



Call for Papers

Document Recognition and Retrieval XIX

Part of the IS&T/SPIE 24th Annual Symposium on Electronic Imaging 2012 22-26 January 2012 • San Francisco, CA USA

Conference Chairs: **Christian Viard-Gaudin**, Univ. of Nantes (France); **Richard Zanibbi**, Rochester Institute of Technology (USA)

Program Committee: **Gady Agam**, Illinois Institute of Technology; **Elisa H. Barney Smith**, Boise State Univ.; **Bill Barrett**, Brigham Young Univ.; **Kathrin Berkner**, Ricoh Innovations, Inc.; **Bertrand Couasnon**, IRISA/INSA Rennes (France), **Hervé Dejean**, Xerox Research (France), **Xiaoqing Ding**, Tsinghua Univ. (China); **David S. Doermann**, Univ. of Maryland/College Park; **Oleg Golubitsky**, Univ. of Western Ontario (Canada); **Jianying Hu**, IBM Thomas J. Watson Research Ctr.; **Laurence Likforman-Sulem**, Telecom ParisTech (France); **Marcus Liwicki**, DFKI (Germany); **Xiaofan Lin**, Vobile Inc.; **Daniel P. Lopresti**, Lehigh Univ.; **Hiroshi Sako**, Hitachi, Ltd. (Japan); **Sargur N. Srihari**, Univ. at Buffalo; **Venkata Subramaniam**, IBM India Research Lab. (India); **Kazem Taghva**, Univ. of Nevada/Las Vegas; **George R. Thoma**, National Library of Medicine; **Alessandro Vinciarelli**, University of Glasgow (United Kingdom); **Berrin Yanikoglu**, Sabanci Univ. (Turkey); **Jie Zou**, National Library of Medicine.

We are pleased to announce the 19th Document Recognition and Retrieval Conference (DRR), to be held on 22-26 January 2012, in San Francisco, CA, USA. DRR is an international conference focused on state-of-the-art research in document recognition and retrieval, for offline, online and web documents. The conference is part of the *Electronic Imaging Symposium*, which brings together researchers from various backgrounds related to electronic imaging for an exciting research event. The conference will include oral/poster presentations, and invited papers. Accepted papers will be published in DRR Proceedings. For the sixth year, the **Best Student Paper** will be selected among papers whose first authors are full-time students. Additional details and updated information of this conference can be found at <http://drr2012.irccyn.ec-nantes.fr/>

We are soliciting papers describing algorithms and systems in all aspects of document recognition and retrieval, for offline, online and Web documents. Specifically, we are encouraging papers enlarging the frontiers and dealing for instance with multi-modal information such as online and offline, or speech and handwriting, etc.

Recognizing handwritten or degraded machine printed documents (e.g. faxed or historical documents) remains a challenging problem. Document recognition and understanding are concerned with fully reconstructing a document in electronic form consistent with the original format (fonts, layout etc.) and capturing the underlying logical structure (e.g. section headings, region types, references, and reading order) of the document. Among the challenges for machine-printed documents are complex layouts (text written on images, complex backgrounds, etc.), degraded and noisy documents, and robust recognition of tables and equations. Handwritten documents with unconstrained writing style pose additional challenges due to increased variability and segmentation ambiguities. Non-textual elements in documents form another class of interesting problems. These include the extraction and recognition of logos and signatures, and the conversion of line drawings in documents from raster to vector format. Web documents pose both similar and new challenges.

One of the primary reasons for digitizing existing paper materials is to simplify the retrieval and organization of information. In this regard we are particularly interested in papers which address any of the following issues: retrieval in the presence of noise; retrieval based on sketches, images, tables, diagrams or other non-linguistic objects that appear in the document; retrieval based on text appearing with non-standard alignment, in images or graphics; recognition and tagging of mathematical arrays and equations which serve as indicators of subject content or methodology used in the document; novel methods for retrieval and organization of information based on text or other information in a document.

Papers are solicited in, but not limited to, the following areas:

Document Recognition

- Text recognition: machine-printed, handwritten documents; paper, tablet, camera, video sources
- Writer/style identification, verification, adaptation
- Graphics recognition: vectorization (e.g. for line-art, maps and technical drawings), signature, logo and graphical symbol recognition, figure, chart and graph recognition, diagrammatic notations (e.g. music, mathematical notation)
- Document layout analysis and understanding: document and page region segmentation, form and table recognition, document understanding through combined modalities (e.g. speech and images)
- Evaluation: performance metrics, document degradation models
- Additional topics: document image filtering, enhancement and compression, document clustering and classification, machine learning (e.g. integration and optimization of recognition modules), historical and degraded document images (e.g. fax), multilingual document recognition, web page analysis (including wikis and blogs)

Document Retrieval

- Indexing and Summarization : (noisy) text documents (messages, blogs, etc.), imaged documents, entity tagging from OCR'ed text, text categorization
- Query Languages and Modalities: Content-Based Image Retrieval (CBIR) for documents, keyword spotting, non-textual query-by-example (e.g. tables, figures, math), querying by document geometry and/or logical structure, approximate string matching algorithms for OCR'ed text, retrieval of noisy text documents (melbW•xkGLGMGVWLTMBxkMPLVLTbTbBbDTiB