A Real-time Portable Sign Language Translation System

Lih-Jen kau, Wan-Lin Su, Pei-Ju Yu, Sin-Jhan Wei

Department of Electronic Engineering, National Taipei University of Technology
No.1, Sec. 3, Chung-Hsiao E. Rd., Taipei 10608, Taiwan, R.O.C.
Email: ljkau@ntut.edu.tw, pauline282937@gmail.com, est8660734@hotmail.com, weisinjhan@gmail.com

In this paper, a wireless hand gesture recognition glove is proposed for real-time translation of Taiwanese sign language. To discriminate between different hand gestures, we have flex and inertial sensors embedded into the glove so that the three most important parameters, i.e., the posture of fingers, orientation of the palm, and motion of the hand, defined in Taiwanese Sign Language can be recognized without ambiguity. The finger flexion postures acquired by flex sensors, the palm orientation acquired by G-sensor, and the motion trajectory acquired by gyroscope are used as the input signals of the proposed system. The input signals will be acquired and examined periodically to see if it is a legal sign language gesture or not. Once the sampled signal can last longer than a predefined clock cycles, it is regarded as a valid gesture and will be sent to cell phone via Bluetooth for gesture discrimination and speech translation. With the proposed architecture and algorithm, the accuracy for gesture recognition is quite satisfactory. As we can see in experiments that an accuracy rate up to 94% on sensitivity for gesture recognition can be achieved which justifies the superiority of the proposed architecture.

Keywords—*Flex sensor; Gesture recognition; Inertial sensor; Portable glove; Real-time translator; Sign language; Smart phone.*

I. INTRODUCTION

Sign language has been the most natural and expressive way for deaf-mute to communicate with each other for a long time. To help normal people get familiar with deaf-mute and know what they want to express, translating sign language into voice or the way normal people might understand has been an important issue. In order to discriminate between different gestures in sign language, some researches have been initiated and devoted to the design of a hand gesture recognition system [1]-[3]. Among which, the results can be approximately divided into two categories, i.e., inertia sensor-based and video-based systems. In general, inertia sensor-based systems are also wearable or portable [1]-[2]. The system in [1] is composed of a gesture recognizing glove terminal and a mobile device center, both of which communicate via the Bluetooth wireless signal. The gesture recognizing glove is made of a self-designed circuit module with flex sensors that can detect which finger joints are bent [1]. Upon acquiring the data that is encoded, it will be transmitted wirelessly to a mobile device. However, a large amount of sign language can't be recognized since the glove in [1] can just detect whether the finger is straight or curve. Also the accuracy of hand gesture recognition is not really satisfactory since certain sign language can also have a large amount of continuous hand gesture. In other words, a sign language recognizing system need some different inertial sensors to recognize much more gestures, just like the system proposed in [2]. Nevertheless, a much better algorithm is still required for the system in [2] so that the starting posture and ending posture of each sign language can be recognized.

The second kind of architecture is video-based systems. In a video-based system, usually a hand tracking as well as 3D motion trajectory recording mechanism are required for the recognition and matching of a valid sign language gesture [3]. For example, the video-based system in [3] is composed of a hand tracking camera Kinect and a 3D motion trajectory database, both of which will be matched after comparing the detected 3D motion trajectory with the sign language in database, and the results will be demonstrated on the screen of laptop or desktop computer. With the use of the tracking camera Kinect, some useful information like the depth and the color of images can be easily acquired. Besides, the user doesn't have to wear any device if the video-based system is used [3]. However the realization of hand gesture recognition is really difficult since the complex background and illumination conditions may affect the hand tracking accuracy. In addition, the system is impractical and inconvenient for deaf-mute to use in their daily life when they go out since the demonstration of the sign language is on the screen of laptop or desktop computer [3].

Aimed to design a real-time and portable sign language translating system, we propose in this paper a wireless hand gesture recognition glove system that uses a smart phone as the platform for text to sign language matching database and the text to speech translating device. The inertial (G-sensor and gyroscope) and flex sensors will be used as the sensing components of the proposed system to generate original input signals. The periodically acquired sensor signals will be stored in a queue first, and then the stored signals will be processed by the proposed algorithm to calculate the repeated times of an identical hand gesture in the sampled sequence. If the sampled posture can last longer than a predefined clock cycles, the signal will be regarded as a valid gesture; otherwise the state machine will be reset to the initial state since a different posture has been detected. Moreover, the signal will be encoded and sent to cellphone via Bluetooth signal once a valid hand gesture is detected.

In the cellphone, the received signal, an encoded digital sequence, will go through a text matching process. To do this, the encoded sequence will be compared with the sign language database via table lookup to find out the meaning of this gesture. After the text matching process, the text will be translated to speech by Google translator. The speaker of the cellphone will send out a voice so that normal people can hear and understand what deaf-mute want to express. With the proposed algorithm and architecture, the accuracy of gesture recognition can be quite promoted since we check the times of repeated posture and reset to the initial state once the times of repeated posture can't last for a period longer than the predefined clock cycles. Moreover, the proposed system is friendly to use since we are using inertial sensors and cell phone for gesture detection and discrimination, the user just need to wear the glove and has their smart phone as a speaker, which is easy and convenient for them to talk with normal people by using their most natural and familiar sign language.

II. Taiwanese Sign Language

In this section, an overview of the Taiwanese Sign Language [4] as well as two basic definitions used in this paper will be introduced.

Posture: A posture is a specific meaning of finger flexural state observed at some time instance.

Gesture: A gesture is a sequence of postures connected by motions over a short time span.

Usually a word in Taiwanese Sign Language can be defined as a hand gesture consisting of one or more postures sequentially occurring in a continuous time. In [4], an analysis as well as an image-based translation system for Taiwanese Sign Language is proposed. Four parameters, i.e., the posture, position, orientation, and motion, are also defined in [4] for sign language recognition. In addition, fifty-one fundamental postures (as shown in Fig. 1) and twenty-two typical positions (as shown in Fig. 2) are defined for Taiwanese Sign Language in [4].

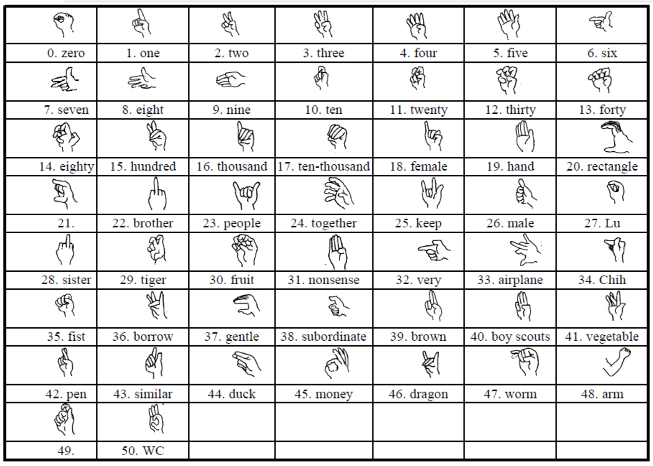


Fig. 1. Fifty-one fundamental postures in Taiwanese Sign Language defined in [4].

1. In front of the body	7. Nose	13. Neck	19. Under the arm
2. Above the head	8. Mouth	14. Shoulder	20. Arm
3. In front of the face	9. Jaw	15. Heart	21. The back of the hand
4. Top of head	10 Temple	16. Breast	22. Wrist
5. Brow	11. Ear	17. Waist	
6. Eye	12. Cheek	18. Leg	

Fig. 2. Twenty-two typical positions in Taiwanese Sign Language defined in [4].

It is noted that there are six different orientations of palm in Taiwanese Sign Language: up, down, right, left, ahead, and back. The six different orientations of palm can also diversify and promote the amount of Sign Language, since the same posture (Fig. 3.(a)) with different orientation of palm can have quite different meaning of sign language (Fig. 3.(b), (c), (d), (e), (f)).

In addition to the finger postures and the orientation of the palm, the motion trajectory of a hand gesture also plays an important role during the classification process [4], since identical posture (Fig. 3.(a)) and orientation of palm (Fig. 3. (d), (e)) can also have different meaning in sign language.

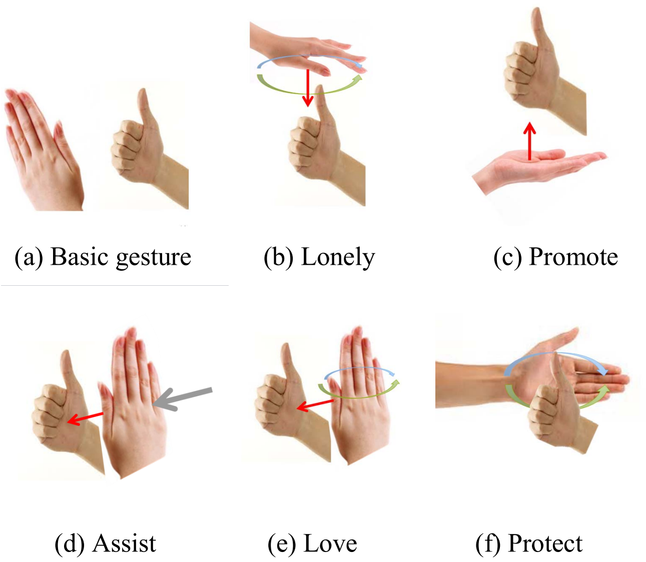


Fig. 3. (a) Take a basic gesture with two postures in Taiwanese Sign Language as an example. (b) Lonely gesture with down orientation of palm. (c) Promote gesture with up orientation of palm. (d) Assist gesture with ahead orientation of palm. (e) Love gesture with ahead orientation of palm. (f) Protect gesture with back orientation of palm.

To integrate the above four parameters (posture, position, orientation, and motion) into recognition process, we must propose a hardware system to acquire important signal.

III. Proposed Hardware And Signal Acquisition System

In this section, the proposed hardware as well as signal acquisition process will be introduced. In order to recognize the posture of fingers, orientation of palm, motion and position of the hand, we use the flex and the inertial sensors (G-sensor and gyroscope) as the major components for input signals generation in the proposed hardware system.

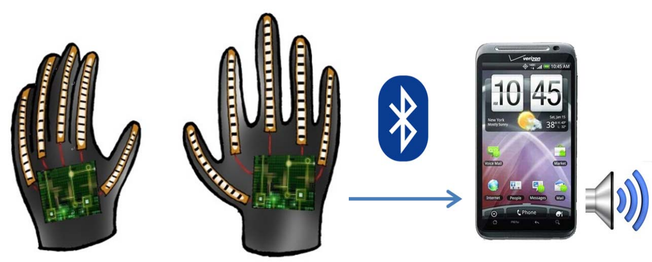
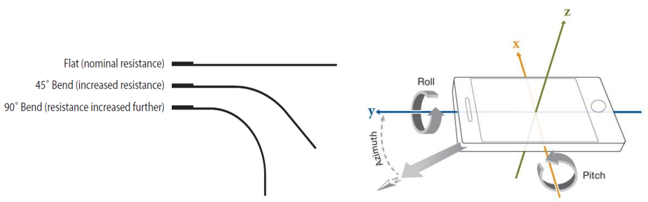


Fig. 4. System architecture of the proposed real-time portable sign language translation system.



(a) Finger flexion detection (b) Orientation convention

Fig. 5. (a) Finger flexion detection via the flex sensor. (b) Motion and palm orientation detection via inertial sensors.

A. The Flex sensors

As mentioned, the first step to recognize hand gesture is to detect the posture from 51 fundamental postures shown in Fig. 1 [4]. In order to know the state of finger flexion in Fig. 5.(a), we develop a wireless hand gesture recognition glove with 10 flex sensors attached to each finger [5]. The output of each flex sensor will be sampled continuously and periodically as the input signals of the proposed system. To simplify the complexity of circuit design, each flex sensor is concatenated with a fixed resister, and then we measure the voltage between the flex sensor (R_F) and the fixed resistor (R_B) given by (1).

$$V_o = V_i \frac{R_F}{R_F + R_B} \tag{1}$$

where R_F is the variable flexural resistor values depending upon the finger flexion, R_B is the value of the fixed resistor which is concatenated with the flexural resistor. Since the measured R_F value is 10kΩ when the flex sensor is in its flat state (in Fig. 5.(a)) [5], we set the R_B value to 10kΩ for all the ten fixed resistors. The higher bend state of flex sensor, the larger R_F value of flexural resistance. The largest R_F value 20kΩ occurs when it is in 90^o bend state (as in Fig. 5.(a)) [5]. In most of the cases, a 45^o angle can be measured when the finger curved. This is the most nature and closed state when used to recognize the finger posture of sign language. The measured R_F value of 45^o bend state (in Fig. 5.(a)) is around 15kΩ in our experiment [5]. Since the input voltage V_i offered from general MCU sensor is 3.3V, the output voltage V_o of 45^o bend state is around 1.98V as standard value of the proposed system. If the voltage V_o is over 1.98V, the output signal will be recognized as the bend state. To simplify the analysis complexity, each finger will be encoded with one bit, i.e., a 0 or a 1, and each individual hand will be encoded with five bits respectively. After knowing every state of 10 finger flexion, we can recognize the hand posture of sign language.

B. The inertial sensors (G-sensor and Gyroscopes)

The second step for hand gesture recognition is to detect the orientation from six different orientations of palm: up, down, right, left, ahead, and back. Since the same posture we recognized in first step (Fig. 3.(a)) may have different meaning in sign language (Fig. 3.(b), (c), (d), (e), (f)), we then have a G-sensor embedded to the glove and placed in the position back of the hand so that the orientation of the palm can be detect. The output signals of G-sensor are three-dimensional data sequence along the X-axis, Y-axis, Z-axis (Fig. 5.(b)), we can easily discriminate between different

Table I. The palm orientation of each Sign Language.

Orientations	Down	Up	Back	Ahead	Right	left
Sign Language	Lonely	Promote		Assist Love	Protect	
three-dimensional data	Z > 170 X < 100 Y < 100	Z < -180 X < 100 Y < 100	Z < 150 X < 100 Y > 200	Z < 150 X < 100 Y > 200	Z < 150 X > 230 Y < 100	Z < 150 X< -220 Y < 100

sign language via the sign and magnitude of the three dimensional data sequence as in Table I.

However there are still some words with same posture and orientation that can't be distinguished (just like Assist and Love in Table I). Therefore, the motion trajectory of hand gesture plays an important role, and will be the third step of classification in the proposed system. In order to distinguish Assist and Love in sign language, we then attach gyroscopes to the glove so we can get the angle of azimuth, pitch and roll. It is noted that the azimuth rotates around the Z-axis, the pitch rotates around the X-axis, and the roll rotates around the Y-axis. The pitch, which indicates the angle between the Y-axis and the Ground, will be selected as one of the feature in the third step for sign language discrimination. Since the sign language of Assist has different forward motion from the sign language of Love, we can easily discriminate the two words from each other by using the pitch angle.

IV. SIGNAL SAMPLING AND RECOGNITION

As mentioned, the classification process in the proposed system can be divided into three steps. During the three steps, the signal will be acquired and sampled continuously and periodically as in Fig. 6.

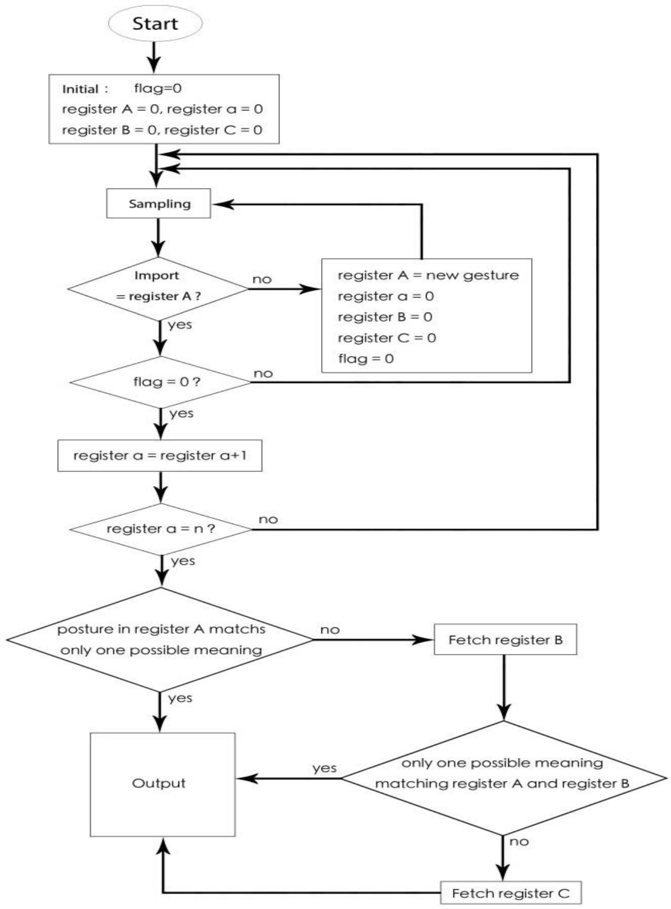


Fig. 6. Flow chart for gesture recognition.

Fig. 6 indicates the flow chart of signal sampling and gesture recognition process. Register(A) records the former posture, Register(a) records how long does a posture persist, Register(B) records the orientation of palm, Register(C) records the motion trajectory, flag indicates whether the gesture had been sent out or not. The initial value of Register(A), Register(a), Register(B), Register(C) and flag are 0. When we start to detect the gesture, we acquire the signal of posture, orientation of palm and motion trajectory. After the sampling and the thresholding process, we can get a digit signal with 20 bits, the first 10 bits will be stored in Register(A), the consecutive 6 bits will be stored in Register(B), and the last 4 bits will be stored in Register(C).

When it comes to next clock cycle, we then compare the first 10 bits to the content in Register(A). If the 10 bits signal is not equal to the content of Register(A), it means that this posture is different from the former one, and we will make Register(a) return to 0, which means recalculating the lasting time for this new posture. Besides, Register(B) and Register(C) should also be reset to 0, since the posture recognition doesn't finish yet. The flag is also reset to 0, which means this gesture has not been transmitted yet.

On the contrary, if the 10 bits signal is equal to that in Register(A), it means that the posture has been remained. After we make sure that this gesture has not been transmitted yet, we then calculate how long the posture is remained. Every clock cycle, as long as the new 10 bits signal equal to Register(A), we make the Register(a)=Register(a)+1, and then compare the Register(a) to the number n. If Register(a) is smaller than n, it means that the posture has not remained for a long time. The posture is very likely to be an interim gesture. The system will go on sampling again. As long as the content of Register(a) is greater than the predefined number n, it means that the posture had remained for a sufficiently long time, and it is very likely that it is the posture that the user wants to express. After we make sure that the posture in Register(A) matches only one possible meaning, we then send out the signal and set the flag to be 1.

On the contrary, if the posture in Register(A) matches many different meaning, we will fetch 6 bits in Register(B), which means the state of the palm orientation as in Table III. We then make sure that the gesture detected in Register(A) and Register(B) matches only one meaning. If the outcome is true, we then send out the signal. If the outcome is false, we will fetch another 4 bits in Register(C). The 4 bits in Register(C) means the motion trajectory of the gesture as in the last column of Table II, which can help us easily discriminate the gesture that have different meaning. Finally, we will send out the signal that the user wants to express.

Table II. The meaning of a 20 bits digital signal

Register A		Register B		Register C	
Posture		Orientation of palm		motion trajectory	
10 bits		6 bits (two hands)		4 bits (two hands)	
Left hand	5 bits	down	001	None	00
		up	110		
		back	010	Clockwise	01
Right hand	5 bits	ahead	101		
		right	100	Counterclockwise	10
		left	011		

V. EXPERIMENT

In this section, the performance of the proposed classification architecture and algorithm are to be analyzed. We find five subjects wearing the proposed hand gesture recognition glove with their smart phone as the speaker. In addition, five different kinds of sign language words including Lonely, Promote, Assist, Love, and Protect will be evaluated by the five subjects each with 50 tests. To evaluate the performance of the proposed approach, the *Sensitivity* (as defined in (2)) will be used as the metric to evaluate the recognition accuracy.

$$\text{Sensitivity} = \frac{TP}{TP+FN}$$

(2)

where TP denotes the true positive, and FN denotes the false negative. As can be seen in Table II, the results obtained from the five subjects with different kinds of sign language words are quite satisfactory. The performance on Sensitivity can be as high as up to 94.56%.

Table III. The performance of the proposed architecture.

user	User A		User B		User C		User D		User E	
Sample	50		50		50		50		50	
	TP	FN	TP	FN	TP	FN	TP	FN	TP	FN
Lonely	46	4	48	2	49	1	46	4	47	3
Promote	49	1	47	3	49	1	46	4	49	1
Assist	47	3	47	3	45	5	48	2	42	8
Love	49	1	49	1	48	2	47	3	47	3
Protect	49	1	47	3	46	4	47	3	48	2
Sensitivity	0.96		0.952		0.948		0.936		0.932	
average	0.9456									

VI. CONCLUSION

We propose in this paper a real-time and portable system for Taiwanese sign language translation. The proposed translating system is realized on a wearable glove with flex sensors and inertial sensors. The flex sensors can detect the posture of the fingers while the inertial sensors can detect the orientation of the palm and the motion of the hand. With the proposed architecture, the gesture will be encoded and sent to the cell phone via Bluetooth for recognition process. By looking up the sign language database, we can translate the sign language into speech or text. With the proposed approach, an accuracy performance up to 94.56% on the sensitivity can be obtained when a set of five sign language words with 1250 tests are estimated by five subjects.

VII. REFERENCE

[1] Celestine. Preetham, Girish. Ramakrishnan, Sujan. Kumar, Anish. Tamse, "Hand Talk-Implementation of a Gesture Recognizing Glove," *Texas Instruments India Educators' Conference (TIIEC),* IEEE, 2013.

[2] Xu. Zhang, Xiang. Chen, Yun. Li, Vuokko. Lantz, Kongqiao. Wang, Jihai. Yang "A framework for hand gesture recognition based on accelerometer and EMG sensors." *IEEE Transactions on Systems, Man and Cybernetics*, pp. 1064-1076, 2011.

[3] Xiujuan. Chai, Guang. Li, Yushun. Lin, Zhihao. Xu, Yili. Tang, Xilin. Chen, and Ming. Zhou, "Sign Language Recognition and Translation with Kinect," *IEEE International Conference on Automatic Face and Gesture Recognition*, 2013.

[4] Rung-Huei Liang, Ming Ouhyoung, "A Real-time Continuous Gesture Recognition System For Sign Language," *IEEE International Conference on Automatic Face and Gessture Recognition,* pp. 558-567, Japan, 1998.

[5] Flex Sensor Datasheet https://www.sparkfun.com/datasheets/Sensors/Flex/FLEXSENSOR%28REVA1%29.pdf