AMSE JOURNALS-AMSE IFRATH Publication –2014-Series: Modelling C; Vol. 75
Special Handicap; pp 43-54
Submitted Sept. 2014; Accepted Oct. 31, 2014

A New Approach for Automatic Transformation of Pedagogic Images into Tactile Images

Yong Chen, Zehira Haddad, Jaime Lopez Krahe
THIM (EA 4004 CHART), University Paris 8,
2 rue de la Liberté, 93526 Saint Denis, France
(ychen09@etud.univ-paris8.fr, {zehira.haddad-bousseksou,jlk}@univ-paris8.fr)

Abstract

This paper presents an automatic transcription approach of educational images into tactile images. The proposed approach consists of several work phases. In each phase, some methods of digital image processing have been applied in order to arrive to the automation of the transcription process. The obtained tactile images are of good quality, which encourages us to continue this work and to adopt this methodology in order to extend it to other types of images.

Key words

Accessibility, pattern recognition, tactile transcription.

1. Introduction

A great number of organizations and associations work on the development of adapted interfaces for blind or visually impaired people. The main objective of these works is to facilitate the access to knowledge, learning, training and education of people with disabilities. Among the existing works, we cite the development of haptic visualization systems for blind or visually impaired people [1], access to graphical web documents using a mouse with a feedback force [2] and the work of Way which concerns the automate transcription of the tactile images [3].

So far, converting images into tactile documents was done manually; this work is slow, tedious and expensive. For this, persons are formed to achieve this work. Thus, the produced tactile maps would be easily exploited by the blind or visually impaired people. The objective of the proposed work is to facilitate the conversion process by automating it. This automation will be achieved by using the current advances in the field of digital image processing and more specifically that of pattern recognition. Indeed, with the development of the technologies in this field, it is possible to propose solution to the addressed problem. However, despite the very good result of the current research, it is impossible to have a success rate of 100% for any type of images. Take for example the automatic character recognition, even if the performance of today's software is very high, there may be also errors in some cases. So, in our work, we decided to insert a tool which permits the verification and validation by the users. The aim of the development of this validation tool is to ensure the maximum reliability of the system. Given that the produced documents by our system will be exploited by people with special needs, this system must respect the various constraints related to the concerned disability.

In the end, the proposed approach allows us to consider a wider application with other educational documents like maps, geometric figures, statistical graphs, analytical curves, amongst others. All this allows us ultimately to open direct access to a large part of digital documents which can be found on the internet.

In order to present our work, we organized the paper as follows: after describing the overall framework of this application, we present in Section 2 the various steps of the proposed process and also the different techniques which we used. In Section 3, we present and discuss the results for each step of the proposed process. Section 4 shows experimental results and Section 5 concludes the work.

2. Adopted methodology

We present in Figure1, the different steps of the automatic transcription process. Firs, the proposed approach applies a threshold to the gray level of the image in order to have a binary image which has only two gray levels: 0 (black) and 1 (white). Indeed, the treated educational images have the contents fairly simple. Therefore, the thresholding step permits to minimizing the loss of semantic information which can affect the image. In the case of color images, a segmentation step is applied in order to separate the different regions and then assign to each of them an appropriate texture.

The second step is dedicated to separate the various shapes contained in the image. This separation is achieved through the analysis step by connected components [4]. For each shape, a feature vector which uniquely characterizes this shape is determined. This vector represents the characteristic properties which can be used to recognize and identify each shape easily. This automatic shape recognition step is finally done by using a fuzzy clustering algorithm [5,6,7]. For each detected shape, we use the Braille text to describe it in the legend. Otherwise, to treat the text region in the image, we propose to apply an automatic process to recognize it. This process involves different steps including: detection, extraction and automatic character recognition. The detected text will be finally automatically transcribed in Braille.

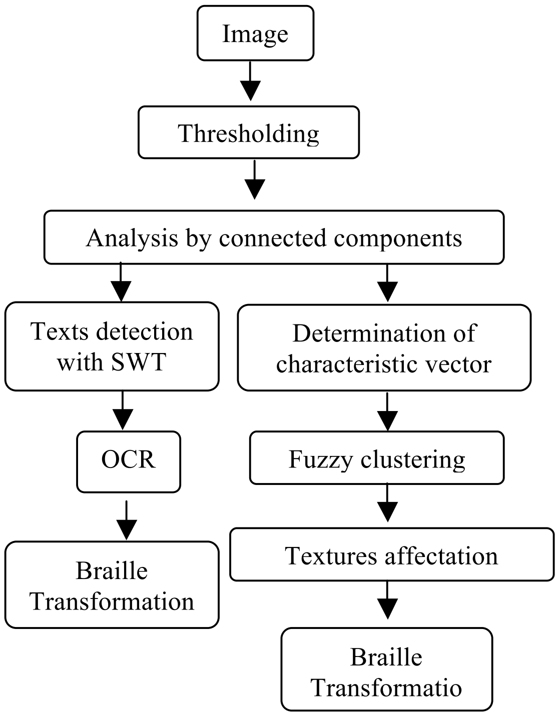

Fig. 1. Flowchart description of the proposed approach.

2.1 Thresholding

The first step of the proposed approach is segmentation by thresholding. In our case, the images are simple, we decided thus to make the image into binary in order to simplify the different pixel values. This step reduces them into two values 0 and 1. Indeed, we can thus manage better the content of the image.

2.2 Analysis by connected components

This step represents the second level of segmentation. It consists in finding the connected pixels called connected components. This permits to extract the shapes in the image. Each separated object is uniquely identified by an appropriate label. The main advantage of this segmentation is to analyze the content of the image in order to identify and extract the interesting objects.

2.3 Shapes recognition

In some educational images (like for example astronomical images), we can find usually some specific shapes which correspond to certain information of the image. For this image example, the most often found shapes are circles and ellipses which correspond to the different planets and the sun. We propose in the context of this work, an automatic detection and an intelligent recognition of different geometrical shapes in the image. To each specific shape, corresponds some specific properties which characterize this shape. We then propose to use a vector which contents these shape's parameters. We called this vector the characteristic vector. For this, in order to recognize a shape and distinguish it from the others, we used the following mathematical parameters:

- The number of vertices which is given by the Harris detector [8]. Figure 2 shows the peaks of some simple geometric shapes. These peaks are represented in this figure by red crosses.
- The ratio between the concave surface and the convex surface A/Ach. In Figure 3, we represent the convex portion and the concave portion of several geometrical shapes.
- The eccentricity mathematics: this is a real positive parameter characteristic for ellipse.
- The ration between the concave surface and the bounding box surface A/Ab. Figure 4 show the difference of them.
- The ratio between the squared perimeter of the convex hull and its surface Pch^2/Ach.
- The ratio between the perimeter of shape and the perimeter of its convex hull P/Pch.

The following figures illustrate the previous concepts used in the determination of the feature vector:

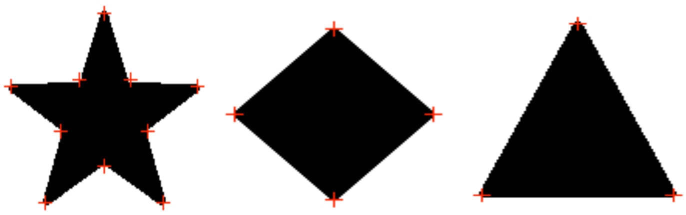

Fig. 2. The vertices of different shapes

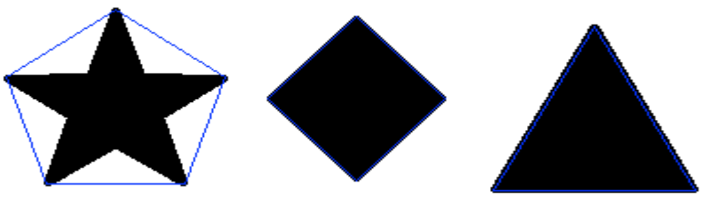

Fig. 3. The difference between concave and convex surface.

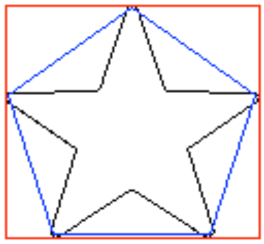

Fig.4. Representation of the convex hull (blue line) and bounding box (red line).

Each shape corresponds to a characteristic vector with parameter values. Table 1 shows the ideal values for each shape.

	Number of Vertices	Eccentricity	A/A_{ch}	A/A_b	P_{ch}^2/A_{ch}	P/P_{ch}
Triangle	3	X	1	0.5	X	1
Square	4	0	1	1	X	1
Rectangle	4	X	1	1	X	1
Rhombus	4	X	1	0.5	X	1
Circle	0	0	1	$\pi/4$	4π	1
Ellipse	0	X	1	$\pi/4$	X	1
Star	10 (5 on the convex hull)	X	0.6	0.4	X	X

Table 1. Mathematical features corresponding to different forms

Then, based on the calculation of these parameters, a feature vector representative of the shape is extracted. Thus, we apply for each detected region in the segmentation step, the fuzzy classification algorithm [5] in order to automatically recognize the shape [9, 10]. So, each geometric form can be automatically identified in the image and transmitted with a specific tactile code described in the legend.

2.4 Textures affectation and consideration tactile constraints

Once the different shapes and objects in the image have been identified, the next step is to fill textures for each shape. Studies have shown that blind and visually impaired persons have problems to distinguish a great number of textures in the same image. To overcome this problem, we limited the number of used textures at 5 [11]. We can manually choose the texture and objects to fill and also limit the number of textures. The shapes will be filled by the color black. Aiming to allow blind persons to easily distinguish these textures with their touch, we chose very different textures in terms of touch. Indeed, unlike the vision, the sense of touch is quite limited in the 2D exploration. With the vision, we can easily get a great amount of information in an immediate way, while the sense of touch is relatively limited. Indeed, the assimilation of information transmitted by touch is not so evident. The passing from a textural area to another is not easy to understand for a blind person. For this, we proposed to accentuate and enhance the contours of the different textural zones. The aim of this work is to offer a better exploration and a better understanding of tactile image for the blind and visually impaired persons.

In this work, the automatic extraction of contour is archived by the method of Sobel [12]. We also introduced the simplification of contours in order to facilitate the exploration by the blind persons.

2.5 Detection of the text zone and its transformation into Braille

In an educational image, there are not just shapes and objects. There is also text which describes the different objects of the image. So, in order to make a tactile image transcription, we must detect the text areas and recognize the text in order to translate it into Braille.

The used methodology to achieve this transcription is to locate firstly the different text zones. This is done by using the SWT (Stroke Width Transform) algorithm [13, 14]. This transform proposes to capture the main feature of the text (stroke width of the letters) and then to use this feature to distinguish the text areas from the non text areas. The developed system can easily and efficiently find the text areas even with the presence of various constraints like the size change or the direction change of the text, and also the use of several different fonts.

Then, we propose the second step which is the extraction of these founded text areas, this work is done by using the approach of Boris Epshtein [15]. The extracted text areas are then submitted to automatic characters recognition software which is the Tesseract OCR [16]. The choice of this OCR has several reasons; including its efficiency and the fact that it is open source software. After detection, extraction and text recognition, the last step is the Braille transformation of text. This involves translating each alphabet letter into Braille letter [17].

2.6 Arrow detection

Certain images contain some very special element like the arrows. This element has a very small size that it cannot be detected by the fuzzy method, but it returns very important information corresponding to the directional information which is essential for understanding the image. So our goal is to detect these arrows and transform them in bigger and clearer elements in order to make them more easily distinguished by blind people.

We propose to use the correlation method for the detection of this element. For this purpose, we use a template image and then calculate the correlation between the template image and the original image. The maximum of correlation correspond to the maximum similarity between the template and the image. The areas which have the max correlation value are the positions of the arrows. We choose to apply a rotation of the template by scanning all the possible angles. Then, we calculate the correlation for each rotated template in order to find the arrows of all the possible directions in the image. The same applies to the size of the arrow in order to detecting the arrows of different sizes. Thus, the areas related to the maximum correlation value correspond to the positions of the arrows. When the positions of arrows are detected, we propose to replace the original arrow by a bigger arrow, so the blind people can distinguish more easily the direction of this arrow.

2.7 Tactile exploration constraints

During this phase, we had to respect the various constraints related to the tactile image exploration. One of these constraints is to set a certain size and certain interline of the Braille text. The lack of space which there may be in this kind of images, involves the proposition of put the text in a "legend" card. So for every image, two tactile images are generated: the first image contains the various objects represented by different textures. Also, near each texture will be introduced a key that will refer to the Braille text that will be in the second "legend" card.

3. Experimental results

For our experiments, we develop a specific application which includes all the process steps described above. We apply this application on different educational images. The following figures present the results of the different application steps.

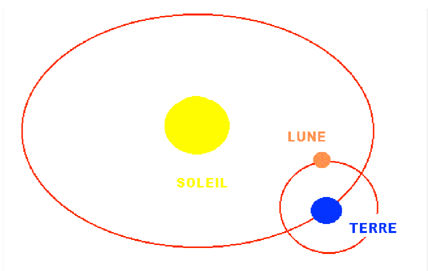

Fig. 5. Original image: the Sun-Earth-Moon system

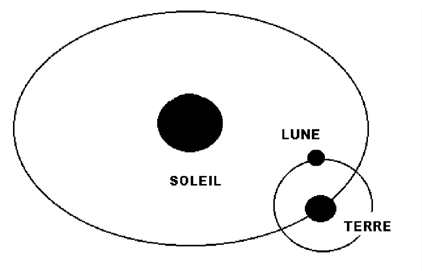

Fig. 6. Binary Image

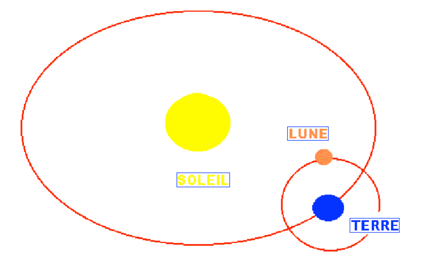

Fig. 7. Text location in the image

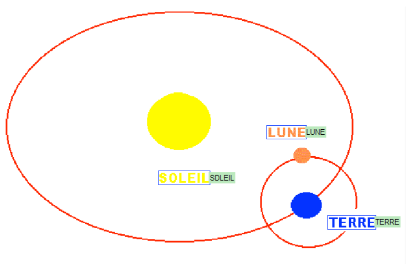

Fig. 8. Automatic Character Recognition

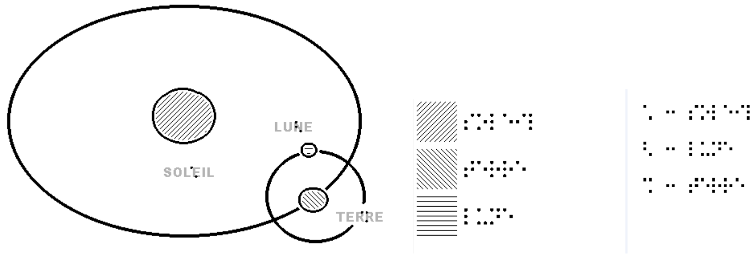

Fig. 9 . Tactile image: the Sun-Earth-Moon system and his legend in Braille

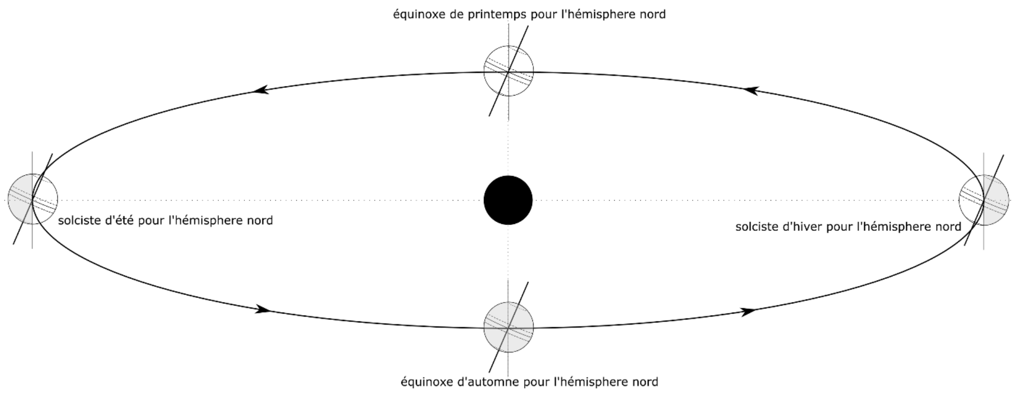

Fig. 10. Original image: the four seasons

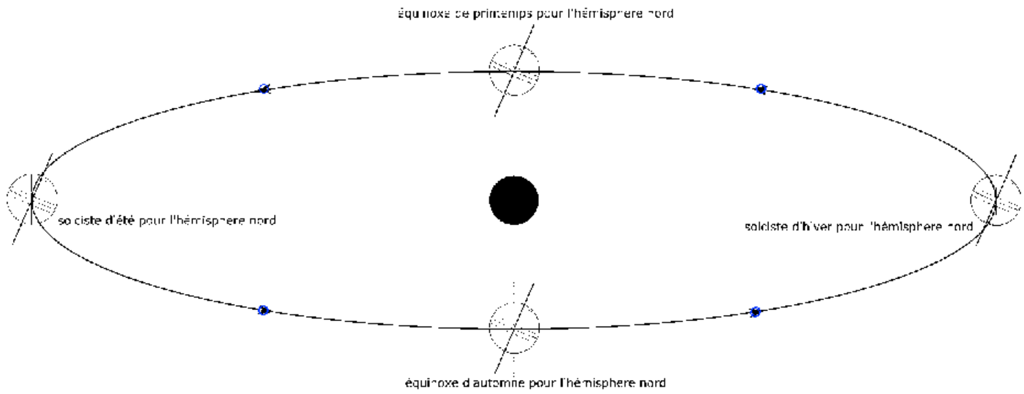

Fig. 11. Arrow position detection

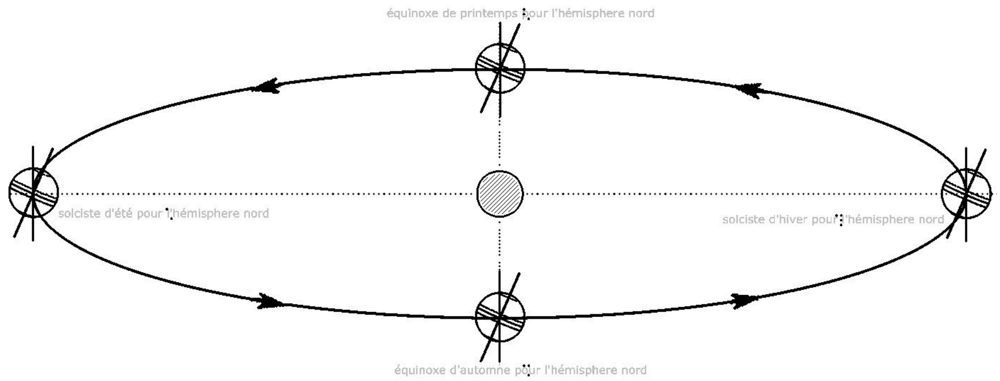

Fig.11. Tactile image: the four seasons

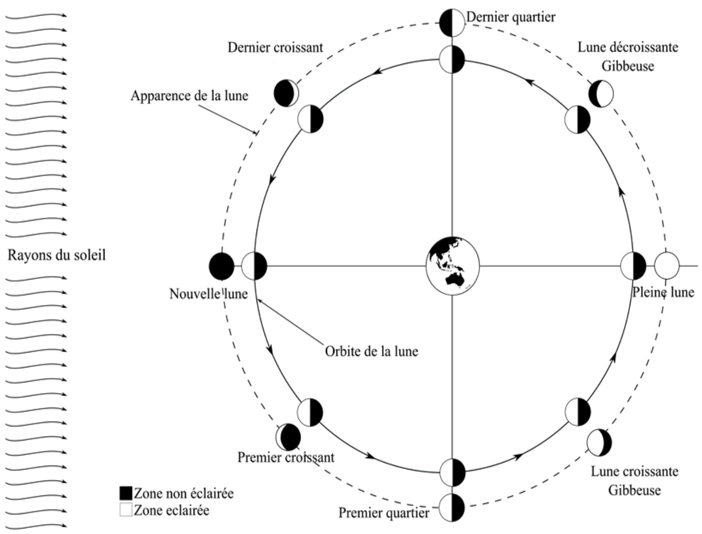

Fig.12. Original image: the moon cycle

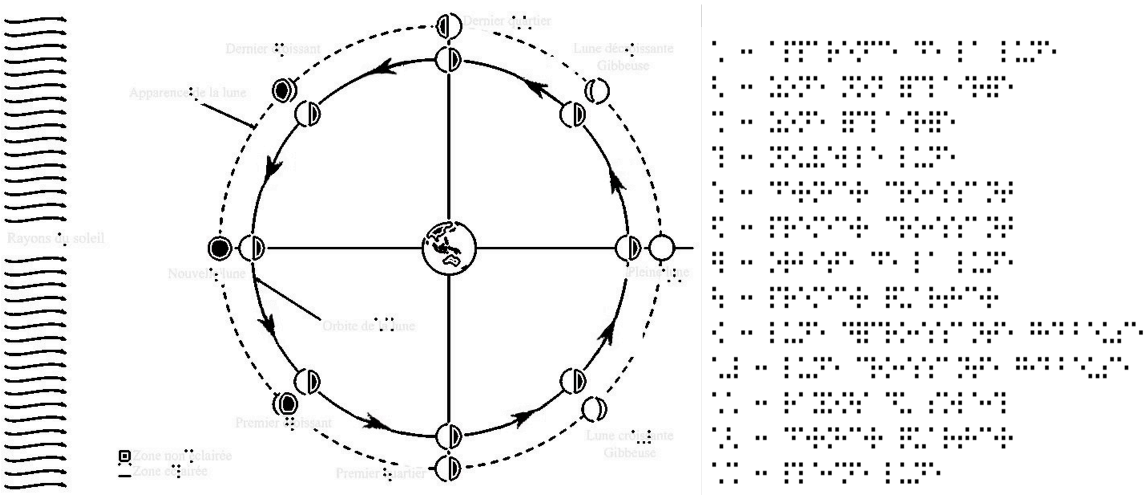

Fig. 13. Tactile image: the moon cycle and its legend in Braille

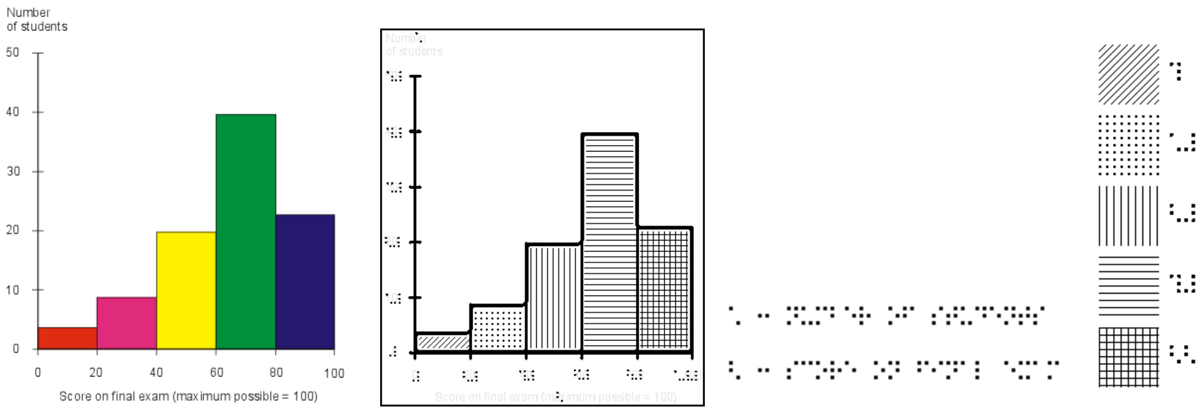

Fig. 14 Example of the histogram image

For few images set, we noted a few errors due to the limit of the used OCR. To remedy that, we implement a verification and correction step by the user. This is an easy way to correct the errors. The following figure illustrates this stage:

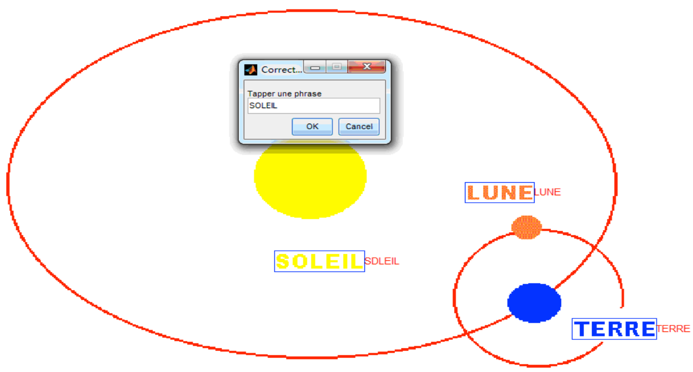

Fig. 15. The correction step

Conclusion

This paper presents a contribution towards an automatic transformation of pedagogic images into tactile images. So far, this process is purely manual, which represents several constraints related to the need of human and financial resources. We developed an application that allows to apply special textures to different image regions so they are well distinguishable by touch. This application also allows text transcription in Braille and automatic recognition of different geometric forms. For this purpose, we took into account some precautions related to the concerned disability. These precautions include the limitation of the number of textures used, the smoothing and enhancement of the different contours and the standardized size of the Braille font. It is true that in the presented work, the complete automation of the process does not give a success rate of 100% for all the tested images. However, to overcome this limitation, we propose to introduce user verification and validation. A preliminary evaluation by a blind student indicates that the proposed tactile transcription results are satisfactory.

As future work, we propose to complete our experimental study by collaborating with the blind or visually impaired peoples and to add specified dictionaries in the OCR. This work may have different applications, such as the accessibility in learning or the documentation access in general. In addition, we can use it to treat images from map libraries by generating toponymic databases in order to address and localize.

References

1. J. P. Fritz, K. E. Barner, "Design of a haptic visualisation system for people with visual impairments",IEEE Transactions on Rehabilitaion Engineering, pages 372–384,1999
2. N. Baptiste-Jessel, B. Tornil, B. Encelle,"Using SVG and a force feedback mouse to enable blind people to access 'Graphical' Web based documents".Computers Helping People with Special Needs (pp. 228-235). Springer Berlin Heidelberg, 2004V
3. T. P. Way, K. E. Barner, "Automatic visual to tactile translation II. Evaluation of the tactile image creation system". Rehabilitation Engineering, IEEE Transactions on, 5(1), 95-105, 1997
4. C. A. Bouman, "Connected Component Analysis", Digital Image Processing - January9, 2012
5. Manuel J. Fonseca, Joaquim A. Jorge, "Using Fuzzy Logic to Recognize Geometric Shapes Interactively". *Fuzzy Systems, 2000. FUZZ IEEE 2000.* The Ninth IEEE International Conference on 7-10 May 2000, San Antonio, TX, USA.
6. G. Nagalakshmi, S. Jyothi, "A Survey on Pattern Recognition using Fuzzy Clustering Approaches", *International Refereed Journal of Engineering and Science*, ISSN (Online) 2319-183X, (Print) 2319-1821, Volume 2, Issue 5(May 2013), PP.24-31
7. Sushmita Mitra, Sankar K. Pal, "Fuzzy sets in pattern recognition and machine intelligence", Fuzzy Sets and Systems, Volume 156, Issue 3, 16 December 2005, Pages 381 – 386
8. J. Lopez-Krahe, T. Alamo-Cantarero, E. Davila-Gonzalez, "Discrete line and circles: application to pattern recognition". ENST. D ISSN 0751-1345
9. J. Llados, E. Marti, J.Lopez-Krahe,"A Hough-based method for hatched pattern detection in maps and diagrams". ICDAR '99, Proceedings of the Fifth International Conference on Document Analysis and Recognition, pp. 479, 20-22 September 1999, Bangalore, India.
10. C. Harris, M. Stephens, "A Combined Corner and Edge Detector", Proceedings of The Fourth Alvey Vision Conference, pp. 147-151, 1988, Manchester, UK.

11. M. BRIS ," Recommandations pour la transcription de documents, documentation du service des documents adaptés pour les déficients visuels", INSHEA, Suresnes, 2004.
12. K. Engel, " Real-time volume graphics", pp. 112–114, 2006.
13. C. Jung, Q. Liu, J. Kim, "A stroke filter and its application for text localization", *PRL* vol 30(2), 2009.
14. A. Ikica, P. Peer,"SWT voting-based color reduction for text detection in natural scene images", *EURASIP Journal on Advances in Signal Processing*, 2013:95
15. B. Epshtein, E. Ofek, Y. Wexler, "Detecting Text in Natural Scenes with Stroke Width Transform", *Computer Vision and Pattern Recognition* (CVPR), 2010 IEEE Conference on 13-18 June 2010, San Francisco, CA,USA.
16. Kay, Anthony (July 2007). "Tesseract: an Open-Source Optical Character Recognition Engine". *Linux Journal*. Retrieved 28 September 2011.
17. "Code braille français uniformisé pour la transcription des textes imprimés", Deuxième édition, Septembre 2008.