Web-Cam based Assistive Text and Product Label Read from Hand-Held Objects for Visual Impaired

Janardhana K.L[1] Bhuvana Suganthi.D[2]
[1]Student [2]Associate Professor
[1,2]Department of Electronics & Communication Engineering
[1,2]AMC Engineering College, Bangalore, India

Abstract— we propose a camera-based assistive content perusing structure to visually impaired persons read content marks and item bundling from hand-held articles in their day by day lives. To disconnect the article from jumbled foundations or other encompassing items in the camera view, we first propose a proficient and successful movement based technique to characterize an area of interest (ROI) in the video by requesting that the client shake the item. This strategy separates moving item district by a blend of-Gaussians-based foundation sub-footing technique. In the extricated ROI, content confinement and acknowledgment are led to gain content data. To consequently confine the content districts from the item ROI, we propose a novel content limitation calculation by learning slope elements of stroke introductions and disseminations of edge pixels in an Adaboost model. Content characters in the limited content locales are then binarized and perceived by off-the-rack optical character acknowledgment delicate product. The perceived content codes are yield to visually impaired clients in discourse. Execution of the proposed content limitation calculation is quantitatively assessed on ICDAR-2003 and ICDAR-2011 Robust Reading Datasets. Trial results show that our calculation accomplishes the condition of expressions of the human experience. The Proof of idea model is additionally assessed on a dataset gathered utilizing ten visually impaired persons to assess the viability of the framework's equipment. We investigate client between face issues and evaluate power of the calculation in extricating and perusing content from various items with complex foundations.

Key words: Assistive Gadgets, Visual Deficiency, Dispersion of Edge Pixels, Hand-Held Articles, Optical Character Acknowledgment (OCR), Stroke Introduction, Content Perusing, Content District Restriction

I. INTRODUCTION

Of the 314 million outwardly hindered individuals around the world, 45 million are visually impaired. Indeed, even in a created nation like the U.S., the 2008 National Health Interview Survey reported that an expected 25.2 million grown-up Americans (more than 8%) are visually impaired or outwardly weakened. This number is expanding quickly as the person born after WW2 era ages. Late improvements in Computer vision, computerized cameras, and convenient PCs make it possible to help these people by creating camera-based items that consolidate PC vision innovation with other existing business items such optical character acknowledgment (OCR) frameworks.

Perusing is clearly crucial in today's general public. Printed content is all over the place as reports, receipts, bank proclamations, eatery menus, classroom freebees, item bundles, directions on prescription containers, and so on. Keeping in mind optical guides, video magnifiers, and screen per users can dazzle clients and those with low vision to get to archives, there are couple of gadgets that can give great access to basic hand-held questions, for example, item bundles, and protests printed with content, for example, physician recommended prescription jugs. The capacity of individuals who are visually impaired or have huge visual debilitations to peruse printed names and item bundles will upgrade autonomous living and cultivate monetary and social independence.

Today, there are as of now a couple of frameworks that have some guarantee for convenient use, however they can't deal with item marking. For instance, compact standardized tag per users intended to visually impaired individuals distinguish diverse items in a broad item database can empower clients why should daze access data about these items [22] through discourse and Braille. Be that as it may, a major restriction is that it is hard for visually impaired clients to discover the position of the standardized identification and to accurately point the scanner tag peruser at the standardized tag. Some perusing assistive frameworks, for example, pen scanners may be utilized in these and comparative circumstances. Such systems coordinate OCR programming to offer the capacity of checking and acknowledgment of content and some have incorporated voice yield. In any case, these frameworks are for the most part intended for and perform best with report pictures with basic foundations, standard textual styles, a little scope of text dimensions, and very much sorted out characters instead of business item boxes with numerous ornamental examples. Most best in class OCR programming can't straightforwardly handle scene pictures with complex foundations.

Various convenient perusing associates have been planned particularly for the outwardly weakened; KReader Mobile keeps running on a wireless and permits the client to peruse mail, receipts, fliers, and numerous different archives. Nonetheless, the record to be perused must be almost level, set on a reasonable, dim surface (i.e., a non jumbled foundation), and contain generally message. Besides, KReader

Fig. 1: Case of Printed Content from Hand-Held Items with Numerous Hues, Complex Foundations, or Non-Level Surfaces.

Versatile precisely peruses dark print on a white foundation, yet has issues perceiving hued content or content on a shaded foundation. It can't read content with complex foundations, content imprinted on barrels with distorted or inadequate pictures, (for example, soup jars or pharmaceutical jugs). Besides, these frameworks require a visually impaired client to physically confine territories of interest and content areas on the articles as a rule.

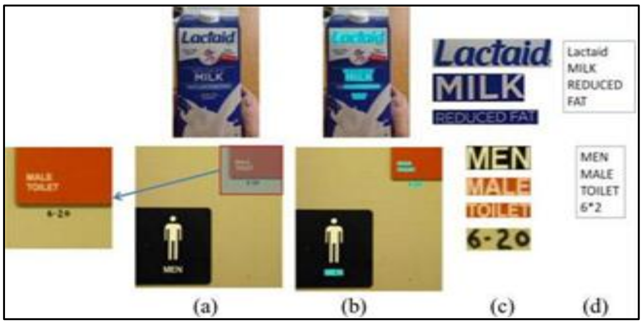

Fig. 2: Two case of content limitation and acknowledgment from camera caught pictures (Top) Milk box (Base) Men lavatory signage (a) camera-caught pictures (b) Localized content areas (set apart in blue) (c) Text locales trimmed from picture (d) Text codes perceived by OCR. Content at the upper right corner of base picture is appeared in an amplified callout.

In spite of the fact that various perusing collaborators have been composed particularly for the outwardly debilitated, as far as anyone is concerned, no current perusing colleague can read content from the sorts of testing examples and foundations found on numerous regular business items. As appeared in Fig. 1, such content data can show up in different scales, text styles, hues, and introductions. To help blind persons to peruse content from these sorts of hand-held articles, we have considered a camera-based assistive content perusing outline work to track the object of enthusiasm inside the camera view and concentrate print content data from the item. Our proposed calculation can successfully handle complex foundation and different examples, and concentrate content data from both hand-held articles and close-by signage, as appeared in Fig. 2.

In assistive perusing frameworks for visually impaired persons, it is exceptionally trying for clients to position the object of enthusiasm inside the focal point of the camera's perspective. Starting now, there are still no acknowledge capable arrangements. We approach the issue in stages. To ensure the hand-held item shows up in the camera view, we utilize a camera with adequately wide edge to oblige clients with just surmised point. This may frequently bring about other content items showing up in the camera's perspective (for instance, while shopping at a grocery store). To extricate the hand-held article from the camera picture, we build up a movement based strategy to acquire a locale of interest (ROI) of the item. At that point, we perform content acknowledgment just in this ROI.

It is a testing issue to consequently confine items and content ROIs from caught pictures with complex back-grounds, since content in caught pictures is in all probability encompassed by different foundation anomaly "clamor," and content characters generally show up in various scales, text styles, and hues. For the content introductions, this paper expect that content strings in scene pictures keep roughly even arrangement. Numerous calculations have been created for restriction of content locales in scene pictures. We separate them into two classes: guideline based and learning-based.

While catching pictures of the hand-held article, the visually impaired client first keeps the item still and after that daintily shakes the item for 1 or 2 second, Here, we apply the effective numerous Gaussian-blend based BGS strategy to identify the article district while blind client shakes it. More subtle elements of the calculation can be found in. Once the object of interest is extricated from the camera picture, the framework is prepared to apply our programmed content extraction calculation.

II. AUTOMATIC TEXT EXTRACTION

We plan a learning-based calculation for programmed limitation of content districts in image. In request to handle complex foundations, we propose two novel component maps to concentrates content elements in light of stroke introductions and edge disseminations, separately. Here, stroke is characterized as a uniform area with limited width and huge degree. These component maps are consolidated to fabricate an Adaboost-based content classifier. An example of content strokes demonstrating connections between stroke introductions and angle introductions at pixels of stroke limits. Blue bolts indicate the stroke introductions at the segments and red bolts signify the inclination introductions at pixels of stroke limits.

A. Block Diagram of the Text Extraction System

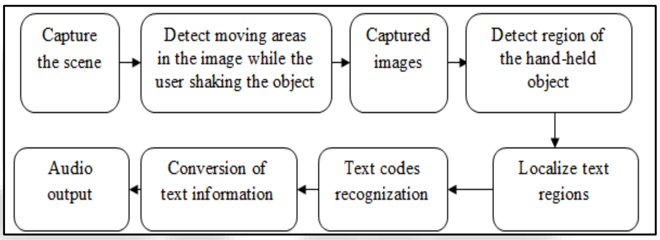

Fig. 3: Text Extraction system

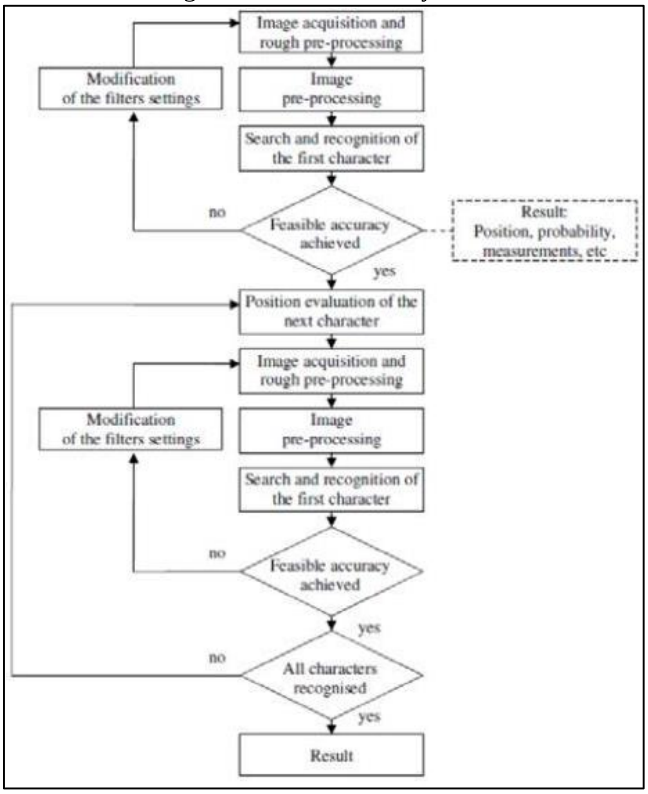

Fig. 4: Flowchart of the Text Extraction System

Fig.3 demonstrates the square graph of the Text extraction framework, how to distinguish areas in a picture that contain content. This is a typical assignment performed on unstructured scenes. Unstructured scenes are pictures that contain undetermined or irregular situations. For instance, you can distinguish and perceive message naturally from caught video to caution a driver around a street sign. This is not the same as organized scenes, which contain known situations where the position of content is known in advance. Dividing content from an unstructured scene enormously assists with extra errands, for example, optical character acknowledgment (OCR). The mechanized content discovery calculation in this illustration recognizes countless district competitors and continuously expels those less inclined to contain content. Fig.4. flowchart demonstrates the orderly execution of picture handling and Text extraction process.

Basic steps involved in Image to Text extraction consists of the following steps:

1) Step 1: Detect Candidate Text Regions using MSER

The MSER feature detector works well for finding text regions. It works well for text because the consistent color and high contrast of text leads to stable intensity profiles. Use the `detect MSER features` function to find all the regions within the image and plot these results. Notice that there are many non-text regions detected alongside the text.

2) Step 2: Remove Non-Text Regions based on Basic Geometric Properties

Although the MSER algorithm picks out most of the text, it also detects many other stable regions in the image that are not text. You can use a rule-based approach to remove non-text regions. For example, geometric properties of text can be used to filter out non-text regions using simple thresholds. Alternatively, you can use a machine learning approach to train a text vs. non-text classifier. Typically, a combination of the two approaches produces better results. This example uses a simple rule-based approach to filter non-text regions based on geometric properties.

There are several geometric properties that are good for discriminating between text and non-text regions including:

- Aspect ratio
- Eccentricity
- Euler number
- Extent
- Solidity

Use regionprops to measure a few of these properties and then remove regions based on their property values.

3) Step 3: Remove Non-Text Regions Based On Stroke Width Variation

Another common metric used to discriminate between text and non-text is stroke width. *Stroke width* is a measure of the width of the curves and lines that make up a character. Text regions tend to have little stroke width variation, whereas non-text regions tend to have larger variations.

To help understand how the stroke width can be used to remove non-text regions, estimate the stroke width of one of the detected MSER regions. You can do this by using a distance transform and binary thinning operation. In the images shown above, notice how the stroke width image has very little variation over most of the region. This indicates that the region is more likely to be a text region because the lines and curves that make up the region all have similar widths, which is a common characteristic of human readable text. In order to use stroke width variation to remove non-text regions using a threshold value, the variation over the entire region must be quantified into a single metric as follows:

4) Step 4: Merge Text Regions for Final Detection Result

At this point, all the detection results are composed of individual text characters. To use these results for recognition tasks, such as OCR, the individual text characters must be merged into words or text lines. This enables recognition of the actual words in an image, which carry more meaningful information than just the individual characters. For example, recognizing the string 'EXIT' vs. the set of individual characters {'X','E','T','I'}, where the meaning of the word is lost without the correct ordering.

One approach for merging individual text regions into words or text lines is to first find neighboring text regions and then form a bounding box around these regions. To find neighboring regions, expand the bounding boxes computed earlier with regionprops. This makes the bounding boxes of neighboring text regions overlap such that text regions that are part of the same word or text line form a chain of overlapping bounding boxes. Now, the overlapping bounding boxes can be merged together to form a single bounding box around individual words or text lines. To do this, compute the overlap ratio between all bounding box pairs. This quantifies the distance between all pairs of text regions so that it is possible to find groups of neighboring text regions by looking for non-zero overlap ratios. Once the pair-wise overlap ratios are computed, use a graph to find all the text regions "connected" by a non-zero overlap ratio.

Use the bboxOverlapRatio function to compute the pair-wise overlap ratios for all the expanded bounding boxes, then use `graph` to find all the connected regions. The output of conncomp is indices to the connected text regions to which each bounding box belongs. Use these indices to merge multiple neighboring bounding boxes into a single bounding box by computing the minimum and maximum of the individual bounding boxes that make up each connected component.

5) Step 5: Recognize Detected Text using OCR

After detecting the text regions, use the ocr function to recognize the text within each bounding box. Note that without first finding the text regions, the output of the OCR function would be considerably more noisy. This example showed you how to detect text in an image using the MSER feature detector to first find candidate text regions, and then it described how to use geometric measurements to remove all the non-text regions. This example code is a good starting point for developing more robust text detection algorithms. Note that without further enhancements this example can produce reasonable results for a variety of other images, for example, posters.jpg or licensePlates.jpg, Handicap.jpg.

Two datasets are utilized to assess our calculation. In the first place, the ICDAR Robust Reading Dataset is utilized to assess the proposed content restriction calculation. The ICDAR-2003 dataset contains 509 normal scene pictures altogether. Most pictures contain indoor or open air content signage. The picture resolutions range from 640 × 480 to 1600 × 1200. Since design examination in view of adjoining character gathering can just handle content strings with three or more character individuals, we overlook the pictures containing just ground truth content areas of under three content characters. Subsequently, 488 pictures are chosen from this dataset as testing pictures to assess our limitation calculation.

To promote comprehend the execution of the model sys-tem and build up an easy to understand interface, taking after Human Subjects Institutional Review Board endorsement, we enrolled ten visually impaired persons to gather a dataset of perusing content close by held items.

Fig. 5: Examples of visually impaired persons catching pictures of the item in their grasp.

The equipment of the model framework incorporates a Logitech web camera with self-adjust, which is secured to the nose scaffold of a couple of shades. The camera is associated with a HP smaller than normal portable PC by a USB association. The tablet plays out the handling and gives sound yield. With a specific end goal to stay away from genuine blocking or aural diversion, we would pick a remote "open" style Bluetooth earpiece for showing discovery results as discourse yields to the visually impaired explorers in a full model usage.

The visually impaired client wore the camera/shades to catch the picture of the items in his/her hand, as showed in Fig.5, the determination of the caught picture is 960 × 720. There were 14 trying articles for every individual, including basic supply boxes, solution bottles, books, and so forth. They were required keep their head (where the camera is settled) stationary for a few moments and hence shake the item for an extra couple of seconds to recognize the district of object of interest. Every article was then pivoted by the client a few times to guarantee that surfaces with content subtitles are uncovered and caught. We physically extricated 116 caught pictures and marked 312 content areas of primary titles.

III. PROTOTYPE SYSTEM EVALUATION

The programmed or automated ROI identification and content limitation calculations were autonomously assessed as unit tests to guarantee viability and heartiness of the entire framework. We thusly assessed this model arrangement of assistive content perusing utilizing pictures of hand-held articles caught by ten visually impaired clients in individual.

Two alignments were connected to get ready for the framework test. In the first place, we trained visually impaired clients to place hand-held item inside the camera view. Since it is troublesome for visually impaired clients to point their held articles, we utilized a camera with a sensibly wide edge. In future frameworks, we will include finger guide discovery and following toward adaptively train blind clients to point the article. Second, in a pertinent visually impaired assistive framework, a content confinement calculation may incline toward higher review by relinquishing some exactness. We balanced the parameters of our content limitation calculation and acquired another gathering of assessment results, as exactness 0.48, review 0.72, f - measure 0.51. The higher review guarantees a lower miss (false negative) rate. To sift through false positive restriction, we could assist utilize some post preparing calculation in view of scene content acknowledgment or lexical investigation. This work will be done in future work.

Next, we assessed the client caught dataset of article content. The dataset was physically commented on by marking the locales of the object of interests and the content areas inside the object of interest districts. In our calculation assessment, we characterized a locale as accurately identified if the proportion of the cover zone of a distinguished district and its ground truth area is no under 3/4. Tests demonstrated that 225 of the 312 ground truth content areas were hit by our confinement calculation. By utilizing the same assessment measures as above tests, we got accuracy 0.52, review 0.62, and f - measure 0.52 on this dataset. The accuracy is lower than that on the Robust Reading Dataset. The pictures in the client caught dataset have lower resolutions and more minimized dispersion of content data, so they create low-quality edge maps and content limits, which result in disgraceful spatial designs and content structure highlights.

OCR is connected to the confined content areas for character and word acknowledgment. Fig. 16 demonstrates a few case of content limitation and acknowledgment of our proposed structure. We take note of that the acknowledgment calculation may not effectively and totally yield the words inside confined areas. Extra spelling remedy is likely required to yield precise content data. Our content perusing framework burns through 1.87 s all things considered perusing content from a camera-based picture. The framework productivity can and will be enhanced by parallel preparing of content extraction and gadget info/yield, i.e., discourse yield of perceived content and limitation of content areas in the following picture are performed at the same time.

IV. PROGRESSIVE STREAMLINING CALCULATION

The purpose behind utilizing of this calculation is that the fundamental OCR alone is not appropriate to peruse powerful and bended content. Our calculation will utilize design character acknowledgment and various leveled streamlining to perform better acknowledgment of the printable writings. Another explanation behind utilizing this calculation is a direct result of its high level of precision in acknowledgment alongside its rate. For our situation the pictures are taken by a visually impaired client, there is a decent probability for a mutilated picture. These are prerequisites that should be fulfilled by our calculation, stableness to mistakes in perceived characters, rapid, effectively prepared and tuned.

By thinking about above necessities, design acknowledgment is best fitted calculation to create the normal yield. It has high acknowledgment speed furthermore it is steady for minor blunders in the picture yet little bending will bring about negative acknowledgment of characters. In example acknowledgment an item with required character is chosen out of the first picture and contrasted and every one of the examples in the database. The example with least contrasts from unique picture is considered as result. While looking at both the examples, it is conceivable that one of the examples ought to be moved vertically or on a level plane in any event. Furthermore the acknowledgment time relies on upon size and the turn of the example. The significant obstacle in utilizing design calculation for content perusing on clamor and contorted pictures is the more twisting of the characters, it makes direct correlations of example inconceivable.

To determine these issues, taking after methodologies are recommended to utilize: (i) Template Distortion while contrasting with the perceived character, (ii) changing the settings of the pre-process channels persistently and dissecting the outcomes by considering the last acknowledgment results. Both of the above proposals build the acknowledgment time. With a specific end goal to enhance pace of layout looking that has a place best with the character on the picture various leveled probabilistic coordinating is utilized. The above strategy comprises of taking after strides: when contrasting the layouts and character, utilizing unequivocal calculation just part of the format positions are checked in light of the picture and point number (determination) of formats and inquiry territory is changed at first one after one and afterward many steps till it utilizes every one of the focuses.

As it were, streamlining calculation seeks the greatest match (ideal) of the quality capacity; it mirrors the decency in the correspondence between the format and the character by current point number (resolution).Powel's strategy with advancement was chosen as an improvement calculation. The accompanying qualities can be changed without reliance from each other by determination: 1) Translations along the tomahawks x and y, 2) interpretation in the course (x, y), (- x, y) and so on 3) Rotation around the tomahawks x and y 4) Rotation around the z-hub 5) scaling i.e. the Translation along the z-pivot. So the enhancement by every point number (determination) is done in 8-dimensionsal space. Taking into account advancement calculation, the area of the format for the following emphasis is characterized in light of value foundation, which is ascertained as beneath:

$$\boxed{QC_1 = \sum_{i=0}^{N} d_i^k}$$

$$\boxed{QC_2 = \sum_{i=0}^{N} \frac{n_e}{n_t} \cdot \frac{n_e}{n_c}}$$

N - Amount of the layout focuses, di - separation from ith purpose of the format to the closest purpose of the perceived character, k - ½, 1 or 2. Ne - measure of the layout guides correspondent toward the purposes of the perceived character, Nt - the aggregate sum of the format focuses, Nc - measure of the purposes of the picture region which is limited by the connected format.

V. TEST AND RESULTS

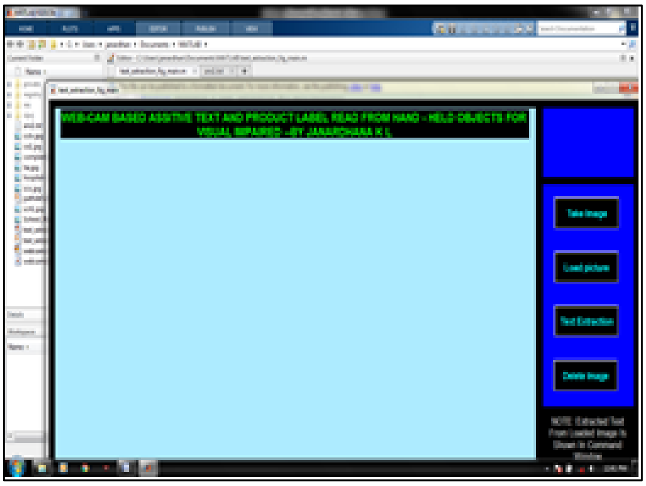

Fig. 6: GUI windows for Text extraction

Fig. 7: loaded image for text extraction

Fig. 8: MSER Regions

Fig. 9: Removal of non-text regions based on Geometric properties

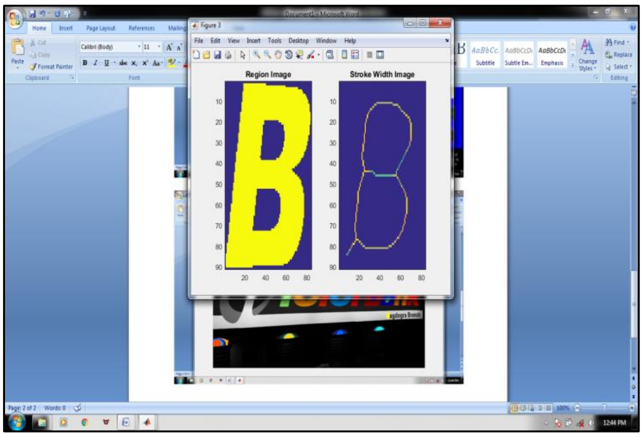

Fig. 10: Region and Stroke width images

Fig. 11: Removal of non content locales in view of stroke width variety

Fig. 12: Expansion of bounding boxes text

Fig. 13: Detected resultant text

In the figures above 6 to 13 (6, 7, 8, 9, 10, 11, 12 and 13), shows how the orderly content extraction is done from hand-held or caught picture that is to recognize areas in a photo it incorporates content. This is an ordinary undertaking performed on unstructured scenes. Unstructured scenes are pictures that contain undetermined or self-decisive circumstances. For example, you can perceive and see message typically from inspired video to caution a driver around a street sign. This is not the same as formed scenes, which contain known circumstances where the position of substance is known starting at this point. Portioning content from an unstructured scene basically assists with extra errands, for instance, optical character affirmation (OCR). The automated substance ID figuring in this delineation recognizes endless zone contenders and constantly ousts those less slanted to contain content.

Ex: Loaded image is icici bank.

A. *Speech Output Generation*

The perceived content from OCR is composed on a content document. It is given as contribution to the discourse motor. Speech engine converts the texts from the file and store it into an array and after that it will be compared to the library and then audio is generated based on the output.

The detected text and audio model shown in Fig.14.

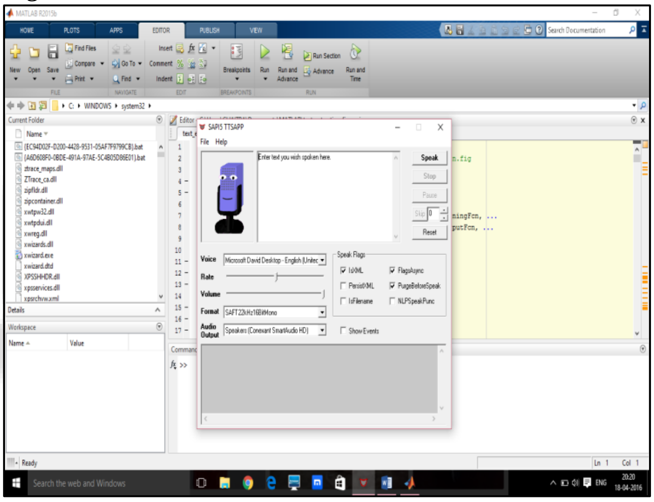

Fig. 14: Text to speech out model

VI. Conclusion and Future Work

In this paper, we have portrayed a model framework to peruse printed content available held articles for helping blind persons. Keeping in mind the end goal to tackle the basic pointing issue for visually impaired clients, we have proposed a movement based strategy to recognize the object of interest, while the visually impaired client essentially shakes the article for two or three seconds. This strategy can successfully recognize the object of enthusiasm from foundation or different items in the cam-time view.

To concentrate content areas from complex foundations, we have proposed a novel content confinement calculation in light of models of stroke introduction and edge circulations. The comparing highlight maps gauge the worldwide auxiliary component of content at each pixel. Piece designs extend the proposed highlight maps of a picture patch into a component vector. Contiguous character gathering is performed to figure applicants of content patches arranged for content order. An Adaboost learning model is utilized to confine content in camera-based pictures. Off-the-rack OCR is utilized to perform word acknowledgment on the limited content areas and change into sound yield for visually impaired clients.

Our future work will extend our restriction calculation to process content strings with characters less than three and to outline more vigorous square examples for content element extraction. We will likewise extend our calculation to handle non flat content strings. Moreover, we will address the critical human interface issues connected with perusing content by visually impaired clients.

ACKNOWLEDGMENT

The creators might want to thank the mysterious analysts for their productive remarks and quick proposals that enhanced the nature of this original copy.

REFERENCES

[1] Chucai Yi, Yigli Tian. Aries Arditi, “Portable Camera-Based Assistive Text and Product Label Reading From Hand-Held Objects for Blind Persons,” In IEEE Transactions on mechatronics, Vol. 19, No. 3, June 2014.

[2] L. Ma, C. Wang, and B. Xiao, “Text detection in natural images based on multi-scale edge detection and classification,” in Proc. Int. Congr. Image Signal Process., 2010, vol.4, pp. 1961–1965.

[3] C. Yi and Y. Tian, “Text string detection from natural scenes by structure based partition and grouping,” IEEE Trans. Image Process., vol. 20, no. 9, pp. 2594–2605, Sep. 2011.

[4] X. Chen and A. L. Yuille, “Detecting and reading text in natural scenes,” In CVPR, Vol. 2, pp. II-366 – II-373, 2004.

[5] X. Chen, J. Yang, J. Zhang and A. Waibel, “Automatic detection and recognition of signs from natural scenes,” In IEEE Transactions on image processing, Vol. 13, No. 1, pp. 87-99, 2004.

[6] “KReader Mobile User Guide”, knfb Reading Technology Inc, http://www.knfbReading.com

[7] J. Liang, D. Doermann, and H. Li, Camera-based analysis of text and documents: a survey. In International Journal of Document Analysis and Recognition (IJDAR), No. 2-3, pp. 84-104, 2005.

[8] N. Nikolaou and N. Papamarkos, “Color Reduction for Complex Document Images,” In International Journal of Imaging Systems and Technology, Vol.19, pp.14-26, 2009.

[9] ScanTalker, Bar code scanning application to help Blind Identify over one million products,http://www.freedomscientific.com/fs_news/PressRoom/en/2006/ScanTalker2-Announcement_3-30-2006.asp

[10] N. Giudice and G. Legge, “Blind navigation and the role of technology,” in The Engineering Handbook of Smart Technology for Aging, Disability, and Independence, A. A. Helal, M. Mokhtari, and B. Abdulrazak, Eds. Hoboken, NJ, USA: Wiley, 2008

[11] K. Kim, K. Jung, and J. Kim, “Texture-based approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 12, pp. 1631–1639, Dec. 2003.

[12] B. Epshtein, E. Ofek, and Y. Wexler, “Detecting text in natural scenes with stroke width transform,” in Proc. Comput. Vision Pattern Recognit., 2010, pp. 2963–2970.

[13] World Health Organization. (2009). 10facts about blindness and visual impairment. Available: www.who.int/features/factfiles/blindness/blindness_facts/en/index.html

[14] Advance Data Reports from the National Health Interview Survey (2008). [Online]. Available: http://www.cdc.gov/nchs/nhis/nhis_ad.htm

[15] International Workshop on Camera-Based Document Analysis and Recognition (CBDAR 2005, 2007, 2009, 2011). [Online]. Available: http://www.m.cs.osakafu-u.ac.jp/cbdar2011/

[16] M. Shi, Y. Fujisawab, T. Wakabayashia, and F. Kimura, “Handwritten numeral recognition using gradient and curvature of gray scale image,” Pattern Recognit., vol. 35, no. 10, pp. 2051–2059, 2002

