Document Recognition and Retrieval XXI
Wednesday - Thursday 5 - 6 February 2014

Conference Sessions
At A Glance

<table>
<thead>
<tr>
<th>Session</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plenary Session and Society Award Presentations</td>
<td></td>
</tr>
<tr>
<td>Symposium Demonstration Session</td>
<td></td>
</tr>
<tr>
<td>Plenary Session and Conference Award Presentations</td>
<td></td>
</tr>
<tr>
<td>Interactive Paper Session-Wednesday</td>
<td></td>
</tr>
<tr>
<td>1: Handwriting</td>
<td></td>
</tr>
<tr>
<td>2: Form Classification</td>
<td></td>
</tr>
<tr>
<td>3: Invited Presentation I</td>
<td></td>
</tr>
<tr>
<td>4: Text Recognition</td>
<td></td>
</tr>
<tr>
<td>5: Handwritten Text Line Segmentation</td>
<td></td>
</tr>
<tr>
<td>6: Invited Presentation II</td>
<td></td>
</tr>
<tr>
<td>7: Layout Analysis</td>
<td></td>
</tr>
<tr>
<td>8: Information Retrieval</td>
<td></td>
</tr>
<tr>
<td>9: Data Sets and Ground-Truthing</td>
<td></td>
</tr>
<tr>
<td>Panel Discussion: Data Sets and Ground-Truthing</td>
<td></td>
</tr>
</tbody>
</table>

Important Dates

<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract Due:</td>
<td>22 July 2013</td>
</tr>
<tr>
<td>Author Notification:</td>
<td>7 October 2013</td>
</tr>
<tr>
<td>Manuscript Due Date:</td>
<td>25 November 2013</td>
</tr>
</tbody>
</table>

Conference Cosponsors

Sponsored by:

Google

Conference Committee

Conference Chairs
Bertrand Couasnon, Institut National des Sciences Appliquées de Rennes (France)
Eric K. Ringger, Brigham Young Univ. (United States)

Program Committee
Gady Agam, Illinois Institute of Technology (United States)
Sameer K. Antani, National Library of Medicine (United States)
Elisa H. Barney Smith, Boise State Univ. (United States)
Kathrin Berkner, Ricoh Innovations, Inc. (United States)
Hervé Déjean, Xerox Research Ctr. Europe Grenoble (France)
Xiaofan Lin, Tsinghua Univ. (China)
David S. Doermann, Univ. of Maryland, College Park (United States)
Oleg D. Golubitsky, Google Waterloo (Canada)
Jianying Hu, IBM Thomas J. Watson Research Ctr. (United States)
Ergina Kavallieratou, Univ. of the Aegean (Greece)
Christopher Kermorvant, A2iA SA (France)

Program Committee continued...
Laurence Likforman-Sulem, Telecom ParisTech (France)
Xiaofan Lin, A9.com, Inc. (United States)
Marcus Liwicki, Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (Germany)
Daniel P. Lopresti, Lehigh Univ. (United States)
Umapada Pal, Indian Statistical Institute (India)
Sarfarz N. Srivani, Univ. at Buffalo (United States)
Venkata Subramaniam, IBM India Research Lab. (India)
Kazem Taghva, Univ. of Nevada, Las Vegas (United States)
George R. Thomas, National Library of Medicine (United States)
Christian Viard-Gaudin, Univ. de Nantes (France)
Berrin Yanikoglu, Sabanci Univ. (Turkey)
Richard Zanibbi, Rochester Institute of Technology (United States)
Jie Zou, National Library of Medicine (United States)

TUESDAY 4 FEBRUARY

Session Plen1:
Plenary Session and Society Award Presentations

Tuesday 4 February 2014
8:30 AM - 9:50 AM
Location: Continental Ballroom 5
Using fMRI To Reverse Engineer the Human Visual System *(Plenary Presentation)*  
Paper 9011-301  
Time: 8:30 AM - 9:15 AM  
Author(s): Jack L. Gallant, Univ. of California, Berkeley (United States)  
Add To My Schedule

Symposium Demonstration Session

Tuesday 4 February 2014  
5:30 PM - 7:30 PM  
Location: Continental Ballroom 4

A symposium-wide demonstration session will be open to attendees 5:30 to 7:30 pm Tuesday evening. Demonstrators will provide interactive, hands-on demonstrations of a wide-range of products related to Electronic Imaging.

WEDNESDAY 5 FEBRUARY

Session Plen2:  
Plenary Session and Conference Award Presentations

Wednesday 5 February 2014  
8:30 AM - 9:50 AM  
Location: Continental Ballroom 5

Integrated Imaging: Creating Images from the Tight Integration of Algorithms, Computation, and Sensors *(Plenary Presentation)*  
Paper 9011-300  
Time: 8:30 AM - 9:15 AM  
Author(s): Charles A. Bouman, Purdue Univ. (United States)  
Add To My Schedule

Session PWed:  
Interactive Paper Session-Wednesday

Wednesday 5 February 2014  
9:50 AM - 11:10 AM  
Location: Continental Ballroom and Golden Gate Level Foyers

Interactive papers will be placed on display after 8:30 am on Wednesday. An interactive paper session, with authors present at their papers, will be held Wednesday morning, 9:50 to 11:10 am. Refreshments will be served.

Two-stage approach to keyword spotting in handwritten documents  
Paper 9021-24  
Time: 9:50 AM - 11:10 AM  
Author(s): Mehdi Haji, IMDS Software (Canada), Concordia Univ. (Canada); Mohammad R. Ameri, Concordia University (Canada); Tien D. Bui, Ching Y. Suen, Concordia Univ. (Canada); Dominique Ponson, IMDS Software (Canada)  
Add To My Schedule

Extraction and labeling high-resolution images from PDF documents  
Paper 9021-25  
Time: 9:50 AM - 11:10 AM  
Author(s): Suchet K. Chachra, Zhiyun Xue, Sameer Antani, Dina Demner-Fushman, George R. Thoma, National Library of Medicine (United States)  
Add To My Schedule

Structure analysis for plane geometry figures  
Paper 9021-26  
Time: 9:50 AM - 11:10 AM  
Author(s): Tianxiao Feng, Xiaojing Lu, Lu Liu, Keqiang Li, Zhi Tang, Peking Univ. (China)  
Add To My Schedule

On-line signature verification method by Laplacian spectral analysis and dynamic time warping  
Paper 9021-27
Session 1:
Handwriting

Wednesday 5 February 2014
11:10 AM - 12:30 PM
Location: Golden Gate 8
Session Chair: Eric K. Ringger, Brigham Young Univ. (United States)

Writer identification on historical Glagolitic documents
Paper 9021-1
Time: 11:10 AM - 11:30 AM
Author(s): Stefan Fiel, Fabian Hollaus, Melanie Gau, Robert Sablatnig, Technische Univ. Wien (Austria)

Probabilistic modeling of children's handwriting
Paper 9021-2
Time: 11:30 AM - 11:50 AM
Author(s): Mukta Puri, Sargur N. Srihari, Univ. at Buffalo (United States); Lisa Hanson, Bureau of Criminal Apprehension Laboratory (United States)

Variational dynamic background model for keyword spotting in handwritten documents
Paper 9021-3
Time: 11:50 AM - 12:10 PM
Author(s): Gaurav Kumar, Univ. at Buffalo (United States); Safwan Wshah, Xerox Corp. (United States); Venu Govindaraju, Univ. at Buffalo (United States)

Boosting bonsai trees for handwritten/printed text discrimination
Paper 9021-4
Time: 12:10 PM - 12:30 PM
Author(s): Yann Ricquebourg, Christian Raymond, Baptiste Poirriez, IRISA / INRIA de Rennes (France); Aurélie Lemaitre, IRISA, Univ. de Rennes 2 (France); Bertrand Couăsnor, IRISA / INRIA de Rennes (France)

Lunch Break 12:30 PM - 2:10 PM

Session 2:
Form Classification
Wednesday 5 February 2014
2:10 PM - 2:50 PM
Location: Golden Gate 8

Session Chair: Gady Agam, Illinois Institute of Technology (United States)

Form similarity via Levenshtein distance between ortho-filtered logarithmic ruling-gap ratios
Paper 9021-5
Time: 2:10 PM - 2:30 PM
Author(s): George Nagy, Rensselaer Polytechnic Institute (United States); Daniel P. Lopresti, Lehigh Univ. (United States)

Form classification and retrieval using bag of words with shape features of line structures
Paper 9021-6
Time: 2:30 PM - 2:50 PM
Author(s): Florian Kleber, Markus Diem, Robert Sablatnig, Technische Univ. Wien (Austria)

Session 3:
Invited Presentation I

Wednesday 5 February 2014
2:50 PM - 3:40 PM
Location: Golden Gate 8

Session Chair: Bertrand Coüasnon, Institut National des Sciences Appliquées de Rennes (France)

OCR for Google Books (Invited Paper)
Paper 9021-7
Time: 2:50 PM - 3:40 PM
Author(s): Ray Smith, Ashok Popat, Google (United States)

Coffee Break 3:40 PM - 4:10 PM

Session 4:
Text Recognition

Wednesday 5 February 2014
4:10 PM - 5:10 PM
Location: Golden Gate 8

Session Chair: Sameer Antani, National Library of Medicine (United States)

Utilizing web data in identification and correction of OCR errors
Paper 9021-8
Time: 4:10 PM - 4:30 PM
Author(s): Kazem Taghva, Shivam Agarwal, Univ. of Nevada, Las Vegas (United States)

How well does multiple OCR error correction generalize?
Paper 9021-9
Time: 4:30 PM - 4:50 PM
Author(s): William B. Lund, Eric K. Ringger, Brigham Young Univ. (United States); Daniel D. Walker, Microsoft Corp. (United States)

Video text localization using wavelet and shearlet transforms
Paper 9021-10
Time: 4:50 PM - 5:10 PM
Author(s): Purnendu Banerjee, Bidyut B. Chaudhuri, Indian Statistical Institute (India)

THURSDAY 6 FEBRUARY

Show All Abstracts
Session 5:
Handwritten Text Line Segmentation

Thursday 6 February 2014
8:50 AM - 9:30 AM
Location: Golden Gate 8

Session Chair: Elisa H. Barney Smith, Boise State Univ. (United States)

A Markov chain based line segmentation framework for handwritten character recognition
Paper 9021-11
Time: 8:50 AM - 9:10 AM
Author(s): Yue Wu, Shengxin Zha, Huagu Cao, Daben Liu, Raytheon BBN Technologies (United States); Premkumar Natarajan, The Univ. of Southern California (United States)

Add To My Schedule

Handwritten text segmentation using blurred image
Paper 9021-12
Time: 9:10 AM - 9:30 AM
Author(s): Aurélie Lemaitre, IRISA, Univ. Rennes 2 (France); Jean Camillerapp, IRISA / INRIA Rennes (France); Bertrand Coüasnon, IRISA / INRIA de Rennes (France)

Add To My Schedule

Session 6:
Invited Presentation II

Thursday 6 February 2014
9:30 AM - 10:20 AM
Location: Golden Gate 8

Session Chair: Eric K. Ringger, Brigham Young Univ. (United States)

What makes big visual data hard? (Invited Paper)
Paper 9021-13
Time: 9:30 AM - 10:20 AM
Author(s): Alexei (Alyosha) Efros, Univ. of California, Berkeley (United States)

Add To My Schedule

Coffee Break 10:20 AM - 10:50 AM

Session 7:
Layout Analysis

Thursday 6 February 2014
10:50 AM - 12:30 PM
Location: Golden Gate 8

Session Chair: Daniel P. Lopresti, Lehigh Univ. (United States)

Optical music recognition on the International Music Score Library Project
Paper 9021-14
Time: 10:50 AM - 11:10 AM
Author(s): Christopher S. Raphael, Rong Jin, Indiana Univ. (United States)

Add To My Schedule

Document flow segmentation for business applications
Paper 9021-15
Time: 11:10 AM - 11:30 AM
Author(s): Daher Hani, Abdel Belaïd, LORIA (France), Univ. de Lorraine (France)

Add To My Schedule

LearnPos: a new tool for interactive learning positioning
Session 8:
Information Retrieval

Thursday 6 February 2014
2:10 PM - 3:30 PM
Location: Golden Gate 8
Session Chair: Xiaofan Lin, A9.com, Inc. (United States)

Document page structure learning for fixed-layout e-books using conditional random fields
Paper 9021-16
Time: 11:30 AM - 11:50 AM
Author(s): Cérès Carton, IRISA / INRIA Rennes (France); Aurélie Lemaître, IRISA / Univ. Rennes 2 (France); Bertrand Coúasnon, IRISA / INRIA de Rennes (France)

Automatic comic page image understanding based on edge segment analysis
Paper 9021-17
Time: 12:10 PM - 12:30 PM
Author(s): Xin Tao, Zhi Tang, Canhui Xu, Peking Univ. (China)

Lunch Break 12:30 PM - 2:10 PM

Session 9:
Data Sets and Ground-Truthing

Thursday 6 February 2014
4:00 PM - 4:20 PM
Location: Golden Gate 8

Scalable ranked retrieval using document images
Paper 9021-19
Time: 2:10 PM - 2:30 PM
Author(s): Rajiv Jain, Douglas Oard, David Doermann, Univ. of Maryland, College Park (United States)

A contour-based shape descriptor for biomedical image classification and retrieval
Paper 9021-20
Time: 2:30 PM - 2:50 PM
Author(s): Daekeun You, Sameer Antani, Dina Demner-Fushman, George R. Thoma, National Library of Medicine (United States)

Semi-automated document image clustering and retrieval
Paper 9021-21
Time: 2:50 PM - 3:10 PM
Author(s): Markus Diem, Florian Kleber, Stefan Fiel, Robert Sablatnig, Technische Univ. Wien (Austria)

Fast structural matching for document image retrieval through spatial databases
Paper 9021-22
Time: 3:10 PM - 3:30 PM
Author(s): Hongxing Gao, Maçal Rusiñol, Dimosthenis Karatzas, Josep Lladós, Univ. Autònoma de Barcelona (Spain)
Session Chairs: Bertrand Couasnon, Institut National des Sciences Appliquées de Rennes (France); Eric K. Ringger, Brigham Young Univ. (United States)

The Lehigh Steel Collection: a new open dataset for document recognition research
Paper 9021-23
Time: 4:00 PM - 4:20 PM
Author(s): Barri Bruno, Daniel P. Lopresti, Lehigh Univ. (United States)

Add To My Schedule

Session Panel:
Panel Discussion: Data Sets and Ground-Truthing

Thursday 6 February 2014
4:20 PM - 5:00 PM
Location: Golden Gate 8