DEVELOPMENT OF RECOGNITION SYSTEM FOR TEMPLATE-MARKED DIGITS IN BILLETS

Jihun Ko *
YoungKow Lee, SungHoo Choi, and Sang Woo Kim **,[1]
JongHak Lee ***

* *Agency for Defense Development (ADD), Rep. of Korea*
** *Department of Electronics and Electrical Engineering*
Pohang University of Science and Technology (POSTECH)
Pohang, Rep. of Korea
*** *POSCO Technology Laboratory, Rep. of Korea*

Abstract: This paper presents new solutions to recognize template-marked digits in billets. Since these digits can be easily blurred, systematic preprocessing methods and recognition algorithms with better generalization properties are required. In this paper, preprocessing methods using the local threshold and skew correction are proposed to extract features with good quality from captured billet images. And multi-class Support Vector Machine(SVM) based algorithms which select the final class among potential classes are applied. Experimental results show that the proposed approach has higher recognition rate than that of the previous method using the K-Nearest Neighborhood(K-NN) technique for template-marked digits in billets.

Keywords: Pattern recognition, Identification, Image processing

1. INTRODUCTION

In the steel making process, the recognition system is needed to identify current plates with marked codes. Existing recognition process burdens workers with heavy and repeated tasks. Thus, the modified recognition system is required. Moreover, this system can be used to install the database for managing materials. Especially, there are three kinds of marked digits such as auto-marked, template-marked and handwritten digits in billets. They are easy to mix one another. Thus, the recognition system for billets which are produced in various continuous-casting plants plays an important role. Three kinds of marked digits is shown in Figure 1. It is relatively easy to recognize auto-marked digits. However, it is

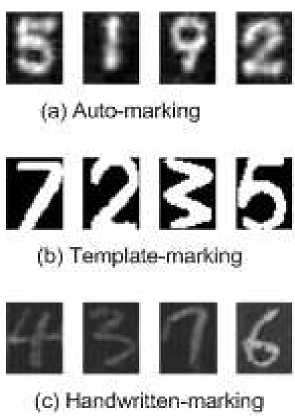

Fig. 1. Three kinds of Billet Digit Data

[1] This work was supported by the Ministry of Education of Korea through its Brain Korea (BK) 21 program.

difficult to recognize template-marked digits in billets, because they are easily contaminated and blurred. The original templates used in billets are shown in Figure 2. In this paper, techniques and

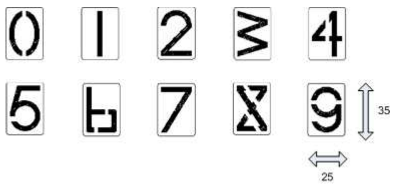

Fig. 2. The basic form of template-marked digit

algorithms are shown to improve the recognition rate of template-marked digits.

Typically preprocessing, normalization, feature extraction, classification, and postprocessing operations are required in the handwriting recognition (Bunke, 2003) (Bunke and P. Wang, 1997). It is shown the process of the recognition system fot auto-marked digits of billets in Figure 3. Especially, two preprocessing techniques which

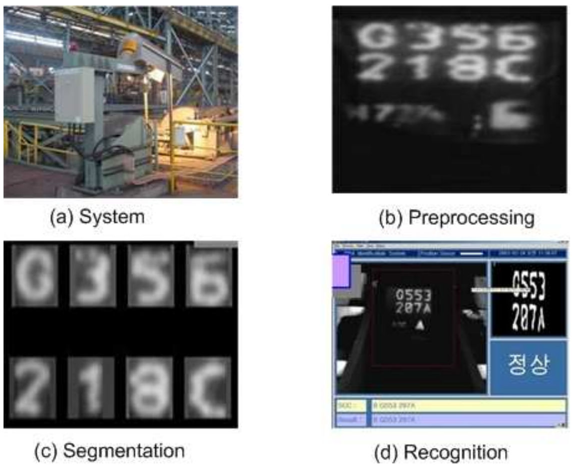

Fig. 3. The process of Billet recognition system

can extract features with good quality from the billet images are proposed. One is the technique which converts the gray-level image to the binary image. A number of techniques for this step are known from (S. Mori and Yamada, 1999), (O'Gorman and R. Jasturi, 1995). The other is the technique which corrects the skewed digits. Since the digits are not perfectly aligned in billets, they bring bad results in the segmentation process. Methods for skew angle estimation are based on horizontal projection profiles: the handwriting's contour and other quantities (M. Morita and Sabourin, 1999).

In the recognition system for template-marked digits, its algorithm with generalization ability is required since the digits can be often blurred and distorted. In many recognition systems such as printed character recognition system, K-NN method has been widely used because of easy realization, fast recognition speed and high recognition rate. Neural network and fuzzy system have been also used (J. Liu, 2002)(C. Suen and Leung, 2003). Recently, the SVM derived from statistical learning theory has been widely applied to pattern recognition and regression estimation. This is well known as a learning machine with good generalization ability in recognition systems for handwritten characters. Since this approach was originally devised for binary classification problems, multi-class classification problems are often solved using voting schemes based on the combination of binary decision functions. Now there are some methods to solve N-class pattern recognition problem. One approach is one-to-one strategy, where $\frac{N(N-1)}{2}$ binary classifiers are constructed - separating each pair of classes. Another approach is to build a hierarchy or tree of binary classifiers such as Decision Directed Acyclic Graph (DDAG)-based SVMs (Ken-ich Maruyama and Nakano, 2002). The other approaches are shown in many researches: (Schwenker, 2000) (Hyun-Chul Kim and Bang, 2002).

In this paper, modified algorithms for optimal threshold and skew correction are proposed to extract digits with good quality. And two modified multi-class SVM structures for practical classification problems are discussed. There are one-to-one structure and DDAG-type dual-tournament structure which use potential classes. They consist two procedures. First of all, potential classes as candidates is selected. And then the final class among the selected classes are recognized. They also represent similar recognition rates. But the recognition speed of dual-tournament structure is faster than that of one-to-one structure. Therefore DDAG-type dual-tournament structure is suitable for on-line system.

This paper is composed as follows. Two proposed preprocessing techniques are described in section 2. The existing multi-class SVM structures are briefly reviewed in section 3. The proposed multi-class SVM structures are presented in section 4. The performance of these algorithms is shown in section 5.

2. PREPROCESSING

2.1 Optimal global threshold

It is shown the cropped image of a billet in Figure 4. A billet image is composed of 10-template-marked and 2-handwritten digits. In the preprocessing, the region of digits from a billet image is roughly cropped and then this image is filtered to remove the noise. The gray-level image is also converted to the binary image. This process is required to extract digits exactly. By the

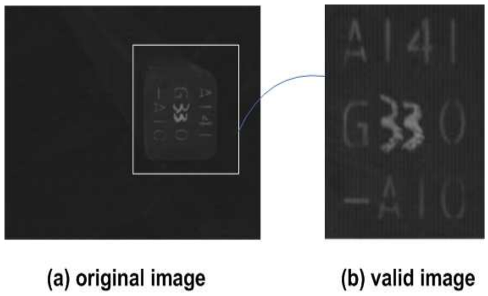

Fig. 4. Cropping process of a template-marked billet image

way, handwritten digits are often brighter than template-marked digits. If the degrees of brightness between template-marked and handwritten digits are large, template-marked digits become black by the global threshold of the gray-level image. Only handwritten digits appear as shown in Figure 5(a). To make up for this drawback,

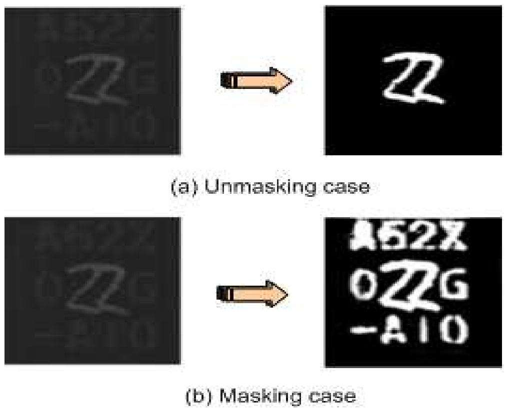

Fig. 5. Binarization

the masking method is used. The local threshold of each template-marked region is calculated by masking another region. And then the lowest threshold is selected. Binarization is taken by this value. The result of masking method is shown in Figure 5(b)

2.2 Skew correction

The binary image made by proposed masking technique should be divided into 3-row by 4-column digits. If the marked digits are not aligned with the coordinate system, they are not easily segmented and recognized. In this paper, The modified skew correction algorithm is proposed. The first row of the gray-level image is extracted by the area of the first row in the binary image. And then the local threshold values of the each four digits are calculated in the first row of the gray-level image. They are applied to the calculation of the skew angle. It is calculated by each position of first-row digits. The skewed digits are corrected by this value. The result of skew correction is shown in Figure 6.

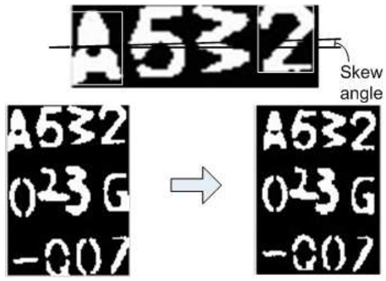

Fig. 6. Correction for the skew of digits

3. MULTI-CLASS SVM

3.1 Support Vector Machine (SVM)

The SVM is a powerful binary classification technique proposed by Vapnik et al (Ken-ich Maruyama and Nakano, 2002)(F. Wang and Schomaker, 2000). In this algorithm, an input vector is mapped into a high-dimensional space by a non-linear function, and then a linear discriminant function is searched in the high-dimensional feature space. The learning based on the margin maximization is done by solving the following quadratic programming optimization problem:

$$\max \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) \quad (1)$$

$$subject\ \ to \quad \sum_{i=1}^{N} \alpha_i y_i = 0, \quad 0 \leq \alpha_i \leq C$$

,where $\mathbf{x}_i$ denotes the learning data, N denotes the number of the learning data, α denotes the lagrange multiplier for $\mathbf{x}_i$ and $y_i \in \{-1, \ +1\}$ denotes the label of $\mathbf{x}_i$. The support vectors satisfy the condition: $0 < \alpha_i < C$, where C is the parameter which penalizes the learning errors.

The decision function for an input vector $\mathbf{z}$ is defined as

$$f(\mathbf{z}) = \sum_{i \in SV} \alpha_i y_k K(\mathbf{z}, \mathbf{x}_i) - b \qquad (2)$$

,where $\mathbf{x}_i$ denotes a learning vector, SV denotes the set of support vectors and $K(\mathbf{z}, \mathbf{x}_i)$ is a kernel function. The class is determined by the sign of function $f(\mathbf{z})$. From several kernel functions, Gaussian kernel is applied in this paper.

$$K(\mathbf{x}, \mathbf{y}) = \exp(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{\sigma^2}) \qquad (3)$$

Since the SVM has good generalization, it has been used in pattern recognition. But the SVM is originally devised for binary classification, it must be extended to solve multi-class classification such as numeral, character and face recognition. Three strategies are constructed below (Zhao Bin and Shao-wei, 2000).

3.2 One-to-one Structure

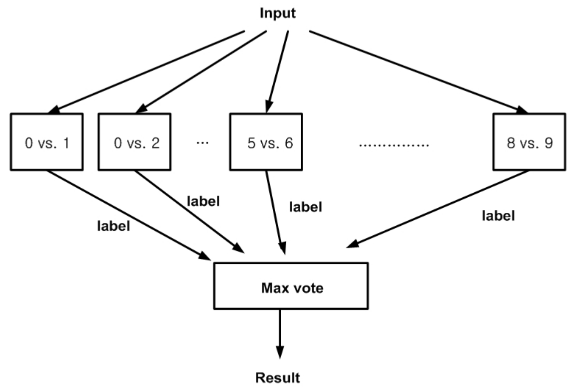

Fig. 7. One-to-one SVM

The structure of one-to-one SVM is shown in Figure 7. Becuase a SVM is constructed by training two classes, there are $N(N-1)/2$ SVMs for N-class problem. All SVMs are used to classify test samples. Supposed a SVM is constructed by the i-th and j-th class, d_{ij} is defined below:

$$d_{ij}(\mathbf{x}) = \begin{cases} +1, & \mathbf{x} \in class\ i \\ -1, & \mathbf{x} \in class\ j \end{cases} \tag{4}$$
$$and$$
$$d_{ij}(\mathbf{x}) = -d_{ji}(\mathbf{x})$$
$$where \quad i,j = 1,2,\cdots,N \quad (i \neq j)$$

Let $S_i(\mathbf{x})$ denote the vote of each class for the input sample $\mathbf{x}$

$$S_i(\mathbf{x}) = \sum_{i,j}(d_{ij}(\mathbf{x})+1)/2 \tag{5}$$
$$where \quad i,j = 1,2,\cdots,N \quad (i \neq j)$$

The input sample belongs to the i-th class if there exists

$$class \quad of \quad \mathbf{x} = \arg_i \max S_i(\mathbf{x}) \tag{6}$$

3.3 One-to-rests Structure

If the j-th class is all classes except the i-th class, the samples of the i-th class are labeled as '+1', and the samples of the j-th class are labeled as '-1' to train samples. For the N-class problem, N SVMs are needs.

For the input sample $\mathbf{x}$,

$$d_i(\mathbf{x}) = \mathbf{w}_i \cdot \mathbf{x} + b_i \tag{7}$$
$$where \quad i,j = 1,2,\cdots,N$$

The input sample belongs to the i-th class if there exists

$$class \quad of \quad \mathbf{x} = \arg_i \max d_i(\mathbf{x}) \tag{8}$$

3.4 Hierarchical Structure

The SVM which basically treats a two-class problem is extended to multi-decision levels with the idea of the decision tree. A SVM is the same as each branch of the tree. A classified class is denoted as a leaf. 'Divided and Conquer' is applied to this structure. Because the SVM of the high-level classifies more samples and the SVM of low-level classifies fewer samples, they are classified more precisely. But this structure is affected by the initial grouping.

4. PROPOSED MULTI-CLASS SVM

As mentioned before, one-to-one SVM usually uses the maximum voting to decide the final class. When some classes have similar votes, the maximum voting method has the risk that causes the wrong classification. It is necessary to revaluate potential classes with similar votes. But, K-NN and existing multi-class SVMs are not suited to template-marked digits with partially erased parts. Therefore, two algorithms which select potential classes are proposed: Potential one-to-one SVM and Dual-DAG SVM.

4.1 Potential one-to-one SVM (PSVM)

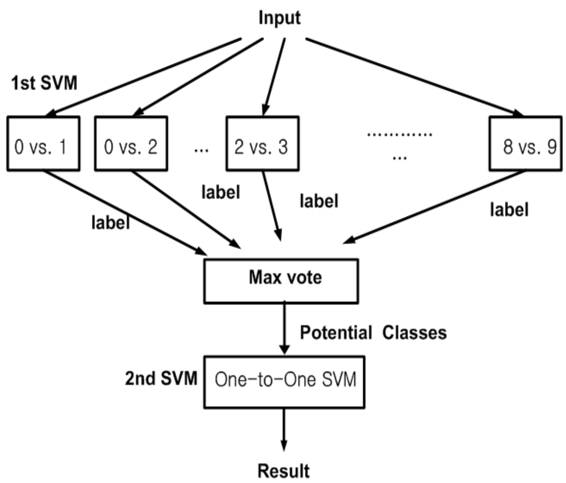

Fig. 8. Potential one-to-one SVM

The structure of Potential one-to-one SVM is shown in Figure 8. It selects potential classes using all possible matches. The classes with more than six votes are regarded as potential classes in this structure. Since each potential class is revaluated using the SVM with different kernel, this structure can find the final class more exactly. But, it needs many calculations. If the N_p out of N are selected as potential classes, $\frac{N(N-1)}{2} + \frac{N_p(N_p-1)}{2}$ SVMs are required. Thus, It's not good for the fast recognition application.

4.2 Dual-DAG SVM (DDSVM)

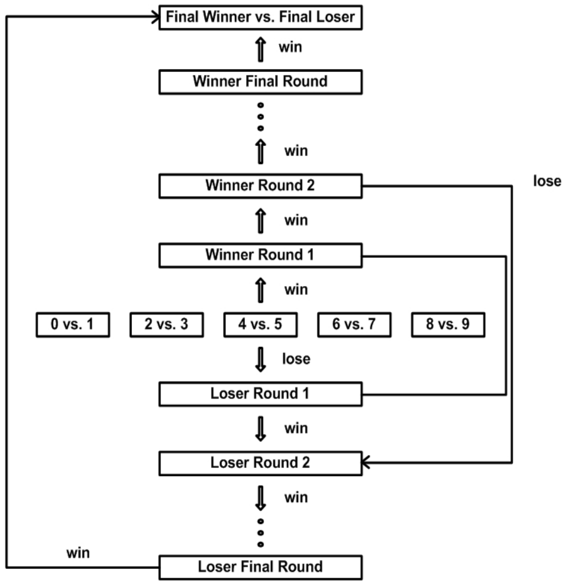

Fig. 9. Dual-DAG SVM

As mentioned before, the recognition speed of PSVM is slower than that of the conventional algorithm due to revaluating the potential classes. To improve the recognition speed, tree structure can be used. The structure of DDSVM is shown in Figure 9. Especially, DDAG is attractive.The DDAG eliminates a class at each step from the first step to the final step (Ken-ich Maruyama and Nakano, 2002). Since a loser class at match can not be the final class again, the match structure of the DDAG must be designed well. It can shorten the recognition time, but cannot select potential classes. Because reasonable method are required, It's proposed the DDSVM that does not eliminate a loser class at once but give one more chance to a loser class. The performance of the DDSVM is less affected than that of DDAG by the initial array of matches. The classes in the final winner round and loser round become potential classes. One of them is decided as the result of the recognition.

5. EXPERIMENTAL RESULTS

5.1 Recognition of segmented digits

The template-marked digits in billet images acquired from the continuous-casting plant of POSCO are used. The number of classes is 10. The data set used to evaluate the performance of the algorithms consists of 2000 template-marked digits. The normalized digits are represented as a 16 $\times$ 16 matrix. The data set consists of each set of 1,000 training samples and 1,000 test samples. The training set has been used to design the parameters of the algorithms. The test set has been used to evaluate the performance of the algorithms. For this data set, training condition and experimental results are as follows. The kernel parameters, C is 1000 or 100 and σ is 10, are used. Experimental results is shown in Table 5.1. The recognition rates for one-to-one SVM, PSVM and DDSVM are similar. K-NN method shows less performance than the others. In spite of good recognition rates for the other digits, K-NN causes errors in the recognition of number 7. PSVM shows the best classification performance. DDSVM also shows the fast recognition speed and the acceptable recognition rate in the recognition system.

Table 1. Recognition performance

Method	Recog. rate(%)	Recog. speed (s/digit)	Worst rate(Num.)
K-NN	97.2	0.03	81 (7)
DDSVM	98.3	0.21	88 (8)
OneSVM	98.5	0.34	90 (8)
PSVM	98.9	0.35	93 (8)

where OneSVM is one-to-one SVM

5.2 Recognition of Billet Digit Data

The proposed algorithms, PSVM and DDSVM are applied to the recognition system. The data set used to evaluate the performance of the algorithms. It consists of template-marked digits preprocessed and segmented 500 billet images by the proposed methods. It is shown the performance of proposed preprocessing in Table 2. Segmentation error are reduce using the proposed preprocessing techniques.

Table 2. Segmentation performance

Method	Seg. rate(%)	Num. of error
Conventional preprocessing	94.8	26
Proposed preprocessing	98.8	6

The parameters of the algorithms are same as those in the previous experiment. The recognition rate is calculated by the following equation :

$$R(\%) = \frac{D_{total} - D_{seg.err} - D_{recog.err}}{D_{total} - D_{seg.err}} \times 100. \quad (9)$$

,where D_{total} denotes the number of the total samples, $D_{seg.err}$ denotes the number of the segmentation errors and $D_{recog.err}$ denotes the number of the recognition errors. For this data set, the results are shown in Table 5.2. PSVM show the best classification performance, but it has the slow recognition speed. DDSVM shows the fast recognition speed and the acceptable recognition rate in the recognition system.

Table 3. Recognition performance

Method	Recog. rate(%)	Num. of error	Recog. speed (s/digit)
DDSVM	92.1	39	3.8
PSVM	93	33	4.5

6. CONCLUSION

In this paper, the preprocessing techniques and SVM-based algorithms are proposed to recognize template-marked digits in billets. For systematic preprocessing approaches, threshold decision technique using masking method and skew correction method are also proposed. The threshold is calculated to extract digits from the noisy background by the proposed masking method. The recognition rate of skewed digits in billets are improved by the proposed skew correction method.

In recognition algorithm, multi-class SVM-based algorithms are proposed to improve the generalization property. PSVM and DDSVM are introduced to select potential classes. PSVM improves recognition rate. But, it has slower recognition speed than one-to-one SVM. DDSVM has fast recognition speed and good recognition rate for on-line application. Especially, the performance of this method is not affected by the intial array of classes. Experimental results show proposed algorithms have better performance than existing algorithms.

REFERENCES

Bunke, H. and editors P. Wang (1997). Handbook of character recognition and document image analysis. *World Scientific.*

Bunke, Horst (2003). Recognition of cursive roman handwriting - past, present and future. *7th Int. Conference on Document Analysis and Recognition.*

C. Suen, S. Mori, S. Kim and C. Leung (2003). Analysis and recognition of asian scripts - the state of the art. *In 7th Int. Conference on Document Analysis and Recognition.*

F. Wang, . Vuurpijl and L. Schomaker (2000). Support vector machines for the classification of western handwritten capitals. *In 7th Int. Workshop on Frontiers in Handwriting Recognition* pp. 167–176.

Hyun-Chul Kim, Shanoning Pang, Hong-Mo Je Daijin Kim and Sung Yang Bang (2002). Pattern classification using support vector machine ensemble. *8th Int. Workshop on Frotiers in Handwriting Recognition.*

J. Liu, P. Gader (2002). Neural networks with enhanced outlier rejection ability for off-line handwritten word recognition. *Int. Journal on Document Analysis and Recognition* **35**, 2061–2071.

Ken-ich Maruyama, Minoru Maruyama, Hidetoshi Miyao and Yasuaki Nakano (2002). Handprinted hiragana recognition using support vector machines. *8th Int. Workshop on Frotiers in Handwriting Recognition.*

M. Morita, J. Facon, F. Bortolozzi S. Garnes and R. Sabourin (1999). Mathematical morphology aand weighted least squares to correct handwriting baseline skew. *5th Int. Conference on Document Analysis and Recognition* pp. 430–433.

O'Gorman, L. and editors R. Jasturi (1995). *Document Image Analysis.* IEEE Computer Society Press.

S. Mori, H. Nishida and H. Yamada (1999). *Optical Character Recognition.* John Wiley and Sons.

Schwenker, Friedhelm (2000). Hierarchical support vector machines for multi-class pattern recognition. *4th Int. Conference on knowledge-Based Intelligent Engineering Systems and Allied Technologies.*

Zhao Bin, Liu Yong and Xia Shao-wei (2000). Supprot vector machine and its application in handwritten numerical recognition. *In 7th Int. Workshop on Frontiers in Handwriting Recognition* pp. 720–721.