Hand Gesture Recognition for the Application of Sign Language Interpretation

Resmi George, K Gerard Joe Nigel,

Karunya University,

Coimbatore

Abstract - **Classifying human hand gesture in the context of sign language has been historically dominated by artificial neural networks with varying degree of success. The main objective is to introduce moment invariants as the feature as an alternative machine learning method for hand gesture interpretation in sign language. The features are found to be robust to translation, rotation, scaling, noisy images and capable of achieving good recognition for a small number of training samples. Hence, these feature exhibit user independence. It can be well realized for real-time implementation of gesture based application.**

1. INTRODUCTION

About 22 million people in the world are deaf as well as dumb. The disables group use sign language for communication. [9]Sign language is important in humankind is that showing an increasing research in eliminating barriers faced by differently abled person in with the society. A functioning sign language recognition system will provide opportunity for a mute person to communicate with common people without the need for an interpreter. With the use of Neural Networks in image processing the input image is compared with set of images in the dataset the word corresponds to the matched image will gives the output. Thus human can easily interact with computer.

In earlier there are several method used to implement sign language recognition. Among them Gloves based devices are connected to the main processor with cables that restrict the user's natural ability to communicate. Most of these approaches have been implement to focus on single aspect of gestures such as hand tracking and posture estimation or hand poses. And Classification using uniquely coloured (RGB) gloves or markers on hands or fingers [5,6].Here we are extracting the idea of consumer electronics controlled by hand gesture recognition system which makes human computer interaction possible. Some of the drawbacks of exciting system are most of the system are relied on Template matching [10] and shape descriptors and required more processing time, the user were restricted to wearing gloves[8,9] or markers to increase reliability, and are also required at specific distance from the camera. This paper is distinguished from previous attempts by a few marked differences:

[1] Hand tracking that isolates region of interest (ROI).
[2] Distance from the hand and camera are immaterial.
[3] Extracted Feature is Moment Invariants.

2. SYSTEM ARCHITECTURE

The system camera initially captures a frame every second. In order to identify the hand gesture, skin-like region is segmented. Since RGB domain is found to be helpless in achieving skin segmentation, image is initially converted to binary and later to YCbCr domain. Even though domain has been changed it may result in a noisy image. However noise can be removed during the gesture normalization stage which would use morphological filtering technique using erosion and dilation. The output of this stage is a smoothen region of the hand gesture, which is stored as a logical bitmap image.
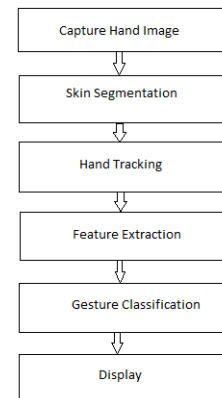

Fig 1

2.1 Hand Region Segmentation

When part of the arm is captured along with the hand as show in the Fig.2, there needs to be further processing to isolate the hand removing arm region from the hand region for effective hand gesture recognition.

Fig [2]. ROI   Fig [3]. Segmentation

3. HAND TRACKING

A user who is giving a command to a hand gesture system may move his/her hand knowingly or unknowingly during the course of the gesture or gestures. It is important for an effective system to track the hand so that finding the ROI will be fast and error free.

Calculating the boundary of hand and determine the centroid point of hand region. Through iteration of hand tracking process, we can obtain the motion trajectory of the hand so-called gesture path from connecting hand centroid points set. When we obtain the hand location from hand tracking procedure, it will enter the state of check start point firstly. If the hand is no motion then it takes the next one, else this point is start point and the pure path is begin recorded. Then, it will stay at the state of path recording until the hand location is not moving .Finally we will enter the state of check end point and obtain the pure gesture path for following recognition.

Fig4.Hand Tracking

4. FEATURE EXTRACTION

There have been many attempts to use a variety of feature for gesture classification. Template matching a moment invariants based feature extraction methods to evaluate their suitability for gesture classification. Image classification is a very mature field today. There are many approaches to finding matches between images or image segments. Starting from the basic correlation approach to the scale-space technique, variety of feature extraction methods with varying success. The moment invariants algorithm has been recognized as one of the most effective methods to extract descriptive feature for object recognition applications and has been widely applied in classification of subjects such as aircrafts, ship, and ground targets. These properties are invariant to rotation, scale, and translation. Let $f(i,j)$ be a point of a digital image of size MxN(i= 1,2,.....,M and j= 1,2,...,N).The 2 D moments and central moments of the order $(p+q)$ of $f(i,j)$ are defined as

$$m_{pq} = \sum_{i=1}^{M} \sum_{j=1}^{N} i^p j^q f(i,j) \text{ and}$$

$$u_{pq} = \sum_{i=1}^{M} \sum_{j=1}^{N} (i - T(i)^p)(j - T(j))^q f(i,j)$$

Where $i = m_{10}/m_{00}$, $j = m_{01}/m_{00}$ from the 2nd order and 3rd order moments,a set of seven moment invariants are derived as follows:

$$\varphi_1 = \eta_{20} + \eta_{02} \quad (6.1)$$

$$\varphi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11} \quad (6.2)$$

$$\varphi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (6.3)$$

$$\varphi_4 = (\eta_{30} - \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (6.4)$$

$$\varphi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} - \eta_{12})^2 -$$
$$(3\eta_{21} - \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})$$
$$[(3\eta_{30} - \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (6.5)$$

$$\varphi_6 = (\eta_{02} + \eta_{20})[(\eta_{30} - \eta_{12})^2 -] + 4\eta$$
$$(\eta_{21} - \eta_{03})^2 (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})$$
$$(6.6)$$

$$\varphi_7 = 3(\eta_{21} + \eta_{03})(\eta_{30} - \eta_{12})[(\eta_{30} - \eta_{12})^2 -$$
$$3(\eta_{30} + \eta_{21})] - (\eta_{02} + \eta_{20})(\eta_{30} - 3\eta_{12})$$
$$[3(\eta_{30} - \eta_{12})^2 - (\eta_{21} - \eta_{03})^2] \quad (6.7)$$

Where $\eta_{pq} = u_{pq}/u_{00}^r$

$r = [(p+q)/2]+1$ and $p+q = 2,3...$

Moments	Figure1	Figure2	Figure3
U1	0.19051	0.1950	0.19729
U2	0.00111	0.0032	0.00199
U3	0.00167	0.00146	0.00276
U4	0.000287	0.00034	0.000570
U5	1.922210	2.2657e-7	7.143e-07
U7	1.7310e-8	1.2424e-7	2.454e-07

Fig 5. moment invariants of different orientation of sign

The parameters are almost same even the sign is scaled, rotated and translated.

5. GESTURE CLASSIFICATION

5.1 Neural Networks

Neural networks has been apply to perform complex functions in numerous application includes: pattern recognition, classification, identification etc..Once implemented, it can compute the output significantly faster than nearest neighbour classifier. Neural networks also have the ability to learn and predict over the time. This

property enables the system to be viewed more as a human-like entity that can actually 'understand' the user, which is also one of the major objectives of our project. This system is also design to capture 1 image frame every 100ms and is then segmented for skin region detection and other pre-processing before the invariant moments are calculated. If any of the static images is captured when the hand is moving, resultant image would be blurred. This will results in an unrecognized hand gesture. The designed neural network is 'back propagation' network in which input vectors (invariant moments of the sample set) and the output target vectors are used to train the network until it can approximate the function between the input and the output.
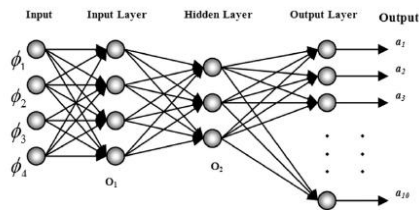

Fig6.Structure of neural networks

6. RESULT AND DISCUSSIONS

The result of the proposed system has been very encouraging. A user may start the gesture recognition process by keeping the hand displaying the palm to camera so that the whole hand in the middle of the frame. The system then captures the ROI as the rectangle encompassing the hand and track the hand. The feature is calculated from the tracked region.

*Moment invariants are invariant to rotation, translation, and scaling.

*The method is susceptible to noise. Most of the noise is filtered at the gesture normalisation process.

*The system is easy to implement and only an insignificant computational effort from CPU.

*The first four moments $\phi_1, \phi_2, \phi_3, \phi_4$ are adequate to represent a gesture uniquely and hence result in a simple feature vector with only four values.

6. REFERENCE

1. SaboohAjaz b, Prashan Premaratne a, Malin Premaratne c 'Hand getsure tracking and recognition system using Lucas-Kanade algorithms for control of consumer electroics ',SciVerse Science Direct,Neurocomputing 116(2013) 242-249

2. Prabin Kumar Bora, S. Padam Priyal "A robust static hand gesture recognition system using geometry based normalization and Krawtchouk moments" Science Direct, Pattern Recognition 46 (2013) 2202-2219

3. Sigal Berman , Helman Stern, Merav Shmueli "Most discriminating segment longest common subsequence algorithm for dymanic hand gesture classification" Pattern Recognition (2013)

4. V .Radhab, M.Krishnavenia "Classifier fusion based on Bayes aggregation method for Indian sign language recognition datasets" ProcediaEngineering 30 (2012) 1110-1118

5. John Mc Donald,Daniel Kelly "Weakly supervised traing of a sign language recognition system using multiple instance lerning density marices" ieee trans on system, man , and cybernetics, vol 41. No.2 April 2011

6. Frank M.CIaramello,Sheila a.Hemamai "A computational intergrity model for assessment and compression of American sign language video" ieee trans on image processing ,vol .20,no.11,nov 2011

7. Michael Greenspan, Hong Li a, "Model based segmentation and recognition of dynamic gesture in continues video stream" pattern recognition 44 (2011) 1614-1628

8. George caridakis,Athenesios Drosopoulos" SOMM : self organising markov map for gesture recognition" Pattern Recognition letter 31 (2010) 52-59