Mexican sign language recognition using normalized moments and artificial neural networks

J-Francisco Solís-V*[a], Carina Toxqui-Quitl[b], David Martínez-Martínez[a], Margarita H-G[a]
[a]Universidad Autónoma del Estado de México, Inst. Literario #100, Toluca, Estado de México, México;
[b]Universidad Politécnica de Tulancingo, Ingenierías #100, Tulancingo, Hidalgo, México

ABSTRACT

This work presents a framework designed for the Mexican Sign Language (MSL) recognition. A data set was recorded with 24 static signs from the MSL using 5 different versions, this MSL dataset was captured using a digital camera in incoherent light conditions. Digital Image Processing was used to segment hand gestures, a uniform background was selected to avoid using gloved hands or some special markers. Feature extraction was performed by calculating normalized geometric moments of gray scaled signs, then an Artificial Neural Network performs the recognition using a 10-fold cross validation tested in weka, the best result achieved 95.83% of recognition rate.

Keywords: Mexican Sign Language Recognition, Normalized Moments, Multi-Layer Perceptron, Computer Vision System.

1. INTRODUCTION

Interest in Sign Language Recognition (SLR) has been increased in last decade[1-5]. Expressing signs was the main way to communicate before human develop oral communication, nevertheless, actually for deaf people community Sign Language (SL) has become the preferred communication way. Sign Language is not universal, many countries or regions have their own language, like American Sign Language (ASL), Mexican Sign Language (MSL), Persian Sign Language (PSL), Chinese Sign Language (CSL) among others.

The need to develop a SLR comes from communication limitations that deaf community faces every day when they have to communicate with not unhearing people, that's why they need a translator to help them in some situations such as institutional transactions or a phone calls, in other words, most people doesn't know the SL, then unhearing people lose privacy because often closest family or friends are the translators.

Communicate using SL leads to expression through the hand shape, movement of arms, head, body and facial expressions, in fact one of the most challenging problem in SLR is getting accurate data to represent signs, that's why advances in SLR has limits, most works in SLR introduce limited results in controlled situations.

Advances in SLR can be classified as computer vision based or electronic systems. Developments that allows a more natural way to communicate using a computer is the artificial vision, because signer is not obligated to wear electronic devices physically connected to the body, nevertheless is more complex than electronic systems. In the other hand, electronic systems sense accurate data about position, orientation and movement of hand shape, despite of obligate to signer to be physically connected to the system.

This work presents a Computer Vision System for the MSL recognition, using a solid white color in background to avoid the use of color gloves or special markers for segmentation process, the work consists in recognize the MSL from a static signs database.

2. MEXICAN SIGN LANGUAGE

2.1 Introduction

Mexican Sign Language (MSL) is the SL mainly use in Mexico's country, most signs are expressed in static way, but some others like 'j' or 'z' are represented by a special hand movement, for this reason this work only use static signs to develop the MSL database, then 26 signs form the complete MSL but 24 static signs were chosen to create de database,

Optics and Photonics for Information Processing VIII, edited by Abdul A. S. Awwal, Khan M. Iftekharuddin, Mohammad A. Matin, Andrés Márquez, Proc. of SPIE Vol. 9216, 92161A © 2014 SPIE
CCC code: 0277-786X/14/$18 · doi: 10.1117/12.2061077

in Figure 1 can be seen the 24 mexican static signs, it can be appreciated that signs like ‘g’ and ‘h’ are similar, same case for ‘m’ and ‘s’ or ‘s’ and ‘t’.

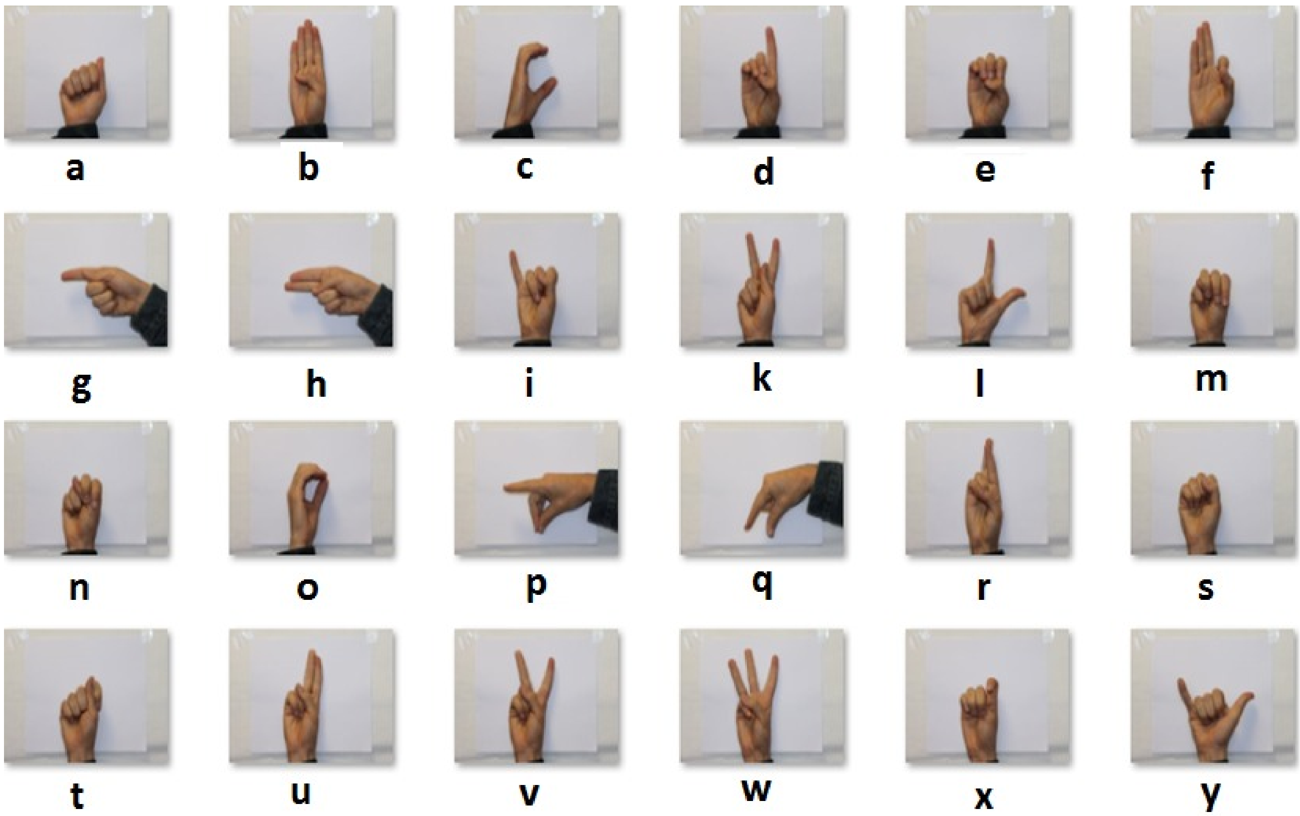

Figure 1. Twenty four static signs of MSL used in this work for SLR.

2.2 Database

As can be seen in Figure 1, database design consists in the use of 24 static signs from MSL, captured using a solid white background color in order to avoid the use of color gloves or special color markers, five different versions of each sign were captured in incoherent light conditions using the flash of a digital camera (CMOS-Cannon EOS REBEL T3), all signs were expressed by one male signer.

Database was specially designed for the recognition of MSL avoiding the use of special markers in the signer body, as well as avoiding to attach some electronic devices to the hands. Keeping in mind that these considerations allows to signer express the signs in a more natural way.

3. MEXICAN SIGN LANGUAGE RECOGNITION

3.1 Feature extraction

Descriptors were computed from normalized moments based on geometric moments[6,7], as the name suggests geometric moments represents statistical information or geometric measures, geometric moments are given by

$$m_{p,q} = \iint x^p y^q f(x,y)dxdy \tag{1}$$

where ‘p’ denotes the order, ‘q’ is commonly called repetition, ‘f(x,y)’ is the image function and ‘m’ represents the moment of the image, for low values of order ‘p’ and repetition ‘q’, geometric moments have intuitive representation, for example geometric moment $m_{0,0}$ calculates the area of a binary object.

Central moments are given by the next expression

$$\mu_{p,q} = \iint (x - x_c)^p (y - y_c)^q f(x,y) dxdy \quad (2)$$

where $x_c = m_{1,0} / \mathrm{m}_{0,0}$ represents the centroid in 'x' and $y_c = m_{0,1} / \mathrm{m}_{0,0}$ the centroid in 'y', so central moments are invariant to translation, central moments are especially useful to compute normalized moments, this relation can be represented as

$$v_{p,q} = \frac{\mu_{pq}}{\mu_{00}^{\omega}} \quad (3)$$

where

$$\omega = \frac{p+q}{2} + 1 \quad (4)$$

Normalized moments v are scale invariant which is helpful for the MSL recognition due to gives robustness to the system.

3.2 Framework

Figure 2 shows the MSL recognition process, first step is to crop the original image, in order to reduce the computational cost, next step consists in isolate the sign or hand shape form the rest of scene (background), and using the blue channel data from RGB space color gives best results, then feature extraction begins, for most tests in this work 64 normalized moments were computed as combinations of $p = 0,1,\ldots,7$ and $q = 0,1,\ldots,7$, once moments are calculated, then a Multilayer Perceptron classify the patterns using the GNU software WEKA[8], with k-fold cross validation where k=10.

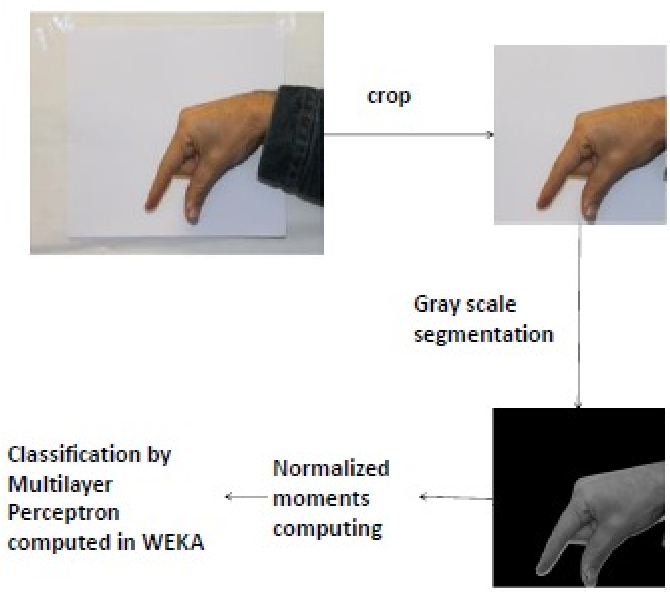

Figure 2. Framework of MSL recognition system.

4. RESULTS

Representing MSL signs represented as binary objects (see Figure 3-a), achieves 93.33% of recognition rate using 64 normalized moments, some transformations were tested such as a x-axe resized object (Figure 3-b), y-axe resized sign (Figure 3-c) or x and y resized sign (Figure 3-d) giving 87.5%, 87.5% and 85.83% of performance.

Best results were achieved by using gray scale signs (Figure 3-e), looking to reduce the features amount, some tests were performed, for $p = 1, 2, ..., 4$ and $q = 0, 1, ..., 7$ general performance achieved 90% (using 32 descriptors) and finally for $p = 0, 1, ..., 7$ and $q = 2, 3, ..., 6$ achieves 95.83% of recognition rate using 40 descriptors.

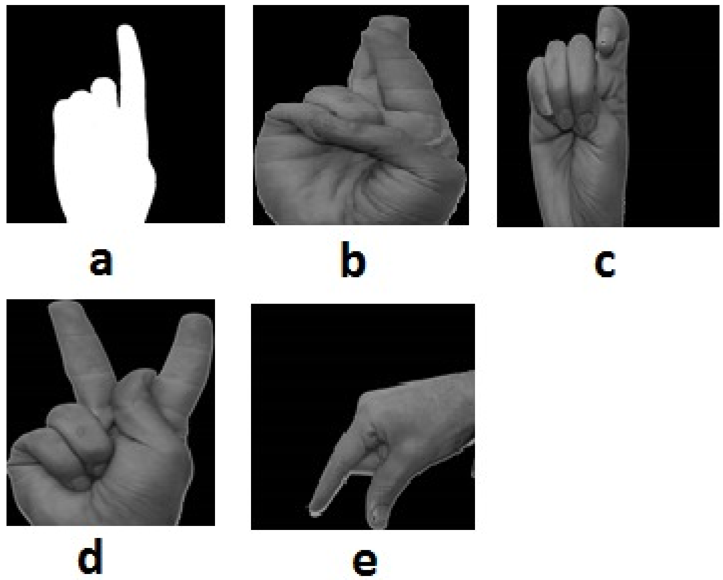

Figure 3. a) Binary sign, b) x-axe resized sign, c) y-axe resized sign, d) x and y axes resized and e) gray scale sign.

5. CONCLUSIONS AND ACKNOWLEDGMENTS

Binary signs perform 90% of recognition rate, the resize transforms did not improve the performance and instead drops about 5%, best results were achieved using 40 features of gray scale signs achieving 95.83% of recognition rate.

Authors would like to thank to ‘Secretaría de Investigación y Estudios Avanzados’ of the ‘Universidad Autónoma del Estado de México’ for the support to the project 3482/2013CHT, authors would like to thank to ‘Universidad Politécnica de Tulancingo’ as well.

REFERENCES

[1] Samir,A., Abull, M. and Tolba, M., “A proposed PCNN features quality optimization technique for pose-invariant 3D Arabic sign language recognition” Applied Soft Computing 13, 1646-1660 (2013).
[2] Fenlon, J., Schembri, A., Rentelis, R. and Cormier, K., “Variation in hand shape and orientation in British Sign Language: The case of the ‘1’ hand configuration” Language & Communication 33, 69-91 (2013).
[3] Karami, A., Zanj, B. and Kiani, A., “Persian sign language (PSL) recognition using wavelet transform and neural networks” Expert Systems with Applications 38, 2661-2667 (2011).
[4] Subha, P. and Balakrishnan, G., “Recognition of Tamil Sign Language Alphabet using Image Processing to aid Deaf-Dumb People” Procedia Engineering 30, 861-868 (2012).
[5] Rozado, D., Rodriguez, F. and Varona, P., “Extending the bioinspired hierarchical temporal memory paradigm for sign language recognition” Neurocomputing 79, 75-86 (2012).
[6] Flusser, J., Suk, T. and Zitová, B., “Moments and Moment Invariant in Pattern Recognition”, Ed. Willey, first ed, (2009).
[7] Hu, M. “Visual pattern recognition by moment invariants”, IRE Trans. Inform. Theory IT-8, 179-187 (1962).

[8] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutmann, P. and Witten, I., "The WEKA Data Mining Software: An Update; SIGKDD Explorations" 2009 < http://www.cs.waikato.ac.nz/ml/weka/>. www.cs.waikato.ac.nz/ml/weka/.