

# An Interactive Smart Glass

Srinivas M<sup>1</sup>, Gangaram D<sup>2</sup>, Varun Reddy R<sup>3</sup>, AtulKewat K<sup>4</sup>

UG students, Dept. of CSE, RRCE, Bengaluru

Dr. UshaSakthivel<sup>5</sup>, HOD& Prof., Dept. of CSE, RRCE, Bengaluru

**ABSTRACT**-Our research goal is to make the world smarter in day to day life. Nowadays, everything in the world is getting smarter i.e. for ex: smart phone, smart TV etc. Our project is mainly based on this concept, nothing but an interactive smart glass. It implies several constraints. First, conversion of image to text and produces output by voice. Second, Object recognition by using IR rays. Third, Enhancement of old document and detect the content present in it. Fourth, Recognition of ancient or hand written text content in the document. Fifth, Navigating the location to the user by GPS mode and displaying it on the screen. All the above constraint can be fulfilled by using Text detection, Enhancement process, Object recognition method embedded in normal glass with android functionality. And Hence, we all this glass as an Interactive Smart Glass where User can interact with it by providing input to it.

**Keywords:**Image Segmentation, Optical character recognition(OCR), Text detection, Object recognition, Thresholding.

## I. INTRODUCTION

Today, there are about 2-3% people of world population are blind and low vision paired people. We all know that blind people have own script language known as Braille language, which is slightly complex to learn. If a person has low vision or who loses his eye sight due to some accident might feel bit difficult to learn the Braille language. So in order to overcome difficulties faced by this special people we have researched and proposed a model known as an Interactive Smart Glass, where user can read the document by capturing the image and detecting the text content in it and output is produced through speech to the user. Secondly, User can also recognize the object present in front of camera(which is present in smart glass) by IR rays for about 5m distance. This smart glass can be used by even normal people, as it helps in navigating the location to the user through GPS and display it on the glass screen. And it helps in enhancing the old document and detecting the content present in it. And also detect the ancient or handwritten text present in the document through text detection process and produces the output as speech to the user. For processing all this we are using android function and RAM for this. And we also use image segmentation, object recognition, text detection method which will be embedded in android environment.

## II. RELATED TECHNOLOGY

Previously, for converting the digital image into text and speech, we need to perform it by using MATLAB concept. This concept is based on embedded system term which will run on the computer system. Even for converting the text to an image we were using Text to Speech synthesis. The above methods take more processing time, so in order to overcome this we are proposing Image to Text and Speech conversion, which will be developed and embedded in android platform which user friendly.

## III. MECHANISM

In order to implement the above proposed work, we need to follow some methods. The methods are as follow:

### Document Enhancement

Here, we concentrate on enhancing the old document and detecting the content present in it. We capture or upload the image which we need to enhance and the image is being processed for enhancing. After a few second, image will be enhanced, which will be better version of previous. After the enhancement process is completed, Text detection and image segmentation will be applied to it. And then content present is detected in it by the method and give the output through voice to the user. During the enhancement process it will undergo following step: De-blurring of the image, Interpolation method, and applying adaptive threshold to that image. All these methods are executed in android application which is developed in android environment

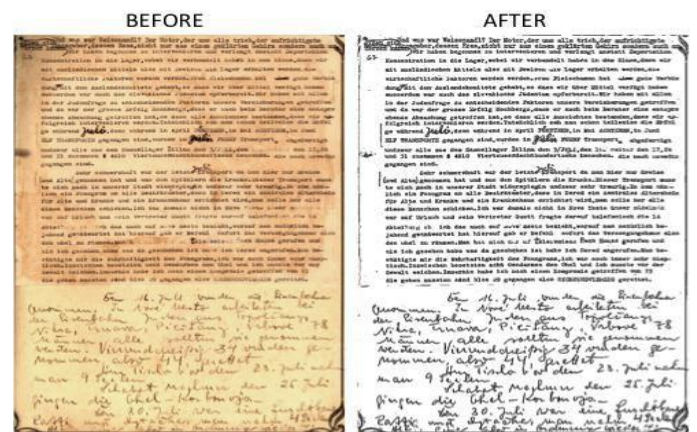


Fig3.1: Document Enhancement

### Object Recognition

Here, we look on the recognition of the object by using Infrared rays. The camera present in the smart glass will capture or scan the scene present in front of it. Through IR rays it detects the object present in it. Infrared rays can

travel up to some extent for about 5m distance. It will recognize the object present and produce output through voice to the user. There are few which will undergo during object recognition. First, Alignment method. secondly, invariants properties will be applied. And finally, Decomposition of parts will be applied. All the methods are done in android environment, where we use RAM of 1GB and Quad-core processor for it.

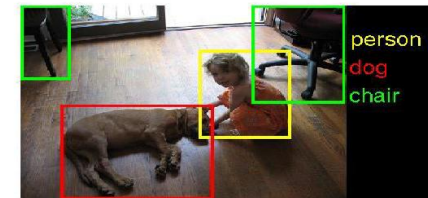


Fig3.2 (a): Object Recognition

The above picture depicts the object recognition. Here, an image is captured and subjected to object recognition which is being recognized by using IR rays. And IR rays detect the object present in front of the camera.



Fig3.2 (b): The device used for object recognition using IR rays.

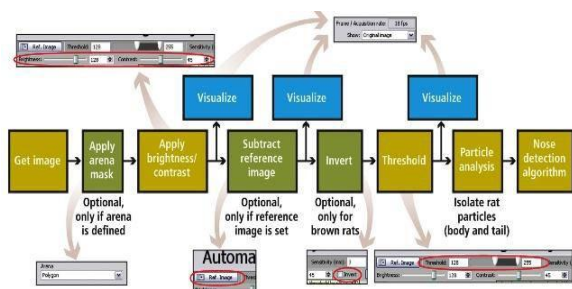


Fig3.2(c): Flow diagram of Object recognition.

**Navigation**

Navigation means showing the direction. Here, the concept of navigation is also applied in our model in order to navigate the place by using WiFi and GPS mode. In this feature if we give input to the smart glass through voice, it will navigate us to the location. The navigation of the place will be displayed on the screen of the glass and voice output

will also be produced to the user. This feature will use Global positioning system for navigating the location. And it will be helpful for all people.



Fig3.3: Navigation

**Text to Speech Conversion**

Here, we mainly concentrate on conversion of text to speech. Firstly, we capture the image through a camera, and then image is being pre-processed where it undergoes, following method: Feature extraction, image segmentation, edge detection, optical character recognition and Text synthesis to speech. After undergoing all this process an image containing text will be synthesized and produce the output as speech to the user. By this it will be able to convert the image to text and speech. And this feature will be more helpful for low pair vision and Blind people to read the books or a document.

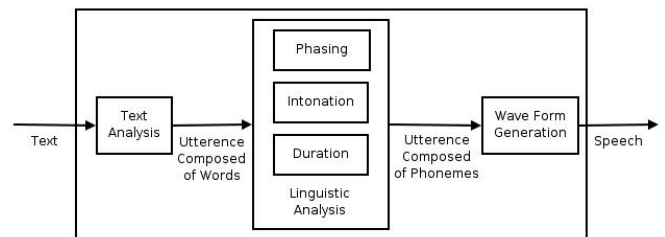


Fig 3.4: Text to Speech Synthesis

In the above fig Text to Speech synthesis flow is shown. Here, Text will be taken as input and it is analyzed and then it is synthesized to speech.

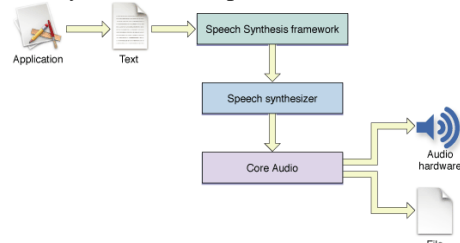


Fig 3.4(a): Text to Speech flow

**3.5 Text Detection**

In this we mainly focus on Text Detection concept. This method is also same as that of object recognition but here it detects only text present in the image.



Fig 3.5(a): Text Detection

As you see in the above picture, the smart glass captures an image and detects the text present in the image. Recognition of the text is done by preprocessing of the image. Preprocessing stage may involve image segmentation, edge detection and text detection methods. All the methods are developed as an android application on android platform and embedded in the smart glass.

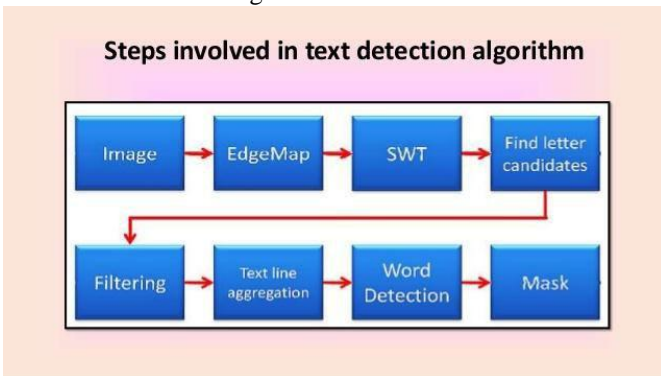


Fig 3.5(b): Steps involved in Text Detection Algorithm

As the image only depicts the flow of the algorithm, where image captured is preprocessed, edge detection method is applied to that image and further filters are applied to that image in order to detect the text present in the image.

**IV. REAL TIME REQUIREMENTS**

In order to build this proposed work, user feedback plays a vital role in it. Lot of people are interested in android, so for this purpose we are using android functionality to build this model. We also require a Quad core processor with 1GB of RAM for speed processing. If the size of RAM is more, then performance of a smart glass will be higher. We also use High Definition of camera and IR Rays for processing some methods in it.

**V. FUTURE ENHANCEMENT**

Efforts are being made for further enhancing of this device. Our new ideas towards enhancing the device. Firstly, 3D object modelling view where you can view the object in 3D view and interact with it by using hand gestures. Secondly,

to view the 2D image into 3D image in High Definition mode.

**VI. CONCLUSIONS**

We have compiled research trends in an Interactive Smart Glass using Google scholar’s searched tool citation found exciting. Here image segmentation, object recognition and text detection feedback have received a lot of attention, and hence further proceeding towards future enhancement of the device which will be useful to society.

**VII. ACKNOWLEDGEMENT**

The documentation and implementation of this paper would not be succeeded without the kind support from Dr. R BalaKrishna, principal, RajaRajeshwari College of Engineering, Bengaluru, who always gives us valuable advice and kind assistance to complete this paper. We are grateful to Dr.UshaSakthivel, Professor and HOD, Department Of Computer Science and Engineering, RajaRajeswari College Of Engineering, Bengaluru, for giving us the great knowledge.

Finally, we would like to thank our guide and parents who constantly supported us since the beginning till the end of the paper.

**REFERENCES**

[1] L. von Ahn and L. Dabbish, -Labeling Images with a Computer Game, *Proc. ACM CHI*, 2004.  
 [2] Airlines.net, <http://www.airliners.net>.  
 [3] J. Amores, N. Sebe, P. Radeva, T. Gevers, and A. Smeulders, -Boosting Contextual Information in Content-Based Image Retrieval, *Proc. Workshop on Multimedia Information Retrieval, in conjunction with ACM Multimedia*, 2004.  
 [4] J. Assfalg, A. Del Bimbo, and P. Pala, -Three-Dimensional Interfaces for Querying by Example in Content-Based Image Retrieval, *IEEE Trans. Visualization and Computer Graphics*, 8(4):305-318, 2002.  
 [5] K. Barnard, P. Duygulu, D. Forsyth, N. de Freitas, D. M. Blei, and M. I. Jordan, -Matching Words and Pictures, *Journal of Machine Learning Research*, 3:1107-1135, 2003.  
 [6] I. Bartolini, P. Ciaccia, and M. Patella, -WARP: Accurate Retrieval of Shapes Using Phase of Fourier Descriptors and Time Warping Distance, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 27(1):142-147, 2005.  
 [7] S. Belongie, J. Malik, and J. Puzicha, -Shape Matching and Object Recognition Using Shape Contexts, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 24(4):509-522, 2002.  
 [8] D. M. Blei and M. I. Jordan, -Modeling Annotated Data, *Proc. ACM Conference on Research and Development in Information Retrieval*, 2003  
 [9] C. Bohm, S. Berchtold, and D. A. Keim, -Searching in High-Dimensional Spaces Index Structures for Improving the Performance of Multimedia Databases, *ACM Computing*

- Surveys*, 33(3):322-373, 2001.
- [10] J. Carballido-Gamio, S. Belongie, and S. Majumdar –Normalized Cuts in 3-D for Spinal MRI Segmentation, *IEEE Trans. Medical Imaging*, 23(1):36–44, 2004.
- [11] G. Carneiro and N. Vasconcelos, –Minimum Bayes Error Features for Visual Recognition by Sequential Feature Selection and Extraction, *Proc. Canadian Conference on Computer and Robot Vision*, 2005.
- [12] C. Carson, S. Belongie, H. Greenspan, and J. Malik, –Blobworld: Image Segmentation Using Expectation-maximization and Its Application to Image Querying, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 24(8):1026-1038, 2002.
- [13] A. Chalechale, G. Naghdy, and A. Mertins, –Sketch-Based Image Matching Using Angular Partitioning, *IEEE Trans. Systems, Man, and Cybernetics*, 35(1):28–41, 2005.
- [14] E. Y. Chang, K. Goh, G. Sychay, and G. Wu, –CBSA: Content-based Soft Annotation for Multimodal Image Retrieval Using Bayes Point Machines, *IEEE Trans. Circuits and Systems for Video Technology*, 13(1):26–38, 2003.
- [15] C.-c. Chen, H. Wactlar, J. Z. Wang, and K. Kiernan, –Digital Imagery for Significant Cultural and Historical Materials - An Emerging Research Field Bridging People, Culture, and Technologies, *International Journal on Digital Libraries*, 5(4):275–286, 2005.
- [16] J. Chen, T.N. Pappas, A. Mojsilovic, and B. Rogowitz, –Adaptive image segmentation based on color and texture, *Proc. IEEE International Conference on Image Processing*, 2002.
- [17] Y. Chen and J. Z. Wang, –A Region-Based Fuzzy Feature Matching Approach to Content-Based Image Retrieval, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 24(9):1252–1267, 2002.
- [18] Y. Chen and J. Z. Wang, –Image Categorization by Learning and Reasoning with Regions, *Journal of Machine Learning Research*, 5:913-939, 2004.
- [19] Y. Chen, J. Z. Wang, and R. Krovetz, –CLUE: Cluster-Based Retrieval of Images by Unsupervised Learning, *IEEE Trans. Image Processing*, 14(8):1187–1201, 2005.
- [20] D. Comaniciu and P. Meer, –Mean Shift: A Robust Approach Toward Feature Space Analysis, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 24(5), 603–619, 2002.
- [21] I. J. Cox, M. L. Miller, T. P. Minka, T. V. Papatomas, and P. N. Yianilos, –The Bayesian Image Retrieval System, PicHunter: Theory, Implementation, and Psychophysical Experiments, *IEEE Trans. Image Processing*, 9(1):20–37, 2000.
- [22] A. Csillaghy, H. Hinterberger, and A.O. Benz, –Content Based Image Retrieval in Astronomy, *Information Retrieval*, 3(3):229–241, 2000.
- [23] S. J. Cunningham, D. Bainbridge, and M. Masoodian, –How People Describe Their Image Information Needs: A Grounded Theory Analysis of Visual Arts Queries, *Proc. ACM/IEEE Joint Conference on Digital Libraries*, 2004.
- [24] C. Dagli and T. S. Huang, –A Framework for Grid-Based Image Retrieval, *Proc. IEEE International Conference on Pattern recognition*, 2004.
- [25] Y. Deng, B. S. Manjunath, C. Kenney, M. S. Moore, and H. Shin, –An Efficient Color Representation for Image Retrieval, *IEEE Trans. Image Processing*, 10(1):140–147, 2001.
- [26] M. N. Do and M. Vetterli, –Wavelet-Based Texture Retrieval Using Generalized Gaussian Density and Kullback-Leibler Distance, *IEEE Trans. Image Processing*, 11(2):146–158, 2002.
- [27] A. Dong and B. Bhanu, –Active Concept Learning for Image Retrieval in Dynamic Databases, *Proc. IEEE International Conference on Computer Vision*, 2003.
- [28] Y. Du and J. Z. Wang –A Scalable Integrated Region-Based Image Retrieval System, *Proc. IEEE International Conference on Image Processing*, 2001.
- [29] P. Duygulu, K. Barnard, N. de Freitas, and D. Forsyth. Object recognition as machine translation: Learning a lexicon for a fixed image vocabulary. In *Seventh European Conference on Computer Vision*, pages 97–112, 2002.
- [30] Global Memory Net, <http://www.memorynet.org>.
- [31] A. W. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, –Content-Based Image Retrieval at the End of the Early Years, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(12):1349–1380, 2000.