CONFERENCE REPORT: ICDAR'07

George Nagy
DocLab
Electrical, Computer, and Systems Engineering
Rensselaer Polytechnic Institute
Troy, NY 12180 USA

Summary

The Ninth Conference took place August 27-29, 2007 in the rarefied atmosphere of the Shangri-La Hotel in Katmandu. It attracted 2740 participants, including 2480 teleattendants. The program committee selected 185 papers from 640 submissions. Thirty were presented live on-site. The three invited speakers gave cautiously optimistic assessments of the direction and momentum of the discipline. The president of the IAPR presented several awards for distinguished service or contributions, but one of the customary prizes was again withheld for lack of a meritorious nomination.

The thirty-two sessions were grouped according to nine major themes: Text Recognition, Graphics, Devices, Data Mining and Warehousing, Sublimation, Oral Documents, Docunets, Epistechnology, and History of OCR/DIA. The significant paradigm shifts of the last decade are reflected in the distribution of topics. For example, Roman print recognition was the subject of only a single OCR session (perhaps as a consequence of the conference venue), but there were eight sessions on Han and Sanskrit scripts.

Audio recording displaced handwriting and electronic ink as the preferred mode of personal data entry (partly, no doubt, because of sweeping worldwide changes in elementary education policies). The confluence of speech, document, and information retrieval technologies was even more evident than at previous conferences.

Most of the research on graphics recognition has shifted to the translation of maps, drawings and schematics originally produced with recently-deplatformed CAD tools. All current stratigraphic model-entry systems are integrated with IGES or STEP databases and viewers.

New organic semiconductors and polymer-dispersed LCDs are fostering the development of convenient, high-capacity devices for electronic document storage (e.g., the Docupin), conversion (Multipad), and display (Radiotablet and E-scap). Several authors presented data compression methods based on GLZE (Generalized Lempel-Ziv Encoding), which is finally becoming practical on Gigabit ATM networks. Defensive information filtering techniques are assuming a dominant role as a result of the growth of "push" technologies.

Almost all the theoretical contributions were focused on unsupervised learning, with heavy emphasis on task-specific adaptation bootstrapped from a UFL (Universal Format Lexicon). Concern about generalization has given way to improving specialization. A unified graph-probabilistic framework was presented for decision trees, object-oriented ontologies, BBNs, ANNs and HMMs.

A dozen noteworthy papers, most by established researchers, are singled out for attention. Many of the selected communications are the culmination of focused research of many years' duration. Also reviewed are participant reports on several large-scale projects. Some are efforts to reduce document pollution, particularly in the government sector and on global servers. Others can be traced to the Digital Library projects initiated in the last century, when the emphasis was just the opposite.

The program committee deserves credit for organizing the first-ever ICDAR session on the history of OCR and DIA, which befits our status as a mature discipline. One telepresentation correlated the increase in the annual harvest of OCR/DIA papers (from 75 to 1230 in the last twenty-five years) with the decrease in the median annual citation frequency (currently 0.32), and with average content-overlap between papers sharing an author (estimated with standard discourse-analysis tools at 0.68). Statistical analysis also reveals a surprising 13.4-year random-phase periodicity in OCR/DIA publications.

The presentation will conclude with speculation about the advances we can expect, extrapolated from the accomplishments reported here, at ICDAR'17.

0-8186-7898-4/97 $10.00 © 1997 IEEE