CAMERA BASED LABEL READER FOR BLIND PEOPLE

A. SUBBIAH[*], T. ARIVUKKARASU, M. S. SARAVANAN and V. BALAJI

Department of ECE, Aarupadai Veedu Institute of Technology,
CHENNAI (T.N.) INDIA

ABSTRACT

An OCR (Optical Character Recognition) system, which is a branch of computer vision and in turn a sub-class of Artificial Intelligence. OCR is the translation of optically scanned bitmaps of printed or hand written text into audio output by using of Raspberry Pi. OCRs developed for many world languages are already under efficient use. The Moving object region is extracted by a mixture-of-Gaussians-based background subtraction method. A text localization and recognition are conducted to acquire text information. To localize the text regions from the object, a text localization algorithm by learning gradient features of stroke orientations and distributions of edge pixels .Text characters in the localized text regions are then binarized and recognized by off-the-shelf OCR software. The text codes recognized are output to blind users in speech. As the recognition process is completed, the character codes in the text file are processed using Raspberry pi device which recognize character using Adaboost algorithm and python programming, the audio output is listened.

Key words: Label reader, Blind, Camera, OCR.

INTRODUCTION

Optical character recognition (OCR) technology offers blind and visually impaired persons the capacity to scan printed text and then speak it back in synthetic speech or save it to a computer. Little technology exists to interpret graphics such as line art, photographs, and graphs into a medium easily accessible to blind and visually impaired persons. It also is not yet possible to convert handwriting, whether script or block printing, into an accessible medium. There are three essential elements to OCR technology-scanning, recognition, and reading text. Initially, a printed document is scanned by a camera. OCR software then converts the images into recognized characters and words. The synthesizer in the OCR system then speaks the recognized text. Finally, the information is stored in an electronic

_____

[*]Author for correspondence; E-mail: subbiah@avit.ac.in

form, either in a personal computer (PC) or the memory of the OCR system itself. The recognition process takes account of the logical structure of the language. An OCR system will deduce that the word "tke" at the beginning of a sentence is a mistake and should be read as the word "the." OCR's also use a lexicon and apply spell checking techniques similar to those found in many word processors.

All OCR systems create temporary files containing the texts' characters and page layout. In some OCR's these temporary files can be converted into formats retrievable by commonly used computer software such as word processors and spreadsheet and database software. The blind or visually impaired user can access the scanned text by using adaptive technology devices that magnify the computer screen or provide speech or braille output.

The text, which is given as input includes stereotypical forms – such as street signs, hospital signs, and bus numbers –as well as more variable forms such as shop signs, house numbers, and billboards. Our database of city images were taken in ZZZZ partly by normally sighted viewers and partly by blind volunteers who were accompanied by sighted guides (for safety reasons) using automatic camera settings and little practical knowledge of where the text was located in the image. The databases have been labeled to enable us to train part of our algorithm and to evaluate the algorithm performance.

The first, and most important, component of the algorithm is a strong classifier which is trained by the AdaBoost learning algorithm on labeled data. AdaBoost requires specifying a set of features from which to build the strong classifier. This paper selects this feature set guided by the principle of informative features. We calculate joint probability distributions of these feature responses on and off text, so weak classifiers can be obtained as log-likelihood ratio tests. The strong classifier is applied to sub-regions of the image (at multiple scale) and outputs text candidate regions.
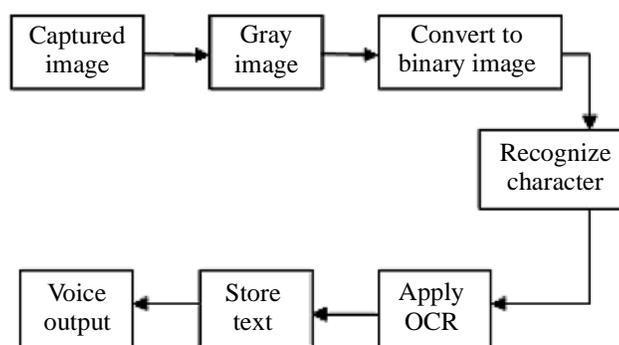
Fig. 1: Block diagram of conversion from image to voice

Proposed system

In this we have proposed a system using Raspberry Pi Model for scanning the images from stereotypical forms – such as street signs, hospital signs, and bus numbers –as well as more variable forms such as shop signs, house numbers, and billboards. Here we are using AdaBoost Algorithm for processing the visual information and converting into audio speech.

ADA Boost algorithm

AdaBoost is discovered in computer science is called the Multiplicative Weights Update Algorithm (MWUA), and it has many applications like from learning theory to combinatorial optimization and game theory. The aim is to Maintain a nonnegative weight for the elements of some set, Draw a random element proportionally to the weights, So something with the chosen element, and based on the outcome of the "something…" Update the weights and repeat.

The "something" is usually a black box algorithm like "solve this simple optimization problem." The output of the "something" is interpreted, and the weights are updated according to the severity. Here one can interpret MWUA as minimizing regret with respect to the best alternative element one could have chosen in hindsight. In fact, this was precisely the technique we used to attack the adversarial bandit learning problem .
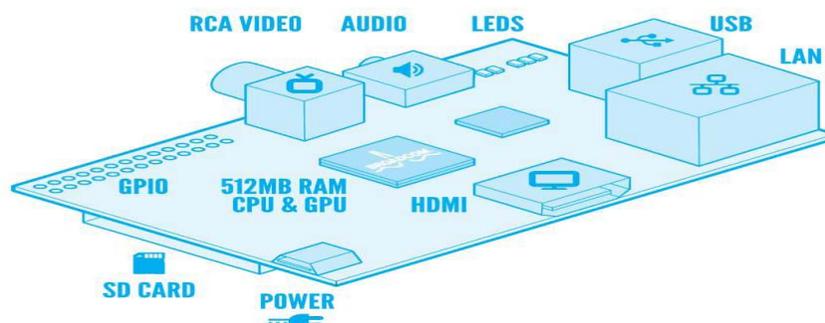
AdaBoost requires a set of classified data with image windows labeled manually as being text or non-text. We performed this labeling for the training dataset and divided each text window into several overlapping text segments with fixed width-to-height ratio 2:1. This lead to a total of 7,132 text segments, which were used as positive examples. The negative examples were obtained by a bootstrap process similar to Drucker et al.[2] First we selected negative examples by randomly sampling from windows in the image dataset. After training with these samples, we applied the AdaBoost algorithm to classify all windows in the training images (at a range of sizes). Those misclassified as text was then used as negative examples for retraining AdaBoost. The image regions most easily confused with text were vegetation, repetitive structures such as railings or building facades, and some chance patterns.

RASPBERRY-PI

The Raspberry-Pi is a credit-card sized single-board computer developed in the UK with the Intention of promoting the teaching of basic computer science in schools. Being so

versatile, it soon became one important tool which can be used in different electronic projects. The Raspberry-Pi has a Broadcom BCM2835 SoC, which includes an ARM1176JZF-S 700 MHz processor and a Videocore 4 GPU. The more recent model (B) has 512 MB of RAM, 2 USB ports and an Ethernet port. It does not have a built-in hard disk, so it relies on an SD card for booting and persistent storage. It also includes a camera board that connects to the CSI 2 camera port on the Raspberry Pi using a short ribbon cable. It provides connectivity for a camera capable of capturing still images or video recordings. The camera connects to the Image System Pipeline (ISP) in the Raspberry Pi's SoC, where the incoming camera data is processed and eventually converted to an image or video on the SD card. This camera module is capable of taking photos up to 5 MP (2592 × 1944 pixels) and can record video at resolutions up to 1080p30 (1920 x 1080 x 30fps).

RESULTS AND DISCUSSION

In this paper, we have used the Adaboost algorithm in Raspberry PI Model for the conversion for Text to audio speech so that the blind can easily understand and read the data which they could not see. The same project can be extended and tested with other OCR Algorithms'.

REFERENCES

1. X. Chen and A. L. Yuille, Ada Boost Learning for Detecting and Reading Text in City Scenes Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition (2004) pp. 366-373.

2. Lukas Neumann Jiri Matas, Real-Time Scene Text Localization and Recognition Published at the 25th IEEE Conference on Computer Vision and Pattern Recognition, CVPR, Providence, RI, USA (2012).

3.   Chucai Yi, Yingli Tian and Aries Arditi, Portable Camera-Based Assistive Text and Product Label Reading From Hand-Held Objects for Blind Persons, IEEE/Asme Transactions on Mechatronics (2013).

4.   N. Rajkumar, M. G. Anand and N. Barathiraja, Portable Camera-Based Product Label Reading For Blind People, Int. J. Engg. Trends Technol., **10(11)**, 521 (2014).