

1. Record Nr.	UNINA9910790976203321
Autore	Chaudhuri B. B (Bidyut Baran)
Titolo	Advances in digital document processing and retrieval // editors, Bidyut Baran Chaudhuri, Swapan Kumar Parui, Indian Statistical Institute, India
Pubbl/distr/stampa	New Jersey : , : World Scientific, , [2014] 2014
ISBN	981-4368-71-7
Descrizione fisica	1 online resource (x, 323 pages) : illustrations
Collana	Statistical science and interdisciplinary research, , 1793-6195 ; ; volume 13 Platinum jubilee series
Disciplina	025.04
Soggetti	Information storage and retrieval systems Documentation - Data processing Digital preservation Document imaging systems Electronic publishing Multimedia systems Electronic records
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Description based upon print version of record.
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	Foreword; Preface; Contents; 1. Document Image Analysis using Markovian Models: Application to Historical Documents; 1.1. Introduction; 1.2. Hidden Markov Random Field Models; 1.2.1. Theoretical foundations; 1.2.1.1. Simulated annealing; 1.2.1.2. Iterated Conditional Modes (ICM); 1.2.1.3. Highest Confidence First (HCF) algorithm; 1.2.1.4. 2D Dynamic Programming; 1.2.2. Application of MRF labelling to handwritten document segmentation; 1.2.2.1. Probability densities; 1.2.2.2. Clique potential functions; 1.2.2.3. Observations; 1.2.2.4. Decoding strategy; 1.2.3. Results; 1.2.3.1. Zone labelling 1.2.3.2. Text line labelling 1.2.3.3. Conclusion; 1.3. Conditional Random Field Models; 1.3.1. Proposed model; 1.3.1.1. Feature functions; 1.3.1.2. Model inference; 1.3.1.3. Parameter learning; 1.3.2.

A two level CRF model; 1.3.2.1. Observation features; 1.3.2.2. Label features; 1.3.2.3. Learning; 1.3.3. Integrating more contextual information; 1.3.3.1. Global feature function; 1.3.3.2. Combination of the information sources; 1.3.3.3. Linear combination of the information sources (impl.2); 1.3.3.4. Combination of the information sources using an MLP (impl.3); 1.3.4. Experiments and results  
 1.4. Conclusions and Outlook  
 Acknowledgments; References; 2. Information Just-in-Time: Going Beyond the Myth of Paperlessness; 2.1. Introduction; 2.2. Information Just-in-Time; 2.2.1. Personal Information Environment; 2.2.2. Hot/Warm/Cold Documents; 2.2.3. Proposed Approach; 2.3. Digital Pen Solution; 2.3.1. Anoto Functionality; 2.3.2. Data Entry Applications; 2.4. iJIT Collaboration Platform; 2.4.1. On-Demand Printing; 2.4.2. Hybrid Document Management System; 2.4.3. Research Notebook Application - iJITNote; 2.4.4. Future Directions; 2.5. Conclusions; Acknowledgments; References  
 3. The Role of Document Image Analysis in Trustworthy Elections  
 3.1. Introduction; 3.2. History; 3.3. Problems with Current Voting Technologies; 3.4. Experimental Approaches to Reliable Processing of Voting Records; 3.4.1. Statistical distribution of mark sense errors; 3.4.2. Unbiased context-free visual auditing based on ballot images; 3.4.3. Homogenous class display; 3.4.4. Unique identification of ballots; 3.4.5. Error characteristics of DRE with VVPAT; 3.4.6. Development of testing procedures for voting systems; 3.4.7. Affordances for voters with disabilities; 3.5. Some Related Efforts  
 3.6. Concluding Remarks  
 References; 4. Information Retrieval from Document Image Databases; 4.1. Introduction; 4.2. Related Work; 4.3. Word Shape Coding; 4.3.1. Word Shape Coding by Character Stroke Categorization; 4.3.2. Word Shape Coding by Character Boundary Extrema; 4.3.3. Word Shape Coding by Character Holes and Reservoirs; 4.4. Document Image Retrieval; 4.4.1. Document Vector Construction; 4.4.2. Document Similarity Measurement; 4.5. Discussions; 4.5.1. Coding Ambiguity; 4.5.2. Coding Robustness; 4.5.3. Document Similarity Measurements; 4.5.4. Coding Scheme Selection; 4.6. Conclusion  
 References

**Sommario/riassunto**

From the participation of researchers in most important international conferences in the field, it is noted that activities in automatic document processing have been continuously growing. This book is an edited volume in Digital Document Processing where the chapters are written by several internationally renowned researchers in the domain. It will be useful for both students and researchers working on various aspects of document image analysis and recognition problems. It contains chapters on topics that are not covered by any textbook, but are more futuristic like "Going beyond the Myth of