Sign language recognition using sensor gloves

Conference Paper · December 2002

DOI: 10.1109/ICONIP.2002.1201884 · Source: IEEE Xplore

2 authors:

Syed Atif Mehdi
Technische Universität Kaiserslautern

Yasir Niaz Khan
University of Lahore

Sign Language Recognition using Sensor Gloves

Yasir Niaz Khan
Syed Atif Mehdi
FAST-National University of Computer and Emerging Sciences, Lahore.
Email: MSCS030@nu.edu.pk, MSCS027@nu.edu.pk

Abstract

This paper examines the possibility of recognizing sign language gestures using sensor gloves. Previously sensor gloves are used in games or in applications with custom gestures. This paper explores their use in Sign Language recognition. This is done by implementing a project called "Talking Hands", and studying the results. The project uses a sensor glove to capture the signs of American Sign Language performed by a user and translates them into sentences of English language. Artificial neural networks are used to recognize the sensor values coming from the sensor glove. These values are then categorized in 24 alphabets of English language and two punctuation symbols introduced by the author. So, mute people can write complete sentences using this application.

Keywords

American Sign language, Neural Network, sensor gloves, language recognition, deaf,

Content Areas

Neural Networks
Human Computer Interaction
Natural language understanding
Machine learning
Artificial Intelligence

1. Sign Languages

Sign language is the language used by deaf and mute people. It is a combination of shapes and movements of different parts of the body. These parts include face and hands. The area of performance of the movements may be from well above the head to the belt level. Signs are used in a sign language to communicate words and sentences to audience.

A gesture in a sign language, is a particular movement of the hands with a specific shape made out of them. Facial expressions also count toward the gesture, at the same time. A posture on the other hand, is a static shape of the hand to indicate a sign.

A sign language usually provides signs for whole words. It also provides signs of letters to perform words that don't have a corresponding sign in that sign language. So, although sentences can be made using the signs for letters, performing with signs of words is faster. The sign language chosen for this project is the American Sign Language.

1.1. American Sign Language

It is the most well documented and most widely used language in the world. American Sign Language (ASL) is a complex visual-spatial language that is used by the Deaf community in the United States and English-speaking parts of Canada. It is a linguistically complete, natural language. It is the native language of many Deaf men and women, as well as some hearing children born into Deaf families. ASL shares no grammatical similarities to English and should not be considered in any way to be a broken, mimed, or gestural form of English.

2. Sensor Gloves

Sensor gloves are normally gloves made out of cloth with sensors fitted on it. Using data glove is a better idea over camera as the user has flexibility of moving around freely within a radius limited by the length of wire connecting the glove to the computer, unlike the camera where the user has to stay in position before the camera. This limit can be further lowered by using a wireless camera. The effect of light, electric or magnetic fields or any other disturbance does not effect the performance of the glove.

We have used 7-sensor glove of 5DT company. It has 7 sensors on it. 5 sensors are for each finger and thumb. One sensor is to measure the tilt of the hand and one sensor for the rotation of the hand. Optic fibers are mounted on the glove to measure the flexure of fingers and thumb. Each sensor returns an integer value between 0 and 4095. This value tells about the bent of the sensor. 0 means fully stretched and 4095 means fully bent. So, we get a range of 7 * 4096 combinations as our input.

3. Previous work

Previously, sensor gloves have been used in games for creating virtual 3D environments. Players can give input to the game using the gloves. Gloves, along with other sensor devices, have also been used in making games. Actions of the experts wearing the sensors are captured and translated into the game to give a realistic look to the game. Sensor gloves have also been used in giving commands to robots. Streams of shapes of the hand are defined and then recognized to control a robotic hand or vehicle.

Glove-TalkII is a system that translates hand gestures to speech through an adaptive interface. Currently, the best version of Glove-TalkII uses several input devices (including a Cyberglove, a ContactGlove, a polhemus sensor, and a foot-pedal), a parallel formant speech synthesizer and 3 neural networks. One subject was trained to use Glove-TalkII. After 100 hours of practice he is able to speak intelligibly. The subject passed through 8 distinct stages while he learned to speak. His speech is fairly slow (1.5~to~3 times slower than normal speech) and somewhat robotic. Reading novel passages intelligibly usually requires several attempts, especially with polysyllabic words. Intelligible spontaneous speech is possible but difficult.

4. Sign Language Recognition

Our system is aimed at maximum recognition of gesture without any training. This makes the system usable at public places where there is no room for long training sessions. The speed of gesture capturing and recognition can be adjusted in the application to incorporate both the slow and fast performers of ASL.

Since a glove can only capture the shape of the hand and not the shape or motion of other parts of the body, e.g. arms, elbows, face, etc. so only postures are taken in this project. Signs for letters 'j' and 'z' are ignored as they involve moving gestures. Two custom signs have been added to the input set. One is for space between words and the other is for full stop. These are not part of sign language, but have been added to facilitate in writing the English equivalent of the sentence being performed.

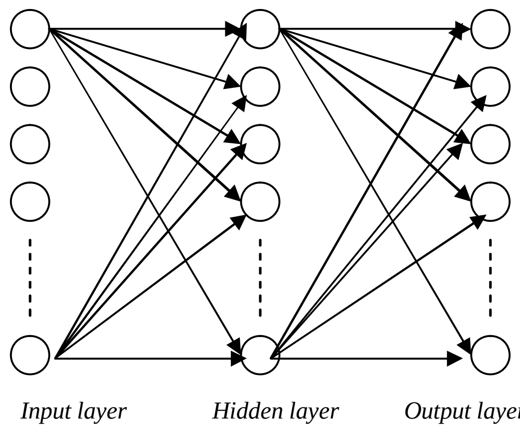

Figure 1: Model of Neural Network used in the project. Input, hidden and output layers contain 7, 54 and 26 neurons (nodes) respectively.

4.1. Neural Network Model

Artificial Neural Network with feed forward and back propagation algorithms have been used. Feed forward algorithm is used to calculate the output for a specific input pattern. Back propagation algorithm is used for learning of the network. Three layers of nodes have been used in the network. First layer is the input layer that takes 7 sensor values from the sensors on the glove. So this layer has 7 sensors. This layer does not do any processing and just passes the values forward.

Next layer is the hidden layer, which takes the values from the input layer and applies the weights on them. This layer has 52 nodes. This layer passes its output to the third layer. The third layer is the output layer, which takes its input from the hidden layer and applies weights to them. There are 26 nodes in this layer. Each node denotes one alphabet of the sign language subset. This layer passes out the final output.

A threshold is applied to the final output. Only the values above this threshold and considered. If none of the nodes give an output above the threshold value, no letter is outputted. If more than one node gives a value above the threshold, no letter is outputted. The activation function used is the sigmoid function. This activation function is applied at both of the processing layer after the weights have been applied. This function is used in processing and learning of the network.

Sampling is done 4 times a second. The user must keep the sign performed for $3/4^{th}$ of a second to get it recognized. This limit can be lowered for faster performers.

4.2. Results

The accuracy rate of the software was found to be 88%. This figure is lower due to the fact that training was done on the samples of people who did not know sign language and were given a handout to perform the signs by reading from it. So, there was great deal of variation in the samples. Some samples even gave completely wrong readings of the sensors. Testing was also done on the same kind of people.

4.3. Problems

One problem that was faced in the project was that some of the alphabets involved dynamic gestures. These may not be recognized using this glove. So these were left out from the domain of the project. Also, some gestures require use of both hands. This requires two sensor gloves.

4.4. Proposed solutions

The problem of dynamic gestures can be resolved using sensors on the arm as well. Sensors would be required at elbow and perhaps shoulder. Hidden Markov Models can be employed to recognize the sequence of readings given by moving hands.

4.5. Future work

As mentioned above, signs of sign languages are usually performed not only with hands but also with facial expressions. One big extension to the application can be use of sensors (or cameras) to capture facial expressions.

Sign languages are also space-dependant. This means that the space (relative to the body) where the gestures are performed also contributes to sentence formation. Sensors would be needed to detect the relative space where the gestures are performed.

Sign languages, as spoken languages, have certain rules of grammar for forming sentences. These rules must be taken into account while translating a sign language into a spoken language. Rules of the targeted spoken language must also be considered into account. In the end, adding a speech engine to speak the translated text would help enhance ease of use.

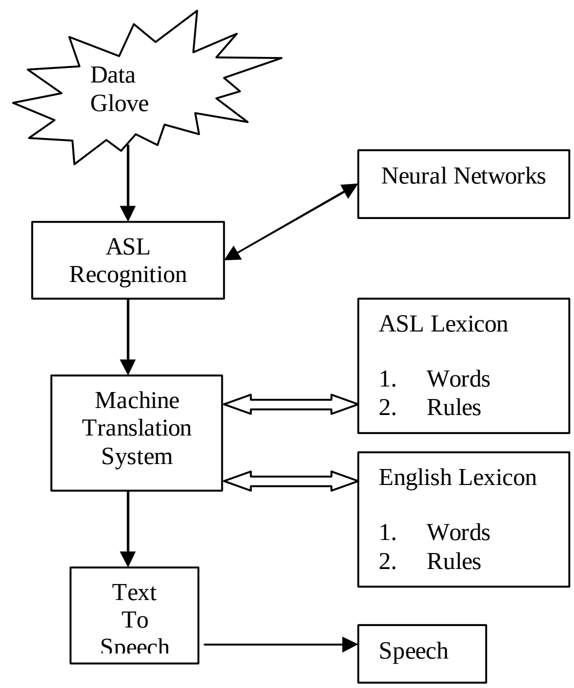

Figure 2: Model of an application that can fully translate a sign language into a spoken language.

5. Conclusion

This project was meant to be a prototype to check the feasibility of recognizing sign languages using sensor gloves. The completion of this prototype suggests that sensor gloves can be used for partial sign language recognition. More sensors can be employed to recognize full sign language.

6. Application

The product generated as a result can be used at public places like airports, railway stations and counters of banks, hotels etc. where there is communication between different people. In addition to this a mute person can deliver a lecture using it.

Assuming the fact that we are able to convert whole of American Sign Language into spoken English, we can manufacture a handy and portable hardware device having this translating system built in as a chip. With the help of this hardware device, which has built in speakers as well, and group of body sensors along with the pair of data gloves a mute person can communicate to any normal person anywhere. A special dress can also be designed having the required number of sensors at appropriate places for this purpose. This will almost bridge the communication gap present between the deaf community and the normal world.

7. Reference

[1] http://www.bconnex.net/~randys/index1.html

[2] http://www.acm.org/sigchi/chi95/Electronic/documnts/papers/ssf_bdy.htm

[3] http://where.com/scott.net/asl/index.html

[4] http://www-white.media.mit.edu/~testarne/asl/asl-tr466/index.html

[5] Charlotte Baker Shenk & Dennis Cokely, *American Sign Language, A teacher's resource text on Grammar and Culture,* Clerc Books Gallaudet University Press, Washington D.C., 1981.