I J C T A, 9(2) 2016, pp. 1119-1123

Identification of the Movements of Human in a Video

[*]Divya M., [*]S. Padmavathi

ABSTRACT

Automatic video surveillance is used for monitoring the behavior of people from distance. Identifying human movement is a major task involved in the process. In this paper the movements of human in a video is identified by using the Global GIST feature. This feature is used to track the corner of the moving body in each frame. The results for various videos involving single or two persons are summarized.

Index Terms: Control Points, GIST Feature, Skeleton Detection.

I. INTRODUCTION

Human movement identification is much important in video surveillance. This movement helps to find the human activity in public or in private area. Identifying clashes between humans in the video automatically will help in handling the situation in a better way. A fight between two humans is identified by the movements of their hands, legs, and head.

In this paper, the human movement is extracted from video the background is estimated using first few frames and the foreground is extracted from the remaining frames using this. Blob analysis is done on the foreground and the GIST features are extracted for the energy levels .These features are used for identifying the human movements. The outer skeleton is detected as the control points are plotted on the corners of the human body.

This paper is organized into six chapters, Section II the related papers are explained in an overview. Section III describes the proposed method in detail. Section IV is the experimental analysis where the outputs are illustrated and explained in detailed. Section V is the conclusion and future work.

II. RELATED PAPERS

In [1], paper the human action recognition is made by using global GIST feature and local patch coding. The work plan of this paper includes:

The human action areas are extracted from each frames by using background subtraction. GIST feature is computed to each of action centric region and represented as m×n as the gist vector.

The action centric region of the m×n gist vector divides the human body into four patches. The edge patch, middle patch, up patch and the below patch. For action recognition the Support vector machine(SVM) is used for recognizing the actions for the given unknown input video.

In [2], paper the human action recognition surveillance based on Spatial Transform in Video Surveillance. This paper includes:

[*] Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetam, Coimbatore, India
E-mails: [1]divyamarudhavanan@gmail.com, [2]s_padmavathi@cb.amrita.edu

The foreground is extracted by using the background subtraction. The body boundaries are extracted and the control points to identify the posture of the body. The Delaunay triangulation technique transforms the extracted body boundary into triangular meshes. We extract the features like area, angle and centroid variations of formulated triangles from the triangulation result.SVM classifier is to classify different postures of the video.

In [3], paper is about the human detection and behaviour understanding. The techniques followed are:

The background subtraction is used for the foreground mask detection. Optical flow corresponds to the motion field. Object classification is done by motion based classification and shape based classification.

For action recognition they have used the template based methods and state space approaches. The purpose of behavior description is to report behaviors of the moving objects.

III. METHODOLOGY

(A) Foreground Mask Region Extraction and Blob Detection

The video sequence is converted to frames and each frame is subtracted from the previous frame to obtain the sequence of foreground.

The subtracted frames should be threshold to remove the small noisy regions from the frames. The larger area regions are identified and labeled. Assuming the major movement in the video, is made by the humans, the labeled regions corresponds to the humans. To have the clear information of the humans from each frames dilation and erosion is done to remove the unwanted noise from the frames. The blob analysis identifies the number of persons in the frames and labels the region of the human structure.

(B) Global GIST Feature Descriptor

GIST feature is like a dictionary, which stores the feature information and compares it with each frame. The Gabor filter bank is applied and the energy levels are calculated. The orientation, histogram, intensity, width and the height of each frame is mapped with a threshold value after the blob detection.

The pyramid feature calculates the position of the object. And these energy levels are stored as Bag-of-features. The Gabor function is defined as: [1]

$$G(X,Y)=\exp((x1^2+Y1^2)/2s^2(S1))\cos(2\eth(FxX1+FyY1) \quad \text{Equation.1}$$

where X1=x cosQs + y sinQs; Y1=x sinQs + y cosQs and (FX,FY) the frequency of sinusoidal component s standard deviation of gauss function, S is the number of scales, Qs is the orientations of scale S.

The GIST feature helps to plot the control points on the proper edges of the body, even if the body moves flexibly. This helps to improve in plotting the points with good accuracy even in complex videos.

(C) Skeleton Detection

The skeleton detection is used for tracking the human body movement. This is achieved by finding the corners on the edges and marking the control points. It is to extract a region-based shape feature representing the general form of an object. It is a common pre-processing operation in raster-to-vector conversion or in pattern recognition.

(D) Control Point

The control points are plotted on the body structure. The head of the body is found using viola jonas algorithm. The corner and edge are detected from other parts of the body. The overall frame work is given in Fig. 1.

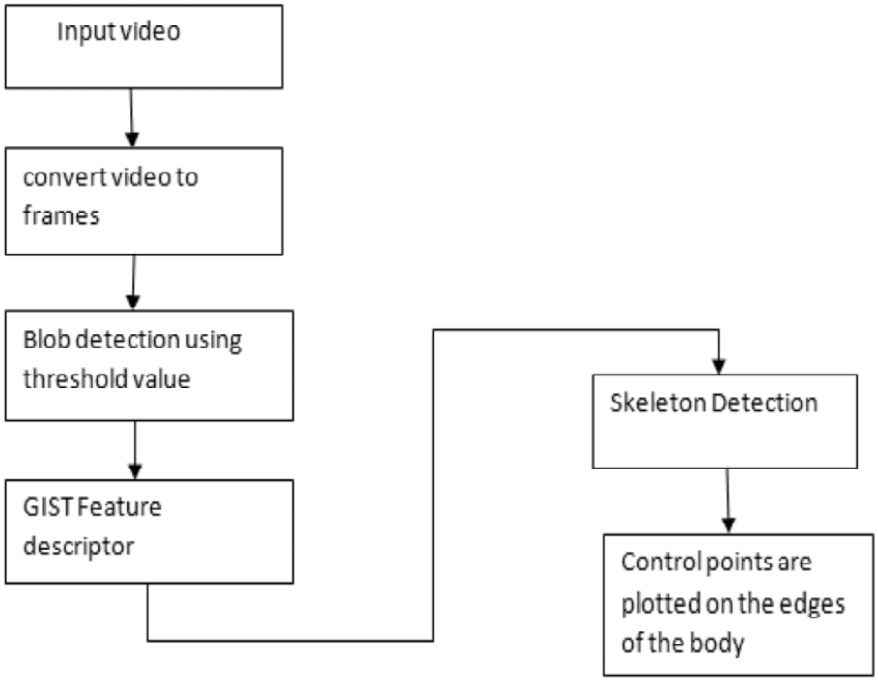

Figure 1: Frame Work

IV. RESULT ANALYSIS

The UT dataset is used for experimentation. The data set involves videos of the actions like fighting, kicking, pushing, hugging, and handshaking.

The video is converted to frames. These frames subtracts with the previous frame for the foreground mask.

Figure 2: Video to Frame Conversion

The larger region areas are labeled which are used in blob Analysis. The major movements in the video which is assumed as humans. The blob identifies the two persons on the frames differentiating it with green and red in color.

The Gist descriptor is to compare the features within the video to converted frames. The GIST is generated in row and columns. The red line is plotted in the X and Y axes using the boundary extraction.

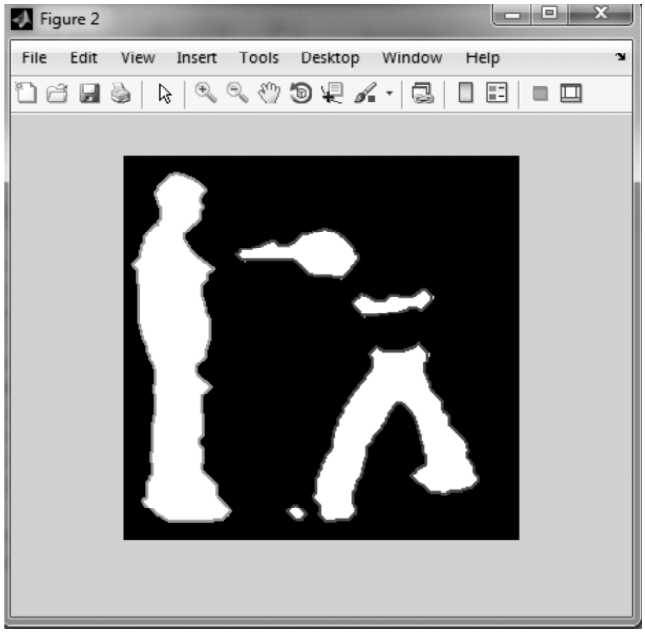

Figure 3: Blob Detection

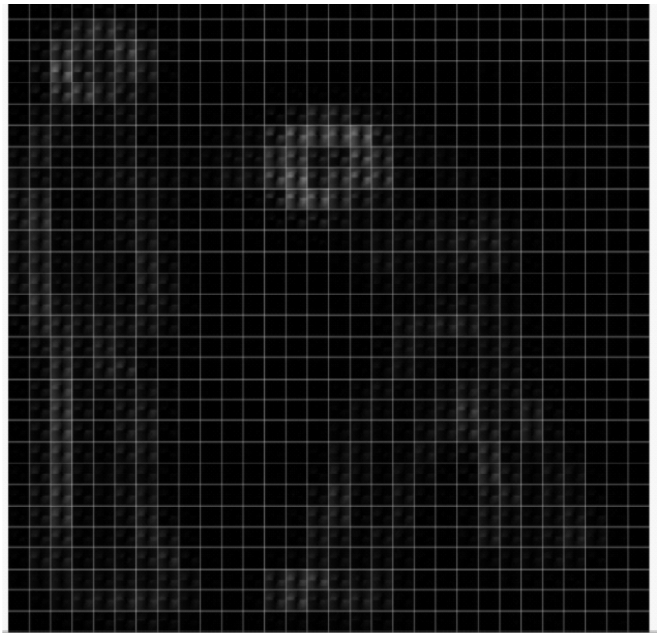

Figure 4: Gist Feature Descriptor

From the output of the Gist feature descriptor the Skeleton detection is formed**.**

The control points are plotted with the help of skeleton detection. The red dots on the corners of the human body are the control points.

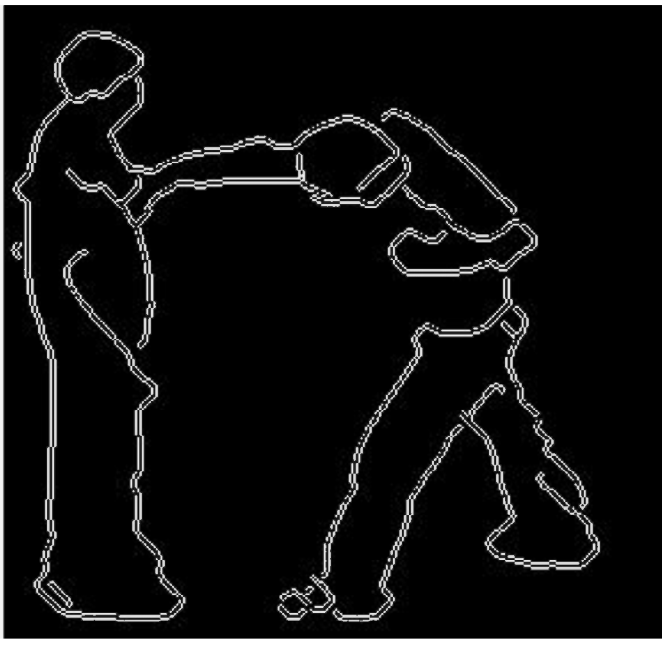

Figure 5: Skeleton Detection

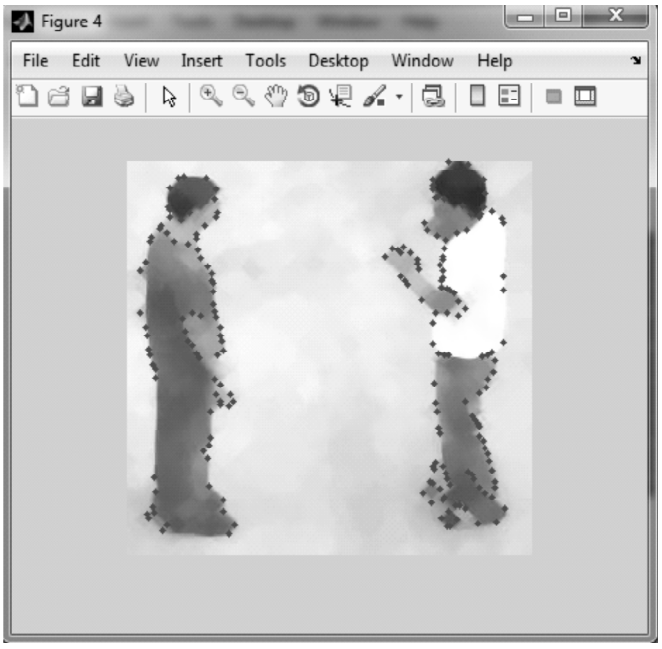

Figure 6: Control points

V. CONCLUSION

The human movement identification is a challenging part and they are to be trained with the large number of datasets. The dataset involve action of single or two persons. The Gist feature descriptor and skeleton detection has the good efficiency to find the flexible movements of body in the unknown video. In the future work is to identify the clash and non-clash between two people or many people and indicating a censor as violence on the particular clash video.

REFERENCES

[1] Yangyang Wang1, Yibo Li2 and Xiaofei Ji2,"Human Action Recognition Based on Global Gist Feature and Local Patch Coding", *International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol. 8, No. 2* (2015).

[2] A. Vasanth Kumar, A. D. Saravanan and S. R. Vinotha ," Human Activity Recognition Based on Spatial Transform in Video Surveillance ", *International Conference on Computational Techniques and Artificial Intelligence* (2011).

[3] Xiaofei Ji, Student Member, IEEE, and Honghai Liu, Senior Member, *IEEE ,"* Advances in View-Invariant Human Motion Analysis: A Review", *IEEE Transactions on Systems, Man, and Cybernetics—Part c: Applications and Reviews, Vol. 40, No. 1,(2010).*

[4] Paulo Peixoto", Jorge Batista1, Helder J. Araujo2, " Real-time human activity monitoring exploring multiple vision sensors", *Robotics and Autonomous Systems 35*(2001).

[5] Daniel S. Chivers and A. Ardeshir Goshtasby," Human Action Recognition in Videos via Principal Component Analysis of Motion Curves", *Ipcv* (2012).

[6] Z.A. Khan, W. Sohn," Real Time Human Activity Recognition System based on Radon Transform", IJCA Special Issue on "*Artificial Intelligence Techniques - Novel Approaches & Practical Applications", AIT,* (2011).

[7] Matteo Munaro, Gioia Ballin, Stefano Michieletto, Emanuele Menegatti," 3D Flow Estimation for Human Action Recognition from Colored Point Clouds", *Computer Vision and Image Understanding, vol. 115, no. 2,* (2013).

[8] Alessandro Perina a,*, Marco Cristani b,1, Vittorio Murino b,2," Learning natural scene categories by selective multi-scale feature extraction", Image and Vision Computing, (2010).

[9] Xiaofei Ji, Honghai Liu," Advances in View-Invariant Human Motion Analysis: A Review", IEEE Transaction on System, Man, And Cybernetics—Part C: Applications And Reviews, VOL. 40, NO. 1, (2010).

[10] Paulo Peixoto", Jorge Batista1, Helder J. Araujo2," Real-time human activity monitoring exploring multiple vision sensors", Robotics and Autonomous Systems (2001).