Gesture Recognition Based On Hand Postures And Trajectories By Using Dataglove: A Fuzzy Probability Approach – A Review

Gesture Recognition Based On Hand Postures And Trajectories By Using Dataglove: A Fuzzy Probability Approach – A Review

Khairunizam Wan[1,2], Nazrul Hamizi Bin Adnan[1,3], Shahriman AB[1,2], Siti Khadijah Za'ba[1,2], Mohd Azri Abd Aziz[1,2], and Zulkifli Md. Yusof[4]

[1]Advanced Intelligent Computing and Sustainability (AICoS) Research Group
[2]School Of Mechatronic Engineering, Universiti Malaysia Perlis, 02600 Arau, Malaysia
[3]Institut Kemahiran MARA Beseri Perlis, Malaysia
[4]Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia

Abstract— **Many deaf people communicate with hearing people either through interpreter or text writing. However, some of them are able to use sign language (SL). SL is a language which determines a form of nonverbal communication in which visible bodily actions are used to communicate particular messages, either in place of speech or together and in parallel with spoken words. Nowadays, many facilities for the deaf and could not speak person was created. Among them are quite popular right now is Data glove. A device capable of recording hand movements, both the position of the hand and its orientation as well as finger movements; it is capable of simple gesture recognition and general tracking of three-dimensional hand orientation.**

Keywords-component; sign language; SL; nonverbal communication; Data Glove; Gesture Recognition and General tracking of three-dimensional hand orientation

I. INTRODUCTION

A **wired glove** (sometimes called a "dataglove" or "cyberglove") is an input device for human–computer interaction worn like a glove. Various sensor technologies are used to capture physical data such as bending of fingers. Often a motion tracker, such as a magnetic tracking device or inertial tracking device, is attached to capture the global position/rotation data of the glove. These movements are then interpreted by the software that accompanies the glove, so any one movement can mean any number of things. Gestures can then be categorized into useful information, such as to recognize Sign Language or other symbolic functions. Expensive high-end wired gloves can also provide haptic feedback, which is a simulation of the sense of touch. This allows a wired glove to also be used as an output device. Traditionally, wired gloves have only been available at a huge cost, with the finger bend sensors and the tracking device having to be bought separately.

Many researchers have been working on the recognition of various sign language and gestures [1]. It also involves the study of specialization in hand posture recognition with no constrain on the shapes because the human hand gestures or shape is a complex articulated object consist of many connected parts and joints. Hand gestures can be classified into two categories: 1) static hand gestures which rely only on the information about the flexure angles of the fingers and 2) dynamic hand gestures which rely not only on the fingers' flex angles but also on the hand trajectories and orientations. Dynamic hand gestures can be further divided into two subclasses. The first subclass consists of hand gestures involving hand movements, and the second subclass consists of hand gestures involving only the fingers' movements without changing the position of the hands [2]. This paper surveys studies on gesture recognition techniques by using Data glove. The organization of this paper is as follows; Section 2 surveys methods used for hand gesture recognition. Approaches used for SL recognition researches are reviewed in Section 3. Section 4 states the conclusion.

II. HAND GESTURE RECOGNITION

The first step in SLs systems is to detect and track both hands, however this is a complex task because hands may occlude each other's and/or face [3]. Gesture recognition is important, because it is a useful communication tool between humans and computer. It makes many models and recognizes gesture has attracted the attention of many researchers. The recognition of continuous gestures suffers greatly from the existences of non-gesture hand motions. The given gestures can start at any moment in an input sequence

A. Hidden Markov Models (HMM) Hand Gesture Recognition

This property makes HMMs appear to be an ideal approach for hand gesture recognition [5]–[8]. The price paid for the efficiency in this case is that we have to collect a great amount of data and a lot of time is required to estimate corresponding parameters in HMMs. HMM is a statistical model where the distributed initial points work well and the output distributions are automatically learned by the training process. In addition, HMM is widely used in hand writing, speech and character recognition [4][9]. Another advantage of HMM is that it is capable of modelling spatio-temporal time series where the same gesture can differ in shape and duration. There are two major problems arising in real-time gesture recognition system for continuous gesture to extract meaningful gesture. The first problem is the segmentation that means how to determine when a gesture starts and when it ends from hand motion trajectory, which is due to the intermediate movement of the hand between two gestures. The second problem is that the same gesture varies in shape and duration [10].

The most important thing in HMM hand gesture recognition is what the input features are that best represent the characteristics of the moving hand gesture. It considers a planar hand gesture in front of a camera and use 8-directional chain codes as input vectors as an example. For training an HMM network, a simple context modelling method is embedded for training on a "left-to-right" HMM model. This model is applied to drawing and editing specified graphic elements [11]. In particular, the proposed system consists of three main stages; an automatic hand segmentation and tracking, feature extraction and classification [12].

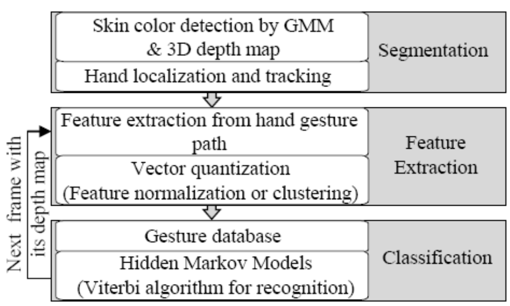

Figure 1. Simplified structure showing the main computational modules for isolated gesture recognition system by HMM.

B. Spatio-Temporal Hand Gesture Recognition

Several different types of recurrent networks and learning algorithms have been proposed in the past several years. A brief overview for these recurrent networks can be found in [13]-[15]. One may find that although recurrent networks can be trained to deal with those tasks it is by no means the easiest way because it usually takes a lot of time to train a recurrent network. Actually the first task-spatial-temporal pattern recognition does not necessarily require a recurrent network. The simplest way to recognize spatio-temporal patterns is first to turn the temporal sequence into a spatial pattern and then to employ the template match technique. Patterns are identified by comparing the input pattern to a list of stored pattern representations. The stored pattern representations are the templates. An input pattern to be recognized is passed to a comparator which performs a similarity measure between it and each of a set of pre stored template patterns; the comparison that produces the best match is deemed to be the recognized pattern.

Spatio-temporal signal processing may be classified into three tasks: (1) spatio-temporal pattern recognition: To produce a particular output pattern when a specific input sequence is seen. For example, this is appropriate for speech recognition problems and spatio-temporal hand gesture recognition problems; (2) sequence reproduction: In this case one tries to generate the rest of a sequence when part of the sequence is seen; and (3) temporal association: A particular output sequence must be produced in response to a specific input sequence [16].

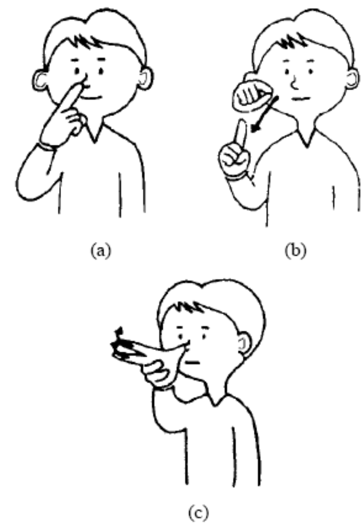

Figure 2. Examples of sign words in (Taiwanese Sign Language) TSL: (a) static sign word "I"; (b) dynamic sign word "one month"; and (c) dynamic sign word "mouse" [2].

C. Pseudo Two Dimension Hidden Markov Model (P2-DHMMs)

Many approaches to gesture recognition have been developed. A large variety of techniques have been used for modeling the hand. An approach based on the 2D locations of fingertips and palms was used by Davis and Shah [17]. N.D.Binh, E.Shuichi and T. Ejima use a Kalman filter and hand blobs analysis for hand tracking [18]. Bobick and Wilson developed dynamic gestures, which have been handled using framework [19]. Previous attempts to develop hand gesture recognition systems employed geometric feature-based methods, template-based methods, and active statistic models [20]. Given the success of HMMs in speech, it is also used successfully in hand gesture recognition systems [21]. Motivated by the desire to provide users with a capable gesture recognition system, Pham The Bao, Nguyen Thanh Binh and Tu Duy Khoa developed Tower method to obtain hand region and use P2D HMMs to recognize hand gesture [22]. A gesture is a specific combination of hand position, orientation, and flexion observation at some time instance. The system indentifies a gesture based upon the temporal sequence of hand regions in the image frame. The output of hand tracking process is the input of the recognition process. The hand region can be extracted by Tower Tracking Method.

After obtaining the hand region, a P2D HMMs are used to recognize the gesture.

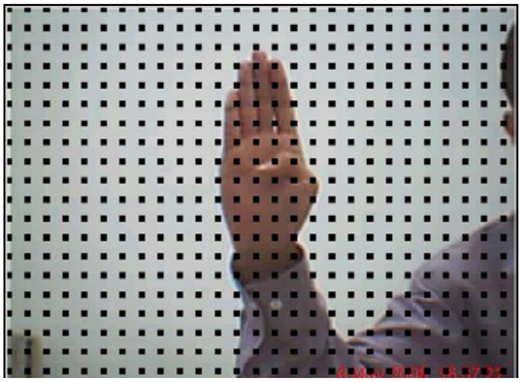

Figure 3. Generating towers in the image [22]

P2-DHMMs are made up of five levels of states, one for each significant facial region in which the input frontal images are sequenced: forehead, eyes, nose, mouth and chin. Each of P2-D HMMs has been trained by coefficients of an artificial neural network used to compress a bitmap image in order to represent it with a number of coefficients that is smaller than the total number of pixels. These systems require a robust algorithm especially for human face recognition under different lighting conditions, facial expressions and orientations [23].

III. ARTIFICIAL INTELLIGENT FOR HAND GESTURE RECOGNITION

The most important American Sign Language (ASL) feature is hand shape, which comes from the study of structural linguistics. Stokoe (1960) [24] used a structural linguistic framework to analyze ASL sign formulation. Most of the hand shapes represent the ASL alphabet, basic numbers, and some particular words. ASL has 36 hand shapes.

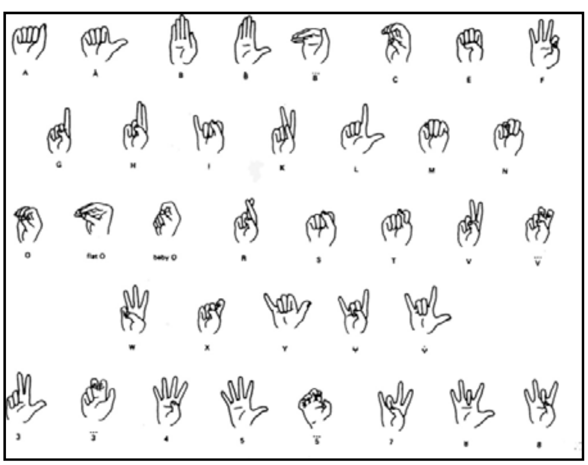

Figure 4. ASL hand shapes.

American Sign Language (ASL) is the language most commonly used by the American deaf community for interpersonal communication. This is a language acquired naturally by deaf infants, if they are exposed to it early, in a manner similar to hearing children acquiring speech [26]. But normal speech acquisition is at best difficult if not impossible for most deaf person [27], [28], [29]. The ASL code is naturally different from the phonetic code of the spoken language [25]. It uses hand, body, and facial expressions to make signs and convey ideas or feelings. While sign language may be expressed manually, inclusion of other components makes it rich, pleasing, and informational, similar to intonation, etc., in spoken languages. Further, syntactic aspects can be represented via non manual representations.

A. Artificial Neural Networks

An artificial neural network (ANN), usually called neural network (NN), is a mathematical model or computational model that is inspired by the structure and/or functional aspects of biological neural networks. A neural network consists of an interconnected group of artificial neurons, and it processes information using a connectionist approach to computation. In most cases an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase. Modern neural networks are non-linear statistical data modelling tools. They are usually used to model complex relationships between inputs and outputs or to find patterns in data. Artificial neural networks are parallel computer algorithms which have been used for pattern recognition. These recognition algorithms are robust when variation in input data is present [30]. Neural networks perform mapping functions from an n-dimensional input space to an m-dimensional output space. They learn this mapping from sample inputs through a "learning" process. During the learning or training process, through repeated presentations, the network develops associations between sample inputs and outputs [30]. Once the network is fully trained, it retains the learned mapping function.

Since 1990, Artificial Neural Networks have been used widely for solving engineering and industrial problems. Because of the popularity of ANNs, sign language researchers have applied this algorithm to solve their problems. Kramer and Leifer developed an ASL finger spelling system using a Cyberglove with the use of a neural network for feature classification and sign recognition by Kramer and Leifer, 1990 [31]; Kramer, 1996 [32]. Murakami and Taguchi (1991) [33] established a recurrent neural network method to recognize 110 distinct Japanese Sign Language signs. Waldron and Kim (1995) [34] used a neural network method to recognize 14 ASL signs using a different network for hand shape and hand orientation and position.

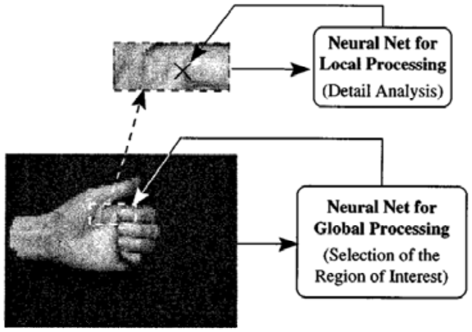

Figure 5. Hierarchical processing: a first neural network computes a coarse estimate of finger tip position based on global image data.

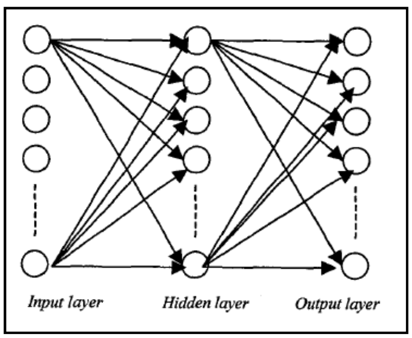

Figure 6. Model of Neural Network

B. *Fuzzy Inference Systems*

Fuzzy sets were introduced in 1965 by Zadeh [35] for representing vagueness in everyday life, providing an approximate and effective means for describing the characteristics of a system that is too complex or ill-defined to be described by precise mathematical statements. In a fuzzy approach the relationship between elements and sets follows a transition from membership to non membership that is gradual rather than abrupt. A fuzzy system implements a function (usually nonlinear) of *n* variables, given by a linguistic description of the relationship between those variables. Figure 7 illustrates the architecture of standard fuzzy systems. The fuzzificator computes the membership degrees of the crisp input values to the linguistic terms (fuzzy sets) associated to each input linguistic variable. The rule base contains the inference rules that associate linguistic terms of input linguistic variables to linguistic terms of output linguistic values. The information manager is responsible for searching in the rule base which rules are applicable for the current input. The inference machine determines the membership degrees of the output values in the output sets, by the application of the rules selected in the rule base. The defuzzificator gives a single output value as a function of the output values and their membership degrees to the output sets. Applications were found in control systems [36], decision making [37], expert systems [38], etc.

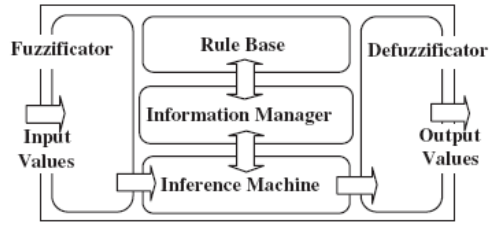

Figure 7. A standard fuzzy system

However, many approximate methods do not produce a single final result, presenting several alternative solutions to a single problem (e.g., the different classes to which a given input may belong). Among them, several fuzzy rule based methods for pattern recognition [39, 44], fuzzy relations, fuzzy clustering, fuzzy neural systems [45] were developed, with applications to signature verification [46], and face recognition [47], for example. Now, let consider a hypothetical data glove with 19 sensors, as shown in Fig. 8. The fingers are labelled as: F1 (little finger), F2 (ring finger), F3 (middle finger), F4 (index finger) and F5 (thumb). The joints in the fingers are labelled as J1 (the knuckle), J2 and J3, for each finger. A separation between two fingers is labelled as Sij to indicate that it is a separation between the fingers Fi and Fj.

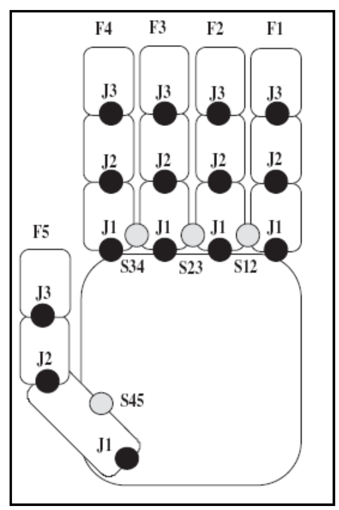

Figure 8. Localization of sensors in the data glove [48]

Since any movement can be represented as a sequence of frames, a hand movement using a data glove is represented as a sequence of hand configurations, one for each discrete time instant [48].

IV. CONCLUSION

In writing this paper, we have presented several methods that can be used to produce a prototype that is useful for hand gesture and recognition. From the literature suggests that the use of artificial neural network (ANN) is very broad as well as the use of Fuzzy Inference System. There are many kinds of methods can be used to ensure a successful study of hand gestures. Our study will be more focusing on the specialization of *Fuzzy Probability Approach*. This project was meant to be a prototype to check the feasibility of recognizing sign languages using sensor gloves / data glove. The completion of this prototype suggests that sensor gloves can be used for partial sign language recognition. 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. In the end, adding a speech engine to speak the translated text would help enhance ease of use.

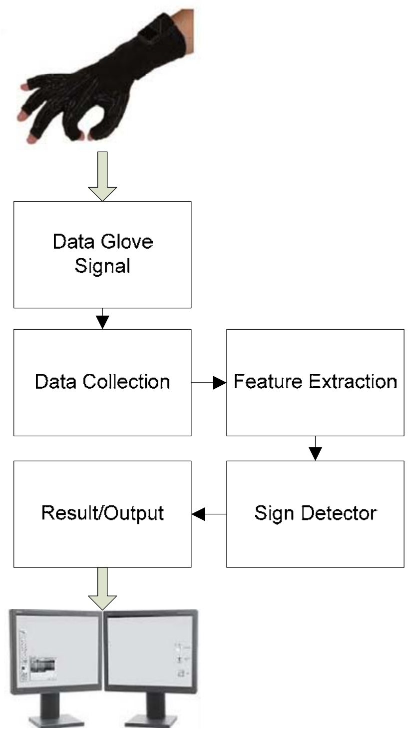

Figure 9. Overall structure of the system

ACKNOWLEDGMENT

This works is supported by Majlis Amanah Rakyat (MARA) and Universiti Malaysia Perlis (UNIMAP) by the fundamental research grant scheme (FRGS 9003-00214) awarded by the Ministry of Higher Education, Malaysia.

REFERENCES

[1]. Cemil Oz, Ming C. Leu (2011) “American Sign Language word recognition with a sensory glove using artificial neural networks” Engineering Applications of Artificial Intelligence 24, 1204–1213.

[2]. Mu-Chun Su (2000) “A Fuzzy Rule-Based Approach to Spatio-Temporal Hand Gesture Recognition” IEEE Transactions On Systems, Man, And Cybernetics, VOL. 30, NO. 2, MAY 2000.

[3]. Sara Bilal, Rini Akmeliawati, Momoh Jimoh El Salami, Amir A. Shafie, (2011) “Vision-based Hand Posture Detection and Recognition for Sign Language-A Study” International Conference on Mechatronics (ICOM) 2011.

[4]. R. Rebiner, “A tutorial on hidden Markov models and selected applications in speech recognition,” Proc. IEEE, vol. 77, no. 2, pp. 257–285, 1989.

[5]. H. Liang and M. Ouhyoung, “A sign language recognition using hidden Markov model and context sensitive search,” in ACM Symp. Virtual Reality Software Technology, 1996, pp. 59–66.

[6]. K. Grobel and M. Assan, “Isolated sign language recognition using hidden Markov models,” in IEEE Int. Conf. Systems, Man, Cybernetics, 1997, pp. 162–167.

[7]. Y. Nam and K. Y. Wohn, “Recognition of space-time hand-gestures using hidden Markov model,” in ACM Symp. Virtual Reality Software Technology, 1996, pp. 51–58.

[8]. C. Vogler and D. Metaxas, “Adapting hidden Markov models for ASL recognition by using three-dimensional computer vision methods,” in IEEE Int. Conf. Systems, Man, Cybernetics, 1997, pp. 156–161.

[9]. M. Elmezain, A. Al-Hamadi, and B. Michaelis, “Real-Time Capable System for Hand Gesture Recognition Using Hidden Markov Models” in Stereo Color Image Sequences, The Journal of W S C G’08, Vol. 16(1), pp. 65-72, 2008.

[10]. Mahmoud Elmezain, Ayoub Al-Hamadi, J¨org Appenrodt, and Bernd Michaelis (2008) “A Hidden Markov Model-Based Isolated and Meaningful Hand Gesture Recognition” World Academy of Science, Engineering and Technology 41 2008, 393-400.

[11]. Byung-Woo Min, Ho-Sub Yoon, Jung Soh, Yun-Mo Yang, Toshiaki Ejima (1997) “Hand gesture recognition using hidden Markov models” IEEE International Conference on 1997, 4232 - 4235 vol.5

[12]. Mahmoud Elmezain, Ayoub Al-Hamadi, Saira Saleem Pathan, Bernd Michaelis (2009) “Spatio-Temporal Feature Extraction-Based Hand Gesture Recognition for Isolated American Sign Language and Arabic Numbers” Proceedings of the 6th International Symposium on Image and Signal Processing and Analysis (2009), 254-259

[13]. S. Haykin, Neural Networks, A Comprehensive Foundation, Macmillan College Publishing Company, Inc.,1994.

[14]. C. -T. Lin and C. S. George Lee, Neural Fuzzy Systems “A Neuro-Fuzzy Synergism to Intelligent Systems”, Prentice-Hall International, Inc., 1996.

[15]. J. -S. R. Jang, C. -T. Sun, and E. Mizutani, “Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Learning”, Prentice-Hall International, Inc., 1997

[16]. Mu-Chun Su, Hi Huang, Chia-Hsien Lin, Chen-Lee Huang, and Chi-Da Lin (1998) “Application of Neural Networks in Spatio-temporal Hand Gesture Recognition” IEEE 1998, 2116-2121.

[17]. J. Davis and M. Shah. “Recognizing hand gestures”. In Proceedings of European Conference on Computer Vision, ECCV: 331-340, 1994.

[18]. N.D.Binh, E.Shuichi and T.Ejima, “Real-time hand tracking and hand recognition system,” ICCI 2006, 5th IEEE International Conference, 2006.

[19]. D. J. Turman and D. Zelter, “Survey of glove-based input”, IEEE Computer Graphics and Application 14:30-39, 1994.

[20]. J.Yang, Y.Xu, and C.S. Chen, “Gesture Interface: Modeling and Learning,” IEEE International Conference on Robotics and Automation, Vol.2, pp. 1747-1752, 1994.

[21]. R. Lockton. A.W. Fitzgibbon, “Real-Time gesture recognition using deterministic boosting,” Preceedings of British Machine Vision Conference, 2002.

[22]. Pham The Bao, Nguyen Thanh Binh, Tu Duy Khoa (2009), “A New Approach To Hand Tracking And Gesture Recognition By A New Feature Type And HMM” Sixth International Conference on Fuzzy Systems and Knowledge Discovery, 2009

[23]. Vitoantonio Bevilacqua, Lucia Cariello, Gaetano Carro, Domenico Daleno, Giuseppe Mastronardi (2008), “A face recognition system based on Pseudo 2D HMM applied to neural network coefficients”, Soft Comput (2008) 12:615–621

[24]. Stokoe, W.C., (1960). Sign language structure: an outline of the visual communication system of the American deaf. Studies Linguistics: Occasional Papers (No. 8). Buffalo: Dept. of Anthropology and Linguistics, University of Buffalo.

[25]. M. B. Waldron and **S.** Kim, “Technology for ASL recognition,” in Proc. Conf Technol. Persons Disabilities, Los Angeles, CA, 1993, pp. 152-158.

[26]. R. P. Meier and E. L. Newport, “Out of the hands of the babes: On a sign advantage in language acquisition,” Language, vol. 66, pp. 1-23, 1990.

[27]. J. D. Schein, “At Home Among Strangers”. Washington, D.C.: Gallaudet Univ. Press, 1989.

[28]. M. B. Waldron, “Identification of nonlinear controls in the developmental model of motor speech in the hearing and the deaf,” Perceptual Motor Skills, vol. 54, pp. 3-10. 1982.

[29]. L. A. Petitto and P. F. Marentette, “Babbling in the manual mode: Evidence for the ontogeny of language,” Sci. Rep., vol. 251, pp. 1493-1496, Mar. 1991.

[30]. R. C. Eberhart and R. W. Dobbins, “Neural Network PC Tools: A Practical Guide”. San Diego, CA Academic, 1990.

[31]. Kramer, J., Leifer, L.J., 1990. “A ‘Talking Glove’ for Nonverbal Deaf Individual”. Technical Report CDR TR 1990 0312, Center for Design Research, Stanford University.

[32]. Kramer, J., 1996. “The Talking-Glove: Hand-Gesture-to-Speech Using an Instrumented Glove and a Tree-Structured Neural Classifying Vector Quantizer”, Ph.D. Thesis, Stanford University.

[33]. Murakami, K., Taguchi, H., 1991. Gesture recognition using recurrent neural networks. In: CHI’91 Conference Proceedings, Human Interface Laboratory, Fujitsu Laboratories, ACM, pp. 237–242.

[34]. Waldron, M.B., Kim, S., 1995. Isolated ASL recognition system for deaf persons. IEEE Transactions on Rehabilitation Engineering 3 (3), 261–271.

[35]. L. A. Zadeh. Theory of approximate reasoning. In J. Hayes, D. Michie, and L. I. Mikulich, editors, Machine Intelligence, pages 149–194. Ellis Horwood, 1970.

[36]. G. Chen and T. T. Pham. Fuzzy Sets, Fuzzy Logic, and Fuzzy Control Systems. CRC Press, Boca Raton, 2001.

[37]. C. Carlsson and R. Fuller. Fuzzy Reasoning in Decision Making and Optimization. Springer, Heidelberg, 2002.

[38]. W. Siler and J. J. Buckley. Fuzzy Expert Systems and Fuzzy Reasoning. John Wiley & Sons, New York, 2004.

[39]. S. Mitra and S. K. Pal. Fuzzy sets in pattern recognition and machine intelligence. Fuzzy Sets and Systems, 156:381–386, 2005.

[40]. R. E. Moore. “Methods and Applications of Interval Analysis” SIAM, Philadelphia, 1979.

[41]. X. Ou, X. Chen and J. Yang. “Gesture recognition for remote collaborative physical tasks using tablet PCs”. In Proc. of the 9th IEEE Intl. Conf. on Computer Vision, Work. on Multimedia Tech. in E-Learning and Collaboration, Nice, 2003.

[42]. S. K. Pal. Fuzzy “Image Processing and Recognition: Uncertainty Handling and Applications”. Intl. Journal Image Graphics, 1:169–195, 2001.

[43]. S. K. Pal and A. Skowron (1999). “Rough-Fuzzy Hybridization: A New Trend in Decision Making”. Springer, Heidelberg, 1999.

[44]. W. Pedrycz. Fuzzy sets in pattern recognition: Accomplishments and challenges. Fuzzy Sets and Systems, 90:171–176, 1997.

[45]. C. T. Lin and C. S. G. Lee. Neural Fuzzy Systems: A neurofuzzy synergism to intelligent systems. Prentice Hall, Upper Saddle River, 1996.

[46]. M. Hanmandlu, M. H. M. Yusof, and V. K. Madasu. “Offline signature verification and forgery detection using fuzzy modelling”. Pattern Recognition, 38(3):341–356, 2005.

[47]. K. Kwak and W. Pedrycz. “Face recognition using a fuzzyfisherface classifier”. Pattern Recognition, 38(10):1717–1732, 2005.

[48]. Benjam´ın R. Callejas Bedregal, Grac¸aliz P. Dimuro, Antˆonio C. Rocha Costa “Interval Fuzzy Rule-Based Hand Gesture Recognition” Universidade Federal do Rio Grande do Norte, Natal, Brazil.