Correction of Signboard Distortion by Vertical Stroke Estimation

Jun Sik Lim[1] , In Seop Na[1*] and Soo Hyung Kim[1]
[1] School of Electronics & Computer Engineering, Chonnam National University
Gwangju, 500-757 - Korea
[e-mail: , ljs-32@hanmail.net, ypencil@hanmail.netshkim@jnu.ac.kr]
*Corresponding author: In Seop Na

Received February 14, 2013; revised April 26, 2013; accepted May 20, 2013; published September 30, 2013

Abstract

In this paper, we propose a preprocessing method that it is to correct the distortion of text area in Korean signboard images as a preprocessing step to improve character recognition. Distorted perspective in recognizing of Korean signboard text **may** cause of the low recognition rate. The proposed method consists of four main steps and eight sub-steps: main step consists of potential vertical components detection, vertical components detection, text-boundary estimation and distortion correction. First, potential vertical line components detection consists of four steps, including edge detection for each connected component, pixel distance normalization in the edge, dominant-point detection in the edge and removal of horizontal components. Second, vertical line components detection is composed of removal of diagonal components and extraction of vertical line components. Third, the outline estimation step is composed of the left and right boundary line detection. Finally, distortion of the text image is corrected by bilinear transformation based on the estimated outline. We compared the changes in recognition rates of OCR before and after applying the proposed algorithm. The recognition rate of the distortion corrected signboard images is 29.63% and 21.9% higher at the character and the text unit than those of the original images.

Keywords: Korean Signboard, Text Area, Correcting perspective distortion, Components detection, Text boundary estimation

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MEST)(2012-047759) and this research was supported by the MSIP(Ministry of Science, ICT&Future Planning), Korea, under the ITRC(Information Technology Research Center) support program (NIPA-2013-H0301-13-3005) supervised by the NIPA(National IT Industry Promotion Agency).

http://dx.doi.org/10.3837/tiis.2013.09.014

1. Introduction

Studies on the development of various applications running in mobile devices equipped with a high-resolution camera to provide much information to the mobile users are being performed [12,14,15]. As a typical example, content including the name and brief description of shops or institutions can be provided instantly to foreign tourists by recognizing the text of signboard images and converting them to different languages using language translators [13].

Existing studies on the correction of text image distortion are based on analyses on the outline of text images and the formation of text [2,4,6,10,16]. The approaches to distortion correction includes the method to estimate a perspective projection matrix from a vanishing point [2,7,11]; the method to perform affine transformation by analyzing the formation of the text line [8]; and the method to recognize distorted texts without correcting the distortion using the structural invariables of the text [4].

The extraction of the boundary quadrilateral may not be accurate and the boundary quadrilateral may not always be visible in the complete text area. Horizontal and vertical vanishing points were calculated using the composition of text instead of the text boundary to overcome these problems [2,3]. They assume uniform spacing between text lines to compute the vertical vanishing point. Although the performance has improved, these assumptions are difficult to apply to non-document scene text. However, the performance of detection of those invariants, such as ascender and descender, highly depends on the length of text lines, hence this method may fail when a text line is too short.

Korean signboard images are composed of only a few characters, as shown in **Fig. 1.** In addition, it is difficult to detect the outline and correct the image distortion. Thus, we have examined distorted Korean signboard images to resolve these problems. In general, the camera viewpoint causes a perspective distortion of the signboard images, as shown by the dashed line in **Fig. 1.**

Fig. 1. Korean signboard images where perspective distortion are shown by dashed lines

We measured the variation of the slopes of vertical line components in the signboard images and estimated trapezoidal outlines of the text region to estimate the distortion. Our method can estimate the trapezoidal outline correctly at least two vertical line components exist in the signboard text.

The remainder of this paper is organized as follows. Section 2 describes the proposed method, including how to detect vertical line components and how to correct the distortions. Section 3 provides experimental results. Section 4 discusses and summarizes the unique characteristics and limitations of this study.

2. Proposed Method

Fig. 2 shows the flow chart of the proposed system. The input image is a binary image resulting from a text extracting module in Park [5] to the text region of a color signboard image shot by a mobile phone. There are only a few characters in Korean signboard images and their size and between-characters gaps are irregular. Thus, it is very difficult to analyze the structure of the text region. Korean characters consist of a combination of five basic patterns, '—' (horizontal), '|' (vertical), '/' (diagonal), '\' (inverse diagonal) and 'O' (round). For example, the character "성" in Figure 1 consists of '/', '\', '—', '|' and 'O'. Here, we pay attention only to the vertical line component '|', because it appears relatively more frequently in Korean text than in other languages and the trend of their slopes can be used to describe the amount of image distortion. **Fig. 2** shows the flow chart of the proposed system. The input image is a binary image resulting from a text extracting module in Park [5] to the text region of a color signboard image shot by a mobile phone.

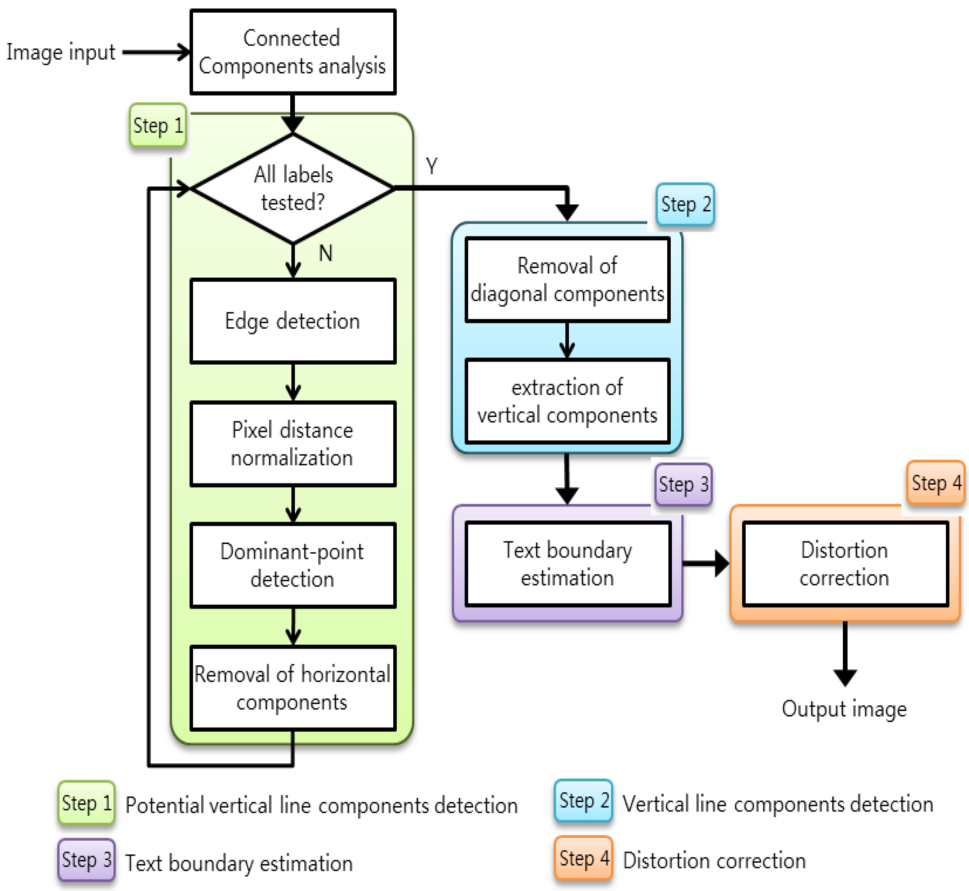

Fig. 2. Flowchart of the proposed system

The proposed system consists of four stages, potential vertical line components detection, vertical line components detection, text boundary estimation and distortion correction. First, potential vertical line components detection consists of four steps, including edge detection for each connected component, pixel distance normalization in the edge, dominant-point detection in the edge and removal of horizontal components. Second, vertical line components detection is composed of removal of diagonal components and extraction of vertical line components. Third, the outline estimation step is composed of the left and right boundary line detection. Finally, distortion of the text image is corrected by bilinear transformation based on the estimated outline.

2.1. Potential Vertical Line Components Detection

Given a binarized text region of a signboard image, as in **Fig. 3(a),** the characters in the text image are divided into letters using connected components analysis. For each of the divided letters, edges are detected by applying a Canny mask **Fig. 3(b).** Distance-based normalization is performed to remove all other points, except for the central point of an m by m sized mask, by applying the mask to the edge. This normalization can remove noise generated in the shooting of the image and the binarization process, while preserving the straightness of the edges. In our experiments, the size of m is empirically set to 5. Thus, the distance between points on the edge is normalized at regular intervals **Fig. 3 (c)** and is defined as $\mathrm{N} = \{\mathrm{n_0}, \mathrm{n_1}, \cdots, \mathrm{n_k}\}$. Next, the dominant points in the edge are found and each line component between two dominant points is further considered to detect potential vertical line components (candidates of vertical line components) in the normalized edge. To do this, let the set of the dominant points (pixels) in the edge of each letter and the connections between $\mathrm{n_{i-1}}$ and $\mathrm{n_i}$ be chain-coded along the edge. Only the 0th to 7th codes were used from the East through North to West of the 16 directions to remove the repeated codes. For more detail, let $\mathrm{d(n_{i-1}, n_i)}$ be the direction code between two points from $\mathrm{n_{i-1}}$ to $\mathrm{n_i}$ with condition of $\mathrm{d(n_{i-1}, n_i)} = \mathrm{d(n_i, n_{i-1})}$. We consider $\mathrm{n_i}$ to be a dominant point if the direction from or to $\mathrm{n_i}$ changes suddenly, as Eq. 1.

$$|d(n_{i-1}, n_i) - d(n_i, n_{i+1})| \geq 2 \qquad (1)$$

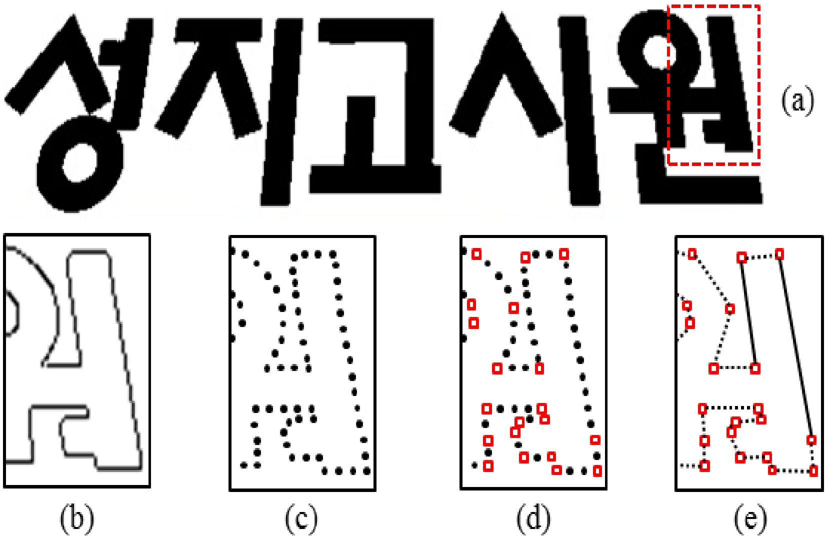

Fig. 3. Extraction steps of potential vertical line components in a Korean signboard image. (a) distorted input image with dashed line for partial image, (b) edge detected by Canny operator in the partial image, (c) Distance-based normalization with 5 by 5 mask, (d) dominant points marked with red squares, (e) potential vertical line components in solid line

This normalization can remove noise generated in the shooting of the image and the binarization process, while preserving the straightness of the edges. Every edge is automatically divided at the dominant point **Fig. 3(d).** Now, let the set of line components between two dominant points be $S = \{s_0, s_1, \cdots, s_m\}$, then S includes vertical, horizontal and diagonal components. To preserve vertical line components, whilst removing horizontal components, the length of each individual component is calculated and the linear equation (l_i) is found using two dominant points of the component (two end pixels of the component). Then, any components meeting the condition, as shown in Eq. 2, are considered as potential vertical line components (V^C).

$$V^C = \left\{s_i \;/\; s_i \subseteq S \,, \mathrm{length}(s_i) > \frac{H}{5} \,, @\, l_i \cap (y = H) \in I \,,\, l_i \cap (y = 1) \in I\right\} \quad \textbf{(2)}$$

Where I is the input image, $\mathrm{length}(s_i)$ is the length of the component and H is the height of the image. In addition, the linear equation should cross the top ($y = H$) and the bottom ($y = 1$) of the input image, because the vertical line components in Korean text images generally take place from top to bottom of the image, despite distortion. Here, we eliminate small components with the scondition $\mathrm{length}(s_i) > \frac{H}{5}$.

Now, we have the potential vertical line components without horizontal ones **Fig. 3(e).** Each linear equation of the potential vertical line components is defined as $l_j = y = a_j x + b_j$ with the slope angle of θ_j in image coordinates (x and y) where $j = 1, \ldots, n$ and a set of linear equations for the potential vertical components is defined as $L = \{l_1, \cdots, l_n\}$ (see **Fig. 4).**

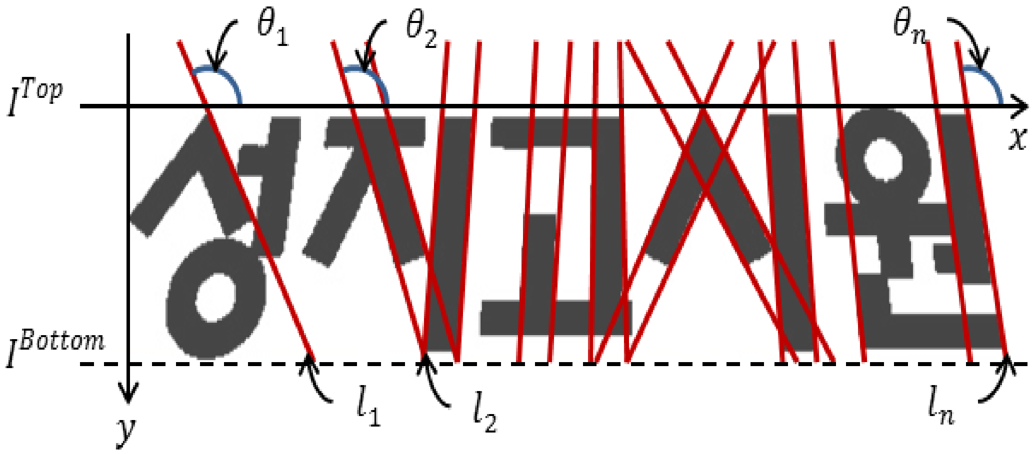

Fig. 4. Example of potential vertical line components with slope angles

2.2. Vertical line components detection

Generally, the slope angles of the vertical line component increases from left to right in the distorted text region. However, the diagonal components are located in irregular directions. Thus, our goal here is to find the regression line for genuine vertical line components, as the solid regression line in **Fig. 5.**

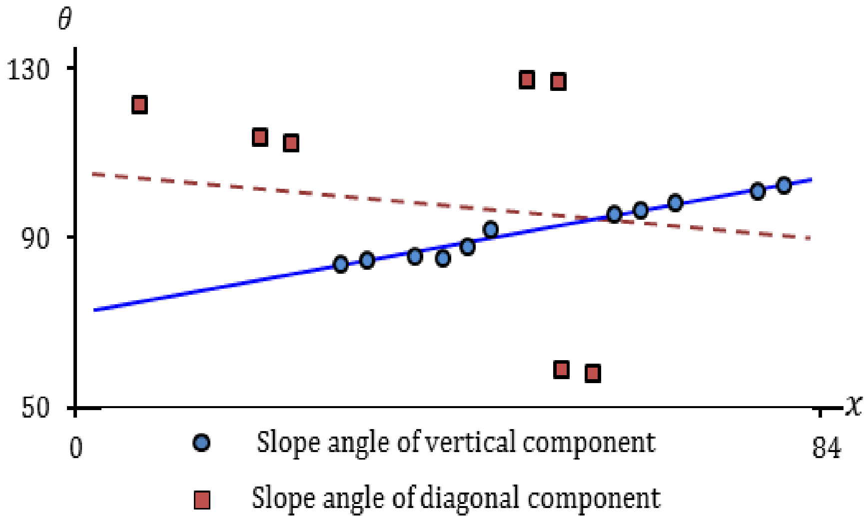

Fig. 5. Regression lines of slope angles of potential vertical line components. The dashed line and the solid line present the regression for all potential vertical line components, including genuine vertical line components, respectively.

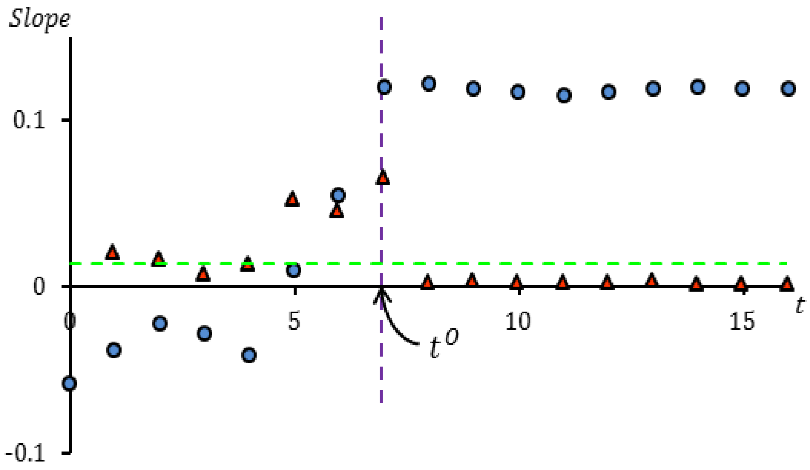

Fig. 6. Finding an optimal iteration index (t^0) to remove diagonal components. The circular shape is the slope of the regression line of all the potential vertical line components, while the triangular shape is the slope difference between two iterations (t and t-1). Vertical dashed lines and horizontal dashed lines indicate an optimal iteration index (t^0) and the average slope difference, respectively.

The flowing is the process of finding the regression line. Step 1. A linear equation of the slope angles (θ) of all the potential vertical line components is estimated using a LLS(linear least squares) as the dotted regression line in **Fig. 5.** To remove diagonal components, a potential vertical line component causing the maximum error with perspective to the regression line is eliminated. Step 2. While repeating Step 1, the proposed method saves the current iteration index (t), the slope angle of the estimated linear equation at t, as well as the difference between the slope angle of currently estimated linear equation and the one at the previous iteration (t-1). Step 3. After iterations are terminated, the method calculates the average of slope differences saved at every iteration. Step 4. Finally, the method pops the slope difference of two linear equations in the order of the storage and compares it to the average slope difference, until it is larger than the average slope difference at an optimal iteration (t^0).

Fig. 7. Genuine vertical line components in red solid lines after removing diagonal components

Now, all of the eliminated components under the optimal iteration are considered diagonal components **Fig. 6.** The execution of this process can effectively remove the diagonal components **Fig. 7.**

2.3 Text Boundary Estimation

The LLS is a well-known method to estimate parameters that fit a linear function to a set of points, while minimizing the sum of squared errors with the difference between the actual data points and the function [1]. In our experiments, we estimated a linear equation, $g(x)=\theta=\alpha x+\beta$, of the slope angles of the set of linear equations (L) of potential vertical line components, as shown in **Fig. 4**. Parameters of the linear equation, α and β, are found when the sum of squared error, $E(\alpha,\beta)=\sum(\theta_j-(\alpha x_j+\beta))^2$ at $y=H$ is minimized. We took the partial derivative of E with respect to α and β, and set it equal to 0, to find the two parameters.

If E is arranged for α,

$$\frac{\partial E(\alpha,\beta)}{\partial\alpha}=\frac{\partial}{\partial\alpha}\sum(\theta_j-(\alpha x_j+\beta))^2 \tag{3}$$

$$=\sum 2(\theta_j-(\alpha x_j+\beta))(-x_j)$$

$$=\sum 2\alpha x_j^2-2(\theta_j-\beta)x_j$$

$$=0$$

$$\because \alpha\sum x_j^2+\beta\sum x_j=\sum(x_j\theta_j)$$

If this is arranged for β,

$$\frac{\partial E(\alpha,\beta)}{\alpha\beta}=\frac{\partial}{\partial\beta}\sum(\theta_j-(\alpha x_j+\beta))^2 \tag{4}$$

$$=\sum 2(\theta_j-(\alpha x_j+\beta))(-1)$$

$$=\sum 2\beta-2(\theta_j-\alpha x_j)$$

$$=0$$

$$\because \alpha\sum x_j+n\beta=\sum\theta_j$$

α and β are approximated by solving simultaneous equations in Eq. 3 and Eq. 4.

$$\alpha=\frac{\sum\theta_j\sum x_j-n\sum(x_j\theta_j)}{(\sum x_j)^2-n\sum x_j^2)} \tag{5}$$

$$\beta=\frac{\sum\theta_j(x_j)^2-n\sum x_j\sum(x_j\theta_j)}{n^2\sum(x_j)^2-n\sum(x_j)^2} \tag{6}$$

Some Korean characters, such as '승' and '증', do not include vertical line components and other characters, such as '성' and '지', contain vertical line components only on their right side. If these characters appear in the beginning or end of the text region, it becomes difficult to estimate the outline of the distorted text region, because it is not guaranteed that the vertical line components lie on the left and right boundaries of the text region. Therefore, we should estimate virtual vertical line components at the left and right boundaries.

Especially, left or right boundary lines to cover the entire text region have the minimum or maximum slope angle. Therefore, if a line $\theta = \alpha x + \beta$ passes the pixel on the text with the minimum or maximum slope angle (θ), it becomes the left or the right boundary line of the text region. At this time, let the pixels $P_l(x_l, y_l)$ and $P_r(x_r, y_r)$ be the left and right boundary. However, since $\theta = \alpha x + \beta$ is an equation at $y = H$, α and β have to recalculated at different $y = y_l$ or y_r. Here, we can arrange x using $l_j = y = a_j x + b_j$ as Eq. 7. Thus, slope angle θ for any pixel on the text can be calculated using $g'(x, y) = \theta = \alpha' x + \beta'$, after replacing α and β with Eq. 8 and 9.

$$x = \frac{y - b_j}{a_j} \tag{7}$$

$$\alpha' = \frac{\sum\theta_j \sum(\frac{y-b_j}{a_j}) - n\sum((\frac{y-b_j}{a_j}) * \theta_j)}{(\sum(\frac{y-b_j}{a_j}))^2 - n\sum(\frac{y-b_j}{a_j})^2)} \tag{8}$$

$$\beta' = \frac{\sum\theta_j \sum(\frac{y-b_j}{a_j})^2 - n\sum(\frac{y-b_j}{a_j})\sum((\frac{y-b_j}{a_j}) * \theta_j)}{n^2\sum(\frac{y-b_j}{a_j})^2 - n(\sum(\frac{y-b_j}{a_j}))^2} \tag{9}$$

Now, we are able to calculate the reference point $P_l(x_l, y_l)$ with the minimum slope angle for the left hand boundary of the text region. Similarly, the reference point with the right hand boundary is also calculated at $P_r(x_r, y_r)$ (Eq.(10)). Where I_B is the set of black pixels of I.

$$\begin{aligned} P_l(x_l, y_l) &= \arg\min\ g'(x, y),\ (x, y) \in I_B \\ P_r(x_r, y_r) &= \arg\max\ g(x, y),\ (x, y) \in I_B \end{aligned} \tag{10}$$

The top side line and bottom side line of the text region are acquired directly from the horizontal line of the top and bottom of the text region. Four sides of the boundary can be defined as Eq.(11).

$$\begin{aligned} &top\_side &&: \{y = H\} \\ &bottom\_side &&: \{y = 0\} \\ &left\_side &&: \{y = \tan(g'(x_l, y_l))(x - x_l) + y_l\} \\ &right-side &&: \{y = \tan(g'(x_r, y_r))(x - x_r) + y_r\} \end{aligned} \tag{11}$$

2.4. Distortion Correction

Our final goal is to correct the distortion of the text region. We now have four boundaries (trapezoid outline) that tightly enclose the region, as shown in **Fig. 8(a).** A bilinear transformation [9] is applied to the text region with the trapezoidal outline to correct the distortion.

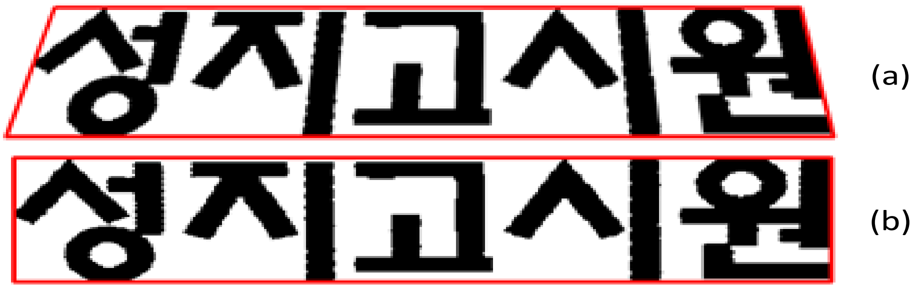

Fig. 8. (a) The trapezoid outline estimated by the proposed method, (b) The results without distortion

A bilinear transformation [21] is applied to the text region with the trapezoidal outline to correct the distortion using Eq.(12). Here (u, v) is the coordinate of a point in the input image and (x, y) is the coordinate of the point in the output image. (See Fig.9.)

$$\begin{aligned} u_{01} &= u_0 + dx(u_1 - u_0) \quad , \quad & u_{32} &= u_3 + dx(u_2 - u_3) \\ v_{03} &= v_0 + dy(v_3 - v_0) \quad , \quad & v_{12} &= v_1 + dy(v_2 - v_1) \\ u &= u_{01} + dy(u_{32} - u_{01}) \quad , \quad & v &= v_{03} + dx(v_{12} - v_{03}) \end{aligned}$$
$$, \text{where } \; dx = \frac{x}{u_2 - u_3}, dy = \frac{y}{v_3 - v_1} \qquad \textbf{(12)}$$

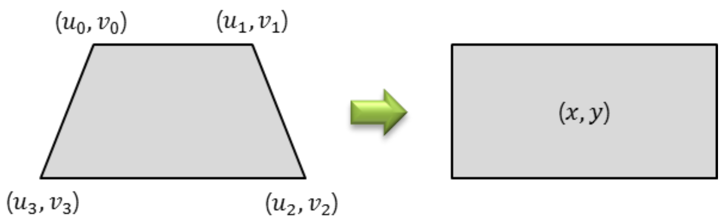

Fig. 9. Example of bilinear transformation

3. Experiment Results

We assume that the signboard images include only Korean characters. Testing datasets are composed of two image sets: artificially distorted text images and real distortion signboard images. The former has been made with 40 text images; each image in the testing dataset consists of 3 to 6 characters generated from the word processing software. The perspective distortion was added to the each image with steadily increasing angle from 5° to 25° by both 5° on the left and right. Therefore, 25 different combinations of the distortion have been generated from each image. Thus, there were 40×25=1000 artificially distorted images. In contrast, the latter consists of 189 real signboard images (consisting of 936 characters) including 90 images used in Park11 and an additional 99 real images collected in this research. Their texts were already binarized using a method presented in Park [5].

We performed two different experiments. First, we applied our method to the first dataset to correct the manually generated distortion and calculate the similarity between the revised images and the original images. The Dice coefficient $= 2(I_o \cap I_t)/(I_o \cup I_t)6$ was used to measure the similarity, where I_o is the original image and I_t the revised image by the proposed method. Second, we applied the proposed method to the real signboard images to correct the distortion and measured the OCR accuracy before and after the distortion correction, using the character recognition method proposed by Park [5]. The recognition rates were measured for both the signboard and character. A signboard is recognized correctly when all the characters in the signboard were recognized correctly. For the first data set of 1000 images, the proposed method shows a similarity from 96.08% to 97.53%, with a standard deviation of 0.37%, despite a variety of distortion angles, from 5° to 25° on both the left and right side **Fig. 10.**

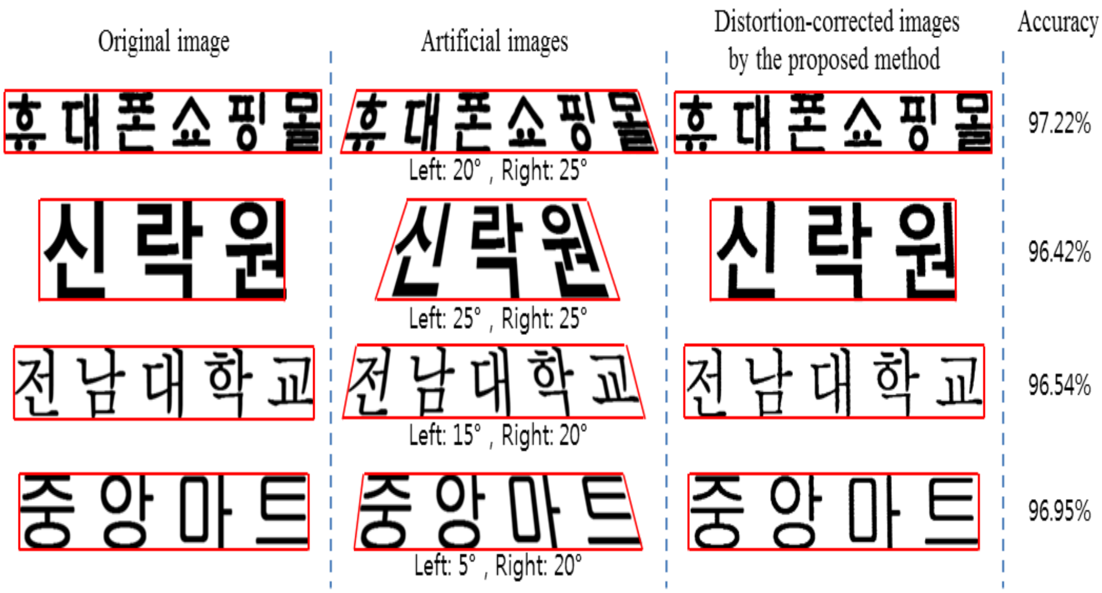

Fig. 10. Example of synthetic dataset. The manually generated original images and the artificially distorted images using various angles are shown in the left and middle column; distortion-corrected images using the proposed correction method are depicted in the right hand column.

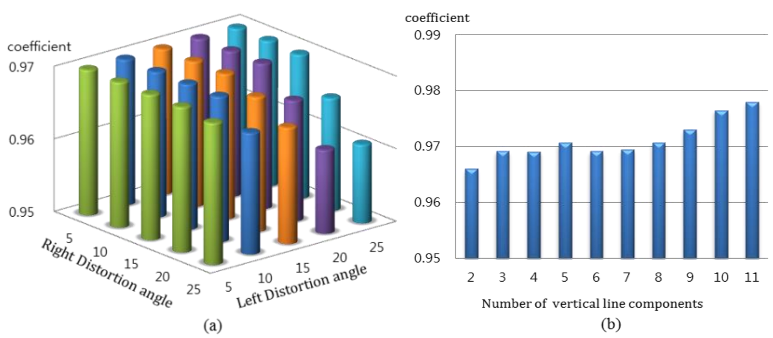

Fig. 11. Similarities between the original images and their distortion corrected images of (a) various distortion angles and (b) of various numbers of vertical line components within the text.

Fig. 11(a) shows the change of similarity as a function of the distortion angle. As can be seen from this graph, the similarity degrades, as the distortion angle increases. The degree of similarity is proportional to the number of vertical line components, as **Fig. 11(b). Fig. 11(b)** shows the relationship between the number of vertical line components and the similarity. The average similarity of the proposed method for all the data set was 97.02%.

Table 1. Comparison of recognition rates before and after distortion correction

Unit	Total No	Before correction	After correction
Signboard	189	10(5.3%)	66 (34.9%)
Character	936	529(56.5%)	734(78.4%)

Table 1 show that the OCR accuracy of the character recognition method proposed by Park [5] for the real signboard images in the second dataset and their distortion corrected images. The recognition rate for the signboard unit was enhanced from 5.26% to 34.73% after correcting the distortion, while that for the character unit was increased from 56.52% to 78.42%.

Fig. 12 shows examples of four signboard images with the estimated outline and distortion-corrected images obtained by the proposed method.

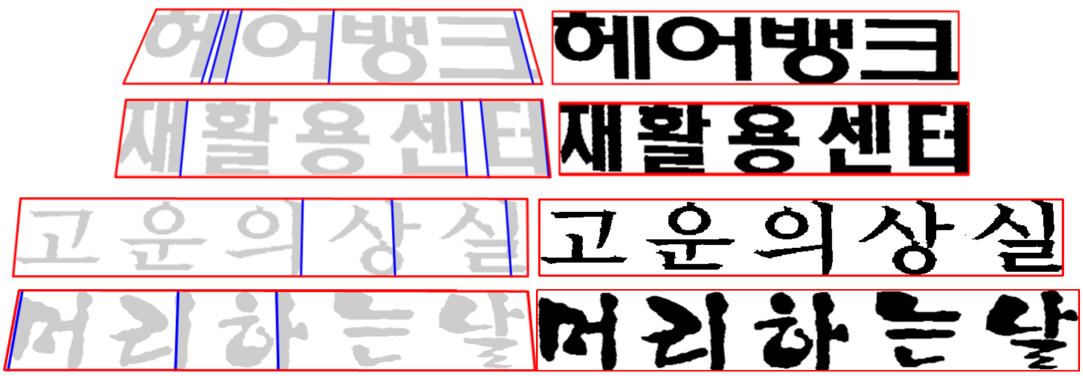

Fig. 12. Examples of the estimated outlines and distortion correction results. Four images in the first column show the estimated outlines on the signboard images along with the vertical components extracted from the text, and the images in the second column show the distortion-corrected images by the proposed method.

4. Conclusion

The early studies assume that the document image is long and composed of a number of text lines. Thus, this assumption makes it hard to apply the methods directly to the Korean signboard images, where only a few characters appear in a single text line. Based on the observation of the Korean signboard images, we determined that the vertical lines of characters in the Korean signboard images are able to describe the amount of distortion. From this observation, we try to find the vertical line components and estimated the trapezoid outline of the image using the slopes of the line components.

This paper suggests a new method to correct the distortion of the text regions in Korean signboard images. The proposed method estimates the trapezoid outline of the text regions by analyzing the vertical line components of the characters in the signboard images. We performed two experiments with two data sets, manually distorted text images and real signboard images collected using smart phones, to evaluate the performance of the proposed method. The similarity between the manually distorted images and their revised images was 97.02% in the Dice coefficient. The recognition rate for the signboard unit was enhanced from 5.26% to 34.73% after correcting the distortion, while that for the character unit increased from 56.52% to 78.42%. These results show that the proposed method effectively corrects the distorted text regions in Korean signboard images to improve their recognition.

Although the results are encouraging, the proposed method has several limitations. First, it focuses only on vertical distortion, because the camera viewpoint was limited at the middle just below the signboards. We have collected a variety of images from different camera viewpoints to diversify the range of distortion. Second, as shown in the experimental results, the ability of the distortion correction depends on the number of the characters in the text, because the greater the number of characters in a line of text, the greater the number of vertical line components we can extract from them. There should be at least two vertical line components to estimate the trapezoidal outline of the text. This paper proposes a new method to correct the distortion of the text region within signboard images shot by smart phones.

References

[1] O. Bretscher, "Linear Algebra With Applications," Prentice Hall, (1997)

[2] P. Clark and M. Mirmehdi, "Recognizing text in real scenes," *International Journal on Document Analysis and Recognition (IJDAR)*, 4(4) (2004), 243–257.
http://dx.doi.org/doi:10.1007/s10032-001-0072-2.

[3] P. Clark and M. Mirmehdi, "On the Recovery of Oriented Documents from Single Images," in *Proc. of 4th IEEE Conf. on Advanced Concepts for Intelligent Vision Systems*, (2001), 190–197.

[4] C. Li, X. Ding, and Y. Wu, "Automatic text location in natural scene images," in *Proc. of International Conference on Document Analysis and Recognition*, (2001), 1069-1073.
http://dx.doi.org/doi:10.1109/ICDAR.2001.953950.

[5] J. H. Park, G. S. Lee, J. S. Lim, S. H. Kim, H. J. Yang, M. H. Lee and S. T. Hwang, "Automatic detection and recognition of Korean text in outdoor signboard images," *Pattern Recognition Letters*, 31 (2010), 1728-1739.
http://dx.doi.org/doi:10.1016/j.patrec.2010.05.024.

[6] P. Pan, Y. Zhu, J. Sun, S. Naoi, Recognizing Characters with Severe Perspective Distortion Using Hash Tables and Perspective Invariants, in *Proc. of 11th International Conference on Document Analysis and Recognition*, 548-552, (2011).
http://dx.doi.org/doi:10.1109/ICDAR.2011.116.

[7] L. Li, C. L. Tan, "Recognizing Planar Symbols with Severe Perspective Deformation," *Pattern Analysis and Machine Intelligence*, 32(4) (2010), 755–762.
http://dx.doi.org/doi:10.1109/TPAMI.2009.196.

[8] G. K. Myers, R. C. Bolles, Q. T. Luong, J. A. Herson and H. B. Aradhye, "Rectification and recognition of text in 3-d scenes," *International Journal on Document Analysis and Recognition (IJDAR)*, 7(2-3) (2004), 147-158.
http://dx.doi.org/doi:10.1007/s10032-004-0133-4.

[9] R. Woods, R. Gonzalez and S. Eddins., "Digital image processing using MATLAB," *Prentice Hall*, (2004)

[10] X. C. Yin, H. W. Hao, J. Sun, S. Naoi, "Robust Vanishing Point Detection for Mobile Cam-Based

Documents," in *Proc. of 11th International Conference on Document Analysis and Recognition*, (2011), 136-140.
http://dx.doi.org/doi:10.1109/ICDAR.2011.36.

[11] Biyi Lin, Yide Wang, Bruno Feuvrie and Qingyun Dai, "A Novel Power Amplifier Linearization Technique based on Pre-distortion Principle and Wiener model," *International Journal of Advanced Science and Technology*, 22, 59 (2010).

[12] S. Lu and C. L. Tan, "Camera text recognition based on perspective invariants," in *Proc. of 6th International Conference on Pattern Recognition*, (2006) 2, 1042–1045; Hong Kong, China.
http://dx.doi.org/doi:10.1109/ICPR.2006.351.

[13] G. K. Myers, R. C. Bolles, Q. T. Luong, J., "Rectification and recognition of text in 3-d scenes," *Int. J. Doc. Anal. Recog.*,7, 147-158, (2005).
http://dx.doi.org/doi:10.1007/s10032-004-0133-4.

[14] C. M. Thillou, B. Gosselin, Natural Scene Text Understanding, "Vision Systems: Segmentation and Pattern Recognition," *I-Tech*, Vienna (2007), pp.307-333.

[15] Y. Watanabe, K. Sono, K. Yokomizo, and Y. Okada, Translation camera on mobile phone, in *Proc. of 2003 International Conference on Multimedia and Expo*, (2003) July 2, 177-180; Baltimore, USA.
http://doi.ieeecomputersociety.org/10.1109/ICME.2003.1221582.

[16] Shijian Lu, Ben M. Chen, Chi Chung Ko," Perspective rectification of document images using fuzzy set and morphological operations", Image Vision Comput. 23(5) (2005),541-553.
http://dx.doi.org/doi:10.1016/j.imavis.2005.01.003 .

JunSik Lim received the B.S, M.S degree in Computer Science from Chonnam National University, Korea. He is currently a Ph.D. student at School of Engineering and Computer Engineering, Chonnam National University, Korea. His research interests are pattern recognition, document image processing, medical image processing.

In Seop Na received his B.S., M.S. and Ph.D. degree in Computer Science from Chonnam National University, Korea in 1997, 1999 and 2008, respectively. Since 2012, he has been a research professor in Department of Computer Science, Chonnam National University, Korea. His research interests are image processing, pattern recognition, character recognition and digital library.

Soo Hyung Kim received his B.S. degree in Computer Engineering from Seoul National University in 1986, and his M.S. and Ph.D degrees in Computer Science from Korea Advanced Institute of Science and Technology in 1988 and 1993, respectively. From 1990 to 1996, he was a senior member of research staff in Multimedia Research Center of Samsung Electronics Co., Korea. Since 1997, he has been a professor in the Department of Computer Science, Chonnam National University, Korea. His research interests are pattern recognition, document image processing, medical image processing, and ubiquitous computing.