

Automatic Classification of Hand Gesture Contours Based on Human Computer Interaction

Ms. Nayan S. Mane¹, Prof. Suresh S. Rode²

PG Student, Electronics & Telecommunications Department, D.Y. Patil School of Academy, Pune, India¹

Assistant Professor, Electronics & Telecommunications Department, D.Y. Patil School of Academy, Pune, India²

Abstract: The accurate classification of hand gestures in the development of novel hand gesture based systems designed for human computer interaction (HCI) and for human alternative and augmentative communication (HAAC) is decisive. A complete vision-based system consisting of hand gesture acquisition, segmentation, filtering, representation, and classification is developed for classification of hand gestures. The algorithms in the subsystems are formulated or selected to optimally classify hand gestures. By using a histogram thresholding algorithm grey scale image of a hand gesture is segmented. For effectively remove background and object noise in the segmented image a morphological filtering approach is designed. The contour of a gesture is represented by a localized contour sequence whose samples are the perpendicular distances between the contour pixels and the chord connecting the end-points of a window centred on the contour pixels. Gesture similarity is determined by measuring the similarity between the localized contour sequences of the gestures. To measure the similarity between the localized contour sequences the linear alignment and nonlinear alignment are developed. Experiments and evaluations on a subset of American Sign Language (ASL) hand gestures show that, by using nonlinear alignment, no gestures are misclassified by the system. Additionally, it is also estimated that real-time gesture classification is possible through the use of a high-speed PC, high-speed digital signal processing chips, and code optimization.

Keywords: Gesture, augmentative, Human Computer Interaction, VHGC System.

I. INTRODUCTION

This paper describes the design and implementation of a vision-based hand gesture classification (VHGC) system which can be used for novel human-computer-interaction (HCI) applications and for human alternative and augmentative communication (HAAC) applications. The main approaches for analysing and classifying hand gestures include glove-based techniques and vision-based techniques for HCI and HAAC applications. In the glove-based techniques sensors are used to measure the positions of the fingers and the position of the hand in real-time. However, gloves tend to be quite expensive and the weight of the glove as well as the cables of the associated measuring equipment hinders free movement of the hand. The vision-based techniques can be divided into the three-dimensional (3-D) and the two-dimensional (2-D) approaches and usually glove-free. In the 3-D approach, gesture classification is based upon the parameters of a 3-D model of the human hand. Gesture classification is based upon the parameters of an image of the gesture in the 2-D approach. Because 3-D hand models are quite complex, the classification of gestures from parameters derived from 3-D models is computationally extensive making real-time classification difficult. The 2-D models are relatively less complex than the 3-D models. However, 2-D models do not carry the finger movement and finger position information required for the classification of complex dynamic gestures. Therefore, the 2-D approach is restricted to the less complex problem of classifying well-defined static gestures. References contain notable work in addressing issues and developing methodologies to solve problems related to vision based techniques.

II. SYSTEM DESCRIPTION AND OPERATIONAL FLOWCHART

The VHGC system is shown in Fig. 1. The relatively simple system consists of a single video camera to acquire an image of a gesture and an HP 586-200 MHz personal computer for processing and classifying the gesture. The Video Savant real-time video software for Windows NT is used for the real-time acquisition, sampling, quantizing, and storing of the gestures. Static hand gestures are represented by their contours (discrete boundaries) and the similarity between gestures is determined by measuring the similarity between contour representations. An operational flowchart of the processing steps, which will be described in the subsequent sections, is shown in Fig. 2. The algorithms for the segmentation, filtering, representation, and classification steps were developed using the C++ programming language. In operation, hand gestures are formed between the front of the video camera and a uniform background in a laboratory with florescent lights in the ceiling. No additional illumination sources are used. The individuals, however, are instructed to keep the hands approximately parallel to the camera lens in order to maintain the gesture shape.

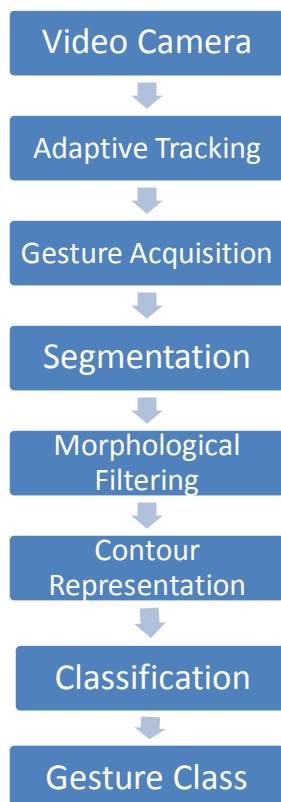


Fig. 1: Flowchart of the process

A. Gesture Segmentation

The goal of gesture segmentation is to extract the hand gesture as accurately as possible from the background of the image. That is, the segmented hand gesture should not have parts of the background due to under segmentation nor should it have parts of the hand deleted due to over segmentation. In general, the selection of an appropriate segmentation algorithm depends largely on the type of images and the application areas. Because the laboratory environment (lighting conditions) of the VHGC system is fixed, an autonomous segmentation algorithm gives good segmentation results in the laboratory environment and is also computationally simple can be selected.

B. Morphological Filtering

A close examination of the segmented gesture images revealed that the segmentation was seldom perfect. The background may have 1s (background noise) and the gesture may have 0s (object noise). The background noise and object noise can cause problems in extracting the contour of the gesture, especially when they are close to the contour. It is, therefore, desirable to decrease the background noise and object noise prior to the extraction of the gesture contour. A morphological filtering approach using a sequence of dilation and erosion operations was developed to obtain a smooth, closed, and complete contour of a gesture.

C. Contour Representation

A careful examination of the filtered gestures reveals that what distinguishes one gesture form another gesture is the shape of the contour. Therefore, the contour can be used as a basis for the discrimination of the hand gestures. The localized contour sequence (LCS), which has been proven to be a very effective representation of contours [10], is selected to represent the gesture contours. A contour tracking algorithm is used to track the contour of a gesture in the clockwise direction and the contour pixels are numbered sequentially starting from the arbitrarily selected contour pixel.

D. Classification Results

The amplitude and duration normalized LCSs of the 100 gesture images in the database were computed. That is, each gesture was represented by its corresponding normalized LCS. The LCSs were normalized to have a duration equal to 516 which was the average duration of the 100 LCSs. LCSs of the ten gestures in Fig. 3. In order to robustly evaluate the performance, a random sampling approach was used to generate multiple classification trials. For each trial, a reference LCS for each class was randomly selected from the ten LCSs of the class.

The remaining nine LCSs formed the test set for each class. Therefore, the total number of LCSs tested in each trial was $(9 \times 10) = 90$.

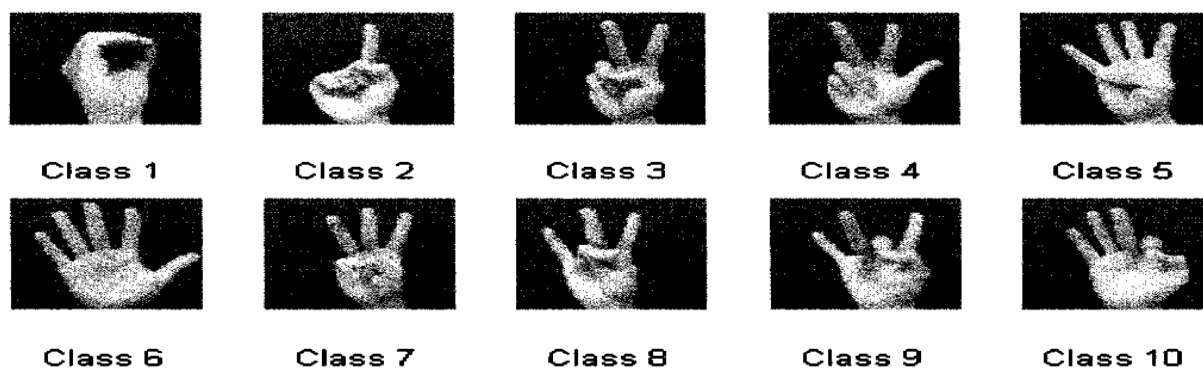


Fig. 3. Examples of one gesture from each gesture class

III. CONCLUSION

The goal of the paper was to develop a complete system capable of robustly classifying hand gestures for HCI and HAAC applications. From a visual analysis of hand gestures, it was determined that essential shape information for discriminating gestures was in the boundary of the gestures. Therefore, a contour and vision based classification approach was formulated. The relatively simple system consisted of a video camera, video capturing software, and a personal computer. For flexible operation, no constraint other than holding the gesture approximately parallel to the camera lens was imposed. The processing steps to classify a gesture included gesture acquisition, segmentation, morphological filtering, contour representation, and alignment based classification. Rather than forming an arbitrary set of gestures, the database for off-line evaluation consisted of the gestures for numbers 0 through 9 of the ASL. These gestures were selected because they are typical of the hand gestures that can be used for HCI and HAAC applications. The ten-class database consisted of ten example gestures for each class. The Otsu algorithm was selected to autonomously segment the gesture images and a morphological filtering approach was developed to remove background and object noise. The contour of a gesture was represented by the LCS and a linear alignment and a nonlinear alignment method were formulated to determine the similarity between two LCSs. The classification results showed that no misclassifications were obtained using nonlinear alignment even though the within-class variations were high because the gestures were formed by individuals not trained in ASL and with few constraints. The performance of the nonlinear alignment method was superior to that of the linear alignment method because, unlike linear alignment which simply uniformly expands or compresses the duration of a sequence, nonlinear alignment optimally aligns the samples of two sequences to minimize the dissimilarity between the sequences. Nonlinear alignment is clearly the better choice for classifying hand gestures because of the inherent nonlinear distortions that can be expected in the contours of the gestures.

IV. FUTURE SCOPE

This method can be used for one degree of accuracy which can be improved up to 0.1 degree or even less than that. So that while movement of fingers and hand gestures we can perform very small actions such as wiring connection etc. Hand gestures can be controlled by either using simple micro-controller or by the use of Arduino board. This will make easy for interfacing with it. For the tracking we have used MATLAB to track the different color but we can make it more precise and accurately measuring in real time by the use of Open CV for the proper application of real time processing.

REFERENCES

- [1] J. Lee and T. L. Kunii, "Model-based analysis of hand posture," IEEE Comput. Graph. Appl., pp. 77–86, Sept. 1995.
- [2] B. Moghaddam and A. Pentland, "Probabilistic visual learning for object recognition," IEEE Trans. Pattern Anal. Machine Intell., vol. 19, pp.696–710, July 1997.
- [3] T. Staner, J. Weaver, and A. Pentland, "Real-time American sign language recognition using desk and wearable computer based video," IEEE Trans. Pattern Anal. Machine Intell., vol. 20, pp. 1371–1375, Dec. 1998.
- [4] Y. Cui and J. Weng, "A learning-based prediction-and-verification segmentation scheme for hand sign image sequence," IEEE Trans. Pattern Anal. Machine Intell., vol. 21, pp. 798–804, Aug. 1999.
- [5] V. I. Pavlovic, R. Sharma, and T. S. Huang, "Visual interpretation of hand gestures for human computer interaction: A review," IEEE Trans. Pattern Anal. Machine Intell., vol. 19, pp. 677–694, July 1997.
- [6] D. J. Sturman and D. Zeltzer, "A survey of glove-based input," IEEE Comput. Graph. Appl., vol. 14, pp. 30–39, Jan. 1994.