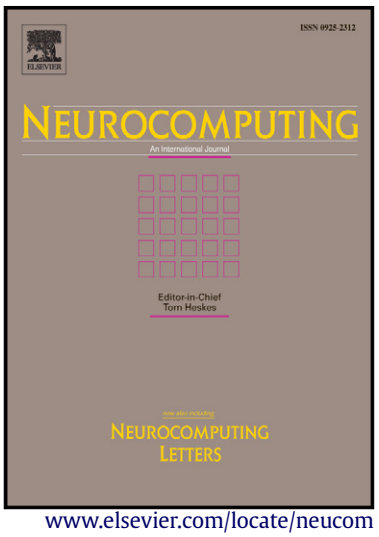
Author's Accepted Manuscript



www.elsevier.com/locate/neucom

Active contour based optical character recognition for automated scene understanding

Joanna Isabelle Olszewska

PII: S0925-2312(15)00213-1
DOI: http://dx.doi.org/10.1016/j.neucom.2014.12.089
Reference: NEUCOM15180

To appear in: *Neurocomputing*

Received date: 1 June 2014
Revised date: 5 November 2014
Accepted date: 1 December 2014

Cite this article as: Joanna Isabelle Olszewska, Active contour based optical character recognition for automated scene understanding, *Neurocomputing,* http://dx.doi.org/10.1016/j.neucom.2014.12.089

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 galley proof before it is published in its final citable 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

Active Contour Based Optical Character Recognition for Automated Scene Understanding

Joanna Isabelle Olszewska

School of Computing and Technology, University of Gloucestershire
The Park, Cheltenham, GL50 2RH, United Kingdom
jolszewska@glos.ac.uk

Abstract

In this paper, we present a new optical character recognition (OCR) approach which allows real-time, automatic extraction and recognition of digits in images and videos. Our method relies on active contours in order to robustly extract optical characters from real-world visual scenes. The detected character recognition is based on template matching. Our developed system has shown excellent results when applied to the automated identification of team players' numbers in sport datasets and has outperformed state-of-the-art methods.

Key words: Pattern Recognition, Information Fusion, Fast Segmentation, Scene-Text Localization and Extraction, Active Contours, Snake Evolution, Chromaticity, Online Optical Character Recognition, Sport Events, Image and Video Datasets, Automated Scene Understanding

Accepted manuscript

1. Introduction

Scene text extraction in images and videos is of prime importance for automated scene understanding and video content analysis [1]. Besides the perceptual content of the images and videos analyzed with vision-based techniques like object detection [2], object recognition [3], or tracking [4], the semantic content of the images and videos provides inestimable information in terms of scene objects and their relations [5], [6], useful for applications such as automated comments production and media archiving.

In particular, sport events' refereeing and analysis rely on the identification of team players based on the recognition of the numbers on their uniform rather than on their face recognition. Indeed, face detection techniques [7] are not intrinsically adapted to identify a player whose back is turned to the camera, in which case his face is poorly or not visible at all.

Automatic extraction of the scene text involves the use of an optical character recognition (OCR) system to detect and recognize the textual characters from a given image.

In this work, we focus on the automatic player identification in images of any type for sport scene analysis, based on the detection and recognition of the player's jersey number, and therefore, on the development of a full, efficient OCR system for this purpose.

OCR major phases are (i) character extraction and (ii) character recognition. In the first step, the system localizes and extracts the character by detecting its geometrical features like edges [8] or color features [9], or both [10]. In the second step, character recognition is usually performed by matching [11] or by using classifiers, e.g. AdaBoost [12]. However, these existing OCR systems are mainly applied to recognize license plate numbers or handwritten characters, whereas player number recognition presents additional challenges. Indeed, the foreground, i.e. the character, could be highly skewed with respect to the camera, or the background, i.e. the jersey, could be folded so that part of the number could be hidden. Moreover, sport images are often blurred, since cameras or players or both are quickly moving.

Most of the existing OCR approaches developed to identify numbers on team player's uniform exploit the temporal redundancy of a character across several frames and thus are limited to analyze only videos [13], [14], [15], [16], [17], [18] and not suited for tasks such as still image dataset retrieval. Other works [19] use both facial and textual cues, but present a low computational speed.

Hence, variations of scene text due to differences in size, style, font, color, orientation, and alignement, as

Download English Version:

https://daneshyari.com/en/article/6865792

Download Persian Version:

https://daneshyari.com/article/6865792

Daneshyari.com