Static Hand Gesture Recognition based on DWT Feature Extraction Technique

Rajashree Patil
PG Student
Department of Electronics and Telecommunication
RSCOE, Savitribai Phule Pune University

Shailaja Patil
Professor
Department of Electronics and Telecommunication
RSCOE, Savitribai Phule Pune University

Abstract

Every day number of images are generated which implies the necessity to classify, organise and access them using an easy, faster and efficient way. The classification of images into semantic categories is a challenging and important problem now days. Thus hand gesture image detection and recognition is also difficult task. In this paper a novel system which can be used for sign language recognition and interaction with an application or videogame via hand gestures is introduced. This proposed system elucidate framework for vision based recognition of American sign language alphabets from static hand gesture images using Discrete Wavelet Transform based feature extraction technique and Minimum distance vector classifier Daubenchies Wavelet basis function is selected for performing three level two dimensional wavelet decomposition of input static hand gesture image and 7 features are extracted for each image. In the training stage feature vectors are calculated for every training images creating database for classification. In the testing stage input image features are compared with stored trained image features using minimum distance vector classifier.

Keywords: Discrete wavelet transforms (DWT), hand posture, human computer interaction, object detection, object recognition, Scale invariant feature transform (SIFT)

I. INTRODUCTION

Gesture is defined as a meaningful physical movement of the fingers, hands, arms, or face with the purpose to convey information for human computer interaction. Gestures can be static (posture) or dynamic (sequence of postures). Static gestures require less computational complexity rather than dynamic gestures which are complex and for that it is suitable for real time environments. Gesture recognition, needs a good interpretation of the hand movement so as to have effectively meaningful commands. Gesture recognition is the process by which gestures made by the user are made known to the system. For recognizing the gestures, there are several algorithms that are available. For Human Computer interpretation systems there are two commonly approaches.

First is Data glove approach. These methods employs mechanical or optical sensors attached to a glove that transforms finger flexion's into electrical signals to determine the hand posture. Using this method the data is collected by one or more data- glove instruments which have different measures for the joint angles of the hand and degree off freedom (DOF) that contain data position and orientation of the hand used for tracking the hand. However, this method requires the glove must be worn and a wear some device with a load of cables connected to the computer, which will hampers the naturalness of user-computer interaction. Second approach is vision based approach. These techniques based on the how person realize information about the environment. These methods usually done by capturing the input image using camera(s). In order to create the database for gesture system, the gestures should be selected with their relevant meaning and each gesture may contain multi samples for increasing the accuracy of the system. Vision Based hand gesture recognition approaches can be 1) Appearance Based approaches. These approaches use features extracted from visual appearance of the input image model the hand, comparing these modeled features with features extracted from input camera(s) or video input. 2) 3D Model Based approaches these Model based approaches depend on the kinematic hand DOF's of the hand. These methods try to infer some hand parameters like, pose of palm, joint angles from the input image, and make 2D projection from 3D hand model economical algorithms.

Hand Gestures offer a natural and intuitive communication modality for human–computer interaction. Efficient human computer interfaces (HCIs) need to be developed to allow computers to visually acknowledge in real time hand gestures. However, vision-based hand gesture recognition is challenging problem due to complexness of hand gestures, which are wealthy in diversities due to high degrees of freedom (DOF) concerned by the human hand. In order to successfully fulfill their role, the hand gesture HCIs has to meet in terms of time period performance, recognition accuracy, transformations and cluttered background. To satisfy these needs, many gesture recognition systems used the assistance of colored markers or information gloves to form the task easier [1]. However, using of markers and gloves restrict the user's convenience. In this paper, we concentrate on bare hand gesture recognition without any markers and gloves.

According to the psycho visual studies human visual system processes an image in a multi scale manner. The multi resolution representation gives a hierarchical framework for analyzing the information content of images at various resolutions, to get the

details of different physical structures of the image (scene). These details analyzed at different resolutions are regrouped into a pyramidal structure called a pyramidal transform. To achieve this many types of techniques were developed, including wavelets, Gaussian, and Laplacian pyramids [2]. The use of Wavelet Transform (WT) as a framework of multi resolution signal decomposition for texture description was first suggested by Mallat [3]. The WT can be designed as a pyramid or a tree structure. However, pyramidal algorithm has down sampling which in turn saves a large amount of computational time.

This paper is organized as follows: section II represents the literature survey over the various methods presented for hand gesture image recognition and detection; Section III describes the proposed approach and its system block diagram is depicted. In section IV we are presenting the system implementation and results achieved. Finally is described in section V.

II. LITERATURE SURVEY

This section we are presenting the different methods for hand gesture recognition &its related Work. Sebastiean Marcel, Oliver Bernier, Jean Emmanuel Viallet and Danieal Collobert have proposed the same using Input-output Hidden Markov Models [4]. Xia Liu and Kikuo Fujimura have proposed the hand gesture recognition using depth data [5]. A method is presented for recognizing hand gestures by using a sequence of real-time depth image data acquired by an active sensing hardware. Hand posture and motion information extracted from a video is represented in a gesture space which consists of a number of aspects including hand shape, location and motion information. For hand detection, many approached uses color or motion information [6, 7]. Attila Licsar and Tamas Sziranyi have developed a hand gesture recognition system based on the shape analysis of the static gesture [8], for human computer interaction purposes. Our appearance-based recognition uses modified Fourier descriptors for the classification of hand shapes. As always found in literature, such recognition systems consist of two phases: training and recognition. In our new practical approach, following the chosen appearance-based model, training and recognition is done in an interactive supervised way: the adaptation for untrained gestures is also solved by hand signals. In [9] B. Stenger presented a practical method for hypothesizing hand locations and subsequently recognizing a discrete number of poses in image sequences. In a typical setting the user is gesturing in front of a single camera and interactively performing gesture input with one hand. The approach is to identify likely hand locations in the image based on discriminative features of color and motion. A set of exemplar templates is stored in memory and a nearest neighbor classifier is then used for hypothesis verification and poses estimation. Byung-Woo Min, Ho-Sub Yoon, Jung Soh, Yun-Mo Yangc and Toskiaki Ejima have suggested the method of Hand Gesture Recognition using Hidden Markov models [10].

In this paper consideration is a planar hand gesture in front of a camera and use 8-directional chain codes as input vectors. For training an HMM network, a simple context modeling method is embedded for training on a “left-to-right” HMM model. This model is applied to drawing and editing specified graphic elements. The overall objective is to recognize 12 different dynamic gestures. Another very important method is suggested by Meide Zhao, Francis K.H. Quek and Xindong Wu [11]. They have used AQ Family Algorithms and R-MINI Algorithms for the detection of Hand Gestures. A recursive induction learning scheme in the RIEVL algorithm is designed to escape local minima in the solution space. There is another efficient technique which uses Fast Multi-Scale Analysis for the recognition of hand gestures as suggested by Yikai Fang, Jian Cheng, Kongqiao Wang and Hanqing Lu [12], but this method is computationally expensive. Chris Joslin etc. all have suggested the method for enabling dynamic gesture recognition for hand gestures [13]. Rotation invariant method is widely used for texture classification and recognition. Timi Ojala etc. all have suggested the method for texture classification using Local Binary Patterns [14].

III. PROPOSED APPROACH FRAMEWORK AND DESIGN

A. Aim of Proposed System

The main objective of this work is to design &implement new approach for the static hand gesture recognition. The proposed system is based on simple and effective DWT based feature extraction. Detecting and tracking human hand in a cluttered background will increase the performance of hand gesture recognition in terms of accuracy and speed.

B. Proposed System Architecture

In this section proposed system architecture is described in the following figure.

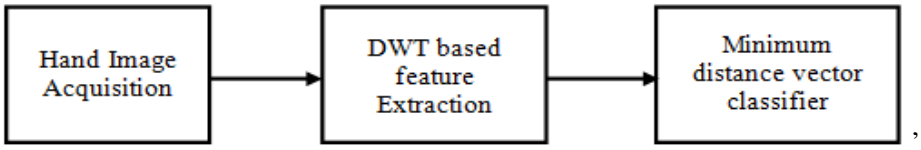


Fig. 1: Proposed System Architecture

The proposed system has two stages training and testing. In the training stage feature vector is calculated for every input image to create database. In the testing stage Input image feature vector is calculated & compared with the train images feature vectors which is stored in database.

1) Stage 1: Hand Image Acquisition

It contains Training & Testing image dataset used for Hand gesture recognition for practical experiments. Sebastien Marchel database is used. This database is benchmark database in the field of hand gesture recognition. The database contains 6 categories of statics hand gestures in total: A, B, C, Five Point & V performed by different people on complex & uniform background.

2) Stage 2: DWT based Feature Extraction

Discrete wavelet transform (DWT) decomposition (pyramid structured) is performed by passing the original image first through the low-pass and high-pass decomposition filters to generate four lower resolution components: one low–low (LL1) sub-image, which is the approximation of the original image and is also called smooth image, and three detailed sub-images, which represent the horizontal (LH1), vertical (HL1), and diagonal directions (HH1) of the original image. The sub-band LL1 alone is further decomposed to obtain (LL2, LH2, HL2 and HH2) the next coarse level of discrete wavelet coefficients; similarly, further decomposition of LL2 is done to obtain the next coarse level. This decomposition process continues up to third level as shown in fig.

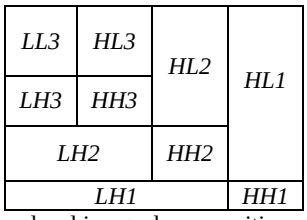


Fig. 2: Three level image decomposition using DWT

In wavelet transform scaling property is used to create different lengths of wavelet functions by compressing or dilating the mother wavelet in order to capture different frequency resolutions in the entire sample data. Whereas, translation property is used to translate or move every generated wavelet function over the entire sample data, in order to capture the spatial localization information. These two important properties have actually explained the capability of multi resolution analysis in wavelet transforms. For every image feature vector is obtained. It contains only seven wavelet coefficient as diagonal detailed coefficient is not considered. Only Energy feature is calculated from original image as well as from approximation and detail sub-images of every level of decomposition using the following formula,

$$\text{Energy} = \sum_{g=0}^{G-1} H^2(g)$$

Where

g is the gray level value (i.e. g = 0, 1...G-1)

G is the highest gray level value.

H (g) is the probability of certain pixel occurring in an image

(i.e. $H(g) = \frac{ng}{N}, n$ g is the of pixels of value g in an image and N is the number of all pixels in an image .

3) Stage 3: Classification by using minimum distance vector

The test image is decomposed using DWT and a set of wavelet statistical features are extracted and then compared with the corresponding feature values of all the classes stored in the features library using a distance vector formula as given in.

$$D(M) = \sqrt{\sum_{j=1}^{N} [fj(X) - fj(M)]^2}$$

Where,

N is the number of features in feature vector.

$fj(X)$ Represents the jth feature of the test sample X.

$fj(M)$ represents the jth feature of the Mth class in the feature library.

Then the input test image is classified to the category of train set image whose feature vector distance is minimum among all available distances of the train set image when compared with input test image.

IV. RESULT ANALYSIS

In this section we are discussing the practical environment, scenarios, performance metrics used etc.

A. Input

Static hand gesture image dataset acquired from digital camera is used for practical experiments. Sebastien Marchel database is used. This database is benchmark database in the field of hand gesture recognition. The database contains 6 categories of static hand gestures in total as: A, B, C, Five Point & V which are performed by different people on complex & uniform background.

B. Hardware and Software Configuration

Hardware Requirements:

- Processor : Pentium IV 2.6 GHz
- RAM : 512 MB DD RAM
- Monitor : 15" COLOR
- Hard Disk : 20 GB

Software Requirements:

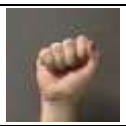
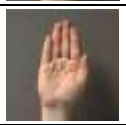
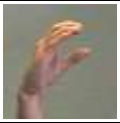
- Front End : Matlab
- Tools Used : Matlab 2012
- Operating System : Windows XP

C. Results

Following result analysis table shows the performance comparison of existing [15] method based on scale invariant feature extraction (SIFT) and the DWT based feature extraction method which is implemented in this paper.

Table -1

Performance with input image test data set

Hand Posture	*Number of Testing image*	*Correct*		*Accuracy*	
		SIFT feature extraction Method	*DWT feature extraction Method*	*SIFT feature extraction Method*	*DWT feature extraction Method*
	*350*	*145*	*280*	*41.42 %*	*80%*
	*350*	*135*	*272*	*38.57%*	*77.71%*
	*350*	*118*	*268*	*33.71%*	*76.57%*
	*350*	*125*	*300*	*35.71%*	*85.71%*
	*350*	*160*	*255*	*45.71%*	*72.55%*
	*350*	*148*	*285*	*42.28%*	*81.42%*

V. CONCLUSION

In this paper a native static hand gesture recognition System that consist of three module 1) Input Image Acquisition by Data base 2) Feature Extraction by using Discrete wavelet transform techniques 3) Image classification & hand gesture recognition by minimum distance vector method is implemented. The recognition accuracy positive prediction, negative prediction is calculated. Experiments show that the system has achieved satisfactory Performance classification accuracy of 79.05% under variable scale, orientation and illumination conditions and cluttered background showing improved results than Existing SIFT feature extraction method.

REFERENCES

[1] A. El-Sawah, N. Georganas, and E. Petriu (Aug. 2008) "A prototype for 3-D hand tracking and gesture estimation," IEEE Trans. Instrum. Meas., vol. 57, no. 8, pp. 1627–1636.

[2] Mandriota, M. Nitti, N. Ancona, E. Stella, A. Distante (2004) " Filter based feature selection for rail defect detection," Machine Vision and Applications, vol.1, pp. 179–185.

[3] S. Mallat, "A theory for multi resolution signal decomposition: The Wavelet representation(1989)" IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 11, No. 7, pp 674-693.
[4] Sebastian Marcel, Oliver Bernier, Jean Emmanuel Viallet and Daniel Collobert.(2000) "Hand Gesture Recognition using Input –Output Hidden Markov Models", Proc. of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 456 – 461.
[5] Xia Liu and Kikuo Fujimura. ,"Hand Gesture Recognition using Depth Data"(2004) Proc. of the Sixth IEEE International conference on automatic Face and Gesture Recognition, pp. 529-534.
[6] L. Bretzner, I. Laptev, and T. Lindeberg.(2003) "Hand Gesture using multi-scale color features, hierarchical models and particle filtering", Proc. of the Fifth International conference on Automatic Face and Gesture Recognition, pp. 423- 428.
[7] V. Pavlovic, et al. (1997) "Visual Interpretation Of Hand Gesture For Human-Computer Interaction: A Review", IEEE Trans. On Pattern anal. Mach. Intel. 19(7), pp 677-695.
[8] Attila Licsar and Tamas Sziranyi. (2002)"Supervised training based hand gesture recognition system", Proc. of the 16th International Conference on Pattern Recognition, Vol. 3, pp 30999 – 31003.
[9] E.Stergiopoulou and N.Papamarkos(2006.) "A New Technique on Hand Gesture Recognition", Proc of the IEEE International Conference on Image Processing, 2657-2660.
[10] Byung-Woo Min, Ho-Sub Yoon, Jung Soh, Yun-Mo Yangc and Toshiaki Ejima(1997) "Hand Gesture Recognition Using Hidden Markov Models", Proc. of the IEEE International conference on Systems, Man and Cybernetics, vol 5, pp. 4232 -4235.
[11] Meide Zhao, Francis K.H. Quek, Member, IEEE, and Xindong Wu, Senior Member, IEEE(November 1998) "RIEVL: Recursive Induction Learning in Hand Gesture Recognition", IEEE Transactions on Pattern Analysis and machine intelligence, vol. 20, no. 11.
[12] Yikai Fang, Jian Cheng, Kongqiao Wang and Hanqing Lu.(2007) "Hand Gesture Recognition Using Fast Multi-scale Analysis", Proc. of the Fourth International Conference on Image and Graphics, pp 694-698.
[13] Chris Joslin, Ayman El-Sawah, Qing chen, Nicolas Georganas.(2005) "Dynamic Gesture Recognition", Proc. of the Instrumental and Measurement Technology Conference, pp 1706-1710.
[14] Timi Ojala, Matti Pietikainen and Topi Maenpaa.(2002) "Multi-resolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 24, pp. 971-987.
[15] Nasser H. Dardas and Nicolas D. Georganas, Fellow, IEEE,(Nov2011.) "Real-Time Hand Gesture Detection and Recognition Using Bag-of-Features and Support Vector Machine Techniques", IEEE Transaction On Instrumentation and Measurement, VOL. 60, NO. 11, pp3592-3607.