# An Image Processing Based Visual Compensation System for Vision Defects

Chin-Lun Lai Communication Engineering Department Oriental Institute of Technology Taipei, Taiwan fo001@mail.oit.edu.tw

Shu-Wen Chang Ophthalmology Department Far Eastern Memorial Hospital Taipei, Taiwan swchang@mail.femh.org.tw

***Abstract—This paper is intended to propose a userfriendly, portable, and integrated visual function aid system that can cooperate multiple visual compensation subsystems for vision defects (in a broad sense, referring to people with color blindness, low vision, visual field defects, … etc.). With the proposed real-time image processing algorithms, the system is able to detect and construct the user's perceptual model, and then meet their vision requirements by compensating the input visual information thus to fit their vision models. Experimental results show that the proposed system structure and algorithms deed help those people with vision defects in communicating with the external world, including enhancing the reading speed as well as perceiving the image detailed. Thus obtain a practical and affordable system to improve their life quality.***

***Keywords- Assistive system, visual perception compensation,vision model construction , low vision, colorblind, Image processing***

## I. INTRODUCTION

For people get heritage, illness or weakness because of ageing or injured, it comes out the lost of vision sense functions and thus make them communicate with outside world difficultly. It is meaningful and encouraged to develop assistive tools to let these people live as normal ones. Among these auxiliary systems, the common ones adopt the computer based CCTV structure to help people perceiving the image content much easily and fast[1]-[5], while some of them use cladding or other techniques to modulate the spectra of the incident light to help people with colorblind or color weakness to recognize the correct colors[6]-[11]. Although there exist specific solutions to help these disabled people with various conditions (ex. colorblind, low vision, or other weakness), an efficient, effective, and practical one is not available until now. In this paper, an image processing based portable assistant is proposed to integrate multiple assistive functions for the visually impaired to enhance/recover their visual perception, thus help them to live with better life quality. In the first research phase, two assistive functions, for chromopia and low vision, are integrated into the proposed aid unit. This paper is organized as follows. In section II the system structure and principals are described. The experimental methods and results are given in section III, while the conclusions and further works are given in the last section

In order to achieve the goal of retrieving the disabled visual functions, visual condition of the defected perception should be known first. Then the defects can be made up via using the auxiliary tools. Thus, the proposed assistive system functions can be divided into two parts, named as the visual model measurement and the information compensation processing core. That is, by adopting the visual perception measurement function, personal perception condition of the assistive system user can be measured and modeled before usual use. After that, the captured video stream can be adjusted adequately according to the model parameters, follow by projecting to the HMD (head mounted display) device to replace the original images thus compensate the user's defected visual functions. Figures 1 shows the structure of the proposed assistant system while the corresponding function flowchart is given in figure 2. The proposed system comprises a HMD, a mini laptop, and two web cameras mounted on the HMD. The whole processing process is described as follows. First, simple experiment test is designed to evaluate the user's perception in color and contrast discrimination. The series tests are undergone in the laptop platform and will project specific image patterns onto the HMD. When viewing the patterns, user can interact with the system by keypad to respond his visual perception. Once the measurement process is completed, the personal visual model can be built according to the response parameters. Two function flowcharts relative to color discrimination and contrast sensitivity are shown in figures 3 and 4 respectively. Once the impaired vision model is known, the disabled functions, theoretically, can be compensated by simply adopting the inverse-like functions. That is, the input image sequence, captured by the web cameras, is not presented directly to the user via HMD, but modified first by the compensation process core and then redirect to the HMD to replace the original one such that the user's perceptual function can be enhanced. The compensation processes include linear/nonlinear image scaling, edge and contrast enhancement, color coordinate transformation, and histogram modifications, and the compensation process is described as follows.

Compensation for color perception can be used to address the following two categories: monochromatic color blindness (red-green color recognition obstacle, yellow-blue color recognition obstacle) and weak color level (red,

green and blue) identification, and is depicted in figure 5. Patients with redgreen or yellow-blue color blindness will mistake different colors for the same one. Therefore, we transfer, through color level conversion, different colors to the color gamut that patient can identify. At first, we convert RGB colors into HSV colors for color gamut transfer and separate adjustment of hues and saturation as described as follows

$$' HHH = + \Delta ; ' SSS = + \Delta ; ' VV = \quad (1)$$

where $' H$ , $' S$ , and $' V$ indicate the corrected hue, saturation, and brightness values to replace the original mistaken color's attributes HSV respectively, and $H \Delta / S \Delta$ is the hue/saturation displacement value. The correction of saturation is summarized as follows: As there are high-saturation colors, mid-saturation colors and low-saturation colors, we reserve the high-saturation area and use the mid-saturation area for replacement with the originally mistaken saturation to help users identify the difference. On the other hand, since every individual has different color discrimination abilities. With the visual model measurement system discussed above, we may obtain the subject's ability to discriminate colors and, through grayscale adjustments, enable color levels of input images match the measurement results. In addition, as brightness affects grayscale discrimination, the system will adjust brightness information when the image is either excessively dark or bright to enhance color discrimination as described in the following equation

$$' HH = ; 'SSS = +\Delta ; ' V V V = +\Delta = (2)$$ where $VV +\Delta$ indicates the brightness correction value through histogram equalization (the adjustment will be the most evident as it is excessively dark or bright) which is defined as where max V is the maximum brightness level, j n is the j-th cumulative frequency and n is the total frequencies. Figures 6-7 show the correction examples for color blind and weak color perception defects respectively.

.2 Compensation for Low Vision Patients suffer from low vision (subnormal vision) mainly because they lose the ability to discriminate highfrequency information of images (i.e. the ability of image detail discrimination). There are two kinds of solutions. One is to transfer image detail signals, by image enlargement, to a frequency area where the patient's vision is more sensitive; the other, through enhancement of the said image contrast, to make the patient more likely detect image details. Therefore, this system provides image enlargement and image

## II. LITERATURE REVIEW

Several investigations on summarization of documents and image processing were done separately throughout years. But a limited amount of research were done to in order to help visually impaired students.

### *A. Summarization tools*

For text summarization a huge number of researches were done successfully in the past. Among them, the research done by SRA International and Department of Defense in the United States of America [1], they describe a trainable and scalable summarization system which utilizes features derived from information retrieval, information extraction, and NLP techniques and on-line resources.

Eduard Hovy and Chin-Yew Lin of Information Sciences Institute of the University of Southern California [2], implemented a summarization tool named "SUMMARIST" to create summaries of arbitrary text in English and selected other languages. It was included language-specific techniques of parsing and semantic analysis, and was combined robust Natural Language Processing with symbolic world knowledge, derived from WordNet.

The survey done by Vishal Gupta of Computer Science & Engineering department, in Panjab University Chandigarh, India [3], presents a number of techniques used for text summarization and extraction. Also, he has pointed out the main functionalities of some successive summarizers, such as, Newsblaster, Automatic text summarization system [4], The trainable document summarizer [5] and the ANES [6] text summarization system.

### *B. Diagram Extraction*

So many researches were carried out for image extraction and description, but very few were done aiming visually impaired students. And some researches were conducted to identify and describe flow charts.

On-line handwritten flowchart recognition, beautification, and editing system [7] was conducted in order to identify handwritten flow charts. Flow chart symbol identification is done using loops. Shapes are identified and inside details are extracted. As this research is done for handwritten flow charts, they have modified the flow charts using beautification tool.

The research, A sign reading system for the visually impaired [8], done by American University of Sharjah, that aims at helping the visually impaired by locating indoor signs and reading their content out loud, thus guiding them toward their destination. The process is to capture the signs in public places using a camera and to extract the text and to process the image and to read it aloud in order to give directions.

Text/Image Region Separation for Document Layout Detection of Old Document Images using Non-linear Diffusion and Level Set, research [9] discusses how to identify images out of a document and to extract the features of the image. Profiling

or morphological operations are used to separate diagrams from text. They have introduced a nonlinear diffusion method known as edge enhancement diffusion (EED). The result was that any diagram was separated from a document successfully.

The research, Textline detection in degraded historical images [10], which was conducted by EURASIP Journal on Image and Video Processing has been carried out to extract text from degraded images. Textline detection and binarization were the main techniques used.

*C. Graph information extraction tool based on the image*

According to the past literature reviews related to the graph information extraction, got to know that some components are not that much implemented.

"Graph-based representations and techniques for image processing and image analysis" [11] is one of the research that is based on graph extraction, they have mentioned more theoretical techniques than applying practically. They explained the graph partitioning greedy algorithm for colour image segmentation. Image segmentation is the process of partitioning an image into a set of non-intersecting regions such that each region is homogeneous and the transition from one region to another is sharp. And also it has described a novel fusion of color-based segmentation and depth from stereo that yields a graph representing every object in the scene. Like that all the theoretical parts have mentioned in this research. But the way of extracting the graph information and storing the information has not been mentioned practically. "Graphs for image processing, analysis and pattern recognition" [12] also another thesis that describe the theory behind the graph information extraction. It has described single graph methods like segmentation or labeling graph-cuts, graphs pattern recognition. Under graph matching it has described the graph or sub graph isomorphism, error tolerant graph matching and more. But it has not been mentioned the practical situation.

"Extraction of graph information based on image contents and the use of ontology" [13] is one of the research based on the graph information extraction. It describes the way of analyzing X, Y axis and the more attention to extracting information from the graph.

*D. AI Chatbot*

AI Chatbot will try to understand the query and provide a definitive answer. There will be four main units to the system working together to understand the question and return an appropriate answer: Generic question construction - capable of taking a natural language question and making it more generic. Generic answer construction [14] - capable of taking a generic question template and providing a generic answer template. Generic answer population - capable of taking a generic answer template and populating it with information from the database to form an answer [15]. Information extraction - capable of finding information through structured or unstructured websites, and storing that information in a database.

## III. METHODOLOGY

Proposed solution addresses a tool to extract and summarize the research methodology for the visually impaired students. This project has very important research areas like, Natural Language Processing, Machine learning, Image Processing, Text to Speech conversion and web application development. This research is conducted on the abovementioned research areas to achieve the project objectives.

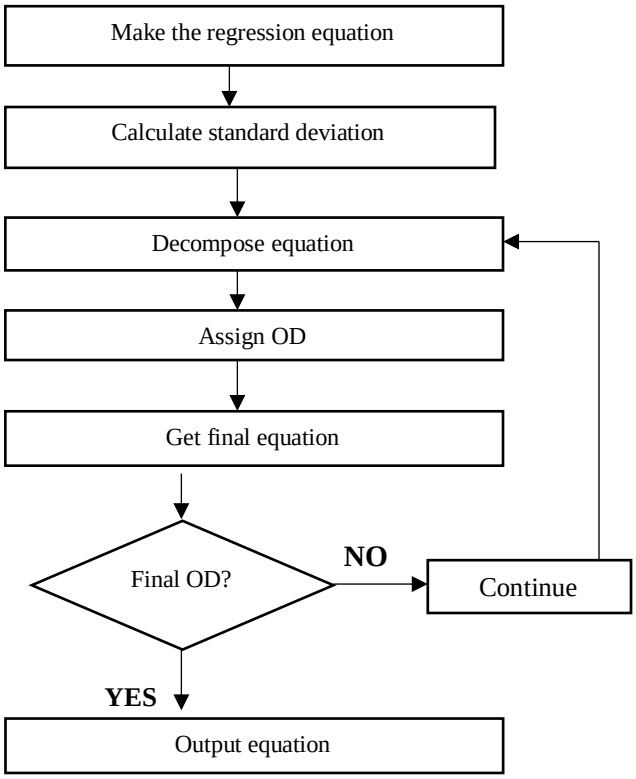



*Figure 2: system diagram*

*1.Text Summarization*

In a computer, a text is represented as a string of characters. The first step in processing natural language is to break up this character string into units that will provide the basis for all our Natural Language Processing by Tokenizing. Grammatical words such as prepositions and determinants do not make much difference for the meaning in the text. These are also called stop words and are usually removed in tasks such as information retrieval.

Secondly, we may want to consider two or more words as being one unit since they have the same root or stem. For

example, ‘descendant’ and ‘descending’ can be assimilated to ‘descend’. Combining words having the same root into one stem is called ‘Stemming’. The remaining words after removing the stop words are stemmed using the well-known and widely used Porter Stemmer. The stemming output will be gained as a sentence matrix.

Finally, the tool generates a summary based on the rules derived from supervised machine learning algorithms which are, Sci-Kit, Pagerank and Textrank algorithms. The text summarization approach also has a polynomial time complexity which returns the output at a reasonable time.

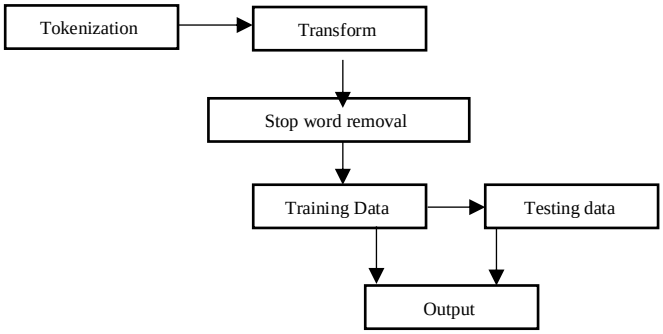



*Fig. 3. Image processing flow chart*

*2. Flow chart extraction and Describing tool*

As the input of the system is a pdf document, in order to extract images out of it, we have used pdf2image library files to convert pdf to image format. Image processing techniques and library files have been used to identify and extract flow charts out of the document.

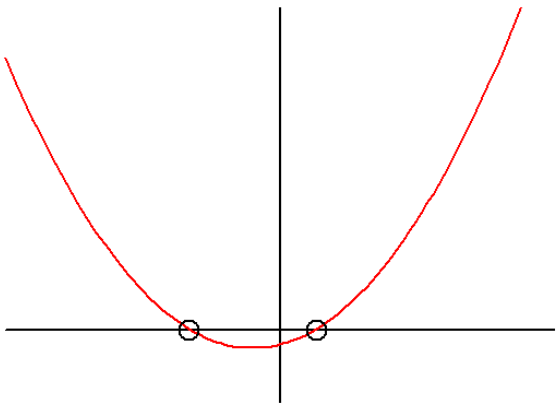
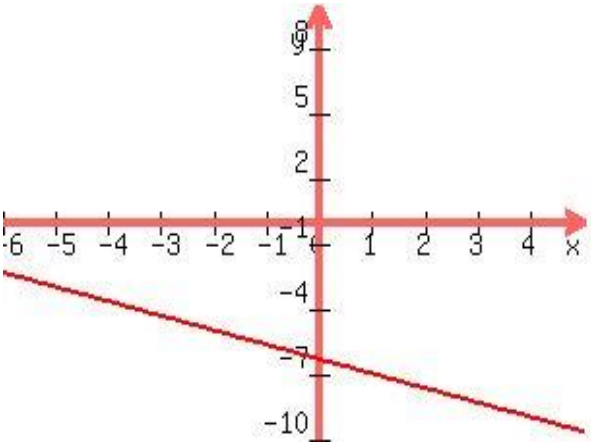


*Fig. 3. Analysis of flow charts*

*3. Graph extraction and summarization*

The PDF document is the input for the system. Before the identification of graphs, the system extracts images from the PDF as PNG files, whether they are referenced by pages or not. In order to do that, we have used MuPDF 1.16.0 library with PyMuPDF 1.16.0 Python bindings. After that system identifies




the Cartesian graph images out of all the extracted images and extract the category of the graph and other relevant information. For this task, we have used trained advanced algorithm to recognize any Cartesian graph patterns by using the Convolution Neural Network (CNN) in deep learning. Using PyPDF2 1.26.0 library, the system will extract the other information related to the Cartesian graph. The final output will generate as a PDF report, by including all the above-extracted graph information.

Fig.4: Graph Analysis

## IV. RESULTS AND DISCUSSION

*A.Text summarization tool*

The design of a tool of the text summarizer is of great importance in the current world which is so filled with data. It would reduce the pain visually impaired people suffer when using e-documents in their day to day life. The text summarization part was developed based on tokenization, stopword removal, sentence segmentation, stemming and vectorization. It generates a summary based on the rules derived from a supervised machine learning algorithm, The Sci-Kit algorithm. The text summarization approach also has a polynomial time complexity which returns the output at a reasonable time.

In order to receive a precise summary, the Textrank algorithm is mentioned when developing the text summarizing tool. Here, we have used some special features and gained results, such as, removal of punctuation and special characters, create a word vector each with size 100 of each sentence, check similarity between sentences and making a word graph by applying Pagerank algorithm.

Table I shows the accuracy, precision and recall of the text summarization at different approaches in percentage.

*B. Flow Chart Extraction and Describing Tool*

This section of the research produces an application to identify flow charts out of a research paper related to Information Technology and to extract the content of the flow chart. Basically, image processing techniques are used in order to accomplish this task. First as a pdf document is given as the input to the system, this pdf is broken down as pages and these pages are converted into images using pdf2image library files. As the second step, these images are processed, and flow chart diagrams are identified.

*Fig. 3. Identification of flow charts*

Next, these identified images are extracted and processed to get the content of them.

The content of the image is extracted in the order of the flow chart using pytesseract library files. Recurrent Neural Network(RNN) algorithms are used for text extraction and identification. Dependent and independent nodes are identified and extracted in the order of the flow chart. Decision nodes are identified separately and mentioned in the output document. The output will be a pdf document with the information of the flow chart.

By this section of the system, a visually impaired student will be able to get to know whether there are flow charts in the research paper. If there are any flow charts they will be extracted, and the content will be given as that the visually impaired student will be able to imagine the flow chart through a pdf document. The student will be able to imagine the flow chart as the details are given in the exact order of the flow chart with the decision nodes.

Even if the visually impaired students get the chance to read a research paper using other tools, the student will not be able to identify flow charts and will not be able to get to know about what is written in the flow chart. Therefore, this section of the system will be a great help for the visually impaired students to do their studies.

*C. AI Chatbot*

The design of an AI Chatbot to interact with the user to simulate human conversation via text or audio messages to communicate with the application. A Generative based Conversational AI model will be used for the purpose of question answering and navigation throughout the web application. The model will take the user input and understand the intentions of the user's message , determine what type of response message is required, and follow correct grammatical and lexical rules while forming the response.

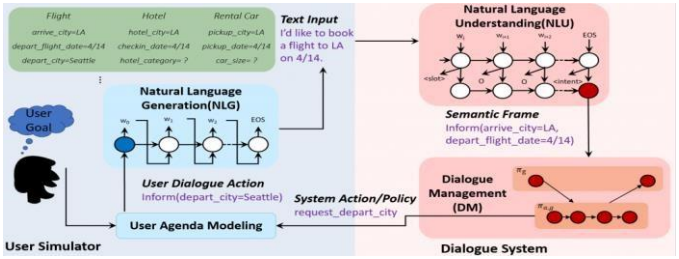



*Fig. 6. The figure shows how Deep Learning based chatbot work internally.*

*D. Graph information extraction tool based on the image*

A graph is an effective form of data representation used to summarize complex information. Understanding the graph details means an important task. When we speak about visually handicapped students, graph information extraction become a dream. Because without seeing the graph image, it is very difficult to identify the details behind the graph. In this section represent fully descriptive information related to the "extraction of the graph information" based on image contents.

Initially system will identify the curve graph images in the given pdf by using graph-based image segmentation algorithm. After that it will automatically convert the graph image into binary image by using thresholding .It will extract all the X, Y coordinates (use OCR and the mask creator to identify the coordinates) related to each curve graphs and listed in a separate excel sheet by using "xlswrite ()" function. It will process all the extracted information and finally give a complete explanation about the behavior of the graph with important details (ex: - minimum, maximum, slope, intercept, etc.)

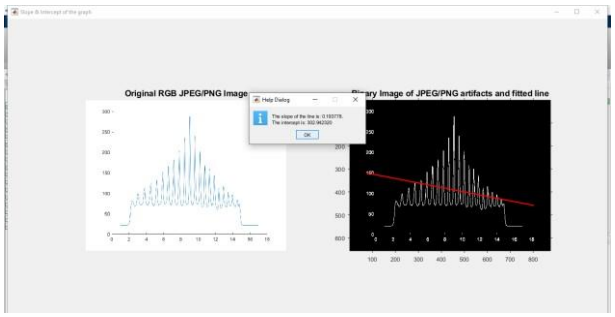



*Fig. 7. Original image to binary image conversion and identify the slope and the intercept.*

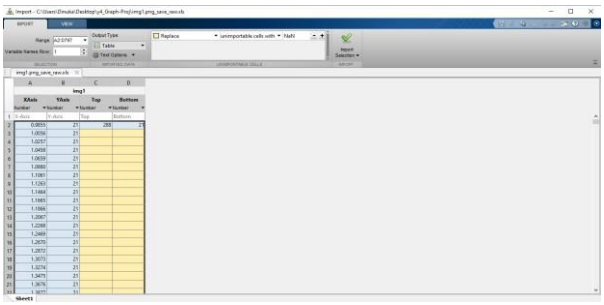


*Fig. 8. Extracted graph information.*

An identifying graph information is also become a huge task for not blind people. So final output become voice converted as well as the summarized documentation with including all the summarized details about the graphs. Therefore, people can easily get the idea about each and every graphs in detailed manner.

## V. CONCLUSION AND FUTURE DIMENSIONS

### *A.Conclusion*

Outcome of this research produce an application to summarize research papers was built mainly focusing on visually impaired students. The whole solution as a system provides an application to select a research paper and get the summarization of it with a description about flow charts and graphs if present. This task will be done within a few minutes and with a high accuracy.

The summarization will be presented in voice and as a document which contains the summarized information in point form. Visually impaired students will be able to use this as a

mobile application and a desktop application.

In addition, this application can be used by ordinary students to get a summary of a research paper very easily within few minutes without reading the paper.

### *B.Future Dimensions*

Currently system is developed to summarize the methodology, but the authors hope to develop the application to summarize the whole research paper.

This application summarizes only the flow charts, it can be developed to identify and summarize all other types of diagrams related to the field of IT.

Currently system efficiently summarizes only research articles related to Information Technology. We expect to expand the range of this tool to summarize research articles related to any field.

## ACKNOWLEDGMENT

The authors would gratefully acknowledge Prof. Samantha Thelijjagoda, who lead us by giving suggestions, encouraging, and helping in coordinating the research activities. Furthermore, we thank Mr. Upul Weerasinghe (Assistant Director of Central Bank, Colombo) and Mr. Sanka who provided us with information on the needs of visually impaired students by giving us necessary ideas. We would appreciate the guidance given by the panel of lecturers who gave us their valuable ideas and time.

## REFERENCES

[1] Chinatsu Aone, Mary Ellen Okurowski, James Gorlinsky, ”Trainable, scalable summarization using robust NLP and machine learning” [Online] Available: https://aclanthology.info/pdf/C/C98/C98-1009.pdf. [Accessed: 13- Feb- 2019].

[2] ]Eduard Hovy , Chin-Yew Lin, Automated text summarization in SUMMARIST [Online] Available: https://apps.dtic.mil/dtic/tr/fulltext/u2/a631227.pdf. [Accessed: 14-Feb-2019].

[3] Vishal Gupta, A survey of text summarization extractive techniques, 2010. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.348.3840&rep = rep1&type=pdf#page=104. [Accessed: 13-Feb- 2019].

[4] H. P. Edmundson.,” New methods in automatic extracting”, Journal of the ACM, 16(2):264-285, April 1969.

[5] J. Kupiec, J. Pedersen, and F. Chen, “A trainable document summarizer”, In Proceedings of the 18th ACMSIGIR Conference, pages 68-73, 1995

[6] Ronald Brandow, Karl Mitze, and Lisa F. Rau. “Automatic condensation of electronic publications by sentence selection. Information Processing and Management”, 31(5):675-685,1995.

[7] Hidetoshi Miyao and Rei Maruyama Faculty of Engineering, Shinshu University, Japan, ”On-line handwritten flowchart recognition, beautification, and editing system”.

[8] Mahmoud Younes, Basma AlMoshaigih, Michel Pasquier, Ghassan Qadah, “A sign reading system for the visually impaired”, International Conference on Frontiers in Handwriting Recognition,2012.

[9] Sachin Kumar Sa *, Parvathy Rajendrana , Prabaharan Pb , K P Somana, Text/Image Region Separation for Document Layout Detection of Old Document Images using Non-linear Diffusion and Level Set [online]

Available:https://ac.els-cdn.com/S1877050916314776/1-s2.0S1877050916314776-main.pdf?_tid=fb1ca9ae-6f82-4172-8241b9197b8cd4d5&acdnat=1551705164_01be6405467f516266fbe755c6dce5 36.[Accessed:4-March-2019].

EURASIP Journal on Image and Video Processing2017**2017**:82,

[10] "Textline detection in degraded historical document images" [online] Available:https://jivpeurasipjournals.springeropen.com/articles/10.1186/s13640-017-0229-7.
[Accessed: 4-March-2019].

[11] Graph-based representations and techniques for image processing and image analysis. A. Sanfeliua, R. Alquézarb, J. Andradea, J. Climentc, F. Serratosad and J.Vergésa. [Online] Available:
https://pdfs.semanticscholar.org/2a15/d3f6127b8ff8d0a301b371641728e3178967.pdf [Accessed: 17- Feb- 2019].

[12] Graphs for image processing, analysis and pattern recognition.
FlorenceTupinFlorence.Tupin@telecom-paristech.fr http://perso.telecom-paristech.fr/˜tupin.. [Online] Available:
https://perso.telecom-paristech.fr/tupin/cours/AIC/graphes.pdf
[Accessed: 17-Feb-2019].

[13] Extraction of graph information based on image contents and the use of ontology. Sarunya Kanjanawattana1 and Masaomi Kimura2 1Graduate School 2 Information Science and Engineering Shibaura Institute of Technology, 3-5-7 Koto-ku Toyosu, Tokyo 135-8548, Japan[Online] Available: https://files.eric.ed.gov/fulltext/ED571596.pdf . [Accessed: 17-Feb-2019].

[14] Bhagwat, Vyas Ajay, "Deep Learning for Chatbots" (2018).Master's Master's Projects. 630
[Online]Available:https://scholarworks.sjsu.edu/etd_projects/630/. [Accessed: 30- Mar-2019].

[15] Ravi, R. (2018). Intelligent Chatbot for Easy Web-Analytics Insights. 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI). [Online] Available:
https://ieeexplore.ieee.org/document/4462971 [Accessed: 30 Apr-2019].

[16] Wesley T. Chuang , Jihoon Yang, Extracting sentence segments for text summarization: a machine learning approach. [Online]Available: https://wtlab.um.ac.ir/images/elibrary/text_summarization/Extractin%20Sentence%20Segments%20for%20Text%20Summarization.pdf.
[Accessed: 15-Feb-2019]