ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering
(An ISO 3297: 2007 Certified Organization)
Website: www.ijircce.com
Vol. 4, Issue 12, December 2016

Text and Label Reading Using Robust Text Localization

Shradha Chalikwar, A. B. Manwatkar, Sinu Nambiar

M.E Student, Dept of Computer Engineering, Indira College of Engineering and Management, Savitribai Phule Pune University, Maharashtra, India.

Assistant Professor, Dept. of Computer Engineering, Indira College of Engineering and Management, Savitribai Phule Pune University, Maharashtra, India.

Assistant Lecturer, Dept. of Computer Engineering, Indira College of Engineering and Management, Savitribai Phule Pune University, Maharashtra, India.

ABSTRACT: Currently in day to day life, in order to read the text label from packaged hand-held object for the blind person is very difficult, in market camera based product used for text reading framework for identifying the text. Text is localized by effective motion method to define ROI (Region of interest) in image by asking user to shake the object i.e. In ROI, text localization and recognition are conducted to obtain textual information. Different types of methods are used to localize the text for example: Mixture of Gaussian subtraction method, Adaboost model. Scope of the amendment is to achieve better accuracy in proposed algorithms. At present, SWT operator is being used to fine the value of stroke width for each image pixel. Operator is local and data dependent which makes it robust and fast.

KEYWORDS: Text localization, SWT, Edge Detection, ROI, Adaboost model.

I. INTRODUCTION

Consider an example from our day to day life, suppose removing text and sign, then how difficult it would be to get information; so it is an important aspect to make visually impaired people's life easier. For visually impaired people, text reading is important aspect of their life which decides the quality of life. Identification of text from natural image is important for visually impaired person. Main aspect behind that to try and help those people who are suffering in their daily routine..

People with clear vision can use facilities like computer, internet and education soft-ware. However, for visually impaired people can have less access to any of these digital facility. As a result of this, they cannot enrich their knowledge with these advancement facilities. So it impacts economic, commercial, education ventures in the society. One of the way to shorten gap is to develop such system which is economical and facilitate them to communicate freely with widely used information infrastructure.

By using camera based products and combining computer vision technology with other commercial products such optical character recognition (OCR) systems help these individual to trace information. There are various product use by visually impaired person in their daily routine. A number of portable reading assistants have been designed specifically for the visually impaired, K-Reader Mobile runs on a cell phone and allows the user to read mail, receipts and many other documents. However, the document to be read must be nearly flat, placed on a clear, dark surface (i.e a non-cluttered back-ground), and contain mostly text Mobile accurately reads black print on a white background. Furthermore,systems require visually impaired user to manually localize areas of interest and text regions on the objects in most cases. Here the text information can appear in multiple scales, fonts, colors, and orientations. To assist visually impaired persons to read text from these kinds of hand-held objects, It has conceived of a camera-based assisted text reading framework to track the object of interest within the camera view. And extract print text information from the object.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering
(An ISO 3297: 2007 Certified Organization)
Website: www.ijircce.com
Vol. 4, Issue 12, December 2016

II. RELATED WORK

In [2] they research on "Robust Text Detection in Natural Scene Images". Their approaches for scene textual content detection can roughly be categorized into three companies: sliding window established procedure linked factor head quartered ways and hybrid approaches. Sliding window founded ways, often referred to as region-cantered ways. Stable totally regions (MSERs) established ways, which can be categorized as related component founded methods but utilizing MSERs as personality candidates. Connected aspect head quartered methods extract persona candidates from pictures by means of connected aspect evaluation followed by means of grouping persona candidates into text; extra checks may be carried out to take away false positives. MSER-head quartered scene text detection system. MSER founded methods have suggested promising efficiency on the largely used IC-DAR 2011 robust studying competitors database.In [3] has research on text and label reading concept. Their research work on Detecting and Reading Text in Natural Scenes. Text detection and Reading in Natural Scenes which was based on text detection from city scenes natural images like garden, flowers name, any daily product etc. The databases have been named to empower dazzle individuals after that in the first place, and most essential, part of the calculation is a solid classifier which is prepared by the Ad-Boost learning calculation on named information. Ad-Boost requires determining an arrangement of components from which to construct the solid classifier. They chooses the list of capabilities guided by the rule of instructive components that compute joint likelihood disseminations of these element reactions on and off con-tent, so feeble classifiers can be gotten as log-probability proportion tests. The Solid classifier is connected to sub-areas of the picture (at various scale) and yields con-tent hopeful districts. The application of there work are regularly between 2-5 false encouraging points in pictures of 2,048 x 1,536 pixels. The second segment is an augmentation and binary calculation that follows up on the content area competitors. The expansion and binary calculation takes the content locales as sources of info, augments these areas, in order to incorporate content that the solid classifier did not recognize, and binaries them (in a perfect world, so that the content is white and the foundation is dark). The third part is an OCR programming program which follows up on the binarized areas (the OCR programming gave far more regrettable execution when connected specifically to the picture). The OCR programming either verifies that the areas are content, and understands them, or rejects the locale as content.In[4]who work on Machine Vision and Intelligent Systems Group under University of Oulu, FINLAND. They research on edge factor which is important to detect text. Paper is "Edge-Based Method for Text Detection from Complex Document Images". This paper is base on Detection of textual content from documents wherein textual content is embedded in difficult colored and textured backgrounds is an awfully challenging obstacle. This paper, advocate an easy texture-founded procedure based on aspect understanding for this assignment. The efficiency of process is compared to that received with the aid of a method centered on the discrete cosine change into for text localization in compressed digital video. This paper, demonstrate that straightforward texture measures established on side knowledge provide very valuable knowledge for textual content detection from problematic file pictures.In [5]. work on "Automatic Detection and Recognition of Signs From Natural Scenes". In that research they has research on multi resolution and multi scale area detection, adaptive shopping, color analysis, and affine rectification in a hierarchical framework for signal detection, with specific emphases at every phase to manage the text in exclusive sizes, orientations, color distributions and backgrounds.The affine rectification use to recover deformation of the text regions caused by an inappropriate camera view angle. The procedure can significantly improve text detection rate and optical character recognition (OCR) accuracy. Instead of using binary information for OCR, we extract features from an intensity image directly. They propose research work on local intensity normalization method to effectively handle lighting variations, followed by transform method to obtain local features, and finally a linear discriminant analysis (LDA) method for feature selection. They have applied the approach in developing a Chinese sign translation system, which can automatically detect and recognize Chinese signs as input from a camera, and translate the recognized text into English. Signs are everywhere in our lives. A sign is an object that suggests the presence of a fact. It can be a displayed structure bear-ing letters or symbols, used to identify something or advertise a place of business. It can also be a posted notice bearing a designation, direction, safety advisory, or command. The advantage of the DCT (Discrete Cosine Transform) region investigation strategy is that DCT coefficients can be acquired specifically from a JPEG or MPEG picture, while then again the wavelet change can give more steady components contrasted and DCT method

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering
(An ISO 3297: 2007 Certified Organization)
Website: www.ijircce.com
Vol. 4, Issue 12, December 2016

III. PROPOSED ALGORITHM

A. *DESIGN CONSIDERATIONS:*

- Scene Capture.
- Text Localization
- Audio Output

B. *DESCRIPTION OF THE PROPOSED ALGORITHM:*

The scene seize component collects scenes containing objects of curiosity in the form of pix or video. In our proposed mannequin the camera hooked up to a pair of sunglasses. The information processing aspect is used for deploying our proposed algorithms, including, Object-of-interest detection to selectively extract the image of the object held by the user from the cluttered background or other neutral objects in the camera view. Text localization to obtain image regions containing text, and text recognition to transform.Image-based text information into readable codes. Text to audio conversion: once the audio output component is to inform the user of recognized text codes. Bluetooth earpiece with mini microphone is employed for speech output. In the propose work text detection and label reading using SWT (Stroke Width Transform) method and canny edge detection method.

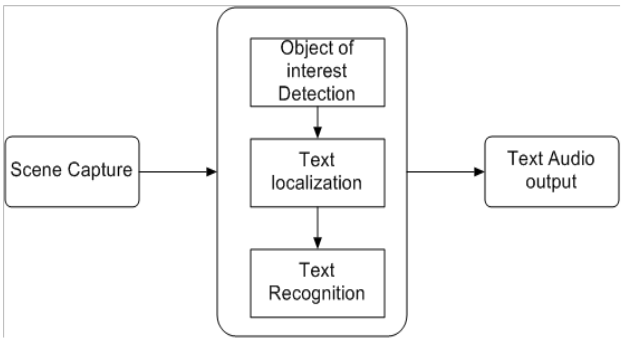

Fig Proposed System Architecture

Stroke width Transform (SWT):-The Stroke Width Transform to be (SWT) is a local image operator which computes per pixel the width of the obviously stroke containing the pixel. The output of the SWT is an snapshot of measurement equal to the size of the enter image the place each element includes the width of the stroke associated with the pixel. In determine(a).Can outline a stroke to be a contiguous a part of an image that varieties a band of a almost steady width. It does now not expect to know the exact width of the stroke however as a substitute get better it. The preliminary price of every element of the SWT is ready to. As a way to get better strokes, first compute edges within the snapshot using canny area detector. After that, a gradient course dp of every e edge pixel p is considered fig of SWT. If p lies on a stroke boundary, then dp ought to be roughly perpendicular to the orientation of the stroke. Then follow the ray r=p+n*dp, n>0 until another edge pixel q is found. consider then the gradient direction dq at pixel q. If dq is roughly opposite todp(dq= -dp± π/6)element sof the SWT output image corresponding pixel along segment[p,q] is assign width |p-q| unless it already has a lower value.Otherwise, if the matching pixel $\boldsymbol{q}$ is not found, or if $\boldsymbol{d_q}$is not opposite to $\boldsymbol{d_p}$, the ray is discarded.as show in fig of SWT

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering
(An ISO 3297: 2007 Certified Organization)
Website: www.ijircce.com
Vol. 4, Issue 12, December 2016

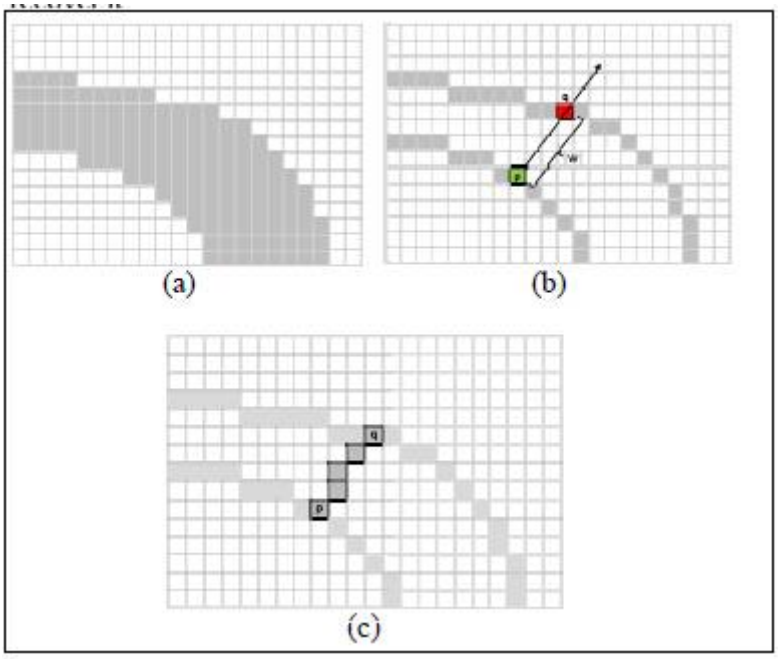

Fig of Stroke width Transform

The SWT values more complex situations, like corners will not be true stroke widths after the first pass. Therefore it pass along each non discarded ray again, compute median SWT value ***m*** of all its pixels, and then set all the pixel of ray with SWT values above ***m*** to be equal to m

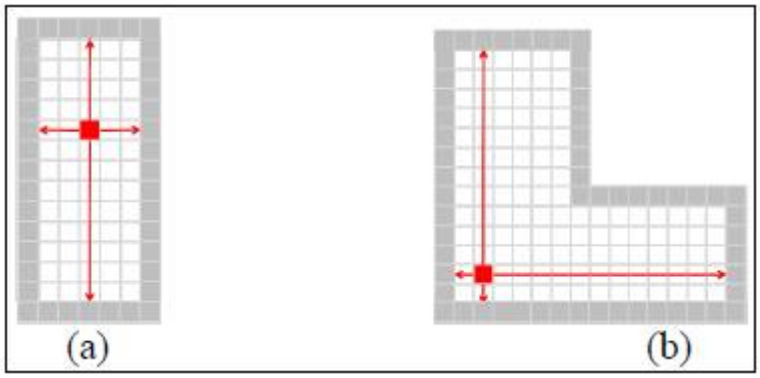

Fig of Example of SWT

Above figure shows filling pixels with SWT values. (a) An example red pixel is filled with minimum between the lengths of vertical and horizontal rays passing through it proper stroke width value is stored. (b) An example red pixel stores the min between two rays lengths; this is not the true stroke width this shows the necessity of the second pass.

Canny Edge Detection:-The purpose of edge detection almost always is to greatly scale down the quantity of knowledge in an image, while keeping the structural properties for use for extra snap-shot processing. Several algorithms exists, and this worksheet makes a specialty of a specified one developed with the aid of John F. Canny (JFC) in 1986 .Although it's really historic, it has come to be one of the most usual area detection. The aim of JFC was to develop an algorithm that's finest in terms of the following standards:

1. Detection: The chance of detecting real edge features will have to be maximized at the same time the likelihood of falsely detecting non-edge features must be minimized. This corresponds to maximizing the sign-to-noise ratio.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering
(An ISO 3297: 2007 Certified Organization)
Website: www.ijircce.com
Vol. 4, Issue 12, December 2016

2. Localization: The detected edges must be as close as possible to the actual edges.

3. No of responses: One actual part will have to now not result in a couple of detected aspect (one could argue that this is implicitly incorporated within the first requirement).

Canny edge Detector is perfect for a designated category of edges (often called step edges)it's apparent from that an image of the gradient magnitudes in most cases indicate the edges quite certainly.

IV. PSEUDO CODE

To implement the text detection system uses two method are compared to the stroke orientation to stroke width transform and the edge distribution of pixel to canny edge detection. These two method were choose because of their generally good performance. They improve text detection performance and give better quality result. In that system is work focus on only text detection. For text to speech part using Microsoft lib.Following are the steps of algorithm.

1. Capture image of the hand held object by portable camera.
2. Image capturing exactly localize text on object by using SWT method.
3. After this, conversion of image into grey scale takes place.
5. Firstly it maps edges text by using canny edge detection algorithm. Canny edge detector returns the co-ordinates of the text area.
6. Using this co-ordinate, cropping the text region from the image takes place.
7. Hereafter text detector using ASPOSE library is applied.
8. These steps returns the letter identification. Then these letters are grouped together into text.
9. Detected word is given to the text to speed converter. Text to speech converter converts text into respective audio format

V. EXPERIMENTAL RESULTS

Text classification based on the Cascade-Ad-boost classifier plays an principal role in text area localization. To evaluate the effectiveness of the textual content classifier, It performed first on collection of images on the dataset of pattern patches, in which the patches containing textual content are positive samples and people without textual content are negative samples. Each patch was assigned a prediction rating by the text classifier; a better rating indicates a higher chance of textual content information.

It define the proper constructive cost because the ratio of correct optimistic predictions to the total quantity of confident samples. Similarly, the false optimistic cost is the ratio of correct confident predictions to the complete quantity of positive predictions. The following graph shows the precession and recall value of sampleimage.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering
(An ISO 3297: 2007 Certified Organization)
Website: www.ijircce.com
Vol. 4, Issue 12, December 2016

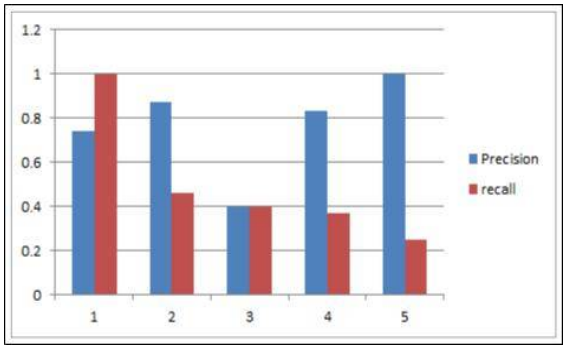

Fig of result of precision and recall of image data

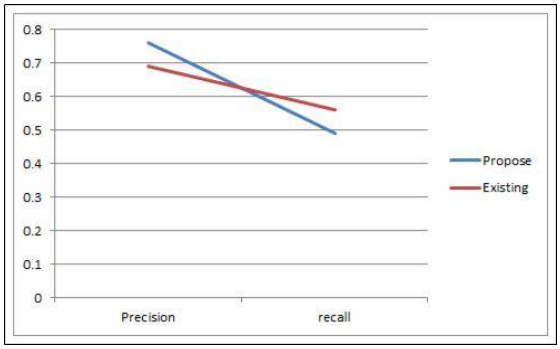

Fig of result analysis of existing & proposed sys.

According to existing system experimental result propose system is gives better result as show in graph. This graph describes the sample images and recall value that indicate image accuracy in percentage.

Method	Precision	Recall
Our	0.76	0.49
Chueai	0.69	0.56
Hinnerk Beck	0.62	0.67
AlexChen	0.60	0.60

Fig of: Different method Result Table

In this system managed the Through utilizing the equal analysis measures as above experiments, existing procedure obtained precision 0.52, recall o.62, this dataset. In that utilizing ICDAR 2013 dataset the precision is diminish than that on the strong reading dataset.The photos in the person-captured dataset have slash resolutions and extra compact distribution of text information, in order that they generate low-first-rate area maps and text boundaries, which result in unsuitable spatial layouts and textual content constitution aspects.According to propose system method good result as compare other result. In the propose system using ICDAR 2015 dataset reading challenges refer. Precision value is 0.76 and recall value is 0.49.The stroke width transform (Text detection) graph shows the SWT text detection result.

VI. CONCLUSION AND FUTURE WORK

The proposed work is applicable for visually impaired people, which helps them to read the labels of product. The proposed system work is more efficient and effective than existing system. In a propose work using a stroke width transform text detection method, and canny edge detection method for calculate edge base. It provides better accuracy for text detection.

REFERENCES

1. Chucai Yi and YingLiTian, ,Portable Camera-based Assistive "Text and Product Label Reading from Hand-held Objects for Blind Persons"IEEE/ASME Transactions
2. Xu-Cheng Yin, Xuwang Yin, Kaizhu Huang, and Hong-Wei Hao,Robust Text Detection in Natural Scene Images.
3. Xiangrong Chen1 Alan L. Yuille1 , Detecting and Reading Text in Natu-ral Scenes2 Departments of Statistics1, Psychology2, University of Califor-nia, Los Angeles, Los Angeles,CA90095
4. MattiPietikainen and Oleg Okun Machine Edge-Based Method for Text Detection from Complex Document Images Vision and Intelligent Systems Group Infotech Oulu and Department of Electrical Engineering, P.O.Box 4500 FIN-90014 University of Oulu, FINLAND.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering
(An ISO 3297: 2007 Certified Organization)
Website: www.ijircce.com
Vol. 4, Issue 12, December 2016

5. M SharmilaKumariAkshatha Local Features based Text Detection Techniques in Document Images.
6. X. Chen, J. Yang, J. Zhang, and A. Waibel, Automatic detection and recognition of signs from natural scenes,IEEETrans.Image Process.,vol. 13, no. 1, pp.8799, Jan.2004.
7. Anish Kumar J Assistive Text Reading System using Gradient and Stroke Orientation for Blind Person,
8. Advance Data Reports from the National Health Interview Survey (2008). International Workshop on Camera-Based Document Analysis and Recognition (CBDAR 2005, 2007, 2009, 2011). [Online]: Available: http://www.m.cs.osakafuu. ac.jp/cbdar2011
9. Shehzad Muhammad and Hanif, Lionel Prevost Text Detection and Localization in Complex Scene Images usin Constrained AdaBoost Algorithm Institute des Systems Intelligentset de Robotique, CNRS, UMR .

BIOGRAPHY

Shradha Vitthalrao Chalikwar, M.E Student, Department of Computer Engg, Indira College of Engineering and Management, Savitribai Phule Pune University, Maharashtra, India.