Lecture Notes in Computer Science 8357

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison
Lancaster University, Lancaster, 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, Zürich, Switzerland
John C. Mitchell
Stanford University, Stanford, CA, USA
Moni Naor
Weizmann Institute of Science, Rehovot, Israel
Oscar Nierstrasz
University of Bern, Bern, Switzerland
C. Pandu Rangan
Indian Institute of Technology, Madras, India
Bernhard Steffen
TU Dortmund University, Dortmund, Germany
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

For further volumes:
http://www.springer.com/series/7412

Masakazu Iwamura · Faisal Shafait (Eds.)

Camera-Based Document Analysis and Recognition

5th International Workshop, CBDAR 2013
Washington, DC, USA, August 23, 2013
Revised Selected Papers

Springer

Editors
Masakazu Iwamura
Graduate School of Engineering
Osaka Prefecture University
Osaka
Japan

Faisal Shafait
The University of Western Australia
Crawley, WA
Australia

ISSN 0302-9743 ISSN 1611-3349 (electronic)
ISBN 978-3-319-05166-6 ISBN 978-3-319-05167-3 (eBook)
DOI 10.1007/978-3-319-05167-3
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014933578

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

LNCS Sublibrary: SL6 – Image Processing, Computer Vision, Pattern Recognition and Graphics

© Springer International Publishing Switzerland 2014
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, 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.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

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 has led the community to recognize document analysis and recognition of digital camera images as a promising and growing sub-field of Document Analysis and Recognition. Constraints imposed by the memory, processing speed, and image quality are leading to new interesting open problems that cannot be directly resolved by traditional techniques.

To cater for the demands of camera-based document processing, the idea of a new satellite workshop of International Conference on Document Analysis and Recognition (ICDAR) was conceived by Prof. Koichi Kise. Together with Prof. 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, South Korea. The workshop was very well received by the community and hence it was held in 2007 (Curitiba, Brazil), 2009 (Barcelona, Spain), and 2011 (Beijing, China) with the corresponding ICDAR conferences. It is our pleasure to hold the Fifth International Workshop on Camera-Based Document Analysis and Recognition (CBDAR 2013) in Washington D.C., USA, following the success of the past four workshops. The workshop is aimed to provide an opportunity to researchers and developers from various backgrounds to exchange their ideas and explore new research directions through the presentation of recent research activities and discussions.

In the eight years since the first CBDAR was held, the situation surrounding the CBDAR field has been evolving. New technologies have brought a shift in the paradigm from static camera-captured scene image reading to real-time video-based OCR using cameras on wearable devices, possibly complementing the camera input with other sensors (e.g., eye tracking). Such sensors and recent technologies have the potential to understand a user's behavior, habit, and thought, as well as improve user experience while reading.

The program of CBDAR 2013 was organized in a single-track one-day workshop. It consisted of two oral sessions and one poster session. In addition to that, a keynote talk was given by Dr. Kai Kunze from Osaka Prefecture University. Finally, a panel discussion on the state of the art and new challenges was organized as the concluding session of CBDAR 2013.

After the workshop, authors of selected contributions were invited to submit expanded versions of their papers for this edited volume. The authors were encouraged to include the ideas and suggestions that arose during the discussions at the workshop. Thus, this volume contains refereed and improved versions of papers presented at CBDAR 2013. We intend to give a snapshot of state-of-the-art research in the field of camera-based document analysis and recognition.

Finally, we would like to sincerely thank those who are helping to ensure this workshop is a success: Dr. David Doermann (ICDAR General Chair), Prof. Daniel

Lopresti (ICDAR Executive Co-chair), Prof. Apostolos Antonacopoulos (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; IAPR for its sponsorship of the workshop.

The Sixth International Workshop on Camera-Based Document Analysis and Recognition (CBDAR 2015) is planned to be held in Tunis, Tunisia.

December 2013

Masakazu Iwamura
Faisal Shafait

Contents

Text Detection and Recognition in Scene Images

Spatially Prioritized and Persistent Text Detection and Decoding 3
Hsueh-Cheng Wang, Yafim Landa, Maurice Fallon, and Seth Teller

A Hierarchical Visual Saliency Model for Character Detection in Natural Scenes 18
Renwu Gao, Faisal Shafait, Seiichi Uchida, and Yaokai Feng

A Robust Approach to Extraction of Texts from Camera Captured Images . . . 30
Sudipto Banerjee, Koustav Mullick, and Ujjwal Bhattacharya

Scene Text Detection via Integrated Discrimination of Component Appearance and Consensus 47
Qixiang Ye and David Doermann

Accuracy Improvement of Viewpoint-Free Scene Character Recognition by Rotation Angle Estimation 60
Kanta Kuramoto, Wataru Ohyama, Tetsushi Wakabayashi, and Fumitaka Kimura

Sign Detection Based Text Localization in Mobile Device Captured Scene Images 71
Jing Zhang and Rangachar Kasturi

Font Distribution Observation by Network-Based Analysis 83
Chihiro Nakamoto, Rong Huang, Sota Koizumi, Ryosuke Ishida, Yaokai Feng, and Seiichi Uchida

Camera-Based Systems

Dewarping Book Page Spreads Captured with a Mobile Phone Camera 101
Chelhwon Kim, Patrick Chiu, and Surendar Chandra

A Dataset for Quality Assessment of Camera Captured Document Images. . . . 113
Jayant Kumar, Peng Ye, and David Doermann

A Morphology-Based Border Noise Removal Method for Camera-Captured Label Images. 126
Mengyang Liu, Chongshou Li, Wenbin Zhu, and Andrew Lim

Robust Binarization of Stereo and Monocular Document Images Using Percentile Filter 139
Muhammad Zeshan Afzal, Martin Krämer, Syed Saqib Bukhari, Mohammad Reza Yousefi, Faisal Shafait, and Thomas M. Breuel

Hyperspectral Document Imaging: Challenges and Perspectives 150
Zohaib Khan, Faisal Shafait, and Ajmal Mian

Mobile Phone Camera-Based Video Scanning of Paper Documents 164
Muhammad Muzzamil Luqman, Petra Gomez-Krämer, and Jean-Marc Ogier

Real-life Activity Recognition – Focus on Recognizing Reading Activities . . . 179
Kai Kunze

Author Index 187