PROCEEDINGS
IS&T/SPIE
Electronic
Imaging
Science and Technology

Document Recognition and Retrieval XVII

Laurence Likforman-Sulem
Gady Agam
Editors

19–21 January 2010
San Jose, California, United States

Sponsored and Published by
IS&T—The Society for Imaging Science and Technology
SPIE

Cosponsored by
Institut TELECOM (France)

Volume 7534

Proceedings of SPIE, 0277-786X, v. 7534

The papers included in this volume were part of the technical conference cited on the cover and title page. Papers were selected and subject to review by the editors and conference program committee. Some conference presentations may not be available for publication. The papers published in these proceedings reflect the work and thoughts of the authors and are published herein as submitted. The publishers are not responsible for the validity of the information or for any outcomes resulting from reliance thereon.

Please use the following format to cite material from this book:
Author(s), "Title of Paper," in *Document Recognition and Retrieval XVII*, edited by Laurence Likforman-Sulem, Gady Agam, Proceedings of SPIE-IS&T Electronic Imaging, SPIE Vol. 7534, Article CID Number (2010).

ISSN 0277-786X
ISBN 9780819479273

Copublished by
SPIE
P.O. Box 10, Bellingham, Washington 98227-0010 USA
Telephone +1 360 676 3290 (Pacific Time) · Fax +1 360 647 1445
SPIE.org
and
IS&T—The Society for Imaging Science and Technology
7003 Kilworth Lane, Springfield, Virginia, 22151 USA
Telephone +1 703 642 9090 (Eastern Time) · Fax +1 703 642 9094
imaging.org

Printed in the United States of America.

Paper Numbering: Proceedings of SPIE follow an e-First publication model, with papers published first online and then in print and on CD-ROM. Papers are published as they are submitted and meet publication criteria. A unique, consistent, permanent citation identifier (CID) number is assigned to each article at the time of the first publication. Utilization of CIDs allows articles to be fully citable as soon they are published online, and connects the same identifier to all online, print, and electronic versions of the publication. SPIE uses a six-digit CID article numbering system in which:

- The first four digits correspond to the SPIE volume number.
- The last two digits indicate publication order within the volume using a Base 36 numbering system employing both numerals and letters. These two-number sets start with 00, 01, 02, 03, 04, 05, 06, 07, 08, 09, 0A, 0B … 0Z, followed by 10-1Z, 20-2Z, etc.

The CID number appears on each page of the manuscript. The complete citation is used on the first page, and an abbreviated version on subsequent pages. Numbers in the index correspond to the last two digits of the six-digit CID number.

Contents

Conference Committee

Symposium Chair

Jan P. Allebach, Purdue University (United States)

Symposium Cochair

Sabine Süsstrunk, Ecole Polytechnique Fédérale de Lausanne (Switzerland)

Conference Chairs

Laurence Likforman-Sulem, Telecom ParisTech (France)
Gady Agam, Illinois Institute of Technology (United States)

Program Committee

Apostolos Antonacopoulos, University of Salford (United Kingdom)
Elisa H. Barney Smith, Boise State University (United States)
Kathrin Berkner, Ricoh Innovations, Inc. (United States)
Xiaoqing Ding, Tsinghua University (China)
David S. Doermann, University of Maryland, College Park (United States)
Oleg D. Golubitsky, The University of Western Ontario (Canada)
Jianying Hu, IBM Thomas J. Watson Research Center (United States)
Xiaofan Lin, Vobile, Inc. (United States)
Marcus Liwicki, Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (Germany)
Daniel P. Lopresti, Lehigh University (United States)
Hiroshi Sako, Hitachi, Ltd. (Japan)
Lambert R. B. Schomaker, University of Groningen (Netherlands)
Sargur N. Srihari, University at Buffalo (United States)
Venkata Subramaniam, IBM India Research Laboratory (India)
Kazem Taghva, University of Nevada, Las Vegas (United States)
George R. Thoma, National Library of Medicine (United States)
Christian Viard-Gaudin, Université de Nantes (France)
Alessandro Vinciarelli, Idiap Research Institute (Switzerland)
Berrin Yanikoglu, Sabancı University (Turkey)
Jie Zou, National Library of Medicine (United States)

Introduction

This volume brings together the papers presented at the Document Recognition and Retrieval XVII (DRR) conference held January 2010 in San Jose as part of the IS&T/SPIE Electronic Imaging symposium. The DRR conference aims at presenting approaches which may improve the extraction of information from various types of documents: scanned, web-based, on-line, or camera-based documents. These approaches may concern the early stages of information extraction such as image enhancement, binarization, or document structure extraction. They may also concern higher stages of information extraction such as OCR, handwriting recognition, or symbol recognition. When these stages are achieved or when an electronic version is already available, information extraction and retrieval techniques can be applied. Related studies include document or writer/signature authentication. The 36 articles included in this volume deal with those topics for documents as various as historical documents, handwritten mails, maps or medical articles, and for documents written in diverse languages.

Our two invited speakers this year are Craig Knoblock from the University of Southern California and Hiroshi Sako from Hitachi Japan. Craig Knoblock will give a talk on the processing of maps and map discovery. Hiroshi Sako will speak about advanced automated teller machines (ATMs) which include intelligence for preventing different types of attacks.

This year we continue the tradition of giving the Best Student Paper Award to a paper whose lead author is a full-time student. We gratefully acknowledge Institut TELECOM and Telecom ParisTech for sponsoring this award.

We wish to thank the program committee of DRR for providing the three required reviews and the additional reviewers for assisting in the review process. Many thanks go to the SPIE conference organizers for their help in organizing the conference. At last, we thank all the participating authors of the papers for their contributions to this conference and we welcome any feedback and suggestions.

We encourage you to actively use this conference as an opportunity for discussions with your colleagues in this unique interdisciplinary framework, and we wish you a very fruitful stay in San Jose and California.

Laurence Likforman-Sulem
Gady Agam