Recognition of Shorthand Writing using Neural Networks

[1]Diana M. Vázquez, [1]Karla L. Segovia and [2]Roberto A. Vázquez
[1]Unidad Profesional Interdisciplinaria de Ingeniería y Tecnología Avanzadas –IPN
Departamento de Biónica
Av. Instituto Politécnico Nacional No. 2580
Ciudad de México, 07340, México.
Contact: dvazqueze@ipn.mx, ksegovia@ipn.mx

[2] Centro de Investigación en Computación – IPN
Av. Juan de Dios Batíz, esquina con Miguel Otón de Mendizábal
Ciudad de México, 07738, México.
Contact: ravem@ipn.mx

Abstract. Shorthand is any system of rapid handwriting which can be used to transcribe the spoken word as fast as people speak. In this paper is described a simple but efficient method for the recognition of shorthand signs. This method is based on a new technique for obtaining invariant descriptions from shorthand signs. The technique proposed is invariant to scaling, translations and deformations. In this technique, signs are divided into sub-regions, and those sub-regions are then described in terms of their geometric moments of order zero. Finally, the invariant description already obtained is normalized in order to train a neural network. The proposal is tested using a bank of 1100 signs written by 10 different people.

1 Introduction

Character recognition [6, 7] is an interesting task for many researches. Nowadays, OCR (Optical Character Recognition) systems recognize without difficulty printed characters, but handwriting is still a difficult task due to different ways of writing the same character. Due to OCR can not recognize handwriting, new tools knowing as ICR (Intelligent Character Recognition) were developed. An ICR [8] can recognize handwriting under affine transformations as rotation, translation, deformations and scale change, and also can recognize printed characters.

Shorthand is any system of rapid handwriting which can be used to transcribe a spoken word. Shorthand systems use special signs to represent phonemes, words and phrases. There are many different shorthand systems currently in use. The most popular ones include: Pitman shorthand method and Gregg shorthand methods. In Mexico the Pitman method was adapted to Spanish by Luis E. Maumejan [5] in 1903 and has 20 basic signs, see Table 1. Actually, the Pitman method adapted to Spanish is used in Mexico.

The use of shorthand methods combined with technology is a powerful tool, useful for every one who wants to write as fast as he speaks, and then having the possibility

© A. Gelbukh, R. Monroy. (Eds.)
Advances in Artificial Intelligence Theory
Research on Computing Science 16, 2005, pp. 161-170

of recovering what he wrote, converted into printed characters. However, when the signs are noisy and with a lack of uniformity, classification is a very difficult problem, because decision regions are hard to define in an optimal way.

First thing that has to be clear is that an OCR can not be used for recognizing shorthand signs due to OCR can not recognize handwriting signs. On the other hand, an ICR can recognize handwriting signs, but the recognition of shorthand signs is a difficult task; this because Pitman's system has basically two signs and the orientation and thickness of these signs are associated with only one sound. In other words, an ICR is not useful for recognizing shorthand signs due to this tool is invariant under rotations and the same sign rotated in different degrees has several associated sounds, see Table 1.

In this paper recognition of shorthand signs is proposed, the efforts are aimed to develop a simple but efficient technique for extracting an invariant descriptor of these sings. This technique is invariant to scaling, translations under x axis and deformations of the signs. First, a sign is obtained and divided into sub-regions, and then the geometric moment of order zero of each sub-region is computed. After that, the invariant description obtained is normalized and transformed into its binary version with the propose of train a neural network for its further recognition. Finally, the identity of the sign is determined by using the neural network already trained. The technique proposed is tested using 100 images of each sign wrote by 10 different people.

The rest of the paper is organized as follows. In section 2, the classification tools used in this research are described. In section 3, the proposal is described in detail. In section 4, numerical examples to better follow the functioning of the proposal are given. In section 5, experiment results with signs and words written by 10 different people are provided, while in section 6, conclusions and directions for further research are given.

2 Neural networks

Since the rebirth of neural networks, several models inspired in neurobiological process have emerged in the last years. Such models are often dedicated and incorporate some existing clustering or classification algorithm. Advantages of neural networks are adaptability, robustness, and easy of implementation. In this paper two neural networks are investigated and compared on their classification capabilities.

2.1 Self-Organizing Maps

Kohonen's Self-Organizing Map (SOM) [1] is one of the most popular artificial neural network algorithms. Kohonen, who presents an unsupervised learning network as an explanation of the existence of ordered maps in the brain, orders the output units such that not only the weight vectors to a winning unit am affected, but those to its neighbors as well, see figure 1(a). SOM learn to classify input vectors according to how they are grouped in the input space. They differ from competitive layers in that neighboring neurons in the self-organizing map learn to recognize neighboring sections of the input space. Thus, self-organizing maps learn both the distribution (as do competitive layers) and topology of the input vectors they are trained on.

The basic idea of a SOM is to map the data patterns onto an *n*-dimensional grid of neurons. That grid forms what is known as the output space, as opposed to the input space where the data patterns are. This mapping tries to preserve topological relations, i.e., patterns that are close in the input space will be mapped to neurons that are close in the output space, and vice versa.

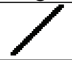
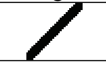
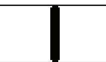
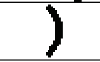
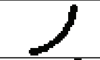
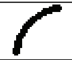
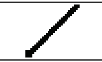
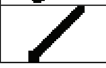
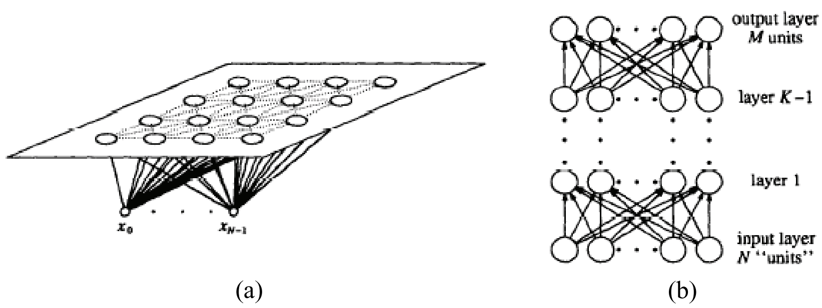
Sign	Sound	Sign	Sound	Sign	Sound	Sign	Sound
	(re)		(rre)		(ne)		(ñe)
	(que, ke)		(gue)		(pe)		(be)
	(te)		(de)		(me)		(es, ex, ez)
	(fe)		(ve)		(er)		(el)
	(le)		(ye, lle)		(che)		(je, ge)

Table 1. Signs used in Pitman's method. Sign (che) is equal to sign (re) but (che) is identified by the direction in which it is written (right-up to left-down); the same to (je, ge) and (rre).

2.2 Feed-Forward network

The general topology of what we call a K-layer' feed-forward network is depicted in figure 1(b). The weights and biases of each unit j in layer 1 describe a hyperplane in N-dimensional space. This hyperplane divides the input patterns in two classes, indicated by the activation value. In a feed-forward network, a unit in layer 2 represents a convex region in hyperspace, enclosed by the hyperplanes determined by the units of layer 1. When three layers are used, any arbitrary shape in N-dimensional space can be enclosed. For detail see [2].



(a) (b)

Figure 1. (a) Typical Kohonen network. (b) General topology of a K-layer feed-forward network.

3 The proposal

One important problem to solve in character recognition is the recognition of shorthand signs even if they are noisy or with lack of uniformity. Signs, written by people, often present distortions or small spurious, see Table 2. The technique proposed in this paper can recognize shorthand signs even if those signs present deformations.

The technique is based on dividing a sign in sub-regions. An invariant description of a given sign is computed using the information of each sub-region.

To compute the invariant description of each sign, we get an image of the sign and then proceed as follows:

```
1.  Apply a standard threshold [2] to get a binary image.
2.  Separate  the sign form its background
3.  Obtain the invariant description.
```

Let **I** an image of length $m \times n$, $\mathbf{x} \in \mathbb{N}^m$ and $\mathbf{y} \in \mathbb{N}^n$ vectors used for isolating a sign are given by:

$$x_l = \sum_{k=0}^{n} i_{lk} \quad , \quad y_l = \sum_{k=0}^{m} i_{kl} \tag{1}$$

xmin = l where the component $x_l \neq 0$ and l is the minimum, xmax = l where the component $x_l \neq 0$ and l is the maximum, ymin = l where the component $y_l \neq 0$ and l is the minimum, ymax = l where the component $y_l \neq 0$ and l is the maximum.

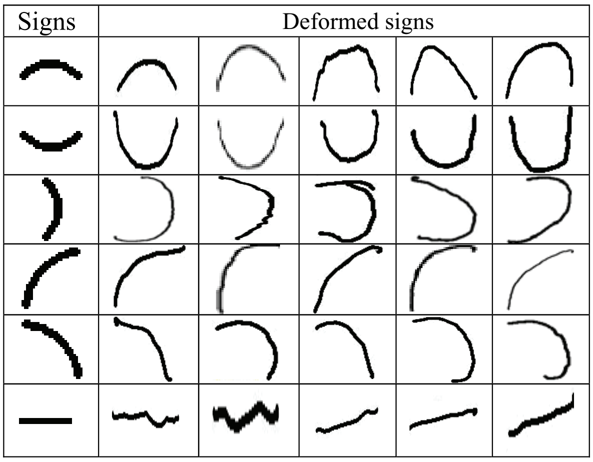


Table 2. Some distorted versions of shorthand signs.

In order to separate the sign from its background, the next steps are applied:

```
1.  Compute vectors x, y using equation 1
2.  For x and y compute xmin, xmax, ymin and ymax.
```

Finally, the sign is centered in a rectangle with vertexes (xmin,ymin) and (xmax,ymax).

3.1 Invariant descriptions of shorthand signs

After a sign is enclosed by the rectangle, the region obtained is divided into 25 sub-regions, see figure 2, and then an invariant description of the sign is obtained.

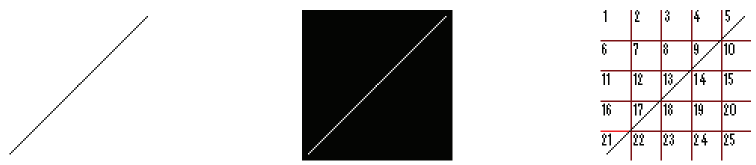
Invariant feature $\mathbf{r} \in \mathbb{N}^{25}$ is obtained computing the geometric moment of order zero [4] over each sub-region as follow:

$$r_i = \sum_x \sum_y x^0 y^0 \mathbf{R}_i(x, y) \tag{2}$$

where $\mathbf{R}_i$ is a sub-region of the enclosed sign.

Finally, the feature vector is normalized and transformed into its binary feature vector version. This transformation is given by:

$$\bar{r}_i = \begin{cases} 1, & \text{if} \quad r_i \geq \dfrac{\max(\mathbf{r})}{3} \\ 0 & \text{otherwise} \end{cases} \tag{3}$$



(a) (b) (c)

Figure 2. (a) Shorthand sign. (b) Shorthand sign enclosed by a rectangle. (c) Region divided into 25 sub-regions.

Thickness of the sign is defined by a threshold. If the density of $\mathbf{R}_i$ is less than the threshold the sign is thin, other case is thick.

In order to determined the associated sound to the sign, the image is divided in five regions: S_A, S_E, S_I, S_O, S_U. The position of the sign in the image (before of being isolated) determine its associated sound, see figure 3. If the center of mass of sign$\left(\bar{x}, \bar{y}\right) \in \mathrm{S}_{\mathrm{A}}$, the associated sound is (a) and so on. Just for remember center of mass is given by:

$$\bar{x} = \frac{m_{10}}{m_{00}}, \ \bar{y} = \frac{m_{01}}{m_{00}} \tag{4}$$

where $m_{10} = \sum_x \sum_y x f(x, y)$ and $m_{01} = \sum_x \sum_y y f(x, y)$.

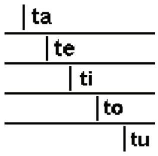


Figure. 3. Sound associated to each region.

4 Numerical examples

To better understanding of the proposal, let us look the next numerical examples. Suppose we want to obtain the invariant vectors from sign and its deformed versions showing in table 2.

First, an invariant vector using equation 2 is obtained. Next, the invariant vector already obtained is normalized and transform into its binary vector using equation 3.

Sign	Invariant vector (using eq. 2)	Normalizad vector (using eq. 3)
	$\mathrm{r} = [0\ 10\ 19\ 12\ 0\ 1\ 9\ 0\ 7\ 1\ 7\ 0\ 0\ 0\ 8\ 6\ 0\ 0\ 0\ 6\ 7\ 0\ 0\ 0\ 7]$	$\overline{\mathrm{r}} = [\ 01110\ 01010\ 10001\ 10001\ 10001]$
	$\mathrm{r} = [\ 0\ 11\ 21\ 20\ 0\ 0\ 9\ 0\ 4\ 8\ 5\ 3\ 0\ 0\ 5\ 10\ 0\ 0\ 0\ 7\ 16\ 0\ 0\ 0\ 7]$	$\overline{\mathrm{r}} = [\ 01110\ 01001\ 00000\ 10001\ 10001]$
	$\mathrm{r} = [0\ 26\ 29\ 23\ 0\ 11\ 1\ 0\ 5\ 14\ 10\ 0\ 0\ 0\ 9\ 11\ 0\ 0\ 0\ 10\ 12\ 0\ 0\ 0\ 0]$	$\overline{\mathrm{r}} = [\ 01110\ 10001\ 10001\ 10001\ 10000\ ]$
	$\mathrm{r} = [0\ 7\ 21\ 21\ 11\ 7\ 15\ 0\ 0\ 13\ 14\ 0\ 0\ 0\ 10\ 10\ 0\ 0\ 0\ 8\ 9\ 0\ 0\ 0\ 4]$	$\overline{\mathrm{r}} = [\ 01111\ 11001\ 10001\ 10001\ 10000\ ]$
	$\mathrm{r} = [6\ 16\ 11\ 0\ 0\ 11\ 0\ 7\ 7\ 0\ 11\ 0\ 0\ 14\ 2\ 11\ 0\ 0\ 0\ 10\ 9\ 0\ 0\ 0\ 13]$	$\overline{\mathrm{r}} = [\ 11100\ 10110\ 10010\ 10001\ 10001\ ]$

As you can appreciate, although the signs appear with small spurious and deformations, the normalized vectors obtained are very close.

5 Experimental results

In this section, the proposal is tested with 100 images of each signs written by 10 different people. A SOM and feed-forward neural network were trained using the feature vector of each sign obtained with the proposal. In a second set of experiments, a feed-forward neural network was trained using the well-known first four Hu descriptors, invariant to translations, rotations and scaling, computed over the signs. In addition, commercial OCR's and ICR's were tested in the recognition of shorthand sings. Finally, 20 words written by 5 different people were use to test the performance of the technique, example of this words are given in Table 5. This experiment consists on translated some words written using signs into Spanish.

5.1 Training of the networks

Only the eleven thin signs whose images are shown in table 1 were used to train the networks. To each sign the feature vector was obtained as in section 3.

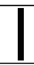
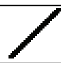
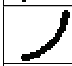
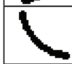
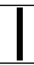
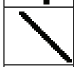
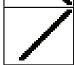
Sign												other
	8%	10%	0%	10%	3%	3%	0%	4%	0%	0%	0%	62 %
	0%	96%	1%	0%	0%	1%	0%	1%	0%	0%	0%	1%
	15%	2%	54%	0%	4%	8%	12%	0%	5%	0%	0%	0%
	0%	0%	0%	95%	0%	0%	0%	0%	0%	0%	0%	5%
	34%	10%	18%	0%	28%	0%	9%	0%	0%	0%	0%	1%
	0%	39%	0%	0%	0%	61%	0%	0%	0%	0%	0%	0%
	0%	0%	29%	0%	1%	1%	69%	0%	0%	0%	0%	0%
	0%	1%	0%	0%	1%	0%	0%	87%	5%	6%	0%	0%
	0%	0%	1%	0%	0%	0%	2%	0%	90%	1%	5%	2%
	0%	0%	0%	0%	0%	0%	0%	0%	0%	97%	1%	2%
	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	97%	2%
% of Recognition		**71%**										

Table 3. Percentage of classification using a feed-forward network.

Feed-forward neural network. It was trained using two layer; input layer with 25 neurons and output layer with four neurons. Training was performed using hardlim function and 50 epochs.

Self-organizing map. A 2-dimensional SOM network of length 25x11 was used. Initial weights were initializing to 0.5. Kohonen learning rate was set to 0.01 and the conscience learning rate was set to 0.001. Training was performed using Euclidian distance and 500 epochs.

5.2 Results

Using the feed-forward neural network combined with the proposal the performance obtained was of 71 %, Table 3 summarizes the classification results.

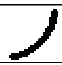
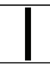
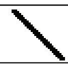
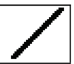
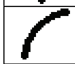
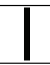
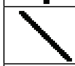
Sign											
	85%	0%	5%	5%	0%	0%	1%	4%	0%	0%	0%
	0%	83%	3%	0%	1%	11%	0%	1%	1%	0%	0%
	0%	0%	84%	0%	0%	0%	11%	0%	5%	0%	0%
	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%
	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%
	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%
	1%	0%	0%	0%	0%	0%	99%	0%	0%	0%	0%
	0%	0%	0%	0%	1%	0%	0%	98%	0%	1%	0%
	0%	0%	0%	0%	1%	0%	3%	0%	92%	4%	0%
	0%	0%	0%	0%	0%	0%	2%	0%	0%	98%	0%
	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	99%
% of Recognition		**95%**									

Table 4. Percentage of classification using SOM.

In general, deformed instances of the signs were incorrectly classified, due to distortions and similarity between the feature vector obtained and the feature vector of each sign. In some cases the sign was not classified into any of the existing classes.

The performance of the proposal combined with SOM was of 95%, Table 4 summarizes the classification results. Majority of the deformed instances of the signs were correctly classified. Bad classification of the instances is due to deformation of the signs. These deformation causes that the feature vector was more similar to other class than its own class, i.e. the signs were classified in the classes of their most similar signs already learned.

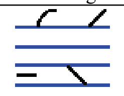
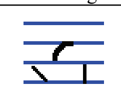
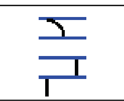
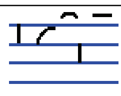
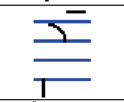
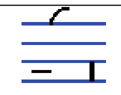
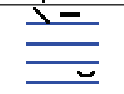
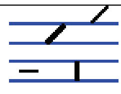
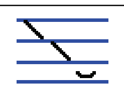
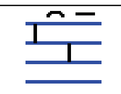
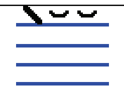
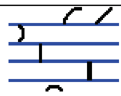
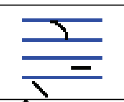
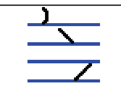
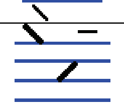
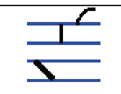
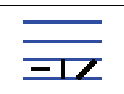
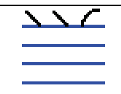
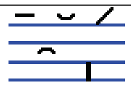
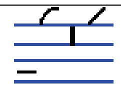
Spanish word	Pitman signs	% of recognition	Spanish word	Pitman signs	% of recognition
Colabora		100%	Pollito		100%
Tuerto		100%	Telematica		80%
Tuerca		80%	Colado		100%
Pagano		100%	Corredora		100%
Pepino		80%	Tematica		100%
Banana		100%	Estimuladora		80%
Puerco		100%	Aspero		100%
Barrica		100%	Botella		100%
Cotorro		100%	Papaya		100%
Caminadora		60%	Coladera		100%

Table 5. Percentage of translation from pitman signs to Spanish word.

Performance drastically decreases when Hu descriptors were used. This is due to those signs are the same sign but in different rotation and Hu descriptors are invariant to rotation therefore, the description obtained from some signs, for example (re, que, te and pe) or (fe, le, ne, me, er, el, es), is the same or the descriptions obtained are very close.

Three OCR were tested: ABBYY FineReader 7.0 Professional Edition, Microsoft Office Document Imaging, Iread Forms. Neither of them could recognize shorthand signs and handwriting. Soft writing 4.1 is an ICR which can recognize handwriting but in the recognition of shorthand signs had a low accuracy, signs (re, que, te and pe) were classified into the same class.

In last experiment 90% of the words were correctly recovered and translated. Incorrect translations were due to some signs incorrectly classified; Table 5 summarizes the classification results.

6 Conclusion and ongoing research

A new technique for obtaining invariant descriptors of shorthand signs has been proposed. The proposal is invariant to scaling, translation and deformations. The proposal has been tested with 100 images of each sign written by 10 different people.

Proposal was tested with two neural networks. The performance obtained using a feed-forward network was of 71 %, while using a self-organizing map the performance obtained increase to 95 %.

In addition, a comparison among the proposal, Hu descriptors, and others commercial OCR's was performed. Hu descriptors, OCR and ICR tools in this kind of signs were useless. This is due to those tools are invariant to rotation and the descriptions obtained from some signs were the same.

The results obtained (95%) support the effectiveness of the proposal and the advantage is clear against Hu descriptors, OCR and ICR tools. Translation of words written using pitman signs was also performed with an accuracy of 90%.

Nowadays, we are testing the proposal using associative memories in order to generalize the method for the recognition of handwriting.

References

1. T. Kohonen. Self-organized formation of topologically correct feature maps. *Biological Cybernetics,* 43:59-69, 1982.
2. F. Rosenblatt. Principles of Neurodynamics: Perceptrons and the theory of brain mechanisms. Spartan, Washington, DC, 1962.
3. N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on SMC*, 9(1):62-66. 1979.
4. M. K. Hu. Visual pattern recognition by moment invariants*. IRE Transactions on Information Theory*, 8:179-187. 1962.
5. Luis E. Maumejan. Manual de Fonografia. Escuela de taquigrafía, 1903.
6. P. Patrick van der Smagt. A comparative study of neural networks algorithms applied to optical character recognition*. Proc. of International conference on industrial and engineering applications of artificial intelligence and expert systems*, 2:1037-1044, 1990.
7. D. Mehr and S. Richfield. Neural net applications to optical character recognition. *IEEE First International Conference on Neural Networks*, 1987.
8. Ning Sun et al. Intelligent recognition of character using associative matching technique. *Proc. of Pacific Rim International conference on Artificial Intelligence*, 1:546-551, 1990.