# SPEECH

## TABLE OF CONTENTS

### CELP CODING

**Chair:** Jean-Pierre Adoul, University of Sherbrooke *(CANADA)*

**4KBPS Improved Pitch Prediction CELP Speech Coding with 20ms Frame**
Masahiro Serizawa, Kazunori Ozawa, NEC Corporation *(JAPAN)*

**A Low-Complexity Toll-Quality Variable Bit Rate Coder for CDMA Cellular Systems**
Peter Kroon, Michael Recchione, AT&T Bell Laboratories *(USA)*

**Toll Quality 16 kb/s CELP Speech Coding with Very Low Complexity**
Juin-Hwey Chen, AT&T Bell Laboratories *(USA)*

**CELP Coding Using Trellis-Coded Vector Quantization of the Excitation**
Andrei Popescu, Nicolas Moreau, Télécom Paris, Claude Lamblin, CNET/LAA/TSS/CMC *(FRANCE)*

**Interpolating the History Improved Excitation Coding for High Quality CELP Coding**
Per Hedelin, Thomas Eriksson, Chalmers University of Technology *(SWEDEN)*

**Fast Stochastic Codebook Search Through the Use of Odd-Symmetric Crosscorrelation Basis Vectors**
Cheung-Fat Chan, City University of Hong Kong *(HONG KONG)*

**Improvements of Background Sound Coding in Linear Predictive Speech Coders**

**Improved CS-CELP Speech Coding in a Noisy Environment Using a Trained Sparse Conjugate Codebook**
Akitoshi Kataoka, Sachiko Hosaka, NTT Human Interface Labs; Jotaro Ikedo, NTT Wireless Systems Laboratories; Takehiro Moriya, Shinja Hayashi, NTT Human Interface Labs. *(JAPAN)*

**CELP Coding Based on Mel-cepstral Analysis**
Kazuhito Koishida, Keiichi Tokuda, Takao Kobayashi, Satoshi Imai, Tokyo Institute of Technology *(JAPAN)*

**An Embedded Scheme for Regular Pulse Excited (RPE) Linear Predictive Coding**
Shude Zhang, Gordon Lockhart, University of Leeds *(UK)*

### RECOGNITION: LARGE VOCABULARY

**Chair:** Michael Picheny, IBM *(USA)*

**Performance of the IBM Large Vocabulary Continuous Speech Recognition System on the ARPA Wall Street Journal Task**

**New Developments in the Lincoln Stack-Decoder Based Large Vocabulary CSR System**
Douglas B. Paul, MIT Lincoln Laboratory *(USA)*

**Large Vocabulary Continuous Speech Recognition Using Word Graphs**
Xavier Aubert, Philips GmbH Research Laboratories - Aachen, Hermann Ney, Aachen University of Technology *(GERMANY)*

**Reducing Word Error Rate on Conversational Speech from the Switchboard Corpus**
P. Jeanrenaud, E. Eide, U. Chaudhari, J. McDonough K. Ng, M. Siu, H. Gish, BBN Systems and Technologies *(USA)*

**Golden Mandarin (III) -- A User-Adaptive Prosodic - Segment-Based Mandarin Dictation Machine for Chinese Language with Very Large Vocabulary**
Ren-Yuan Lyu, National Taiwan University; Lee-Feng Chien, Academia Sinica; Shiao-Hong Hwang, Hung-Yun Hsieh, Rung-Chuian Yang, Bo-Ren Bai, Jia-Chi Weng, Yen-Ju Yang, Shi-Wei Lin, National Taiwan University; Keh-Jiann Chen, Chiu-Yu Tseng, Lin-Shan Lee, Academia Sinica *(REPUBLIC OF CHINA)*
Complete Recognition of Continuous Mandarin Speech for Chinese Language with Very Large Vocabulary but Limited Training Data 61
Hsin-min Wang, Jia-lin Shen, Yen-Ju Yang, National Taiwan University; Chiu-yu Tseng, Academia Sinica, Lin-shan Lee, National Taiwan University (REPUBLIC OF CHINA)

Developments in Continuous Speech Dictation Using the ARPA WSJ Task 65
J.L. Gauvain, L. Lamel, M. Adda-Decker, LIMSI-CNRS (FRANCE)

Recent Improvements to the ABBOT Large Vocabulary CSR System 69
M.M. Hochberg, Cambridge University; S.J. Renals, University of Sheffield; A.J. Robinson, G. D. Cook, Cambridge University (ENGLAND)

The 1994 HTK Large Vocabulary Speech Recognition System 73
P. C. Woodland, C. J. Leggetter, J. J. Odell, V. Valitchev, S.J. Young, Cambridge University (UK)

Tangerine: A Large Vocabulary Mandarin Dictation System 77
Yuqing Gao, Hsiao-Wuen Hon, Zhiwei Lin, Gareth Loudon, S. Yogananthan, Baosheng Yuan, National University of Singapore (SINGAPORE)

Concept-Based Speech Translation 97
L. Mayfield, M. Gavaldà, W. Ward, A. Waibel, Carnegie Mellon University (USA)

PhoneBook: A Phonetically-Rich Isolated-Word Telephone-Speech Database 101
John F. Pitrelli, Cynthia Fong, Suk H. Wong, Judith R. Spitz, Hong C. Leung, NYNEX Science & Technology, Inc. (USA)

CTIMIT: A Speech Corpus for the Cellular Environment with Applications to Automatic Speech Recognition 105
Kathy L. Brown, E. Bryan George, Lockheed Sanders, Inc. (USA)

Toward Movement-Invariant Automatic Lip-Reading and Speech Recognition 109
Paul Duchnowski, University of Karlsruhe (GERMANY); Martin Hunke, Carnegie Mellon University (USA); Dietrich Bütching, Uwe Meier, University of Karlsruhe (GERMANY); Alex Waibel, Carnegie Mellon University (USA)

Some Results with a Trainable Speech Translation and Understanding System 113
V.M. Jiménez, A. Castellanos, E. Vidal, Universidad Politécnica de Valencia (SPAIN)

A Continuous Speech Recognition System Using Finite State Network and Viterbi Beam Search for the Automatic Interpretation 117
Nam-Yong Han, Hoi-Rin Kim, Kyu-Woong Hwang, Young-Mok Ahn, Joon-Hyung Ryoo, ETRI (KOREA)

ASR SYSTEM & CORPORA

Chair: John Bridle, Diagem (USA)

WSJCAMO: A British English Speech Corpus for Large Vocabulary Continuous Speech Recognition 81
Tony Robinson, Jeroen Fransen, David Pye, Jonathan Foote, Steve Renals, Cambridge University (UK)

Voice Across Hispanic America: A Telephone Speech Corpus of American Spanish 85
Yeshwant Muthusamy, Edward Holliman, Barbara Wheatley, Texas Instruments; Joseph Picone, Mississippi State University; John Godfrey, University of Pennsylvania (USA)

Implementation of the POW (Phonetically Optimized Words) Algorithm for Speech Database 89
Yeonja Lim, Youngjiik Lee, ETRI (KOREA)

Microsoft Windows Highly Intelligent Speech Recognizer: Whisper 93
Xuedong Huang, Alex Acero, Fil Alleva, Mei-Yuh Hwang, Li Jiang, Milind Mahajan, Microsoft Corporation (USA)

ROBUST SPEECH RECOGNITION

Chair: Richard Stern, Carnegie Mellon University (USA)

Robust Speech Recognition Based on Stochastic Matching 121
Ananth Sankar, SRI International; Chin-Hui Lee, AT&T Bell Laboratories (USA)

On the Robustness of Linear Discriminant Analysis as a Preprocessing Step for Noisy Speech Recognition 125
Olivier Siohan, CRIN-CNRS & INRIA - Lorraine (FRANCE)

A Maximum Likelihood Procedure for a Universal Adaptation Method Based on HMM Composition 129
Yasuhiro Minami, Sadaoki Furui, NTT Human Interface Laboratories (JAPAN)

A Fast and Flexible Implementation of Parallel Model Combination 133
M. J. F. Gales, S. J. Young, Cambridge University (U.K.)
Multivariate-Gaussian-Based Cepstral Normalization for Robust Speech Recognition 137
Pedro J. Moreno, Bhiksha Raj, Evandro Gouvea, Richard M. Stern, Carnegie Mellon University (USA)

Robust Speech Recognition in Noise Using Adaptation and Mapping Techniques 141
Leonardo Neumeyer, Mitchel Weintraub, SRI International (USA)

Noisy Speech Recognition Using Robust Inversion of Hidden Markov Models 145
Seokyong Moon, Jenq-Neng Hwang, University of Washington (USA)

Rapid Environment Adaptation for Robust Speech Recognition 149
Keizaburo Takagi, Hiroaki Hattori, Takao Watanabe, NEC Corporation (JAPAN)

Noise Estimation Techniques for Robust Speech Recognition 153
H.G. Hirsch, C. Ehrlicher, Aachen University of Technology (GERMANY)

Pole-Filtered Cepstral Mean Subtraction 157
Devang Naik, Rutgers University (USA)

LANGUAGE MODELING

Chair: Roni Rosenfeld, Carnegie Mellon University (USA)

Language Model Adaptation via Minimum Discrimination Information 161
P. Srinivasa Rao, Michael D. Monkowski, Salim Roukos, IBM T. J. Watson Research Center (USA)

Clustering Word Category Based on Binomial posteriori Co-occurrence Distribution 165
Masafumi Tamoto, Takeshi Kawabata, NTT Basic Research Labs (JAPAN)

Language Modeling by Variable Length Sequences: Theoretical Formulation and Evaluation of Multigrams 169
Sabine Deligne, Frédéric Bimbot, Télécom Paris (FRANCE)

An Integrated Grammar/Bigram Language Model Using Path Scores 173
Harvey Lloyd-Thomas, Jerry H. Wright, Ensigna Limited; Gareth J.F. Jones, University of Cambridge (UK)

Discourse Structure for Multi-Speaker Spontaneous Spoken Dialogs: Incorporating Heuristics into Stochastic RTNs 177
Sheryl R. Young, Carnegie Mellon University (USA)

Improved Backing-Off for M-Gram Language Modeling 181
Reinhard Kneser, Philips GmbH Research Laboratories, Hermann Ney, RWTH Aachen, University of Technology (GERMANY)

QWI: A Method for Improved Smoothing in Language Modeling 185
G. Bordel, I. Torres, Universidad del Pais Vasco; E. Vidal, Universidad Politécnica de Valencia (SPAIN)

Using a Stochastic Context-Free Grammar as a Language Model for Speech Recognition 189
Daniel Jurafsky, University of California at Berkeley; Chuck Wooters, Department of Defense; Jonathan Segal, University of California at Berkeley; Andreas Stolcke, SRI International; Eric Fosler, University of California at Berkeley; Gary Tajchman, Voice Processing Corporation; Nelson Morgan, University of California at Berkeley (USA)

Improved Language Modeling by Unsupervised Acquisition of Structure 193
Klaus Ries, Finn Dag Bua, University of Karlsruhe (GERMANY); Ye-Yi Wang, Carnegie Mellon University (USA); Alex Waibel, University of Karlsruhe (GERMANY)

Understanding Referring Expressions in a Person-Machine Spoken Dialogue 197
Claudia Pateras, Gregory Dudek, Renato De Mori, McGill University (CANADA)

USE OF KNOWLEDGE IN ASR

Chair: Jim Glass, Massachusetts Institute of Technology, LCS (USA)

Analysis of Acoustic-Phonetic Variations in Fluent Speech Using TIMIT 201
Don X. Sun, State University of New York (USA), Li Deng, University of Waterloo (CANADA)

Analyzing Weaknesses of Language Models for Speech Recognition 205
Joerg P. Ueberla, DRA Malvern (UK)

A Hidden Markov Model with Optimized Inter-Frame Dependence 209
F. J. Smith, J. Ming, P. O'Boyle, A.D. Irvine, The Queen's University (N. IRELAND)
On the Use of Scaler Quantization for Fast HMM Computation
Shigeki Sagayama, Satoshi Takahashi, NTT Human Interface Laboratories (JAPAN)

Large-vocabulary Speech Recognition in Specialized Domains
Haakon Chevalier, Chuck Ingold, Carol Kunz, Chip Moore, Crispen Roven, Jon Yamron, Bradley Baker, Paul Bamberg, Sarah Bridle, Tracy Bruce, Amy Weader, Dragon Systems, Inc. (USA)

Understanding and Improving Speech Recognition Performance through the Use of Diagnostic Tools
Ellen Eide, Herbert Gish, Philippe Jeanrenaud, Angela Mielke, BBN Systems and Technologies (USA)

Phrase Bigrams for Continuous Speech Recognition
Egidio P. Giachin, CSELT (ITALY)

Using Explicit Segmentation to Improve HMM Phone Recognition
Carl D. Mitchell, Mary P. Harper, Leah H. Jamieson, Purdue University (USA)

Viterbi Algorithm for Acoustic Vectors Generated by a Linear Stochastic Differential Equation on Each State
Marco Saerens, Université Libre de Bruxelles, (BELGIUM)

Non Deterministic Stochastic Language Models for Speech Recognition
G. Riccardi, E. Bocchieri, R. Pieraccini, AT&T Bell Laboratories (USA)

END OF PROGRAM

TOPICS IN SPEECH CODING

Chair: Peter Kroon, AT&T Bell Laboratories (USA)

Improving 16 kb/s G.728 LD-CELP Speech Coder for Frame Erasure Channels
Craig R. Watkins, Juin-Hwey Chen, AT&T Bell Laboratories (USA)

Reconstruction of Missing Packets for CELP-Based Speech Coders
Aamir Husain, Vladimir Cuperman, Simon Fraser University (CANADA)

A Robust Variable-Rate Speech Coder
A. Shen, B. Tang, A. Alwan, G. Pottie, University of California - Los Angeles(USA)

Wideband Speech Coding Using Multiple Codebooks and Glottal Pulses
C. McElroy, B.P. Murray, A.D. Fagan, University College - Dublin (IRELAND)

Speech Coding Using ISI Coded Quantization
Nam Phamdo, SUNY; Cheng-Chieh Lee, University of Maryland; Rajiv Laroia, AT&T Bell Laboratories (USA)

New Techniques for Multi-prototype Waveform Coding at 2.84 kb/s
I.S. Burnett, G.J. Bradley, University of Wollongong (AUSTRALIA)

Quantization of Non-Linear Predictors in Speech Coding
Jes Thyssen, Henrik Nielsen, Tele Danmark Research, Steffen Duus Hansen, Technical University of Denmark (DENMARK)

A Fast Robust Stochastic Algorithm for Vector Quantizer Design for Nonstationary Channels
B. Kövesi, S. Saoudi, J.M. Boucher, ENST/Bretagne, (FRANCE); Z. Reguly, Technical University of Budapest (HUNGARY)

Voice Quality of Interconnected PCS, Japanese Cellular, and Public Switched Telephone Networks
Spiros Dimolitsas, Franklin L. Corcoran, Channasandra Ravishanker, COMSAT Laboratories; Marion Baraniecki, INTELSAT (USA)

Objective Speech Measure for Chinese in Wireless Environment
K.H. Lam, O.C. Au, C.C. Chan, K.F. Hui, S.F. Lau, Hong Kong University of Science & Technology (HONG KONG)

WORDS SPOTTING, REJECTION, AND TOPIC IDENTIFICATION

Chair: Jay Wilpon, AT&T Bell Laboratories (USA)

A Training Procedure for Verifying String Hypotheses in Continuous Speech Recognition
R.C. Rose, B.H. Juang, C.H. Lee, AT&T Bell Laboratories (USA)

Robust Utterance Verification for Connected Digits Recognition
Mazin G. Rahim, Chin-Hui Lee, Biing-Hwang Juang, AT&T Bell Laboratories (USA)
A Hybrid Wordspotting Method for Spontaneous Speech Understanding Using Word-Based Pattern Matching and Phoneme-Based HMM
Hiroshi Kanazawa, Mitsuyoshi Tachimori, Yoichi Takebayashi, Toshiba Corporation (JAPAN) 289

Acoustic and Language Modeling of Human and Nonhuman Noises for Human-to-Human Spontaneous Speech Recognition
T. Schultz, I. Rogina, University of Karlsruhe (GERMANY) and Carnegie Mellon University (USA) 293

LVCSR Log-Likelihood Ratio Scoring for Keyword Spotting
Mitchel Weintraub, SRI International (USA) 297

Keyword Spotting Using Supervised/Unsupervised Competitive Learning
Chakib Tadj, Franck Poirier, Télécom Paris (FRANCE) 301

A Continuous Density Neural Tree Network Word Spotting System
Stephen V. Kosonocky, IBM T.J. Watson Research Center; Richard J. Mammone, Rutgers University (USA) 305

Video Mail Retrieval: The Effect of Word Spotting Accuracy on Precision
G.J.F. Jones, J.T. Foote, K. Sparck Jones, S.J. Young, Cambridge University (UK) 309

Improved Topic Spotting through Statistical Modelling of Keyword Dependencies
Jerry H. Wright, Michael J. Carey, Eluned S. Parris, Ensigna Limited (UK) 313

Topic Focusing Mechanism for Speech Recognition Based on Probabilistic Grammar and Topic Markov Model
Takeshi Kawabata, NTT Basic Research Labs (JAPAN) 317

SPEAKER RECOGNITION
Chair: S. Parthasarathy, AT&T Bell Laboratories (USA) 317

The Influence of Noise on the Speaker Recognition Performance Using the Higher Frequency Band
Shoji Hayakawa, Fumitada Itakura, Nagoya University (JAPAN) 321

Measuring Fine Structure in Speech: Application to Speaker Identification
C.R. Jankowski, Jr., T.F. Quatieri, D.A. Reynolds, MIT Lincoln Laboratory (USA) 325

The Effects of Telephone Transmission Degradations on Speaker Recognition Performance
D.A. Reynolds, M.A. Zissman, T.F. Quatieri, G.C. O'Leary, B.A. Carlson, MIT Lincoln Laboratory (USA) 329

Covariance Estimation Methods for Channel Robust Text-Independent Speaker Identification
Michael Schmidt, Herbert Gish, Angela Mielke, BBN Systems and Technologies (USA) 333

Channel and Noise Compensation for Text Dependent Speaker Verification Over Telephone
William Y. Huang, ITT Aerospace Communications; Bhaskar D. Rao, University of California (USA) 337

Testing with the Yoho CD-ROM Voice Verification Corpus
Joseph P. Campbell, Jr., U.S. Department of Defense (USA) 341

An Orthogonal Polynomial Representation of Speech Signals and Its Probabilistic Model for Text Independent Speaker Verification
Chi-Shi Liu, Ministry of Transportation and Communications; Hsiao-Chuan Wang, National Tsing Hua University (TAIWAN); Frank K. Soong, AT&T Bell Laboratories (USA); Chao-Shih Huang, Ministry of Transportation and Communications (TAIWAN) 345

Text-Dependent Speaker Verification Using Data Fusion
Kevin R. Farrell, Dictaphone Corporation (USA) 349

Neural Net Approaches to Speaker Verification: Comparison with Second Order Statistic Measures
M. Mehdi Homayounpour, CNRS/URA, (FRANCE); Gérard Chollet, IDIAP (SWITZERLAND) 353

A Subword Neural Tree Network Approach to Text-Dependent Speaker Verification
Han-Sheng Liou, Richard J. Mammone, Rutgers University (USA) 357

RECOGNITION: FEATURE ANALYSIS
Chair: Shigeki Sagayama, NTT (JAPAN) 357

Statistical Modeling of Speech Feature Vector Trajectories Based on a Piecewise Continuous Mean Path
Mark M. Thomson, University of Auckland (NEW ZEALAND) 361
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trace-Segmentation of Isolated Utterances for Speech Recognition</td>
<td>365</td>
</tr>
<tr>
<td>Euvaldo F. Cabral, Jr., University of São Paulo, (BRAZIL); Graham D. Tattersall, University of East Anglia (UK)</td>
<td></td>
</tr>
<tr>
<td>Optimal Linear Feature Transformations for Semi-Continuous Hidden Markov Models</td>
<td>369</td>
</tr>
<tr>
<td>E. Günter Schukat-Talamazzini, Joachim Hornegger, Heinrich Niemann, Universität Erlangen-Nürnberg (GERMANY)</td>
<td></td>
</tr>
<tr>
<td>Use of Generalized Dynamic Feature Parameters for Speech Recognition: Maximum Likelihood and Minimum Classification Error Approaches</td>
<td>373</td>
</tr>
<tr>
<td>C. Rathinavelu, L. Deng, University of Waterloo (CANADA)</td>
<td></td>
</tr>
<tr>
<td>A Statistical Pattern Recognition Approach to Robust Recursive Identification of Non-stationary AR Model of Speech Production System</td>
<td>377</td>
</tr>
<tr>
<td>Milan Ž. Marković, Institute of Applied Math and Electronics, Branko D. Kovačević, University of Belgrade, Milan M. Milosavljević, Institute of Applied Math and Electronics (YUGOSLAVIA)</td>
<td></td>
</tr>
<tr>
<td>The NPSpeech Activity Detection Algorithm</td>
<td>381</td>
</tr>
<tr>
<td>Joseph Pencak, Douglas Nelson, Department of Defense (USA)</td>
<td></td>
</tr>
<tr>
<td>Improved Speech Modeling and Recognition Using Multi-dimensional Articulatory States as Primitive Speech Units</td>
<td>385</td>
</tr>
<tr>
<td>L. Deng, J. Wu, H. Sameti, University of Waterloo (CANADA)</td>
<td></td>
</tr>
<tr>
<td>Speech Analysis Based on Malvar Wavelet Transform</td>
<td>389</td>
</tr>
<tr>
<td>Christophe Ris, Vincent Fontaine, Henri Leich, Faculté Polytechnique de Mons (BELGIUM)</td>
<td></td>
</tr>
<tr>
<td>Magnitude Spectral Estimation via Poisson Moments with Application to Speech Recognition</td>
<td>393</td>
</tr>
<tr>
<td>Samel Celebi, Jose C. Principe, University of Florida (USA)</td>
<td></td>
</tr>
<tr>
<td>Stochastic Perceptual Models of Speech</td>
<td>397</td>
</tr>
<tr>
<td>Nelson Morgan, University of California at Berkeley, (USA); Hervé Bourlard, Faculté Polytechnique (BELGIUM); Steven Greenberg, University of California at Berkeley; Hynek Hermansky, Oregon Graduate Institute, Su-Lin Wu, University of California at Berkeley (USA)</td>
<td></td>
</tr>
</tbody>
</table>

**TOPICS IN NOISE AND RECOGNITION**

**Chair:** Yariv Ephraim, George Mason University (USA)

- Auditory Scene Analysis and Hidden Markov Model Recognition of Speech in Noise | 401 |
  P.D. Green, M.P. Cooke, M.D. Crawford, University of Sheffield (UK)

- Speech Enhancement Based on Temporal Processing | 405 |
  Hynek Hermansky, Eric A. Wan, Carlos Avendano, Oregon Graduate Institute of Science & Technology (USA)

- A Comparative Study of Mel Cepstra and EIH for Phone Classification under Adverse Conditions | 409 |
  Sumeet Sandhu, Oded Ghitza, AT&T Bell Laboratories (USA)

- Supplementary Orthogonal Cepstral Features | 413 |
  Khaled T. Assaleh, Motorola, GSTG (USA)

- Subband Analysis for Robust Speech Recognition in the Presence of Car Noise | 417 |
  Engin Erzin, Bilkent University; A. Enis Çetin, Koç University; Yasemin Yardımcı, Boğaziçi University (TURKEY)

- Robust Speech Feature Extraction Using SBCOR Analysis | 421 |
  Shoji Kajita, Fumitada Itakura, Nagoya University (JAPAN)

- Methods for Improved Speech Recognition Over Telephone Lines | 425 |
  Alfred Hauenstein, Erwin Marschall, Siemens AG (GERMANY)

- New HOS-Based Parameter Estimation Methods for Speech Recognition in Noisy Environments | 429 |
  Asunción Moreno, Sergio Tortola, Josep Vidal, José A. R. Fonollosa, Universitat Politècnica de Catalunya (SPAIN)

- Noise Compensation for Speech Recognition in Car Noise Environments | 433 |
  Ruikang Yang, Petri Haavisto, Nokia Research Center (FINLAND)

- Speech Recognition in Impulsive Noise | 437 |
  S.V. Vaseghi, B.P. Milner, University of East Anglia (UK)
Chair: Robin Rohlicek, BBN, Inc. (USA)

Speaker-independent Phone Modeling Based on Speaker-dependent HMMs' Composition and Clustering
Tetsuo Kosaka, Shoichi Matsunaga, ATR Interpreting Telecommunications Research Laboratories; Mikio Kuraoka, Toyohashi University of Technology (JAPAN)

Using Morphology Towards Better Large-Vocabulary Speech Recognition Systems
P. Geutner, Universitat Karlsruhe (GERMANY)

Optimal Splitting of HMM Gaussian Mixture Components with MMIE Training
Yves Normandin, Centre de Recherche informatique de Montréal (CANADA)

Dictionary Learning: Performance through Consistency
Tilo Sloboda, Universitat Karlsruhe (GERMANY)

Incremental MAP Estimation of HMMs for Efficient Training and Improved Performance
Yoshihiko Gotoh, Brown University (USA); Michael M. Hochberg, Cambridge University (UK); Daniel J. Mashao, Harvey F. Silverman, Brown University (USA)

Discrete MMI Probability Models for HMM Speech Recognition
J.T. Foote, Cambridge University (UK)

Global Discrimination for Neural Predictive Systems Based on N-Best Algorithm
Abdelhamid Mellouk, LRI, UA 410 CNRS; Patrick Gallinari, LA FORIA, UA CNRS 1095 (FRANCE)

Enhancement of Discriminative Capabilities of HMM Based Recognizer through Modification of Viterbi Algorithm
Jianming Song, The University of Woolongong (AUSTRALIA)

A Generalization of the Baum Algorithm to Functions on Non-linear Manifolds
D. Kanevsky, IBM T.J. Watson Research Center (USA)

Data-Driven Codebook Adaptation in Phonetically Tied SCHMMs
Thomas Kemp, Universitat Karlsruhe (GERMANY)

Chair: Thomas E. Tremain, U.S. Department of Defense (USA)

NATO STANAG 4479: A Standard for an 800 bps Vocoder and Channel Coding in HF-ECCM System
B. Mouy, P. de La Noue, G. Goudezeune, Thomson CSF-RGS (FRANCE)

Harmonic and Noise Coding of LPC Residuals with Classified Vector Quantization
Masayuki Nishiguchi, Jun Matsumoto, Sony Corporation (JAPAN)

Progress Towards a New Government Standard 2400 bps Voice Coder

Variable Dimension Spectral Coding of Speech at 2400 bps and Below with Phonetic Classification
Amitava Das, Allen Gersho, University of California - Santa Barbara (USA)

Spectral Excitation Coding of Speech At 2.4 kb/s
V. Cuperman, P. Lupini, B. Bhattacharya, Simon Fraser University (CANADA)

A Robust 2400 bps Subband LPC Vocoder
P. A. Laurent, P. de La Noue, Thomson CSF-RGS (FRANCE)

Band-Widened Harmonic Vocoder at 2 to 4 kbps
Gao Yang, G. Zanellato, Lernou & Hauspie Speech Products; H. Leich, Faculté Polytechnique de Mons (BELGIUM)

A Speech Coder Based on Decomposition of Characteristic Waveforms
W. Bastiaan Kleijn, Jesper Haagen, AT&T Bell Laboratories (USA)

Speech Compression Using Pitch Synchronous Interpolation
R. Taori, R.J. Sluijter, E. Kathmann, Philips Research Laboratories (THE NETHERLANDS)

Pitch-Synchronous Multi-Band (PSMB) Speech Coding
Haiyun Yang, Soo-Nege Koh, Pratap Sivaprakasapillai, Nanyang Technological University (SINGAPORE)
RECOGNITION: MODELING STRUCTURES

Chair: Hsiao-Wuen Hon, Apple ISS Research Centre, University of Singapore (SINGAPORE)

Four-Level Tied-Structure for Efficient Representation of Acoustic Modeling 520
Satoshi Takahashi, Shigeki Sagayama, NTT Human Interface Laboratories (JAPAN)

Application of Clustering Techniques to Mixture Density Modelling for Continuous-Speech Recognition 524
Christian Dugast, Peter Beyerlein, Reinhold Haeb-Umbach, Philips Research Laboratories (GERMANY)

Context Dependent Phonetic Duration Models for Decoding Conversational Speech 528
Michael D. Monkowski, Michael A. Picheny, P. Srinivasa Rao, IBM - TJ Watson Research Center (USA)

A Unified Way in Incorporating Segmental Feature and Segmental Model into HMM 532
Jun He, Henri Leich, Faculté Polytechnique de Mons (BELGIUM)

Experimental Evaluation of Segmental HMMs 536
Wendy J. Holmes, Martin J. Russell, DRA Malvern (UK)

Improved Acoustic Modeling for Speech Recognition Using 2D Markov Random Fields 540
Helmut Lucke, ATR/ITL (JAPAN)

Structured Markov Models for Speech Recognition 544
F. Wolfertstetter, G. Ruske, Munich University of Technology (GERMANY)

Robust Parametric Modeling of Durations in Hidden Markov Models 548
David Burshtein, Tel-Aviv University (ISRAEL)

Improved Decision Trees for Phonetic Modeling 552
Roland Kuhn, Ariane Lazarides, Yves Normandin, Julie Brousseau, CRIM (CANADA)

High Speed Speech Recognition Using Tree-Structured Probability Density Function 556
Takao Watanabe, Koichi Shinoda, Keizaburo Takagi, Ken-ichi Iso, NEC Corporation (JAPAN)

RECOGNITION: SEARCH TECHNIQUES

Chair: Al Alewa, Microsoft Corporation (USA)

A Fast Segmental Viterbi Algorithm for Large Vocabulary Recognition 560
P. Laface, C. Vair, Politecnico di Torino, L. Fissore, CSELT (ITALY)

Searching with a Transcription Graph 564
Z. Li, P. Kenny, D. O'Shaughnessy, Université du Québec (CANADA)

On the Use of Stochastic Inference Networks for Representing Multiple Word Pronunciations 568
Renato De Mori, Charles Snow, Michael Galler, McGill University School of Computer Science (CANADA)

A Tree Search Strategy for Large-Vocabulary Continuous Speech Recognition 572
P.S. Gopalakrishnan, L.R. Bahl, IBM; R.L. Mercer, Renaissance Technologies (USA)

Lattice-Based Search Strategies for Large Vocabulary Speech Recognition 576
F. Richardson, M. Ostendorf, Boston University; J.R. Rohlicek, Bolt Beranek & Newman, Inc. (USA)

On Using a priori Segmentation of the Speech Signal in an N-Best Solutions Post-processing 580
T. Moudenc, D. Jouvet, J. Monné, France Télécom (FRANCE)

Time-Synchronous Continuous Speech Recognizer Driven by a Context-Free Grammar 584
Tohru Shimizu, ATR/ITL; Seikou Monzen, Yamagata University; Harald Singer, Shoichi Matsunaga, ATR/ITL (JAPAN)

Language Model Representations for Beam-Search Decoding 588
Giuliano Antoniol, Fabio Brugnara, Mauro Cettolo, Marcello Federico, I.R.S.T. (ITALY)

A Lower-Complexity Viterbi Algorithm 592
Sarvar Patel, Bellcore (USA)

Efficient Search Using Posterior Phone Probability Estimates 596
Steve Renals, University of Sheffield, Mike Hochberg, University of Cambridge (UK)
PROSODY FOR SYNTHESIS & RECOGNITION

Chair: Yoshinori Sagisaka, AT&T Bell Laboratories (USA)

Timing Patterns in Fluent and Disfluent Spontaneous Speech
Douglas O'Shaughnessy, Université du Québec (CANADA)

Stochastic Modeling of Pause Insertion Using Context-Free Grammar
Shigeru Fujio, Yoshinori Sagisaka, Norio Higuchi, ATR-ITL (JAPAN)

Automatic Classification of Pitch Movements via MLP-Based Estimation of Class Probabilities
Louis F. M. ten Bosch, Institute for Perception Research, (THE NETHERLANDS)

On the Effects of Speech Rate in Large Vocabulary Speech Recognition Systems
Matthew A. Siegler, Richard M. Stern, Carnegie Mellon University (USA)

A Prosodic Model of Mandarin Speech and Its Application to Pitch Level Generation for Text-to-Speech
Shaw-Hwa Hwang, Sin-Horng Chen, National Chiao Tung University (REPUBLIC OF CHINA)

Prosodic Cues to Word Usage
Karen Ward, David G. Novick, Oregon Graduate Institute of Science & Technology (USA)

Automatic Prosodic Segmentation by $F_0$ Clustering Using Superpositional Modeling
Mitsuru Nakai, Tohoku University, Harald Singer, Yoshinori Sagisaka, ATR-ITL; Hiroshi Shimodaira, JAIST (JAPAN)

Duration Modeling in Large Vocabulary Speech Recognition
Anastasios Anastasakos, Northeastern University, Richard Schwartz, Han Shu, BBN Systems and Technologies (USA)

Speaker-Independent Automatic Classification of Thai Tones in Connected Speech by Analysis-Synthesis Method
Siripong Potisuk, Mary P. Harper, Jackson T. Gandour, Purdue University (USA)

SPEECH SYNTHESIS & PRODUCTION

Chair: Kathleen Cummings, Georgia Institute of Technology (USA)

Speech Synthesis System Based on a Variable Decimation/Interpolation Factor
F. M. Giménez de los Galanes, Universitat Politècnica de Madrid; M. H. Savoji, Universidad de Cantabria; J. M. Pardo, Universitat Politècnica de Madrid (SPAIN)

Automatic Speech Synthesiser Parameter Estimation Using HMMs
R. E. Donovan, P. C. Woodland, Cambridge University (UK)

Speaker Modification with LPC Pole Analysis
Janet Slifka, Systems Research Laboratories, Timothy R. Anderson, Armstrong Laboratory (USA)

Synthesizing Styled Speech Using the Klatt Synthesizer
Janet C. Rutledge, Northwestern University; Kathleen E. Cummings, Daniel A. Lambert, Mark A. Clements, Georgia Institute of Technology (USA)

Acoustical Measurements of the Vocal Tract Area Function: Sensitivity Analysis and Experiments
Hani Yehia, Nagoya University; Masaaki Honda, NTT Basic Research Laboratories; Fumitada Itakura, Nagoya University (JAPAN)

Shape-Invariant Pitch-Synchronous Text-to-Speech Conversion
Eduardo R. Banga, Carmen García-Mateo, Universidad de Vigo, (SPAIN)

Speech Parameter Generation from HMM Using Dynamic Features
Keiichi Tokuda, Takao Kobayashi, Satoshi Imai, Tokyo Institute of Technology (JAPAN)

A Source Generator Based Modeling Framework for Synthesis of Speech Under Stress
Sahar E. Bou-Ghazale, John H. L. Hansen, Duke University (USA)

MBE Synthesis of Speech Coded in LPC Format
K. F. Lam, C. F. Chan, City Polytechnic of Hong Kong (HONG KONG)

Modeling Speech Production Using Yee's Finite Difference Method
Kathleen E. Cummings, Georgia Institute of Technology; James G. Maloney, Georgia Technical Research Institute; Mark A. Clements, Georgia Institute of Technology (USA)
SPEAKER ADAPTATION

Chair: C.H. Lee, AT&T Bell Laboratories (USA)

Batch, Incremental and Instantaneous Adaptation Techniques for Speech Recognition 676
G. Zavaliagkos, Northeastern University; R. Schwartz, J. Makhoul, BBN Systems and Technologies (USA)

Speaker Adaptation Using Combined Transformation and Bayesian Methods 680
Vassilios Digalakis, Leonardo Neumeyer, SRI International (USA)

Rapid Speaker Adaptation Using Model Prediction 684
S. M. Ahadi, P. C. Woodland, Cambridge University (UK)

Speaker Adaptation Based on Transfer Vector Field Smoothing Using Maximum a posteriori Probability Estimation 688
Masahiro Tonomura, Tetsuo Kosaka, Shoichi Matsunaga, ATR Interpreting Telecommunications Research Labs (JAPAN)

Experiments Using Data Augmentation for Speaker Adaptation 692
Jerome R. Bellegarda, Apple Computer Inc.; Peter V. de Souza, David Nahamoo, Mukund Padmanabhan, Michael A. Picheny, Lalit R. Bahl, IBM (USA)

Vector-Field-Smoothed Bayesian Learning for Incremental Speaker Adaptation 696
Jun-ichi Takahashi, Shigeki Sagayama, NTT Human Interface Laboratories (JAPAN)

A Speaker Adaptation Technique Using Linear Regression 700
S.J. Cox, University of East Anglia (UK)

Speaker Adaptation Based on Spectral Normalization and Dynamic HMM Parameter Adaptation 704
Ming-Whei Feng, GTE Laboratories, Inc. (USA)

On-line Bayes Adaptation of SCHMM Parameters for Speech Recognition 708
Qiang Huo, Chorkin Chan, University of Hong Kong (HONG KONG)

Iterative Self-Learning Speaker and Channel Adaptation under Various Initial Conditions 712
Yunxin Zhao, University of Illinois at Urbana-Champaign (USA)

SPECTRAL QUANTIZATION

Chair: Costas Xydeas, University of Manchester (UK)

Fast and Low-Complexity LSF Quantization Using Algebraic Vector Quantizer 716
Minjie Xie, Jean-Pierre Adoul, University of Sherbrooke (CANADA)

Low Cost Vector Quantization Methods for Spectral Coding in Low Rate Speech Coders 720
H.R. Sadegh Mohammadi, W.H. Holmes, University of New South Wales (AUSTRALIA)

Matrix Product Quantization for Very-Low-Rate Speech Coding 724
Stefan Bruhn, Technical University of Berlin (GERMANY)

An Intrinsically Reliable and Fast Algorithm to Compute the Line Spectrum Pairs (LSP) in Low Bit Rate CELP Coding 728
A. Gooalic, S. Saoudi, ENST-Bretagne (FRANCE)

Spectral Dynamics is More Important Than Spectral Distortion 732
H. Petter Knagenhjelm, W. Bastiaan Kleijn, AT&T Bell Laboratories (USA)

Efficient Quantization of LSF Parameters Using Classified SVQ Combined with Conditional Splitting 736
Dong-il Chang, Young-kwon Cho, Souguil Ann, Seoul National University (KOREA)

Efficient Coding of LSP Parameters Using Split Matrix Quantisation 740
C.S. Xydeas, C. Papanastasiou, University of Manchester (UK)

How Good Is Your β? Observations on VQ Training Ratios 744
John S. Collura, Thomas E. Tremain, U.S. Department of Defense (USA)

Variable Rate Spectral Quantization for Phonetically Classified CELP Coding 748
Roar Hagen, Chalmers University of Technology; Erdal Paksoy, Allen Gersho, University of California (USA)

Optimal Distortion Measures for the High Rate Vector Quantization of LPC Parameters 752
William R. Gardner, University of California-San Diego; Bhaskar D. Rao, Qual Comm, Inc. (USA)
SPEECH ANALYSIS

Chair: Paul Mermelstein, INRS-Télécom (FRANCE)

Harmonics Tracking and Pitch Extraction Based on Instantaneous Frequency  
Toshihiko Abe, Takao Kobayashi, Satoshi Imai, Tokyo Institute of Technology (JAPAN)

Decomposition of Speech Signals into Deterministic and Stochastic Components  
C. d'Alessandro, LIMSI-CNRS (FRANCE); B. Yegnanarayana, Indian Institute of Technology (INDIA); V. Darsinos, University of Patras (GREECE)

Modeling and Processing Speech with Sums of AM-FM Formant Models  
Shan Lu, Peter C. Doerschuk, Purdue University (USA)

On the Statistical Properties of Line Spectrum Pairs  
J.S. Erkelens, P.M.T. Broersen, Delft University of Technology (THE NETHERLANDS)

Individual Variations in Glottal Characteristics of Female Speakers  
Helen M. Hanson, Harvard University (USA)

A Robust Method for Determining Instants of Major Excitations in Voiced Speech  
B. Yegnanarayana, Indian Institute of Technology (INDIA); R.L.H.M. Smits, Institute for Perception Research (THE NETHERLANDS)

Interpolation of LPC Spectra via Pole Shifting  
Vladimir Goncharoff, Maureen Kaine-Krolak, University of Illinois at Chicago (USA)

Speech Formant Frequency and Bandwidth Tracking Using Multiband Energy Demodulation  
Alexandros Potamianos, Petros Maragos, Georgia Institute of Technology (USA)

Nonlinear Prediction for Speech Coding Using Radial Basis Functions  
Fernando Díaz-de-María, Universidad de Cantabria; Aníbal R. Figueiras-Vidal, Universidad Politécnica de Madrid (SPAIN)

Recognition of Unvoiced Stops from Their Time-Frequency Representation  
Maria Rangoussi, Anastasios Delopoulos, National Technical University of Athens (GREECE)

SPEECH ENHANCEMENT & NOISE REDUCTION

Chair: John H.L. Hansen, Duke University (USA)

Speech Enhancement Based on Masking Properties of the Auditory System  
Nathalie Virag, Swiss Federal Institute of Technology (SWITZERLAND)

Optimizing Speech Enhancement by Exploiting Masking Properties of the Human Ear  
A. Akbari Azirani, R. Le Bouquin Jeanès, G. Faucon, Université de Rennes I (FRANCE)

A Spectrally-Based Signal Subspace Approach for Speech Enhancement  
Yariv Ephraim, Harry L. VanTrees, George Mason University (USA)

Real-Time Implementation of HMM-Based MMSE Algorithm for Speech Enhancement in Hearing Aid Applications  
H. Sheikzadeh, Université de Waterloo; R.L. Brennan, Unitron Industries, Ltd.; H. Sameti, University of Waterloo (CANADA)

New Methods for Adaptive Noise Suppression  
Levent Arslan, Alan McCree, Vishu Viswanathan, Texas Instruments (USA)

Single-Sensor Speech Enhancement Using a Soft-Decision/Variable Attenuation Algorithm  
E. Bryan George, Lockheed Sanders, Inc. (USA)

Speech Enhancement Using a Ternary-Decision Based Filter  
T.S. Sun, S. Nandkumar, J. Carmody, J. Rothweiler, A. Goldschen, N. Russell, S. Mpasi, P. Green, Martin Marietta Laboratories (USA)

Signal Modeling Enhancements for Automatic Speech Recognition  
Zaki B. Nossair, Peter L. Silsbee, Stephen A. Zahorian, Old Dominion University (USA)

Co-Channel Speaker Separation  
David P. Morgan, E. Bryan George, Texas Instruments; Leonard T. Lee, Lockheed Sanders, Inc.; Stephen M. Kay, University of Rhode Island (USA)

Speech Enhancement Based on the Generalized Dual Excitation Model with Adaptive Analysis Window  
Chang D. Yoo, Jae S. Lim, Massachusetts Institute of Technology (USA)
SPECIAL TOPICS IN SPEECH RECOGNITION

Chair: K. Paliwal, Griffith University

Foreign Accent Classification Using Source Generator Based Prosodic Features 836
John H. L. Hansen, Levent M. Arslan, Duke University (USA)

Automatic Transcription of Unknown Words in a Speech Recognition System 840
R. Haeb-Umbach, P. Beyerlein, E. Thelen, Philips Research Laboratories-Aachen (GERMANY)

An Evaluation of an Adaptive Multichannel System for Speech Enhancement with Automatic Phase Alignment 844
Silvana L. do N. Cunha Costa, Benedito G. Aguiar Neto, Universidade Federal da Paraiba (BRAZIL)

Knowing Who to Listen to in Speech Recognition: Visually Guided Beamforming 848
Udo Bub, Martin Hunke, Alex Waibel, Carnegie Mellon University (USA)

An N-Best Strategy, Dynamic Grammars and Selectively Trained Neural Networks for Real-Time Recognition of Continuously Spelled Names Over the Telephone 852
Jean-Claude Junqua, Stephane Valente, Speech Technology Laboratory, (USA); Dominique Fohr, Jean-François Mari, CRIN/INRIA, (FRANCE)

Language Models for a Spelled Letter Recognizer 856
Martin Betz, Hermann Hild, Universitat Karlsruhe (GERMANY)

Hands Free Continuous Speech Recognition in Noisy Environment Using a Four Microphone Array 860
D. Giuliani, M. Matassoni, M. Omologo, P. Svaizer, IRST (ITALY)

A New Method for Automatic Generation of Speaker-Dependent Phonological Rules 864
Toru Imai, Akio Ando, Eiichi Miyasaka, NHK Science & Technology Research Labs (JAPAN)

Enhancing Automatic Speech Recognition with an Ultrasonic Lip Motion Detector 868
David L. Jennings, Dennis W. Ruck, AFIT/ENG (USA)

Classification and Clustering of Stop Consonants via Nonparametric Transformations and Wavelets 872
Basilis Gidas, Brown University; Alejandro Murua, University of Chicago (USA)