Self-Similar Human Action Recognition

Anju Marium Abraham[1] Smita C Thomas[2]
[1,2]Mount Zion College of Engineering, Kadamannitta, Pathanamthitta, Kerala, India

Abstract— The dynamic of an human action is recognized by the system. In this software we first train the system for identifying the human face, sound generation and action recognition based on action recognition technique. HOG pattern generation, Edge based matching and Skin colour segmentation are the algorithms used here. We train the system for identifying the human face, skin and action, the system extract all the features and store it in the database.it is applicable in the pharmacy management system for managing the pharmacy details such as medicine details, user details etc. it is mainly useful for dumb people who wants to know more about medicines etc. The proposed system make use of silhouettes, tracked points and histogram of gradient.

Key words: Human Action Recognition

I. INTRODUCTION

Action recognition is active area of research.it attracted the attention of researchers. Action recognition is applicable on Automatic video indexing and archiving, video surveillance, human computer interaction, augmented reality, user interface design. There are many problem such as large variation in performing an action by different people, varying the postures, execution speed, variation in the sequences, distracting background motion. Aftereffect of the current methods restricted and simplified scenarios with simple backgrounds, simpler kinematic action classes, static cameras or limited view variations Various approaches have been proposed for action recognition, classified on the basis of representation, they can be categorized as: time evolution of human silhouettes, space-time shapes, dense trajectories and local 3D patch analysis generally coupled with some machine learning techniques. All these works rely on effective feature extraction. These feature extraction methods divided into the following four categories: motion based [1], appearance based, space-time volume based [2]–[3], and space-time interest points or local features based Motion based methods generally compute optical flow from a given action sequence, followed by appropriate feature extraction[1]. However, these methods are known to be very sensitive to noise and easily lead to inaccuracies. Appearance based methods are prone to differences in appearance between the training dataset and the testing sequences. Volume or shape based methods mostly require highly detailed silhouette extraction, which may not be possible in real-world noisy video dataset.

II. PROPOSED SYSTEM

The aim of this project is to create a method to recognize hand gestures on the basis of pattern recognition technique, employing histograms of local orientation and visual surveillance system. The orientation histogram will be used as a feature vector for gesture classification and interpolation. In practice, it involves the construction of human detectors, where the detectors search given images for human and localize them.

A human action recognition system has a combination of two key factors: a feature extraction algorithm that transformers image regions to feature vectors, and a detector that uses the computed features to make human/non-human decisions. This thesis targets general purpose human detectors that do not make strong contextual assumptions. More robust discriminant image descriptors simplify the classification task allowing humans to be discriminated more easily with less training data and less computational costs.

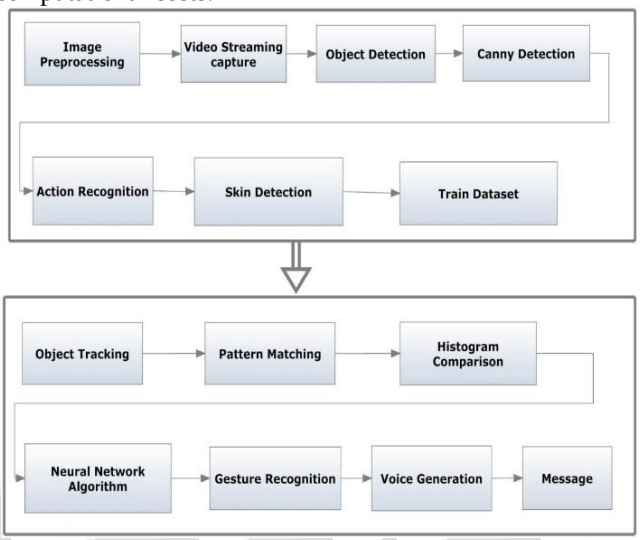

Fig. 2: System Architecture

Applications such as recognizing particular static hand signal, require a huge description of the shape of the input object than image provide. If the hand signals are in predetermined set, and the camera views a close-up of the hand, we may use an example-based approach, combined with a simple method top analyze hand signals called orientation histograms. This applications involve two phases; training and running. In the training phase, the user shows the system one or more action of a specific hand shape. The computer forms and stores the corresponding orientation histograms. In the execution phase, the system compares the histogram of the current image with each of the stored templates and selects the category of the closest match, or interpolates between templates, as appropriate. This method should be robust against small differences in the size of the hand but probably would be sensitive to changes in hand orientation. Face monitor module communicate with the camera/Web cam connected and captures the image in accordance with user input to system. If a person get into the area and makes some hands gestures like signs in front of the camera, the application should detect the type of the gesture, and raise an event, that is, when a hands gesture is detected, the application could perform different actions depending on the type of the gesture. That is, each alphabet may be assigned a unique gesture. In this system, the image of the hand is captured using a simple web camera. The acquired image is then processed and some features are extracted. These features are then used as input to a classification algorithm for recognition. The recognized gesture may then be used to generate speech. Very little attempts have been made in the

past to recognize the gestures made using hands but with limitations of recognition rate and time. This system is on the basis of ASL (American Sign Language). In ASL each alphabet of English vocabulary, A-Z, is assigned a unique gesture.

III. LITERATURE SURVEY

Human recognition has continued to be an active area of research and has thus rightfully attracted much attention from the researchers over the years. Important application domains, such as automatic video indexing and archiving, video surveillance, human-computer interaction, augmented reality, user interface design, and human factors would benefit immensely from a robust and efficient solution to this problem.

In this system we capture live video stream and pass it as input and process it to detect humans and gestures that occur in the video. The two main processes done in this project are Gesture recognition using Contour-based Object Detection and Skin recognition algorithm. The increase in human-machine interactions in our daily lives has made user interface technology progressively more important. Physical gestures as intuitive expressions will greatly ease the interaction process and enable humans to more naturally command computers or machines. Gesture recognition can be seen as a way for computers to begin to understand human body language, thus building a richer bridge between machines and humans than primitive text user interfaces or even GUIs (graphical user interfaces),which still limit the majority of input to keyboard and mouse. Gesture recognition enables humans to interface with the machine (HMI) and interact naturally without any mechanical devices. Using the concept of gesture recognition, it is possible to point finger at the computer screen so that the cursor will move accordingly. This could potentially make Conventional input devices such as mouse, key boards and even touch-screens redundant. A gesture may be defined as a movement, usually of hand or face that expresses an idea, sentiment or emotion e.g. rising of eyebrows, shrugging of shoulders is some of the gestures we use in our day to day life. Sign language is a more organized and defined way of communication in which every word or alphabetic assigned some gesture. In American Sign Language (ASL) each alphabet of English vocabulary, A-Z, is assigned a unique gesture. Sign language is mostly used by the deaf, dumb or people with any other kind of disabilities. With the rapid advancements in Technology, the use of computers in our daily life has increased manifolds. Our aim is to design a Human Computer Interface(HCI) system that can understand the sign language accurately so that the signing people may communicate with the non signing people without the need of an interpreter. It can be used to generate speech or text. Unfortunately, there has not been any system with these capabilities so far. A huge population in India alone is of the deaf and dumb. It is our social responsibility to make this community more independent in life so that they can also be apart of this growing technology world. In this work a sample sign language has been used for the purpose of testing .No one form of sign language is universal as it varies from region to region and country to country and a single gesture can carry a different meaning in a different part of the world. Various available sign languages are American Sign Language (ASL), British Sign Language (BSL), Turkish Sign Language (TSL),Indian Sign Language (ISL) and many more. There are a total of 26 alphabets in the English vocabulary. Each alphabet may be assigned a unique gesture. In our project, the image of the hand is captured using a simple web camera. The acquired image is then processed and some features are extracted. These feature are then used as input to a classification algorithm for recognition. The recognized gesture may then be used to generate speech. Few attempts have been made in the past to recognize the gestures made using hands but with limitations of recognition rate and time. In our project we are using one of the ASL (American Sign Language). In ASL each alphabet of English vocabulary, A-Z, is assigned a unique gesture.

This system uses motion detection as its first step, and then does some interesting routines with the detected object – hands gesture recognition. For a video objects, depending on the threshold value setting of Canny filtering, object segmentation is performed, all the features points found in the bounding box are used as a matching template in searching for a corresponding object bounding box in the next frame. The center of the bounding box is then used to indicate the location (coordinate) of the object. A search range is defined and all pixels within the search area are compared. The gravity center of the best matched points is considered as the location of the object frame. The bounding box of the object is updated accordingly. Note that this matching process is the most computationally intensive module of the object tracking system. When somebody gets into the area and makes some hands gestures like signs in front of the camera, the application should detect the type of the gesture, and raise an event, that is, when a hands gesture is detected, the application could perform different actions depending on the type of the gesture. That is, each alphabet may be assigned a unique gesture. In this system, the image of the hand is captured using a simple web camera. The acquired image is then processed and some features are extracted. These features are then used as input to a classification algorithm for recognition. The recognized gesture may then be used to generate speech. Few attempts have been made in the past to recognize the gestures made using hands but with limitations of recognition rate and time. In our project we are using the ASL (American Sign Language). In ASL each alphabet of English vocabulary, A-Z, is assigned a unique gesture.

Deaf Mute Communication Interpreter- A Review [4] : This paper aims to cover the various prevailing methods of deaf-mute communication interpreter system. The two broad classification of the communication methodologies used by the deaf –mute people are - Wearable Communication Device and Online Learning System. Under Wearable communication method, there are Glove based system, Keypad method and Handicom Touch-screen. All the above mentioned three sub-divided methods make use of various sensors, accelerometer, a suitable micro-controller, a text to speech conversion module, a keypad and a touch-screen. The need for an external device to interpret the message between a deaf –mute and non-deaf-mute people can be overcome by the second method i.e online learning system. The Online Learning System has different methods. The five subdivided methods are- SLIM module, TESSA, Wi-See

Technology, SWI_PELE System and Web-Sign Technology. An Efficient Framework for Indian Sign Language Recognition Using Wavelet Transform [5]:The proposed ISLR system is considered as a pattern recognition technique that has two important modules: feature extraction and classification. The joint use of Discrete Wavelet Transform (DWT) based feature extraction and nearest neighbour classifier is used to recognize the sign language. The experimental results show that the proposed hand gesture recognition system achieves maximum 99.23% classification accuracy while using cosine distance classifier. Hand Gesture Recognition Using PCA in [6]: In this paper authors presented a scheme using a database driven hand gesture recognition based upon skin color model approach and thresholding approach along with an effective template matching with can be effectively used for human robotics applications and similar other applications.. Initially, hand region is segmented by applying skin color model in YCbCr color space. In the next stage thresholding is applied to separate foreground and background. Finally, template based matching technique is developed using Principal Component Analysis (PCA) for recognition. Hand Gesture Recognition System for Dumb People [7]: Authors presented the static hand gesture recognition system using digital image processing. For hand gesture feature vector SIFT algorithm is used. The SIFT features have been computed at the edges which are invariant to scaling, rotation, addition of noise.

IV. Conclusion

The most developing area in the field of research is action recognition. There are a lot methods, such as automatic video indexing and archiving, video surveillance, human-computer interaction, augmented reality, user interface design, and human factors would benefit immensely from a robust and efficient solution to this problem. The aim of this project is to create a method to recognize hand gestures on the basis of pattern recognition technique, employing histograms of local orientation and visual surveillance system. The orientation histogram will be used as a feature vector for gesture classification and interpolation. In practice, it involves the construction of human detectors, where the detectors search given images for human and localize them.

Acknowledgement

We would like to thank, first and foremost, Almighty God, without his support this work would not have been possible. We would also like to thank all the faculty members of Mount zion college of engineering, for their immense support.

References

[1] A. A. Efros, A. C. Berg, G. Mori, and J. Malik, "Recognizing action ata distance," in Proc. 9th IEEE ICCV, Oct. 2003, pp. 726–733.umber 5 october 2016. (journal style)
[2] A. Yilmaz and M. Shah, "Actions sketch: A novel action representation," in Proc. IEEE CVPR, Jun. 2005, pp. 984–989.
[3] M. Grundmann, F. Meier, and I. Essa, "3D shape context and distance transform for action recognition," in Proc. 19th ICPR, Dec. 2008, pp. 1–4.
[4] Sunitha K. A, Anitha Saraswathi.P, Aarthi.M, Jayapriya. K, Lingam Sunny, "Deaf Mute Communication Interpreter- A Review", International Journal of Applied Engineering Research, Volume 11, pp 290-296 , 2016.
[5] Mathavan Suresh Anand, Nagarajan Mohan Kumar, Angappan Kumaresan, " An Efficient Framework for Indian SignLanguage Recognition Using Wavelet Transform" Circuits and Systems, Volume 7, pp 1874-1883, 2016.
[6] Mandeep Kaur Ahuja, Amardeep Singh, "Hand Gesture Recognition Using PCA", International Journal of Computer Science Engineering and Technology (IJCSET), Volume 5, Issue 7, pp. 267-27, July 2015.
[7] Sagar P.More, Prof. Abdul Sattar, "Hand gesture recognition system for dumb people",

