GENERATING NEW SAMPLES FROM HANDWRITTEN NUMERALS BASED ON POINT CORRESPONDENCE

Mori, Minoru; Suzuki, Akira; Shio, Akio; Ohtsuka, Sakuichi

Generating New Samples from Handwritten Numerals Based on Point Correspondence

Minoru Mori, Akira Suzuki, Akio Shio, and Sakuichi Ohtsuka
NTT Cyber Space Laboratories
1-1 Hikari-no-oka, Yokosuka-shi, Kanagawa, 239-0847 Japan
E-mail: mmori@marsh.hil.ntt.co.jp

This paper describes a character generation method based on point correspondence between patterns. The number of training samples used in constructing a recognition dictionary strongly affects its recognition performance. Unfortunately, it's so time-consuming to gather large new samples that it is more useful to generate new samples from a few original ones. The character generation method proposed herein is based on the point correspondence between each sample and the template derived from all samples. The proposed method can automatically generate new samples that appear to be written naturally and extends the handwriting deformation seen in the original samples. Initial experiments show that using the samples so generated can improve the recognition performance.

1 Introduction

In character recognition, many training samples are needed to construct a useful recognition dictionary; increasing the number used increases the recognition performance since the training samples can cover wider variations in handwriting character deformation. Unfortunately, constructing a high performance recognition dictionary needs a very large number of training samples; the number of training samples available and the variations of the handwriting deformation contained in these training samples are limited. It goes without saying that the best approach is gathering many new samples written by different people, but this is too time-consuming and expensive.

To avoid this problem, the generation of samples has become a key research topic [1,2,3]. As for handwritten characters, several deformation models for sample generation have been proposed [4,5,6]. Unfortunately, these top-down models don't reflect the deformation tendencies seen in handwritten characters.

In this paper, we propose a generation method that is based on the point correspondence technique and that well reflects handwriting deformation trends as derived from training samples. We describe recognition experiments that assess a dictionary containing original and generated samples. Section 2 introduces the algorithm of the character generation method. Section 3 describes recognition experiments using the dictionary holding the original training and generated samples and discusses the results. Finally, we summarize this paper and list future works in section 4.

In: L.R.B. Schomaker and L.G. Vuurpijl (Eds.), Proceedings of the Seventh International Workshop on Frontiers in Handwriting Recognition, September 11-13 2000, Amsterdam, ISBN 90-76942-01-3, Nijmegen: International Unipen Foundation, pp 281-290

2 Character Generation Method

2.1 Conventional Methods for Generating Samples

Several methods have been proposed for generating new samples from off-line data. Ishii applied the perturbation operation to standard patterns [4,5]. Examples of using nonlinear transformation for deforming patterns include barrel, pincushion, and trapezoidal transformations [7]. As for linear transformation, skew has been proposed [8]. However, since these conventional methods involve top-down models, they always don't generate natural-looking samples in terms of deformation, and don't reflect the tendency of handwriting deformation derived from the actual samples collected.

Methods that assess on-line character data have been proposed for analyzing the tendency of handwriting deformation and finding the correspondence between patterns easily [12]. However, on-line data is not always effective as the training samples used to construct OCR recognition dictionaries. Moreover, many on-line samples are needed the collection of which is time-consuming and expensive.

2.2 Proposed Generation Method Based on Point Correspondence

We propose a method of automatically generating extended off-line character samples by using, for each category, the pattern correspondence between the training samples and the template pattern. The template pattern is derived from all training samples of the same category. Because the method uses the correspondence between each sample and the template, it can generate samples that reflect the deformation in the training samples and generate samples inside and/or outside the class boundaries for the category. We also use the point correspondence based on the distance between the pixels of a pattern and the pixels of the template in forming the pattern correspondence. The process of generating new samples is as follows.

Step 1. Set the template as the standard point.

Step 2. Determine the pattern correspondence between a sample and the template.

Step 3. Set a standard displacement path for each pixel pairing between the sample and the template.

Step 4. Create new displacement paths by varying the standard ones with a variation parameter.

Step 5. Map each pixel of the newly generated sample by using the displacement paths.

2.3 Template as the Standard Point

We create the template pattern, for each category, as the standard point by computing the average of all samples of the same category as follows. Let b^c_{xyk} be the value ($= 0$ or 1) of each coordinate (x, y) of the k-th sample in the c-th category. Let bn^c_k be the number of black pixels in the k-th sample of the c-th category. First, the mean value on each coordinate $\bar{b}^c_{xy}$ is given by $\sum_{k=1}^{N^c} b^c_{xyk}/N^c$; N^c is the number of the samples of the c-th category. Next, the mean number of black pixels in sample $\bar{bn}^c$ is given by $\sum_{k=1}^{N^c} bn^c_k/N^c$. Then, the $\bar{bn}^c$-th largest value in $\bar{b}^c_{xy}$ is set as the threshold α^c. Finally, the template $temp^c_{xy}$ is obtained using the following equation.

$$temp^c_{xy} = \begin{cases} 1 & (\,\bar{b}^c_{xy} \geq \alpha^c) \\ 0 & (\,\bar{b}^c_{xy} < \alpha^c) \end{cases} \tag{1}$$

An example of a template made from 100 samples of the handwritten numeral 5 is shown in Figure 1(a).

2.4 Point Correspondence

Next, the correspondence of each sample to the template is found using a point correspondence technique. Point correspondence using the distance between the pixels of the patterns is sometimes used for on-line character recognition[11]. Here we propose a point correspondence algorithm that allows for the inevitable imbalance in the pixel number between the sample and the template in each local area. In a pre-processing operation, the template and each training sample are skeletonized. Let $S = (s_1, ..., s_i, ..., s_m)$ be the pixel of the skeletonized sample and $T = (t_1, ..., t_j, ..., t_n)$ be the pixel of the skeletonized template.

Step 2-1. Calculate distance ds from each s to all t and distance dt from each t to all s.

Step 2-2. For each s_i, find t_j with minimum ds_i. For this t_j, find $s_{i'}$ with minimum dt_j. If $s_i = s_{i'}$, set the correspondence between the s_i and the t_j as the initial state.

Step 2-3. For each unmatched s_i, find t_j with minimum ds_i. If this t_j is unmatched, find $s_{i'}$ with minimum dt_j among unmatched s. If $s_i = s_{i'}$, set the correspondence between the s_i and this t_j.

Step 2-4. Find the correspondence between unmatched pixels in t_j and s_i by the same way of **Step 2-3** from T to S.

Step 2-5. Repeat **Step 2-3** (from S to T) and **Step 2-4** (from T to S) L times from the minimum distance until the L-th distance between S and T. Here we choose $L = 4$.

Step 2-6. Each s that remains unmatched is mapped to t that is nearest among t within the 8-neighborhood of the s of interest. In the same way, each remaining unmatched t is mapped to appropriate s.

Step 2-7. The pixels (except the skeleton pixels) of each training sample are mapped onto the template by shifting them by the same distance as that between the nearest skeleton pixel pair.

Figure 1(b) illustrates the correspondence result for one sample and the template of numeral 5. The correspondence result appears suitable even though some local areas exhibit an imbalance in pixel number.

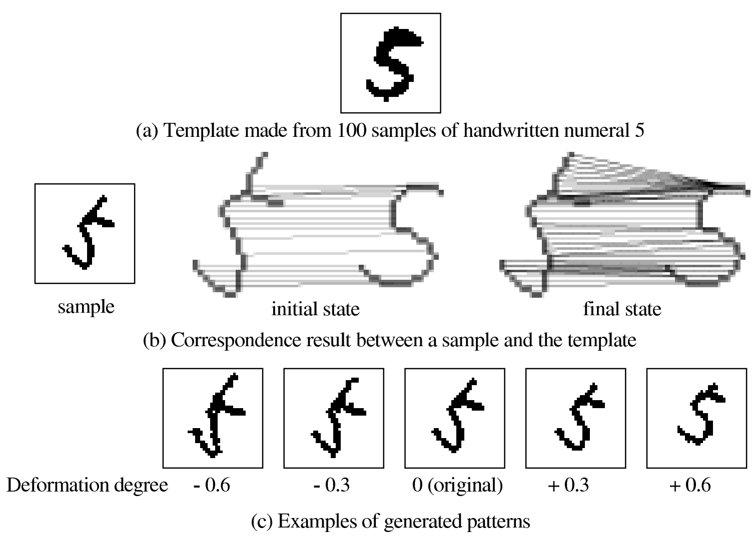

Figure 1: Examples of the template, correspondence result, and generated patterns

2.5 Generation of Deformed Patterns

The standard paths for each pixel pairing are varied by a parameter to yield new displacement paths. The new displacement paths create a new deformed sample by moving each of the original pixels to new each coordinate, respectively. Figure 1(c) illustrates patterns generated by deforming the sample shown in Figure 1(b) by the method mentioned above. Plus values indicate distortion towards the template. Using a value of 1 generates a pattern that is virtually identical to the template. Minus values yield distortion away from the template. The deformed samples look like natural handwriting.

3 Recognition Experiments

3.1 Character Data, Feature, and Discriminant Function

Samples for training and testing were collected from the data of numerals handwritten in a totally unconstrained manner, and were digitized into 40×64 pixels. In pre-processing we normalized each sample in terms of position and size. The mean of the sample was shifted to the center of the rectangle. Sample size was normalized to 32×32 pixels. As the training data set, we used $1,000$ samples per category. For the test data set, we used another $1,000$ samples per category. Examples are shown in Figure 2. As the feature, we used the chain-code feature [10] which is considered effective for character recognition. The chain-code feature is obtained by storing the Freeman chain-code extracted from the character contour on local blocks obtained by partitioning the sample. The feature has 5 blocks $\times$ 5 blocks $\times$ 4 directions $=$ 100 dimensions. As the discriminant function, we used the nearest neighbor (NN) method [9], one of the non-parametric methods. The NN consumes large memory space, but the dictionary is simple to construct and it is easy to see the influence of adding the generated samples to the original training samples.

3.2 Experimental Results

(a) Recognition Result Using Original Training Samples

First, we examined the relation between the number of the training samples and the recognition rate. Figure 3 shows the recognition rate for the whole test data set using the dictionary constructed by using only the original training samples. The vertical axis shows the recognition rate. The horizontal axis shows the number of training samples per category. These results indicate that the recognition rate increases as the number of the training sample increases.

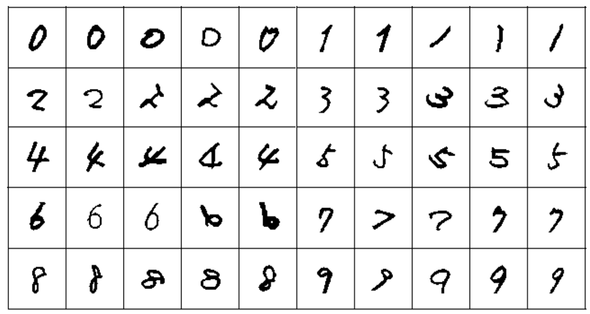

Figure 2: Examples of training samples that are normalized in terms of position and size.

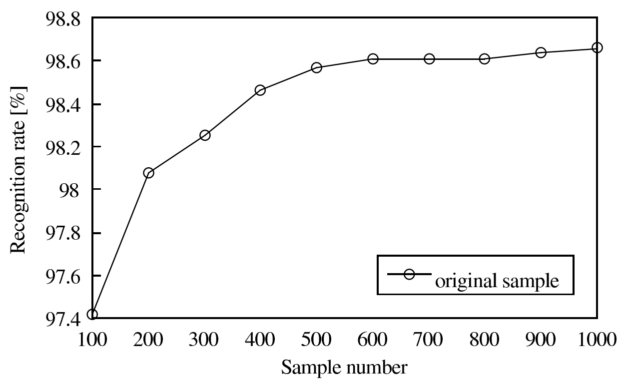

Figure 3: Recognition results using the dictionary constructed by original training samples.

(b) Recognition Result Using Original and Generated Samples

Next, we evaluated the effect of adding newly generated samples. We generated one sample from each of the 100 training samples. Each displacement path was varied by the parameter, which was varied in the range of $+1.0 \sim -1.0$. Examples of the original training samples and some generated samples are shown in Figure 4. Prior to adding generated samples to the original ones, we conducted an experiment in which training samples were recognized using the dictionary holding only generated samples. Generated samples that degraded recognition performance were removed; the remaining samples were used in the subsequent experiment. Figure 5 shows the recognition rate using the dictionary constructed by adding the generated samples to the set of 100 original samples. The horizontal axis shows the deformation degree. The results show that a slight degree of distortion can improve recognition performance. Large levels of distortion (both positive and negative) lower the recognition rate.

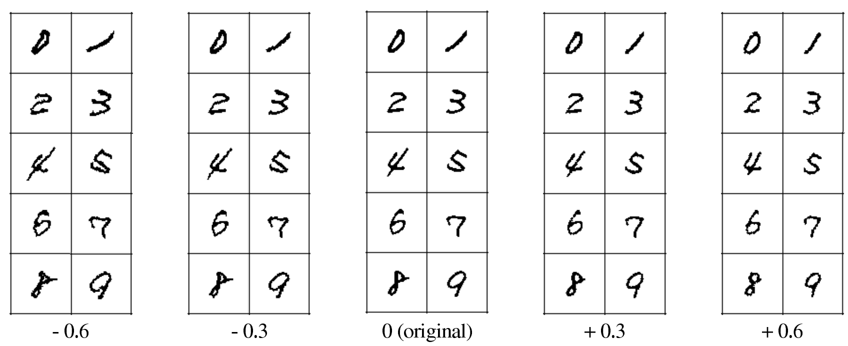

Figure 4: Examples of the original and generated samples. Each value shows deformation degree.

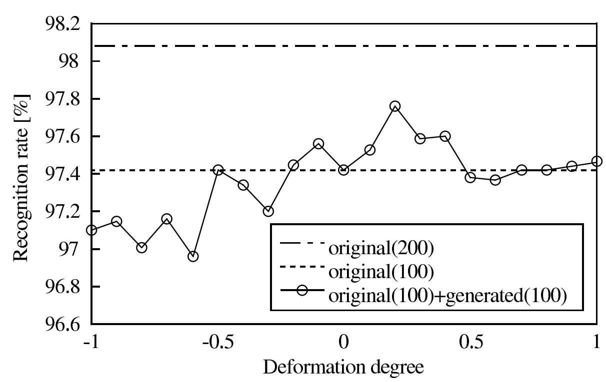

Figure 5: Recognition results using dictionaries holding original and generated samples.

Results of similar experiments using 200, 500, and 1,000 samples as original training samples are shown in Figure 6. The circles (original + generated) show the best recognition result obtained in each experiment. Please note that the number of generated samples never exceeds the number of the original samples due to the removal of some generated samples in preliminary experiments. The horizontal axis shows the number of original samples per category in log scale. The improvements of recognition performance due to the addition of newly generated samples do not, at first glance, seem to be large. However, the addition yields 40 $\sim$ 50% of the improvement achieved by doubling the number of original samples.

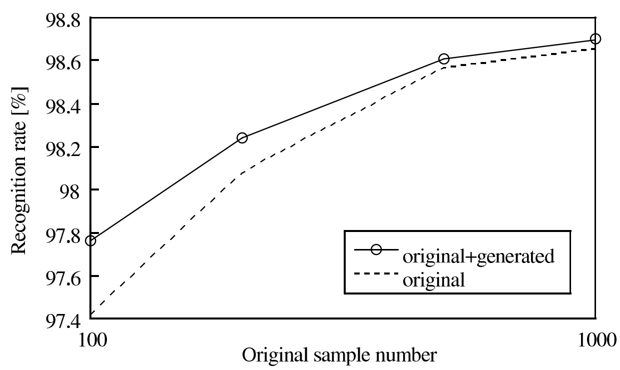

Figure 6: Recognition results using each number of original and generated samples.

3.3 Discussion

Figure 5 shows that deforming the samples away from the template can degrade the recognition performance. These reasons are as follows. All the samples that are deformed towards the template lie inside the class boundary in the feature space and some may describe the boundary in more detail. On the other hand, some of the samples deformed away from the template may cross the true class boundary and so lead to recognition error. That is, a generated sample that contributes to the correct recognition of original samples does not always contribute to the recognition of test samples. Similarly, a generated sample that causes error in the recognition of original samples does not always cause the erroneous recognition of test samples. We need to improve the method of selecting samples to be used in training.

Also, we can see from Figure 5 that slight levels of distortion can enhance the recognition performance. Unfortunately, we note that there are limits on the number of samples that can be generated and the range of deformation that is useful. To overcome these limits, we should study iterative generation methods or a method to obtain multiple templates to yield a wider range of deformation.

4 Conclusion

We have introduced a method of generating new samples from original training samples to enhance character recognition performance. Our proposal method is based on point correspondence using the distance between the pixels of a pattern and those of the corresponding template; it can generate new samples from off-line characters automatically. Generated samples are seen as being naturally deformed and reflect the handwriting deformation derived from the original samples. Initial recognition experiments using the dictionary conducted by adding generated samples to the original ones show that the addition of generated samples is effective in improving the recognition rate.

More work is needed to find a more effective method of selecting generated samples to be used in training, a method of creating multiple templates for deforming patterns in more ways, and to analyze the relation between the behavior of generated samples in the pattern space and that in the feature space.

Acknowledgments

We would like to thank Dr. Kazuhiko Yamamori of Media Processing Project Manager for his encouragement and support.

References

1. H. S. Baird, "Document image defect models and their uses," In *Proc. of the 2nd ICDAR*, pp. 662-667, 1993.
2. T. K. Ho and H. S. Baird, "Large-Scale Simulation Studies in Image Pattern Recognition," *IEEE Trans. Pattern Analysis & Machine Intelligence*, vol. 19, no. 10, pp. 1,067-1,079, 1997.
3. I. T. Phillips, "How to Extend and Bootstrap an Existing Data Set with Real-life Degraded Images," In *Proc. of 5th ICDAR*, pp. 689-692, 1999.
4. K. Ishii, "Generation of Distorted Characters and Its Applications," *Systems, Computers and Controls*, vol. 14, no. 6, pp. 19-27, 1983.

5. K. Ishii, "Design of a Recognition Dictionary using Artificially Distorted Characters," *Systems and Computers in Japan*, vol. 21, no. 9, pp. 35-45, 1990.
6. N. Hagita, M. Sawaki, and K. Ishii, "A Model for Selectively Increasing Learning Sample Number in Character Recognition," In *Proc. of IS&T/SPIE Document Recognition III*, vol. 2660 pp. 235-242, 1996.
7. M. Shiono, "A Method for Generation of Handwritten Style Character Pattern Using Nonlinear Geometrical Distortion," *IEICE Trans.*, vol. J74-D-II, no. 2, pp. 209-219, 1991. (In Japanese)
8. R. G. Casey, "Moment normalization of handprinted characters," *IBM Journal R&D*, vol. 14, no. 5, p. 548, 1970.
9. T. M. Cover and P. E. Hart, "Nearest Neighbor Pattern Classification," *IEEE Trans. Information Theory*, vol. 13, no. 1, pp. 21-27, 1967.
10. F. Kimura, K. Takashima, S. Tsuruoka, and Y. Miyake, "Modified Quadratic Discriminant Functions and the Application to Chinese Character Recognition," *IEEE Trans. Pattern Analysis & Machine Intelligence*, vol. 9, no. 1, pp. 149-153, 1987.
11. T. Wakahara, A. Suzuki, N. Nakajima, S. Miyahara, and K. Odaka, "On-Line Cursive Kanji Character Recognition as Stroke Correspondence Problem," In *Proc. of 3rd ICDAR*, pp. 1,059-1,064, 1995.
12. S. Kurakake and K. Ishii, "Evaluation of Handprinted Character Recognition Systems Using Distorted Characters," *IEICE Trans.*, vol. J74-D-II, no. 9, pp. 1,208-1,216, 1991. (In Japanese)