Sign Language Recognition using SVM

Thesis submitted in partial fulfillment of the requirements for the award of degree of

Master of Engineering

in

Computer Science and Engineering

Submitted By

Shama

(Roll No. 801232024)

Under the supervision of:

Mr. Karun Verma

Assistant Professor

COMPUTER SCIENCE AND ENGINEERING DEPARTMENT

THAPAR UNIVERSITY

PATIALA – 147004

July 2014

CERTIFICATE

I hereby certify that the work which is being presented in the thesis entitled, "*Sign Language Recognition using SVM*", in partial fulfillment of the requirements for the award of degree of Master of Engineering in *Computer Science and Engineering* submitted in Computer Science and Engineering Department of Thapar University, Patiala, is an authentic record of my own work carried out under the supervision of *Karun Verma* and refers other researcher's work which are duly listed in the reference section.

The matter presented in the thesis has not been submitted for award of any other degree of this or any other University.

Signature: Shama

(Shama)

This is to certify that the above statement made by the candidate is correct and true to the best of my knowledge.

Karun

(Karun Verma)

Assistant Professor

Countersigned by

(Dr. Deepak Garg)

Head

Computer Science and Engineering Department

Thapar University

Patiala

(Dr. S. K. Mohapatra)

Dean (Academic Affairs)

Thapar University

Patiala

ACKNOWLEDGEMENT

First of all I would like to express my deep gratitude to my supervisor Mr. Karun Verma, Assistant Professor, Computer Science and Engineering Department, who gave me a chance to explore skills in the field of Natural Language Processing and spend sufficient time in exploring terms used in my thesis work. He gave me suggestions so that my research was always in progressive mode.

I am also thankful to Mr. Deepak Garg, H.O.D., Computer Science and Engineering Department, who always considered the problem and provided each and every facility related to my thesis work.

Also, I would like to thank the members of the Computer Science and Engineering Department for their encouragement and support during preparation of this thesis report.

Most importantly, I would like to thank my parents and the almighty for showing me the right direction out of blue, stay calm in oddest of times and keep moving even at times when there was no hope.

Shama

(Shama)

801232024

ACRONYMS

ISL Indian Sign Language

ASL American Sign Language

SVM Support Vector Machine

SE Structure Element

ABSTRACT

There is one of the natural language called sign language. The sign language is for persons who are challenging in hearing and speech, but all people cannot understand this language. So, there is a need of system that can understand what a challenging person wants to say. For this purpose, a system is developed in which sign is captured and processed. The processing includes use of gaussian filter and bicubic interpolation. The work is done in MATLAB. In this system, a person stands in front of camera and makes sign. The captured RGB image is processed. This image is further used to detect edge using canny edge detector and improved sobel edge detector. Improved sobel edge detector uses dilation and erosion operations. The two images produced after edge detectors are fused together. This data is used to extract features. Zone based approach is used to extract features. Further mean and standard deviation are applied on the data in zone based approach. Then Support Vector Machine is used for training of the system.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction

Sign language is a language for the people who are challenging in hearing and speech. Despite common misconceptions, sign languages are complete natural languages, with their own syntax and grammar. Signs are made of units referred to as cheremes. A sign is made of cheremes and any two signs can be differentiated using at least one chereme.

Indian sign language used by people in India is of mainly three types. It is further divided into sub categories. The hierarchy of ISL is shown in below Figure 1.1

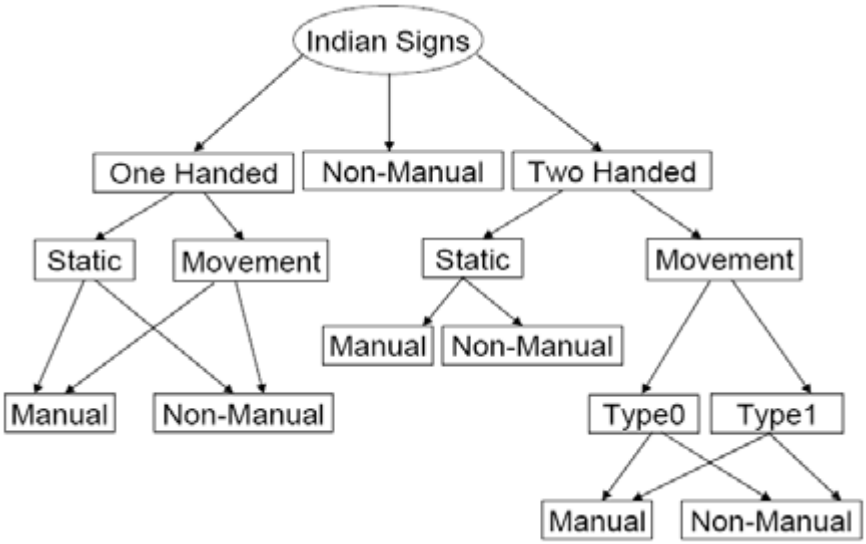

Figure1.1 ISL Type Hierarchy

It is understandable that, information access for a deaf in places like railway stations, banks and hospitals is quite hard. There is need of a system in which they can communicate directly with everyone. There is need of Sign Language Recognition system that can understand signs made by challenging people. In such a system, a signer can make a sign and the other person can understand. That means if the other person does not know the sign language even then communication is possible.

The main objective of Sign Language Recognition is interpretation of images which describes hand sign created by user. Sign language recognition system deals with capturing of image through electronic device. The input mode is video sensor device such as webcam. The Sign Language recognition system is having goal to analyze image and give as output the corresponding text. The main problem occurs in system is recognition of images. It is difficult to work on images. Therefore image preprocessing is done.

Motivation

According to the literature survey, sign recognition leads the thesis to work on different techniques for image analysis. An experimental approach is needed to develop, to compare and to evaluate the performance of distinct features of different signs. In the practical image acquisition systems noise is produced commonly in images. So noise removal is one of the phases in the system. Recognition of signs on the basis of different shapes is the goal of this project. The interest of the thesis is to evaluate the accuracy of sign recognition.

Research Scope

The research in Sign Recognition generally includes sign detection and then feature extraction. The idea arises from recognition systems for English alphabets and words. This thesis concern Punjabi vowels and English alphabets.

All the algorithms related to image processing will not be possible to be included in this thesis. The thesis is containing the research on sign language recognition using hand signs with black background.

Background Information

There are four image types and their explanation according to Matlab is given below:

- **Indexed image:** This type of image is having two matrixes. One matrix is of image data and other is of a colormap. Image matrix is having values that are indexes into the colormap. The colormap matrix represents the colors in the image.

- **Intensity image:** The values represent intensities with some range. It is single matrix. Every element of this matrix corresponds to one image pixel.

- **Binary image:** Each pixel in this type of image assumes one value out of two discrete values. The storage is done as 2-dimensional matrix having values 0's and 1's.

- **RGB image:** In this image, each pixel color is represented as a set of three values. These values represent red, green and blue intensities. These intensities make the color. These values are directly stored in image array.

Image processing: It can also be considered as a form of Signal processing. Image processing includes image for:

- Discrete representation: Visual information is converted into discrete form so that computer can process that information.

- Processing: Image processing deals with improvement of image quality and compression of data.

- Analysis of image: Image is analyzed for feature extraction, to quantify shapes and for registration and recognition.

Edge Detection:

The area of image having significant change in intensity or contrast is called edge. For edge detection, there is need of locating area having strong intensity contrasts.

Edge detection is classified into two categories [1] and explained below:

1. Derivative approach: In this approach edges are detected by taking derivative which is followed by thresholding. Some operators are also having noise cleaning scheme.
   Examples: Roberts operator, Sobel operator, Prewitt operator and Canny operator etc.
2. Pattern fitting approach: A small neighborhood is analyzed by applying edge approximating functions over it. Parameters of best fitting function

and their properties are determined. Presence of edge is decided on this information. These are also called edge filters.

Edge Detection usage

- Reduce unnecessary information in the image.
- Extract important features of an image:

  Corners

  Lines

  Curves

- Recognize objects, boundaries, segmentation.

Support Vector Machines:

Support Vector Machine is one of the supervised learning model having associated learning algorithms. SVMs are used for analysis of data and recognition of patterns, and are applied for classification and regression analysis. SVM belongs to broad family of kernel based learning algorithms.

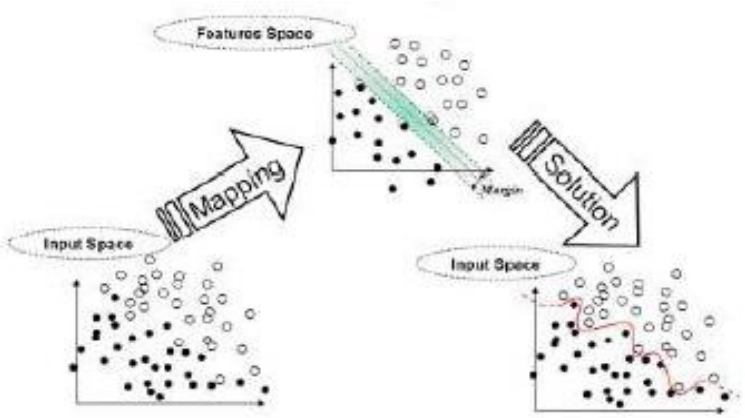

Figure1.2: The SVM Algorithm

SVMs maximize the margin around separating hyperplane. The decision function specified with the support vectors, which are subset of training samples.

Support Vectors: The data points which are closest to the decision surface are called support vectors. It is the very difficult task to classify them.

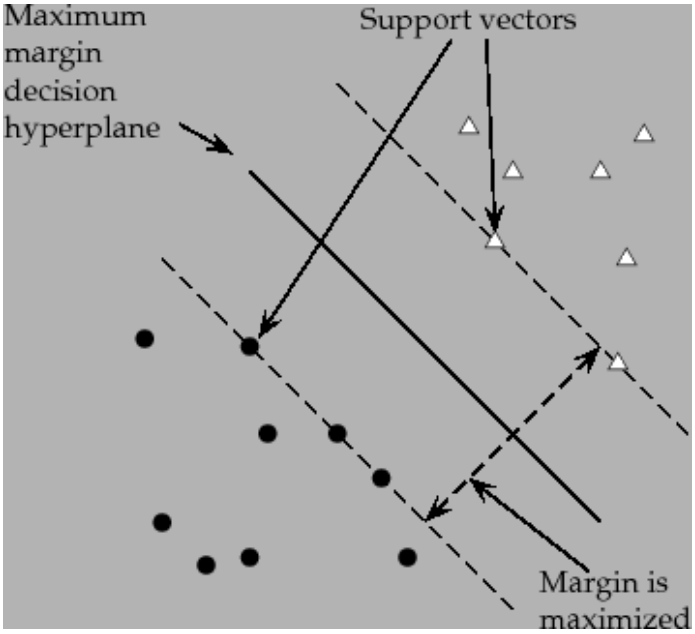

Figure 1.3: Support Vectors

There are mainly two types of SVMs:

1. Linear Support Vector Machine
2. Non Linear Support Vector Machine

Fundamental idea of SVM is optimal hyper-plane for linear separable patterns. In case of not linearly separable patterns it does transformations of original data map into new space that is kernel function.

Linear SVM: In Linear SVM, training data of two classes can be separated in two hyper-planes with no points between them. Then their distance is maximized. There is a region called margin which is bounded by them. Linear classifiers depend on dot product between data point's vectors.

Let us assume a set of training samples, each marked as belonging to two categories. SVM training algorithm is used to build a model. This model assigns new samples into a category that makes it a binary linear classifier and it is non-probabilistic.

Non Linear SVM: There are training data sets that are not linearly separable. So, non linear SVM is used for them. In this, non linear kernel function replaced dot product. Some kernels are described below:

Polynomial: Polynomial kernels are having form of $k(z_i, z_j) = (z_i.z_j+1)^d$ where k means kernel function. If d=1 then it is linear kernel. If d=2 then it is quadratic kernel.

Radial Basis: It can be recognized as squared Euclidean distance between two support vectors. As the distance increases its value decreases and the value ranges between zero and one.

Sigmoid**:** It is originated from neural network.

Chapter 2

State Of Art in Sign Recognition

Linguistic work on ISL began during late 1970s. Before that, the existence of ISL was not acknowledged. In 1977 a survey was conducted and it was revealed that ISL is a complete natural language instigated at the Indian subcontinent.

Real time system of Hidden Markov Model (HMM) is used to recognize sentence level American Sign Language [2]. In this system user wear solid colored gloves. The information of shape, trajectory, and orientation is used as input of HMM. Then HMM uses this information for recognizing the signed words. There is no need to invoke complex models of hands, as the rates of error are low on training set as well as on independent test set through hidden Markov models. While the results produced are not sufficient for a working system for ASL recognition, but the approach is promising. The recognition rate is high for training data which shows that the HMM topologies are good and the models are also converging.

A gesture recognition system that used Hidden Markov Models is developed [3]. The system recognizes gestures interactively and online learning of new gestures is performed. The model is updated with each example it recognizes. The implementation of system is done as an interactive interface for robotic system and is programmed by example. With this approach a system is built and tested that recognizes letters from the sign language alphabets by Virtual technologies called Cyberglove. The system works by taking gestures from user. Then it segments data into distinct gestures and classification of each gesture is done. If the system is certain about gesture then immediate action is performed. Otherwise the system queries the user for classification. The symbols of encoded gesture are added by the system to its proper gesture model. The work is done mainly in C. The number of gestures that can be classified is limited by number of observable symbols that are generated by preprocessor.

A system to recognize word level Chinese sign language is developed [4]. The system is based on data glove. There are four modules of the system: frame extraction, feature extraction, learning, and recognition. Neural network and HMM hybrid

method are used for learning and recognizing modules. The hybrid approach is effective for sign language word. But for sentence level sign language others methods like context information, fusion of gestures and other statistical methods can be considered.

Orientation histogram based hand gesture recognition system is developed [5]. The system uses visual information without any special glove for gesture. There are three steps in this system. First step is using a search algorithm for hand area. In this algorithm, a hand block is found and segmented efficiently from monochrome input images. If the hand area is extracted successfully then feature vectors representing the hand shape are analyzed with the orientation histogram scheme. For the moving hand, feature vectors are obtained by motion estimations. Third and the last step is to recognize hand gesture by using feature vectors of hand shape and movements of hand. But there is need for decision of key frame, and training methods for modeling.

The system for gesture recognition using data gloves and neural networks for virtual reality applications is proposed [6]. In this system, for virtual reality the hand gestures are used as human and computer interactions. The specific hand gestures like index finger and fist are defined for the application. Data glove is used as input that provides eighteen measurement values for different finger point angles. Three standard three-layered back-propagation networks are used for experiment. Logistic function used is following:

The system is basically about controlling a program using sign language. The commands such as rotate robot hand and grab object are assigned to some gestures. The system can be navigated in the virtual reality to manipulate objects there.

A System of gesture based interaction and communication with automated hand gesture contours classification is proposed [7]. A complete vision based system is developed to classify hand gestures. The input image is gray scale. Segmentation of gray scale image is done using Otsu segmentation algorithm. Noise of background and noise of object is removed from segmented image using morphological filtering. The gesture's contour is represented with the help of localized contour sequence. Their samples are having distances that are perpendicular. This distance is between the chord and contour pixels that connect end points on contour pixels of window

centered. The similarity between localized contour sequences is used to determine gesture similarity. The similarity between localized contour sequences is measured by linear alignment and nonlinear alignment. In this system, it is shown that on some subset of American Sign Language hand gestures if non linear alignment is used then gestures are not misclassified by system. The system is not used to classify dynamic gestures that are represented by sequence of static gestures.

A system to recognize 3 D arm movements is developed by using fuzzy rule based approach [8]. The system works for Taiwanese Sign Language. The input image goes for preprocessing steps. The machine learning is done by using Hyper Rectangular Composite Neural Networks (HRCNN). HRCNN integrate the rule based neural networks. A two layer HRCNN is generated by supervised decision directed learning (SDDL) algorithm. In this system, accumulative similarities between arm movements in vocabulary with unknown type of arm movements are compared. On the basis of this comparison unknown arm movements are classified to their corresponding arm movement in the vocabulary. The system is not made practical speaking.

There is an online communication system for sign language [9]. There are three parts of the system. First is an editor to create the sign language animation sequence. The virtual avatar with 39 joints and 86 segments is used for animating the human model. Second part is for the user, a viewer applet to view online the sign language animation files. The third part provides users the service to upload their own sign language animation series. Also the users can learn from website online.

A system for gesture recognition by using 3D data is developed [10]. A dense range image is generated from a scene using 3D sensor. The system does not rely on color information of the scene. The system recognizes complex static hand postures. The image undergoes arm segmentation then estimation of hands pose. Then 3D feature extraction is done and then gesture classification is done using these features. Testing is done by using small number of images and test cases, and it is not real time application.

The system produces signed sentences by an avatar [11] from text. French sign language is considered in this system. Written or verbal messages are translated into

signed sentences. The focus is mainly on facial region of the avatar. Facial region considers facial expressions which are non manual signs.

An automatic sign language conversion system is developed [12]. Support vector machine is used as classifier that used a Gaussian kernel. Image preprocessing techniques are applied on input images. All outline images are obtained and are used to calculate Fisher score. In this process firstly parameter is extracted from outline i.e. chain code. Chain code is used as input for the creation of Hidden Markov Model (HMM). Then Fisher score is calculated from gradient and it is of a logarithm that is of the observation symbol probability distribution. In SVM, the distance between the score of two sequences is used as separation between patterns.

There is a cross modal system that translates Hindi strings to Indian Sign Language. The system is for use at reservation counters in Indian railways. The system is named as INGIT [13]. Construction Grammar approach is used to handle formulaic inputs. These inputs are in form of construction lexicon and single constituents as well as larger phrases which are having semantic mappings at each level. Construction grammars, irrespective of their structural complexity, commit themselves to parity of linguistic expressions. These grammars treat equally the general patterns of language that is responsible for more idiomatic and compositional utterances. Meaning mappings are used for constructions. These mappings are bidirectional. These mappings are used for production as well as for parsing. The domain specific construction grammar converts input into a thin semantic structure. The input to ellipsis resolution is this thin semantic structure. After this a saturated semantic structure is obtained. An ISL tag structure is generated on the basis of type of utterance. It is generated with the help of ISL generator. Then HamNoSys converter uses this ISL tag for generating graphical simulation. HamNoSys is a sign notation system and it is an editor that provides graphical images for phonological parameters of a sign. The system works with different phases. This system contains Input parser, Ellipsis Resolution and ISL Generator. In Input parser spoken language strings are transcribed. There are some sentences that does not contain subject or object directly then Ellipsis Resolution is used. It saturates the semantics of the event. In ISL generator smaller units are grouped together and sentences are created using bottom up approach. HamNoSys is used to write signs. The HamNoSys editor provides a set

of graphical images for most of phonological parameters of a sign. For instantiating the signs graphical simulation modules are constructed. Graphical simulator is used that converts ISL tags into HamNoSys. The HamNoSys are displayed on a graphical simulator. The limitation of this system is that in present scenario facial expressions are not supported in this system.

There is system based on Gabor Wavelet transform that is developed in China [14]. In this combination of hand and facial information is used. Another system developed is Continuous sign language recognition for Taiwanese sign language. In this continuous sign language is segmented into isolated sign segments. These segments are then interpreted by Product HMMs if these are signs. Gloves are used while making signs.

There is an ongoing study that aims to teach Sign Language, using humanoid robot [15], to hearing impaired child. Currently its demo is in Turkish Sign Language. It is for children that do not know how to read and write, and are unfamiliar with sign language. The system is used to teach Sign Language to hearing impaired children using non-verbal communication. It is imitation based interaction system that is having games between a child and a humanoid robot. With the help of body, face gestures and hand movements, the robot expresses a word in Sign language from a set of chosen words by comprehending the word. A relevant feedback is given by the child to the robot. Child shows color flashcards with the illustration of word. If flashcard and word are matched then the robot pronounces the word and continues same way.

A system uses video of full gesture of ISL and converts it into text [16]. Those words are processed with simple grammar rules to form meaningful sentences. From the video frames are extracted. The key extraction simplified algorithm is applied to get frames of interest. Now from these frames features are extracted. Shape based and geometrical features are used. Shape signatures and their Fourier descriptors are used for feature extraction. After features extraction, the system is trained by learning key extracted with General purpose Fuzzy MinMax neural network. When the training is done then testing phase comes. Testing goes through same process and gestures are matched and the result is corresponding text form. This text is further used to create meaningful sentences using grammar. This system got 92.92% accuracy in gesture identification and 100% accuracy in sentence formation by using identified gestures.

There is another system that recognizes ISL and translates it into text [17]. Three phases are used: training, testing and recognition. A new feature vector is created by using Hu invariant moment and structural shape descriptors. For training and recognition multi class support vector machine is used. The system gives 96% recognition rate with 720 images. The system need to focus on independence of signer, large vocabulary systems in isolated as well as continuous recognition tasks.

(IJIP)A system to automatically recognize gestures from video is proposed [18]. The system converts words and sentences of ISL into text and voice. Image processing and artificial intelligence techniques are used. In this edge detection, wavelet transforms and image fusion techniques are used for segmenting shapes in video. For feature extraction it uses Fourier descriptors and uses principal component analysis for feature set optimization and reduction. Input video is compared with database of extracted features using trained fuzzy inference system. 80 gestures from 10 different signers are used with 91% accuracy. The system can be used for training new features as it is robust. The system need to be developed for working in real time.

An approach for low cost sign language recognition is proposed [19]. Different methods are used like hu moments, contour computation, defects computation and convex computation. It is found that Hu moment classification is best approach among others as per requirement. There is system that deals with image having multiple text strings [20]. It uses morphology for extraction of text.

A project called talking hands is developed [21]. In this project sensor glove are used to capture signs of American Sign Language. Then translate that signs into sentences of English language. Sensor values coming from sensor gloves are recognized with the help of artificial neural networks. The values are then categorized in 2 punctuation symbols and 24 English alphabets. With this application, challenging people can write complete sentences.

Sign language recognition using Gabor wavelet transform is developed [22]. The method for sign language recognition works with facial information fusion. Segmentation on the basis of skin tone is done. For hand detection the gesture features are collected and for face the Gabor features are collected. Then facial information

and hand information are combined together. The features are normalized and then further used for recognition.

A signer independent system for Chinese sign language recognition is developed [23]. The system combines feature maps that are self organizing and hidden Markov models. The comparison of results of HMM based system is done. The results of combination of feature maps and hidden Markov models are 5 % more accurate than simple HMM based system. A recognition system, which is self adjusting, is further applied to combined system for the improvement of the results.

A sign recognition system that is viewpoint independent is proposed [24]. In this two, after time wraping samples of same sign are taken as input to a stereo vision system. There is matrix related to two views and that matrix must be unique. Recognition task which is temporal spatial is converted into verification task with stereo vision system. The proposed system reaches both viewpoint and temporal invariance.

Chapter 3

Problem Statement

There is need of a system that can understand signs made by challenging people. The computer system is used so that if a person standing in front of camera makes sign then another person can understand. In this system, a signer can make a sign. On the other side, the other person can know the meaning of that sign. That means if the other person does not know the sign language even then communication is possible. When the person makes sign then it gets recorded and matched with stored signs i.e. pattern matching is done and results are displayed.

Now the problem exists is about creating signs and working on images. A person can use simple background and gloves. But another question arises is that what if it's about real time environment? So the focus will be to recognize signs without using gloves. Input to the system will be images and output will be accuracy. A computer system can understand only binary numbers i.e. 0s and 1s. So it is needed to convert image in a way that computer can understand and its originality remain as of requirements. In this system, its shape should be preserved.

Chapter 4

Research Problem Solving

According to problem statement there is need of system that process images to get the required output. Here is a block diagram showing the working of the process.

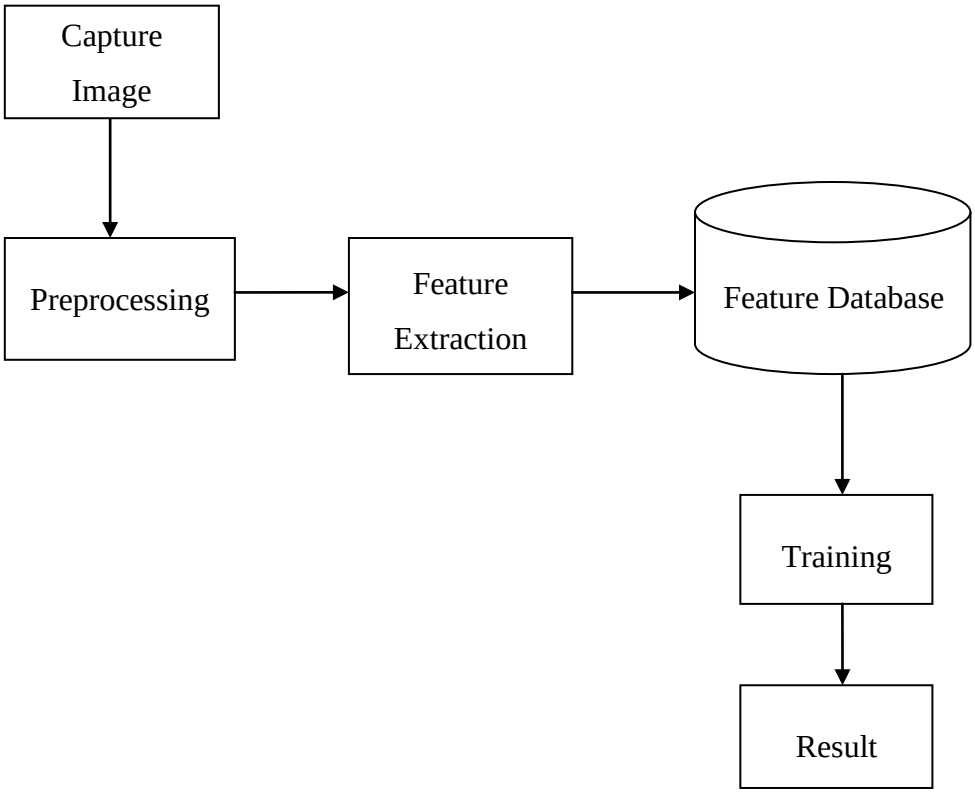

Figure 4.1: Basic Process of Sign Language Recognition

4.1 Image Data Collection

Sign images are collected using webcam. Signs are made using black background. The database is having 23 classes with 10 samples each, i.e. there are 230 total images. Some of sample images are shown below:

ਦੁਲਾਵਾਂ

ਕੰਨਾ

ਲਾਂ

ਸਿਹਾਰੀ

ਟਿੱਪੀ

ਬਿੰਦੀ

A

There are images of Punjabi vowels as well as English alphabets. Some images are shown above.

4.2 Image Processing

Image processing can also be considered as a form of Signal processing. In this, input is an image and output can be a new image itself or a set of parameters related to this image.

4.2.1 Image Scaling

The original RGB image is captured see Figure 4.1. The RGB image has 3 color planes for each frame. The RGB image is then converted into grayscale image. Weighted sum of these R, G and B components result into grayscale image.

Figure 4.3 Original Image

Then image scaling is applied over image i.e. resize the image. Bicubic interpolation is used for image scaling. Bicubic interpolation is cubic interpolation in two dimensions. Bicubic Interpolation considers 16 pixels i.e. 4 by 4 neighborhoods. The image size is changed to [250 250]. Bicubic interpolation is used as it provides smoother surfaces than bilinear interpolation and nearest-neighbor interpolation.

4.2.2 Smoothing

Due to modifications, the resultant image may be having noise and blur. Therefore filter is used to reduce the effect of noise and blur. Gaussian low pass filter is used here. The result of Gaussian filer is reduction of high-frequency components of

Fourier transform of image; so it is called low pass filter.

Further effect of Gaussian filter is that it helps in the improvement of results of edge detection algorithm.

Figure 4.4 Gray scale image

4.3 Canny Edge Detection

In an image, edge is a local feature that separates two regions within a neighborhood. This separation is based on the difference of intensity in the image. Edge is detected in this phase. Before applying canny edge detection, a threshold value is calculated. The threshold function is based on Otsu's method. Otsu's method chooses threshold for minimizing interclass variance between black and white pixels.

Canny edge detector is used to detect a range of edges in images using multi stage algorithm. Canny edge detector is first derivative having noise cleaning. The accuracy of edge detection is improved by using noise smoothing. Approximation of first derivative of Gaussian is done for optimal result by canny edge detector. The result is blurred version of original image.

Another stage is to find the intensity gradient of the image. It uses four filters to detect vertical, diagonal and horizontal edges in blurred image. Then one stage is non maximum suppression, it is an edge thinning technique. Then canny operator trace edges through threshold.

Differential edge detection can also be used to obtain edges. Figure 4.5 shows the result.

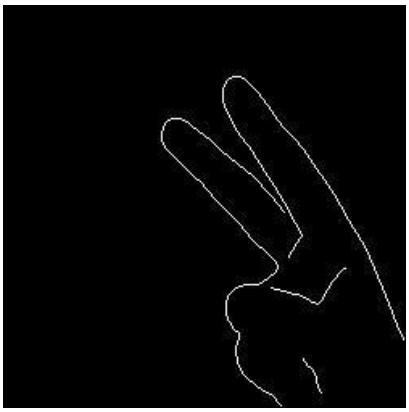

Figure4.5 Canny edge detection

4.4 Improved Sobel Edge Detection

The processed image of processing phase is used to detect edges again using canny edge detector. Improved sobel edge detector [25] is combination of sobel operator and open close functions of dilation and erosion.

Dilation and erosion are morphological operations that process image based on shapes. These operations apply structure elements to an input image and create same size output image. In output image the value of each pixel is obtained by comparison of corresponding pixel of input image and its neighbors.

Dilation is used to add pixels to the object boundaries in image and Erosion removes pixels from object boundaries. The decision on number of pixels to be added or deleted depends upon size and shape of structuring element.

The SE is a matrix containing only 0s and 1s and it can have any arbitrary shape and size. The pixel value 1 defines the neighborhood. Two structuring elements SE1 and SE2 with 3 by 3 and 5 by 5 are created, respectively and shown in below table.

Neighborhood SE1:	0	1	0
	1	1	1
	0	1	0

Neighborhood SE2:	0	0	1	0	0
	0	1	1	1	0
	1	1	1	1	1
	0	1	1	1	0
	0	0	1	0	0

Table4.1 structuring elements

Dilated and eroded images are obtained with these two structuring elements as shown in Figure 4.6.

Figure4.6 Dilated and Eroded image

To smooth the image open and close operation are performed. Open operation is applying erosion followed by dilation. Close operation is applying dilation followed by erosion. Then sobel edge detection operator is applied on the resulting image. Then edge thinning is done. This whole process is improved sobel edge detection. The image obtained is Figure 4.7.

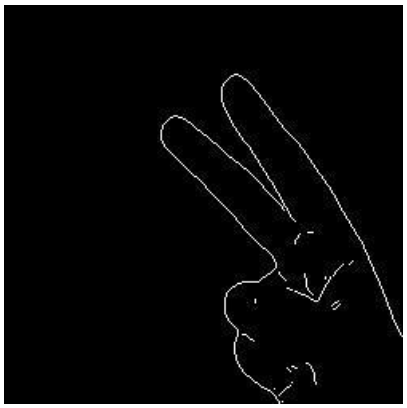

Figure4.7 Improved sobel edge detected image

4.5 Wavelet Image Fusion

Image fusion is combination of information from two or more images that produce result as single image. The images obtained from improved sobel edge detection and canny edge detection is fused together using wavelet image fusion [26]. Block diagram of image fusion is Figure 4.8 The new image is having more information after image fusion.

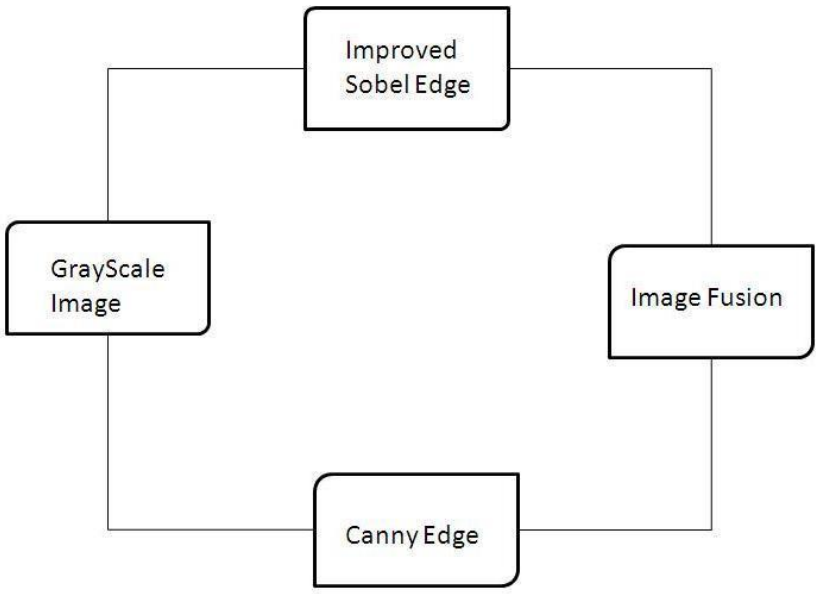

Figure 4.8 Block diagram of image fusion

Further the fusion can be seen in Figure 4.9. The haar wavelet level 2 is used for image fusion.

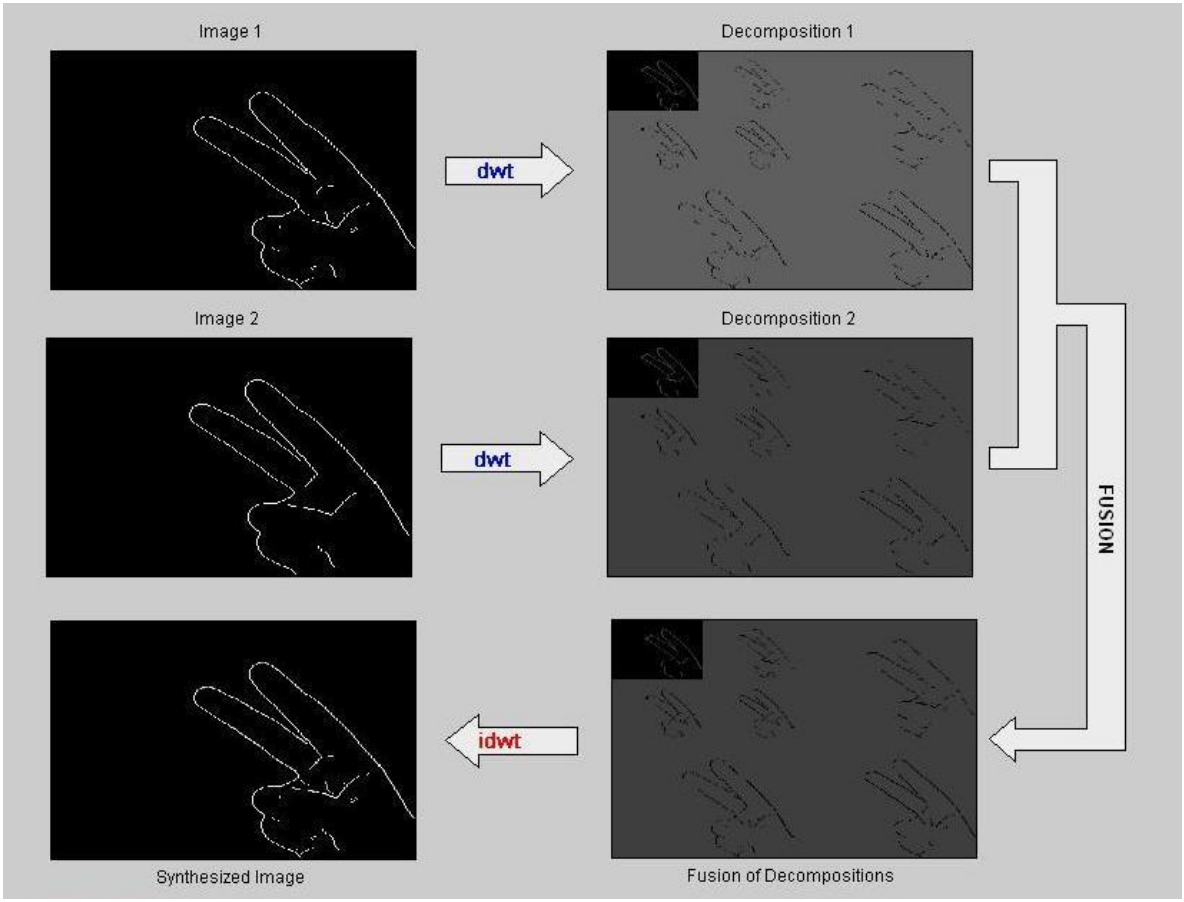

Figure 4.9 Image fusion

4.6 Feature Extraction

Feature Extraction is a form of dimensionality reduction which is used in image processing and pattern recognition. When input data to be processed is very large then this data is transformed into reduced features set, this process is called feature extraction.

Features to be extracted should be chosen carefully, so that relevant information can be collected.

Image is segmented into regions each of which is either object or part of object. Feature extraction is determining properties and attributes associated with an object.

The output of canny edge detection is used to get features. The technique used is Zone based approach [27]. It works as shown in following steps:

1. Image of size [250 250] is divided by into 100 zones. Each zone is having size 25 by 25.

2. The mean and standard deviation for each zone are calculated. Each image is having 100 features correspondingly.

3. Features for all the samples are extracted and stored in three data sets. One data set is Data- Mean. Data-Mean is having all the zonal means of image. Data-Mean stores these features as:

$\{M_1 ; M_2 ; M_3 ; M_4 ; \ldots\ldots;M_{100} \}$

The other data set is Data-Std. Data-Std stores all the zonal standard deviations as:

$\{\sigma_1 ; \sigma_2 ; \sigma_3 ; \sigma_4 ; \ldots\ldots;\sigma_{100} \}$

Third data set stores mean as well as standard deviation one after other i.e. mean followed by standard deviation. It is named as Data-MeS and data is stored in the following manner:

$\{M_1,\sigma_1 ; M_2,\sigma_2 ; M_3,\sigma_3 ; M_4,\sigma_4 ;\ldots\ldots\ldots; M_{100}, \sigma_{s100} \}$

4.7 Pattern Recognition using SVM

SVM is used for image recognition. Support vector machine is supervised learning model that is having excellent ability to find solution for classification problems. SVM works on the concept of decision planes which define decision boundaries. Support vector machine is having statistical classifiers. SVM can perform linear as well as non linear classification. Non linear classification is performed using kernel trick that implicitly mapped inputs to high dimensional feature spaces.

SVM creates hyperplane or set of hyperplanes that is used for regression, classification or other tasks. Many different hyperplanes may do data classification. The best hyperplane among various hyperplanes is selected that represents largest margin or separation among two different classes. So hyperplan is chosen so that

distance from hyperplan to nearest data point is maximized. In other words, SVM by maximizing the margin between separating plan and data one can found the optimal separating hyperplan. SVM with grid and different kernel functions is applied. The kernel functions are linear, polynomial, sigmoid and radial basis. The accuracy obtained by kernel functions is shown in following tables.

Three data sets are taken Data-Mean for mean values of features. Data-Std represents standard deviation and Data-MeS stores mean as well as standard deviation values. Log c and log g are parameters representing logarithmic values of cost and gamma respectively.

Here Table 4.2 is linear function, Table 4.3 is polynomial function, Table 4.4 is radial basis function and Table 4.5 is sigmoid function.

Data Set	**Linear function**		
	Log c	**Log g**	**Accuray**
Data-Mean	**-5**	**-7**	**96.347**
Data-Std	**-5**	**-7**	**95.4338**
Data-MeS	**-5**	**0.5**	**95.8159**

Table 4.2 Parameters of Linear Function for different feature sets

Data Set	Polynomial function		
	Log c	Log g	Accuracy
Data-Mean	32	-13	96.347
Data-Std	-5	-7	95.8904
Data-MeS	-5	-7	95.8159

Table 4.3 Parameters of Polynomial function for different feature sets

Data Set	Radial function		
	Log c	Log g	Accuracy
Data-Mean	8	-13	96.347
Data-Std	2	-15	95.8904
Data-MeS	2	-15	95.8159

Table 4.4 Parameters of Radial function for different feature sets

Data Set	Sigmoid function		
	Log c	Log g	Accuracy
Data-Mean	32	-13	96.347
Data-Std	32	-15	95.4338
Data-MeS	32	-15	95.3975

Table 4.5 Parameters of Sigmoid function for different feature sets

Chapter 5

Conclusion and Future Research

Conclusion:

In this thesis, sign images are analyzed. Bicubic interpolation is used for image scaling instead of bilinear interpolation because bicubic provides smooth results as compared to bilinear interpolation.

Then canny edge is used to detect edges. It is optimal than other edges present there. Improved edge detection is also applied and both edge detection images are fused together. From implementation, the accuracy is obtained that is shown in tables of implementation chapter .On the basis of accuracy, the conclusion can be drawn that accuracy for mean data set is high as compare to standard deviation as well as combination of both mean and standard deviation.

The accuracy for mean of features is calculated as 96.347%.

Future Scope:

Further the system can be converted to sign to speech. The system is able to recognize signs. So further these signs produced text. By using resulting text, it can be converted to speech. One person will make sign and another will listen. The system can be further improved by increasing size and number of training samples. With this more signs can be recognized.

It can also be made working without using any background, a system that will work in real time. In real environment communication can take place with ease. So there can be a system in which communication takes place using computer system. There will be a great change with this system in present scenario for the challenging persons.

REFERENCES

[1]B. Chanda and D. D. Majumder, *Digital image processing and analysis*. PHI Learning Pvt. Ltd., 2004.

[2] T. Starner and A. Pentland, "Real-Time American Sign Language Recognition From Video Using Hidden Markov Models," *Proc. Int',l Symp. Computer Vision,* Coral Gables, Fla., 1995.Los Alamitos, Calif.: IEEE CS Press.

[3] C. Lee and Y. Xu "Online, interactive learning of gestures for human/robot interfaces", *IEEE International Conference on Robotics and Automation*, vol. 4, pp.2982 -2987 1996.

[4] Jiangqin, W., Wen, G., Yibo, S. and Bo, P. ,"A simple sign language recognition system based on data glove", In *Proceedings of the Fourth International Conference on Signal Processing*, 1998.

[5] Hyung-Ji Lee and Jae-Ho Chung, "Hand gesture recognition using orientation histogram", Tencon 99. *Proceedings of the IEEE Region 10 Conference,* South Korea, pp. 1355-1358 Vol. 2, Dec 1999.

[6] J. Weissmann, R. Salomon, "Gesture recognition for virtual reality applications using data gloves and neural networks", *Neural Networks,* pp. 2043-2046 , 1999.

[7] Gupta, Lalit. Suwei Ma, "Gesture-Based Interaction and Communication: Automated Classification of Hand Gesture Contours", *IEEE Transactions on Systems, man and Cybernetics,* PartC: Applications and Reviews, Vol.31, No. 1, February 2001.

[8] M.-C. Su, Y.-X. Zhao, H. Huang and H.-F Chen, "A Fuzzy Rule-Based Approach to Recognizing 3-D Arm Movements," *IEEE Trans. Neural Systems Rehabilitation Eng.*, vol. 9, no. 2, pp. 191-201, June 2001.

[9] Siau, Tan Long, and Ling Li. "An on-line sign language communication system", *Web Information Systems Engineering, 2001, Proceedings of the Second International Conference on*. Vol. 1., IEEE, 2001.

[10] N. Aifanti S. Malassiotis, M. G. Strintzis, “A gesture recognition system using 3d data,” *Proceedings of the First International Symposium on 3D Data Processing Visualization and Transmission*, IEEE, 2002.

[11] Losson, Olivier, and Brigitte Cantegrit, "Generation of signed sentences by an avatar from their textual description", *Systems, Man and Cybernetics, 2002 IEEE International Conference on*. Vol. 3., IEEE, 2002.

[12] Travieso, Carlos M., Jesús B. Alonso, and Miguel A. Ferrer. "Sign language to text by SVM", *Signal Processing and Its Applications, Proceedings, Seventh International Symposium on*. Vol. 2 . IEEE, 2003.

[13] Amitabh Mukherjee, Achla Rana. M Purushottam Kar, Madhusudan Reddy, “Ingit: Limited domain formulaic translation from Hindi strings to Indian sign language,” Indian Institute of Technology Kanpur, India 2011.

[14] Shih-Chung, Hsu-Hung-Wei, Lin Shin-Han Yu, Chung- Lin Huang, Hau-Wei Wang, “Vision-based continuous sign language recognition using product hmm,” IEEE, 2011.

[15] Kose, Hatice, Rabia Yorganci, and Itauma I. Itauma. "Humanoid robot assisted interactive sign language tutoring game." *Robotics and Biomimetics (ROBIO), 2011 IEEE International Conference on*. IEEE, 2011.

[16] Pravin R.Futane, Dr. Rajiv V. Dharaskar, “Video gestures identification and recognition using fourier descriptor and general fuzzy minmax neural network for subset of indian sign language,” *Hybrid Intelligent Systems (HIS), 12th International Conference on*. IEEE, 2012.

[17] Dixit, Karishma, and Anand Singh Jalal, "Automatic Indian Sign Language Recognition System", *Advance Computing Conference (IACC), 2013 IEEE 3rd International*. IEEE, 2013.

[18] E.Kiran Kumar, P.V.V.Kishore, P.Rajesh Kumar, S.R.C.Kishore, “Video audio interface for recognizing gestures of Indian sign language,” *International Journal of Image Processing (IJIP)*, Volume (5) : Issue (4) : 2011.

[19] Matheesha Fernando, Janaka Wijayanayaka, ”Low cost approach for Real Time Sign Language”, *IEEE 8th International Conference on Industrial and Information Systems*, Sri Lanka, Aug. 18-20, 2013.

[20] Wang, Yuming, and Naoki Tanaka, "Text string extraction from scene image based on edge feature and morphology", *Document Analysis Systems, The Eighth IAPR International Workshop on*. IEEE, 2008.

[21] Mehdi, Syed Atif, and Yasir Niaz Khan. "Sign language recognition using sensor gloves." *Neural Information Processing, 2002. ICONIP'02. Proceedings of the 9th International Conference on*. Vol. 5. IEEE, 2002.

[22] Zhishen Huang, Dongyang Jiang, Wentao Zhao ,”Study of Sign Language Recognition Based on Gabor Wavelet Transforms”, International Conference On Computer Design And Appliation, IEEE, 2010.

[23] Fang, Gaolin, Wen Gao, and Jiyong Ma. "Signer-independent sign language recognition based on SOFM/HMM." *Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems. Proceedings. IEEE ICCV Workshop on*. IEEE, 2001.

[24] Qi Wang, Xilin Chen1, Liangguo Zhang, Chunli Wang, Wen Gao, “viewpoint invariant sign language recognition”, School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China, IEEE, 2005.

[25] Deng, Caixia, Weifeng Ma, and Yin Yin, "An edge detection approach of image fusion based on improved Sobel operator", *Image and Signal Processing (CISP), 2011 4th International Congress on*. Vol. 3, IEEE, 2011.

[26] Pajares, Gonzalo, and Jesús Manuel de la Cruz. "A wavelet-based image fusion tutorial." *Pattern recognition* 37.9 (2004): 2004

[27] D.H. Manjaiah H.N Ashoka, Rabindranath Bera, “Zone based feature extraction and statistical classification technique for kannada handwritten numeral recognition,” Sikkim Manipal Institute of Technology, Sikkim Manipal University, Sikkim, India, IEEE 2012.

LIST OF PUBLICATIONS

Accepted Paper

1. Shama, Karun Verma, "Sign Recognition Using Support Vector Machine", will be presented in *International Conference on Emerging Research in Computing, Information, Communication and Applications, ERCICA* 2014 held at Nitte Meenakhsi Institute of Technology, Bangalore, India from August 01 to August 02, 2014.