TEXT EXTRACTION USING MATLAB

Extraction of text using Edge Detection & Connected Components

[1]Sukanta Kumar Tulo, [2]N. Ravi Kumar, [3]Sumit Kumar Singh, [4]Jagmeet Singh
[1]Assistant Professor, [2]Student, [3]Student, [4]Student
[1,2,3,4] Electronics & Communication Engineering,
[1,2,3,4] G.I.E.T Gunupur, Odisha, India

Abstract : **This project extracts the text from image and recognizes the text written. Text Extraction and recognition in general have quite a lot of relevant application for automatic indexing or information retrieval such document indexing, content-based image retrieval and recognition which further opens up the possibility for more improved and advanced systems. This example shows how to detect regions containing text in an image. It is a common task performed on unstructured scenes, for example when capturing video from a moving vehicle for the purpose of alerting a driver about a road sign. Segmenting out the text from a cluttered scene greatly helps with additional tasks such as optical character recognition (OCR).**
The automated text detection algorithm in this example starts with the detection of MSER regions and a large number of text region candidates and progressively removes those less likely to contain text. To highlight this algorithm's flexibility, it is applied to images containing a road sign, a poster.
The resulting multiple segmentation hypotheses are post-processed by a connected component analysis and a gray scale consistency constraint algorithm. Finally, they are processed by OCR software. A selection algorithm based on language modeling and OCR statistics chooses the text result from all the produced text strings.

Keywords - Text Extraction, Canny Edge Detection, OCR, Segmenting, Mser Regions

I. INTRODUCTION

Text detection and recognition in images is a research area which attempts to develop a computer system with the ability to automatically read from images the text content visually embedded in complex backgrounds. As an example of the general object recognition issues, this computer system, should answer two typical questions "Where & What": "where is a text string?" and "what does the text string say?" in an image. In other words, using such a system, text embedded in complex backgrounds can be automatically detected and each character or word can be recognized.

As computer, compress technology, storage media and high speed communication skill are developed dramatically; digital video has become one of the most important elements in many applications such as education, news and games. Multimedia data are also getting bigger than before. In order to extract and search important information from a huge amount of data, we need to extract text from images. Text is obviously an important element in images. So extracting text appears as a key clue for understanding contents of image and for instance for classifying automatically some images. Text detection and recognition has been identified as one of the key components for the retrieval and analysis system. Text detection and recognition can be used in many applications, summarization, video surveillance and security, multilingual video information access, etc.

1. MAIN CONCEPT:

Text extraction in image consists in three steps. The first one is to find text region in original images. Then the text needs to be separated from background. And finally a binary image has to be produced (for example, text is white and background is black). The term text detection here means the distinguishing the letters or the characters from the image part. This is the process of determining whether a given part or part of the image is a text or some other figures. Text detection generally can be classified into two categories:

1.1. BOTTOM-UP METHODS:

They segment images into regions and group "character" regions into words. . The input image is segmented based on the monochromatic nature of the text components using a split-and-merge algorithm. Segments that are too small and too large are filtered out. After dilation, motion information and contrast analysis are used to enhance the extracted results. The methods, to some degree, can avoid performing text detection. Due to the difficulty of developing efficient segmentation algorithms for text in complex background, the methods are not robust for detecting text in many camera-based images and videos.

1.2. TOP-DOWN METHODS:

They first detect text regions in images using filters and then perform bottom-up techniques inside the text regions. These methods are able to process more complex images than bottom-up approaches. Top-down methods are also divided into two categories

1.2.1 -Heuristic methods: they use heuristic filters.
Heuristic top-down algorithms first detect text blocks in images using heuristic filters, then segment them into text and background regions. In other words, the first part of the algorithms aims at detecting text regions in images and the second part can be regarded as applying a bottom-up method on a local image. This enables the algorithms to process complex images but difficulties are encountered in both the detection and segmentation stages.

1.2.2- Machine learning methods: they use trained filters.
Most of the heuristic top-down methods employ manually designed heuristic features and usually perform fast detection but are not very robust when the background texture is very complex. As an alternative, a few systems considered machine learning tools to perform the text detection. These systems extracted wavelet or derivative features from fixed-size blocks of pixels and classified the feature vectors into text or non-text using artificial neural networks. However, since the neural network based classification was applied to all the possible positions of the whole image, the detection system was not efficient in terms of computation cost and produced unsatisfactory false alarm and rejection rates.
Here we are using heuristic method of text extraction. This method of text extraction can be performed in two different approaches. Each of both uses the characteristics of artificial text.

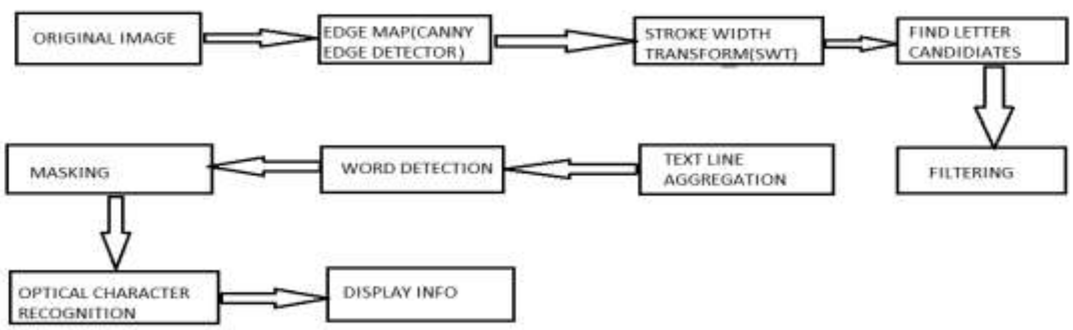
2. Algorithm & Flow Chart:

We are using edge detection approach as it is simple to implement and is also works almost efficiently. The algorithm as well as the flow chart for our software is shown below:

Algorithm:

Step 1: Load the image.
Step 2: Detect MSER Regions present in the image.
Step 3: Use Canny Edge Detector to Further Segment the Text.
Step 4: Filter Character Candidates Using Connected Component Analysis.
Step 5: Filter Character Candidates Using the Stroke Width Image.
Step 6: Determine Bounding Boxes With Enclosed Text Regions.
Step 7: Perform Optical Character Recognition on Text Region for the extraction of text.

Flow Chart:



FLOW CHART

3. OCR (Optical character recognition):

Optical character recognition (OCR) addresses the problem of reading optically processed characters and has become one of the most successful applications of technology in the field of pattern recognition and artificial intelligence
The performances of OCR systems closely rely on the quality of the targeting documents. Figure illustrates the performance of some typical commercial OCR software in function of page quality and resolution. These statistical data come from the ISRI 1995 annual test of OCR accuracy . The different curves in the figure represent the performance (in terms of word/character recognition rate) of different software. . Here the page quality only measures white papers, in which there are almost no background or just dots noise. We can see that the performance of OCR software drop abruptly because of low quality of page or poor resolution. This drawback of current OCR software raises an important question in the context of content-based indexing and retrieval. The question is: what kind of performance can lead to a good retrieval

4. MSER REGIONS:

In computer vision, maximally stable extremal regions (MSER) are used as a method of blob detection in images. This technique was proposed by Matas et al. to find correspondences between image elements from two images with different viewpoints. This method of extracting a comprehensive number of corresponding image elements contributes to the wide-baseline matching, and it has led to better stereo matching and object recognition algorithms.

4.1. EXTREMAL REGIONS:

Extremal regions in this context have two important properties, that the set is closed under...

1. Continuous transformation of image coordinates. This means it is affine invariant and it doesn't matter if the image is warped or skewed.
2. Monotonic transformation of image intensities. The approach is of course sensitive to natural lighting effects as change of day light or moving shadows.

4.2. USE IN TEXT DETECTION :

The MSER algorithm has been used in text detection by Chen by combining MSER with Canny edges. Canny edges are used to help cope with the weakness of MSER to blur. MSER is first applied to the image in question to determine the character regions. To enhance the MSER regions any pixels outside the boundaries formed by Canny edges are removed. The separation of the later provided by the edges greatly increases the usability of MSER in the extraction of blurred text. An alternative use of MSER in text detection is the work by Shi using a graph model. This method again applies MSER to the image to generate preliminary regions. These are then used to construct a graph model based on the position distance and color distance between each MSER, which is treated as a node. Next the nodes are separated into foreground and background using cost functions. One cost function is to relate the distance from the node to the foreground and background. The other penalizes nodes for being significantly different from its neighbor. When these are minimized the graph is then cut to separate the text nodes from the non-text nodes.To enable text detection in a general scene, Neumann uses the MSER algorithm in a variety of projections. In addition to the greyscale intensity projection, he uses the red, blue, and green color channels to detect text regions that are color distinct but not necessarily distinct in greyscale intensity. This method allows for detection of more text than solely using the MSER+ and MSER- functions discussed above.

4.3. GRAY LEVEL TRANSFORMATIONS

An image processing system that looks at every input pixel gray level and generates a corresponding output gray level according to a fixed gray level map is referred to as a basic gray level transformation. An image can be represented by a two-dimensional function, f(x, y), where x and y are special co-ordinates, and the amplitude of f at any point is the intensity or gray level of that image at that point. The RGB image when grayscaled, the resulting image is a form of black and white image. This method is used for contrast stretching. Because enhancement of the image at any point depends only at the gray level at that point. This method is the simplest of al the image enhancement techniques.

II. PICTORIAL ANALYSIS WITH FIGURES

Fig:1

Fig:2

Fig:3

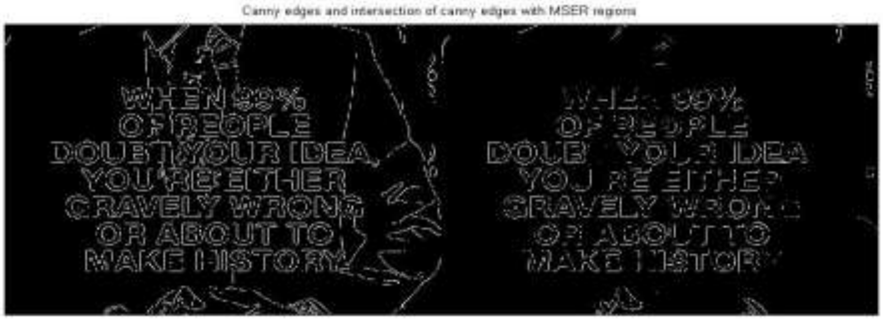


Fig:4

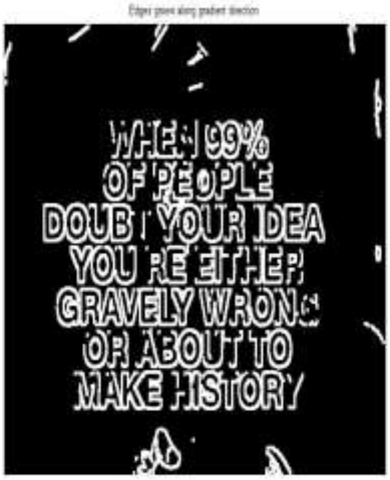


Fig:5

Fig:6

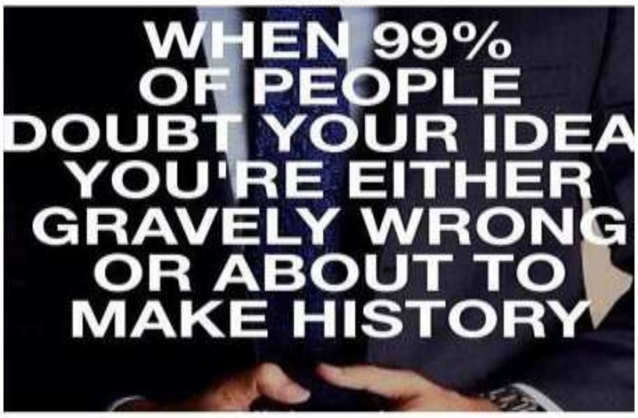
Fig7

Fig:8

Fig:9

Fig:10

Text region



RESULT: WHEN 99%
OF PEOPLE
DOUBT YOUR IDEA
YOU'RE EITHER
GRAVELY WRONG
OR
ABOUT TO
MAKE HISTORY

III. CONCLUSION

In many ways the result of these experiments are both surprisingly good and surprisingly bad. For images without definite edges the program may not work properly. But it will work perfectly for image texts which have prominent edge. This can be extended by including Text Recognition to it. Also extending it to video, real time operation, the program can work surprisingly well and useful

- segment and recognize characters in image have been achieved and extracted.
- Segmentation accuracy from the experiment is 100%.
- Segmentation using connected components is best method to segmenting the image.
- The result during real application may be lower due to limited set of picture used in experiment.

Advantages:

1) Tilt text is detected.
2) High accuracy in natural scene.
3) Requires less text extraction database
4) Most relevant and accurate data is retrieved from the web.

Disadvantages:

1) Handwritten text cannot be accurately recognized.

Application:

1) Analysis of documents can be easily done.
2) Industrial automation.

REFERENCES

[1] Sunil Kumar Rajat Gupta Nitin Khanna, Santanu Chaudhury and Shiv Dutt Joshi. Text extraction and document image segmentation using matched wavelets and mrf model. IEEE TRANSACTIONS ON IMAGE PROCESSING, 16:1–8, 2007.
[2] IEEE Li Linlin Chew Lim Tan Senior Member IEEE Shijian Lu, Member. Documentimage retrieval through word shape coding. pages 11–15, 2008.
[3] Anand Mishra Karteek Alahari C. V. Jawahar. Top-down and bottom-up cues for scene text recognition. pages 1–3.
[4] Yonatan Wexler Boris Epshtein, Eyal Ofek. Detecting text in natural scenes with stroke width transform. pages 2–5.
[5] Qiuhau Lin Tong Zhang, Derek Ma. Mobile camera based text detection and translation. IEEE, 7(2):87–105, 2010.