Malay Sign Language Gesture Recognition System

Tan Tian Swee, Sh-Hussain Salleh, A.K. Ariff, Chee-Ming Ting, Siew Kean Seng,
and Leong Seng Huat

Center for Biomedical Engineering
Universiti Teknologi Malaysia
81310 UTM Skudai
Johor, Malaysia
E-mail: tantianswee@hotmail.com, amarism@yahoo.com, hussain@fke.utm.my, eddie_siew@yahoo.com,
shleongst@yahoo.com.sg

Abstract*—This paper describes hardware design and sensors setting and configuration for Malay Sign Language Gesture Recognition systems. A set of sensors consists of accelerometers and flexure sensors has been setup to capture the movement or gesture of shoulder, elbow, wrist, palm and fingers. This project is based on the need of developing an electronic device that can translate the sign language into speech (sound) in order to enable communication to take place between the mute and deaf community with the common public. Hence, the main objective of this project is to develop a system that can convert the sign language into speech so that deaf people are able to communicate efficiently with normal people. This Human-Computer Interaction system is able to recognize 25 common words signing in Bahasa Isyarat Malaysia (BIM) by using the Hidden Markov Models (HMM) method. Both hands are involved to perform the BIM with all the sensors connect wirelessly to a PC with a Bluetooth module. This project aims to capture the hand gestures which involve multiple axis of movement. Altogether 24 sensors have been setup in different hand locations to capture hand and wrist movement in different directions.

***Keywords*— Sign Language Recognition, HMM, Speech Synthesis**

I. *Introduction*

Sign language is a communication skill that uses gestures instead of sound to convey meaning - simultaneously combining hand shapes, orientation and movement of the hands, arms or body, and facial expressions to express fluidly a speaker's thoughts [1][2]. Sign languages are commonly developed in deaf communities, which can include interpreters and friends and families of deaf people as well as people who are deaf or hard of hearing themselves [3].

In Malaysia, Bahasa Isyarat Malaysia (BIM) is the sign language that is commonly used by the deaf community [4]. Therefore the design of the sign language recognition system will be based on the BIM, in order to suit the local people best as well as to benefit them.

Figure 1 shows the possible gesture that hands can make. The hand gesture for sign language is not limited to wrist and finger movement alone, but also involves the movement of elbow and arm of hand in different axis as shown in Table 1 [6]. Most of the projects to capture hand gesture is done using data glove [7-9]. But it is hard to capture or detect the complicated movement by flexture sensors alone. Thus, accelerometer 2-axis and 3-axis were installed at the wrist, elbow and shoulder to capture the different axis of movements vertically, horizontally or diagonally [10].

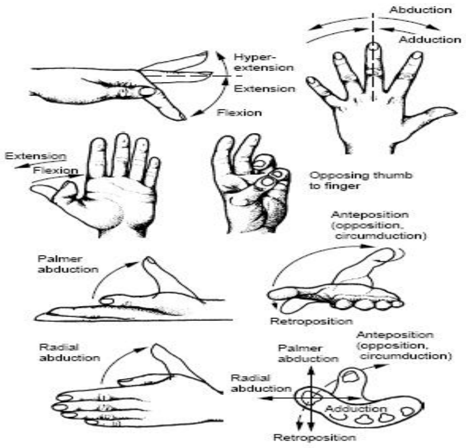
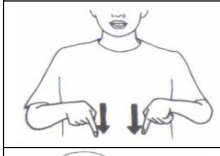
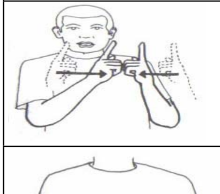
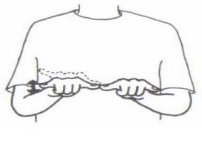
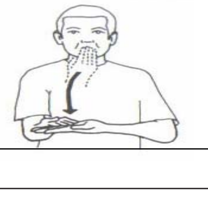


Figure 1: Possible Gesture that Hands can make in Sign Language [5]

Table 1: Example of Malay Sign Language that involved the movement of different hand axis [6]

Gesture Picture	Description
	HARI INI (TODAY) Gesture Movement For 'Hari Ini' (Today) Sign : Both hand indicating sign "D" with the index fingers pointing down simultaneously at the abdomen level.
	JUMPA (MEET) Gesture movement for 'Jumpa' (Meet) sign : Both hands indicating sign "D" with index fingers pointing up and move slowly and meet at the middle of the chest.
	JURUTERA (ENGINEER) Gesture movement for 'Jurutera' (Engineer) sign : Both hand indicating sign "Y" where both thumb connecting each other at the chest level and moving the fingers of right hand up and down several times while the both thumbs stay balance.
	KHABAR BAIK (FINE) Gesture movement for 'Khabar Baik' (Fine) sign : Both hands indicating sign "5" with right hand put at the lip then move down and finally land on the left palm.

1-4244-1355-9/07/$25.00 @2007 IEEE

II. *System Designed and Implementation*

Most of the sign languages are derived from palm sign (in alphabet signing) at different locations of the upper body and the movement of palm sign such as twisting the palm sign, or sliding the palm sign in different locations. For instance, the sign "hari ini" is the sliding of both hands with the sign "D" with the index fingers pointing down simultaneously at the abdomen level.

Thus, in this project, a pair of data gloves is implemented to capture the movement or gesture of a user's fingers . Then, accelerometers are put at that location to capture the gestures.

The structure of the system is shown in Figure 2. It digitized the sign language gesture using flexure sensors and accelerometers using PIC's ADC module. Then, the digitized gesture is transferred to a computer via the Bluetooth module and recognized using the Hidden Markov Model.

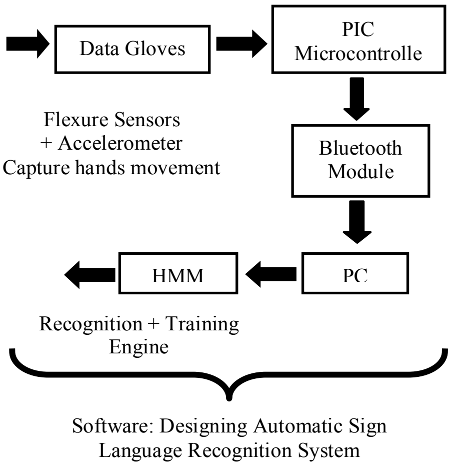


Figure 2: Block Diagram of Wireless Bluetooth Data Gloves Sign Language Recognition System

2.1 Sensors configuration and gesture digitized module

Hand movement is the most complicated movement and always involves twisting the hand and normally in 3-axis. Meanwhile, the elbow and shoulder movement normally involve the 2-axis in vertical and horizontal sliding. Thus, as the user performs sign languages with the data gloves, all the sensors will provide measurement of hand gesture in analogue signals. The analogue signals are then fed to the PIC microcontroller for analogue to digital signal conversion. At the same time, the PIC also collects all gestures data and send it wirelessly to the PC via Bluetooth.

3.1.1 Sensors module

Sensors modules are built to capture information of hand's movement in terms of electric signal (analogue voltage). The outputs of the sensors are connected to the microcontroller in the control module. Two types of sensors have been applied in the project.

3.1.2 Flexure sensors

To capture the degree bend of the fingers. The sensors are attached in a pair of gloves sewn with fabrics.

3.1.3 Accelerometer

To capture the movement of wrists, elbows and shoulders. 3-axis accelerometers are attached at the wrist while 2-axis accelerometers at the elbow and shoulder.

There are altogether 5 flexure sensors and 3 accelerometers on each hand as shown in Figure 4. The flexure sensors function as variable resistance sensor that change resistant according to the sensor's bending.

Each sensor has the resistance ranged from 10kΩ to 30kΩ. 10kΩ means fully stretched and 30kΩ means fully bent. By using the voltage divider concept, the output voltage range is from 1.35 volt to 2.5 volt.

Meanwhile, accelerometer is a small (18mm x 18mm), low power device to measure the static acceleration of gravity in the tilt sensing application.

The 3-axis accelerometer is used at wrist for full capture of the hand movement changes while 2 axis accelerometers are attached at the elbow and shoulder. The output voltage range is from 1 volt to 2.2 volt.

Figure 3: Flexure Sensors setting in dataglove

Figure 4: Overall Hardware Setting for Capturing Hand Gesture.

2.2 Microcontroller and Bluetooth

All output signals generated from flexure sensors and accelerometers are in analogue form. These signals need to be digitized before it can be transmitted to the computer. Therefore, microcontroller PIC18F4550 is used as the main controller to the hardware and at the same time digitizes all analogue signals from the sensors. There are 12 sensor outputs on each hand connected to PIC18F4550. PIC18F4550 will only start to convert the input signal once it gets starting signal from the PC. When the conversion is done, it will send the data to the computer via Bluetooth module. The digitized signal will then be stored in a text file (.txt) (as shown in Figure 5 below).

Firstly, the input data of the text file will be plotted on graph for screening purpose in the filtering stage. Then the start and end point of the data will be detected to locate the important information of the 24 channels input signal.

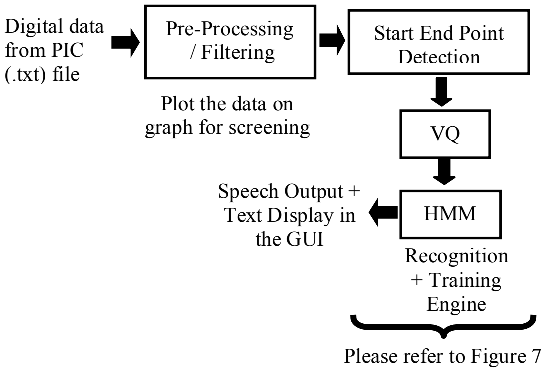


Figure 5: Block diagram of software part

III. *Gesture Recognition and Training Engine*

3.1 Vector Quantization (VQ)

As there are 24 input signals from the sensor, the system will require high computation time to process all the signals simultaneously. Thus, a data compression technique, vector quantization (VQ) is used to reduce transmission and storage cost while the quality is preserved. Vector Quantization (VQ) has been shown to be very useful in compressing data for many applications [11]. Data compression using VQ has received considerable attention due to its compression ratio and rather simple structure [12].

In the VQ process, a codebook is usually trained to minimize the quantization error for data from an individual speaker [13]. The input vectors are then individually quantized to the closest codewords in the codebook built. Compression is gained by using the indices of codewords for transmission and storage [14].

In this project, a codebook with the size of 128 is built to represent all the training signals. For every signal, the indices of the codewords will be used as the input to the HMMs.

IV. *Hidden Markov Models (HMMs)*

The topology HMM model used in this project is a five state HMM as shown in figure 6 [15-16]. This topology of five states was considered sufficient for the most complex signs in BIM. In our approach, each sign is modeled with one HMM.

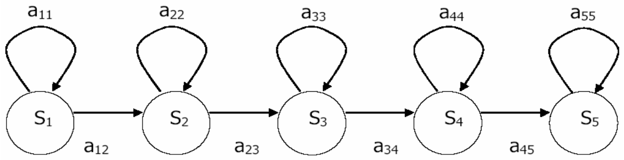


Figure 6: The 5-state HMM used for recognition

A. Training / Learning Phase

The main purpose of learning phase is to create database for the recognition engine. 20 samples of each BIM sign are collected. For this project, only one signer is used for creating signer dependent system. Signer dependent system usually offers higher accuracy compare to signer independent system. This is because 25 BIM signs have been used in this project to make 25 HMM models by using 500 files (25 signs x 20 samples for each sign).

By using the training software developed by the Centre of Biomedical Engineering (CBE), the input data can be trained to produce the HMM models. In the training state, 128 codewords are built. The input data is clustered to 128 classes by using VQ [17]. VQ is used to produce the closest codeword from the codebook for each gesture observation. VQ is required to map each continuous observation vector into a discrete codebook index.

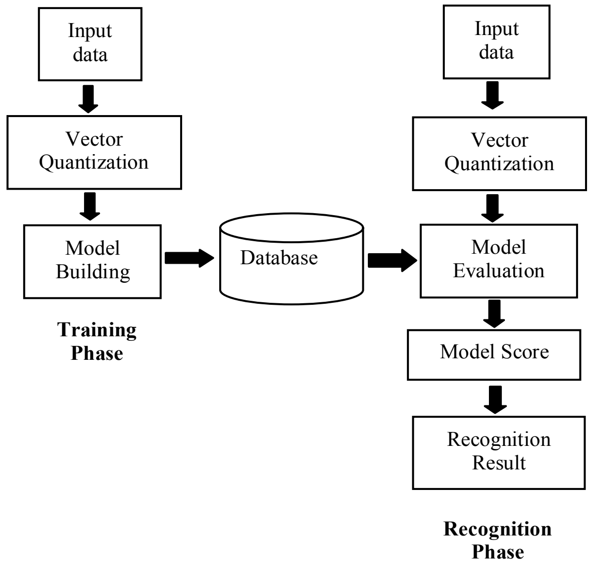


Figure 7: Block Diagram of HMM Recognition and Training Engine

B. Real-time Recognition Phase

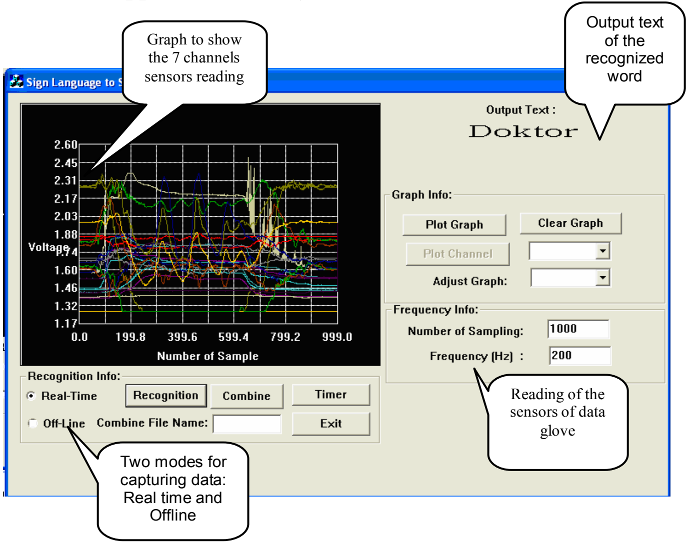
The real time recognition takes place when the user wearing data glove performs a certain sign and it is able the get the output speech immediately.

The system is quite similar with the learning phase. When the user performs a sign language, it will only go through the same process of training phase without creating the database.

After the model evaluation, the system will get the model score as the output of evaluation. Once the score is matched, then the result of that particular BIM sign will appear as output in text form in the GUI with speech as shown in Figure 8. The speech of each BIM sign is pre-recorded and will only plat out if the score of that sign is matched.

V. *result and Discussion*

This project is the first version of BIM Sign Language to Speech Machine that has the capability in verifying approximately 25 words using 2 data gloves plus accelerometers. Users can utilize it as a learning and training tool to learn BIM Sign Language. The sign recognizer engine is still in computer-based as a real sign language to speech machine needs a standalone system which is applicable in daily life.



Output Text : Doktor

● Real-Time

○ Off-Line

Combine File Name:

Number of Sampling: 1000

Frequency (Hz) : 200

Figure 8: The GUI of Malay Sign Language to Speech

Acknowledgment

This research project is supported by CBE (Center of Biomedical Engineering) at Universiti Teknologi Malaysia and funded by Minister of Higher Education (MOHE), Malaysia under grant "Support Vector Machines (SVM) As a Large Margin Classifier For Machine Learning Algorithm" Vot 78029.

References

[1] Mehdi, S.A.; Khan, Y.N. "Sign language recognition using sensor gloves" Proceedings of the 9th International Conference on Neural Information Processing, 2002. ICONIP '02. Volume 5, 18-22 Nov. 2002 Page(s):2204 - 2206 vol.5

[2] W.Kadous. Recognition of Australian Sign language using instrumented glove. Master Thesis. University of New South Wales. October 1995.

[3] H. Liang and M. Ouhyoung, A sign language recognition system using Hidden Markov Model and Context Sensitive Search, *ACM VRST,* 1996.

[4] Malaysian Sign Language http://www.epekak.net.my

[5] David Joel Sturman. Whole-hand Input. . PhD. Thesis. University of New South Wales. October 1992.

[6] Persekutuan Orang Pekak Malaysia, "Bahasa Isyarat Malaysia", Penerbit Persekutuan Orang Pekak Malaysia, 2000.

[7] David J. Sturman, David Zeltzer. A survey of glove-based input, Computer Graphics and Applications. IEEE, Volume 14, Jan. 1994 Page(s):30 – 39

[8] T Kuroda, Y Tabata, A Goto, H Ikuta and M Murakami.
Consumer price data-glove for sign language recognition.
Kyoto University Hospital/Kyoto College of Medical
Technology/AMITEQ Corp., Tokyo/Teiken Limited, Osaka,
JAPAN, 2004. [3]

[9] Kramer & Larry J. Leifer. The Talking Glove – a speaking aid for non vocal deaf and deaf blind individual. Technical report CDR TR 1990 0312 center for Design Research. Stamford University 1990.

[10] J. Jang, W. B. Park. Signal processing of the accelerometer for gesture awareness on handheld. Robot and Human Interactive Communication, 2003. Proceedings. ROMAN 2003. The 12th IEEE International Workshop on. 31 Oct.-2 Nov. 2003 Page(s):139 – 144

[11] Corral, J. A., Guerrero, M. and Zufiria, P. Image Compression via Optimal Vector Quantization: A Comparison between SOM, LBG abd K-Means Algorithms. *Proceedings of IEEE World Congress on Computational Intelligence*. 27 Jun- 2 Jul 1994. USA: IEEE, 1994. 4113 -4118 vol.6.

[12] Huang, C. M. and Harris, R. W. A Comparison of Several Vector Quantization Codebook Generation Approaches. *IEEE Transactions on Image Processing*. Jan 1993. 2(1): 108 –112.

[13] Zhang, Y. *et al.* Speaker-Independent Isolated Word Recognition using Multiple hidden Markov Models. IEE Proceedings- Vision, Image and Signal Processing. Jun 1994. 141(3): 197 –202.

[14] He, J., Liu, L. and Palm, G. A New Codebook Training Algorithm for VQbased Speaker Recognition. *Proceedings of 1997 IEEE International Conference on Acoustics, Speech, and Signal Processing*. 21-24 Apr 1997. Germany: IEEE, 1997. 1091 -1094 vol.2.

[15] Rabiner L. R. *A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition*. Proceedings of the IEEE. Volume : 77 Issue:2, Feb 1989, 257 – 286.

[16] *Rabiner, L. and Juang, B.H. An Introduction to Hidden Markov Models, IEEE ASSP Magazine*. 4-16.

[17] Huang, C. M. and Harris, R. W. A Comparison of Several Vector Quantization Codebook Generation Approaches. *IEEE Transactions on Image Processing*. Jan 1993. 2(1): 108 –112.