

Hand Gesture Recognition From Video

Priya Gairola¹, Sanjay Kumar²

Computer Science and Engineering, Uttarakhand Technical University, Dehradun, India

Assistant professor, Department of Computer Science and Engineering, Uttarakhand Technical University Dehradun, India

Abstract: Gestures are the very expressive and meaningful body motions. Gesture recognition provides meaningful expressions of motion by a human involving the hands, arms, face, head, and body. Human motion capture is used both when the subject is viewed as a single object and articulated motion with a number of joints. For example, Speech and handwriting, here gestures vary between individuals, and even we can say for the same individual between different instances. In this paper we have implemented the hand gesture recognition from video using point matching technique. Gesture recognition typically requires the use of different imaging and tracking devices or gadgets.

Keywords: Indian Sign Language (ISL), Hand posture, Key point Matching, Sign/hand gesture recognition, feature extraction, gesture determination

1. Introduction

Gestures are ambiguous and incompletely specified. For example, there are different ways to indicate the concept "stop," one can use gestures such as a raised hand with palm facing forward or, an exaggerated waving of both hands over the head. Similarly, speech and handwriting, gestures vary between individuals, and even for the same individual between different instances.

The meaning of a gesture can be dependent on the following:

- Spatial information: (where it occurs)
- Pathic information: (the path it takes)

Gestures can be:

- Static (the user assumes a certain pose or configuration)
- Dynamic (with prestroke, stroke, and poststroke phases).

1.1 Structure of Gesture Recognition

1. Initialization
2. Tracking
3. Pose Estimation
4. Recognition

1. Initialization

It deals with Preprocessing of data, Camera calibration, Adaptation to scene characteristics, Model initialization, It Covers the action needed to ensure a system commences its operation with a correct interpretation.

2. Tracking

Establishing coherent relations of subject between frames, to prepare data for POSE ESTIMATION, Can be done in LOW LEVEL & HIGH LEVEL.

- LOW LEVEL (such as edges).
- HIGH LEVEL (such as hand & head)

3. Pose Estimation

How the human body is configured in a given scene? Various levels of accuracy is required in pose estimation, the precise pose in terms of positions, orientation, width.

4. Recognition

Recognition provides meaningful expressions of motion by a human; Recognition is static or dynamic (i.e. whether the recognition is based on one or more frames).

1.2 Type of Gesture

- 1)Hand And Arm Gesture: Recognition of hand poses, sign languages, and entertainment applications (allowing children to play and interact in virtual environments)
- 2)Hand and Face Gesture: Shaking of head, direction of eye gaze, raising the eyebrows, opening the mouth to speak, winking, flaring the nostrils, looks of surprise, happiness, disgust, fear, anger, sadness, contempt, etc.
- 3)Body Gesture: Involvement of full body motion, such as tracking movements of two people interacting outdoors, analyzing movements of a dancer for generating matching music and graphics and recognizing human gaits for medical rehabilitation and athletic training.

2. Literature Survey

We have studied many previous works done in this field by different researchers. There are many approaches that were followed by different researchers like vision based, data glove based, Artificial Neural Network, Fuzzy Logic, Genetic Algorithm, Hidden Markov Model, Support Vector Machines etc. Some of the previous works are given below: There are different tools for gesture recognition:

1. Based on the approaches ranging from statistical modeling
2. Computer vision
3. Pattern recognition
4. Image processing,
5. Connectionist systems, etc

Most of the problems have been addressed based on statistical modeling, such as:

1. HMMs
2. PCA
3. Kalman filtering
4. Advanced particle filtering
5. Condensation algorithms
6. FSM

2.1. HMM

A time-domain process demonstrates a Markov property if the conditional probability density of the current event, given all present and past events, depends only on the j th most recent event.

First Order Markov Process:

If the current event depends solely on the most recent past event this is a useful assumption to make, when considering the positions and orientations of the hands of a gesturer through time.

The HMM is a double stochastic process governed by:

1. An underlying Markov chain with a finite number of states
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.

2.2 PCA

The Eigen faces method is based on the statistical representation of the face space. It finds the principal components (Karhunen–Loeve expansion) of the facial image distribution, or, the eigenvectors of the covariance matrix of the set of face images. These eigenvectors representing a set of macro features (that are generated *a posteriori* on a statistical basis) characterizing the face, constitute the Eigen faces. Fisher's linear discriminant separates the eyebrows, eyes, nostrils, mouth, cheeks, and chin.

The model employs the following:

1. An improved variation of adaptive Hough transform for geometrical shape parameterization, involving curve detection based on ellipse containing the main (oval) connected component of the image (related to cheeks and chin detection)
2. Minima analysis of feature candidates corresponding to the low-intensity regions of the face (extracting eyes, nostrils, and mouth)
3. Template matching for inner facial feature localization, using an appropriate binary mask on an area restricted by the eyes (extracting upper eyebrow edges that may not otherwise be uniformly described by a geometric curve)
4. Dynamic deformation of active contours for inner face contour detection
5. Projective geometry properties for accurate pose determination, along with analysis of face symmetry properties for determination of gaze direction. The skin like regions in the image are detected from the hue saturation-value (HSV) color space representation

2.6 FSM Approach

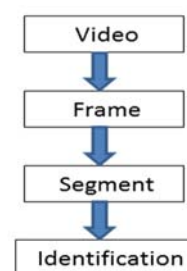
FSM has been effectively employed in modeling human gestures. Computer vision and pattern recognition techniques, involving feature extraction, object detection, clustering, and classification, have been successfully used for many gesture recognition systems. In the FSM approach:

1. A gesture can be modeled as an ordered sequence of states in a spatio-temporal configuration space.
2. The number of states in the FSM may vary between applications.
3. The gesture is recognized as a prototype trajectory from an unsegmented, continuous stream of sensor data constituting an ensemble of trajectories.

The trajectories of the gestures are represented as a set of points (e.g., sampled positions of the head, hand, and eyes) in a 2-D space. The training of the model is done **off-line**, using many possible examples of each gesture as training data, and the parameters (criteria or characteristics) of each state in the FSM are derived. The recognition of gestures can be performed **online** using the trained FSM. When input data (feature vectors such as trajectories) are supplied to the gesture recognizer, the latter decides whether to stay at the current state of the FSM or jump to the next state based on the parameters of the input data.

If it reaches a final state, we say that a gesture has been recognized. The state-based representation can be extended to accommodate multiple models for the representation of different gestures, or even different phases of the same gesture. If more than one model (gesture recognizer) reaches their final states at the same time, we can apply a winning criterion to choose the most probable gesture. The concept of motion energy has been used to extract the temporal signature of hand motion from a limited set of dynamic hand gestures. This is subsequently analyzed and interpreted by a deterministic FSM.

3. Proposed Work

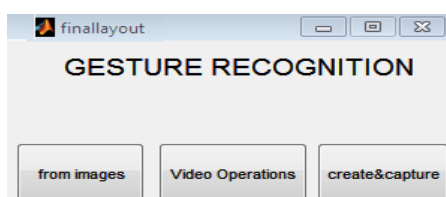


1. From images
We have two options:
 - Load image
 - Find match
- Video operations
 - Capture video
 - Snapshot
 - Segmentation
- Create capture
 - Capture gesture from video

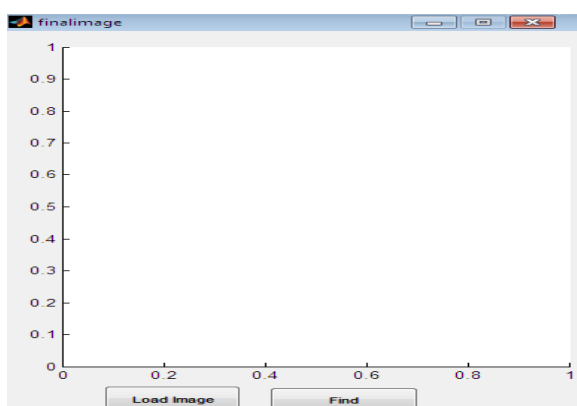
The proposed algorithm can be explained in following manner:

- Step [1] Extraction of the frame from the video
- Step [2]: Obtain the number of frames in the video. With this the total time of the video gets divided and time of image is known.
- Step [3]: Capture the object occurring in this time span and save in the folder as frame1
- Step [4]: Loop through the movie, writing all frames out. Each frame will be in a separate file with unique name.
- PART [2] Adjusting the size of each image
- Step [5]: once the frame is obtained, the part of image which does not carry any information is selected.
- Step [6]: This part of image is excluded by cropping the rest of the part of image.

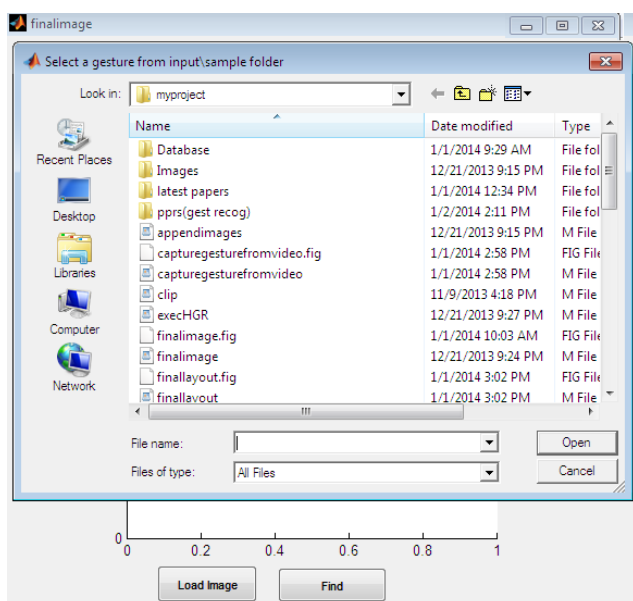
Final layout



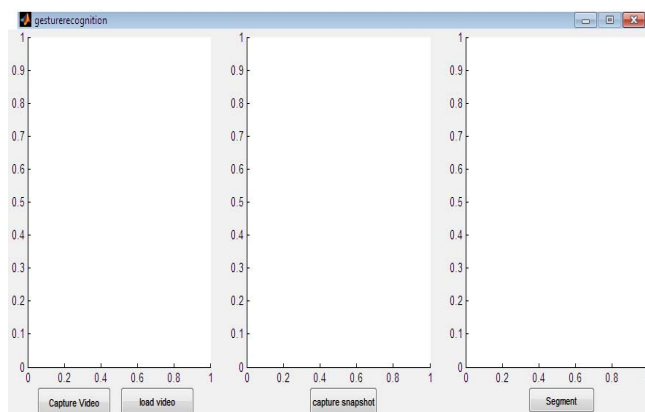
From image



Load image



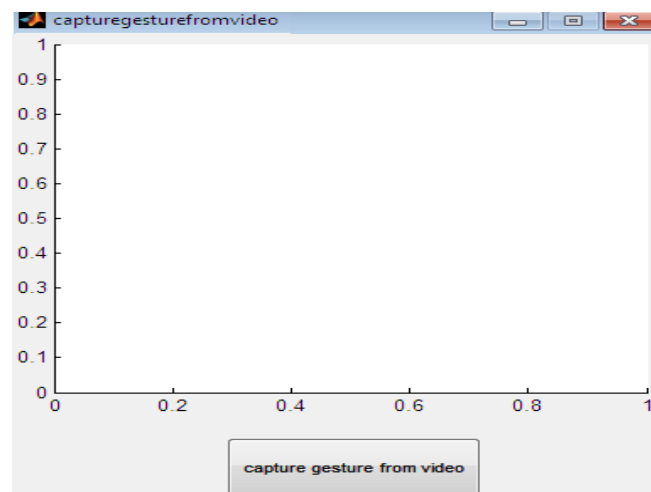
Video operation



Segmentation and conversion into grey scale



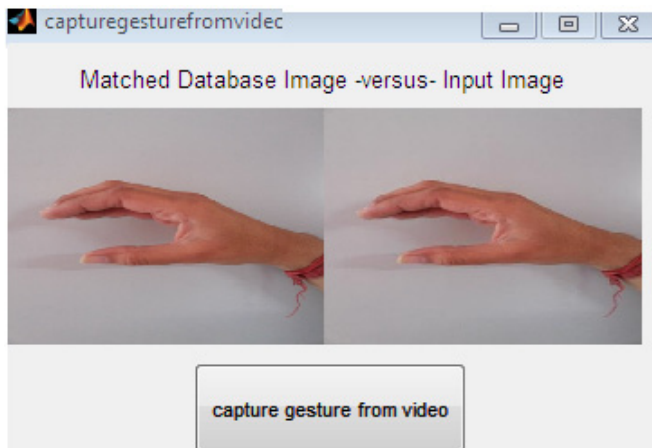
Capture gesture from video



Gesture recognised from video



Gesture recognized from video



Gesture recognized from video



4. Conclusion and Future Work

The proposed algorithm shows very good results as it shows 100% accuracy i.e. it extracts and compares each frame of given video with database. With our algorithm we were able to decode a video successfully with frames. It will benefit human motion capture. The proposed system shows good performance with the database of American Sign Language. In future the algorithm can be applied to live images i.e. real time. The system can also be tested on other databases and with two hand signs. It will benefit human motion capture.

References

- [1] Abdunnaser Diaf(1) and Rachid Benlamri, University of Windsor, Windsor, ON N9B 3P4, Canada , An Effective View-based Motion Representation for Human Motion Recognition
- [2] Chris Joslin, Ayman El-Sawah, Qing Chen, Nicolas Georganas Discover Laboratory, University of Ottawa, SITE, 800 King Edward Avenue, Ottawa, Ontario, K1N 6N5 {joslin, aelsawah, qchen georganas}@discover.uottawa.ca Dynamic Gesture Recognition, IMTC 2005 – Instrumentation and Measurement Technology Conference Ottawa, Canada, 17-19 May 2005
- [3] Daniel Wood Methods for Multi-touch Gesture Recognition

- [4] George Caridakis, Kostas Karpouzis, Nasos Drosopoulos and Stefanos Kollias Image, Video and Multimedia Systems Laboratory National Technical University of Athens gcari, kkar pou, ndroso, skollias@image.ntua.gr

ADAPTIVE GESTURE RECOGNITION IN HUMAN COMPUTER INTERACTION

- [5] German Torres, University of Kansas ‘Gesture Recognition Using Motion Detection’ germantf@ku.edu
- [6] J. Rekha, 2J. Bhattacharya and 3S. Majumder Scientist, 3Scientist & Head, Surface Robotics Laboratory, Central Mechanical Engineering Research Institute (CMERI-CSIR) Shape, Texture and Local Movement Hand Gesture Features for Indian Sign Language Recognition
- [7] Joyeeta Singhal¹, Karen Das², ^{1,2}Department of Electronics and Communication Engineering, ‘Hand Gesture Recognition Based on Karhunen-Loeve Transform’ Mobile & Embedded Technology International Conference 2013
- [8] Kiran S. Bhat Carnegie Mellon University, kiranb@cs.cmu.edu Mahesh Sapharishi Carnegie Mellon University Pradeep Khosla Carnegie Mellon
- [9] Mark Bayazit, Alex Couture-Beil, Greg Mori School of Computing Science Simon Fraser University Burnaby, BC, Canada Real-time Motion-based Gesture Recognition using the GPU
- [10] Mark Stamp Associate Professor, Department of Computer Science San Jose State University, A Revealing Introduction to Hidden Markov Models, September 28, 2012
- [11] Michel Valstar, Maja Pantic and Ioannis Patras Delft University of Technology Electrical Engineering, Mathematics and Computer Science Man-Machine Interaction Group, Delft, the Netherlands Motion History for Facial Action Detection in Video
- [12] Mohamed Alsheakhali, Ahmed Skaik, Mohammed Aldahdouh, Mahmoud Alhelou, Computer Engineering Department, The Islamic University of Gaza, ‘Hand Gesture Recognition System’ 2011
- [13] Rafiqul Zaman Khan¹ and Noor Adnan Ibraheem, Department of Computer Science, A.M.U., Aligarh, India

COMPARATIVE STUDY OF HAND GESTURE RECOGNITION SYSTEM

- [14] Raymond Lockton and Andrew W. Fitzgibbon, Department of Engineering Science University of Oxford, ‘Real-time gesture recognition using deterministic boosting’
- [15] Ross Cutler University of Maryland College Park, Maryland rgc@cs.umd.ed Matthew Turk Microsoft Research Redmond, Washington mturk@microsoft.com View-based Interpretation of Real-time Optical Flow for Gesture Recognition
- [16] Ruize Xu, Shengli Zhou, and Wen J. Li, Fellow, IEEE, ‘MEMS Accelerometer Based Nonspecific-User Hand Gesture Recognition’ IEEE SENSORS JOURNAL, VOL. 12, NO. 5, MAY 2012
- [17] Sonal Singh, Alpika Tripathi, ‘Gesture Recognition Using Wrist Watch’ International Journal of Scientific & Engineering Research, Volume 4, Issue 6, June-2013 984 ISSN 2229-5518 IJSER © 2013

- [18] Sushmita Mitra, *Senior Member, IEEE, and Tinku Acharya, Senior Member, 'Gesture Recognition: A Survey' IEEE A Survey of Computer*
- [19] Taher, Khadhraoui LSTS Laboratory, National School of Engineers of Tunis (ENIT), Tunisia, 'Gesture Determination for Hand Recognition'
- [20] Thomas B. Moeslund and Erik Granum, 'Vision-Based Human Motion Capture', *Laboratory of Computer Vision and Media Technology, Aalborg University*, accepted September 27, 2000
- [21] *University, pkhosla@cmu.edu, Motion Detection and Segmentation Using Image Mosaics*