1st International Conference on Sensing Technology
November 21-23, 2005 Palmerston North, New Zealand

Recognition of Vietnamese sign language using MEMS accelerometers

The Duy Bui and Long Thang Nguyen

College of Technology,
Vietnam National University, Hanoi
{duybt,longnt}@vnu.edu.vn

Abstract

In this paper, we present work on understanding Vietnamese sign language through the use of Micro Electronic Mechanical System (MEMS) accelerometers. The system consists of six ADXL202 accelerometers for sensing the hand posture, a BASIC Stamp microcontroller and a PC for data acquisition and recognition of sign language. The data is then transformed to relative angles between fingers and the palm. Each character is recognized by a fuzzy rule based classification system, which allows the concept of vagueness in recognition. In addition, a set of Vietnamese spelling rules has been applied to improve the classification results.

Keywords: MEMS sensors, Vietnamese Sign Language, sign language recognition, human computer interaction

1 Introduction

At present, most of the interaction between humans and computers is done via text, mouse and keyboard. Together with the rapid development of advanced sensor technology, researchers are increasingly paying attention to making this interaction more adaptive, flexible, human-oriented and especially, more affordable. For example, MEMS sensors such as accelerometers, geophones and gyros, thanks to their small size and weight, modest power consumption and cost, and high reliability, allow the development of human computer interaction systems with portability and affordability.

One of the active areas in the application of advanced sensor technology is the investigation of hand shape and gesture recognition. Besides the application in human computer interaction, the recognition of hand shape and gesture can also be useful in many other interesting application such as robotics, virtual reality, tele-manipulation, tele-presence, and sign language translation.

In this paper, we discuss the application of MEMS accelerometers for Vietnamese Sign Language (VSL) recognition. We focus on the recognition of postures that represent the 23 Vietnamese based letters. The data obtained from the sensing device is transformed to relative angles between fingers and the palm. Each character is recognized by a fuzzy rule based classification system, which allows the concept of vagueness in recognition. In addition, a set of Vietnamese spelling rules has been applied to improve the classification results. The recognition rate is high even when the postures are not performed perfectly, e.g. the finger is not bended completely or the palm is not straight.

In principle, VSL is based on the well-established American Sign Language (ASL). According to ASL dictionary [1], four components are used to describe a sign: hand shape, location in relation to the body, movement of the hands, and orientation of the palms. A popular concept in sign language, "posture", is formed by the hand shape (position of the fingers with respect to the palm), the static component of the sign and the orientation of the palm. The alphabet in ASL, which consists of 26 unique distinguishable postures, is used to spell names or uncommon words that are not well defined in the dictionary. This well-established ASL alphabet contains some signs with high amount of finger occlusion. Therefore, it is very hard to use vision-based systems in the recognition task.

In literature, there are numbers of work on gesture recognitions based on sensing gloves. For example, in order to enter ASCII characters to a computer, Grimes [2] developed the Data Entry Glove using switches and other sensors sewn to the glove. Kramer [3] used a look-up table with his patented CyberGlove to recognize the 26 letters of the alphabet. Alternatively, Erenshteyn [4] used a method involving coded output, such as Hamming, Golay, and other hybrid codes together with the CyberGlove. Zimmerman invented the VPL Data Glove [5] in order to recognize postures in different sign languages. For example, a set of 51 basic postures of Taiwanese Sign Language was solved by Liang [6] with probability models; and 36

ASL postures were able to be recognized with this glove by the work of Waldron [7] with a two-stage neural network. Those mentioned gloves, however, are very expensive. A more-affordable option was proposed by Kadous [8]. This is a system for Australian Sign Language based on Mattel's Power Glove. However, because of a lack of sensors on the pinky finger, the glove could not be used to recognize the alphabet hand shapes. With accelerometers at fingertips, Pister [9] developed a text editor where each hand gesture refers to a letter of the alphabet. Thorough reviews of the use of gloves for gesture recognition can be found in papers by Sturman [10] and Watson [11].

Recently, Hernandez et. al. [12] proposed a system called The Accele Glove, a whole-hand input device using MEMS accelerometers to manipulate three different virtual objects: a virtual hand, icons on a virtual desktop and a virtual keyboard using the 26 postures of the American Sign Language (ASL) alphabet. When using this device as finger spelling translator, a Multi-Class Pattern recognition algorithm is applied [13]. Firstly, the data are collected and analyzed 'off line' on a PC. The obtained data are transformed to vectors in the posture space then divided into subclasses. By, this way, it is possible to apply simple linear discrimination of the postures in 2D space, and Bayes' Rule in those cases where classes had features with overlapped distributions. This algorithm can be implemented as a sequence of 'if-then-else' statements in the micro controller, allowing a real time processing. Since the application of this device has potential, in this work, we aim to develop a similar device for recognition of Vietnamese Sign Language. In addition to the five sensors as in the Accele Glove, we placed one more sensor on the back of the hand to improve the recognition process. In addition, we use a completely different method for the classification process leading to very promising results.

2 Vietnamese alphabet system

During the 17th century, Roman Catholic missionaries introduced a Latin-based orthography for Vietnamese. To day it becomes Vietnamese national language. Vietnamese alphabet is listed below:

A Ă Â B C D Đ E Ê G H I K L M N O Ô Ơ P Q R S T U Ư V X Y

The letters J, W and Z are also used, but only in foreign loan words. In addition, Vietnamese is a tonal language with 6 tones. These tones are marked as follows: level, high rising , low (falling), dipping rising, high rising glottalized, low glottalized.

Since Vietnamese alphabet system is more complicated than English alphabet system, more signs are required for VSL in comparison with ASL. However, it is possible to implement finger spelling of Vietnamese words similar to the ASL system. The VSL consists of 23 based letter and some addition signs for the accents and the tones. The 23 based letter are:

A B C D Đ E G H I K L M N O P Q R S T U V X Y

In this paper, we concentrate only on the recognition of these based letters.

3 The sensing device

One of the most successful MEMS sensors in the market is ADXL202 accelerometers from Analog Devices [14]. The ADXL202 are low cost, low power, complete 2-axis accelerometers on a single IC chip with a measurement range of ±2 g. The ADXL202 can measure both dynamic acceleration (e.g., vibration) and static acceleration (e.g., gravity). The accelerometer is fabricated by surface micromaching technology. It is composed of a small mass suspended by springs. Capacitive sensors distributed along two orthogonal axes (X and Y) provide a measurement proportional to the displacement of the mass with respect to its rest position. Because the mass is displaced from the center, either due to acceleration or due to an inclination with respect to the gravitational vector g, the sensor can be used to measure absolute angular position. The outputs are digital signals whose duty cycles (ratio of pulse width to period) are proportional to the acceleration in each of the 2 sensitive axes. The output period is adjustable from 0.5 ms to 10 ms via a single resistor (R_{SET}). If a voltage output is desired, a voltage output proportional to acceleration is available from the X_{FILT} and Y_{FILT} pins, or may be reconstructed by filtering the duty cycle outputs. The bandwidth of the ADXL202 may be set from 0.01 Hz to 5 kHz via capacitors C_X and C_Y. The typical noise floor is 500 $\mu g/\sqrt{Hz}$ allowing signals below 5 mg to be resolved for bandwidths below 60 Hz.

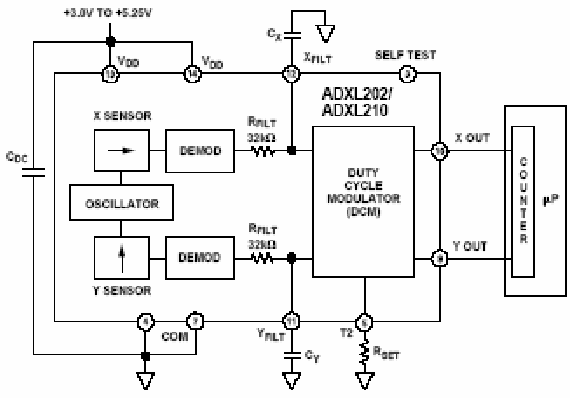

Figure 1: Function block diagram of ADXL202 [14]

Our sensing device consists of six ADXL202 accelerometers attached on a glove, five on the fingers and one on the back of the palm. The Y-axis points toward the fingertip, providing a measure of joint flexion. The X-axis can be used to extract information of hand roll or yaw, or individual finger abduction. Data is collected by measuring the duty cycle of a train of pulses of 1kHz. When a sensor is in its horizontal position, the duty cycle is 50%. When it is tilted from +90° to –90° the duty cycle varies from 37.5% (.375 msec) to 62.5% (.625 msec), respectively. In our device, the duty cycle is measured using a BASIC Stamp microcontroller. The Parallax BASIC Stamp module is a small, low cost general-purpose I/O computer that is programmed in a simple form of BASIC [15]. The pulse width modulated output of the ADXL202 can be read directly of the BASIC Stamp module, so no ADC is necessary. Twelve pulse widths are read sequentially by the micro controller, beginning with the X-axis followed by the Y-axis, thumb first. The data then is sent through the serial port to a PC for further analyses.

Figure 2: Sensing Glove with six accelerometers and Basic Stamp microcontroller.

4 Data processing

Our sensing glove produces the raw data represented as a vector of 12 measurements, two axes per finger and the last two axes for the palm:

$$D=[x_t\ y_t\ x_i\ y_i\ x_m\ y_m\ x_r\ y_r\ x_p\ y_p\ x_{pa}\ y_{pa}]$$

t= thumb, i=index, m=medium, r=ring, p=pinky, pa=palm.

We do not follow the method proposed in [12] to recognize the postures by extracting the features directly from the raw data. There are two reasons for this. The first reason is that the raw data is the pulse widths which relate to the rolling or flexing angles through cosine functions. Because the cosine function itself is not linear, the sum of pulse widths measured on fingers does not represent the hand shape accurately. The second reason is that the sum is not a good function to extract the feature of hand shape. A lot of different hand shapes can result in the same feature extracted by the sum function.

Differently, we first convert our data to the angles. After that, we subtract the x and y values of the fingers to the x and y values of the palm, respectively. Note that our sensing device has one more sensor than the Accele Glove, which measure the rolling and flexing angle of the palm. By processing the data this way, we convert the raw data into the relative angles between the fingers and the palm. We will do the classification based on the x and y value of the palm and the relative angles between the fingers and the palm.

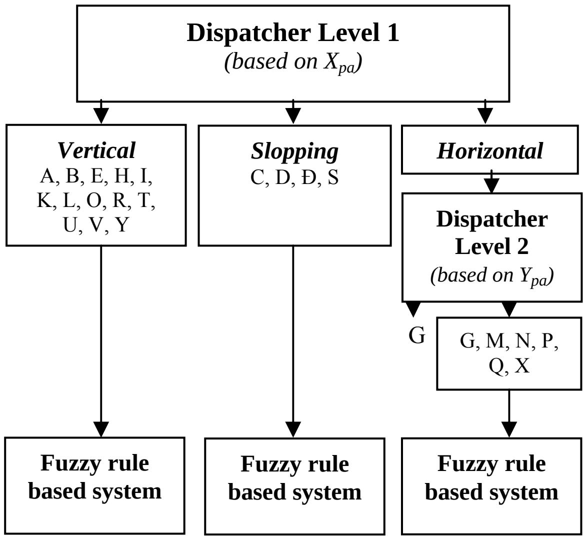

Figure 3: An overview of the classification system.

5 Classification

The classification system can be seen in Figure 3. First of all, we use the x value of the palm to divide the postures into three categories: "Vertical" which consists of the postures of letter 'A', 'B', 'E', 'H', 'I', 'K', 'L', 'O', 'R', 'T', 'U', 'V' and 'Y' ; "Sloping" which consists of the postures of letter 'C', 'D', 'Đ' and 'S' ; and "Horizontal" which consists of the postures of letter 'G', 'M', 'N', 'P', 'Q', 'X'. After the postures are divided into three categories, we use three fuzzy rule based system to perform further classification.

Human beings often need to deal with input that is not in precise or numerical form. Inspired by that observation, Zadeh [16] developed a fuzzy set theory that allows concepts that do not have well-defined sharp boundaries. In contrast to the classical set theory in which an object can only be a full member or a full non-member of a set, an object in fuzzy set theory can possess a partial membership of a fuzzy set. A fuzzy proposition of the form "if x is A" is partially satisfied if the object x (usually crisp value x) is partial membership of the fuzzy set A. Based on that, fuzzy logic was developed to deal with fuzzy "if-then" rules where the "if" condition of the rules is a boolean combination of fuzzy propositions. When the "if" condition is partially satisfied, the conclusion of a fuzzy rule is drawn based on the degree to which the condition is satisfied.

We have found that the concept of fuzzy set is well suited for the problem of posture classification because the posture is normally defined in a vague way, e.g. "the index finger bends a little bit". Moreover, with a fuzzy rule based system, the classification can be solved by a set of rules in natural language which look like:

if all fingers bend maximally **then**
it is the posture of letter 'A'
if all fingers does not bend **then**
it is the posture of letter 'B'
...

We model the level of bending or flexing of the fingers by five fuzzy sets (Figure 4): **Very-Low**, **Low**, **Medium**, **High**, **Very-High.**

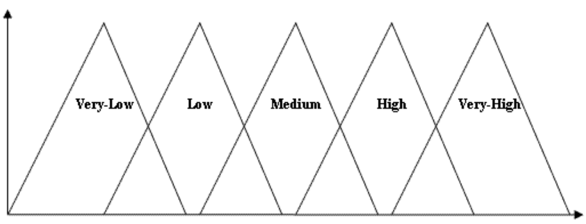

Figure 4: The five fuzzy sets representing the level of bending or flexing of the fingers.

The fuzzy classification rules look like:

if thumb finger's bending is **Low**
and index finger's bending is **Very-Low**
and medium finger's bending is **Very-Low**
and ring finger's bending is **Very-Low**
and pinky finger's bending is **Very-Low then**
the posture is recognized as letter 'B'

We have created 22 fuzzy rules to classify Vietnamese sign language postures. The posture of letter 'G' is recognized directly with the use of the y value of the palm. With these fuzzy rules, the classification process is done as follows. Every time we receive the data from the sensing device, we first verify if the hand is at static position by comparing with previous data. We wait until the hand stops moving to start the recognition process. The pre-processed data is used to calculate the "membership values" – the degree to which the data belongs to the fuzzy sets. We then calculate the degree to which the current data set matches each of the 22 fuzzy rules. The matching degree is calculated as the product of the membership values to which the data belongs to the fuzzy sets in the rules. Finally, the data set is recognized by the rule with the highest matching degree.

The recognition process is enhanced by the use of Vietnamese spelling rules. Vietnamese has a very special characteristic, that is all the words are monosyllabic. Moreover, Vietnamese spelling rule are very strict, e.g., except "NG", "NGH", "NH", "PH", "TH", "CH", "TR" and "GH", a consonant can not be followed by another consonant; or besides a set of combination of vowels, a vowel cannot be followed by a vowel. Taking advantage of these rules, when we recognized word formed by postures of letters, we can eliminate many misclassifications.

6 Result

The system was implemented in C++, and was tested using a total of 200 samples to measure the recognition rate. Twenty out of the 23 letters reached a 100% recognition rate after the application of Vietnamese spelling rules. The three letters 'R', 'U', and 'V' produce ambiguity as the data represented these letters is similar. Before applying Vietnamese spelling rules, the recognition rate is 90%, 79% and 93% respectively. However, after applying Vietnamese spelling rules, the recognition rate has increased significantly, which is 94%, 90% and 96% respectively.

The novelty of our system is that the postures can also be recognized even when they are not performed perfectly, e.g. the finger is not bended completely or the palm is not straight. This is because we carry out the classification on the relative angles between the fingers and the palm instead of the classification on the raw data as in [12]. This is also the result of the fuzzy rule based system, which allows the concept of vagueness in recognition.

7 Conclusion

In this paper, we presented work on understanding Vietnamese sign language through the use of Micro Electronic Mechanical System (MEMS) accelerometers. The system consists of six ADXL202 accelerometers for sensing the hand posture, a BASIC Stamp microcontroller and a PC for data acquisition and recognition of sign language. The classification process is done by a fuzzy rule based system on the pre-processed data. In addition, we have applied a set of Vietnamese spelling rules in order to improve the classification results. We have achieved very high recognition rates. Moreover, the postures can also be recognized even when they are not performed perfectly.

Some potential applications for our Sensing Glove are: a wireless wearable mouse pointing device, a wireless wearable keyboard, hand motion and gesture recognition tools, virtual musical instruments, computer sporting games, and work training in a simulated environment.

In the future, we want to place more sensor of different types into our sensing device in order to recognize more complex form of the Vietnamese Sign Language, as well as to recognize gestures for other human computer interaction applications.

References

[1] E. Costello, "Random House Webster's Concise American Sign Language Dictionary". Random House N.Y., 1999.

[2] G. Grimes. US Patent No. 4414537, November 1983.

[3] J. Kramer and L. Leifer, "The Talking Glove: An Expressive and Receptive 'Verbal' Communication Aid for the Deaf, Deaf-Blind, and Nonvocal". *SIGCAPH 39*, pp.12-15 (spring 1988).

[4] R. Erenshteyn et. al., "Distributed Output Encoding for Multi-Class Pattern Recognition". *Proceedings of the Intl. Conf. On Image Analysis and Processing,* 1999. 229-234.

[5] T. Zimmerman "Optical Flex Sensor". US Patent No. 4,542,291. 1987

[6] R. Liang, M. Ouhyoung. "A Real-time Continuous Gesture Recognition System for Sign Language". *Proceeding of the Third IEEE Int. Conf. On Automatic Face and Gesture Recognition,* 1998. pp 558-567

[7] M.B. Waldron et al. "Isolated ASL Sign Recognition System for Deaf Persons". *IEEE Trans. On Rehabilitation Engineering*, vol 3 No. 3 September 1995.

[8] M.W. Kadous "Grasp: Recognition of Australian sign language using instrumented gloves". Master's thesis, University of New South Wales, Oct. 1995

[9] J. Perng, B. Fisher, S. Hollar, K.S.J. Pister, "Acceleration Sensing Glove (ASG)," *ISWC International Symposium on Wearable Computers*, San Francisco, October 18-19, 1999

[10] D.J. Sturman and D. Zeltser "A Survey of Glove-Based Input". *IEEE Computer Graphics & Applications.* 1994, pp 30-39.

[11] R. Watson. "A Survey of Gesture Recognition Techniques". Technical Report TCD-CS- 93-11. Department of Computer Science, Trinity College, Dublin. July 1993.

[12] J. Hernandez, N. Kyriakopoulos, R. Lindeman. "The AcceleGlove a Whole Hand Input for Virtual Reality". *SIGGRAPH 2002*. San Antonio Texas July 21-26, 2002.

[13] J. Hernandez, R. Lindeman, N. Kyriakopoulos, "A Multi-Class Pattern Recognition System for Practical Finger Spelling Translation," *Fourth IEEE Int'l Conference on Multimodal Interfaces (ICMI'02)*, Oct. 14-16, 2002, Pittsburgh, PA, USA, pp. 185-190.

[14] Analog Devices Inc, www.analog.com.

[15] Parallax Inc, www.parallax.com

[16] Zadeh, L. A. (1965). Fuzzy sets. *Information and Control,* 8:358-353.