

CURVELET TRANSFORM AND HMM CLASSIFIER BASED SIGN LANGUAGE RECOGNITION SYSTEM

Suresh Anand M,
Department of Computer Science and Engineering,
Sri Sai Ram Engineering College,
Chennai, India
suresh.anandm@rediffmail.com

Mohan Kumar N,
Department of Electronics and Communication Engineering,
SKP Engineering College,
Tiruvannamalai, India
nmkphdju@gmail.com

Abstract: A communication tool in the form of sign language is required for deaf and dumb persons as there is no oral communication possible between them. They perform the conversion of sign languages into voice/text. Recently, many algorithms are developed for this purpose. An Indian Sign Language Recognition (ISLR) system is presented in this paper. It uses curvelet transform based entropy features for the recognition, and the transform is applied only to the segmented hand region. Then, the features of each sign of English alphabets are modelled by a classifier network called Hidden Markov Models (HMM). The system gives an average accuracy of 82.95% using 3rd level features which can help to reduce the communication gap between deaf-dumb and normal people in the world.

Keywords: Curvelet Transform, Sign Language, Entropy Features, HMM Classifier.

I. INTRODUCTION

Sign language using hand gesture is one of the emerging technologies for hearing impaired community. It helps deaf and dumb persons to understand the meaning of languages in the world by the use of a communication system using hand gesture images. Some of the previously available SLR systems are reviewed here. The sign language recognition based on K-Nearest Neighbour (KNN) search is discussed in [1]. The input hand gesture images are preprocessed, and then edge detection is done to detect the shape of the alphabets, and the discontinuities present between the edges are joined. At last the fuzzy rule is applied to recognize the sign language.

HMM classifier model is most widely used in speech recognition and recently in the hand gesture recognition [2]. Literally, it is perfect for visual recognition of hand gestures to identify the English alphabets. Then a method of HMM based system is implemented to recognize the sentences in American Sign Language (ASL).

A method of automatic Chinese SLR is discussed in [3]. It uses the surface Electromyographic (sEMG) and portable Accelerometer (ACC) sensors for its operation. From the extracted features of sEMG and ACC data, the basic components like movement and hand shape orientation are modeled. These features are learned by component classifiers. The process of speech reading of British Sign Language (BSL) with a functional magnetic resonance imaging is used

in [4]. At the temporal-parietal-occipital junction, a higher activation is given at BSL processing. Distinct regions that are present between the temporal cortexes are not only sensitive but also differentially sensitive to the different articulators within the signed language.

A sequential data classification method is used for the glove based Arabic SLR system in [5]. A preprocessing and feature extraction techniques are used in order to emphasize and capture the temporal dependence of the data. Then a Modified KNN (MKNN) is used for the classification process. The vision based sign language recognition system is explained in [6] for the translation of signs to English. It is an automated method in which the signs are breaking up into convenient subunits. A boosting algorithm is used to learn the weak classifiers subset that represents the combinations of subunits and the features for each sign.

An automatic system of identifying and training signs from continuous sentences of sign language is explained in [7]. The multiple learning density matrix algorithms are used to extract the signs of noisy and weak supervision text translation from full sentences. The isolated samples are extracted in order to train the hand posture classifiers and the spatiotemporal gestures. Translator system of ISL using a Gesture recognition algorithm is explained in [8]. The system translates the ISL gestures including alphabets, numbers, and phrases into English. Using the combinational and template matching algorithm, the system performs the data acquisitions followed by pre-processing for the recognition of the sign.

The pipeline process system that allows the users to attain the Italian language text translation in the virtual avatar signing animation is discussed in [9]. It gives a brief explanation about the translation system architecture and the integration of resource languages in it. A model of similarity measurement for the vision and sign language semantics from the video content in Chinese sign language is explained in [10]. The volume local binary patterns are defined as the distance between the vision components of the model. For measuring the shape and size, the contour is used whereas for measuring the movement, location and quantizing the trajectory is used in the sign language semantic.

The binary representation of UP & DOWN positions of a set of 32 signs of the five fingers is explained in [11]. The pre-processing for the static images using the feature point method is done and 10 number of each sign images are used for training. For the identification of the finger tip position and the conversion of image into text the IP technique over the static images. For the extraction of the exemplars from signer independent hidden Markov models, affinity propagation is used in [12]. The method is used to shape the adaptive training vocabulary without any human intervention. The grouping of iterative vector field smoothing is done from the collected sign gestures vocabulary, and is used to generate the maximum posterior.

II. RESEARCH FRAMEWORK

The framework of Indian sign language recognition system is shown in Figure 1. The Indian SLR system consists of the following modules; pre-processing by thresholding, feature extraction by curvelet transform and recognition by HMM.

A. Preprocessing

The Indian sign languages are recognized by using hand gesture images. At first, the input images are pre-processed by colour space conversion from RGB

to gray scale. Then a simple thresholding approach is applied to separate the foreground from the background region. Hence, the hand region is separated from the background.

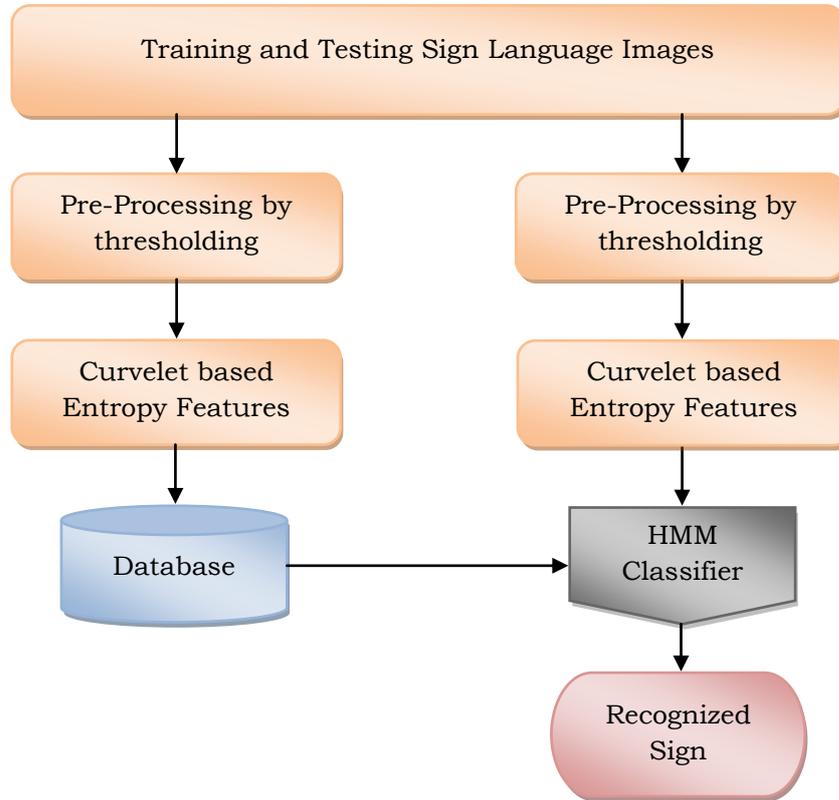


Fig. 1 Indian Sign Language Recognition system

B. Curvelet Based Entropy Feature Extraction

From the segmented hand region, the entropy features are extracted after decomposing it by curvelet transform algorithm. The multi-scale representation of the image is obtained by Curvelets, a non-adaptive technique which can represent edges and curves singularities effectively. Curvelet captures point discontinuities into linear structures. From the decomposed sub-bands obtained from the curvelets, the entropy feature known as the log entropy is extracted. It characterizes the information's in an image which is uncertainty in nature.

As an image could be viewed as a combination of different blocks, the information's available in each block is described by its entropy. The intensity entropy of an image $f(x, y)$ with dimension M by N is defined as

$$En = - \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} p_{ij} \log p_{ij} \quad (1)$$

C. Classification by HMM

It is the final stage of sign language recognition system by curvelets. After the features are extracted from the hand gesture image by using the curvelet transform, the classification is done to make the final decision by HMM classifier. HMM is applied in many pattern recognition approaches such as handwriting, speech, and hand gesture recognition. The first step for the classification is to divide the dataset for each class into two sets. Then train the HMM per class using one set. The remaining set is used for testing which is based on the likelihood of each model. They are related to the use of Markov process rather than independent of each other.

III. RESULTS AND DISCUSSION

The evaluation of the sign language recognition system using the curvelet based entropy features and HMM classifier is completed using hand gestures images which are acquired using 5 megapixels Logitech C170 digital camera. The vertical and horizontal distance between camera and person is fixed at 105 cm and 80 cm respectively. In order to differentiate the hand gesture from the background, all the persons are asked to wear a black jacket. 100 hand gesture images per alphabet are captured from different individuals. Figure 2 shows hand gesture of English alphabets in the database and the outputs of preprocessing steps are shown in Figure 3.

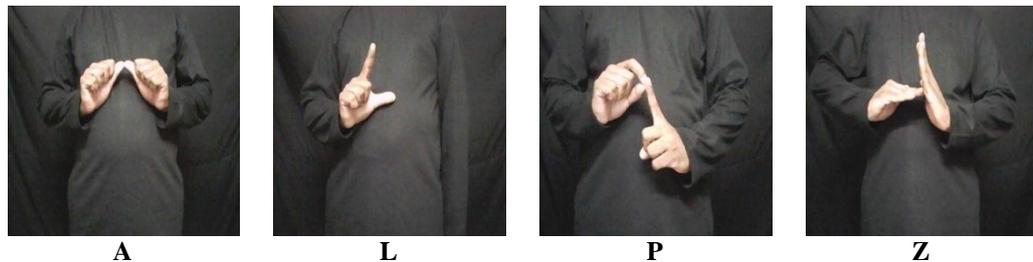


Fig. 2 Sample images in the database

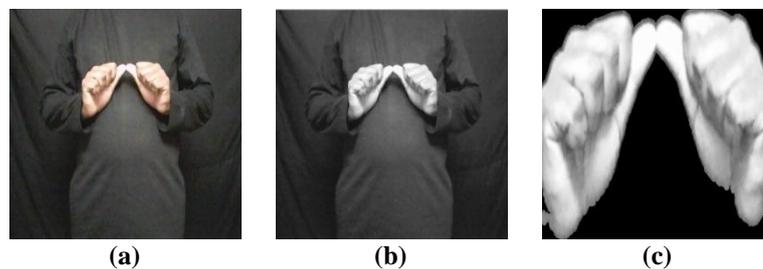


Fig. 3 (a) Input hand gesture Image (b) RGB to Gray converted image (c) Segmented image

The performance of the system is measured by classification accuracy which is defined as the ratio between correctly classified hand gestures images and tested images of the same Alphabet. It is already stated that curvelet is a multi-scale representation. Hence, the system is analyzed from 1-5 scale of features. Figure 4 shows the average accuracies at each scale.

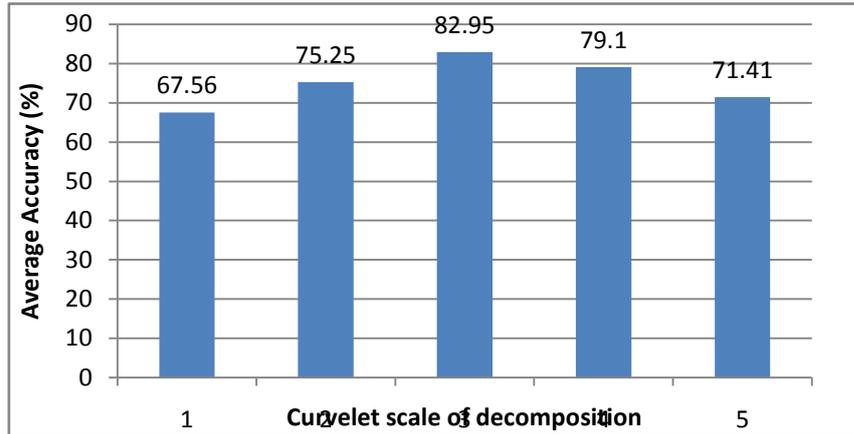


Fig. 4 Average Accuracies vs. Curvelet scale of decompositions

It is observed that the ISLR gives maximum average accuracy at 3rd scale of curvelet features with HMM classifier. Table 1 shows the accuracies of Indian SLR system for each English alphabet.

TABLE 1 Performance of Indian SLR system at 3rd level of decomposition

Alphabets	Recognition Accuracy (%)	Alphabets	Recognition Accuracy (%)
A	86	N	84
B	85	O	86
C	85	P	89
D	80	Q	80
E	78	R	84.8
F	78.4	S	90
G	77.6	T	77
H	91	U	82
I	84	V	81
J	80	W	77.4
K	83.6	X	84
L	80	Y	78.6
M	88.2	Z	86.2
Average		82.95	

IV. CONCLUSION

In this paper, an approach for ISLR is proposed by curvelet based entropy features and the HMM classifier. Hand gesture images are used to find the sign of one of the 26 English alphabets. It is achieved by designing three modules; pre-processing by thresholding, feature extraction by curvelet transform and recognition by HMM. The evaluation of the system is carried on the collected hand gesture image samples for all the 26 alphabets. By using this system, a maximum

of 82.95% overall accuracy is obtained from the hand gesture images while using 26 English alphabets with HMM classifier.

REFERENCES

- [1]. E.A. Kalsh and N.S. Garewal, "Sign Language Recognition system", International journal of computational engineering research, Vol. 3, No. 6, 2013, pp. 16-21.
- [2]. T. Starner and A. Pentland, "Real-time American Sign Language recognition from video using hidden Markov models", In Motion-Based Recognition, Springer Netherlands, 1997, pp. 227-243.
- [3]. Y. Li, X. Chen, X. Zhang, K. Wang and Z.J. Wang, "A sign-component-based framework for Chinese sign language recognition using accelerometer and sEMG data", IEEE transactions on biomedical engineering, Vol. 59, No. 10, 2012, pp. 2695-2704.
- [4]. C.M. Capek, D. Waters, B. Woll, M. MacSweeney, M.J. Brammer, P.K. McGuire and R. Campbell, "Hand and mouth: Cortical correlates of lexical processing in British Sign Language and speech reading English", Journal of Cognitive Neuroscience, Vol. 20, No. 7, 2008, pp. 1220-1234.
- [5]. N. Tubaiz, T. Shanableh and K. Assaleh, "Glove-based continuous Arabic sign language recognition in user dependent mode", IEEE Transactions on Human-Machine Systems, Vol. 45, No. 4, 2015, pp. 526-533.
- [6]. J. Han, G. Awad, and A. Sutherland, "Boosted subunits: a framework for recognising sign language from videos", IET Image Processing, Vol. 7, No. 1, 2013, pp. 70-80.
- [7]. D. Kelly, J. Mc Donald, and C. Markham, "Weakly supervised training of a sign language recognition system using multiple instance learning density matrices", IEEE Transactions on Systems, Man, and Cybernetics, Vol. 41, No. 2, 2011, pp. 526-541.
- [8]. P.C. Badhe, and V. Kulkarni, "Indian sign language translator using gesture recognition algorithm", IEEE International Conference on Computer Graphics, Vision and Information Security, 2015, pp. 195-200.
- [9]. D. Barberis, N. Garazzino, P. Prinetto, G. Tiotto, A. Savino, U. Shoaib, and N. Ahmad, "Language resources for computer assisted translation from Italian to Italian sign language of deaf people", In Proceedings of Accessibility Reaching Everywhere AEGIS Workshop and International Conference, Brussels, 2011, Belgium.
- [10]. L.C. Wang, R. Wang, D.H. Kong, and B.C. Yin, "Similarity assessment model for Chinese sign language videos", IEEE Transactions on Multimedia, Vol. 16, No. 3, 2014, pp. 751-761.
- [11]. P.S. Rajam, and G. Balakrishnan, "Real time Indian sign language recognition system to aid deaf-dumb people", IEEE 13th International Conference on Communication Technology, 2011, pp. 737-742.
- [12]. Y. Zhou, X. Chen, D. Zhao, H. Yao, and W. Gao, "Adaptive sign language recognition with Exemplar extraction and MAP/IVFS", IEEE signal processing letters, Vol. 17, No. 3, 2010, pp. 297-300.
- [13]. E. Candes, L. Demanet, D. Donoho and L. Ying, "Fast discrete curvelet transforms". Multiscale Modeling & Simulation, Vol. 5, No. 3, 2006, pp. 861-899.