



# A robust static hand gesture recognition system using geometry based normalizations and Krawtchouk moments

S. Padam Priyal\*, Prabin Kumar Bora

Department of Electronics and Electrical Engineering, Indian Institute of Technology Guwahati, Guwahati 781039, Assam, India

## ARTICLE INFO

### Article history:

Received 5 August 2011

Received in revised form

28 December 2012

Accepted 30 January 2013

Available online 8 February 2013

### Keywords:

Hand extraction

Hand gesture

Krawtchouk moments

Minimum distance classifier

Rotation normalization

Skin color detection

View and user-independent recognition

## ABSTRACT

Static hand gesture recognition involves interpretation of hand shapes by a computer. This work addresses three main issues in developing a gesture interpretation system. They are (i) the separation of the hand from the forearm region, (ii) rotation normalization using the geometry of gestures and (iii) user and view independent gesture recognition. The gesture image comprising the hand and the forearm is detected through skin color detection and segmented to obtain a binary silhouette. A novel method based on the anthropometric measures of the hand is proposed for extracting the regions constituting the hand and the forearm. An efficient rotation normalization method that depends on the gesture geometry is devised for aligning the extracted hand. These normalized binary silhouettes are represented using the Krawtchouk moment features and classified using a minimum distance classifier. The Krawtchouk features are found to be robust to viewpoint changes and capable of achieving good recognition for a small number of training samples. Hence, these features exhibit user independence. The developed gesture recognition system is robust to similarity transformations and perspective distortions. It can be well realized for real-time implementation of gesture based applications.

© 2013 Elsevier Ltd. All rights reserved.

## 1. Introduction

Human–computer interaction (HCI) is an important activity that forms an elementary unit of intelligence based automation. The very common HCI is based on the use of simple mechanical devices such as the mouse and the keyboard. Despite familiarity, these devices inherently limit the speed and naturalness of the interaction between man and the machine. The ultimate goal of HCI is to develop interactive computer systems that are non-obtrusive and emulate the ‘natural’ way of interaction among humans. The futuristic technologies in intelligent automation attempt to incorporate communication modalities like speech, hand writing and hand gestures with HCI. The development of hand gesture interfaces finds successful applications in sign-to-text translation systems, robotics, video/dance annotations, assistive systems, sign language communication, virtual reality and video based surveillance.

The hand gesture interfaces are based on the hand shape (*static gesture*) or the movement of the hand (*dynamic gesture*). The HCI interpretation of these gestures requires proper means by which the dynamic and/or static configurations of the hand could be

properly defined to the machine. Hence, computer vision techniques in which one or more cameras are used to capture the hand images have evolved. The methods based on these techniques are called *vision based* methods. The availability of fast computing and the advances in computer vision algorithms have led to the rapid growth in the development of vision based gestural interfaces. Many reported works on static hand gesture recognition have also focused in incorporating the dynamic characteristics of the hand. However, the level of complexity in recognizing the hand posture is comparatively high and recovering the hand shape is difficult due to variation in size, rotation of the hand and the variation of the viewpoint with respect to the camera.

The approaches to hand shape recognition are based on the 3D modeling of the hand or using 2D image models like the image contour and the silhouette. The computational cost in fully recovering the 3D hand/arm state is very high for real-time recognition and slight variations in the model parameters greatly affect the system performance [1]. By contrast, the processing of 2D image models involves low computational cost and high accuracy for a modest gesture vocabulary [1]. Thus, the 2D approaches are well pertinent for real time processing. The general approach to vision based hand shape recognition is to extract a unique set of visual features and match them to a pre-defined representation of the hand gesture. Therefore, the important factor in developing the gesture recognition system is the accurate representation of the hand shapes. This step is usually known as the *feature extraction* in pattern recognition.

\* Corresponding author. Tel.: +91 361 258 2502; fax: +91 361 258 2542.

E-mail addresses: s.priyal@iitg.ernet.in, padampriyal@gmail.com (S. Padam Priyal), prabin@iitg.ernet.in (P.K. Bora).

The features are derived either from the spatial domain or from the transform domain representation of the hand shapes. The extracted features describing the hand gestures can be classified into two groups: contour based features and the region based features. The contour based features correspond to the information derived only from the shape boundary. Common contour based methods that are used for shape description are Fourier descriptors, shape signatures, curvature scale space and chain code representations [2]. Hausdorff distance [3] and shape context [2] are correspondence based matching techniques in which the boundary points are the features representing the shape. The region based features are global descriptors derived by considering the entire pixels constituting the shape region. The common region based methods include moment functions and moment invariants, shape matrices, convex hulls and medial axis transforms [2]. Similarly, the spatial-domain measures like the Euclidean distance, the city-block distance and the image correlation are used for region based matching in which the pixels within the shape region are considered as features.

The efficiency of these features is generally evaluated based on the compactness in representation, the robustness to spatial transformations, the sensitivity to noise, accuracy in classification, low computational complexity and the storage requirements [2]. In this context, the moments based representations are preferred mainly due to their compact representation, invariance properties and the robustness to noise [4]. The moments also offer the advantages of reduced computational load and database storage requirements. Hence, the moment functions are among the robust features that are widely used for shape representation and find successful applications in the field of pattern recognition which involves archiving and fast retrieval of images [5].

Recently the discrete orthogonal moments like the Tchebichef moments and the Krawtchouk moments were introduced for image analysis [6,7]. It is shown that these moments provide higher approximation accuracy than the existing moment based representations and are potential features for pattern classification. Hence, in this work we have proposed the classification of static hand gestures using Krawtchouk moments as features. The objective of this work is to study the potentiality of the Krawtchouk moments in uniquely representing the hand shapes for gesture classification. Hence, the experiments are performed on a database consisting of gesture images that are normalized for similarity variations like scaling, translation and rotation. The performance of the Krawtchouk moments is compared with the geometric and the Zernike moments based recognition methods.

The other main issues considered in developing the gesture recognition system are the (i) identification of the hand region and (ii) normalization of the rotation changes.

The identification of the hand region involves separating the hand from the forearm. The lack of gesture information in the forearm makes it redundant and its presence increases the data size. In most of the previous works, the forearm region is excluded by either making the gesturers to wear full arm clothing or by limiting the forearm region into the scene while acquisition. However, such restrictions are not suitable in real-time applications. The orientation of the acquired gesture changes due to the angle made by the gesturer with respect to the camera and vice-versa.

This work concentrates on vision based static hand gesture recognition considering the afore-mentioned problems. In [8], the Krawtchouk moments are introduced as features for gesture classification. The performance of Krawtchouk moments is compared with that of a few other moments like the geometric, the Zernike and the Tchebichef moments. It is shown that the Krawtchouk moments based representation of hand shapes gives high recognition rates. The analysis is performed on hand regions that are manually extracted and corrected for rotation changes.

This work presents a detailed gesture recognition system that evaluates the performance of the Krawtchouk moment features on a database that consists of 4230 gesture samples. We propose novel methods based on the anthropometric measures to automatically identify the hand and its constituent regions. The geometry of the gesture is characterized in terms of the abducted fingers. This gesture geometry is used to normalize for the orientation changes. These proposed normalization techniques are robust to similarity and perspective distortions. The main contributions in this work are:

1. A rule based technique using the anthropometric measures of the hand is devised to identify the forearm and the hand regions.
2. A rotation normalization method based on the protruded/abducted fingers and the longest axis of the hand is devised.
3. A static hand gesture database consisting of 10 gesture classes and 4230 samples is constructed.
4. A study on the Krawtchouk moment features in comparison to the geometric and the Zernike moments for viewpoint and user invariant hand gesture recognition is performed.

The rest of the paper is organized as follows: Section 2 presents a summary of the related works in static hand gesture recognition. Section 3 gives the formulation for Krawtchouk and other considered moments. Section 4 provides an overview of the proposed gesture analysis system in detail. Experimental results are discussed in Section 5. Section 6 concludes the paper mentioning the scope for future work.

## 2. Summary of the related works

The primary issues in hand gesture recognition are: (i) hand localization, (ii) scale and rotational invariance and (iii) viewpoint and person/user independence. Ong and Ranganath [9] presented a thorough review on hand gesture analysis along with the insight into problems associated with it.

Earlier works assumed the gestures to be performed in a uniform background. This required a simple thresholding technique to obtain the hand silhouette. For a non-uniform background, skin color detection is the most popular and the general method for hand localization [1,10–14]. The skin color cues are combined with the motion cues [10,12,13] and the edge information [14] for improving the efficiency of hand detection. Segmented hand images are usually normalized for the size, orientation and illumination variations [15–17]. The features can be extracted directly from the intensity images or the binary silhouettes or the contours.

In [18,19], the *orientation histograms* are derived from the intensity images. These histograms represent summarized information on the orientations of the hand image and are shown to be illumination invariant. They are however, rotation variant. Triesch et al. [20] have classified hand postures using *elastic graph matching*. The system is designed to efficiently identify gestures in a complex background. It is sensitive to geometric distortions that arise due to the variations in the hand anatomy and the viewpoint. The advantages of the matching procedure are its user independence and the robustness to complex environments. In [1,15], *local linear embedding* (LLE) is introduced to map the high dimensional data to a low dimensional space in such a way as to preserve the relationship between neighboring points. Each point in the low dimensional space is approximated by a linear combination of their neighbors. The approach is invariant to scale and translation but sensitive to rotation. Just et al. [21] used the

Download English Version:

<https://daneshyari.com/en/article/531071>

Download Persian Version:

<https://daneshyari.com/article/531071>

[Daneshyari.com](https://daneshyari.com)