

Applying Hidden Markov Model Technique in CSMMI for Action and Gesture Recognition

Naeem Ahmed
IV Sem, M.tech DCN branch
T.John Institute of Technology
Bangalore, India

Mrs. Champa H
Assistant professor, Dept. of ECE
T. John Institute of Technology
Bangalore, India

Abstract:- Gesture recognition is with the goal of interpreting Human gestures via mathematical algorithms. Gestures can be commonly originated from the face or hand and from any bodily motion or state. Here the main theme focuses on the hand gesture recognition. To interpret sign language many approaches have been used by the usage of cameras and computer vision algorithms. However, the identification and recognition of posture and human behavior, gesture-capture devices, tracking algorithms for capturing of motion, extraction of features, and classification algorithms are subjects of gesture recognition. Many approaches have been used by the usage of computer vision algorithms and cameras. Gesture recognition is the way through which computers are made to understand human body language and hence builds bridge between machines and humans than the text user interfaces or graphical user interfaces, which even limits the input given by keyboard and mouse. Full body gesture recognition technique has wide verities of applications in research field including dance gesture recognition. Though it is a challenging task, there is a rapid growth in this domain and now by using approach called Class-Specific Maximization of Mutual Information a sub modular technique aims at learning a compact and discriminative wordbook for every category. The previous dictionary-based algorithms depends on shared wordbook for all of the different categories but here there is a tendency to combine the mutual data of both interclass and interclass into a single objective for the optimization of class-specific wordbook.

I. INTRODUCTION

Gesture recognition is becoming increasingly important as a mode of communication, in addition to the more common visual and oral modes. It is being investigated as an alternative method of computer input for people with severe speech and motor impairment.

An initial model-based approach to estimate the kinematic motion of the arm from acceleration measurements is shown by the chosen measurement schemes. Recently, sparse representations for human action recognition are receiving an increasing attention. The theory of "Sparse representation for dictionary learning frameworks" was proposed using K-SVD shared dictionary (one dictionary for all classes), class-specific dictionary (one dictionary per class) and concatenated dictionary (concatenation of the class-specific dictionary).

Wrong outputs are the common false results on gesture recognition. A gesture may seem fine in development but is found to trigger accidentally during an initial deployment of the interface, restarting development and increasing cost and time consumption taken for the project development. K-SVD, a technique for false positive prediction and prevention has used. The K-SVD only focuses on minimizing the reconstruction error and it is not clear from, how to optimize the learned dictionaries. The learned dictionary via K-SVD may be not compact and discriminative. Computer recognition of hand gestures may provide a more natural-computer interface, allowing people to point, or rotate a CAD model by rotating their hands. Hand gestures can be classified in two categories: static and dynamic.

1.1 Objective and Scope

The introduction of a new method called Class Specific Maximization of Mutual Information (CSMMI) which is used to learn a compact and discriminative dictionary for each class. CSMMI not only discovers the latent class-specific dictionary items that best discriminates different actions, but also captures unique dictionary items for a specific class. One of the common approaches for dictionary optimization is to use information theory and it shows promising results for action and gesture recognition.

1.2 Applications

The concept involving Action and Gesture recognition system can be used in many of the areas, Such as:

Man and machine interface: Using hand gestures to control the computer functions. Example of this, this has been implemented in this project as it controls various keyboard and mouse functions using gestures alone.

For gaming technology: Gestures are used in interactions within video games to try and to make the game player's experience more interactive.

Control of mechanical systems (such as robotics): Using the hand to remotely control a manipulator.

As a remote control: Remote control with the wave of a hand for various devices is possible. The signal will indicate the desired response, and also shows which device to be controlled.

Recognition of Sign language: Some types of gesture recognition software can transcribe the symbols represented from sign language into text.

Control through facial gestures: Controlling of a computer through facial gestures is a useful application for gesture recognition system for the users who may not physically be able to use a mouse or keyboard.

Eye tracking: This may be of use for controlling cursor motion or focusing.

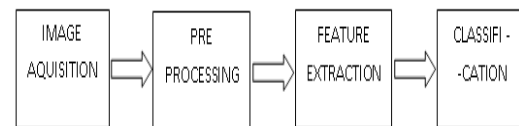
II.EXISTING SYSTEM

The traditional dictionary-based algorithms typically learn a shared dictionary for all of the classes, error in recognizing a Gesture is common problem for the interfaces that depends on gesture recognition. Gesture may look fine in development but is found to trigger accidentally during an initial deployment of the interface, again restarting the entire development are increasing in expense. Hear a technique called K-SVD is introduced in order to detect wrong prediction and prevention that can be used interactively avoided. The K-SVD only focuses on minimizing the reconstruction of errors and it is not so clear from, how to optimize the learned dictionaries. The learned dictionary through K-SVD may be not being compact and discriminative.

The tendency is to scale back the computational complexness of CSMMI, now by introducing a novel sub modular technique, which is one among the necessary contributions. Additionally this contributes a progressive end to-end system for action and gesture recognition, with feature extraction and learning initial wordbook per each category. CSMMI through sub modularity, and classification supported reconstruction errors.

III.PROPOSED SYSTEM

In this research, I mainly focus on the research of segmentation and extracting of the fingers in gesture recognition. We create an application by using Matlab to extract the feature of a specific gesture (which is the gesture "one") in different natural environments. Our aim is to extract the feature of the specific gesture "one" in different natural environments. The proposed Concept is named as Hidden Markov Model Technique (HMM) for action and gesture recognition. HMM have 4 steps feature extraction and representation, learning the initial class specific dictionaries, CSMMI and classification. But this is the only focus on the shared dictionary while this work explores the relationship between intra-class and inter-class MI for video-based recognition.



Block diagram of HGR System

Hand gesture is one of human body languages which are popularly used in our daily life. It is a communication system that consists of hand movements and facial expressions via communication by actions and sights. Gestures have vivid, concise and intuitive features that it is very worthy to be researched in Human-Computer Interaction. For example, a same gesture may present different meanings in different cultures. The research of gestures can help people to distinguish different gesture cultures.

3.1.Hidden Markov Models for modeling and recognizing gesture under variation

Hidden Markov models (HMMs) is a popular technique for recognizing human hand and movement gestures in a many of the applications and sensor configurations, as discussed in the applications above. Among their benefits is the fact that the gesture models may be trained automatically using series of examples of the gesture class, and the fact that the trained models encode the variation present in the set of examples. On the surface, applying HMMs to the task of recognizing gestures from video as input is no different than applying HMMs to any other kind of signal: features are computed at each time step, example sequences of the features are stored, and models trained on the examples are later matched to a novel input feature sequence. A naive application of HMMs to recognize gestures from video might treat the collection of image pixel values at each time step as the feature vector. Besides being computationally daunting, this approach suffers from the fact that it would take a great many examples to span the space of variation present in the raw appearance of a human gesture, particularly if multiple viewing conditions and multiple users are considered.

The HMM is governed by:

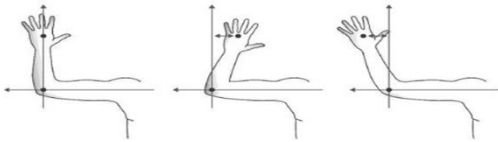
1) The Markov chain process with a finite number of states and

2) A set of random functions, each associated with one state.

In discrete time instants, the process is in one of the states and generates an observation symbol according to the random function corresponding to the current state. Each transition between the states has a pair of probabilities, this has been defined as follows:

1) Transition probability, which provides the probability for undergoing the transition;

2) Output probability, which defines the conditional probability of emitting an output symbol from a finite given state.



Simple Man-machine interfaces used in HMM

IV. METHODOLOGY

Five fundamental steps have to be performed in Action and Gesture Recognition:

The GUI has option to select between Action or Gesture, because Action requires Video as the input and Gesture requires An image.

Step I. Read the Input Video; it should be in .avi Format.

The system will use a single, colour camera mounted above a computer screen. The output of the camera will be displayed on the monitor.

Step II. The Algorithm will then convert the video into frames.

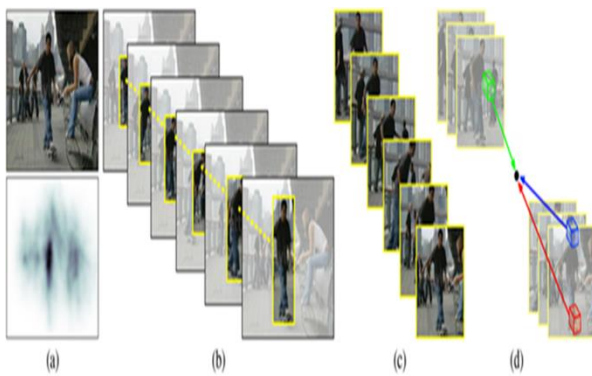
Step III. Now by using Hidden Markov models (HMM), the previous frames are compared with the next frames.

Step IV. Clustering: Hear grouping of related gestures and actions are done and background is subtracted.

Step V. The Gesture is being recognized and compared for some Classes using CSMMI technique, and the Recognized Gestured is identified and output is displayed.

Step VI. For Gesture recognition image is given as input and same process is followed.

V. EXPECTED RESULT



- Background Subtraction
- Converting Video into Frames
- Detecting Action
- Change in the action.

VI. CONCLUSION

From the above project, the algorithms which first of all shows the extracting the moving objects from a video feed and second to successfully Recognize hand gestures demonstrated by the given object. The recognition algorithm (HMM) is simple and easy for the implementation, and for understanding. Overall the design met initial expectation in many aspects. It detected hand motion and gesture.

REFERENCES

- [1] Q. Qiu, Z. Jiang, and R. Chellappa, "Sparse dictionary-based Representation and recognition of action attributes," in *Proc. IEEE ICCV*, Nov. 2011, pp. 707–714.
- [2] T. Guha and R. K. Ward, "Learning sparse representations for human action recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 8, pp. 1576–1588, Aug. 2012.
- [3] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing over complete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
- [4] R. Fergus, L. Fei-Fei, P. Perona, and A. Zisserman, "Learning object categories from Google's image search," in *Proc. IEEE ICCV*, vol. 2, Oct. 2005, pp. 1816–1823.

WEBSITES

1. <http://www.mathworks.com>
2. <http://www.ieee.org>
3. <http://www.wikipedia.com>
4. <http://www.science.gov>
5. www.Google.com