Analysis and Recognition of Asian Scripts - the State of the Art

Ching Y. Suen[1], Shunji Mori[2], Soo H. Kim[3] and Cheung H. Leung[4]

1. CENPARMI, Concordia University
Montreal, Quebec H3G 1M8, Canada

2. Graduate School of Computer Science and Engineering
University of Aizu, Aizuwakamatsu-shi 965-8580, Japan

3. Dept. Computer Science, Chonnam National University
500-757 Kwangju, Korea

4. Department of Electrical and Electronic Engineering
University of Hong Kong, Pokfulam Rd., Hong Kong

suen@cenparmi.concordia.ca, s-mori@u-aizu.ac.jp, shkim@chonnam.ac.kr, chleung@eee.hku.hk

Abstract

This paper summarizes the research activities of the past decade on the recognition of handwritten scripts used in China, Japan, and Korea. It presents the recognition methodologies, features explored, databases used, and classification schemes investigated. In addition, it includes a description of the performance of numerous recognition systems found in both academic and industrial research laboratories. Recent achievements and applications are also presented. A list of relevant references is attached together with our remarks on this subject.

1. Introduction

Many languages are written and spoken in Asia. They vary considerably in shape and sound. Some look like more European than the others. More details can be found in books related to Asian linguistics and culture. This paper will look at only three languages that are related by shapes, and/or syntax and semantics: Chinese, Korean, and Japanese [G1]. Both Chinese and Japanese contain thousands of Chinese characters. In Japanese, they are called *Kanjis* (*Kan* = Han dynasty, *ji* = character). However, in addition to Kanjis, Japanese also use many phonetic symbols called Kanas, to increase its own vocabulary and to absorb foreign words. Chinese characters were also used in Korea until its King, Sejong of the Yi dynasty invented a new phonetic alphabet in 1443, called 'Hangul', to replace the complicated Chinese characters. A Hangul character is a syllabic composed of two or three graphemes: a consonant, a vowel, and optionally another consonant. Every Hangul character must contain at least two parts, the first consonant and the vowel, and the first consonant is always on the left or on top of the vowel. The optional last consonant is always below the first consonant and the vowel. A set of 11,172 legitimate characters can be generated, but only 2,350 classes among them are chosen as the Korean Standard.

Due to large vocabularies and complexity of the characters, recognition of Asian scripts is known to be difficult [G1, G2]. This paper attempts to summarize the prominent activities in this area in recent years and the challenges lying ahead.

2 Recognition of Chinese scripts

2.1 Introduction

Many scientists have pursued research on handwritten Chinese character recognition for more than 30 years. A comprehensive review paper was published in 1984 [C1]. Two other surveys followed, one in 1993 [C2], and the other in 1996 [C3]. While [C2] covered a wide range of sources, [C3] concentrated mainly on research conducted in Japan.

A large portion of the research work falls into two main groups. The first group of methods extract feature vectors from training samples, compute the statistical distributions, and classify an unknown input by extracting its feature vector and estimating the most likely character class to which it belongs. The second group of methods recognize a character by precisely matching the individual strokes between the input character and a number of template characters. Other research studies include neural networks, combination of multiple experts, stroke extraction and analysis, recognition of character strings with touching characters, etc. Applications include bank cheque processing, postal address recognition, document data extraction, etc. A short report on this subject is given below. Due to the vast amount of published works, the selectio materials is rather difficult and it is inevitable that a lo

Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003)
0-7695-1960-1/03 $17.00 © 2003 **IEEE**

good works cannot be included here due to the limitation of space.

2.2 Feature extraction and classification

Since Chinese characters are made up of strokes in four main directions, most methods extract features related to stroke position and direction. For example, as illustrated in Fig. 1, the peripheral feature corresponds to the distance AB from a point on the character frame to the first black pixel, while CD corresponds to the distance between the first black-to-white and the second white-to-black transitions. All such distances are effective in representing the external shape and internal structure of a character.

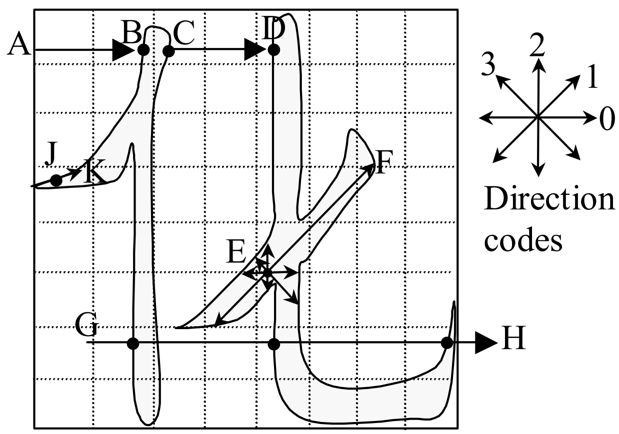


Figure 1: Illustration of feature extraction

Another useful feature is the direction contributivity density. Referring to stroke pixel E in Fig. 1, the distances along the eight directions from E to the boundary pixels of character strokes are used as features. If the distance along direction EF is large and the distances along other directions are small, it implies E belongs to a slant stroke.

The crossing count feature records the number of white-to-black transitions. For example, at pixel G in Fig. 1, horizontal scanning along direction GH yields a count of 3. Scanning along the vertical and slant directions are also performed, and the collected counts represent the density of strokes along various directions viewed from position G.

The tangential direction of a boundary pixel represents the local stroke direction. For example, at pixel J in Fig. 1, the direction is JK. It can be coded by direction codes 0 to 3 as given in the Figure. Coding can be done by either selecting the code with the nearest direction as JK, or resolving JK into two components along the two nearest adjacent code directions.

After getting the feature values at various positions on the character plane, they are assembled into a feature vector as follows. For the peripheral features, the character plane is partitioned into horizontal and vertical bands (with or without overlapping), and the feature values at each band are summed up (with or without positional weightings). For example, in Fig. 1, eight horizontal and eight vertical bands are shown, resulting in a 32-dimensional feature vector for the peripheral stroke distances such as AB, and another 32-dimensional vector for inter-stroke distances such as CD. For the other features (direction contributivity density, crossing count, tangential stroke direction), the character plane is divided into rectangular cells (with or without overlapping). Within each cell, the feature values with the same directional attributes are added up (again with or without positional weightings), resulting in a feature vector of a few dimensions. The overall feature vector is obtained by simply concatenating the vectors from all the cells. The dimension of the resulting overall feature vector is usually large, and may be within the range of a few tens to over a thousand. There are many different combinations and variations to the above methods, resulting in feature vectors that represent both local and global properties.

The feature vector of the unknown input character is compared with that of the training set and some kind of distance measure is used to select the most likely character category. For example, the distance can be city block distance, Mahananobis distance, Euclidean distance, etc. If the dimension of the feature vector is too large, it can be reduced by transformations such as Karhunen-Loeve transformation and retaining those components with large eigenvalues. Due to the large number of Chinese character classes, the classification task may take a long time, and some methods are used to speed up the process. For example, in Ref. [C4], city block distance is used to quickly select the most likely 30 candidate classes from among the 3,000 character classes, and then an asymmetric Mahalanobis distance is used for fine classification. In Ref. [C5], a clustering procedure is adopted to group similar character classes into non-disjoint groups, and a 5-stage classification process, each matching with different features, is used to successively narrow down the probable candidate classes.

2.3 Recognition rates achieved on various databases

In order to compare the performance of different algorithms, it is necessary to test the algorithms on the same database. In Japan, there are standard databases for this purpose. The ETL8 database consists of 881 classes of handwritten Chinese characters and 75 classes of Hirakana characters, with 160 samples per class. On the other hand, the ETL9 database consists of 2,965 Chinese and 71 Hirakana character classes, with 200 samples per class. The characters in these two databases are generally not poorly written. In Taiwan, the ITRI (Industrial Technology Research Institute) database consists of 5,401 handwritten Chinese character classes with 200 samples per class. Besides these databases, there are also other databases collected by individual research groups. Recognition rates reported for tests on the ETL9 database are generally and reach 99.4% in [C4]. For the ITRI database with 5,401 character classes, a recognition rate of 89.0% has been reported [C5]. Higher recognition rates for this database has also been reported, but the number of character classes used was only 2,000.

2.4 Structural matching

Although feature extraction methods coupled with statistical techniques give quite good recognition results, there is still room for improvement, especially for highly cursive and distorted handwriting. The approach of structural matching tries to fill this gap. It aims at precisely matching individual strokes between the unknown input character and stored templates. The input character S is optimally modified to match with the reference character R. The proposed matching techniques are generally iterative optimization processes, and to alleviate the problem of being trapped in local minima, a coarse-to-fine strategy is usually adopted. During the initial phases, large or global scale structures are matched. In successive iterations, the scale is gradually reduced so that finer details are matched.

In Ref. [C6], a global affine transformation (GAT) is applied to the input character S so that the mean of the nearest neighbour inter-point distances between S and the reference character R is minimized. The required GAT is obtained by an iterative procedure. After the application of the GAT, a local affine transformation (LAT) is applied to each pixel of S so that it matches better with R. To prevent excessive distortions, the objective function that is being minimized tries to enforce continuity in the LATs applied to nearby pixels, i.e., LATs applied to nearby pixels within the same neighbourhood should be about the same. The size of neighbourhood is gradually decreased in successive iterations so that finer and finer details are matched. The method has been applied to the matching of handwritten Chinese characters, and actual recognition experiments are underway. The author proposed two different approaches for recognition. The first is to normalize an input character with respect to the reference template by distorting the input with the above matching process. Feature vectors are then extracted from the input character and used for recognition with the conventional statistical pattern recognition techniques. It is expected that features extracted after the distortion normalization procedure will be more stable and hence more reliable for classification. The second approach is to use the residual distance between the GAT/LAT-adjusted input character and the reference template for classification. Both approaches can be combined in a multi-expert system to get better performance.

In Ref. [C7], the input and template characters are systematically distorted to match with each other. An energy function is defined to guide the iterative movements. The function consists of two terms, one measures how close the two patterns are and the other measures the amount of distortion. Hence by minimizing the sum of these two terms, the two patterns are driven to match with one another without excessive distortions. To alleviate the problem of getting trapped in local minima, a size parameter is used for controlling the neighbourhood of interaction. Initially, a large neighbourhood is used with the effect that structures are aligned in a coarse or global manner. The size of neighbourhood is gradually reduced in successive iterations so that finer and finer details are aligned. Satisfactory matching is obtained even for complex character patterns. After matching, the corresponding stroke segments in the two characters are identified, and features such as connectivity, curvature, direction, degree of match, and pre-classification score are used for recognition. The algorithm was tested on a private database with a recognition rate of 96.1%. The computation complexity is proportional to the square of the number of stroke segments and is relatively efficient compared with other algorithms such as relaxation labelling.

The method of relaxation matching is adopted in [C8]. The probability P(i,m) of stroke segment i in the input character matching with segment m in the reference character is updated in successive iterations. A compatibility function is used to measure how good segment i and its neighbour j in the input character match with segment m and its neighbour n in the template character. If the matching of i to m is consistent (or inconsistent) with the matching of j to n, P(i,m) is strengthened (or weakened). In each iteration, P(i,m) is updated according to the support from all the neighbours j and n. Hence mutually consistent matches will dominate after some iterations. Merging and selection procedures are used to refine the matches so that multiple segments in the input character can be matched with a single segment in the template and vice versa. Recognition experiments were carried out on a 2,000 class subset of the ITRI database. A recognition rate of 93.8% was obtained. Relaxation procedures generally require a high computation complexity, being proportional to the fourth power of the number of segments to be matched.

In [C9], a model-based stroke matching algorithm is proposed. The strokes of each reference character are extracted from the pen trajectories of an on-line character input system. A database of stroke attributes and inter-stroke relations for each reference character is built up manually. The information includes stroke type, stroke length, orientation, tolerance for variations in length and orientation, and twelve types of inter-stroke relations. Stroke types include horizontal, vertical, dot, slash, hook, etc. Inter-stroke relations include information such as the start point of stroke A is close to the end point of stroke B, the start point of stroke C is near the mid-portion of stroke D, etc. In actual recognition, the unknown input character is pre-processed to give a graph of line segments. An elaborate search algorithm is used to obtain from the input character candidate strokes that correspond to strokes in the reference character. An objective function is defined that depends on the distance between the input stroke and reference stroke, as well as on the compatibility between this pair of strokes and other related stroke pairs. The algorithm tries to search f solution with minimum cost such that the constraints imposed by the types of relations are satisfied. Experiments on a database of 783 character classes with 200 samples per class produced an error rate of 1.5% and a rejection rate of 0.6%.

2.5 Hybrid method

A hybrid method combining distortion matching and statistical classification was proposed in [C10]. The input image is preprocessed to give four image planes each of dimension 16×16 and each contains stroke pixels in one of the four main directions only. An energy functional is defined which includes a distortion function as a parameter. A distortion function is to be sought such that it distorts the input image planes to match with the templates. The energy function includes two terms: one measures the degree of match and the other measures the smoothness of the distortion function. The idea is to seek a distortion function which gives good matches but at the same time only distorts the input image planes in a smooth manner. An iterative minimization procedure is used to obtain the distortion function. The distortion function is used as a 'normalization' transformation applied to the input image planes. Several distance measures are proposed for statistical classification. One of the proposed distances is the Euclidean distance between the 'normalized' input image planes and the template planes divided by a variance term. The variance is derived during the training phase from the distances between 'normalized' training characters and the template. Experiments on the ETL8B database gave a recognition rate of 96.7%.

2.6 Applications

Applications include postal address recognition, bank cheque amount recognition, recognition of handwritten characters in printed forms, etc. Compared with the recognition of isolated characters, the recognition of handwritten character strings is much more challenging. The main problem lies in the segmentation of touching characters. An approach to tackle the problem is to analyze the connected components and individual strokes [C11]. Components whose rectangular bounding boxes overlap substantially are likely to belong to the same character, for example, boxes A and B in Fig.2. Strokes that touch unnaturally are likely to belong to different adjacent characters. For example, noting that it is an exceptionally long and unnatural stroke may segment CDE in Fig. 2. In Ref. [C11], some heuristics are proposed to decide whether each stroke component should be grouped with which other components.

However, in many situations, treating segmentation as an isolated process is prone to errors. One solution is to merge segmentation with the recognition process. The input character string is over-segmented, i.e., whenever there is a possibility of splitting, the character is temporarily split. For example, the character in Fig. 1 can be split into two connected components: the left component and the right component. In the recognition stage, these two components are tested in various combinations to see whether each forms a character by itself, or forms a character by combining with its adjacent component on its left or right. The confidence score in each case is recorded and an optimal search is performed to obtain the recognized string with the highest overall score. Besides merging segmentation with recognition, the processes can also merge with grammatical or lexical parsing. For example, in postal address recognition, each address is constrained by the rules of the lexicon [C12]. The address consists of the name of a province, a city, a village and a street. Given the name of a province, the alternatives for the name of the city is limited. The name of a city outside the given province is not allowed. This type of constraint helps to narrow down the number of possibilities and can be used to screen out erroneous paths during the searching stage. In the case of bank cheque amount recognition, some grammatical rules can be imposed [C13], e.g. the string 'thousand hundred' is forbidden.

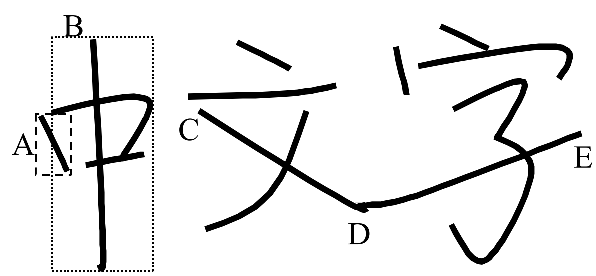


Figure 2: Illustration of segmentation of character strings

Despite these elaborate proposals, the recognition of connected strings is still far from successful. This is because any single error in the recognition or segmentation of the character string, which may consist of over ten characters, will give a wrong overall result.

2.7 Current problems and future directions

For legibly written Chinese characters, the method of feature extraction and statistical classification can achieve very good results. However, when there is too much distortion, the features extracted become unstable enough for reliable classification. Structural matching aims at precisely matching the individual strokes between an unknown input and the stored templates, but the current methods are still not satisfactory. This is because characters are defined by the spatial relations between strokes instead of the absolute positions of individual strokes. Hence the distortion in terms of spatial distance can be extremely large even for character samples within the same class. The distance between two characters of the same class can easily exceed that between characters of different classes and this creates problems for classification. Hence a structural matching algorithm should adopt a distance measure in terms of spatial relations instead of measures based on spatial distances. Unfortunately, it is very difficult to incorporate such a relational dist measure into an objective function in the optimization process. More research is still needed. If spatial relations are to be used for structural matching, it is necessary to implement a database of human knowledge about the writing rules of each character. For example, stroke A is a short horizontal stroke and must be above stroke B, strokes C and D must intersect one another, etc. The manual process is very tedious and some semi-automatic algorithm is desirable. For example, as Chinese characters are built up from radicals which in turn are built from strokes, one way is to manually define the stroke relations for each radical, and then define the radical relations for each character. Some manpower can be saved in this way.

As explained previously, for applications such as postal address and bank cheque amount recognition, the difficulty is in segmenting touching characters in character strings. The current approach of using a soft segmentation strategy in which segmentation is merged with recognition and grammatical or lexical parsing is appropriate. More work is still needed for improvement.

To exploit the different algorithms available for character recognition, a multi-expert approach looks promising. Finally, since human beings have superior performance in terms of character recognition, it is worthwhile to investigate the psychological processes of human beings in character recognition hoping that new discoveries can open up new strategies for computer recognition of handwritten characters [C14].

3 Recognition of Korean Scripts

3.1 Introduction

Like Chinese character recognition, it is a challenge to recognize Korean characters due to the large number of classes. In addition, the degree of inter-class similarity is relatively high because 2,350 (or maximum of 11,172) common characters are generated from only 24 graphemes. Due to these circumstances, most researchers in Korea believe that the problem of Korean character recognition is more difficult than that of Chinese character recognition. A nice review on the techniques for Korean character recognition can be found from [K1, K2] which summarize the research activities before 1998.

3.2 Current Research Activities

Most recent methods for Hangul character recognition can be classified into either holistic or segmentation-based approach. In a holistic method, a character itself is a recognition unit and therefore graphemes do not have to be segmented for character recognition. In a segmentation-based method, on the other hand, a grapheme is the recognition unit and all the graphemes should be segmented from a character prior to their recognition. Unfortunately, both approaches suffer from the chicken-egg paradox. Because of the cursive variation of handwritings, it is difficult to segment graphemes and it is also difficult to recognize a whole character without grapheme segmentation. Apart from these two major approaches, on-line recognizers have also been tried on off-line data. They convert the character image (off-line) into a temporal data (on-line) by estimating the stroke sequence, and then apply a high-performance on-line classifier upon the converted data. The success of this approach depends on the consistency of stroke sequencing.

(1) Holistic approach

Most of the holistic approaches extract a global feature as illustrated in Figure 3, and adopt a multi-stage strategy to alleviate the difficulty of large number of character classes. In this strategy, a rough classification or pre-classification is performed first, and then a final decision is made by statistical pattern analyses such as neural networks, stroke matching, hidden Markov models (HMM), nonlinear pattern matching, k-nearest neighbor algorithm, and so forth [K3-K5]. There are many advantages with the holistic approach: the recognition algorithm is quite simple, it can be useful for the applications where the set of character classes can be restricted, such as courtesy amount recognition, postal automation, and so on.

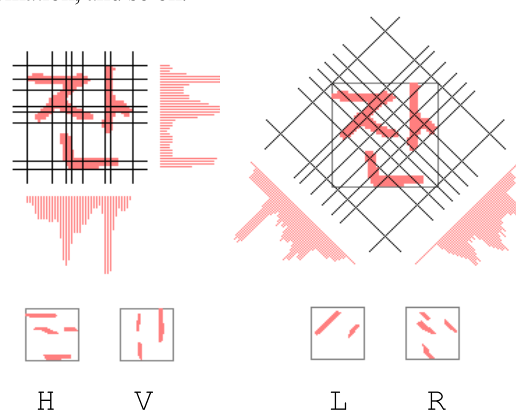


Figure 3: Extraction of a mesh feature for holistic approach [K5]

Although the holistic approaches brought some success, they have severe drawbacks also. The large number of Hangul character classes makes the statistical approaches infeasible and intractable. The computational complexity for model construction and recognition is linearly proportional to the number of character classes. Moreover, the existence of many character classes of similar shapes makes the holistic approaches infeasible because the approaches generally utilize the global statistical features of input character patterns. The holistic approaches also have a tendency to get confused by the local shape distortions caused by touching graphemes.

Recognition accuracy of state-of-the-art systems belonging to the holistic approach varies from 80 to 86%, measured with public or private databases.

Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003)
0-7695-1960-1/03 $17.00 © 2003 **IEEE**

(2) Segmentation-based approach

Another majority of the previous research is the segmentation-based approach. If the graphemes are segmented successfully from a character, their recognition would be a mere job because the shape of graphemes is simple and the number of grapheme classes are comparatively small. However, the problem of grapheme segmentation has been proven to be another difficulty – it is more difficult than the problem of character segmentation in English word recognition because of the 2-D nature of Korean character structure. One of the methods to alleviate this problem is to get feedback from a recognition engine during the grapheme segmentation. These approaches can be formulated as a problem that finds the best segmentation candidate with the highest recognition score among all possible grapheme segmentation candidates. In this approach, all candidates for grapheme segmentation compete and the best one is chosen [K6-K8] – see Figure 4 as an example.

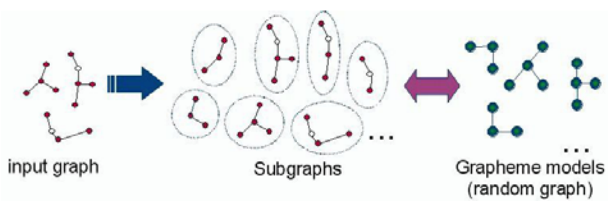


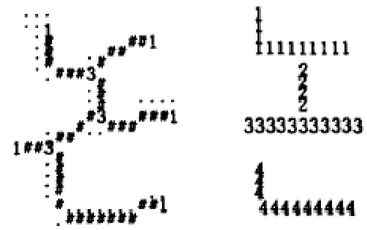
Figure 4: Grapheme segmentation with subgraph matching [K6]

The segmentation-based method is intuitive in the sense that it reflects the principle of Korean character generation. In addition, this method is complete, i.e., it can recognize characters of all classes in the set of 11,172 Korean characters. There are some disadvantages too. The segmentation-based algorithm becomes complex due to the grapheme segmentation. This approach is more sensitive to the noises, such as invalid holes or salt-and-pepper noises, which distort the structure of character strokes. The algorithm is not flexible for the change of recognition target, where the set of character classes under consideration can be confined to a subset of 2,350 Korean Standard classes. Since this approach is effective only for the recognition of Korean characters, it is difficult to extend it to the recognition of multi-lingual characters.

Recognition accuracy of state-of-the-art systems belonging to the segmentation-based approach varies from 78 to 86 percents, measured with a few public or private databases.

(3) On-line approach

Motivated from the fact that the recognition performance of on-line classifier is generally higher than that of off-line ones, a few methods have adopted on-line classifiers [K9-K10]. First they extract temporal data by applying thinning operation, stroke extraction, and a set of stroke ordering rules – refer Figure 5. After the character image is converted into a temporal data, a high performance on-line classifier is applied. In this approach, the stroke ordering should be consistent. Also, the on-line classifier should be trained with the patterns converted from the off-line samples, rather than with the patterns obtained through pen-based tablet devices, to maximize its recognition accuracy.



(a) (b)

Figure 5: An example of stroke sequencing [K10]: (a) after preprocessing; (b) result of stroke sequencing

It is reported that the recognition accuracy of on-line method is around 75%, which is lower than those of the previous two approaches.

3.3 Application Areas

As mentioned earlier, off-line recognition of Korean characters is regarded as a very difficult problem due to a large number of character classes and a high degree of inter-class similarity. State-of-the-art systems have accuracy below 85%. Because of its low recognition accuracy, there have been only a few successful applications during the last two decades.

Currently, the most active area of research and development is postal automation. Utilizing the domain knowledge that the number of Korean characters used for the address is less than 500 and the recognition results can be verified with respect to a lexicon, a special-purpose recognition engine can be developed to make it applicable to real-world environment [K11, K3].

Another successful application is forms processing. There is a lot of demand from the industry to customize their Electronic Document Management System (EDMS) with a form processing solution to handle application forms, receipts, tax forms, credit-card slips, and so on. Some commercial toolkits for form processing and handwriting recognition are available in the market, and several System Integration (SI) companies provide customizing service with these toolkits.

3.4 Problems and Challenges

(1) Character Database

There are two public databases of off-line Korean character samples, PE92 [K12] and KU-1 [K13]. PE92 contains sets of 2,350 characters written by 554 persons, while KU-1 consists of 1,000 sets of 1,000 frequently used characters written by a few thousand persons. Images in both databases are scanned at a resolution of 300 dpi (dots per inch) with 256 gray scales. Problems of these two databases are (1) the number of samples per class is too small to be useful for algorithm development, (2) patterns are written unnaturally, i.e., every sample is written within a fixed-size box. More large-scale databases of samples collected from real-world documents should be constructed in the future to make a practical algorithm.

(2) Lexicon-driven Word Recognition

Although the accuracy of character-level recognition technology does not fulfill the requirement for practical applications, it is possible to make a word-level recognition system whose recognition accuracy is high enough for commercial products in a restricted domain where it is possible to use a lexicon. Use of lexicon means that the input words for the recognizer should be one of the words in a pre-defined lexicon or dictionary. Applications of lexicon-driven word recognition include postal automation, recognition of courtesy amounts in bank slips, recognition of a word from a small vocabulary, and so on [K14-K15].

(3) Shape decomposition

The success of segmentation-based approach for handwritten Hangul recognition highly depends on the success of grapheme segmentation. Due to cursive variation of handwritings, it is difficult to segment graphemes exactly in spite of the use of grapheme recognition during the segmentation. Even worse, the thinning algorithm, used for reducing the stroke width and finding some primitive segments from the strokes, generally distorts the shape of character especially around the corners and junctions. A possible solution for this problem is to decompose the character strokes into a set of building blocks in a consistent way [K17], and then apply a method of combining these blocks to form a grapheme or a character.

(4) Feature Design

A number of features have been proposed for the off-line recognition of handwritten Korean characters. It seems that there still remains much room for further improvement in feature extraction [K17]. Various types of features, psychological as well as perceptual, global as well as local, statistical as well as structural features, should be designed and compared to each other to provide a number of choices for various recognition methodologies.

3.5 Remarks

To develop an algorithm for off-line recognition of handwritten Korean characters, we believe that the following five requirements should be considered. Relative importance of each requirement depends on the target application.

High performance - the algorithm to be developed should outperform existing ones in terms of speed and recognition accuracy.

Coverage - it is desirable that the method can recognize all the characters in the set of 2,350 classes in Korean Standard, or all the 11,172 classes, if possible. In this regard, the segmentation-based approach is more attractive than the holistic one.

Extensibility - it is desirable that the recognition engine can be fused into a multi-lingual classifier, in which more than two kinds of languages appear. In Korea, most documents in daily life are basically multi-lingual, i.e., numerals, English, or Chinese characters, as well as Korean ones are mixed. In this viewpoint, the holistic method is more promising than the other methods.

Flexibility - only 16 Korean characters are used to form the courtesy amounts in bank slips, and less than 500 characters are used for address representation. So, it is desirable that one can utilize this domain knowledge into the recognizer and expect some degree of improvement of performance proportional to the amount of class reduction.

Post-processing scheme - to improve the recognition accuracy of character classification, a post-processing scheme could be applied to the result of recognition. Therefore, it is desirable for the character recognizer to output a list of recognition candidates, along with some scores or confidence measures, suitable for probabilistic processing.

4 Script recognition in Japan

Script recognition is a very active field in Japan. Due to the limitation of space, we can only describe the works of only a couple of active researchers in this paper, but we shall provide appropriate references to other works. Based on the components of an OCR system, we shall describe progress made in preprocessing and discrimination in the past 3 years.

4.1 Preprocessing (normalization)

First we mention the progresses in preprocessing. Recently the principle of the recognition used in OCR is focused on the inner product in a feature space. Therefore, normalization is essential in the field of Kanji recognition for position and size. However, this conventional one is not enough and so non-linear methods and so-called "perturbation methods" have been developed [J27, J32, J4, J21]. Nevertheless these techniques can deal with only a limited range of affine transformation. This leads us naturally to Wakahara's work in this field.

4.2 Research Groups

4.2.1 Wakahara's work

Wakahara conducted large scale experiments for characters used daily in Japan [J22]: Kanji, Kana, and alpha-numerics, based on his normalization techniques, called Global Affine Transformation (GAT) [J31] in addition to Local Affine Transformation (LAT). The essential point of GAT is adaptive normalization. That is, the shape of an input 2D image is adaptively normalized to give the best match with each candidate's 2D reference images. More specifically the skeletons of both input and reference images are used and defined as sets of 2D loci vectors of points.

The overall recognition system is constructed in three stages: (i) rough classification by the basic OCR, (ii) adaptive normalization by GAT, and (iii) Re-classification by the basic OCR. For the basic OCR so-called "extended peripheral direction contributivity (ePDC) is used [J17]. The ePDC feature based OCR achieved around 95% recognition rate for the public handwritten Kanji character database ETL9B [J16]. The system provides K candidates in the first stage of rough classification, where the distance is denoted as D1 (k), $1 \leq k \leq K$. Then at the second stage, these K candidates are adaptively normalized and the distance DNN (k) is obtained. Finally they are fed to the ePDC again and the distance D2(k) is given. Thus the total distance D(k) is defined as

$$D(k) = [\alpha D1(k) + (1 - \alpha)D2(k)] + \beta DNN(k), \quad (2)$$

where $\alpha \in [0, 1]$ and $\beta > 0$. These parameters are established by experimenting a part of the database ETL9B.

4.2.2 Experimental Results

The ePDC was trained in advance using a private handwritten Japanese character database of 3,201 categories including Kanji, Kana, and alpha-numerics, written in the square-style by 1,350 people for each category. The degree of shape variation or distortion per category is almost the same as the ETL9B. However, the number of samples per category is much larger than that of ETL9B, i.e. l, 350 vs. 200. As the test data, the 28, 694 newly gathered images of totally unconstrained handwritten Japanese characters in sciences written on white paper by 300 people are used. The results are analyzed according to the shape complexity among Kanji, kana, and alpha-numerics as follows:

The overall recognition rate is 87.4%. The original recognition rate was 85.8% in which the GAT was not applied. Normalization does not always give a better result for every category. That is, both "positive" and "negative" effects appear. As expected, GAT is effective for Kana and alpha-numerics in particular. The time to recognize one character is 0.12 seconds with GAT and 0.01 seconds without GAT on a Pentium III 650 MHz.

4.2.3 Miyake-Kimura group's work

Miyake-Kimura group conducted a large-scale Kanji recognition experiment using ETL9 database [J3]. So far they continue to revise the record of recognition rate for Kanji constructing high performance OCR systems for ETL8 database [J30]. They challenged to revise it by improving their system in pre-processing. The motivation is that Kanji characters have some groups of mutually similar shapes. Furthermore for complex Kanji characters, their preprocessing seems to be rough because of drastic reduction in the dimension with strong blurring although it is very effective keeping the topological nature of 2D plane. They used directional feature to introduce blurring effect.

In the detection of direction, both the chain code and gradient method were used. Therefore, the resolution of the new system is increased to 49 x 49, although the final size is kept the same as before. The improvements for these changes are summarized as follows:

	Chain code (%)	Gradient (%)
5 X 5 GF	99.14	99.29
13 x 13 GF	99.23	99.31
5 x 5 GF + 13 x 13 sampling size	99.03	99.21
31 x 31 GF + 49 x 49 sampling size	99.26	99.34

Furthermore, they tried the improvement in terms of normalization by constructing a variance absorbing covariance matrix.

4.2.4 Experimental results

The improvements are shown below:

	Old Covariance Matrix	New Covariance Matrix
Chain code	99.07	99.22
Gradient	99.26	99.38
Gradient + 49 X 49 sampling size +31 X 31 GF	99.31	99.41

For Hiragana, the recognition rate is 98.4%. The processing speed is quite high, 14.2 chars/s for the new sys-tem. The old system has 20.8% chars/s. As the test procedure the rotation method is used. The ETL9B database consists of 200 samples/category. So the samples are divided into 10 subsets. When testing some particular subset, the remaining subsets are used as learning samples. The procedure repeated 10 times and the average was taken as the results of 10 tested subsets.

4.3 Other groups

Apart from the above groups, several others have been quite active, making valuable contributions to Japanese script recognition. However, due to space limitation, only som them will be listed below.

(A) Aso-Nemoto's group have achieved high recognition rate of 99.31% using compound Mahalanobis function for improving the recognition in a similar character sets [J24] in 1996. They also conducted experiments on Hiragana using both ETL8 and ETL9B databases constructing elastic models [J14] based on global deformation and character generating parts.

First ETL8B database was used in part of Hiragana set. The Hiragana set consists of 71 characters, each having 160 samples. However, Dakuon and Handakuon characters were excluded and so the remaining 48 kinds characters were used in the experiment. Here Dakuon and Handakuon respectively denote voiced consonants and the p-sound in the Kana syllabary. Their shapes have common structure with double points or small circles placed on the original shapes. So these attached marks create problems in the recognition. The average recognition rate is 99.10%.

The part of Hiragana set of ETL9B database was used as mentioned above. Testing was done the same as that adopted in Kimura-Miyake group. The result is 98.20%.

(B) Sakoe's group proposed a new direction to elastic matching related to normalization. They have applied Dynamic Programming (DP) to the elastic matching and succeeded in both speech and character recognition. Sakoe is known as an originator who first applied DP to elastic matching in the recognition of Hiragana characters. In addition to the brightness /gray-level matching, directional feature matching was considered.

The Hiragana set of ETL8B was used excluding Dakuon (attaching double points) and Handakuon (attaching a circle) The Hiragana set has 160 characters per category and so the even numbered characters were used as test data and the rest as training data to construct the template pattern, using average values. The highest recognition rate of 93.9% was obtained by adjusting one parameter.

(C) Nemoto-Kato Group conducted a large sale experiments on ETL9B database improving multi-layered perceptron using squared connections [J29]. The motivation is to obtain so-called locally excited mechanism, which is required in detailed classification after rough classification. So far they have developed Exclusive Learning neural network (ELNET) for this purpose [J18]. Using a three-layered perceptron in which the connections between the hidden layer and output layer are squared and activated by a Gaussian function.

They used Hiragana set of ETL9B database. The full samples of Hiragana set were used in the experiment. The input to the perceptron is directional feature vector with 196 dimensions as used in [J23]. The result is 97.02% with 120 neurons in the hidden layer. The categories giving less than 95% are listed below.

Category	Recognition rate (%)
ぎ	92
な	94
ぱ	94
ぴ	94
ぺ	93
ぽ	87
さ	94
ば	84
び	90
ぶ	92
ぼ	93
ら	94

They also conducted a large-scale experiment using ETL9B database. In the rough classification HLVQ (Hierarchical Learning Vector Quantization) method was used [J30]. For the detailed classification the proposed method was used. HLVQ method gives 99.75% for the accumulated recognition rate for including the 30th ranked candidates. They obtained the highest recognition rate of 97.04% at 6 candidates in the rough classification.

(D) Kato-Nemoto group investigated the effect of quadratic compound functional. The motivation is to improve recognition of similar characters. Compound function was considered to differentiate between mutually similar characters and some functions have been proposed [J5, J3, J25]. All these methods were proved to be effective. However, the group constructing Quadratic Compound Maharanobis Function (QCMF) tried further improvement. Further more they devised approximation method to reduce calculation load for matrix in QCF. The recognition results are almost the same as the original QCF. From ETL9B database, 15 pairs, i.e., 30 kinds of characters were chosen as similar characters. The best results vary from 92.19% to 95.50%.

(E) Recently support vector machine (SVM) [J28] has become quite active. The basic structure is very simple as McCulloch Pitts model. It has high generalization ability and local maxima can be avoided. It is extended to non-linear SVM and is implemented by employing kernal function simply. Therefore, it was applied to Hiragana characters successfully, and using the subspace method in Hilbert space, Tsuda obtained a recognition rate of 96.36% [Jl].

(F) Nakano's group [J12] conducted recognition experiment on Hiragana characters using the new handwritten character database, called JEITA-HP. The number of the categories is 74. Each category has 1,156-1,158 samples. They adopte vs-l approach in which Max Win Algorithm [J6] and DDAG [J15] and obtained a recognition rate of 94%.

(G) An extensive comparative study on the evaluation of prototype learning algorithms has been conducted by Liu and Nakajima [J10]. They tested 11 prototype learning algorithms, including their own algorithms. They focused on the prototype learning of k-NN classifier based on parametric optimization. The compared algorithms are LVQ2.1 [J7], LVQ3 [J8], the minimum classification error (MSE) [Jll], the generalized LVQ (GLVQ) [J19], a new algorithm MAXP1 devised by them, and so on. The 11 algorithms are tested in handwritten numeral recognition on CENPARMI database and seven algorithms are tested in handwritten Chinese character recognition on the ETL8B database. MCE, the GLVQ and the MAXP1 gave the best performances.

(H) Fujisawa's group reported their work for Japanese mail address reading developed by his group for long time [J9]. This is a historical paper of mail addressing reading system. They emphasized on lexicon-driven segmentation and recognition of handwritten character strings for characters that are used in Japan, in particular Kanji characters, although Japanese address includes Hiragana, Katakana, and alpha-numerics. The lexicon used has very large vocabulary. It contains 111, 349 address phrases from all over Japan. Out of a very large database, 3,589 handwritten letters were picked up for testing, and the highest correct rate of 88.38% is obtained. Tightening the acceptance condition, the error rate can be lower than 1% keeping with recognition rate as is, while the correct rate is still as high as 83.68%.

4.4. Applications

4.4.1 Business card reader [J2]

Business card readers started to appear in the Japanese market in 1990. The first reader was made by Seiko-Epson with price tag of 350,000 yens. Later on, software packages were developed by Media-Drive and other companies like Ricoh, AISOFT, NEC, and Fujitsu Middle Ware, and pushed the price down to the 10,000 yens range (with a scanner) recently, making it popular for personal use. The performances of the readers in the market appeared in NIKEI PC journal [J33]. The CPU of the PC used in the test is Pentium III with 733 MHz and OS of Windows Me. Six readers were tested. The average results are shown below.

Correct reading accuracy of the readers in % in Japanese market

Co. name	name	pos. code	addr.	e-mail addr.	phone.Fax number
78.7	89.7	81.6	85	81.2	90.9

where co. pos. addr. and num. are company, postal, address and other pertinent information respectively.

On the average, the correct reading accuracy is 85% and so it is not so good as compared with the recognition rate of 98% by a commercial OCR reader. The reasons are due to the special properties of business cards, which do not appear in general documents like:

Logos that make the layout analysis difficult.

* Such items as address and phone number are printed in very small sizes.

In cases of e-mail and home page addresses, the characters are very small and they touch each other.

4.4.2 Character recognition in natural environments

Studies on character recognition in scenes began recently. The quality of characters in scenes is usually poor in resolution and they also suffer from background noise. Sometimes they are blurred, because they are taken by CCD cameras directly by humans [J20, J13]. On the other hand, the development of cellular phone makes possible to send scenes easily, so the function of character recognition can be attached to a cellular phone. This will open a broad area of character recognition. Eventually we will be requested to make an intelligent robot which can read any characters captured by camera eyes.

4.5 Concluding remarks

As seen, considerable efforts have been made to improve the recognition algorithms. For example, Kimura-Miyake group achieved very high performance of 99.41% recognition rate for the full set of ETL9B. On the other hand, Wakahara obtained 87.4% recognition rate. What can we see from the difference between the two results? In case of Wakahara, the test data is totally unconstrained which were written in free space without any guide. Naturally we need to classify the data quality and also research target. First concerning the research target, there are two classes roughly speaking; segmented character recognition class and character string recognition. For the latter, postal card recognition is a typical example. In that case, Fujisawa’s group has achieved an 88.38% recognition rate. In Japan, however, people have the habit of writing strings of characters like box arrays. Therefore, the research on segmented character recognition has its value even in practice. In any case, common and public databases are much needed so that we can compare each recognition algorithm and measure progress made in the character recognition technology. In this sense, so far ETL8B and ETL9B have been extensively used in Japan. Now, Hewlett-Packard released JEITA-HP database. It is much bigger than ETL9B and will be used extensively. In fact, Nakano's group used it and obtained 94.2% of recognition rate. On the other hand, Tsuda obtained 96.36% recognition rate using ETL9B based on the same method of SVM. However, there is big difference in test samples between Tsuda and Nakano, i.e., 40 vs 580 respectively. So we can guess that this is one major reason of the difference of recognition rates. We would like to know how the recognition rates change when the test sample size increa

If we have big databases, we can conduct such interes experiments. However, for this purpose we must need tremendous samples. Actually the variation is infinite in principle. Therefore, automatic data generation is tried recently. In this respect NTT has very big database of Kanji, Kana, and alpha-numerics, but it is not made public. We hope that it will be made public in near future to contribute to the research of character recognition. However, we need to note that in such higher accuracy a serious problem is hidden. That is, the performance is not uniform. Sometimes it can read a very distorted sample, which is hard for humans, but on the other hand it can misrecognize a very neat sample of even simple structure such as "ば as explained before. Unfortunately such kind of mistake is not rare and is a very big problem both theoretically and in practice. It is so-called non-excusable error. Hence it is very important to set an appropriate rejection zone so that it can be recovered in some another way.

From theoretical point of view, therefore, we have to consider the current character recognition algorithms analytically. The basic principle of the recognition mentioned above is "inner product" as a measure of similarity. It is well known that such similarity is very sensitive to transformation such as affine one. We can estimate that any shifts of double points on the right top the body of some Japanese characters will fail the OCR system. For such basic problem the reader can visit Pavlidis' home page in which the speech in his K. S. Fu Award is found [J26]. This is really impressive and readable. Another point to be considered is learning. More and more training data are required, which can turn into a major bottleneck. For example, in the case of SVM, we have to deal with a matrix having the number of the training samples as it dimension. However, the manner of the learning in artificial neural network is very different from humans. Usually humans can learn to recognize a character when shown only one template. In this sense Wakahara's and Aso-Nemoto's works of constructing template to be deformed is a good direction, although both methods require considerable training. Anyway humans also need learning. Children can learn to recognize characters at the age of three or four. Before that age children seem to have learned to recognize and differentiate the basic structure of shapes, based on which children can learn to recognize many more characters very easily. However, it is said that it is very difficult to learn Chinese characters. However, this is not for reading, but it is for writing.

Now let us observe what humans do in terms of recognition. When humans look at a character, they feel something like thin or thick, cursive or straight, soft or hard, and so on. In particular, for Chinese characters we can feel aesthetic one for handwritten shape with blush, in particular. In fact it is art. On the other hand, we are requested to make a system that is sensitive to such subtle features. Humans can throw away the detailed features knowing that they exist and what they are exactly. Humans have astonishing ability of analysis and description. That approach in research is called structural analysis. This is old and has been studied by many researchers. However, in the field of character recognition this approach has almost disappeared. As far as we know the last one on Chinese character recognition was conducted Suen's group [J11]. Perhaps this is due to the success of feature matching with blurring and functional analysis approach and revival of neural network. Concerning Japanese characters, recognizing Hiragana is difficult because it is cursive and it is hard to segment strokes in some characters such as "れ" and "ね". So it seems to be easy to treat such characters as a whole, i.e., template matching. So now it is the time for us to get more involved in the structural analysis approach in handwriting recognition so that we can extend its application to broader areas such as document understanding and recognition of characters, which occur in natural environments and scenes.

Acknowledgement

S. Mori would like to thank Professor K. Yamamoto for his help in providing valuable materials to this paper. Thanks are also due to the Natural Sciences and Engineering Research Council of Canada for financial support through a research grant awarded to C. Y. Suen.

References - General

[G1] C. Y. Suen, S. Mori, H.-C. Rim, and P. S. P. Wang, "Intriguing aspects of oriental languages," Int. J. of Pattern Recognition & Artificial Intelligence, vol. 12, pp. 5-29, 1998.

[G2] H. Fujisawa, J. H. Kim, and C.-L. Liu (Guest Eds.), Special Issue: Oriental Character Recognition, Int. J. of Computer Processing of Oriental Languages, vol. 15, June 2002.

References on Chinese scripts

[C1] S. Mori, K. Yamamoto, and M. Yasuda, "Research on machine recognition of handprinted characters," IEEE Trans. on PAMI, vol. 6, no. 4, pp. 386-405, 1984.

[C2] T. H. Hildebrandt and W. Liu, "Optical recognition of handwritten Chinese characters: advances since 1980," Pattern Recognition, vol. 26, no. 2, pp. 205-225, 1993.

[C3] M. Umeda, "Advances in recognition methods for handwritten Kanji characters," IEICE Trans. inf. & syst., vol. E79-D, no. 5, pp. 401-410, 1996.

[C4] N. Kato, M. Suzuki, S. Omachi, H. Aso, and Y. Nemoto, "A handwritten Character recognition system using directional element feature and asymmetric Mahalanobis distance," IEEE Trans. on PAMI, vol. 21, no. 3, pp. 258-262, 1999.

[C5] Y.Y. Tang, L.T. Tu, J. Liu, S.W. Lee, W.W. Lin, and I.S. Shyu, "Offline recognition of Chinese handwriting by multifeature and multilevel classification," IEEE Trans. on PAMI, vol. 20, no. 5, pp. 556-561, 1998.

[C6] T. Wakahara and Y. Kimura, "Toward robust handwritten Kanji character recognition," Pa

Recognition Letters, vol. 20, pp. 979-990, 1999.
[C7] C.H. Leung and C.Y. Suen, “Matching of Complex Patterns by Energy Minimization,” IEEE Trans. Systems, Man and Cybernetic, Part B, vol. 28, pp. 712-720, 1998.
[C8] F.H. Cheng, "Multi-stroke relaxation matching method for handwritten Chinese character recognition," Pattern Recognition, vol. 31, no. 4, pp. 401-410, 1998.
[C9] C.L. Liu, I.J. Kim, and J.H. Kim, "Model-based stroke extraction and matching for handwritten Chinese character recognition," Pattern Recognition, vol. 34, pp. 2339-2352, 2001.
[C10] Y. Mizukami, "A handwritten Chinese character recognition system using hierarchical displacement extraction based on directional features," Pattern Recognition Letters, vol. 19, pp. 595-604, 1998.
[C11] L.Y. Tseng and R.C. Chen, "Segmenting handwritten Chinese characters based on heuristic merging of stroke bounding boxes and dynamic programming," Pattern Recognition Letters, vol. 19, pp. 963-973, 1998.
[C12] J. Xue, X. Ding, C. Liu, R. Zhang, and W. Qian, "Location and interpretation of destination addresses on handwritten Chinese envelopes," Pattern Recognition Letters, vol. 22, pp. 639-656, 2001.
[C13] M.L. Yu, P.C.K. Kwok, C.H. Leung, and K.W. Tse, "Recognition of Chinese bank cheque amounts," International Journal on Document Analysis and Recognition, vol. 3, Issue 4, pp. 207-217, 2001.
[C14] Y.Y. Yang, "Adaptive recognition of Chinese characters: imitation of psychological processing in machine recognition," IEEE Trans. SMC, Part A, vol. 28, no. 3, pp. 253-265, 1998.

References on Korean scripts

[K1] S.W. Lee and H.S. Park, “A Survey on Korean Character Recognition Researches,” Proc. 1st Workshop on Character Recognition, pp. 3-46, Choongbuk University, 1993. (in Korean)
[K2] S.H. Kim, S.H. Jeong and I.S. Oh, “A Survey on the Off-line Recognition of Handwritten Korean Characters,” Proc. Conf. Korean Information System Society, 25(2), pp. 396-398, Ajou University, Oct. 1998. (in Korean)
[K3] S.H. Jeong, K.T. Lim and Y.S. Nam, “A Combination Method of Two Classifiers Based on the Information of Confusion,” Proc. 8th Int. Workshop on Frontiers in Handwriting Recognition, pp. 519-523, Niagara-on-the-Lake, 2002.
[K4] H.H. Song and S.W. Lee, “A Self-Organizing Neural Tree for Large-Set Pattern Classification,” IEEE Trans. Neural Networks, 9(3), pp. 369-380, 1998.
[K5] S.H. Kim, "Performance Improvement Strategies on Template Matching for Large Set Character Recognition," Proc. 17th International Conference on Computer Processing of Oriental Languages, pp. 250-253, Hong Kong, April 1997.
[K6] K.W. Kang, Hierarchical Hangul Character Recognition with Stochastic Relationship Modeling and Candidate Pruning, Ph.D. Dissertation, KAIST, Korea, 2002.
[K7] H.Y. Kim and J.H. Kim, “Heirarchical Random Graph Representation of Handwritten Characters and Its Application to Hangul Recognition,” Pattern Recognition, 34, pp. 187-201, 2001.
[K8] Y.S. Hwang and S.Y. Bang, “Recognition of a Handwritten Korean Character by Combining Segments Using Constraints Satisfying Graph,” Proc. 6th Int. Workshop on Frontiers in Handwriting Recognition, pp. 515-526, Taejon, Korea, 1998.
[K9] H. Kwak and K. Chung, “Recovery of Temporal Information from Static Images of Hangul Handwriting based on Its Structural Information,” Ibid, pp. 527-536, Taejon, Korea, 1998.
[K10] H.J. Kim and P.K. Kim, “Recognition of Off-line Handwritten Korean Characters,” Pattern Recognition, 29(2), pp. 245-254, 1996.
[K11] S.H. Kim, "Recognition of Korean Address Strings by Tight-Coupling of Minimum Distance Classification and Dictionary-Based Post-Processing," International Journal of Computer Processing of Oriental Languages, Vol. 12, No. 2, pp. 207-221, Dec. 1998.
[K12] D.H. Kim, Y.S. Hwang, S.T. Park, E.J. Kim, S.H. Park and S.Y. Bang, “Handwritten Korean Character Image Database PE92,” Proc. 2nd Int. Conf. Document Analysis and Recognition, Japan, pp. 470-473, 1993.
[K13] J.S. Park, H. J. Kang and S.W. Lee, “Automatic Quality Measurement of Gray-Scale Handwriting Based on Extended Average Entropy,” Proc. 15th International Conference on Pattern Recognition, Barcelona, Spain, pp. 426-429, 2000.
[K14] S.H. Kim, S. Jeong, and C.Y. Suen, "A Lexicon Driven Approach for Optimal Segment Combination in Off-line Recognition of Unconstrained Handwritten Korean Words," Pattern Recognition, Vol. 34, No. 7, pp. 1437-1447, July 2001.
[K15] Y.H. Jin, I.J. Kim, J.H. Kim and Y.W. Choi, Lexicon Driven Handwritten Word Recognition Using Grapheme Combination Type,” Proc. Conf. Korean Information Processing Society – Fall 1999,” 6(2), pp. 1-8, 1999. (in Korean)
[K16] Jeong-Sun Park and Il-Seok Oh, “Shape Decomposition and Stroke Extraction of Character Patterns,” Proc. ICPR, pp. 411-414, Quebec city, Canada, Aug. 2002.
[K17] K.H. Hong, E.H. Jeong and B.C. Ahn, "Feature Extraction of Off-line Handwritten Characters Based on Optical Neural Field," Journal of Korea Information Processing Society, Vol. 6, No. 12, pp. 3530-3538, 1999. (in Korean)

References on Japanese scripts

[J1] Koji Tsuda, “Subspace method in the Hilbert space,” IEICE Trans. D-II, I82-D-II(4): 592-599, May 1999.
[J2] Mar. 2002. Report on Input System of Recognition type, The committee of Input System of Recognition type, EITA published.

[J3] T. N. and T. Wakabayashi, F. Kimura, and Y. Miyake. High accuracy in similar characters using compound function. IEICE Trans. D-II, J83-D-II(2): 623-633, Feb. 2000.

[J4] T. Ha and H. Bunke. Off-line, handwritten numeral recognition by perturbation method. IEEE Trans. Pattern Anal. Machine IntelL, 19: 535-539, 1997.

[J5] T. lijima. Theory of Pattern Recognition. Morikita Publishing Ltd., 1989. See for English, S. Mori, H. Nishida, and H. Yamada, 'Optical Character Recognition," Wiley 1999.

[J6] B. Huang and S. Katagiri. Discriminant learning for minimization error classification. IEEE Trans. Signal Process., 40(12): 3043-3054, 1995.

[J7] T. Kohonen. Improved versions of learning vector quntiza-tion. Proc. of the IJCNN, 1: 545-550, 1990.

[J8] C. Liu, I. Kim, and J. Kim. High accuracy handwritten Chinese character recognition by improved feature matching method. Proc. of 4th ICDAR, Ulm, Germany, pages 1033-1037, 1997.

[J9] C. Liu, M. Koga, and H. Fujisawa. Lexicon-driven seg-mentation and recognition of handwritten character strings for Japanese address reading. IEEE Trans. Pattern Anal, and Machin Intel, 24(11): 1425-1437, Nov. 2002.

[J10] C. Liu and M. Nakagawa. Evaluation of prototype learning algorithms for nearest-neighbor classifier in application to handwritten character recognition. Pattern Recognition, 34: 601-615, 2001.

[J11] S. Lu, Y. Ren, and C. Suen. Hierarchical attributed graph representation and recognition of handwritten Chinese characters. Pattern Recognition, 24(7): 617-632, 1991.

[J12] K. Maruyama, M. Maruyama, H. Miyao, and Y. Nakano. Handprinted hiragana recognition using support vector machines. Proc. of the 8th IWFHR, pages 55-60, 2002.

[J13] Niimura and et al. Recognition characteristics of characters in scene changing distances between objects and camera. IEICE PRMU2001-97, pages 21-28, 2001.

[J14] T. K. S. Omachi and H. Aso. Handprinted character recognition using elastic models. IEICE Trans. D-II, J83-D-H(12): 2578-2586, Dec. 2000.

[J15] J. Platt, N. Clistiani, and J. Shawe-Taylor. Large mapping dag's for multi-class classification, pages 547-553, 2000.

[J16] T. Saito, H. Yamada, and K. Yamamoto. On the database et!9 of handprinted characters in jis Chinese characters and its analysis. IEICE Trans. J68-D, pages 757-767, 1985.

[J17] H. Sakano and N. Miyamoto. Handprinted character recognition by extended peripheral direction contributivity. Natl. Conv. Rec. IEICE Japan, page 556, 1995.

[J18] K. Saruta, N. Kato, M. Abe, and Y. Nemoto. High accuracy recognition of et!9b using exclusive learning network-ii (elnet-ii). IEICE Trans. Inf. & Syst., E79-D(5): 516-522, May 1996.

[J19] A. Sato and K.Yamada. A formalization of learning vector quantization using a new misclassification measure. Proc.14th IJCPR, Brisbane, 1: 423-429, 1998.

[J20] Sawaguchi, Yamamoto, and Kato. Comparative study of low resolution character recognition using moving images. IEICE Spring Annual Convention, D-12-15, 2001.

[J21] P. Simard, Y. Le Cun, and J. Denker. Efficient pattern recognition using a new transformation distance. Advances in Neural Information Processing Systems, 5: 50-58, 1993. Morgan Kaufmann.

[J22] T. Wakahara. Handwritten Japanese character recognition using adaptive shape normalization by global affine transformation. Int. Journal of Comp. Processing of Oriental Languages, 15(2): 117-131, 2002.

[J23] N. Sun, M. Abe, and Y. Nemoto. A handwritten character recognition system by using improved directional element feature and subspace method. IEICE Trans., J78-D-H(6): 922-930, Jun. 1995.

[J24] M. Suzuki, S. Oumachi, N. Kato, H. Aso, and Y. Nemoto. IEICE Trans. D-II, J80-D-II(10): 2752-2750, Oct. 1996.

[J25] M. Suzuki, S. Oumachi, N. Kato, H. Aso, and Y. Nemoto. High accuracy recognition method for similar characters using compound Mahalanobis function. IEICE Trans. D-II, J80-D-n(10): 2752-2760, Oct. 1997.

[J26] T.Pavlidis. http://theopavlidis.com.

[J27] J. Tsukumo and H. Tanaka. Classification of handprinted Chinese characters using nonlinear normalization methods. Proc. 9th Int. Conf. on Pattern Recognition, pages 168-171, Nov. 1988.

[J28] V. Vapnik. Statistical Leaning Theory. Wiley, New York, 1998.

[J29] Y. Waizumi, N. Kato, and Y. Nemoto. High accuracy recognition for handwritten characters using multi-layered per-ceptron with squared connections. IEICE Trans. D-II, J83-D-H(10): 1969-1976, Oct. 2000.

[J30] T. Wakabayashi, Y. Deng, S. Tsuruoka, F. Kimura, and Y. Miyake. High accuracy in handwritten kanji characters by nonlinear normalization and compressing feature volume. IEICE Trans. D-II, J79-D-H(5): 765-774, May 1996.

[J31] T. Wakahara and K. Odaka. Adaptive normalization of handwritten characters using global/local affine transformation. IEEE Trans. Pattern Anal. Machine Intell, 20: 1332-1341, 1998.

[J32] H. Yamada, K. Yamamoto, and T. Saito. A nonlinear normalization method for handprinted kanji character recognition - line density equalization. Pattern Recognition, 23: 1023-1029, 1990.

[J33] Yamazaki. Performance evaluation of 6 business card management software packages. NIKEI PC, Aug. 20th 2001.

0-7695-1960-1/03 $17.00 © 2003 IEEE