Contemporary Engineering Sciences, Vol. 8, 2015, no. 32, 1499 - 1508
HIKARI Ltd, www.m-hikari.com
http://dx.doi.org/10.12988/ces.2015.59278

An Automated Method for Identification of Vowels on the Sign Language

J.C. Martínez-Perales

Instituto Politécnico Nacional
Escuela Superior de Cómputo
Av. Juan de Dios Batíz, esq. Miguel Othón de Mendizábal
Ciudad de México 07738, México

R. Flores-Carapia

Instituto Politécnico Nacional
Centro de Innovación y Desarrollo Tecnológico en Cómputo
Av. Juan de Dios Batíz, esq. Miguel Othón de Mendizábal
Ciudad de México 07738, México

B. Luna-Benoso

Instituto Politécnico Nacional
Centro de Innovación y Desarrollo Tecnológico en Cómputo
Av. Juan de Dios Batíz, esq. Miguel Othón de Mendizábal
Ciudad de México 07738, México

Abstract

This paper presents a methodology that allows the identification of the vocal in sign language. For this purpose, the methodology is divided into three modules: 1) image segmentation: by methods in the spatial domain, 2) feature extraction: using moments of Hu, and 3) classification: through an associative memory model based on cellular automata.

Keywords: Digital image processing, methods in the spatial domain, segmentation images, and sign language

1. Introduction

About 70 million deaf people in the world use sign language as their native or main language [14]. The difficulties to which the deaf-mute people face in daily life are quite difficult due to the total lack of understanding or meaning of the sign language they use. Sign language is a means of communication for people with hearing defciency, where words and sentences are represented by hand gestures, and have grammatical structures perfectly defined. Due to the difficult communication between a deaf person handling the sign language with people who don't know the language, some doctors have assisted with people who know the language to handle the translation process with the goal to study in depth the deaf culture with psychotherapeutic purposes [9]. A system capable of performing automatic recognition of sign language without the need of an individual translator can provide an excellent tool for deaf people to communicate with people who don't know the language, and thus achieve a better way life [6]. This paper proposes a methodology to carry out the identification of vowels in Mexican Sign Language (Figure 1).

Figure 1: Representation of the vowels in the Mexican sign language LSM.

2. Proposed Method

Before entering into detail with the proposed model, it is necessary to mention the definition of what a digital image is. A digital image is a two-dimensional function $f(x, y)$ of the light intensity (brightness) in a point in space, where (x, y) coordinates of that point [5]. Since a digital image is a function $f(x, y)$ discretized in both spatial coordinates and in brightness, often usually it represented as a two dimensional matrix $F_{ij} = (f_{ij})_{H \times W}$ where H and W represent the size of the image and $f_{ij} = f(x_i, x_j)$ (Figure 2).

The method described, is divided into three modules: 1) segmentation of the signal carried by the individual, 2) feature extraction and 3) classification. Then each of the modules are shown.

1) Segmentation of the sign. For this stage are considered as follows:

Step 1. The image is decomposed into three Red, Green and Blue (RGB) and the red component was considered because human skin is clearer in this component (figure 3).

Step 2. The image is considered grayscale from red plane, that is, the red values component is copied in the green and blue component, thereby obtaining the grayscale image.

Step 3. The image histogram was obtained with gray levels in the range $[0, 255]$, where the histogram is a discrete function $H[k]$ that represents the number of colors in each gray level ($k = 0, \ldots, 255$) [5].

Step 4. Given an image $f(x, y)$ and two variables u and v, binarization by thresholding defined is used as follows:

$$bin_{ij} = \begin{cases} 0 & if \quad u \leq f_{ij} \leq v \\ 255 & if \quad f_{ij} < u \quad o \quad f_{ij} > v \end{cases}$$

It is observed in Figure 4, the histogram shows two peaks at 0 and 255, to remove these peaks were considered as thresholds $u = 20$ and $v = 185$.

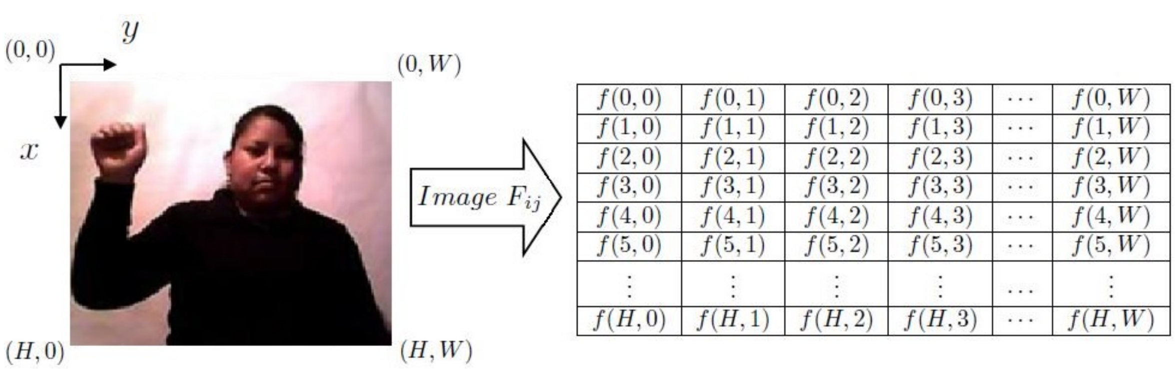

Figure 2: Definition of an image

Once the binarized image was obtained, we proceeded to locate the part corresponding to the hand, therefore, it is considering the following:

Step 5. Consider a mask size x into another mask size y as shown in Figure 5. Consider two variables u and v, to locate the area corresponding to the hand image, is applied the following function: For each point (i, j) on the image, whether M_y^{ij} mask size and center at (i, j), the function is defined T_{ij} function is defined at each point (i, j) as

$$T_{ij} : M_y^{ij} \to [0, 255]$$

where

$$T_{ij}(r, s) = \{ \begin{array}{ccc} f_{rs} & if & A_x \geq u \ y \ A_y - A_x \leq v \\ 255 & if & A_x < u \ o \ A_y - A_x > v \end{array}$$

with A_x and A_y enclosed areas by the size masks x and y respectively.

Step 6. Once located the section of the hand corresponding to the binary image, the operator was applied morphological erosion. If A and B are subsets of $\mathbb{Z}^2$, then *erosion* of A by B, denoted by $A \ominus B$ is the *Minkowski subtraction* of A and B; ie: The morphological erosion is defined as [8, 11, 12]:

$$A \ominus B = \{\mathbf{x} \in \mathbb{Z}^2 | \mathbf{x} + \mathbf{b} \in A \ for \ each \ \mathbf{b} \in B\}$$

Step 7. morphological dilation operator applies. If A and B are subsets of $\mathbb{Z}^2$, then dilation of A by B, denoted by $A \oplus B$, is the *Minkowski sum* of A and B; this is [12, 4, 1, 13]:

$$A \oplus B = \{\mathbf{a} + \mathbf{b} | \mathbf{a} \in A \ y \ \mathbf{b} \in B\}$$

Figure 6 show the process of applying image erosion segmented hand of a neighborhood Moore, and subsequently applying dilation 6 times the same neighborhood. Finally, shows the image of the segmented hand color.

2) Feature extraction: In this module the characteristics were extracted to make the image segmented 1), to thereby Hu moments invariants used. Hu invariant moments or just moments of Hu are numerical properties obtainable from a given image from the pixels that compose it. Hu moments are invariant with respect to translations, rotations and escalations. To define Hu moments, consider the following definitions pre [5, 7].

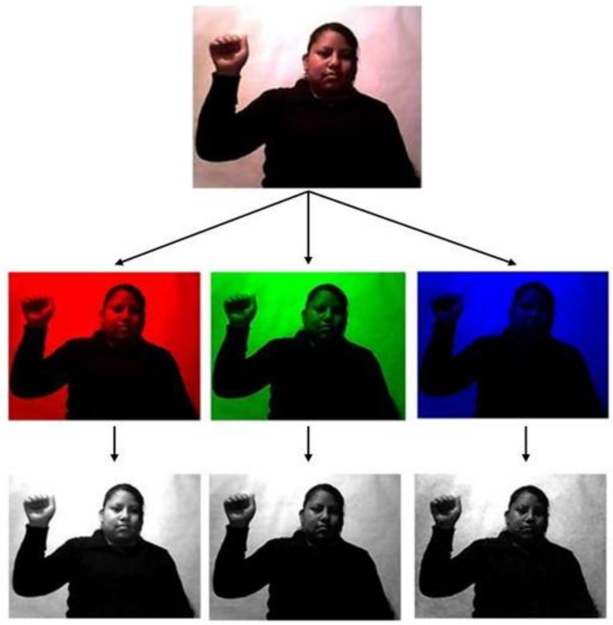

Figure 3: Decomposition of an image in its three components RGB and grayscale

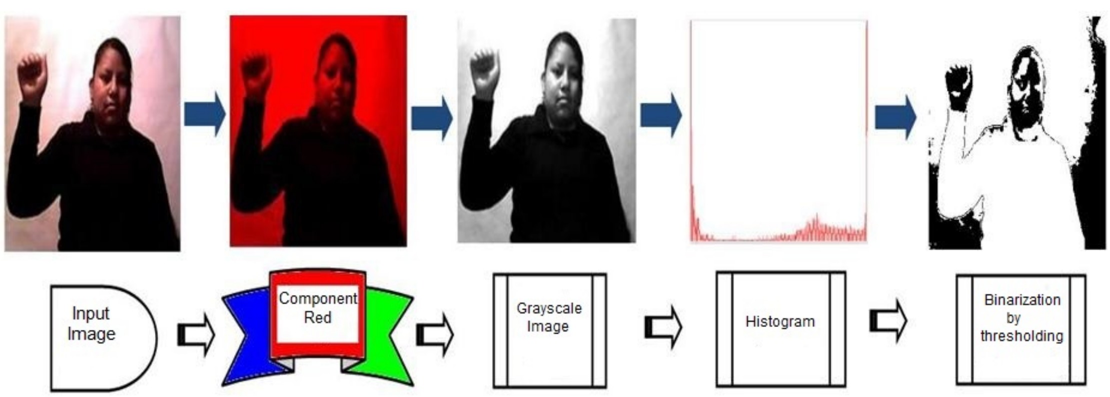

Figure 4: Displays the first 4 steps applied to a color image

Statistical moments are defined as $m_{rs} = \sum_{i,j \in REG} i^r j^s$ where $r, s \in \mathbb{N}$ REG assembly pixels within the region.

The center of gravity of a region defined by the coordinates $(\bar{i}, \bar{j})$ where $\bar{i} = \frac{m_{10}}{m_{00}}$ and $\bar{j} = \frac{m_{01}}{m_{00}}$. From the center of gravity, the central moments defined as $\mu_{rs} = \sum_{i,j \in REG} (i - \bar{i})^r (j - \bar{j})^s$.

From the previous concepts, 7 Hu moments are defined as follows:

$\Phi_1 = \eta_{20} + \eta_{02}$
$\Phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$
$\Phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$
$\Phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$
$\Phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$
$\Phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$
$\Phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$

where

$$\eta_{rs} = \frac{\mu_{rs}}{\mu_{00}^t} \text{ y } t = \frac{r+s}{2} + 1.$$

3) Classification. Associative memories are mathematical models which objective is retrieve complete patterns from input patterns. The operation of the associative memories is divided into two phases: learning stage where the memory is generated; and recovery phase stage where the associative memory operating [10, 2]. Then the learning phase and recovery shown for the partnership model based on cellular automata [3].

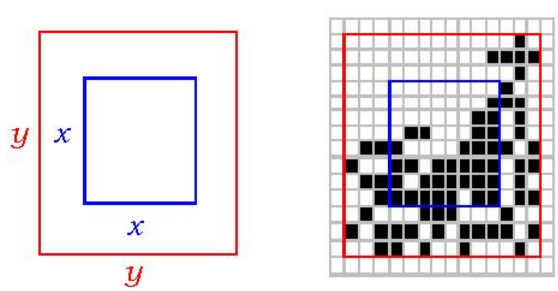

Figure 5: Masks of size x and y

Consider the CA $\mathcal{Q} = (\mathcal{L}, \mathcal{S}, \mathcal{N}, f_{\mathcal{Q}})$ and $\mathcal{W} = (\mathcal{L}, \mathcal{S}, \mathcal{N}', f_{\mathcal{W}})$ with $\mathcal{N}' = IJ$, and $f_{\mathcal{Q}} : \mathcal{N} \to S$, $f_{\mathcal{W}} : \mathcal{N}' \to S$ defined as follows:

$$f_{\mathcal{Q}}(v^{(i,j)}) = \begin{cases} 1 & \text{if} \quad (i,j) \in \mathcal{L}_{FS} \\ \\ 0 & \text{if} \quad (i,j) \notin \mathcal{L}_{FS} \end{cases}$$

$$f_{\mathcal{W}}(v_{(i,j)}) = \begin{cases} 1 & \text{in position } (i+1, j) \quad \text{if } (i, j-1) = 1 \\ \\ 1 & \text{in position } (i, j-1) \quad \text{if } (i+1, j) = 1 \end{cases}$$

We define the Associative CA (ACA) in its learning phase as

$$\mathcal{W} * \mathcal{Q} = (\mathcal{L}, \mathcal{S}, \mathcal{N}, f_A) \tag{1}$$

The recovery phase for the ACA makes use of the composition of erosions and dilations CA. The algorithm which defines the phase of recovery is shown in algorithm 1.

3. Conclusions

This paper presents a methodology that allows the identification of vowels in sign language. For this methodology is divided into three modules: image segmentation, extraction characteristics and classification by means of cellular automata.

Acknowledgements. The authors would like to thank the Instituto Politécnico Nacional (Secretaría Académica, EDI, EDD, COFAA, SIP, ESCOM and CIDETEC), the CONACyT and SNI for their economical support to develop this work.

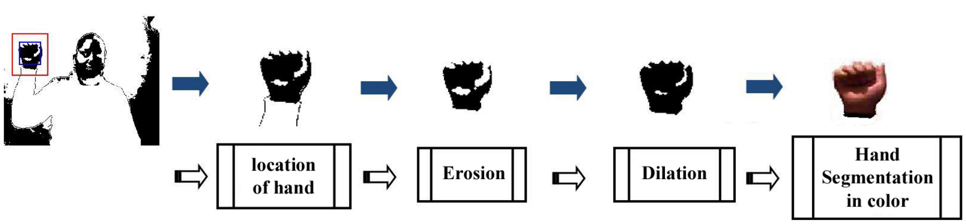

Figure 6: Process segmentation hand to locate the sign made

Algorithm 1 ACA in recovery phase

Require: Fundamental set $FS = \{(\mathbf{x}^\mu, \mathbf{y}^\mu) | \mu = 1, 2, \ldots, p\}$; structuring element B; integer value ne (number of erosions); integer value nd (number of dilations); pattern to recovery $\tilde{\mathbf{x}} \in A^n$

Ensure: Recovery pattern $\tilde{\mathbf{y}} \in A^m$

1. Building the Learning ACA for FS.
2. Applying ne times the cell erosion $\mathcal{E}$ with the structuring element B to the initial configuration of learning ACA. This is, applied to the configuration of the ACA, $\mathcal{E} * \mathcal{E} * \cdots * \mathcal{E}$, ne times.
3. Applying nd times the cellular dilation with the structuring element $\mathcal{D}$ to configuration obtained in point 2. This is, applied to the configuration obtained in point 2, $\mathcal{D} * \mathcal{D} * \cdots * \mathcal{D}$, nd times.
4. For the input pattern $\tilde{\mathbf{x}} \in An$ will get the output pattern $\tilde{\mathbf{y}} \in A^m$ applying:

for $i = 1 \to m$ **do**
    $\tilde{y}_i = 1$
    for $j = 1 \to n$ **do**
        if $\neg(\tilde{x}_j = 0 \wedge (2j-1, 2i-2) = 1)$ **then**
            if $\neg(\tilde{x}_j = 1 \wedge ((2j-2, 2i-2) = 1 \vee (2j-1, 2i-2) = 1))$ **then**
                $\tilde{y}_i = 0$
                $Break$
            end if
        end if
    end for
end for

References

[1] E. Aptoula, S. Lefèvre, A comparative study on multivariate mathematical morphology, *Pattern Recognition*, **40** (2007), 2914-2929. http://dx.doi.org/10.1016/j.patcog.2007.02.004

[2] Roberto A. Vázquez, Humberto Sossa, Behavior of morphological associative memories with true-color image patterns, *Neurocomputing*, **73** (2009), no. 1-3, 225-244.
http://dx.doi.org/10.1016/j.neucom.2009.09.004

[3] B. Luna-Benoso, R. Flores-Carapia, C. Yañez-Márquez, Associative Memories Based on Cellular Automata: An Application to Pattern Recognition, *Applied Mathematical Science*, **7** (2013), no. 18, 857-866.

[4] E. R. Dougherty, R. A. Lotufo, *Hands-om Morphological Image Processing*, Vol. TT59, Spie press, 2003.
http://dx.doi.org/10.1117/3.501104

[5] R. Gonzalez, J. Woods, *Digital Image Processing*, 3rd ed., Prentice Hall, New Jersey, 2008.

[6] M. Krishnaveni, V. Radha, Classifier fusion base don Bayes aggregation method for Indian sign language datasets, *Procedia Engineering*, **30** (2012), 1110-1118.
http://dx.doi.org/10.1016/j.proeng.2012.01.970

[7] José Antonio Martín H., Matilde Santos, Javier de Lope, Orthogonal variant moments features in image analysis, *Information Sciences*, **180** (2010), no. 6, 846-860.
http://dx.doi.org/10.1016/j.ins.2009.08.032

[8] G. Matheron, *Random Sets and Integral Geometry*, Wiley, New York, 1975.

[9] A. Porter, Sign-language interpretation in psychotherapy with deaf patients, *American Journal of Psychotherapy*, **53** (1999), no. 2, 163-176.

[10] Israel Román-Godínez, Itzamá López-Yáñez, Cornelio Yáñez-Márquez, Classifying Patterns in Bioinformatics Databases by Using Alpha-Beta Associative Memories, *Biomedical Data and Applications*, **224** (2009), 187-210.
http://dx.doi.org/10.1007/978-3-642-02193-0_8

[11] J. Serra, *Image Analysis and Mathematical Morphology*, Vol. I, Academic Press, London, 1982.

[12] J. Serra, *Image Analysis and Mathematical Morphology*, Vol. 11 Theoretical Advances, Academic Press, London, 1988.

[13] J.H. Sossa-Azuela, C. Yáñez-Márquez, J. L. Díaz de León S, Computing geometric moments using morphological erosion, *Pattern Recognition*, **34** (2001), no. 2, 271-276.
http://dx.doi.org/10.1016/s0031-3203(99)00213-7

[14] World Federation of the Deaf. www.wfdeaf.org

Received: October 12, 2015; Published: November 21, 2015