

Recognition and Classification of Sign Language for Spanish

Griselda Saldaña González¹, Jorge Cerezo Sánchez¹, Mario Mauricio Bustillo Díaz², Apolonio Ata Pérez²

¹ Universidad Tecnológica de Puebla,
Ingeniería en Tecnologías para la automatización, Puebla,
Mexico

² Benemérita Universidad Autónoma de Puebla,
Facultad de Ciencias de la Computación, Puebla,
Mexico

{griselda.saldana, jorge.cerezo}@utpuebla.edu.mx, bustillo1956@hotmail.com, apolonio.ata@gmail.com

Abstract. In this paper it is presented a computational system for recognition and classification of letters of the sign language in Spanish, designed for helping deaf-mute people to communicate with other persons. A low-cost glove that captures the hand movements has been constructed. This one contains an accelerometer for each finger which allows detecting its position by using an acquisition data board. Sensor information is sent wirelessly to a computer having a software interface, developed in LabVIEW, in which the symbols dataset is generated. For the automatic recognition of letters we have applied a statistical treatment to the dataset obtaining accuracy greater than 96% independently of the user.

Keywords. Signs language, machine learning, glove.

1 Introduction

Recognition of signs has been the focus of several research areas such as human-computer interaction, virtual reality, tele-manipulation and images processing. Another area of application is sign language interpretation [1]. Among the types of gestures, sign language is one of the most structured; usually each gesture is associated to a predefined meaning. However, the application of strong context rules and grammar makes sign language more difficult to recognize [22]. According to sensing technology used to capture gestures, there are two main approaches for sign recognition. One based on vision techniques [16], in which hand movement is followed and

the corresponding sign is interpreted [23, 18] and another based on gloves [15], with sensors that capture the movement and rotation of hands and fingers [9]. Other methods include Leap Motion [11], or Kinect sensors [17].

Corresponding to the approach based on vision; in [20] a method to convert the Indian Sign Language (ISL), hand gestures into appropriate text message is presented. The hand gestures are captured through a webcam and the corresponding frames are segmented considering features such as number of fingers and the angle between them. Trigueiros et al. [25], used vision based technique for recognition of Portuguese language. For their implementation, hand gesture was captured in real time.

SVM algorithm is used for classification purpose. In this system vowels are recognized with accuracy of 99.4% and consonants are recognized with 99.6% accuracy. In [3], a real-time method for hand gesture recognition is presented. The hand region is extracted from the background, then the palm and fingers are segmented to detect and recognize the fingers. A rule classifier is applied to predict the labels of hand gestures. Computer vision based techniques have the potential to provide more natural and non-contact solutions, and are based on the way human beings perceive information about their surroundings [21].

The main drawback is in the acquisition process due to many environmental apprehensions such as the place of the camera, background condition and

lightning sensitivity [14], in addition accuracy and processing speed are challenging.

Leap Motion controller is a small USB device that using monochromatic IR cameras and infrared LEDs, observes a roughly hemispherical area, to a distance of about 1 meter. The LEDs generate pattern-less IR light and the cameras generate almost 200 frames per second [26]. P. Karthick et al. [10] used a model that transform Indian sign language into text using a leap controller. The Leap device detects the data like point, wave, reach, grab which is generated by a leap motion controller. Combination of DTW and IS algorithm are used for conversion of hand gesture into text. Neural network was used for training the data.

In [6] a leap motion controller is used for recognition of Australian sign language. Leap motion controller senses the hand movement and convert that hand movement into computer commands. Artificial neural network is used for training symbols. The disadvantage of that system was low accuracy and fidelity. With the emergence of RGB-D (color images and depth maps synchronized) and capture devices, using mainly the Microsoft Kinect sensor; the gesture recognition field had a great push forward [12]. In [5] a Microsoft kinect was used to recognize American Sign Language (ASL). Depth camera is Kinect sensor used to detect ASL alphabet. Distance adaptive scheme was used for feature extraction. Support vector machine and RF classifier algorithm used for classification purpose. Training of data was done using neural network.

The accuracy of the system was 90%. In [2] a 3D trajectory description of one sign language word is used and matched it against a gallery of trajectories. Another work presented in [7] used an RGB-D image from the Microsoft Kinect sensor to recognize the letters of the manual alphabet, known as fingerspelling. These works used data from a point cloud and required further processing for hand detection before actually detecting gestures. The Leap Motion skip this step, because already handles the detection by itself.

Recognition based on sensors such as accelerometers and gyroscopes offer the following advantages: a) because movement sensors are not affected by the surrounding, recognition is

more adequate than recognition based in vision in complex surroundings b) they are joined to a user, this allows a bigger coverage, and c) the signs can be acquired wirelessly [13]. Gloves have been successfully used for the recognition of signs in previous works [4, 24], in [1] a system for the recognition of the 23 letters of the Vietnamese language is presented; this system uses a glove with accelerometers MEMS, whose data is transformed to relative angles between the fingers and the hand palm. For the recognition of the letters, it uses a classification system based in fuzzy logic.

In [27] a glove based in accelerometers and myoelectric sensors is reported, its elements allow it to automatically detect the initial and final point of two significative segments of the symbols by the intensity of the myoelectric sensors. To obtain the final result, it uses decision trees and hidden models of Markov. The functionality of the system is shown by the classification of the 72 symbols of Chinese sign language. [8] presents a framework for Sign Language Gesture recognition using an accelerometer glove. The evaluation of the solution presents the results of the gesture recognition attempt by using a selected set of sign language gestures with a described method based on Hidden Markov Model (HMM) and parallel HMM approaches, achieving a 99.75% recognition accuracy.

In this work, the implementation of a training system for the sign language of the Spanish alphabet for deaf-mute people is presented. It consists of a glove-like device with an accelerometer connected to each finger. The outputs of the sensors go through an acquisition board that sends the data wirelessly to a computer where an interface in LabVIEW resides. The collected data are kept in a sign database in which, differently to [9], the information is classified using a statistic method.

Once the signs are discriminated without ambiguity, the system can be used for the training of deaf-mute people who can make each of the Spanish alphabet letters from another interface in LabVIEW, and confirm if they are doing it right. The rest of this document is organized in the following way. In section 2, a description of

the system is presented, making emphasis on the implementation of the glove and the sensors functioning. In section 3, data classification mechanism is presented. In section 4, the tests carried out to the system and some of the results obtained are presented, and finally in section 5 conclusions and future work are presented.

2 System Description

The system consists of three elements, a glove instrumented with analog accelerometers that can send information wirelessly, and two programs in LabVIEW, the first one for the samples capture and the second one for people's training in sign making. The programs count with a graphic interface which are intuitive and allow any user to interact with the system.

2.1 Glove Construction

The glove design is based on the use of accelerometers, in this case ADXL335 since they are low cost and consume little power. These accelerometers give a measurement of the fingers position in three axes with a serial format (x, y, z), the glove accelerometers provide raw data which is sent to the acquisition board in a vector format and it is sent to the central computer through an Xbee device, Figure 1.



Fig. 1. Glove structure

2.2 Samples Capture

The computer has a LabVIEW program used to capture the data corresponding to each of the Spanish alphabet letter and to storage them in a database. In order to do that, a group of 25 deaf-mute people was used to make each letter 50 times. The user interface made for the samples capture is shown in Figure 2.

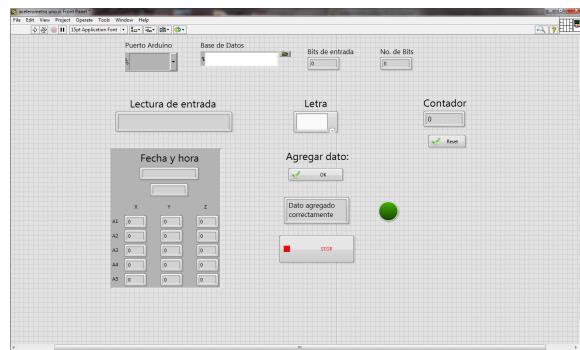


Fig. 2. User interface for the data capture

This data is used offline to make a classification process in which each subclass corresponds to an specific letter. For the online operation, the user to be trained, accesses to another user interface in which he is notified if he is adequately making each letter. The user executes a letter and then the reading of the glove data is made, the information is compared with the information obtained by the training system. Once each X, Y, Z accelerometer reading is recognized, the corresponding letter is shown on the screen, this way, the user can corroborate if he makes it adequately and repeat the process with a new sign. If the user wants to, he can proceed to form a word.

3 Classification

Once the data is captured, we can use it to build a classification model that could be used later to identify a sign and automatically associate it to a determined letter.

The X, Y, Z readings that were obtained from each of the five accelerometers are used as characteristics for the construction of the

Table 1. Samples quantity for each alphabet sign

Letter	Samples	Letter	Samples	Letter	Samples
a	58	i	51	q	50
b	51	j	48	r	65
c	57	k	48	s	66
d	51	l	49	t	56
e	50	m	47	u	53
f	96	n	59	v	51
g	51	o	56	w	51
h	48	p	66	x	50
				y	50

classification model. Particularly, we have experimented with the three following classifiers:

- (a) J48: It is a decision tree type classifier: Algorithm J48 is an implementation of algorithm C4.5, one of the data mining algorithms most used in several applications.
- (b) SMO: It stands for the English words "Sequential minimal optimization", and is an algorithm used to solve the problem of quadratic programming that arises during the training of Support Vector Machines. It was invented in 1998 by John Platt [19] and it is broadly used nowadays.
- (c) The multilayered perceptron is an artificial neuronal network formed by multiple layers; this allows solving problems that are not linearly separable, which is the main limitation of the perceptron (also named simple perceptron).

The results obtained in the experiments are shown in the following section.

4 Tests and Results

In this section, characteristics associated to the training corpus, as well as the evaluation methodology and the results obtained are described.

4.1 Data Set

Table 1, shows the number of samples that have been taken for each of the signs considered in the training corpus. The minimum number of samples was of 47 for letter 'm', and the maximum number of samples was of 96 for letter 'f'. The median of the samples was of 55.12. In total, 1,378 samples were taken.

4.2 Evaluation Methodology

The evaluation process considers the use of the training corpus to validate the exactitude of the letter identification by using the three models of automatic classification suggested before.

Every set of samples of each letter is divided in 10 partitions and ten interactions are executed using 90% of the data for training and the remaining 10%, for tests in a process named 10-fold cross validation and leave-one out. The results obtained in the three classifiers, as well as the discussion of those results, are presented in the following section.

4.3 Obtained Results

Table 2 shows the results obtained for each of the classifiers. As it can be observed, the classifier based in multilayered perceptron is the one that obtains the best results, with an exactitude superior to 97%. Out of the 1,378 classified samples, only a total of 36 instances are classified incorrectly, this resulted in an error of 2.61%.

Table 2. Comparison of results obtained among the three classifiers

	J48		SMO		Multilayered Perceptron	
Classified instances	Quantity	Percentage	Quantity	Percentage	Quantity	Percentage
Correctly	1,227	89.04%	1,276	92.60%	1,342	97.39%
Incorrectly	151	10.96%	102	7.40%	36	2.61%

In fact, it surpasses in 5 percent points the SMO classifier and in 8 percent points the J48 classifier. These results show that the exactitude degree is high and sufficient for the classification process of letter based on sign language.

It is necessary however to perform an analysis of the execution times necessary for each algorithm to construct the classification model in order to verify its use pertinence in real time systems. Table 3 shows the results.

Table 3. Comparison of the construction time of the classification model

	J48	SMO	Multilayered Perceptron
Time (seconds)	0.19	2.45	17.86

As it can be observed, the exactitude level is inversely proportional to the necessary execution times to construct the model. Actually, the nearly 18 necessary seconds used by the classifier based in the multilayered perceptron does not turn out to be prohibitive for the construction of a classification model. As a matter of fact, the evaluation time of the test instances are of thousands of seconds for any of the three classifiers tested.

5 Conclusions and Future Work

This work presents a glove based in accelerometers that allow the training of deaf-mute people for the making of Spanish alphabet letters. The data management has been carried out statistically; this gives precision to the process of classification, makes the system independent from the user and allows the sign detection even if signs are not correctly made. Experiments carried out with three methods of automatic classification show that the precision obtained in the sign identification is

higher than 89%. It is the algorithm based in neuronal networks using a multilayered perceptron the one that particularly obtained the best result, with exactitude of 97%.

Acknowledgements

The authors thank PRODEP for the support for the execution of this work.

References

1. Bui, T. D., Nguyen, & Thuy, X. (2011). Recognizing postures in vietnamese sign language with mems accelerometers. *IEEE Sensors Journal*, volume 7, pp. 707–712.
2. Chai, X., Li, G., Lin, Y., Xu, Z., Tang, Y., Chen, X., & Zhou, M. (2013). Sign language recognition and translation with kinect. *International Conference on Automatic Face and Gesture Recognition*, IEEE.
3. Chen, Z., Kim, J., Liang, J., Zhang, J., & Yuan, Y. (2014). Real-time hand gesture recognition using finger segmentation. *The Scientific World Journal*, pp. 9.
4. Dipietro, L., Sabatini, A. M., & Dario, P. (2008). A survey of glove-based systems and their applications. *IEEE Trans. Systems, Man, and Cybernetics, Part C*, Vol. 38, No. 4, pp. 461–482.
5. Dong, C., Leu, M. C., & Yin, Z. (2015). American sign language alphabet recognition using microsoft kinect. *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, IEEE, Boston, MA, USA.
6. Ellen-Potter, L., Araullo, J., & Carter, L. (2013). The leap motion controller: A view on sign language. *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration, OzCHI '13*, ACM, New York, NY, USA, pp. 175–178.

7. **Estrela, B., Cámera-Chávez, G., Campos, M. F., Schwartz, W. R., & Nascimento, E. R. (2013).** Sign Language Recognition using Partial Least Squares and RGB-D Information. *Workshop de Visão Computacional (WVC)*, Rio de Janeiro.
8. **Galka, J., Masior, M., Zaborski, M., & Barczewska, K. (2016).** Inertial motion sensing glove for sign language gesture acquisition and recognition. *IEEE Sensors Journal*, Vol. 16, No. 16, pp. 6310–6316.
9. **Hernandez-Rebollar, J. L., Lindeman, R. W., & Kyriakopoulos, N. (2002).** A multi-class pattern recognition system for practical finger spelling translation. *Proceedings of the 4th IEEE International Conference on Multimodal Interfaces, ICMI '02*, IEEE Computer Society, Washington, DC, USA, pp. 185–190.
10. **Karthick, P., Prathiba, N., Rekha, V. B., & Thana-laxmi, S. (2014).** Transforming indian sign language into text using leap motion. *International Journal of Innovative Research in Science, Engineering and Technology*, Vol. 3, No. 4, pp. 5.
11. **Khan, F. R., Ong, H. F., & Bahar, N. (2016).** A sign language to text converter using leap motion. *International Journal on Advanced Science, Engineering and Information Technology*, Vol. 6, No. 6, pp. 1089–1095.
12. **Khelil, B. & Amiri, H. (2016).** Hand gesture recognition using leap motion controller for recognition of arabic sign language. *Proceedings of Engineering & Technology (PET)*, 3rd International Conference on Automation, Control, Engineering and Computer Science (ACECS'16), pp. 233–238.
13. **Kim, J.-H., Thang, N. D., & Kim, T.-S. (2009).** 3-d hand motion tracking and gesture recognition using a data glove. *IEEE International Symposium on Industrial Electronics (ISIE)*, IEEE, Seoul, Korea.
14. **Lokhande, P., Prajapati, R., & Pansare, S. (2015).** Data gloves for sign language recognition system. *International Journal of Computer Applications; National Conference on Emerging Trends in Advanced Communication Technologies (NCETACT-2015)*, pp. 11–14.
15. **Mäntyärví, J., Kela, J., Koripää, P., & Kallio, S. (2004).** Enabling fast and effortless customization in accelerometer based gesture interaction. *Proceedings of the 3rd International Conference on Mobile Ubiquitous Multimedia, MUM'04*, ACM, pp. 25–31.
16. **Mitra, S. & Acharya, T. (2007).** Gesture recognition: A survey. *Transactions on Systems, Man, and Cybernetics*, Vol. 37, No. 3, pp. 311–324.
17. **Oszust, M. & Wysocki, M. (2013).** Polish sign language words recognition with kinect. *The 6th International Conference on Human System Interaction (HSI)*, IEEE, Sopot, Poland, pp. 219–226.
18. **Park, J. W., Hyun, S. D., & Lee, C. W. (2008).** Real-time finger gesture recognition. *HCI*, Korea, pp. 847–850.
19. **Platt, J. C. (1999).** Fast training of support vector machines using sequential minimal optimization. In Schölkopf, B., Burges, C. J. C., & Smola, A. J., editors, *Advances in Kernel Methods*. MIT Press, Cambridge, MA, USA, pp. 185–208.
20. **Shangeetha, R. K., Valliammai, V., & Padmavathi, S. (2012).** Computer vision based approach for indian sign language character recognition. *International Conference on Machine Vision and Image Processing (MVIP)*, IEEE, pp. 181–184.
21. **Simion, G., Gui, V., & Oteșteanu, M. (2011).** A brief review of vision based hand gesture recognition. *Proceedings of the 10th WSEAS International Conference on Circuits, Systems, Electronics, Control, Signal Processing, and Proceedings of the 7th WSEAS International Conference on Applied and Theoretical Mechanics, CSECS/MECHANICS'11*, World Scientific and Engineering Academy and Society (WSEAS), Stevens Point, Wisconsin, USA, pp. 181–188.
22. **Starner, T. & Pentland, A. (1997).** Real-time american sign language recognition from video using hidden markov models. In **Shah, M. & Jain, R.**, editors, *Motion-Based Recognition*, volume 9 of *Computational Imaging and Vision*. Springer Netherlands, Dordrecht, pp. 227–243.
23. **Tan, U.-X., Veluvolu, K. C., Latt, W. T., Shee, C. Y., Riviere, C., & Ang, W.-T. (2008).** Estimating displacement of periodic motion with inertial sensors. *IEEE Sensors Journal*, Vol. 8, No. 8, pp. 1385–1388.
24. **Tian-Swee, T., Salleh, S. H., Ariff, A. K., Ting, C. M., Kean-Seng, S., & Seng-Huat, L. (2007).** Malay sign language gesture recognition system. *International Conference on Intelligent and Advanced Systems, ICIAS*, IEEE, Kuala Lumpur, Malaysia.
25. **Trigueiros, P., Ribeiro, F., & Reis, L. P. (2014).** Vision-based portuguese sign language recognition system. In **Rocha, Á., Correia, A. M., Tan, F. B., & Stroetmann, K. A.**, editors, *New Perspectives in Information Systems and Technologies*, volume 1. Springer International Publishing, pp. 605–617.

26. **Weichert, F., Bachmann, D., Rudak, B., & Fisseler, D. (2013).** Analysis of the accuracy and robustness of the leap motion controller. *Sensors*, Vol. 13, No. 5, pp. 6380–6393.

27. **Zhang, X., Chen, X., Li, Y., Lantz, V., Wang, K., & Yang, J. (2011).** A framework for hand gesture recognition based on accelerometer and emg sensors. *Transactions on Systems, Man, And Cybernetics. Part A*, Vol. 41, No. 6, pp. 1064–1076.

Article received on 29/08/2016; accepted on 09/10/2016.
Corresponding author is Griselda Saldaña González.