ORIGINAL RESEARCH PAPER

Electronics Science

Volume : 6 | Issue : 11 | November 2016 | ISSN - 2249-555X | IF : 3.919 | IC Value : 74.50

Portable Camera- Based Assistive Text Reading Framework for Blind People by Effective Motion Based Method

KEYWORDS

Camera; Framework; Region of Interest; ICDAR; Gaussians; Algorithm; Robust Reading Datasets.

G. Nandagopal

Assistant Professor, Department of Electronics Science, Mohamed Sathak College of Arts and Science, Sholinganallur

ABSTRACT We propose a camera-based assistive text reading framework to help blind persons read text labels and product packaging from hand-held objects in their daily lives. To isolate the object from cluttered backgrounds or other surrounding objects in the camera view, we first propose an efficient and effective motion based method to define a region of interest (ROI) in the video by asking the user to shake the object. This method extracts moving object region by a mixture-of-Gaussians-based background subtraction method. In the extracted ROI, text localization and recognition are conducted to acquire text information. To automatically localize the text regions from the object ROI, we propose a novel text localization algorithm by learning gradient features of stroke orientations and distributions of edge pixels in an Adaboost model. Performance of the proposed text localization algorithm is quantitatively evaluated on ICDAR-2003 and ICDAR-2011 Robust Reading Datasets. Experimental results demonstrate that our algorithm achieves the state of the arts.

Introduction

The purpose of this project is camera-based assistive text reading framework to read text labels especially product label reading from hand held objects. So this technology essentially helps the people like physically challenged. Most essentially blind peoples can use this technology to recognize the hand held product. For that they need only Tablet or a Laptop and Software tool to process it. Here we are using software tool called MATLAB which essentially helps to convert the text label into audio. The first step is placing the product label in front of the camera which captures the image and after multiple process it gives a audio about the hand held product. To isolate the object from cluttered backgrounds or other surrounding objects in the camera view, we first propose an efficient and effective motion based method to define a region of interest (ROI). This method extracts moving object region by a mixture-of-Gaussians-based background subtraction method. It is important to keep in mind that enhancement is a very subjective area of image processing. By this multiple process we can convert a text label into an audio.

Materials and Methods

Text localization is then performed on the camera based image. The Cascade Ada boost classifier confirms the existence of text information.

Cascade Classifier

The work with a cascade classifier includes two major stages: training and detection. Detection stage is described in a documentation of object detect module of general Open CV documentation. Documentation gives some basic information about cascade classifier.

Training Data Preparation

For training we need a set of samples. There are two types of samples: negative and positive. Negative samples correspond to non object images. Positive samples correspond to images with detected objects. Set of negative samples must be prepared manually, whereas set of positive samples is created using opencv_create samples utility.

Character Detection

We use Open CV (open source computer vision)library to process the images so that features for each letter could be extracted.

Image Capturing

First, we get the frames continuously from the camera and send it to the process. Once the object of interest is extracted from the camera image using cascade classifier, subsequent process can be done using following steps.

Conversion to Gray Scale

Conversion to grey scale can be done in Open CV. The reason we had to convert our image to gray scale was because threshold could be applied to monochrome pictures only In an 8 bit image each pixel is represented by one number from 0 to 255 where 0 is black and 255 is white. The extremism is searched over the whole array, selected ROI (in case of Image) or, if mask is not NULL Convert Scale has several different purposes and thus has several synonyms. It copies one array to another with optional scaling, which is performed first, and/or optional type conversion, performed and get the binary output.

Existing System for Text Reading

Today, there are already a few systems that have some promise for portable use, but they cannot handle product labelling. For example, portable bar code readers designed to help blind people identify different products in an extensive product database can enable users who are blind to access information about these products. Most state-of-the-art OCR software cannot directly handle scene images with complex backgrounds. 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, fliers, and many other documents. Such as text information can appear in various scales, fonts, colours, and orientations. Sliding window based methods, also known as region-based methods.

This method uses a sliding window to search for possible texts in the image and then use machine learning techniques to

Volume : 6 | Issue : 11 | November 2016 | ISSN - 2249-555X | IF : 3.919 | IC Value : 74.50

identify the text. This method limits the search of text to a subset of image rectangles. So, it reduces the number of subsets checked for the presence of text.

Connected Components (cc) Based methods
It extracts the character candidates from an image by connected component analysis, followed by grouping character candidates into text; additional checks may be performed to remove false positives.

A) Method based on Edges
This method is based on the factor like edge of character; edge is reliable feature of the text regardless of color/intensity, layout, orientations, etc. As the text region has high contrast to it's background, the edges of character can be easily detected. There are two steps used in this method: first, an edge extraction algorithm (such as canny edge detector) is used to get the edges and second, smoothing algorithm or morphology is used for edges connections and obtaining a full character boundary.

B) Method based on Color
In this method, color clustering is done by categorizing the pixels with same or similar colors and forming a candidate region. Then the candidate regions are analyzed and the CC is estimated. The main challenge of this method is the degree of clustering.

C) The Method based on the Combining of Edges and Color
Some methods combine Method 1 and Method 2, which detects both edges and colors of the text. This method has achieved better results by combining both features together than using these features separately.

Hybrid Methods
This method uses a region detector to detect text candidates and extracts connected components as character candidates by local binearization; non-characters are eliminated with a Conditional Random Fields model, and characters can finally be grouped into text. But it is Hard-to-segment the texts.

Proposed System
To automatically localize the text regions from the object ROI, we propose a novel text localization algorithm by learning gradient features of stroke orientations and distributions of edge pixels in an Adaboost model. Exploits higher order properties of text, Uses exhaustive search for pruning Estimate class conditional probabilities of ERs using trained classifier. Detects char in low quality, low resolution, strong noises, low contrast non-text MSERs are efficiently removed.

Optical Character Recognition(OCR)
OCR stands for optical character recognition, a wonderful and marvelous technology. It enables you to convert previously printed text material into information your computer can understand, without having to retype it. That is, if you scan a typewritten page, you would have a picture of the typewritten page, not the typed letters themselves. But with OCR software, you can scan the printed page and the software figures out what characters those little shapes are supposed to be, and it turns the scanned image into real text that you can put into your word processor!

Since the OCR software almost always makes some mistakes, you should then pass the text through a spelling checker, Then you can edit the text as much as you like, or use it as a story in a newsletter, or add it to your thesis paper. Some people use this technology effectively right now, and some complain about the efficiency of it- it isn't really much faster than typing-but OCR software is getting better all the time. OCR also refers to a special font that an OCR system can recognize easily you know, those typefaces that look like a computer did it.

Result

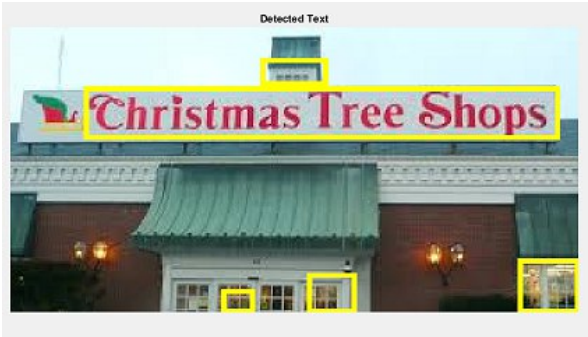

Fig 1: Detected text

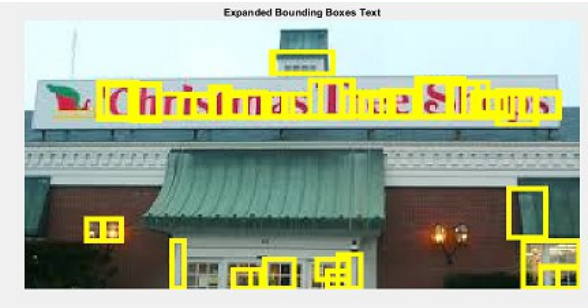

Fig 2: Expanded bounding boxes text

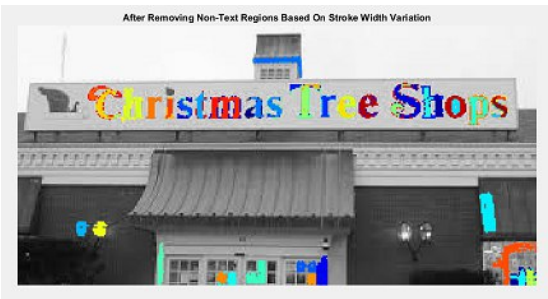

Fig 3:After removing Non-Text

Based On Geometric Properties

Conclusion
In this paper, we have described a prototype system to read printed text on hand-held objects for assisting blind persons. In order to solve the common aiming problem for blind users, we have proposed a motion-based method to detect the object of interest, while the blind user simply shakes the object for a couple of seconds. This method can effectively distinguish the object of interest from background or other objects in the camera view. To extract text regions from complex backgrounds, we have proposed a novel text localization algorithm based on models of stroke orientation and edge distributions. The corresponding feature maps estimate the global structural feature of text at every pixel. Block patterns project the proposed feature maps of an image patch into a feature vector. Adjacent character grouping is performed to calculate candidates of text patches prepared for text classification. An Adaboost learning model is employed to localize text in camera-based images. Off-the-shelf OCR is used to perform word recognition on the localized text regions and

Volume : 6 | Issue : 11 | November 2016 | ISSN - 2249-555X | IF : 3.919 | IC Value : 74.50

transform into audio output for blind users. Our future work will extend our localization algorithm to process text strings with characters fewer than three and to design more robust block patterns for text feature extraction. We will also extend our algorithm to handle non horizontal text strings. Furthermore, we will address the significant human interface issues associated with reading text by blind users.

References

1. World Health Organization. (2009). 10 facts about blindness and visual impairment [Online]. Available: www.who.int/features/factfiles/blindness/blindness_facts/en/index.html
2. Advance Data Reports from the National Health Interview Survey (2008). [Online]. Available: http://www.cdc.gov/nchs/nhis/nhis_ad.htm
3. International Workshop on Camera-Based DocumentAnalysis and Recognition (CBDAR 2005, 2007, 2009, 2011). [Online]. Available: http://www.m.cs.osakafu-u.ac.jp/cbdar2011
4. X.Chen and A. L. Yuille, "Detecting and reading text in natural scenes," inProc.Comput.VisionPatternRecognit.,2004,vol.2,pp.II-366-II-373.
5. X. Chen, J. Yang, J. Zhang, and A. Waibel, "Automatic detection and recognition of signs from natural scenes," IEEE Trans. Image Process., vol. 13, no. 1, pp. 87-99, Jan. 2004.
6. D. Dakopoulos and N. G. Bourbakis, "Wearable obstacle avoidance electronic travel aids for blind: A survey," IEEE Trans. Syst., Man, Cybern., vol. 40, no. 1, pp. 25-35, Jan. 2010.
7. Yi and Y. Tian, "Assistive text reading from complex background for blind persons," in Proc. Int. Workshop Camera-Based Document Anal. Recognit., 2011, vol. LNCS-7139, pp. 15-28.
8. S. Kumar, R. Gupta, N. Khanna, S. Chaudhury, and S. D. Joshi, "Text extraction and document image segmentation using matched wavelets and MRF model," IEEE Trans Image Process., vol. 16, no. 8, pp. 2117-2128, Aug. 2007.
9. 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.
10. P. Shivakumara, T. Phan, and C. L. Tan, "A gradient difference based technique for video text detection," in Proc. Int. Conf. Document Anal. Recognit., 2009, pp. 66-70.
11. S. Shoval, J. Borenstein, and Y. Koren, "Auditory guidance with the Navbelt: A computerized travel for the blind,"IEEETrans.Syst.,Man,Cybern. C. Appl. Rev., vol. 28, no. 3, pp. 459-467, Aug. 1998.
12. C.Stauffer andW. E. L. Grimson,"Adaptive backgroundmixture models forreal-timetracking,"presentedatthelEEEComput.Soc.Conf.Comput. Vision Pattern Recognit., Fort Collins, CO, USA, 1999.