Classification of Hand Gestures Using Gabor Filter with Bayesian and Naïve Bayes Classifier

Tahira Ashfaq, Khurram Khurshid

Electrical Engineering Department, Institute of Space Technology
1, Islamabad Highway, Islamabad, 44000, Pakistan

Abstract—**A hand Gesture is basically the movement, position or posture of hand used extensively in our daily lives as part of non-verbal communication. A lot of research is being carried out to classify hand gestures in videos as well as images for various applications. The primary objective of this communication is to present an effective system that can classify various static hand gestures in complex background environment. The system is based on hand region localized using a combination of morphological operations. Gabor filter is applied to the extracted region of interest (ROI) for extraction of hand features that are then fed to Bayesian and Naïve Bayes classifiers. The results of the system are very encouraging with an average accuracy of over 90%.**

Keywords—Human Computer Interaction; Hand Segmentation; Gesture recognition; Gabor Filter; Bayesian and Naïve Bayes classifiers; Feature Extraction; Image Processing

I. INTRODUCTION

These days artificial intelligence is being commonly used in various daily life applications. Image processing and machine learning play a vital role in artificial intelligence applications. Systems like face recognition, facial expression recognition and gesture recognition involve multiple steps of image segmentation and processing and research in this domain continues to prosper to find out more suitable solutions for forthcoming challenges and problems.

Gesture recognition is considered as quite a difficult job due to nature of challenges that hinder the state of the art systems get 100% accurate results. Different techniques like Support Vector Machines (SVM), Neural Networks (NN) and Hidden Markov Models (HMM) are used by researchers for classification of gestures in various classes. Hand gesture classification has various applications, like in Human Computer Interactive applications, gesture recognition and classification allows users to interact efficiently with computers using easy to use/understand hand gestures. William1995 was one of the first ones that established the use of hand gesture recognition for controlling household machines. Ayesha2010 defines that in gesture recognition main goal is to create a system which can recognize precise human gestures for relaying useful information.

The hand gestures can be categorized as static hand gestures and dynamic hand gestures. Static hand gestures are basically still images having a certain position and features of hand. The hands can be differentiated by different hand position features in which finger positions are determined and features are found for the precise finger, thumb and palm pattern. Thus a static hand gesture is signified by an individual frame or image.

Dynamic hand gestures on the other hand can be differentiated through the starting and last stroke movement of moving hand gesture. A video is the best example of dynamic hand gesture. Mainly two different approaches are used for hand gesture-tracking named as glove based systems and vision based systems for hand gesture analysis. In the first approach, sensors and other instruments are attached with user's hand in subjective analysis. In order to analyse data, Magnetic field tracker device and data glove or body suits are used. Extracted features from different human body positions such as the angle and rotation are essential for examination in such systems. Vision-based approach is the substitute to glove based technique because normally it does not need sophisticated or even small external equipment apart from a camera. These systems work using various image processing and machine learning algorithms. In any system however, the 3 important steps include hand region segmentation, feature extraction and then final classification & recognition.

II. PROPOSED SYSTEM

In this paper, we present a scheme for static gesture recognition using vision based approach. Our proposed system can be summarized in the following steps. In the first step pre-processing is done, where initial pre-filtering for noise removal is performed followed by conversion of images from RGB to LAB color space. Then the third dimension grey level is converted to binary using global threshold and is mapped to extract just the desired hand region. For removal of undesirable areas, combination of different morphological operations is applied on the image. The resulted image is then used to extract the edges of the segmented region. Once the hand region is segmented, we can extract the certain features of hand which are then used in the classification module for gesture mapping. The classification is done by using Bayesian and naïve Byes classifier and their performance is compared in the end.
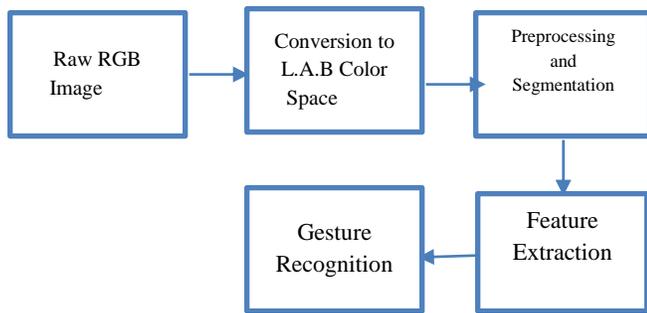
Fig. 1.   System Block Diagram

The procedures included in above block diagram are discussed in detail in the following section.

A. *Preprocessing*

1) Conversion to L.A.B Space
We use RGB color images as input which is converted to LAB color space by analysing the pixel's color and differentiating on its basis. Skin color has certain properties that can be extracted easily via color information. In contrast to RGB and CMYK, L.A.B color is designed for accurate estimations and therefore better results in case of image segmentation.

Luminance L* is separated by color space from two different color components (a* and b *). This is because human perception of lightness is almost matched by the 'L' component. Accurate colored balance correction can be made by modifying output curves in 'a' and 'b' components, or for the adjustment of lightness and contrast using the 'L' component.

2) Thresholding
Automatic grey thresholding of third component 'b' is performed in L.A.B. color space.

$$I(:,:,3)$$: I represents the input image

The obtained gray image is converted to binary using the below mentioned threshold. The resultant binary image can be mapped to original image to view the segmented region.

$$J(x,y) = \begin{cases} I(x,y), & Where \ O(x,y) \sim= 0 \\ 0, & otherwise \end{cases}$$

J in above equation represents the output of segmented image. O is the binary image and if it has a nonzero pixel value then the same spatial coordinates intensity of I will be assigned to the output pixel $J(x,y)$. Thus segmented skin area can be extracted through this mapping.

B. *Segmentation*

For segmentation, we have to perform the following operations.

1) Morphological Operations
Opening is performed in order to remove narrow joints and noise. In Opening we first perform erosion by using 5×5 structuring element.

$$A \Theta B = \bigcap_{b \in B} A_{-b}$$

Here erosion is used to remove the extra region. Where A represents the image and B denotes the structuring element. In the above equation, structuring elements are applied on the image A. If the 'ON' pixels of a structuring element hit the pixel image where the pixel value is more likely to be 0, then the center value of mask is shifted towards the new image. Whenever there is a missing 'ON' pixel in the structure element, either structure element or mask will be shifted towards the next pixel.

After erosion we dilate image by using 5×5 structuring element in order to fill the disturbed region.

$$A \oplus B = \bigcup_{b \in B} A_{b}$$

Here A is represents the gray scale image, and B again denotes the structuring element.

To recap, dilation adds pixels to the object boundary while erosion removes object pixels from boundary.

2) Canny edge detector
After morphological operations, we need to detect the boundary of the region. For that we tried various known filters like Sobel, Prewitt, Canny. We finalized Canny edge detection as it gave more desirable results as compared to the other two options. After this stage, we get a well detected hand region and its boundary.

C. *Feature Extraction*

Once the hand region is extracted, we have to extract good features that have more discriminative power. These precise features are used for classifying various hand gestures so that the problems like interclass similarity and intra class variability can be catered for easily. Extraction of the these best features is definitely a challenging task. Various researchers have used various features for classifying gestures based in internal information of the object in image. A good feature set highly effects the performance of a classifier.

For our technique, we use Gabor filter based features that tell us about the nature of gesture using its pattern and orientation. The advantage we have using Gabor filter is that we get large variance between the features of different classes which minimizes the chances of error if used properly in a classifier.

1) Gabor Filter
Gabor filter has been used in literature for various applications including satellite image segmentation, urban zone detection [17], document and camera images for edge detection etc. In spatial domain, a 2D Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave. The impulse response of Gabor filter is defined by a sinusoidal wave. Sinusoidal wave is known as a plane wave for 2D Gabor filters, multiplied by a Gaussian function. Because of the multiplication-convolution property (Convolution theorem), the Fourier transform of a Gabor filter's impulse response is the convolution of the Fourier

transform of the harmonic function and the Fourier transform of the Gaussian function. Gabor filter has a real and an imaginary component representing orthogonal directions. The two components may be formed into a complex number or used individually.

Complex equation is given below

$$g\left(x, y; \lambda, \theta, \psi, \sigma, \gamma\right) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right)\exp\left(i\left(2\pi\frac{x'}{\lambda} + \psi\right)\right)$$

Real part is as follows

$$g\left(x, y; \lambda, \theta, \psi, \sigma, \gamma\right) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x'}{\lambda} + \psi\right)$$

And imaginary part is as follows

$$g\left(x, y; \lambda, \theta, \psi, \sigma, \gamma\right) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right)\sin\left(2\pi\frac{x'}{\lambda} + \psi\right)$$

Where

$$x' = x\cos\left(\theta\right) + y\sin\left(\theta\right)$$
$$y' = -x\sin\left(\theta\right) + y\cos\left(\theta\right)$$

In the above equation

The wavelength of sinusoidal factor is represented by λ.

θ is the orientation of the normal to the parallel stripes of the Gabor function, ψ is the phase offset, σ is the sigma/standard deviation of Gaussian envelope, and γ is the spatial aspect ratio that specifies the Ellipticity of the support for Gabor function.

Parameter	Symbol	Values
Orientation	Θ	$\{0, \frac{\Pi}{8}, \frac{1\Pi}{8}, \frac{2\Pi}{8}, \frac{3\Pi}{8}, \frac{4\Pi}{8}, \frac{5\Pi}{8}, \frac{6\Pi}{8}, \frac{7\Pi}{8}\}$
Wavelength	λ	$\{4, \sqrt[4]{2}, 8, \sqrt[8]{2}, 16\}$
Phase	Ψ	$\{0, \frac{\Pi}{2}\}$
Gaussian envelope	γ	$\gamma = \lambda$

We apply this filter to each hand image to find out the subjected feature values from the image.

D. Classification

Classification means to segment regions of interest into various groups called classes. Mainly there are two types of classification, Supervised and Unsupervised.

In unsupervised classification (also known as clustering in pattern recognition) the segmented groups are without labels, while in supervised classification we need some training samples that are helpful in classifying the unknown data or test cases. The grouping or matching is based on certain features that are extracted from the objects. These features should be discriminating in nature to give a correct and meaningful classification of the data. In our case, we give the Gabor Filter based features to two classifiers (Bayesian and Naïve Bayes) for deciding the class of the query gesture.

1) Bayesian classifier

Bayesian classifier is based on calculating posterior probabilities using Bayes theorem. Posterior probabilities which are also known as class density estimates are calculated as below:

$$P\left(H \mid X\right) = \frac{P\left(X \mid H\right)P\left(H\right)}{P\left(X\right)}$$

In the above expression, posterior probability is denoted by $P\left(X \mid H\right)$ in which X is conditioned on H. The prior probability of X is denoted by $P\left(X\right)$. Classifier's accuracy highly depends on this parameter.

2) Naïve Bayes Classifier

The other technique that we tried was the Naïve Bayes Classification. This classifier basically estimates the class of unknown data item using probability models. Using Bayes theorem, posterior probability can be calculated using the following expression

$$P\left(C_i \mid X\right) = \frac{P\left(X \mid C_i\right)P\left(C_i\right)}{P\left(X\right)}$$

Here $P\left(C_i\right)$ is known as class prior probabilities and it can be estimated by the following expression.

$$P\left(C_i\right) = \frac{S_i}{S}$$

Where S_i represents the selected training sample, which has class C_i and S denotes the total number of training samples.

III. EXPERIMENTAL RESULTS

Now for testing the above explained methodology, we have used data with complication or real background. The data is acquired by a 7 mega pixel camera. We have a total of ten hand gestures that we are classifying. In our data set, we have collected 18 images of each hand gesture in which 5 are used for training and 13 for final testing. To justify the best accuracy of our proposed result, we have compared our technique results with a similar work done by nawazish et.al (2013) [18] that used Neural Network Based hand gesture recognition of dumb people. In that paper, the authors worked on a technique using neural networks with the combination of mean and entropy feature. In the paper he extracted features using block base and used 4 blocks of the image and extracted the feature separately. After extracting the features he used an array to store the results of all blocks and for classification he used neural networks. The results are given below in table 1. The table shows that our method of Gabor Filter based feature extraction and Bayesian Classification works well and classifies the gestures 90% of the time while the method of Nawazish achieved a classification rate of 88.25%. Step wise results of various stages of our technique are shown in figure 2. Two examples are depicted here in detail.
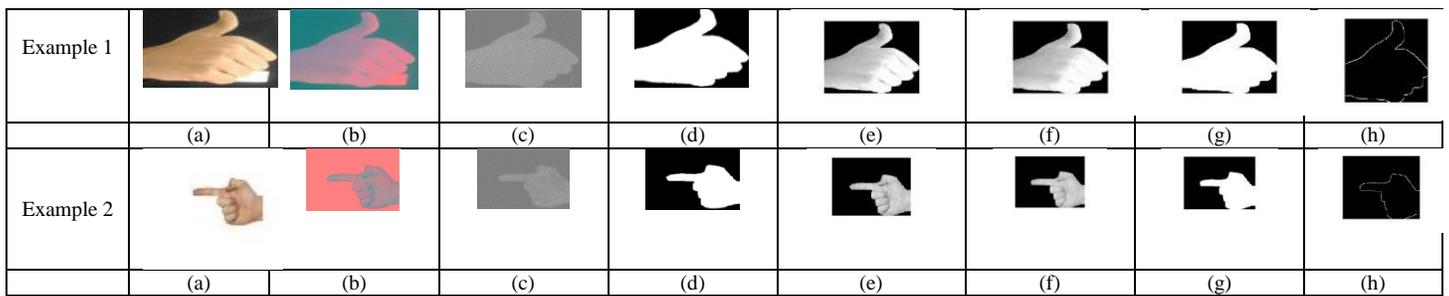
Fig. 2.   Segmentation Results (a) Original image (b) LAB color space image (c) third color component of Lab image *A (:,:, 3)* (d) binary Image (e) Binary mapping (f) Morphological Operation opening (g) Final Binary image (h) Segmented Hand

TABLE I.   COMPARISON OF RESULTS

Gestures	Comparison of various techniques		
	Nawazish et.al (2013)	*Naïve Bayes*	*Bayesian Classifier*
Stop	89.0	90.12	90.53
Pointing	85.2	87.20	88.01
Self-Pointing	87.34	89.01	89.85
Drinking	90.05	90.12	91.01
Take Care	89.66	90.00	91.00
…..	…..	…..	…..
Total	88.25	89.29	90.08

IV.   CONCLUSION AND FUTURE WORK

In this paper, we briefly presented a technique for gesture classification in static images using Gabor Filter based features and Bayesian classifier. Our proposed technique shows promising results on our data set. The main point is that the hand segmentation part in our technique plays very important role. If the segmentation is good, the results of gesture classification improve also. The classification accuracy is obtained with the help of Bayesian and Naïve Bayes classifier. In future we intend to extend this work on dynamic gestures as well. Human Computer Interaction and Sign Language Recognition are two of the most prime applications of this research. We also need to compare our method with more state of the art methods and improve it further to incorporate more gestures.

ACKNOWLEDGMENT

REFERENCES

[1]   Ahmad Yahya Dawod, Junaidi Abdullah and Md. Jahangir Alam (2010) *"Adaptive Skin Color Model for Hand Segmentation"* 2010 International Conference on Computer Applications and Industrial Electronics (ICCAIE 2010), December 5-7, 2010, Kuala Lumpur, Malaysia.

[2]   Huang D, Xiang C and Ge SS (2007) "*Feature extraction for face recognition using recursive Bayesian linear discriminate*". In International symposium on image and signal processing and analysis, Istanbul, Turkey, pp 356–361.

[3]   Kwok Wai Wong, kin-man lam and wan-chi Siu (2003) *"an efficient color compensation scheme for skin Color segmentation* "0-1803-7161-3/03\$11.00 02003 LEEE.

[4]   Wei-Hua, Rewwang and Chun liangtung (2008) *"Dynamic Hand Gesture Recognition Using* Hierarchical Dynamic Bayesian Networks Through Low-Level Image Processing" Proceedings of The Seventh International Conference on Machine Learning and Cybernetics, Kunming, 12-15 july2008.

[5]   Reflex-UL-Zaman Khan and Noor Adnan Ibraheem (2012) "*hand gesture recognition: a literature review*" International Journal of Artificial Intelligence & Applications (IJAIA), Vol.3, No.4, July 2012.

[6]   Swapnil D. Badgujar, Gourab Talukdar, Omkar Gondhalekar, Mrs. S.Y. Kulkarni (2014)" *Hand Gesture Recognition System, "* International Journal of Scientific and Research Publications, Volume 4, Issue 2, February 2014 1 ISSN 2250-3153.

[7]   Deepika Tewari, Sanjay Kumar Srivastava, (2012) a Visual Recognition of Static Hand Gestures in Indian Sign Language based on Kohonen Self- Organizing Map Algorithm‖, International Journal of Engineering and Advanced Technology (IJEAT), Vol.2, Dec 2012, pp. 165- 170.

[8]   ShrikantKulkarni, Manoj H, Sumam David, VenugopalaMadumbu and Y Senthil Kumar (2011) "Robust Hand Gesture Recognition System Using Motion Templates" 2011 11th International Conference on ITS Telecommunications 978-1-61284-671-2/11/\$26.00.

[9]   Wenjun Tan, Chengdong Wu, Shuying Zhao and Shuo Chen (2009) "*Hand Extraction Using* Geometric Moments Based on Active Skin Color Model" 978-1-4244-4738-1/09/\$25.00 ©2009 IEEE.

[10] Surve Pranjali and Prof. Ubale V.S (2015) "*Recognition of Hand Gesture using PCA, Image   Hu's Moment and SVM*" IPASJ International Journal of Electronics & Communication (IIJEC). Volume 3, Issue 4, April 2015   ISSN 2321-5984.

[11] M. A. MONI and A B M Shawkat Ali (2009) "*HMM based Hand Gesture Recognition: A Review on Techniques and Approaches*" 978-1-4244-4520-2/09/\$25.00 ©2009 IEEE.

[12] Mrsayeshabutalia, Msdivya shah and Dr. R. vdharaskar (2010) "*Gesture recognition system*" ©2010 International Journal of Computer Applications (0975 – 8887) Volume 1– No. 5.

[13] Doe Hyung Lee and Kwang Seok Hong (2011) "*Game Interface using Hand Gesture Recognition*" Computer Sciences and Convergence Information Technology (ICCIT), 2010 5th International Conference on Issue Date: Nov. 30 2010-Dec. 2 2010On pages (s): 1092 - 1097Location: Seoul Print ISBN: 978-1-4244-8567-3References Cited: 10INSPEC Accession.

[14] Xiao long Teng, Bin Wu and Weiwei Yu, Chongqing Liu (2005) "*A hand gesture recognition system based on local linear embedding"* Journal of Visual Languages and Computing 16 (2005) 442–454, www.elsevier.com/locate/jvlc 1045-926X/\$.

[15] Eshed Ohn-Bar and Mohan Manubhai Trivedi (2014)," *Hand Gesture Recognition in Real-Time for Automotive Interfaces: A Multimodal Vision-based Approach and Evaluations*" To Appear in IEEE Transactions on Intelligent Transportation Systems 2014.

[16] Pavlo Molchanov, Shalini Gupta, Kihwan Kim, and Kari Pulli, (2015) *"Multi-sensor   System for   Driver's Hand-Gesture Recognition"* NVIDIA Research, Santa Clara, California, USA. IEEE International Conference on Automatic Face and Gesture Recognition (FG2015), May 2015.

[17] Hareem Iftikhar, Khurram Khurshid, Fusion of Gabor Filter and Morphological Operators for the Detection of Settlement Zones in Google Earth Satellite Images, IEEE International Conference on Signal and Image Processing Applications (ICSIPA), Nov 2011, Malaysia

[18] *Nawazish,* Neural Network Based Hand Gesture Recognition, First International Conference on Technology for Helping People with Special Needs, Saudi Arabia, 2013.