

Preface

**Faisal Shafait · Dimosthenis Karatzas ·
Seiichi Uchida · Masakazu Iwamura**

Published online: 3 May 2015
© Springer-Verlag Berlin Heidelberg 2015

Document Image Analysis (DIA) was established to address the need for the automated analysis of scanned paper documents, but has evolved to incorporate a great variety of text containers including digital-born (Web and email) images, scene images, and video sequences. These types of text containers introduce a range of new challenges for reading systems, stemming from perspective distortions, a large variety of fonts and backgrounds, anti-aliasing, and compression artefacts. Addressing the need of reading textual content in such domains entails a paradigm shift in the DIA field that calls for a change in the definition of “document” to include complex unconstrained text containers, as opposed to the traditional scanned paper document. Robust Reading refers to the research area dealing with the interpretation of written communication in unconstrained settings.

Robust Reading is generally perceived as the detection and recognition of textual information in scene images, but in the wider sense it refers to techniques and methodologies that have been developed specifically for complex text containers beyond scanned paper documents. Robust Reading is

at the meeting point between camera-based document analysis and scene interpretation and serves as common ground between the DIA community and the wider computer vision community.

To gauge research progress in the field of text recognition in unconstrained environments, the ICDAR Robust Reading Competitions (held in 2003, 2005, 2011, 2013, 2015) have become the defacto international benchmark and provide standardized datasets, evaluation protocols, and software for the scientific community. The first two editions of the competition in 2003 and 2005 focused on detection and recognition of scene text. Research in this topic was in its infancy at that stage, so the competitions did not receive large participation and were discontinued. However, research in this direction gained momentum in the following years owing largely to the rapid proliferation of smartphone devices having built-in cameras. The ground-truthed datasets developed for ICDAR 2003 competition became a yardstick to evaluate algorithms for scene text detection and recognition. The ICDAR Robust Reading competition was revived in 2011 to provide a standardized platform for researchers for evaluation of their algorithms. Besides scene text detection and recognition, an additional dimension of extracting text from Born-Digital images (e.g. images in spam emails, or advertisements on Web pages) was added. This time, the competition received tremendous success with participation from 36 teams in different challenges and tasks. Hence, the competition was held again in 2013 including reading text from videos as an additional challenge receiving 58 public submissions and more than 1800 private submissions during the year after the competition. The 2015 edition of the competition, ongoing at the time of writing, includes evaluation of end-to-end text recognition systems keeping in view the maturity of this research field.

F. Shafait (✉)
National University of Science and Technology,
Islamabad, Pakistan
e-mail: faisal.shafait@seecs.edu.pk

D. Karatzas
Computer Vision Centre, Universitat Autònoma de Barcelona,
Barcelona, Spain
e-mail: dimos@cvc.uab.es

S. Uchida
Kyushu University, Fukuoka, Japan
e-mail: uchida@ait.kyushu-u.ac.jp

M. Iwamura
Osaka Prefecture University, Osaka, Japan
e-mail: masa@cs.osakafu-u.ac.jp

The IJDAR Special Issue on Robust Reading provides a snapshot of the state-of-the-art algorithms. A call for papers was circulated to solicit original works on topics related to the detection, extraction, and recognition of textual content in complex unconstrained text containers. In response to the call, 18 submissions were received for this special issue. After two review rounds, four papers were accepted for publication in this Special Issue leading to an acceptance rate of 22%. In selecting the papers for this special issue from the wide range of work being done in this area, we focused on research topics that we believe have the broadest range of applicability to the community working in Robust Reading. The paper by Changsong Liu et al. discusses a model-free dewarping approach for rectification of camera-captured document images, which is a key prerequisite for accurate recognition of such documents. Yirui Wu et al. present a pre-processing approach for recognition of text in videos. The presented method per-

forms thinning that preserves visual topology of characters in videos to achieve higher text recognition rates. The paper by Shijian Lu et al. presents a scene text extraction technique that automatically detects and segments text based on edges from scene images. Finally, Muhammad Fraz et al. demonstrate that the colour information within an image is good enough to distinguish text regions from the surrounding noise and effectively utilizing this information results in a significant performance improvement for the recognition of characters and words.

We would like to thank the authors who contributed to this special issue and express our gratitude to the reviewers who did a tremendous job within tight deadlines. Finally, we are indebted to Springer support staff as well as the IJDAR Editors-in-Chief—Prof Koichi Kise, Prof Daniel Lopresti, and Prof Simone Marinai—for their constant support and understanding in completing this effort.