An algorithm based on visual perception for handwriting comparison

Antonio Parziale, Stanislao Davino, Angelo Marcelli

An algorithm based on visual perception for handwriting comparison

Antonio Parziale, Stanislao Davino and Angelo Marcelli

Natural Computation Lab,
Department of Information Engineering, Electrical Engineering and Applied Mathematics
University of Salerno
Via Giovanni Paolo II, 132
84084, Fisciano (SA), ITALY

Abstract. We propose an algorithm based on a model of visual perception that is meant to reflect the human judgment about the similarity of handwritten samples. The algorithm builds upon the Fuzzy Feature Contrast model and proposes an implementation of such a model in the domain of handwriting. The algorithm has been validated on the RIMES dataset, by comparing its performance with those of a panel of human experts. The experimental results show that the performance of the proposed algorithm is almost indistinguishable from the expert one and therefore may be a viable tool for handwriting comparison.

1. Introduction

Human perception of shape similarity has been the subject of many studies in diverse fields, such as experimental psychology, neuroscience, visual perception and artificial intelligence, just to mention a few. In such a context, the similarity/dissimilarity between stimuli seems to play a key role in defining the categories the organism needs to properly behave in an ever-changing environment. A concise yet comprehensive survey of the current opinions on the subject from such a point of view may be found in (Blough, 2001).

In the realm of handwriting analysis and recognition, when dealing with the problem of building a machine capable of comparing the shape of handwriting sample, the usual assumption is that the similarity of the samples is reflected by the distance of their representations in a feature space: the closer the representations are, the more similar the samples are. The approach makes two fundamental assumptions: the samples can be represented by values of a few distinctive characteristics, or features, and their distance in the feature space reflects their similarity (Plamondon & Shrihari, 2000). Although this approach has lead to applications for handwriting analysis and recognition that are routinely used, such as signature verification, OCR, postal address recognition and check processing, in case of cursive handwriting it provides solutions that are in contrast with the results of experiments with human beings. The book by Pekalska and Duin discusses the limitations of such an approach, argues that the notion of similarity is more fundamental than that of a feature or a class, introduces similarity representations and methodologies to deal with them (Pekalska & Duin, 2005).

Among the theoretical approaches to similarity, the *Feature Contrast model* originally introduced by Tversky accounts for several characteristics of similarity data that contradict the metric assumption discussed above, mainly asymmetry and angular inequality (Tversky, 1977). The model, however, assumes that each sample is described by a binary vector, each element of which represents whether or not the corresponding features is present in the sample. Santini and Jain have extended the original model into the *Fuzzy Features Contrast model*, by proposing to represent the sample by a fuzzy feature vector, each element of which represents the fuzziness of the presence of the corresponding features in that sample, thus allowing to deal with cases when features enumeration is either impractical or impossible (Santini & Jain, 1999).

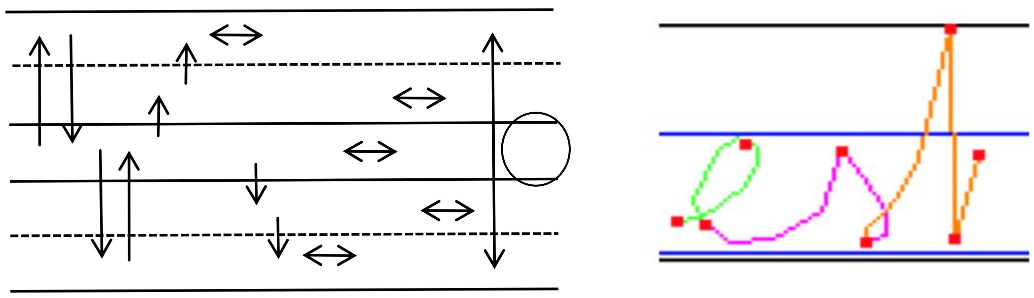
Along this line of investigation, and assuming that cursive handwriting is a sequence of strokes as suggested by many studies on handwriting generation, in Section 2 we propose an implementation of the Fuzzy Feature Contrast model for cursive handwriting that builds upon a novel set of features to describe both the shapes and the spatial arrangement of the strokes, and an algorithm for evaluating the similarity between two cursive fragments. In Section 3 we presents the results of two experiments performed to validate the model and to assess its performance. Eventually, we discuss the experimental results and outline possible directions for future investigations.

2. Ink Similarity

Most of the features used by the Ink Similarity algorithm are based upon measures and classification related to single strokes. The Fuzzy Feature Contrast Model assumes that the feature vectors of the two stimuli have the same dimension; for this reason, to generate “global” features vectors of the same dimension starting from “local” features vectors of different dimension, we decide to aggregate “local” strokes information to obtain “global” information. For each handwritten fragment image, the algorithm creates a vector of 54 elements that holds 3 different types of features: *Zone*, *Curvature* and *Shape* features, as described below.

2.1 Zone features

The sequence of strokes extracted from the ink is preliminarily partitioned in three subsequences: the first one containing the strokes that span over the first 30% of the horizontal size of the word, the second containing the strokes that span over the following 40% of the word, and the third containing the remaining strokes. Each stroke is then classified depending on which part of the word layout it occupies. A zoning algorithm, based on the histogram of the horizontal projection, evaluates the size of the center, the bottom and the upper zones of the word layout. We have defined 15 zone features for each stroke, depending on the position of the stroke in the zone and the way they are drawn: Ascender (Up/Down), Descender (Down/up), Upper (Center/Bottom/Top), Lower (Center/Top/Bottom), Center (Center/Upper/Lower), Pipe and Loop. As each features is evaluated for the three subsequences, we have 45 zone features in total. Fig. 1a) shows the features from left to right, and Fig. 1b) the partition of the sequence of strokes of the word "est" and the zones detected by the zoning algorithm. For each of them, the Ink Similarity counts how many times the *i*-th zone feature appears in the sequence, and then calculate the ratio between such a count and the total number of strokes. This number is then used as a *crisp* value of a linear membership function to obtain the fuzzy value.



a) b)

Figure 1 The "zone" features. a) The features: their labels depend on both the writing direction (represented by the arrow) and the zones of the stroke extremes. The arrows are not representative of the actual shape of the strokes. b) the partitioning into subsequences: begin (green), middle (magenta) and end (yellow)

2.2 Curvature features

As in the zone features case, the curvature features are extracted from subsequences of the sequence of strokes extracted from the ink, but in this case there are five subsequence, each one containing the strokes that span over the 20% of the word horizontal size, from the beginning to the end. For each subsequence we compute the average of the curvature maxima of the strokes, obtaining 5 curvature features. Such a count, normalized as above, is then used as *crisp* value of a linear membership function.

2.3 Shape features

In contrast to the previous ones, shape features do not build up on stroke features, but have the purpose of describing the entire shape of the word. We adopted the 4 word features proposed in (Powalka & al, 1997): Middle, Middle-Upper, Middle-Lower and Upper-Middle-Lower. From the output of the zoning algorithm, and for each word feature, we compute two parameters: the vertical size of the center zone, (*Width)* and the distance between the middle line of the center zone and the bottom line of the bottom zone (*YPos*). Their values are then used as *crisp* value for two membership functions, $F_{YPos}(YPos)$ and $F_{Width}(Width)$ as in fig. 2, whose outputs are eventually combined to obtain the feature score as $min(F_{YPos}(YPos), F_{Width}(Width))$. Eventually, from each score, the degree of ownership of the word's shape features is computed by means of a linear membership function.

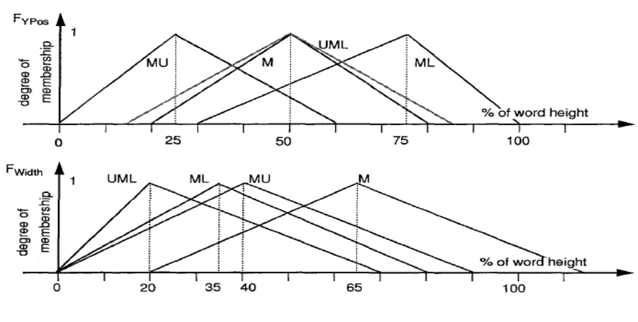


Figure 2. The membership function for the shape features.

2.4 Similarity measure

Given two handwritten words, each one represented by the feature vector described above, we assume as similarity measure between them the following:

$$S(a,b) = \frac{f(A \cap B)}{f(A \cap B) + \alpha f(A - B) + \beta f(B - A)}$$

where:

- a and b represent the image of two cursive handwritten words;
- A and B represent the fuzzy features vectors associated to a and b as described before;
- $A \cap B$ represents the intersection between the two fuzzy vectors. The resulting vector represents the common features between a and b;
- A - B and B - A represent the complements between A and B. The resulting vectors represent the distinct features between a and b;
- $f(FV)$ is the saliency function that associates to an entire feature vector FV a single number; in our implementation we choose the function f as: $f(FV) = \sum_{i=1}^{54} FV_i$
- α and β are two weights that model the imbalance of the judgment of inequality that is typical of human judgment.

3. Experimental results

In order to validate the proposed model, we have performed a set of experiments on the RIMES dataset, a publicly available dataset largely used for performance evaluation of handwriting analysis and recognition systems (Grosicki & al., 2008). From the data set, we have selected 10 images of the bigram “en” and 10 images of the word “es” as Reference Set (RS) and again 10 images of the bigram “en” and 10 images of the bigram “es”, different from the previous ones, as Test Set (TS).

In each experiment, 1 image of TS and 5 images of RS, randomly selected but the same for all the subjects, were shown to each subject, and he/she was asked to rate the similarity between the Test image and each of the Reference one. The rating was reported by using a 5-point scale, ranging from 1, the most similar, to 5, the least similar. Figure 3 shows the GUI designed for the experiments with the human subjects. The same task was assigned to the algorithm, so as to have, for each image of TS, the ranks of 17 subjects and the rank of the algorithm.

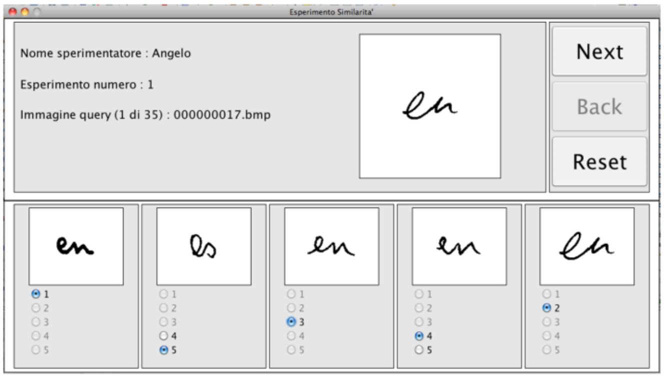
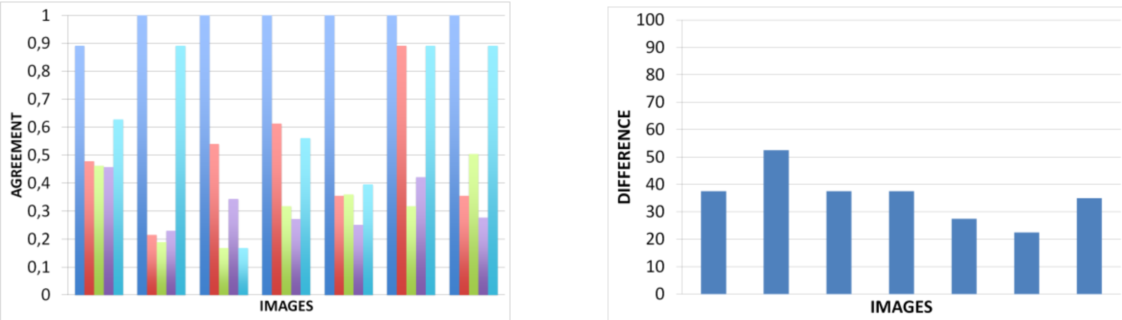


Figure 3. The GUI used during the experiments. The subjects were requested to rank the similarity between each of the Reference images shown in the bottom pane with the Test image shown on the top pane by using a 5 point scale, with 1 representing the most similar Reference image and 5 the least similar one.

Then, for each image of TS, we measures the difference D in the ranks by the formula:

$$D = \sum_{i=1}^{5} d_i * w_i$$

where d_i is the difference between the position of the *i-th* image of RS in the two ranks and w_i = 1 for the top/bottom position of the rank, 0.5 for the following/preceding one and 0.25 for the middle position. The weights have been fixed so as to ensure that very similar/dissimilar feature plays a major role in the final judgment, as in the case of human perception of similarity. Figure 4a) reports the level of agreement between the subjects for 10 of the images of TS, while figure 4b) reports, for the same images, the value of *D* between the rank provided by the proposed algorithm and the one obtained by combining the ranks of the subjects by the Borda count method (deBorda, 1781), as to represent the "mean" behavior of the human subjects.



a) b)

Figure 3. The experimental results: a) the agreement on the ranking between the subjects: each of the 5 bars refers to the agreement on the corresponding position of the rank; b) the difference between the algorithm and the "mean" subject. For sake of legibility, the figure reports the results on a few images selected as representative of the performance on the whole Test set.

Conclusions

We have presented an algorithm to evaluate the similarity between handwritten words that builds upon a fuzzy computational model proposed to account for the visual perception of similarity. Such a general model has been customized by adopting a suitable set of features to represent the distinctive feature of handwriting and its performance compared with that of human subject in a similarity evaluation task.

The experimental results allow for the following preliminary conclusions:

- the ranks of the human subjects are more similar as with regards to the most/least similar shapes than with regards the shapes that are somehow in between these extreme cases;
- the difference between the rank of the algorithm and those of the subjects may vary, but even in the worst case such a difference is slightly bigger than 0.5, meaning that at most two images were ranked differently and that those images were not ranked as the top or the bottom ones;
- the performance of the algorithm depends to a limited extent from the images, suggesting that it is a robust implementation of the fuzzy computational model it builds upon.

According to those results, the proposed algorithm seems to implement an agent whose behavior resembles that of the human subjects in that:

- its judgment is very similar to those of the subjects as with regard to the most/least similar samples;
- the differences between the rank of the algorithm and those of the subjects are very similar to the differences within the subjects, so as to make the proposed algorithm indistinguishable from any of the subject.

In the future, we will perform further experiments, on larger data sets, including longer words, considering different membership function and different implementations for the similarity measures, in order to ascertain the performance of the proposed algorithm with different handwriting styles, as well as its robustness with respect to the actual values of its parameters.

References

Blough, D.S. (2001). The Perception of Similarity, in: R.G. Cook (Ed.), *Avian Visual Cognition*, [On-line]. Available: www.pigeon.psy.tufts.edu/avc/dblough/

Plamondon, R. & Shrihari, S.N. (2000), Online and Off-line Handwriting Recognition: A Comprehensive Survey, *IEEE Trans. on Patt. Anal. Mach. Intell.*, 22(1), pp. 63-84.

Pekalska, E. & Duin, R.P.W. (2005). The dissimilarity representation for pattern recognition. Foundations and applications. Singapore: World Scientific Press, 2005.

Tversky, A. (1977). Features of Similarity, *Psychological Review*, 84(4), pp. 327-352.

Santini, S. & Jain, R. (1999). Similarity Measures. *IEEE Trans. on Patt. Anal. Mach. Intell.*, 21(9), pp. 871-883.

Powalka, R.K, & al. (1997). Word Shape Analysis for a Hybrid Recognition System, *Pattern Recognition,* 30(3), pp. 421-445.

Grosicki, E. & al. (2008). RIMES evaluation campaign for handwritten mail processing. *Proc. of the Int. Conf. on Frontiers in Handwriting Recognition*, Montreal, Canada, pp. 1- 6.

de Borda, J.C. (1781). Mémoire sur les élections au scrutin. Histoire de l'Académie Royale des Sciences, Paris.