

Binarization of Color Document Images via Luminance and Saturation Color Features

Chun-Ming Tsai and Hsi-Jian Lee

Abstract—This paper presents a novel binarization algorithm for color document images. Conventional thresholding methods do not produce satisfactory binarization results for documents with close or mixed foreground colors and background colors. Initially, statistical image features are extracted from the luminance distribution. Then, a decision-tree based binarization method is proposed, which selects various color features to binarize color document images. First, if the document image colors are concentrated within a limited range, saturation is employed. Second, if the image foreground colors are significant, luminance is adopted. Third, if the image background colors are concentrated within a limited range, luminance is also applied. Fourth, if the total number of pixels with low luminance (less than 60) is limited, saturation is applied; else both luminance and saturation are employed. Our experiments include 519 color images, most of which are uniform invoice and name-card document images. The proposed binarization method generates better results than other available methods in shape and connected-component measurements. Also, the binarization method obtains higher recognition accuracy in a commercial OCR system than other comparable methods.

Index Terms—Color document, color feature, decision-tree, luminance, name-card, saturation, uniform invoice.

I. INTRODUCTION

PATTERN recognition is the study of how machines can observe the environment, learn to distinguish patterns of interest from their background and make sound and reasonable decisions about the categories of the patterns [1]. Document image analysis (DIA) is one of the applications of pattern recognition. It aims to convert document images to symbolic forms for modification, storage, retrieval, reuse transmission [2]. Most existing research in digital image processing deals only with monochrome images. Some of this research can be extended to the processing of color images in a straightforward fashion [3]. We also extend conventional image processing methods to analyze color document images.

Because of the advancement of color printing technology, color document images are employed increasingly. Documents can be digitized and then recognized via optical character recognition (OCR) techniques. For most OCR engines, character features are extracted and trained from binary character images. It is relatively difficult to obtain satisfactory binarization images from various kinds of document images. Moreover, it is difficult to extract text from these color document images. Several color

document image properties may yield unsatisfactory binarization results. Two of these are described as follows: foreground colors are close to or mixed with background colors in the gray level histogram [Fig. 1(a) and (b)] and Fig. 2 background contains mostly gray levels and foreground contains varying colors with very few pixels [Fig. 1(c) and (d)]. To binarize these color document images, document layout, intensity distributions, and color features may be employed. Herein, color feature is applied for binarization.

Segmentation is the first stage of color document image analysis. The methods can be classified into five groups [4]–[6]: 1) feature-space-based techniques; 2) image-domain-based techniques; 3) physics-based techniques; 4) fuzzy set techniques; and 5) hybrid techniques. The feature-space-based techniques are based on color features such as luminance, saturation and hue. The techniques can be further divided into clustering, adaptive k -mean clustering and histogram thresholding. These methods do not take into consideration spatial locations of pixels [4]. The image-domain-based techniques prefer using spatial grouping. These algorithms usually include split-and-merge, region growing, edge-based and neural network-based classification techniques [4]. If the objects portrayed in the color images are affected by highlights, shadowing and shadows, the former two groups of techniques are prone to segmentation errors. All these phenomena cause the appearance of color to change drastically. To overcome this drawback, models of the physical interaction are introduced in the segmentation algorithms. This motivates the physics-based techniques [4]. The fuzzy set techniques use the fuzzy theory to segment color images. The hybrid techniques combine some of the above methods to segment color images. In this paper, we will extend conventional histogram thresholding techniques, which are feature-space-based techniques, to process color document images.

Histogram thresholding is a well-known technique for gray level image segmentation. Many thresholding methods have been proposed. Thresholding techniques can be classified into three classes, which are global, local and hybrid. Global thresholding methods find a threshold from the information of an entire image. Sahoo *et al.* [7] evaluated more than twenty global thresholding methods via uniformity and shape measures. They concluded that Otsu's class separability method [8], Tsai's moment preserving method [9] and Kapur *et al.*'s entropy method [10] are satisfactory. Lee and Chung [11] compared five global thresholding methods employing the error probability, shape uniformity measures as the criteria. They concluded that the methods of Otsu and Kittler and Illingworth [12], [13] yielded relatively acceptable results. Otsu's method assumed that the thresholded images have two normal

Manuscript received November 10, 2000; revised June 18, 2001. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jezekiel Ben-Arie.

The authors are with the Department of Computer Science and Information Engineering, National Chiao Tung University, Hsinchu 30050, Taiwan, R. O. C. (e-mail: chunming@csie.nctu.edu.tw; hjlee@csie.nctu.edu.tw).

Publisher Item Identifier S 1057-7149(02)01321-0.

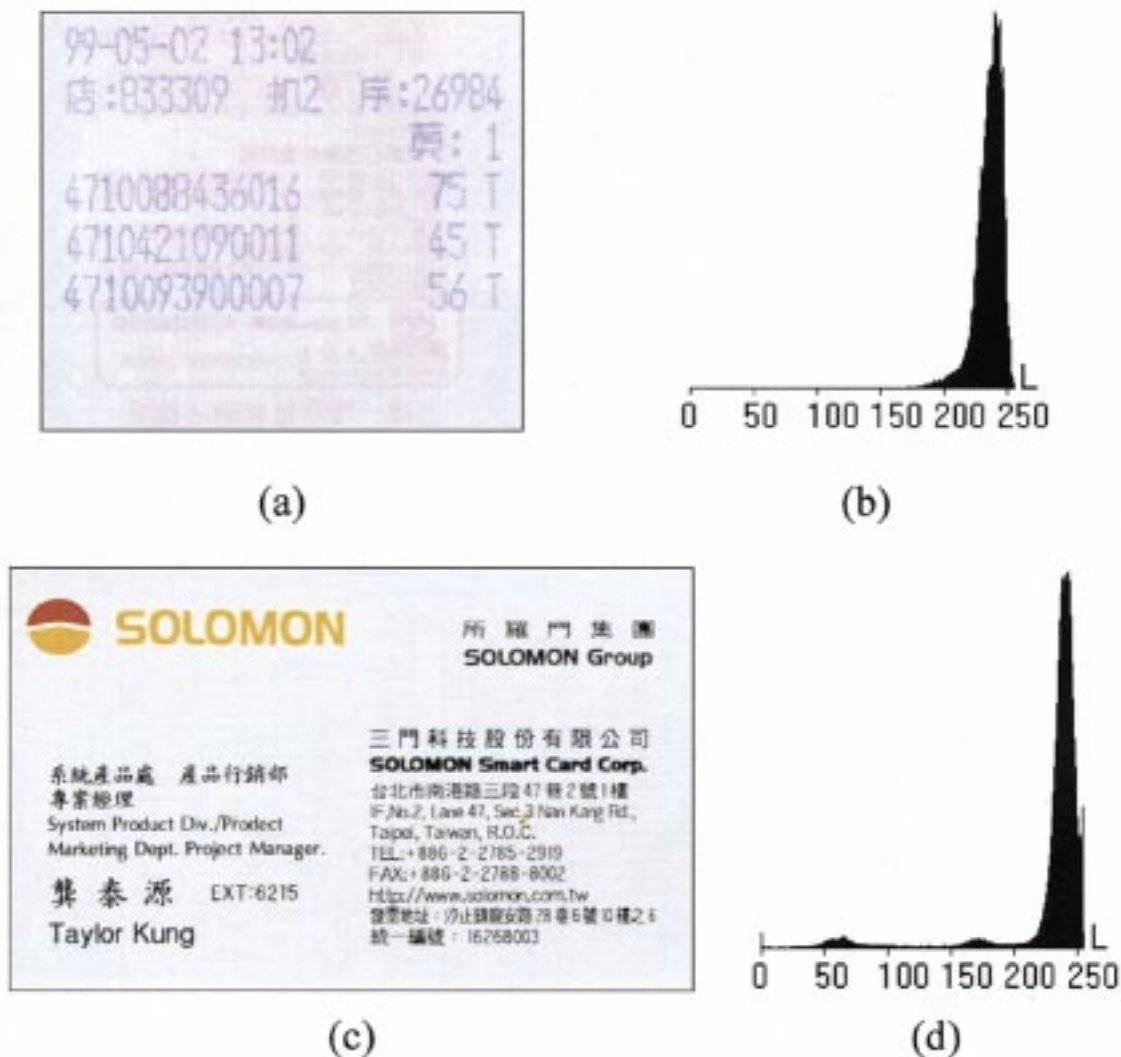


Fig. 1. Two properties of color document images that render binarization unsatisfactory. (a) A source image with inseparable foreground colors and background colors in the gray level histogram. (b) Inseparable gray level histogram. (c) A source image takes a mostly gray level population in the background and varying colors with very few pixels in the foreground. (d) Background-major histogram.

distributions with similar variances. The threshold is selected by partitioning the image pixels into two classes at the gray level that maximizes the between-class scatter measurement. Kittler and Illingworth's methods assumed that the thresholded images have two normal distributions with distinct variances. They derived a criterion function that minimizes the average pixel classification error rate. Notably, when the foreground and background intensities can be separated clearly, the global methods yield good threshold results the execute time is fast. However, there are many shortcomings for global methods. 1) When images have highly unequal population sizes, global methods tend to split the larger mode into two halves [12], [13]. 2) If the gray levels of foreground and background are inseparable, global methods cannot find acceptable thresholds. 3) Global methods cannot handle images with gradually decaying background or with texture. In short, global methods neglect the spatial relationships among pixels. This paper intends to solve the first two problems.

Local thresholding methods determine an individual threshold for each pixel according to the intensities of its own and its neighbor. Trier and Taxt [14] and Trier and Jain [15] evaluated 11 popular local methods for map images. Their experimental results demonstrated that when using the OCR engine of Trier and Jain, the methods of Niblack [16] and Bernsen [17] produced better and faster OCR results. Based upon local means and standard deviations, Niblack's method defined varying threshold values upon the image. Bernsen's method detected the lowest and the highest gray level, G_{low} and G_{high} , in a $w \times w$ square window centered at (x, y) . The threshold of pixel (x, y) is defined as $T(x, y) = (G_{low} + G_{high})/2$. In sum, local methods yield better experimental results, but are slower than global methods.

Hybrid thresholding methods combine both global and local methods. Liu and Li [18] proposed a 2-D Otsu thresholding method, which performs much better than the 1-D Otsu method does, when images are corrupted by noise. Their method calcu-

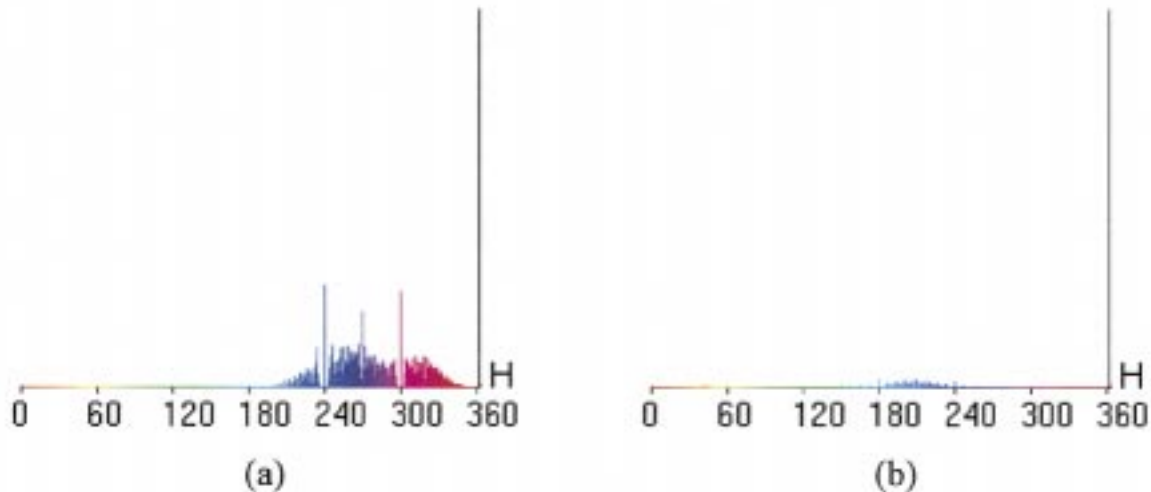


Fig. 2. Two examples in which hue cannot be used for binarization properly. (a) Hue histogram of Fig. 1(a) and (b) hue histogram of Fig. 1(c).

lates the local average gray level within a limited window. They constructed a 2-D histogram, in which the x -axis and the y -axis are the gray value and the local average gray level, respectively. The optimal threshold vector (s, t) is selected at the maximum between-class variance. Gong *et al.* [19] proposed a fast 2-D Otsu method to accelerate Liu and Li's method. Via this method, the computation complexity can be decreased from $O(G^4)$ to $O(G^2)$, where G is the number of gray levels. Both the computation time and the memory space are reduced greatly, while the segmentation quality is maintained. In addition, Tseng and Lee [20] proposed a document image binarization method, which is a two-layer block extraction method. In the first layer, dark backgrounds are extracted by locating connected-components. In the second layer, background intensities are determined and removed from each component. In comparison, hybrid methods have better experimental results, but are also slower than global methods.

Conventional thresholding methods are performed on gray level images. However, we aim to process color document images. Since the color pixels are represented by RGB or some transformation of RGB , their histograms are represented by a 3-D array. Selecting thresholds from the histograms is not easy and the computational time is high. One way to solve this problem is to develop efficient storing and processing methods to handle information in the 3-D color space; some are reported in [21], [22]. Another way is to project the 3-D space onto a lower dimensional (2-D or 1-D) [5], [20], [23] and [24]. Cheng and Sun [5] proposed a two-phase color image segmentation algorithm, which extends the histogram to a homogeneity histogram. The homogeneity is defined as a composition of two components: standard deviation and discontinuity of the intensities. In the first phase, uniform regions are identified via multilevel thresholding on the homogeneity histogram. In the second phase, the hue histogram is analyzed in each uniform region. Their method is a kind of hybrid thresholding. Tseng and Lee [20] transformed a color document image into a luminance (intensity) image. Celenk [23] proposed a simple dimensionality reduction approach that constrains the shape of clusters in the cylindrical coordinates of cube-root

systems and use a 1-D thresholding scheme in human color perception (L : lightness, H : hue, C : chroma). However, the overall computation time is high. Tseng *et al.* [24] use only hue information and suggest a circular histogram thresholding of such attributes.

In contrast to the above projection methods focusing on the hue feature [5], [23] and [24], we present in this paper a new binarization method for color document images via luminance and saturation features. Hue is a useful attribute that provides sufficient information for color segmentation [24], but has many disadvantages as follows. 1) Hue is meaningless when luminance is very low or high [24], [25]. 2) Hue is unstable when saturation is very low [5], [25], [26]. 3) Hue cannot distinguish small color change [5]. 4) the effective ranges of hue have to be determined [24]; that is, we have to determine which region in an image is achromatic and which region is chromatic an accurate achromatic area is difficult to establish. 5) In real-time applications, speed is important, hue needs more time to be transformed from the RGB color space than luminance and saturation do. 6) It is difficult to use hue to threshold color document images with inseparable foreground colors and backgrounds colors. For example, in Fig. 1(a), the foreground digits and reverse text appear in the foreground. These have hue in the range from blue to red. In Fig. 1(c), the major part of the foreground is text, which has achromatic colors. The hue histograms of Fig. 1(a) and (c) are shown in Fig. 2(a) and (b), in which the x -axis represents the hue degree and the y -axis represents the number of pixels for each degree. It is improper to use hue to threshold these document images.

The reasons that components from different color spaces are not used are summarized below. First, the RGB and XYZ color spaces are nonuniform chromaticity scale. An adequate segmentation result depends on segmentation techniques by detecting similarity among the attributes of image pixels [24]. However, the similarity measure (Euclidean distance) between two colors in RGB or XYZ do not necessarily reflect the visual separation between these two colors [6]. We can apply the thresholding method for each component and combine the results. However, when an image with inseparable foreground and back-

ground colors, each color component cannot produce proper results. Also, it is difficult to find a strategy to combine the results from different color components. Second, the *CMY* color space is used in connection with generating hardcopy output. The inverse transform from *CMY* to *RGB* is generally of no practical interest. When we apply the thresholding method for each component, their results are improper. Third, in the *HSI* color space, the problems of using hue have been described in the preceding paragraph. The intensity is defined as $I = (R + G + B)/3$. A pure red color (255, 0, 0) and a pure green color (0, 255, 0), for example, have the same intensity value I (85) and cannot be distinguished. It is important to differentiate these two colors in general applications. Fourth, in the *YIQ* color space, the luminance Y and color information (*Inphase* and *Quadrature*) are decoupled. Luminance is proportional to the amount of light perceived by the eye [25]; therefore, it can be regarded as intensity. For the pure red and green colors, the corresponding luminance values are different, 76.245 and 149.685, respectively, and can be separated [20]. Fifth, transformations to other perceptually based spaces such as CIE Lab and CIE Luv take much computation time [6].

However, no single thresholding scheme gives satisfactory segmentation results on a variety of images [2]. Particularly, when the gray levels of both foreground and background are inseparable, all of the above methods fail to provide good and efficient threshold results. Therefore, the objective of this investigation is to propose a method that can process color document images either via conventional threshold methods or with the two additional properties. These include foreground colors that are close to or mixed with background colors in the gray level histograms and the background color has chiefly a gray level population. This paper presents a thresholding scheme to be applicable on a variety of images, including uniform invoices, name-cards, calligraphy documents, advertisements, magazines, newspapers, brochures and so on.

Fig. 3 illustrates the system flow diagram proposed. The color document image is initially transformed into the luminance and saturation spaces. Then, the luminance histogram is computed and a Gaussian smoothing filter is employed to remove unreliable peaks and valleys upon the histogram. The threshold candidates are then selected from the histogram. Next, the entire image distribution is analyzed to extract statistical features to be applied in a decision tree. The tree decides if luminance, saturation, or both features will be employed to binarize color document images. Finally, the color features binarize the images.

The remainder of this paper is organized as follows. Section II describes the *LS* color features, which are employed to binarize color documents. Section III presents the decision tree, which decides whether saturation, or luminance or both will be employed. Section IV explains the color image thresholding algorithm. Section V provides and discusses the experimental results. Finally, Section VI includes the conclusion and suggestions for future works.

II. COLOR FEATURES

A robust document reading system must process color, gray-level and binary images. Generally, several techniques,

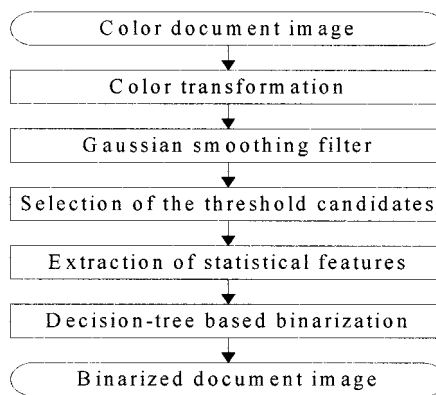


Fig. 3. Flow diagram of the proposed system.

including document layout analysis, character segmentation and character recognition, are developed to manipulate binary images. Hence, to binarize color and gray-level images, a binarization module must be applied. Notably, when document images have inseparable gray levels both in the foreground and background, the global, local and hybrid thresholding methods fail to provide good binarization results. For example, via the thresholding methods of Gong, Otsu, and Kittler, Fig. 4(b)–(d) depict poor binary images of Fig. 4(a). Within a histogram, the Otsu and Kittler methods tend to divide the larger distribution into two halves [Fig. 4(e) and (f), respectively]. Herein, to binarize color document images, color features, saturation and luminance, are applied.

A. *LS* Color Features

The *LS* color features are employed from the *LHS* color model [27]–[29]. This model describes colors in a more readily understood fashion. *Luminance*, *Lum*, is the brightness of the color, similar to gray level. *Saturation*, *Sat*, is a measure of the amount of white within the color, such as, pink is red with more white, that is, it is less saturated than a pure red. Increasing or decreasing *Lum* makes colors lighter or darker, respectively. When *Sat* is decreased, the colors become grayer [30]. The luminance [25], [27]–[33], *Lum*, which is defined by

$$Lum = 0.299R + 0.587G + 0.114B. \quad (1)$$

In this formula, $Lum = 0$ denotes black and $Lum = 255$ white. The weights reflect the eye’s brightness sensitivity to the *RGB*. The saturation, *Sat*, which is defined [25], [27]–[33] by

$$Sat = 1 - 3 \frac{\min(R, G, B)}{(R + G + B)}. \quad (2)$$

Within this formula, the lowest and highest saturation are 0 and 255, respectively.

B. *Negatively Scaled Saturation*

Herein, a negatively scaled version of saturation is defined as follows:

$$Sat = 3 \frac{\min(R, G, B)}{(R + G + B)} * 255. \quad (3)$$

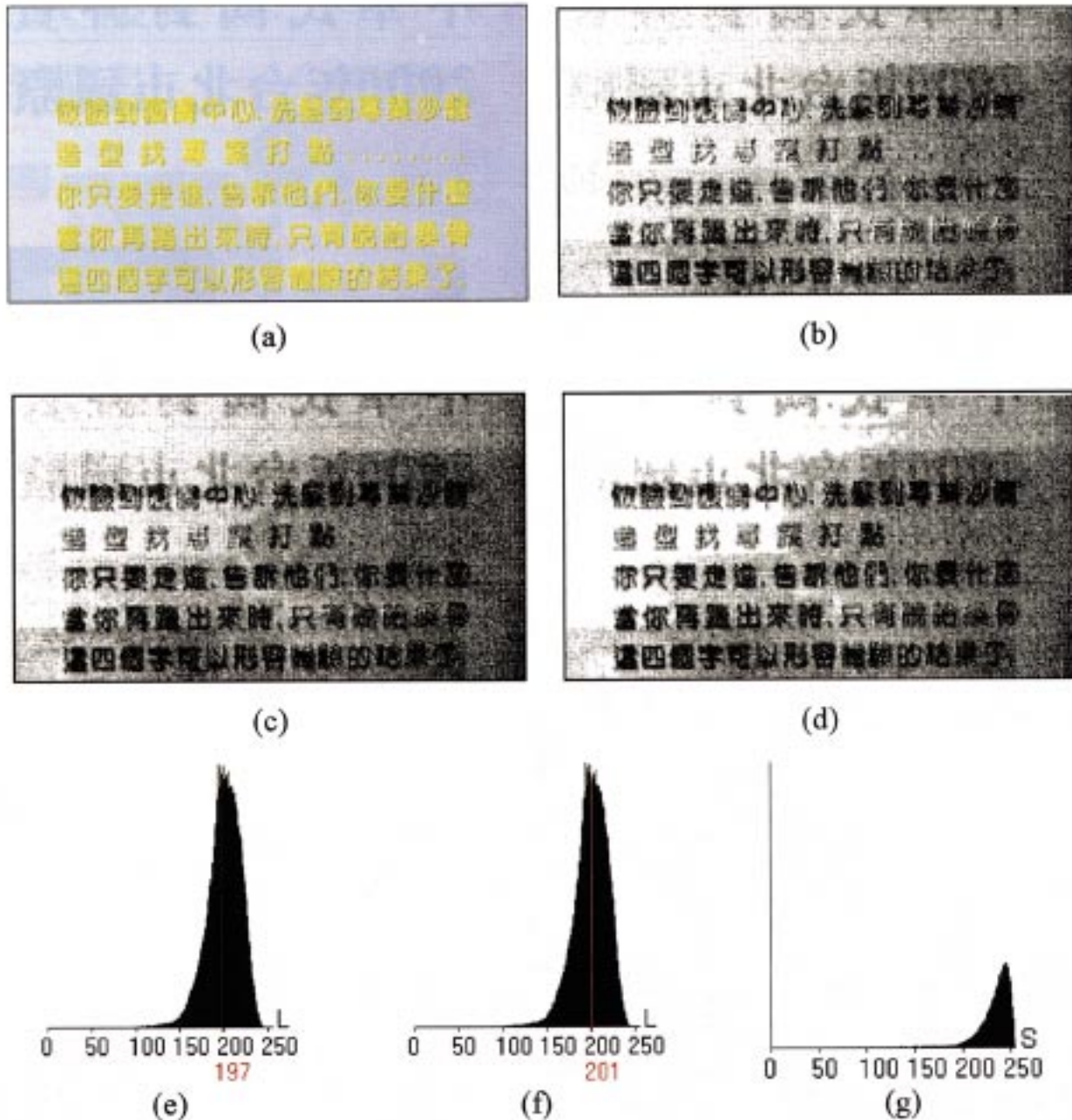


Fig. 4. Binarization via Gong's, Otsu's, and Kittler's thresholding methods. (a) Source image, (b) Gong's 2-D Otsu method, (c) Otsu's method, (d) Kittler's method, (e) Inseparable luminance histogram threshold by Otsu's threshold value (197), (f) threshold by Kittler's threshold value (201), and (g) separable negatively scaled saturation histogram of Fig. 4(a). Notably, there is a large peak on the left side.

Within this formula, the highest and lowest saturation are 0 and 255, respectively. For pure black, that is, all RGB are 0, the saturation is undefined and thus, fixed at zero herein. Negatively scaled saturation has two advantages: 1) producing an image that correlates closely with luminance when we display saturation and 2) employing the same conventional thresholding methods as luminance does. Furthermore, the saturation histogram, like the luminance histogram, can be computed.

C. Usage of Negatively Scaled Saturation

Saturation provides a measure of the degree to which a pure color is diluted by white light [25]. There are two cases herein. First, negatively scaled saturation is employed for thresholding

when the luminance distributions of the foreground and the background are inseparable. For example, in Fig. 4(e) indicate that the luminance histogram of Fig. 4(a) is inseparable and in Fig. 4(g) reveals that the negatively scaled saturation histogram of Fig. 4(a) is separable. Notably, there is a large peak on the left side of Fig. 4(g).

Second, both luminance and negatively scaled saturation are applied for thresholding when the entire image has large luminance variance and the background contains the most population, also with a large variance. Fig. 5 illustrates the application of both of the luminance and saturation features. Fig. 5(a) depicts the luminance histogram with large variance and the background contains the most population, also with a large variance. Fig. 5(b) indi-

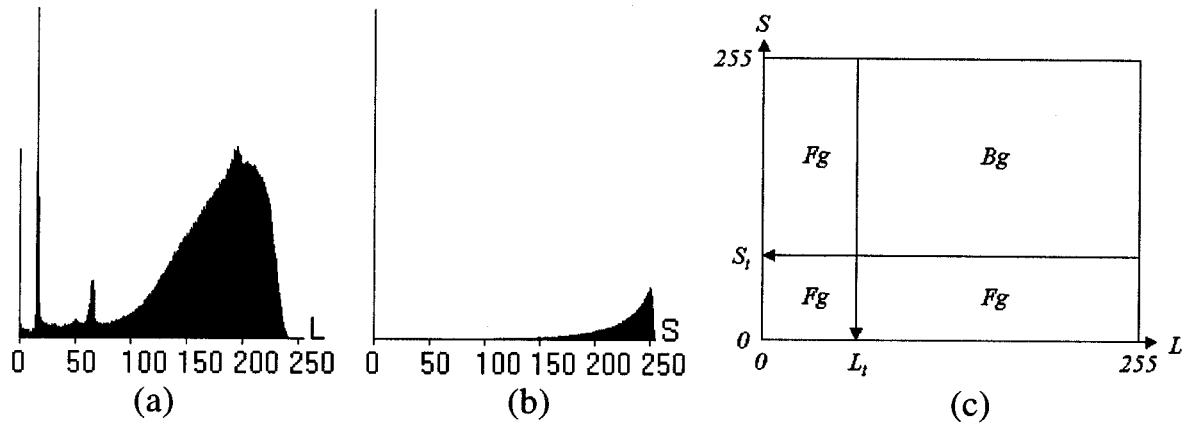


Fig. 5. Example that uses both luminance and saturation features. (a) Not easy separable luminance histogram, (b) separable negatively scaled saturation histogram, and (c) diagram to explain the use of both luminance and saturation features. F_g denotes foreground, B_g background, L luminance axis, S saturation axis and L_t and S_t is the threshold of the luminance and saturation, respectively.

cates that the negatively scaled saturation histogram can be separated into two distributions. However, a few foreground pixels contain higher saturation (from S_t (saturation threshold) to 255), which represents an unsaturated area. Typically, these pixels have a lower luminance, in fact less than L_t (luminance threshold). In this instance, if both features are employed, the foreground and background can be separated. Fig. 5(c) confirms this, in which F_g denotes foreground, B_g is background, L is luminance axis and S is saturation axis. In the following discussion, saturation denotes negatively scaled saturation.

III. DECISION FLOW FOR EITHER SATURATION OR LUMINANCE OR BOTH

In order to binarize color document images by the decision-based method, statistical image features are extracted from the luminance distribution. These features are used to determine a path in the decision tree. Here, three points are assumed. First, color distributions in document images are distributed normally. Second, the majority of distribution is located in the background of the image and third, the background has higher luminance levels than the foreground does. To compute the luminance for an input color document image, first, (1) is used. Then, the luminance histogram $H_L(x)$ is computed, where the number of pixels with the various levels of brightness from black to white is shown. The histogram represents a probability distribution of the brightness levels. Finally, to obtain reliable peaks and valleys, a Gaussian smoothing filter is applied to smooth the original histogram, thus removing unreliable peaks and valleys.

A. Gaussian Smoothing Filter [34]–[37]

The Gaussian convolution of a luminance histogram $H_L(x)$ depends upon both x and σ_g , namely, the Gaussian standard deviation. The convolution function $S_{HL}(x, \sigma_g)$ is provided by

$$\begin{aligned} S_{HL}(x, \sigma_g) &= \int_{-\infty}^{\infty} H_L(u)g(x-u, \sigma_g) \\ &= \int_{-\infty}^{\infty} H_L(u) \frac{1}{\sqrt{2\pi}\sigma_g} e^{-\frac{(x-u)^2}{2\sigma_g^2}} du \quad (4) \end{aligned}$$

where “*” denotes the convolution operator and $g(x-u, \sigma_g)$ is the Gaussian function. The degree of smoothing is controlled by the standard deviation of the Gaussian function. The larger the standard deviation σ_g is, the smoother the function $S_{HL}(x, \sigma_g)$ is. Notably, in [36],[37] the smoothing parameter is predetermined. However, in our proposed method, the standard deviation is decided automatically.

Standard deviation σ_g is defined herein based upon the majority of the widths within the luminance histogram. In a histogram $H_L(x)$, if $H_L(x) < H_L(x-1)$ and $H_L(x) < H_L(x+1)$, then luminance x is a valley. The highest point between two successive valleys is a peak, which identifies a distribution. Therefore, the widths between two successive valleys are computed and thus, the maximum width, W_{\max} , among the widths is determined. Next, the width histogram for all peaks from zero to W_{\max} is computed. After the highest point in the width histogram is located, which is regarded as the standard deviation σ_g of the widths, (4) is employed to convolute the histogram $H_L(x)$ that provides smoothing histogram.

B. Selection of Threshold Candidates

After the small peaks and valleys have been removed, the average differences are employed as the first derivation with which to determine the major peaks and valleys. The average difference in point x is defined by

$$S'_{HL}(x) = \frac{1}{\sigma_g - 1} \sum_{i=1}^{\sigma_g - 1} \frac{S_{HL}(x+i) - S_{HL}(x-i)}{2 \times i}. \quad (5)$$

A peak is defined as a positive to negative crossover in the first derivation of the smoothed histogram. Furthermore, a valley is defined as a negative to positive crossover. All peaks and valleys from the first derivation of the smoothed histogram were discovered. In cases where the peaks and valleys are too close, they will be removed if the distance between a valley and a peak is less than the standard deviation σ_g . The remaining valleys are the candidate threshold values.

C. Features Employed

The background of a document image is defined as the region with the maximum population. To decide whether to em-

ploy saturation, or luminance, or both features for binarization, four statistical features were defined, based upon the luminance distribution in the decision tree.

- 1) The luminance variance, σ^2 , of the entire image: This feature describes whether the distribution of the entire image is wide or narrow.
- 2) The ratio R_{FB} of the population of the foreground to the background: The *color population* $CP(i)$ for color, i , is defined as

$$CP(i) = \sum_{j=\nu(i)}^{\nu(i+1)} S_{HL}(j) \quad (6)$$

where $\nu(i)$, $i = 0, \dots, n$, are the valleys of the luminance histogram. Each color, i , has a *left valley* $\nu(i)$ and a *right valley* $\nu(i+1)$, which are threshold candidates. Then, the ratio R_{FB} is defined as

$$R_{FB} = \frac{(N - \max CP(i))}{\max CP(i)} \quad (7)$$

where N is the total number of pixels within the whole color image and $\max CP(i)$ is the background population, which has a left valley $\nu(i)$ and a right valley $\nu(i+1)$. If R_{FB} is greater than a threshold, the image is dubbed *foreground-significant*, otherwise *background-major*.

- 3) The luminance variance, σ_b^2 , of the background distribution: This feature describes whether the background distribution is concentrated within a specific range or not.
- 4) The total number, N_{LL} , of pixels that range between the lowest luminance 0 and *low-level luminance*: The *low-level luminance* is 60. Notably, within human perception, pixels with a luminance of less than 60 are black. This feature describes whether the number of lower luminance pixels is large or small. In sum, to illustrate these four features, Fig. 6 contains an example of the luminance histogram to explain the features used.

D. Decision Tree

After the features are extracted from the color distribution, a decision tree is used to decide if luminance, saturation, or both features are to be applied. Fig. 7 illustrates the diagram of the decision tree, which displays five thresholding cases. Case A uses saturation, if the luminance variance, σ^2 , of the entire image is small. In Case B, luminance is adopted if the ratio R_{FB} of the foreground to background color is significant. Case C also uses luminance for thresholding, if the luminance variance, σ_b^2 , of the background color is small. Case E uses saturation for thresholding if the total number of pixels, N_{LL} , of the lower image foreground is small. Finally, Case E uses both luminance and saturation for thresholding. These five thresholding cases are described in detail in the following section.

IV. COLOR IMAGE THRESHOLDING

A. Luminance-Based Thresholding

In Fig. 7, there are two cases that apply luminance for thresholding, which are Cases B and C. Images of the former contain two properties. They are 1) the luminance variance, σ^2 , of the

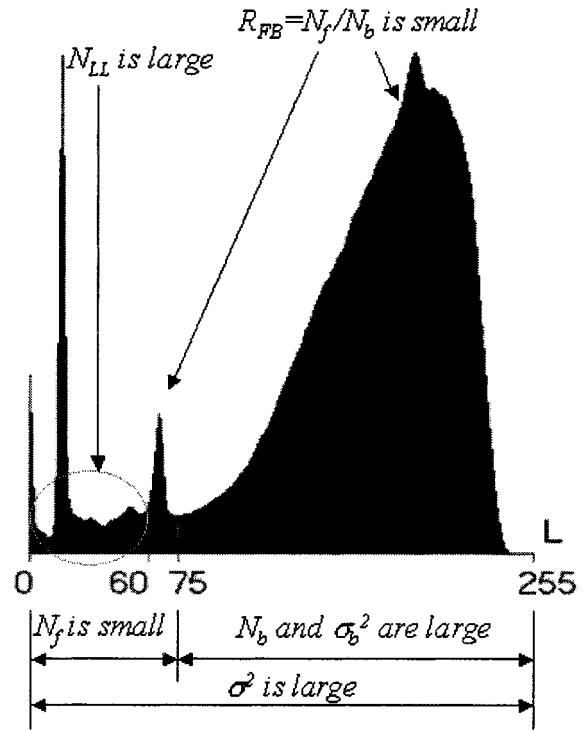


Fig. 6. An example of the luminance histogram to explain the four features used. N_f and N_b are the total numbers of foreground and background pixels, respectively.

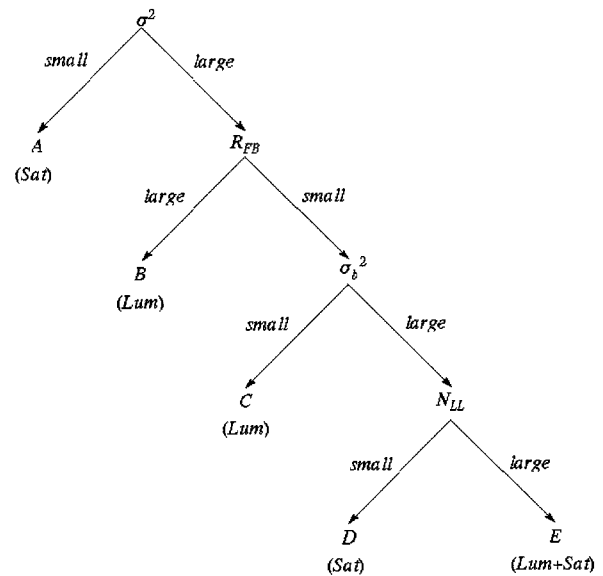


Fig. 7. Diagram of the decision tree to determine whether either saturation feature (denoted as *Sat*), or luminance feature (*Lum*) or both for binarization are to be applied.

entire image is large and 2) the foreground colors are significant, that is, R_{FB} is large. These properties suggest that the background and the foreground are separable, but the optimum threshold values cannot be obtained easily between foreground and background. Fig. 8 illustrates two examples for Case B. If Otsu's method is employed, owing to improper thresholds, bad binarization results will be obtained. In the following section, a foreground-significant thresholding algorithm is presented to binarize images of this case.

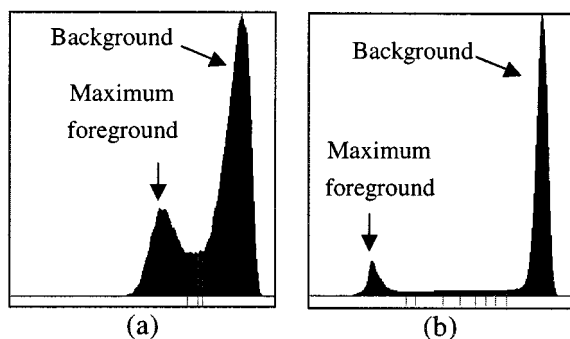


Fig. 8. Two luminance smoothed distributions of Case B. Their foreground and background are separable, but the optimum threshold values cannot be obtained easily between foreground and background. (a) Three candidate thresholds and (b) eight candidate thresholds.

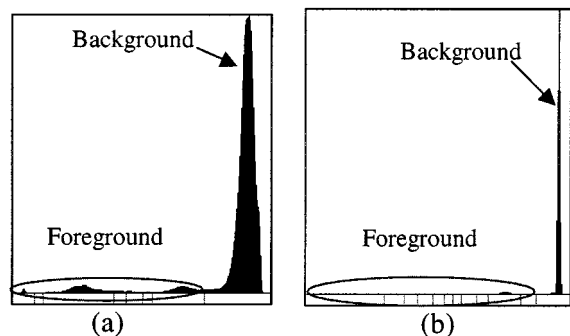


Fig. 9. Two examples of Case C. The foreground and background are separable and the background is concentrated within a certain range. (a) Seven candidate thresholds and (b) 11 candidate thresholds.

Images within Case C possess three properties. These are 1) the luminance variance, σ^2 , of the entire image is large; 2) the background colors are prominent, that is, R_{FB} is low; 3) the luminance variance, σ_b^2 , of the background is low. These properties imply that the background and the foreground are separable and that the background color distribution is concentrated within a particular range. Fig. 9 displays two luminance histograms for this case. If Otsu's method is applied, due to improper thresholds, bad empirical results will be obtained. In Section IV-A2, a background-major thresholding algorithm is presented to binarize the images of this case.

1) *Foreground-Significant Thresholding*: Because of the two properties of Case B, stated previously, conventional thresholding methods fail to produce satisfactory threshold results. Therefore, to partition Case B into two subclasses, another feature was defined. This feature is the *distance*, d_{FB} , between the right valley of the *maximum foreground*, which is the region with the second highest population, that is, the second $CP(i)$, defined in Section III-C and the left valley of the background.

In Case B1, the distance, d_{FB} , is small [Fig. 8(a)]. It indicates that the maximum foreground and the background are close. The luminance threshold value is defined as the middle between the maximum foreground and the background. In Case B2, the distance, d_{FB} , is large [Fig. 8(b)]. It implies that the maximum foreground and the background are distant. Notably, in the middle between the maximum foreground and the background, many small color distributions exist, which are regarded

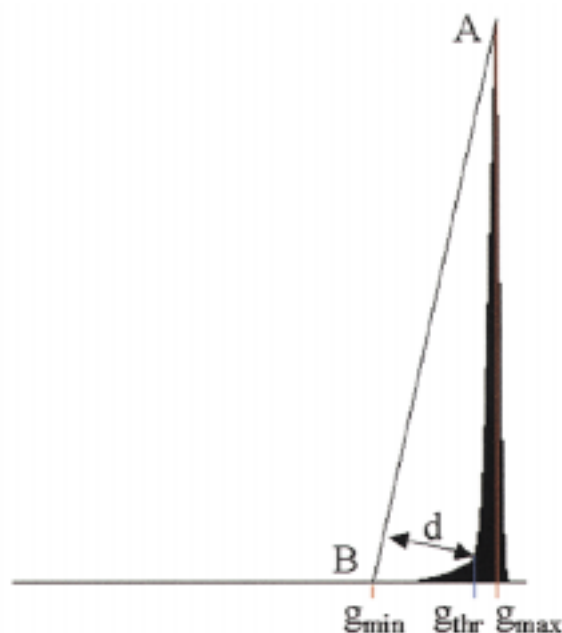


Fig. 10. An example for the gray level triangle thresholding method, which be used in our saturation-based thresholding method.

as foreground. In this manner, the left valley of the background, the threshold, separates foreground and background.

2) *Background-Major Thresholding*: In Case C, the background and the foreground are separable and the background is concentrated with a distinct range. Since the foreground and background are separated by left valley of the background, the left valley becomes the threshold value. If the luminance value of a pixel is greater than the threshold value, then the pixel is held as white, otherwise, it is black.

B. Saturation-Based Thresholding

According to our analyses, Cases A and D apply saturation for thresholding. Images of Case A have one property. Restated, the luminance variance, σ^2 , of the entire image is small, which implies that the background and the foreground of the domain cannot be separated easily. A saturation-based thresholding method is proposed for this case.

Images within Case D have four primary properties. These are 1) the luminance variance, σ^2 , of the entire image is large; 2) the background luminance is substantial, that is, R_{FB} is small; 3) the luminance variance, σ_b^2 , of the background is large; and 4) the total number of pixels, N_{LL} , of the lower luminance is small. Therefore, the background and the foreground that contain these properties in the luminance domain cannot be separated easily. An alternate saturation-based thresholding method is recommended for this case.

Preprocessing of the saturation-based thresholding methods are similar to the decision flow provided in Section III. First, the histogram $H_S(x)$ is computed for the saturation. Second, via the Gaussian smoothing filter, the saturation histogram $H_S(x)$ is smoothened to remove unreliable peaks and valleys. Third, to select candidate peaks and valleys, the threshold selection algorithm was employed. Last, compute the saturation variance, σ_s^2 , of the entire image.

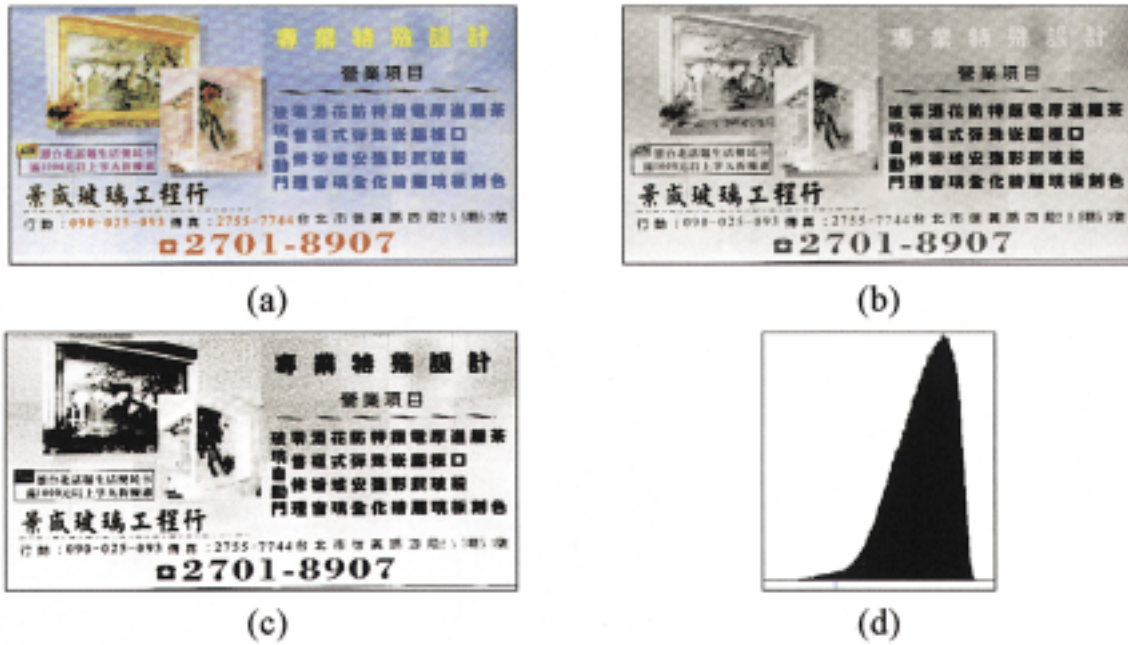


Fig. 11. Both luminance and saturation. (a) Source color document image. (b) Source image in the luminance domain. (c) Source image in the saturation domain. (d) Luminance histogram of the pixels, which have saturation values greater than the saturation threshold value.

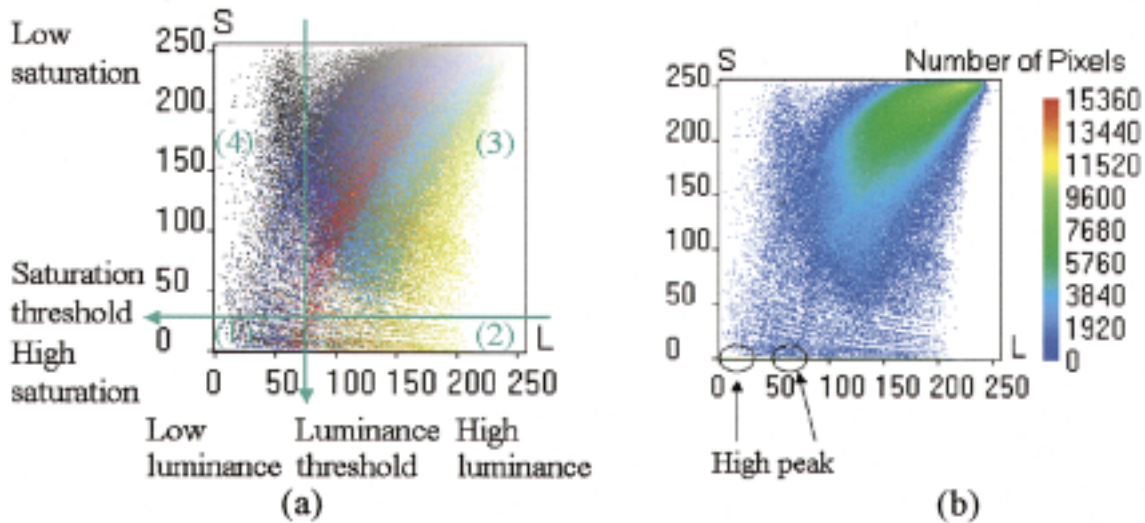


Fig. 12. Example for using luminance and saturation features. (a) An *LS* color image of Fig. 11(a). (b) The color histogram of Fig. 11(a), in *LS* space is shown by 2-D histogram. The number of pixels is represented by different colors. The less the number of pixels is, the closer is the color to blue. The more the number of pixels is, the closer is the color to red.

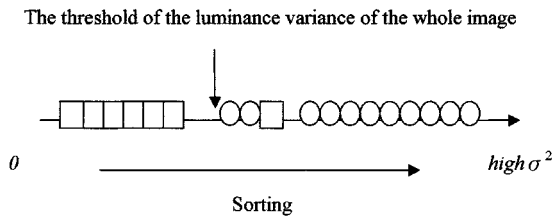


Fig. 13. Luminance variance threshold selection □ denotes a good threshold result by the saturation-based thresholding and O denotes a bad one.

The saturation-based thresholding method contain herein is based upon the gray level triangle method [38], which contains three chief properties. First, it determines the threshold value

TABLE I
NUMBER OF IMAGES OF THE FIVE
DECISION CASES

	Case A	Case B		Case C	Case D	Case E
		B1	B2			
Numbers	98	96	146	70	40	49
Features	<i>Sat</i>	<i>Lum</i>	<i>Lum</i>	<i>Lum</i>	<i>Sat</i>	<i>Lum+Sat</i>

near the foot of the background. Second, the area with large gray level variations is classified as the foreground. Third, this mechanism can be implemented easily. Fig. 10 depicts the method to locating the foot of a single peak gray level histogram. Notably, a line is constructed between the highest peak, *A* and the lowest

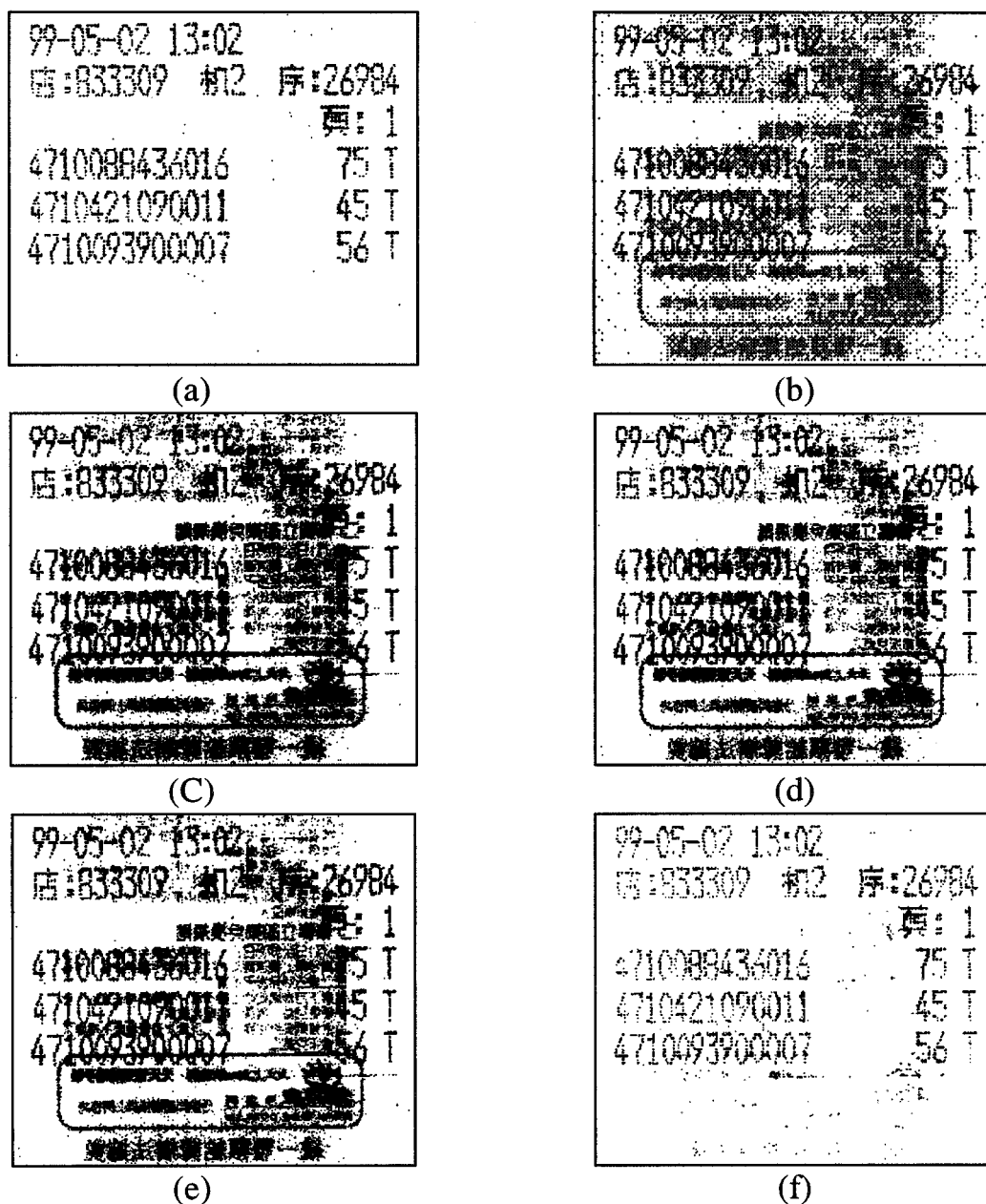


Fig. 14. Binarization results of Fig. 1(a). (a) Our saturation-based binarization method (Cases A and D) and other luminance-based thresholding method. (b)–(f) Tseng, Gong, Otsu, Kittler, and triangle luminance-based binarization methods, respectively.

value, B , located between the gray value g_{max} and g_{min} in the x -axis. For each x -axis value from g_{min} to g_{max} , the distance, d , from the top of the histogram to the connecting line from A to B was calculated. The x -axis value with the maximum distance is the optimal threshold value, g_{thr} .

As stated previously, to binarize images, two saturation-based thresholding methods are proposed. The conventional triangle method is effective only when the foreground pixels produce a weak peak in the histogram. The triangle method produces poor experimental findings when the foreground pixels display pronounced peaks. To partition the two methods, two features are defined, which are the number of peaks and the saturation histogram variance. That is, to obtain the saturation threshold value, S_t , the first triangle thresholding method is employed when the saturation variance, σ_s^2 , of the entire image is small or

the peak number is one. Otherwise, the second triangle thresholding method is applied.

The procedure for the first triangle thresholding method is as follows.

- 1) Locate the highest unsaturated position, S_{max} and the lowest saturated position, S_{min} , of the saturation distribution.
- 2) Compute the distance, d , between the line and the histogram $H_S(s)$ for values of s from $S_{min} + 2\sigma_s$ to $S_{max} - 2\sigma_s$.
- 3) Select the optimal threshold value S_t , as the x -axis value with the maximum distance.

The procedure of the second triangle thresholding method is as follows.



(a)



(b)



(c)



(d)



(e)



(f)

Fig. 15. Binarization results of Fig. 1(c). (a) Our luminance-based binarization method (Cases B and C). (b)–(f) Tseng, Gong, Otsu, Kittler, and triangle luminance-based binarization methods, respectively.

- 1) Determine the highest unsaturated position, S_{\max} , of the highest peak from unsaturated 255 decreasingly.
- 2) Determine the highest saturated position, S_{\min} , of the first peak from saturated 0 increasingly.
- 3) Construct a line between the peak of the histogram at S_{\max} and the lowest value S_{\min} , which is represented as

$$As + BH_S(s) + C = 0 \quad (8)$$

where A , B and C are coefficients and are given by

$$A = -\frac{H_S(S_{\max}) - H_S(S_{\min})}{S_{\max} - S_{\min}} \quad (9)$$

$$B = 1 \quad (10)$$

$$C = -A \cdot S_{\min} - H_S(S_{\min}). \quad (11)$$

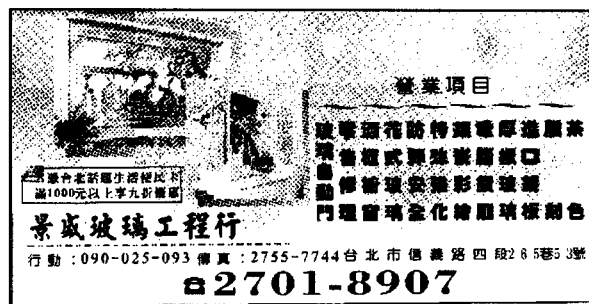
- 4) If the histogram $H_S(S_{\max})$ is greater than $H_S(S_{\min})$, then the distance, d , between the line and the histogram $H_S(s)$ is computed for all values of s from $LOW_{S_{\max}}$ to $S_{\max}-1$. Otherwise, s is computed from $S_{\min}+1$ to $UP_{S_{\min}}$, where $LOW_{S_{\max}}$ is the left valley of the highest unsaturated peak and $UP_{S_{\min}}$ is the right valley of the first saturated peak. The distance, d , between the line and a point $(s_0, H_S(s_0))$ is computed by

$$d = \frac{As_0 + BH_S(s_0) + C}{\sqrt{A^2 + B^2}}. \quad (12)$$

The best saturation threshold value, $S_t = s_0$, occurs where the distance between $H_S(s_0)$ and the line is maximal.



(a)



(b)



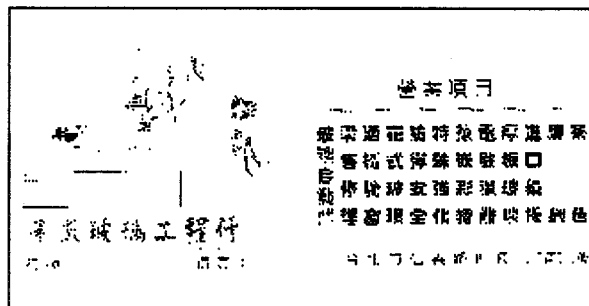
(c)



(d)



(e)



(f)

Fig. 16. Binarization results of Fig. 11(a). (a) Our luminance and saturation-based binarization method (Case E). (b)–(f) Tseng, Gong, Otsu, Kittler, and triangle luminance-based binarization methods, respectively.

C. Luminance and Saturation-Based Thresholding

Case E contains images that employ both luminance and saturation for thresholding. These images have four properties: 1) the luminance variance, σ^2 , of the entire image is large; 2) The background colors are pronounced, that is, R_{FB} is small; 3) The luminance variance, σ_b^2 , of the background is large; and 4) the total number of pixels, N_{LL} , of the lower luminance is large. These properties indicate that, within the luminance domain, the background and the foreground are not separated easily. Fig. 11 illustrates the reasons for application of both luminance and saturation for thresholding. Fig. 11(a) displays a source color document image. Fig. 11(b) and (c) depict the images in the luminance and the saturation domains. The luminance histogram has been shown in Fig. 5(a). Due to the proximity of the foreground and background colors, they are not separated easily. Fig. 11(d) demonstrates the luminance histogram of those pixels that have saturation values greater than the saturation threshold value. Therefore, if the saturation-based thresholding is applied to threshold these images, those foreground

colors will be missed. There are many colors that have low luminance and high saturation values. If we only use luminance or saturation to threshold these images, the results are unsatisfactory. We define an *LS color image* to explain this view, which is described as follows.

An *LS color image* is defined as a 2-D color function, which contains 256×256 pixels. A pixel with a value of color, c , in coordinate (l, s) is denoted as $ls(l, s, c)$, where l is from zero to L in the luminance domain, s is from 0 to S in the saturation domain and c is a function of *RGB*. For example, Fig. 12(a) illustrates the *LS color image* of Fig. 11(a). The luminance and saturation threshold values divide the *LS color image* into four quadrants. For example, quadrant 1 contains high saturation and low luminance pixels. The other quadrants can be defined similarly. Better experimental results are obtained via both luminance and saturation and by filling Quadrant 3 with white and the remaining with black. The LS color histogram is shown in Fig. 12(b), in which the number of pixels is represented by different colors. The less the number of pixels is, the closer is the color to blue. The more the number of pixels is, the closer is the

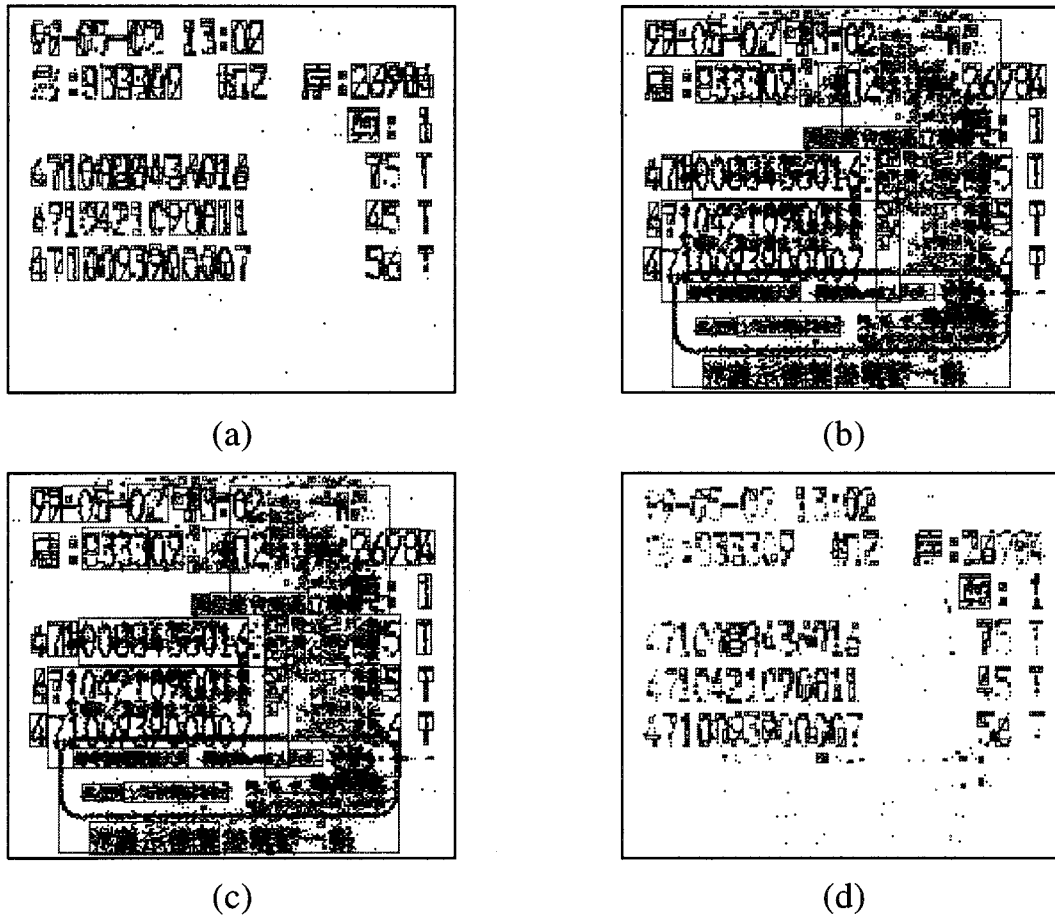


Fig. 17. Comparison with the results using counts of connected-components to evaluate the binarization of Fig. 1(a). (a) Our saturation-based thresholding, (b) Otsu's method, (c) Kittler's method, and (d) the triangle method.

color to red. In this figure, we show that it is proper to employ both luminance and saturation for thresholding.

Luminance and saturation-based thresholding is summarized as follows:

- 1) establish the luminance threshold value as the left valley of the background;
- 2) determine the saturation threshold value through the method introduced in Section IV-B;
- 3) if both the luminance and saturation values are greater than their threshold values, then output white otherwise output black.

V. EXPERIMENTAL RESULTS AND DISCUSSION

The binarization algorithm proposed for color document images was implemented as a Windows-based application on a Pentium II-350 PC. All test documents were scanned at a resolution of 300 dots per inch (dpi). The test documents and the threshold conclusions were saved as true color and binary images, respectively.

Our test color document images consist of 519 different color document images, which included images from uniform invoices, name-cards, calligraphy documents, advertisements, magazines, newspapers and brochures. Several fixed parameters were obtained experimentally and employed herein. Furthermore, to determine the size of the four features, such

TABLE II
TOTAL NUMBER OF CONNECTED-COMPONENTS, THE RATIOS OF SMALL, MEDIAN AND LARGE CONNECTED-COMPONENTS, THE SHAPE MEASURE AND THRESHOLD VALUES FOR FIGS. 1(A), 1(C) AND 11(A). F1: FEATURES, F2: FIGURES, NCC: NUMBER OF C.C., SCCR: SMALL C.C. RATIO, MCCR: MEDIAN C.C. RATIO, LCCR: LARGE C.C. RATIO, TV: THRESHOLD VALUE, L_t: LUMINANCE THRESHOLD, AND S_t: SATURATION THRESHOLD

F1	F2	Method	NCC	SCCR	MCCR	LCCR	SM	TV	
								L _t	S _t
Sat	1(a)	Our	638	0.0501	0.9498	0.0000	0.9851		237
		Otsu	699	0.5193	0.4692	0.0114	0.3299	232	
		Kittler	699	0.5193	0.4692	0.0114	0.3299	232	
		Triangle	934	0.1991	0.8009	0.0000	0.4159	217	
Lum	1(c)	Our	356	0.0056	0.9944	0.0000	0.7420	194	
		Otsu	482	0.1597	0.8402	0.0000	0.6278	157	
		Kittler	474	0.3924	0.6076	0.0000	0.3030	218	
		Triangle	371	0.2075	0.7925	0.0000	0.3224	216	
Sat + Lum	11(a)	Our	580	0.2503	0.7445	0.0052	0.9080	75	4
		Otsu	3371	0.5322	0.4666	0.0012	0.5753	122	
		Kittler	5044	0.4615	0.5378	0.0006	0.5973	134	
		Triangle	476	0.2495	0.7505	0.0000	0.1629	21	

as the luminance variance, σ^2 , of the entire image within the decision tree, 25 color documents for each class in the training set were employed.

The procedure applied to compute the threshold of the luminance variance, σ^2 , is as follows. 1) Compute the luminance variance, σ^2 , for all training images. 2) Employ the

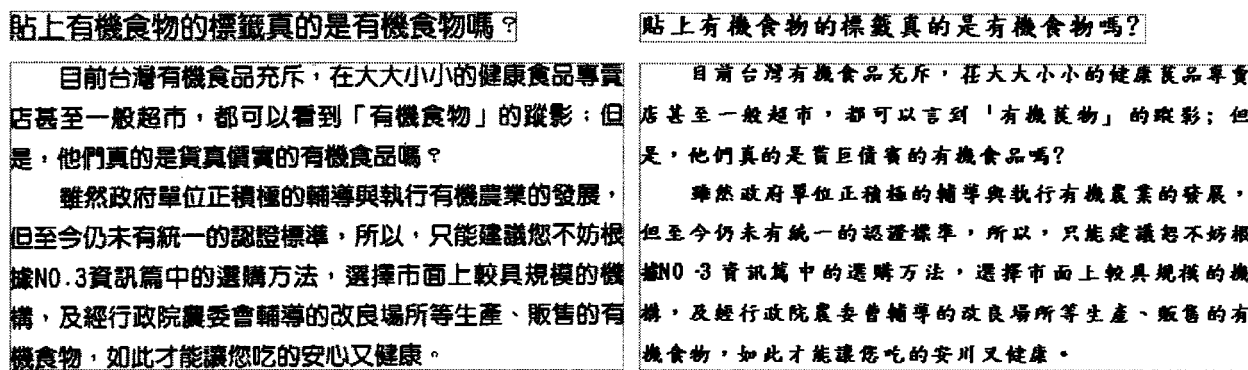


Fig. 18. An example of our binarization result from an OCR system. (a) Original document after layout analysis and (b) results after the OCR process. The recognition rate is 99.5%.

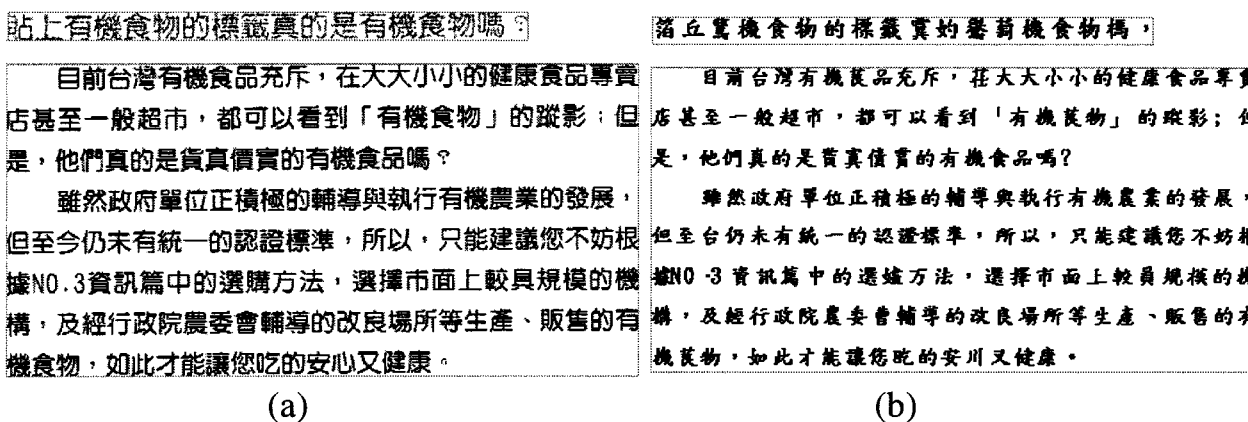


Fig. 19. An example of Otsu binarization result from an OCR system. (a) Original document after layout analysis and (b) result after the OCR process. The recognition rate is 88.5%.

saturation-based thresholding method to binarize all training images. 3) If the binarization results are satisfactory, the results were recorded as good. Otherwise, they were recorded as bad. 4) Sort the luminance variances, σ^2 , of all the training images in increasing order. 5) Set the lower bound of the bad binarization results as the threshold of the luminance variance, σ^2 . Fig. 13 illustrates the diagram of the training procedure. Herein, the luminance variance threshold was 400. To obtain parameter values in other decision flows, similar procedures were applied. Some examples are presented as follows.

According to the decision tree, Table I confirms the total images of the five cases employed in the experiments. Each case has a distinct ratio of images, which represent an appropriate classification of the decision tree. Furthermore, to threshold images of various classes varying features were applied.

Fig. 14 presents the binarization results of Fig. 1(a) via our saturation-based thresholding (Cases A and D) and other luminance-based thresholding methods. Fig. 1(a) has many reverse texts appear in the foreground. Fig. 14(b)–(f) confirm the experimental results of Tseng’s, Gong’s, Otsu’s, Kittler’s, as well as the triangle methods. For the latter three methods, the selected luminance threshold values were 232, 232, and 217, respectively. Tseng and Gong’s methods are hybrid thresholding methods. For our method, the saturation threshold value was 237. Our method successfully extracted the foreground text in the original color document image [Fig. 1(a)]. Others product

TABLE III
RECOGNITION RESULTS FOR BINARY IMAGES OBTAINED
BY DIFFERENT METHODS

	Niblack	Gong	Otsu	Kittler	Triangle	Our
Average recognition rate	66.89%	69.5%	59.3	53.07%	46.63%	80.06%

many noises, even, extract the reverse text. Hence, in color document images, which have inseparable foreground and background in luminance, that our method is better than other luminance-based thresholding methods.

Fig. 15 presents the binarization of Fig. 1(c) via our luminance-based thresholding method (Cases B and C) and other luminance-based thresholding method. Fig. 1(c) has a few text characters whose luminance values are proximal to background values. Fig. 15(a)–(f) depict the results of ours, Tseng’s, Gong’s, Otsu’s, Kittler’s, as well as the triangle methods. For the latter three cases, the selected luminance threshold values were 157, 218, and 216, respectively. Otsu’s method fails to consider the population ratio of the foreground and background, which is very small. If the text with luminance values proximal to the background is processed via Otsu’s method, then the result are unsatisfactory. Kittler’s method fails to consider the foreground, which is not a normal distribution. That is, it produces both noises and blurred characters. Notably, when a weak peak is produced by the foreground pixels in the histogram, the triangle

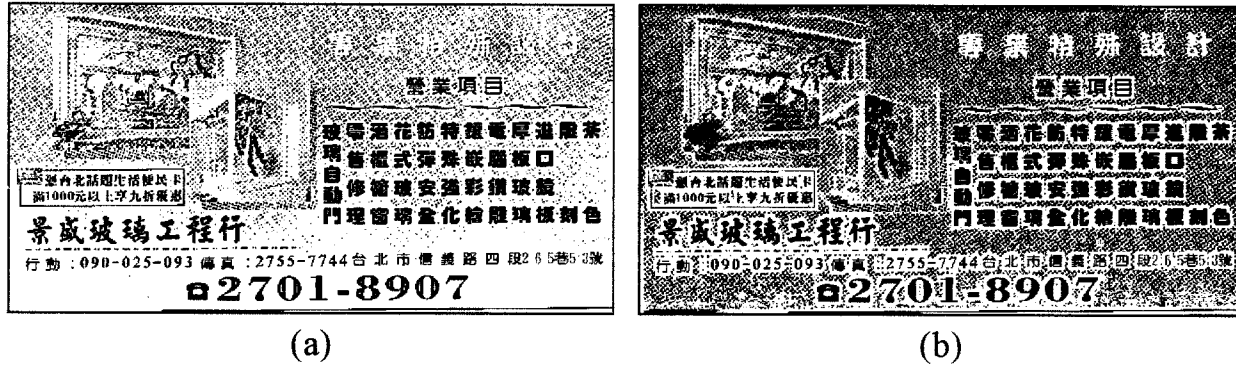


Fig. 20. An example to show that our method has better results than other local thresholding methods have. (a) Local Niblack's method and (b) Local Bernsen's method.

method is effective. However, in this case, the foreground pixels have several small peaks and scattered. Thus, rendering the binarization results unsatisfactory. The dark background cannot be properly determined in the first layer of Tseng's method. Thus, the binarization result is unsatisfactory. The hybrid Gong method, which employs both global gray values and local average gray values, yields best result. However, it generates several small holes in the character strokes. A comparison indicates that our luminance-based thresholding method binarization extracted characters the same as Gong's does, extracted more characters than Otsu's does and produce less noise than the other methods do. The advantages of our method are more text is extracted and a filter is not required to remove noise.

Fig. 16 presents the binarization of Fig. 11(a) via our luminance and saturation-based thresholding method (Case E) and other luminance-based thresholding method. Fig. 11(a) has many text characters with high luminance and high saturation, as well as many text characters with low luminance and high saturation. Fig. 16(b)–(f) depict the experimental results of Tseng's, Gong's, Otsu's, Kittler's, as well as the triangle methods. Again, for the latter three cases, the selected luminance threshold values were 122, 134, and 21, respectively. These methods produce unsatisfactory results for color document images, which have an inseparable foreground and background. Furthermore, when the character luminance is very proximal to the background, Tseng' and Gong's method produce unsuitable product. The luminance and saturation threshold values of our method were 75 and four, respectively. Therefore, as in the above figures, our method produces a satisfactory product.

From the above, our system produces a better product than other methods do regarding images have inseparable foreground and background. As well, to binarize color document images, both luminance and saturation can be used.

To evaluate the performance of the thresholding methods, uniformity and the shape measure are used broadly [7], [11], [36], [37], [39]. The uniformity measure (UM) adopted from Levine and Nazif [39] is defined as

$$UM = 1 - \frac{\sigma_w^2}{C} \quad (13)$$

where σ_w^2 denotes the within-class variance of the given threshold value and C is the normalization factor which limits

TABLE IV
EXECUTION TIME OF DIFFERENT BINARIZATION METHODS

Binarization methods	Document types	Average image size (pixels)	Average execution time (second)
Otsu's method	Uniform invoice	487x420	0.07
	Name-card	1029x625	0.18
	Calligraphy document	1370x500	0.20
	Advertisement	1290x641	0.23
	Magazine	840x408	0.10
Niblack's method	Newspaper	1078x1038	0.29
	Uniform invoice	487x420	17.38
	Name-card	1029x625	53.53
	Calligraphy document	1370x500	54.02
	Advertisement	1290x641	79.80
Our method	Magazine	840x408	32.79
	Newspaper	1078x1038	105.49
	Uniform invoice	487x420	0.09
	Name-card	1029x625	0.16
	Calligraphy document	1370x500	0.18
	Advertisement	1290x641	0.36
	Magazine	840x408	0.09
	Newspaper	1078x1038	0.24

the maximum value of UM to 1. Also, it is based upon the region and its grayscale range. Furthermore, to measure object shape within an image, the shape measure, SM , is used. The detail of shape measure can be found in [7]

The performance of the thresholding upon color document images cannot be discriminated easily with the uniformity measurement. To extract text from these images and to evaluate binarization, the shape measure and the small, median and large connected-components ratios are employed, respectively. Here, a connected-component is a maximal 8-connected pixel with equal foreground color.

A *connected-component* is classified as *small*, *median* and *large* by the following rules. A connected-component is *small* if the number of pixels NP in the component is less than four and greater than one, *large* if NP is greater than image area multiplied by a constant, k , which was predetermined as 0.0065 and otherwise, it is *median*. Furthermore, a *connected-component ratio* is the number of connected-components divided by the total number of connected-components. One of the main objectives of a color document analysis system is extracting text from color document images, which include binarization, segmentation, recognition, and applications. In binarization, if the luminance threshold value is high, the output may produce a few large connected-components. However, if the luminance

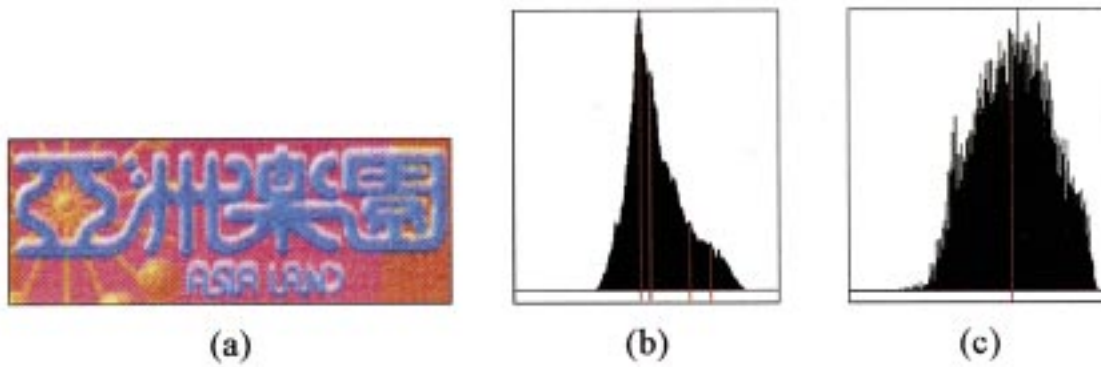


Fig. 21. Example 1, which is unsolvable by ours as well as other conventional methods. (a) A color image with both inseparable luminance and saturation, (b) luminance histogram, and (c) saturation histogram.

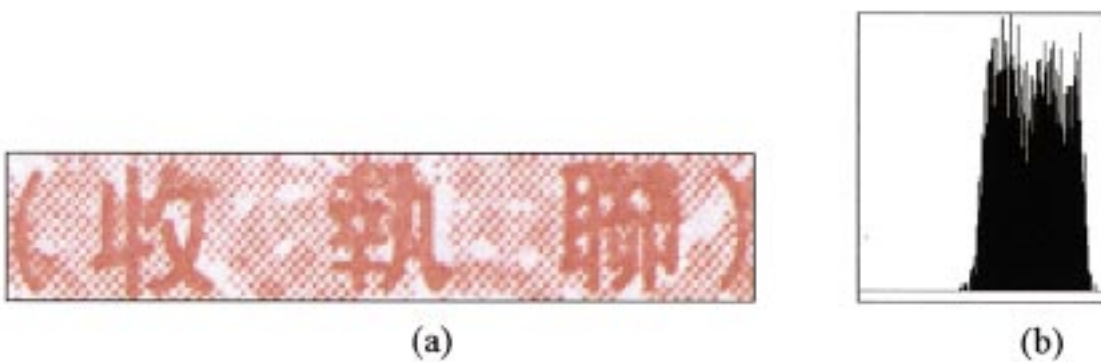


Fig. 22. Example 3, which is unsolvable by ours as well as other conventional methods. (a) Color half-tone image and (b) luminance histogram.

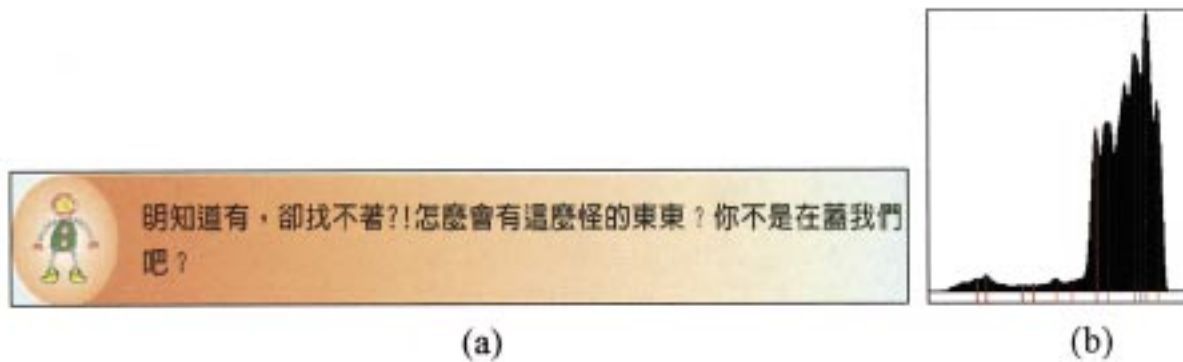


Fig. 23. Example 2, which is unsolvable by ours as well as other conventional methods. (a) Color image with gradually decaying background and (b) luminance histogram.

threshold value is low, segmentation may produce many small connected-components.

Fig. 17 depicts an example for the connected-components of Fig. 1(a), in which text on the reverse may appear on the front, especially with thin paper. The luminance values of the back text may be similar to that of the front. The methods of both Otsu and Kittler produce many large connected-components. This indicates that these two methods cannot process color document images, which have inseparable foreground and background colors. Similarly, the triangle method produces many small connected-components, which indicates an improper lu-

minance threshold value. In this instance, our method produces only a few small connected-components.

The total number of connected-components, the ratios of small, median and large connected-components, the shape measure and threshold values for Figs. 1(a), 1(c), and 11(a) are listed in Table II. From this table, the score of shape measure for our method is larger than Otsu's, Kittler's, and the triangle methods have. Regarding the small and median ratios of the connected-component, our method yields fewer small and more median connected-components. According to evaluation, our method has better performance than the others.

To give an objective comparison criterion, all binary images generated by each method were fed into a commercial OCR system produced by Newsoft Technology Corporation. Table III summarizes those results. From this table, our method obtains higher recognition accuracy than other methods. The result of our method is shown in Fig. 18 while Otsu's method in Fig. 19.

Compound document images contain graphic or textual content along with pictures [40]. They are common in magazines, brochures, web sites, newspapers, uniform invoices, name-cards, etc. For an application system to extract characters from name-card images, we first use the proposed binarization method to threshold input images. Before OCR, pictures, and graphics are removed by use of their properties such as the distributions, sizes and aspect ratios of connected components.

A general approach to compression of a compound image contains three major steps [41]: 1) image segmentation into the regions of similar data type; 2) selection of the best compression algorithm for each region; and 3) bit allocation among various regions. The segmentation approaches used in compound document compression can be grouped into three classes [41]: 1) object-based segmentation; 2) layer-based segmentation; 3) approximate block-based segmentation. Our proposed binarization method can be used in the first stage of compound image segmentation.

Fig. 20 depicts the results of other local thresholding methods. From this, our product [Fig. 16(a)] is better than Niblack's and Bernsen's are, in color documents with inseparable foreground and background. The proposed method attempts to improve the binarization obtained by global thresholding techniques, as local thresholding methods were time consuming.

Table IV shows the execution time of the methods employed. Our method is faster than the local Niblack's method, equal to Otsu's via luminance, but slower than Otsu's via both luminance and saturation.

Figs. 21–23 presents several bad thresholding results. That is, Fig. 21(a) is a color image with both inseparable luminance and saturation. Fig. 21(b) and (c) are its luminance and saturation histograms, respectively. Also, Fig. 22(a) is a color image with a gradually decaying background. Fig. 22(b) is the luminance histogram of the former figure, which has many background peaks. Similarly, Fig. 23(a) contains a color half-tone image, with its luminance histogram provided in Fig. 23(b). This histogram has many small peaks, which the proposed method cannot smooth easily.

VI. CONCLUSIONS AND FUTURE WORKS

This study has presented a new histogram-based thresholding method for thresholding color document images. A decision tree was employed to decide if luminance, saturation, or both features should be applied to threshold document images. The proposed method is trained on 150 different color documents and tested on an additional 519 color documents. The performance analysis indicated that our method is very efficient and effective in color documents with foreground and background colors either close or mixed. In the objective performance analysis by OCR system, our method obtains higher recognition accuracy

than other methods. Also confirmed is that the conventional thresholding method, when used for luminance, can also be used for saturation. In addition, via saturation or both luminance and saturation, a binarization method can be implemented. The measurement of the connected-component ratio can be used in color document images. Finally, the shape measure can also be used in color document images, when their foreground and background are proximal or mixed.

To obtain satisfactory binarization results and processing speed, future works should focus on the following: 1) Optimize our program; 2) process the color document images with background distribution in a lower luminance level; and 3) process the images with inseparable luminance and saturation, with gradually decaying background, with texture and finally, with half-tone colors.

ACKNOWLEDGMENT

The authors wish to thank the Associate Editor of the Transactions, Prof. Jezekiel Ben-Arie and anonymous reviewers who made many useful comments. Their help is gratefully appreciated.

REFERENCES

- [1] A. K. Jain, P. W. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, pp. 4–37, Jan. 2000.
- [2] G. Nagy, "Twenty years of document image analysis in PAMI," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, pp. 38–62, Jan. 2000.
- [3] G. Sharma and H. J. Trussell, "Digital color imaging," *IEEE Trans. Image Processing*, vol. 6, pp. 901–932, July 1997.
- [4] L. Lucchese and S. K. Mitra, "Advances in color image segmentation," in *Proc. GLOBECOM '99*, vol. 4, Dec. 5–9, 1999, pp. 2038–2044.
- [5] H. D. Cheng and Y. Sun, "A hierarchical approach to color image segmentation using homogeneity," *IEEE Trans. Image Processing*, vol. 9, pp. 2071–2082, Dec. 2000.
- [6] S. Wesolkowski, R. D. Dony, and M. E. Jernigan, "Global color image segmentation strategies: Euclidean distance vs. Vector angle," in *Proc. 1999 IEEE Signal Processing Society Workshop on Neural Networks for Signal Processing IX*, 1999, pp. 419–428.
- [7] P. K. Sahoo, S. Soltani, and A. K. C. Wong, "A survey of thresholding techniques," *Comput. Vis., Graph., Image Process.*, vol. 41, pp. 233–260, 1988.
- [8] N. Otsu, "A thresholding selection method from gray-scale histogram," *IEEE Trans. Syst., Man, Cybern.*, vol. 9, pp. 62–66, 1979.
- [9] W. H. Tsai, "Moment-preserving thresholding: a new approach," *Comput. Vis., Graph., Image Process.*, vol. 29, pp. 377–393, 1985.
- [10] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram," *Comput. Vis., Graph., Image Process.*, vol. 29, pp. 273–285, 1985.
- [11] S. U. Lee and S. Y. Chung, "A comparative performance study of several global thresholding techniques for segmentation," *Comput. Vis., Graph., Image Process.*, vol. 52, pp. 171–190, 1990.
- [12] J. Kittler and J. Illingworth, "On threshold selection using clustering criteria," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-15, pp. 652–655, 1985.
- [13] —, "Minimum error thresholding," *Pattern Recognit.*, vol. 19, no. 1, pp. 41–47, 1986.
- [14] Q. D. Trier and T. Taxt, "Evaluation of binarization methods for document images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, pp. 312–315, Mar. 1995.
- [15] Q. D. Trier and A. K. Jain, "Goal-directed evaluation of binarization method," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, pp. 1191–1201, 1995.
- [16] W. Niblack, *An Introduction to Digital Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1986, pp. 115–116.
- [17] J. Bernsen, "Dynamic thresholding of gray-level images," in *Proc. 8th Int. Conf. Pattern Recognition*, Paris, 1986, pp. 1251–1255.

[18] J. Z. Liu and W. Q. Li, "The automatic thresholding of gray-level pictures via two-dimensional otsu method" (in Chinese), *Acta Automat. Sinica*, vol. 19, pp. 101–105, 1993.

[19] J. Gong, L. Li, and W. Chen, "Fast recursive algorithms for two-dimensional thresholding," *Pattern Recognit.*, vol. 31, no. 3, pp. 295–300, 1998.

[20] Y. H. Tseng and H. J. Lee, "Document image binarization by two-layer block extraction and background intensity determination," *Comput. Vis., Graph., Image Process.*, 2000.

[21] B. Schacter, L. Davis, and A. Rosenfeld, "Scene Segmentation by Cluster Detection in Color Space," *Comput. Sci. Center, Univ. Maryland, College Park, Tech. Rep. 424*, 1975.

[22] A. Sarabi and J. K. Aggarwal, "Segmentation of chromatic images," *Pattern Recognit.*, vol. 13, no. 6, pp. 417–427, 1981.

[23] M. Celenk, "A color clustering technique for image segmentation," *Comput. Vis., Graph., Image Process.*, vol. 52, pp. 377–393, 1990.

[24] D. C. Tseng and C. H. Chang, "Color segmentation using UCS perceptual attributes," *Proc. Nat. Sci. Council A*, vol. 18, no. 3, pp. 305–314, 1994.

[25] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Reading, MA: Addison-Wesley, 1992, pp. 418–419.

[26] T. Y. Ohta, Kanade, and T. Sakai, "Color information for region segmentation," *Comput. Vis., Graph., Image Process.*, vol. 13, pp. 222–241, 1980.

[27] R. N. Strickland, C. S. Kim, and W. F. McDonnell, "Digital color image enhancement based on the saturation component," *Opt. Eng.*, vol. 26, no. 7, pp. 609–616, July 1987.

[28] B. A. Thomas, R. N. Strickland, and J. J. Rodriguez, "Color image enhancement using spatially adaptive saturation feedback," in *Proc. Int. Conf. Image Processing*, vol. 3, Santa Barbara, CA, Oct. 26–29, 1997, pp. 30–33.

[29] C. C. Yang and J. J. Rodriguez, "Efficient luminance and saturation processing techniques for color images," *J. Vis. Commun. Image Represent.*, vol. 3, no. 3, pp. 263–277, Sept. 1997.

[30] D. Hearn and M. P. Baker, *Computer Graphics, C Version*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1997.

[31] S. E. Umbaugh, *Computer Vision and Image Processing: A Practical Approach Using CVIPtools*. Englewood Cliffs, NJ: Prentice-Hall, 1998.

[32] Foley, F. Van Dam, Feiner, and Hughes. *Computer Graphics, Principles and Practice*, 2nd ed. Reading, MA: Addison-Wesley, 1996.

[33] G. Wyszecki and W. S. Stiles, "Color science—Concepts and methods," in *Quantitative Data and Formulae*. New York: Wiley, 1982.

[34] A. Witkin, "Scale-space filtering," in *Proc. 7th Int. Joint Conf. Artificial Intell.*, Karlsruhe, Germany, 1983, pp. 1019–1021.

[35] ———, "Scale space filtering: a new approach to multi-scale description," in *Image Understanding*, S. Ullman and W. Richards, Eds. Norwood, NJ: Ablex, 1984, pp. 79–95.

[36] D. M. Tsai, "A fast thresholding selection procedure for multimode and unimodal histograms," *Pattern Recognit. Lett.*, vol. 16, pp. 653–666, 1995.

[37] C. C. Chang and L. L. Wang, "A fast multilevel thresholding method based on lowpass and highpass filtering," *Pattern Recognit. Lett.*, vol. 18, pp. 1469–1478, 1997.

[38] G. W. Zack, W. E. Rogers, and S. A. Latt, *Automat. Meas. Sister Chrom. Exch. Freq.*, vol. 25, no. 7, 1977.

[39] M. D. Levine and A. M. Nazif, "Dynamic measurement of computer generated image segmentations," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-7, no. 2, pp. 155–164, 1985.

[40] R. L. D. Queiroz, Z. Fan, and T. D. Tran, "Optimizing block-thresholding segmentation for multilayer compression of compound images," *IEEE Trans. Image Processing*, vol. 9, pp. 1461–1471, Sept. 2000.

[41] A. Said and A. Drukarev, "Simplified segmentation for compound image compression," in *Proc. Int. Conf. Image Processing*, vol. 1, 1999, pp. 229–233.



Chun-Ming Tsai was born in Taipei, Taiwan, R.O.C. in 1963. He received the B.E. degree in Department of Industrial Education from National Taiwan Normal University, Taipei, in 1988. He completed forty credits course of study in Department of Computer and Information Science from National Chiao Tung University (NCTU), Hsinchu, Taiwan, in 1994 and received the M.S. degree from the Department of Computer Science and Information Engineering from NCTU, in 1996, where he is currently pursuing the Ph.D. degree in Department

of Computer Science and Information Engineering, National Chiao Tung University.

Since 1987, he has been a Teacher with the Information Department, National Tamshui Commercial Industrial Vocational Senior High School, Taipei. His research interests include color image processing, computer vision, pattern recognition, document analysis, computer graphics, VLSI design, and mobile computing.



Hsi-Jian Lee received the B.S., M.S., and Ph.D. degrees in computer engineering from National Chiao Tung University, Hsinchu, Taiwan, in 1976, 1980, and 1984, respectively.

Since August 1981, he has been with National Chiao Tung University as a Lecturer, Associate Professor and Professor. He was the Chairman of the Department of Computer Science and Information Engineering from August 1991 to July 1997. From January 1997 to July 1998, he was a Deputy Director of Microelectronic and Information Research Center (MIRC). Since August 1998, he has been the Chief Secretary to the president. He was the editor-in-chief of the *International Journal of Computer Processing of Oriental Languages* (CPOL) and associate editor of the *International Journal of Pattern Recognition and Artificial Intelligence and Pattern Analysis and Applications*. His current research interests include document analysis, optical character recognition, image processing, pattern recognition, digital library, and artificial intelligence.

Dr. Lee was the president of the Chinese Language Computer Society (CLCS), Program Chair of the 1994 International Computer Symposium and the Fourth International Workshop on Frontiers in Handwriting Recognition (IWFHR) and was the General Chair of the Fourth Asia Conference of Computer Vision (ACCV), in January 2000.