Australian Journal of Basic and Applied Sciences, 10(12) July 2016, Pages: 182-187

AUSTRALIAN JOURNAL OF BASIC AND APPLIED SCIENCES

ISSN:1991-8178 EISSN: 2309-8414
Journal home page: www.ajbasweb.com

Performance Analysis Of Optical Character Recognition On Text Images

T. Maheshwary, ME II year, Dr. M. Arulselvi

Assistant Professor, Dept. of Computer science & Engineering Annamalai University, India.

Address For Correspondence:
T. Maheshwary, Assistant Professor, Dept. of Computer science & Engineering Annamalai University, India.

ARTICLE INFO

Article history:
Received 26 April 2016
Accepted 21 July 2016
Published 30 July 2016

Keywords:
OCR (Optical Character Recognition), Computer Vision tool box, performance measures.

ABSTRACT

Detection of text from natural images receives a growing attention because of potential applications in image retrieval, robotics and intelligent transport system. Camera-based document analysis becomes a real possibility with the increasing resolution and availability of digital cameras. The objective is to develop a system that detect and recognize the text encountered in natural scenes. Challenging part of natural scene text detection systems is that, sometimes scene images containing varying in text due to the style, size, alignment, orientation, complex background etc. Many research methods presented over the detection of text from the natural images previously and still many researches are going in this same area. The main goal of this work is to implement Optical Character Recognition algorithm to recognize text from natural scenes. A comparison has been made to measure the performance of OCR on different types of natural images.

INTRODUCTION

The current work over the research field called retrieval of contents from the images recognized different range of application areas where the need of automated text extraction is required from the natural scene images. Recently the new application is developed from the mobile banking which is provided by particular bank to their customers with aim of facilitating their banking users to execute their transactions by sending just image of their cheque of passbook to the main server of bank. Another application of it is tourist guide, which helps the tourist to understand the different language written display boards as image text translation systems to help the visually impaired people and also tourists. The above applications are based on concept of text extraction from the scene images. This algorithm efficiently detects, localize as well as extract the text related information available in natural scene images. One of the text detection is based on Optical Character Recognition. First of all, in order to understand challenges of this in text detection (Xu-chengYin, X.-C. Yin, 2000).

In this work, OCR algorithm is used to detect the text from natural scene images. The first OCR equipment installed at the Reader’s Digest in 1954 was employed to convert typewritten reports into computer readable punched cards. In the early 1960’s, commercial OCR algorithm were developed for recognizing a small amount of letters and symbols specially designed for machine reading. Several years later, IBM exhibited its IBM 1287 system which was able to recognize regular machine printed characters. Toshiba also developed a postal code numbers recognizer that had hand-printed character processing capability. In the middle of the 1970’s, some prototype OCR algorithms were able to process large printed and hand-written character sets in poor quality documents. OCR algorithm is available as software packages for home usage after 1986 since the prices of related hardware were getting cheaper.OCR algorithm has many advantages over this problem. OCR engine is loaded with about 675 languages and over 4000 font type.

To Cite This Article: T. Maheshwary, ME II year, Dr. M. Arulselvi., Performance Analysis Of Optical Character Recognition On Text Images. ***Aust. J. Basic & Appl. Sci.,*** *10(12): 182-187, 2016*

Overview Of Ocr:

Text Extraction is done by OCR. It is usually performed in the preprocessing stage of different image processing related application such as optical character recognition (OCR). There are two basic types of core OCR algorithm, which may produce a ranked list of candidate characters. Matrix matching involves comparing an image to a stored glyph on a pixel-by-pixel basis; it is also known as "pattern matching", "pattern recognition", or "image correlation”. This relies on the input glyph being correctly isolated from the rest of the image, and on the stored glyph being in a similar font and at the same scale. This technique works best with typewritten text and does not work well when new fonts are encountered. This is the technique the early physical photocell-based OCR implemented, rather directly (Malyan, R.R., *et al*., 1989).

Feature extraction decomposes glyphs into "features" like lines, closed loops, line direction, and line intersections. These are compared with an abstract vector-like representation of a character, which might reduce to one or more glyph prototypes (Jianli Liu, *et al*., 1993). General techniques of feature detection in computer vision are applicable to this type of OCR, which is commonly seen in "intelligent" handwriting recognition and indeed most modern OCR software. Nearest neighbor classifiers such as the k-nearest neighbors algorithm are used to compare image features with stored glyph features and choose the nearest match.

Optical Character Recognition technology recognizes the text from the images automatically. It supports different types of image formats like JPG, PNG, BMP, GIF, TIFF and multi-page PDF files. OCR involves analysis of the captured or scanned images and then translate character images into character codes, so that it can be edited, searched, stored more efficiently, displayed on-line, and used in machine processes. Scanned images can easilyextract that text with the help of different OCR Tools (Yi, C. and Y. Tian, 2013). It works with images that almost consist of text in it . The output of a tool is based on the type of input image. Achieving 100% accuracy is not possible, but it is better to have something rather than nothing. To improve accuracy most of the OCR tools use dictionaries, recognizing individual characters then it try to recognize entire words that exist in the selected dictionary.

Implementation:

This work is implemented by using the following modules and shown in Fig 1.

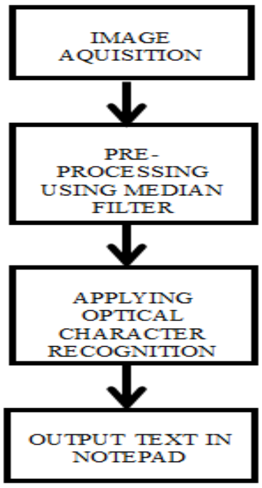


Fig. 1: Shows the flowchart for OCR algorithm

A. Image Acquisition:

Image acquisition in image processing can be broadly defined as the action of retrieving an image from some source, usually a hardware-based source, so it can be passed through whatever processes need to occur afterward. Performing image acquisition in image processing is always the first step in the workflow sequence because, without an image, no processing is possible. The image that is acquired is completely unprocessed and is the result of whatever hardware was used to generate it, which can be very important in some fields to have a consistent baseline from which to work. One of the ultimate goals of this process is to have a source of input that operates within such controlled and measured guidelines that the same image can, if necessary, be nearly perfectly reproduced under the same conditions so anomalous factors are easier to locate and eliminate (Mori, S., *et al*., 1992). This OCR doesn’t have any limitations on Sizes, resolutions & pixel size.

B. Pre-Processing Using Median Filter:

These are the pre-processing steps often performed in OCR. Pre-processing images commonly involves removing low-frequency background noise, normalizing the intensity of the individual particles images, removing reflections, and masking portions of images. Image pre-processing is the technique of enhancing data images prior to computational processing. Image pre-processing can significantly increase the reliability of an opticalinspection. Several filteroperations which intensify or reduce certain image details enable an easier or faster evaluation. Users are able to optimize a camera image with just a few clicks.

C. Pre-Processing:

(i) RGB to Gray Scale Conversion:

Humans perceive color through wavelength-sensitive sensory cells called cones. There are three different types of cones, each with a different sensitivity to electromagnetic radiation (light) of different wavelength. One type of cone is mainly sensitive to red light, one to green light, and one to blue light. By emitting a controlled combination of these three basic colors (red, green and blue), and hence stimulate the three types of cones at will, we are able to generate almost any perceivable color. This is the reasoning behind why color images are often stored as three separate image matrices; one storing the amount of red (R) in each pixel, one the amount of green (G) and one the amount of blue (B). We call such color images as stored in an RGB format. In gray scale images, however, we do not differentiate how much we emit of the different colors, we emit the same amount in each channel. What we can differentiate is the total amount of emitted light for each pixel; little light gives dark pixels and much light is perceived as bright pixels. When converting an RGB image to gray scale, we have to take the RGB values for each pixel and make as output a single value reflecting the brightness of that pixel. One such approach is to take the average of the contribution from each channel: (R+B+C)/3. However, since the perceived brightness is often dominated by the green component, a different, more "human-oriented", method is to take a weighted average, e.g.: 0.3R + 0.59G + 0.11B.

(ii)Applying Median Filter:

The median filter is a nonlinear digital filtering technique, often used to remove noise. Such noise reduction is a typical pre-processing step to improve the results of later processing (for example, edge detection on an image). Median filtering is very widely used in digital image processing because, under certain conditions, it preserves edges while removing noise (Malyan, R.R., *et al*., 1989). Median filtering is one kind of smoothing technique, as is linear Gaussian filtering. All smoothing techniques are effective at removing noise in smooth patches or smooth regions of a signal, but adversely affect edges. Often though, at the same time as reducing the noise in a signal, it is important to preserve the edges. However, its performance is not that much better than Gaussian blur for high levels of noise, whereas, for speckle noise and salt and pepper noise (impulsive noise), it is particularly effective. Because of this, median filtering is very widely used in digital image processing.

Gray Scale to Binary Image Conversion:

Binary images are also called Bi-level images, which contains pixel value 0 and 1.So 1 bit is sufficient to represent the pixel values. In Image processing Binary images are encountered in many ways. The Binary image is created from gray scale image (8 bits 28=256) using a threshold process. The pixel value is compared with threshold value. If the pixel value of the gray scale image is greater than the threshold value, the pixel value in the binary image is considered as one. Otherwise the pixel value is zero. In addition, image processing operations produce binary images at intermediate stages.

In photography and computing, a gray scale digital image is an image in which the value of each pixel is a single sample, that is, it carries only intensity information. Images of this sort, also known as black-and-white, are composed exclusively of shades of gray, varying from black at the weakest intensity to white at the strongest. Gray scale images are distinct from one-bit bi-tonal black-and-white images, which in the context of computer imaging are images with only the two colors, black, and white (also called bi-level or binary images). Gray scale images have many shades of gray in between. Gray scale images are often the result of measuring the intensity of light at each pixel in a single band of the electromagnetic spectrum (e.g. infrared, visible light, ultraviolet, etc.), and in such cases they are monochromatic proper when only a given frequency is captured. But also they can be synthesized from a full color image; see the section about converting to gray scale.

Pre-processing images commonly involves removing low-frequency background noise, normalizing the intensity of the individual particles images, removing reflections, and masking portions of images. Image pre-processing is the technique of enhancing data images prior to computational processing. Image pre-processing can significantly increase the reliability of an optical inspection. Several filter operations which intensify or reduce certain image details enable an easier or faster evaluation. Users are able to optimize a camera image with just a few clicks.

D. Optical Character Recognition Process:

OCR Recognize the text using Optical Character Recognition. txt = ocr(I) returns an ocr text object containing the recognized text, the location of the text within I, and a metric indicating the confidence of the

Australian Journal of Basic and Applied Sciences, 10(12) July 2016, Pages: 182-187

recognition result. Confidence values range between 0 and 1 and should be interpreted as probabilities.

The general steps in the OCR process are as follows

- Lines detection and removal: This step is required to improve page layout analysis, to achieve better recognition quality for underlined text, to detect tables, etc.
- Page layout analysis (also called "zoning"). The OCR algorithm must detect the positions and types of all important areas in the image.
- Detection of text lines and words: Sometimes it is not an easy task because of different font sizes and small spaces between words.

E. Output Text In Notepad:

After the extraction the output text is shown on notepad. Any word related software can be used to display the extracted texts. The performance measure can becalculatedusing these formulas.

F. Performance Measures:

To verify whether an OCR tool has converted all the characters available in the input image correctly or not. Character accuracy (CA) and Character error rate (CER) are

calculated. To calculate CA and CER following formulae are used (Lynch, M.R., P.J. Rayner, 1989).

Character accuracy (CA) = (a/n) *100

(Where a is Total number of characters in the resultant text document n is Total number of characters in the input image)

Character error rate = 100-CA.

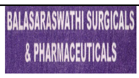
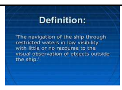
Experimental Analysis:

The comparison of performance parameters on different images are shown below(*Table-1*).

Table 1: shows performance of OCR on different images

S.no	Images	Time Taken(in sec)	Accuracy %	Error %
1		12	98	2.0
2		13	98	2.0
3		11	95	5.0
4		15	99	1.0
5		10	98	2.0
6		11	99	1
7		12	99	1
8		14	99	1

Australian Journal of Basic and Applied Sciences, 10(12) July 2016, Pages: 182-187

9		9	98	2
10		14	96	4
11		11	98	2.0
12		9	100	0
13		16	95	5.0
14		10	98	2.0
15		12	98	2.0

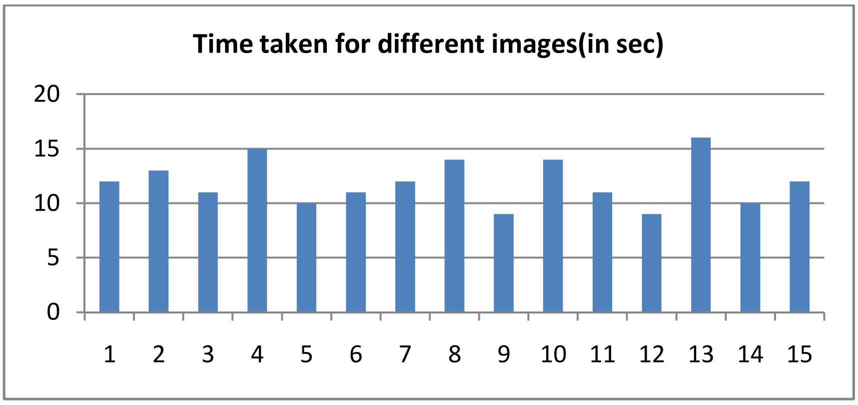


Fig. 2: shows the graph for time takenby the algorithm for different images

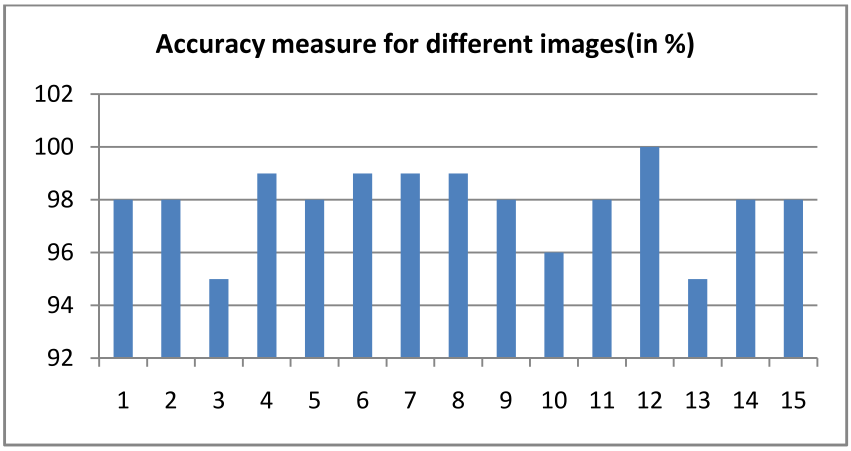


Fig. 3: shows the graph for accuracy level for different images.

Australian Journal of Basic and Applied Sciences, 10(12) July 2016, Pages: 182-187

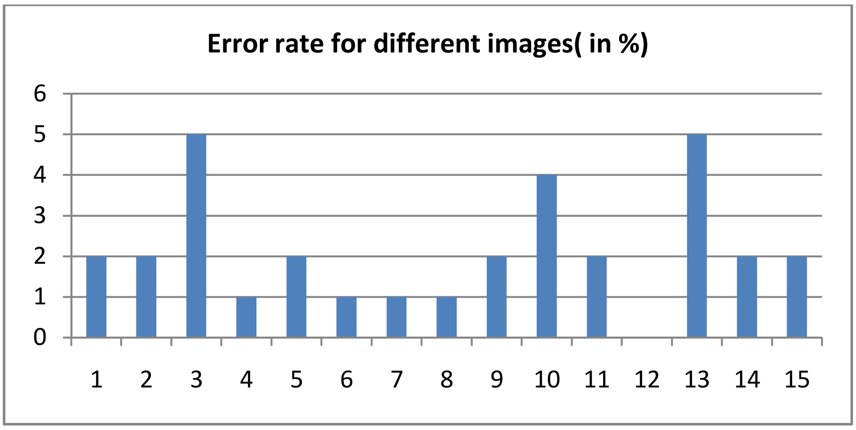


Fig. 4: shows the graph for error rate for different images

Conclusion:

In this work different types of text containing images are compared the accuracy, time and error rate are evaluated using OCR technique. OCR accuracy is good for the images which are taken with good resolution. But some texts are not recognized due to lack of clarity and resolution. And also tilted images are not recognized properly. In future care must be taken while pre-processing the image with different techniques.

REFERENCES

Chen, X., J. Yang, J. Zhang and A. Waibel, 2004. “Automatic detection andrecognition of signs from natural scenes,” *In IEEE Transactions on image processing*, 13(1): 87-99.

Elliman, D.G., 1994. "Peeling potatoes with a cheese grater [handwritten document OCR," *Handwriting Analysis and Recognition: A European Perspective, IEE European Workshop on*, pp: 14/1-14/4, 12-13.

Epshtein, B., E. Ofek and Y. Wexler, 2010. “Detecting text in natural scenes with stroke width transform,” in *Proc. IEEE Conf. CVPR,San Francisco, CA, USA*, pp: 2963-2970.

Jianli Liu, Nugent, J.H., D.G. Bowen, J.E. Bowen, 1993. "Intelligent OCR editor,"*Electrical and Computer Engineering, 1993. Canadian Conference*on, pp: 9-11.

Kim, K., K. Jung and J. Kim, 2003. “Texture-base approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm,*” IEEE Trans. Pattern Anal.Mach. Intell*., 25(12): 1631-1639.

Liu, Q., C. Jung and Y. Moon, 2006. “Text Segmentation based on StrokeFilter,” *Proceedings of International Conference on Multimedia*, pp: 129-132.

Lynch, M.R., P.J. Rayner, 1989. "Optical character recognition using a new connectionistmodel," *Image Processing and its Applications, Third InternationalConference* on,pp: 63-67.

Malyan, R.R., S. Sunthankar, H. Teranchi, A. Yeghiazarian, 1989. "Perception of multiauthorhandprinted text," *Character Recognition and Applications, IEE Colloquium*, pp: 8/1-8/3.

Mori, S., C.Y. Suen, K. Yamamoto, 1992. "Historical review of OCR research anddevelopment," *Proceedings of the IEEE* , 80(7): 1029-1058.

Sunil Kumar, Rajat Gupta, NitinKhanna, 2007. “Text extraction and document image segmentationUsing matched wavelets and mrf model”, *ieee transactions on image processing*, 16(8).

Vijayarani, Dr. S. and Ms. A.Sakila, 2015"Performance Comparison Of OCR Tools," *International Journal of UbiComp (IJU)*, 6: 3.

Xu-chengYin, X.-C. Yin, 2000. “Robust Text Detection In Natural Scene Images,” *IEEE Trans. Image Process*., 9(1): 147-156.

Yi, C. and Y. Tian, 2011.“Text string detection from natural scenes by structure-based partition and grouping,” *IEEE Trans. Image Process.*, 20(9): 2594-2605.

Yi, C. and Y. Tian, 2013. “Text extraction from scene images by character appearance and structure modeling,” *Comput. Vis. Image Und*., 117(2): 182-194.

Yin, X., X.-C. Yin, H.-W. Hao and K. Iqbal, 2012. “Effective text localization in natural scene images with MSER, geometry-based grouping and ada boost,” *in Proc. Int. Conf. Pattern Recognit.*,Tsukuba, Japan, pp: 725-728.