unconstrained environment. Initially the research focused to detect the object in video frames and then object tracking became another issue of concern

The problems with tracking are extraction of the features, representation of the body in motion, study of the dynamic shape of the object and estimate the motion of the body. [10] These problem are exploited and enhanced algorithms are continuously being proposed.

One of the optimized technique for tracking the body in motion is to discriminate between the classes of multiple object in motion and between object and it's background. [11] The researcher faced the challenge to develop the algorithms for tracking objects in unconstrained videos due to continuous change in environment.

Another system was the described [12] that focused on the development of a robot capable of detecting landmines, toxic

gases, fire and positions of the heat radiating life forms. Thus the system will perform multiple functions and hence is an intelligent multipurpose Warfield surveillance and scouting robot. The robot recognizes every face that has been already updated in the database. It easily updates a new person in the database by scanning and taking 20 photos. The person is associated with a certain ID and the ID is associated with the person's name so when the person appears in front of the webcam the system detects facial contours using the photos in the database for reference and displays the name of the person. This can be used for hostage situations, finding fallen soldiers on the battlefield, retrieving stolen objects, surveillance in areas with gas leaks or radiation, to detect landmines and various applications.

C. Analyzing Gestures of Body Motion

The author [13] proposed a new method for Multi-criteria collaborative filtering (MC-CF) system which reduces the dimensions of the fames in consideration applying various classification and soft computing techniques. The surveillance systems aims to improve the accuracy in the perception about the features extracted to analyze the behavior and prediction of the behavior of a suspect based on the features extracted. Such a real time system needs to have a high precision value and should be able to overcome the limitations of multiple dimensions, illumination conditions and uncertainty of the predictive behaviour. So to overcome these limitations, the proposed approach applied HOSVD for dimensionality reduction, ANFIS with subtractive clustering for knowledge discovery and classification methods for classification tasks. The approaches to the proposed method were used in two phases, offline and online. The offline phase uses a model with both the HOSVD and ANFIS including subtractive clustering for the tracking of the body in various dimensions and under various lighting conditions. For predicting the new user class in the online phase, SVM, K-NN and FBNN were used.

The proposed technique [14] analyzes human behavior and detects any kind of abnormality or suspicious gesture to predict if the individual is a criminal. The main objective of the research is to provide scientific evidence to support the efficiency of human detection systems and completely identify the criminal before suspect commits any crime. The technique employed the signal detection theory. The technique identifies the suspect into four categories: a hit, a correct rejection, a false alarm and a miss. The researcher took only one type of crime into consideration i.e. theft but still the criminals give a sign of their ill intentions when trying to be normal using which the observer can identify the suspicious behavior and hence improve the capability by training the system again. This led to the generation of false alarms.

To improvise upon the limitations of research another technique was proposed [6] with an intelligent camera driven data where the surveillance can be done in three steps: generating an alarm for the suspicious activities being captured, store the videos which have a high rate of abnormality and retrieve the data from the database after a crime has occurred. The proposed algorithm works on a database that is created with all the listed abnormal behavioral activities in the form a descriptive information. Thus if any kind of abnormality is deciphered by the smart camera then it checks for such abnormalities at various locations (multisite association analysis) and finally system generates a more accurate and reasonable pre-alarm. This proposed technique improves the smart monitoring performance of surveillance systems and stores the relevant surveillance video data thereby saving the storage space. Due to the efficient utilization of storage space, sometimes the relevant information might be lost which can lead to missing of a true positive in the detection of a threat. To overcome the inefficiency due to loss of data for the threat detection a new real-time scheme was proposed.

The author [15]describes a real-time problem identification at an airport and analyzing the high stake behavior using an integrated scientific approach. Since airports have a very stressful environment so optimizing the safety and security at the airport is a major task by eradicating the possibility of false positives. The various categories of threats were detected and deduced by the BDOs which included transporting weapons, IEDs or other dangerous or illegal objects/ substances through the airport, those with fake IDs or invalid IDs. An iALERT model proposed is to be registered with the system users so that the security system is not hacked as such security alert systems are prone to cyber-attacks.

III. DESCRIPTIVE ANALYSIS

Various researchers have worked on video frames to detect and classify the human body. The body motion analysis includes the extraction of features for the motion tracking, segmentation of the video frames to interpret the motion and study the gestures based on the movement. The various approaches proposed by the researches are summarized in Table 1.

TABLE 1 REVIEW OF TECHNIQUES USED

Title and Author	Work Done	Techniques used
Example Based Object Detection in Images by Components (C. Papageorgiou and T. Poggio, 2001)[2]	• detecting the constituent components • Combining the components using a classifier • Adaptive Combination of Classifiers (ACC) and Voting Combination of Classifiers (VCC). • Full body detection system as baseline	• Harr wavelets • SVM

Survey on Video Object Detection & Tracking (Dixit et al. ,2016) [5]	• Division of image in foreground and background pixels	• Contour tracking • Background subtraction • Kalman filters
A continuous object boundary detection and tracking scheme for failure-prone sensor networks (Imran et al., 2017) [7]	detect boundary of object by: • diagram-based network clustering • failure detection • recovery scheme	• Voronoi-based node clustering • Java simulator
Exploring Techniques for Vision Based Human Activity Recognition: Methods, Systems, and Evaluation (Xu et al , 2013) [8]	• features are extracted in three levels ➢ individual object features ➢ object features ➢ relationship between the two • study of the Motion-history image (MHI) and Motion Energy Image (MEI)	• background subtraction • binarization • Hidden Markovs Model • Hidden Semi Markovs Model
Dynamic, data-driven processing of multispectral video streams(Li et al, 2017) [9]	• Framework for processing of multi spectral video frames using Lightweight Dataflow techniques • Selection of band subset using pixel level fusion	• Multi spectral background subtraction • Multi spectral Foreground binarization
Visual perception-based criminal identification: a query-based approach (Singh et al, 2017)[10]	• Query based approach for reduction of search space • Two Datasets are considered i.e. viewed and forensic sketches • Measurement of approximation of age, height and concept of trusted features	• Dynamic decision tree data structure • MATLAB toolkit
Human action recognition based on discriminant body regions selection (Mliki et al., 2018) [11]	• Study of body skeleton to analyze human action • Offline phase to select region of interest and online phase to classify action • Cumulative skeletonized image matrix	• mRMR (Minimum Redundancy Maximum Relevance) SFS (Sequential Forward Selection) and SFFS (Sequential Floating Forward Selection) • SVM classifier
Military Robot for Reconnaissance and Surveillance using Image Processing (Zubair et al., 2017) [12]	• intelligent multipurpose Warfield surveillance and scouting robot • the system detects facial contours using the photos in the database for reference and displays the name of the person	• MATLAB toolkit
Using behavioral indicators to help detect potential violent acts: a review of the science base.(Davis et al., 2013) [13]	Uses multiple techniques for behaviour detection • Polygraph • Speech analysis • EEG • Voice stress • Gait analysis • Facial Expression	• ANFIS • HOSVD • SVM • K-NN
Who ' s the criminal ?" Early detection of hidden criminal intentions - influence of nonverbal behavioral cues, theoretical knowledge, and professional experience.(Frey 2014) [14]	Detection of suspect based on certain frames and check for • A hit • A miss • A false hit • A false miss All of these result in generation of an alarm with can be false or true	• Behavior monitoring cameras
Smart Monitoring Cameras Driven Intelligent Processing to Big Surveillance Video Data (Zhen Feng et al. , 2017) [6]	• generating alarm for captured suspicious activities • store videos with high abnormality rate • retrieve the data from the database after a crime has occurred	• multi-site monitoring cameras
Observe, Target, Engage, Respond (OTER): High-stake behaviour analysis using an integrated , scientific approach within an airport context (Lansley et al., 2017) [15]	• a real time problem identification using integrated scientific approach • proposed iALERT model for capturing of behavioral data	• real time truth detection using holistic approach

IV. RESEARCH GAPS

In recent years, various researchers have proposed work in the field of individual detection based on the video and visual images. Video analytics have played a major role in the various application including defense activity, surveillance activities and so on. Object detection, motion tracking and study of gestures in body motion use video analytics to monitor and track each individual for identification. The process of tracking is a challenging task due to various constraints in the environment. Various researchers have experienced many challenges in the security surveillance techniques. These challenges are the issue of concern and a future aspect for the

researchers in tracking any kind of suspicious behavior to monitor a suspect or a criminal.

1. One of the major challenge faced by [1] is that the change in appearance affects the view point and this change affects geometric properties. Due to detection of an object in various environments and with the unconstrained background, extraction of the geometric features of human body that is hand, leg, face, gesture, etc. is difficult.

2. Another problem for the researcher [7] is to develop the algorithms for tracking of a body motion in unconstrained environment but due to the dynamic shape and texture of the human body, this is challenging. This makes the study of the gestures of the moving body to analyze the behavior of an individual more difficult.

3. Due to efficient utilization of storage space in the system proposed by [10], the surveillance cameras stores only those video frames which detect changes in body gestures or track any kind of suspicious motion. But due to this, the relevant information might be lost which can lead to missing of a true positive in the detection of a threat. This result in the generation of false positives which may cause havoc in a real-time situation. This is a major concern for security.

V. CONCLUSION

The current research proposes to examine several recommendation systems available for recognition of the various features from the image in motion and analyzing the behavioral change in the video frames obtained. The allegations of the recognition systems and diverse techniques employed for the effective recognition of the patterns or any specific features of the image will be performed to accurately identify the behavior by image recognition and the other imaging techniques and detect a specific pattern in the video frame. Despite the valuable knowledge in the existing body of literature, this study utilizes the research gaps for the development of a model for accurate analysis and performance of the recommendation system.

REFERENCES

[1] A. Jalal, S. Kamal, and D. S. Kim, "Detecting complex 3D human motions with body model low-rank representation for real-time smart activity monitoring system," *KSII Transactions on Internet and Information Systems*, vol. 12, no. 3, pp. 1189–1204, 2018.

[2] C. Papageorgiou and T. Poggio, "Example Based Object Detection in Images by Components," *IEEE transactions on pattern analysis and machine intelligence*, pp. 1–35, 2001.

[3] A. Sargano, P. Angelov, and Z. Habib, *A Comprehensive Review on Handcrafted and Learning-Based Action Representation Approaches for Human Activity Recognition*, vol. 7, no. 1. 2017.

[4] T. B. Moeslund, A. Hilton, and V. Krüger, "A survey of advances in vision-based human motion capture and analysis," *Computer Vision and Image Understanding*, vol. 104, no. 2–3 SPEC. ISS., pp. 90–126, 2006.

[5] A. Dixit, M. Verma, and K. Patidar, "Survey on Video Object Detection & Tracking," *International Journal of Current Trends in Engineering & Technology*, vol. 02, pp. 264–268, 2016.

[6] Z. Shao, J. Cai, and Z. Wang, "Smart Monitoring Cameras Driven Intelligent Processing to Big Surveillance Video Data," *IEEE Transactions on Big Data*, vol. 7790, no. c, pp. 1–1, 2017.

[7] S. Imran and Y. B. Ko, "A continuous object boundary detection and tracking scheme for failure-prone sensor networks," *Sensors (Switzerland)*, vol. 17, no. 2, 2017.

[8] X. Xu, J. Tang, X. Zhang, X. Liu, H. Zhang, and Y. Qiu, "Exploring Techniques for Vision Based Human Activity Recognition: Methods, Systems, and Evaluation," *Sensors*, vol. 13, no. 2, pp. 1635–1650, 2013.

[9] H. Li *et al.*, "Dynamic, data-driven processing of multispectral video streams," *IEEE Aerospace and Electronic Systems Magazine*, vol. 32, no. 7, pp. 50–57, 2017.

[10] A. K. Singh and G. C. Nandi, *Visual perception-based criminal identification: a query-based approach*, vol. 29, no. 1. 2017.

[11] H. Mliki, R. Zaafouri, and M. Hammami, "Human action recognition based on discriminant body regions selection," *Signal, Image and Video Processing*, vol. 12, no. 5, pp. 1–8, 2018.

[12] Z. Ghouse, N. Hiwrale, and N. Ranjan, "Military Robot for Reconnaissance and Surveillance using Image Processing," *International Research Journal of Engineering and Technology (IRJET)*, vol. May, pp. 1767–1769, 2017.

[13] P. K. Davis, *Using behavioral indicators to help detect potential violent acts: a review of the science base*. 2013.

[14] C. I. Frey, *" Who ' s the criminal ?" Early detection of hidden criminal intentions - influence of nonverbal behavioral cues, theoretical knowledge, and professional experience*. 2014, p. 114.

[15] C. A. Lansley, A. J. Garner, D. E. Archer, R. Dimu, C. Blanariu, and S. Losnita, *Observe, Target, Engage, Respond (OTER): High-stake behaviour analysis using an integrated , scientific pproach within an airport context*, vol. 20, no. March. 2017.