

# ICASSP-94

VOLUME 5

## I

### IMAGE AND MULTIDIMENSIONAL SIGNAL PROCESSING

1994  
IEEE INTERNATIONAL CONFERENCE  
ON  
ACOUSTICS, SPEECH AND  
SIGNAL PROCESSING

APRIL 19-22, 1994  
ADELAIDE CONVENTION CENTRE  
ADELAIDE, SOUTH AUSTRALIA



IEEE

94CH3387-8

Sponsored by  
THE INSTITUTE OF ELECTRICAL AND  
ELECTRONICS ENGINEERS,  
SIGNAL PROCESSING SOCIETY

## ICASSP-94 Proceedings in 6 Volumes:

Went  
TA  
365  
I102  
1994  
5

|           |                     |                                                                                 |
|-----------|---------------------|---------------------------------------------------------------------------------|
| Volume 1. | S <sub>1</sub>      | Speech Processing 1                                                             |
| Volume 2. | S <sub>2</sub> AUVN | Speech Processing 2<br>Audio<br>Underwater Acoustics<br>VLSI<br>Neural Networks |
| Volume 3. | D                   | Digital Signal Processing                                                       |
| Volume 4. | S                   | Statistical Signal and Array Processing                                         |
| Volume 5. | I                   | Image and Multidimensional Signal Processing                                    |
| Volume 6. | P                   | Plenary and Special Sessions                                                    |

Cover Design: David Jones Art and Design, Adelaide, South Australia  
Typesetting: Musictype Pty. Ltd., Adelaide, South Australia

Additional copies may be ordered from:

IEEE Service Center  
445 Hoes Lane  
P.O. Box 1331  
Piscataway, NJ 08855-1331

Copyright and Reprint Permission: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 27 Congress Street, Salem, MA 01970. For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, P.O. Box 1331, Piscataway, NJ 08855-1331. All rights reserved. Copyright © 1994 by the Institute of Electrical and Electronics Engineers, Inc.

IEEE Catalog No. 94CH3387-8  
ISBN 0-7803-1775-0 (softbound)  
ISBN 0-7803-1776-9 (casebound)  
ISBN 0-7803-1777-7 (microfiche)  
Library of Congress Catalog Card No. 84-645139 (serial)

# VOLUME 5

## IMAGE & MULTIDIMENSIONAL SIGNAL PROCESSING

### Table of Contents

#### IMAGE ANALYSIS I

Chairperson: Ed Delp, *Purdue University, (USA)*

- Unsupervised Segmentation of Radar Images Using Wavelet Decomposition and Cumulants** V-1  
Jean-Marc Boucher, Stéphane Pleihers, *ENSTB, (FRANCE)*
- Morphological Scale-Space Fingerprints and their Use in Object Recognition in Range Images** V-5  
Paul Jackway, Wageeh Boles, Mohamed Deriche, *Signal Processing Research Centre, Queensland University of Technology, (AUSTRALIA)*
- Detection of Point Targets in High Resolution Synthetic Aperture Radar Images** V-9  
Ying Wang, Rama Chellappa, Qinfen Zheng, *University of Maryland, (USA)*
- An Energy Minimization Approach to Building Detection in Aerial Images** V-13  
Santhana Krishnamachari, Rama Chellappa, *University of Maryland, (USA)*
- Use of M-Band Wavelet Transform for Multidirectional and Multiscale Edge Detection** V-17  
Turgut Aydin, Yücel Yemez, Bülent Sankur, Emin Anarim, Oktay Alkin, *Boğaziçi Üniversitesi, (TURKEY)*
- A Wavelet Based Mammographic System** V-21  
Andrew Laine, *University of Florida*, Mike Lewis, *The Athena Group*, Fred J Taylor, *University of Florida, (USA)*
- Orientation Selective Operators for Ridge, Valley, Edge and Line Detection in Imagery** V-25  
Jianxin Hou, Roberto H Bamberger, *Washington State University, (USA)*
- Periodicity Estimation in Textured Images Using a ML Approach** V-29  
Jorge Marques, *INESC/Technical University of Lisbon, (PORTUGAL)*
- Detecting Scene Changes and Activities in Video Databases** V-33  
Robert Hsu, Hiroshi Harashima, *University of Tokyo, (JAPAN)*
- On Vector Quantization for Fast Facet Edge Detection** V-37  
M Y Jaisimha, Jill R Goldschneider, Alexander E Mohr, Eve A Riskin, Robert M Haralick, *University of Washington, Seattle, (USA)*
- A Contour-Based Part Segmentation Algorithm** V-41  
Mohammed Bennamoun, *Signal Processing Research Centre, Queensland University of Technology, (AUSTRALIA)*

#### IMAGE FILTERING AND ENHANCEMENT

Chairperson: P Combettes, *City University of New York (USA)*

- An Image Processing Algorithm for a Super High Definition Imaging Scheme with Multiple Different-Aperture Cameras** V-45  
Takahiro Saito, Takashi Komatsu, *Kanagawa University*, Kiyoharu Aizawa, *The University of Tokyo, (JAPAN)*

- An Improvement on the Reception of PAL-M Television Signal Using an Additional Simple Delay Filter** V-49  
Yuzo Iano, *University of Campinas, (BRAZIL)*

- Image Visual Quality Restoration by Cancellation of the Unmasked Noise** V-53  
Benoit Macq, *Université Catholique de Louvain, (BELGIUM)*; Marco Mattavelli, *Inst. Polyt. Fed. de Lausanne, (SWITZERLAND)*; Olivier Van Calster, Emmanuel van der Plancke, Serge Comes, *Université Catholique de Louvain (BELGIUM)*; Wei Li, *Inst. Polyt. Fed. de Lausanne (SWITZERLAND)*

- Processing Incomplete and Uncertain Data Using Subspace Filtering** ~~WITHDRAWN~~  
Carl-Fredrik Walin, Lars Rindsson, *Linköping University, (SWEDEN)*

- A 2-D IIR Neural Hybrid Filter for Image Processing** V-57  
Mitsuji Muneyasu, Satoshi Tsujii, Takao Hinamoto, *Hiroshima University, (JAPAN)*

- Adaptive  $\alpha$ -Trimmed Mean Filters with Excellent Detail-Preserving** V-61  
Akira Taguchi, *Musashi Institute of Technology, (JAPAN)*

- Design of Optimal Median-Type Filters under Structural Constraints** V-65  
Bing Zeng, *Hong Kong University of Science & Technology, (HONG KONG)*

- Statistical Analysis of the Median Based Multi-Shell Order-Statistics Filters** V-69  
J S Jimmy Li, Anand Ramsingh, *University of Canterbury, (NEW ZEALAND)*

- GLMOS and a Comparative Study of Nonlinear Filters** V-73  
Hamid Rabiee, R L Kashyap, *Purdue University, (USA)*

- Statistical Morphological Filters for Binary Image Processing** V-77  
Carlo Regazzoni, *University of Genova, (ITALY)*; Anastasios N Venetsanopoulos, *University of Toronto, (CANADA)*; Gianluca Foresti, Gianni Vernazza, *University of Genova, (ITALY)*

- A Human-Machine Interactive System for Efficient Image Restoration** V-81  
Hong Tang, *Australian National University (AUSTRALIA)*

- Blind Superresolving Image Recovery from Blur-Invariant Edges** V-85  
Kazuki Nishi, *University of Electro-Communications*, Shigeru Ando, *University of Tokyo, (JAPAN)*

#### IMAGE DISPLAY

Chairperson: Ping Wah Wong, *Hewlett-Packard Laboratories, (USA)*

- Colour Quantization of Images Based on Human Vision Perception** V-89  
Navin Chaddha, Wee-Chiew Tan, Teresa H Y Meng, *Stanford University, (USA)*

|                                                                                                                                |              |
|--------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>Colour Quantization Error in Terms of Perceived Image Quality</b>                                                           | <b>V-93</b>  |
| Alain Tremeau, <i>URA 842 Lab TSI, Maurice Calonnier, Institut Textile de France, Bernard Laget, URA 842 Lab TSI, (FRANCE)</i> |              |
| <b>Greedy Tree Growing for Colour Image Quantization</b>                                                                       | <b>V-97</b>  |
| Tsann-Shyong Liu, Long-Wen Chang, <i>Tsing Hua University, (TAIWAN ROC)</i>                                                    |              |
| <b>Designing the Colour Palette for Textile Materials Printing Process</b>                                                     | <b>V-101</b> |
| Gabriel Marcu, Kansei Iwata, <i>Graphica Computer Corporation, (JAPAN)</i>                                                     |              |
| <b>Halftoning Technique Using Genetic Algorithm</b>                                                                            | <b>V-105</b> |
| Naoki Kobayashi, Hideo Saito, <i>Keio University, (JAPAN)</i>                                                                  |              |
| <b>Conversion of Scanned Documents to the Open Document Architecture</b>                                                       | <b>V-109</b> |
| Gary Farrow, Costas Xydeas, John Oakley, <i>Manchester University, (UK)</i>                                                    |              |
| <b>Error Diffusion with Dynamically Adjusted Kernel</b>                                                                        | <b>V-113</b> |
| Ping Wah Wong, <i>Hewlett-Packard Laboratories, (USA)</i>                                                                      |              |
| <b>A Distortion Measure for Image Artifacts Based on Human Visual Sensitivity</b>                                              | <b>V-117</b> |
| Shanika Karunasekera, Nick Kingsbury, <i>Cambridge University, (UK)</i>                                                        |              |
| <b>Application of Directional Statistics in Vector Direction Estimation</b>                                                    | <b>V-121</b> |
| Nikos Nikolaidis, Ioannis Pitas, <i>University of Thessaloniki, (GREECE)</i>                                                   |              |
| <b>Generating Non-Gaussian Random Fields for Sea-Surface Simulations</b>                                                       | <b>*</b>     |
| Garry N Newsam, Michael Wegener, <i>Defence Science Technology Organisation, (AUSTRALIA)</i>                                   |              |
| <b>IMAGE ANALYSIS II</b>                                                                                                       |              |
| <b>Chairperson:</b> Rosalind Picard, <i>MIT Media Laboratory, (USA)</i>                                                        |              |
| <b>On-Line Cursive Handwriting Recognition Using Speech Recognition Methods</b>                                                | <b>V-125</b> |
| Thad Starner, <i>MIT Media Lab, John Makhoul, Richard Schwartz, George Chou, Bolt Beranek and Newman Inc, (USA)</i>            |              |
| <b>A New Word Ordering for Image Similarity</b>                                                                                | <b>V-129</b> |
| Rosalind W Picard, Fang Liu, <i>MIT Media Laboratory, (USA)</i>                                                                |              |
| <b>Filter Estimation Maximization Algorithm for Image Segmentation</b>                                                         | <b>V-133</b> |
| Hocine Cherifi, <i>TSI (CNRS 842), Richard Grisel, ICPI - L2S2, (FRANCE)</i>                                                   |              |
| <b>Recognition of Space Curves Based on the Dyadic Wavelet Transform</b>                                                       | <b>V-137</b> |
| Quang Minh Tieng, Wageeh Boles, <i>Signal Processing Research Centre, Queensland University of Technology, (AUSTRALIA)</i>     |              |
| <b>Modelling and Classification of Shapes in Two-Dimensions Using Vector Quantization</b>                                      | <b>V-141</b> |
| Simon Lee, Brian Lovell, <i>University of Queensland/CSSIP, (AUSTRALIA)</i>                                                    |              |
| <b>Heuristic Image Decoding Using Separable Source Models</b>                                                                  | <b>V-145</b> |
| Anthony C Kam, <i>Caliper Corp, USA; Gary E Kopec, Xerox PARC, (USA)</i>                                                       |              |
| <b>Supervised Hidden Markov Modelling for On-line Handwriting Recognition</b>                                                  | <b>V-149</b> |
| Jerome R Bellegarda, David Nahamoo, Krishna Nathan, Eveline Bellegarda, <i>IBM T J Watson Research Centre, (USA)</i>           |              |
| <b>Correlation Filters for Texture Recognition and Applications to Terrain-Delimitation in Wide-Area Surveillance</b>          | <b>V-153</b> |
| Hemant Singh, <i>Aware Inc, Abhijit Mahalanobis, Hughes Missile Systems Co., (USA)</i>                                         |              |

## COMPUTED IMAGING I: SYNTHETIC APERTURE

**Chairperson:** Mehrdad Soumekh,  
*State University of New York at Buffalo, (USA)*

|                                                                                                               |              |
|---------------------------------------------------------------------------------------------------------------|--------------|
| <b>Optimal Configuration and Weighting of Nonuniform Arrays According to a Maximum ISLR Criterion</b>         | <b>V-157</b> |
| Carlo Boni, Mario Richard, <i>Alenia S.p.A., Sergio Barbarossa, University of Rome "La Sapienza", (ITALY)</i> |              |

|                                                                                           |              |
|-------------------------------------------------------------------------------------------|--------------|
| <b>Target Shifts Due to Modelling Assumptions in Inverse Synthetic Aperture Radar</b>     | <b>V-161</b> |
| Hyeokho Choi, David C Munson Jr, <i>University of Illinois at Urbana-Champaign, (USA)</i> |              |

|                                                                                                                                                   |              |
|---------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>Comparative Study of Some Algorithms for Terrain Classification Using SAR Images</b>                                                           | <b>V-165</b> |
| Zied Belhadj, <i>S2HF - IRESTE, Ali Saad, LATI - IRESTE, Safwen El Assad, J Saillard, S2HF - IRESTE, Dominique Barba, LATI - IRESTE, (FRANCE)</i> |              |

|                                                                                                                                                              |              |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>Automatic Target Detection in Dynamic Clutter from Incoherent ISAR Data</b>                                                                               | <b>V-169</b> |
| Mehrdad Soumekh, <i>State University of New York at Buffalo, Michael Pollock, Robert Dinger, Naval Command, Control and Ocean Surveillance Center, (USA)</i> |              |

|                                                                                                     |              |
|-----------------------------------------------------------------------------------------------------|--------------|
| <b>Imaging of Multitargets with ISAR Based on the Time-Frequency Distribution</b>                   | <b>V-173</b> |
| Aiyuan Wang, Yinfang Mao, Zongzhi Chen, <i>Institute of Electronics Academia Sinica, (PR CHINA)</i> |              |

|                                                                                                                      |              |
|----------------------------------------------------------------------------------------------------------------------|--------------|
| <b>ISARLAB: A Radar Signal Processing Tool</b>                                                                       | <b>V-177</b> |
| Brett Haywood, Anthony Zyweck, Ross Kyprianou, <i>Defence Science and Technology Organisation/CSSIP, (AUSTRALIA)</i> |              |

|                                                                                                                             |              |
|-----------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>A New Methodology for Fourier Synthesis. Fourier Interpolation and Reconstruction via Shannon-type Techniques: FIRST</b> | <b>V-181</b> |
| André Lannes, Éric Anterrieu, Sylvie Roques, Géraldine Fitoussi, <i>CNRS/OMP, (FRANCE)</i>                                  |              |

|                                                                              |              |
|------------------------------------------------------------------------------|--------------|
| <b>Synthetic Aperture Technique Used in Ultrasonic Intravascular Imaging</b> | <b>V-185</b> |
| Jiang Hui, Hou Chao-huan, <i>Academia Sinica, (PR CHINA)</i>                 |              |

|                                                                                                                             |              |
|-----------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>Vector Quantization of Raw SAR Data</b>                                                                                  | <b>V-189</b> |
| Jean-Marie Moureaux, Patricia Gauthier, <i>IS3-CNRS, Michel Barlaud, Pascale Bellemain, Aerospatiale (Cannes), (FRANCE)</i> |              |

|                                                                                                                                              |              |
|----------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>A Real Time Processor for the Australian Synthetic Aperture Radar</b>                                                                     | <b>V-193</b> |
| Nick Stacy, Michael Burgess, J J Douglass, Marshall Muller, Murray Robinson, <i>Defence Science and Technology Organisation, (AUSTRALIA)</i> |              |

|                                                                                                      |              |
|------------------------------------------------------------------------------------------------------|--------------|
| <b>Sign Language Image Processing for an Intelligent Communications by a Communication Satellite</b> | <b>V-197</b> |
| Yoshinao Aoki, Shin Tanahashi, Jun Xu, <i>Hokkaido University, (JAPAN)</i>                           |              |

## MOTION ESTIMATION

**Chairperson:** Rama Chellappa, *University of Maryland, (USA)*

|                                                                           |              |
|---------------------------------------------------------------------------|--------------|
| <b>A New Approach to Motion Estimation for ISAR Imaging</b>               | <b>V-201</b> |
| Stephen Simmons, Robin Evans, <i>University of Melbourne, (AUSTRALIA)</i> |              |

|                                                                                                                             |              |
|-----------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>Motion Compensated Video Sequence Interpolation Using Digital Image Warping</b>                                          | <b>V-205</b> |
| Jacek Nieweglowski, Timo Moisala, <i>Tampere University of Technology, Petri Haavisto, Nokia Research Center, (FINLAND)</i> |              |

**Motion Estimation Using Multiple Image Sensors** V-209  
Kiyoharu Aizawa, Ken-ichi Iwata, *University of Tokyo*, Takahiro Saito, *Kanagawa University*, Mitsutoshi Hatori, *University of Tokyo*, (JAPAN)

**Spanning the Gap Between Motion Estimation and Morphing** V-213  
Michele Covell, Margaret Withgott, *Interval Research Corporation*, (USA)

**A New Technique for Block-Based Motion Compensation** V-217  
Shinichi Koza, *NEC Corporation*, (JAPAN); Sanjeev Kulkarni, *Princeton University*, (USA)

**An Algorithm for Simultaneous Motion Estimation and Scene Segmentation** V-221  
Michael M Chang, *University of Rochester*, M Ibrahim Sezan, *Eastman Kodak Co*, A Murat Tekalp, *University of Rochester*, (USA)

**Digital Video Standards Conversion in the Presence of Accelerated Motion** V-225  
Andrew J Patti, *University of Rochester*, M Ibrahim Sezan, *Eastman Kodak Co*, A Murat Tekalp, *University of Rochester*, (USA)

**Nonuniform Image Motion Estimation Using Kalman Filtering** V-229  
Nader Namazi, Pablo Penafiel, Chieh-Min Fan, *The Catholic University of America*, (USA)

**Intensity Scale Invariant Motion Estimation with Rotation and Spatial Scaling Information** V-233  
Colin Bussiere, Dimitrios Hatzinakos, *University of Toronto*, (CANADA)

**Restoration of Low Bit Rate Compressed Images using Mean Field Annealing** V-237  
James C Brailean, *Motorola*, Taner Özcelik, Aggelos K Katsaggelos, *Northwestern University*, (USA)

**Adaptive Multi-Feature Motion Estimation** V-241  
Regis Crinon, Wojciech Kolodziej, *Oregon State University*, (USA)

**Replacement Noise in Image Sequences - Detection and Interpolation by Motion Field Segmentation** V-245  
Robin Morris, William Fitzgerald, *Cambridge University*, (UK)

#### IMAGE CODING 0 AND QUANTIZATION

**Chairperson:** Tor Ramstad, *Norwegian Institute of Technology*, (NORWAY)

**Minimum Generalised Quadratic Error Quantization for Image and Video Coding** V-249  
John Princen, *AT & T Bell Laboratories*, (USA); Ming H Chan, *Telecom Australia*, (AUSTRALIA)

**Entropy-Constrained Predictive Trellis Coded Quantization: Application to Hyperspectral Image Compression** V-253  
Glen P Abovseman, Michael W Marcellin, Bobby R Hunt, *University of Arizona*, (USA)

**Image Coding Using Adaptive Recursive Interpolative DPCM with Entropy-Constrained Trellis Coded Quantization** V-257  
Eric Gifford, Bobby R Hunt, Michael Marcellin, *University of Arizona*, (USA)

**Deterministic Annealing for Trellis Quantizer and HMM Design Using Baum-Welch Re-Estimation** V-261  
David Miller, Kenneth Rose, *University of California at Santa Barbara*, Philip A Chou, *Xerox Palo Alto Research Center*, (USA)

**Optimal Entropy Constrained Scalar Quantization for Exponential and Laplacian Random Variables** V-265  
Gary J Sullivan, *PictureTel Corp*, (USA)

**Lattice Vector Quantization of Image Wavelet Coefficient Vectors Using a Simplified Form of Entropy Coding** V-269  
Andrew Woolf, Glynn Rogers, *CSIRO Division of Radiophysics*, (AUSTRALIA)

**Object-Oriented Video Coding Employing Dense Motion Fields** V-273  
Christoph Stiller, *Aachen University of Technology*, (GERMANY)

**3D Contour Image Coding Based on Morphological Filters and Motion Estimation** V-277  
Chuang Gu, *Swiss Federal Institute of Technology*, (SWITZERLAND)

**An Analogue Interpretation of Compression for Digital Communication Systems** V-281  
John M Lervik, Tor A Ramstad, *Norwegian Institute of Technology*, (NORWAY)

**A Fast Algorithm for Region-Oriented Texture Coding** V-285  
Marco Cermelli, Fabio Lavagetto, Matteo Pampolini, *DIST - University of Genova*, (ITALY)

**Address Predictive Colour Quantization Image Compression For Multimedia Applications** V-289  
Lai-Man Po, Wen-Tao Tan, Chi-Ho Chan, *City Polytechnic of Hong Kong*, (HONG KONG)

#### COMPUTED TOMOGRAPHY

**Chairperson:** David Munson, *University of Illinois*, (USA)

**Genetic Algorithms for Neuromagnetic Source Reconstruction** V-293  
Paul S Lewis, John Mosher, *Los Alamos National Laboratory*, (USA)

**Image Reconstruction of Contour Data Using a Backpropagation Neural Network** V-297  
Karim Faez, *Amirkabir University of Technology*, (IRAN); Mohamed Kamel, *University of Waterloo*, (CANADA)

**Sampling of Two-Dimensional Signals Below Nyquist Density with Applications to Computer Aided Tomography** V-301  
Kai-Kou Roger Yu, Sze-Fong Yau, *Hong Kong University of Science and Technology*, (HONG KONG)

**A Fast Tomographic Reconstruction Algorithm in the 2-D Wavelet Transform Domain** V-305  
Laure Blanc-Féraud, Pierre Charbonnier, Pierre Lobel, Michel Barlaud, *Université de Nice-Sophia Antipolis*, (FRANCE)

**Tomographic Reconstruction of Time-Varying Object from Linear Time-Sequential Sampled Projections** V-309  
Ying Ha Chiu, Sze Fong Yau, *Hong Kong University of Science & Technology*, (HONG KONG)

**The Reconstruction of Subsurface Property Maps using Projection onto Convex Sets** V-313  
Alberto Malinverno, David Rossi, *Schlumberger-Doll Research*, Michael Daniel, *MIT*, (USA)

**Simultaneous Confidence Intervals for Image Reconstruction Problems** V-317  
Yong Zhang, Alfred O Hero III, W L Rogers, *University of Michigan*, (USA)

**Iterative Reconstruction of Multidimensional Objects Buried in Inhomogeneous Elastic Media** V-321  
Tarek Habashy, *Schlumberger-Doll Research*, Eveline Bellegarda, *IBM T J Watson Research Center*, (USA)

#### IMAGE CODING I

**Chairperson:** Henrik Sorensen, *University of Pennsylvania*, (USA)

**Discrete Multichannel Orthogonal Transforms** V-325  
Ioannis Pitas, Anestis Karasiris, *University of Thessaloniki*, (GREECE)

**An Edge Classification Based Approach to the Post-Processing of Transform Coded Images** V-329  
John McDonnell, Robert Shorten, Anthony D Fagan, *University College Dublin, (IRELAND)*

**Block Predictive Transform Coding of Still Images** V-333  
Jianzhong Huang, *Georgia Institute of Technology*, Sam Liu, *Hewlett-Packard Laboratories, (USA)*

**Image Coding with Discrete Cosine Transforms using Efficient Energy-Based Adaptive Zonal Filtering** V-337  
Alessandro Palau, Gagan Mirchandani, *University of Vermont, (USA)*

**A New Method for Block Effect Removal in Low Bit-Rate Image Compression** V-341  
Jiebo Luo, Chang Wen Chen, Kevin J Parker, *University of Rochester*, Thomas S Huang, *University of Illinois at Urbana-Champaign, (USA)*

**Fast Segmented Image Coding using Weakly Separable Bases** V-345  
Wilfried Philips, *University of Gent*, Charilaos Christopoulos, *ETRO/IRIS-VUB, (BELGIUM)*

**Estimation of Aliasing Error in Layered Coding System** V-349  
Masahiro Iwahashi, *Nagaoka University of Technology*, Koichi Ohya, *Graphics Communication Laboratories*, Noriyoshi Kambayashi, *Nagaoka University of Technology, (JAPAN)*

**Two-Stage Discrete Cosine Transform and Its Application on Image Compression** ~~WITHDRAWN~~  
Neng-Chung Hu, Hsien-Wen Chen, *National Taiwan Institute of Technology, (TAIWAN)*

**Design of Subband Filters for Image Coding Based on Perceptual Criteria** ~~WITHDRAWN~~  
V Ralph Algazi, *Forrest, Blumhew, Mow, Azfar Najmi, University of California, Davis, (USA)*

**Subband Coding of Colour Images with Limited Palette Size** V-353  
Patrick Waldemar, Tor Audun Ramstad, *Norwegian Institute of Technology, (NORWAY)*

#### IMAGE MODELLING AND ANALYSIS

**Chairperson:** Vinod Chandran, *Signal Processing Research Centre, Queensland University of Technology, (AUSTRALIA)*

**Inversion of Large-Support Ill-conditioned Linear Operators Using a Markov Model with a Line Process** V-357  
Mila Nikolova, Ali Mohammad-Djafari, Jérôme Idier, *LSS-ESE, (FRANCE)*

**Estimation of  $q$ -th Order Fractal Dimensions** V-361  
Daniele Giusto, *University of Cagliari*, Stefano Fioravanti, *University of Genoa, (ITALY)*

**On the Distribution of the DCT Coefficients** V-365  
Thierry Eude, *La3i-LCIA, INSA Rouen*, Richard Grisel, *ICPI-L2S2*, Hocine Cherifi, *TSI (CNRS 842)*, Roland Debric, *La3i-LCIA, INSA Rouen, (FRANCE)*

**Self-Similarity Modeling for Interpolation and Extrapolation of Multi-Viewpoint Image Sets** V-369  
Takeshi Naemura, Hiroshi Harashima, *The University of Tokyo, (JAPAN)*

**Stochastic Modelling and Estimation of Multispectral Image Data** V-373  
Richard Schultz, Robert L Stevenson, *University of Notre Dame, (USA)*

**Approximate Covariance Functions for Grey-Level Gibbs Random Fields** V-377  
Ibrahim M Elfadel, *MIT Research Laboratory of Electronics, (USA)*

**Cluster-Based Probability Model Applied to Image Restoration and Compression** V-381  
Kris Popat, Rosalind W Picard, *MIT Media Laboratory, (USA)*

**Covariance Matrix Matching for Multi-Spectral Image Classification** V-385  
Paul J Whitbread, *Defence Science and Technology Organisation, (AUSTRALIA)*

**Tracking Subspace Representations of Face Images** V-389  
Hsi-Jung Wu, Dulce Ponceleon, Katherine Wang, James Normile, *Apple Computer Inc, (USA)*

**Texture Class Assignment in TEXSCALE: An Evaluation Study** V-393  
Jane You, *University of South Australia*, Harvey A Cohen, *La Trobe University, (AUSTRALIA)*

**A CAD Driven Multiscale Approach to Automated Inspection** V-397  
Daniel Tretter, Khalid Khawaja, Charles Bouman, Anthony Maciejewski, *Purdue University, (USA)*

#### VIDEO CODING I

**Chairperson:** Kiyoharu Aizawa, *University of Tokyo, (JAPAN)*

**Adaptive Post-Processing Algorithms for Low Bit Rate Video Signals** V-401  
Tsann-Shyong Liu, *Institute of Computer Science, (TAIWAN ROC)*; Nikil Jayant, *AT&T Bell Laboratories, (USA)*

**3D Subband Coder for Very Low Bit Rates** V-405  
Weng Leong Chooi, King Ngi Ngan, *Monash University, (AUSTRALIA)*

**Performance Evaluation of Video Coding Schemes Working at Very Low Bit Rates** V-409  
Laura Contin, Stefano Battista, *CSELT, (ITALY)*

**Simultaneous 3-D Motion Estimation and Wire-Frame Model Adaptation Including Photometric Effects for Knowledge-based Video Coding** V-413  
Gözde Bozdağı, *Bilkent University, (TURKEY)*; A Murat Tekalp, *University of Rochester, (USA)*; Levent Onural, *Bilkent University, (TURKEY)*

**Rate-Distortion Analysis of Variable Block Size VQ-Based Motion Compensated Video Codecs** V-417  
Sam Liu, *Hewlett-Packard Laboratories, (USA)*

**Use of Steerable Viewing Window (SVW) to Improve the Visual Sensation in Face to Face Teleconferencing** V-421  
Liyanage C De Silva, Kiyoharu Aizawa, Mitsutoshi Hatori, *University of Tokyo, (JAPAN)*

**Two-Layered DCT Based Coding Scheme for Recording Digital HDTV Signals** V-425  
Jae Hyun Kim, Gooman Park, *Samsung Electronics Co Ltd, (KOREA)*

**Noise Reduction for MPEG Type of Codec** V-429  
Li Yan, *AT&T Bell Laboratories, (USA)*

**The Application of Subband Coding in MPEG for Prioritized ATM Networks** V-433  
Brian DeCleene, Henrik Sorensen, *University of Pennsylvania, (USA)*

**An Investigation of JPEG Image and Video Compression Using Parallel Processing** V-437  
Gregory W Cook, Edward J Delp, *Purdue University, (USA)*

**Effective Nearly Lossless Compression of Digital Video Sequences via Motion Compensated Filtering** V-441  
Balas K Natarjan, Vasudev Bhaskaran, *Hewlett Packard Labs, (USA)*

## IMAGE RESTORATION AND ESTIMATION

**Chairperson:** Trevor Cole, *University of Sydney, (AUSTRALIA)*

**Blind Deconvolution for Multidimensional Images** V-445  
R P Millane, *Purdue University, (USA)*; P J Bones,  
H Jiang, *University of Canterbury, (NEW ZEALAND)*

**Iterative Adaptive  $I_p$  Restoration of Blurred Images** V-449  
Wai Ho Pun, Brian D Jeffs, *Brigham Young University, (USA)*

**Applying Generalised Cross-Validation to Image Restoration** V-453  
Robert Whatmough, *Defence Science and Technology Organisation, (AUSTRALIA)*

**On the Equivalence of Normalized Convolution and Normalized Differential Convolution** V-457  
Carl-Fredrik Westin, Klas Nordberg, Hans Knutsson, *Linköping University, (SWEDEN)*

**Sparse Colour and Grey Scale Image Restoration Using a Morphological Method** V-461  
Alan Harvey, *Royal Melbourne Institute of Technology, (AUSTRALIA)*; Harvey Cohen, *La Trobe University, (AUSTRALIA)*

**Image Reconstruction From Zeros of the Z-transform** V-465  
Charles Parker, Brenda Satherley, Philip J Bones, *University of Canterbury, (NEW ZEALAND)*

**A Subspace Decomposition Method for Point Source Localization in Blurred Images** V-469  
Metin Gunsay, Brian D Jeffs, *Brigham Young University, (USA)*

**A Fast Parallel Projection Algorithm for Set Theoretic Image Recovery** V-473  
P L Combettes, H Puh, *City University of New York, (USA)*

## VIDEO CODING II

**Chairperson:** Murat Tekalp, *University of Rochester, (USA)*

**Frequency Scalable Video Coding Using the MDCT** V-477  
Andrew Johnson, *Telecom Australia (AUSTRALIA)*; John Princen, *AT & T Bell Laboratories, (USA)*; Ming Chan, *Telecom Australia, (AUSTRALIA)*

**Hybrid Video Coding for Low Bit-Rate Applications** V-481  
Feng-Ming Wang, *C-Cube Microsystems, (USA)*; Sam Liu, *Hewlett Packard Laboratories, (USA)*

**A Novel Tree-Structured Video Coder** V-485  
Francesco De Natale, *University of Genoa, (ITALY)*; Giuseppe Desoli, *PhD Engineering, Daniele Giusto, University of Genoa, (ITALY)*

**Library-based Image Coding** V-489  
Nuno Vasconcelos, Andrew Lippman, *MIT Media Laboratory, (USA)*

**Rate and Resolution Scalable Subband Coding of Video** V-493  
David Taubman, Avideh Zakhor, *University of California at Berkeley, (USA)*

**Subband Video Coding with Temporally Adaptive Motion Interpolation** V-497  
Jungwoo Lee, Bradley Dickinson, *Princeton University, (USA)*

**Efficient Prediction of Uncovered Background in Interframe Coding Using Spatial Extrapolation** V-501  
André Kaup, Til Aach, *RWTH Aachen, (GERMANY)*

**Rate Buffered Fractal Video** V-505  
David Wilson, Jeremy Nicholls, Donald M Monro, *University of Bath, (UK)*

## IMAGE PROCESSING AND ANALYSIS

**Chairperson:** W Boles, *Queensland University of Technology, (AUSTRALIA)*

**Importance of Feature Selection in Principal Component Face Recognition** **WITHDRAWN**  
Tanweer Kabir, David Rees, *CSIRO, (AUSTRALIA)*

**Fast Design of Optimal Array Filters** V-509  
Phillip Musumeci, *Australian Defence Force Academy, (AUSTRALIA)*

**3-D Band-Limitation by Motion Adaptive Spatial Filtering** V-513  
Hyun-Soo Kang, *Korea Advanced Institute of Science and Technology, (KOREA)*; Jong-Hun Kim, *DAEWOO Elec. Co. Ltd, (KOREA)*; Lee, Seong-Dae Kim, *Korea Advanced Institute of Science and Technology, (KOREA)*

**The Automatic Generation of 3D Object Model from Range Image** V-517  
Wentao Zheng, Hiroshi Harashima, *University of Tokyo, (JAPAN)*

**Real-Time Determination of the Signal-to-Noise Ratio of Partly Coherent Seismic Time Series** V-521  
Peter K Møller, *Technical University of Denmark and Signal-Data, (DENMARK)*

**A Focusing Algorithm for the Optical Array Imaging System** V-525  
Osamu Ikeda, *Takushoku University, (JAPAN)*

**Robust N-Dimensional Orientation Estimation Using Quadrature Filters and Tensor Whitening** V-529  
Hans Knutsson, Mats Andersson, *Linköping University, (SWEDEN)*

**Global Non-Linear Multigrid Optimization for Image Analysis Tasks** V-533  
Jean-Marc LaFerté, Patrick Pérez, Fabrice Heitz, *INRIA - Rennes, (FRANCE)*

**Maximum Likelihood Scale Estimation for a Class of Markov Random Fields** V-537  
Charles A Bouman, *Purdue University, (USA)*; Ken Sauer, *University of Notre Dame, (USA)*

**Divergence Penalty for Image Regularization** V-541  
Joseph A O'Sullivan, *Washington University in St Louis, (USA)*

**New Prospects in Line Detection for Remote Sensing Images** V-545  
Nicolas Merlet, *Hebrew University of Jerusalem, (ISRAEL)*; Josiane Zerubia, *INRIA, (FRANCE)*

**Bayesian Restoration of Millimeter Wave Imagery** V-549  
B R Hunt, D DeKruger, *University of Arizona, (USA)*

## IMAGE CODING II

**Chairperson:** Abdesselam Bouzerdoum, *University of Adelaide, (AUSTRALIA)*

**Residual Image Coding Using Mathematical Morphology** V-553  
Josep R Casas, Luis Torres, *Universitat Politècnica de Catalunya, (SPAIN)*

**Performance Comparison of Different Wavelet Functions in Compression and Reconstruction of Images** \*  
Prakash Kumar Pati, Sumana Gupta, *Indian Institute of Technology, (INDIA)*

**Fractal Image Compression without Searching** V-557  
Donald Monro, Stuart Woolley, *University of Bath, (UK)*

**On the Convergence of Fractal Transforms**  
 Bernd Hürtgen, Thomas Hain, *Aachen University of Technology, (GERMANY)*

V-561

## VECTOR QUANTIZATION

Chairperson: Eve Riskin, *University of Washington, (USA)*

**New Improved Collage Theorem with Applications to Multiresolution Fractal Image Coding**  
 Geir E Øien, *Rogaland University Centre, (NORWAY)*; Zachi Baharav, *Technion - Israel Institute of Technology, (ISRAEL)*; Skjalg Lepsoy, *Consensus Analysis A/S, (NORWAY)*; E Karnin, *IBM Israel Science and Technology, David Malah, Technion - Israel Institute of Technology, (ISRAEL)*

V-565

**Image Coding with Overlapped Projection and Pyramid Vector Quantization** V-593  
 Rosa Lancini, *CEFRIEL-COGEFO, Emanuele Marconetti, Stefano Tubaro, Politecnico di Milano, (ITALY)*

**Overlapped Adaptive Partitioning for Image Coding Based on the Theory of Iterative Functions Systems**  
 Emmanuel Reusens, *Swiss Federal Institute of Technology, (SWITZERLAND)*

V-569

**Segmentation Based Coding of Textures Using Stochastic Vector Quantization** V-597  
 Luis Torres, Josep R Casas, S de Diego, *Universitat Politècnica de Catalunya, (SPAIN)*

### Position-Dependent Encoding

John G Apostolopoulos, Aleksandar Pfajfer, Hae Mook Jung, Jae S Lim, *Massachusetts Institute of Technology, (USA)*

V-573

**Image Coding using Pyramid Vector Quantization of Subband Coefficients** V-601  
 Ely Tsern, Teresa Meng, *Stanford University, (USA)*

### 3-D Image Coding Based on Affine Transform

Toshiaki Fujii, Hiroshi Harashima, *University of Tokyo, (JAPAN)*

V-577

**Vector Quantization over a Noisy Channel Using Soft Decision Decoding** V-605  
 Mikael Skoglund, Per Hedelin, *Chalmers University of Technology, (SWEDEN)*

### A Study of Convex Coders With An Application to Image Coding

Kohtaro Asai, *Mitsubishi Electric Corp, (JAPAN)*; Nguyen T Thao, *Hong Kong University of Science and Technology, (HONG KONG)*; Martin Vetterli, *University of Southern California/Columbia University, (USA)*

V-581

**Predictive Mean Search Algorithm for Vector Quantization of Images** V-609  
 Kwok-Tung Lo, *Hong Kong Polytechnic, (HONG KONG)*; Jian Feng, *University of New South Wales, (AUSTRALIA)*

### An Extension to the Analytical Gabor Expansion with Applications in Image Coding

Andreas Teuner, *University of Duisburg, (GERMANY)*; Per Asbeck Nielsen, Bedrich J Hosticka, *Fraunhofer Institute of Microelectronic Circuits and Systems, (GERMANY)*

V-585

**An Efficient Neural Prediction for Vector Quantization** V-613  
 Roberto Fioravanti, *PhD Engineering, Genova, Stefano Fioravanti, Università di Genova, Daniele D Giusto, Università di Cagliari, (ITALY)*

### Comparison of "Wavelet" Filters and Subband Analysis Structures for Still Image Compression

James Andrew, Philip Ogunbona, Frank Paoloni, *University of Wollongong, (AUSTRALIA)*

V-589

**Next-State Functions for Finite-State Vector Quantization** V-617  
 Nasser M Nasrabadi, Syed A Rizvi, *State University of New York at Buffalo, (USA)*

**Lapped RVQ and Alphabet and Entropy Constraints** V-621  
 Francesco Nesci, Faouzi Kossentini, Mark J T Smith, *Georgia Institute of Technology, (USA)*

**One-Pass Adaptive Universal Vector Quantization** V-625  
 Michelle Effros, *Stanford University, Philip A Chou, Xerox Palo Alto Research Center, Robert M Gray, Stanford University, (USA)*

**Image Reconstruction using Vector Quantized Linear Interpolation** V-629  
 Sheila S Hemami, Robert M Gray, *Stanford University, (USA)*

\* Manuscript unavailable for publication.  
 Late papers may be included in the rear of Volume 6.

## ON-LINE CURSIVE HANDWRITING RECOGNITION USING SPEECH RECOGNITION METHODS

Thad Starnert†, John Makhoul, Richard Schwartz, and George Chou

BBN Systems and Technologies  
70 Fawcett Street, Cambridge, MA 02138  
Email: Makhoul@bbn.com

### ABSTRACT

A hidden Markov model (HMM) based continuous speech recognition system is applied to on-line cursive handwriting recognition. The base system is unmodified except for using handwriting feature vectors instead of speech. Due to inherent properties of HMMs, segmentation of the handwritten script sentences is unnecessary. A 1.1% word error rate is achieved for a 3050 word lexicon, 52 character, writer-dependent task and 3%-5% word error rates are obtained for six different writers in a 25,595 word lexicon, 86 character, writer-dependent task. Similarities and differences between the continuous speech and on-line cursive handwriting recognition tasks are explored; the handwriting database collected over the past year is described; and specific implementation details of the handwriting system are discussed.

### 1. INTRODUCTION

Traditionally, the first step in handwriting recognition is the segmentation of words into component characters [1]. However, in modern continuous speech recognition efforts, phonemes are not segmented before training or recognition. Instead, segmentation occurs simultaneously with recognition. If such a system could be adapted for handwriting, the very difficult and time consuming issue of segmentation could be avoided. This paper presents an approach for the automatic recognition of *on-line cursive handwriting* (using input from a pentop computer) by using continuous speech recognition methods. Specifically, the use of hidden Markov models and statistical grammars is explored. We show that, with essentially no modification, a speech recognition system can perform accurate on-line handwriting recognition, with the input features being those of the writing instead of speech.

Hidden Markov models have intrinsic properties which make them very attractive for handwriting recognition. For training, all that is necessary is a data stream and its transcription (the text matching the handwriting). The training process automatically aligns the components of the transcription to the data. Thus, no special effort is needed to label training data. Segmentation, in the traditional sense, is avoided altogether. Recognition is performed on another data stream. Again, no explicit segmentation is necessary.

†currently with the MIT Media Lab.

The segmentation of words into characters or even sentences into words occurs naturally by incorporating the use of a lexicon and a language model into the recognition process. The result is a text stream that can be compared to a reference text for error calculation.

Section 2 discusses the similarities of speech and handwriting recognition tasks and provides some background on techniques. Section 3 describes an initial 3050 word, 52 symbol, writer dependent experiment. Section 4 discusses a more ambitious 25,595 word, 86 symbol, writer dependent system involving multiple writers. Section 5 examines the results of this experiment and discusses future work.

### 2. SIMILARITIES OF ON-LINE HANDWRITING RECOGNITION TO SPEECH RECOGNITION

On-line handwriting is very similar to the problem of continuous speech recognition. On-line handwriting can be viewed as a signal (x,y coordinates) over time, just like in speech. The items to be recognized are well-defined (usually the alphanumeric characters) and finite in number, as are the phonemes in speech. The shape of a handwritten character depends on its neighbors. Spoken phonemes also change due to coarticulation in speech. In both cases, these basic units form words and the words form phrases. Thus, language modeling can be applied to improve recognition performance.

In spite of these similarities, handwriting recognition has some basic differences to speech recognition. Unlike continuous speech, word boundaries are usually distinct in handwriting. Thus, words should be easier to distinguish. However, in cursive writing the dots and crosses involved in the characters "i", "j", "x", and "t" are not added until after the whole word is written. Thus, all the evidence for a character may not be contiguous. Additionally, in words with multiple crossings ("t" and "x") and/or dottings ("i" and "j") the order of pen strokes is ambiguous. Even so, with the many parallels between on-line writing and speech, speech recognition methods should be applicable to on-line handwriting recognition. Since hidden Markov models currently constitute the state of the art in speech recognition, this method also seems a likely candidate for handwriting recognition.

There has been some interest in the use of HMMs for on-line handwriting recognition (see, for example, [2,3]). How-

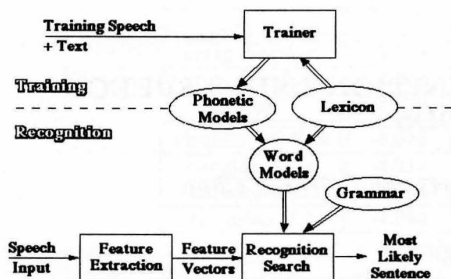


Figure 1: BYBLOS speech system.

ever, the few studies that have used HMMs have dealt with small vocabularies, isolated characters, or isolated words. In this study, our objective is to deal with continuous cursive handwriting and large vocabularies (thousands of words), using a speech recognition system and language models.

### 3. INITIAL SYSTEM

In the initial system, the BBN BYBLOS Continuous Speech Recognition system [4] (see Figure 1) was used without modification on an on-line cursive handwriting corpus created from prompts from the ARPA Airline Travel Information Service (ATIS) corpus [5]. These full sentence prompts (approximately 10 words per sentence) were written by a single subject. These sentences were then reviewed (verified) to make sure that the prompts were transcribed correctly. After verification, these sentences were separated into a set of 381 training sentences and a mutually exclusive set of 94 test sentences. The lexicon for this task included 3050 words, where lowercase and capitalized versions of a word are considered distinct.

For this initial system there were 54 characters: 52 lower and upper case alphabetic, a space character, and a "backspace" character. The backspace character is appended onto words that contain "i", "j", "x", or "t". This character models the space the pen moves after finishing the body of the word to add the dot or the cross when drawing one of these characters.

The data was acquired using a Momenta pentop which stored the script in a simple time series of x and y coordinates at a sampling rate of 66 Hz. The handwriting data is sampled continuously in time, except when the pen is lifted (Momenta pentops provide no information about pen movement between strokes). Because we wanted to use our speech recognition system with no modification, we decided to simulate a continuous-time feature vector by arbitrarily connecting the samples from pen-up to pen-down with a straight line and then sampling that line ten times. Thus, the data effectively became one long criss-crossing stroke for the entire sentence, where words run together and "i" and "j" dots and "t" and "x" crosses cause backtracing over previously drawn script (see Figure 2).

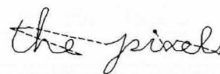


Figure 2: Connecting strokes.

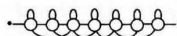


Figure 3: 7-state HMM used to model each character.

For each sample point, an analysis program computed a two-element feature vector: the writing angle at that sample and the change in the writing angle [2]. These time series of feature vectors were then fed into the BYBLOS system. For this task, BYBLOS quantizes the feature vectors for a sentence into 64 different clusters. These new time series are then used with their respective sentence transcriptions to train HMMs representing the script characters (note that the alignment of the clusters with the sentence transcriptions occurs automatically in this process). A 7-state HMM model was chosen to represent each symbol (see Figure 3). Since the penning of a script letter often differs depending on the letters written before and after it, additional HMMs are used to model these contextual effects [5]. Adjacent effects between two letters (bilets) are modeled as well as three letter (trilet) contexts. In a given set of sentences there may be many trilets, up to the number of symbols cubed. However, in English only a subset of these are allowed. In the ATIS task there are 3639 different trilets in the training sentences.

A statistical grammar can also be used to improve recognition performance. For this experiment, a bigram grammar (to relate pairs of words) was created using a larger set of 17209 sentences from the ATIS corpus (the 94 test sentences were not included). The resultant grammar has a perplexity of 20. Table 1 shows the word error rates for this task when doing recognition using context without the grammar (perplexity = 3050), using the grammar without context, and using both context and the grammar. Word error rate is measured as the sum of the percentage of words deleted, the percentage of words inserted, and the percentage of words that are substituted for other words in the set of test sentences.

Table 1: ATIS 3050 word, writer-dependent test results

|                 | context +<br>no gram. | no context +<br>gram. | context +<br>gram. |
|-----------------|-----------------------|-----------------------|--------------------|
| word error rate | 4.2%                  | 2.2%                  | 1.1%               |

As can be seen from the table, both context and a grammar are very powerful tools in aiding recognition. With no grammar but with context an error rate of 4.2% was observed. When the grammar was added and context not used, the error rate dropped to 2.2%. However, the best result used both context and a grammar for an word error rate of 1.1%. Of interest is the factors of 2 relating the error rates shown. Similar factors of 2 have also been observed

Glass, the U.K.'s largest pharmaceutical concern, advanced 23 W # 14.13.

Figure 4: Writing from subject aim.

in the research on the speech version of this corpus. With the best (1.1%) word error rate, only 10 errors occurred for the entire test set. Experimentation was suspended at this point since so few errors did not allow any further analysis of the problems in our methods.

The above experiments demonstrated the potential utility of speech recognition methods, especially the use of HMMs and grammars, to the problem of on-line cursive handwriting recognition. Based on these good preliminary results, we embarked on a more ambitious task with a larger vocabulary and more writers.

#### 4. WALL STREET JOURNAL: A 25,000 WORD TASK

Recently, we have collected cursive written data using text from the ARPA Wall Street Journal task (WSJ) [6], including numerals, punctuation, and other symbols, for a total of 88 symbols (62 alphanumeric, 24 punctuation and special symbols, space, and backspace). The prompts from the Wall Street Journal consist mainly of full sentences with scattered article headings and stock listings (all are referred to as sentences for convenience). We have thus far collected over 7000 sentences (175,000 words total or about 25 words/sentence) from 21 writers on two GRiD Convertible pentops. See Figure 4 for an example of the data collected (this sentence was taken from a test set). The writers were gathered from the Cambridge, Massachusetts area and were mainly students and young professionals. Several non-native writers were included (writers whose first working language was not English). While the handwriting input was constrained, the rules given the subjects were simple: write the given sentence in cursive; keep the body of a word connected (do not lift the pen in the middle of a word); and do crossings and dottings after completing the body of a word. However, since many writers could not remember how to write capital letters in cursive, great leniency was allowed. Furthermore, apostrophes were allowed to be written both in the body of the word, or at the end of the word like a cross or dot. For example, the word "don't" could be written as "dont" followed by the placement of the apostrophe or "don", apostrophe, and "t". Overall, this task might be best described as "pure cursive" in the handwriting recognition literature.

For the purposes of this experiment, punctuation, numerals, and symbols are counted as words. Thus, ".", ",", "0", "1", "\$", "{", etc., are each counted as a word. However, apostrophes within words are counted as part of that word. Again, a capitalized version of a word is counted as distinct from the lowercase version of the word. While these standards may artificially inflate the word error rates, they are a simple way to disambiguate the definition of a word.

In addition to the angle and delta angle features described in the last section, the following features were added: delta x, delta y, pen up/pen down, and  $\text{sgn}(x - \max(x))$ . Pen up/pen down is 1 only during the ten samples connecting one pen stroke to another; everywhere else it is 0.  $\text{Sgn}(x - \max(x))$  is 1 only when, at that time, the current sample is the right-most sample of the data to date. Also, two preprocessing steps were used on the subjects' data. The first was a simple noise filter which required that the pen traverse over one hundredth of an inch before allowing a new sample. The second step padded each pen stroke to a minimum size of ten samples.

At the time of this writing, samples from six subjects were used for writer dependent experiments. Three fourths of a subject's sentences were used for training with the remaining fourth used for testing (see Table 2). A lexicon of 25,595 words was used since it spanned all of the data. A bigram grammar was created from approximately two million Wall Street Journal sentences from 1987 to 1989 (not including the sentences used in data collection). The results of the writer dependent tests are shown in Table 3. Substitution, deletion, insertion, and the total word error rates are included. Table 4 shows estimated character recognition error rates for each class of character: alphabetic, numeral, and punctuation and other symbols. The sum of the substitution and deletion error rates for each class is represented in this table since insertions are not directly attributable to a particular class of character. However, the total character error shown incorporates insertion errors since these errors are distributed over the entire set of classes. On average, the test sets consist of 1.9% numerals, 4.1% punctuation and other symbols, and 94% alphabets. Both aim and shs are non-native writers. A test experiment was performed without a grammar (but with context) on subject shs resulting in an error rate approximately four times the previous error rate. This result was the same ratio seen in the ATIS task.

Table 2: Division of subjects' sentences into training and test

| subject | test              |                  |
|---------|-------------------|------------------|
|         | # train sentences | # test sentences |
| aim     | 423               | 141              |
| dsf     | 404               | 135              |
| rgb     | 437               | 146              |
| shs     | 423               | 141              |
| slb     | 411               | 137              |
| wcd     | 314               | 105              |

Table 3: WSJ 25,595 word, writer dependent word errors

| subject | Subst. | Delet. | Insert. | Total |
|---------|--------|--------|---------|-------|
| aim     | 2.7%   | 0.4%   | 1.4%    | 4.5%  |
| dsf     | 3.6%   | 0.4%   | 1.2%    | 5.2%  |
| rgb     | 3.3%   | 0.5%   | 1.7%    | 5.5%  |
| shs     | 2.0%   | 0.2%   | 0.7%    | 2.9%  |
| slb     | 2.9%   | 0.1%   | 1.3%    | 4.3%  |
| wcd     | 2.1%   | 0.4%   | 0.5%    | 3.0%  |
| ave.    | 2.8%   | 0.3%   | 1.1%    | 4.2%  |

Table 4: Estimated character error rates for alphabetic, numerals, and symbols

| subject | Est.<br>num. | Est.<br>sym. | Est.<br>alpha. | total |
|---------|--------------|--------------|----------------|-------|
| aim     | 7.1%         | 4.7%         | .47%           | 1.4%  |
| dsf     | 8.3%         | 8.6%         | .78%           | 1.9%  |
| rgb     | 3.2%         | 11.%         | .77%           | 1.8%  |
| shs     | 4.6%         | 7.8%         | .26%           | 0.80% |
| slb     | 7.2%         | 7.1%         | .64%           | 1.7%  |
| wcd     | 5.4%         | 5.7%         | .47%           | 1.0%  |
| ave.    | 6.2%         | 7.5%         | .57%           | 1.4%  |

## 5. EXPERIMENT ANALYSIS AND FUTURE DIRECTIONS

These results are quite startling when put in context. The BYBLOS speech system was not significantly modified for handwriting recognition, yet it handled several difficult handwriting tasks. Furthermore, none of the BYBLOS automatic optimization features were used to improve the results of any writer (or group of writers). No particular stroke order was enforced on the writers for dottings and crossings (besides being after the body of the word), and there are known inaccuracies in the transcription files. Note that a significantly larger error rate was observed for numerals and symbols than for alphabetic. Even with all insertion errors added to the estimate of the alphabetic error, the error rates for numerals and symbols are still significantly higher. One improvement may be to specifically train on common digit strings such as "1989", "80286", and "747" (presently, "1989" is recognized as four separate words instead of the more salient whole). Also, apostrophes are handled incorrectly by expecting only the intra-word stroke version. By expecting both standard stroke orders in words with apostrophes, the system can increase the recognition accuracy of these words significantly. By fixing these simple problems and using BYBLOS's optimizing features, a 10-50% reduction in word error rate may occur.

In this experiment we used a large number of training sentences per writer. Supplying such a large amount of training text may be tiring for just one writer. However, there is some evidence that not as many training sentences per writer are needed for good performance. Furthermore, if good word error rates for the cursive dictation task can be assured, a writer may be willing to spend some time writing sample sentences. A possible compromise is to create a writer independent system which can then be adapted to a particular writer with a few sample sentences. With this level of training it may be possible to relax the few restrictions made on the writers in this experiment.

Future experiments will be directed at further reduction of the error rates for the writer dependent task. In addition, writer independent and writer adaptive systems will be attempted. Scalability of the number of training sentences will be addressed along with possible changes to the BYBLOS system to better accommodate handwriting.

## 6. CONCLUSION

We have shown that a HMM based speech recognition system can perform well on on-line cursive handwriting tasks without needing segmentation of training or test data. On a 25,595 word, 86 symbol, writer dependent task over six writers, an average of 4.2% word error rate and an average of 1.4% character error rate was achieved. With some simple tuning, significant reduction in these error rates is expected. These findings suggest that HMM-based methods combined with statistical grammars will prove to be a very powerful tool in handwriting recognition.

## 7. ACKNOWLEDGMENTS

The authors wish to thank Long Nguyen and George Zavaliagos for their help with the BYBLOS system, Tavenner Hall and Brenda Pendleton for their assistance in verifying data, and the Vision & Modeling Group, MIT Media Lab for use of their facilities.

## 8. REFERENCES

- [1] C. Tappert, C. Suen, and T. Wakahara. The State of the Art in On-Line Handwriting Recognition *IEEE T. Pat. Anal. & Mach. Int.*, pp. 787-808, August 1990.
- [2] R. Nag, K. H. Wong, F. Fallside, Script Recognition using Hidden Markov Models. In *Proc. ICASSP*, pp. 2071-2074, Tokyo, Japan, 1986.
- [3] K. Nathan, J. Bellegarda, D. Nahamoo, E. Bellegarda. On-Line Handwriting Recognition Using Continuous Parameter Hidden Markov Models. In *Proc. ICASSP*, pp. V-121-124, Minneapolis, MN, 1993.
- [4] F. Kubala, A. Anastasakos, J. Makhoul, L. Nguyen, R. Schwartz, G. Zavaliagos. Comparative Experiments on Large Vocabulary Speech Recognition To be presented at *ICASSP*, Adelaide, Australia, 1994.
- [5] MADCOW Multi-Site Data Collection for a Spoken Language Corpus. *Proc. DARPA Speech and Natural Language Workshop*, pp. 7-14, Harriman, NY, Morgan Kaufmann Publishers, 1992.
- [6] R. M. Schwartz, Y. L. Chow, O. A. Kimball, S. Roucos, M. Krasner, and J. Makhoul Context-Dependent Modeling for Acoustic-Phonetic Recognition of Continuous Speech. *Proc. ICASSP*, pp.1205-1208, Tampa, FL, March 1985.
- [7] D. Paul The Design for the Wall Street Journal-based CSR Corpus. *Proc. DARPA Speech and Natural Language Workshop*, pp. 357-360, Morgan Kaufmann Publishers, 1992.