Text-Line Extraction and Character Recognition of Document Headlines With Graphical Designs Using Complementary Similarity Measure

Minako Sawaki and

Norihiro Hagita, *Member, IEEE Computer Society*

Abstract—A method for recognizing characters on graphical designs is proposed. A new projection feature that separates text-line regions from backgrounds, and adaptive thresholding in displacement matching are introduced. Experimental results for newspaper headlines with graphical designs show a recognition rate of 97.7 percent.

Index Terms—Character recognition, OCR, character segmentation, projection feature, displacement matching, adaptive thresholding.

———————————— ✦ ————————————

1 INTRODUCTION

THE optical character reader (OCR) plays an important role as a convenient device for transforming paper documents into a computer-readable format. One problem encountered by the OCR is that headlines or emphasized strings are often printed on top of graphical designs. For example, graphical designs are often used in Japanese newspapers to draw the reader's attention to certain articles. However, recognition of such decorated characters suffers from several problems:

1) Characters often have a variety of graphical designs such as black-plain or white-plain, outline, and textured characters (Figs. 1a-g).
2) Backgrounds also have a variety of graphical designs such as white-plain, black-plain, or graphical designs over all or part of the headline (Figs. 1a-g). Graphical designs sometimes change horizontally or vertically (Figs. 1e and 1f).
3) Text-lines can be printed horizontally or vertically (Figs. 1a-g). A headline sometimes consists of several text-lines (Figs. 1d and 1e).

To solve these problems, several methods for removing the graphical designs before recognition have been studied. Sakou et al. [1] focused on differences in the texture properties between the character and background regions. Okamoto et al. [2] and Liang et al. [3] proposed methods to remove background texture iteratively using morphological techniques. Ozawa et al. [4] attempted to separate character and background regions using gray-scale images. However, these methods often extract incomplete regions of character parts, that is, character strokes are eliminated or background designs are left in segmented character images. This occurs because the stroke configurations of complicated Kanji characters can be quite similar to textured backgrounds.

This paper proposes a method for recognizing characters without removing the graphical designs. Japanese newspaper headlines are selected as an example of this task, as headlines usually include keywords for queries, and their recognition is especially significant. The proposed method allows text-line regions to be separated from backgrounds and is independent of the language of the document. We have already proposed a robust recognition method for characters with graphical designs and degraded characters [5]. It utilizes a binary image as a feature vector and the complementary similarity measure as a discriminant function. To apply this method to the recognition of headlines, the number of text-line regions and character heights or widths must be extracted before recognition. A text-line region is, in this paper, a rectangular region that contains all characters in a text-line and its height or width is the average height, or average width, of all characters in the text-line, as shown in Fig. 5a.

First, a new projection value is introduced for separating the text-line regions from backgrounds in the headline images. The number of text-line regions and the averaged character heights are extracted from a local distribution of the projection values. Next, characters in the extracted text-line regions are recognized by displacement matching. When conventional methods recognize individual characters with no graphical designs, they select cut positions of individual characters in the text-line region. Features more effective than projection profiles of black or white pixels have been proposed for selecting the cut positions, such as the second differential of the marginal density [6], the upper- and lower-contour shapes of the characters [7], and the break cost [8]. However, these methods cannot readily handle headline images with graphical designs. Therefore, we perform recognition without selecting cut positions. Displacement matching is applied for this purpose.

Kovalevsky formulated a recognition algorithm based on displacement matching for text-lines [9]. Casey and Nagy [10] also applied this algorithm to the segmentation and classification of composite character patterns. In displacement matching, a character candidate window is matched against reference patterns in a dictionary using a similarity measure while being shifted pixel-by-pixel along the horizontal axis. In Kovalevsky's algorithm, a directed graph is obtained from the distribution of the similarity measure and window location. Recognition involves finding the maximal path of the directed graph. Our character recognition method is based on Kovalevsky's algorithm, and high recognition reliability is achieved using the complementary similarity measure which is sensitive to position translation and adaptive thresholding against the degree of degradation, instead of finding the maximal path.

We introduce the complementary similarity measure in Section 2, and a new projection measure for extracting the number of text-line regions and the averaged character heights in Section 3. Section 4 describes our method of recognizing individual characters in the text-line region by displacement matching and learning an adaptive threshold value. Section 5 presents experimental results. Section 6 presents conclusions.

2 COMPLEMENTARY SIMILARITY MEASURE

We take a brief look at the complementary similarity measure [5]. An input character is binarized and then normalized in size. The normalized character pattern ($n = N \times N$ pixels) is expressed as an n-dimensional binary feature vector. Now, let

$$F = \left(f_1, f_2, \ldots, f_i, \ldots, f_n \right) \text{ (where } f_i = 0 \text{ or } 1\text{)}$$

be the feature vector of an input character and

$$T = \left(t_1, t_2, \ldots, t_i, \ldots, t_n \right) \text{ (where } t_i = 0 \text{ or } 1\text{)}$$

be the feature vector of a binary reference pattern. The complementary similarity measure S_c of F to T is defined as

————————————————

• *M. Sawaki is with NTT Basic Research Laboratories, 3-1 Morinosato-Wakamiya, Atsugi, Kanagawa, 243-0198, Japan.*
   E-mail: minako@apollo3.brl.ntt.co.jp
• *N. Hagita is with NTT Communication Science Laboratories, 2-4 Hikaridai, Seika-cho, Soraku-gun, Kyoto, 619-0237, Japan.*
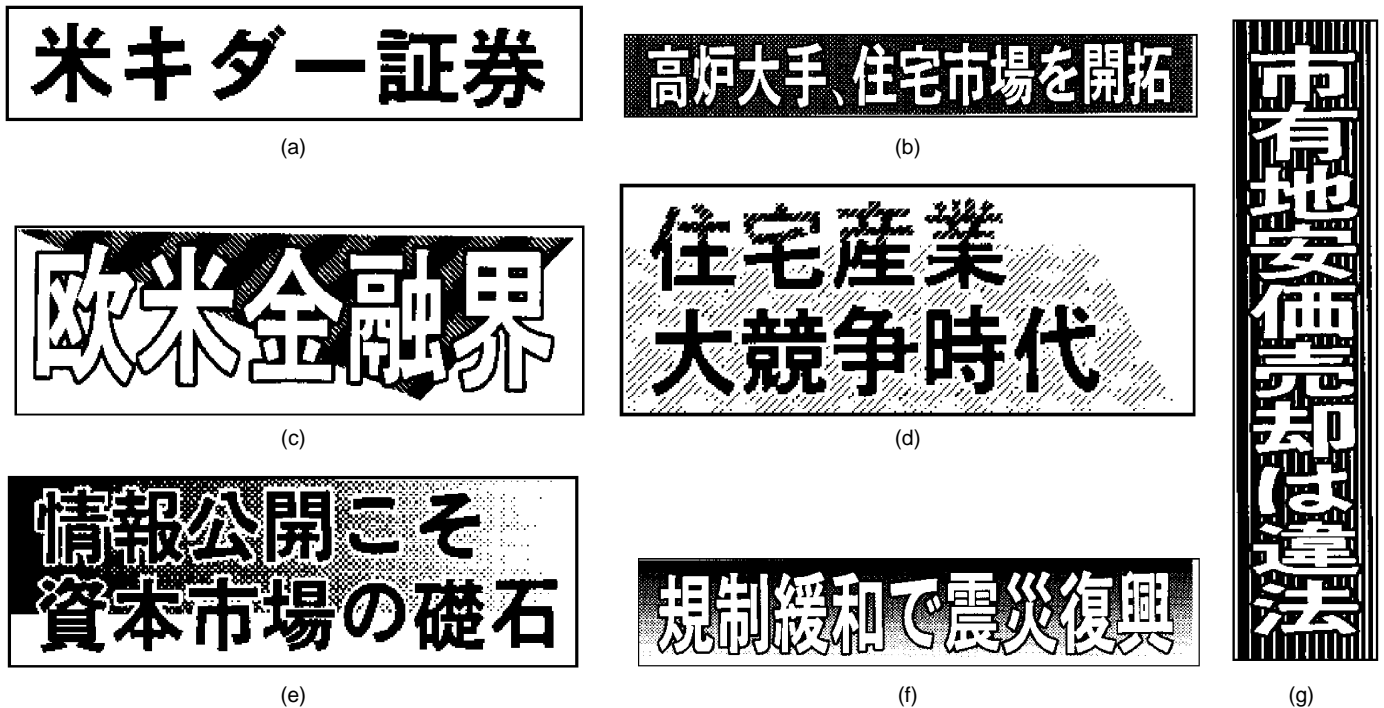   *E-mail: hagita@cslab.kecl.ntt.co.jp.*

Fig. 1. Examples of headline images. (a) Black-plain characters on white-plain background. (b) White-plain characters on textured background. (c) Outline characters on textured background. (d) Textured characters on textured background. (e) Black-plain characters on textured background. (f) White-plain characters on textured background. (g) White-plain characters on textured background.

$$S_c(\boldsymbol{F}, \boldsymbol{T}) = \frac{a \cdot e - b \cdot c}{\sqrt{T \cdot (n - T)}} \tag{1}$$

where

$$a = \sum_{i=1}^{n} f_i \cdot t_i, \quad b = \sum_{i=1}^{n} (1 - f_i) \cdot t_i, \quad c = \sum_{i=1}^{n} f_i \cdot (1 - t_i),$$

$$e = \sum_{i=1}^{n} (1 - f_i) \cdot (1 - t_i), \tag{2}$$

$$T = \|\boldsymbol{T}\|, \qquad a + b + c + e = n. \tag{3}$$

The value of S_c varies in the range

$$-\sqrt{T(n - T)} \le S_c \le \sqrt{T(n - T)}.$$

When an input character $\boldsymbol{F}$ comes from $\boldsymbol{T}$ with deletion noise, c becomes zero, and when $\boldsymbol{F}$ is derived from $\boldsymbol{T}$ with additive noise, b becomes zero. In both cases, provided that the term $b \cdot c = 0$, the similarity values of these degraded patterns are still high. This complementary relationship makes this measure robust against noise. The complementary similarity measure S_c has the following characteristic against the reverse contrast pattern $\boldsymbol{F}^c$ of $\boldsymbol{F}$.

$$S_c(\boldsymbol{F}^c, \boldsymbol{T}) = -S_c(\boldsymbol{F}, \boldsymbol{T}). \tag{4}$$

Equation (4) shows that by using the absolute value of the complementary similarity measure as a discriminant function, input characters may be recognized regardless of character color. In this paper, we use the measure $|S_c|$ (absolute value of S_c, $0 \le |S_c| \le \sqrt{T(n - T)}$) as a discriminant function in character recognition, since newspaper headlines can have different character colors. The dimension of the feature vector is $n = 32 \times 32 = 1,024$ pixels, which is large enough to express even complicated Kanji characters.

Fig. 2 shows recognition rates for the complementary similarity measure under the assumption that character heights or widths are correctly detected [5]. ETL2 (571 categories) was used, which is one of the standard machine-printed Kanji character databases offered by the Electrotechnical Laboratory. It is noted that the reference patterns were black-plain characters on white-plain backgrounds. The measure achieves over 98 percent recognition accuracy with printed Kanji characters with graphical designs.

3 TEXT-LINE REGION EXTRACTION BASED ON COMPLEMENTARY SIMILARITY MEASURE

This section introduces a new projection value for extracting text-line regions [11]. The value stems from the complementary relationships between characters and backgrounds in terms of black and white runs.

We assume that headline images are extracted from the newspaper, and that skews are corrected before text-line region extraction. This assumption is reasonable given conventional methods of skew correction and text and headline image extraction [12].

Japanese newspaper headlines fall into five types in terms of character parts and backgrounds. Table 1 shows the occurrence rates of these types for two of the main newspapers in Japan. Type V is seldom used, because it is not legible.

With no graphical designs (Type I), the projection profile of black or white pixels enables us to detect the text-line regions [12]. However, the projection profile technique is not applicable for headlines with graphical designs (Types II-IV). Therefore, we developed an alternative measure which estimates the number of text-lines and the averaged character heights of text-line regions that suits Types II-IV as well as Type I. This measure focuses on the complementary relationship between characters and backgrounds. That is, backgrounds usually are white-plain or black-plain for textured characters, while they usually contain texture or the reverse contrast color for black-plain and white-plain characters. These relationships are mixed for outline characters. Since textured

	Black-plain character	Textured character	Character with textured backgrounds	Outline character	Reverse contrast character
Examples					
Recognition rates	99.68%	99.79	99.72	98.74	99.82

Fig. 2. Recognition rates for characters with graphical designs [5] (Test samples in gothic style characters in newspapers in ETL2).

TABLE 1
TYPES OF HEADLINE IMAGES IN JAPANESE NEWSPAPERS (ASAHI NEWSPAPER AND NIKKEI NEWSPAPER)

Type	Characters	Backgrounds	Occurrence (%) Asahi	Nikkei
I	Black-plain	White-plain	43.5	62.9
	White-plain	Black-plain	1.6	0.5
II	Black-plain	Textured	15.4	13.9
	White-plain	Textured	22.9	17.8
	Black and White characters	Textured	2.3	1.0
III	Textured	White-plain	11.1	2.0
	Textured	Black-plain	0.3	
IV	Outline	Textured	2.3	2.0
V	Others		0.6	None
	Total number of headlines		306	202

backgrounds can change gradually along the horizontal or vertical axis, as shown in Figs. 1e and 1f, the number of text-lines and their character heights are locally estimated using the projection profile in a local rectangle window, which is shifted pixel-by-pixel along the direction of the text-line.

In this paper, we will explain the algorithm for a headline with horizontal text-lines. For a headline with vertical text-lines, the scanning direction is vertical.

Let G ($N_x \times N_y$ pixels) be the input headline image and G_w ($g_w(u, y)$; $W \times N_y$ pixels; $u = 1, 2, ..., W$; $y = 1, 2, ..., N_y$) be a rectangular window that can include at least one character. The projection value $p(y)$ enhancing the difference between the character parts and backgrounds for Types I-IV is defined by:

$$p(y) = \frac{a_p \cdot e_p - b_p \cdot c_p}{\sqrt{r_T(r - r_t) \cdot r_x(r - r_x)}},\qquad(5)$$

where

$$a_p = \sum_{u=1}^{W-1} g_w(u, y) \cdot g_w(u+1, y),$$

$$b_p = \sum_{u=1}^{W-1} (1 - g_w(u, y)) \cdot g_w(u+1, y),$$

$$c_p = \sum_{u=1}^{W-1} g_w(u, y) \cdot (1 - g_w(u+1, y)),$$

$$e_p = \sum_{u=1}^{W-1} (1 - g_w(u, y)) \cdot (1 - g_w(u+1, y)),\qquad(6)$$

$$r_T = a_p + b_p,\qquad r_X = a_p + c_p,\qquad(7)$$

$$a_p + b_p + c_p + e_p = r.\qquad(8)$$

This projection profile makes use of the complementary relationship of these four parameters (a_p, e_p, b_p, c_p). They correspond to the four possible changes of black and white pixels (black-to-black, white-to-white, white-to-black, and black-to-white) that can occur along the scan direction of the text-line. Using these four parameters, $p(y)$ is invariant for character color.

$p(y)$ lies in the range $-1 \le p(y) \le 1$. $a_p \cdot e_p$ corresponds to the product of the total of the black run-length and the total of the white run-length, and $b_p \cdot c_p$ corresponds to the square of line complexity in each scanning line. In general, $a_p \cdot e_p$ takes higher values in character parts and lower ones in black-plain or white-plain backgrounds. Also, $b_p \cdot c_p$ takes higher values for the textured backgrounds and lower ones for plain characters. As a result, $p(y)$ takes higher values in character parts than in backgrounds for Types I-IV. Since $p(y)$ is the same expression as four-fold point correlation, a special case of the complementary similarity measure, $p(y)$ comes from the complementary similarity measure.

The projection profile is averaged to avoid notches. The projection axis is divided into N sections and projection values $p(y)$ are averaged in each section. A group of high projection values in the vertical projection profile is then selected as a candidate of a text-line region. Its range h corresponds to the character height at location G_w. When the headline consists of two text-lines, the vertical projection profile generates two groups with high values. More specifically, when the averaged projection value exceeds 30 percent of the maximum of the projection profile, the section is determined to be a text-line region. This threshold was defined empirically. In practice, the text-line regions and character heights fluctuate somewhat over all G_w due to the different character heights at each location and the graphical designs. To avoid these notches, the text-line region and its character height are averaged over all G_w.

Fig. 3 shows an example of the averaged projection profile of $p(y)$ for G_w and the projection profile of black pixels without smoothing at $x = 1$. The two groups of high values in the averaged projection profile of $p(y)$ in Fig. 3b show the existence of two text-line regions and their character heights, although the projection values of black pixels in Fig. 3c do not indicate the text-line regions
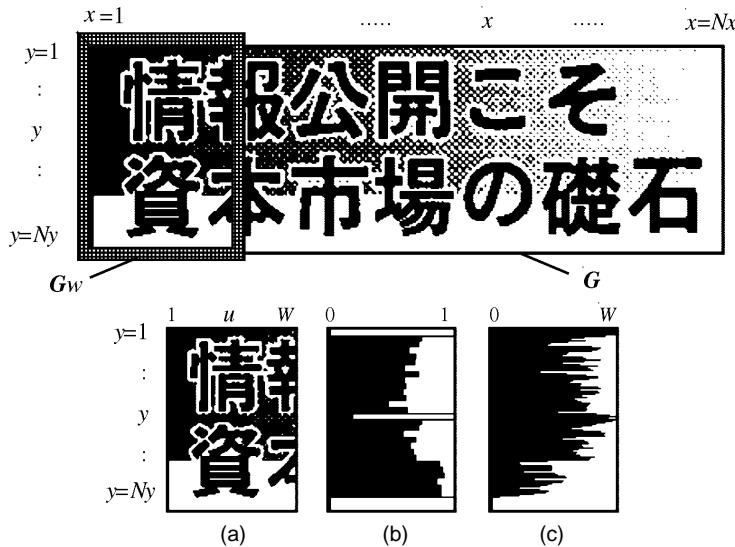
Fig. 3. Observation window for region extraction and projection profile of *p(y)*.

well. Fig. 4 shows the local distributions of $p(y)$ for four types of headline images at $x = 1$. This figure shows that with the proposed measure, the number of text-lines and the character heights are estimated for headline images of Types I-IV.

4 CHARACTER RECOGNITION BASED ON COMPLEMENTARY SIMILARITY MEASURE

4.1 Displacement Matching

Since it is difficult to select cut positions of individual characters even with the proposed measure $p(y)$, displacement matching is applied to the extracted text-line region for character recognition. When estimating text-line positions, text-line candidates are obtained for each observation window and the candidate positions are averaged. However, for segmenting individual characters, this averaging effect is not available because a text-line has only one character in the vertical direction and the segmentation fails easily. Though displacement matching may require much processing time, this method seldom overlooks existing characters in a text-line.

One of the problems of displacement matching is that extra candidates may be encountered; that is, false categories may be selected

Fig. 4. Local distributions of *p(y)* for various graphical designs. *h*: character height for each text-line region. (a) Type I. (b) Type II. (c) Type III. (d) Type IV.

at locations where no correct character category is located [9]. In the proposed method, the complementary similarity measure and adaptive thresholding suppress these spurious candidates. The complementary similarity measure is sensitive to position translation and takes high values when the input character is matched against the reference pattern of its character category while taking low values for the other categories. These properties hold even for characters with graphical designs or noise. Therefore, whenever the square window is located at a correct character position, the similarity between the window and a reference pattern of the correct category is maximal; local peaks are observed in the distribution of maximum similarity value only at correct character positions.

The extracted text-line region with height h is normalized to yield normalized text-line region L with height N. An $N \times N$-pixel square window is selected for matching normalized text-line region L, since the binary reference patterns consist of $N \times N$ pixels. The square window is shifted pixel-by-pixel along the direction of the text-line and the window's contents, input feature vector F, are compared to the reference patterns using the complementary similarity measure. The binary reference patterns are made of character patterns without graphical designs in advance. As the binary pattern matching method is used, the similarity decreases significantly when the input is shifted in the vertical direction. This shift may be due to fluctuation in text-line extraction. The amount of pixel shift that can be tolerated mainly depends on character stroke width because the similarity value is strongly related to the number of pixels to which the input pattern and reference pattern overlap. For example, the stroke width of the reference patterns we use is within six pixels. Therefore, it is expected that vertical shifts of a few pixels (less than six) can be tolerated when these reference patterns are used.

Fig. 5 shows an example of the distribution of maximum complementary similarity value and the maximum conventional similarity value at each position along the horizontal axis for the normalized text-line region. Fig. 5b and Fig. 5d show that the complementary similarity measure has dominant peaks around the left-most side of F of the exactly correct category, while the conventional similarity measure $S(F, T) = a/\sqrt{F \cdot T}$, $F = \|F\|$ has few peaks.
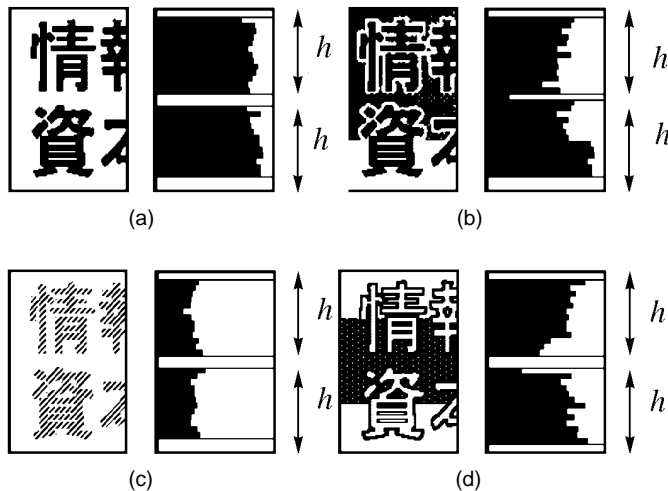
4.2 Adaptive Thresholding

When the maximum $|S_c|$ at each position exceeds the threshold, the recognized category and its position are determined. To recognize
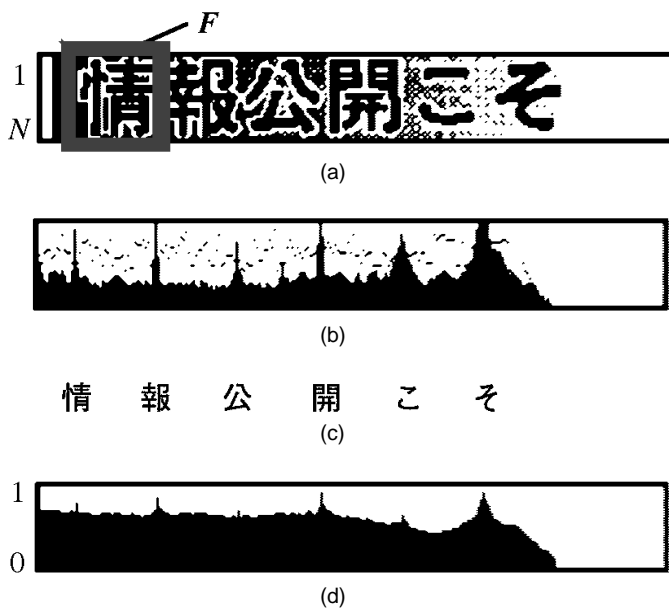
Fig. 5 Character recognition by displacement matching. (a) Normalized text-line region L. (b) Maximum $|S_c|$ (solid bars) and estimated thresholds $Th(T_i, Z)$ (dot plots). (c) Recognized categories. (d) Maximum S.

characters with high precision, the most appropriate threshold should be determined using the degree of degradation in F and the reference pattern of the category with the maximum $|S_c|$ value. When F is located at the correct character position during displacement matching, $|S_c|$ of F to the reference pattern of the correct category decreases according to the degree of degradation while $|S_c|$ of F to the reference patterns of the other categories decrease more rapidly. Therefore, by assuming that the square window F at each position contains the reference pattern having the maximal similarity value, the degree of degradation for F and the adaptive threshold against the degree can be estimated.

The thresholds for different degradation degrees are determined in advance by a learning process for each reference pattern. Let T_i be a reference pattern of the ith category without noise, and

Z_i and Z be a noisy image of T_i and the number of black pixels in Z_i, respectively. Also, let $Th(T_i, Z)$ be a threshold value for T_i when the number of black pixels of an input pattern is Z. A noisy image Z_i is synthesized from T_i and a random dot image [5]. Similarity values $|S_c(Z_i, T_j)|(j \neq i)$ are calculated for different Z. Fig. 6 shows an example of the threshold-to-degradation table, that is the relationship between the maximum $|S_c(Z_i, T_j)|$ for $j = $ 報 and Z. For comparison, Fig. 6 also shows $|S_c(Z_j, T_j)|$ for $j = $ 報. In order to eliminate false recognition candidates, the maximum $|S_c(Z_i, T_j)|$ is determined as the threshold $Th(T_j, Z)$ for the reference pattern T_j. For $|S_c|$, we select $\max[Th(T_j, Z), Th(T_j, n - Z)]$ as the threshold. Other reference patterns showed the same tendency as 報.

5 EXPERIMENTAL RESULTS

5.1 Headline Data

As test data, we used 50 headlines, Types II-IV, in Japanese newspapers (25 horizontal and 25 vertical text-lines) which contained 529 characters. They were gathered by using three binarization thresholds, level 1 (low), level 2 (fine), and level 3 (high). The number of text-lines in one headline was either one or two. The character font in the headlines was Gothic. Reference patterns without graphical designs were extracted manually from 121 headline images at level 2. Fifty headlines out of 121 were the same as the test data; the remaining 71 were also gathered to increase category number. As the aspect ratio of characters differs with text-line direction, the reference patterns were stored in either a horizontal text-line dictionary or a vertical text-line dictionary according to the direction of the text-line of the headline. The number of reference patterns was 913 (500 categories) for the horizontal text-line dictionary and 988 (525 categories) for the vertical text-line dictionary. The numbers of reference patterns per category in the horizontal dictionary and the vertical text-line dictionary were in the range of one to 32 and one to 23, respectively. The horizontal text-line dictionary occupied 139 KB while the vertical text-line dictionary took 150 KB.
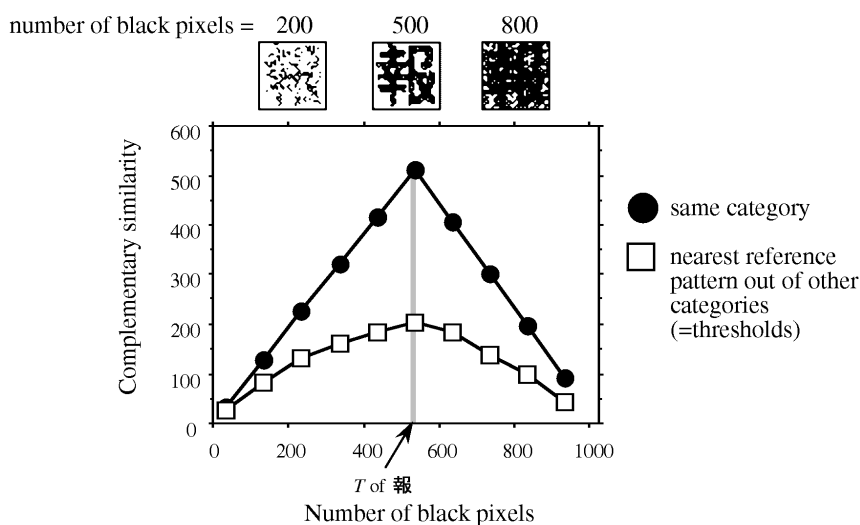

Fig. 6. Relationship between complementary similarity and number of black pixels (Example for reference pattern 報).

TABLE 2
RESULTS OF TEST-LINE EXTRACTION

Threshold level for binarization	# of data	level 1 (low)	level 2 (fine)	level 3 (high)
horizontal text-lines	25	25	25 (7)	23
vertical text-lines	25	22	25 (7)	22

(): conventional method

TABLE 3
RECOGNITION RESULTS

Threshold level for binarization	level 1 (low)	level 2 (fine)	level 3 (high)
Recognition rates (conventional method)	81.3 %	97.7 (17.0)	82.4

TABLE 4
RECOGNITION RESULTS FOR IMAGES IN FIG. 1 (LEVEL 2)

Image number	recognition results
(b)	高炉大手筑住宅市場1を1開拓
(c)	住宅産業題大競争時代例
(d)	欧米金融ト
(e)	情報公開こそ資本市場、の礎石
(f)	規制「緩和で震「災復興
(g)	市有地安価売却は違法

5.2 Text-Line Extraction

Experimental results for text-line extraction are listed in Table 2. Table 2 shows the numbers of images whose all-text-lines were correctly extracted. We determined that text-lines in which all characters were misrecognized as NOT correctly extracted. For images at level 2, all text-lines were correctly extracted, while a few errors occurred for images at levels 1 and 3. One of the reasons for error was that for the images with striped graphical designs whose direction matched the projection direction, the projected value is similar to that for the character regions and for the background regions. The average processing time was 10.2 seconds per one headline image on a Sun Sparcstation 10.

For comparison, a conventional method was applied to the test data with level 2 scanning. Graphical designs were removed using the method in [2] and text-lines in the resulting images were extracted with commercial OCR software. The method in [2] extracts character strokes based on morphological approach. The commercial OCR extracts text-lines using projection features of black pixels. This table shows that the proposed projection value achieves much higher text-line extraction rates than the conventional method. The error of the conventional method comes from the confusion of text regions with backgrounds and the confusion of text and graphic regions.

5.3 Character Recognition

Character recognition was conducted for the test headline images including images with text-line extraction errors. In recognition, the horizontal (vertical) text-line dictionary was used when the width of the input headline image was larger (smaller) than its height. When recognized characters overlapped each other, the category, which had the highest similarity, was kept and the others were removed from the recognition results.

The recognition results for 50 test headline images are shown in Table 3. Table 3 shows that the recognition rate for level 2 is 97.7 percent and the recognition rates for levels 1 and 3 are 81.3 percent and

82.4 percent, respectively. The number of extra candidates are 40 (level 1), 18 (level 2), and 43 (level 3), respectively. Recognition results for the images in Fig. 1 are listed in Table 4. The average processing time was 39.2 seconds per headline image on the Sun Sparcstation-10.

For comparison, a conventional method was applied to the test data with level 2 scanning. Graphical designs were removed using the method in [2] and the resulting images were recognized with commercial OCR software. This OCR software did not learn the reference patterns we obtained from headline images. The recognition rate was 17.0 percent. This shows that the proposed method achieves much higher recognition rates than the conventional method, since it is noted that our test data did not include headline images in Type I but in Types II-IV.

Recognition errors in our method fall into two main sources: unsuccessful extraction of text-line regions and adaptive thresholding. Errors in adaptive thresholding occurred when the maximum similarity of the correct category was less than the estimated threshold. The complementary similarity measure achieves over 98 percent recognition accuracy against both learning and test samples when the character height is known [5]. The accuracy of headline image recognition will be improved by eliminating these two sources of recognition error.

6 CONCLUSIONS

We proposed a method for text-line extraction and character recognition of Japanese newspaper headlines with graphical designs. A new projection measure extracts text-line regions by focusing on the complementary relationship between characters and backgrounds. Characters in the text-lines are recognized by the complementary similarity measure and displacement matching. Spurious candidates in displacement matching are suppressed by adaptive thresholding. Experimental results for 50 newspaper headlines showed that this method achieves a high recognition rate of 97.7 percent, which is higher than the 17.0 percent of a conventional method. Improving text-line extraction and thresholding are future tasks.

REFERENCES

[1] H. Sakou, H. Matsushima, and M. Ejiri, "Texture Discrimination Using Self-Organized Multiresolution Filtering," *IEICE Trans. D-II,* vol. J73-D-II, no. 4, pp. 562-573, 1990 (in Japanese).

[2] M. Okamoto and H. Hayashi, "Character Extraction From Headlines With Background Patterns by Using Shrinking/Expanding Methods," *IEICE,* Technical Report PRU90-151, pp. 47-54, 1991 (in Japanese).

[3] S. Liang and M. Ahmadi, "A Morphological Approach to Text String Extraction from Regular Periodic Overlapping Text/Background Images," *Computer Vision, Graphics, and Image Processing,* vol. 56, no. 5, pp. 402-413, Sept. 1994.

[4] H. Ozawa and T. Nakagawa, "A Character Image Enhancement Method From Characters With Various Background Images," *Proc. Second Int'l Conf. Document Analysis and Recognition,* pp. 58-61, Tsukuba, Japan, Oct. 1993.

[5] M. Sawaki and N. Hagita, "Recognition of Degraded Machine-Printed Characters Using a Complementary Similarity Measure and Error-Correction Learning", *IEICE Trans. Information and Systems,* vol. E79-D, no. 5, pp. 491-497, 1996.

[6] S. Kahan and T. Pavlidis, "On the Recognition of Printed Characters of Any Font and Size", *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 9, no. 3, pp. 274-287, 1987.

[7] R. Fenrich, "Segmentation of Automatically Located Handwritten Words," *Proc. Second Int'l Workshop on Frontiers in Handwriting Recognition,* pp. 33-44, Chateau de Bonas, France, 1991.

[8] S. Tsujimoto and H. Asada, "Major Components of a Complete Text Reading System," *Proc. IEEE,* vol. 80, no. 7, pp. 1,133-1,149, 1992.

[9] V.A. Kovalevsky, *Image Pattern Recognition.* Berlin: Springer, 1980.

[10] R.G. Casey and G. Nagy, "Recursive Segmentation and Classification of Composite Character Patterns," *Proc. Sixth Int'l Conf. Pattern Recognition,* pp. 1,023-1,026, 1982.

[11] M. Sawaki and N. Hagita, "Text-Line Extraction and Character Recognition of Japanese Newspaper Headlines With Graphical Designs," *Proc. 13th Int'l Conf. Pattern Recognition,* Track C, pp. 73-78, 1996.

[12] T. Akiyama and N. Hagita, "Automated Entry System for Printed Documents," *Pattern Recognition,* vol. 23, no. 11, pp. 1,141-1,154, 1990.