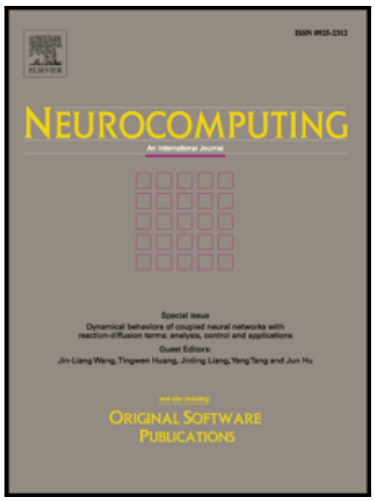
Accepted Manuscript



Scene Text Segmentation using Low Variation Extremal Regions and Sorting Based Character Grouping

Matko Saric

PII: S0925-2312(17)30837-8
DOI: 10.1016/j.neucom.2017.05.021
Reference: NEUCOM 18434

To appear in: *Neurocomputing*

Received date: 24 May 2016
Revised date: 14 March 2017
Accepted date: 10 May 2017

Please cite this article as: Matko Saric, Scene Text Segmentation using Low Variation Extremal Regions and Sorting Based Character Grouping, *Neurocomputing* (2017), doi: 10.1016/j.neucom.2017.05.021

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Scene Text Segmentation using Low Variation Extremal Regions and Sorting Based Character Grouping

Matko Saric

FESB, University of Split, Croatia

Abstract

Extraction of textual information from natural scene images is a challenging task due to imaging conditions and diversity of text properties. Segmentation of scene text is important step in the pipeline that significantly affects the final recognition performance. In this paper I propose a new scene text segmentation method. Firstly, a novel approach for character candidates generation based on extremal regions (ERs) is introduced. Subpaths having low area variation are extracted from ER tree. Instead of using minimum variation criterion for selection of character candidates, position of ER in extracted subpath is used as criterion for that purpose. Each subpath is represented by one ER that is sent to SVM-based classification step. After that a novel method for character candidates grouping is used to discard non-character objects that are wrongly classified as characters. Proposed approach estimates vertical positions of the lines by sorting y coordinates of region centroids and checks spatial relation of adjacent regions in the line. This step enhances precision significantly and has lower computational complexity compared to hierarchical clustering methods. Finally, the last step is restoration of character ERs erroneously eliminated by SVM classifier where text layout properties are exploited to correct false negative classifications. Experimental results obtained on the ICDAR 2013 dataset show that the proposed character candidates generation method efficiently prunes repeating regions and achieves character recall rate superior to recently published ER based method. Proposed segmentation algorithm obtains competitive performance compared to state-of-the-art methods.

Keywords: Extremal regions, scene text segmentation, Character grouping

ACCEPTED MANUSCRIPT

دریافت فوری ← متن کامل مقاله

ISIArticles
مرجع مقالات تخصصی ایران

✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
✓ امکان دانلود نسخه ترجمه شده مقالات
✓ پذیرش سفارش ترجمه تخصصی
✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
✓ دانلود فوری مقاله پس از پرداخت آنلاین
✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات