

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Alfred Kobsa

University of California, Irvine, CA, USA

Friedemann Mattern

ETH Zurich, Switzerland

John C. Mitchell

Stanford University, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz

University of Bern, Switzerland

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

TU Dortmund University, Germany

Madhu Sudan

Microsoft Research, Cambridge, MA, USA

Demetri Terzopoulos

University of California, Los Angeles, CA, USA

Doug Tygar

University of California, Berkeley, CA, USA

Gerhard Weikum

Max Planck Institute for Informatics, Saarbruecken, Germany

Masakazu Iwamura Faisal Shafait (Eds.)

Camera-Based Document Analysis and Recognition

4th International Workshop, CBDAR 2011
Beijing, China, September 22, 2011
Revised Selected Papers

Volume Editors

Masakazu Iwamura
Osaka Prefecture University
Graduate School of Engineering
Dept. of Computer Science and Intelligent Systems
1-1 Gakuencho, Naka Sakai, Osaka 599-8531, Japan
E-mail: masa@cs.osakafu-u.ac.jp

Faisal Shafait
German Research Center for Artificial Intelligence (DFKI GmbH)
Multimedia Analysis and Data Mining Competence Center
Trippstadter Str. 122, 67663 Kaiserslautern, Germany
E-mail: faisal.shafait@dfki.de

ISSN 0302-9743
ISBN 978-3-642-29363-4
DOI 10.1007/978-3-642-29364-1
Springer Heidelberg Dordrecht London New York

e-ISSN 1611-3349
e-ISBN 978-3-642-29364-1

Library of Congress Control Number: 2012934970

CR Subject Classification (1998): I.4, I.5, I.7, I.2.7, H.3, H.5, H.4

LNCS Sublibrary: SL 6 – Image Processing, Computer Vision, Pattern Recognition, and Graphics

© Springer-Verlag Berlin Heidelberg 2012

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The pervasiveness and wide-spread availability of camera phones and hand-held digital still/video cameras have led the community to recognize document analysis and recognition of digital camera images as a promising and growing field of research. Constraints imposed by the memory, processing speed and image quality are leading to new interesting open problems which cannot be directly resolved by traditional techniques.

To cater to the demands of camera-based document processing, the idea of a new satellite workshop of the International Conference on Document Analysis and Recognition (ICDAR) was conceived by Koichi Kise. Together with David Doermann, he took the responsibility of organizing the first workshop on Camera-Based Document Analysis and Recognition as a satellite workshop of ICDAR 2005 in Seoul, Korea. The workshop was very well received by the community and hence it was held again in 2007 (Curitiba, Brazil), and 2009 (Barcelona, Spain) with the corresponding ICDAR conferences. Following the success of the past three workshops, the 4th International Workshop on Camera-Based Document Analysis and Recognition (CBDAR 2011) was held in Beijing, China, quite successfully with 68 participants. The workshop aimed to provide an opportunity to researchers and developers from various backgrounds to exchange their ideas and explore new research directions through presentations of the latest research activities and discussions.

In six years since the first CBDAR was held, the situation surrounding the CBDAR field has evolved. Taking photos/videos and uploading them to photo-sharing sites or one's blog have become more popular. New recognition services of scene text such as Evernote and Google Goggles are now available. Furthermore, a huge number of real images are available in Google Books and Street View. Needless to say that computer performance has been improved. Thus, it is high time to discuss and explore new research directions. This book contains refereed and improved versions of papers presented at CBDAR 2011 and is intended to give a snapshot of the state-of-the-art research in the field of camera-based document analysis and recognition.

The program of CBDAR 2011 was organized in a single-track one-day workshop. It comprised of two oral sessions and one poster session. In addition to that, two keynote talks were held by speakers from industry: Qiong Liu from FXPAL Inc. and Alessandro Bissacco from Google Inc. Finally, a panel discussion on the state of the art and new challenges was organized as the concluding session of CBDAR 2011.

After the workshop, authors of selected papers were invited to submit expanded versions of their papers for this edited volume. The authors were encouraged to include ideas and suggestions that arose in the panel discussions of the workshop. This volume is organized in three sections, reflecting the workshop session topics.

Finally, we would like to sincerely thank those who helped to make CBDAR 2011 a successful event: all paper authors, workshop attendees, Cheng-Lin Liu (ICDAR Executive Chair), Koichi Kise (ICDAR Workshop Chair) and other ICDAR organizers for their generous support, the members of the Program Committee and additional reviewers for reviewing and commenting on all of the submitted papers, and FXPAL, Google and DFKI for their financial support as sponsors of the workshop.

The 5th International Workshop on Camera-Based Document Analysis and Recognition (CBDAR 2013) is planned to be held at Washington DC, USA.

December 2011

Masakazu Iwamura
Faisal Shafait

Table of Contents

Text Detection and Recognition in Scene Images

Multi-script and Multi-oriented Text Localization from Scene Images ...	1
<i>Thotreingam Kasar and Angarai G. Ramakrishnan</i>	
Assistive Text Reading from Complex Background for Blind Persons ...	15
<i>Chucaí Yi and Yingli Tian</i>	
A Head-Mounted Device for Recognizing Text in Natural Scenes	29
<i>Carlos Merino-Gracia, Karel Lenc, and Majid Mirmehdi</i>	
Text Detection of Two Major Indian Scripts in Natural Scene Images ...	42
<i>Aruni Roy Chowdhury, Ujjwal Bhattacharya, and Swapan K. Parui</i>	
An Algorithm for Colour-Based Natural Scene Text Segmentation	58
<i>Chao Zeng, Wenjing Jia, and Xiangjian He</i>	
Recognizing Natural Scene Characters by Convolutional Neural Network and Bimodal Image Enhancement	69
<i>Yuanping Zhu, Jun Sun, and Satoshi Naoi</i>	

Camera-Based Systems

PaperUI	83
<i>Qiong Liu and Chunyuan Liao</i>	
Decapod: A Flexible, Low Cost Digitization Solution for Small and Medium Archives	101
<i>Faisal Shafait, Michael Patrick Cutter, Joost van Beusekom, Syed Saqib Bukhari, and Thomas M. Breuel</i>	
A Method for Camera-Based Interactive Whiteboard Reading	112
<i>Szilárd Vajda, Leonard Rothacker, and Gernot A. Fink</i>	
Border Noise Removal of Camera-Captured Document Images Using Page Frame Detection	126
<i>Syed Saqib Bukhari, Faisal Shafait, and Thomas M. Breuel</i>	

Datasets and Evaluation

An Image Based Performance Evaluation Method for Page Dewarping Algorithms Using SIFT Features	138
<i>Syed Saqib Bukhari, Faisal Shafait, and Thomas M. Breuel</i>	

NEOCR: A Configurable Dataset for Natural Image Text Recognition	150
<i>Robert Nagy, Anders Dicker, and Klaus Meyer-Wegener</i>	
The IUPR Dataset of Camera-Captured Document Images	164
<i>Syed Saqib Bukhari, Faisal Shafait, and Thomas M. Breuel</i>	
Author Index	173