SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

July 2016, Vol. 59 072101:1–072101:14
doi: 10.1007/s11432-016-5575-z

A method for automatically translating print books into electronic Braille books

Taekeun PARK[1], Jungil JUNG[2] & Jinsoo CHO[3]*

[1]*Department of Applied Computer Engineering, Dankook University, Gyeonggi-do* 448-701*, Republic of Korea;*
[2]*Department of IT Convergence Engineering, Gachon University, Gyeonggi-do* 461-701*, Republic of Korea;*
[3]*Department of Computer Engineering, Gachon University, Gyeonggi-do* 461-701*, Republic of Korea*

Received November 10, 2015; accepted December 29, 2015; published online June 16, 2016

Abstract In this paper, a method for automatically translating scanned images from print books into electronic Braille books is proposed with the objective of reducing the amount of time and cost required for producing Braille books. The proposed method consists of processes for identifying character and image areas in a scanned image, automatically translating characters and images into Braille and tactile graphics, respectively, and positioning Braille and tactile graphics into an electronic Braille page. Experimental results show that the proposed method drastically reduces the time required to translate a print book into an electronic Braille book. Despite the drastic reduction in translation time, the method proposed in this paper does not compromise the ability to recognize information for the visually impaired compared to manually produced Braille books, demonstrating its feasibility in practical applications. Therefore, the proposed method is expected to significantly reduce the time and cost required for producing Braille books, and provide more reading materials for the visually impaired, making significant contributions to enhancing their knowledge and welfare.

Keywords Braille, tactile graphic, electronic Braille book, automatic translation, the visually impaired

Citation Park T, Jung J, Cho J. A method for automatically translating print books into electronic Braille books. Sci China Inf Sci, 2016, 59(7): 072101, doi: 10.1007/s11432-016-5575-z

1 Introduction

Braille books have long been used as a major medium for providing information and knowledge to the visually impaired. A typical process for authoring a Braille book involves Braille translators and tactile graphics specialists translating a print book. The production process includes translating characters and images into Braille and tactile graphics in digital form, respectively, and printing them out with a Braille printer. Braille characters are small rectangular cells that contain tiny palpable raised dots, and a number of character recognition and other software tools are available for automatically translating Braille characters. Tactile graphics, including tactile graphs, tactile diagrams, and tactile pictures, are images that also use palpable raised dots, and they have played an important role in providing the visually impaired with visual information that cannot be expressed with Braille characters or audio. Because it has not been possible to automatically translate images into tactile graphics, tactile graphics specialists have had to perform manual translation with the aid of image editing software. This translation process

* Corresponding author (email: jscho@gachon.ac.kr)

is very time-consuming and labor-intensive, significantly limiting the availability of tactile graphics in Braille books. The high cost of producing tactile graphics has become a major cause for hindering the expansion of visual contents in Braille books. Consequently, only about 30% of the demand for Braille books is being met [1], undermining opportunities for the visually impaired to receive education and to acquire information and knowledge.

In this paper, a method is proposed for automatically translating print books into electronic Braille books based on techniques for categorizing and analyzing images scanned from print books. The objective of this study is to reduce the time and cost required for producing Braille books. The advantages of the proposed method are as follows: (1) it provides automatic and efficient classification of characters and images by minimizing errors in detecting graphics from images scanned from print books; (2) it determines graphic complexity based on the techniques used by tactile graphics specialists and how the visually impaired recognize information; and (3) it efficiently automates and expedites the Braille book production process so that knowledge and information can be provided in greater quantity and diversity to the visually impaired at a lower cost.

This paper is constructed as follows. Section 2 surveys related work, Section 3 explains the proposed method, Section 4 shows experimental results, and Section 5 concludes the study.

2 Related work

In general, the technology for automatically translating images scanned from a print book into an electronic Braille book involves identifying characters and images, translating characters into Braille, and translating images into tactile graphics. This section discusses previous studies related to each aspect of the technology.

First, let us examine previous studies on identifying character and image areas. Wong [2] and Phan [3] used varying brightness levels in pixel units to identify character and image areas. However, the method does not consistently yield high performance for various types of color images because it sets a fixed threshold value for the character extraction process, and image areas with substantial variations in brightness, such as graphs and diagrams, are incorrectly extracted as character areas. Jang [4] used linking elements and a median filter to identify characters and non-characters. An area expansion filter and several constraints were also used to modify incorrectly identified elements. However, as with using brightness variation, the method uses fixed threshold values and is not quite effective for extracting character areas from an image. Other methods for identifying characters and images use various techniques, including the neural network for learning [5–9], edges [10–12], color clustering [13–15], and a combination of texture and linking elements [16–19]. All of these methods contain problems, such as cumbersome pre-processing for learning, setting threshold values for separating character pixels from graphic pixels, a wide range of color variation at the boundary, and a substantial amount of calculation required for character extraction.

Second, the most commonly used current technology for translating characters into Braille is based on optical character recognition (OCR), which has been widely studied and is actively used as the core technology for delivering text as speech in screen reader systems for the visually impaired. This paper does not contain any additional research on OCR and uses the open source engine Tesseract OCR [20] to recognize character. However, because both Korean and English characters are automatically translated into Braille in our study, we have developed a Korean Braille translation module in accordance with the Braille translation rule [21], in addition to the English Braille translation module.

Finally, there have been several automatic and semi-automatic methods for translating images into tactile graphics. Way [22] first presented a basic concept for automatically translating natural images such as photographs into a tactile form. Rotard [23] proposed a method for analyzing simple graphics such as graphs and shapes, displayed on a web browser, and semi-automatically converting them into tactile graphics based on predetermined threshold values for binarization. Ladner [24] proposed a semi-automatic image simplification method for translating graphical images (i.e., charts, graphs, and diagrams) into tactile graphics. Petit [25] also proposed a semi-automatic method for translating simplified scientific

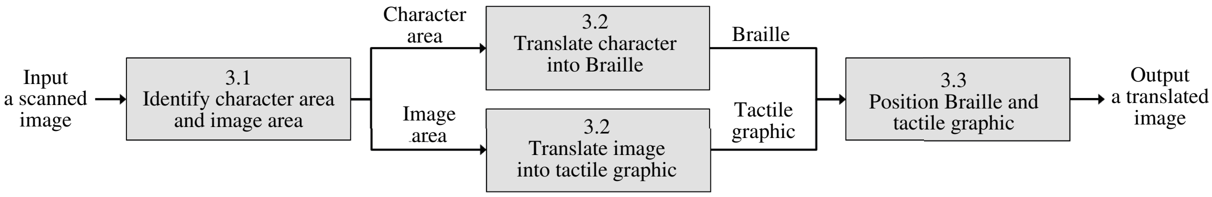


Figure 1 Overall process of the proposed method for automatically translating a print book into an electronic Braille book.

graphical images into tactile graphics. Takagi [26, 27] proposed methods for extracting and classifying solid and broken lines contained in mathematical graphs and automatically translating them into tactile graphics. Chen [28] proposed an automated system for creating tactile maps from simple hand-drawn maps. However, because these methods are only applicable to particular types of images (e.g., graphical images or photographs), they pose significant limitations for translating general print books that contain various types of images.

To solve the problems in previous studies and efficiently automate the electronic Braille book production process, this paper proposes a new method for automatically translating print books into electronic Braille books.

3 Automatic translation of print book into electronic Braille book

In this section, a method is proposed for automatically translating a print book into an electronic Braille book based on in-depth analysis of the processes employed by professional Braille translators and tactile graphics specialists. As shown in Figure 1, the overall process of the proposed method consists of the following processes: the identification of character and image areas in a scanned image, the automatic translation of characters and images into Braille and tactile graphics, and the positioning of Braille and tactile graphics.

3.1 Identification of character and image areas in scanned image

In order to translate scanned images from a print book into an electronic Braille book, characters and graphics in a scanned image must be first identified to determine which specific areas should be translated into Braille or tactile graphics. The objective of this subsection is to resolve the problems found in conventional methods [2, 3] by extracting the characteristics of variations in brightness while adapting to the character size in images. The process of identifying characters and images in a scanned image consists of the following steps: the labeling and filtering of a scanned image, the determination of candidate character areas, and the determination of character and graphic labels.

3.1.1 *Labeling and filtering of scanned image*

In order for characters and images to be effectively identified, regional data of the elements contained in a scanned image are collected through grayscaling, binarization, labeling, and filtering. Grayscaling and binarization are performed in order to isolate elements from the background. For labeling, a linked pixel is defined as a single element (a label hereunder) and data such as the width and the height of the element are stored to be used as basis data for determining character and image areas. To increase the performance of identifying character and image areas, image areas that display drastic brightness variations similar to character areas, such as graphs, tables, and boundary boxes, are detected and categorized in advance with label's black pixel density. Then, the preliminary filtering process is performed to identify character and image areas based on label's black pixel density as follows:

$$\mathrm{label}_i = \begin{cases} \text{image}, & \text{if } (\mathrm{bp}/\mathrm{ap} < T_d), \\ \text{character}, & \text{otherwise}, \end{cases} \tag{1}$$

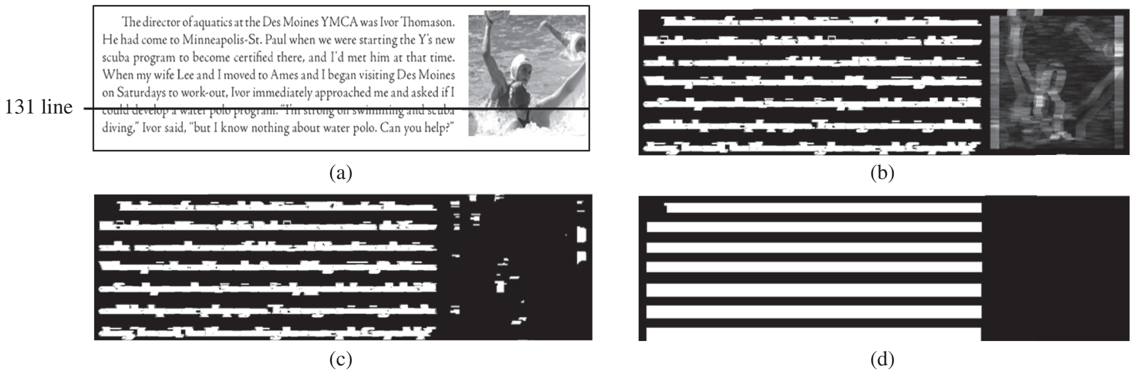


Figure 2 Example of extracting a character area from a scanned image. (a) Scanned image; (b) MGD image; (c) binary MGD image; (d) MGD image after noise removal and morphology.

where label and i denote a label and a label index, respectively, and ap and bp indicate label's total number of pixels and the number of black pixels in the label area, respectively. T_d represents the threshold value (0.15) for determining characters and images, which is derived from the experimental results of comparing the black pixel densities of various character area samples with those of various graph, table, and boundary box samples that can be incorrectly identified as character areas.

3.1.2 *Determination of character area candidates*

Character area candidates in a scanned image are determined based on the maximum gradient difference (MGD) image [3], which is effective for extracting continuously positioned character areas.

First, a simple mask is applied to image's scan lines, as shown in Figure 2(a), to calculate the amount of brightness variation:

$$G(x,y) = f(x,y) - f(x-1,y),\ (1 \leqslant x \leqslant w,\ 1 \leqslant y \leqslant h), \tag{2}$$

where $G(x,y)$ denotes the change in brightness from the adjacent pixel, $f(x,y)$ represents pixel's brightness value at position (x,y), and w and h are the width and the height of the scanned image, respectively. In general, there is a significant variation in brightness between character and background pixels, and a less variation between pixels of identical types (e.g., between images).

Second, an MGD image of the scanned image, as shown in Figure 2(b), is created as follows:

$$\mathrm{MGD}(x,y) = \max(G(x_i,y)) - \min(G(x_i,y)),\ \left(1 \leqslant y < h,\ \frac{n}{2} \leqslant x < w - \frac{n}{2},\ -\frac{n}{2} \leqslant i < \frac{n}{2}\right), \tag{3}$$

where $\mathrm{MGD}(x,y)$ is the MGD within the calculation range, and i denotes the position in MGD in n. In particular, n, which defines the calculation range of $\mathrm{MGD}(x,y)$, indicates the character width and is adaptively determined as follows, to resolve the problems in conventional methods [2, 3] for which the character width has to be determined manually:

$$n = \frac{1}{\mathrm{NLT}} \sum_{k=1}^{\mathrm{NLT}} L_k W,\ \text{where NLT} = \text{count}\left(\left(\frac{1}{\mathrm{NL}} \sum_{j=1}^{\mathrm{NL}} L_j W\right) < L_j W\right), \tag{4}$$

where NL denotes the total number of labels, and L_jW and j represent the horizontal length and the index of the corresponding label, respectively. count indicates the number of L_jWs that satisfy the condition within the parentheses, and NLT is the number of labels greater than the average horizontal length of the label. L_kW and k denote the horizontal length and the index of the label, respectively, that satisfies the count condition.

Third, in the MGD image created above, there are high and low MGDs for character and image area candidates, respectively. Hence, binarization is performed with the Otsu algorithm [29] to highlight the characteristics of the characters and to minimize the number of image area candidates, as shown in Figure 2(c).

Finally, noise is removed from the MGD image to determine character area candidates. At this stage, the MGD image mostly consists of MGDs for character area candidates. However, because there can be residual MGDs due to the noise generated during image scanning or due to the drastic variations in brightness in the image area candidates, the following condition is applied to the MGD image's scan line to eliminate such MGDs.

$$\mathrm{MGD}(x,y) = \begin{cases} \mathrm{MGD}(x,y), & \mathrm{length}(\mathrm{MGD}(x,y)) < 2n, \\ 0, & \text{otherwise}, \end{cases} \tag{5}$$

where n indicates the character width in (4), length is the length of an MGD that is not 0. In (5), noise is removed based on the fact that a character does not generally exist on its own but within a string (two or more characters). After scattered MGDs are combined with the morphology operation based on the $n \times n$ square filter, the final character area candidates can be determined with the starting and end points of each of the combined MGD areas, as shown in Figure 2(d).

3.1.3 *Determination of character and graphic labels*

Labels and character area candidates are compared to determine if the labels that belong to the character area candidates are character labels and those that do not are graphic labels. Algorithm 1 displays the pseudo code for this process, where *label* and i indicate label set and index, respectively. Each label has attributes for the region (start_P, end_P, $P = x$ or y) and the type. Labels for characters and images can be identified by the processes described in Subsections 3.1.1–3.1.3.

Algorithm 1 Pseudo code for determining character and graphic labels.

```
function dectectLabelType(label, i)
for y = label[i].start_y; y < label[i].end_y; y++
    for x = label[i].start_x; x < label[i].end_x; x++
        if label[i].position is in text block then
            label[i].type = character;
        else
            label[i].type = graphic;
        end if
    end for
end for
end function
```

3.2 Automatic translation of characters and images into Braille and tactile graphics

Character areas with character labels and image areas with image labels identified in Subsection 3.1 are automatically translated into Braille and tactile graphics, respectively.

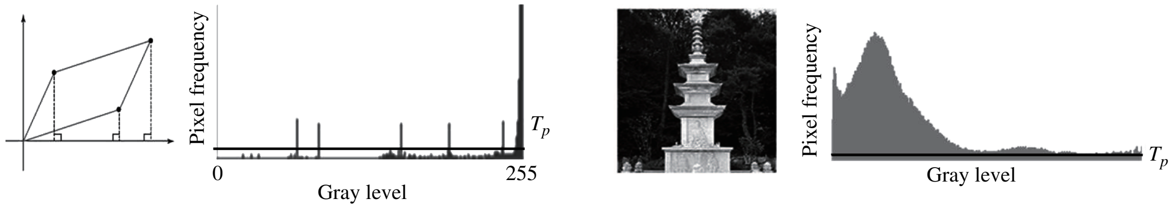
3.2.1 *Translation of character into Braille*

As mentioned earlier, character areas are translated into Braille with the open source Tesseract OCR engine [20] and an English Braille translation module. We have also developed a Korean Braille translation module in accordance with the Braille translation rule [21].

3.2.2 *Translation of images into tactile graphics*

The process of automatically translating image into tactile graphic consists of the following three steps: image complexity classification, tactile graphic translation of low- or high-complexity image, and image simplification.

Step 1. Image complexity classification. In general, tactile graphics specialists classify graphical images such as graphs and diagrams with clear and simple boundaries as low-complexity images, which



(a) (b)

Figure 3 Characteristics of low- and high-complexity image. (a) Example of low-complexity image and its gray level histogram; (b) example of high-complexity image and its gray level histogram.

can be easily translated into tactile graphics. On the other hand, images with continuous and diverse distributions of color and brightness are classified as high-complexity images, which require complex procedures and a substantial amount of time for translation. Accordingly, a method for classifying image complexity based on the characteristics of input image's gray level distribution is proposed as the first step in automatic tactile graphics translation.

Figure 3(a) shows a graphical image classified by a tactile graphics specialist as a low-complexity image with a gray level distribution discretely concentrated around a few levels. In comparison, the image shown in Figure 3(b) is classified as a high-complexity image with a complex object and background that results in a gray level distributed continuously at various levels. However, a low-complexity image scanned from a textbook usually shows a continuous distribution of gray levels with low pixel frequencies close to 0, caused by the noise created during the scanning process. This can obscure the criteria for distinguishing between low- and high-complexity images. Therefore, values below the pixel frequency threshold (TP) are first eliminated from the original histogram to clarify the criteria for classifying images based on complexity as follows.

$$H_{\text{modified}} = \begin{cases} H - T_p, & H - T_p > 0, \\ 0, & H - T_p \leqslant 0, \end{cases} \tag{6}$$

where H and H_{modified} represent the original histogram and the modified histogram, respectively. T_P is defined as follows:

$$T_P = \frac{1}{\text{count}_T(H(i))} \sum_{i=0}^{255} H(i), \text{ where } H(i) \geqslant Q_1 - 1.5(Q_3 - Q_1) \text{ and } H(i) \leqslant Q_3 - 1.5(Q_3 - Q_1). \tag{7}$$

$H(i)$ is the pixel frequency of gray level i in the original histogram, Q_1 and Q_3 are the 1st quartile (25th percentile) and the 3rd quartile (75th percentile), respectively, when the pixel frequency values are sorted in the ascending order, and $\text{count}_T()$ is the number of gray levels that are not regarded as outliers. R_H, the variation rate between H and H_{modified}, is calculated as follows.

$$R_H = \frac{\text{count}_R(H) - \text{count}_R(H_{\text{modified}})}{\text{count}_R(H)} \times 100, \tag{8}$$

where count_R indicates the number of gray levels with a non-zero value. We can now define image complexity as follows.

$$\text{Complexity} = \begin{cases} \text{High}, & R_H < T_C, \\ \text{Low}, & \text{otherwise}, \end{cases} \tag{9}$$

where T_C is an experimental threshold value (= 60) determined by five tactile graphics specialists for classifying the image complexities of 100 test images selected from high school science and mathematics textbooks.

Step 2. Tactile graphics translation of low- and high-complexity images. Tactile graphics specialists typically translate a low-complexity image into a tactile graphic by simply drawing a binary image

Table 1 Survey results on tactile graphics recognition characteristics of the visually impaired (15 blind, 10 with low vision)

Item	Recognition characteristics
Image details	Expressing excessive detail can cause confusion in determining the direction and intersection of image outlines. Therefore, the outline in both low-and high-complexity images should be expressed as simply as possible to increase the information recognition capability.
High complexity image with a primary object	For an image containing a primary object, the background and surrounding data should be removed and only the outline of the primary object should be provided to increase recognition capability.
High complexity image without a primary object	For an image without a primary object, such as a landscape, translating the outline does not usually enable the visually impaired to recognize the essential information.

based on the boundary information using image-editing software. We employ a similar approach for low-complexity images by creating binary images simply based on a binarization process [29]. However, according to our survey on the information recognition capability of the visually impaired for tactile graphics (Table 1), lines drawn around a tactile graphic must be simple; otherwise, the information might not be recognized. Accordingly, the binary image of a low-complexity graphic is simplified, as will be described in Step 3, before creating the final tactile graphic.

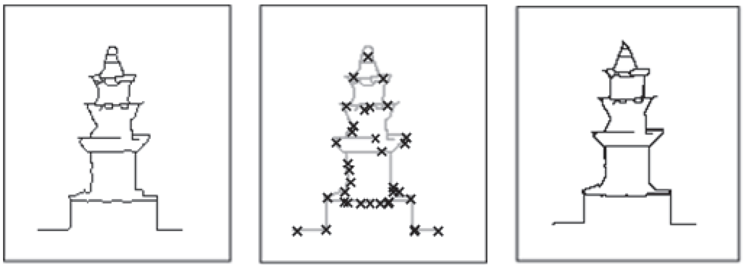
The tactile graphics translation process of a high-complexity image is performed based on primary (central) object detection. As shown in Table 1, the information recognition characteristics of the visually impaired for the tactile graphics of high-complexity images depend on the existence of a primary object and the level of detail in tactile graphics. In general, the information recognition capability of the visually impaired rapidly declines in tactile graphics that contain a mixture of various objects, those without a specific central object (e.g., a landscape picture), and those with a complex expression of a specific object. Accordingly, tactile graphics translation of a high-complexity image is performed by determining a central object and converting it into a binary image according to the following steps.

First, because most tactile graphics specialists have normal vision, they often use color differences to discriminate between objects and background, and mainly translate a primary object usually located in the center of an image into a tactile graphic. Afterwards, color segmentation and labeling processes are performed to extract a central object in a high-complexity image. Color segmentation involves simplifying the color complexity of the image through quantization using peer group filtering (PGF) [30], and the image is segmented into similar colors by measuring color similarities based on the national bureau standards (NBS) distance [31]. By performing labeling for each segmented color and segmenting the image into color elements, all of the elements can be extracted from the image. The image is segmented based on the NBS distance because most images are created by people with normal vision and it would be appropriate to analyze images from their viewpoint.

Second, the central object area is located based on the general principles of photography and the characteristics of triangular composition, and the label with the largest size within the area is defined as a central label. However, even when the position of the central label is within the central-object area, central object may not exist if the horizontal and vertical distributions of the pixels within the central label are significantly narrow or if the distributions are horizontally extensive, as in the case of a background. Therefore, in order to accurately determine whether a central object exists, pixel distribution feature values within the central label that quantitatively reflect the decision criteria of tactile graphics specialists are defined as follows.

$$\begin{cases} V_x(\mathrm{SP}) = \sum((P_x - m_x)^2/w)/(w \times h), \\ V_y(\mathrm{SP}) = \sum((P_y - m_y)^2/h)/(w \times h), \\ V(\mathrm{SP}) = \max(V_x(\mathrm{SP}),\ V_y(\mathrm{SP})), \end{cases} \tag{10}$$

where P_x and P_y indicate pixel positions inside the central label, m_x and m_y are mean values of P_x and P_y, respectively, and w and h denote the width and the height of the central label, respectively. $V_x(\mathrm{SP})$ and $V_y(\mathrm{SP})$ indicate the horizontal and vertical variances, respectively, for pixel positions distributed within



(a) (b) (c)

Figure 4 Example of the image simplification process. (a) Extracted central object contour; (b) extraction of corner points (indicated by $\times$); (c) result of the contour simplification process.

the central label, and $V(\mathrm{SP})$ represents the maximum variance value between $V_x(\mathrm{SP})$ and $V_y(\mathrm{SP})$. In order to compile tactile graphics specialists' decision-making data, decision outcomes of 5 tactile graphics specialists on 100 high-complexity images selected from high school science and mathematics textbooks are analyzed to define the existence of a central object as follows.

$$\text{Central object} = \begin{cases} \text{None}, & \text{if } ((V(\mathrm{SP}) > T_x \text{ and } V(\mathrm{SP}) \equiv V_x(\mathrm{SP})) \text{ or } (V(\mathrm{SP}) < T_{\max})), \\ \text{Exists}, & \text{otherwise.} \end{cases} \tag{11}$$

T_x and $T_{\max}$ represent the horizontal variance threshold and the maximum variance threshold, respectively, obtained from the experiments conducted with tactile graphics specialists for determining the existence of a central object. T_x and $T_{\max}$ have values of 0.06 and 0.01, respectively.

Finally, binarization of the central object in the central label is performed to produce the binary image.

Step 3. Image simplification. The final step in tactile graphics translation for both low- and high-complexity images is image simplification. Taking into account the difficulties faced by the visually impaired in recognizing thick contours with inconsistent directionality, the binary image created in Step 2 is simplified to increase the recognition rate. First, dilation and segmentation are performed on the contour to create a simplified contour. Next, corner points are extracted using the Shi-Tomasi algorithm [32], as shown in Figure 4(b), and the corner points are removed to simplify the image by the following equation, as shown in Figure 4(c).

$$\mathrm{CP} = \begin{cases} \text{Remove}, & \text{if } (C_{\text{curvature}} > T_{\text{curvature}}), \\ \text{Don't Remove}, & \text{otherwise.} \end{cases} \tag{12}$$

CP represents a corner point, and $C_{\text{curvature}}$ denotes the curvature between two lines connected to the corner point. $T_{\text{curvature}}$ is the threshold value for removing the corner point (=30 degree) obtained by analyzing curvatures from a recognition experiment conducted with 10 visually impaired individuals and tactile graphics printed on embossed paper.

3.3 Positioning of Braille and tactile graphic

Because documents are read from the top left-hand corner toward the right one line at a time, the label positions of the Braille and the tactile graphics created in Subsection 3.2 need to be sorted in the ascending order in x and y directions to position each element in the Braille book. In order to increase legibility, tactile graphics are given priority over Braille and positioned as the first element in the ascend-sorted order so that they can always be at the head of a page or a paragraph.

4 Experimental results

To evaluate the performance of the proposed method, a C/C++ program was developed and experiment was conducted with a PC running on a 3.2 GHz CPU. In the experiment, key functions of the proposed

Table 2 Experimental subjects for identifying character and image areas

Document type	Composition	Total number of characters	Total number of images
Article	Korean article (3 pages) English article (3 pages)	9369	395
Textbook	Mathematics book (9 pages) Science book (9 pages)	7716	393

Table 3 Experimental results of identifying character and image areas

Applied method	Document type	Identification accuracy	
		Characters (%)	Images (%)
Wong's method [2]	Article	94.44	40.25
	Textbook	96.07	58.78
	Average	95.26	49.51
Phan's method [4]	Article	96.06	23.04
	Textbook	79.82	50.38
	Average	87.94	36.71
Jang's method [5]	Article	96.99	51.39
	Textbook	96.07	69.95
	Average	96.53	60.67
Proposed method	Article	97.64	80.00
	Textbook	99.26	96.95
	Average	98.45	88.48

method were evaluated in terms of identification of character and image areas, classification of image complexity, and detection of a central object. To assess the efficiency of the developed automatic translation program, the time it takes for non-professionals to use the program was comparatively analyzed with the time it takes for Braille translators and tactile graphics specialist to translate a print book into a Braille book using the conventional manual method. Furthermore, a Braille book created with the proposed method was printed with a Braille printer, and the recognition rate was evaluated among the visually impaired to assess its practical feasibility.

4.1 Evaluation of character and image area identification

In this experiment, accuracies for identifying character and image areas in scanned images were evaluated and compared with the results obtained with other conventional methods (Table 2). Accuracy for identifying character areas is defined as the number of correctly identified and extracted characters among the total number of characters in a scanned image. Accuracy of identifying image areas is defined as the number of correctly identified and extracted images among the total number of images in a scanned image. As shown in Table 3, the average character area identification accuracy was 98.45%, higher than those with conventional methods by up to 10.51%. The average image area identification accuracy was 88.48%, substantially higher than those of conventional methods by 27.81%–51.77%. These results demonstrate that, the proposed method can substantially reduce the problem of incorrectly identifying an image area adjacent to a character area as a character area, which is frequently found in conventional methods. Figure 5 displays step-by-step results of identifying and separating character and image areas from scanned images containing a graph, an illustration, and a photograph using the proposed method. In Column (f), blue boxes indicate character areas, red boxes image areas, and green boxes character areas detected within image areas. Green box areas are removed prior to tactile graphics translation because they can cause confusion for the visually impaired trying to recognize information.

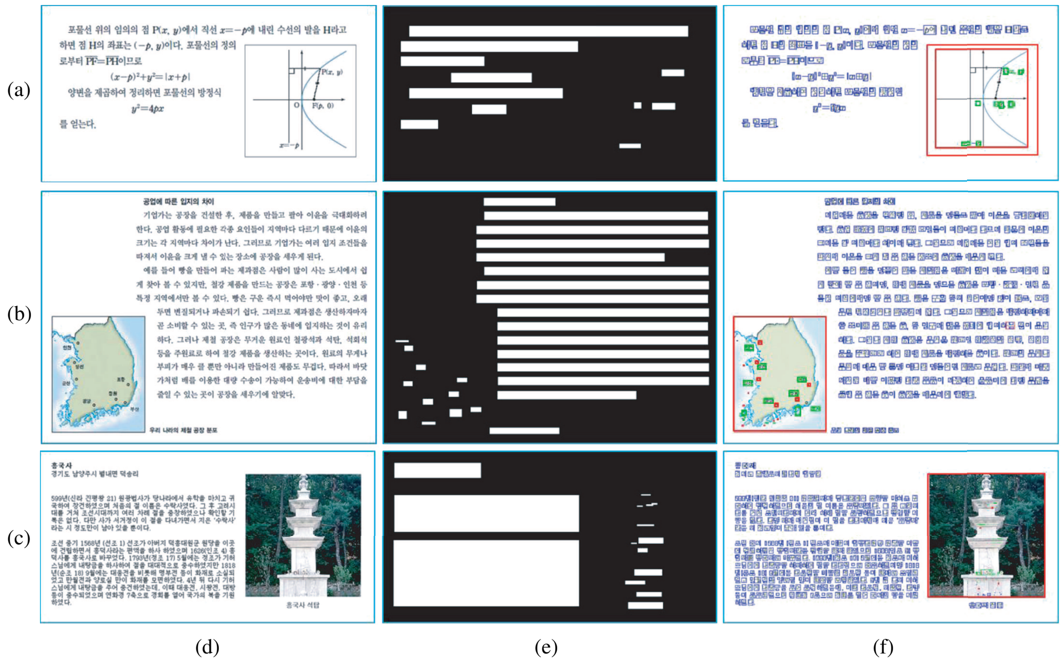


Figure 5 Results of identifying and separating character and image areas from scanned images using the proposed method. Row (a): an image with a graph (a low-complexity image); Row (b): an image with an illustration (a high-complexity image); Row (c): an image with a photograph (a high-complexity image); Column (d): scanned input image; Column (e): results of identifying character areas; Column (f): results of separating image and character areas (blue boxes indicate character areas, red boxes image areas, and green boxes character areas detected within image areas).

Table 4 Experimental results of classifying image complexity

Image type	Complexity	Number of images	Number of correct classification	Accuracy (%)
Graph	Low	50	47	94.00
Shape	Low	50	45	90.00
Illustration	High	50	37	74.00
Photograph	High	50	48	96.00
Total	-	200	177	88.50

Table 5 Experimental results of detecting central object in high-complexity image

Image type	Number of images with a central object	Number of correct detection	Accuracy (%)
Illustration	25	21	84.00
Photograph	25	18	72.00
Total	50	39	78.00

4.2 Evaluation of image complexity classification and central object detection

In this experiment, 200 images (50 graphs, 50 shapes, 50 illustrations, and 50 photographs) were used to evaluate the performances of classifying image complexities and detecting central objects in high-complexity images. Based on the opinions of five tactile graphics specialists, we selected 50 high-complexity images (25 illustrations and 25 photographs) with and without a central object. Table 4 shows the experimental results of classifying image complexity. The accuracies for graphs, shapes, and photographs were high at 94%, 90%, and 96%, respectively, whereas the accuracy for illustrations was relatively lower at 74%. This is because some illustrations consisted of simple gray level values, which increased the histogram variation. Therefore, we believe that in addition to colors, geometric properties of image outlines should be taken into account in order to accurately classify low-complexity images such as graphs and shapes, as well as high-complexity images such as illustrations and photographs. Table 5

Table 6 Comparisons of translation times by the manual method and the proposed method

Image type	Average translation time by manual method (s)			Average translation time by the proposed method (s)		
	Overall	Characters	Images	Overall	Characters	Images
5 low-complexity images with shapes	117.58	28.83	88.75	9.98	9.34	0.64
5 low-complexity images with graphs	152.78	24.93	127.85	10.51	9.89	0.62
5 high-complexity images with illustrations	255.94	31.71	224.23	17.61	11.13	6.48
5 high-complexity images with photographs	262.70	37.05	225.65	28.55	8.65	19.90

shows the experimental results of detecting central objects in high-complexity images. Central objects were correctly detected in illustrations and photographs with accuracies of 84% and 72%, respectively. By automatically extracting central objects from high-complexity images with an accuracy of around 80%, the proposed method can significantly expedite the tactile graphics translation process for various types of images at a lower cost. Until now, Braille books have not contained a wide range of images due to the production process based on manual work, which requires significant amounts of time and labor.

4.3 Evaluation of translation process

To evaluate the efficiency of the proposed translation process, the time required for Braille translators and tactile graphics specialists to manually translate scanned images from print books into electronic Braille pages using image-editing software was compared with the time required for non-professionals to automatically translate the same scanned images into electronic Braille pages using the software developed in this study. Twenty images scanned from a print book consisting of 5 shapes, 5 graphs, 5 illustrations, and 5 photographs were used for the experiment. The translation time was measured from the point where Braille translators and tactile graphics specialists became familiar with the experimental materials to the completion of translation into an electronic Braille page. The time required for translating character areas and image areas were measured separately. As shown in Table 6, compared to the manual method, the proposed method significantly reduced the average translation times for shapes, graphs, illustrations, and photographs by 91.5%, 93.1%, 93.1%, and 89.1%, respectively. These results demonstrate that the proposed method can significantly save the time and labor required for authoring electronic Braille books.

Figure 6 displays tactile graphics translation of low- and high-complexity images. Whereas translation of a low-complexity image only requires binarization, translation of a high-complexity image requires quantization, central object detection, and simplification. Simplification was performed in two iterations so that dense outlines could be minimized to deliver optimal information to the visually impaired, as shown in Columns (f) and (g). Figure 7 displays the final electronic Braille pages created by positioning and combining translated Braille and tactile graphics. Some characters were incorrectly translated into Braille due to the accuracy variation of OCR for the character area, and we intend to improve on this by continuing the research on Braille translation of various languages, equations, and symbols.

4.4 Evaluation of information recognition by the visually impaired

To assess the practical feasibility of the proposed method, electronic Braille pages created with the manual method and the proposed method were printed out with a Braille printer and given to 10 visually impaired individuals (students in a special high school for the blind). The pages were printed on 20 pairs of embossed Braille paper (5 pairs containing shapes, 5 pairs containing graphs, 5 pairs containing illustrations, and 5 pairs containing photographs). Each pair consisted of two electronic Braille pages of an identical image: one created with the manual method and the other with the proposed method. In this experiment, however, only the information recognition of tactile graphics was evaluated to prevent the results from being affected by the functionality of the conventional OCR used to translate characters.

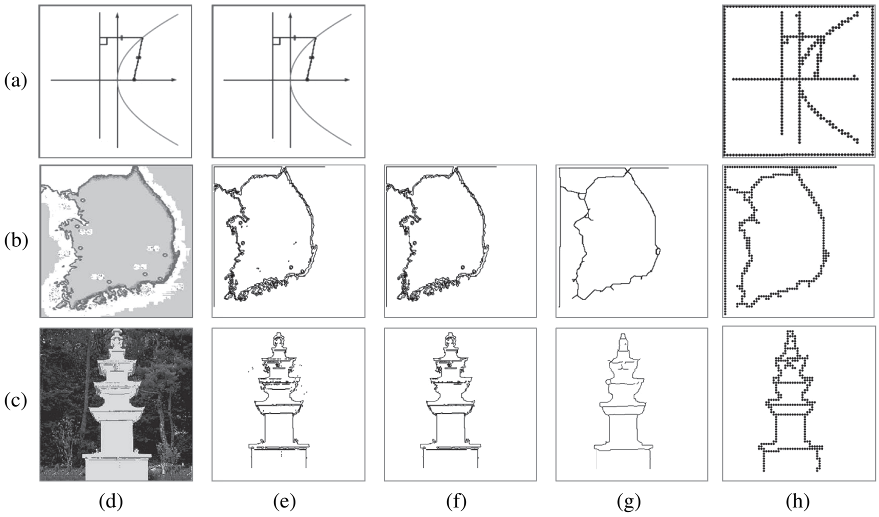


Figure 6 Tactile graphics translation of image areas. Row (a): translation of a graph (a low-complexity image); Row (b): translation of an illustration (a high-complexity image); Row (c): translation of a photograph (a high-complexity image); Column (d): results of quantization; Column (e): results of central object detection; Column (f): results of first simplification iteration; Column (g): results of second simplification iteration; Column (h): final tactile graphics.

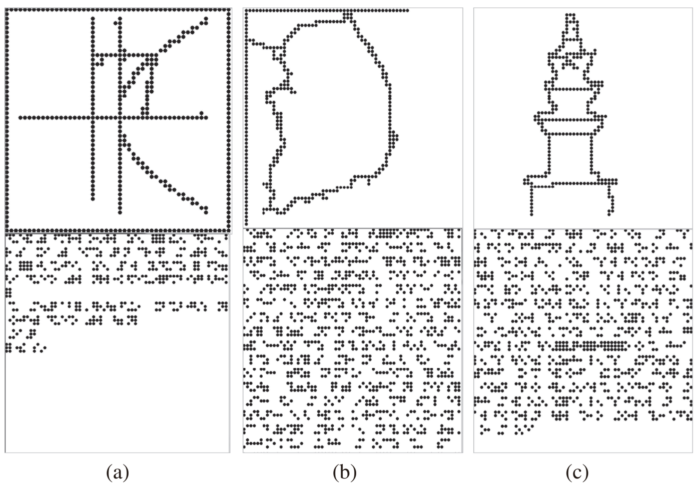


Figure 7 Final electronic Braille pages created after positioning and combining translated Braille and tactile graphics. (a) Electronic Braille page from an image containing a graph; (b) electronic Braille page from an image containing an illustration; (c) electronic Braille page from an image containing a photograph.

Table 7 compares the degrees of information recognition (i.e., recognition points and recognition rates) by the visually impaired. First, the recognition rate was calculated based on how many visually impaired individuals correctly recognized tactile graphics. Second, each visually impaired individual was asked to score tactile graphics on a scale from 0 (unrecognizable) to 5 (completely recognizable). As shown in Table 7, the differences in recognition rates between the tactile graphics created with the proposed method and those with the manual method for graphs, shapes, and illustrations were not significant at 4%, 4%, and 2%, respectively. In comparison, the differences in recognition points for all types of tactile graphics were very similar. These results suggest that the proposed method drastically reduces the time required for creating tactile graphics, while providing almost the same recognition rate for the visually impaired. However, the recognition rate of the tactile graphics created by the proposed method was 12% lower than those created manually for high-complexity images such as photographs. This was mainly caused by the relatively low accuracy for detecting central objects in images, as shown by the experimental results in Subsection 4.2. Therefore, we believe further research is necessary for recognizing central objects to enhance the accuracy of automatically translating complex images into tactile graphics.

Table 7 Comparison of information recognition for tactile graphics created with the manual method and the proposed method

Translation method	Tactile graphic type	Average recognition point	Recognition rate (%)
Manual method	Graph	4.5	90
	Shape	4.4	88
	Illustration	3.6	72
	Picture	2.7	54
Proposed method	Graph	4.3	86
	Shape	4.2	84
	Illustration	3.5	70
	Picture	2.1	42

5 Conclusion

In this paper, a method for automatically translating print books into electronic Braille books was proposed to automate the production of Braille books, which has been performed inefficiently through manual work of Braille translators and tactile graphics specialists. The proposed method consists of the processes for identifying character and image areas, automatically translating characters and images into Braille and tactile graphics, respectively, and positioning and combining translated Braille and tactile graphics into an electronic Braille page. Experimental results show that the proposed method drastically reduces the time required for translating print books into Braille books, while maintaining the information recognition rate for the visually impaired. By facilitating and expediting the processes for authoring electronic Braille books at a lower cost, the proposed method is expected to provide Braille books in greater quantity and diversity to the visually impaired, making significant contributions to enhancing their knowledge acquisition and the level of welfare. In order to provide more accurate information to the visually impaired, we intend to conduct further research on technologies for recognizing complex objects in a diverse range of images.

Acknowledgements The present research was conducted by the research fund of Dankook University in 2013.

Conflict of interest The authors declare that they have no conflict of interest.

References

1 Improvement suggestion for ensuring accessibility to knowledge and information of disabled people. ACRC Technical Report. Korea, 2014
2 Wong E K, Chen M. A new robust algorithm for video text extraction. Patt Recogn, 2003, 36: 1397–1406
3 Shivakumara P, Phan T Q, Tan C L. A gradient difference based technique for video text detection. In: Proceedings of the 10th International Conference on Document Analysis and Recognition, Barcelona, 2009. 156–160
4 Jang D G, Hwang C S. Document image layout analysis using image filters and constrained conditions. KIPS Trans B, 2002, 9: 311–318
5 Chun B T, Bae Y I, Kim T Y. Text extraction in videos using topographical feature of characters. In: Proceedings of the IEEE International Conference on Fuzzy System, Seoul, 1999. 1126–1130
6 Fu X L, Cai L H, Liu Y, et al. A computational cognition model of perception, memory, and judgment. Sci China Inf Sci, 2014, 57: 032114
7 Liu Y, Fu Q, Liu Y, et al. A distributed computational cognitive model for object recognition. Sci China Inf Sci, 2013, 56: 092101
8 Strouthopoulos C, Papamarkos N, Atsalakis A E. Text extraction in complex color document. Patt Recogn, 2002, 35: 1743–1758
9 Chun B T, Song C Y. A method for character segmentation using frequency characteristics and back propagation neural network. J Korea Soc Comput Inf, 2006, 4: 55–60
10 Yuan Q, Tan C L. Text extraction from gray scale document images using edge information. In: Proceedings of the 6th International Conference on Document Analysis and Recognition, Seattle, 2001. 302–306
11 Park J C. Text region detection of various slope and size of text using the adaptive character-edge map. In: Proceedings of KOCON Conference, 2007. 5–9

12 Grover S, Arora K, Mitra S K. Text extraction from document images using edge information. In: Proceedings of the IEEE India Council Conference, Gujarat, 2009. 1–4
13 Kim E J. Character area extraction and the character segmentation on the color document. J Korean Inst Intell Syst, 1999, 9: 444–450
14 Song Y J, Kim K C, Choi Y W, et al. Text region extraction and text segmentation on camera-captured document style images. In: Proceedings of the 8th International Conference on Document Analysis and Recognition, Seoul, 2005. 172–176
15 Kim J S, Kim S H. Three-level color clustering algorithm for binarizing scene text images. KIPS Trans B, 2005, 12: 1–8
16 Jung K C, Han J H. Hybrid approach to efficient text extraction in complex color images. Patt Recogn Lett, 2004, 25: 679–699
17 Jung K C. Hybrid approach of texture and connected component methods for text extraction in complex images. IEEK, 2004, 41: 175–186
18 Huang T J, Tian Y H, Li J, et al. Salient region detection and segmentation for general object recognition and image understanding. Sci China Inf Sci, 2011, 54: 2461–2470
19 Zhou L, Hu D W, Zhou Z T. Scene recognition combining structural and textural features. Sci China Inf Sci, 2013, 56: 078106
20 Smith R. An overview of the Tesseract OCR Engine. In: Proceedings of the 12th International Conference on Document Analysis and Recognition. Washington DC: IEEE, 2007. 629–633
21 Ministry of Culture, Sports and Tourism. Korea Braille Regulations, MCST Regulation 2006-39. Korea, 2006
22 Way T P, Barner K E. Automatic visual to tactile translation-part II: evaluation of the TACTile image creation system. IEEE Trans Rehabil Eng, 1997, 5: 95–105
23 Rotard M, Knödler S, Ertl T. A tactile web browser for the visually disabled. In: Proceedings of the 16th ACM Conference on Hypertext and Hypermedia. New York: ACM, 2005. 15–22
24 Ladner R E, Ivory M Y, Rao R, et al. Automating tactile graphics translation. In: Proceedings of the 7th International ACM SIGACCESS Conference on Computers and Accessibility. New York: ACM, 2005. 150–157
25 Petit G, Dufresne A, Levesque V, et al. Refreshable tactile graphics applied to schoolbook illustrations for students with visual impairment. In: Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility. New York: ACM, 2008. 89–96
26 Chen J J, Nagaya R, Takagi N. An extraction method of solid line graph elements in mathematical graphs for automating translation of tactile graphics. In: Proceedings of the 13th IEEE International Symposium on Advanced Intelligent Systems. Washington DC: IEEE, 2012. 422–427
27 Takagi N, Chen J J. A broken line classification method of mathematical graphs for automating translation into scalable vector graphic. In: Proceedings of the IEEE 43rd International Symposium on Multiple-Valued Logic. Washington DC: IEEE, 2013. 71–76
28 Chen J J, Takagi N. A pattern recognition method for automating tactile graphics translation from hand-drawn maps. In: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics. Washington DC: IEEE, 2013, 4173–4178
29 Chen Q, Zhao L, Lu J, et al. Modified two-dimensional Otsu image segmentation algorithm and fast realization. IET Image Process, 2012, 6: 426–433
30 Deng Y N, Kenney C, Moore M S, et al. Peer group filtering and perceptual color image quantization. In: Proceedings of the IEEE International Symposium on Circuits and Systems. Washington DC: IEEE, 1999. 21–24
31 Gong Y. Advancing content-based image retrieval by exploiting image color and region features. Multimedia Syst, 1999, 7: 449–457
32 Shi J, Tomasi C. Good features to track. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE, 1994. 593–600