Call for Papers and Announcement

Document Recognition and Retrieval XIV

Part of the IS&T/SPIE International Symposium on Electronic Imaging

28 January – 1 February 2007 • San Jose Marriott and San Jose Convention Center •

San Jose, California, USA

Conference Chairs: **Xiaofan Lin**, Hewlett-Packard Labs; **Berrin Yanikoglu**, Sabanci Univ. (Turkey)

Program Committee: **Tim L. Andersen**, Boise State Univ.; **Apostolos Antonacopoulos**, Univ. of Salford (United Kingdom); **Elisa H. Barney Smith**, Boise State Univ.; **Brian D. Davison**, Lehigh Univ.; **Xiaoqing Ding**, Tsinghua Univ. (China); **David S. Doermann**, Univ. of Maryland/College Park; **Jianying Hu**, IBM Thomas J. Watson Research Ctr.; **Matthew F. Hurst**, Intelliseek, Inc.; **Hisashi Ikeda**, Hitachi, Ltd. (Japan); **Tapas Kanungo**, IBM Almaden Research Ctr.; **Daniel P. Lopresti**, Lehigh Univ.; **Thomas A. Nartker**, Univ. of Nevada/Las Vegas; **Sargur N. Srihari**, SUNY/Univ. at Buffalo; **Kazem Taghva**, Univ. of Nevada/Las Vegas**; George R. Thoma**, National Library of Medicine

We are pleased to announce the 14th Document Recognition and Retrieval Conference (DRR), to be held 28 January-1 Feb. 2007, in San Jose, CA, USA. DRR is an international conference for 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. New for this year, **the Best Student Paper will be selected among papers whose lead authors are full-time students.** Additional details and updated information of this conference will be announced at http://www.drr2007.org.

Recognizing handwritten or degraded machine print documents (e.g. faxed and old/historical documents) remains as a challenging problem. Beyond OCR, document recognition includes the recovery of a document's logical structure and format. With successful layout analysis and recognition, document recognition aims to fully reconstruct a document in electronic form, in its original format (fonts, layout etc.). Among the remaining challenges for machine-printed documents are complex layouts (text written on images, complex backgrounds, etc.) and robust recognition of tables and equations, while handwritten documents written with unconstrained writing style pose a challenge. Furthermore, converting line drawings in a document from raster to vector format, creating graphical objects endowed with semantic meaning, is another goal of document recognition. Documents with online handwriting (where the image is accompanied with temporal information, as in Tablet PCs) and Web documents pose both similar and new challenges, as two "new" classes of documents. We are soliciting papers describing algorithms and systems in all aspects of document recognition, for offline, online and Web documents.

Since the primary reason for digitizing existing paper materials is to simplify retrieval and organization of information, we are particularly interested in papers which address any of the following issues: (1) retrieval in the face of corrupted readings of the terms in a document; (2) retrieval based on sketches, images, tables, diagrams or other non-linguistic objects that appear in the document; (3) retrieval based on text appearing with non-standard alignment, in images or graphics; (4) recognition and tagging of mathematical arrays and equations which serve as indicators of subject content or methodology used in the document; (5) novel methods for retrieval and organization of information based on text or other information in a document. Papers addressing retrieval-specific issues are encouraged to use a standard methodology from either statistics (such as the ROC representation) or IR (such as precision versus recall) to assess the effectiveness

of proposed techniques against the endpoint goal of correct recognition and retrieval of the entire document, or a section thereof.

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

Document Recognition

- Machine-print and handwritten text recognition (degraded documents such as faxed or old/historical documents, multilingual documents, etc.)
- Web document recognition and analysis
- Video- and camera-based OCR
- Algorithms, systems, and quality assurance methods towards large-scale digital libraries
- Identification and recognition of tables or equations
- Graphics recognition (for line-art, maps, and technical drawings)
- Document segmentation and layout analysis
- Filtering, enhancement, and compression techniques for document images
- Document degradation models
- Document analysis and synthesis for digital publishing (template reuse and layout generation for new contents)

Document Retrieval

- Recovery and use of logical structure for retrieval
- Information extraction from forms
- Relevance feedback techniques for document retrieval
- Cross-language and multi-lingual retrieval
- Summarization of text documents and imaged documents
- Keyword spotting in document images
- Approximate string matching algorithms for OCR'ed text
- Non-textual retrieval and search, including image and multimedia interfaces for retrieval
- Benchmarking and evaluation issues
- Impact of recognition accuracy on retrieval effectiveness
- Techniques to support spoken language access to document text (audio browsing of document databases)

Note: submissions to Document Recognition and Retrieval XIV should be abbreviated papers (5-7 pages). The paper should informative and address the following questions: i) What is the paper about? ii) What is the original contribution? iii) What is the most closely related work by others and how does this work differs? iv) How can others make use of this work? v) What are the main experimental/theoretical results? If you are qualified and would like to compete for the Best Student Paper, please indicate in the abbreviated paper. Full papers (10-12 pages) will be needed for the final proceedings. Please contact Xiaofan Lin (xiaofan.lin@hp.com) or Berrin Yanikoglu (berrin@sabanciuniv.edu) for questions related to the conference.

Important dates:

17 July 2006	Abstracts due http://www.electronicimaging.org
August 2006	Acceptance notice
6 November 2006	Manuscripts due for volumes available onsite
20 November 2006	Final Summaries to SPIE
28 January—1 February 2007	Symposium Dates (San Jose, California)