

## REVIEW OF TEXT RECOGNITION ALGOROTHS

Петухова К.С.

Науковий керівник – проф. Кузьомін О.Я.

Харківський національний університет радіоелектроніки

(61166, Харків, просп. Науки, 14, каф. Інформатики, тел.)

e-mail: katernyna.petukhova@nure.ua, тел. +380634731403

This work is created to consider various methods of text recognition on images. The aim of it is to categorize and analyze different algorithms. Text recognition can be defined as main technique for image labeling. In recent years machine learning and pattern recognition, which are useful for text extraction, became very popular in computer vision area. That is why a lot of new schemes and algorithms have appeared. There are a lot of different algorithms and some of them can be useful for one purpose and another for second one. That is why it is important to understand and consider variety of all these approaches and methods.

Firstly, let's review top-down, bottom-up and hybrid approaches of text recognition in scene-text images.

The main idea of top-down approach is to separate image to segments, starting from paragraphs and ending with characters, using recursive algorithms. According to comparing analysis [1] the highest precision, which is 91.8%, has algorithm, created by Liu and Samarabandu [2]. Results are very high but this method has errors because of using edge detector, which sometimes extracts unnecessary objects instead of words.

Bottom-up approach provides techniques, which divide images on character level and then aggregate small parts into words and sentences. All algorithms use different frameworks, techniques and function to extract characters and group them into words. For instance, divide image into some areas and assign label to each part [3]. Use two different detection operations for small and big characters [4] etc.

Secondly, let's review segmentation of document images. The main problem of this task are handwritten texts, so the biggest part of mentioned algorithms are aimed at recognition of handwritten characters. In general, segmentation of document images can be divided on 3 steps: text line, word and character segmentation. Sometimes second step needs previously extracted lines from the document, but in another cases this condition is not necessary.

Bolan Su and Shijian Lu propose new technique for text recognition in scene images in their article "Accurate Scene Text Recognition Based on Recurrent Neural Network"[5]. The main idea of their method is to recognize whole word instead of image segmentation to character level.

According to Su and Lu in their paper, a lot of algorithms make errors because of noise, perspective distortion and fonts. Author think that it is better to find word without dividing images to parts, what will decrease amount of fails. Firstly, the text-scene image is converted into unsegmented sequence, which

consists of column vectors of Histogram of Oriented Gradient (HOG). Second step is classifying these vectors to the words using the Recurrent Neural Network (RNN). Instead of simple Neural Network RNN was used, because, as Su and Lu explain, it uses long short-term memory (LSTM) structure for information incorporating. Second argument, presented by authors is absence of necessity to apply image labeling for each single column vector of the sequence. Su and Lu say, that it is very important factor, because of blurring, noise and distortion task of vector labeling becomes impracticable.

For word storing Su and Lu use as known Lexicon. It is list of potential words, which is used for looking for the best condition satisfying words. Algorithm counts possible score for the words and choose the best option if the score is maximum.

This work is helpful for understanding basic principles of text extraction and provide an overview of different text-segmentation algorithms, which use various methods and techniques. It considers, that it is better to use different schemes, because their performance depends on noise, background, light etc. That is why they compare different techniques, but with the same ideas.

### **References:**

1. Sahare P. & Dhok S. B. (2017) Review of Text Extraction Algorithms for Scene-text and Document Images, IETE Technical Review, 34:2, 144-164, DOI:10.1080/02564602.2016.1160805
2. Liu X. and Samarabandu J., "Multiscale edge-based text extraction from complex images," in IEEE International Conference on Multimedia and Expo (ICME), Toronto, ON, Jul. 9A12, 2006, pp. 1721A1724.
3. Lee S. H. and Kim J. H., "Integrating multiple character proposals for robust scene text extraction," Image Vision Comput., Vol. 31, no. 11, pp. 823A840, Nov. 2013.
4. Ezaki N., Bulacu M., and Schomaker L., "Text detection from natural scene images: towards a system for visually impaired persons," in Proceedings of 17th International Conference on Pattern Recognition, Cambridge, Aug. 23A26, 2004, Vol. 2, pp. 683A686.
5. Su B., Lu S. (2015) Accurate Scene Text Recognition Based on Recurrent Neural Network. In: Cremers D., Reid I., Saito H., Yang MH. (eds) Computer Vision -- ACCV 2014. ACCV 2014. Lecture Notes in Computer Science, vol 9003. Springer, Cham