Tamil Sign Language to Speech Translation

S.Sudha
Department of Computer Science and Engineering,
Annamalai University,
Annamalai Nagar, Chidambaram.

S. JothiLakshmi
Department of Computer Science and Engineering,
Annamalai University,
Annamalai Nagar, Chidambaram.

ABSTRACT

The Sign Language Recognition is the most popular research area involving computer vision, pattern recognition and image processing. Many techniques have been developed recently in these fields. Sign languages are used for communication and interface. There are various types of systems available for Sign Language Recognition. Sign Language Recognition is a very challenging research area. In this paper, a system to recognize static gestures representing Tamil words and speech translation has been proposed. The approach used in this paper is robust and efficient for static hand gesture recognition. Sign language has different applications in many domains like HCI (Human Computer Interaction), Robot control, Security, Gaming, Computer vision etc. A different approach is handled in this work for recognizing Tamil Sign Language. It consists of three phases. The first phase is preprocessing where the images are processed through the steps like resizing, gray conversion, filtering for reducing the distortion and black and white conversion. Second phase is feature extraction. Shape descriptors such as solidity, eccentricity, perimeter, convex area, Major axis length, Minor axis length and orientation are applied to the black and white images to extract the features. Third phase is the classification where a Naive Bayesian classifier is used to recognize the signs from trained set of gestures. The features derived are used to train the classifier first and then the testing images have been introduced for classification. The proposed system is able to recognize images with 90% accuracy.

Keywords: Naïve Bayesian classifier, Shape Descriptors, Sign Language Recognition.

1. INTRODUCTION

A sign language (also called signed language) is a language which, instead of acoustically conveys sound patterns, uses manual communication and body language to convey meaning. Sign language is important in humankind that is showing an increasing research interest in eradicating barriers faced by differently abled people in communicating and contributing to the society [1].

Sign language is used as a communication medium among deaf and dumb people to convey the message with each other. A normal person cannot communicate with those people unless and until they are familiar with sign language. Similarly when a deaf and dumb person wants to communicate with a normal person he/she finds it difficult. Sign language is a substitute of speech for deaf and dumb people.

Indian Sign languages are rich, faceted language, and the complexity is beyond current gesture recognition technologies [2]. The interpersonal communication problem between signer and hearing community could be resolved by building up a new communication bridge integrating components for sign [3].There are different sign languages all over the world such as American Sign language (ASL) [4], British Sign Language (BSL), Japanese Sign Language [6] family (Japanese, Taiwanese and Korean Sign Languages), French Sign Language family (French, Italian, Irish, Russian and Dutch Sign Languages), Australian Sign Language [5], etc.

Similarly Indian Sign Languages was also developed for Indian deaf community. It is different in the phonetics, grammar, hand gestures and syntax from other country's sign languages.TSL uses static and dynamic hand gestures, facial expressions, head/body postures, locations of hand with respect to body etc., to represent signs. It is more challenging than other sign languages for the following reasons:

- Most of the signs make use of the gestures of both the hands together.
- Many of the gestures result in occlusions.
- Hand contacts the body for some signs.
- Locations of the hand with respect to body contribute to the sign.

The main aim of this work is to bridge the gap between a normal person and the deaf and dumb people. An approach is proposed such that it will convert the given static gesture into the corresponding Tamil text. Several techniques have been adopted by different researchers to recognize sign languages. But the rate of accuracy differs based on the features and classifiers used. Thus a special classifier and features have been used to recognize static sign of Tamil sign language. This paper is organized as follows: section (2) describes the work related to this paper. Section (3) describes the work proposed. Section (4) deals with a system design for the proposed work with a detailed description of features and classification methods. Experiments and Results are discussed in section (5).Conclusion of this work and perspectives for future work are discussed in section (6)

2. RELATED WORK

In the last two decades, several different methods of potential applications in the advanced gesture interfaces for HCI have been suggested but these differ from each other in their models. Some of them are Neural Network, Hidden Markov Models [7] and Fuzzy Systems [8]. The hand is detected using attributes like its motion and the skin color [9].

Many novel and interesting applications of hand gesture recognition have been introduced in recent years. Generally, such systems are divided into two basic approaches namely glove based and vision based approaches. In glove based analysis, detection of the hand is eliminated by the sensors on the hand and 3D model of the hand is subjected to the virtual world so that body motion can easily be captured. Christopher Lee and Yangsheng Xu [10] developed a glove-based gesture recognition system. On the other hand vision-based analysis is more natural and useful for real time applications. Sushmita Mitra and Tinku Acharya had provided a survey on gesture

recognition, with detailed description on hand gestures and facial expression [11].

Several methods have been developed to recognize signs. Hand gestures are recognized using edge detection algorithm. In some systems neural network approach is used. To recognize the signs, static hand gestures are grabbed and are converted into feature vectors.

3. PROPOSED WORK

An efficient and accurate technique for Tamil sign recognition is presented in this paper. The main aim of this work is to develop a Tamil sign language recognition system for deaf and dumb people. The proposed work consists of three stages. First stage is preprocessing, were in the sample images are processed by using the following steps i) resizing ii) gray conversion iii) filtering iv) black and white conversion. Second stage is the feature extraction, which extracts the required feature vectors from the output obtained from the first stage. Features like solidity, eccentricity, perimeter, convex area, Major axis length, Minor axis length, orientation are used to obtain the shape. Third phase is the classification where a Naive Bayesian classifier is used to recognize the signs from trained set of gestures.

4. METHODOLOGY USED

A vision based analysis is used in this work. Vision based analysis, is based on the way human beings perceive information about their surroundings, yet it is probably the most difficult to implement in a satisfactory way. Several different approaches have been tested so far.

• One is to build a three-dimensional model of the human hand. The model is matched to images of the hand by one and parameters corresponding to palm orientation and joint angles are estimated. These parameters are then used to perform gesture classification.

• Second one to capture the image using a camera then extract some features and those features are used as input in a classification algorithm for classification. In this work we have used second method for modeling the system. Images are captured using a camera and the features are extracted and for the extracted feature a classifier is applied to classify the signs. Fig (1) describes the steps involved in this work.
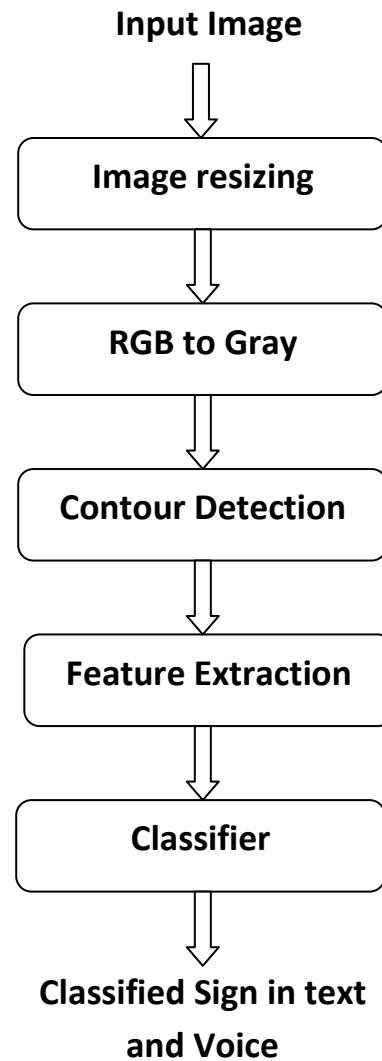
Input Image

⬇

Image resizing

⬇

RGB to Gray

⬇

Contour Detection

⬇

Feature Extraction

⬇

Classifier

⬇

Classified Sign in text and Voice

Fig 1 Overview of the System

The system consists of the following stages

- Input image
- Preprocessing
- Feature extraction
- Classification.

4.1.Image Acquisition

The first stage of any vision system is the image acquisition stage. Static hand gestures and facial gestures were captured using USB connected camera. After the image has been different vision tasks required. Each image represents a unique Tamil sign word. The resolution of the grabbed image is large so it is resized to a resolution of 200 into 200, which is given as input to the next stage of the model i.e., preprocessing. The sample images collected are shown in fig (2).
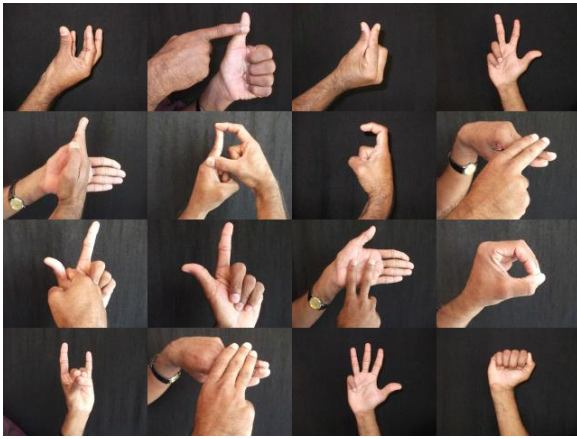
Fig 2 A Collection of Static Gesture Images

4.2.Preprocessing

Preprocessing methods use a small neighborhood of a pixel in an input image to get a new brightness value in the output image. It consists of two steps

- Segmentation
- Gaussian filtering

Segmentation is done to convert gray scale images into binary image. The obtained image has some noise. So it is better to filter those noises using Gaussian filtering approach. This approach helps us to obtain a smooth, closed and complete contour of a gesture. The output obtained at this stage is black and white image which uses the features like RGB to gray conversion, filtering and thresholding. Fig (3) shows the preprocessed image.
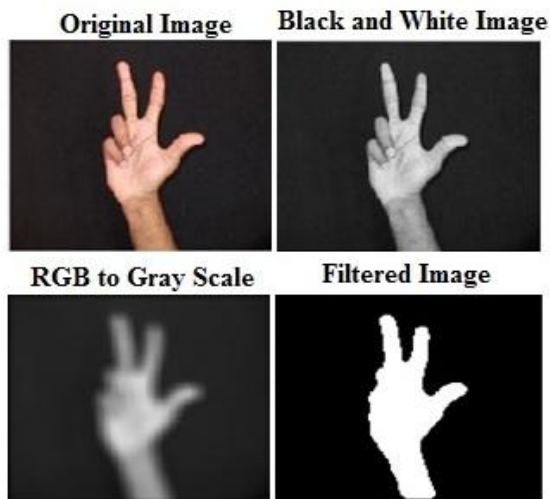

Fig 3 Result of preprocessed image

4.3.Feature extraction

Region-based analysis [12, 13] exploits both boundary and interior pixels of an object. The following are the shape descriptors used as features (a) Solidity (b) Eccentricity (c) Perimeter (d) Convex area (e) Major axis length (f) Minor axis length (g) Orientation. Let us now assess these features

4.3.1.Solidity

A scalar quantity which is defined to be the ratio of area to the convex area of the same object. It is computed as

$$Solidity = \frac{Area}{Convex\ Area}$$

For a solid object or cell, this value is 1.

4.3.2.Eccentricity

A scalar quantity which is defined to be ratio of the major to the minor axis. The value is between 0 and 1. It is given by the equation

$$Eccentricity = \frac{Minor\ length\ axis}{Major\ length\ axis}$$

4.3.3.Perimeter

A scalar quantity which specifies the distance around the boundary of the region. Perimeter is calculated to be the distance between each adjoining pair of pixels around the border of the region. If the image contains discontinuous regions, region props returns unexpected results.

4.3.4.Convex area

Scalar that specifies the number of pixels in 'convex image'. The image is the size of the bounding box of the region. This property is supported only for 2-D input label matrices.

4.3.5.Major axis length

Scalar specifying the length (in pixels) of the major axis of the ellipse that has the same normalized second central moments as the region. This property is supported only for 2-D input label matrices.

4.3.6.Minor axis length

Scalar specifying the length (in pixels) of the minor axis of the ellipse that has the same normalized second central moments as the region. This property is supported only for 2-D input label matrices.

4.3.7.Orientation

Scalar specifying the angle (in degrees ranging from -90 to 90 degrees) between the x-axis and the major axis of the ellipse that has the same second moments as the region. This property is supported only for 2-D input label matrices.

4.4.Classification

Classifier always tries to improve the classification rate by pushing classifiers into an optimized structure [15]. Bayesian theory gives a mathematical calculus of degrees of belief, describing what it means for beliefs to be consistent and how they should change with evidence.

Bayesian Classifier is a fast-supervised classification technique. Bayesian Classifier is suitable for large-scale prediction and classification tasks on complex and incomplete datasets. Naïve Bayesian Classification performs well if the values of the attributes for the sessions are independent. Although this assumption is almost always violated in practice, recent work [16] has shown that naïve Bayesian learning is remarkably effective in practice and difficult to improve upon systematically. We have decided to use the naive Bayesian Classifier in our application, for the following reasons. First, it is efficient for both the training phase and the

recognition phase. Second, its training time is linear in the number of examples and its recognition time is independent of the number of examples. Finally, it provides relatively fine-grained probability estimates that can be used to classify the new session [17]. The computational complexity of Bayesian Classification is fairly low as compared to other classification techniques. Consider a session with f attributes, each with v values. Then with naive Bayesian classifier with e sessions, the training time is $O(ef)$ and hence independent of v.

A **Naive Bayesian** classifier assigns a new observation to the most probable class, assuming that the features are conditionally independent given the class value [14]. It can outperform more sophisticated classification methods by categorizing incoming objects to their appropriate class. The Naive Bayesian classifiers can handle a random number of independent variables whether continuous or categorical.

The proposed classifier Naive Bayes is used to justify the objects using new methods to get a maximum. The features are the parameters extracted from the sign images which are taken from a camera. In each image, a measure of properties is taken to determine the sign in different position. They estimate the probability that a sign belongs to each of the target classes that is predetermined. In the training phase, the training set is used to decide how the parameters must be weighted and combined in order to separate the various classes of signs.

It classifies data in two steps:

Training step: Using the training samples, the method estimates the parameters of a probability distribution, assuming features are conditionally independent given the class.

Prediction step: For any unseen test sample, the method computes the posterior probability of that sample belonging to each class. The method then classifies the test sample according to the largest posterior probability.

Bayes theorem used, takes the equation as given in (1) and (2)

$$P(H|X) = P(X|H)P(H)/P(X) \quad \ldots\ldots (1)$$

It can also be expressed as

$$P(C_i|X) = \frac{P(X|C_i)P(C_i)}{P(X)} \quad \ldots\ldots (2)$$

Where is $P(X)$ constant for all classes, only $P(X|C_i)\,P(C_i)$ need be maximized.

The class-conditional independence assumption greatly simplifies the training step since estimation can be done using one-dimensional class-conditional density for each feature individually. This assumption of class independence allows the Naive Bayesian classifier to better estimate the parameters required for accurate classification while using less training data than many other classifiers. This makes it particularly effective for datasets containing many predictors or features.

5. EXPERIMENTS AND RESULTS

There are no resources to download TSL alphabet image dataset. So after a hard effort in getting the dataset various resources, the Tamil Sign Language database has been made. Factors affecting the recognition result are eliminated by selecting a right set of features. Features are the primary key for this Tamil Sign Language recognition system. Black and

White images are used so that shape descriptors can be applied to extract the following features like solidity, perimeter, Convex area, eccentricity, major axis length, minor axis length and orientation. All the features are interrelated with one another. For example, to measure the eccentricity we need the major axis and minor axis length values so that each feature depends on one another only for the corresponding image and does not depend on features of other images. But in some cases the features like solidity are found to have mere closer values for all set of images even though it does not affect the system performance. Fig (4) describes the output of the proposed work, where the recognized Tamil sign is converted into its corresponding text and voice are displayed.
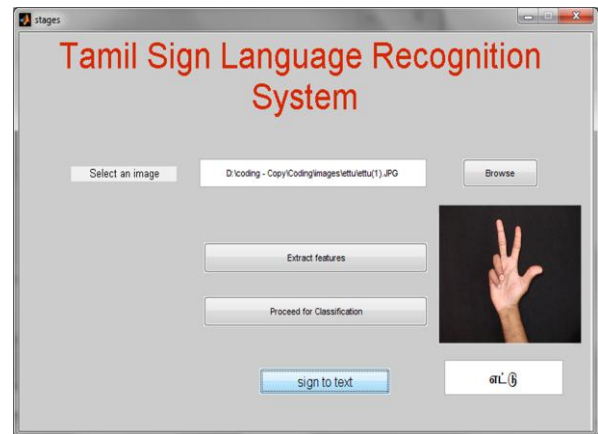

Fig 4 Conversion of Sign to Text

A set of Tamil sign words have been recognized using Mat lab R2009a which is processed in windows 7 Operating System. The proposed work was trained and tested with 41 categories each containing 10 subjects. **Leave-one-out-cross validation** method is used. If model selection and true error estimates are to be computed simultaneously, the data needs to be divided into three disjoint sets.

Training set: a set of examples used for learning, to fit the parameters of the classifier.

Validation set: a set of examples used to tune the parameters of a classifier.

Test set: a set of examples used only to assess the performance of a fully-trained classifier.

The overall accuracy of the proposed method is given below

$$\text{Accuracy} = \frac{\text{No. of Patterns} - \text{No. of False Result Patterns}}{\text{No. of Patterns}} \times 100\%$$

The accuracy rate achieved in this work is 90%.

Table 1 Performance Evaluation

Sign	Valid Recognized sign	Invalid Recognition	Recognition rate (%)
Gjd; (Wednesday)	10	0	100
VO (Seven)	10	0	100
vl;L (Eight)	10	0	100
gpg;utup (February)	10	0	100
QhapW (Sunday)	10	0	100
,e;jpah (India)	9	1	90
,ul;il ,iy (Two leaves)	9	1	90
ldtup (January)	9	1	90
IPd; (June)	9	1	90
flTs; (God)	10	0	100
fUg;G (Black)	8	2	80
nfhiy (Murder)	8	2	80
Nfhop (Hen)	8	2	80
ghdk; (Drink)	7	3	70
tho;f;if (Life)	10	0	100
khL (Cow)	10	0	100
khd; (Deer)	10	0	100
kJiu (Madurai)	7	3	70
khHr; (March)	10	0	100
Nk (May)	10	0	100
Nkhjpuk; (Ring)	8	2	80
KUfd; (Murugan)	7	3	70
ez;gd; (Friend)	10	0	100
etk;gh; (November)	8	2	80
mf;Nlhgh; (October)	8	2	80
Xd;gJ (Nine)	9	1	90
gj;J (Ten)	9	1	90
Ngdh (Pen)	10	0	100
rdp (Saturday)	10	0	100
nrq;fy; (Bricks)	7	3	70
nrg;lk;gh; (September)	10	0	100
nrt;tha; (Tuesday)	10	0	100
rptd; (Sivan)	7	3	70
rptg;G (Red)	8	2	80
#upad; (Sun)	7	3	70
Nfhapy; (Temple)	10	0	100
jpq;fs; (Monday)	10	0	100
jpuhl;ir (Grapes)	7	3	70
tzf;fk; (Welcome)	10	0	100
nts;sp (Friday)	10	0	100
tpahod; (Thursday)	10	0	100
Total		**90.00**	

6. CONCLUSION AND FUTURE WORK

The main aim of this work is to develop a static Tamil sign language recognition system for deaf and dumb people. A large set of data were collected and processed with shape descriptors like Solidity, Eccentricity, Perimeter, Convex area, Major axis length, Minor axis length, Orientation which are classified using Naïve Bayesian classifier. Bayesian Classifier has proved to have the best classification accuracy for static Tamil sign recognition. The results are evaluated and found that Naïve Bayesian classifier works well with 90% of accuracy. In future, the work can be extended to recognize the dynamic signs by using some other classifiers along which will be supplemented with the new features and then the performance of the classification techniques would be evaluated.

7. REFERENCES

[1] Jiong June Phu and Yong Haur Tay "Computer Vision Based Hand Gesture Recognition Using Artificial Neural Network". Faculty of Information and Communication Technology, University Tunku Abdul Rahman (UTAR), MALAYSIA

[2] Tirthankar Dasgupta, Sambit Shukla, Sandeep Kumar, "A Multilingual Multimedia Indian Sign language Dictionary Tool", the 6th Workshop on Asian Language Resources, 2008.

[3] Reza Hassanpour, Stephan Wong Asadollah Shahbahrami, "Vision Based Hand Gesture Recognition for Human Computer Interaction", A Review, International Conference Interfaces and Human Computer Interaction, 2008.

[4] T.Starner, J. Weaver, and A. Pentland, "Real-Time American Sign Language Recognition Using Desk and

Wearable Computer Based Video", IEEE Trans. Pattern Analysis Machine Intelligence, vol.20, no. 12, pp. 1371-1375, Dec. 1998.

[5] M.W. Kadous, "Machine recognition of Australian signs using power gloves: Toward large-lexicon recognition of sign language", Proc. Workshop Integration Gesture Language Speech, pp. 165–174, 1996.

[6] H. Matsuo, S. Igi, S. Lu, Y. Nagashima, Y.akata, and T. Teshima, "The recognition algorithm with noncontact for Japanese sign language using morphological analysis," Proc. Int. Gesture Workshop, 1997, pp. 273–284.

[7] Czapnik Karol, Kasprazak Walodzimierz, Wilkowski Arthur, "Hand Gesture Recognition in Image sequences Using Active Contours and HMMs", 2000.

[8] D. M.Gavrila, "The visual analysis of human movement: A survey", Vol. 47, No. 3, Oct 2005.

[9] Sylvie Gibet and Pierre Francois Marteau, "On Approximation of curvature and velocity for gesture segmentation and synthesis", University de Bretagne France, 2009.

[10] Christopher Lee and Yangsheng Xu, "Online, interactive learning of gestures for human robot interfaces", Carnegie Mellon University, The Robotics Institute, Pittsburgh, Pennsylvania, USA, 1996.

[11] Tinku Acharya, Senior Member, IEEE and Sushmita Mitra, Senior Member, IEEE, "*Gesture Recognition: A Survey*", IEEE Transaction on Systems, Man and Cybernetics- Applications and Reviews, Vol. 37, NO. 3, May 2007.

[12] Dengsheng Zhang and Guojun Lu, "Review of shape representation and description techniques", Pattern Recognition, Vol. 37, 1–19, 2004.

[13] Liviu Vladutu, "Shape Recognition for Irish Sign Language Understanding", International Conference on Simulation, Modeling and Optimization, Vol. 4, 242-247, 2009.

[14] Dymitr Ruta and Bogdan Gabrys "An Overview of Classifier Fusion Methods", Computing and Information Systems, p.1-10, 2000.

[15] Danesh, A., B. Moshiri and O. Fatemi, "Improve text classification accuracy based on classifier fusion methods", Proceeding of the 10th International Conference on Information Fusion, July 9-12, IEEE Computer Society, USA July 9-12, 2007.

[16] P. Domingos, M. Pazzani, Beyond Independence: "Conditions for the Optimality of the Simple Bayesian Classifier", Proceedings of the Thirteenth International Conference on Machine Learning. Bari, Italy: Morgan Kaufmann, (pp. 105-112), 1996.

[17] Elkan. C, "Boosting and Naive Bayesian learning", In proceeding of KDD -97, New Port beach, CA, 1997.