Feature Extraction for Hand Gesture Recognition : A Review

Sana'a Khudayer Jadwaa

Abstract—Body language is an important way of communication among humans, adding emphasis to voice messages or even being a complete message by itself. Human hand has remained a popular choice to convey information in situations where other forms like speech cannot be used. Hand gestures which can represent ideas using unique shapes and finger orientation have a scope for human machine interaction. 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. The performance of a general recognition system first depends on getting efficient features to represent pattern characteristics. There are several methods of representing such gesture trajectory feature. The aim of this paper is to show different features for hand's image used with different approaches that yielding arobust and reliable hand gesture recognition.

Index Terms— Computer vision ,Features for Image,Gestures ,Gesture Recognition,Hand Gesture Recognition,Hand Image Features, , Human Computer Interaction

1 INTRODUCTION

WITHIN the last years, in the field of "Computer Vision", considerable resources have been allocated in the developing face and hand recognition techniques. Being able to recognize hand and face region from static images or video sequences gives us an advantage in application such as teleconferences, telemedicine and in the field of human-computer interaction devices. In particular, the hand gestures are an attractive way of interacting with such systems because these are the natural ways of communication. Another advantage is that the user does not have to learn how to manipulate certain specialized hardware devices[1].Vision-based hand gesture recognition is an active area of research in human-computer interaction (HCI), as direct use of hands is a natural means for humans to communicate with each other and more recently, with devices in intelligent environments[2].Hand gestures involve relative flexure of the user's fingers and consist of information that is often too abstract to be interpreted by a machine[3].Hand gestures provide a separate complementary modality to speech for expressing ones ideas[4]. Information associated with hand gestures in a conversation is degree, discourse structure, spatial and temporal structure, So a natural interaction between humans and computing devices can be achieved by using hand gestures for communication between them. The key problem in gesture interaction is how to make hand gestures understood by computers [4].

Dr. Sana'a khudayerJadwaa
Ph.D inComputer Sciences
Lecturer in Computer Unit, College of Medicine
Al-Nahrain University
Baghdad- Iraq
E-mail: sanaakhudayer@yahoo.com

Gestures are powerful tools of communication among humans and are of particular interest to the deaf and dumb community. The recognition problem is solved through a matching process in which the segmented hand is compared with all the Images in the database. The system's memory stores all the images, distance transform and all information required for comparison[5]. This paper survey recent approaches that are used various features for recognition of static hand gestures. The organization of the rest of this paper is as follows. Section 2 highlights the general gesture recognition system. Section 3 introduces features needed to represent the hand gesture . Section 4 describes various application areas of hand gesturerecognition. Section 5 concludes the paper.

2 GESTURE RECOGNITION SYSTEM

The general gesture recognition process in any kind of system can be broken down into the following components shown in figure (1) .The first stage in the figure, is mostly concerned with the hardware side of the system and how data for the recognition process is gathered (in the form of bitmaps or lists of vertices). The second stage is a pre-processor stage. Here edge-detection can take place as well as smoothing and other filtering processes. This prepares the data for the main computational stage, the feature extraction. Some systems might never use the term feature, but somewhere along the line, they will find a way of quantifying their input. The features of the input are then evaluated in one or more of several possible ways to make a decision about which gesture the system is most likely subjected to in the fourth stage, the evaluation stage. All systems will have a limited set of gestures that they can recognize at any given time [6].

ISSN 2229-5518

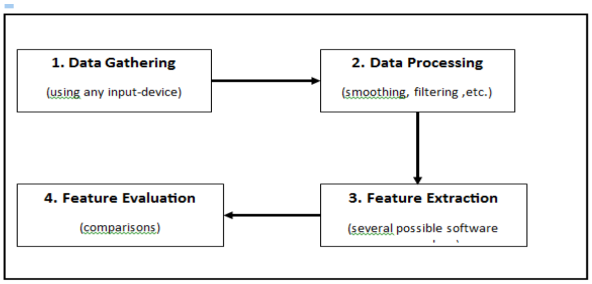


Fig.1. A Generic Gesture Recognition System

3 Feature Extraction

In pattern recognition and in image processing, feature extraction is a special form of dimensionality reduction. When the input data to an algorithm is too large to be processed and it is suspected to be notoriously redundant (e.g. the same measurement in both feet and meters) then the input data will be transformed into a reduced representation set of features (also named features vector). Transforming the input data into the set of features is called feature extraction. If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input.[7].For any computer vision system the notion of a "feature" is important. Computer vision systems extract features from their input data and produce useful information from those features[8]. Castleman [9] defines a feature as follows:

"A feature is a function of one or more measurements, computed so that it quantifies some significant characteristic of the object."

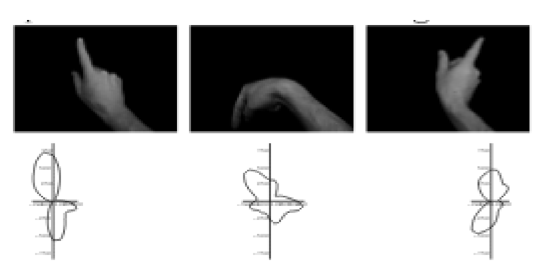
This very general definition does not specify how exactly a feature is computed nor if it describes the whole object to be measured or a part of it[8]. This distinction is made in a further separation into global features and local features. Global features quantify characteristics of the whole image to be measured. An example would be a color histogram of an image which gives important information about the whole image.1 Local features quantify characteristics of a particular region of the object to be measured[8].When selecting features for use in a computer imaging application, an important factor is the robustness of a feature. A feature is robust if it will provide consistent results across the entire application domain. For example, if we are developing a system to work under any lighting conditions, we do not want to use features that are lighting dependent they will not provide consistent results in the application domain. Another type of robustness ,especially applicable to object features, is called RST-invariance, where the RST means rotation, size, and translation. A very robust feature will be RST-invariant, meaning that if the image is rotated, shrunk or enlarged, or translated(shifted left/ right or up/down), the value for the feature will not change[10].

3.1 Feature Representation

A specific image feature, defined in terms of a specific structure in the image data, can often be represented in different ways. For example, an edge can be represented as a boolean variable in each image point that describes whether an edge is present at that point. Alternatively, we can instead use a representation which provides a certainty measure instead of a boolean statement of the edge's existence and combine this with information about the orientation of the edge. Similarly, the color of a specific region can either be represented in terms of the average color (three scalars) or a color histogram (three functions)[7].When a computer vision system or computer vision algorithm is designed the choice of feature representation can be a critical issue. In some cases, a higher level of detail in the description of a feature may be necessary for solving the problem, but this comes at the cost of having to deal with more data and more demanding processing[7].

3.2 Feature Vectors and Feature Spaces

A feature vector is one method to represent an image, or part of an image(an object), by finding measurements on a set of features. The feature vector is an n-dimensional vector that contains these measurements[10]. In some applications it is not sufficient to extract only one type of feature to obtain the relevant information from the image data, instead two or more different features are extracted, resulting in two or more feature descriptors at each image point. A common practice is to organize the information provided by all these descriptors as the elements of one single vector, commonly referred to as a feature vector. The set of all possible feature vectors constitutes a feature space[7].A commonexample of feature vectors appears when each image point is to be classified as belonging to a specific class. Assuming that each image point has a corresponding feature vector based on a suitable set of features, meaning that each class is well separated in the corresponding feature space, the classification of each image point can be done using standard classification method. Another ,and related example, occurs when neural network based processing is applied to images.The input data fed totheneural network is often given in terms of a feature vector from each image point, where the vector isconstructed from severaldifferent feature extracted from the image data. During a learning phase, the networks can itself find which combinations of different features that are useful for solving the problem at hand[7].



3.3 Features for Hand Gesture

ISSN 2229-5518

The aim of feature extraction phase is to find and extract features that can be used to determine the meaning of a given gesture. Ideally such a feature, or a set of such features, should uniquely describe the gesture in order to achieve a reliable recognition. Therefore, different gestures should result in different, good discriminable features. Furthermore, shift and rotation invariant features lead to a better recognition of hand gestures even if the hand gesture is captured in a different angle[11].Some of these features are presented as follows :

3.3.1. Zernike Moments (ZMs)

In [12] they utilized ZMs(Zernike moments) & PZMs (Pseudo Zernike moments) descriptors for obtaining a quantitative shape feature description of the hand for recognizing static gestures[12]. Zernike moments are based on a set of complex polynomials that form a complete orthogonal set over the interior of the unit circle . They are defined as the projection of the image on these orthogonal basis functions [13]. In general, ZMs are better in describing shapes than PZMs, whereas PZMs are less affected by noise than ZMs. The binary hand silhouette is first accommodated with a minimum bounding circle (MBC). The binary hand silhouette is then decomposed into the finger part and the palm part by morphological operations according to the radius of the MBC. After that, the ZMs & PZMs of the finger part and the palm part with different importance, respectively, are computed based on the center of the MBC. Figure (2-a) shows an example of an MBC that accommodates the inside of the binary hand silhouette; Figure (2-b) shows the finger part and palm part of the binary hand silhouette[12].

(a) (b)

Fig.2. (a) MBC that accommodates the binary hand silhouette inside and (b) finger part and palm part of the binary hand silhouette

3.3.2. Local orientation histogram

In[14] they employing histograms of local orientation, they use the orientation histogram as a feature vector for gesture classification and interpolation as shown in figure(3). This method is simple and fast to compute, and offers some robustness to scene illumination changes. The orientation analysis gives robustness to lighting changes; the histogramming gives translational invariance. The advantage of a histogram is that it can be calculated quickly. This simple method will work if examples of the same gesture map to similar orientation histograms, and different gestures map to substantially different histograms.

Fig.3.The hand gestures with its corresponding orientation histograms.

3.3.3. Local Brightness

In [15] the local brightness of each block of the gesture image is used , the input gesture image is divided into 25x25 blocks each of 5x5 block size, and they have calculated the local brightness of each divided block after applying colored segmentation operation using HSV (Hue, Saturation and Value) color model as shown in figure(4) , so, each gesture produces 25x25 feature values called feature vector .After the preparing of the image and segmenting of the hand gesture, a black-white image is created and represented the hand pose inset, the feature extraction phase will start, the overall feature vector size is 625 elements which are the brightness values of each block in the gesture, these features are stored in the database.

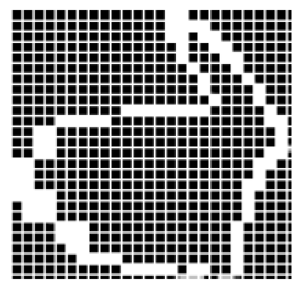


Fig. 4. Features Calculation via Dividing the Gesture.

3.3.4. Binary Object Features

In[16] the height, area, centroid, and distance of the centroid from the origin (top-left corner) of the binary image are used as features. To find the height of a sign from binary image where 0's represent black pixels and 1's represent white pixels, they have used pixel scanning method .The average height of the binary image is calculated as :

$$Heightavg = N/M \qquad (1)$$

Where N is the total number of black pixels (0's) of the image and M is the total number of columns containing at least one black pixel (0). The centroid of a polyhedron is simply the average of the respective coordinates of all the vertices of the polyhedron .For example, if the coordinates are (x1, y1) (x2, y2) ... (xN, yN), then the centroid would be :

$$xc = (x1 + x2 + \cdots + xN)/N \qquad (2)$$

$$yc = (y1 + y2 + \cdots + yN)/N \qquad (3)$$

then the euclidian distance of the centroid from the origin is calculated , Finally, the features collected from the above then they all combined to form a feature vector in the following order: Feature vector, V= [area, average height, x-centroid, ycentroid,centroid-distance].

4 Applications Directions

Gesture recognition has wide-ranging applications [17] such as the following:

_developing aids for the hearing impaired;
_ enabling very young children to interact with computers;
_ designing techniques for forensic identification;
_ recognizing sign language;
_ medically monitoring patients' emotional states or stress levels;
_ lie detection;
_ navigating and/or manipulating in virtual environments;
_ communicating in video conferencing;
_ distance learning/tele-teaching assistance
_ monitoring automobile drivers' alertness/drowsiness levels, etc.

5CONCLUSION

In this paper we have given an idea of static hand gesture recognition with focus on the feature extraction stage and some of hand gesture recognition applications. We have also present some researches that used different hand features for recognition process that are make the recognition process more accurate. The hand gesture recognition process is still in its first steps so it is clear that further research in the areas of feature extraction, classification methods and gesture representation are required to realize the ultimate goal of humans interfacing with machines on their natural terms.

6 REFERENCES

[1] B. Ionescu, D, Coquin, “Static Hand Gesture Recognition System”, U.P.B.Sci. Bull., Series D, Vol. 66, No. 2-4, 2004.

[2] P.Vijaya Kumar1 ,N.R.V.Praneeth and Sudhee, “Hand And Finger Gesture Recognition System for Robotic Application”,International Journal of Computer Communication and Information System (IJCCIS) Vol2. No1. ISSN: 0976–1349 July – Dec 2010

[3] M. Chun,W. Jean and H. Chang,” Static Hand Gesture RecognitionSystem Using a Composite Neural Network”, Department of Electrical Engineering,Tamkang University, Taiwan,R.O,C. IEEE 1996.

[4] P.Garg, N. Aggarwal, and S.Sofat,”Vision Based Hand Gesture Recognition” World Academy of Science, Engineering and Technology49, 2009.

[5] G. Thomas, “A Review of Various Hand Gesture Recognition Techniques”,International Journal of Electrical, Electronics &CommunicationEngineering, VSRD-IJEECE, Vol. 1 (7), 2011 .

[6] S. khudayer ,”Hand Gesture Recognition Algorithm”, Ph.D thesis, University of Technology,2009.

[7] Wikipedia internet Web Site

[8] K. R. Castleman,” Digital Image Processing”, Prentice Hall, Englewood Cliffs, NJ, USA,1996

[9] S. Nowozin, “Object Classification using Local Image Features”, TechnischeUniversit¨at Berlin Sebastian Nowozin May 8, 2006

[10] S.Umbaugh ,”Computer Vision and Image Processing a practical approach using CVIP tools”, Prentice Hall PTR, NJ, USA, 1998.

[11] T.Messer , “Static Hand Gesture Recognition”, Report ,Department of Information ,University of Fibourg ,Switzerland , 2006.

[12] C. Chang, J. Chen, W.Tai, “New Approach for Static GestureRecognition”,Jornal of Information Science and Engineering, Vol 22,pp 1047- 1057,2006.

[13] G. Amayeh, G.Bebis, A. Erol,“Peg-Free Hand Shape Verification UsingHigh Order Zernike Moments”, Computer Vision aboratory, University of Nevada, Reno, NV 89557-2008 .

[14] W. Freeman, M. Roth, “ Orientation Histograms for Hand GestureRecognition”, IEEE Intl. Wkshp. on Automatic Face and Gesture Recognition, Zurich, June, 1995.

[15] M. M. Hasan , P. Misra,” HSV Brightness Factor Matching for Gesture Recognition System” , International Journal of Image Processing (IJIP), Volume (4): Issue (5). 2010.

[16] A. Rahman, A. Ul-Ambia and Aktaruzzaman,” Recognition Static Hand Gestures of Alphabet in ASL” , IJCIT, ISSN 2078-5828 (PRINT), ISSN 2218-5224 (ONLINE), VOLUME 02, ISSUE 01, MANUSCRIPT CODE: 110749 , 2011.

[17] S.Mitra, T. Acharya, “ Gesture Recognition :A Survey”, IEEE Transactionson Systems, Man,and Cybernetics—Partc: Applications and Reviews, Vol .37 No.3, May 2007.