

Proceedings

Ninth International Conference on Document Analysis and Recognition

**Curitiba, Paraná, Brazil
September 23-26, 2007**



Los Alamitos, California
Washington • Tokyo



TABLE OF CONTENTS

VOLUME 1

Welcome from the Honourary Chair
Welcome from the General Chairs
Message from the PC Chairs
Conference Committees
Reviewers

KEYNOTE I

A View on the Past and Future of Character and Document Recognition 1
Hirohichi Fujisawa

SESSION A1: HANDWRITING RECOGNITION (1)

Using Hidden Markov Models as a Tool for Handwritten Text Line Segmentation..... 6
Florence Luthy, Tamas Varga, Horst Bunke

Character Stroke Extraction Based on B-spline Curve Matching by Constrained Alternating Optimization 11
Xiabi Liu, Yunde Jia

Non-Uniform Slant Correction for Handwritten Text Line Recognition..... 16
Roman Bertolami, Seiichi Uchida, Matthias Zimmermann, Horst Bunke

A Method of Annotation Extraction from Paper Documents Using Alignment Based on Local Arrangements of Feature Points..... 21
Tomohiro Nakai, Koichi Kise, Masakazu Iwamura

Learning to Group Text Lines and Regions in Freeform Handwritten Notes..... 26
Ming Ye, Paul Viola, Sashi Raghupathy, Herry Sutanto, Chengyang Li

Character-Stroke Detection for Text-Localization and Extraction 31
Krishna Subramanian, Prem Natarajan, Michael Decerbo, David Castañón

SESSION A2: ON-LINE HANDWRITING RECOGNITION

Kernel Modified Quadratic Discriminant Function for Online Handwritten Chinese Characters Recognition 36
Duanduan Yang, Lianwen Jin

A Minimax Classification Approach to HMM-Based Online Handwritten Chinese Character Recognition Robust Against Affine Distortions 41
Qiang Huo, Tingting He

Online Handwritten Japanese Character String Recognition Incorporating Geometric Context..... 46
Xiang-Dong Zhou, Jin-Lun Yu, Cheng-Lin Liu, Takeshi Nagasaki, Katsumi Marukawa

Minimum Error Discriminative Training for Radical-based Online Chinese Handwriting Recognition	51
<i>Yaodong Zhang, Peng Liu, Frank K. Soong</i>	
Direction Code Based Features for Recognition of Online Handwritten Characters of Bangla	56
<i>Ujjwal Bhattacharya, Bikash K. Gupta, Swapan K. Parui</i>	
HMM-based Online Handwriting Recognition System for Telugu Symbols	61
<i>V. Jagadeesh Babu, L. Prasanth, R. Raghunath Sharma, G.V. Prabhakara Rao, A. Bharath</i>	

SESSION A3: DOCUMENT RETRIEVAL

A Fast Keyword-Spotting Technique	66
<i>Linlin Li, Shijian Lu, Chew Lim Tan</i>	
Locality Sensitive Pseudo-Code for Document Images	71
<i>Kengo Terasawa, Yuzuru Tanaka</i>	
XML Data Representation in Document Image Analysis	76
<i>Abdel Bela d, Ingrid Falk, Yves Rangoni</i>	
Language Models for Handwritten Short Message Services	81
<i>Emmanuel Prochasson, Christian Viard-Gaudin, Emmanuel Morin</i>	
Vector Model Based Indexing and Retrieval of Handwritten Medical Forms	86
<i>Huaigu Cao, Venu Govindaraju</i>	
Document Content Inventory & Retrieval	91
<i>Michael A. Moll, Henry S. Baird</i>	

POSTER SESSION 1

Strategies for Large Handwritten Farsi/Arabic Lexicon Reduction	96
<i>Saeed Mozaffari, Karim Faez, Volker Märgner, Haikal El Abed</i>	
Fusing Asynchronous Feature Streams for On-line Writer Identification	101
<i>Andreas Schlapbach, Horst Bunke</i>	
Writer Identification in Handwritten Documents	106
<i>Imran Ahmed Siddiqi, Nicole Vincent</i>	
Robust Document Warping with Interpolated Vector Fields	111
<i>David C. Schneider, Marco Block, Raúl Rojas</i>	
Simultaneous Layout Style and Logical Entity Recognition in a Heterogeneous Collection of Documents	116
<i>Siyuan Chen, Song Mao, George R. Thoma</i>	
Learning of Pattern-Based Rules for Document Classification	121
<i>Andreas Dengel</i>	
A MSD-HMM Approach to Pen Trajectory Modeling for Online Handwriting Recognition	126
<i>Lei Ma, Yi-Jian Wu, Peng Liu, Frank Soong</i>	
HMM-Based Recognizer with Segmentation-free Strategy for Unconstrained Chinese Handwritten Text	131
<i>Tong-Hua Su, Tian-Wen Zhang, Hu-Jie Huang, Yu Zhou</i>	

Text Line Segmentation of Historical Arabic Documents	136
<i>Abderrazak Zahour, Laurence Likforman-Sulem, Wafa Boussalaa, Bruno Taconet</i>	
Document Images Retrieval Based on Multiple Features Combination	141
<i>Gaofeng Meng, Nanning Zheng, Yonghong Song, Yuanlin Zhang</i>	
How to Reduce the Size of Bank Check Image Archive?	146
<i>Vladimir Shapiro</i>	
A Human-Centric Off-Line Signature Verification System	151
<i>Hanno Coetzer, Robert Sabourin</i>	
A Model-based Book Dewarping Method to Handle 2D Images Captured by a Digital Camera	156
<i>Minghui Wu, Rongfeng Li, Bin Fu, Wenxin Li, Zhuoqun Xu</i>	
Ink Annotations and their Anchoring in Heterogeneous Digital Documents	161
<i>Xin Wang, Sashi Raghupathy</i>	
A Method for the Synthesis of Dynamic Biometric Signature Data	166
<i>C. Rabasse, R.M.Guest, M.C. Fairhurst</i>	
Recognition of Broken Characters from Historical Printed Books Using Dynamic Bayesian Networks	171
<i>Laurence Likforman-Sulem, Marc Sigelle</i>	
An Efficient Word Segmentation Technique for Historical and Degraded Machine-Printed Documents	176
<i>M. Makridis, N. Nikolaou, B. Gatos</i>	
Circular Noises Removal from Scanned Document Images	181
<i>Gaofeng Meng, Nanning Zheng, Yuanlin Zhang, Yonghong Song</i>	
Automatic Extraction of Data from 2-D Plots in Documents	186
<i>Xiaonan Lu, James Z. Wang, Prasenjit Mitra, C. Lee Giles</i>	
An Efficient Thresholding Algorithm for Brazilian Bank Checks	191
<i>Carlos A.B.Mello, Byron L.D.Bezerra, Cleber Zanchettin, Valmir Macário</i>	
Recognition of Fragmented Characters Using Multiple Feature-Subset Classifiers	196
<i>Chien-Hsing Chou, Chien-Yang Guo, Fu Chang</i>	
Web Document Parsing: A New Approach to Modeling Layout-Language Relations	201
<i>Minoru Yoshida, Hiroshi Nakagawa</i>	
Cryptogram Decoding for OCR using Numerization Strings	206
<i>Gary Huang, Erik Learned-Miller, Andrew McCallum</i>	
Knowledge-based Recognition of Utility Map Sub-diagrams	211
<i>Simon J. Hickinbotham, Anthony G. Cohn</i>	
Synthesis of Chinese Character Using Affine Transformation	216
<i>LianWen Jin, XiaoNa Zu</i>	
Adaptive Window Based Uneven Lighting Document Segmentation	221
<i>Guoqing Gu, Wenwen Han</i>	
A Review of Shape Descriptors for Document Analysis	225
<i>O. Ramos Terrades, S. Tabbone, E. Valveny</i>	
A High Performance European OCR System	230
<i>Kai Wang, Qingren Wang</i>	
Automatic Detection of Document Script and Orientation	235
<i>Shijian Lu, Chew Lim Tan</i>	

Dynamic Handwritten Keyword Spotting based on the NSHP-HMM	240
<i>Christophe Choisy</i>	
Contribution of Multiresolution Description for Archive Document Structure Recognition	245
<i>Aurelie Lemaitre, Jean Camillerapp, Bertrand Couasnon</i>	
Iterated Document Content Classification	250
<i>Chang An, Henry S. Baird, Pingping Xiu</i>	
Human Image Preference and Document Degradation Models	255
<i>Chris Hale, Elisa H. Barney Smith</i>	
Text and Layout Information Extraction from Document Files of Various Formats Based on the Analysis of Page Description Language	260
<i>Takashi Hirano, Yuichi Okano, Yasuhiro Okada, Fumio Yoda</i>	
Curved Paper Rectification for Digital Camera Document Images by Shape from Parallel Geodesics Using Continuous Dynamic Programming	265
<i>Katsuhito Fujimoto, Jun Sumb, Hiroaki Takebea, Misako Suwaa, Satoshi Naoi</i>	
Annotated Databases for the Recognition of Screen-Rendered Text	270
<i>Steffen Wachenfeld, Hans-Ulrich Klein, Xiaoyi Jiang</i>	
An EM Based Algorithm for Skew Detection	275
<i>Amir Egozi, Itshak Dinstein</i>	
Automatic Writer Identification Based on a Velocity Profile Approach to Processing of Difficult Handwriting Styles	280
<i>J. Chapran, M. Fairhurst</i>	
A Novel Binarization System for Degraded Document Images	285
<i>Yan Xi, Youbin Chen, Qingmin Liao, Leung Winghong, Fung Shunming, Deng Jiangwen</i>	
A Suffix Tree Based Handwritten Chinese Address Recognition System	290
<i>Yan Jiang, Xiaqing Ding, Zheng Ren</i>	
Language-Based Feature Extraction Using Template- Matching In Farsi/Arabic Handwritten Numeral Recognition	295
<i>Majid Ziaratban, Karim Faez, Farhad Faradji</i>	
Recognition of Character Strings in Low-quality Images Using Character and Inter-character Space Patterns	300
<i>Hiroyuki Ishida, Tomokazu Takahashi, Ichiro Ide, Hiroshi Murase</i>	
Locating Charts from Scanned Document Pages	305
<i>Weihua Huan</i>	
<i>g, Chew Lim Tan</i>	
Binarization of Badly Illuminated Document Images through Shading Estimation and Compensation	310
<i>Shijian Lu, Chew Lim Tan</i>	
A Multi-stream Approach to Off-Line Handwritten Word Recognition	315
<i>Yousri Kessentini, Thierry Paquet, AbdelMajid Benhamadou</i>	
Finger-Written Chinese Characters Recognition using Hierarchical LDA	320
<i>Duanduan Yang, Lianwen Jin, Li-Xin Zhen, Jiang-Cheng Huang</i>	
A Case-based Reasoning Approach for Invoice Structure Extraction	325
<i>Hatem Hamza, Yolande Bela d, Abdel Bela d</i>	

An Approach to Improve Accuracy Rate of On-line Signature Verification Systems of Different Sizes	330
<i>Rodrigo S. A. Araujo, George D. C. Cavalcanti, Edson C. B. C. Filho</i>	

KEYNOTE II

Energy-Based Models in Document Recognition and Computer Vision	335
<i>Yann LeCun, Sumit Chopra, Marc'Aurelio Ranzato, Fu-Jie Huang</i>	

SESSION B1: HANDWRITING RECOGNITION (2)

Learning Handwritten Digit Recognition by the Max-Min Posterior Pseudo-Probabilities Method	340
<i>Xuefeng Chen, Xiabi Liu, Yunde Jia</i>	
Transform-Invariance in Local Averaging Classifier for Handwritten Digit Pattern Recognition	345
<i>Seiji Hotta</i>	
Identification of Latin-based Languages through Character Stroke Categorization	350
<i>Shijian Lu, Linlin Li, Chew Lim Tan</i>	
Layout Analysis of Handwritten Historical Documents for Searching the Archive of the Cabinet of the Dutch Queen	355
<i>Marius Bulacu, Rutger van Koert, Lambert Schomaker, Tijn van der Zant</i>	
Indexing Historical Documents by Word Shape Signatures	360
<i>Josep Lladós, Gemma Sanchez</i>	

SESSION B2: ON-LINE HANDWRITING RECOGNITION (2)

A Novel Approach to On-Line Handwriting Recognition Based on Bidirectional Long Short-Term Memory Networks	365
<i>Marcus Liwicki, Alex Graves, Horst Bunke, Jurgen Schmidhuber</i>	
Combining On-Line and Off-Line Systems for Handwriting Recognition	370
<i>Marcus Liwicki, Horst Bunke</i>	
Text/Non-text Ink Stroke Classification in Japanese Handwriting Based on Markov Random Fields	375
<i>Xiang-Dong Zhou, Cheng-Lin Liu</i>	
Use of a Confusion Network to Detect and Correct Errors in an On-Line Handwritten Sentence Recognition System	380
<i>Solen Quiniou, Eric Anquetil</i>	
Redundant Bit Vectors for Robust Indexing and Retrieval of Electronic Ink	385
<i>Kumar Chellapilla, John Platt</i>	

SESSION B3: PAGE SEGMENTATION

Robust Page Segmentation Based on Smearing and Error Correction Unifying Top-down and Bottom-up Approaches	390
<i>Huaigu Cao, Rohit Prasad, Prem Natarajan, Ehry MacRostie</i>	

Decompose Document Image Using Integer Linear Programming	395
<i>Dashan Gao, Yizhou Wang, Haitham Hindi, Minh Do</i>	
Character Line Segmentation Based on Feature Clustering	400
<i>Yan Xi, Youbin Chen, Qingmin Liao, Leung Winghong, Fung Shunming, Deng Jiangwen</i>	
Document Image Segmentation Using a 2D Conditional Random Field Model	405
<i>Stéphane Nicolas, Julien Dardenne, Thierry Paquet, Laurent Heutte</i>	
Text Segmentation from Complex Background Using Sparse Representations	410
<i>W.M. Pan, T.D. Bui, C.Y. Suen</i>	

SESSION C1: CLASSIFIER

Exploiting Fisher Kernels in Decoding Severely Noisy Document Images	415
<i>Jindong Chen, Yizhou Wang</i>	
K-Nearest Oracle for Dynamic Ensemble Selection	420
<i>Albert Hung-Ren Ko, Robert Sabourin, Alceu de Souza Britto Jr.</i>	
Multi-Objective Optimization for SVM Model Selection	425
<i>C. Chatelain, S. Adam, Y. Lecourtier, L. Heutte, T. Paquet</i>	
A Classifier of Similar Characters using Compound Mahalanobis Function based on Difference Subspace	430
<i>Junichi Hirayama, Hidehisa Nakayama, Nei Kato</i>	
Extraction of Embedded Class Information from Universal Character Pattern	435
<i>Seiichi Uchida, Megumi Sakai, Masakazu Iwamura, Shinichiro Omachi, Koichi Kise</i>	

SESSION C2: ON-LINE HANDWRITING APPLICATIONS

Context-Driven Constraint Multiset Grammars with Incremental Parsing for On-line Structured Document Interpretation	440
<i>Sebastien Mace, Eric Anquetil</i>	
On-Line Handwritten Text Line Detection Using Dynamic Programming	445
<i>Marcus Liwicki, Emanuel Indermuhle, Horst Bunke</i>	
An Incremental On-line Parsing Algorithm for Recognizing Sketching Diagrams	450
<i>Joan Mas, Gemma Sanchez, Josep Lladós, Bart Lamiroy</i>	
New Aspects of InkML for Pen-Based Computing	455
<i>Stephen M. Watt</i>	
Frechet Distance based Approach for Searching Online Handwritten Documents	459
<i>R. Sriraghavendra, K. Karthik, Chiranjib Bhattacharyya</i>	

SESSION C3: COLOR ANALYSIS AND PERFORMANCE EVALUTION

Color Document Synthesis as a Compression Strategy	464
<i>João Marcelo Monte da Silva, Rafael Dueire Lins</i>	
Fast Selection of Small and Precise Candidate Sets from Dictionaries for Text Correction Tasks	469
<i>Stoyan Mihov, Peter Mitankin, Klaus U. Schulz</i>	

Automatic Ground-truth Generation for Document Image Analysis and Understanding	474
<i>Pierre Hérroux, Eugen Barbu, Sébastien Adam, Éric Trupin</i>	
New Metrics for Evaluating Performance in Document Analysis Tasks - Application to the Table Case -	479
<i>Ana Costa e Silva</i>	
Identification of Non-Black Inks using HSV Colour Space	484
<i>Haritha Dasari, Chakravarthy Bhagvati</i>	

SESSION D1: ASIAN CHARACTER RECOGNITION

Curvelet-based Multi SVM Recognizer for Offline Handwritten Bangla: A Major Indian Script	489
<i>B.B. Chaudhuri, A. Majumdar</i>	
Off-Line Handwritten Character Recognition of Devnagari Script	494
<i>U. Pal, N. Sharma, T. Wakabayashi, F. Kimura</i>	
Modular Approach to Recognition of Strokes in Telugu Script	499
<i>Anitha Jayaraman, C. Chandra Sekhar, V. Srinivasa Chakravarthy</i>	
Hidden Markov Models for Online Handwritten Tamil Word Recognition	504
<i>A. Bharath, Sriganesh Madhvanath</i>	
A Two Stage Recognition Scheme for Handwritten Tamil Characters	509
<i>U. Bhattacharya, S. K. Ghosh, S. K. Parui</i>	

SESSION D2: GRAPHICS RECOGNITION

A New Syntactic Approach to Graphic Symbol Recognition	514
<i>Yu Yajie, Wan Zhang, Liu Wenyin</i>	
Extraction of Vectorized Graphical Information from Scientific Chart Images	519
<i>Ruizhe Liu, Weihua Huang, Chew Lim Tan</i>	
Robust Circle Detection	524
<i>Bart Lamiroy, Olivier Gaucher, Laurent Fritz</i>	
Toponym Recognition in Scanned Color Topographic Maps	529
<i>Joachim Pouderoux, Jean-Christophe Gonzato, Aurelien Pereira, Pascal Guitton</i>	
A New Vectorial Signature for Quick Symbol Indexing, Filtering and Recognition	534
<i>Wan Zhang, Liu Wenyin</i>	

SESSION D3: IMAGE PREPROCESSING AND IMAGE CLASSIFICATION

PDE-based Enhancement of Low Quality Documents	539
<i>Ifeoma Nwogu, Zhixin Shi, Venu Govindaraju</i>	
Non-Local Text Image Reconstruction	544
<i>Hiep Quang Luong, Wilfried Philips</i>	
SVM Based Scheme for Thai and English Script Identification	549
<i>S. Chanda, Oriol Ramos Terrades, U. Pal</i>	
A Hybrid System for Robust Recognition of Ethiopic Script	554
<i>Yaregal Assabie, Josef Bigun</i>	

WEB Image Classification Based on the Fusion of Image and Text Classifiers	559
<i>Pedro R. Kalva, Fabricio Enembreck, Alessandro L. Koerich</i>	

POSTER SESSION II

Assessing and Improving the Quality of Document Images Acquired with Portable Digital Cameras	564
<i>Rafael Dueire Lins, Gabriel Pereira e Silva, André Ricardson Gomes e Silva</i>	
A Multi-stage Strategy to Perspective Rectification for Mobile Phone Camera-based Document Images	569
<i>Xu-Cheng Yin, Jun Sun, Satoshi Naoi, Katsuhito Fujimoto, Hiroaki Takebe, Yusaku Fujii, Koji Kurokawa</i>	
Degraded Character Recognition by Complementary Classifiers Combination	574
<i>Jun Sun, Kaizhu Huang, Yoshinobu Hotta, Katsuhito Fujimoto, Satoshi Naoi</i>	
Middle Zone Component Extraction and Recognition of Telugu Document Image	579
<i>Pratap Reddy L, Satyaprasad L</i>	
An SVM-Based High-accurate Recognition Approach for Handwritten Numerals by Using Difference Features	584
<i>Kaizhu Huang, Jun Sun, Yoshinobu Hotta, Katsuhito Fujimoto, Satoshi Naoi</i>	
Retrieval of Handwritten Lines in Historical Documents	589
<i>L.R.B. Schomaker</i>	
Text Line Detection in Unconstrained Handwritten Documents Using a Block- Based Hough Transform Approach	594
<i>G. Louloudis, B. Gatos, C. Halatsis</i>	
Systematic Multi-Path HMM Topology Design for Online Handwriting Recognition of East Asian Characters	599
<i>Shi Han, Ming Chang, Yu Zou, Xinjian Chen, Dongmei Zhang</i>	
Bibliographic Meta-Data Extraction Using Probabilistic Finite State Transducers	604
<i>Martin Kramer, Hagen Kaprykowsky, Daniel Keysers, Thomas Breuel</i>	
Perceptive Vision for Headline Localisation in Bangla Handwritten Text Recognition	609
<i>Aurelie Lemaitre, Bidyut Baran Chaudhuri, Bertrand Couasnon</i>	
Segmentation of Text and Graphics from Document Images	614
<i>S. P. Chowdhury, S. Mandal, A. K. Das, Bhabatosh Chanda</i>	
Layout Based Information Extraction from HTML Documents	619
<i>Radek Burget</i>	
An Overview of the Tesseract OCR Engine	624
<i>Ray Smith</i>	
Symbol Recognition using a 2-class hierarchical model of Choquet Integrals	629
<i>Jan Rendek, Laurent Wendling</i>	
Deriving Symbol Dependent Edit Weights for Text Correction - the Use of Error Dictionaries	634
<i>Christoph Ringlstetter, Ulrich Reffle, Annette Gotscharek, Klaus U. Schulz</i>	
Arabic Handwriting Recognition Using Variable Duration HMM	639
<i>Amlan Kundu, Tom Hines, Jon Phillips, Benjamin D. Huyck, Linda C. Van Guilder</i>	

VOLUME 2

Curvelets Based Queries for CBIR Application in Handwriting Collections	644
<i>G. Joutel, V. Eglin, S. Bres, H. Emptoz</i>	
Two Approaches for Arabic Script Recognition-based Segmentation Using the Hough Transform	649
<i>Sofène Touj, Najoua Essoukri Ben Amara, Hamid Amiri</i>	
A General Method of Segmentation-Recognition Collaboration Applied to Pairs of Touching and Overlapping Symbols	654
<i>Christophe Renaudin, Yann Ricquebourg, Jean Camillerapp</i>	
Extracting “Carbon Copy” Names and Organizations From a Heterogeneous Document Collection	659
<i>Kazem Taghva, Russell Beckley, Jeffrey Coombs</i>	
An Analytic Word Recognition Algorithm Using a Posteriori Probability	664
<i>Tomoyuki Hamamura, Takuma Akagi, Bunpei Irie</i>	
Predictive DP Matching for On-Line Character Recognition.....	669
<i>Daiki Baba, Seiichi Uchida, Hiroaki Sakoe</i>	
Matching Local Descriptors for Image Identification on Cultural Databases.....	674
<i>Eduardo Valle, Matthieu Cord, Sylvie Philipp-Foliguet</i>	
Fast and Accurate Detection of Document Skew and Orientation	679
<i>Shijian Lu, Jie Wang, Chew Lim Tan</i>	
Information Management System Using Structure Analysis of Paper/Electronic Documents and Its Applications	684
<i>Minenobu Seki, Masakazu Fujio, Takeshi Nagasaki, Hiroshi Shinjo, Katsumi Marukawa</i>	
Mathpad: A Fuzzy Logic-Based Recognition System for Handwritten Mathematics	689
<i>John A. Fitzgerald, Franz Geiselbrechtinger, Tahar Kechadi</i>	
Patra: A Novel Document Architecture for Integrating Handwriting with Audio-visual Information.....	694
<i>Gaurav Harit, Vinod Mankar, Santanu Chaudhury</i>	
Integrated Segmentation and Recognition of Mixed Chinese/English Document	699
<i>Yong Xia, Bai-Hua Xiao, Chun-Heng Wang, Ru-Wei Dai</i>	
Appearance Based Models in Document Script Identification.....	704
<i>T. N. Vikram, D. S. Guru</i>	
Logical Structure Analysis for Form Images with Arbitrary Layout by Belief Propagation.....	709
<i>Akihiro Minagawa, Yusaku Fujii, Hiroaki Takebe, Katsuhito Fujimoto</i>	
An Automatic Mathematical Expression Understanding System.....	714
<i>Yu-sheng Guo, Lei Huang, Chang-ping Liu, Xin Jiang</i>	
Modified Dynamic Time Warping for Stroke-based On-line Signature Verification.....	719
<i>Won-Du Chang, Jungpil Shin</i>	
On the Use of Context-Dependent Modeling Units for HMM-Based Offline Handwriting Recognition	724
<i>Gernot A. Fink, Thomas Plotz</i>	
Off-line Signature Verification Using Enhanced Modified Direction Features in Conjunction with Neural Classifiers and Support Vector Machines.....	729
<i>Vu Nguyen, Michael Blumenstein, Vallipuram Muthukumarasamy, Graham Leedham</i>	

A Multi-Agent System for Hand-drawn Diagram Recognition	734
<i>Giovanni Casella, Vincenzo Deufemia, Viviana Mascardi</i>	
Enhancement of Old Manuscript Images.....	739
<i>Aveakash Gupta, Sunil Kumar, Rajat Gupta, Santanu Chaudhury, Shiv Dutt Joshi</i>	
Handwritten Numeral Recognition of Six Popular Indian Scripts	744
<i>U. Pal, T. Wakabayashi, N. Sharma, F. Kimura</i>	
Text Input System using Online Overlapped Handwriting Recognition for Mobile Devices.....	749
<i>Yojiro Tonouchi, Akinori Kawamura</i>	
Machine Dating of Handwritten Manuscripts.....	754
<i>Utpal Garain, S. K Parui, T. Paquet, L. Heutte</i>	
Off-line Handwritten Signature GPDS-960 Corpus.....	759
<i>J. Francisco Vargas, Miguel A. Ferrer, Carlos M. Travieso, Jesús B. Alonso</i>	
Text-Independent Writer Identification and Verification on Offline Arabic Handwriting	764
<i>Marius Bulacu, Lambert Schomaker, Axel Brink</i>	
A New Physically Motivated Warping Model for Form Drop-Out	769
<i>Guy Rosman, Asaf Tzadok, Doron Tal</i>	
Global Contour Lines Reconstruction in Topographic Maps	774
<i>Joachim Pouderoux, Salvatore Spinello</i>	
Combination of OCR Engines for Page Segmentation Based on Performance Evaluation	779
<i>Miquel Ferrer, Ernest Valveny</i>	
An Indexing Method for Graphical Documents	784
<i>S. Tabbone, D. Zuwala</i>	
Document Processing Via Trained Morphological Operators	789
<i>Nina S. T. Hirata</i>	
Content-level Annotation of Large Collection of Printed Document Images	794
<i>C. V. Jawahar, Anand Kumar</i>	
Bibliographic Attributes Extraction with Layer-upon-Layer Tagging	799
<i>Wei Wei, Irwin King, Jimmy Ho-Man Lee</i>	
Guido: a Musical Score Recognition System	804
<i>Mariusz Szwoch</i>	
A Coupled Mean Shift-Anisotropic Diffusion Approach for Document Image Segmentation and Restoration	809
<i>Fadoua Drira, Frank LeBourgeois, Hubert Emptoz</i>	

KEYNOTE III

Google Book Search: Document Understanding on a Massive Scale	814
<i>Luc Vincent</i>	

SESSION E1: WRITER IDENTIFICATION AND VERIFICATION

Towards Explainable Writer Verification and Identification Using Vantage Writers	819
<i>Axel Brink, Lambert Schomaker, Marius Bulacu</i>	

A New Method for Writer Identification and Verification Based on Farsi/Arabic Handwritten Texts	824
<i>F. Shahabi nejad, M. Rahmati</i>	
Off-line Signature Verification System Performance against Image Acquisition Resolution	829
<i>J. Francisco Vargas, Miguel A. Ferrer, Carlos M. Travieso, Jesús B. Alonso</i>	
Writer Identification using Steered Hermite Features and SVM	834
<i>Asim Imdad, Stephane Bres, Veronique Eglin, Hubert Emptoz, Carlos Rivero-Moreno</i>	
On Computing Strength of Evidence for Writer Verification	839
<i>Harish Srinivasan, Shrivardhan Kabra, Chen Huang, Sargur Srihari</i>	

SESSION E2: SYMBOL RECOGNITION

Mathematical Formula Recognition using 2D Grammars	844
<i>Daniel Prusa, Vaclav Hlavac</i>	
A Unified Framework for Symbol Segmentation and Recognition of Handwritten Mathematical Expressions	849
<i>Yu Shi, Hai Yang Li, Frank K. Soong</i>	
Hybrid Mathematical Symbol Recognition using Support Vector Machines	854
<i>Birendra Keshari, Stephen M. Watt</i>	
Automatic Document Logo Detection	859
<i>Guangyu Zhu, David Doermann</i>	
A Bayesian Network Approach to Mode Detection for Interactive Maps	864
<i>Don J.M. Willems, Louis. G. Vuurpijl</i>	

SESSION E3: CAMERA AND VIDEO APPLICATIONS

Robust Binarization for Video Text Recognition	869
<i>Zohra Saidane, Christophe Garcia</i>	
A Parametric Spectral-based Method for Verification of Text in Videos	874
<i>Vassilis Papavassiliou, Themis Stafylakis, Vassilis Katsouros, George Carayannis</i>	
Camera-Based Graphical Symbol Detection	879
<i>Marçal Rusinol, Josep Lladós, Phiippe Dosh</i>	
A Weighted Finite-State Framework for Correcting Errors in Natural Scene OCR	884
<i>Richard Beaufort, Celine Mancas-Thillou</i>	
Exploiting Videotext “Events” for Improved Videotext Detection	889
<i>Hrishikesh B. Aradhya, Gregory K. Myers</i>	

SESSION F1: CHINESE HANDWRITING RECOGNITION

Skew Detection for Chinese Handwriting by Horizontal Stroke Histogram	894
<i>Tong-Hua Su, Tian-Wen Zhang, Hu-Jie Huang, Yu Zhou</i>	
LDA-Based Compound Distance for Handwritten Chinese Character Recognition	899
<i>Tian-Fu Gao, Cheng-Lin Liu</i>	

Building Compact MQDF Classifier for Off-line Handwritten Chinese Characters by Subspace Distribution Sharing	904
<i>Teng Long, Lian-Wen Jin</i>	

Handwritten Chinese Character Recognition Using Modified LDA and Kernel FDA	909
<i>Duanduan Yang, Lianwen Jin</i>	

SESSION F2: LAYOUT ANALYSIS

Example-Based Logical Labeling of Document Title Page Images	914
<i>Joost van Beusekom, Daniel Keysers, Faisal Shafait, Thomas M. Breuel</i>	

A Data Mining Approach to Reading Order Detection	919
<i>Michelangelo Ceci, Margherita Berardi, Giuseppe A. Porcelli, Donato Malerba</i>	

A General Approach for Partitioning Web Page Content Based on Geometric and Style Information	924
<i>Hui Guo, Jalal Mahmud, Yevgen Borodin, Amanda Stent, I.V. Ramakrishnan</i>	

Searching for Tables in Digital Documents	929
<i>Ying Liu, Kun Bai, Prasenjit Mitra, C. Lee Giles</i>	

SESSION F3: APPLICATIONS

A Pen-based Interface for Real-time Document Edition	934
<i>Jose A. Rodriguez, Gemma Sanchez, Josep Lladós</i>	

Computer Assisted Transcription of Handwritten Text Images	939
<i>Alejandro H. Toselli, Veronica Romero, Enrique Vidal, Luis Rodriguez</i>	

A Method for Image Local-difference Visualization	944
<i>Etienne Baudrier, Alain Riffaud</i>	

iGesture: A General Gesture Recognition Framework	949
<i>Beat Signer, Ueli Kurmann, Moira C. Norrie</i>	

SESSION G1: ARABIC RECOGNITION

Combination of HMM-Based Classifiers for the Recognition of Arabic Handwritten Words	954
<i>Ramy Al-Hajj, Chafic Mokbel, Laurence Likforman-Sulem</i>	

A Hybrid Approach for Off-line Arabic Handwriting Recognition Based on a Planar Hidden Markov Modeling	959
<i>Sameh Masmoudi Touj, Najoua Essoukri Ben Amara, Hamid Amiri</i>	

Shape-based Alphabet for Off-line Arabic Handwriting Recognition	964
<i>F. Menasri, N. Vincent, E. Augustin, M. Cheriet</i>	

Comparison of Different Preprocessing and Feature Extraction Methods for Offline Recognition of Handwritten Arabic Words	969
<i>Haikal El Abed, Volker Margner</i>	

SESSION G2: CAMERA-BASED SYSTEMS

Fast Lexicon-Based Scene Text Recognition with Sparse Belief Propagation	974
<i>Jerod J. Weinman, Erik Learned-Miller, Allen Hanson</i>	
Removing Shading Distortions in Camera-based Document Images Using Inpainting and Surface Fitting With Radial Basis Functions.....	979
<i>Li Zhang, Andy M. Yip, Chew Lim Tan</i>	
Segmentation Based Recovery of Arbitrarily Warped Document Images.....	984
<i>B. Gatos, I. Pratikakis, K. Ntirogiannis</i>	
Keyword Spotting and Retrieval of Document Images Captured by a Digital Camera	989
<i>Shijian Lu, Chew Lim Tan</i>	

SESSION G3: APPLICATIONS (2)

Spoken Handwriting Verification using Statistical Models.....	994
<i>Andreas Humm, Rolf Ingold, Jean Hennebert</i>	
An Embedded OCR Software Architecture for Enhancing Portability	999
<i>Seho Kim, Jaehwa Park, Young-Bin Kwon</i>	
Generating Copybooks from Consistent Handwriting Styles	1004
<i>Ralph Niels, Louis Vuurpijl</i>	

POSTER SESSION III

A Feature based on Encoding the Relative Position of a Point in the Character for Online Handwritten Character Recognition	1009
<i>Dinesh M, Murali Krishna Sridhar</i>	
Implementation of Content Analysis System for Recognition of Journals' Table of Contents	1013
<i>Young-Bin Kwon, Jaehwa Park</i>	
Preliminary Experiments in Layout Analysis of Handwritten Letters Based on Textural and Spatial Information and a 2D Markovian Approach	1018
<i>Melanie Lema tre, Emmanuele Grosicki, Edouard Geoffrois, Francoise Preteux</i>	
Elastic Matching of Online Handwritten Tamil and Telugu Scripts Using Local Features.....	1023
<i>L.Prasanth,V.Jagadeesh Babu, R. Raghunath Sharma, G.V.Prabhakara Rao, M. Dinesh</i>	
Usage-oriented Performance Evaluation for Text Localization Algorithms.....	1028
<i>Yichao Ma, Chunheng Wang, Baihua Xiao, Ruwei Dai</i>	
An Effective and Practical Classifier Fusion Strategy for Improving Handwritten Character Recognition.....	1033
<i>Qiang Fu, X.Q.Ding, T.Zh.Li, C.S.Liu</i>	
Using Random Forests for Handwritten Digit Recognition	1038
<i>Simon Bernard, Laurent Heutte, Sebastien Adam</i>	
An approach for Multifont Arabic characters features Extraction based on Contourlet Transform.....	1043
<i>Nadia Ben Amor , Najoua Essoukri Ben Amara</i>	
A Proposition of Retrieval tools for Historical Document Images libraries.....	1048
<i>Nicholas Journet, Remy Mullot, Veronique Eglin, Jean-Yves Ramel</i>	

PRAAD: Preprocessing and Analysis Tool for Arabic Ancient Documents	1053
<i>Wafa Boussellaa, Abderrazak Zahour, Bruno Taconet, Adel Alimi, Abdellatif Benabdelhafid</i>	
The hOCR Microformat for OCR Workflow and Results	1058
<i>Thomas M. Breuel</i>	
OCR Accuracy Improvement through a PDE-Based Approach	1063
<i>Fadoua Drira, Frank LeBourgeois, Hubert Emptoz</i>	
An Efficient Feature Extraction and Dimensionality Reduction Scheme for Isolated Greek Handwritten Character Recognition	1068
<i>G.Vamvakas, B.Gatos, S. Petridis, N.Stamatopoulos</i>	
Using Grammar-based Recognizers for Symbol Completion in Diagrammatic Sketches	1073
<i>Gennaro Costagliola, Vincenzo Deufemia, Michele Risi</i>	
A Visualization Tool to Improve the Performance of a Classifier Based on Hidden Markov Models	1078
<i>Gleidson Pegoretti da Silva, Masaki Nakagawa</i>	
On the Use of Lexeme Features for Writer Verification	1083
<i>Anurag Bhardwaj, Abhishek Singh, Harish Srinivasan, Sargur Srihari</i>	
Incremental Learning of First Order Logic Theories for the Automatic Annotations of Web Documents	1088
<i>Floriana Esposito, Stefano Ferilli, Nicola Di Mauro, Teresa M.A. Basile</i>	
Handwritten Word Recognition using Conditional Random Fields	1093
<i>Shravya Shetty, Harish Srinivasan, Sargur Srihari</i>	
Three Decision Levels Strategy for Arabic and Latin Texts Differentiation in Printed and Handwritten Natures	1098
<i>Mohamed Ben Jlaïel, Slim Kanoun, Adel Mohamed Alimi, Rémy Mullot</i>	
Neural Network for the Recognition of Handwritten Tunisian City Names	1103
<i>Imen Ben Cheikh, Afef Kacem</i>	
On the effects of sampling rate and interpolation in HMM-based dynamic signature verification	1108
<i>M. Martinez-Diaz, J. Fierrez, M. R. Freire, J. Ortega-Garcia</i>	
Skeletonization by Gradient Regularization and Diffusion	1113
<i>F. Le Bourgeois, H. Emptoz</i>	
Use of Legal Amount to Confirm or Correct the Courtesy Amount on Farsi Bank Checks	1118
<i>Majid Ziaratban, Karim Faez, Mehdi Ezoji</i>	
A colour document interpretation: Application to ancient cadastral maps	1123
<i>Romain Raveaux, Jean-Christophe Burie, Jean-Marc Ogier</i>	
Arabic Handwriting Texture Analysis for Writer Identification Using the DWT-Lifting Scheme	1128
<i>Sami Gazzah, Najoua Essoukri Ben Amara</i>	
Recognition, Understanding and Aestheticization of Freehand Drawing Flowcharts	1133
<i>Wioleta Szwoch</i>	
Table Recognition and Understanding from PDF Files	1138
<i>Tamir Hassan, Robert Baumgartner</i>	
Ink Discrimination based on Co-occurrence Analysis of Visible and Infrared Images	1143
<i>Vasiliki Kokla, Alexandra Psarrou, Vassilis Konstantinou</i>	
Segmentation of Very Low Resolution Screen-Rendered Text	1148
<i>Steffen Wachenfeld, Hans-Ulrich Klein, Stefan Fleischer, Xiaoyi Jiang</i>	

Quantile Linear Algorithm for Robust Binarization of Digitalized Letters	1153
<i>Marte Ram rez, Ernesto Tapia, Marco Block, Raul Rojas</i>	
A Shared Parts Model for Document Image Recognition	1158
<i>Mithun Das Gupta, Prateek Sarkar</i>	
Context-Sensitive Error Correction: Using Topic Models to Improve OCR	1163
<i>Michael L. Wick, Michael G. Ross, Erik G. Learned-Miller</i>	
On the Applicability of Off-line Signatures to the Fuzzy Vault Construction	1168
<i>Manuel R. Freire, Julian Fierrez, Marcos Martinez-Diaz, Javier Ortega-Garcia</i>	
Autonomous Text Capturing Robot Using Improved DCT Feature and Text Tracking	1173
<i>Makoto Tanaka, Hideaki Goto</i>	
Aspects of Mathematical Expression Analysis in Arabic Handwriting	1178
<i>Elena Smirnova, Stephen M. Watt</i>	
Online Handwritten Kanji Recognition Based on Inter-stroke Grammar	1183
<i>Kumi Ota, Ryo Yamamoto, Shinji Sako, Shigeki Sagayama</i>	
Physical Structure Segmentation with Projection Profile for Mathematic Formulae and Graphics in Academic Paper Images	1188
<i>Tzu-Yuan Chang, Yusuke Takiguchi, Minoru Okada</i>	
Representing and Characterizing Handwritten Mathematical Symbols through Succinct Functional Approximation	1193
<i>Bruce W. Char, Stephen M. Watt</i>	
Extracting Features of Paper-made Objects Recognized from Origami Books Based on Design Knowledge	1198
<i>Hiroshi Shimanuki, Jien Kato, Toyohide Watanabe</i>	
Word Image Based Latent Semantic Indexing for Conceptual Querying in Document Image Databases	1203
<i>Sameek Banerjee, Gaurav Harit, Santanu Chaudhury</i>	
A Sparse and Locally Shift Invariant Feature Extractor Applied to Document Images	1208
<i>Marc'Aurelio Ranzato, Yann LeCun</i>	
A Novel Hierarchical Classification Scheme for Online Tamil Character Recognition	1213
<i>Suresh Sundaram, A. G. Ramakrishnan</i>	
Extraction of Arabic Words from Complex Color Image	1218
<i>Radwa Fathalla, Yasser El Sonbaty, Mohamed A. Ismail</i>	
Image Pixel Force Fields and Their Application for Color Map Vectorisation	1223
<i>V. Bucha, S. Uchida, S. Ablameyko</i>	
New Strategy for the on Line Handwriting Modeling	1228
<i>Houcine Boubaker, Monji Kherallah, Adel M. Alimi</i>	
On Using Classical Poetry Structure for Indian Language Post-Processing	1233
<i>Anoop M. Namboodiri, P. J. Narayanan, C. V. Jawahar</i>	
On Segmentation of Documents in Complex Scripts	1238
<i>K. S. Sesh Kumar, Sukesh Kumar, C. V. Jawahar</i>	
A Blind Indic Script Recognizer for Multi-script Documents	1243
<i>Peeta Basa Pati, A G Ramakrishnan</i>	
Streaming-Archival InkML Conversion	1248
<i>Birendra Keshari, Stephen M. Watt</i>	

Performance Analysis Framework for Layout Analysis Methods	1253
<i>A. Antonacopoulos, D. Bridson</i>	
Unsupervised Newspaper Segmentation Using Language Context	1258
<i>Ralph Furmaniak</i>	

KEYNOTE IV

Advances in Writer Identification and Verification	1263
<i>Lambert Schomaker</i>	

COMPETITION SESSION

ICDAR 2007 - Arabic Handwriting Recognition Competition	1269
<i>Volker Margner, Haikal El Abed</i>	
ICDAR2007 Page Segmentation Competition	1274
<i>A. Antonacopoulos, B. Gatos, D. Bridson</i>	
ICDAR2007 Handwriting Segmentation Contest	1279
<i>B. Gatos, A. Antonacopoulos, N. Stamatopoulos</i>	

Author Index