

1. Record Nr.	UNINA9910144349303321
Titolo	Document Analysis Systems VI : 6th International Workshop, DAS 2004, Florence, Italy, September 8-10, 2004, Proceedings // edited by Simone Marinai, Andreas Dengel
Pubbl/distr/stampa	Berlin, Heidelberg : , : Springer Berlin Heidelberg : , : Imprint : Springer, , 2004
ISBN	3-540-28640-3
Edizione	[1st ed. 2004.]
Descrizione fisica	1 online resource (XII, 568 p.)
Collana	Lecture Notes in Computer Science, , 0302-9743 ; ; 3163
Disciplina	006.4/2
Soggetti	Pattern recognition Information storage and retrieval Optical data processing Computer simulation Application software Pattern Recognition Information Storage and Retrieval Image Processing and Computer Vision Simulation and Modeling Computer Appl. in Administrative Data Processing
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Bibliographic Level Mode of Issuance: Monograph
Nota di bibliografia	Includes bibliographical references at the end of each chapters and index.
Nota di contenuto	Digital Libraries -- Document Analysis Systems for Digital Libraries: Challenges and Opportunities -- The Trinity College Dublin 1872 Online Catalogue -- DL Architecture for Indic Scripts -- A Semantic-Based System for Querying Personal Digital Libraries -- Toward Personalized Digital Library for Providing "Information JIT" -- Historical Documents -- Tilting at Windmills: Adventures in Attempting to Reconstruct Don Quixote -- A Segmentation-Free Recognition Technique to Assist Old Greek Handwritten Manuscript OCR -- Automatic Metadata Retrieval from Ancient Manuscripts -- A Complete Approach to the Conversion of Typewritten Historical Documents for Digital Archives -- An Adaptive Binarization Technique for Low Quality

Historical Documents -- Segmentation of Handwritten Characters for Digitalizing Korean Historical Documents -- Self-organizing Maps and Ancient Documents -- Enriching Historical Manuscripts: The Bovary Project -- Layout Analysis -- Word Grouping in Document Images Based on Voronoi Tessellation -- Multi-component Document Image Coding Using Regions-of-Interest -- Physical Layout Analysis of Complex Structured Arabic Documents Using Artificial Neural Nets -- An Integrated Approach for Automatic Semantic Structure Extraction in Document Images -- Multi-view hac for Semi-supervised Document Image Classification -- Configurable Text Stamp Identification Tool with Application of Fuzzy Logic -- Layout and Content Extraction for PDF Documents -- Automatic Extraction of Filled-In Items from Bank-Check Images -- Color Documents -- Bleed-Through Removal from Degraded Documents Using a Color Decorrelation Method -- Colour Map Classification for Archive Documents -- Serialized k-Means for Adaptive Color Image Segmentation -- Adaptive Region Growing Color Segmentation for Text Using Irregular Pyramid -- Preprocessing and Segmentation of Bad Quality Machine Typed Documents -- Handwritten Documents -- Ensembles of Classifiers for Handwritten Word Recognition Specialized on Individual Handwriting Style -- Information Retrieval System for Handwritten Documents -- Word-Wise Script Identification from Indian Documents -- Recognizing Freeform Digital Ink Annotations -- Post-processing of Handwritten Pitman's Shorthand Using Unigram and Heuristic Approaches -- Multiscale Handwriting Characterization for Writers' Classification -- Graphics Recognition -- A Hybrid Approach to Detect Graphical Symbols in Documents -- Performance Evaluation of Symbol Recognition -- The Search for Genericity in Graphics Recognition Applications: Design Issues of the Qgar Software System -- Attributed Graph Matching Based Engineering Drawings Retrieval -- A Platform to Extract Knowledge from Graphic Documents. Application to an Architectural Sketch Understanding Scenario -- Internet Documents -- A Graph-Based Framework for Web Document Mining -- XML Documents Within a Legal Domain: Standards and Tools for the Italian Legislative Environment -- Rule-Based Structural Analysis of Web Pages -- Extracting Table Information from the Web -- A Neural Network Classifier for Junk E-Mail -- Document Analysis Systems -- Results of a Study on Invoice-Reading Systems in Germany -- A Document Analysis System Based on Text Line Matching of Multiple OCR Outputs -- DocMining: A Document Analysis System Builder -- Automatic Fax Routing -- Applications -- Contextual Swarm-Based Multi-layered Lattices: A New Architecture for Contextual Pattern Recognition -- Natural Language Processing of Patents and Technical Documentation -- Document Image Retrieval in a Question Answering System for Document Images -- A Robust Braille Recognition System -- Document Image Watermarking Based on Weight-Invariant Partition Using Support Vector Machine -- Video Degradation Model and Its Application to Character Recognition in e-Learning Videos -- Unity Is Strength: Coupling Media for Thematic Segmentation.

---

### Sommario/riassunto

This volume contains papers selected for presentation at the 6th IAPR Workshop on Document Analysis Systems (DAS 2004) held during September 8–10, 2004 at the University of Florence, Italy. Several papers represent the state of the art in a broad range of “traditional” topics such as layout analysis, applications to graphics recognition, and handwritten documents. Other contributions address the description of complete working systems, which is one of the strengths of this workshop. Some papers extend the application domains to other media, like the processing of Internet documents. The peculiarity of this 6th workshop

was the large number of papers related to digital libraries and to the processing of historical documents, a task which frequently requires the analysis of color documents. A total of 17 papers are associated with these topics, whereas two years ago (in DAS 2002) only a couple of papers dealt with these problems. In our view there are three main reasons for this new wave in the DAS community. From the scientific point of view, several research fields reached a thorough knowledge of techniques and problems that can be effectively solved, and this expertise can now be applied to new domains. Another incentive has been provided by several research projects funded by the EC and the NSF on topics related to digital libraries.

---