HANDWRITTEN CHARACTER RECOGNITION USING NEURAL NETWORKS

Philippe GENTRIC *Laboratoires d'Electronique Philips*
22, av. Descartes 94453 Limeil-Brevannes (FRANCE)
Email: gentric@lep.lep-philips.fr

Introduction

The aim of this work is to assess the possibility of having on-line handwriting recognition abilities on a small portable computer. We studied the recognitions rates dependence on the target character set and the incremental training made possible by new neural network techniques. We worked on cross-cultural data gathered from French and Dutch people containing about 70000 characters from 200 peoples, both segmented and unsegmented.

Segmentation and character recognition

In our on-line recognition system the input signal is the pen tip position and 1-bit quantized pressure on the writing surface. Segmentation is performed by building a string of "candidate characters" from the acquired string of strokes. For each stroke of the original data we determine if this stroke does belong to an existing candidate character regarding several criteria such as: overlap, distance (it also provides word-building information) and diacriticity (position and size as for the point on a "i"). Finally the regularity of the character spacing can also be used in a second pass. In case of text recognition, we found that punctuation needs a dedicated processing due to the fact that the shape of a punctuation mark is usually much less important than its position. The punctuation processing is made using an "expert" with rules such as: "a small isolated stroke above the average character height with nothing underneath is a quote". Depending on the results of the character recognition, it may be decided that the segmentation was wrong and that back-tracking on the segmentation with changed decision thresholds is needed. We tested two encoding and two classification methods. As the aim of the writer is the written shape and not the writing gesture it is very natural to build an image of what was written and use this image as the input of a classifier. We found that in this case the best classifier is a locally-connected constrained-weights MLP [Le Cun & al., 1990] Indeed, the use of a RBF-based classifier requires then a dedicated metric and we found that this was much more computer intensive than a perceptron. In the case of direct encoding of dynamic data we found the best classifier to be a Radial Basis Function (RBF) network because it has equivalent performance when compared to a multilayer perceptron (MLP) but it provides user-adaptability.

Constructive Tree Radial-Basis-Function network (CTRBF)

One very important requirement for on-line handwriting is the possibility to adapt to the user. This adaptation must also be performed on-line which is not possible with classical algorithms (such as perceptrons). For this reason, we developed CTRBF, a new neural classification algorithm. Radial-Basis-Function Networks are known to be capable of universal approximation [Park & Sandberg, 1991] and the output of a RBF network can be related to Bayesian properties. One of the most interesting properties of RBF networks is that they provide intrinsically a very reliable rejection of "completely unknown" patterns at variance from MLP. Furthermore, as the synaptic vectors of the input layer store *locations* in the problem space (a pattern), it is possible to provide incremental training by creating a new hidden unit whose input synaptic weight vector will store the new training pattern. Most of the RBF literature [Hartmann & Keeler, 1991; Moody & Darken, 1988] is devoted to the use of optimization techniques in order to compute an optimized set of synaptic weights in the RBF network. Our experience is that for real world problems these methods demand a great amount of computations and in the end produce systems that do not perform much better than CTRBF. The specifics of CTRBF are firstly that a search tree is associated to a hierarchy of hidden units in order to increase the evaluation speed and secondly we developed several constructive algorithms for building the network and tree [Gentric, 1993] .

Experimental results

We found that the recognition (generalization) rate depends on the target character set size as described in figure 1 (these performances are without rejection). For two different target sets (10 digits and 84 ASCII characters) we plot the error-rejection trade-off in figure 2. "1-answer" is when only the most active RBF output unit is taken into account, "multi-answer" means that the second and third best answers are also counted (this is useful when associated with higher level processing such as a lexical checker). One can see that the "multi-answer" performance for 84 classes is the almost the same as for "1-answer" performance on 10 classes. The recognition speed is 10 characters per second on a Intel 486 computer.

figure 1

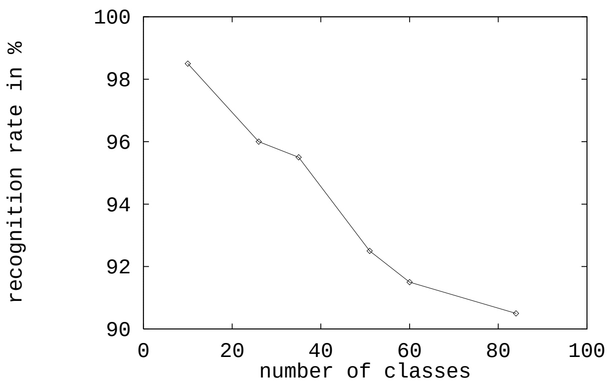
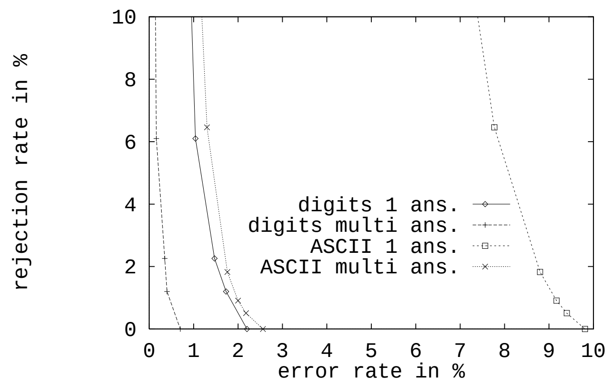


figure 2



The cross-cultural variability exists but is not easy to demonstrate. In the following example we use 30000 French character and 3000 Dutch character sets each split into training and test sets. If we learn the system with French and Dutch data we get 92% recognition rate in average on both test sets. If we train the system using the French training set the results on French and Dutch test sets are listed in table 1. Then the network is incrementally trained using the Dutch training set and tested on the Dutch test set (last line of table 1). Table 2 gives the results obtained when the nationalities are exchanged. Our conclusion is that our Dutch data was collected from "less cooperative" persons than the French data, but it may also be suspected that a multi-cultural handwriting recognition system would have a lower performance than a system specialized on a given population (nationality, education, etc...). More detailed results also indicate that this depends on the target character set.

One important part of this project was to assess the efficiency of "customizing" a handwriting recognition engine for one particular person. We first made an engine trained using 30000 upper case and lower case characters which had a 92% recognition rate. Then we used data for specific persons for incremental training. The first trials we made gave very deceiving results because we were only using a few tens of character per person. Table 3 gives the results obtained by performing incremental training with about 4000 characters per person. The persons used for the test where NOT represented in the original data. Obviously the incremental training efficiency depends on each individual but it is very efficient. On the other end collecting 4000 characters takes several hours ...

Conclusion

In conclusion, we report conditions under which user adaptable handwriting recognition can be reasonably implemented using neural networks. This study suggests that in a real application due to the dependence on the target set and culture it would be important to use specialized character recognition engines for each sub-task.

conditions	error	rejection	1st choice	2nd choice	3rd choice
test on French data	1.1	0.0	94.7	3.6	0.6
test on Dutch data	9.5	0.0	74.7	11.2	4.6
test on Dutch data after incremental training	2.4	0.0	88.9	6.9	1.8

Table 1: Incremental training with Dutch data

conditions	error	rejection	1st choice	2nd choice	3rd choice
test on Dutch data	2.9	0.0	89.3	6.6	1.2
test on French data	22.9	0.0	65.7	8.3	3.1
test on French data after incremental training	1.1	0.0	94.5	3.6	0.7

Table 2: Incremental training with French data

	before incremental training	after incremental training
person 1	92%	95%
person 2	80%	97%
person 3	92%	98%
person 4	88%	95%

Table 3: Incremental training for 1 person (4000 characters) "multi-answer" recognition rates

Acknowledgments

We would like to thank Hans DOLFING from Philips Research Laboratories in Eindhoven for his collection of Dutch data and his essential work on the software integration of the recognition engines on the PC platform.

References

• Y. Le Cun et al., Neural Information Processing Systems, Vol. 2, D. Touretzky (editor) Morgan Kaufmann (1990)
• P. Gentric, Constructive Methods For A New Classifier Based On A Radial-basis-function Neural Network Accelerated By A Tree, Proc. International Workshop on Artificial Neural Networks, Barcelona (1993).
• E.Hartmann, J.D.Keeler, Predicting the Future: Advantage of Semilocal Units, *Neural Computation* 3 (1991)
• J. Moody and C. Darken, Learning with localized receptive fields, *Proceedings of the 1988 Connectionist Models Summer School*, ed. D. Touretzky, Morgan Kaufmann Publishing, San Mateo, CA (1988).
• J.Park, I.W. Sandberg, Universal Approximation Using Radial-Basis-Function Networks, *Neural Computation* 3 (1991)