PROCEEDINGS OF SPIE

Imaging and Printing in a Web 2.0 World II

Qian Lin
Jan P. Allebach
Zhigang Fan
Editors

26–27 January 2011
San Francisco, California, United States

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

Volume 7879

Proceedings of SPIE, 0277-786X, v. 7879

SPIE is an international society advancing an interdisciplinary approach to the science and application of light.

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 publisher is 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 *Imaging and Printing in a Web 2.0 World II*, edited by Qian Lin, Jan P. Allebach, Zhigang Fan, Proceedings of SPIE Vol. 7879 (SPIE, Bellingham, WA, 2011) Article CID Number.

ISSN 0277-786X
ISBN 9780819484161

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

Copyright © 2011, Society of Photo-Optical Instrumentation Engineers and The Society for Imaging Science and Technology.

Copying of material in this book for internal or personal use, or for the internal or personal use of specific clients, beyond the fair use provisions granted by the U.S. Copyright Law is authorized by the publishers subject to payment of copying fees. The Transactional Reporting Service base fee for this volume is $18.00 per article (or portion thereof), which should be paid directly to the Copyright Clearance Center (CCC), 222 Rosewood Drive, Danvers, MA 01923. Payment may also be made electronically through CCC Online at copyright.com. Other copying for republication, resale, advertising or promotion, or any form of systematic or multiple reproduction of any material in this book is prohibited except with permission in writing from the publisher. The CCC fee code is 0277-786X/11/$18.00.

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

vii *Conference Committee*

SESSION 1 WEB PRINTING AND PUBLISHING I

7879 02 **Web-based magazine design for self publishers** [7879-01]
A. Hunter, D. Slatter, D. Greig, Hewlett Packard Labs. (United Kingdom)

7879 03 **Improve artwork designs through data ranking system** [7879-02]
W. Wang, R. Muzzolini, Shutterfly, Inc. (United States)

7879 04 **DOM-based print-link detection for web article extraction** [7879-03]
S. Liu, S.-H. Lim, J. Liu, Hewlett-Packard Co. (United States)

SESSION 2 WEB PRINTING AND PUBLISHING II

7879 06 **A web-based troubleshooting tool to help customers self-solve color issues with a digital printing workflow** [7879-05]
H. J. Santos-Villalobos, Purdue Univ. (United States); V. Loewen, Hewlett-Packard Co. (United States); M. Lehto, J. Allebach, Purdue Univ. (United States)

7879 07 **Language-based color editing for mobile device** [7879-06]
Y. Zhao, R. Bala, K. M. Braun, Z. Langford, R. J. Rolleston, M. T. Stevens, Xerox Corp. (United States)

7879 08 **Personalized imaging: moving closer to reality** [7879-07]
H. Ding, Purdue Univ. (United States); R. Bala, Z. Fan, Xerox Corp. (United States); C. A. Bouman, J. P. Allebach, Purdue Univ. (United States)

SESSION 3 WEB DESIGN AND LAYOUT ANALYSIS

7879 09 **Document similarity measures and document browsing** [7879-08]
I. Ahmadullin, Purdue Univ. (United States); J. Fan, N. Damera-Venkata, S. H. Lim, Q. Lin, J. Liu, S. Liu, E. O'Brien-Strain, Hewlett-Packard Labs. (United States); J. Allebach, Purdue Univ. (United States)

7879 0A **Adaptive removal of background and white space from document images using seam categorization** [7879-09]
C. Fillion, Z. Fan, Xerox Corp. (United States); V. Monga, The Pennsylvania State Univ. (United States)

7879 0B **Aesthetic role of transparency and layering in the creation of photo layouts** [7879-10]
M. V. Ortiz Segovia, Purdue Univ. (United States); N. Damera-Venkata, E. O'Brien-Strain, J. Fan, S. H. Lim, S. Liu, J. Liu, Q. Lin, Hewlett-Packard Labs. (United States); J. P. Allebach, Purdue Univ. (United States)

7879 0C **Automatic picture orientation detection based on classifier combination** [7879-11]
Y. Sun, C. Liu, X. Ding, Tsinghua Univ. (China); Z. Fan, F. Tse, Xerox Corp. (United States)

SESSION 4 DIGITAL LIBRARIES/INFORMATION SHARING

7879 0D **Whiteboard sharing: capture, process, and print or email** [7879-12]
M. Gormish, B. Erol, D. G. Van Olst, T. Li, A. Mariotti, Ricoh Innovations, Inc. (United States)

7879 0E **Building a print on demand web service** [7879-13]
P. Reddy, Hewlett-Packard Labs. (United States); B. Rozario, S. Dudekula, A. D. V, Hewlett-Packard Labs. India (India)

7879 0G **iULib: where UDL and Wikipedia could meet** [7879-15]
Y. Tian, T. Huang, W. Gao, Peking Univ. (China)

SESSION 5 WEB CONTENT ANALYSIS AND UNDERSTANDING I

7879 0H **Book Widget: embedding automated photo-document publication on the Web and in mobile devices** [7879-04]
E. O'Brien-Strain, Hewlett-Packard Labs. (United States); A. Hunter, Hewlett-Packard Labs. (United Kingdom); J. Liu, Q. Lin, D. Tretter, Hewlett-Packard Labs. (United States); J. Wang, Hewlett-Packard Labs. China (China); X. Zhang, P. Wu, Hewlett-Packard Labs. (United States)

7879 0I **Semantic photo books: leveraging blogs and social media for photo book creation** [7879-17]
M. Rabbath, P. Sandhaus, OFFIS e.V. (Germany); S. Boll, Univ. of Oldenburg (Germany)

7879 0J **Automatic image selection scheme utilizing comments for insertion of images into weblogs** [7879-18]
T. Konno, E. Myodo, K. Takagi, R. Kawada, KDDI R&D Labs., Inc. (Japan)

7879 0K **Title identification of web article pages using HTML and visual features** [7879-19]
J. Fan, Hewlett-Packard Labs. (United States); P. Luo, Hewlett-Packard Labs. (China); P. Joshi, Hewlett-Packard Labs. (United States)

7879 0L **Creating 3D realistic head: from two orthogonal photos to multiview face contents** [7879-20]
Y. Lin, Tsinghua Univ. (China); Q. Lin, F. Tang, L. Tang, S. Lim, Hewlett-Packard Labs. (United States); S. Wang, Tsinghua Univ. (China)

SESSION 6 WEB CONTENT ANALYSIS AND UNDERSTANDING II

7879 0M **Mobile multimedia understanding applications: an overview (Invited Paper)** [7879-21]
X. Lin, Vobile, Inc. (United States)

7879 0N **Learning object detectors from online image search** [7879-22]
F. Tang, D. R. Tretter, Hewlett-Packard Labs. (United States)

7879 0O **Image categorization for marketing purposes** [7879-23]
M. I. Almishari, Univ. of California, Irvine (United States); H. Lee, N. Gnanasambandam, Xerox Corp. (United States)

7879 0P **Text extraction from web images** [7879-24]
C. Liu, C. Yang, X. Ding, Tsinghua Univ. (China); J. Fan, Hewlett-Packard Labs. (United States)

7879 0Q **Web image annotation using two-step filtering on social tags** [7879-25]
S. Cho, J. Cha, H. Byun, Yonsei Univ. (Korea, Republic of)

Author Index

Conference Committee

Symposium Chair

Sabine Süsstrunk, École Polytechnique Fédérale de Lausanne (Switzerland)

Symposium Cochair

Majid Rabbani, Eastman Kodak Company (United States)

Conference Chairs

Qian Lin, Hewlett-Packard Laboratories (United States)
Jan P. Allebach, Purdue University (United States)
Zhigang Fan, Xerox Corporation (United States)

Program Committee

Patricia Albanese, Rochester Institute of Technology (United States)
Kathrin Berkner, Ricoh Innovations, Inc. (United States)
Susanne C. J. Boll, Carl von Ossietzky Universität Oldenburg (Germany)
Guotong Feng, Ricoh Innovations, Inc. (United States)
Jerry J. Liu, Hewlett-Packard Laboratories (United States)
Jiebo Luo, Eastman Kodak Company (United States)
Robert J. Rolleston, Xerox Corporation (United States)
David N. Slatter, Hewlett-Packard Laboratories (United Kingdom)
Yonghong Tian, Beijing University (China)
Wiley H. Wang, Shutterfly, Inc. (United States)
Shengjin Wang, Tsinghua University (China)

Session Chairs

1 Web Printing and Publishing I
Qian Lin, Hewlett-Packard Laboratories (United States)

2 Web Printing and Publishing II
Kathrin Berkner, Ricoh Innovations, Inc. (United States)

3 Web Design and Layout Analysis
Jerry J. Liu, Hewlett-Packard Laboratories (United States)

4 Digital Libraries/Information Sharing
Wiley H. Wang, Shutterfly, Inc. (United States)

5 Web Content Analysis and Understanding I
Zhigang Fan, Xerox Corporation (United States)

6 Web Content Analysis and Understanding II
Jan P. Allebach, Purdue University (United States)