Indian Sign Language Gesture recognition

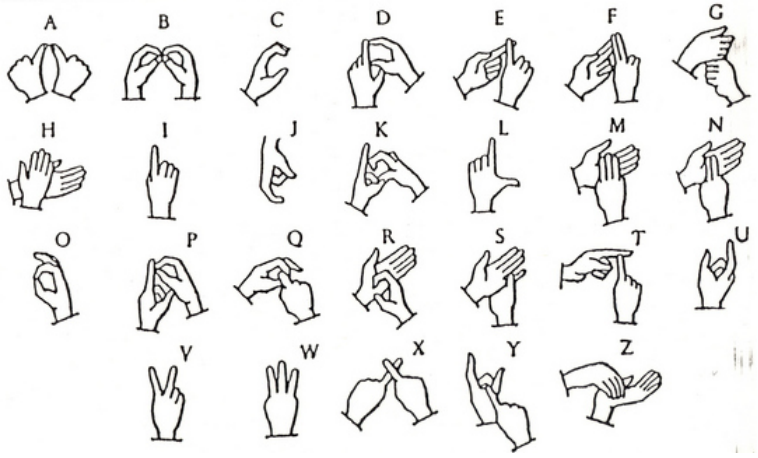
Sanil Jain(12616) and K.V.Sameer Raja(12332)

March 16, 2015

1 Objective

This project aims at identifying alphabets in Indian Sign Language from the corresponding gesture. Gesture recognition and sign language recognition has been a well researched topic for American Sign Language but has been rarely touched for its Indian counterpart. We aim to tackle this problem but instead of using high end technology like gloves or kinect for gesture recognition, we aim at recognition from images(which can be obtained from say webcam) and then use computer vision techniques and machine learning techniques for extracting relevant features and subsequent classification.

Figure 1: Sign Language Alphabets



2 Motivation

Communication is one of the basic requirement for survival in society. Deaf and dumb people communicate among themselves using sign language but normal people find it difficult to understand their language. Extensive work has been

done on American sign language recognition but Indian sign language differs significantly from American sign language.ISL uses two hands for communicating(20 out of 26) whereas ASL uses single hand for communicating. Using both hands often leads to obscurity of features due to overlapping of hands. In addition to this, lack of datasets along with variance in sign language with locality has resulted in restrained efforts in ISL gesture detection. Our project aims at taking the basic step in bridging the communication gap between normal people and deaf and dumb people using Indian sign language. Effective extension of this project to words and common expressions may not only make the deaf and dumb people communicate faster and easier with outer world, but also provide a boost in developing autonomous systems for understanding and aiding them.

3 Dataset

No official datasets were available for Indian Sign Language and the few videos we found on internet were taken by people just describing what it looked like and not the ones that actually spoke the language. So we went to **Jyori Badhir Vidayalaya**, a school for deaf in a much remote section of Bithoor, Kanpur for collecting videos of the students there making gestures for different alphabets. We made around a minute length video of every alphabet taken from over 15 students using a 30 fps camera, which roughly evaluates to around 1800 frames per alphabet. We wanted to take numerals as well per decided against it as the principal told that gestures for 0,2,3 is same as that for o,v,w and the interpretation is context dependent.
Note :- We already mentioned that gestures in Indian Sign Language alphabets are subject to regional variance and even there we found that there existed multiple signs(both dynamic and static) for some alphabets. We asked the students to use the most commonly used static sign for such alphabets.

4 Approach

In this project we are targeting static gestures First we would need to extract the relevant information from the image.

- The first task would be to separate the skin pixels from the non skin pixels as the latter don't contribute to the gesture. We consider this using the approach described in [4] in which the author distinguishes skin pixels from non skin pixels constructing YIQ and YUV color models from RGB model and placing constraints on values of I and $\theta = \tan^{-1}(V/U)$.

- All the skin captured need not be relevant as elbows and other parts of skin can often come in the image. So given the skin segmented image, we need to extract the hand features in this image for better gesture recognition. Some techniques for hand feature extraction are suggested in [3]. We hope to accomplish that using opencv library in python.

- A few alphabets use single hand, and rest use double hand and we intend to first classify an image as using single or two hands and then train on these two classes separately.
- Algorithms like PCA,Isomap etc can be applied on these features to get the most significant features which can then be fed into machine learning models for training.
- Since using two hands often results in occlusion, one approach we are considering is using the HOG features extracted from the skin segmented image for training as described in [2].

Once the relevant features are extracted, we can process them using following machine learning models

- Multiclass Support Vector Machines
- Artificial Neural Networks [1]
- Random Forests

Many alphabets in two hands are very close in gesture(m,n and r,e and f) and we might need to come up with heuristics for them (based on hand features like number of fingers etc). And due to this closeness, unsupervised approaches like Kmeans clustering may not work well, when two or more clusters overlap quite significantly.

References

[1] Justin K. Chen, Debabrata Sengupta, R. R. S. Sign language gesture recognition with unsupervised feature learning, 2011.

[2] Neha V. Tavari, P. A. V. D. Indian sign language recognition based on histograms of oriented gradient. *International Journal of Computer Science and Information Technologies 5*, 3 (2014), 3657–3660.

[3] Padmavathi . S, Saipreethy.M.S, V. Indian sign language character recognition using neural networks. *IJCA Special Issue on Recent Trends in Pattern Recognition and Image Analysis, RTPRIA* (2013).

[4] Xiaolong Teng, Biani Wu, W. Y., and Liu, C. A hand gesture recognition system based on local linear embedding, April 2005.