Content identification from regular scene pictures: towards a framework for visually impaired people

[1]A Deepika, [2]Vijaya Bhasker Reddy
[1]Assistant Professor, [2]Assistant Professor,
[1]Electronics and Communication Engineering,
[1]K.G.Reddy College of Engineering and Technology, Hyderabad, India.

Abstract : Here a framework has been proposed that peruses the content experienced in regular scenes with the mean to give help to the visually impaired people. This paper portrays the framework plan and assesses a few character extraction techniques. Programmed content acknowledgment from normal pictures gets a developing consideration due to potential applications in picture recovery, mechanical autonomy and clever transport framework. Camera-based archive examination turns into a genuine plausibility with the expanding determination and accessibility of advanced cameras. Notwithstanding, on account of a visually impaired individual, finding the content is the primary imperative issue that must be tended to, in light of the fact that it can't be expected that the obtained picture contains just characters. At to start with, our framework tries to discover in the picture regions with little characters. At that point it zooms into the discovered regions to retake higher determination pictures vital for character acknowledgment. In the present paper, It has been proposed four character-extraction techniques in light of associated segments. The execution of the diverse techniques relies upon character measure. In the information, greater characters are more common and the best extraction technique turns out to be the succession: Sobel edge discovery, Otsu binarization, connected component extraction and rule-based connected component filtering.
IndexTerms - Sobel edge discovery, Otsu binarization, connected component extraction and rule-based connected component filtering.

1. INTRODUCTION

Consistently, the quantity of outwardly weakened people is expanding because of eye sicknesses diabetes, activity accid ents and different causes. There are around 200,000 people with gained visual impairment in Japan. Consequently PC applications that offer help to the outwardly disabled people have turned into an essential subject. We have officially built up a pen - based character input framework for daze people utilizing a PDA [1]. On this framework, individuals with gained visual impairment recall the shape and composing request of Japanese characters and they can utilize this framework as a scratch pad and as an E-mail terminal whenever, anyplace. This application basically fills in as specialized device. In any case, such a gadget does not take care of the greater part of the issues experienced by a visually impaired individual willing to go outside unaccompanied.
At the point when an outwardly impeded individual is strolling around, it is imperative to get content data which is available in the scene. For instance, a 'stop' sign at an intersection without acoustic flag has a critical importance. As a rule, way finding into a man-made condition is helped extensively by the capacity to peruse signs. For instance, if the billboard of a store can be perused, the shopping wishes of the visually impaired individual can be fulfilled less demanding.

The exploration on content extraction from normal scene

Pictures have been developing over the most recent quite a while [2]. Numerous techniques have been proposed in light of edge recognition [3], binarization [4], spatial-recurrence picture examination [5] and scientific morphology operations [6]. Yang has proposed a sign acknowledgment and interpretation framework for voyagers [7]: characters are removed from pictures with Chinese sign sheets and meant English. There are likewise other parallel research endeavors to build up a scene-content perusing framework for the outwardly hindered [8]. Every one of these frameworks make apparent that the content regions can't be splendidly extricated from the picture since normal scenes comprises of complex items, once in a while profoundly finished, structures, trees, window edges et cetera, offering ascend to false content location and misses. The initial phase in building up our content perusing framework is to address the issue of content recognition in regular scene pictures. In this paper, we portray the framework plan and propose four content extraction techniques in view of associated parts.

Most investigations depend on a solitary technique for content discovery. We found that the viability of various strategies firmly relies upon character measure. Since in characteristic scenes the watched characters may have generally unique sizes, it is in this way hard to extricate all content zones from the picture utilizing just a solitary technique. This is particularly the case for this present reality pictures obtained by an outwardly hindered individual. Under the imagined utilization conditions, the camera disposition will be substantially less compelled than is the situation in current benchmark databases. We test the exactness of the proposed character extraction strategies on a recently accessible benchmark dataset gathered with the event of the ICDAR 2003 Robust Reading Competition. We likewise assess how the individual techniques can be joined for enhancing execution.

2. DESIGN OF THE SYSTEM

Figure 1 demonstrates the general setup of our proposed framework. The building components are the PDA, the CDD-camera and the voice combine r. Zooming, skillet tilt movement and auto-center are fundamental capacities required for the CCD-camera. Finding scene content includes two situations. To start with, in the 'stroll around mode', the camera which is put on the client's shoulder obtains a picture of the scene

naturally and after that the look for content ranges is performed utilizing strategies intended for little characters. On the off chance that a territory is distinguished, the camera zooms in to get a more point by point picture on which extraction techniques for extensive characters are utilized. These higher determination characters are then perceived and read out to the visually impaired pers on by means of a voice synthesizer. Obviously, a look adjustment work is required in this mode, with the end goal that the framework does not lose the objective competitor character range while the client is strolling. In this paper, nonetheless, we expect that the client is stopping when the pictures are caught

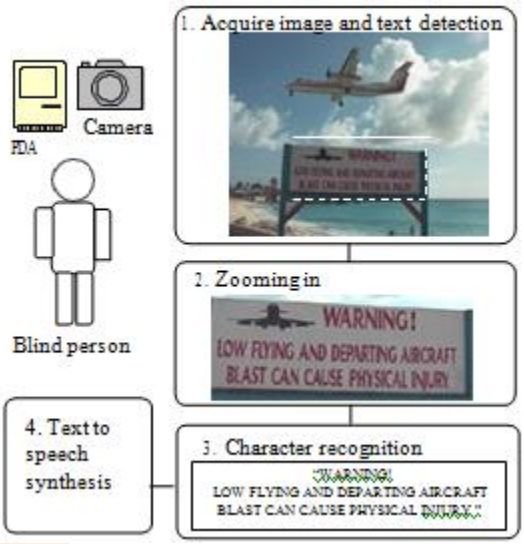
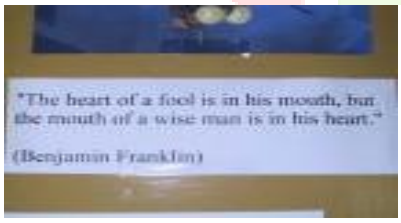


Figure1 System configuration (walk-around mode)

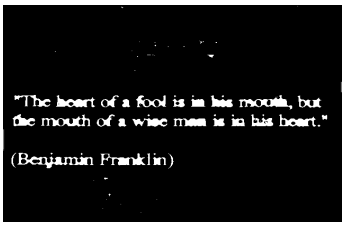
In a moment mode, the framework is utilized for perusing an eatery menu or a book cover. In this situation, the client can think about where the content is approximately and he/she can utilize the camera as a hand scanner. For this situation, picture determination should be higher than in the 'stroll around mode' since it is normal that the pictures will contain many characters.

3. EXTRACTION OF SMALL CHARACTERS USING NUMERICAL MORPHOLOGY OPERATIONS

The primary strategy we propose focuses on the little characters (not exactly around 30 pixels in stature) and it depends on scientific morphology operations. We utilize a changed best cap handling [6]. As a rule, top-cap differentiate upgrade is performed by ascertaining the contrast between the first picture and the picture got subsequent to applying the opening picture on the first picture. As a result, the best cap operation is appropriate when the pixels of the content characters have intension esteems which are adequately not quite the same as the foundation. For example, Gu [6] utilizes the contrast between shutting operation and the first picture for content recognition when character pixels have bring down qualities than the foundation (for light content on a dim foundation). This strategy is exceptionally powerful, nonetheless it turn out to be computationally costly if an expansive channel is utilized as a part of request to extricate extensive characters. We built up an invariant technique relevant to little characters. We utilize a circle channel with a span of 3 pixels and we take the contrast between the end picture and the opening picture. The sifted pictures are binarized and after that associated segments (CoCos) are removed (Fig.2b). This technique distinguishes associated content territories containing a few little characters. As western content comprises of series of characters that are generally on a level plane set, we take on a level plane long territories ($1 < \text{width/stature} < 25$) from the yield picture as the last hopeful content locales (Fig.2c).



a. Original image



b. Difference between image versions subjected to the closing and opening

c. extracted characters are found by a mask operator between the original image and the bounding

Figure2 Extraction of small character text using morphological operation

4. THREE EXTENSIVE TECHNIQUES FOR EXTRACTING CHARACTERS

Here three extraction techniques for extensive characters (more than around 30 pixels in tallness) has been proposed. The initial two depend on Sobel edge discovery and the third depends on RGB shading data. In the general framework, these strategies ought to be utilized in the wake of zooming into the territories at first found by the morphological operations. Every strategy extricates associated segments that speak to competitor content zones. Choice principles in view of the sizes and relative situating of these territories are a while later used to prune the quantity of potential outcomes and lessen the vast number of false hits.

Character extraction from the edge picture

In this strategy, Sobel edge recognition is ap utilized on each shading channel of the RGB picture. The three edge pictures are then joined into a solitary yield picture by taking the greatest of the three edge esteems relating to every pixel. The yield picture is binarized utilizing Otsu's strategy [9] lastly CoCos are separated.

This strategy will come up short when the edges of a few characters are lumped together into a solitary substantial CoCo that is disposed of by the determination rules. This regularly happens when the content characters are near each other or when the foundation isn't uniform.

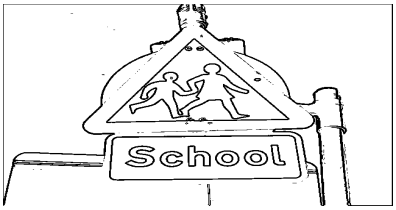
Character extraction from the turn around edge picture This technique is reciprocal to the past one; the paired picture is switched before associated part extraction. It will be successful just when characters are encompassed by associated edges and the internal ink range isn't broken (as on account of boldface characters).

Colour Shading based character extraction

The three strategies proposed up to this point utilize morphological and edge data for content discovery. In any case, shading data is likewise imperative, on the grounds that, normally, related characters in a content have practically a similar shading for a given occurrence experienced in the scene. The initial step is to disentangle the shading space and we decrease it to 8 hues by the accompanying strategy. We are applying Otsu binarization autonomously on the three RGB shading channels. Every pixel would now be able to have just 23 = 8 conceivable mixes of shading esteems. We isolate the 8 double pictures , then we separate and select CoCos on every one freely .

a. Original image

b. Edge based method



c. Reverse image of edge

d. colour based

Figure 3 Example of an image with medium-size character

Associated segment determination rules

It can be seen that, up to now, the proposed techniques are extremely broad in nature and not particular to content identification. Of course, a significant number of the removed CoCos don't really contain content characters. Now straightforward principles are utilized to sift through the false recognitions. We force limitations on the angle proportion and range size to diminish the quantity of non-character hopefuls. In

Fig. 4, Wi and Hi are the width and height of an extracted range; delta x and delta y are the separations between the focuses of gravity of every territory. Angle proportion is processed as width/height.

An essential perception is that, for the most part, content characters don't seem alone, however together with different characters of comparative measurements and ordinarily consistently put in a flat string. We utilize the accompanying tenets to additionally dispose of from all the distinguished CoCos those that don't really compare to content characters (Fig.4):

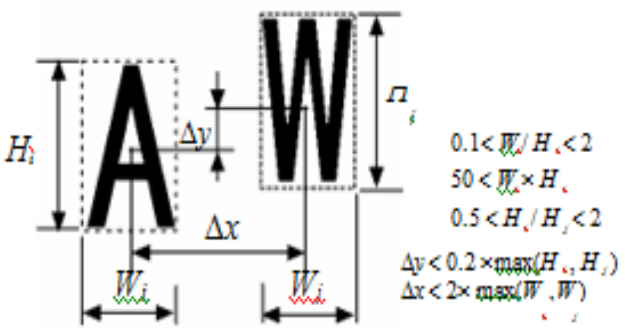


Figure 4 Character strings and rules

The framework experiences all blends of two CoCos and just those agreeing to all the choice standards prevail with regards to turning into some of the last proposed content locale.

Figure 5 Final result for the example given in Fig. 3

5. CONCLUSION

In this paper, the design has been presented of a scene-text detection module within a reading system for visually impaired persons. As the first step in the development of this system, four connected-component-based methods for text detection have been implemented and evaluated. The most effective proves to be the sequence: Sobel edge detection, Otsu binarization, connected-component extraction and rule-based connected-component selection. A high recall rate can be achieved by collecting all the candidate text areas proposed by the four individual methods. However, current results are not enough for practical use. Future work will focus on new methods for extracting small text characters with higher accuracy.

REFERENCES

[1] N. Ezaki, K. Kiyota, H. Takizawa and S. Yamamoto, "Pen-based Ubiquitous Computing System for Visually Impaired Person", human-computer interaction, Theory and Practice (PartII), Volume2, 2003, pp.48-52.

[2] D. Doerma nn, J. Liang, and H. Li, "Progress in Camera-
Based Document Images Analysis ", *Proc.of the ICDAR*, 2003, pp. 606-616.

[3] T. Yamaguchi, Y.Nakano, M. Maruyama, H. Miyao and T.Hananoi, "Digit Classification on Signboards for Telephone Number Recognition", *Proc.of the ICDAR*, 2003, pp.359-363.

[4] .Matsuo , K.Ueda and M.Umeda, "Extraction of Character String from Scene Image by Binarizing Local Target Area", *T-IEE Japan*, Vol. 122-C(2), 2002, pp.232-241.

[5] Y. Liu, T. Yamamura, N. Ohnishi and N. Sugie, "Extraction of Character String Regions from a Scene Image", *IEICE Japan*, D-II, Vol. J81, No.4, 1998, pp.641-650.

[6] L. Gu, N. Tanaka, T. Kaneko and R.M. Haralick, "The Extraction of Characters from Cover Images Using Mathematical Morphology", *IEICE Japan*, D-II, Vol. J80, No.10, 1997, pp. 2696-2704.

[7] J. Yang, J. Gao, Y. Zhang, X. Chen and A. Waibel, "An
Automatic Sign Recognition and Translation System", *Proceedings of the Workshop on Perceptive User Interfaces (PUI'01)*, 2001, pp. 1-8.

[8] A. Zandifar, R. Durais wami, A. Chahine, and L. Davis, "A Video Based Interface to Textual Information for the Visually Impaired", *IEEE 4th ICMI*, 2002, pp.325-330.