2011 International Conference on Document Analysis and Recognition

Effects of Generating a Large Amount of Artificial Patterns for On-line Handwritten Japanese Character Recognition

Bin Chen, Bilan Zhu, Masaki Nakagawa
Department of Computer and Information Sciences
Tokyo University of Agriculture and Technology
Tokyo, Japan
e-mail: 5001034305@st.tuat.ac.jp

Abstract— **This paper describes effects of a large amount of artificial patterns to train an on-line handwritten Japanese character recognizer. In general, the more learning patterns employed for training pattern recognition systems, the higher recognition rate is obtained. In reality, however, the existing pattern samples are not enough, especially for languages of a large character set. Therefore, for on-line handwritten Japanese character recognition, we construct six linear distortion models and combine them with a nonlinear distortion model to generate a large amount of artificial patterns. We apply the method for the TUAT Nakayosi database and train a recognizer while evaluate the effects for the TUAT Kuchibue database with the remarkable effects of improving recognition accuracy.**

Keywords-online handwriting recognition; artificial patterns; linear distortion models; Nonlinear distortion model;

I. INTRODUCTION

Researches on on-line handwritten Japanese character recognition have been all the while pursuing high recognition performance which can be accepted by users of real applications. No matter which method we select, a great number of training samples are generally demanded to obtain a better parameter set of a recognizer due to the fact that there are various writing habits and styles, though it is time consuming and costly to prepare a large database. Namely, samples for recognizer training are limited.

To solve this problem, several works are proposed by transforming character patterns and produce artificial patterns. Ha et al [1] deform a character pattern image according to deformation models to produce pattern variations for off-line handwritten numeral recognition. Leung et al [2] generate a huge number of virtual training samples artificially for off-line handwritten Chinese characters recognition which demonstrates remarkable effectiveness.

In this paper, we extend the method of character shape transformation and use an effective linear distortion models (LDMs) to generate a great deal of artificial patterns, with which we train a handwritten Japanese character recognizer and combine them with nonlinear distortion model (NLDM). We compare the results with the recognizer trained without artificial characters. Experimental results on the TUAT Kuchibue and Nakayosi databases [3][4] demonstrate the superiority of our proposed method.

The rest of this paper is organized as follows: Section 2 describes an overview of our proposed method. Section 3 details the six LDMs and combinations of LDMs with NLDM. Section 4 presents the experimental results. Section 5 draws our concluding remarks.

II. OVERVIEW

Our approach is based on the modeling of the rules of writing characters beautifully and human writing habits.

One key idea of the transformation approach is to conform to the rules of writing characters beautifully. As far as calligraphy is concerned, characters should be written by following several types of transformation. Fig. 1 shows examples of transformation and the print types.

Fig.1 Handwritten styles (upper) and printed styles (lower)

In addition, we can get inspiration from casual writing habits as shown in Fig 2. Some people fail to write characters beautifully because of their habits. In fact, we often face with these transformations in daily life.

Fig.2 Real handwritten samples

To simulate these styles of transformations, we construct 6 LDMs. In the following section we will detail the proposed LDMs.

III. DISTORTION MODELS

A. 6 Linear Distortion Models (LDMs)

The trajectories of on-line handwritten character patterns are distorted with the six models included four basic and two combined models. Basic models are composed of rotation, shear, shrink and perspective; combined models consist of shrink with rotation and perspective with rotation. The six LDMs are given in the following.

The rotation model is given in Eq.1:

$$\begin{cases} x' = x\cos\theta - y\sin\theta \\ y' = x\sin\theta + y\cos\theta \end{cases} \tag{1}$$

where (x', y') indicates the new coordinate transformed by the rotation model. θ denotes the angle of rotation.

The shear model can be divided into two types in X-direction and in Y-direction according to the direction of shearing. Shear model in X-direction and that in Y-direction are given in Eq.2 and Eq.3:

$$\begin{cases} x' = x + y\tan\theta \\ y' = y \end{cases} \tag{2}$$

$$\begin{cases} x' = x \\ y' = x\tan\theta + y \end{cases} \tag{3}$$

where (x', y') is the new coordinate transformed by the shear model. θ denotes the angle of shear .

The shrink model is also divided into two types in X-direction and in Y-direction according to the direction of shrinking. Shrink model in X-direction and that in Y-direction are shown in Eq.4 and Eq.5:

$$\begin{cases} x' = x \\ y' = y\left(\sin(\pi/2 - \theta) - \left((x\sin\theta)/100\right)\right) \end{cases} \tag{4}$$

$$\begin{cases} x' = x\left(\sin(\pi/2 - \theta) - \left((y\sin\theta)/100\right)\right) \\ y' = y \end{cases} \tag{5}$$

where (x', y') denotes the new coordinate produced by the shrink model. θ is indicator of changing degree.

The definition of the perspective model is similar to the shrink model, except the transformation function of the perspective model. The perspective model in X-direction and that in Y-direction are given in the Eq.6 and Eq.7:

$$\begin{cases} x' = 2/3\ \left(x + 50\cos\left(4\theta * \left((x - 50)/100\right)\right)\right) \\ y' = 2/3\, y\left(\sin(\pi/2 - \theta) - \left((x\sin\theta)/100\right)\right) \end{cases} \tag{6}$$

$$\begin{cases} x' = 2/3x\left(\sin(\pi/2 - \theta) - \left((x\sin\theta)/100\right)\right) \\ y' = 2/3\left(y + 50\cos\left(4\theta * \left((y - 50)/100\right)\right)\right) \end{cases} \tag{7}$$

where (x', y') indicates the new coordinate transformed by the perspective model. θ denotes an indicator of changing degree.

The shrink with rotation model consists of two phases. First we do the shrink transformation according to the shrink model; further, we realize the rotation transformation based on the result produced from the first phase. Because of the same process with shrink, the perspective with rotation model is not described here.

The trajectories of training samples are distorted with the above functions, which greatly enlarges the size of the training set.

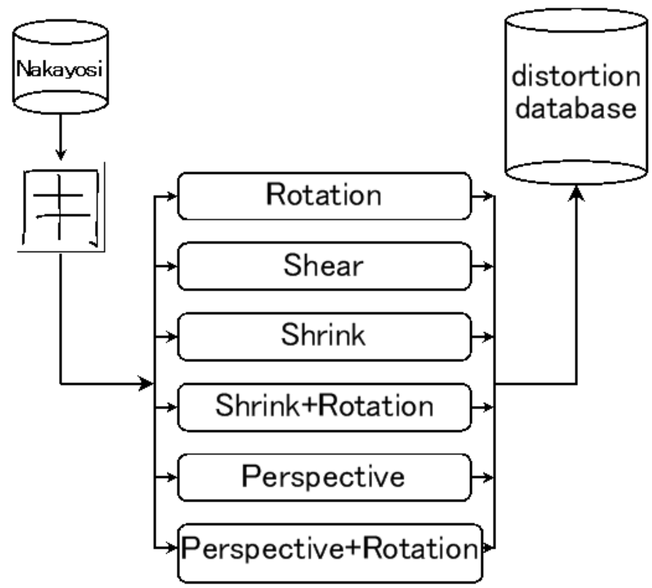

Fig.3 Flowchart of 6 LDMs transformation process

For the rotation model and shear model, θ denotes the rotation angle and the shear angle, respectively. For the shrink model and perspective model, θ is an indicator of changing degree. The flowchart of the transformation process is shown in Fig 3.

With the above-mentioned six LDMs, the number of pattern samples in the database is enlarged from 20 to 160 times. For the rotation, we obtain artificial samples 20 times and 40 times by changing θ from -10 degree to 10 degree within every 1 degree step and within every 0.5 degree step, respectively. Similarly, for the shear in X-direction we can also enlarge the database 20 times and 40 times by changing the angle θ from -10 degree to 10 degree within every 1 degree step and within every 0.5 degree step, respectively. For the shear in Y-direction, we apply the same method to enlarge database 20 times and 40 times. So, we can totally enlarge the database 40 times and 80 times by the shear model. Like the shear model, both the shrink and perspective models can also extend the database 40 times and 80 times by changing θ from -10 degree to 10 degree within every 1 degree step and within every 0.5 degree step, respectively. For the shrink with rotation model, we change θ from -10 degree to 10 degree within every 1 degree step and within every 0.5 degree step, and for each θ we transform each sample by the shrink model firstly, and then rotate it by two rotation degrees θ and -θ. Hence we totally enlarge the database 80 times and 160 times. The process of perspective with rotation is the same as the shrink with rotation model and we obtain the same amount of enlargement.

B. Lueng's Non-Leaner Distortion Model (NLDM)

We also realized the Non-linear distortion model (NLDM) defined by Lueng et al. [2]. We replace 6 LDMs in the transformation process by Lueng's NLDM as shown in Fig.4

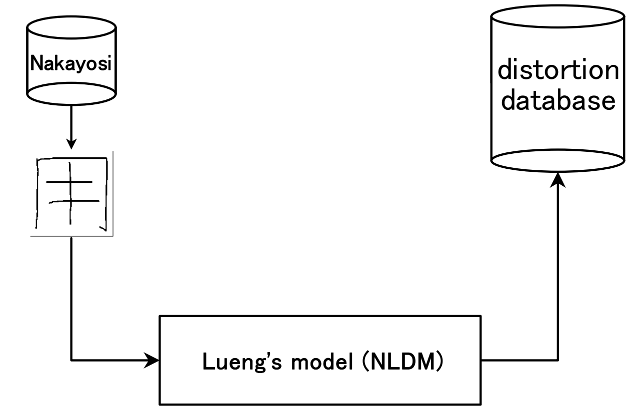

Fig.4 Flowchart of Lueng's NLDM transformation process

C. Combined Distortion Model

The above mentioned For 6 LDMs and NLDM both of the two distortion models enlarged learning patterns original data effectively, we try to explore further effect by combining LDMs and NLDMa better performance form the two distortion models. In this paper we add a NLDM process after applying our 6 LDMs. In another word, we apply 6 LDMs, to get all distorted patterns. Second, we apply NLDM for every distorted sample to get a further distorted sample. In this way, we get a set of further distorted samples. It is showed in Fig.5.

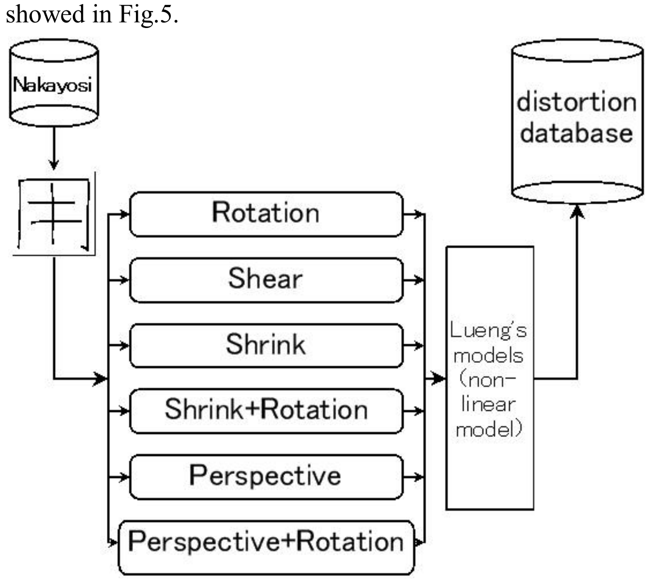

Fig.5 Flowchart of Component Model transformation process

IV. EXPERIMENTS

To compare the performance of recognizer training with and without artificial patterns, we experiment on the TUAT HANDS Nakayosi and Kuchibue databases of online handwritten Japanese characters [4]. The Kuchibue database contains the handwritten samples of 120 writers, 11,962 patterns per writer covering 3,356 character classes. There are 11,951 patterns for 3,345 classes per writer. The Nakayosi database contains the samples of 163 writers, 10,403 patterns covering 4,438 classes per writer [3][4]. Summary of the databases are shown in Table 1.

Table.1 Summary of Kuchibue and Nakayosi

	Kuchibue	Nakayosi
Number of writers	120	163
Number of classes	3,356	4,438
Number of patterns	1,435,440	1,695,689

The Nakayosi database is used to generate artificial patterns and to train the character recognizer, and the Kuchibue database is used for testing. The character recognizer adopts a linear-chain MRF model with weighting parameters optimized by CRFs to recognize character patterns [5]. The flow chart of the online recognizer is shown in Fig.6.

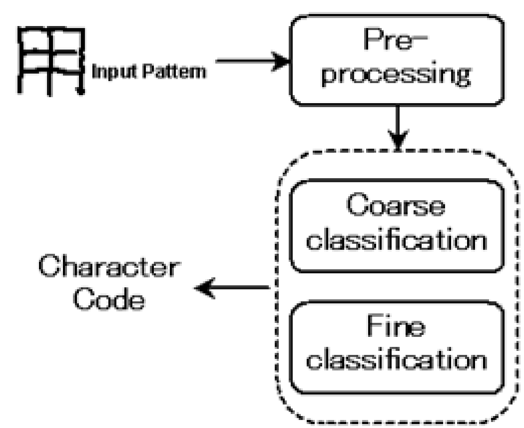

Fig.6 Flowchart of online recognizer

The Japanese character set consists of different types of characters: symbols, numerals, upper case English, lower case English, upper case Greek, lower case Greek, hiragana, katakana and Kanji characters of Chinese origin. With the above 9 character sub groups, we get the overall performance on the Kuchibue database.

In order to describe this experiment, we use a transformation function with three variables DM(m, t, Δ) to show the transformation method. The first parameter m is the ID of a distortion model, the second parameter t is the effect of enlargement, and the third parameter Δ is interval degree.

For example, DM(1, 20, 1) denotes a transformation method that applies the rotation model to enlarge the Nakayosi database to 20 times by changing θ from -10 degree to 10 degree within every 1 degree step.

For parameter Δ we try two values 0.5 and 1. Therefore, we get two groups of θ.

$\theta = \pm1,\pm2,\ldots,\pm10.(\ \Delta=1).$

$\theta = \pm0.5,\pm1,\ldots,\pm9.5,\pm10.(\ \Delta=0.5)$

Consequently, we test the twelve transformation methods as shown in Table 2.

Table.2 Transformation methods

Model mark	Model info	Enlarge ratio	Interval degree
DM(1,20,1)	Rotation	20.00	1.00
DM(1,40,0.5)	Rotation	40.00	0.50
DM(2,40,1)	Shear	40.00	1.00
DM(2,80,0.5)	Shear	80.00	0.50
DM(3,40,1)	Shrink	40.00	1.00
DM(3,80,0.5)	Shrink	80.00	0.50
DM(4,40,1)	Perspective	40.00	1.00
DM(4,80,0.5)	Perspective	80.00	0.50
DM(5,80,1)	Shrink+Rotation	80.00	1.00
DM(5,160,0.5)	Shrink+Rotation	160.00	0.50
DM(6,80,1)	Perspective+Rotation	80.00	1.00
DM(6,160,0.5)	Perspective+Rotation	160.00	0.50

Fig.7 illustrates some transformations of a character pattern with maximum distortion degree.

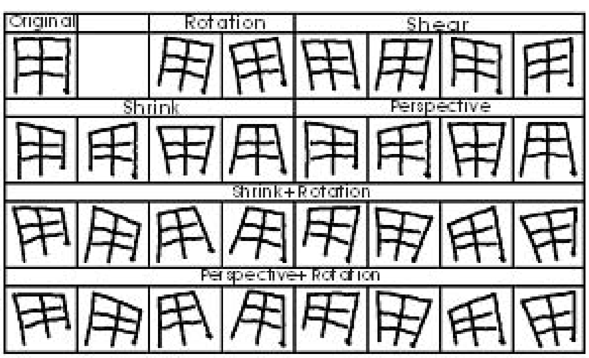

Fig.7 Illustration of 6 LDMs

NLDM defined by Lueng et al. is achieved via tow mapping functions [2]. It considers both shearing and local resizing. We generate artificial patterns with 3 kinds of enlargement ratio, 40 times, 80 times and 160 times. It is marked with A, B, C. Fig.8 illustrates some random transformations of NLDM.

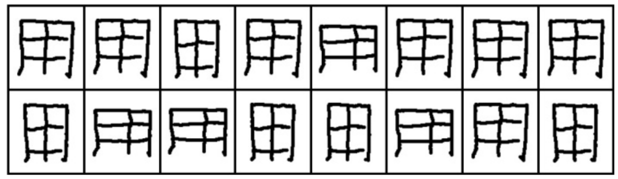

Fig.8 Illustration of nonlinear transformations

We combine LDMs and NLDM. Considering that skew is modeled in our 6 LDMs, we realize NLDM without skew distortion.

Fig.9 illustrates some typical transformations of a character pattern by NLDM, without skew distortion.

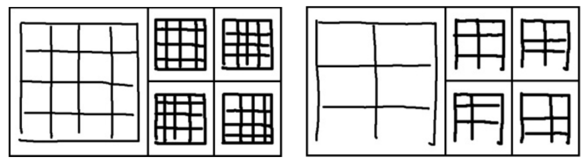

Fig.9 Illustration of Nonlinear model, without skew distortion.

In training the recognizer, the samples were processed iteratively for ten cycles. The system was implemented in MS Visual C++ 2008 with 64 bit compiler, on 64bit Windows7 and tested on two PCs. One is Intel Dual Core2 Quad Q9550 CPU with 6GB memory, and another is Intel Xeon W5590 CPU with 24GB memory. The recognition results are shown in Table 3. We can see that the recognition accuracy has been improved by training the recognizer by diversified patterns compared to the baseline recognizer without using them. It shows also that enlarging the database simply by change interval degree cannot produce higher recognition rate.

Table.3 Recognition rate for each character sub-group, tested on LDMs

	Symbol	Numeral	Upper English	Lower English	Upper greek	Lower greek	Hira gana	Kata kana	Kanji
Original	71.73	98.20	96.92	94.23	79.75	91.33	85.39	83.01	94.84
DM(1,20,1)	73.05	98.30	97.62	94.65	84.75	92.71	85.77	83.50	95.64
DM(1,40,0.5)	72.95	98.30	97.69	94.81	82.50	92.13	85.81	83.69	95.22
DM(2,40,1)	73.14	98.50	97.54	94.76	83.75	92.38	85.72	83.65	95.65
DM(2,80,0.5)	72.91	98.60	97.65	94.96	83.25	92.42	85.73	83.31	95.59
DM(3,40,1)	73.23	98.60	97.88	94.81	83.00	92.79	85.54	83.22	95.55
DM(3,80,0.5)	72.74	98.50	97.81	94.62	83.75	92.79	85.55	83.10	95.49
DM(4,40,1)	73.41	98.30	97.73	95.19	84.25	92.79	85.96	83.86	95.72
DM(4,80,0.5)	73.01	98.10	97.73	94.88	84.50	92.83	85.71	83.57	95.66
DM(5,80,1)	73.74	98.10	97.19	95.07	84.75	93.42	85.18	83.43	95.75
DM(5,160,0.5)	73.83	98.50	97.35	94.92	85.25	93.08	85.92	83.52	95.70
DM(5,80,1)	74.26	98.20	97.42	95.12	85.75	93.63	85.86	83.59	95.87
DM(5,160,0.5)	74.03	98.40	97.58	95.00	86.50	93.33	86.17	83.83	95.81

From our experiments, The most important result is the recognition accuracy is improved from 94.84% to 95.87% for Kanji characters in Kuchibue database, where we applied the perspective model, and enlarged original database to 80 times.

Table.4 Recognition rate for each character sub-group, tested on NLDM

	Symbol	Numeral	Upper English	Lower English	Upper greek	Lower greek	Hira gana	Kata kana	Kanji
Original	71.73	98.20	96.92	94.23	79.75	91.33	85.39	83.01	94.84
NLDM A	71.82	98.60	97.27	94.58	83.25	93.04	85.40	82.81	95.61
NLDM B	72.82	98.50	97.73	94.35	83.00	93.00	85.28	82.65	95.71
NLDM C	73.54	98.40	97.88	94.73	83.25	93.54	85.57	82.70	95.82

The results of non-linear distortion model show that, in a certain range, the more artificial patterns are generated, the higher recognition rate is achieved.

Table.5 Recognition rate for each character sub-group, tested on combined distortion models

	Symbol	Numeral	Upper English	Lower English	Upper greek	Lower greek	Hira gana	Kata kana	Kanji
Original	71.73	98.20	96.92	94.23	79.75	91.33	85.39	83.01	94.84
6 LDMs	74.92	98.40	97.62	94.88	79.00	93.50	85.13	84.02	95.79
combined	73.18	98.40	97.58	94.38	78.75	93.00	85.13	83.73	95.94

In this experiment, first we apply all of the 6 LDMs, to get all distorted patterns. Second, we apply NLDM for every distorted sample to get a further distorted sample. In this way, we get a set of further distorted samples. Recognition rate is showed in Table.5.

V. CONCLUSION

We have presented an effective approach to enhance the accuracy of online handwriting Japanese recognition by using transformed training samples generated by 6 linear distortion models and combination with non-linear distortion model. With experiments on 9 character sub groups of Kuchibue database, the recognition accuracies are improved for most of the sub groups and models, which demonstrates the effectiveness of our approach. With this approach we have improved recognition accuracies without taking more time in recognition process.

Our six LDMs has improved the recognition rate to all the character groups and achieved 95.87% recognition rate for Kanji. We also tested NLDM defined by Lueng. It showed that, in a certain range, the more artificial patterns are generated, the higher recognition rate is achieves. Moreover, the combined model achieved 95.94% recognition rate for Kanji.

Acknowledgements

This work is being supported by the R&D fund for “Development of pen & paper based user interaction” under Japan Science and Technology Agency.

REFERENCES

[1] T. M. Ha and H. Bunke, “Off-line, Handwritten Numeral Recognition by Perturbation Method, IEEE Transactions on Pattern Analysis and Machine Intelligence”, vol. 19, no. 5, pp. 535-539, 1997.

[2] K.C.Leung and C.H.Leung, “Recognition of Handwritten Chinese Characters by Combining Regularization, Fisher’s Discriminant and Transformation Sample Generation. In 10th International Conference of Document Analysis and Recognition” (ICDAR), pp.1026-1030, 2009.

[3] S. Jaeger and M. Nakagawa, “Two On-Line Japanese Character Databases in Unipen Format. In Sixth International Conference on Document Analysis and Recognition” (ICDAR), Seattle, USA, pp. 566-570, 2001.

[4] M. Nakagawa and K. Matsumoto, “Collection of on-line handwritten Japanese character pattern databases and their analyses”, IJDAR 2004 vol.7,pp. 69–81

[5] B. Zhu and M. Nakagawa, “A MRF Model with parameters optimization by CRF for on-line recognition of handwritten Japanese characters”, to appear in Proc. Document Recognition and Retrieval XVIII (DRR), USA, January 2011