IORD Journal of Science & Technology
E-ISSN: 2348-0831
Volume 3, Issue 3 (MAR-APRIL 2017) PP71-75
IMPACT FACTOR 1.719
www.iord.in

"Study Of Stroke Technique for Recognition Of Nature Scene Text"

Author1.(Mr.Vineetkumar S. Tiwari)
Vidarbha Institute of Technology,
Nagpur,India

Authors 2nd. (Pravin Kulurkar)
Vidarbha Institute of Technology,
Nagpur,India

***Abstract*— Recognizing scene text is a challenging problem, even more so than the recognition of scanned documents. Given the rapid growth of camera-based applications readily available on mobile phones, understanding scene text is more important than ever. One could, for instance, foresee an application to answer questions such as, "What does this sign say?". This is related to the problem of Optical Character Recognition (OCR), which has a long history in the computer vision community. However, the success of OCR systems is largely restricted to text from scanned documents. Scene text exhibits a large variability in appearances, and can prove to be challenging even for the state-of-the-art OCR methods. Many scene understanding methods recognize objects and regions like roads, trees, sky in the image successfully, but tend to ignore the text on the sign board. Our goal is to fill this gap in understanding the scene. Scene text recognition has inspired great interests from the computer vision community in recent years. Here propose a novel scene text recognition method integrating structure guided character detection and linguistic knowledge and use part-based tree structure to model each category of characters so as to detect and recognize characters simultaneously. Since the character models make use of both the local appearance and global structure information's, the detection results are more reliable. For word recognition, we combine the detection scores and language model into the posterior probability of character sequence from the Bayesian decision view. The final word recognition result is obtained by maximizing the character sequence posterior probability via Viterbi algorithm. Experimental results on a range of challenging public data sets demonstrate that the proposed method achieves state-of the-art performance both for character detection and word recognition.**
Keywords— (Text Detection system, StrokeWidth Transforms, Component Generation)

Introduction

Scence text localization and recognition (also known as photo OCR or text-in-the-wild problem) is an open problem with many interesting applications, ranging from helping the visually impaired, language translation with automatic input of text written in an unknown script, to indexing large image and video databases by their textual content (e.g., Google Street View, Flickr, etc.). Unlike traditional printed document OCR, no scene text recognition methods has yet achieved sufficient accuracy for practical applications—the winning method in the most recent ICDAR 2013 contest was able to localize only 66 percent words correctly despite the fact that the dataset is not fully realistic—the word orientations are only horizontal, they occupy a significant part of the image, there is no perspective distortion or significant noise. In the competition, text recognition was evaluated in an artificial setup where the words are manually localized by a human annotator. The winner was able to correctly recognize 82 percent of such "cut-out" words; end-to-end text recognition combining both localization and recognition was not included because of too few potentially participating methods. Text localization can be computationally very expensive because in an image of N pixels generally any of its 2N subsets can correspond to text. Generally, two approaches to deal with this issue exist in the literature. The methods in the first group exploit a sliding-window approach to localize individual characters or whole words drawing inspiration from other object detection problems where this approach has been has been successfully applied .Strengths of such methods include robustness to noise and blur, since features aggregated over the whole region of interest are exploited. The main drawback is that the number of rectangles that needs to be evaluated grows rapidly when text with different scale, aspect, rotation and other distortions has to be found.The second, recently more popular approach is based on localizing individual characters as connected components using local properties of an image such as color, intensity or stroke-width. The complexity of these methods does not depend on the parameters of the text as characters of all scales and orientations can be detected in one pass and the connected component representation also provides a character segmentation which can be exploited in an OCR stage. The biggest disadvantage of such methods is a dependence on the assumption that a character is a connected component, which is brittle—a change in a single image pixel introduced by noise can cause an unproportional change in the connected component size, shape or other properties, which subsequently may affect its classification.

1. Text Recognization

Text detection and recognition in natural scene images has applications in computer vision systems such as license plate detection, automatic street sign translation, image retrieval and help for visually impaired people. Scene text, however, has complex background, image blur, partially occluded text, variations in font-styles, image noise and varying illumination.Hence scene text recognition is a challenging computer vision problem. This work addresses the problem of dictionary driven end-to-end scene text recognition, which is divided into a text detection problem and a text recognition problem. For text detection an AdaBoost sliding window classifier is used to detect text in multiple scales. Text information extracted from scene images is often the key clue for better performance of scene understanding and image retrieval. However, the clutter background and variations, which are intrinsic in scene images, make the natural scene character recognition task rather complicated. To overcome these disadvantages, we propose a novel approach for character recognition task in natural scene images. In the method, character classes are described by groups of local features using a probabilistic model. Structures of characters are represented by mutual positions of local features. For model learning, parameter estimating is done through expectation-maximization in a weak-supervised manner. Experiment results over datasets which includes both synthetic and authentic data demonstrate the validity of the approach.

.

1. Problem Statement

Problem statement involves text localization in a natural scene image. Once the text is localized the next step is to extract the meaningful text from the text localized region. Neural network / SVM based classifier can be used for text recognition.

Aim of Paper

The main aim of this thesis is to propose a dictionary driven end-to-end scene text recognition system, achieving competitive performance. Hence, a system is proposed consisting of a sliding window classifier which detects text regions. In these regions Connected Components (CCs) are labeled by machine learning models and grouped to textlines. Textlines are split into words, which are classified by the proposed text recognition system to predict the word for the cropped patch.

2. Literature Survey

Scene Text Detection and Recognition: Recent Advances and Future Trends As a product of human abstraction and manipulation, text in natural scenes directly carries high level semantics. This property makes text present in natural images and videos a special, important source of information. The rich and Received accepted month precise information embodied in text can variety of vision-based applications, such as image search [1], target geolocation [2], human-computer interaction [3], robot navigation [4] and industrial automation [5].

Professor Latika R. Desai1 , Miss. Poonam B. Kadam , Professor Swati Shinde **"Review on Text Detection Methodology from Images".** Number of approaches for text detection in images has been proposed into the past. Automatic detection and translation of text in images done using different techniques proposed. These methods aim to detect the characters based on general properties of character pixels. The distribution of edges, color is used in many text detection methods also for low resolution document are processed by perticuler method.

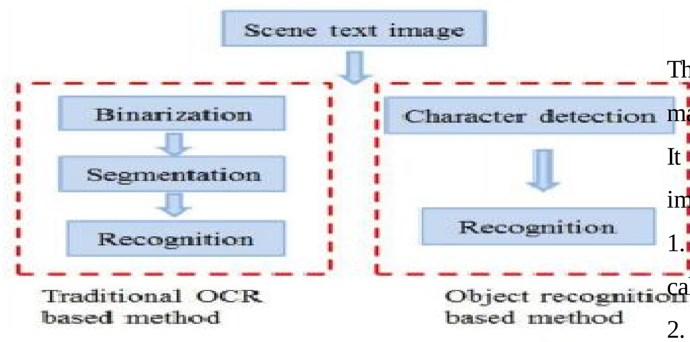
Rampurkar Vyankatesh Vijaykumar, Gyankamal J.Chhajed, Sahil Kailas Shah **"Review on Text String Detection from Natural Scenes"** In this paper, Basilios Gatos proposes a novel methodology for text detection in natural scene images [2]. The proposed methodology is based on an efficient binarization and enhancement technique followed by a suitable connected component analysis procedure. Image binarization successfully processes natural scene images having shadows, non-uniform illumination, low contrast and large signal-dependent noise. Connected component analysis is used to define the final binary images that mainly consist of text regions.

3. Proposed Work

1. Character Detection

In the first stage, character candidates are efficiently detected as external regions selected in a two-stage classification process, operating on a coarse Gaussian scale space pyramid and on multiple image projections

IORD Journal of Science & Technology
E-ISSN: 2348-0831
Volume 3, Issue 3 (MAR-APRIL 2017) PP71-75
IMPACT FACTOR 1.719
www.iord.in



The traditional OCR-based method and object recognition-based method

4. Stroke Width Transforms

A stroke in the image is a continuous band of a nearly constant width. The Stroke Width Transform (SWT) is a local operator which calculates for each pixel the width of the most likely stroke containing the pixel. First, all pixels are initialized with ∞ as their stroke width. Then, we calculate the edge map of the image by using the canny edge detector. We consider the edges as possible stroke boundaries, and we wish to find the width of such stroke.

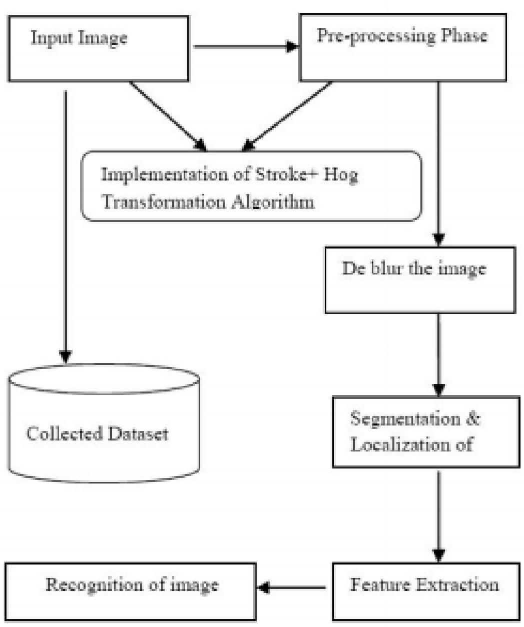
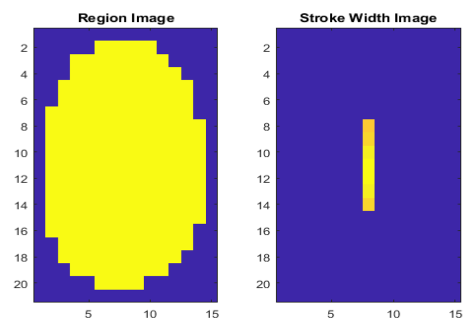


Fig3: Step of Stroke width transforms method

5. Implementation of Stroke Width Transform is as follows

The SWT Text Detector application is designed to locate and mark the regions of an image that are suspected to contain text. It returns an image of the same size as the input image. The implementation of the application contains several parts:

1. The stroke width transform: edge detection and stroke width calculation.
2. Removing stray lines from the SW map.
3. Finding letter candidates: finding the connected components and detecting the components with the features of a letter.
4. Grouping the letters into regions of text.



In the images shown above, notice how the stroke width image has very little variation over most of the region. This indicates that the region is more likely to be a text region because the lines and curves that make up the region all have similar widths, which is a common characteristic of human readable text.

6. Text Segmentation and Localization

In this phase, we have to segment the characters in the image by bounding box method or by bounding the characters by edges and after that finally we are in position to normalize the image to find out correct characters present in the input images and localize the text in the input image.

7. Recognition:

On the basis of above all procedure of the proposed algorithm finally the recognition and text localization of input images will be done in this phase.

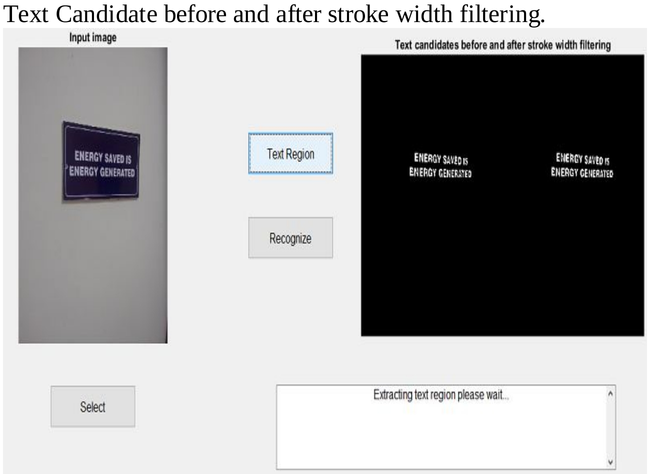
IORD Journal of Science & Technology
E-ISSN: 2348-0831
Volume 3, Issue 3 (MAR-APRIL 2017) PP71-75
IMPACT FACTOR 1.719
www.iord.in

Final Result

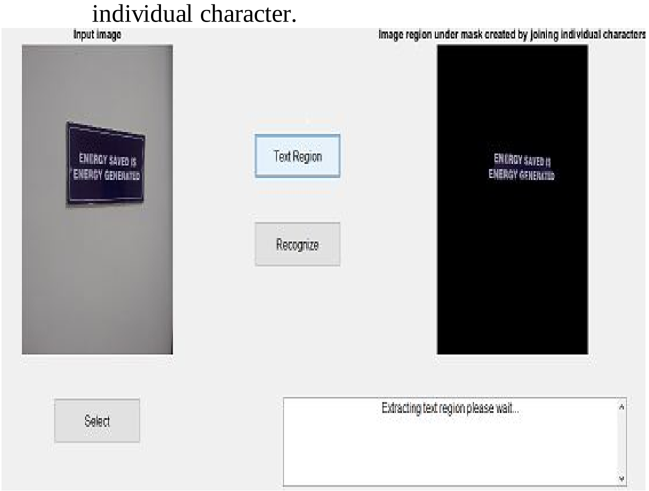
The Final Result Generated as follows

1. **Home Page**

Text Candidate before and after stroke width filtering.

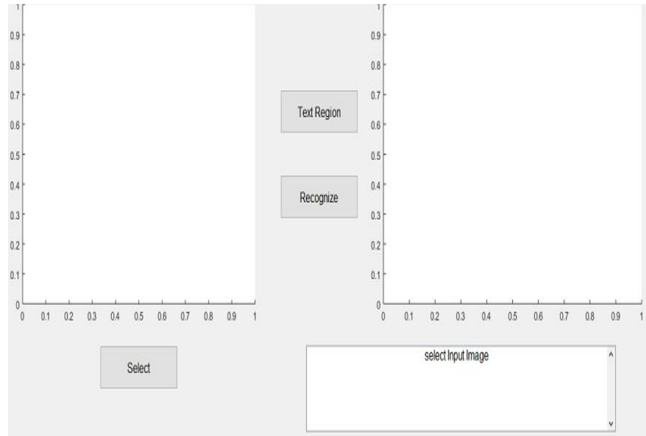


2. Image Region under mask created by joining individual character.

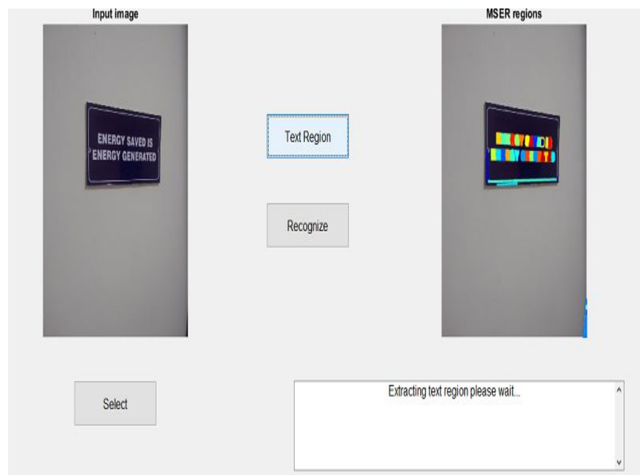


Maximally Stable Extremal Regions (MSER) is a feature detector; Like the SIFT detector, the MSER algorithm extracts from an image I a number of co-variant regions, called MSERs. An MSER is a stable connected component of some level sets of the image I. Optionally, elliptical frames are attached to the MSERs by fitting ellipses to the regions.

3. **Initial Stage**

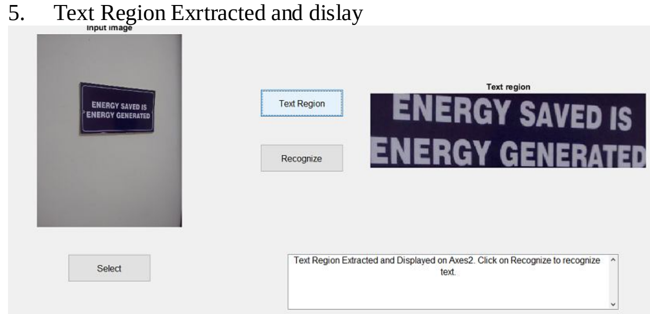


4. **MSER Region**



The MSER algorithm extracts from an image a number of co-variant regions. this algorithm is to match MSERs to establish correspondence points between images. First MSER regions are computed on the intensity image (MSER+) and on the inverted image (MSER-). Measurement regions are selected at multiple scales.

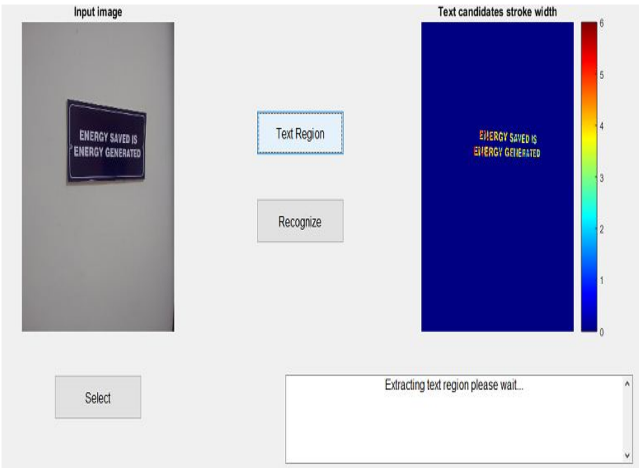
5. Text Region Exrtracted and dislay



A methodology to detect and extract text regions from low resolution natural scene images is presented. The texture features are then

IORD Journal of Science & Technology
E-ISSN: 2348-0831
Volume 3, Issue 3 (MAR-APRIL 2017) PP71-75
IMPACT FACTOR 1.719
www.iord.in

obtained on every 50×50 block of the processed image and potential text blocks are identified using newly defined discriminant functions. Further, the detected text blocks are merged and refined to extract text regions



CONCLUSIONS

In this work we show how to leverage on the idea of the recovery of stroke width for text detection. We define the notion of a stroke and derive an efficient algorithm to compute it, producing a new image feature. Once recovered, it provides a feature that has proven to be reliable and flexible for text detection. Unlike previous features used for text detection, the proposed SWT combines dense estimation (computed at every pixel) with non-local scope (stroke width depends on information contained sometimes in very far apart pixels). Compared to the most recent available tests, our algorithm reached first place and was about 15 times faster than the speed reported there. The feature was dominant enough to be used by itself, without the need for actual character recognition step as used in some previous works [3]. This allows us to apply the method to many languages and fonts.

REFERENCES

[1]. S. Tsai, H. Chen, D. Chen, G. Schroth, R. Grzeszczuk, and B. Girod. Mobile visual search on printed documents using text and low bit-rate features. In Proc. of ICIP, 2011.

[2]. D. B. Barber, J. D. Redding, T. W. McLain, R. W. Beard, and C. N. Taylor. Vision-based target geo-location using a fixed-wing miniature air vehicle. Journal of Intelligent and Robotic Systems, 47(4):361–382, 2006.

[3]. B. Kisacanin, V. Pavlovic, and T. S. Huang. Real-Time Vision for Human-Computer Interaction. Springer, Heidelberg, 2005.

[4]. G. N. DeSouza and A. C. Kak. Vision for mobile robot navigation: A survey. IEEE Trans. PAMI, 24(2):237–267, 2002.

[5] D. Karatzas, F. Shafait, et al., "ICDAR 2013 robust reading competition," in Proc. 12th Int. Conf. Document Anal. Recog., 2013, pp. 1484–1493.

[6] A. Bissacco, M. Cummins, et al., "PhotoOCR: Reading text in uncontrolled conditions," in Proc. IEEE Int. Conf. Comput. Vis., 2013, pp. 785–792.

[7] K. Wang, B. Babenko, and S. Belongie, "End-to-end scene text recognition," in Proc. IEEE Int. Conf. Comput. Vis., 2011, pp. 1457–1464.

[8] J. J. Lee, P.-H. Lee, et al., "AdaBoost for text detection in natural scene," in Proc. Int. Conf. Document Anal. Recog., 2011, pp. 429–434.

[9] P. Viola and M. J. Jones, "Robust real-time face detection," Int. J. Comput. Vis., vol. 57, no. 2, pp. 137–154, 2004. [6] B. Epshtein, E. Ofek, and Y. Wexler, "Detecting text in natural scenes with stroke width transform," in Proc. IEEE Conf. Comput. Vis. Pattern Recog., 2010, pp. 2963–2970.