

# CONFERENCE REPORT ON ICDAR '07

George Nagy

Professor Emeritus, Rensselaer Polytechnic Institute

## OUTLINE:

VENUE

INVITED SPEAKERS

AWARDS

SESSION TOPICS

NOTABLE CONTRIBUTIONS

LARGE-SCALE PROJECTS

OUTLOOK

# ICDAR CHRONOLOGY

<u>YEAR</u>	<u>LOCATION</u>	<u>ATTENDANCE</u>	<u>PRESENTATIONS</u>
1991	St. Malo	200	95
1993	Tsukuba	450	210
1995	Montreal	425	277
1997	Ulm	380	237
1999	Bangalore	360	250
2001	Seattle	500	275
2003	Shanghai	810	408
2005	Delft	450	225

## ICDAR '07

Date: August 27-29, 2007 (Mon-Wed)  
Location: Shangri La, Katmandu  
Attendance: 60 physical, 780 virtual  
Submissions: 640  
Accepted: 42 physical, 115 remote via 2-way HDTV  
Proceedings: On Docupin  
Conf. Language: English & Chinese  
Social Functions: Mediated by virtual reality terminals  
Invited Speakers: 5 (funded by NATO)  
Awards: 4

# INVITED SPEAKERS

(all second generation)

Prof. Tin Kam Ho

Web-wide Voting Network for Weak Classifiers

Dr. Tapas Kanungo

Multimedia Document Devastation Models

Dr. Rainer Hoch

High-speed Context Switches

Dr. Abdelwahab Zramdini

Web Typography

Prof. Omid Kia

Processing Compressed Multimedia Documents

# AWARDS

## **Godfried Toussaint Award for Applications of Computational Geometry to OCR**

Horst Bunke

A Parallel Algorithm for Dynamic Ray Shooting at  
Targets in Parse Trees for Error Correction

## **Mindy Bokser Prize for Discretion in Publication**

shared by three anonymous authors from an OCR company  
that did not wish to be named

## AWARDS (Cont'd)

**Rejean Plamondon Prize for the Best Analytical  
Representation of Cursive Script**

was not awarded (for the third time in a row)

**Herbert Freeman Prize for Best Name Placement**

P.S.P. **W**ang (with Sh. **X**ia, K. **Y**amamoto, & K. **Z**uyev)

**Theodosius Pavlidis Medal for Information Theoretic  
Analysis of Hellenic Bar Codes**

Eugene P. Andrews

## NINE TOPICS (32 sessions)

### I. Text Processing

Han-based scripts (4)

Sanskrit-based (3),

Hangul (1),

Roman (1),

Other (1)

Unicode-64: 24-bit symbol identifier, 12-bit typeface id, 8-bit type-size, 8-bit vertical offset, 8-bit color, 4-bits free

Little on hand print and cursive since so many schools stopped teaching it in favor of elocution for speech input

Distributed OCR (a cottage industry on the Web)

Low-resolution technicolor webscript

## II. Graphics

### Synthetic (graphics recognition)

Mostly computer-produced utility maps and engineering drawings from deplatformed CAD tools

Source identification for 2-D bar codes

### Natural (picture processing)

Caption generator for recreational digital photography;  
PortraitPort to geneological databases,  
Image-content based retrieval (mainly sunsets).



### III. Devices

#### Storage

Docupin, DocuPinCushion (DPC)

750 MB/DP, 18GB/DPC

Not addressable

#### E-Scap

Cholesteric polymer-dispersed hybrid substrate

10 micron thick (100 sheets per cm)

Ten megacell array provides 300 dpi resolution

10\*\*6 read-write cycles

Ecologically benign (silicon, not carbon, based)

No color so far

## I/O

### Radiotablet

Lambertian high-contrast *reflective* display

Satellite cyberband channel

Pentameric (soft, 4mm thick elastomere) vary-keyboard

100 MB flash-cache

Docupin port

### Prisco (Printer-scanner-copier)

Reads and writes paper and E-scap

Docupin port

SCRAPPY port (Segmented Colophonic Reversible  
Petabus Parallel Ynterface)

1 GB internal storage

expected to displace most printers, scanners, fax  
machines and copiers within five years

### Docupeeper

## IV. Data Mining and Warehousing

Search Engines

Filters

Sound

"Find what I want, not what I ask"

Push-pull filters for Web interface

## V. Sublimation

Contextual interpolation and extrapolation

Autonomous correspondence agents (>640 on Turing Test)

Condensers (semantic compression)

- typical home page  $\Rightarrow$  56 bytes

Unifying Theory for object-oriented ontologies,  
decision trees, BBNs, ANNs, and HMMs.

## **VI. Oral Documents**

Speech-to-print

Print-to-speech

Convergence of speech and document image processing -  
began in 1990's with adoption of HMM in OCR

## **VII. Docunets**

Generalized LZW: transparent external pointers  
i.e., distributed files with automatic link insertion

## VIII. Epistepchnology

"Unsupervised" learning and Clustering

Bayesian belief networks, Neural Nets and Genetic Algorithms

(training versus tracking and adaptation)

Rule-based (expert systems)

Some examples:

Self-corrective recognition

(update character-shape models)

D/A reconstruction with random-phase sampling

Adaptive noise and degradation models

Adaptive language models

(update n-gram statistics, lexicon, word transitions)

Adaptive segmentation and alignment

Adaptive script recognition (string averages)

Adaptive vectorization (constraints  $\Rightarrow$  parameters)

## IX. History of OCR & DIA

*Cultivating self-awareness is appropriate for a maturing discipline.*

Where do we publish?

1913 Proceedings of the IEEE

1952 IEEE-EC (C)

1955 IEEE-IT

1966 Computer (Computer Group News)

1968 Pattern Recognition (Pergamon)

## Journals (cont'd)

- 1971 IEEE-SMC
- 1972 CGVIP (Academic)
- 1975 J. Imaging Technology (SPIE)
- 1979 IEEE-PAMI
  
- 1980 Pattern Recognition Letters (IAPR/Elsevier)
- 1983 Image and Vision Computing (Elsevier)
- 1985 Computer Vision (Kluwer)
- 1987 IJPRAI (World Scientific)
- 1988 Machine Vision and Applications (Springer)
  
- 1991 Pattern Recognition and Image Analysis (Pleiades)
- 1991 Mathematical Imaging and Vision (Kluwer)
- 1992 IEEE-IP
- 1992 Electronic Imaging (SPIE)
- 1995 Multimedia Tools and Applications (Kluwer)
- 1998 Document Analysis and Recognition (Kluwer)

## Where do we meet?

1976	ICPR	1983	IGS
1981	ICIAP	1988	FHW
1986	SPIE	1990	SSPR
1990	CNED/COFED	1991	MVA
1991	ICDAR	1992	DIMUND
1992	SDAIR	1994	ACV
1994	ICIP	1994	DAS
1995	SDIUT	1995	GREC

Annual yield of papers on OCR & DIA 75  $\Rightarrow$  1230 in 25 years  
(1963 bibliography on OCR had 900 entries)

Statistics estimated with standard discourse analysis tools:

Median annual citation frequency: 0.32

Content overlap (shared authors): 0.68

13.4 year random phase periodicity



## 15 NOTABLE ICDAR '07 PAPERS

Henry Baird and Larry Spitz:  
128-Language OCR

Luigi Cordella and Gabriella Sanniti de Baja:  
19-37-79-133 Weighted Distance Map for HP Digits

Andreas Dengel (Minister of German PTT):  
Error Detection and Correction in  
Government-Industry Communications

David Doermann & Junichi Kanai:  
Compression of Compressed Documents

Dov Dori (President-elect, IAPR):  
In Situ Classification of Navajo Arrowheads

Richard Duda, Peter Hart, David Stork:  
Pattern Recognition and Scene Analysis  
(Synopsis of 2nd Edition, to be ready by 2008).

Robert Haralick:  
All the World's Papers on Docupin &  
A Contravariant Tensor Noise Model for Bleed-through

Jonathan Hull:  
Paragraph Recognition Using Viterbi Syntax

Anil Jain:  
Gabor Filters for On-line Script Recognition

Rangachar Kasturi:  
Extraction of Political Boundaries from  
Large-Scale Map of Antarctica

Gary Kopec:

A Communications Approach to Decoding Dante's Inferno

Louisa Lam & Ching Suen:

Fractals and Wavelets for Circuit Diagram Conversion

Seon-Whang Lee:

Neural Networks, Dynamic Programming,  
Arithmetic Coding and Simulated Annealing for  
Context-sensitive Script Recognition

Juergen Schuermann:

From Atoms to Goethe (in 3 volumes)

Karl Tombre:

Rapid Prototyping from Rough Sketches:  
a Lathe with a PaperPort

## LARGE-SCALE PROJECTS

### Cultural:

variants, glosses, critiques, commentaries, interpretations, exegeses, reviews, dissertations, concordances, ....

Western Poetry: *The International Poetry Collection*

Chinese Classics: *Buddhist Writings, Book of Changes (I-Ching), Book of Odes (Shi-Ching), Dynastic Histories (219 volumes), Tale of the Water Margins, The Romance of the Three Kingdoms, The Dream of the Red Chamber, The Monkeys, Strange Stories, ...*

### Hindu Classics:

*The Vedas, Brahma-sutra, Bhagavad-Gita, Puranas, Shankutala, Panchatantra, The Clay Cart, Brihatkatha, Kathasaritzagara, Ramayana, Mahabharata, ...*

# LARGE-SCALE PROJECTS

## (cont'd)

### **Bureaucratic:**

Document Declassification (based on 1994 Executive Order)  
(yearly increase in number of classified documents drops to 35%)

NIST METTREC: expert system for decoding US Gov. documents

US DOE Nuclear Data Depository (Yucca Mountain waste burial project revived: now scheduled to begin in 2030.)

L/C begins compilation of a comprehensive OCR/DIA collection  
(about 7000 titles)

## PROPHECIES FOR ICDAR '17

Results of *First Decennial Document Census* (2010) published

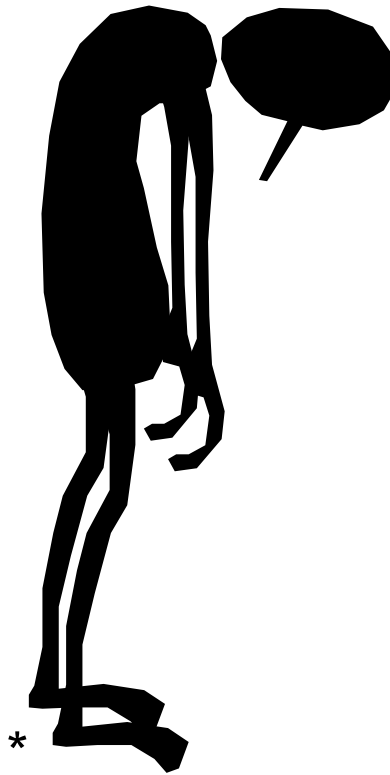
OCR and DIA will be integrated into Windows/2017++

Speech dialog fades as computer penetration becomes universal

FDA will approve foreign language implants

Experimentation on eye-bypass for human document acquisition

Walt Disney will invest \$100,000,000 into Book- to-Virtual Reality Converter and release VR version of *Gone with the Wind*



**AUF WIEDERSEHEN!**