CHAPTER 15

SEGMENTATION OF TEXT IN WEB IMAGES

A. Antonacopoulos and D. Karatzas

PRImA (Pattern Recognition and Image Analysis) Group,
Department of Computer Science, University of Liverpool,
Liverpool, L69 7ZF, United Kingdom
E-mail:aa@csc.liv.ac.uk
URL: http://www.csc.liv.ac.uk/~prima

The abstract should summarize the context, content and conclusions of the paper in less than 200 words. It should not contain any references or displayed equations. Typeset the abstract in 9 pt Times roman with baselineskip of 11 pt, making an indentation of 1.5 pica on the left and right margins.

1. First-Order Heading

The first paragraph should not have the first line indented. The style to use is “Text first”. All subsequent paragraphs in the same section (or subsection) should be indented (use the “Text” style).

The paper size used in this template is A4. The top margin is set at 6.8 cm, the bottom margin at 4.8 cm, the left/right margins at 4.79 cm. There should be different headers at the top of even and odd pages (as shown in this template).

Headings should be as shown (do not use fourth-level headings) with the first letter of each word capitalised. The ‘References’ heading is not numbered.

References are formatted as shown at the end of this document. Citations are entered in the text as superscripts, starting at number 1 in order of appearance in the text. Note the examples in the text for cases when more than one citation[1] is made at the same point.[1,3,4] For instance, if two and above citations are made, separate the numbers by a comma. If a range of numbers is used, separate[4–7] the numbers by an *en-dash* (see below). In case of punctuation marks, such as comma or dot, list citations after the punctuation mark.[3]

Please do not use bookmarks, special fields and dynamically-generated citations unless you have tested them thoroughly.

Please ensure the quotation marks are paired correctly, e.g "good quotes" rather than ``bad quotes''.

Use a hyphen (-) for compound words (e.g. 'two-dimensional'), an *en-dash* (–)(can be generated by ALT+0150) to link numbers, nouns or names (e.g. 220–240 Volts, electron–positron collisions, Einstein–Rosen–Podolsky paradox), and an *em-dash* (—) (can be generated using ALT+0151) to link sentences or clauses—this is what we would regard as a 'normal' dash.

The standard abbreviations are:

- Equation(s) - Eq./Eqs.
- Figure(s) - Fig./Figs.
- Reference(s) - Ref./Refs.
- Section(s)- Sec./Secs.

Please spell in full if any of the above is the first word of the sentence. Also, use a fixed space (CTRL+SHIFT+SPACE) between the reference and identifiers (e.g. Fig.~1.1), etc.

Latin words are in italized, e.g. *et al.*, *a priori*, *in situ* etc.

1.1. *Second-Order Heading*

Please note that the text from here on is for *illustration* purposes only. Printed text is represented by Braille characters. Each character is constructed as a set of six points arranged in two columns of three, as it can be seen in Fig. 1. This arrangement is called a Braille *cell*. Each point position is identified by a number and can be either raised (a protrusion) or flat. A raised point is also called a Braille *dot*.

The dimensions of a Braille dot have been set according to the tactile resolution of the fingertips of person. In theory, the dimensions of Braille dots must fall within certain bounds as indicated in Fig. 2. The horizontal and vertical distance between dots in a character, the distance between cells representing a word and the inter-line distance are also specified by the Library of Congress. In practice, although most instances of Braille are close to these standards, there are slight variations between Braille produced by different devices.

Each Braille character is a unique combination of dots within the cell. So far, only Braille cells with six points have been mentioned. There is, however, an eight-point version of Braille cells whose use is not very widespread compared to that of the six-point ones. The two extra points allow for the representation of a greater variety of printed symbols. It has mainly been used in computer programming where a variety of symbols is used and each must be represented unambiguously by a single Braille character (see Sec. 1.1). The larger number of possible dot combinations has also been used for the representation of Japanese characters[1]. Since from the recognition point of view the principle is the same, in the rest of this chapter Braille characters will be assumed to be composed of six points, unless otherwise explicitly stated.

1.1.1. Third-Order Heading

In *grade 1* Braille, each printed letter corresponds to one Braille character as can be seen in Fig. 3. Punctuation marks are also represented by one Braille character with the exception of the opening quotes and the exclamation mark which are represented by the same Braille character (see Fig. 3). In this case, while reading, the interpretation is made according to the context.

The combination of dots in each Braille character is designed in such a way that symbols are distinguished from certain others by a translation of the same dot arrangement (e.g. the letter “d” and the full stop “.”). Furthermore, the symbols that are represented by only one dot are likely to be found next to others with dots in other positions which will help define the position for the reader.

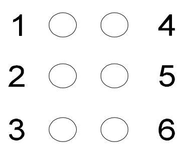

Fig. 1. The Braille cell.

It is not possible to represent the large number of existing symbols by a combination of dots in a single Braille character. Therefore, some symbols (e.g. the asterisk) are represented by two Braille characters. Furthermore, the digits “1”, “2”, ..., “9”, “0” are represented by the Braille characters corresponding to the range of letters “a”,...,“j” preceded by the *numeral* symbol as can be seen in Fig. 3. Still, grade 1 Braille is quite straightforward to interpret.

The numeral symbol is an example of a *mode* symbol. Another mode symbol is that used to indicate the capitalisation of a letter (Braille character composed of dot six only). It is mainly used in North American Braille and precedes the symbol of the character to be capitalised. Similarly, mode symbols are used in many languages to denote that a particular accent should be placed on the character that follows.

1.2. *Another Second-Order Heading*

Most Braille documents are formed by embossing the dots on the back side of the medium sheet so that they can be read from the facing side. Most often the medium is thick paper in order to withstand the embossing. Another, less common, medium is a transparent plastic sheet (similar to an overhead projector transparency). In addition to the embossing production technique, a process of forming dots on paper from drops of hot-melt material has also been reported.[2] This chapter, however, will concentrate on the embossed-paper Braille documents because they essentially constitute the body of the information available in Braille form.

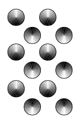

Fig. 2. Two cells on different sides of an inter-point Braille document. When the figure caption extends beyond one line, it should be justified instead of centere, like in this example.

One way of embossing the sheet of paper is by hand using a Brailler, which resembles a typewriter with a key corresponding to each of the six points (plus a space key). To produce a dot, the user applies force to the appropriate key which embosses the paper. For mass production, plates are formed which are then pressed against the paper sheets to form pages of Braille. Furthermore, as it has already been mentioned, there are Braille printers, connected as peripherals to computers, which emboss the dots on paper according to the information sent in electronic form.

The size of the page varies between documents produced by different means. In addition, the colour of the paper varies, as it does not play any role in conveying the written information. The majority of Braille documents, however, are either buff-coloured or white. In addition, their surface usually has a low gloss finish.

The Braille characters are embossed in lines from the top of the document to the bottom much like printed documents. Braille documents can have the dots embossed on one side or on both. As the volume of Braille documents is quite large, it is advantageous to use both sides of the sheet. This is achieved by carefully embossing the dots of one side between the dots embossed on the other, as can be seen in Fig. 2. Because the points of Braille cells in one side are interleaved with the points of the cells of the other, these double-sided documents are also referred to as *inter-point*.

2. Another First-Order Heading

For a better analysis of the different aspects of the automatic Braille reading problem and for ease of comparison between current solutions, the stages of the automatic Braille reading process are examined individually. The first stage to be considered is the acquisition of the image of the tactile surface of the document. This is followed by various enhancements performed on the image prior to the extraction of any information. In the third stage, the Braille dots are identified in the image.

The two stages that follow are concerned with the segmentation and recognition of Braille characters. The possibilities for verification of the recognition results are examined next and, finally, the main uses of the encoded Braille characters are described.

2.1. *Another Second-Order Heading*

At this stage, the regions in the image corresponding to Braille dots are identified. In a binary image, dots can be identified as small connected components. In the case of the twin shadow approach of Hentzschel,[11] a Braille dot corresponds to two such components.

It should be noted that this situation is valid only when single-sided Braille documents are encountered. In this case, the above approaches are able to identify dots from either the dark (shadow) or the light areas in the image which, after thresholding, are represented as connected components. In the special case of transparent Braille documents, Nitta *et al.*[1] identify connected components having the form of circular arcs. These are then processed so that complete circles are obtained. The centre of each of these circles is identified as the location of the Braille dot.

The two stages that follow are concerned with the segmentation and recognition of Braille characters. The possibilities for verification of the recognition results are examined next and, finally, the main uses of the encoded Braille characters are described.

It should be noted that this situation is valid only when single-sided Braille documents are encountered. In this case, the above approaches are able to identify dots from either the dark (shadow) or the light areas in the image which, after thresholding, are represented as connected components. In the special case of transparent Braille documents, Nitta *et al.*[1] identify connected components having the form of circular arcs. These are then processed so that complete circles are obtained. The centre of each of these circles is identified as the location of the Braille dot.

The algorithm proceeds as follows:

```
for each dark region
    if a light region exists above then
        a depression is found
    else
        if a light region exists below then
            a protrusion is found
            check regions and mark the appropriate ones as "used"
```

Having located the lines of Braille characters, the next step is to determine the positions of the characters themselves inside each line. In order to do so, the positions of character columns having at least one dot are first identified from the horizontal projection-profile. The space between columns is then examined to determine whether a column is the left or the right one of a Braille character. The spacing between adjacent columns can vary beyond the obvious intra-cell and inter-cell values.

This is due to the fact that a Braille character may have dots in only one column. Hence the distances between columns may correspond to one of the following:

(i) space between the two columns of the same character (*a* in Fig. 3),
(ii) space between a *right* column and the *left* one of the next character in the word, i.e. the space between adjacent Braille cells (*b* in Fig. 3),
(iii) space between a *right* column and a *right* columned character (*c* in Fig. 3),
(iv) space between a *left* columned character and the *left* column of the next character (*d* in Fig. 3),
(v) space between a *left* columned character and a *right* columned one (*e* in Fig. 3), and
(vi) space between words (e.g. *f* in Fig. 3), which varies according to the number of space characters present.

To segment the characters in a line, the distance between consecutive columns is examined. Accordingly, each column is characterised as a left or a right one. If the distance between two columns is greater than the sum of the inter-character distance and the width of a Braille character (space character in this particular case), it is determined as inter-word distance and the number of Braille space characters is calculated. Special consideration is given to the situation where two single-columned characters are encountered. Ambiguity can arise if they are both left-columned or both right-columned (distance *c* or *d* in Fig. 4).

3. Concluding Remarks

A comparison of all methods on the same image data is however needed in order to obtain an objective measure of recognition performance and to progress forward. Apart from a high recognition performance, the time taken for the completion of the reading process is of significant importance. Repeated image accesses and the need for time-consuming operations can render an approach less practical than others.

For example, dependence on skew correction, even for very small angles, is not desirable. The use of computationally intensive image transforms should also be kept to a minimum. Finally, it is advantageous to obtain a representation of the dot regions, in terms of their co-ordinates and size, and use this through the various stages instead of having to search and index through the volume of the image data.

References

1. S. Nitta, J. Oshima and M. Okhawa, Recognition algorithm on transparent Braille, *Proc. 28th SICE Annual Conference*, Matsuyama, Japan, 1989, 1017–1020.
2. M. Benjelloun, V. Devlaminck, F. Wauquier, P. Altmayer and J. P. Dubus, Application du traitement d'images à la conception d'un système de transmission de relief Braille en texte noir, *Traitement du Signal* **6** (1989) 291–300.
3. T. Otake, T. Saito and Y. Yonezawa, Braille printer using hot-melt material, *J. Microcomputer Applications* **13** (1990) 123–131.
4. F. Germagnoli and G. Magenes, Computerised system for helping blind people to learn Braille code, *Proc. 15th Annual Int. Conf. on Engineering in Medicine and Biology*, 1993, 1318–1319.
5. O. Sueda, Braille translation by microcomputer and a paperless Braille dictionary, in *Computerised Braille Production Today and Tomorrow*, eds. D. W. Croisdale, H. Kamp and H. Werner (Springer-Verlag, 1979) 272–289.
6. Royal National Institute for the Blind, *Braille Primer* (RNIB, 1992).

7. W. Slaby, Computerised Braille translation, *Computers for Handicapped Persons, Proc. Conf. on Computers for Handicapped Persons*, eds. A. M. Tjoa, H. Reiterer and R. Wagner (R. Oldenbourg, Vienna, Austria, 1989) 26–34.
8. P. Blenkhorn and T. Maley, Computer Braille code, *Computer Assisted Learning for the Visually Handicapped*, Information Sheet 4 (Research Centre for the Visually Handicapped, Faculty of Education, University of Birmingham, 1985).
9. G. François and P. Calders, The reproduction of Braille originals by means of optical pattern recognition, *Proc. 5th Int. Workshop on Computer Braille Production*, Heverlee, 1985, 119–122.
10. J. P. Dubus, M. Benjelloun, V. Devlaminck, F. Wauquier and P. Altmayer, Image processing techniques to perform an autonomous system to translate relief Braille into black-ink, called: lectobraille, *Proc. IEEE Engineering in Medicine and Biology Society 10th Annual Int. Conf.*, New Orleans, 1988, 1584–1585.
11. T. W. Hentzschel, An optical Braille reading system, M.Sc. Dissertation, Faculty of Technology, University of Manchester, 1992.
12. J. Mennens, L. Van Tichelen, G. François and J. Engelen, Optical recognition of Braille writing, *Proc. 2nd Int. Conf. on Document Analysis and Recognition*, Tsukuba, Japan, Oct. 1993, 428–431.
13. R. T. Ritchings, A. Antonacopoulos and D. Drakopoulos, Analysis of scanned Braille documents, *Proc. Int. Association for Pattern Recognition Workshop on Document Analysis Systems*, Kaiserslautern, Germany, Oct. 1994, 417–424.
14. D. Drakopoulos, Analysis of scanned Braille document, M.Sc. Dissertation, Faculty of Technology, University of Manchester, 1993.