

GESTURE CONTROLLED HOME AUTOMATION FOR DIFFERENTLY CHALLENGED PEOPLE

R.Prabhuraj| B.Saravanakumar

¹(Electronics and Communication Engineering, Anna University, Chennai, India, r.prabhuraj8@gmail.com)

²(Electronics and Communication Engineering, Anna University, Chennai, India Department, bsaravanakumars952@gmail.com)

Abstract— Everyday communication with the hearing population poses a major challenge to those with hearing loss. For this purpose, an automatic American Sign Language recognition system is developed using artificial neural network (ANN) and to translate the ASL alphabets into text and sound. A glove circuit is designed with flex sensors, 3-axis accelerometer and EMG sensors to capture the gestures. The finger bending data is obtained from the flex sensors on each finger whereas the accelerometer provides the trajectories of the hand motion. Some local features are extracted from the ASL alphabets which are then classified using neural network. Finger bending data is transmitted via zigbee and given to driver circuit to control the home appliances.

Keywords— Accelerometer, Artificial Neural Network, Electromyography, Flex Sensors, Sign Language Recognition.

1. INTRODUCTION

Gestures play a major role in the daily activities of human life, in particular during communication providing easier understanding. In other words, Gesture recognition refers to recognizing meaningful expressions of motion by a human, involving the hands, arms, face, head, and/or body. Between all the gestures performed hand gestures plays an important role which helps us to express more in less time. Now a day, Human-Machine interface has gained a lot of research attentions employing hand gestures. There are other applications which could be controlled by a gesture are investigated [1]-[3], [9], [10] such as media players, remote controllers, robots, and virtual objects or environments. Hand gestures particularly sign languages can be interpreted uniquely by investigating the patterns of four basic components such as hand shape also known as hand configuration, hand movement, orientation and classification [4], [5]. Two approaches that are commonly used to recognize the gestures are 1) Vision based systems and 2) Glove based systems. Vision based system provides a more natural and non-contact solutions such as camera to perceive the information of human motions and their surroundings. Because of complex background and occlusions this requires intelligent processing making it difficult to design. Though it is feasible for controlled environments it has lack of accuracy and processing speed. For e.g., the Arabic Sign Language gesture recognition [6] using a spatiotemporal feature extraction scheme provides an accuracy of 93% with its sensitiveness to the use environment such as background texture, color, and lighting [1], [7]. In order to enhance the robust performance of vision-based approaches, some previous studies utilized colour gloves [10] or multiple cameras [11] for accurate hand tracking, segmentation, and recognition. These conditions limit their applications in mobile environment. Absence level American Sign Language recognition system consisting of a wearable hat, mounted with camera achieves 97% accuracy based on real time

Hidden Markov Model. Thus, the combined accelerometer and the vision system

Improve the accuracy in noisy and problematic conditions. Home appliance control system is based on zigbee network technology for transmission of finger bending data from sender to receiver. Gesture data sending and receiving is used for ubiquitous access of appliances and allowing breach control at home.

Glove based systems on the other hand employs sensors attached to the glove captures the movement of the hand and finger and also the rotation. Many efforts have been made to interpret hand gestures, particularly the signals which are changing over time Hidden Markov Model (HMM) is employed as an effective tool in

Most of the works. A system with two data gloves and three position trackers as input devices and a fuzzy decision tree as a classifier is used to recognize the Chinese Sign Language gestures [8]. With a 5113 sign vocabulary, it achieves 91.6% recognition accuracy. The combined accelerometer and the surface electromyographic sensor provides an alternative method of gesture sensing unlike the above two methods mentioned previously. The Kinematic information of the hand and the arm are provided by the accelerometer and also it is capable of distinguishing the hand orientations or movements with different trajectories [12]. EMG Sensor measures the electrical activity provided by the skeletal muscles and the signal contain rich information for the co activation and coordination of multiple muscles associated with different sign gestures. Recent works on ACC and sEMG have demonstrated the improved recognition performances. For instance, it is shown that the sEMG combined with ACC can achieve accuracy in the range of 5–10% for recognition of various wrist and finger gestures [15]. The complementary functionality of both sensors is examined for the recognition of seven isolated words in German sign language [13]. The intrinsic mode entropy is successfully applied on ACC and sEMG data acquired from the dominant hand to recognize 60 isolated signs in Greek sign

language [14]. Recently, a wearable system using body-worn ACC and sEMG sensors is designed [16] to remotely monitor functional activity in stroke.

The goal of this paper is to develop a portable communication system having multiple sensors for American Sign Language Recognition and to transmit these gestures to driver circuit of appliances via zigbee to automate.

2. SYSTEM ARCHITECTURE

The basic concept of proposed system is the idea of developing glove-based automation aid in mobile environment. The architecture of automation system for sign language gesture recognition and controlling the appliances is illustrated (see Fig.1).

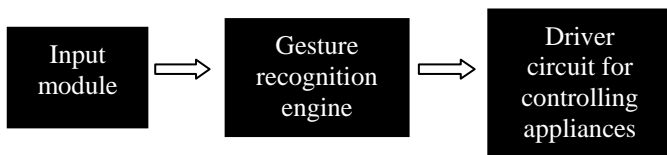


Fig.1. Overview of System Architecture

In addition, the system would synthesize speech using text to speech conversion module. The present work takes the first step toward the development of such interfaces providing minimized communication gap, easier collaboration and ease of sharing ideas & experience. The block diagram of hand gesture recognition and controlling system consists of two main units namely Hand-Gloves and the receiver station (controlling unit) is shown (See fig. 2). The glove circuit is designed with multiple sensors whereas base station consists of controller unit, text to speech conversion module and in the receiver module, driver is placed to control the appliances. The glove circuit consists of one flex sensors for each finger, one 3-axis accelerometer and three sEMG sensors with one being a reference electrode.

The flex sensor produces the change in resistance value depending on the amount of degree bend of each finger and the corresponding hand movement, orientation is reported by the tri-axial accelerometer. The sEMG sensors are used to measure the muscle activity of the hand while performing gestures in terms of electrical signals. The signal from the flex sensor and the EMG sensors are filtered to remove the noise signals and amplified before to be fed into ADC whereas the accelerometer is connected using an interface circuitry. Thus, the captured outputs are fed into the in-built ADC of ARM LPC2148 micro-controller for analog to digital conversion. The output from microcontroller's ADC is interfaced to the PC via RS232 cable and the inputs are classified using MATLAB. The recognized gesture is converted and displayed into corresponding text and speech using text to speech conversion module.

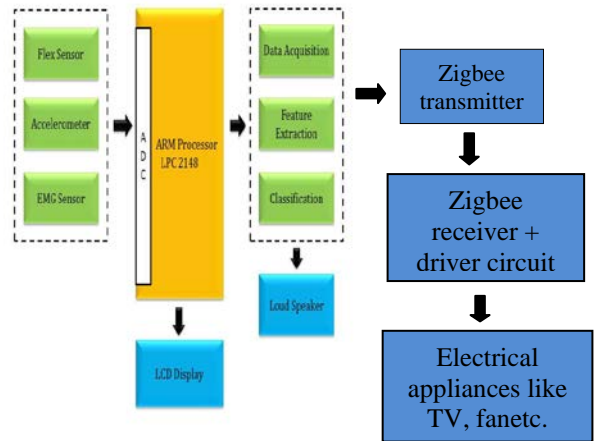


Fig 2: Block Diagram for gesture controlled home automation System

3. INPUT MODULE

Data gloves, a simple electronic glove that transforms the hand and finger movement into real time data for applications. The proposed data glove not only has the information of finger position but the movement, orientation of the hand with electrical signal emanating from the muscle activity as well. The sensors that are selected to be attached to the data glove to capture the hand gestures are: Flex sensors and Accelerometer. Many materials are used for the glove including leather, cotton, and plastic. The cotton gloves proved to be ideal for this application, since the sensor is attached firmly and the glove can easily be removed without destroying the sensors.

3.1 Flex Sensor

Flex sensors are sensors that changes the resistance depending upon on the amount of bend on the sensor. They convert the change in bend to electrical resistance- the more the bend, the more the resistance value. They are usually in the form of a thin strip from 11 – 51 long that vary in resistance. They can also be made uni-directional or bi-directional. The flex sensors are made with the same principle as strain gauges (changing their resistance on the bending occasion), but they have large resistance differences. Inside the flex sensors are the carbon resistive elements within a thin flexible substrate when bent produces a resistance output relative to the bend. Flex sensors works in the principle of voltage divider form and the basic flex sensor circuitis shown (See fig. 3).

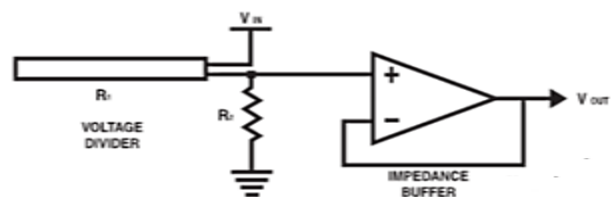


Fig 3: Basic Flex Sensor Circuit

3.2 Accelerometer

Accelerometer is an electromechanical device that will measure acceleration forces. These forces may be static, like the constant force of gravity pulling at your feet, or they could be dynamic - caused by moving or vibrating the accelerometer. The MMA7260QT low cost capacitive micro-machined accelerometer (See fig. 4) features signal conditioning, a 1-pole low pass filter, temperature compensation and g-Select which allows for the selection among 4 sensitivities.

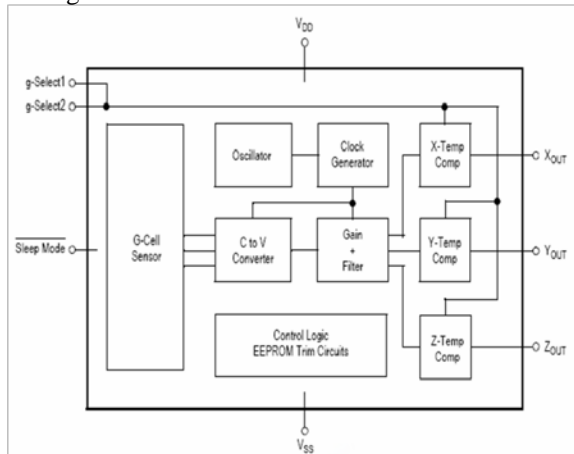


Fig 4: Functional Block diagram for MMA7260QT Accelerometer

The device consists of two surface micro-machined capacitive sensing cells (g-cell) and a signal conditioning ASIC contained in a single integrated circuit package. The sensing elements are sealed hermetically at the wafer level using a bulk micro-machined cap wafer. The g-cell is a mechanical structure formed from semiconductor materials (polysilicon) using semiconductor processes (masking and etching). It can be modeled as a set of beams attached to a movable central mass that move between fixed beams. The movable beams can be deflected from their rest position by subjecting the system to acceleration. As the beams attached to the central mass move, the distance from them to the fixed beams on one side will increase by the same amount that the distance to the fixed beams on the other side decreases. The change in distance is a measure of acceleration. The g-cell beams form two back-to-back capacitors Figure. As the center beam moves with acceleration, the distance between the beams changes and each capacitor's value will change, $(C = A\epsilon/D)$. Where A is the area of the beam, ϵ is the dielectric constant, and D is the distance between the beams. This accelerometer is placed on the back of fore-arm to capture information about hand orientation and trajectories.

3.3 EMG Sensors

Surface EMG is relatively easy to use as compared to other EMG electrodes. This is the reason why it is being extensively used in the control of robotic mechanisms to achieve prosthesis. Other kinds of EMG electrodes (needle and fine wire), when inserted into the skin of the subject,

may effect a twitching sensation and cause him or her to make movements. In order to get the best results from sEMG, it is really important to have a proper understanding of the muscles from which the EMG signal is being extracted. The placement on skin also requires adequate study and requires skin preparation beforehand as well. Three electrodes are used to measure the EMG waves in which two electrodes are fixed on flexor carpi radialis, extensor carpi radialis longus and another one electrode is fixed as reference ground electrode. The flexor carpi radialis originates from the medial epicondyle of the humerus. It inserts at the proximal aspect of the second and third metacarpals. The extensor carpi radialis longus originates from the lateral epicondyle of the humerus and inserts at the proximal aspect of the second metacarpal on the dorsal aspect of the hand. This muscle lies predominantly on the lateral and dorsal aspect of the forearm rather than the ventral aspect like the other two muscles [21]. The signal from the EMG detecting surfaces is gathered with respect to a reference. An EMG reference electrode acts as a ground for this signal. It should be placed far from the EMG detecting surfaces, on an electrically neutral tissue. Electrode 1 and Electrode 2 pick up the EMG waves from the hands. Once the electrode is properly placed and the signal is extracted, noise plays a major role in hampering the recording of the EMG signal. For this purpose, the signal has to be properly filtered, even after differential amplification (See fig. 5). The noise frequencies contaminating the raw EMG signal can be high as well as low. Low frequency noise can be caused from amplifier DC offsets, sensor drift on skin and temperature fluctuations and can be removed using a high pass filter. High frequency noise can be caused from nerve conduction and high frequency interference from radio broadcasts, computers, cellular phones etc. and can be deleted using a low pass filter.

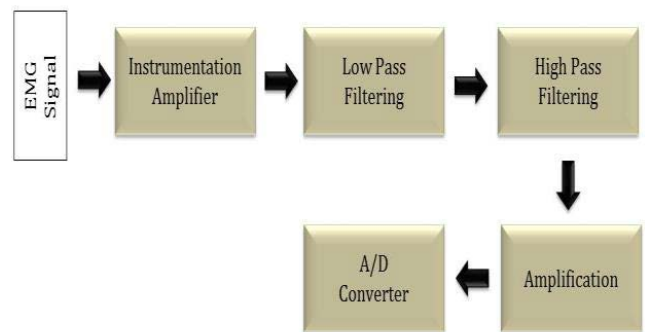
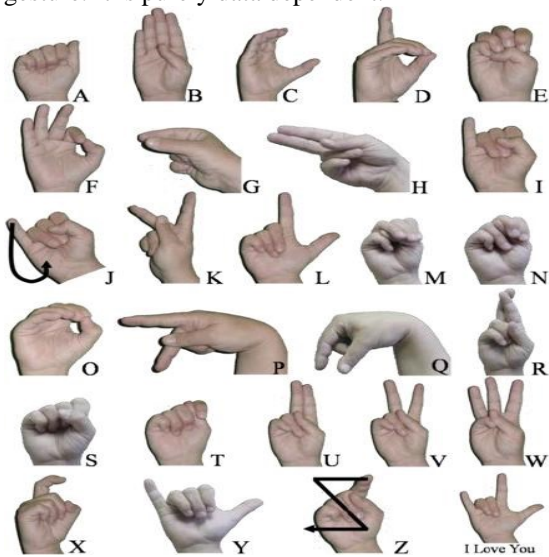


Fig 5: Block diagram for EMG signal conditioning

3.4 Sign Language

In this project all operations were performed on American Sign Language (ASL) (See fig. 6). In the ASL manual alphabet, fingerspelling is used primarily for spelling out names or English terms which do not have established signs. The database consists of 26 ASL alphabets in which the letters j and z are dynamic gestures whereas all the others are static gestures. ASL signs have a number of

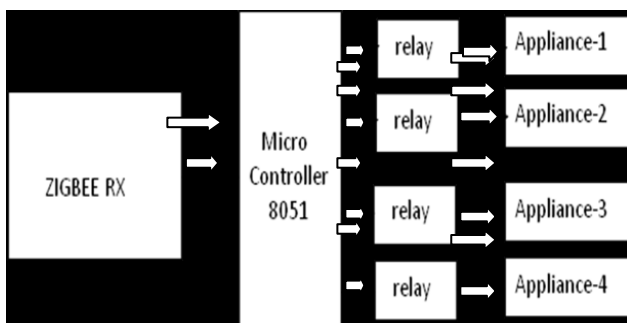
phonemic components, including movement of the face and torso as well as the hands. ASL is not a form of pantomime, but iconicity does play a larger role in ASL than in spoken languages. When communicating with hearing English speakers, ASL-speakers often use what is commonly called Pidgin Signed English (PSE) or 'contact signing', a blend of English structure with ASL. A low-cost hand glove circuit developed with multiple sensors is used to capture the hand gestures performed by the performer. It produces the finger flexion of each finger, the movement and orientation of the hand and the electrical signal from the muscle activities of the hand. The system works online gesture recognition i.e., the real time signal from the gloves is given as an input and the system tells us the matched gesture. It is purely data dependent.



ASL Hand Shapes

3.5 Receiver section

Receiver section consists of zigbee receiver, microcontroller, relays and appliances. Once the gesture commands are recognized, control characters are sent to the specified appliance address through ZigBee communication protocol. Each appliance that has to be controlled has a relay controlling circuit. Control characters corresponding to the recognized commands are then sent serially from the central controller module to the appliance control modules that are connected to the home appliances.



Appliance control module

ZigBee RF communication

Zigbee protocol is the communication protocol that's used in this system. Zigbee offers 250 kbps as maximum baud rate, however, 115200 bps was used for sending and receiving as this was the highest speed that the UART of the microcontroller could be programmed to operate at. For each byte transmitted, there is a start and stop bit. Hence the actual baud rate is:

$$\text{Actual baud rate} = \text{configured baud rate} * (8/10)$$

$$\text{Actual baud rate} = 115200 * (8/10) = 92160$$

ZigBee Key Features:

- Low Power
- Robust
- Mesh Networking
- Interoperability

4. GESTURE RECOGNITION ENGINE

4.1 Feature Extraction

The block diagram of gesture recognition engine using multichannel EMG signal, 3-axis ACC and flex sensors is shown below (See fig. 8). In this study the feature vectors are extracted to develop a national system for writing signs, containing symbols defined by hand shape, hand location, hand movement and hand orientation. The 3-D accelerometer measures the rate of change of velocity along three axes (x, y, z) when hand gestures are performed. Some of the statistical features such as Mean value and the standard deviation (SD) of each axis are extracted. These simple features will be used by the following ANN classifier. The hand-shape recognition depends on finger bending data, which are collected from the custom glove where the term hand shape refers to the shape of the hand while performing the sign. For this purpose we use 5 sensors of the glove one for each of the thumb, index, middle, ring and little finger. Various kinds of features for the classification of the EMG have been considered in the literature [17], [18]. These features have included a variety of time domain, frequency-domain, and time-frequency-domain features. It has been shown that some successful applications can be achieved by time-domain parameters, for example, zero-crossing rate and root mean square (RMS). The autoregressive (AR) model coefficients [19] of the EMG signals with a typical order of 4-6 yield good performance for myoelectric control. Many time-frequency approaches, such as short-time Fourier transform, discrete wavelet transform, and wavelet packet transform, have been investigated for EMG feature extraction. However, time-frequency-domain features require much more complicated processing than time-domain features. In our work the time domain features such as Mean Absolute Value (MAV) and the Root Mean Square (RMS) is calculated.

4.2 Artificial Neural Network Model

A back propagation algorithm is used for training the ANN model. The basic structure and formulation of back propagation is summarized here. Training a neural network involves computing weights so as to get an output response to the input within an error limit. The input and target vectors make up a training pair. The back propagation algorithm includes the following steps:

- (1) Select the first training pair and apply the input vector to the net,
- (2) Calculate the net output,
- (3) Compare the actual output with the corresponding target and find the error,
- (4) Modify the weights so as to reduce the error.

These steps are repeated until the error is within the accepted limits. In Step 2, the output sets for test inputs are calculated. If they are the same as the expected sets within an error range, then it is considered that the net has learned the problem, and the final weights are stored so that they can be reused when needed. The developed ANN has a multi-layer feed forward structure.

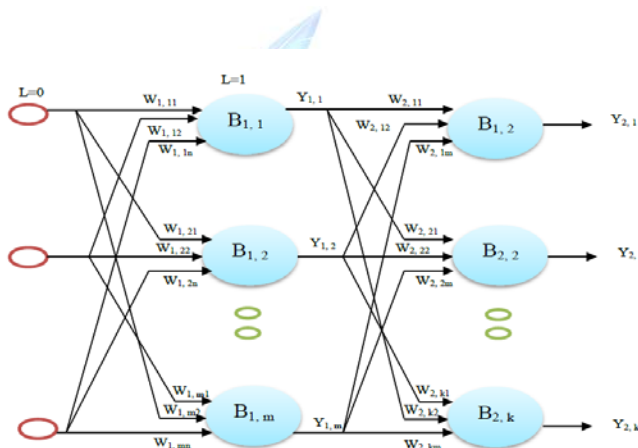


Fig 7: Multi-Layer Feed Forward Network Structure

The variable definitions are given as follows: L=0: input layer; L=1: hidden layer; L=2: output layer; $W1, ji$: weight matrix between the input layer and the hidden layer; $W2,tj$: weight matrix between the hidden and the output layer; $B1,j$: bias values of hidden neurons; $B2,t$: bias values of output neurons.

In the sign created with both hands, the right hand is often more active than the left hand. Along with the hand the speakers also support their sign with the head, eyes and facial expressions. In this paper, the right hand words only are studied. A multi-layer ANN is designed to recognize a set of ASL alphabets. In this study the 26 ASL alphabets are considered for demonstration though it is flexible for any number of words. The speaker's signing with the hand position and finger bending is digitized. The digitized values are shown in the table. 1. The extracted feature

vectors of ASL alphabets are provided as input to the network

For every ASL sign, we use a total of 45 features which are extracted from five different signers: 25 for hand shape, 10 for hand movement (x, y, z) and 10 for EMG feature sequences which differentiates subtle finger configurations. The system is designed to recognize a single word as a whole at one time. A Levenberg–Marquardt back propagation algorithm is used for training. The ANN is trained and tested for two different data sets, single-user data and multi user data. The output file consists of 26 outputs, each representing one word or expression. The training set of alphabets, which can be signed with the right hand is selected from an ASL dictionary.

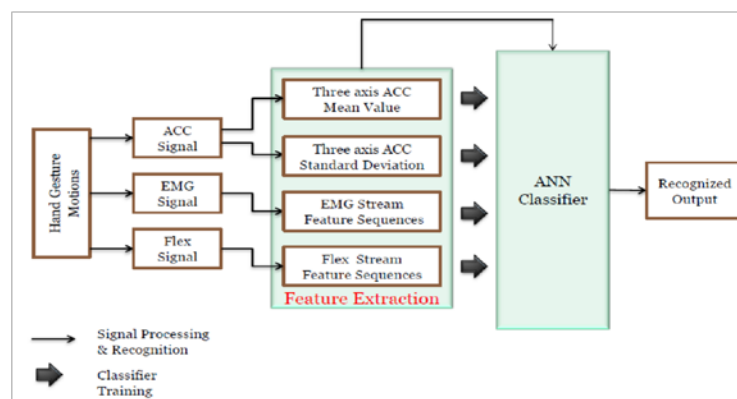
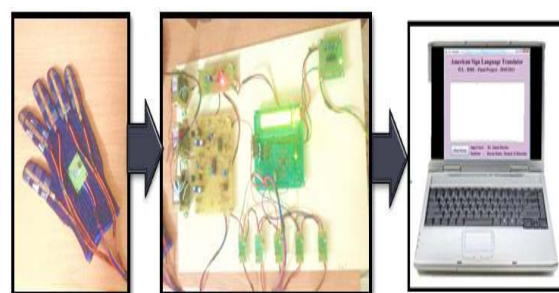


Fig 8: Block Diagram for Feature Extraction and Classification

5. TEST RESULTS

The ASL recognition systems developed is tested with two types of data sets, one with single-user data and the other with multi-user data. In both tests, the ASL recognition system trained with five samples of data with 26 alphabets. At the testing stage, real-time data were used. In total, 130 ASL signs (5 *26) in the training set were used for the test. Both the single-user model and the multiple-user model were tested in two different ways, sequentially and randomly. Sequential testing results were better than random testing results. Test results show that the recognition accuracy of the system is about >95 %. The hardware snapshot of the ASL recognition system is shown below



Hardware Snapshot of gesture controlled automation System

6. CONCLUSION

The goal of this project is to design a useful and fully functional real-world product that efficiently translates the movement of the fingers for the fingerspelling of American Sign Language (ASL). Our motivation is to help differently able people to control the electronic appliances more easily. The gesture control automation system also teaches people to learn the ASL and it uses a glove to recognize the hand positions and outputs the ASL onto a display and control the electronic devices like fan, television etc. The data are collected from the sensory glove and some local features were extracted for every ASL word. Neural networks were used to classify these feature vectors. The system was trained and tested for single and multiple users for a vocabulary of 26 ASL words. The proposed method can also be extended to recognize any number words without modifying the system, just requiring a further training of the network.

7. REFERENCES

- [1] J. Mäntyjärvi, J. Kela, P. Korpipää, and S. Kallio, 2004 —Enabling fast and effortless customisation in accelerometer based gesture interaction, | in Proc. 3rd Int. Conf. Mobile Ubiquitous Multimedia, New York, and pp. 25–31.
- [2] T. Pylvänäinen, 2005, —Accelerometer based gesture recognition using continuous HMMs, | in Proc. Pattern Recog. Image Anal., LNCS 3522, pp. 639–646.
- [3] S. C. W. Ong and S. Ranganath, 2005, —Automatic sign language analysis: A survey and the future beyond lexical meaning, | IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 6, pp. 873–891.
- [4] L. Ding and A. M. Martinez, 2009, —Modelling and recognition of the linguistic components in American sign language, | Image Vis. Comput., vol. 27, no. 12, pp. 1826–1844.
- [5] X. Zhang, X. Chen, Y. Li, V. Lantz, K. Wang, and J. Yang, 2011, | A Framework for Hand Gesture Recognition Based on Accelerometer and EMG Sensors, | IEEE Trans. on Sys. Man and Cybernetics—Part a: Sys. And Humans, Vol. 41, No. 6,
- [6] T. Shanableh, K. Assaleh, and M. Al-Rousan, 2007, —Spatio-temporal feature extraction techniques for isolated gesture recognition in Arabic Sign Language, | IEEE Trans. Syst., Man, Cybern. B, Cybern, vol. 37, no. 3, pp. 641–650.
- [7] S. Mitra and T. Acharya, 2007, —Gesture recognition: A survey, | IEEE Trans. Syst., Man, Cybern. C, Appl. Rev., vol. 37, no. 3, pp. 311–324.
- [8] G. Fang, W. GAO, and D. Zhao, 2004, —Large vocabulary sign language recognition based on fuzzy decision trees, | IEEE Trans. Syst., Man, and Cybern. A, Syst., Humans, vol. 34, no. 3, pp. 305–314.
- [9] T. S. Saponas, D. S. Tan, D. Morris, and R. Balakrishnan, 2008, —Demonstrating the feasibility of using forearm electromyography for muscle–computer interfaces, | in Proc. 26th SIGCHI Conf. Human Factors Comput. Syst., Florence, Italy, pp. 515–524.
- [10] T. Starner, J. Weaver, and A. Pentland, 1998, —Real-time American Sign Language recognition using desk and wearable computer based video, | IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 12, pp. 1371–1375.
- [11] C. Vogler and D. Metaxas, 1999, —ASL recognition based on a coupling between HMMs and 3D motion analysis, | in Proc. 6th Int. Conf. Comput. Vis., Bombay, India, pp. 363–369.
- [12] D. M. Sherrill, P. Bonato, and C. J. De Luca, 2002, —A neural network approach to monitor motor activities, | in Proc. 2nd Joint EMBS/BMES Conf., Houston, TX, vol. 1, pp. 52–53.
- [13] K. R. Wheeler, M. H. Chang, and K. H. Knuth, 2006, —Gesture-based control and EMG decomposition, | IEEE Trans. Syst., Man, Cybern. C, Appl. Rev., vol. 36, no. 4, pp. 503–514.
- [14] V. E. Kosmidou and L. J. Hadjileontiadis, 2009, —Sign language recognition using intrinsic mode sample entropy on sEMG and accelerometer data, | IEEE Trans. Biomed. Eng., vol. 56, no. 12, pp. 2879–2890.
- [15] X. Chen, X. Zhang, Z. Y. Zhao, J. H. Yang, V. Lantz, and K. Q. Wang, 2007, —Hand gesture recognition Research based on surface EMG sensors and 2D-accelerometers, | in Proc. 11th IEEE ISWC, pp. 11–14.
- [16] S. H. Roy, M. S. Cheng, S. S. Chang, J. Moore, G. De Luca, S. H. Nawab, and C. J. De Luca, 2009, —A combined sEMG and accelerometer system for monitoring functional activity in stroke, | IEEE Trans. Neural Syst. Rehabil. Eng., vol. 17, no. 6, pp. 585–594.
- [17] V. E. Kosmidou, L. J. Hadjileontiadis, and S. M. Panas, 2006, —Evaluation of surface EMG features for the recognition of American Sign Language gestures, | in Proc. IEEE 28th Annu. Int. Conf. EMBS, New York, pp. 6197–6200.
- [18] R. N. Khushaba and A. Al-Jumaily, 2007, —Channel and feature selection in multifunction myoelectric control, | in Proc. IEEE 29th Annu. Int. Conf. EMBS, Lyon, France, pp. 5182–5185.
- [19] X. Hu and V. Nenov, 2004, —Multivariate AR modeling of electromyography for the classification of upper arm movements, | Clinical Neurophysiol., vol. 115, no. 6, pp. 1276–1287.
- [20] Lippman R.P., 1987 —An introduction to computing with neural nets, | IEEE ASSP, pp. 2–22.
- [21] Muhammad Zahak Jamal, 2012 —Signal Acquisition Using Surface EMG and Circuit Design Considerations for Robotic Prosthesis, | Intech, pp. 427–448.