利用決策樹分類的技術作名片影像的自動分析

Automatic Analysis of Name Card Contents by Image Processing and Decision-Tree Classification Techniques

Ya-Ru Yen (顏雅茹) and Wen-Hsiang Tsai (蔡文祥)

Department of Computer & Information Science

National Chiao Tung University

Emails: whtsai@cis.nctu.edu.tw & whtsai@cis.nctu.edu.tw

摘　要

在本論文中我們利用影像分析的技巧，提出一些可以自動分析名片影像的方法。影像分析的主要工作，是希望能取出名片有用的資訊。在我們的方法中，有五個階段：基本區塊的抽取、商標的抽取、名片型別的分類、文字列型別的分類，以及名片影像的重建展示。在基本區塊的抽取階段，我們使用邊緣偵測及區塊生長演算法來找出基本區塊，再用矩量保持二值化來減少基本區塊裡的顏色。在商標的抽取階段，我們利用區塊的顏色以及相關資訊，判斷這些區塊為商標或是文字區塊。在名片型別的分類階段，我們利用文字區塊寬與高的比例來區分名片的型別，分成中文名片和英文名片。在文字列型別的分類階段，在中文名片中我們處理九種文字列型別，分別是姓名、公司名稱、電子信箱、網頁位址、行動電話、傳真電話，電話號碼、統一編號、地址。英文名片中的文字列型別除了統一編號之外與中文名片的相同。我們利用適當的決策樹提出對中文名片和英文名片的文字列的型別作分類的分法。最後，我們利用一個適當的技術來壓縮名片影像中的組成成分，以降低儲存空間。良好的實驗結果，顯示出我們所提方法的可行性。

關鍵詞：名片影像、基本區塊、擷取商標、文字列的型態、決策樹分類。

Abstract

Automatic analysis of name card contents using image processing and decision-tree classification techniques is proposed. First, basic blocks in name card images are extracted by edge detection and region growing, and the colors in each basic block are reduced by the moment-preserving thresholding technique into two representative ones. Several effective features are proposed to classify extracted blocks into logo blocks and text blocks. The width/height radios of text blocks are used to classify card types into Chinese and English. Furthermore, adaptive decision-tree methods for classifying text lines both in Chinese and in English name cards are proposed. Nine types of text lines are recognized, including name, title, e-mail, web address, mobile phone number, fax number, phone number, government publications number, and address. Finally, a suitable compression method is employed to reduce the data volumes of the recognized name card contents to save storage space and display time. Good experimental results show the feasibility of the proposed method.

Keywords: name card image, basic block, log extraction, text line type, moment-preserving thresholding, decision-tree classification.

1. Introduction

Nowadays, it is a common way to introduce oneself in social meetings to other people through the use of name cards. To save their meaningful data, scanning name cards and transforming the resulting images into digital form is desired. This leads to the need of name card image analysis [1]. The purpose of such analysis is to extract useful and meaningful components from name card images. A logo is an emblem of a company or organization. Consequently, it is desired to extract logos effectively from name card images. A name card has two sides, with one side composed of characters of the native language and the other usually of English. For efficiency of card content analysis, it is necessary to classify card types automatically at the beginning. In order to extract useful information in name cards, like name and address, we need a good method for segmentation of meaningful text lines. Consequently, it is desired to investigate suitable text line segmentation and classification methods for different card types. Then, it will be easier and faster to recognize characters in the text lines

through an optical character recognition (OCR) system [8-9], according to the attributes of text line types. In this study, it is attempted to propose proper algorithms to extract logos, to classify name card types, and to segment and classify text line types for name card content analysis.

In the proposed method, name cards are scanned into BMP files first. Then edge detection and region growing techniques [2] are employed for background elimination and basic block extraction. Furthermore, a thresholding approach [3] is applied to reduce colors in a basic block into two major colors, one being the foreground color and the other the background color. Based on the use of foreground colors and some effective features [4-6] of logo blocks, logo blocks are extracted effectively. By excluding the logo block, a set of text blocks is obtained. Based on the use of some features of the character types [7], name card types are then determined. Adaptive text block segmentation and text line segmentation [13] of Chinese or English name card images are also conducted, followed by recognition of text line types using adaptive features [10-11] of some Chinese or English characters in the text lines. In this study, text lines in Chinese name card images are classified into nine types, including (1) name line, (2) title line, (3) phone number line, (4) fax number line, (5) mobile number line, (6) Government Publications Number line, (7) e-mail line, (8) web address line, and (9) address line. And text line types in English name cards are classified identically but without the Government Publications Number line. After the above-mentioned processes are completed, all meaningful name card components are extracted. A compression technique [12] is then utilized to compress the component data. Finally, a friendly interface for displaying the resulting name card content is provided. An overall illustration of the proposed system for name card image analysis and display is shown in Fig. 1.

In the remainder of this paper, employed methods for preprocessing name card images are described in Section 2. In Section 3, the proposed methods for logo block extraction and card type classification are described. In Section 4, the employed methods for text line segmentation are introduced first, followed by the descriptions of the proposed method for text line classification. In Sections 5, the proposed classification procedure for English text lines is presented and in Section 6, the employed method for name card content compression is described. Finally, some experimental results are shown in Section 7, with some conclusions given in Section 8.

2. Basic Block Extraction

Basic block is referred to in this study as the smallest unit in the name card image. The proposed steps for extracting basic blocks are described in the following.

A. Edge detection and region growing

Characters are composed of edges. Therefore, the Sobel operator is employed to detect edges in an input name card image as the first step of card content analysis. Then, a region growing algorithm is used to construct basic blocks from the resulting edge points by circumscribing them with rectangular shapes.
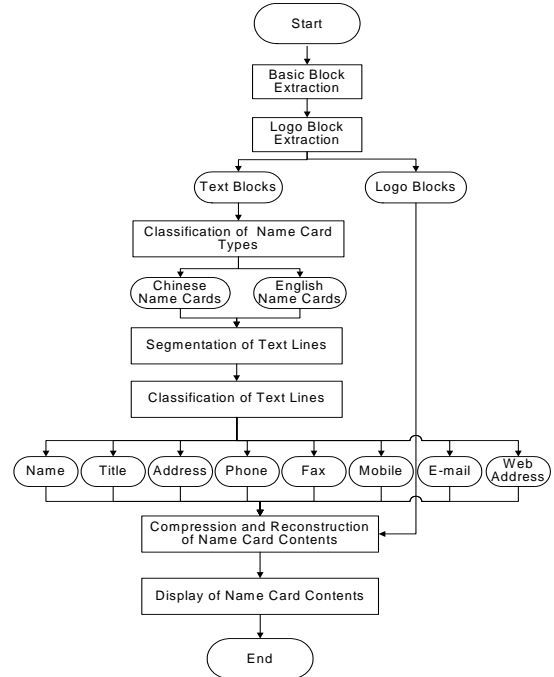

Fig.1. Processes of proposed system for a name card image analysis and display.

B. Basic block merging by geometric position analysis and noise block elimination

After region growing, we have obtained a set of basic blocks. They might overlap or close to one another. Such basic blocks will affect the result of later processing. Therefore, we merge these basic blocks into larger ones according to their geometric positions.

C. Moment-preserving thresholding of basic block images

A basic block in a name card image usually is small in size and includes just a character. It is sufficient to use two colors to represent the colors in a basic block. Each name card image in the RGB color model consists of three independent image planes. We apply the moment-preserving thresholding technique proposed by Tsai [2] to the three color planes respectively to obtain two representative gray values for each plane. We then compose the results to obtain two *representative*

2

colors for each basic block. Each pixel in a basic block is finally assigned a representative color according the Euclidean distance measure.

D. Determination of foreground colors in basic blocks

The foreground pixels of a character are mostly distributed over the central part of a basic block. On the contrary, the background pixels are mostly distributed over the surrounding region of a basic block. Checking the majority color of the surrounding pixels in a basic block, we can determine the *foreground* and *background* colors in a basic block. More specifically, we collect the surrounding pixels in the top and bottom rows and leftmost and rightmost columns of a basic block, and take the majority of the representative colors of these surrounding pixels to be the background color of the basic block. The other representative color of the basic block is then taken to be the foreground color.

3. Logo Block Extraction and Card Type Classification

We describe first the features and the algorithm proposed in the study for basic block classification, then the proposed algorithm for basic block extraction, and finally the proposed algorithm for classification of card types into English or Chinese.

3.1 Features and Algorithm for Classification of Basic Blocks

Based on our observations and experiments, the following features of basic blocks are effective for basic block classifications: foreground color, height, and position. A logo in a name card usually is emphasized with a special color and a larger height, and often appears in one corner of the card. A *text logo* is a set of texts that composes a logo while a *graphic logo* is a graphic picture. The features employed in this study for recognizing graphic and text logos are those as described above, but for text logos two additional properties are used. First, the central points of all the basic blocks in a text logo almost line up. Second, the width of a text logo is usually not larger than half of the width of a name card image. All these features or properties will be referred to integrally as *text logo features*.

To classify basic blocks into logo blocks and others, a procedure of two stages are proposed, namely, coarse classification by *k-means clustering* followed by a detailed classification, called *logo extraction*, which is described in the next section. The first stage is to classify the basic blocks coarsely into two sets, *possible log blocks* and *text blocks*. The input to the proposed algorithm for the first stage is a given set of n basic blocks, each with

a 4-feature vector including the R, G, and B values of the foreground color, and the height of the block. Also, we take k to be 2 for the k-means clustering process used in the algorithm.

Algorithm 1. Coarse classification of basic blocks by k-means clustering.

Step 1: Sort all the feature vectors X_i of the basic blocks by their height values in a decreasing order.
Step 2: Select two initial cluster centers for two types of blocks in the following way:
    2.1 assign the feature vector with the largest height value as the initial center C_1 for the cluster of possible logo blocks;
    2.2 assign the feature vector with the smallest height value as the other initial center C_2 for the cluster of text blocks.
Step 3: Assign each feature vector X_i to the nearest center c_k according to the Euclidean distance computed as $d(X_i, C_i) = |X_i - C_k|$ where $k = 1$ and 2. That is, assign X_i to the C_k, which is closer.
Step 4: Compute the new center for each cluster C_k as the mean of all the feature vectors in C_k.
Step 5: Repeat the previous two steps until the number of iterations reaches a pre-defined value, and take the final two clusters to be the desired sets of *possible logo blocks* and *text blocks*, respectively.

3.2 Logo Block Extraction by Decision-Tree Classification

It is assumed in this study that there is only one graphic logo or one text logo in a name card image. We design a decision-tree classification scheme for extracting logo blocks. In the scheme, we use first two algorithms to confirm the existences of the text logo blocks or the graphic logo block in a name card, respectively. Then, we use a third algorithm for logo block extraction. Here by *text logo blocks*, we mean those basic blocks that compose a text logo. And by a *graphic logo block*, we mean the basic block that is the graphic logo itself.

A. Confirming text logo blocks

The input to the first algorithm for confirming the existence of text logo blocks is a given set S of *possible logo blocks* obtained from Algorithm 1, each with a feature vector $X_i = (x_i, y_i, t_i, b_i)$, where x and y represent respectively the x- and y-coordinates of the central point, and t and b represent respectively the y-coordinates of the block's top and bottom boundaries.

Algorithm 2. Confirming text logo blocks in a name card.

Step 1: Select from S the block B with the largest

height and let its feature vector be $X_B = (x_B, y_B, t_B, b_B)$.

Step 2: Compute the x-coordinate of the vertical centerline L_v of the name card image, and denote it by x_{vcl}.

Step 3: Check each basic block B_i with feature vector $X_i = (x_i, y_i, t_i, b_i)$ in S to see if it satisfies the following conditions:
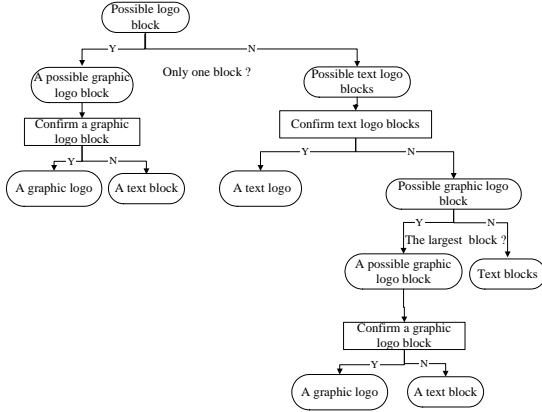   (1) being located roughly on the horizontal

Fig.2. Decision tree for detailed classification of possible logo blocks.

line on which B is located, that is, $t_B \le y_i \le b_B$;
   (2) being located on the same side as B with respect to the centerline of the name card, that is, both $x_B \le x_{vcl}$ and $x_i \le x_{vcl}$, or both $x_B > x_{vcl}$ and $x_i > x_{vcl}$.

Step 4: If all the blocks in S meet the above conditions, then decide that these blocks are text logo blocks that compose a text logo; otherwise, decide them to be possible graphic logo blocks.

B. Confirming a graphic logo block

The input to the second algorithm for conforming the existence of a graphic logo block is a possible logo block B with feature vector $X_B = (x_B, y_B, t_B, b_B)$ and a given set S of n basic blocks, each with a feature vectors $X_i = (x_i, y_i, t_i, b_i)$, where x, y, t, and b are features as described previously.

Algorithm 3. Confirming a graphic logo block in a name card.

Step 1: Collect from S those basic block B_i with feature vector $X_i = (x_i, y_i, t_i, b_i)$ which is roughly located on the horizontal line on which B is located, that is, $t_B \le y_i \le b_B$ is true.

Step 2: Check each of the basic blocks, denoted as B_i, collected in the last step with feature vector $X_i = (x_i, y_i, t_i, b_i)$ to see if they meet the following conditions.
   (1) No basic block B_i is located to the left of B, that is, $x_i \le x_B$ is true *for all i*. This

means that B appears in the top-right or the bottom-right corner of the name card.
   (2) No basic block B_i is located to the right of B, that is, $x_i > x_B$ is true *for all i*. This means that B appears in the top-left or the bottom-left corner of the name card.

Step 3: If either of the above two conditions is satisfied, then decide B to be a graphic logo block. Otherwise, regard it as a text block (*not* a text logo block).

C. Logo block extraction

An overall decision tree for detailed classification of possible logo blocks is presented in Fig. 2. And the third algorithm mentioned previously for logo block extraction is described as follows. The input to the algorithm is the set C1 of possible logo blocks yielded by Algorithm 1.

Algorithm 4. A detailed integrated algorithm for logo block extraction.

Step 1: Compute the number of blocks in C_1. If there is only one block, let the block be a *possible graphic logo block L* and go to Step 4. Otherwise, continue.

Step 2: Apply Algorithm 2 to the blocks in C_1 to decide if they are really text logo blocks. If yes, then merge them into a text logo and exit. Otherwise, continue.

Step 3: Take the block in C_1 with the largest height as a possible graphic logo block L and go to Step 4, regarding the remaining blocks in C_1 as text blocks.

Step 4: Apply Algorithm 3 to L to decide whether L is a graphic logo block or not. If yes, take it as a graphic logo; otherwise, a text block.

3.3 Classification of Name Card Types

To classify the type of a name card from its image, we employ the idea of estimating the width/height ratio range of characters with square shapes. In a Chinese name card image, most text blocks are Chinese characters. On the contrary, most text blocks are English characters in an English name card image. The width/height ratios of Chinese and English characters are useful features for classification of name card types. By our observations, characters with large heights are usually Chinese characters in a Chinese name card image. Therefore, we propose to cluster the text blocks into two clusters by the heights of the text blocks. Then, the blocks in the cluster with larger heights are used to decide the card type into Chinese and English according to a threshold value learned from our experimental experience.

4. Recognition of Text Lines in Chinese Name Cards by Decision Trees

In this section, we describe how we recognize types of text lines in Chinese name card images.

4.1 Segmentation of text blocks into characters

We use the *medium height* of text blocks, which is the medium value of the heights of the blocks, as an estimate of the height of normal text blocks. Note that this medium height is also roughly the height of Chinese characters. Also, we use the method of *x-y* cut to segment text blocks into characters. If the height of a given text block is large than the medium height, then we apply the horizontal *y*-cut operation to segment the block further into "lower" blocks. If the width/height ratio of a text block is large, then we apply the vertical *x*-cut operation to segment the text block further into "thinner" blocks that are mostly character blocks.

4.2 Segmentation of text lines

First, a horizontal cut process is applied to the text blocks. A text line so obtained may consist of several thinner text lines. Second, a vertical cut process is applied to the text line. Then, the text line is divided into two individual parts. Finally, a horizontal cut process is applied again to the two parts, respectively.

At the second stage, we use some conditions based on certain properties of text lines to decide whether the text line has to be cut vertically or not. The properties include the height L_h of the text line, the medium value M_t of the heights of the text blocks in the text line, and the smallest height L_s of the text lines in a name card image. If the text line consists of several thinner text lines, then the value of L_h is certainly much larger than M_t. And the value of L_s is used to avoid cutting e-mail and web address lines. Therefore, the values of M_t and L_s can be used to decide which text lines have to be cut vertically.

4.3 Classification of Chinese Text Lines

Text lines in a Chinese name card image are classified into nine types in this study, including name line, title line, e-mail line, web address line, mobile phone number line, fax number line, phone number line, government publications number line, and address line. On the other hand, we propose the use of the following features for Chinese text line classification.

A. Features for name and title lines

The font sizes of the characters in name and title lines are usually larger than those in other kinds of lines. Therefore, the heights of text lines are effective features. And a Chinese name contains two or three Chinese characters. The position of a name line is usually the closest one to the horizontal centerline of a name card image. And the position of a title line is usually the topmost one in a name card image.

B. Features for e-mail and web address lines

The heights of the text lines are effective features for e-mail and web address lines which usually consist of lowercase English characters, such as "a," "c," "e," "m," "n," "o," "r," "s," "u," "v," "w," and "z." And, there is a special symbol "@" in an e-mail line.

C. Features for other types of text lines

The features proposed in this study for other types of text lines are shown in Table 1. For example, the feature we propose for classification of a mobile phone number is the use of the special Chinese character "行."

To recognize the special characters to classify the text line types shown in Table 1, we propose the use of the following features: the *foreground pixel count*, the *crossing count*, and the *peripheral feature*. The foreground pixel count of a text block is defined as the ratio of the number of the pixels with the foreground color to the size of the character block. The crossing count is defined as the number of transitions from the foreground color to the background color by horizontal- or vertical-directional scanning from top to bottom or from left to right. And the peripheral feature is defined as the number of background pixels by horizontal- or vertical-directional scanning from top to bottom, from bottom to top, from left to right, or from right to left until a foreground pixel appeared.

Table 1 Features (special characters) for several types of Chinese text lines.

	Height of a text line	Most characters	Special characters
mobile phone number	small	Numeral	"行"
fax number	small	Numeral	"傳真," FAX, Fax
phone number	small	Numeral	
government publications number	small	Numeral	"統"
address	small	Chinese	"縣","市"

4.4 Proposed Method for Chinese Text Line Classification

We propose a two-stage method to classify text lines, first coarsely and then in detail.

A. Coarse classification

The first stage is to classify text lines into the four sets coarsely. We classify first text lines automatically into two sets by *K*-means clustering: one with larger heights that constitute a set of possible name and title lines, and the other a set of the remaining text lines. Then, since the possible e-mail and web address set consists of characters with special heights, such as "a" and "c," we can use this property to extract possible e-mail and web address lines from the set of the remaining text lines. Furthermore, most characters in the possible phone number set are numeral one, and most characters in the possible address set are Chinese characters. Therefore, we can discriminate the possible phone number set from the possible address set by the widths of characters. A decision tree for coarse classification of text lines based on these discussions is given in Fig. 3.
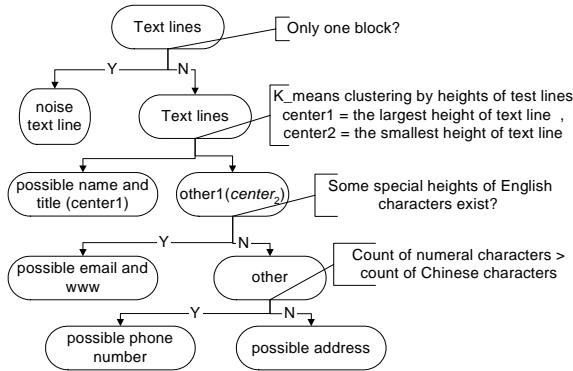

Fig.3. Decision tree for coarse classification of text lines.

B. Detailed classification

We propose a detailed algorithm to classify further the four coarse sets of text lines into the correct text line types, respectively. The name and title lines can be extracted from the possible name and title set. And a decision tree illustrating the detailed classification algorithm for name and title lines are given in Fig. 4.
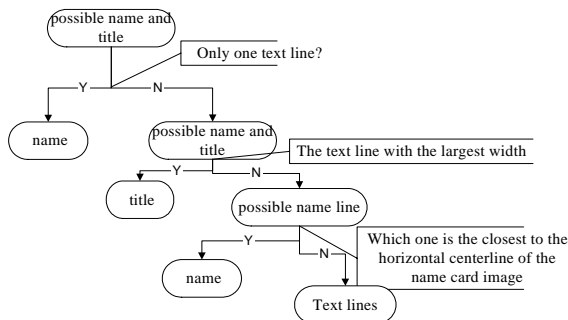

Fig.3. A decision tree for detailed classification of text lines.

According to the Chinese name feature and

the position feature, we can discriminate name lines from title lines. The mobile, fax, government publications, and real phone number lines can be classified from the possible phone number set. By recognizing the special characters specified in Table 1, we can extract respectively the four different number lines from the possible phone number set. Finally, the address lines can be extracted from the possible address set by recognizing the characters "市" and "縣."

5. Recognition of Text Lines in English Name Cards by Decision Trees

5.1 Segmentation of English Text Lines

The process of the horizontal cut of a text block is the same as that for Chinese name card images. About the vertical cut, we apply the vertical cut process to text blocks without the restriction on the width/height ratio of English text blocks, because the widths and heights of characters in English are hard to estimate. For text line segmentation, the process is the same as that for Chinese name cards except that the medium height of text blocks in a text line is not used for English name cards because the shapes of English characters are not regular. Therefore, the width W_L of a text line is estimated instead to decide whether a line has to be cut vertically or not. If the value of W_L is close to the width of the name card image, then this text line may consist of several thinner text lines. But the width of a title line may be close to the width of the name card image, so a predefined threshold value T is proposed to avoid cutting such text lines. That is, the threshold value T is used to decide whether the widths are close enough or not. If so, then the text line should not be cut vertically.

5.2 Classification of English Text Lines

Text lines are classified into eight types in this study, including name line, title line, e-mail line, web address line, mobile phone number line, fax number line, phone number line, and address line. And the features for English text line classification are described as follows.

A. Features for name and title lines

The heights of the text lines are also effective features for English name cards. The types of name and title lines can be discriminated from the other types by the use of the height feature. The position of a name line usually is the closest one to the horizontal centerline of a name card image. And the widths of the title lines are usually close to the width of the name card image.

B. Features for other types of text lines

The features proposed for the other types of

English text lines are shown in Table 2. To recognize the special characters as the features for classifying these text line types, two features are employed in this study, including the crossing count feature and the peripheral feature, both described previously. The foreground feature is not used for recognition of English characters because this feature is not suitable for discrimination of different English characters.

Table 2 Features (special characters) for several types of English text lines.

	Height of a text line	Special characters
e-mail	small	"@"
web address	small	" . "
mobile phone number	small	"Mo"
fax number	small	"FAX", "Fax"
phone number	small	"TE", "Te"
address	small	"Ro", "RD"

5.3 Proposed Method for English Text Line Classification

We classify text lines first into two sets automatically by *k*-means clustering using the heights of the text lines: one set with larger heights including possible name and title lines, and the other including the remaining text lines. According the position feature and the length feature of text lines, we can discriminate name lines from title lines. A decision tree illustrating the classification process for name and title lines are given in Fig. 5. By recognizing the special characters listed in Table 2, we can extract respectively the six different types of lines from the remaining text line set.
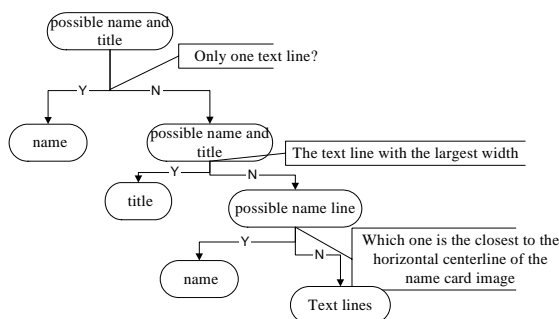

Fig.5. The decision tree illustrating detailed classification for the name and title lines.

6. Data Compression and Name Card Image Reconstruction

Original name card images are colorful. In this study, the colors of the components are reduced to two representative colors. We only keep the indices of the corresponding colors when saving the logo content. Hence, the logo content is binary. In addition, the characters are also binary. We compress these binary data with the popular run-length encoding (RLE) method that is a lossless technique. Advantages of RLE are easy to implement and quick to execute. On the other hand, to achieve the purpose of reconstruction and display of the name card image, it is required to store the attributes of the image components. Through the proposed algorithms described previously, the components in the name card image have already been extracted. The final results of logo blocks and characters are saved into a file in a run-length encoding form while the name card image is reconstructed. Therefore, when the name card image is displayed, each component will be recovered by run-length decoding, and shown at the original relative positions in the name card.

7. Experimental Results

Several name card images were tested in our experiments. We obtained these name card images from an HP ScanJet scanner at 250dpi resolution with true color levels. The proposed algorithms were implemented on a Pentium IV-2.4G PC with 256 MB RAM and software development was conducted by the use of VC++ 6.0 in a Windows 2000 Professional platform. Some experimental results are shown in Figs. 6-8. Tables 3-6 show the recognition rates of our experimental results, which prove the feasibility of the proposed algorithms.

(a)   (b)

Fig. 6 An example of experimental results of logo extraction. (a) The original image. (b) The resulting image after logo extraction.

(a)   (b)

Fig. 7 An example of experimental results of text line classification in a Chinese name card image. (a) The original image. (b) The resulting image after text line classification.

(a)   (b)

Fig. 8 An example of experimental results of text line classification in an English name card image. (a) The original image. (b) The resulting image after text line classification.

Table 3 Recognition rate of logo extraction.

The total number of name card images	100
The total number of logo blocks	99
The number of errors	3
Recognition rate	96.69%

Table 4 Recognition rate of card type classification.

	Chinese name cards	English name cards
The total number of name card images	50	50
The number of errors	2	0
Recognition rate	96%	100%

Table 5 Recognition rate of classification of text line types in Chinese name cards.

	name	title	phone	fax	mo-bil	gpn	e-mail	web	address
No. of name cards	50	50	50	50	50	50	50	50	50
No. of text lines	50	50	57	46	11	19	33	11	51
Errors	4	6	5	6	2	1	3	0	8
Recog	92%	88%	91.2	86.9	81.8	94.7	90.9	100	84.3%

Table 6 Recognition rate of classification of text line types in English name cards.

	name	title	phone	fax	mobil	Ad-dress	e-mail	web
No. of name cards	50	50	50	50	50	50	50	50
No. of text lines	50	48	54	48	9	50	41	11
Errors	4	8	8	7	1	6	3	1
Recognition rate	92%	83.3%	85.1%	85.4%	88.8%	88%	92.6%	90.9%

8. Conclusions

Some algorithms for name card image analysis have been proposed. Different topics including extraction of basic blocks, extraction of logo blocks, classification of name card types, classification of text line types, and reconstruction of name card images were studied. In the phase of extraction of basic blocks, edge detection and region growing were used to extract basic blocks from name card images. Furthermore, A thresholding approach was applied to reduce the colors in a basic block so that we can get two representative colors in the block. Finally, a method for determination of the foreground color in a basic block was proposed. In the phase of logo block extraction, several features were employed and a decision tree was proposed to extract a graphic or text logo from basic blocks. In the phase of name card type classification, the width/height ratios of text blocks were employed. In the phase of text line classification for Chinese name cards, methods were proposed for text block segmentation and text line segmentation first. Also, several effective features were adopted and an algorithm was proposed to classify text lines into nine types. In the phase of text line classification for English name cards, an algorithm was proposed to classify text lines into eight types. In the phase of name card image reconstruction, a suitable compression technique was utilized to compress name card components and a friendly user interface was

designed to display the name card image. The experimental results have revealed the feasibility of the proposed methods.

References

[1] G. Nagy, "Twenty years of document image analysis in PAMI," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 1, pp. 38-62, Jan. 2000.

[2] H.M. Suen and J.F. Wang, "Preprocessing of Color-Printed Document Images for Automatic Character Recognition," Ph. D. Dissertation, Department of Computer Science and Information Engineer, National Cheng Kung University, 1998.

[3] W. H. Tsai, "Moment-preserving thresholding: a new approach," *Computer Vision, Graphics, and Image Processing*, Vol. 29, pp. 377-393, 1985.

[4] H.M. Suen and J.F. Wang, "Segmentation of uniform-coloured text from colour graphics background," *Proceedings of IEEE Vision, Image and Signal*, vol. 144, pp. 317-322, Dec. 1997.

[5] R. Lienhart and A. Wernicke, "Localizing and segmenting text in images and videos," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, pp. 256-268, April 2002.

[6] H.M. Suen and J.F. Wang, "Vision, Image and Signal Processing," *Proceedings of IEEE 2001 International Symposium on Image and Signal Processing and Analysis*, vol. 143, no. 4, pp. 210 -216, Aug. 1996.

[7] C. H. Chen; W. H. Tsai, "A New Decision-Tree Approach to Multi-Type Character Recognition by Direct Use of System Fonts as Reference Characters and Pairs of Image Components for Character Type Classification for Automatic Digital Book Construction," *Proceedings of 2002 Conference on Computer Vision, Graphics and Image Processing*, Taiwan, R.O.C., 2002.

[8] G. Nagy, "Chinese character recognition: a twenty-five-year retrospective," *ICPR 88: 9th Int'l Conf. on Pattern Recognition*, Cambridge, UK, vol. 1, pp. 163-167, 1988.

[9] A. X. Huang, J. Gu, and Y. Wu, "A constrained approach to multifont Chinese character recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 8, pp. 838-843, Aug. 1993.

[10] H. Saiga, Y. Nakamura, Y. Kitamura, and T. Morita, "An OCR system for business cards," *Proceedings of International Conference on Document Analysis and Recognition*, pp. 802-805, Tsukuba Science City, Japan, 1993.

[11] Y.H. Chiou and H.J. Lee, "Recognition of Chinese business cards," *Proceedings of the Fourth International Conference on Document Analysis and Recognition*, vol.2, pp. 1028-1032, Ulm, Germany, Aug. 1997.

[12] S.I. Arazaki, M. Saigusa, S. Hashiguchi, M. Ohki, M. Uchiyama, and F. Itoh, "Image data compression by DCT with adaptive run-length coding," *IEEE Transactions on Consumer Electronics*, vol 37, no. 4, pp. 860-866, Nov. 1991.

[13] J. Ha, R.M. Haralick, and I.T. Phillips, " Recursive X-Y cut using bounding boxes of connected components," *Proceedings of the Third International Conference on Document Analysis and Recognition*, vol.2, pp. 952-955