ENHANCING HANDWRITTEN CHARACTER IMAGES
THANKS TO A RE-SAMPLING PROCESS BASED
ON CONVEX HULL EXTRACTION

B. Gosselin
Signal Processing & Circuit Theory Lab.,
Faculté Polytechnique de Mons, 31 Bd Dolez,
B-7000 Mons, Belgium
Tel : +32 65 37 41 33 ; Fax : + 32 65 37 41 29
e-mail : gosselin@tcts.fpms.ac.be

ABSTRACT

In this paper, we propose a new method that allows, by
finding the convex hull of a character image, to set out in
one pass only, the control parameters of a particular
character distortion process. This character distortion
method can then be applied to normalize the character
image, i.e. to reduce the within-class scatter of images of
handwritten characters, which could lead to a significant
improvement of recognition performance. Many tests have
been performed on unconstrained handwritten uppercase
letters extracted from the NIST3 database [1]. Finally, the
combination of two classifiers, one using the proposed
normalization method, and the other one not, has allow
reducing the overall error rate from 5.24% to 3.88%.

1. INTRODUCTION

As any training process of a statistical pattern
recognition system, the one of an artificial neural network,
such as a multilayer perceptron, for optical character
recognition, require a large database. In the case of
handwritten characters, it is often very difficult and tedious
to collect enough samples so as to complete a convenient
database. Then, although multilayer perceptrons show good
capabilities in performing classification tasks, this may
results in a limitation of their recognition performance.

A common approach that may be investigated in order
to attempt to improve performances of a particular character
recognition system consist in applying local distortions to
the images of the characters of the training database, so as
to increase its diversity. This approach seems to be
particularly suitable as the classifier is based on an artificial
neural network, such as a multilayer perceptron for instance,
as they possess very good properties of learning
discriminant features from samples of patterns.

Several efficient character distortion methods have
already been proposed, but most of them require a
representation of a character thanks to a sequence of
straight lines. This is the case for the methods presented in
[2], [3] or [4]. Such a kind of representation is very difficult
to obtain in case of off-line character recognition, however.
As a matter of fact, the writing process is then unknown,
and there is thus not any temporal information available.

A distortion method that is based only on the bi-
dimensional image of the characters has then been
developed, and has appeared to allow improvement of the
recognition performance [5]. This method is based on four
control parameters, which have to be set out so as to ensure
that the new character images will provide enough new
useful information to the classifier. This implies that
sufficient distortions are applied to the training characters.
However, it may then occur that these distortions, besides
being plausible, causes the classifier to not correctly
recognize characters, although their images present a decent
aspect. Moreover, as the control parameters can vary
randomly within a given range, there is no insurance that the
distortions that the neural network will assimilate during its
training phase will correspond, from a statistical point of
view, to the ones encountered in practice.

Then, we propose a method that allows, by finding the
convex hull of a character image, to set out efficient values
for the four control parameters of the distortion method, so
that it can be applied to normalize a character image, i.e. to
reduce the within-class scatter of images of handwritten
characters, and by doing so, to improve the robustness of
the extracted features.

2. THE DISTORTION PROCESS

The distortion method relies on the application of a bi-
dimensional sampling grid in order to normalize a character
image (figure 1). This sampling grid is completely defined
by four control points, i.e. its four corners.

The simplest way to fix the ordinates of the sampling
points is to perform the computation in two steps. For a
particular sampling point $S(i,j)$, the equation of the straight
line which corresponds to the given value of index \( i \) is first computed, relatively to the ordinates of the four corners of the grid. The offset that corresponds to the current value of index \( j \) is then reported along this straight line (figure 1.b). Once the ordinates of a sampling point are fixed, its binary value is attributed as the one of the closest pixel of the initial image.

\[
\begin{align*}
\mathbf{P} \Rightarrow A P = \lambda_i, AD &\Rightarrow S; PS = \lambda_j, PQ \\
\mathbf{Q} \Rightarrow BQ = \lambda_j, BC
\end{align*}
\]

With:

\[
\lambda_i = \frac{i}{N_s - 1} \quad \& \quad \lambda_j = \frac{j}{N_s - 1}, \quad 0 \leq i, j < N_s
\]

Where \( N_s \) is the total number of sampling points, according to each direction. Of course, it has to be set out so as to avoid under-sampling of the character image. Moreover, it has appeared that over-sampling is necessary in practice, in order to obtain high-quality images for the new character [5].

3. EXTRACTION OF THE CONTROL POINTS

The convex hull of the image of the character is first determined. Each segment of the hull is then assigned a position and an orientation (left or right, vertical, or upper or lower, horizontal). Four sets of segments are so defined. From the first and the latest point of each of these sets, four new segments are then determined (figure 2). Given the orientation of these four new segments, a convex hull of four straight lines can be obtained. Finally, the control points of the normalization process are given by the intersections of these last straight lines.

4. APPLICATION

The proposed character image enhancement method has been tested on a recognition problem of unconstrained handwritten uppercase letters extracted from the NIST3 database [1]. 1324 samples of each class have been taken as training set, while 235 other different samples have been chosen as test set. The feature extraction process, described in [6], did consist in a gray-level representation of the image of the character, thanks to a matrix of reduced dimensions. A discriminant analysis method that relies on the generalized Fisher’s criteria [7] has been applied, in order to select the best components of the feature vector for classification. On the basis of the selected features, a multilayer perceptron with one hidden layer has been trained as a classifier. Without applying the proposed normalization method, an error rate of 5.24 % has been obtained on the test set.

Samples of images of characters, before and after application of the proposed normalization process, are shown at figure 3. It appears that the proposed method may produce significant distortion in some cases (letters F and L). A constraint has been introduced to avoid this, so that the re-sampling process can not be applied as the length of
any of the four sides of the sampling grid is too small. From many empirical tests, the minimal length required has been set out to 1/3 of the initial width or height of the character.

Figure 3 - Samples of characters, before and after application of the proposed normalization method.

This modified normalization method has been applied to the images of the characters of the database, and the same feature extraction process than before has been used. The training of a multilayer perceptron has then led to an overall error rate of 5.74%. Although this overall error rate is slightly higher than before, it is important to observe that, for some particular classes of characters, the error rate has been reduced (figure 4). Moreover, it also appears that, for a given class, the misclassification errors don’t occur at the same samples of characters.

Then, we have tried to set out a method to decide if the normalization process should be applied or not. This method relies on the fact that, assuming some hypotheses, the outputs of a multilayer perceptron trained for classification may be regarded has providing good approximations of the \( a \ posteriori \) probability of each class to be the one of the input pattern \([8]\). According to Bayes’ rule, selecting the class whose \( a \ posteriori \) probability is the highest then performs the best classification. So, the whole database has been divided in two classes, a first one for the characters that are better recognized without applying the normalization process, and a second one for the other characters. A multilayer perceptron has then been trained to classify the characters of the whole database between these two classes, \( i.e. \) to decide to apply or not the normalization process.

This, however, did not led to better error rate. Instead, the combination of two multilayer perceptrons, the first one being trained on the basis of the initial characters, and the second one being trained on the basis of the normalized characters, has allow obtaining a significant reduction in the overall error rate (figure 5). The combination of the two multilayer perceptrons is simply performed through a third perceptron, whose inputs are taken directly from their outputs (figure 6) \([9]\).

Figure 4 - Error rates computed over each class, without (in light gray), and with (in dark gray) application of the normalization process.

Figure 5 - Error Rates: without applying the proposed normalization process (a), when using it (b), and when using the two recognition systems in combination (c).

Figure 6 - Combination of Multilayer Perceptrons.
5. CONCLUSION

A previous approach did consist in applying local distortions to the image of the characters of the training database, so as to increase its diversity. This has led to the development of an original character distortion method, whose parameters may vary randomly, within a given range. By doing so, however, there is no insurance that the distortions that the neural network will assimilate during its training phase will correspond, from a statistical point of view, to the ones encountered in practice.

In this paper, we have presented a new method that allows, by finding the convex hull of a character image, to set out in one pass only, efficient values for the control parameters of the previous character distortion method. This character distortion method can then be applied during the recognition phase, so as to normalize a character image, i.e. to reduce the within-class scatter of images of handwritten characters.

The tests that have been performed on unconstrained handwritten uppercase letters extracted from the NIST3 database have shown that the new approach allows classifying correctly various characters, which were not correctly recognized before. As a result, the combination of two classifiers, one using the proposed normalization method, and the other one not, has allow reducing significantly the overall error rate.

References:


