Hand-written Numeric Character Recognition for Facsimile Auto-dialing by Large Scale Neural Network "CombNET-II"

Hiromitsu KAWAJIRI†, *Takatoshi YOSHIKAWA*†, *Junji TANAKA*††, *Anto Satriyo Nugroho*†††, and *Akira IWATA*†††

†Hypermedia Research Center, Sanyo Electric Co., Ltd.
Anpachi-cho, Anpachi-gun, Gifu, 503-0195, JAPAN
††R&D Headquarters R&D Department, Tottori Sanyo Electric Co., Ltd.
Minamiyoshikata, Tottori, 680-8634, JAPAN
†††Dept. of Electrical and Computer Eng., Nagoya Institute of Technology
Gokiso-cho, Showa-ku, Nagoya, 466-8555, JAPAN

ABSTRACT

Facsimiles have recently become popular even at residences. By using a user-friendly man-machine interface, we have developed a smart facsimile which does not need pressing the number buttons any more when sending a fax. The user just writes the destination fax number on the document which then be recognized by the neural network and is transferred automatically as the number to be dialed. In addition, this facsimile also has Auto-registration function which sets both the speed dial and the feature switch simultaneously just after the user inserted the draft with figures written in a specified format.

We used a comb-structured Large Scale Neural Network which is called CombNET-II, with some modifications in the learning process. In applying this model into the facsimile machine, we added neighboring patterns learning, extended selection learning, and weight convergence term in the learning process to achieve even better performance. These modifications realize a fast, high performance and low cost facsimile using a commonly used 8 bit CPU.

The character recognition performance of this system was evaluated using SANYO Hand-written Character Database. The database contains 7500 samples of hand-written characters including nicely-written characters and roughly-written ones. The system can achieve 99.58% recognition rate of nicely-written characters and 99.34% for the roughly-written ones.

1. INTRODUCTION

Facsimiles have recently become popular even at residences. On October 1997 almost 27.2% of residences have facsimiles, and this figure seems to grow year by year. As it becomes popular, more complicated functions are needed such as hand copy machine, cordless answering phone function besides the facsimile-function itself. To response these needs, we have developed a smart facsimile using a user-friendly man-machine interface. This interface makes the facsimile operation simple even for people who are not familiar with such a tool. User just writes the dial number on the document which will be recognized by an offline character recognizer and is transferred automatically as the number to be dialed.

Offline character recognition techniques have been researched for several years, on many aspects such as printed characters recognition, hand-written characters recognition and kanji characters recognition. Some results have been reported in many papers, however, it seems difficult to implement this techniques into an application, since fast CPU and lots of memory are needed. Here we found that a neural network which has been trained to recognize many characters can be applied into facsimile using 8 bit CPU. Since it needs only small amount of memory, the price of the machine can be lowered.

The neural network we use here based on CombNET-II[2] with some modifications in the learning process. This is a joint research between SANYO and Prof.Iwata's Laboratory of Nagoya Institute of Technology, Nagoya, Japan.

2. CombNET-II

CombNET-II consists of 4 layered network with a comb structure as shown in Fig.1. A vector quantizing network forms the first layer as a stem many 3 layered network modules form 2-4 layer as branches. The number of branch network modules is as many as the number of neurons of the stem network. As an input data is given to the stem network, several (3-5) neurons which give higher matching scores between the input data vector and the synaptic weights are selected. Then the input data is led to the plural branch modules which are connected to the selected neurons in the stem network. Final scores are given by the following criterion

$$Z = (SM)^a(SB)^b \qquad (1)$$

where, SM: the matching score in the stem network
SB: the maximum output score in a branch network module

An Input data is classified into a category which has the maximum final score Z. When the input data is trained as one of the members of the branch modules to be classified, a neuron which corresponds with a correct category gives high score SB so that the final score Z to the correct category becomes large even the matching score SM in the stem network is small. Misclassifications in the stem network would be recovered by the branch network modules.

CombNET-II employs the self-growing neural network learning procedure, an original learning procedure [1,2], for training the stem network as shown in Fig.2, where r_{th} is the threshold of similarity, h_c is dividing potential and h_{th} is the threshold of division. The purpose of the stem network is to divide the input feature space into several sub-spaces in which a restricted numbers of the categories are involved. The self-growing neural network learning procedure constructs such a vector quantizing network.

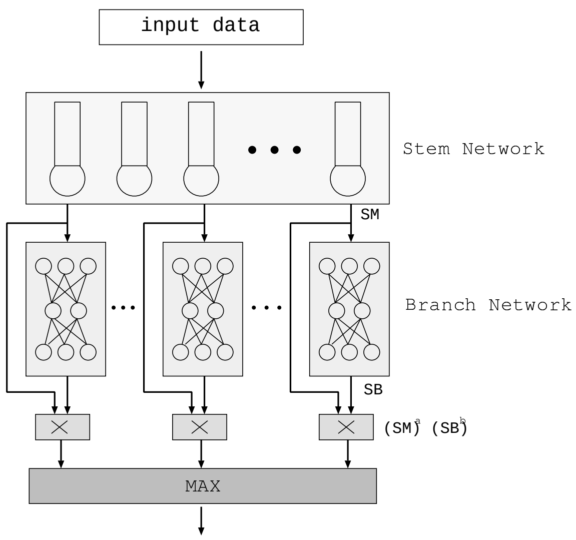

Fig.1: The network structure of CombNET-II

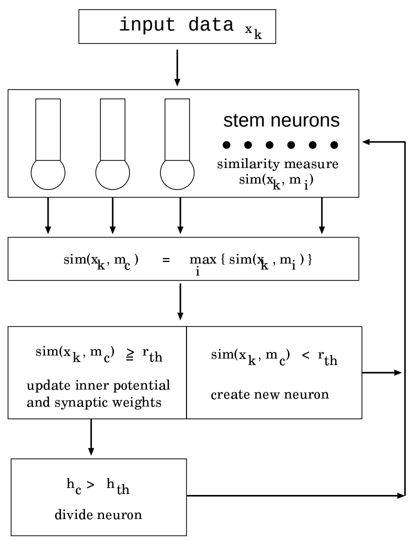

Fig.2: Self-growing NN learning procedure

After learning process of the stem network by these procedures, the input feature space is partitioned into sub-spaces according to the best-matching criteria to the synaptic weight vectors of neurons. The synaptic weight vector becomes a template pattern which represent the common feature of an input data group.

Since the inner potentials of neurons in high density regions of data in the feature space tend to increase, the inner potentials often exceed the threshold of the division h_{th}. Therefore the neurons are frequently created in a region where the many input vectors lie in the feature space. The input data vectors are divided into as many groups as the numbers of the created neurons in the stem network. Since the numbers of categories of input data vectors in each sub-space are restricted by the threshold value of the division, the sizes of the branch networks (numbers of the output neurons) are also restricted.

After training the stem network, all input data are partitioned into category groups according to the best-matching criteria to the synaptic weight vectors of neurons. Then branch network modules are trained for every sub-space to make discriminating boundaries. Back propagation is utilized to train branch modules. Each branch neural network which is 3 layered hierarchical network has a restricted number of output neurons and inter-connections so that it is easy to train.

3. HAND-WRITTEN NUMERIC CHARACTER RECOGNITION

3.1 CHARACTER SEGMENTATION

It is necessary to extract the region of each character(Character Segmentation), since the dial number is assumed to be written in non-frame place. The strokes connection is used to judge the segmentation. A character is not supposed to be connected to the other characters. If two strokes are not connected, it is considered to be 2 characters. However, if the sepparated strokes due to the distance and position characteristic are recognized as one character, the sepparated strokes should be considered to be connected.

3.2 FEATURE EXTRACTION

A neural network to recognize hand-written numeric characters (10 digits and 3 symbols [-,*,#], 13 categories in total) has been implemented by CombNET-II.

After a character region was picked up, a character image was converted to a gray level image of 8×8 pixels. To extract features of patterns, the input patterns of characters are preprocessed using the network as shown in Fig.3. Each unit of the first layer has local connections (3×3 pixel size) from the input layer. The interconnecting weights are shown between the input and first layers in Fig.3. The first layer extracts local feature of line directions in four particular orientations. Each unit of the second layer has local connections (4×4 pixel size) from the first layer. The interconnecting weights are also shown between the first and second layers in Fig.3. The second layer is provided in the network to allow for positional errors of the features extracted by the first layer. $64(4\times4\times4)$ features and $64(8\times8)$ features as first gray level image are used as the input data of CombNET-II.

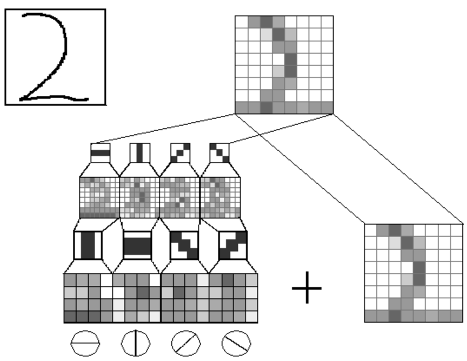

Fig.3: Preprocessing for input patterns of characters

4. MODIFICATIONS IN THE LEARNING PROCESS

4.1 NEIGHBORING PATTERNS LEARNING

As showed in Fig.1, in CombNET-II the most suitable answer should be selected from 3 candidates as a final result. However, to apply this model into Facsimile, the candidates are restricted to one to shorten the processing time.

Since the Stem Network learning process just train the selected group, if a wrong group is selected, it will give a wrong recognition result too.

Here, we take the neighboring patterns in the learning process which make performance of the system better although only the first candidate is used. As a result, the time needed for recognition process is shortened. (equation (1) is not needed anymore)

4.2 EXTENDED SELECTION LEARNING

In Selection Learning of CombNET-II, if the BP output error is too small the weights will not be corrected. It is showed that Selection Learning can shorten the needed time for training and also prevent the system from overlearning.

Here we added, that if the BP output error is too big the weights will not be corrected too. We call this as Extended Selection Learning.

These modifications make a fast and stable learning process which will prevent from the influence of unique pattern.

4.3 WEIGHT CONVERGENCE TERM IN BP LEARNING

Equation (2) shows BP weight correction which is used in CombNET-II. w_{ji} is weight value between neuron i and neuron j. If a big number is included in the absolute value of the weights, it is said that the performance of the neural network will become worse. To keep the absolute value of the weight small, we add weight convergence term as showed in equation (3).

$$\Delta w_{ji}(t) = \eta\delta_j x_i + \alpha\Delta w_{ji}(t-1) \quad (2)$$

$$\Delta w_{ji}(t) = \eta\delta_j x_i + \alpha\Delta w_{ji}(t-1) \underline{-\gamma sgn(w_{ji}(t-1))|w_{ji}(t-1)|^2} \quad (3)$$

where, $sgn(A) = 1(A > 0) or - 1(A < 0)$
η: learning rate, α: momentum, γ: weight convergence rate

This modification improves the recognition performance to unseen characters.

4.4 RESULTS OF MODIFICATIONS

A database of about 13,000 exemplars (about 2,000 people) of hand-written numeric characters collected in Tottori Sanyo Electric Co.,Ltd. is used in the examination. The first about 5,500 exemplars of each character(about 70,000 patterns in total) in the database made up the training data set and the remaining 7,500 exemplars went into the test data set. The test data consists of nicely written characters and roughly written ones. Fig.4 shows the examples of them.

We add the 3 modifications (neighboring patterns learning, extended selection learning and weight convergence term) in the learning process, and the performance of the system is evaluated. Table 1 shows the result.

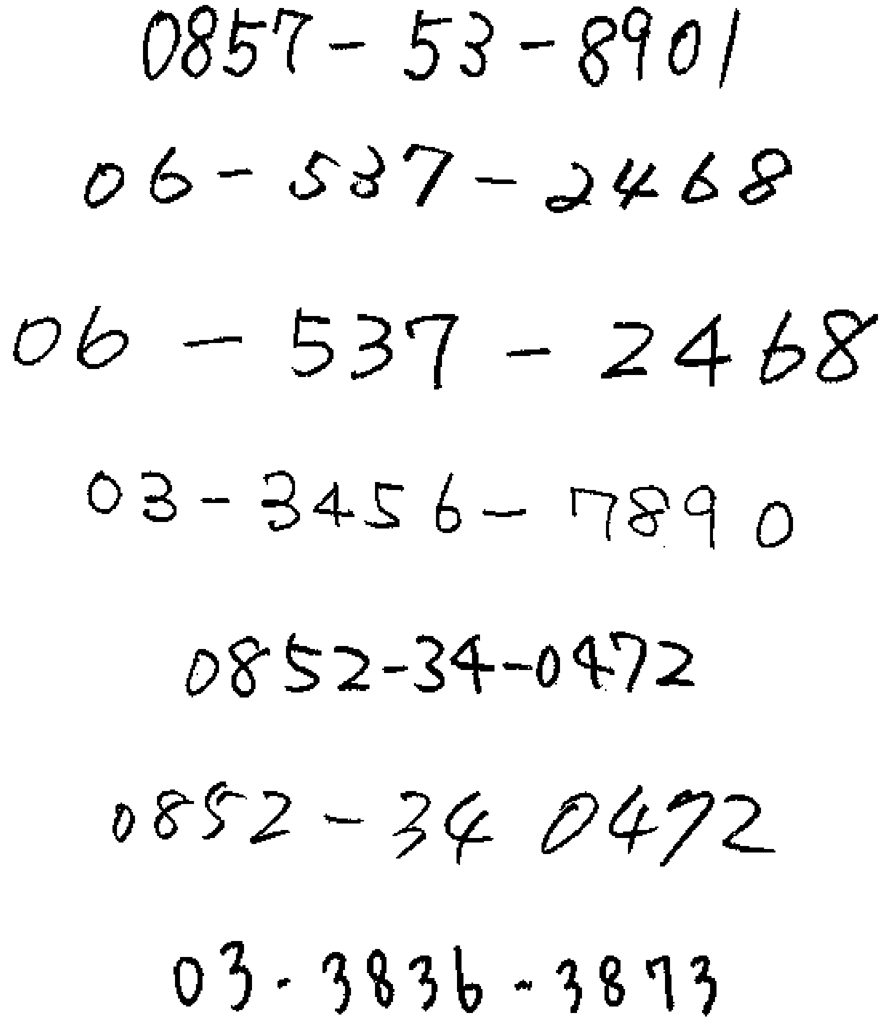

Fig.4: Examples of sample patterns

Table 1: Recognition results

	before modified	after modified
nicely-written characters	99.26%	99.58%
roughly-written characters	98.96%	99.34%

5. FACSIMILE AUTO-DIALING FUNCTION

We have applied hand-written character recognition technique using neural network to implement Auto-dialing function of facsimile. User writes the dial number on the top of the facsimile's draft.(Fig.5) Then, user inserts the draft to the machine and the header will be automatically read by the system. Each number will be recognized, displayed on the LCD and pronounced respectively. After all of the numbers were displayed, it will be pronounced as a confirmation. If the numbers are correct, the user pushes START button, and the machine will dial the displayed number and the draft will be sent.

As default set, after the number has been confirmed, the draft will not be sent until the user pushes the START button. It is possible to set the machine to make the draft sent automatically without pushing the START button. The dial number on the head of the draft will not be sent.

As addition, the facsimile also has Auto-registration function since the system can recognized (*,#), too. Auto-registration function means that if user writes the speed dial number, the telephone number and the number which code the name in a draft with specific format, after the draft is inserted, the facsimile can transfer the data as speed dial recording. Besides the speed dial, it is possible to change the setting using the feature switch of the facsimile.

These functions realized a new and easy to use man-machine interface. This is a great contribution especially to the aged people which are not familiar with mechanical operation.

Fig.5 shows the facsimile with these functions, while the characteristics of the character recognition parts is shown at Table 2.

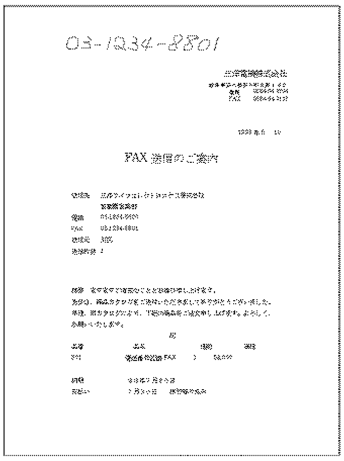

Fig.5: Facsimile draft

Fig.6: Appearance of the facsimile(SFX-81CL)

Table 2: Summary of the character recognition

Categories	digits(0–9) and symbols(-,*,#)
Recognition Rate	nicely:99.58% , roughly:99.34%
Recognition Speed	0.5sec/character
CPU	8bit, 16MHz
Operation	16 bit fixed decimal points (integer:8bit, decimal:8bit)
Neurons	Stem:5, Branch:128(input)-40(hidden)-13(output)
Weights	Branch:5,640
ROM	program:21Kbyte, data:64Kbyte
RAM	32Kbyte

6. CONCLUSIONS

We have successfully applied hand-written character recognition using neural network (CombNET-II) to add a new function to facsimile. This new function will automatically send the draft to the destination fax number which is written by user at the head of the draft.

In this application, we modified the learning process of CombNET-II, i.e.

1. Neighboring Patterns Learning where the neighboring patterns also being trained
2. Extended Selection Learning to select proper patterns to learn fast in BP learning
3. Add the weight convergence term to keep the absolute value of weights small

Our company has collected hand-written character data from about 2,000 people as a database, and the system can achieve 99.26% until 99.58% recognition rate of nicely-written characters and 98.96% until 99.34% recognition rate for the roughly-written ones.

The application of neural network in character recognition made it possible to develop a low cost but high performance facsimile using commonly used low speed CPU.

The facsimile using this new technology(SFX-81CL) has been sold since February of 1997 and has sold out more than 100,000 units.

REFERENCES

[1] K.Hotta, A.Iwata H.Matsuo, and N.Suzumura : "Large Scale Neural Network "CombNET-II"", Tech.and Res.Rep.of IEICEJ, Vol.NC-90,No.34,pp29-36,1990(in Japanese)

[2] A.Iwata, Y.Suwa, Y.Ino and N.Suzumura : "Hand-Written Alpha-Numeric Recognition by a Self-Growing neural Network", Proc.of IJCNN92-Baltimore,accepted, 1992.