

Non-similar candidate removal method for off-line handwritten Korean character recognition

Seon-Hwa Jeong, Yun-Seok Nam, Hye-Kyu Kim
Postal Technology Research Center, ETRI
160 Kajong-Dong, Yusong-Gu, Taejon, 305-600, Korea
E-mail: {sh-jeong, ysnam, hkkim}@etri.re.kr

Abstract

We describe an off-line handwritten Korean character recognition module for real-time address reading. Our module has a two-stage recognition structure. The first recognition aims at reducing target classes for the second recognition at high-speed. We introduced a method to remove non-similar candidates with an input from the candidate set returned by the first recognizer. The optimal candidate set obtained after the operation of the removal method is used to be the target class set for the second recognition and hence should include the class membership of input without fail and a possible small number of candidates in consideration of the whole processing time. The result of experiment done with the PE92 database has proven the superiority of the removal method.

1. Introduction

Our ultimate goal is to develop an off-line handwritten Korean character recognition module for real-time Korean address reading. So, we've considered the following requirements when designing the module. The first is to process an input at high speed and return its membership class with a possible high rank. The second is to return candidates with reliable recognition scores, for example indicating probability, for the next module of address interpretation which generates a valid address based on the results of character recognition and an address lexicon. In this paper, we present some research result made during developing the module meeting such requirements.

In case of Korean, character recognition becomes more challengeable due to the large character set where there are 11,172 characters to the maximum. However, because we are concerned about only the recognition of characters coming from addresses written on mail pieces, the number of characters considered for recognition can be reduced significantly. The analysis of 50,000 live mail images informs us that the number of characters appearing in addresses with the frequency of more than 30 times is 489. With this statistic taken into

consideration, the target class set in our problem has been composed of the 489 characters. Figure 1 shows them. Although the number of target classes is reduced to 489, the problem to recognize them is still difficult because there exist a lot of similar characters in Korean formed by combining 24 kinds of basic graphemes on the two-dimensional space. This difficulty makes the performance of handwritten Korean character recognition poor still to be used in practical applications in spite of many steady studies [1].

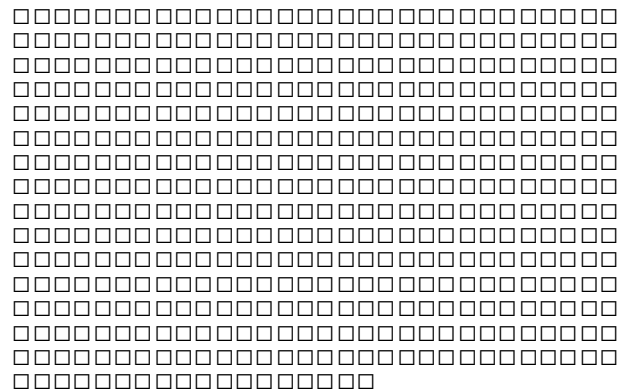


Figure 1. 489 most frequently used characters in Korean addresses

Recently, the studies about classifier combination, a viable alternative to using a single classifier, have been done worldwide actively [2-5]. We also have devised a method combining two recognizers to achieve higher accuracy. Recognition for an input is accomplished two times sequentially. The first recognition aims at reducing the target classes for the second recognition and the goal of the second recognizer is to return the membership class of input with a high reliable recognition score. The proposed method in this paper is related to generate an optimal candidate set from a candidate set returned by the first recognizer. The optimal candidate set is used to be the target class set for the second recognition and hence should contain the class membership of input without fail and a possible small number of candidates. A multi-layer perceptrons network is adopted as the first recognizer and non-similar candidates with an input are removed from a

candidate set returned by the network. To decide whether to remove a candidate or not, some pairwise multi-layer perceptrons networks are used. In Section 2, we will describe the handwritten Korean character recognition module and the proposed method in detail.

2. Handwritten Korean character recognition

2.1. Overview

Figure 2 shows the diagram of our handwritten Korean character recognition module. The input to the module is a binary character image. The output is a set of candidates ranked with their recognition scores. The module starts with feature extraction. Then, the first recognition, the optimal candidate set generation and the second recognition is performed in order. Both of the first and second recognition are based on the character-unit recognition strategy, where each character class is represented by one model and an input is matched to every character model, and use a same feature vector. As the first recognizer, we adopted a multi-layer perceptions network, well-known for its good performance in the pattern classification problem. The network can reduce target classes for the second recognition at high speed and with containing the class membership of an input in a candidate set with a high probability. The second recognizer is under development at present. It should adjust to the dynamical generation of target set according to an input. Therefore, conventional recognition algorithms trained with a fixed target class set are excluded from our interest.

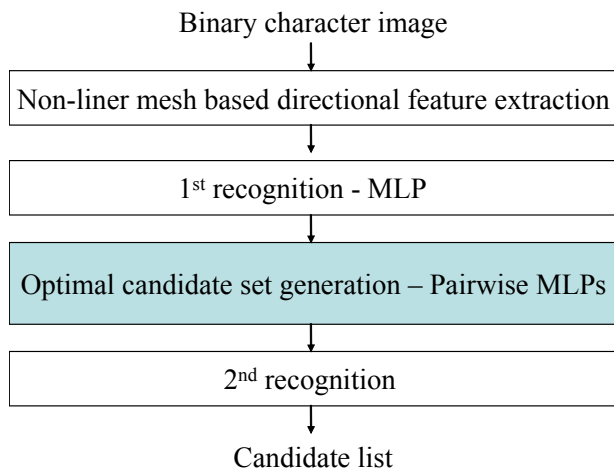


Figure 2. Handwritten Korean character recognition module diagram

2.2 Non-linear mesh based directional feature extraction

Meshing is a key technique to extract the directional features [6]. In this paper, a non-linear meshing method is applied to extract directional features from a binary image - refer to Figure 3(a). Non-linear meshing is equivalent to non-linear shape normalization and can adapt variations in handwritten characters, while its computation complexity is less than that in non-linear shape normalization [7]. Here, we used 6×6 non-linear mesh. The mesh dimension is set through experiment.

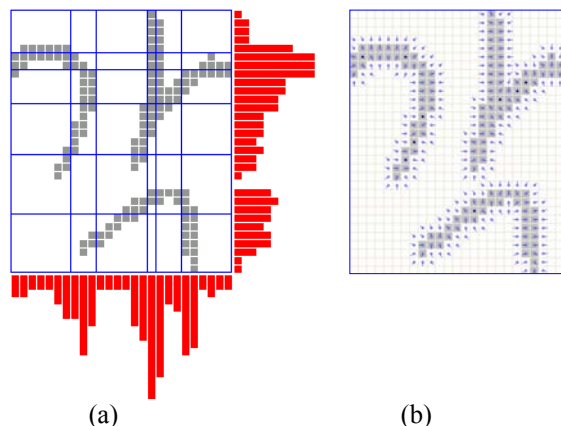


Figure 3. (a) non-linear meshing and (b) directional codes

In each mesh, a histogram of 10 directional codes is obtained. We firstly computed directional values defined in the range of 0° to 359° at pixels and then converted non-zero directional values into one of 8 directional codes, which indicate equal-sized partitioned ranges each. For directional values with zero, we assign them to the 9th code or 10th code in accordance with the pixel value - refer to Figure 3(b). The 9th code represents background and the 10th code foreground. The feature vector is composed of these local histograms. So, 300-dimensional feature vector is extracted from one binary character image. Refer to [8] if obtaining more detailed explanation for our feature extraction.

2.3 1st recognition - MLP

The multi-layer perceptrons network has the structure of three layers: an input layer with the same number of nodes as the dimension of the feature vector, a single hidden layer with 200 nodes, and an output layer of which nodes represent target character classes, where indicating the 489 Korean characters. Refer to Figure 4. One bias term is added to both of the input and hidden layer. The activation function of the hidden and output layer nodes is used to be the sigmoid function. For training the weights in the network, a back-propagation algorithm is used to minimize the mean square error over the training set between the output actually produced and the target

output. Finally, we set learning rate and momentum as 0.7 and 0.1 respectively.

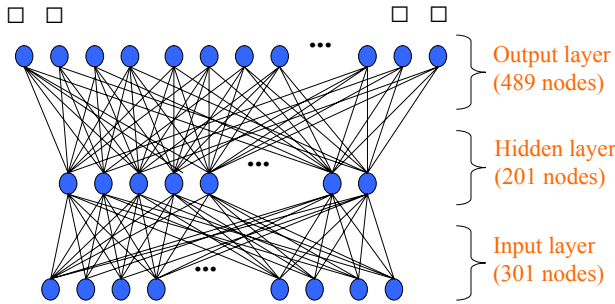


Figure 4. Structure of an MLP for Korean character recognition

2.4 Optimal candidate set generation – Pairwise MLPs

So as most probably to include the class membership of inputs in target class sets for the second recognition, we can consider a lot of candidates from a candidate list returned by the first recognizer. But as the number of target classes is large, the processing time needed in the second recognition increases. So, we should work out a compromise between the two. The proposed method tries to generate an optimal set which includes the class membership without fail and a small number of candidates.

Several methods have been proposed for reducing candidates [9, 10]. They tried to estimate their potential correctness. But, the proposed method removes non-similar candidates with an input based on the results of some pairwise multi-layer perceptrons networks. The pairwise multi-layer perceptrons networks are almost similar with the multi-layer perceptrons network described in the previous section in terms of both of their structure and training method, except the fact that there are only two nodes in their output layer. The two nodes indicate two characters to be discriminated. In the training phase, pairwise multi-layer perceptrons networks for all the pairs made from 489 character classes are trained. The number is 119,316 ($=_{489}C_2$). We have tested them and stored their first-order recognition rates in a table. Depending on the recognition rates, whether the corresponding pairwise multi-layer perceptrons networks work or not is decided. The pairwise multi-layer perceptrons networks with the first-order recognition rate higher than a threshold work. As the value of threshold is high, the confidence level of the candidate removal increases.

Let N candidates denote C_1, C_2, \dots, C_N ranked with their recognition scores. First, to decide whether to remove the first candidate or not, we prepare pairwise

multi-layer perceptrons $PMLP(C_i, C_j), i=2, \dots, N$ of the first candidate and the other candidates. Among these, the networks with the first-order recognition rates higher than a threshold work. When all the working pairwise multi-layer perceptrons networks return C_1 as the first-order candidate, C_1 survives. If even one of them returns C_1 as the second-order candidate, C_1 is removed. Also, the counterparts of C_1 are removed when C_1 is returned as the first-order candidate. Namely, whether to remove C_i and its counterparts C_j s or not is decided depending on the recognition results of $pMLP(C_i, C_j)$, where, C_j s are candidates remaining in the candidate set until then.

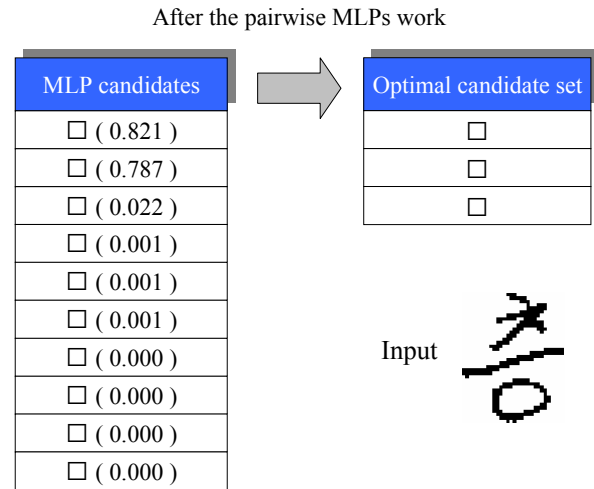


Figure 5. Example of a generated optimal candidate set

3. Experimental results

3.1 Experimental environment

The experiment has been done on a Pentium IV Single Processor 1.3GHz PC with 256MB memory. And to evaluate the multi-layer perceptrons network and the effect of the non-similar candidate removal method using pairwise multi-layer perceptrons networks, we have used a handwritten Korean character database, named PE92 [11]. Among about 100 sets in the database, 70 sets (a total of 34,230 images) were used for training and the rest 13,569 images were used for testing.

3.2 Performance of the MLP

Table 1 shows the cumulative recognition rates of the multi-layer perceptrons network for discriminating the 489 characters. Given more than 20 top candidates, the probability to exist the membership class of input among them is over 99%. The processing time is 1.24ms per character on average. These results prove the validity of

adoption of the multi-layer perceptrons as the first recognizer of which requirement is to reduce target classes for the second recognition at high speed with containing the membership class of input in a candidate set with a high probability.

Table 1. Cumulative recognition rates of the MLP

# of candidate	10	15	20	25	30
Cumul. Rec. Rate (%)	97.61	98.53	99.00	99.23	99.44

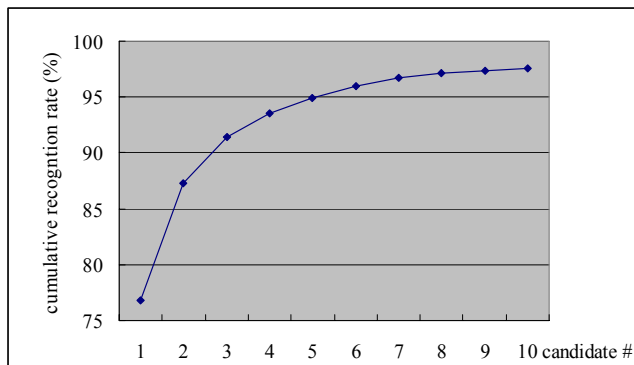


Figure 6. Cumulative recognition rate curve of the MLP

3.3 Performance of the pairwise MLPs

Table 2 shows the performance distribution of 119,316 ($=_{489}C_2$) pairwise multi-layer perceptrons networks. The networks of 85% of them have the first-order recognition rate of 100% and the networks with the first-order recognition rate more than 90% are about 99.6%. This result has supported that candidates can be removed with a good confidence value using the proposed method. The processing time of a pairwise MLP is 0.16ms per character on average.

Table 2. Performance distribution of the pairwise MLPs

1 st recognition rate	Pairwise MLPs #
under 70%	48
70% ~ 80%	66
80% ~ 90%	336
90% ~ 100%	16,999
100%	101,867

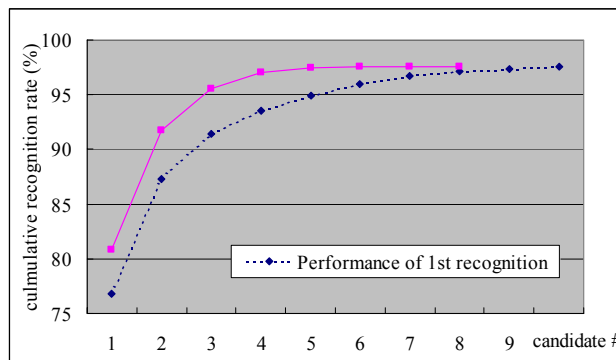
3.4 Effect of the proposed non-similar candidate removal method

Table 3 shows the average numbers of surviving candidates after the proposed method is applied to the candidate sets of the multi-layer perceptrons network. The

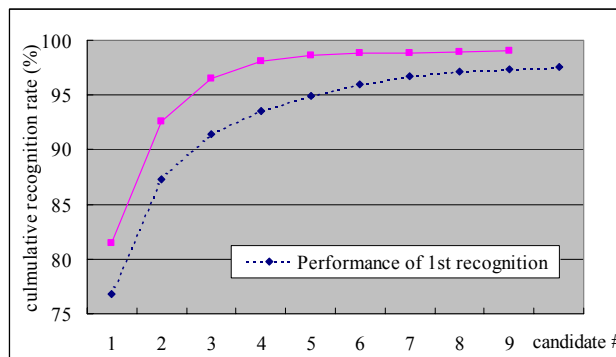
effect of the proposed method seems to be larger when it is applied to the large candidate set. Thirty candidates are reduced to about 6 candidates on average. Here, any drop in the cumulative recognition rates in Table 1 does not happen. Figure 7 presents rather improved cumulative recognition rate curves after the operation of the proposed method. This result comes from the usage of threshold value of 100%. In other words, we let pairwise multi-layer perceptrons networks with the first-order recognition rate of 100% work. So, candidates could be removed with the confidence value of 100%. We could use the threshold value of 100% since the pairwise multi-layer perceptrons networks have good performance as shown in Table 2. If we use a small threshold value, the larger number of candidates can be removed, but the cumulative recognition rates may fall down. So, a trade-off is needed between the cumulative recognition rates and the processing time when deciding the threshold value.

Table 3. Average number of candidates in an optimal set

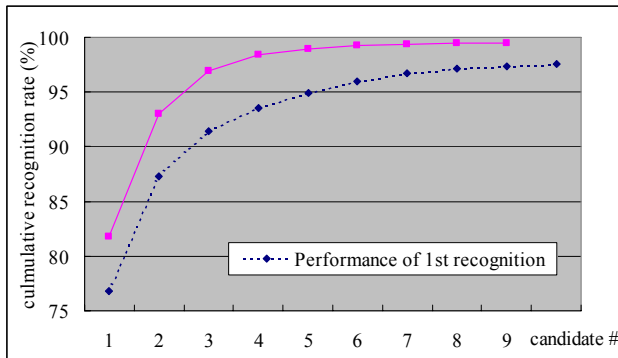
# of MLP candidates	# of surviving candidates after the proposed method is applied
10	4.18
20	5.26
30	5.85



(a) to the top 10 candidates of the MLP



(b) to the top 20 candidates of the MLP



(c) to the top 30 candidates of the MLP

Figure 7. Improved cumulative recognition rate curve after the proposed method is applied

Because our ultimate goal is to develop a real-time character recognition system, the processing time is very important. Until the generation of optimal candidate set, the processing time would be good to be shorter in order to give the second recognizer a possible much processing time. The algorithm of the second recognizer may be time-consuming. Therefore, we've devised the method to generate a possible small target class set for the second recognition. Even though pairwise multi-layer perceptrons networks need a very short processing time, it is desirable that the proposed method should call them a few of times. Table 4 shows the average number of working pairwise multi-layer perceptrons networks when generating an optimal set. As shown, the numbers are not large. So, we do not need much time when generating an optimal set.

Table 4. Average number of working pairwise MLPs

# of MLP candidates	# of working pairwise MLPs
10	10.84
20	28.07
30	46.49

4. Conclusion

In this paper, we have proposed a method to remove non-similar candidates with an input character from a given candidate set. To generate the candidate set, we adopted a multi-layer perceptrons network which has its output nodes corresponding to 489 Korean characters. To decide whether to remove a candidate or not, some pairwise multi-layer perceptrons networks are used. The results of the experiment done with the PE92 database

shows that the proposed removal method can generate an optimal candidate set fast without any drop in the cumulative recognition rates. The processing time until the generation of optimal candidate set is about 3.12ms on average considering 10 candidates. The time is reasonable that our method can be applied to our real-time handwritten Korean character recognition module under development at present.

5. References

- [1] H.Y. Kim, J.W. Suh, J.O. Kwon and J.H. Kim, "Current Hangul recognition activities in Korea," *Proceedings of French-Korean Workshop on Man-Machine Handwritten Comm.*, Paris, France, pp.1-14, 1996.
- [2] L. Yuanxiang, D. Xiaoping, "Multiple candidate characters in the post-processing for off-line handwritten Chinese character recognition," *Proceedings of International conference on Info-tech and Info-net*, Beijing, vol.3, pp.438-443, 2001.
- [3] L.I. Kucheva, "Switching between selection and fusion in combining classifiers: An experiment," *IEEE Transactions on Systems, Man and Cybernetics*, vol.32, no.2, pp.146-156, 2002.
- [4] M. Erp, L. Vuurpijl, and L. Schomaker, "An overview and comparison of voting methods for pattern recognition," *Proceedings of 8th International Workshop on Frontiers in Handwriting Recognition*, Niagara-on-the-Lake, Canada, pp.195-200, 2002.
- [5] A. Bellili, M. Gilloux and P. Gallinari, "An hybrid MLP-SVM handwritten digit recognizer," *Proceedings of 6th ICDAR*, Seattle, USA, pp.28-32, 2001.
- [6] L. Jin, "Handwritten Chinese character recognition with directional decomposition cellular features," *Journal of Circuits, Systems, and Computers*, vol.8, no.4, pp.835-842, 1998.
- [7] X. Gao, L. Jin, J. Yin and J. Hunag, "A new stroke-based directional feature extraction approach for handwritten Chinese character recognition," *Proceedings of 6th ICDAR*, Seattle, USA, pp.635-639, 2001.
- [8] S.H. Lee, H.Y. Kim, K.T. Lim and Y.S. Nam, "Design of a new feature based on direction angle histogram and its application to the recognition of handwritten numerals," *Proceedings on IC-AI*, pp.733-737, 2002.
- [9] E. Ishidera and A. Sato, "A candidate reduction method for handwritten Kanji character recognition," *Proceedings of 6th ICDAR*, Seattle, USA, pp.8-13, 2001.
- [10] C. Lieu and M. Nakagawa, "Precise candidate selection for large character set recognition by confidence evaluation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vo.22, no.6, pp.636-641, 2000.
- [11] D.H. Kim, Y.S. Hwang, S.T. Park, E.J. Kim, S.H. Paek and S.Y. Bang, "Handwritten Korean character

image database PE92," *Proceedings of 2nd ICDAR*,
Tsukuba, Japan, pp.470-473, 1993.