EXTRACTION OF TEXT FOR VISUALLY CHALLENGED PEOPLE USING SUPERPIXEL CLUSTERING AND OPTICAL CHARACTER RECOGNITION

[1]Soundarya.S..,[2]Saravanan.M
Student..,,Professor
Department of ECE,
IFET College of Engineering.,Villupuram..,India.

Abstract: — Focus scene content from picture is used a similar number of uses, in reality. Content distinguishing proof and affirmation in like manner scene can give gainful information for a few applications, for instance, ostensibly crippled individual, vehicle label acknowledgment, present day computerization, zone information etcetera. The pictures are subjected to heaps of hardship, for example, uneven lighting, text dimension, surface, clamor, undetectable. To proposed a calculation to dissect the content in complex scene utilizing superpixel clustering. Edges are distinguished utilizing haar DWT. The associated segments marking examines a picture and gatherings its pixels into segments in view of pixel network and restricting the content. The character extraction is utilized superpixel grouping of associated pixel with comparative highlights. OCR are separated the content in the picture. In this work, an approach has been tried to discrete and see content from scene pictures and change over that apparent substance into talk. This endeavor can be a drawing in control in an apparently tried person's life and can be strong in easing them of their dissatisfaction of not having the capacity to peruse whatever they need, along these lines upgrading the nature of their lives.

IndexTerms -.DWT, OCR, Superpixel clustering, Text to speech**,** SAPI, Visually challenged people

I. INTRODUCTION

Today the most of the information is available on paper or in the form of picture and large information is stored in images. The present innovation is confined to removing content against clean foundations. Content extraction [10] from camera based scene pictures is an extremely troublesome issue since it isn't generally conceivable to accurately characterize the highlights of content in a hued scene picture because of the wide varieties in conceivable arrangements for instance, geometry (area and introduction), shading closeness, textual style and size. Numerous calculation have been produced for remove content from the images for application such as address identification, robot navigation, number recognition and in portable or wearable computers. Text extraction includes discovery, restriction, extraction, improvement and acknowledgment of the content from the given picture.However, be that as it may, varieties of the content because of contrasts in estimate, style, introduction and arrangement and also low picture differentiation and complex foundation make the issue of programmed content extraction to a great degree testing. Research on the area and extraction of the writings in complex foundation has essential criticalness in flow data age [5]. Countless have been proposed to address this issue This text information can provide a much truer form of content–based access to the image and video documents [2] if it can be extracted and harnessed efficiently. This dissertation solves the problem involved in detecting text object in image and video and tracking text event in video.

Outwardly impeded people can read just by utilization of unique applications by them like Braille dialect. The downside of this framework is that each item does not give the content in Braille. In this paper we have proposed an assistive content perusing structure to help outwardly disabled people to peruse writings from different questions in their day by day lives. At first we catch the picture of the required, pre handling is performed on it. Pre handling incorporates steps like dim scale and binarization, protest of intrigue acknowledgment. In the proposed framework we are making the utilization of OTSU calculation to change over the dim scale picture into binarized one. The content areas from the caught picture are then extricated and perceived by utilizing optical character acknowledgment programming (OCR). The principle calculation in OCR specifically MODI is utilized here. This extricated content of different textual styles and sizes at that point can be perceived independently and afterward joined in a word giving its yield as sound utilizing Content to-discourse utilizing the SAPI libraries.

Perusing is exceptionally basic in our day by day lives. Out of 314 million outwardly impeded individuals all around the globe 45 Million are visually impaired and new cases being included every year. Late advancements in PC vision, computerized cameras, and versatile PCs make it attainable to help these people by creating camera-based items that join PC vision innovation with other existing

business items such optical character acknowledgment (OCR) frameworks. Printed content is wherever as reports, receipts, bank articulations, eatery menus, classroom presents, item bundles, pharmaceutical containers and so forth.

There are numerous assistive frameworks accessible today however they have certain issues lessening the possibility for the outwardly tested people. For instance, convenient scanner tag perusers intended to enable visually impaired individuals to recognize diverse items, it empowers the clients who are incognizant in regards to get to data about these items through discourse and Braille. There are frameworks like K Peruser Versatile it keeps running on a mobile phone and enables the client to peruse mail, receipts, fliers, and numerous different records. In any case, the report to be perused must be almost level, set on an unmistakable, dim surface (i.e., a no jumbled foundation), and contain for the most part message. Moreover, K Peruser Portable precisely peruses dark print on a white foundation yet has issues perceiving shaded content or content on a hued foundation. It can't read content with complex foundations. The principle point is to grow such a framework, to the point that will read the writings from complex foundations effectively.

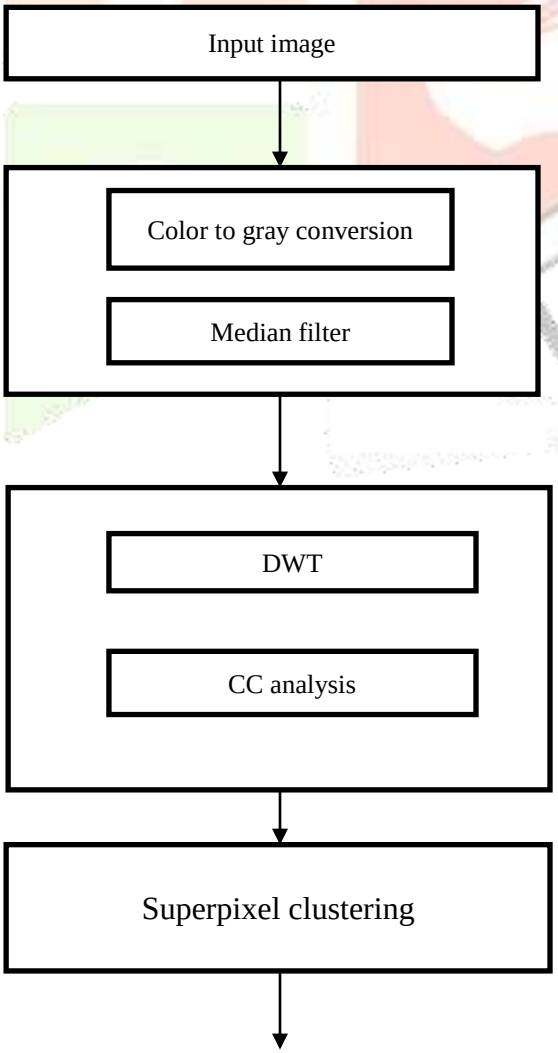
EXISTING METHOD

In prior stage extraction process are utilized to morphological tasks and heuristic separating. Adaboost classifier that decides the contiguousness relationship and group CCs by utilizing their match shrewd relations. At that point we standardize hopeful word districts and decide if every locale contains message or not. Since the scale, skew, and shade of every hopeful can be assessed from CCs, we build up a content/nontext classifier for standardized pictures. The disservice of morphological is obscure tumult over the photo making it less recognizable.

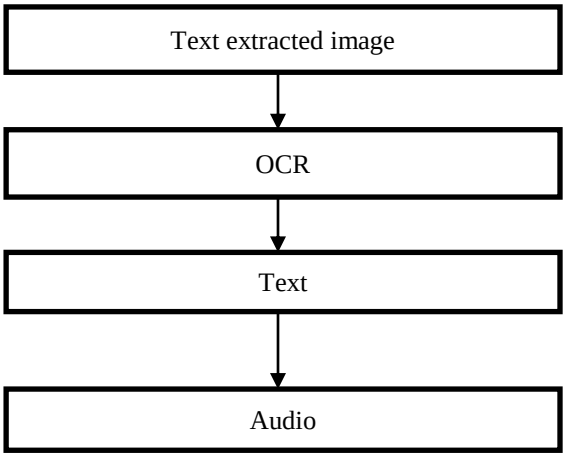
PROPOSED METHOD

The project method has of few steps: pre-processing, localization, extraction and optical character recognition and text to speech.

BLOCK DIAGRAM



.



PRE-PROCESSING

The picture is first transformed from RGB to dim level. In every pixel is triple (RGB) of intensities of red, green and blue. Grayscale is a scope of shades of diminish without obvious shading. Darkest shade is dull and lightest shade is white. Grayscale electronic picture is a photo in which the estimation of each pixel is a solitary illustration the conveys just power information. Middle channel is non straight strategy used to expel noise from images and is also smoothing method.

INPUT IMAGE

COLOR TO GRAY CONVERSION

TEXT LOCALIZATION

Identification and confinement [4] of the content is finished by utilizing extricating edges in the picture. Edge-construct strategies center in light of the high difference between the content and the foundation. The discrete wavelet change is an extremely valuable instrument for flag investigation and picture handling, particularly in multi-determination portrayal. It can break down flag into various segments in the recurrence space. Harr DWT are genuine, orthogonal, and symmetric. Its limit conditions are the least complex among all wavelet-based strategies. The base help property permits self-assertive spatial network interims. It can be used to analyse texture [6] and detect edges of characters. The high-pass filter and the low-pass filter coefficient is simple (either 1 or –1). One-dimensional discrete wavelet change (1-D DWT) breaks down an information grouping into two segments (the normal segment and the detail segment) by counts with a low-pass channel and a high pass channel. Two-dimensional discrete wavelet change (2-D DWT) breaks down an information picture into four sub-groups, one normal part (LL) and three detail segments (LH, HL, HH). In picture handling, the multi-determination of 2-D DWT has been utilized to identify edges of a unique picture. The conventional edge recognition channel [8] can give the comparable outcome also. In any case, 2-D DWT can identify three sorts of edges at any given moment while conventional edge discovery channels can't.

The task for Haar DWT is less difficult than that of some other wavelets. It has been connected to picture preparing particularly in multi determination portrayal [5]. Normal segments are recognized by the LL sub-band, Vertical edges are distinguished by the HL sub-band, Horizontal edges are identified by the LH sub-band, Diagonal edges are recognized by the HH sub-band. Connected components(cc) [1] are marking examines a picture and gatherings its pixels into segments in light of pixel availability, i.e. all pixel in a CC are share comparative pixel power esteems and are somehow associated with each other. The cc's or edges are recognized and afterward combining these, bouncing boxes for content are gotten. CC based strategies utilize a base up approach [7] by gathering little parts into progressively bigger segments until the point when all locales are distinguished in the picture. A geometrical examination is expected to blend the content parts utilizing the spatial game plan of the segments in order to sift through non-content segments and check the limits of the content locales. Associated segments are extricated from every content line area by nearby binarization and after that an associated part investigation approach in light of MRF demonstrate is utilized to shift through non-content segments and confine content lines precisely.

DWT

CC ANALYSIS

TEXT EXTRACTION

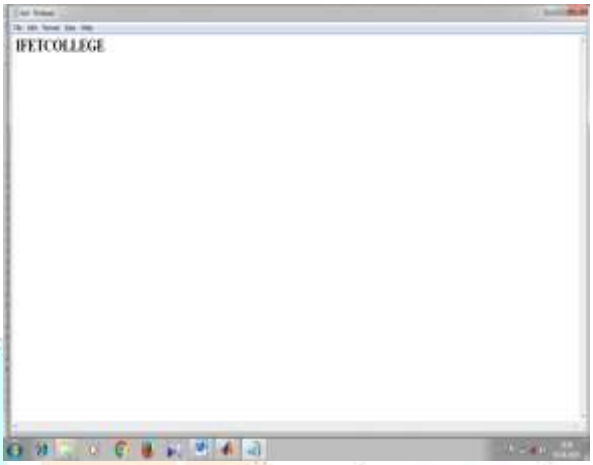
In this step the text pixel [3] are separated from the background pixels by localized region. Morphological sifting is gathering of non-straight activities identified with the shape or morphology of highlights in a picture, for example, limits, skeleton..,etc. the disservice of morphological is obscure commotion over the picture making it less noticeable. By using superpixel clustering to identifying a group of related pixels based on color similarity.

OPTICAL CHARACTER RECOGNITION (OCR)

Recognition of text is done with the OCR. It converts the text in the image into editable/computer readable data. (OCR) is the ID of printed characters utilizing photoelectric gadgets and PC programming. It coverts pictures of composed, manually written or printed content into machine encoded content from examined archive or from subtitle content superimposed on a picture. OCR is utilized as a part of machine process, for example, subjective figuring, machine interpretation, content to discourse, key information and content mining. It is the interpretation of optically filtered bitmaps of printed or composed content into carefully editable

information documents. An OCR encourages the transformation of geometric source protest into a carefully representable character in ASCII or Unicode plan of advanced character portrayal. OCRs are of two sorts: for perceiving printed characters and for written by hand message OCR process.

OCR OUTPUT



The ocr output of text are converted that into the audio [9] This work highlighted the efficacy of personal visual assistance systems in our day-to-day activities for the visually impaired people.

AUDIO CONVERSION

The microsoft Speech SDK 5.1 adds Automation support to the highlights of the past variant of the Speech SDK. would now be able to utilize the Win32 Speech API (sapi) to create discourse applications with Visual Basic ®, ECMA Script and other Automation dialects. All in all adaptations of the API have been planned with the end goal that a product engineer can compose an application to perform discourse acknowledgment and union by utilizing a standard arrangement of interfaces, available from an assortment of programming dialects. Moreover, it is feasible for to create their own Speech Recognition and content to discourse an adjust existing motors to work with SAPI. On a basic level, as long as these motors fit in with the characterized interfaces they can be utilized rather than the Microsoft-provided motors. As a rule the Speech API is an unreservedly redistributable part which can be delivered with any Windows application that desires to utilize discourse technology

There have been two fundamental 'families' of the Microsoft Speech API. SAPI forms 1 through 4 are for the most part like each other, with additional highlights in each more up to date form. SAPI 5 however was a totally new interface.. From that point forward a few sub-renditions of this API have been discharged Rather, every discussion to a runtime part (sapi.dll). There is an API executed by this segment which applications utilize, and another arrangement. The Speech SDK rendition 5.0, consolidating the SAPI 5.0 runtime.

AUDIO MODULE

This module will perceive and separate the content. This will be accomplished utilizing OCR-Optical Character Acknowledgment is the mechanical or electronic transformation of pictures of composed, manually written or printed content into machine-encoded content. We will use here MODI calculation of OCR. This module will get the separated content as an info and it will read out the content utilizing Content To Discourse accessible in the portable. On the off chance that no content is perceived at that point default sound yield will be given. This will be finished utilizing SAPI libraries.

CONCLUSION

To proposed in this work utilizes superpixel bunching for content extraction from scene pictures. After content extraction, they are given as contribution to OCR and after that the outcomes are shown in the scratch pad. From the test outcomes we can see that, the proposed calculation extricate content of various text style, size, and introduction proficiently. This approach fabricates a similitude grid in the wake of utilizing the SLIC superpixel calculation, and after that consolidations these superpixels into a few areas by the grouping calculation with the likeness matrix and to enhance the speed of content extraction utilizing this method. Individuals with

talk adversity individual can utilize along these lines to manage change composed words into vocalization. Trials have been performed to test the content perusing framework and great outcomes have been accomplished.

REFERENCE

[1] Kenji Suzuki, Isao Honda, Noboru Sugie, "Linear-time connected component labeling based on sequential local operations," Elsevier,
Computer vision and image understanding, vol. 89, Issue 1, pp.,1-23, January 2003.

[2] Keechul Jung , Kwang In Kim , Anil K. Jain , "Text information extraction in images and video: A survey," Elsevier ,Pattern Recognition, vol. 37, Issue 5, pp. 977-997, May 2004.

[3] Kumuda T., and L. Basavaraj, "Text extraction from natural scene images using region based methods-A survey," in Proceedings of ACEEE international conference on signal processing, image processing and computer vision, Issn. 22140344, DOI. 03, pp. 412-416, 2014.

[4] Chucai Yi, and Yingli Tian, "Localizing text in scene images by boundary clustering, Stroke segmentation, and string fragment classification," IEEE transactions on image processing, vol. 21, Issue 9, pp. 4256-4268, Sept 2012.

[5] SeongHun Lee, Min Su Cho, Kyomin Jungz, and Jin Hyung Kim, "Scene text extraction with edge constraint and text co-linearity," Pattern recognition (ICPR), 20th international conference on pattern recognition, pp. 3983-3986, Istanbul, Aug 23-26, 2010.

[6] Ranjit G., Anandarup R., and Swapan K. P., "Text extraction from scene images using statistical distribution," 3rd International conference on emerging application technology, 2012.

[7] Hyung il koo, and Duck Hoon Kim, "scene text detection via connected component clustering and nontext filtering," IEEE transactions on image processing, vol. 22, issue. 6, pp. 2296-2305, June, 2013.

[8] S. A. Angadi, and M. M. Kodabagi, "A texture based methodology for text region extraction from low resolution natural scene images," International journal of image processing, vol. 3,Issue. 5, 2010.

[9] Muhammad Shehzad Hanif, Lionel Prevost, "Texture based text detection in natural scene images: a help to blind people with vision and hearing impairments," Proceedings of the conference and workshop on Assistive technology for people with vision and hearing impairments .Granada ,Spain, August 28-31, 2007.

[10] Jian Liang, David Doermann, Huiping Li, "Camera-based analysis of text and documents:A survey," Springer, International Journal of document analysis and recognition (IJDAR), vol. 7, Issue 2, pp. 84-104, July 2005.