ISSN 2250-3153

Human Activity Tracking using Star Skeleton and Activity Recognition using HMM's and Neural Network

Deenbandhu Singh*, Akhilesh Kumar Yadav*, Vivek Kumar *

* Computer Science Department
Babu Banarasi Das University, Lucknow, Uttar Pradesh, India.

Abstract- Human motion detection and analysis is currently an important area of research, motion analysis help us to solve many problems. An Automated Video Surveillance Model is presented in this paper. The model which I proposed is capable of detecting and monitoring people in both environments Indoor and outdoor. The Model is capable to find out the Suspicious and Non-Suspicious activities. It also detects multiple peoples in video and monitoring their activities. Moving targets are detected and their boundaries extracted, we use star skeletonization technique with the adaptive centroid point to create human skeletons. In this paper we use HMM-based methodology for action recognition. In our proposed method, a series of star skeletons is generated according to action over time. Then, time-sequential images frame is converted into a feature vector sequence and the feature vector is translated into sequence of symbols after that we use Neural Network for action recognition which represents the particular action in Suspicious and Non-Suspicious category. We design a codebook of posture, which contains representative star skeletons of various human action types and define a star distance to find out the similarity between feature vectors. Each and every sequence of feature vector is matched against the codebook and Neural Network after that Symbol is assigned to the most similar. Then the time-sequential images or frames are converted to a sequence of symbol posture by this, we easily categorize the actions.

Index Terms- Moving target, Star skeleton, Action recognition, HMM, Neural Network.

1. Introduction

Recognizing human activities from video is important applications of computer vision. This application is very useful for video surveillance, human-computer interface, industry, academia, security agencies, consumer agencies, and the general populace as well [1]. The classification of human actions has remained a challenging problem of the sheer amount of variations in the imaging conditions (view-point, illumination etc.) and attributes of the individual and therefore captures the local appearance and motion information. Methods based on local features or interest points have shown a promising result in action recognition. Most of the approaches described above advocates the use of single features of human action classification [2] [3]. In this paper, we propose a non-model-based algorithm to recognize human activity. In the proposed algorithm, we segment each frame of the input video into foreground and background regions. Morphological operations are applied to the foreground regions to fill holes, remove noise, and ensure connectivity [4]. Background subtraction is a widely used approach from detecting moving objects of videos from static cameras. Thinking of the approach is that of detecting the moving objects of the difference between the current frame and a reference frame, often called the "background model" or "background image". As a baric, the background image must be a representation of the scene of no moving objects and must be kept regularly updated so as to adapt to the varying luminaries conditions and geometry settings [5][6]. We use the Neural Network for classify the particular action [9] [10] [11]. More complex models have extended the concept of "background subtraction" beyond its literal meaning. Background subtraction (BS) is a "quick" way of localizing moving objects. The "star" skeleton consists of only the gross extremities of the target joined to its centroid in a "star" fashion [7] [8]. HMM is a type of stochastic state transit models. By HMMs we generate the codebook and we use Neural Network for action recognizes [12] [13] then we relate them to their use for automatic pattern recognition.

2. Proposed Method

This system architecture consists of three parts:

1. Feature Extraction
2. Mapping features to symbols
3. Action recognition

2.1 Feature Extraction

In the Feature Extraction phase we follow the three steps which are:-

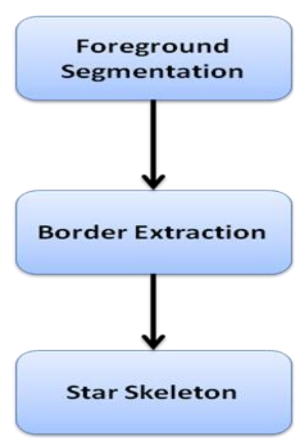


Figure-1 Process flow of Feature Extraction

ISSN 2250-3153

2.1.1 Foreground Segmentation

Foreground Segmentation is a common approach from discriminating moving objects of the Video. By removing the foreground from the background we got the Objects [1].

2.1.2 Border Extraction

For finding the Border of the Object we follow the step of figure-2. According to the algorithm first we Binaries the frame then we perform Dilation and Erosion by this any small holes in the target is removes and it smoothes out any interlacing anomalies [8]. After the target has been cleaned, its outline is extracted and we got the border of the target.

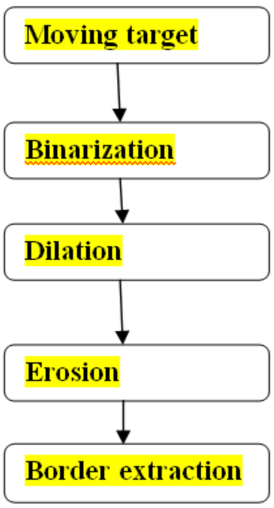


Figure-2 Process flow of border Extraction of target

2.1.3 Star Skeleton

Star skeleton is a very useful approach to find out the stricture of any object. Using the skeletonization approach, no such models are required, so the method can be applied to other objects like animals and vehicles.

- The centroid of the image boundary (X, Y) is determined.

$$X = \frac{1}{N_a} \sum_{i=1}^{N_a} X_i$$

$$Y = \frac{1}{N_a} \sum_{i=1}^{N_a} Y_i$$

Where (X, Y) is the average boundary pixel position, N_a is denoted the number of boundary pixels, pixel on the boundary of the target (X_i, Y_i).

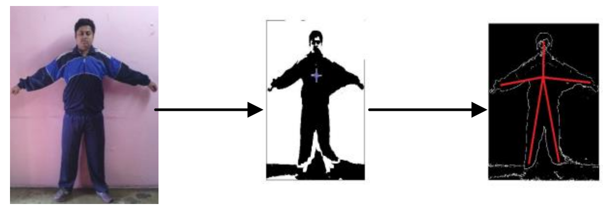
- The distances D from the centroid (X, Y) to each border point (X_i, Y_i) are calculated.

$$D=\sqrt{(X_i - X)^2 + (Y_i - Y)^2}$$

- D_i is expressed as a one dimensional discrete function of D. This function is periodic with period N_a. D_i is smoothed for noise reduction. This can be done using a low pass filtering or linear smoothing filter in the Fourier domain.
- Local maxima of Di are taken as extremal points, and the "star" skeleton is created by joining them to the target centroid (X, Y). By finding zero-crossings of the difference function Local maxima are detected.

$$\delta_i = D_i - D(i - 1)$$

By this "star" skeleton is illustrated.



- Advantages of star skeletonization
  - It is not iterative and is, therefore, computationally cheap.
  - It also explicitly provides a mechanism for controlling scale sensitivity.
  - It relies on no a priori human model.

2.2 Mapping Features to Symbols or Training

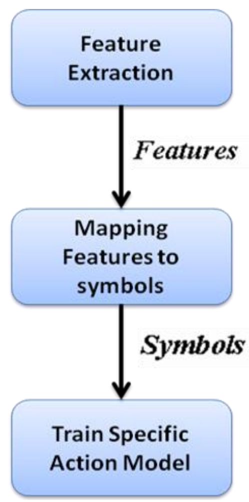


Figure-3 Process flow of border Mapping Features to Symbols or Training

2.2.1 Feature Extraction

A series of Human action postures is composed of time. Boundary shape technique is the effective way to represent the posture of human action. However, to describe a human posture using the whole human contour is inefficient since each border point is too much similar to its neighbor points. Principle Component Analysis techniques are used to reduce the redundancy, due to matrix operations it is computational expensive. On the other hand, to represent posture simple information likes human height and width may be used [7]. To describe a posture consequently features must be extracted.

ISSN 2250-3153

Human skeleton seems to be a good selection. For skeletonization there are many standard techniques such as distance transformation and thinning. These techniques are computationally expensive and highly susceptible to noise in the target boundary.

2.2.2 Mapping features to symbols

HMM is applying to time-sequential video, the symbol sequence is generated by extracted feature sequences of latter action recognition. This is accomplished by Vector Quantization technique.

❖ Hidden Markov Models (HMM)

HMMs are a type of stochastic state transit model. This treat discrete time sequences as the output of a Markov process whose states cannot be directly observed [12] [13].

Suppose that n = 5, then we have to collect statistics for $3^{(5-1)}$ = 81 past histories. Therefore, we will make a simplifying assumption, called the Markov assumption:

For a sequence $\{q_1, q_2 \ldots q_n\}$

$P(q_n|q_{n-1}, q_{n-2} \ldots q_1) = P(q_n|q_{n-1})$.

For first order Markov models we can use these probabilities to draw a probabilistic finite state Automaton. We say that the probability of a certain observation at time n only depends on the observation q_{n-1} at time $n-1$.
An HMM which has 'n' states $Q=\{q_1, q_2, \ldots q_n\}$ and 'm' output symbols $V=\{v_1, v_2, \ldots, v_m\}$.
n=3 and m=3 is shown in Figure. Let the state at time t be s_t.
Now the NxN state transition matrix A is

$A= \{a_{ij} \mid a_{ij} = P_r(s_{t+1} = q_j \mid s_t = q_i)\}$, where *aij* is the probability of transiting from state q_i to state q_j.

The NxM state output probability matrix B is

$B= \{b_j(k) \mid b_j(k) = P_r(v_k \mid s_t = q_j)\}$, where $b_j(k)$ *is* the probability of output symbol v_k at state q_j.

The initial state distribution vector π is

$\pi = \{\pi_i \mid \pi_i = P_r(s_1 = q_i)\}$

Recognition based on HMM involves two parts:
1. Training a model
2. Computing.
By trained each model it's become generates the symbol patterns for its training data. Appling the standard Baum-Welch algorithm we can perform training.

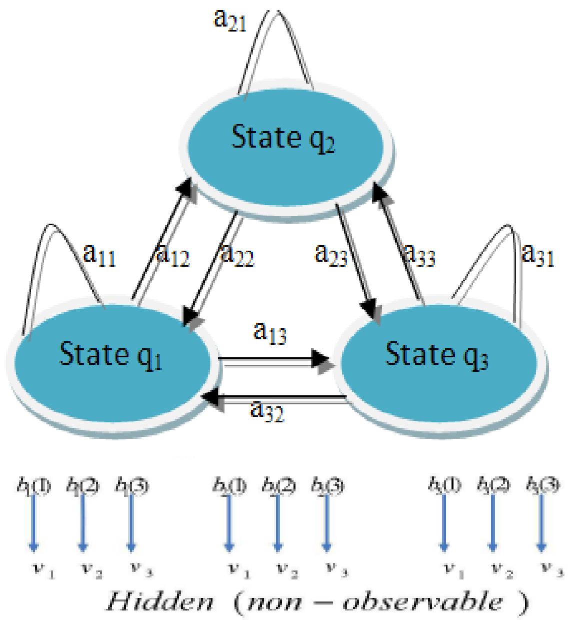


Figure-4 A full connected HMM with three states and three Outputs

❖ Vector Quantization

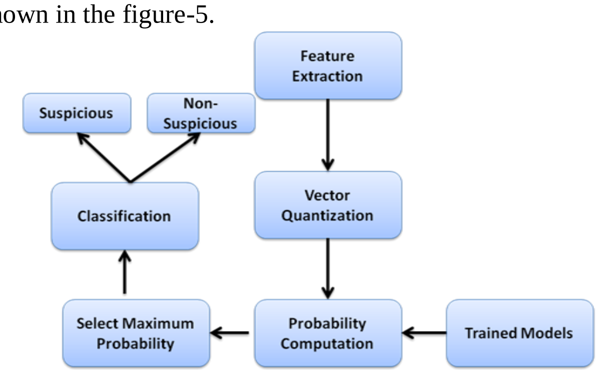
For vector quantization, codeword's $g_j \in R^n$ represents the centers of the clusters of the feature R^n space. It's required in centers of the clusters. Codeword g_j is assigned to symbol v_j. Consequently, the size of the code book is equal to the number of HMM output symbols. Each feature vectors f_i transformed into the symbol that assigned to the codeword nearest to the vector in the feature space. This means f_i is transformed into symbol v_j if $j=\arg\min_j d(f_i, g_j)$ where $d(x, y)$ is the distance between vectors x and y.

2.2.3 Train Specific Action Model

The concept behind using the HMMs is to develop a model for all of the actions that we want to recognize. For each action HMMs give state based representation. The number of states was determined. Further we calculate the probability $P(O|\lambda^i)$ by training each action model. We generate the observation posture sequence O, for probability of model λ^i.

2.3 Action recognition

For the action recognition we use the following model which is shown in the figure-5.



2.3.1 Classification

In the classification phase we us the neural network to separate suspicious and non-suspicious human posture. Figure-6 show the structure of neural network which used for classify the particular action [9] [10].

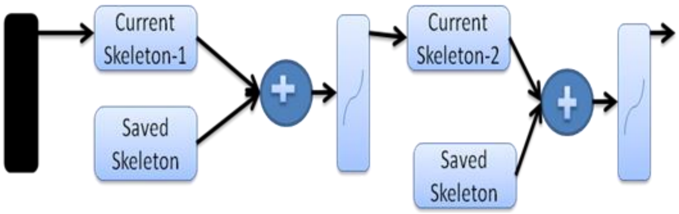


Figure-6 The neural network used for classification

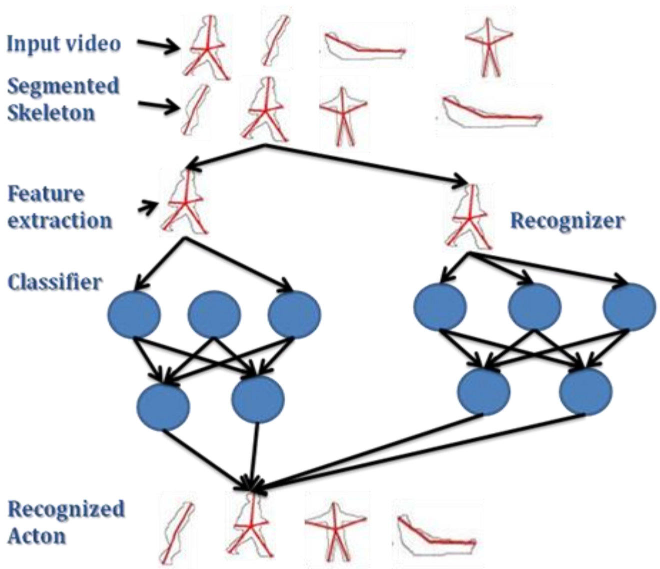
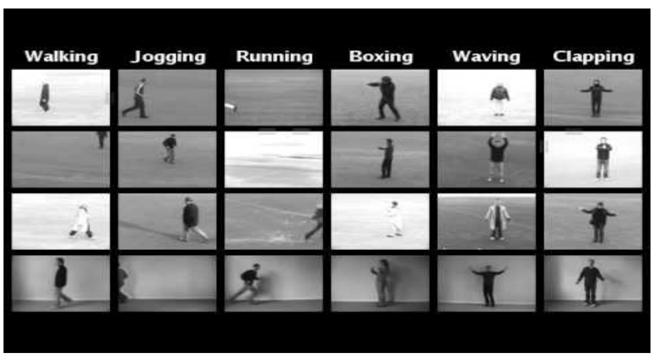


Figure-7 show the process flow of recognizing the specific action

3. Data Set Used

3.1 KTH Data Set:-

In this Data set we have used the 40 different type's video like walking, jogging, running, boxing, waving and clapping etc of 10 subjects. All the videos are capture in the standard condition.



3.2 BBDU Data set

In the BBDU Data set we have 30 videos of 5 subjects. They perform exercise, walking, sit up and sit down etc.

4. Result and Discussion

With the above approach, we have taken 30 videos of KTH data set with 10 subjects and classified 17 videos into suspicious and 13 videos into Non-Suspicious category. In BBDU Data Set we have taken 13 videos with 5 subjects and classified 7 videos into suspicious and 8 videos into Non-Suspicious category. After matching this BBDU Data Set with KTH Data Set, we observed that the results are quite similar under a test condition. Detailed analysis is represented in Matrix Form in Table 1.

Table 1 : Matrix for recognition of testing data of BBDU Data Set

	Sit up	Sit down	Exercise	Walking	Running	Jump
Sit up	12	1	0	0	0	0
Sit down	1	11	0	0	0	0
Exercise	0	0	13	0	0	0
Walking	0	0	0	13	2	0
Running	0	0	0	0	10	0
Jump	0	1	0	0	1	13

5. Conclusion and Future Work

In above proposed model we take the Star Skeleton approach for feature extraction because it is not iterative therefore, it was computationally cheap. It also explicitly provides a mechanism for controlling scale sensitivity. It relies on no a priori human model.HHM and Neural Network is use for recognition which represents the particular action in Suspicious and Non-Suspicious category. In KTH Data Set 98% accuracy is reported in test condition, in our mechanism we have reported 95% accuracy, whose detailed analysis is represented in table-1.In this paper we work only on single subject in different condition but in the feature we take HHM and Neural Network for recognizing various actions of multiple people in a single video.

REFERENCES

[1] Nizar Zarka, Ziad Alhalah and Rada Deeb, "Real-Time Human Motion Detection and Tracking.".

[2] Pavan Turaga, *Student Member*, *IEEE*, Rama Chellappa, *Fellow*, *IEEE*, V. S. Subrahmanian, and Octavian Udrea, "Machine Recognition of Human Activities: A Survey.", VOL. 18, NO. 11.2008.

[3] Matteo Bregonzio, Shaogang Gong and Tao Xiang, "Recognising Action as Clouds of Space-Time Interest Points.",2009.

[4] Jianpeng Zhou and Jack Hoang, "Real Time Robust Human Detection and Tracking System".

[5] Liang Wang, Weiming Hu, Tieniu Tan, "Recent developments in human motion analysis.", Pattern Recognition 36 (2003) 585 – 601.

[6] K. Srinivasan,K. Porkumaran and G. Sainarayanan, "Intelligent human body tracking, modeling, and activity analysis of video surveillance system: A survey" Vol.1 December 2009 pp 1-8.

[7] Hsuan-Sheng Chen, Hua-Tsung Chen, Yi-Wen Chen and Suh-Yin Lee, "Human Action Recognition Using Star Skeleton". VSSN'06, October 27, 2006, Santa Barbara, California, USA. Copyright 2006 ACM 1-59593-496-0/06/0010...$5.00.

[8] Hironobu Fujiyoshi and Alan J. Lipton, "Real-time Human Motion Analysis by Image Skeletonization".

[9] Ani1 K. Jain ,Jianchang Mao and K.M. Mohiuddin, *"Artificial Neural Network: A Tutorial."*. 0018-9162/96/$5.000 1996 IEEE

[10] Heiko Hoffmann, Michael D. Howard, and Michael J. Daily, "Fast pattern matching with time-delay neural networks".

[11] Hazem M. El-Bakry and Nikos Mastorakis, "Fast Human Motion Tracking by using High Speed Neural Networks". ISSN: 1790-5109 ,ISBN: 978-960-474-114-4.

[12] Barbara Resch, "Hidden Markov Models".

[13] Eric Fosler_Lussier, "Markov Models and Hidden Markov Models_ A Brief Tutorial".

AUTHORS

First Author – Deenbandhu Singh, M.tech, Computer Science Department Babu Banarasi Das University, Lucknow, Uttar Pradesh, India deenbandhusingh85@gmail.com

Second Author – Akhilesh Kumar Yadav, M.tech Computer Science Department Babu Banarasi Das University, Lucknow, Uttar Pradesh, India akhi2232232@gmail.com

Third Author – Vivek Kumar, M.tech (Sr. Lecturer), Computer Science Department Babu Banarasi Das University, Lucknow, Uttar Pradesh, India vivek.kumar0091@gmail.com

Correspondence Author Deenbandhu Singh,
deenbandhusingh85@gmail.com
deenbandhusingh13@yahoo.com
Contact no.9454668669, 8604663451