Camera-based Document Recognition Using Hierarchical Classifier

[1]KYE KYUNG KIM, [2]JIN HO KIM, [1]IN HO LEE
[1] Electronics Telecommunications Research Institute, Korea
[2] Kyungil University, Kyungsan, Kyungpook, Korea

Abstract: - This paper is to introduce a camera document recognition using hierarchical classifier. Generally, it is very difficult to recognize camera document image because of inconsistent input condition and camera problem itself. Therefore, Implementation of hierarchical classifier is needed to guarantee recognition performance of camera document image. The camera document goes through four processes such as preprocessing, segmentation and recognition modules. Input document image is evaluated by image enhancement algorithm and local threshold method. Characters are segmented using merging and split method, which is based on structural information of characters. Combined feature extractor and hierarchical MLPs have used to recognize a segmented character. We obtained an encouraging recognition result for camera document by combining image patches. We have experimented with ETRI document database and an encouraging recognition result 94.6% has obtained.

Key-Words: - Camera document recognition; Structural features; Hierarchical classifier

1 Introduction

The recognition of camera-based document has been developed intensively due to wide spread of camera [1-6]. Camera document recognition has applied to real world application fields such as reading book by robot, implementing text DB by camera and recognizing signboard by PDA, *etc*. However, restrictions of illumination, input condition, camera resolution, *etc* have considered to applied camera document recognition to real world.

Generally, document image captured by camera has severe noises due to illumination and camera problem itself such as tilt, distortion [4-6]. Unlike the precisely controlled capturing procedure of the scanner, the camera-based document image is not suitable to recognize characters of document. The camera document image is frequently blurred due to illumination or focus. And also the captured document image is not only tilted but also is distorted in perspective due to the variable viewing angles. All of these handicaps forces more sophisticated preprocessing algorithm, which is obviously one of the most important research themes for camera-based document recognition [4-6]. Especially, binarization and character extraction methods have been studied closely for recognizing camera document. Binarization of camera document is preprocessing for extracting characters from background image. Global thresholding and local thresholding methods have proposed [2-3]. However, it has still problem of extracting characters from background due to illumination effect. The other challenging problems have still occurred to recognize segmented characters because of the distortion of camera, illumination effect and diverse character shape. Therefore, we have tried to get high recognition performance for camera-based document image. Combined feature extraction method and hybrid recognizer are needed to recognize more complicated document recognition such as camera document image.

In this paper, camera-based document image is recognized using hierarchical classifier. At the first, enhancement algorithm is applied to camera document image. Binary image by local thresholding method is used to extract character region. An individual character is segmented by merging and split method, which is based on structural information of a character. And also overlapped character is split in segmentation module. Combined structural and statistical features are extracted to the individual character. At that time, pixel-shifted mesh feature is extracted for character with broken stroke. Hierarchical classifiers are used to recognize camera document images. Type classifier and character classifier of each type are implemented by MLPs. That is for evaluating recognition performance and for reducing recognition error.

2 Configurations

Configuration for the camera-based document recognition is shown in Fig. 1.

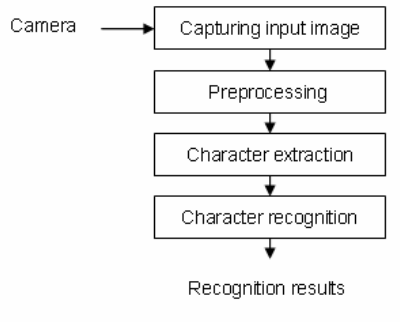


Fig. 1. System configuration for recognizing camera document image.

It is consisted of a camera as input device and six processing modules, which are capturing document image, preprocessing, character extraction and character recognition.

3 Preprocessing

Camera document image is far inferior compared to those captured by scanner in image quality. Image variation, distortion and blur have affected adversely to recognize camera document image. To resolve these kinds of challenging problems, a strong image preprocessing is essential. Example of camera document image is shown in Fig. 2.

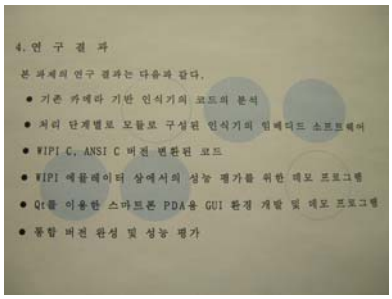


Fig. 2. The example of camera document image.

Gray-level normalization is applied to image patches. It is called histogram stretching. It is an image enhancement algorithm that provides to reduce affect of noise and illumination. The formulas for pixel level after gray level normalization is described as follows.

$$f_1(x,y) = (L-1)\frac{f(x,y)-\min}{\max-\min} \quad (1)$$

where, $\max = \max[f(x,y)]$, *for* $1 \le x \le M$ *and* $1 \le y \le M$

$\min = \min[f(x,y)]$, *for* $1 \le x \le M$ *and* $1 \le y \le M$

, where $f(x, y)$ and $f_1(x, y)$ denote the level of pixel (x, y), the pixel level after image enhancement, respectively. L denotes gray level range of image to be converted and M denotes the height and width of image. **max** and **min** are maximum value and minimum value among pixels in image, respectively.

Local thresholding method is applied to sub document images. Camera document image is binarized using the difference of maximum and minimum intensity value of pixels in sub-window. The sub-window is decided with the virtual height of character string. The local binarization method is compared to pixel-oriented one and global binarization one. It provides better binarization result and high speed processing time.

4 Character segmentation

In character extraction step, character regions are extracted from background in binary image. At the first, small blob is called as noise which is removed and image is also removed. Small blobs may occur in binarizing document image and they are removed by comparing the number of pixels of connected components. The blobs below threshold are removed in this procedure. The threshold is calculated empirically.

An individual character has extracted through extraction of character string and word. Character string is extracted by horizontal projection and word is extracted by vertical projection. Projection features are accumulated pixel sum of connected components of horizontal x direction and vertical y directions, respectively.

An individual character is segmented by follow conditions. In this paper, mixed character, Hangul and non-Hangul are used to recognition target. Non-Hangul is composed of numeral, symbol and English alphabets. Segmentation process has done by analyzing the structural features of characters.

- Condition 1 : $x_s^i - x_{c_s} < y_h$

  If the width among connected components is less than the height of character string, $y_{h,}$

then, connected components are considered as a character.

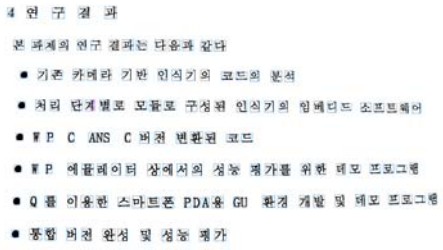
- Condition 2 : $x_e^i - x_{c_s} < y_h + \theta$

  If the width of connected components in virtual character width is less than the height of character string and theta.

  then, connected components are considered as a character.

It is considered as a character if the above two conditions are satisfied. If the first condition is satisfied and the second condition is not satisfied, then it is considered as touching characters. Touching characters are segmented into an individual character by analyzing touching types among six kinds of Hangul and touching types of alphabets. The segmentation processes of characters in sub-document image are shown in Fig. 3.

(a) (b)



(c)

Fig. 3. Example images for extraction of (a) character string, (b) word and (c) individual character.

5 Character recognition

5.1 Feature extraction

For recognition of segmented characters, features from binary character image have extracted. Three kinds of features are extracted including mesh, chain code, and distance of normalized character image as shown in Fig. 4.

Fig. 4. The example of normalized character.

Mesh features by summing foreground pixels in sub-windows of segmented image are extracted from smoothed binary images. Chain code features of *k* directional slopes in sub-windows. Chain code direction feature has been extracted from the contour of a normalized character image. The distance features are distances from the image boundary to the first foreground pixel.

5.2 Character recognition

In character recognition step, we have tried to improve the performance of recognition for mixed Hangul (Korean character) and non-Hangul. MLP classifiers have implemented for type classification and character recognition. Type classifier discriminates six kinds of Hangul and one kind of non-Hangul. Classified character is recognized in each type character recognizer. Six kinds of Hangul type is defined by the structural features of Hangul.

Structure of hierarchical classifier

All characters are grouped into six kinds of character types depend on their shape in Hangul. Other one group is for non-Hangul which is for numeral, symbol and alphabets. Type classifier is shown in Fig.5. Consequently seven independent classifiers are implemented for recognizing characters as shown in Fig. 6 (b). We have tested grouping schemes over 1084 characters including 1000 of Hangul and 84 of non-Hangul.

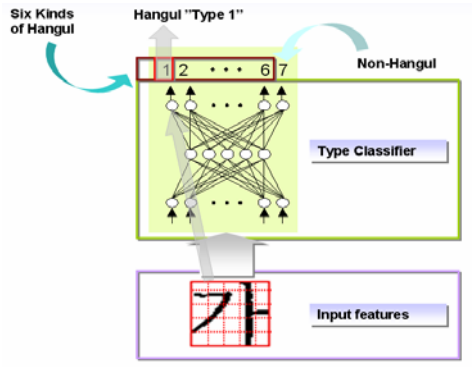
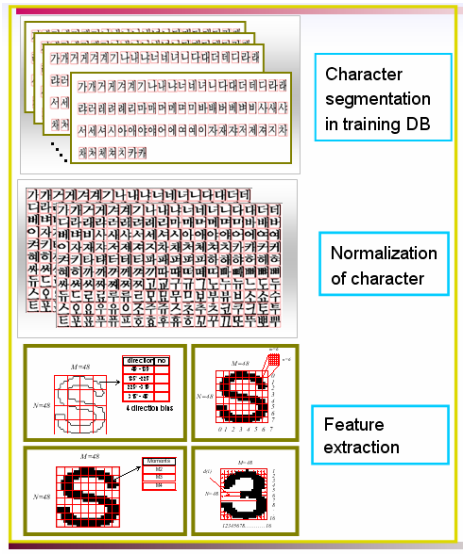
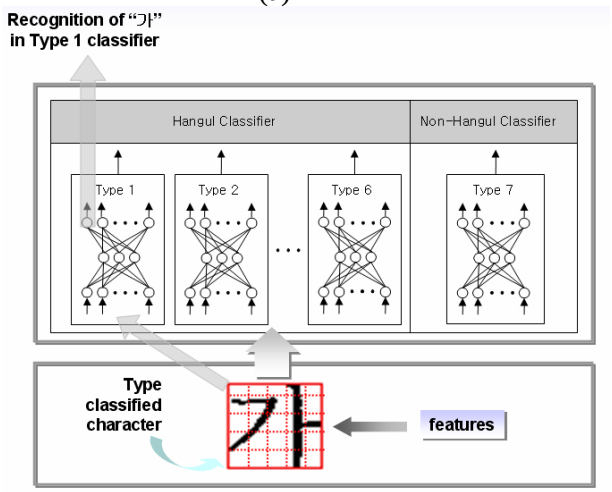


Fig. 5. The example of neural network classifier.

Characters have trained and tested with ETRI document DB. Fig. 6. (a) shows feature extraction processes with training DB that includes character segmentation, normalization and feature extraction in training step. In recognition, hierarchical classifier including six classifiers for Hangul and one classifier for non-Hangul have implemented as shown in Fig. 6 (b).



(a)



(b)

Fig. 6. Recognition processes including (a) feature extraction and (b) implementation of hierarchical classifier.

The overall system of hierarchical classifier recognizes 1084 classes of characters. Each type of character classifier has been trained using its own character type. We have tested hierarchical classifier scheme over seven types of characters. The performance of hierarchical neural network classifier has increased more than other single classifier. And also each dedicated classifier has refined to increase its own performance.

6 Conclusion

This paper described camera document recognition using hierarchical classifier. To recognize camera document, we designed character recognition system with compositing classifiers and features. Enhancement algorithm and local thresholding are applied to camera document image in preprocessing module. An individual character is segmented by merging and split method, which is based on structural information of a character. Hierarchical classifiers including type classifier and each type of character classifier are used to recognize camera document images. Type classifier and each type of character classifier are implemented by MLPs. The proposed method is for evaluating recognition performance and for reducing recognition error.

References:

[1] H. Fujisawa, H. Sako, Y. Okada, and S. W. Lee, "Information capturing camera and developmental issues," *Proc. of the 6th ICDAR*, pp. 205-208, 2001.

[2] M. Seeger and C. Dance, "Binarizing camera images for OCR," *Proc. of the 6th ICDAR*, pp. 54-58, 2001.

[3] L. Fan, L. Fan and C. L. Tan, "Binarizing document image using coplanar prefilter," *Proc. of the 6th ICDAR*, pp. 34-38, 2001.

[4] M. Sawaguchi, K. Yamamoto, and K. Kato, "A proposal of character recognition method for low resolution images by using cellular phone," *Proc. of the 9th Korean-Japan joint workshop FCV*, pp. 216-221, 2003.

[5] K. Wang, J. A. Kangas, and W. Li, "Character segmentation of color images from digital camera," *Proc. of the 6th ICDAR*, pp. 210-214, 2001.

[6] K. K. Kim, S. Y. Chi, Y. K. Chung, and S. K. Park, "Camera document recognition system robust to lighting condition," *The workshop of CVPR*, pp. 90-92, 2002.

[7] J.H. Kim, K.K. Kim and C.Y. Suen, "An HMM-MLP Hybrid Model for Cursive Script Recognition", *Pattern Analysis and Applications*, vol. 3, issue 4, pp. 314-324, 2000.