JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

HUMAN POSE ESTIMATION USING MACHINE LEARNING

Name of Authors: Alefiya Laturwala, Komal Jadhav, Riya Verma, Shikha Bhaskar

Project Guide: Prof. S.V.Wankhade

Department of Computer Engineering
Sinhgad Academy of Engineering, Kondhwa Bk, Pune

Abstract: Computer vision is becoming increasingly important in video surveillance, video retrieval and analysis, and human-computer interaction because of the ubiquitous nature of video data. Due to rising demand for automated analysis of human actions by computers, this project is being undertaken. This paper presents a human activity recognition system based on global feature extractions that can operate in real time. In the proposed method, video sequences are used to compute features of images. A real-time recognition system is the goal. UT Actions such as hugging, handshakes, kicking, punching, and pushing are included in the dataset. Real-time detection of various human activities is made possible through the use of an SVM classifier. The dataset is used to find similar videos based on the actions in the input videos.

Keywords: Support Vector Machine (SVM)

I. INTRODUCTION

Intelligent health care, video surveillance, human-computer interaction, and visual content retrieval systems all benefit from action recognition in videos. People's appearance, lighting changes, and the amount of data generated make video-based real-time human activity recognition a difficult task.

There are three main steps in a real-time human activity recognition system: detection, tracking, and identification. As the number of people using digital video cameras in their daily lives rises, so does the amount of video being created, uploaded, and stored online and in large video data sets.

There are numerous applications for human action recognition including visual surveillance, content-based video retrieval, human-computer interaction, and sports annotation. Visual surveillance systems in large public areas, for example, can automatically extract high-level semantic information from surveillance video if human action recognition is successful.

The tracks of a person's body parts were used as input features in early human action recognition attempts. As a result, most recent research shifts from skeletons to low-level features, such as local features, because full-body tracking from videos is still a difficult problem to solve Recently, the rapid development of depth sensors (such as the Microsoft Kinect) has resulted in sufficient accuracy and low cost for real-time full-body tracking. As a result, we can once again test the viability of activity recognition based on skeleton features. To classify videos of simple periodic actions performed by a single person, algorithms have been proposed in the past. For example, 'pushing' and 'handshakes' are two examples of actions and activities that take place frequently in the real world and are often performed by multiple people (e.g. Complex non-periodic activities, such as interactions between multiple people, will be required for a variety of applications (e.g. automatic detection of violent activities in smart surveillance systems).

II.MOTIVATION

Action recognition in videos has applications in intelligent health care, video surveillance, human computer interaction and visual content retrieval systems. Video based real time human activity recognition is a complex and challenging task due to variation in people's appearance, illumination changes and the amount of data generated. The main step of real time human activity recognition system involves person detection, tracking and recognition

III. REVIEW OF LITERATURE

A real time human activity recognition system based on Radon transform (RT), Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) is presented. Artificial Neural Nets (ANN) is used to recognize different human activities [1].

The data extracted using optical flow is converted to binary image. Then Histogram of Oriented Gradient (HOG) descriptor is used to extract feature vector from the binary images. These feature vectors are given as training features to Support Vector Machine (SVM) classifier to prepare a trained model [2].

In this paper, video based action recognition is performed on KTH dataset using four combinations of two feature descriptors and two classifiers. The feature descriptors used are Histogram of Oriented Gradient Descriptor (HOG) and 3-dimensional Scale Invariant Feature Transform (3D SIFT) and classifiers used are Support Vector Machine (SVM) and K Nearest Neighbour (KNN). Features are extracted from frames of training videos using descriptor and clustered to form Bag-of-words model. [3].

This approach predicts human actions using temporal images and convolutional neural networks (CNN). CNN is a type of deep learning model that can automatically learn features from training videos. Although the state-of-the-art methods have shown high accuracy, they consume a lot of computational resources. Another problem is that many methods assume that exact knowledge of human positions [4].

The central idea of principal component analysis (PCA) is to reduce the dimensionality of a data set consisting of a large number of interrelated variables, while retaining as much as possible of the variation present in the data set. This is achieved by transforming to a new set of variables, the principal components (PCs), which are uncorrelated, and which are ordered so that the first few retain most of the variation present in all of the original variables [5].

Human Activity Recognition Using an Ensemble of Support Vector Machines is employed to improve the classification performance by fusing diverse features from different perspectives. The Dempster-Shafer fusion and product rule from the algebraic combiners have been utilized to combine the outputs of single classifiers [6].

Human motion capture continues to be an increasingly active research area in computer vision with over 350 publications over this period. A number of significant research advances are identified together with novel methodologies for automatic initialization, tracking, pose estimation, and movement recognition. Recent research has addressed reliable tracking and poses estimation in natural scenes. Progress has also been made towards automatic understanding of human actions and behavior [7].

IV. PROPOSED SYSTEM

The proposed work is a two-person interaction-based, video-based system for identifying human activity. This paper presents a human activity recognition system based on global feature extractions that can operate in real time. In the proposed method, video sequences are used to compute features of images. A real-time recognition system is the goal. Actions like hugging and shaking hands are included in this dataset as well as kicking, punching, and pushing. Using this, all features used for classifying two-person interactions are represented by a body-pose feature in the context of identifying interaction activities via Support Vector Machine (SVM). User is the only module in the system. The system responds to a user's request.

V. PROPOSED SYSTEM ARCHITECTURE

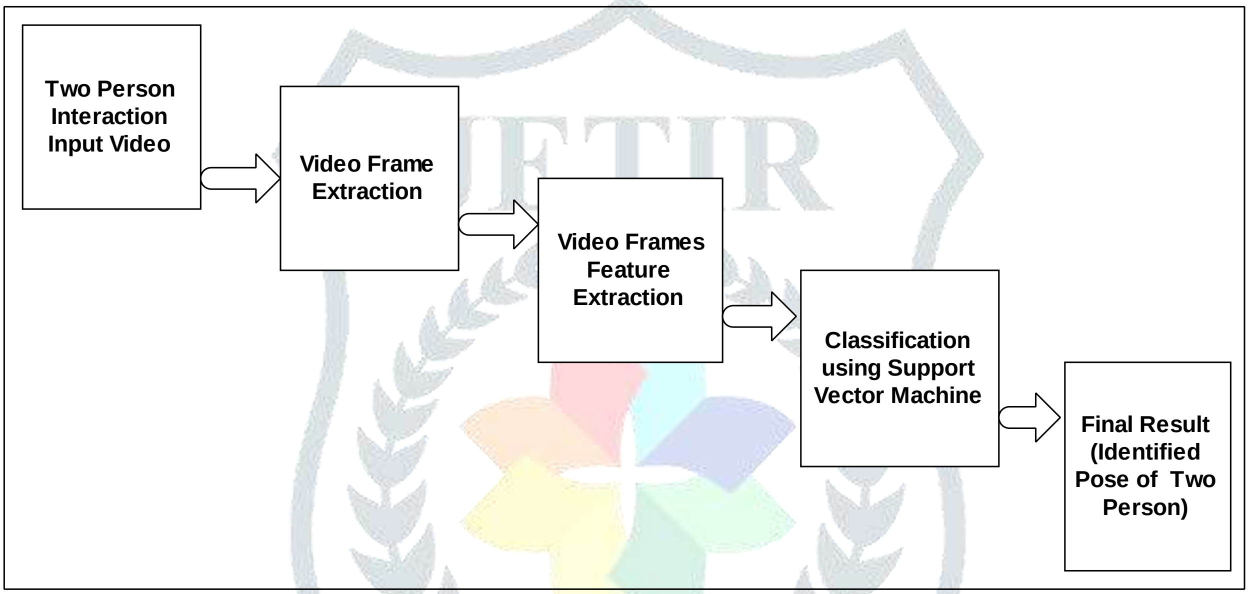

Fig Proposed System Architecture

VI.APPLICATIONS

This project makes the activity recognition techniques more valuable and widely used in diversified applications of our daily lives. In this section, we focus on four dominant applications, including

- Surveillance environments
- Entertainment environments
- Healthcare systems
- In sports

VII.CONCLUSION

In this project, a feature representation and activity recognition system for video retrieval system human activity recognition is proposed. The selection of key frames to represent a sequence of activity significantly reduced the computational complexity. To determine a person's activity, the feature extraction algorithm is implemented. As a result, our human activity recognition system operates more effectively. Classifier support vector machine is used to perform recognition (SVM). For detecting human activity, the SVM classifier achieves the best recognition results. The database is searched for videos that are similar to the input video.

VIII.ACKNOWLEDGEMENT

We would like to express our special thanks to our guide Prof. S.V.Wankhade (Sinhgad Academy Of Engineering ,Pune) who helped us in every possible way and resolved the problems we faced.

IX. REFERENCES

[1] 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.

[2] Jagadeesh B, Chandrashekar M Patil, “Video Based Action Detection and Recognition Human using Optical Flow and SVM Classifier”, IEEE International Conference on Recent Trends in Electronics Information Communication Technology, May 20-21, 2016, India

[3] Aishwarya Budhkar, Nikita Patil, “Video-Based Human Action Recognition: Comparative Analysis of Feature Descriptors and Classifiers”, International Journal of Innovative Research in Computer and Communication Engineering, Vol. 5, Issue 6, June 2017

[4] ChengbinJin, ShengzheLi, Trung Dung Do, Hakil Kim,” Real-Time Human Action Recognition Using CNN Over Temporal Images for Static Video Surveillance Cameras”, Information and Communication Engineering, InhaUniversity,Incheon,Korea

[5] Jolliffe, I.T., “Principal component analysis”, Springer Series in Statistics, 2nd ed., Springer, 2002.

[6] E. Mohammadi, Q.M. Jonathan Wu, M. Saif,” Human Activity Recognition Using an Ensemble of Support Vector Machines”, 2016 IEEE International Conference on High Performance Computing Simu- lation (HPCS), July 2016

[7] T. B., Hilton, A., and Kruger, V., “A survey of advances in vision-based human motion capture and analysis “, Computer Vision and Image Understanding, vol. 104, pp. 90-126, 2006.