

## **Text Extraction in Natural Scene Images**

Shahina k m

Shahinakm2@gmail.com

**Abstract:** Text extraction in scene images is a process of extracting textual information from natural scene images. Various researches have been carried out over the past years to find an efficient and accurate text recognition system. Text recognition in scene images has an inevitable role in different application areas such as artificial intelligence, robotics and automation system. Currently Different techniques are proposed from various research areas such as image processing, graph theory and neural networks for text detection and recognition in natural scene images. In this paper different algorithms and methods applied for text extraction in the literature are summarized. A general frame work for text extraction is also presented in this paper.

**Keywords:** Optical Character Recognition, Region Extraction, Text extraction

### **1 INTRODUCTION**

Text extraction in natural scene images is an important research area in computer vision. Text detection and extraction is challenging due to different factors such as complex background, low quality, distortion of the image, different size, color, font, language and arrangement orientations of the text to be extracted. And large number of techniques and algorithms are proposed to overcome these challenges. Text extraction is widely applied in recognizing signs in driver assisting systems, indexing multimedia achieves and various automation systems. Recently, visual content analysis plays a vital role in computer vision including event detection [18], action detection [17] and object detection [16]. Text can be considered as a special kind of object and widely applied in mobile based text recognition, navigation, multimedia retrieval and so on. Conventional method uses optical character recognition(OCR) which is based on the assumption that black text is printed on white background. Which does not consider the complexity of background image, variations in color, font, size and alignments.

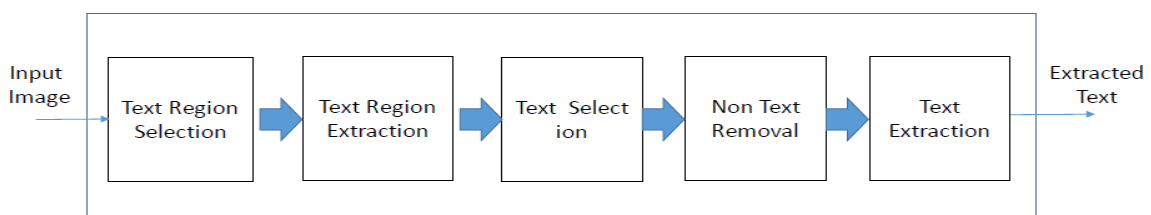
## 1.1 General Framework of Text Extraction

General procedure for the extraction of text mainly includes four steps. First one is text region detection, which find out the local text region within the natural scene images. Then the detected region is extracted and this region may contain overlapped areas and non-text objects. So, the next step is to remove the non-text region. And finally, the text is extracted and various enhancement techniques are applied to produce an accurate and clearly defined text.

For extracting the local text region, different technologies such as Sliding Window method, Nonlinear Niblack method, Stroke Width Transformation(SWT), Maximally Stable External Region(MSER), Non-Maximal Suppression(NMS) are used. Then different classifiers such as SVM, cascade of Adaboost and Convolutional Neural Network classifiers are used for the removal of non-text elements.

## 2 Significance of text extraction

The rapid enhancement in computer vision also increases the significance of text extraction in natural images. A few years back, various image processing techniques such as Optical Character Detection, edge detection and texture detections are used. But today researchers attempted to use graph theory and neural network concept for text extraction. Which helps to extract information from complicated, blurred images. A study of different methods for text recognition is summarized in table 1.



**Fig 1.** General Block diagram of Text Extraction from natural scene images

## 2.1 Review of Text Extraction

**Table 1:** Summary of Text Extraction techniques

Author	Region extraction method	Text and Non-text classification technique	Advantage
Ye et al (2003) [9]	Edge features and Morphology operation	SVM classifier and Wavelet features	Robust to language, font, color and size
Chan and Yuille (2004) [1]	Sliding Window method	Adaboost classifier containing 79 features	High speed and detect text of all form
Zhu et al (2007) [10]	Nonlinear Niblack method	Adaboost classifier containing 12 CC features	90% accuracy on unconstrained city features
Epstein et al (2010) [2]	Stroke Width Transform	Connected Component analysis	Simple and which detect text in many fonts and languages
Newmann and Matas (2011) [4]	Maximally stable external regions(MSER)	Effective pruning with exhaustive enumeration	Exploits higher order properties of text
Wang et al (2012) [6]	CNN with two convolution layers	Non-maximal suppression	Achieve accuracy of 83.9 on ICDAR 2003 data set
Pan et al (2011) [5]	Niblack and region-based method	Integrating region based and robust CC method	Perform well on multilingual text images
Yao et al (2012) [8]	Sliding Window Transformation	CC analysis and CC chain analysis	Detect text of arbitrary orientation
Newman and Mats (2015) [3]	Stroke Support Pixels and MSER	Grab cut iterative segmentation	Robust to blur, illumination, color and texture variation
Xiabing et al (2015) [7]	Multi-layer segmentation	Higher order conditional random field	82% precision for text from natural scene images
Pan and Hou (2009) [11]	Text region detector using local binarization approach	Conditional random field model	Better performance for unconstrained scene text localization
Yin et al (2013) [12]	Maximally stable external region method	Single link clustering algorithm and Distance metric learning algorithm	Fast and accurate extraction for low quality images

Liu et al (2014) [13]	Stroke edge method	Graph cut algorithm	High accuracy for all images especially in poor quality
Winger et al (2014) [14]	Logical stroke width method	Thresholding local variance	Efficient for low contrast LCD displays
Song et al (2017) [15]	Text edge box and conventional neural network	Hierarchical recursive method of CC analysis and Random field-based analysis	Robust detection of text in complex scenes

## Conclusion

Text extraction can be defined as the process of extracting text from natural scene images. Researchers proposed several novel text extraction methods to extract text from images having text with variabilities in font, color, language, and orientations. A general framework for text extraction is presented in this paper. Over fifteen text extraction methods, region text extraction technique, Text and non-text classification method and advantages are also reviewed in this paper.

## References

1. Chen XR, Yuille AL (2004) Detecting and Reading Text in Natural Scenes. Proc. IEEE Conf. On Computer Vision and Pattern Recognition, pp II366-II373
2. Epshtein B, Ofek E, Wexler Y (2010) Detecting Text in Natural Scenes with Stroke Width Transform. In: Proceedings IEEE Conf. on Computer Vision and Pattern Recognition, pp 2963–2970
3. Neumann L, Matas J (2015) Efficient Scene Text Localization and Recognition with Local Character Refinement. IEEE Conf. on Computer Vision and Pattern Recognition
4. Neumann L, Matas J (2011) Text Localization in Real-world Images using Efficiently Pruned Exhaustive Search. In: Proceedings of the 11th International Conference on Document Analysis and Recognition, PP 687–691
5. Pan YF, Hou XW, Liu CL (2011) A Hybrid Approach to Detect and Localize Texts in Natural Scene Images. IEEE Trans Image Process 20(3):800–813

6. Wang T, Wu DJ, Coates A, Ng AY (2012) End-to-End Text Recognition with Convolutional Neural Networks. In: Proceedings of the 21st International Conference on Pattern Recognition, pp 3304–3308
7. Wang XB, Song YH, Yuan LZ, Xin JM (2015) Natural scene text detection with multi-layer segmentation and higher order conditional random field-based analysis.
8. Yao C, Bai X, Liu WY, Ma Y, Tu ZW (2012) Detecting Texts of Arbitrary Orientations in Natural Images. In: Proceedings IEEE Conf. on Computer Vision and Pattern Recognition, pp 1083–1090
9. Ye QX, Gao W, Wang WQ, Zeng W (2003) A robust text detection algorithm in images and video frames. In: Proceedings of the 2003 Joint Conference of the 4th International Conference on Information, Communications and Signal Processing and 4th Pacific-Rim Conference on Multimedia, vol 2, pp 802–806
10. Zhu KH, Qi FH, Jiang RJ, Xu L (2007) Automatic character detection and segmentation in natural scene images. *J Zhejiang Univ Sci A* 8(1):63–71
11. Yi-Feng Pan, Xinwen Hou, Cheng-Lin Liu (2009) Text Localization in Natural Scene Images based on Conditional Random Field. In: Proceedings of the 10th International Conference on Document Analysis and Recognition
12. Xu-Cheng Yin, Xuwang Yin, Kaizhu Huang, and Hong-Wei Hao (2013) Robust Text Detection in Natural Scene Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol: 36, PP 970 - 983
13. Xiaopei Liu, Zhaoyang Lu, Jing Li, Wei Jiang (2014) Detection and Segmentation Text from Natural Scene Images Based on Graph Model. *Wseas transactions on Signal Processing*, Vol 10, PP 124-135
14. Lowell I Winger, Jhon A Robinson, M Ed Jernigan (2014) Low-Complexity Character Extraction in Low-contrast Scene Images. *Int. J. Patt. Recogn. Artif. Intell.* PP:113-135.
15. Xiaobing Wang<sup>1</sup>, Yonghong Song<sup>1</sup>, Yuanlin Zhang<sup>1</sup>, Jingmin Xin<sup>1</sup> (2017) A hierarchical recursive method for text detection in natural scene images. *Multimedia Tools Appl* 76:26201–26223
16. Cabrera CR, Sastre RJ, Rodriguez JA, Bascon SM (2012) Surfing the point clouds: Selective 3D spatial pyramids for category-level object recognition. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp 3458–3465.
17. Wang S, Yang Y, Ma ZG, Li X, Pang CY, Hauptmann AG (2012) Action recognition by exploring data distribution and feature correlation. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp 1370–1377.
18. Yang Y, Ma ZG, Xu ZW, Yan SC, Hauptmann AG (2013) How Related Exemplars Help Complex Event Detection in Web Videos? In: Proceedings of IEEE International Conference on Computer Vision, PP 2104–2111.