

Newsletter 07

Hello everybody,

Well, ICDAR'93 is over and I first will use the opportunity to thank the organizers once again. The conference was interesting and well organized and I think we all appreciated our stay in Tsukuba. So congratulations to our Japanese friends for a successful conference.

Many things were discussed in Tsukuba: scientific matters, organizational matters and friendly matters (conferences are also a good place to get acquainted with the real people behind the names in the bibliography lists;-). I am sure that these different matters will come up again in the coming months, also in this mailing list. By the way, many of you in this mailing list will receive this 7th issue of TC10 news as the first you get, as one consequence of ICDAR was that I extended my mailing list !!! So welcome to the club, as we say here in France !

But I am not going to write a report here about ICDAR'93 (I haven't enough time for that just now...). I will just mention some of the information which was given there, for the benefit of those who couldn't attend.

ICDAR'95 was announced (the call for papers was available in Tsukuba). It will be in Montreal, August 14-16 1995. It is noteworthy to know that IJCAI will be in Montreal the week after, so this is a good opportunity to combine two interesting conferences. The call for papers for ICDAR'95 will probably be available in electronic format; I will put it in this mailing list as soon as I get it. Note that paper submission deadline is December 1st, 1994, so we have time yet!

There was also a meeting to decide where ICDAR'97 should be. There were several candidates for the organization, and after a vote among all attendees of the meeting, the result was that ICDAR'97 will be held in Ulm, Germany.

We also had a TC10 meeting to discuss the planned workshop on graphics recognition. Here are the minutes of this meeting.

TC10 meeting, October 21st, Tsukuba.

Present: L. Cordella, D. Dori, O. Hori, R. Kasturi, J. Kreich, G. Monagan, B. Pasternak, K. Tombre, K. Tomiyama, K. Yamamoto

Karl Tombre presents a draft of the call for papers for the TC10 workshop. This workshop is planned to be held just before ICDAR'95. R. Kasturi is willing to host the workshop at Penn State University.

Remarks were made by several people that this workshop must not drain away good papers from the main ICDAR conference. A discussion follows on the best way to organize the workshop. A consensus is reached not to call for plain papers but rather to define a number of panels on specific topics. For each of these topics, there could be 1 or 2 main paper presentations giving the state of the art and stating the open questions for the topic, followed by a number of shorter presentations where each participant is invited to contribute by proposing solutions to some of the questions or presenting results of his/her work. Each session would be concluded by a panel discussion. Therefore, the topics should first be defined (and every proposal is welcome); a call for contribution by short abstracts of remarks would then follow. We plan also time for working groups to discuss in detail specific problems.

Each session should last at least 2 hours. All in all, it is decided that a 2-days workshop is enough and the dates of August 10-11 are considered to be good. The workshop should be limited to about 75 persons.

The program committee will modify the call-for-papers draft to reflect this consensus; then the call for papers will be sent out. Discussions are being held with IAPR authorities for the workshop to be an official IAPR event.

So you will see the call for papers soon in this mailing list, I think.

Just before leaving for ICDAR, I got the following information by mail from R.M. Haralick; this can also be of interest to many of you:

RMH> From haralick@horus.ee.washington.edu Tue Oct 12 09:38:34 1993
RMH> Message-Id: <9310120836.AA03777@horus.wtc.washington.edu>
RMH> Content-Length: 1645
RMH> From: haralick@horus.wtc.washington.edu
RMH> To: tombre@myrtille.loria.fr
RMH> Subject: UW CDROM
RMH> Date: Tue, 12 Oct 93 01:36:41 PDT

RMH> Perhaps you can include this in your next email distribution.
RMH> UW English Document Image Database I

A Database of Document Images for OCR Research

2 CD-ROM set containing 1147 document page images from English Scientific and Technical Journals having

- Binary images scanned from 1st and other generation photocopies
- Binary and grayscale images scanned directly from technical journals
- Synthetic noise-free images generated from LaTeX files
- Document images from UNLV ISRI database
- All document images zoned and tagged
- Software for OCR performance evaluation
- Software for simulation of photocopy degradation
- Text ground truth generated from two independent data-entry operators followed by three independent verifications

Each document page has associated with it

- Text ground truth data for each text zone
- Bounding box information for each zone on the page
- Coarse level attributes for each document page
- Finer level attributes (such as font size, alignment etc.) for each zone
- Qualitative information on the condition of each page

AT & T Bell Labs degraded character images database

Price \$200 plus \$10 shipping and handling.
 Make your P.O. out to
 Intelligent Systems Laboratory
 Dept. of Electrical Engineering, FT-10
 University of Washington
 Seattle, WA 98195
 Attention: Dr. Robert M. Haralick

Phone: (206) 685-4974
 FAX: (206) 543-3842
 e-mail: haralick@ee.washington.edu

Please make checks payable to Intelligent Systems Laboratory, Univ. of Washington.

 Then, there is the announcement of a new information server and mailing list service for document understanding, the DIMUND information server, which several of you have probably already seen:

 Attached is an announcement which is being released in the next several days about the new Document Information Server and Mailing List serving document researchers. We hope to expand the contents of the server via submissions from the community.

I would like to encourage you to submit any "information" which may be interest to the document understanding community. This includes TR's, Pre-prints of papers, Calls for Papers, Conference Announcements, Database Information, Public Domain Source Code, Bibliographic Information, etc.

Please take subscribe to the mailing list and pass this announcement on to your colleagues.

Thanks,

Dave

*****_

ANNOUNCING
The DIMUND Document Information Server

Gopher Server and Mailing List

*****_

This system is under development as a repository for Document Understanding information and resources. The site is housed with the Document Processing Group at the University of Maryland and available via internet at the node dimund.umd.edu, port 70.

Please take a moment to subscribe to the mailing list and explore the core resources on the server.

Some of the information available on this server includes:

- Research Announcements, Calls for Papers and Conference Programs
- Document-List Digest Archive
- On-line access to a document understanding bibliography including feature extraction, OCR, on-line recognition, text/graphics discrimination, signature verification and related document problems.
- A contribution directory for Document Understanding Resources
- source code
- technical reports
- database information
- Archives of document understanding internet utilities

***** Submissions of all relevant information are strongly encouraged. *****

SYSTEM ACCESS

Access to this information is available via

Gopher:

An interactive information server available at dimund.umd.edu, port 70
To add DIMUND Gopher to an existing system, use the following link:

Name=DIMUND Document Information Server
Type=1
Port=70
Path=1/
Host=dimund.umd.edu

Telnet:

Access is available on a temporary basis via telnet to dimund.umd.edu.
Use "gopher" as a login.

Anonymous FTP:

Access to the selected files on the server by connecting to dimund.umd.edu. Upload capabilities are also available.

Email:

Email Access to selected files stored on the server is available through the document server (Document-Server@dimund.umd.edu).
The DOCUMENTS mailing list is also processed by the server

Gopher access is the most complete and easiest to use. Please take time to browse through the system. Although the core resources will be compiled by the University of Maryland, the success of such a system depends on user support, contributions and advertising.

For additional help getting started, send a mail message containing the following body

HELP

to document-server@dimund.umd.edu

CONTRIBUTIONS

Contributions can be sent to documents@dimund.umd.edu in the format specified in the Gopher file

USING_THE_SYSTEM/SUBMISSION_INSTRUCTIONS

or uploaded via anonymous FTP.

The format can also be obtained via email by sending a mail request to Document-Server@dimund.umd.edu containing the single line

HELP SUBMISSIONS

MAILING LIST

A Document Understanding Mailing List (DOCUMENTS) is also available as part of the document server.

Send a mail message to Document-Server@dimund.umd.edu containing the following body

SUBSCRIBE DOCUMENTS <first name> <last name>

For additional help with the mailing list or online bibliography searches, send a mail message containing the following body

HELP MAIL-LIST
HELP BIBSEARCH

to Document-Server@dimund.umd.edu

QUESTIONS and COMMENTS

Direct all questions to gopher@dimund.umd.edu

Moderated by: Dr. David Doermann

| Document Processing Group
| Phone: (301) 405-4526 | Center For Automation Research
| Fax: (301) 314-9115 | University of Maryland
| Email: gopher@dimund.umd.edu | College Park, MD 20742
| USA

So any of you having Internet access and some gopher software on your site can add this entry to your direct gopher pointers (that's at least what I did...)

Well, this mail has already become quite long. Thanks for reading to the end. As they say, we will be back on the air as soon as there are fresh news to broadcast ;-)

Your TC-10 chairman,

--
Karl Tombre - INRIA Lorraine / CRIN-CNRS --- Email: Karl.Tombre@loria.fr
Post: Batiment LORIA, BP 239, 54506 Vandoeuvre CEDEX, France --or--
615 rue du jardin botanique, BP 101, 54602 Villers CEDEX, France
Phone: +33 83.59.20.71 --- Fax: +33 83.41.30.79 --or-- +33 83.27.83.19