Extracting Graphical Information from Arabic e-Documents for Visually Impaired People

Abdul Khader Jilani Saudagar

Abstract—Information graphics play a crucial role in a variety of multimodal documents, especially when the concern is about Arabic language which is widely used by many people in the world for communication. Alas, visually impaired people who utilize screen reader software's to steer through such electronic documents have restricted access to the graphics. This research work facilitates blind users to put on access to information graphics that materialize on web pages of Arabic websites and electronic documents. The interface is put into operation as a browser adjunct that instigate by a key hit combination. The result of this research work is a textual synopsis of the graphic, the basic content of which is the hypothesized message of the graphic creator. The textual synopsis of the graphic is then passing on to the end user by screen reader software. The carried work has the following benefits 1) not entails any action on the side of the website developer, and 2) giving the end user with the message that individual can gain the knowledge from viewing the graphics. Overall this work results in a system interface which is easy to handle and an effective way of communicating the informational content of graphics to the visually impaired people in Kingdom of Saudi Arabia.

Index Terms—Accessibility, Information Graphics, Multimodal Documents, Visually Impaired.

I. INTRODUCTION

Accessibility of electronic documents such as newspapers, reports etc refers to access information by all individuals, regardless of technological or physical discrimination [1]. Accessibility for web based interaction with blind users refers to making the content perceivable, operable, understandable, robust and easy to navigate and interact with. It is treated as a technical construct that allows the use of assistive technologies such as screen readers etc., in order to give necessary access to interface elements of a system. Due to their efficacy in illustrating quantitative data and the relations amid them, information graphics such as bar charts, line graphs are extensively used and these image appearances give stern information access ultimatums for visually impaired users. This work mainly focuses on Arabic information graphics from well-liked media and electronic documents which are accessed by worldwide web (WWW). The greater part of such graphics is aimed to convey a message. The primary goal of [2] was to develop mechanisms for providing visually impaired persons with alternative access to the content of information graphics. Especially, individuals with sight impairments have deficit access to multimodal documents.

The World Health Organization (WHO) estimated that there were 161 million persons worldwide with visual impairment, in the year 2002 [3, 4]. The majority of them resided in developing countries, including Saudi Arabia [5].

Various projects in the past have strived to make images accessible to visually impaired users by regenerating the image in an different medium, such as sound [6-8], touch [9] or a amalgamation of the two [6, 9]. Extended plan inference techniques[10] are deployed successfully in recognizing the intention of natural language utterances to the novel realm of information graphics. The research work done by Kurze [11] generates verbal description from text using the graphic properties to convey the content of a graphic by taking into consideration style of the graphics, data sets, labels on X-axis, Y-axis etc. Ferres et al. [12], presented the inspectGraph system, which has the aim of providing accessibility for blind users to the graphs published in Statistics Canada’s “The Daily.” Yu and Brewster [13] investigate the progress and utility of web-based multimodal graphs which use haptic devices and auditory output to convey the contents of the graphs to the users. In [14], the technical details and approaches are described thoroughly for the problem of generating a graphical element required for realizing the intended message of a graphic. Different systems [15-19] are developed for identifying the intended message of graphics. The text summary is also generated by successfully extracting the text caption from graphics and giving input to the screen reader. Image Processing is used for analyzing the graphic’s image file and producing an XML representation [20]. Many text extraction algorithms exists in the past and one such algorithm is discussed in [21]. A text extraction algorithm based on thinning [22-24] and dataset comparison technique is explained in [25]. Once the message has been inferred by the Bayesian inference system [26], each category of high-level goal is represented by a node in the tree, and instances of these goals appear as children with inhibitory links [27]. Each goal is broken down further into subtasks and a tree is built dynamically. An XML representation of the visual graphic [28] that specifies the components of the graphic such as the number of bars and the heights of each bar etc. A Bayesian system [29] for recognizing the intended message and content identification rules for line graphs is presented in [30].

Generally visually impaired users browse for particular items on the page using the TAB key [31]. The user interface of the system is implemented as a Browser Helper Object (BHO) [32] that runs inside the Microsoft Internet Explorer. A set of propositions that capture information that someone might extract by looking at a simple bar chart is proposed in [33]. A better overall experience for handling pie charts, grouped bar charts and composite graphics is explained in [34-37]. How to present bar charts using a tactile presentation...
is explored in [38]. A system [39] to collect the drawings from visually impaired people by finger movement on Braille devices or touch devices, such as tablets is proposed. When the drawing data is collected, the system will automatically generate the graphical XML data, which are easily accessed by applications and web services. The graphical XML data are stored locally or remotely. The work [40] suggests a method to extract implicit information of Bar chart, Pie chart, Line chart and math’s graph components of an electronic document and present them to vision impaired users in audio format. The proposed frame work methodology is efficient and faster when compared to the earlier work [41, 42]. In addition, the lack of guidance as to how information graphics is properly presented to visually impaired users is still a limitation [43].

Apart from this, the above discussed systems are only developed for graphics in English electronic media and as such no work has done for the graphics in Arabic documents which is most widely used language in the Arabian Peninsula. This research is independent with respect to others in order to achieve the goals. The overall aim is to generate a succinct coherent summary of a graphic contained in Arabic electronic documents which can hypothesize to the visually blind people through JAWS screen reader software.

II. METHODOLOGY

The following are the major phases outlined to achieve the research objectives. The first phase entails browser extension which let visually impaired users to access textual precis of information graphics. The second phase entails implementation issues that come across in research work. Image Processing Module (IPM) is the third phase and the next phase concerns with generating the summary and the last phase guidelines preparation. All steps presented below are inter-related that will lead to the design and implementation of proposed system.

A. Phase I: The Browser Extension

Since the research work mainly focuses visually impaired people, the users must be able to launch the application from the keyboard. Intended for browser extension the research work is carried on distinctively for JAWS 14.0 and Internet Explorer 9.0 or later. When finding the way through a web page, JAWS clients have many alternatives. So other than the existing navigation commands in JAWS, CONTROL+Z is selected as the key coalition for launching the project interface. However, this work requires a very taut coupling flanked by the system and the web browser, because the system needs to determine the currently focused graphic in the web browser. The apt level of integration can be attained by implementing the user interface as a Browser Helper Object for Internet Explorer. If the item in focus emerges to be a graphic containing a bar chart, the system then attempts to surmise the intended message of the bar chart. The resultant textual summary is presented to the user in a new window. The text in that new-fangled window is then read to the user by JAWS as depicted in Fig. 1.

B. Phase II: Implementation Issues

In this work the BHO is operating inside Internet Explorer, and in order to avoid the confusion that the user may try to launch the system to process a graphic that is not a bar chart. As a result, when the CONTROL+Z keystroke is detected, the BHO runs a pithy image processing algorithm to determine whether the chosen graphic has particular attributes that identify it as a possible bar chart, such as whether the graphic has 20 or smaller number of gray levels, and with a common commencement row or column. In case, if the graphic does not resemble to be a bar chart, the message

 "$\text{"في الرسم المحدد لا يظهر انه مخطط شريطي" "} "$\text{"}"

"The focussed graphic does not appear to be a bar chart," is read to the user by JAWS.

Users hits CONTROL+Z upon JAWS arrive at the graphic, the BHO is activated. If the graphic materializes to be a bar chart, the BHO will insert that graphic into the tab order of the page, and will append

 "$\text{"} "$\text{"}"

"This focused graphic appears to be a bar chart” to the alt text.

C. Phase III: Processing the Image

The Image Processing Module IPM is responsible for analyzing the graphic’s image[20, 26] file and producing an XML representation containing information about the components of the information graphic. The processing of an image can be done as follows which is depicted in Fig. 2.

The purpose of image segmentation is to bunch pixels into salient image regions, i.e., regions analogous to individual surfaces, objects, or natural parts of objects.
In the text extraction and recognition step each individual segment received from step 1, is tested for text regions and object regions. Only the shots with text regions are considered further for text extraction. In this step we will start reading the pixel from the initial position \((x_1, y_1)\) for the RGB values=(0,0,0) and copy those pixels and create another image which consists of only text with white background. And this process is repeated for every segment which has text region. We use the following data sets as shown in figure 3 for Arabic text extraction and recognition. Arabic character set has 28 characters and each character is written in any one of the four forms such as standalone, initial, medial and terminal. The blank fields for the characters in Fig. 3 (a) represents those forms are not available for that characters. Few characters can be omitted due to similarity and the character set used in this implementation is as shown in figure 3 (b). This work uses the same methodology as used in [25] for text extraction and recognition process.

In the next phase the shots with objects are considered and based on their boundary information, color etc., the bars in the graphic are compared and the necessary information is stored in a text file. The system identifies a number of ways in which a unit in a bar chart becomes salient. For a bar chart, the depiction includes the number of bars, the tags of the axes, and facts associated with every bar in the graph for example the label, altitude, color of the bar, and so forth [2]. This work currently handles only electronic images formed with a prearranged set of fonts and no overlapping characters.

### D. Phase IV: Generating the Summary

Image Processing Module results is an XML file that describes the graphic and all of its components. The XML file is analyzed to build the appropriate Bayesian network. Each category of high-level goal is represented by a node in the network as shown in Fig. 4 and instances of these goals appear as children with inhibitory links capturing their mutual exclusivity. Each goal is broken down further dynamically into subtasks in order to append the nodes to the network.

![Fig. 4 Bayesian Network](image)

### E. Phase V: Guidelines Preparation

Image Processing Module (IPM) extracts the text and based on the appearance of the text “1 الكل” the extracted text is added under the node “caption” in the Bayesian network, followed by the text extracted horizontally is added under the X-axis node and vertically extracted text is added under the node of Y-axis. The IPM module process the graph and based on the different colors appearing adjacent to one another count the number of bars, and for each bar its height, hexadecimal color code and title appearing adjacent to it and constructs the Bayesian Network. Conditional probability tables in Bayesian Network capture the probability of each of the values of a child node given the value of its parent(s). A spreadsheet is constructed based on the Depth First Search (DFS) technique and record the values into the spreadsheet as shown in Table 1. This spread sheet can be given as an input to one of the screen reader software’s so that the visually impaired people understand the intended message in the graphic.

### III. RESULTS

The research work is conducted on 50 randomly selected graphics from electronic media. This work is implemented in Java using NetBeans IDE 8.0.1 and mySQL database. Intel i5 x 32 processor with 8 GB RAM and 1 TB hard disk is used for testing purpose. The input image and dataset samples are in jpeg format. For example, if the user bumps into the graphic shown in Fig. 5, the user thumps CONTROL+Z and the summary of the graphic come into sight in a new window. For this picky graphic, this work generates the summary as shown in Table 1.
IV. CONCLUSION

The developed framework has the following benefits:

- High user satisfaction.
- Less complexity.
- Natural Language Processing not required.
- The proposed framework is efficient and faster when compared to the earlier work [41, 42].

This work can be extended to the other types of charts like, pie chart, line chart, column chart, area chart etc.

ACKNOWLEDGMENT

I are very much thankful to Al Imam Mohammad Ibn Saud Islamic University for providing an environment to carry out this research successfully. This research is financially supported by Scientific Research Deanship, Al Imam Mohammad Ibn Saud Islamic University, Riyadh, Saudi Arabia with fund number 350912.
A. K. J. Saudagar received his Bachelor of Engineering B.E, Master of Technology M.Tech and Doctor of Philosophy PhD in Computer Science & Engineering in 2001, 2006 and 2010 respectively. His area of interests are: Artificial Image Processing, E-Commerce and Information Technology.

He has 6 years of teaching experience at both undergraduate (UG) and postgraduate (PG) level and presently working as a Assistant Professor in Department of Information Systems, College of Computer & Information Sciences, Al-Imam Muhammad Ibn Saud Islamic University , Riyadh, The Kingdom of Saudi Arabia.

Dr. Saudagar has published a number of research papers in National, International Conferences and International Journals. He is associated as member with various professional bodies like IACSIT, IAENG, ISTE etc and working as Editorial Board member, Reviewer for many international Journals.