February 2013. Vol. 2, No.1 ISSN 2305-1493
International Journal of Scientific Knowledge
Computing and Information Technology

www.ijsk.org

ALGORITHM PROTOTYPE FOR REGULAR FEATURE EXTRACTION OF ARABIC BRAILLE SCRIPTS

Dr. Zainab [I]. Authman1, Zamen F. Jebr[2]
Collage of science-university of Basrah[1],
University of Thi-qar[2]
,

Abstract:

Braille is one of the important means of written informational communication between visually-impaired and sight people, so it gains the research interest. This paper describes a new technique for recognizing Braille cells in Arabic single side Braille document. The mean challenge in (OBR) system -which is resolved here- is how to binaries Braille image So we suggested two algorithms two binaries Braille sheets based on morphological operations. We apply Opening on green Braille and closing in yellow Braille sheets. Optical Arabic Braille Recognition system has two main tasks, first is to recognize printed Braille cells, and second is to convert them to regular text**.** We introduce in this search, automatic system which converts printed Braille Arabic to binary digital format to perform the first task from recognition system. The system provide high accuracy and fast result.

Keywords: *image processing, morphological operation, Braille, optical recognition*

1.Introduction:

Witness the first decade of the twenty-first century a major development in the field of information processing of all kinds, one of the most important developments is a technology that recognition and processing patterns and digital images, has invested this development in the field of education through the use of image processing techniques and patterns within educational institutions, especial-ly institutions with special needs such as the blind and visually impaired.

Visually impaired people are part of the society, and can play an integral role in its prosperity. Therefore, it has been a must to provide those people with means and systems through which they may communicate with the world. These systems should depend on the sense of hearing or touching. Many systems have been de-veloped to achieve this purpose; the most famous system that is based on touching is Braille system. Braille is a writing system that enables visually people to read and write through touch using a series of raised dots to be read with their fingers.

The Braille is a system of touch reading and writing in which raised dots represent the letters of alphabet. Braille also contain equivalents for punctuation marks and provides symbols to show letter groupings. Braille is read by moving the hand or hands from left to right along each line. Both hands are usually in-volved in the reading process, and reading is generally done with the index fingers [C.Ng ,1999].

Braille was first introduced in 1825 by French scientist Louis Braille. He took a secret code devised for the military and saw in the basis for written com-munication for blind individuals. The original military code was called night writ-ing and was used by soldiers to communicate after dark. It was based

on a twelve-dot cell, two dots wide by six dots high. Each dot or combination of dots within the cell stood for a letter or a phonetic sound. The problem with the military code was that the human fingertip could not feel all the dots with one touch [Hentzchel,T.W., 1995].

Louis Braille created a reading method based on a cell of six dots. This crucial improvement meant that a fingertip could encompass the entire cell unit with one impression and move rapidly from one cell to the next. The system of embossed writing invented by Louis Braille gradually came to be accepted throughout the world as fundamental form of written communication for blind individuals, and it remains basically invented it [C. Ng ,1999].

A Braille character (or Braille cell), is a rectangular array of six points ar-ranged in two columns of three row as can be seen in Figure 1. Each of the points in a cell can be either raised (a bump) or flat. A raised point will be referred to here as a dot. On the document, the Braille characters are embossed from left to right and from top to bottom, much like characters in ordinary documents[Ritchings R.T. ,1995].

The Braille code system has become widely used by several communities because of its simplicity, comfortable for blinds to use it when read and write. Braille was applied or translated into several languages including Arabic language .

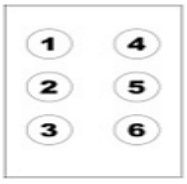

Figure 1, Braille cell

An obvious, perhaps, but significant characteristic of Braille documents is the absence of any information visible in a color contrasting the background. The only information recorded on a Braille page is in terms of the protrusions created by embossing the card. The fact that Braille documents are not intended to convey any visual information also has repercussions on the quality of card used to produce them. It is not uncommon for the card to be of low (visual) quality, with visible grain and imperfections (dark and light regions). This fact can affect the recognition of Braille documents by visual means [A.Antonacopulos , 2004].

Every Braille cell dot's can be set or cleared, giving 64 possible combination in six-dot and 256 combinations in eight-dot Braille. Each of the 64 different Braille characters can correspond to an ordinary character (grade 1 Braille) or to a group of characters (grade 2 Braille)[Ritchings R.T. ,1995,Mennens , J., 1994].

2.Interest of optical Braille recognition

Nowadays, lack or problem of vision has been an important obstacle to access to printed contents and to the information society. For this reason, some people have tried to achieve that blind people are able to access to the printed culture for example Louis Braille understood importance of a communication code. Globally, an estimated 40 to 45 million people are blind and 135 million have low vision according to the World Health Organization (WHO) [http//www.who,2005] and this number grows every year.

A Braille Optical Character Recognizer is interesting due to the following reasons [Nestor Falcon ,2005]:

- It is an excellent communication tool for sight people (who do not know Braille) with the blind writing.
- It is a cheap alternative Braille to Braille copy machine instead of the current complex devices which use a combination of heat and vacuum to form Braille impressions.
- Braille writing is read using finger so is necessary touch the document, for this reason the book after many

readings is possible has been deteriorated.

- It is interesting to store a lot of document of blind authors which were written in Braille and were never converted to digital information.
- Braille recognition system offers a better integration of blind people to the "information society".

The structure of this paper is the next: In the section 3 we will explain Literature review and In section 4 we describe the characteristic of data base created. After that section 5 we explain methodology and techniques that we are used it in recognition system. The last section 6 include results and conclusion

3.Literature Review

In this section presented a brief explanations of attempts had been made to optically recognize embossed ***Arabic Braille writing*** using various methods.

In 2007, AbdulMalik Al-Salman and his team in their work [AbdulMalik Al-salman , 2007] design system to recognize Arabic Braille writing. First, an image of a Braille document page is obtained using flatbed scanner. Second, the image is converted to a gray color. Following that any white or black frames are cropped. The image is then thresholded using two threshold values so that only three classes of region exist: dark, light and background. Having labeled each of the different types of regions, an initial identification of Braille dotes is performed, Braille cell dot's is convert to (4×4) array to reduce number of dot dedicate for each dot in Braille cell. Finally, Braille cells are then recognized. The system apply on brown and white Braille document single and double side.

In 2010, Saad D. Al-Shamma and Sami Fathi in their work [Saad D.Al-shamma ,2010] design system to recognize Arabic Braille writing. First, The test image has scanned by using HP Scan jet djf2200 scanner with horizontal and vertical resolution 200 dpi, bit depth 24, dimension 1621X2248 and the image store in JPEG format. Second, image convert to gray color. Following image convert to binary using global threshold method. canny method is used to finds edges by looking for local maxima of the gradient of input image. The filtering function removes from a binary image all connected components (objects) that have fewer than specified number of pixels, producing another binary image that filtered from micro objects. This system apply on brown Braille ,single side Braille document.

In 2010, Eman Kaise in here work [Zainab A , 2010] design system to recognize Arabic Braille writing. First, an image of a Braille document page is obtained using flatbed scanner. Second, the image is converted to a gray color. They use median filter to remove salt and pepper noise result from optical scan. The image is then thresholded using one threshold value get it manually and enter to the system so that only two classes of region exist: black regions represent background and white region represent Braille dots (recto). Having labeled each of the different types of regions, an initial identification of Braille cell dot's is convert to (3×3) pixels array to reduce number of pixels dedicate for each raise dot in Braille cell. Finally, Braille cells are then recognized. The system applied on yellow Braille document with single side.

4-Database

A big database has been created in order to check the global system with as many characters as we could. This database provides single side Braille documents with color green and yellow which have dots in one side of sheet. The next table (see Table 1) give a full explanation of this database.

Braille sheets	*25 green sheets, 25 yellow sheets*
Digital format	*Color*
Resolution	*100 dpi(horizontal and vertical)*

Image size	*(274 kb)in yellow sheets & (336) in green sheets*
Image format	*Jpg*
Braille type	*Single side – grade 1*
Document size	*29.5 cm. (horizontal)×30.5cm(vertical)*

*5-**System Overview***

The Braille recognition system consist of five operations as illustrate in figure 2:

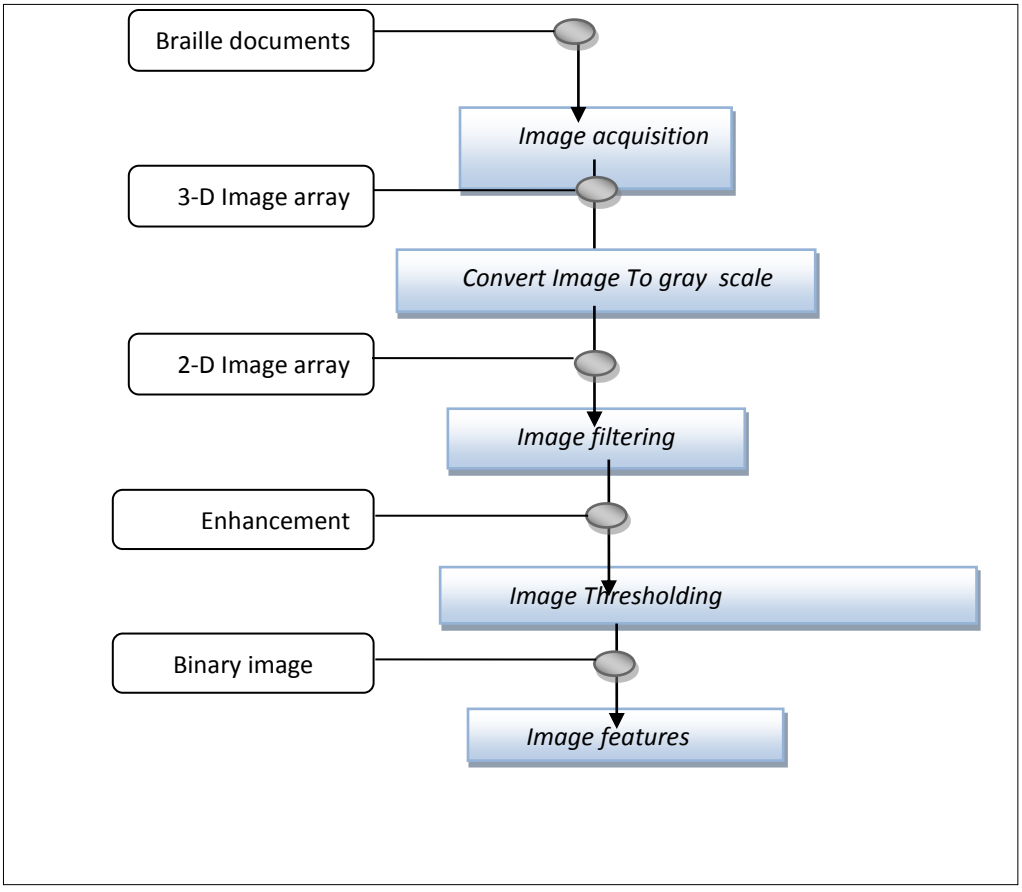

Figure 2 system overview

Image Acquisition

The test image has scanned by using HP Scan Deskjet F2400 scanner with horizontal and vertical resolution 100 dpi, bit depth 24, and the image store in JPEG format. Figure 3 illustrate sub-image of Braille sheets with colors green and yellow.

Figure3 a) yellow Braille sheet b) green Braille sheet

5.2

Convert image to grayscale

Inside a computer system, colored images are stored in 3-D arrays while gray level images are stored in 2-D**.**arrays. Dealing with 2-D arrays is much easier and faster. Therefore, the second step in the proposed system converts colored scanned images to gray level so that any pixel value in the image falls within the range 0- 255 as could be seen in Figure4.

Figure4 a) yellow Braille sheet after convert to gray scale b) green Braille sheet after convert to gray scale

There is difference in gray levels of two Braille sheets images is clearly (green and yellow) because difference in distribution of lighting level in two image where yellow sheet appear more light from green sheet so this will effect on threshold for this image so we use two threshold algorithms for binaries it.

5.3 Image filtering

During image acquisition, impulse noise was introduced in the image. These noises generally manifest themselves as random fluctuation in gray-level values superimposed upon the "ideal" gray-level value, and it usually has a high spatial frequency. Therefore, a low-pass spatial Gaussian filter is applied to the image to attenuate the high spatial frequency noise from the image while at the same time preserving the detailed edge information of the Braille dots [Agui T. 1994].

5.4 ***Image Thresholding***

Since Braille pages contains dots (foreground) and page (background) only, analysis a binaries image is much simpler than that of gray-scale images. For a simple global threshold where the image histogram is bimodal or has easily identifiable peaks and valleys, the selection of the threshold value is straightforward. However, the digitized Braille page images are noisy, and there are considerable spread in gray level values, the selection of threshold value problematic as it illustrate in the histogram of Braille images in figure 5 [F.Zahedi ,1993].

To cope with significant variation in lightness across the whole image, we proposed to use the following two algorithm (algorithm1 and algorithm2) to threshold green and yellow Braille pages images.

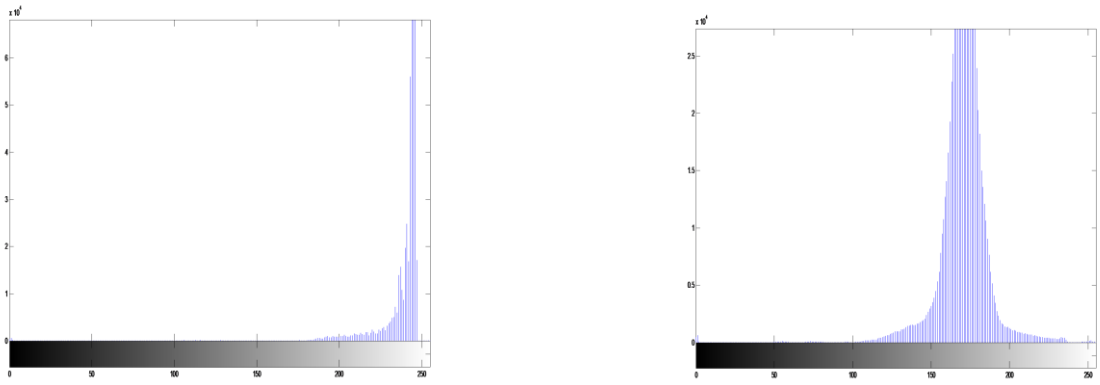

Figure 5 a)histogram yellow Braille image b) histogram green Braille image

Algorthim1

We suggest to use local thresholding method on green Braille page because global thresholding methods fail when the background illumination is uneven, as was noticed in figure4 (b). A common practice in such situations is to preprocess the image to compensate for the illumination problems and then apply a global threshold to the preprocessed image. Apply opening operation on image help us to compensate for non uniform illumination. Opening operation are use to remove small bright details while leaving the overall gray levels and larger bright features relatively undisturbed. Because opening suppresses bright details smaller than the structural element it is used for image smoothing and noise removal [Rafael C.Gonzalez , 2003].

Algorithm 1 threshold green Braille images pages

Step1*: smoothing image by apply opening*

1-set structural element (z)and must be disk shape and size=4. which is define in domain D_z
2 –Rotate the structural element (z) generated in the previous step about its origin.
3-Translate the structural element to all location in the image f(x,y)
4- At each translated location, the rotated structural element values are subtracted from the image(f) pixel value and the minimum is computed result in R1 image:

$$R1(x,y)=\min\{f(x+x',y+y') \mid (x',y') \in D_z\}$$

5- Translate the structural element to all location in the image R1(x,y).
6-At each translated location, the rotated structural element values are added to the image(R1) pixel value and the maximum is computed result in R2 image:

$$R2(x,y)=\max\{R1(x-x',y-y') \mid (x',y') \in D_z\}$$

7-Subtract the last result image from the original:

$$R=(f-R2)$$

Step2*: Apply global threshold (iterative procedure)*

1-Select midpoint between minimum and maximum intensity values in the R image.
2- Segment the image using T to produce two group of pixels G1>T and G2<T.
3-Compute the average intensity values μ_1 *and* μ_2 *for pixels in regions G1 and G2 respectively.*
4-Compute new threshold value :

$$T=1/2(\mu_{1+}\ \mu_2)$$

5-Repeat step2 through 4 until the difference in T in successive iterations is smaller than predefined parameter T_o.

* A procedure consist of step 1 & step 2 called local thresholding method.

Algorthim2

We suggest to use local adaptive thresholding method on yellow Braille page.. Apply closing operation on image help us to compensate for non uniform illumination. closing operation are use to suppresses dark details small than the structuring element. Because closing suppresses bright details smaller than the structural element it is used for image smoothing and noise removal [Rafael C.Gonzalez , 2003]. local adaptive method works by dividing the image into32×32 pixel region (the size of window is experimentally derived) and assesses whether region contains whole dots or just background. This assessment equivalent comparison of sets of ranges of gray levels observed in that region [R.C.Gonzalez ,2002].

Algorithm 2 *threshold yellow Braille images pages*

Step1: *image smoothing by apply closing*

1. *Set structural element (z)and must be disk shape and size=4 which is define in domain* D_z .
2. *Rotating the structural element (z) generated in the previous step about its origin.*
3. *Translate the structural element to all location in the image f(x,y)*
4. *At each translated location, the rotated structural element values are added to the image(f) pixel value and the maximum is computed result in R1 image:*

$$R1(x,y)=max\{f(x-x',y-y')+ \mid (x',y') \in D_z\}$$

5. *Translate the structural element to all location in the image R1(x,y).*
6. *At each translated location, the rotated structural element values are subtracted from the image(R1) pixel value and the minimum is computed result in R2 image:*

$$R2(x,y)=min\{R1(x+x',y+y') \mid (x',y') \in D_z\}$$

7. *Subtract the last result image from the original:*

$$R=(f-R2)$$

Step2:*threshold resulted image using local adaptive method*

*%%segment the image R into(32*32 pixel regions);*

set m =32;n=32;

set P=1;q=1;T=false;

while T

iF m>num-of rows in image R

T=True;

Break;

Else

Z=0.5(maximum (R(1 to m,1 to n))+minimum(R(1 to m,1 to n)))*

For i from p to m

For j from q to n

If R(i,j)>=Z

Rnew(i,j)=1;

Else

Rnew(i,j)=0;

update q with value (n+1);

update n with value (n+32);

End {if};

update P with value (m+1);

update m with value (m+32);

End {for}

End {for}

End {if}

End{while}

We display result of apply this two algorithms on Braille sheets image in the figure (6).

a b

Figure 6 a) yellow Braille image after apply algorithm 2 b) green Braille image after apply algorthim1

5.5 Feature extraction and dots centres's detection

The main function of this phase is to extract the Braille dots from the binarised image. This includes boundary detection where we search from coordinates that busy by Braille dots and find maximum row, minimum row and maximum column ,minimum column for this dot and determine the center of it according to following equations.

$$cenX=(maxX-minX)/2+minx$$

$$cenY=(maxY-minY)/2+minY$$

After that we apply shrinking mechanism, where each dot in Braille cell will be represent in one pixel and it is placed in center location of dot in last cell in Braille image and so on for each cell dot's in binary image as in figure7. This procedure gave us two benefits:

- Make dots of Braille cell in more regular shapes.
- Reduce search time of dot in second task from (OBR) system.

Now, we ended the first task of (OBR) system and Braille sheet image became officinal to the second task of Braille Recognition system (translate Braille cells in to equivalent text).

Figure7 Braille image after apply shrinking mechanism

6. Results and Conclusion

In this paper we have explained the development of an automatic system for recognizing printed Braille cells. It has been divided in different modules for each part of the image processing. For achieving this system, local and adaptive thresholding has been used ,and shrinking mechanism is added to the system for make Braille cells shape's more regular to relent the next task of (OBR) system. This process has an efficiency and it take only 17 sec. to recognize green Braille sheets and 21 sec. in recognize yellow Braille sheets.

References

A. Antonacopulos and D .Bridson ."**A Roubust Braille Recognition System**".Document Analysis system VI ,2004.

AbdulMalik Al-Salman et al**,"An Arabic Optical Braille Recognition System"**,ICTA'07, Hammamet, Tunisia, April, 2007.

Active on April 1st 2005.

Agui T. and Nagao T., "**Computer Image Processing and Recognition** ". Tokyo: Shoho-do, 1994.

C. Ng and V. Lau, "**Regular feature extraction for recognition of Braille** ",ICCIMA'99,1999.

F.Zahedi and R. Thomas, "**Hybrid Image Segmentation within a Computer Vision Hierarchy**", IJEEE, 30, pp. 57-64, 1993.

Hentzchel,T.W.,and P.Blenkhorm**."An Optical Reading Systems for Embossed Braille Characters using a Twin Shadows Approach**", Journal of Microcoputer Applications,pp.341-345.1995.

http://www.who.int Active on April 1st 2005.

Mennens, J., van Tichelen, L., François, G., Engelen, J.J.,**"Optical Recognition of Braille Writing Using Standard Equipment** ". IEEE Trans. on Rehabilitation Engineering, vol. 2 no. 4 December (1994) 207–212.

Nestor Falcon et al ,"**Image Processing techniques for Braille writing recognition"**. EUROCAST 2005, LNCS 3643, pp. 3643, pp.379-385, 2005.

R.C. Gonzalez and R.E. Woods, "**Digital Image Processing**". Prentice Hall,2nd edition,2002.

Rafael C. Gonzalez, Richard E. Woods, Steven L., Digital Image Processing Using MATLAB(R)". Prentice Hall,2003.

Ritchings, R.T., Antonacopoulos, A., Drakopoulos, D., "**Analysis of Scanned Braille Documents**", In: Dengel, A., Spitz, A.L. (eds.): Document Analysis Systems, World Scientific Publishing Company (1995) 413–421.

Saad D. Al-Shamma and Sami Fathi, "**Arabic Braille Recognition and Transcription into Text and Voice**". 2010 5th Cairo International Biomedical Engineering Conference Cairo, Egypt, December 16-18, 2010 .

Zainab A. and Eman Q. , "Automatic Recognition for Arabic Braille Scripts", 2010.

نموذج خوارزمي لاستخراج الميزات المنتظمة في مخطوطات برايل العربية

نظام برايل هو احد الوسائل المهمة للاتصال معرفياً بين الأشخاص المكفوفين والأشخاص المبصرين ولذلك عُنيَ هذا الموضوع باهتمام الباحثين. هذا البحث يصف تقنية جديدة لتميز خلايا مخطوطات برايل العربي ذات الوجه الواحد. الصعوبة في أنظمة تمييز برايل – والتي عولجت في هذا البحث- هو كيفية تحويل صورة مخطوطة برايل إلى صيغة الصورة الثنائية لذا قمنا باقتراح خوارزميتين لتحويل وثائق برايل إلى الصيغة الثنائية (على ورق برايل closing) على ورق برايل ذات اللون الأخضر وعملية (openingبالاعتماد على العمليات التركيبية للصورة حيث استخدمنا عملية (الأصفر نظام التميز الضوئي لمخطوطات برايل يتضمن مهمتين أساسيتين، المهمة الأولى هي تميز خلايا مخطوطات برايل والمهمة الثانية تحويل هذه الخلايا إلى نص. وقد عرضنا في هذا البحث تقنيات آلية لتحويل مخطوطات برايل العربية إلى صيغة رقمية لغرض تنفيذ المهمة الأولى من نظام التميز. التقنيات قدمت نتائج سريعة ودقيقة في تميز خلايا مخطوطات برايل العربي.