Pattern Recognition Letters 35 (2014) 178–185

Contents lists available at SciVerse ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Training of an on-line handwritten Japanese character recognizer by artificial patterns

Bin Chen *, Bilan Zhu, Masaki Nakagawa

Department of Computer and Information Sciences, Tokyo University of Agriculture and Technology, Tokyo, Japan

ARTICLE INFO

Article history:
Available online 27 July 2012

Keywords:
On-line handwriting recognition
Artificial patterns
Linear distortion models
Non-linear distortion model
Combined model
Pattern selection strategy

ABSTRACT

This paper presents effects of a large amount of training patterns artificially generated to train an on-line handwritten Japanese character recognizer, which is based on the Markov Random Field model. In general, the more training patterns, the higher the recognition accuracy. In reality, however, the existing pattern samples are not enough, especially for languages with large sets of characters, for which a higher number of parameters needs to be adjusted. We use six types of linear distortion models and combine them among themselves and with a non-linear distortion model to generate a large amount of artificial patterns. These models are based on several geometry transform models, which are considered to simulate distortions in real handwriting. We apply these models to the TUAT Nakayosi database and expand its volume by up to 300 times while evaluating the notable effect of the TUAT Kuchibue database for improving recognition accuracy. The effect is analyzed for subgroups in the character set and a significant effect is observed for *Kanji*, ideographic characters of Chinese origin. This paper also considers the order of linear and non-linear distortion models and the strategy to select patterns in the original database from patterns close to character class models to those away from them or vice versa. For this consideration, we merge the Nakayosi and Kuchibue databases. We take 100 patterns existed in the merged database to form the testing set, while the remaining samples to form the training set. For the order, linear then non-linear distortions produce higher recognition accuracy. For the strategy, selecting patterns away from character class models to those close to them produce higher accuracy.

1. Introduction

Research on on-line handwritten Japanese character recognition has pursued recognition accuracy high enough to be accepted by users of real applications (Plamondon and Srihari, 2000; Liu et al., 2004). To deal with the problem that character patterns are often distorted, there are four main methods. One is to decrease distortion by non-linear normalization (Yamada et al., 1984; Tsukumo and Tanaka, 1988; Liu et al., 2003) or try to remove distortion by reverse distortion in a normalization step (Wakahara and Odaka, 1996; Satoh et al., 1999). The second is to improve discriminant functions such as MQDF for off-line recognition (Kimura et al., 1987), HMM (Jaeger et al., 2001) or MRF for on-line recognition (Zhu and Nakagawa, 2011). The third is to select or extract stable features (Liu and Zhou, 2006). The fourth is to train classifiers by an increased amount of training patterns (Smith et al., 1994). To collect training patterns is very costly, however, so that artificial pattern generation has been used (Ha and Bunke, 1997; Mori et al., 2000; Leung et al., 1985; Leung and Leung, 2009).

This paper focuses on artificial pattern generation. In general, the more training patterns are employed, the higher the recognition accuracy is achieved. In reality, however, the existing pattern samples are not enough, especially for languages with large sets of characters, for which a higher number of parameters need to be adjusted. Thus, we consider artificial pattern generation. Several works have been proposed to transform character patterns in accordance with some models and produce artificial patterns. Ha and Bunke (1997) used the concept of perturbation due to writing habits and instruments for off-line handwritten numeral recognition, where they proposed six types of linear distortion models to reverse an input image back to its standard form to solve the problem of patterns variation. Mori et al. (2000) proposed a character pattern generation method based on point correspondence between patterns. Leung et al. (1985) and Leung and Leung (2009) generated a huge number of training samples artificially in accordance with a non-linear distortion model for off-line handwritten Chinese characters recognition, which demonstrates that applying distorted sample generation is effective in addition to regularization of class covariance matrices and feature dimension reduction, when the dimension of the feature vector is high while the number of training samples is not sufficient. Velek et al. (2002) proposed a method to generate brush-written off-line patterns from on-line

* Corresponding author.
E-mail addresses: 50010834305@st.tuat.ac.jp, hertz158123@gmail.com (B. Chen), zhubilan@cc.tuat.ac.jp (B. Zhu), nakagawa@cc.tuat.ac.jp (M. Nakagawa).

0167-8655/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.patrec.2012.07.012

patterns. Postal address recognition had problems reading characters written with a traditional brush for new year cards, since the amount of training patterns was limited for such patterns.

In this paper, we consider on-line pattern generation for on-line handwritten Japanese character recognition. We propose six types of linear distortion models (LDMs) as proposed by Ha and Bunke (1997) and use them to generate a great deal of artificial patterns, with which we train a handwritten Japanese character recognizer. Then, we combine LDMs with non-linear distortion model (NLDM) proposed by Leung et al. (1985) and Leung and Leung (2009) to obtain combined distortion models (CDMs) and generate artificial patterns again to train the above recognizer.

Here it is worth noting that the basic LDMs proposed by Ha and Bunke (1997) were applied in preprocessing to reverse an input image back to its standard form; they were applied to just numerical patterns; and they were employed in recognition stage so that additional recognition time was incurred. On the other hand, we employed them for pattern generation so that the recognition time is not affected. Moreover, Leung et al. (1985) and Leung and Leung (2009) proposed the non-linear distortion model for off-line Chinese character recognition while we combined them with LDMs for on-line recognition.

This paper is an extension to the conference papers (Chen et al., 2010, 2011), which reported the increase of recognition rate by employing the proposed method to generate artificial patterns for training. This paper shows them in more detail and considers effects of selecting the combination sequence for CDMs and original pattern selection strategy. There are two combination sequences: LDMs then NLDM and NLDM then LDMs. Moreover, there are two original pattern selection strategies: selecting patterns in the original database, from patterns close to character class models to those away from them and vice versa. These two combination sequences and two original pattern selection strategies are combined pairwisely. For this consideration, we merge the Nakayosi and Kuchibue databases, and take 100 patterns in the merged database to form the testing set, while the remaining samples to form the training set. Moreover, we also attempt to find a generating method with relatively less real patterns employed while increasing recognition accuracy efficiently. The detailed performance evaluations and discussions will be presented that show the effectiveness of the proposed method.

The rest of this paper is organized as follows: Section 2 describes basic ideas of our proposed method. Section 3 briefly describes databases, pattern transformation. Section 4 introduces our recognition classifier that we used. Section 5 details 12 LDMs, NLDM, and CDMs and experimental results for increasing their recognition accuracy. Section 6 presents experiments on the two combination sequences and two original pattern selection strategies. Section 7 describes the results and analysis. Section 6 draws our concluding remarks.

2. Basic ideas

Our approach to generating artificial patterns is based on the observation of how people write and deform character patterns. First, people try to write characters beautifully in accordance with the rules of calligraphy. As far as calligraphy is concerned, characters should be written by following several types of distortion, different with printed type. Fig. 1(a) shows calligraphy styles corresponding to printed types. Samples of shear along the X-direction and Y-direction and shrink toward four directions are shown.

In real handwriting, some people fail to write characters neatly because of their habits as shown Fig. 1(c). Distortions include not only shear and shrink shown in Fig. 1(c) but also perspective. Its models are similar to the shrink models, but they are different in

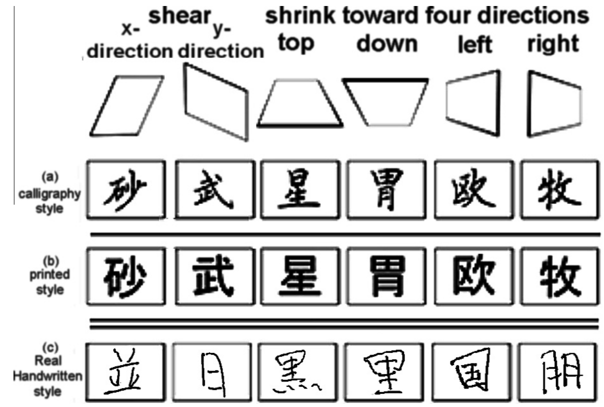

Fig. 1. Handwriting styles based on calligraphy (a), printed (b) and real handwritten styles (c).

keeping the balance between left and right, or up and down as shown in Fig. 2(a), so that the center of a character pattern is shifted toward the narrowing direction in perspective models while it remains in shrink models. In fact, we often face these distortions in daily life. These distortions can be modeled by the shear model, shrink models, perspective models, and their combinations.

Not just these, however, since some people deform the balance between radicals as shown in Fig. 2(b), which is different from the perspective distortion, since shrunk radicals, keep either their height or width. Specifically, people deform patterns non-linearly, which is why non-linear normalization works. Moreover, configuration of the writer's arm, hand, pen, and paper may produce rotated patterns as shown in Fig. 2(c). This type of distortion is modeled by rotation.

Patterns generated from these models are expected to simulate real handwritten samples. In the following sections we will detail the distortion models.

3. Database, and pattern transformation

3.1. Databases

An on-line handwritten character pattern is composed of a sequence of strokes and each stroke is composed of a time-sequence of coordinates sampled from a tablet or touch sensitive device. TUAT HANDS Nakayosi and Kuchibue databases of on-line handwritten Japanese characters patterns (Nakagawa and Matsumoto, 2004) are applied in this experiment. The Kuchibue database contains the patterns of 120 writers: 11,962 patterns per writer covering 3356 categories. Excluding the JIS level-2 Kanji characters, there are 11,951 patterns for 3345 JIS level-1 categories (including 2965 Kanji characters and 380 non-Kanji symbols), which are frequently used in recognition experiments. The Nakayosi database contains the samples of 163 writers, 10403 patterns covering 4438 classes adding frequently used JIS level-2 categories per writer (Nakagawa and Matsumoto, 2004; Jaeger and Nakagawa, 2001).

3.2. Pattern transformation

LDM proposed by Ha and Bunke (1997) and NLDM proposed by Leung et al. (1985) and Leung and Leung (2009) have been applied to transform off-line patterns where each black pixel is moved according to the LDM or NLDM. We can also apply them to transform our on-line patterns. Fig. 3(a) shows an example for pattern transformation. An original on-line pattern has a sequence of coordinates $\{(x_1,y_1), (x_2,y_2), \ldots, (x_i,y_i), \ldots, (x_p,y_p)\}$. We apply LDM, NLDM or CDM to move each coordinate (x_i,y_i) to a new coordinate (x_i^1, y_i^1)

Fig. 2. Handwriting styles based on shrink, perspective, non-linear distortion and rotation.

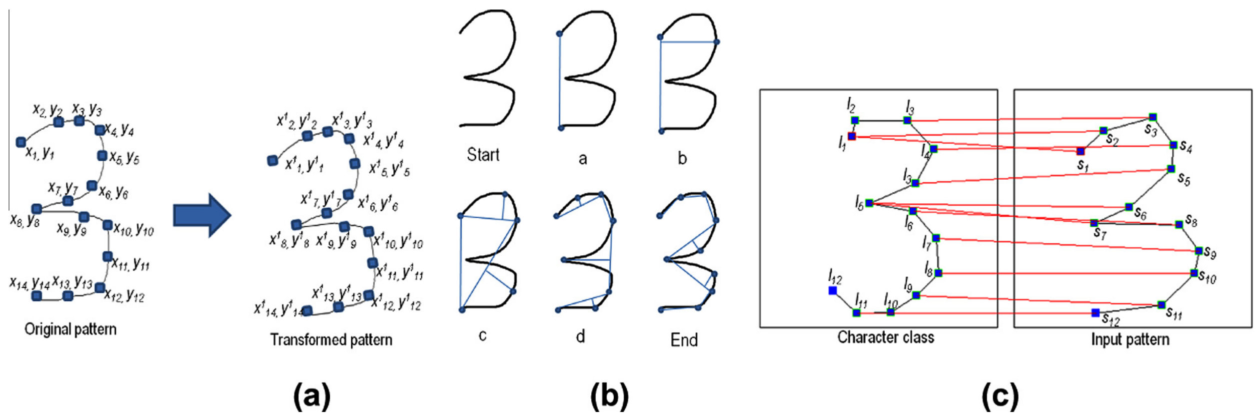

Fig. 3. Pattern transformation, feature points extraction and labeling.

so as to generate a new pattern with a sequence of coordinates $\{(x_1^1,y_1^1),(x_2^1,y_2^1),\ldots,(x_i^1,y_i^1),\ldots,(x_p^1,y_p^1)\}$. At the microscopic level, our distortion models change point coordinates, distances between points, their directions, and eventually change the shape of a pattern at the macroscopic level. By the transformations, however on-line patterns do not lose the dynamic pen movement information such as stroke trajectories, pen's up and pen's down information, as well as pen's speed, although pen movement information are changed according to positional changes of point coordinates.

4. On-line handwritten character recognition system employed

We adopt a linear-chain Markov random field MRF model with weighting parameters optimized by CRFs to recognize character patterns (Zhu and Nakagawa, 2011). Here we summarize the system. It extracts feature points along the pen-tip trace from pen-down to pen-up and sets each feature point from an input pattern as a site and each state from a character class as a label. It uses the coordinates of feature points as unary features and the differences in coordinates between the neighboring feature points as binary features.

An input pattern is normalized linearly by converting the pen-tip trace pattern to a standard size, preserving the horizontal-vertical ratio. After the normalization, we extract feature points using the method by Ramer (1972). First, the start and end points of every stroke are picked up as feature points. Then, the most distant point from the straight line between adjacent feature points is selected as a feature point if the distance to the straight line is greater than a threshold value. This selection is done recursively until no more feature points are selected. The feature point extracting process is shown in Fig. 3(b). The extracted feature points stand for the structure of a pattern. They are effective and more efficient to process than all the pen-tip points. We set feature points from an input pattern as sites $\boldsymbol{S} = \{s_1, s_2, s_3, \ldots, s_I\}$ and states of a character class C as labels $\boldsymbol{L} = \{l_1, l_2, l_3, \ldots, l_J\}$. The system recognizes the input pattern by assigning labels to the sites to make the matching between the input pattern and each character class C such as $\boldsymbol{F} = \{s_1 = l_1, s_2 = l_1, s_3 = l_3, \ldots, s_{11} = l_9, s_{12} = l_{11}\}$ as shown in Fig. 3(c). More details are available in the work of Zhu and Nakagawa (2011).

5. Distortion models and their evaluation

We use distortion models to generate a large amount of artificial patterns form on-line handwritten samples to train the character recognizer. The Japanese character set consists of different types of characters: symbols, numerals, upper case Roman letters, lower case Roman letters, upper case Greek, lower case Greek, *hiragana*, *katakana*, and *Kanji* characters of Chinese origin. With the above nine character subgroups, we apply these models to the TUAT Nakayosi database and obtain the overall performance on the Kuchibue database in our experiments described in this section. To investigate completely the effect of the distortion model on different structures of characters, we present the effects of different distortion models on different subgroups of characters.

5.1. 6 Linear distortion models (LDMs)

The trajectories of on-line handwritten character patterns are distorted by the six models including four basic and two combined models. Basic models are composed of rotation, shear, shrink and perspective; combined models consist of shrink plus rotation and perspective plus rotation. The six LDMs are given below.

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 after converted by the rotation model. θ denotes the angle of rotation.

Shear is the translation along an axis by an amount that increases linearly with another axis. The shear model can be divided into two types in accordance with the direction of shearing in direction of x-axis or y-axis. Shear models in x-axis and y-axis directions are shown in Eq. (2) and Eq. (3), respectively:

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

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

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

The shrink model and perspective model are both similar to the two types of shear model in two different distortion directions. The two types of shrink model in x-axis and y-axis are described in Eqs. (4) and (5), respectively. Those of the perspective model are shown in Eqs. (6) and (7), respectively. Where the (x', y') and θ denote distorted coordinate and the degree that the character pattern is distorted by the different distortion model, respectively.

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

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

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

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

The combined distortion model is combined in accordance with the sequence of two basic distortions to the character pattern. Therefore, the shrink plus rotation model first uses the shrink model to realize the shrink distortion of a character pattern; second, the rotation distortion is used on the basis of the results generated from the first phase. Because it has the same process as the shrink plus rotation model, the perspective plus rotation model is not described here.

In the above, θ in different distortion models denotes different things. In the rotation and shear models, θ denotes the rotation angle and the shear angle, respectively. For the shrink model and perspective model, θ is an indicator of changing degree. Furthermore, different changing step of θ affects the amount of the distorted character patterns directly. To express clearly, we use a labeling with three variables DM(m,t,Δ) to label each distortion model with different parameters. The first parameter m is the ID of a distortion model, the second parameter t is the number of enlargements after distortion, and the third parameter Δ indicates changing step of θ. There are six distortion models, rotation, shear, shrink, perspective, shrink plus rotation, and perspective plus rotation with IDs from 1 to 6, respectively in LDMs. For example, DM(1,20,1) denotes using the first distortion model i.e., rotation, to enlarge the Nakayosi database up to 20 times by changing θ from $-10°$ to $10°$ within every $1°$ step.

By using the rotation model, we obtain 20 or 40 times as many patterns in a database by changing θ from $-10°$ to $10°$ with every $1°$ step or with every $0.5°$ step and denote the model as DM(1,20,1) or DM(1,40,0.5), respectively. By using the shear model in the x-axis direction, we also enlarge the database 20 or 40 times as in the rotation model; similarly in the y-direction, we also enlarge the database 20 or 40 times. Therefore, we totally enlarge the database 40 times or 80 times in both x-direction and y-direction by the shear model and denote the model as DM(2,40,1) or DM(2,40,0.5), respectively. Like the shear model, both the shrink and perspective models also enlarge the database 40 or 80 times and we denote them as DM(3,40,1) or DM(3,80,0.5), and DM(4,40,1) or DM(4,80,0.5), respectively. As for the shrink plus rotation model, we change θ from $-10°$ to $10°$ with every $1°$ step or with every $0.5°$ step, and for each θ we transform each sample by the shrink model first and then rotate it by two rotation degrees θ and $-\theta$. Hence we totally enlarge the database 80 times or 160 times and denote the model as DM(5,80,1) or DM(5,160,0.5). The process of perspective plus rotation is the same as the shrink plus rotation model, and we enlarge the database 80 or 160 times and denote the model as DM(6,80,1) or DM(6,160,0.5), respectively.

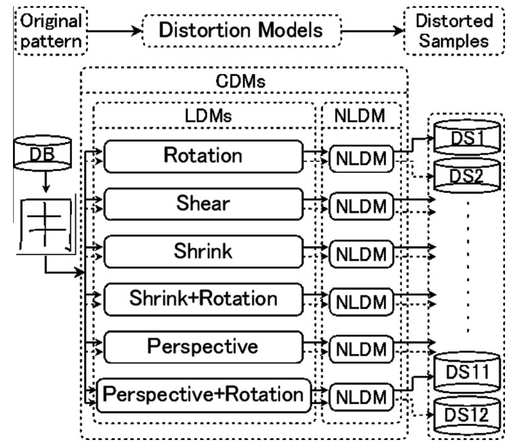

Fig. 4. Process of using distortion models.

We use the 12 LDMs (six types, two different steps) described above to generate a large amount of artificial patterns from patterns in Nakayosi and then use them to train the on-line recognizer. For the evaluation of different LDMs, we use the Kuchibue database as the testing database. Fig. 4 shows the process of using distortion models to generate artificial patterns, solid-line arrows indicate models whose interval degree is set 1, dotted-line arrows indicate models which's interval degree is set 0.5. In this experiment just 12 LDMs are used to generate 12 distorted sample sets. These distorted sample sets are used to train the on-line recognizer, obtain corresponding accuracies of these 12 LDMs. Generated artificial patterns are shown in Fig. 5(a). Table 1 presents the recognition accuracy of the recognizer trained by artificial patterns generated by LDM for each subgroup. We consider the recognition accuracy of the recognizer trained by original patterns in Nakayosi as the baseline.

Compared to the baseline, all the LDMs to all the subgroups except DM(4,80,0.5) to Hiragana improve the accuracy of character recognition. Specifically, the recognition accuracy is improved from 94.84% to 95.87% for Kanji characters when we apply DM(4,80,0.5). Compared 2 LDMs with different interval degree, in Table 1, it also shows that enlarging the database simply by changing the interval degree cannot produce higher recognition accuracy.

5.2. Non-linear distortion models (NLDMs)

There are three main components (NLDM, shear coefficients k_1 and k_2, and constants c_1 and c_2) for horizontal shifting in Leung's NLDM, as presented in Eqs. (8) and (9).

$$\begin{cases} u = h_x(x, y) = w_n(a_1, b_1 9x)) + k_1 y + c_1 \\ v = h_y(x, y) = w_n(a_1, b_1(y)) + k_2 x + c_2 \end{cases} \tag{8}$$

$$\begin{cases} w_1(a, t) = \frac{1-e^{-at}}{1-e^{-a}} \\ w_2(a, t) = \begin{cases} 0.5w_1(a, 2t), & 0 \leqslant t \leqslant 0.5 \\ 0.5 + 0.5w_1(-a.2(t - 0.5)), & 0.5 < t \leqslant 1 \end{cases} \end{cases} \tag{9}$$

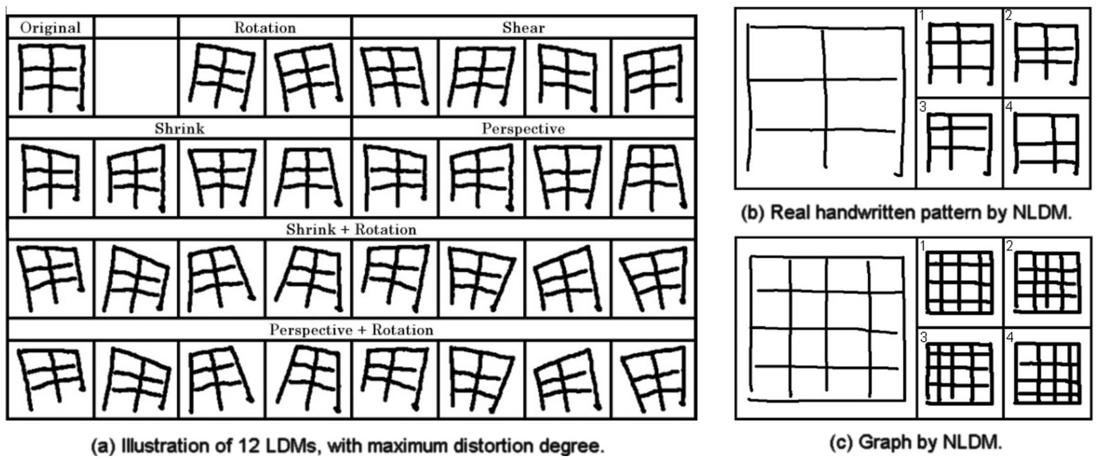

Fig. 5. Illustration of 12 LDMs, with maximum distortion degree.

Table 1
Recognition accuracies of the recognizer trained by artificial patterns generated by LDMs for each subgroup (%).

	Symbol	Numeral	Upper roman	Lower roman	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(6,80,1)	74.26	98.20	97.42	95.12	85.75	93.63	85.86	83.59	95.87
DM(6,160,0.5)	74.03	98.40	97.58	95.00	86.50	93.33	86.17	83.83	95.81
Average	73.36	98.37	97.60	94.90	84.33	92.86	85.74	83.52	95.64

Table 2
Recognition accuracy for each subgroup, trained by artificial patterns distorted by NLDM with shear part (%).

	Symbol	Numeral	Upper roman	Lower roman	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(7,40,1)	71.82	98.60	97.27	94.58	83.25	93.04	85.40	82.81	95.61
DM(7,80,1)	72.82	98.50	97.73	94.35	83.00	93.00	85.28	82.65	95.71
DM(7,160,1)	73.54	98.40	97.88	94.73	83.25	93.54	85.57	82.70	95.82

where $w_n(a_1, b_1(x))$ and $w_n(a_1, b_1(y))$ are warping functions, $b_1(x)$ is a linear function to map x to the range of $[0,1]$, k_1 and k_2 are random slopes of shearing, and c_1 and c_2 are constants such that the center of the character remains unchanged. c_1 and c_2 are not considered here. Two non-linear warping functions w_1 and w_2 are implemented in order to produce more kinds of variations shown in Eq. (9). The choice between w_1 and w_2 is selected randomly with a certain pre-defined probability p. According to Eqs. (8) and (9), each parameter set $\{p, a, t, k_1, k_2\}$ can generate one distorted artificial pattern from one origin pattern in database. Therefore, to enlarge the database 40, 80, or 160 times, we randomly produce 40, 80, or 160 parameter sets for each original pattern in the database, and similar to labeling in Section 5.1, denote the model as DM(7,40,$*$), DM(7,80,$*$) or DM(7,160,$*$), respectively. Because there is no θ, the third parameter in previous labels is set as $*$. Fig. 5(b) shows distorted patterns (right) from the original pattern (left) by NLDM. Fig. 5(c) depicts distortions using a graphical pattern. The distortion (1) expands the center while (2) reduces it. The distortion (3) shifts the pattern to top-left while (4) shifts it to the bottom-right.

To evaluate the NLDM, we use DM(7,40,$*$), DM(7,80,$*$), and DM(7,160,$*$) to generate artificial patterns as the same process as using 12 LDMs, shown in Fig. 4 and train the on-line recognizer, and we present the results in Table 2. Table 2 shows that NLDM effectively improves the recognition accuracy of the MRF recognizer, especially for character subgroups, typified by *Kanji*.

5.3. Combined distortion models(CDMs)

According to the above results, both distortion models LDMs and NLDM improve character recognition accuracy. Naturally, we want to know whether the recognition accuracy is improved further by combining the two distortion models. As mentioned above, there are 12 LDMs. The process of combing each of them to NLDM is the same; therefore, here we just take the rotation model as an instance. As described in Section 5.1, the rotation model includes two types: DM(1,20,1) and DM(1,40,0.5). We combine each type with NLDM and denote them as DM(1-N,20,1) and DM(1-N,

Table 3
Recognition accuracy for each subgroup tested on CDMs applying Leung's NLDM (%).

	Symbol	Numeral	Upper roman	Lower roman	Upper Greek	Lower Greek	Hira gana	Kata kana	Kanji
(a) Recognition accuracy for each subgroup tested on CDMs applying Leung' s NLDM (%)									
Original	71.73	98.20	96.92	94.23	79.75	91.33	85.39	83.01	94.84
DM(1-N,20,1)	70.05	98.00	96.38	92.15	74.75	91.79	83.84	81.40	95.15
DM(1-N,40,0.5)	72.83	98.40	97.65	94.38	79.00	93.08	86.07	83.06	95.87
DM(2-N,40,1)	72.34	98.70	97.46	94.31	77.25	92.97	85.72	83.22	95.90
DM(2-N,80,0.5)	72.70	98.90	97.65	94.46	79.25	92.92	85.71	83.75	95.94
DM(3-N,40,1)	72.50	98.60	97.58	94.31	82.00	92.58	85.78	82.72	95.90
DM(3-N,80,0.5)	72.22	98.60	97.35	94.23	79.00	93.25	85.67	82.94	95.86
DM(4-N,40,1)	71.72	98.70	97.54	94.27	79.75	93.29	85.63	83.13	95.92
DM(4-N,80,0.5)	72.68	98.70	97.46	94.38	78.25	93.38	85.54	83.10	95.90
DM(5-N,80,1)	72.74	98.50	97.46	94.19	80.25	92.92	85.87	83.03	95.93
DM(5-N,160,0.5	72.80	98.70	97.50	94.35	79.00	93.08	85.54	83.10	95.88
DM(6-N,80,1)	72.62	98.60	97.50	94.35	80.50	92.88	85.55	83.30	95.95
DM(6-N,160,0.5)	72.63	98.90	97.88	94.69	79.75	92.96	85.76	83.03	95.90
(b) Recognition accuracy for each subgroup tested on CDMs applying NLDM without shear (%)									
Original	71.73	98.20	96.92	94.23	79.75	91.33	85.39	83.01	94.84
DM(1-N',20,1)	72.78	98.40	97.69	94.23	80.25	93.38	85.49	82.94	95.77
DM(1-N',40,0.5)	72.80	98.60	97.73	94.58	79.25	92.92	85.28	82.85	95.86
DM(2-N',40,1)	72.49	98.50	97.42	94.88	79.75	92.46	85.63	83.42	95.89
DM(2-N',80,0.5)	73.12	98.60	97.31	94.15	78.00	93.13	85.72	83.33	95.93
DM(3-N',40,1)	72.73	98.50	97.50	94.58	79.75	92.79	85.45	83.08	95.88
DM(3-N',80,0.5)	72.70	98.30	97.88	94.38	79.75	93.29	85.52	83.36	95.92
DM(4-N',40,1)	72.90	98.40	97.54	94.69	82.00	93.04	85.78	83.35	95.92
DM(4-N',80,0.5)	72.84	98.40	97.54	94.42	79.50	92.54	85.80	83.49	95.95
DM(5-N',80,1)	72.76	98.30	97.65	94.62	80.00	92.96	85.67	83.43	95.92
DM(5-N',160,0.5)	73.29	98.20	97.50	94.65	79.25	93.29	85.57	83.37	95.92
DM(6-N',80,1)	72.90	98.40	97.81	94.62	78.50	93.38	85.80	83.05	95.95
DM(6-N',160,0.5)	73.20	98.80	97.81	94.46	79.50	93.00	85.82	83.28	95.94

40,0.5) as shown in Fig. 4. It is necessary to address the NLDM model in this situation that distorts just once each of the distorted patterns produced from the previous distortion model. As mentioned above, k_1 and k_2 are random slopes of shearing, and c_1 and c_2 are not considered here. Similarly, distortion models after combinations of the other 10 LDMs with NLDM are denoted as DM(2-N,40,1) and DM(2-N,80,0.5), DM(3-N,40,1) and DM(3-N, 80,0.5), DM(4-N,40,1) and DM(4-N,80,0.5), DM(5-N,80,1) and DM(5-N,160,0.5), and DM(6-N,80,1) and DM(6-N,160,0.5). Similar to previous experiments, we produce artificial patterns by using the CDMs to train the on-line recognizer and present the recognition accuracy for each subgroup in Table 3(a).

Since Leung's NLDM contains both pure NLDM and shear distortion, shear distortion will be executed twice in combining shear distortion LDMs with NLDM. For this reason, we remove the shear part in Leung's NLDM and make similar experiments. The results are shown in Table 3(b). To draw a distinction from CDMs using NLDM with a shear part, we denote CDMs without a shear part by changing *N* into *N'* in their labels.

Compared with the 12 LDMs shown in Table 1 by average recognition accuracy, the combined distortion models with shear parts produce higher recognition accuracy for *Kanji*. In general, even higher recognition accuracy is achieved by CDMs without a shear part.

To generate more artificial patterns for training the on-line recognizer, we enlarge the database by 300 times by three approaches. As showing in Fig. 4, processes alone solid-line arrows are used in this experiment. In the first approach, each pattern in database is distorted by CDMs: DM(1-N,20,1), DM(2-N,40,1), DM(3-N,40,1), DM(4-N,40,1), DM(5-N,80,1), and DM(6-N,80,1). Then patterns 300 times as large as the original patterns are obtained. We denote them collectively as DM(9,300,1). Similarly, the second approach to produce artificial patterns using the CDMs without a shear: DM(1-N',20,1), DM(2-N',40,1), DM(3-N',40,1), DM(4-N',40,1), DM(5-N',80,1), and DM(6-N',80,1), denoted collectively as DM(10,300,1). The last approach straightforwardly uses LDMs models: DM(1,20,1), DM(2,40,1), DM(3,40,1), DM(4,40,1), DM(5,80,1), and DM(6,80,1), denoted collectively as DM(8, 300,1). We produce artificial patterns by the above three approaches to train the on-line recognizer and show the result in Table 4. The testing database is the same as the previous experiments. From Table 4, both of the two combined models with or without a shear part improve recognition accuracy for *Kanji* compared with using six LDMs alone. Among them, DM(10,300,1) produces the best recognition accuracy: 95.94% for *Kanji*.

6. Combination sequence and original pattern quality to artificial pattern

In this section, we investigate the influence of the combining LDMs and NLDM, i.e., LDMs then NLDM and vice versa, denoted as L-NL and NL-L, respectively. CDM without a shear part: DM(10,300,1) belongs to L-NL. Similarly, we change the combining order of LMDs and NLDM inversely and obtain DM (11,300,1), which belongs to NL-L.

In this section, we investigate the influence of the combining LDMs and NLDM, i.e., LDMs then NLDM and vice versa, denoted

Table 4
Recognition accuracy for each character subgroup, tested on CDMs (%).

	Symbol	Numeral	Upper roman	Lower roman	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(8,300,1)	74.92	98.40	97.62	94.88	79.00	93.50	85.13	84.02	95.79
DM(9,300,1)	73.16	98.50	97.58	94.92	78.75	93.50	85.07	84.03	95.79
DM(10,300,1)	73.18	98.40	97.58	94.38	78.75	93.00	85.13	83.73	95.94

as L-NL and NL-L, respectively. CDM without a shear part: DM(10,300,1) belongs to L-NL. Similarly, we change the combining order of LMDs and NLDM inversely and obtain DM (11,300,1), which belongs to NL-L.

To evaluate the effect of the distortion models in the previous sections, we have used artificial patterns generated from entire patterns in the Nakayosi database to train the on-line recognizer. To test the distortion models completely universally, we change the original patterns used to produce artificial patterns. We merge Nakayosi and Kuchibue and divide it into two sets. The first set consisting of 100 patterns per character class is used as testing data. The remaining patterns per character class in all character classes form the second set used for producing artificial patterns.

Here, we compare two strategies to select original patterns used for generating artificial patterns. The first strategy prioritizes patterns of better quality. The second strategy prioritizes patterns of worse quality. We can use our MRF model to match the feature points of each pattern with the states of its corresponding character class model and obtain a similarity for each one. Here, our MRF model was probabilistically trained from learning patterns as shown in (Zhu and Nakagawa, 2011). To realize the above two strategies, patterns per character class are evaluated by the similarities to their respective corresponding character class models and are sorted. Patterns that have little similarity to character class models are considered of worse quality, while those having higher similarity are considered of better quality. Since similarity is always minus, and the bigger the similarity, the better the quality of a pattern to its character class model. When we use DM(10,300,1), we select the top 5% patterns, and then those from the top 10% to 100% patterns with step 10% in the ascending or descending quality in the distorted set to produce artificial patterns then train the on-line recognizer. Similarly, for using two original pattern selection strategies to generate artificial patterns by DM(11,300,1). The results of these two CDMs for every character subgroup are shown in Fig. 6.

In investigating the influence of combining sequence between LDMs and NLDM when the same strategy of selecting original patterns to generate artificial patterns is used, compared with the highest recognition accuracy of every character subgroup between DM(10,300,1) and DM(11,300,1), DM(10,300,1) obtains higher recognition accuracy for *Kanji* and apparent higher improvement accuracy than DM(11,300,1). Second, recognition accuracy over

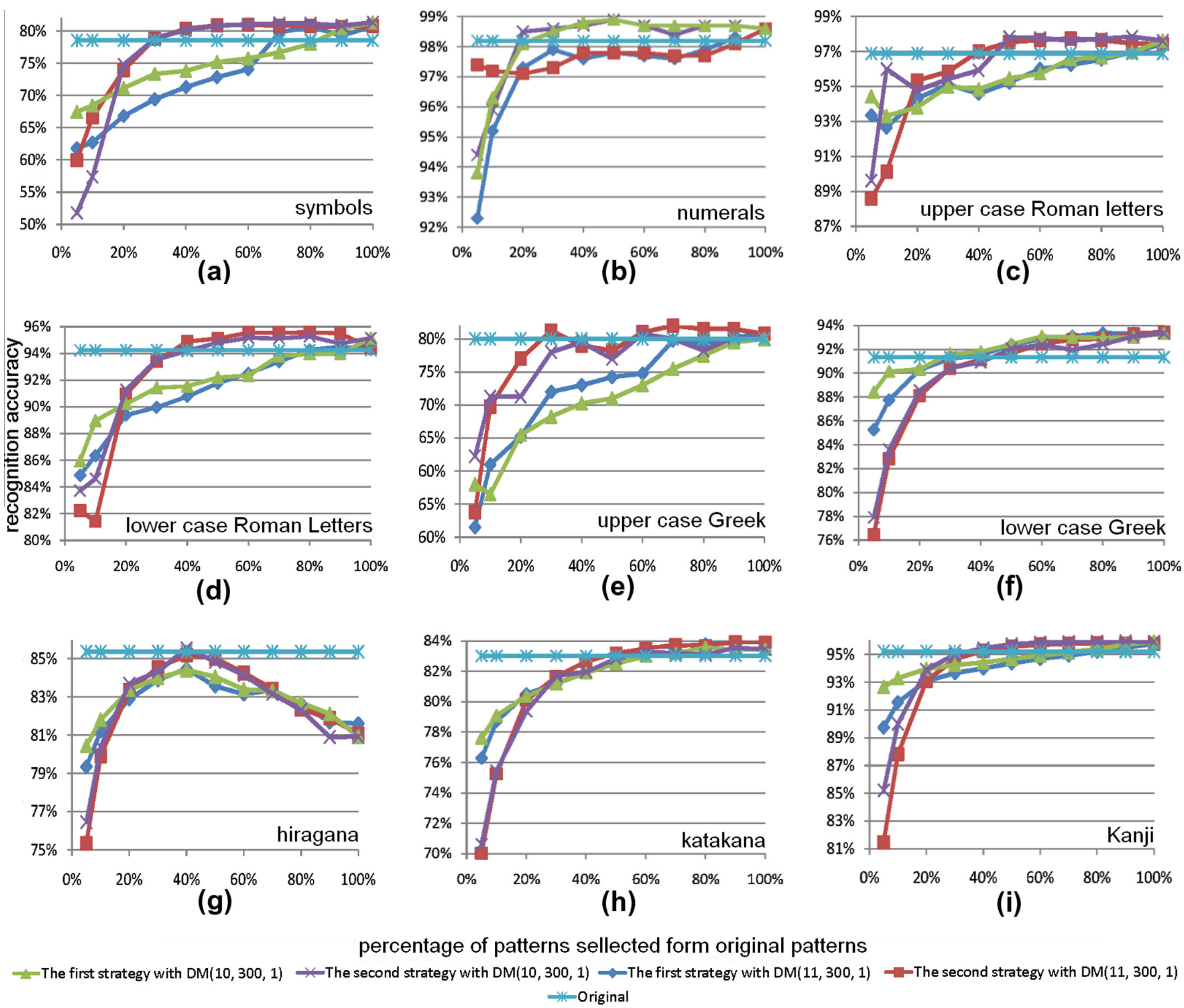

Fig. 6. Recognition accuracy for each character subgroup, tested on CDMs with different data selection strategy, on DM(10,300,1) and DM(11,300,1).

the baseline is obtained before whole original samples are used. Specifically, when using the second strategy of selecting original patterns to generate artificial patterns, recognition accuracy over the baseline can be obtained by even a small quantity of patterns than using the first strategy. Higher recognition accuracy of *Kanji* is obtained by using the second strategy. Finally, the highest recognition accuracy, 95.94% for *Kanji*, is obtained by using the second strategy of selecting 90% original patterns to generate artificial patterns by DM(10,300,1) distortion model.

7. Conclusion

We have presented an effective approach to enhance the accuracy of on-line handwriting Japanese recognition by using a large amount of artificial patterns generated by 12 linear distortion models and combination with a non-linear distortion model. With experiments on nine character subgroups of the Kuchibue database, the recognition accuracies are improved for most of the subgroups, which demonstrates the effectiveness of our approach. Our 12 LDMs improved the recognition accuracy of all the character subgroups and achieved 95.87% recognition accuracy for *Kanji*. Moreover, the distortion model DM(10,300,1), which is, produces the best recognition accuracy, 95.94% for *Kanji*.

As for the combination sequence and strategy to supply training patterns, training patterns from patterns close to prototypes to those away from them and vice versa are distorted by CDMs with two types of sequences: LMDs then NLDM or vice versa. Considering this, we merge the Nakayosi and Kuchibue databases, and take 100 patterns in the merged database form the testing set, while the remaining samples from the training set. DM(10,300,1) obtains higher recognition accuracy for *Kanji* and apparent higher improvement accuracy than DM(11,300,1). Second, Higher recognition accuracy of *Kanji* is obtained, when using the second strategy. Finally, the highest recognition accuracy, 95.94% for *Kanji*, is obtained by using the second strategy of selecting 90% of original patterns to generate artificial patterns by DM(10,300,1) distortion model. This implies that using distant patterns is more efficient for MRF based recognizers.

Though meaningful results have been obtained, there are some unsolved problems to consider for the future research. Artificial pattern generation by pen movement models rather than a pattern deformation model is one direction of our future research. Deformation in the unit of a stroke or radical can be produced and is reflected in character patterns.

References

Chen, B., Zhu, Bilan, Nakagawa, M., 2010. Effects of a large amount of artificial patterns for on-line handwritten japanese character recognition. In: Cina-Japan-Kerea Joint workshop on Pattern Recognition (CJKPR), pp. 90–93.

Chen, B., Zhu, Bilan, Nakagawa, M., 2011. Effects of generating a large amount of artificial patterns for on-line handwritten Japanese character recognition. In: Proc. 11th International Conference of Document Analysis and Recognition (ICDAR), pp. 663–667.

Ha, T.M., Bunke, H., 1997. Off-line, handwritten numeral recognition by perturbation method. IEEE Trans. Pattern Anal. Machine Intell. 19 (5), 535–539.

Jaeger, S. Nakagawa, M. 2001. Two on-line Japanese character databases in unipen format. In: Proc. 6th International Conference on Document Analysis and Recognition (ICDAR), pp. 566–570.

Jaeger, S., Manke, S., Reichert, J., Waibel, A., 2001. On-line handwriting recognition: the NPen++ recognizer. Int. J. Document Anal. Recognit. 3 (3), 169–180.

Kimura, F. et al., 1987. Modified quadratic discriminant functions and the application to Chinese character recognition. IEEE Trans. Pattern Anal. Machine Intell. 9 (1).

Leung, K.C., Leung, C.H., 2009. Recognition of handwritten Chinese characters by combining regularization, Fisher's Discriminant and Transformation Sample Generation. In: Proc. 10th International Conference of Document Analysis and Recognition (ICDAR), pp. 1026–1030.

Leung, C.H., Cheung, Y.S., Chan, K.P., 1985. A distortion model for chinese character generation. In: Proceeding of the 1985 IEEE International Conference on Systems, Man and Cybernetics, pp. 38–41.

Liu. C.L., Zhou, X.D., 2006. Online Japanese character recognition using trajectory-based normalization and direction feature extraction. In: Proc. 10th IWFHR, pp. 217–222.

Liu, C.L., Sako, H., Fujisawa, H., 2003. Handwritten Chinese Character Recognition Alternatives to Nonlinear Normalization. In: Proc. 7th International Conference of Document Analysis and Recognition (ICDAR), pp. 524–528.

Liu, C.L., Jaeger, S., Nakagawa, M., 2004. On-line recogntion of Chinese characters: the state of the art. IEEE Trans. Pattern Anal. Machine Intell. 26 (2), 198–213.

Mori, M., Suzuki, A., Shio, A., Ohtsuka, S., 2000. Generating new samples from handwritten numerals based on point correspondence. In: Proc. 7th Int. Workshop on Frontiers in Handwriting Recognition, pp. 281–290.

Nakagawa, M. Matsumoto, K., 2004. Collection of on-line handwritten Japanese character pattern databases and their analyses. In: IJDAR2004, vol. 7, pp. 69–81.

Plamondon, R., Srihari, S.N., 2000. On-line and off-line handwirting recognition: a comprehensive survey. IEEE Trans. Pattern Anal. Machine Intell. 22 (1), 64–82.

Ramer, U., 1972. An iterative procedure for the polygonal approximation of plan closed curves. Comput. Graphics Image Process. 1, 244–256.

Satoh, S., Aso, H., Miyake, S., Kuroiwa, J., 1999. Pattern recognition system with top-down process of mental rotation. Lect. Notes Comput. Sci. 1606, 816–825.

Smith, S.J., Bourgion, M.O., Sims, K., voorthees, H.L., 1994. Handwritten character classification using nearest neighbor in large database. IEEE Trans. Pattern Anal. Machine Intell. 16 (9), 915–919.

Tsukumo, J., Tanaka, H., 1988. Classification of handprinted Chinese characters using non-linear normalization and correlation methods. In: Proc. Ninth International Conference on Pattern Recognition, pp. 168–171.

Velek, O., Liu, C.-L., Nakagawa, M., 2002. Vector-to-image transformation of character patterns for on-line and off-line recognition. In: International Journal of Computer Processing of Oriental Languages, vol. 15, no. 2, pp. 187–209.

Wakahara, T., Odaka, K., 1996. Adaptive handwrittern character shape normalization using global affine transformation. IEICE, 356.

Yamada, H., Saito, T., Yamamoto, K., 1984. Line density equalization - a nonlinear normalization for correlation method. Trans. IEICE Japan J67-D (11), 1379–1383.

Zhu, Bilan, Nakagawa, M., 2011. On-line handwritten Japanese characters recognition using a mrf model with parameter optimization by CRF. In: Proc. 11th International Conference of Document Analysis and Recognition (ICDAR), pp. 603–607.