Scientific Journal of Impact Factor (SJIF): 4.14
e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 3, Issue 6, June -2016

PORTABLE CAMERA-BASED ASSISTIVE TEXT AND PRODUCT LABEL READING FROM HAND-HELD OBJECTS FOR BLIND PERSONS

Jay Narayanrao Paul[1], Dr.V.N.Nitnaware[2]

[1, 2] *Student of Dept. of ENTC Engineering, Dr.D.Y.Patil School of Engineering Academy , Ambi , Pune*

Abstract — *This paper provides a digital camera-based totally product records reader to help blind individuals to examine statistics of the goods. digital camera acts as main vision in detecting the label photo of the product then image is processed internally and separates label from image through using MATLAB and eventually identifies the product call and identified product records is said thru the optical person popularity (OCR).We suggest a digital camera-based assistive text analysing framework to help blind men and women examine text labels and product packaging from hand-held gadgets in their each day lives. To isolate the object from cluttered backgrounds or other surrounding items inside the digital camera view, we first advise an green and effective movement based technique to define a location of interest (ROI) inside the video with the aid of asking the consumer to shake the object. This technique extracts moving item location by using a mixture-of-Gaussians-based historical past subtraction approach. Inside the extracted ROI, textual content localization and popularity are carried out to acquire text records. To automatically localize the textual content regions from the item ROI, we advocate a singular textual content localization algorithm by way of studying gradient features of stroke orientations and distributions of side pixels in an Adaboost version. Text characters inside the localized text regions are then binaries and recognized by off-the-shelf optical individual popularity software.*

The recognized textual content codes are output to blind customers in speech. Performance of the proposed textual content localization set of rules is quantitatively evaluated on ICDAR-2003 and ICDAR-2011 sturdy reading Data-sets. Experimental outcomes display that our set of rules achieves the state of the arts. The evidence-of-concept prototype is also evaluated on a data-set accumulated using ten blind folks to evaluate the effectiveness of the gadget's hardware. We explore person interface problems and determine robustness of the set of rules in extracting and analysing text from exceptional items with complex backgrounds.

Keywords*-ROI, optical character recognition (OCR), text area localization, text extraction.*

I. INTRODUCTION

Reading is manifestly critical in today's society. Published textual content is anywhere in the shape of stories, receipts, financial institution statements, eating place menus, study room handouts, product programs, commands on remedy bottles, etc. And whilst optical aids, video magnifiers, and display readers can assist blind users and people with low vision to get right of entry to files, there are few devices which could provide proper access to common hand held objects inclusive of product applications, and items printed with textual content such as prescription medicinal drug bottles. The capability of people who are blind or have huge visible impairments to examine revealed labels and product programs will enhance unbiased residing and foster financial and social self-sufficiency. Nowadays, there are already a few structures which have a few promises for transportable use, but they can't handle product labeling. For example, transportable bar code readers designed to assist blind people discover distinct merchandise in an in depth product database can enable customers who are ignorant of get entry to information approximately those merchandise [22] thru speech and braille. However a massive hassle is that it's far very tough for blind users to locate the location of the bar code and to properly point the bar code reader on the bar code. latest traits in pc imaginative and prescient, digital cameras, and transportable computers make it possible to help those people by way of growing digital camera-based products that combine laptop imaginative and prescient technology with other present industrial products such optical individual reputation (OCR) systems.

However, the record to be examined must be nearly flat, positioned on a clean, dark surface, and contain often textual content. Moreover, KReader cell correctly reads black print on a white history, but has troubles spotting coloured textual content or text on a colored historical past. It cannot study text with complex backgrounds, text printed on cylinders with warped or incomplete images (which includes soup cans or medicinal drug bottles).

Moreover, these structures require a blind user to manually localize regions of interest and text regions on the items in most instances. despite the fact that a number of analyzing assistants have been designed particularly for the visually impaired, to our knowledge, no existing studying assistant can examine textual content from the forms of hard patterns and backgrounds found on many regular industrial products.

II. LITERATURE REVIEW

1. X. Chen and A. L. Yuille, "Detecting and reading text in natural scenes," in *Proc. Comput. Vision Pattern Recognit.*, 2004, vol. 2, pp. II-366–II-373.

This paper gives an algorithm for detecting and reading text in natural images. The algorithm is intended for use by blind and visually impaired subjects walking through city scenes. We first obtain a dataset of city images taken by blind and normally sighted subjects. From this dataset, we manually label and extract the text regions. Next we perform statistical analysis of the text regions to determine which image features are reliable indicators of text and have low entropy (i.e. feature response is similar for all text images).We obtain weak classifiers by using joint probabilities for feature responses on and off text. These weak classifiers are used as input to an AdaBoost machine learning algorithm to train a strong classifier. In practice, we trained a cascade with 4 strong classifiers containg 79 features. An adaptive binarization and extension algorithm is applied to those regions selected by the cascade classifier. Commercial OCR software is used to read the text or reject it as a non-text region. The overall algorithm has a success rate of over 90%.

2. Automatic detection and recognition of signs from natural scenes

In this paper, we present an approach to automatic detection and recognition of signs from natural scenes, and its application to a sign translation task. The proposed approach embeds multiresolution and multiscale edge detection, adaptive searching, color analysis, and affine rectification in a hierarchical framework for sign detection, with different emphases at each phase to handle the text in different sizes, orientations, color distributions and backgrounds. We use affine rectification 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.
We propose a local intensity normalization method to effectively handle lighting variations, followed by a Gabor transform to obtain local features, and finally a linear discriminant analysis (LDA) method for feature selection. We 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.

3. WEARABLE OBSTACLE AVOIDANCE ELECTRONIC TRAVEL AIDS FOR BLIND: A SURVEY

The last decades a variety of portable or wearable navigation systems have been developed to assist visually impaired people during navigation in known or unknown, indoor or outdoor environments. There are three main categories of these systems: electronic travel aids (ETAs), electronic orientation aids (EOAs), and position locator devices (PLDs). This paper presents a comparative survey among portable/wearable obstacle detection/avoidance systems (a subcategory of ETAs) in an effort to inform the research community and users about the capabilities of these systems and about the progress in assistive technology for visually impaired people. The survey is based on various features and performance parameters of the systems that classify them in categories, giving qualitative-quantitative measures. Finally, it offers a ranking, which will serve only as a reference point and not as a critique on these systems.

4. Texture-based approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm

The current paper presents a novel texture-based method for detecting texts in images. A support vector machine (SVM) is used to analyze the textural properties of texts. No external texture feature extraction module is used, but rather the intensities of the raw pixels that make up the textural pattern are fed directly to the SVM, which works well even in high-dimensional spaces. Next, text regions are identified by applying a continuously adaptive mean shift algorithm (CAMSHIFT) to the results of the texture analysis.
The combination of CAMSHIFT and SVMs produces both robust and efficient text detection, as time-consuming texture analyses for less relevant pixels are restricted, leaving only a small part of the input image to be texture-analyzed.

5. Context-based Indoor Object Detection as an Aid to Blind Persons Accessing Unfamiliar Environments

Independent travel is a well-known challenge for blind or visually impaired persons. In this paper, we propose a computer vision based indoor way finding system for assisting blind people to independently access unfamiliar buildings. In order to find different rooms (i.e. an office, a lab, or a bathroom) and other building amenities (i.e. an exit or an elevator), we incorporate door detection with text recognition. First we develop a robust and efficient algorithm to detect doors and elevators based on general geometric shape, by combining edges and corners.
The algorithm is generic enough to handle large intra-class variations of the object model among different indoor environments, as well as small inter-class differences between different objects such as doors and elevators. Next, to

distinguish an office door from a bathroom door, we extract and recognize the text information associated with the detected objects. We first extract text regions from indoor signs with multiple colors.
Then text character localization and layout analysis of text strings are applied to filter out background interference. The extracted text is recognized by using off-the-shelf optical character recognition (OCR) software products. The object type, orientation, and location can be displayed as speech for blind travelers.

III. SURVEY OF PROPOSED SYSTEM

This paper offers a prototype gadget of assistive textual content reading. The system framework includes three functional additives: scene capture, statistics processing, and audio output. The scene capture component collects scenes containing gadgets of interest in the form of pix or video. In our prototype, it corresponds to a camera connected to a couple of sunglasses. The facts processing thing is used for deploying our proposed algorithms, which includes 1) item-of-interest detection to selectively extract the image of the item held via the blind consumer from the cluttered historical past or different neutral gadgets in the digital camera view; and a couple of) text localization to reap picture areas containing text, and text popularity to convert picture-based totally text statistics into readable codes. We use a mini pc as the processing tool in our modern-day prototype machine.
The audio output component is to inform the blind consumer of identified textual content codes. A Bluetooth earpiece with mini microphone is employed for speech output. This simple hardware configuration guarantees the portability of the assistive text analyzing machine. Fig. four depicts a piece flowchart of the prototype gadget. A body collection V is captured via a digital camera worn by using blind customers, containing their handheld items and cluttered history.

VI. METHDOLOGY

1. OBJECT REGION DETECTION

To ensure that the hand-held object appears in the camera view, we employ a camera with a reasonably wide angle in our prototype system (since the blind user may not aim accurately). However, this may result in some other extraneous but perhaps text-like objects appearing in the camera view for example, when a user is shopping at a supermarket).
To extract the hand-held object of interest from other objects in the camera view, we ask users to shake the hand-held objects containing the text they wish to identify and then employ a motion-based method to localize the objects from cluttered background. Background subtraction (BGS) is a conventional and effective approach to detect moving objects for video surveillance systems with stationary cameras.

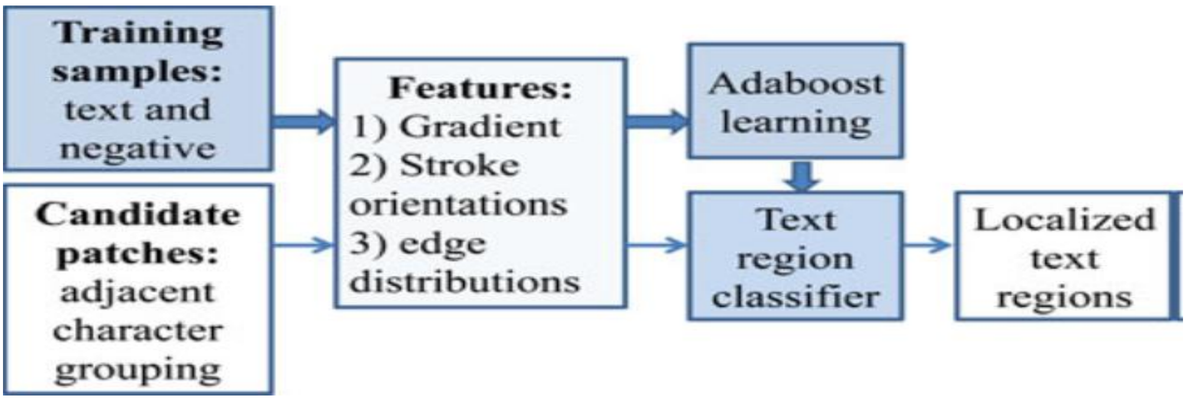


Fig. Diagram of the proposed Adaboost-learning-based text region localization Algorithm by using stroke orientations and edge distributions.

To detect moving objects in a dynamic scene, many adaptive BGS technique have been developed. Stauffer and Grimson modeled each pixel as a mixture of Gaussians and used an approximation to update the model. A mixture of K Gaussians is applied for BGS, where K is from 3 to 5. In this process, the prior weights of K Gaussians are online adjusted based on frame variations. Since background imagery is nearly constant in all frames, a Gaussian always compatible with its subsequent frame pixel distribution is more likely to be the background model.

2. AUTOMATIC TEXT EXTRACTION

We design a learning-based algorithm for automatic localization of text regions in image. In order to handle complex backgrounds, we propose two novel feature maps to extracts text features based on stroke orientations and edge distributions, respectively. Here, stroke is defined as a uniform region with bounded width and significant extent. These feature maps are combined to build an Adabost based text classifier

Text Stroke Orientation

Text characters consist of strokes with constant or variable orientation as the basic structure. Here, we propose a new type of feature, stroke orientation, to describe the local structure of text characters. From the pixel-level analysis, stroke orientation is perpendicular to the gradient orientations at pixels of stroke boundaries. To model the text structure by stroke orientations, we propose a new operator to map a gradient feature of strokes to each pixel. It extends the local structure of a stroke boundary into its neighborhood by gradient of orientations. We use it to develop a feature map to analyze global structures of text characters.

3. TEXT RECOGNITION AND AUDIO OUTPUT

Text recognition is performed by off-the-shelf OCR prior to output of informative words from the localized text regions. A text region labels the minimum rectangular area for the accommodation of characters inside it, so the border of the text region contacts the edge boundary of the text character. However, our experiments show that OCR generates better performance if text regions are first assigned proper margin areas and binarized to segment text characters from background.

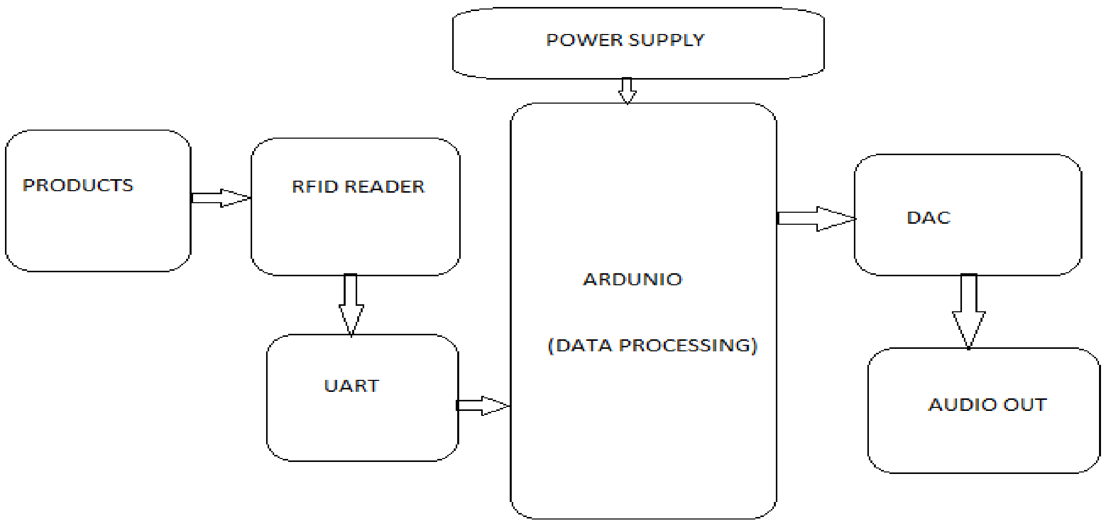
Thus, each localized text region is enlarged by enhancing the height and width by 10 pixels, respectively, and then, we use Otsu's method to perform binarization of text regions, where margin areas are always considered as background. We test both open- and closed-source solutions that allow the final stage of conversion to letter codes (e.g. OmniPage, Tesseract, ABBYReader).

Datasets

Two datasets are used to evaluate our algorithm. First, the ICDAR Robust Reading Dataset is used to evaluate the proposed text localization algorithm. The ICDAR-2003 dataset contains 509 natural scene images in total. Most images contain indoor or outdoor text signage. The image resolutions range from 640 × 480 to 1600 × 1200.

Since layout analysis based on adjacent character grouping can only handle text strings with three or more character members, we omit the images containing only ground truth text regions of less than three text characters. Thus, 488 images are selected from this dataset as testing images to evaluate our localization algorithm.

IV. SYSTEM ARCHITECTURE



V. CONCLUSION

We've described a prototype device to examine published textual content handy-held objects for assisting blind men and women so one can remedy the common aiming problem for blind users, we've got proposed a motion-primarily based method to come across the item of interest, even as the blind user actually shakes the item for multiple seconds. This method can correctly distinguish the object of interest from historical past or other items inside the camera view. To extract text regions from complicated backgrounds, we've proposed a unique text localization set of rules based totally on fashions of stroke orientation and facet distributions. The corresponding function maps estimate the worldwide structural function of text at each pixel. Block patterns mission the proposed feature maps of a photograph patch right into a feature vector. Adjacent man or woman grouping is performed to calculate applicants of textual content patches prepared for text classification. An advert enhance getting to know version is employed to localize text in digital camera-based totally photographs. Off-the-shelf OCR is used to carry out phrase recognition on the localized textual content areas and transform into audio output for blind customers. Our destiny work will extend our localization set of rules to system text strings with characters fewer than 3 and to layout more strong block patterns for text characteristic extraction. we are able to also enlarge our algorithm to deal with non-horizontal textual content strings. Moreover, we are able to deal with the extensive human interface troubles associated with studying textual content by using blind users

VI. REFERENCES

[1] "10 facts about blindness and visual impairment", World Health Organization: Blindness and visual impairment, 2009. www.who.int/features/factfiles/blindness/blindness_facts/en/index.html

[2] Advance Data Reports from the National Health Interview Survey, 2008. http://www.cdc.gov/nchs/nhis/nhis_ad.htm.

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

[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] D. Dakopoulos and N. G. Bourbakis, "Wearable obstacle avoidance electronic travel aids for blind: a survey," In *IEEE Transactions on systems, man, and cybernetics*, Vol. 40, No. 1, pp. 25-35, 2010.

[7] B. Epshtein, E. Ofek and Y. Wexler, "Detecting text in natural scenes with stroke width transform," In *CVPR*, pp. 2963-2970, 2010.

[8] Y. Freund and R. Schapire, "Experiments with a new boosting algorithm," In *Int. Conf. on Machine Learning*, pp.148–156, 1996.

[9] N. Giudice and G. Legge, Blind navigation and the role of technology, in The engineering handbook of smart technology for aging, disability, and IEEE/ASME Transactions on Mechatronics 10

[10]X. Yang, S. Yuan, and Y. Tian, Recognizing Clothes Patterns for Blind People by Confidence Margin based Feature Combination, In *ACM Multimedia*, pp. 1097-1100, 2011.

[11] A. Huete, J. Victores, S. Martinez, A. Gimenez, and C. Balaguer. Personal Autonomy Rehabilitation in Home Environment by a Protable Assistive Robot. In *IEEE Trans. Systems, Man, and Cybernetics, Part C: Applications and Reviews*, Vol. 42, No. 4, pp. 561-570, 2011.

[12] ICDAR 2011 Robust Reading Competition: A. Shahab, F. Shafait, and A. Dengel. ICDAR Robust Reading Competition Challenge 2: Reading Text in Scene Images. In Proceedings of ICDAR, pp. 1491-1496, 2011.

[13] K. Kim, K. Jung, and J. Kim, "Texture-based approach for text detection in images using SVM and continuously adaptive mean shift algorithm," In *IEEE Trans. on PAMI*, Vol.25, No.12, pp.1631-1639, 2003.

[14] "KReader Mobile User Guide", knfb Reading Technology Inc, http://www.knfbReading.com

[15] S. Kumar, R. Gupta, N. Khann