# TABLE OF CONTENTS

## Volume I

### SP-L1: VOICE CONVERSION AND MORPHING ALGORITHMS FOR TTS SYSTEMS

**SP-L1.1: NON-PARALLEL TRAINING FOR VOICE CONVERSION BY MAXIMUM LIKELIHOOD CONSTRAINED ADAPTATION**  
Athanasios Mouchtaris, Jan Van der Spiegel, University of Pennsylvania, United States; Paul Mueller, Corticon Inc., United States

**SP-L1.2: SPEAKING STYLE ADAPTATION USING CONTEXT CLUSTERING DECISION TREE FOR HMM-BASED SPEECH SYNTHESIS**  
Junichi Yamagishi, Makoto Tachibana, Takashi Masuko, Takao Kobayashi, Tokyo Institute of Technology, Japan

**SP-L1.3: HIGH QUALITY VOICE MORPHING**  
Hui Ye, Steve Young, Cambridge University, United Kingdom

**SP-L1.4: ALGORITHM AMALGAM: MORPHING WAVEFORM BASED METHODS, SINOISIDAL MODELS AND STRAIGHT**  
Hideki Kawahara, Hideki Banno, Toshio Irino, Wakayama University, Japan; Parham Zolfaghari, NTT Communication Science Laboratories, Japan

**SP-L1.5: VOICE CHARACTERISTICS CONVERSION FOR TTS USING REVERSE VTLN**  
Matthias Eichner, Matthias Wolff, Rüdiger Hoffmann, Dresden University of Technology, Germany

**SP-L1.6: VOICE CONVERSION THROUGH TRANSFORMATION OF SPECTRAL AND INTONATION FEATURES**  
Dimitrios Rentzos, Saeed Vaseghi, Qin Yan, Brunel University, United Kingdom; Ching-Hsiang Ho, Fortune Institute of Technology, Taiwan

### SP-L2: MODELING APPROACHES IN SPEAKER RECOGNITION

**SP-L2.1: DISCRIMINATIVE TRAINING FOR SPEAKER IDENTIFICATION BASED ON MAXIMUM MODEL DISTANCE ALGORITHM**  
Q. Y. Hong, S. Kwong, City University of Hong Kong, Hong Kong SAR of China

**SP-L2.2: PARAMETER SHARING AND MINIMUM CLASSIFICATION ERROR TRAINING OF MIXTURES OF FACTOR ANALYZERS FOR SPEAKER IDENTIFICATION**  
Hiroyoshi Yamamoto, Yoshikoko Nankaku, Nagoya Institute of Technology, Japan; Chiyomi Miyajima, Nagoya University, Japan; Keiichi Tokuda, Tadashi Kitamura, Nagoya Institute of Technology, Japan

**SP-L2.3: DISCOVERING RELATIONS AMONG DISCRIMINATIVE TRAINING OBJECTIVES**  
Qi Li, LcT, Inc., United States

**SP-L2.4: DISENTANGLING SPEAKER AND CHANNEL EFFECTS IN SPEAKER VERIFICATION**  
Patrick Kenny, Pierre Dumouchel, Centre de recherche informatique de Montreal, Canada

**SP-L2.5: GENERALIZED LOCALLY RECURRENT PROBABILISTIC NEURAL NETWORKS FOR TEXT-INDEPENDENT SPEAKER VERIFICATION**  
Todor Ganchev, Nikos Fakotakis, Dimitris Tasoulis, Michael Vrahatis, University of Patras, Greece
SP-L2.6: DISCRIMINATION POWER WEIGHTED SUBWORD-BASED SPEAKER VERIFICATION
Siu-Man Chan, Man-Hung Siu, Hong Kong University of Science and Technology, Hong Kong SAR of China

SP-L3: DISTRIBUTED SPEECH RECOGNITION

SP-L3.1: SOFT DECODING STRATEGIES FOR DISTRIBUTED SPEECH RECOGNITION OVER IP NETWORKS
Antonio Cardenal-López, Laura Docio-Fernández, Carmen García-Mateo, University of Vigo, Spain

SP-L3.2: THE ETSI EXTENDED DISTRIBUTED SPEECH RECOGNITION (DSR) STANDARDS: SERVER-SIDE SPEECH RECONSTRUCTION
Tenkasi Ramabadran, Motorola, United States; Alexander Sorin, IBM, Israel; Michael McLaughlin, Motorola Labs, United States; Dan Chazan, IBM, Israel; David Pearce, Motorola, United Kingdom; Ron Hoory, IBM, Israel

SP-L3.3: A SUBVECTOR-BASED ERROR CONCEALMENT ALGORITHM FOR SPEECH RECOGNITION OVER MOBILE NETWORKS
Zheng-Hua Tan, Paul Dalsgaard, Børge Lindberg, Aalborg University, Denmark

SP-L3.4: A COMPLEXITY REDUCTION OF ETSI ADVANCED FRONT-END FOR DSR
Jin-Yu Li, University of Science and Technology of China, China; Bo Liu, Ren-Hua Wang, Li-Rong Dai, iFlytek Speech Lab, University of Science and Technology, China

SP-L3.5: ROBUST SPEECH RECOGNITION TECHNIQUES EVALUATION FOR TELEPHONY SERVER BASED IN-CAR APPLICATIONS
Lionel Delphin-Poulat, France Télécom R&D, France

SP-L3.6: EFFICIENT AND ROBUST DISTRIBUTED SPEECH RECOGNITION (DSR) OVER WIRELESS FADING CHANNELS: 2D-DCT COMPRESSION, ITERATIVE BIT ALLOCATION, SHORT BCH CODE AND INTERLEAVING
Wei-hao Hsu, Lin-shan Lee, National Taiwan University, Taiwan

SP-L4: HIGHER-LEVEL KNOWLEDGE IN SPEAKER RECOGNITION

SP-L4.1: HIGH-LEVEL SPEAKER VERIFICATION USING SUPPORT VECTOR MACHINES
William Campbell, Joseph Campbell, Doug Reynolds, Doug Jones, Timothy Leek, MIT Lincoln Laboratory, United States

SP-L4.2: USING HAAR TRANSFORMED VOCAL SOURCE INFORMATION FOR AUTOMATIC SPEAKER RECOGNITION
Nengheng Zheng, P. C. Ching, Chinese University of Hong Kong, Hong Kong SAR of China

SP-L4.3: TEXT-INDEPENDENT SPEAKER RECOGNITION BY COMBINING SPEAKER-SPECIFIC GMM WITH SPEAKER ADAPTED SYLLABLE-BASED HMM
Seiichi Nakagawa, Wei Zhang, Mitsuo Takahashi, Toyoashi University of Technology, Japan

SP-L4.4: APPLYING ARTICULATORY FEATURES TO TELEPHONE-BASED SPEAKER VERIFICATION
Ka-Yee Leung, Man-Wai Mak, Hong Kong Polytechnic University, Hong Kong SAR of China; Sun-Yuan Kung, Princeton University, United States

SP-L4.5: SPEAKER IDENTIFICATION USING SUPRA-SEGMENTAL PITCH PATTERN DYNAMICS
Farhad Farahani, Panayiotis Georgiou, Shrikanth Narayanan, University of Southern California, United States

SP-L4.6: IMPROVEMENT OF SPEAKER RECOGNITION BY COMBINING RESIDENTIAL AND PROSODIC FEATURES WITH ACOUSTIC FEATURES
Shi-Han Chen, Hsiao-Chuan Wang, National Tsing Hua University, Taiwan
SP-L5: PITCH AND TONE BASED SPEECH ANALYSIS

SP-L5.1: PITCH PREDICTION FROM MFCC VECTORS FOR SPEECH RECONSTRUCTION
Xu Shao, Ben Milner, University of East Anglia, United Kingdom

SP-L5.2: ALGORITHM FOR AUTOMATIC GLOTTAL WAVEFORM ESTIMATION WITHOUT THE RELIANCE ON PRECISE GLOTTAL CLOSURE INFORMATION
Elliot Moore, Mark Clements, Georgia Institute of Technology, United States

SP-L5.3: TONE RECOGNITION WITH FRACTIONIZED MODELS AND OUTLINED FEATURES
Ye Tian, Jian-Lai Zhou, Min Chu, Eric Chang, Microsoft Research Asia, China

SP-L5.4: EXTRACTION OF PITCH IN ADVERSE CONDITIONS
Mahadeva Prasanna S. R., Yegnanarayana B., Indian Institute of Technology, Madras, India

SP-L5.5: WEIGHTED AUTOCORRELATION-BASED F0 ESTIMATION FOR DISTANT-TALKING INTERACTION WITH A DISTRIBUTED MICROPHONE NETWORK
Luca Armani, Maurizio Omologo, ITC-irst, Italy

SP-L5.6: A NOVEL METHOD FOR COMPUTATION OF PERIODICITY, APERIODICITY AND PITCH OF SPEECH SIGNALS
Om Deshmukh, Jawahar Singh, Carol Espy-Wilson, University of Maryland, College Park, United States

SP-L6: FEATURE ANALYSIS FOR SPEECH RECOGNITION

SP-L6.1: NON-UNIFORM SPEAKER NORMALIZATION USING AFFINE-TRANSFORMATION
Bharath Kumar SV, General Electric - Global Research, India; Umesh S., Rohit Sinha, Indian Institute of Technology, India

SP-L6.2: PRODUCT OF POWERSPECTRUM AND GROUP DELAY FUNCTION FOR SPEECH RECOGNITION
Donglai Zhu, Kuldip. K Paliwal, Griffith University, Australia

SP-L6.3: THE ETSI EXTENDED DISTRIBUTED SPEECH RECOGNITION (DSR) STANDARDS: CLIENT SIDE PROCESSING AND TONAL LANGUAGE RECOGNITION EVALUATION
Alexander Sorin, IBM Labs, Israel; Tenkasi Ramabadran, Motorola Labs, United States; Dan Chazan, Ron Hoory, IBM Labs, Israel; Michael McLaughlin, David Pearce, Motorola Labs, United States; Fan Wang, IBM Labs, China; Yuxin Zhang, Motorola Labs, China

SP-L6.4: ROBUST SPEECH FEATURE EXTRACTION BY GROWTH TRANSFORMATION IN REPRODUCING KERNEL HILBERT SPACE
Shantanu Chakrabarty, Yunbin Deng, Gert Cauwenberghs, Johns Hopkins University, United States

SP-L6.5: DIMENSIONALITY REDUCTION USING MCE-OPTIMIZED LDA TRANSFORMATION
Xiao-Bing Li, Jin-Yu Li, Ren-Hua Wang, University of Science and Technology of China, China

SP-L6.6: SPEECH FEATURE EXTRACTION METHOD REPRESENTING PERIODICITY AND APERIODICITY IN SUBBANDS FOR ROBUST SPEECH RECOGNITION
Kenataro Ishizuka, Noboru Miyazaki, NTT Corporation, Japan

SP-L7: QUANTIZATION TECHNIQUES IN SPEECH CODING

SP-L7.1: LOW-COMPLEXITY PREDICTIVE TRELLIS CODED QUANTIZATION OF WIDEBAND SPEECH LSF PARAMETERS
Yongwon Shin, Samsung Electronics Co. Ltd., Republic of Korea; Sangwon Kang, Hanyang University, Republic of Korea; Thomas R. Fischer, Washington State University, United States; Changyong Son, Yongbeom Lee, Samsung Advanced Institute of Technology, Republic of Korea
SP-L7.2: MULTIPLE FRAME BLOCK QUANTISATION OF LINE SPECTRAL FREQUENCIES USING GAUSSIAN MIXTURE MODELS
Kuldip K Paliwal, Stephen So, Griffith University, Australia

SP-L7.3: VARIABLE-DIMENSION QUANTIZATION OF SINUSOIDAL AMPLITUDES USING GAUSSIAN MIXTURE MODELS
Jonas Lindblom, Per Hedelin, Chalmers University of Technology, Sweden

SP-L7.4: ON SPLIT QUANTIZATION OF LSF PARAMETERS
Fredrik Nordén, Aalborg University, Denmark; Thomas Eriksson, Chalmers University of Technology, Sweden

SP-L7.5: IMPROVED QUANTIZATION STRUCTURES USING GENERALIZED HMM MODELLING WITH APPLICATION TO WIDEBAND SPEECH CODING
Ethan Duni, Anand Subramaniam, Bhaskar Rao, University of California, San Diego, United States

SP-L7.6: WAVEFORM QUANTIZATION OF SPEECH USING GAUSSIAN MIXTURE MODELS
Jonas Samuelsson, Royal Institute of Technology (KTH), Sweden

SP-L8: ACOUSTIC MODELING: NEW SEARCH FEATURES AND SUPERVISED TRAINING

SP-L8.1: EFFECTS OF TRANSCRIPTION ERRORS ON SUPERVISED LEARNING IN SPEECH RECOGNITION
Ram Sundaram, Conversay, United States; Joseph Picone, Mississippi State University, United States

SP-L8.2: COMBINATION OF HIDDEN MARKOV MODELS WITH DYNAMIC TIME WARPING FOR SPEECH RECOGNITION
Scott Axelrod, Benoit Maison, IBM T. J. Watson Research Center, United States

SP-L8.3: JOINT DECODING FOR PHONEME-GRAPHEME CONTINUOUS SPEECH RECOGNITION
Matthias Magimai-Doss, Samy Bengio, Herve Bourlard, Dalle Molle Institute for Artificial Intelligence, Switzerland

SP-L8.4: A LOCALLY WEIGHTED DISTANCE MEASURE FOR EXAMPLE BASED SPEECH RECOGNITION
Mathias De Wachter, Kris Demuynck, Patrick Wambacq, Dirk Van Compernolle, Katholieke Universiteit Leuven, Belgium

SP-L8.5: LIGHT SUPERVISION IN ACOUSTIC MODEL TRAINING
Long Nguyen, Bing Xiang, BBN Technologies, United States

SP-L8.6: LIGHTLY SUPERVISED ACOUSTIC MODEL TRAINING USING CONSENSUS NETWORKS
Langzhou Chen, Lori Lamel, Jean-Luc Gauvain, LIMSI-CNRS, France

SP-L9: ROBUST FEATURES FOR SPEECH RECOGNITION

SP-L9.1: SPECTRAL ENTROPY BASED FEATURE FOR ROBUST ASR
Hemant Misra, Shajith Ikbal, Hervé Bourlard, Hynek Hermansky, IDIAP, Switzerland

SP-L9.2: HIGHER ORDER CEPSTRAL MOMENT NORMALIZATION (HOCMN) FOR ROBUST SPEECH RECOGNITION
Chang-wen Hsu, Lin-shan Lee, National Taiwan University, Taiwan

SP-L9.3: ROBUSTNESS OF SPEECH RECOGNITION USING GENETIC ALGORITHMS AND A MEL-CEPSTRAL SUBSPACE APPROACH
Sid-Ahmed Selouani, Université de Moncton, Canada; Douglas O’Shaughnessy, INRS-EMT, Canada

SP-L9.4: PHASE AUTOCORRELATION (PAC) FEATURES IN ENTROPY BASED MULTI-STREAM FOR ROBUST SPEECH RECOGNITION
Shajith Ikbal, Hemant Misra, Hervé Bourlard, Hynek Hermansky, IDIAP, Switzerland
SP-L9.5: CEPSTRAL GAIN NORMALIZATION FOR NOISE ROBUST SPEECH RECOGNITION

Shingo Yoshizawa, Noboru Hayasaka, Naoya Wada, Yoshikazu Miyanaga, Hokkaido University, Japan

SP-L9.6: ROBUST SPEECH RECOGNITION USING CEPSTRAL DOMAIN MISSING DATA TECHNIQUES AND NOISY MASKS

Hugo Van hamme, Katholieke Universiteit Leuven, Belgium

SP-L10: MULTICHANNEL SPEECH ENHANCEMENT

SP-L10.1: OPTIMAL BLIND SEPARATION OF CONVOLUTIVE AUDIO MIXTURES WITHOUT TEMPORAL CONSTRAINTS

Kostas Kokkinakis, Asoke K. Nandi, University of Liverpool, United Kingdom

SP-L10.2: MICROPHONE ARRAY POST-FILTER FOR SEPARATION OF SIMULTANEOUS NON-STATIONARY SOURCES

Jean-Marc Valin, Jean Rouat, François Michaud, University of Sherbrooke, Canada

SP-L10.3: OVERDETERMINED BLIND SEPARATION FOR CONVOLUTIVE MIXTURES OF SPEECH BASED ON MULTISTAGE ICA USING SUBARRAY PROCESSING

Tsuyoki Nishikawa, Hiroshi Abe, Hiroshi Saruwatari, Kiyohiro Shikano, Nara Institute of Science and Technology, Japan

SP-L10.4: SPEECH ENHANCEMENT BASED ON A COMBINED MULTI-CHANNEL ARRAY WITH CONSTRAINED INTERACTIVE AND AUDITORY MASKED PROCESSING

Xianxian Zhang, John H. L. Hansen, Kathryn Arehart, University of Colorado, Boulder, United States

SP-L10.5: MULTIPLE-MICROPHONE TIME-VARYING FILTERS FOR ROBUST SPEECH RECOGNITION

Calvin Lai, Parham Aarabi, University of Toronto, Canada

SP-L10.6: NOISE SUPPRESSION FOR AUTOMOTIVE APPLICATIONS BASED ON DIRECTIONAL INFORMATION

Martin Fuchs, Tim Haulick, Gerhard Schmidt, Temic SDS, Germany

SP-L11: LANGUAGE MODELING AND SEARCH

SP-L11.1: META-DATA CONDITIONAL LANGUAGE MODELING

Michiel Bacchiani, Brian Roark, AT&T Labs - Research, United States

SP-L11.2: EXACT TRAINING OF A NEURAL SYNTACTIC LANGUAGE MODEL

Ahmad Emami, Frederick Jelinek, Johns Hopkins University, United States

SP-L11.3: DEVELOPMENT OF THE 2003 CU-HTK CONVERSATIONAL TELEPHONE SPEECH TRANSCRIPTION SYSTEM

Gunnar Evermann, H. Y. Chan, Mark J. F. Gales, Thomas Hain, Xunying Liu, David Mrva, Lan Wang, Phil Woodland, Cambridge University, United Kingdom

SP-L11.4: VOCABULARY-INDEPENDENT SEARCH IN SPONTANEOUS SPEECH

Frank Seide, Peng Yu, Chengyuan Ma, Eric Chang, Microsoft Research Asia, China

SP-L11.5: CROSS-LINGUAL LATENT SEMANTIC ANALYSIS FOR LANGUAGE MODELING

Woosung Kim, Sanjeev Khudanpur, Johns Hopkins University, United States

SP-L11.6: THE USE OF A LINGUISTICALLY MOTIVATED LANGUAGE MODEL IN CONVERSATIONAL SPEECH RECOGNITION

Wen Wang, SRI International / Purdue University, United States; Andreas Stolcke, SRI International, United States; Mary Harper, Purdue University, United States
SP-P1: SPEECH CODING FOR NETWORKS / SINGLE-CHANNEL SPEECH ENHANCEMENT

SP-P1.1: A STUDY OF DESIGN COMPROMISES FOR SPEECH CODERS IN PACKET NETWORKS
Roch Lefebvre, Philippe Gournay, University of Sherbrooke, Canada; Redwan Salami, VoiceAge Corporation, Canada

SP-P1.2: IMPROVEMENT ISSUES ON TRANSCODING ALGORITHMS: FOR THE FLEXIBLE USAGE TO THE VARIOUS PAIRS OF SPEECH CODEC
Jin-Kyu Choi, Chang-Heon Lee, Hong-Goo Kang, Young-Cheol Park, Dae Hee Youn, Yonsei University, Republic of Korea

SP-P1.3: A SCALABLE SPEECH AND AUDIO CODING SCHEME WITH CONTINUOUS BITRATE FLEXIBILITY
Balázs Kövesi, Dominique Massaloux, Aurélien Sollaud, France Télécom R&D, France

SP-P1.4: A MULTIPLE DESCRIPTION SPEECH CODER BASED ON AMR-WB FOR MOBILE AD HOC NETWORKS
Hui Dong, Allen Gersho, Jerry Gibson, Vladimir Cuperman, University of California, Santa Barbara, United States

SP-P1.5: ON THE ARCHITECTURE OF THE CDMA2000® VARIABLE-RATE MULTIMODE WIDEBAND (VMR-WB) SPEECH CODING STANDARD
Milan Jelinek, University of Sherbrooke, Canada; Redwan Salami, VoiceAge Corporation, Canada; Sassan Ahmadi, Nokia, Inc., United States; Bruno Bessette, Philippe Gournay, Claude Laflamme, University of Sherbrooke, Canada

SP-P1.6: A BIT-RATE/BANDWIDTH SCALABLE SPEECH CODER BASED ON ITU-T G.723.1 STANDARD
Sung-Kyo Jung, Kyung-Taek Kim, Hong-Goo Kang, Yonsei University, Republic of Korea

SP-P1.7: A TWO-STEP NOISE REDUCTION TECHNIQUE
Cyril Plapous, Claude Marro, Laurent Mauuary, France Télécom R&D - DIH/IPS, France; Pascal Scalart, ENSSAT - LASTI, France

SP-P1.8: ON THE DECISION-DIRECTED ESTIMATION APPROACH OF EPHRAIM AND MALAH
Israel Cohen, Technion-Israel Institute of Technology, Israel

SP-P1.9: EMPLOYING LAPLACIAN-GAUSSIAN DENSITIES FOR SPEECH ENHANCEMENT
Saeed Gazor, Queen’s University, Canada

SP-P1.10: ROBUST ADAPTIVE KALMAN FILTERING-BASED SPEECH ENHANCEMENT
Marcel Gabrea, École de Technologie Supérieure, Canada

SP-P1.11: A NOISE ESTIMATION ALGORITHM WITH RAPID ADAPTATION FOR HIGHLY NON-STATIONARY ENVIRONMENTS
Sundarrajan Rangachari, Philipos Loizou, Yi Hu, University of Texas, Dallas, United States

SP-P1.12: LOW DISTORTION SPEECH DENOISING USING AN ADAPTIVE PARAMETRIC WIENER FILTER
Ningping Fan, Siemens Corporate Research, United States

SP-P2: SPEAKER ADAPTATION

SP-P2.1: PERFORMANCE COMPARISONS OF ALL-PASS TRANSFORM ADAPTATION WITH MAXIMUM LIKELIHOOD LINEAR REGRESSION
John McDonough, Alex Waibel, University of Karlsruhe, Germany

SP-P2.2: ADAPTIVE TRAINING USING STRUCTURED TRANSFORMS
Kai Yu, Mark J. F. Gales, Cambridge University, United Kingdom
SP-P2.3: MPE-BASED DISCRIMINATIVE LINEAR TRANSFORM FOR SPEAKER ADAPTATION
Lan Wang, Phil Woodland, Cambridge University, United Kingdom

I - 321

SP-P2.4: A STUDY OF VARIOUS COMPOSITE KERNELS FOR KERNEL EIGENVOICE SPEAKER ADAPTATION
Brian Mak, James Kwok, Simon Ho, Hong Kong University of Science and Technology, Hong Kong SAR of China

I - 325

SP-P2.5: FEATURE SPACE GAUSSIANIZATION
George Saon, Satya Dharanipragada, Daniel Povey, IBM T. J. Watson Research Center, United States

I - 329

SP-P2.6: ONLINE SPEAKER CLUSTERING
Daben Liu, Francis Kubala, BBN Technologies, United States

I - 333

SP-P2.7: PRIOR KNOWLEDGE GUIDED MEL BASED MODEL SELECTION AND ADAPTATION FOR NONNATIVE SPEECH RECOGNITION
Xiaodong He, Microsoft, United States; Yunxin Zhao, University of Missouri-Columbia, United States

I - 337

SP-P2.8: ENROLLMENT IN LOW-RESOURCE SPEECH RECOGNITION SYSTEMS
Sabine Deligne, Satya Dharanipragada, IBM T. J. Watson Research Center, United States

I - 341

SP-P2.9: AN INVESTIGATION INTO FRONT-END SIGNAL PROCESSING FOR SPEAKER NORMALIZATION
S. Umesh, Rohit Sinha, Indian Institute of Technology, Kanpur, India; Bharath Kumar SV, General Electric - Global Research, India

I - 345

SP-P2.10: EIGEN-MLLRS APPLIED TO UNSUPERVISED SPEAKER ENROLLMENT FOR LARGE VOCABULARY CONTINUOUS SPEECH RECOGNITION
Xavier Aubert, Philips Research Laboratories Aachen, Germany

I - 349

SP-P2.11: SPEAKER INDEXING AND ADAPTATION USING SPEAKER CLUSTERING BASED ON STATISTICAL MODEL SELECTION
Masafumi Nishida, Chiba University, Japan; Tatsuya Kawahara, Kyoto University, Japan

I - 353

SP-P2.12: EIGENSPACE-BASED MLLR WITH SPEAKER ADAPTIVE TRAINING IN LARGE VOCABULARY CONVERSATIONAL SPEECH RECOGNITION
Vlasios Doumpiotis, Yonggang Deng, Johns Hopkins University, United States

I - 357

SP-P3: TOPICS IN SPEAKER AND LANGUAGE RECOGNITION

SP-P3.1: PARAMETERIZATION OF THE SCORE THRESHOLD FOR A TEXT-DEPENDENT ADAPTIVE SPEAKER VERIFICATION SYSTEM
Nikki Mirghafori, ICSI, United States; Matthieu Hébert, Nuance Communications, Canada

I - 361

SP-P3.2: DESPERATELY SEEKING IMPostORS: DATA-MINING FOR COMPETITIVE IMPOSTOR TESTING IN A TEXT-DEPENDENT SPEAKER VERIFICATION SYSTEM
Matthieu Hébert, Nuance Communications, Canada; Nikki Mirghafori, ICSI, United States

I - 365

SP-P3.3: A MULTIMEDIA APPROACH FOR AUDIO SEGMENTATION IN TV BROADCAST NEWS
Luis Perez-Freire, Carmen García-Mateo, University of Vigo, Spain

I - 369

SP-P3.4: THE ELISA CONSORTIUM APPROACHES IN BROADCAST NEWS SPEAKER SEGMENTATION DURING THE NIST 2003 RICH TRANSCRIPTION EVALUATION
Daniel Moraru, CLIPS-IMAG, France; Sylvain Meignier, Corinne Fredouille, Laboratoire Informatique d'Avignon (LIA), France; Laurent Besacier, CLIPS-IMAG, France; Jean-François Bonastre, Laboratoire Informatique d'Avignon (LIA), France

I - 373

SP-P3.5: ENHANCEMENT OF MISMATCHED CONDITIONS IN SPEAKER RECOGNITION FOR MULTIMEDIA APPLICATIONS
Waleed Fakhr, Ahmed Abdelsalam, Nadder Hamdy, Arab Academy for Science & Technology, Egypt

I - 377

xxi
SP-P3.6: LANGUAGE BOUNDARY DETECTION AND IDENTIFICATION OF MIXED-LANGUAGE SPEECH BASED ON MAP ESTIMATION
Chi-Jui Shia, Yu-Hsien Chiu, Jia-Hsin Hsieh, Chung-Hsien Wu, National Cheng-Kung University, Taiwan

SP-P3.7: FUSING LANGUAGE IDENTIFICATION SYSTEMS USING PERFORMANCE CONFIDENCE INDEXES
Jorge Gutiérrez, Jean-Luc Rouas, Régine André-Obrecht, IRIT - UMR 5505 CNRS INPT UPS, France

SP-P3.8: CONFIDENCE MEASURES IN MULTIPLE PRONUNCIATIONS MODELING FOR SPEAKER VERIFICATION
Mohamed Faouzi BenZeghiba, Hervé Bourlard, IDIAP, Switzerland

SP-P3.9: IDENTIFYING IN-SET AND OUT-OF-SET SPEAKERS USING NEIGHBORHOOD INFORMATION
Pongtep Angkititrakul, John H. L. Hansen, University of Colorado, Boulder, United States

SP-P3.10: BENEFITS OF PRIOR ACOUSTIC SEGMENTATION FOR AUTOMATIC SPEAKER SEGMENTATION
Sylvain Meignier, Laboratoire Informatique d'Avignon (LIA), France; Daniel Moraru, CLIPS-IMAG, France; Corinne Fredouille, Laboratoire Informatique d'Avignon (LIA), France; Laurent Besacier, CLIPS-IMAG, France; Jean-François Bonastre, Laboratoire Informatique d'Avignon (LIA), France

SP-P3.11: LANGUAGE IDENTIFICATION USING PARALLEL SYLLABLE-LIKE UNIT RECOGNITION
Nagarajan Thangavelu, Hema Murthy, Indian Institute of Technology, Madras, India

SP-P3.12: A PITCH SYNCHRONOUS FEATURE EXTRACTION METHOD FOR SPEAKER RECOGNITION
Samuel Kim, Yonsei University, Republic of Korea; Thomas Eriksson, Chalmers University of Technology, Sweden; Hong-Goo Kang, Dae Hee Youn, Yonsei University, Republic of Korea

SP-P4: TOPICS IN SPEECH UNDERSTANDING SYSTEMS

SP-P4.1: BOOTSTRAP ESTIMATES FOR CONFIDENCE INTERVALS IN ASR PERFORMANCE EVALUATION
Maximilian Bisani, Hermann Ney, RWTH Aachen, Germany

SP-P4.2: A DETECTION BASED APPROACH TO ROBUST SPEECH UNDERSTANDING
Kuansan Wang, Microsoft Research, United States

SP-P4.3: ROBUST MULTIMODAL UNDERSTANDING
Srinivas Bangalore, Michael Johnston, AT&T Labs - Research, United States

SP-P4.4: A DISTRIBUTED FRAMEWORK FOR ENTERPRISE LEVEL SPEECH RECOGNITION SERVICES
Iker Arizmendi, AT&T Labs - Research, United States; Richard Rose, McGill University, Canada

SP-P4.5: AUTOMATIC LEARNING OF INTERPRETATION STRATEGIES FOR SPOKEN DIALOGUE SYSTEMS
Christian Raymond, Frédéric Béchet, Renato De Mori, CNRS / University of Avignon, France; Géraldine Damnati, France Télécom R&D, France; Yannick Estève, Université du Maine, France

SP-P4.6: UNSUPERVISED AND ACTIVE LEARNING IN AUTOMATIC SPEECH RECOGNITION FOR CALL CLASSIFICATION
Dilek Hakkani-Tür, Gokhan Tur, Mazin Rahim, Giuseppe Riccardi, AT&T Labs - Research, United States

SP-P4.7: PUBLIC SPEECH-ORIENTED GUIDANCE SYSTEM WITH ADULT AND CHILD DISCRIMINATION CAPABILITY
Ryuichi Nisimura, Akinobu Lee, Hiroshi Saruwatari, Kiyohiro Shikano, Nara Institute of Science and Technology, Japan
SP-P4.8: EXTENDING BOOSTING FOR CALL CLASSIFICATION USING WORD CONFUSION NETWORKS
Gokhan Tur, Dilek Hakkani-Tür, Giuseppe Riccardi, AT&T Labs - Research, United States

SP-P4.9: DIALOG TRAJECTORY ANALYSIS
Alicia Abella, Jerry Wright, Allen Gorin, AT&T Labs - Research, United States

SP-P4.10: IMPROVING PHONEME RECOGNITION OF TELEPHONE QUALITY SPEECH
Qiang Huang, Stephen Cox, University of East Anglia, United Kingdom

SP-P4.11: AUTOMATIC INDEXING OF KEY SENTENCES FOR LECTURE ARCHIVES USING STATISTICS OF PRESUMED DISCOURSE MARKERS
Hiroaki Nanjo, Tasuku Kitade, Tatsuya Kawahara, Kyoto University, Japan

SP-P4.12: SPEECH-ACTIVATED TEXT RETRIEVAL SYSTEM FOR MULTIMODAL CELLULAR PHONES
Shin-ya Ishikawa, Takahiro Ikeda, Kiyokazu Miki, Fumihiro Adachi, NEC Corporation, Japan

SP-P5: TOPICS IN SPEECH CODING

SP-P5.1: NOISE-DEPENDENT POSTFILTERING
Volodya Grancharov, Jonas Samuelsson, W. Bastiaan Kleijn, Royal Institute of Technology (KTH), Sweden

SP-P5.2: A DATA MINING APPROACH TO OBJECTIVE SPEECH QUALITY MEASUREMENT
Wei Zha, Wai-Yip Chan, Queen's University, Canada

SP-P5.3: ADAPTIVE TIME-SEGMENTATION FOR SPEECH CODING WITH LIMITED DELAY
Christoffer A. Rødbro, Aalborg University, Denmark; Jesper Jensen, Richard Heusdens, Delft University of Technology, Netherlands

SP-P5.4: COMBINED ESTIMATION/CODING OF HIGHLAND SPECTRAL ENVELOPES FOR SPEECH SPECTRUM EXPANSION
Yannis Agiomyrgiannakis, Yannis Stylianou, Foundation of Research and Technology Hellas, Greece

SP-P5.5: AUTOMATICALLY DERIVED UNITS FOR SEGMENT VOCODERS
V. Ramasubramanian, Thippur V. Sreenivas, Indian Institute of Science, India

SP-P5.6: MULTISENSOR MELPE USING PARAMETER SUBSTITUTION
Kevin Brady, Thomas Quatieri, Joseph Campbell, William Campbell, Michael Brandstein, Clifford Weinstein, MIT Lincoln Laboratory, United States

SP-P5.7: EFFICIENT SPECTRUM CODING FOR SUPER-WIDEBAND SPEECH AND ITS APPLICATION TO 7/10/15 KHZ BANDWIDTH SCALABLE CODERS
Masahiro Oshikiri, Hiroyuki Ehara, Koji Yoshida, Matsushita Electric Industrial Co., Ltd., Japan

SP-P5.8: ENHANCED STANDARD COMPLIANT DISTRIBUTED SPEECH RECOGNITION (AURORA ENCODER) USING RATE ALLOCATION
Naveen Srinivasamurthy, Antonio Ortega, Shrikanth Narayanan, University of Southern California, United States

SP-P5.9: WIDEBAND AUDIO OVER NARROWBAND LOW-RESOLUTION MEDIA
Heping Ding, National Research Council of Canada, Canada

SP-P5.10: PREDICTING FOREGROUND SH, SL AND BNM DAM SCORES FOR MULTIDIMENIONAL OBJECTIVE MEASURE OF SPEECH QUALITY
D. Sen, University of New South Wales, Australia

SP-P5.11: NOISE REDUCTION ON SPEECH CODEC PARAMETERS
Herve Taddei, Christophe Beaugeant, Mickael de Meuleneire, Siemens AG, Germany
SP-P5.12: LOW-COMPLEXITY MULTI-RATE LATTICE VECTOR QUANTIZATION WITH APPLICATION TO WIDEBAND TCX SPEECH CODING AT 32 KBIT/S
Stéphane Ragot, University of Sherbrooke, Canada; Bruno Bessette, Roch Lefebvre, University Sherbrooke, Canada

SP-P6: FEATURE ANALYSIS FOR ASR, TTS, AND VERIFICATION

SP-P6.1: A MODEL-BASED TONE LABELING METHOD FOR MIN-NAN/TAIWANESE SPEECH
Wei-Chih Kuo, Yih-Ru Wang, Sin-Horng Chen, Chiao Tung University, Taiwan

SP-P6.2: AN AUTOMATIC PROSODY LABELING SYSTEM USING ANN-BASED SYNTACTIC-PROSODIC MODEL AND GMM-BASED ACOUSTIC-PROSODIC MODEL
Ken Chen, Mark Hasegawa-Johnson, Aaron Cohen, University of Illinois at Urbana-Champaign, United States

SP-P6.3: VARIATIONAL BAYESIAN FEATURE SELECTION FOR GAUSSIAN MIXTURE MODELS
Fabio Valente, Christian J. Wellekens, Institut Eurecom, France

SP-P6.4: APPLICATION OF THE MODIFIED GROUP DELAY FUNCTION TO SPEAKER IDENTIFICATION AND DISCRIMINATION
Rajesh Hegde, Hema Murthy, Indian Institute of Technology, India; V. Ramana Rao Gadde, Star Laboratory, SRI International, United States

SP-P6.5: TOWARDS MULTILINGUAL SPEECH RECOGNITION USING DATA DRIVEN SOURCE/TARGET ACOUSTICAL UNITS ASSOCIATION
Rania Bayeh, University of Balamand, Lebanon; Shiuan-Sung Lin, Gerard Chollet, École Nationale Supérieure des Télécommunications, France; Chafic Mokbel, University of Balamand, Lebanon

SP-P6.6: A MULTI-PASS LINEAR FOLD ALGORITHM FOR SENTENCE BOUNDARY DETECTION USING PROSODIC CUES
Dagen Wang, Shrikanth S. Narayanan, University of Southern California, United States

SP-P6.7: FRACTIONAL FOURIER TRANSFORM FEATURES FOR SPEECH RECOGNITION
Ruhi Sarikaya, Yuqing Gao, George Saon, IBM T. J. Watson Research Center, United States

SP-P6.8: JOINT FREQUENCY DOMAIN AND RECONSTRUCTED PHASE SPACE FEATURES FOR SPEECH RECOGNITION
Andrew Lindgren, Michael Johnson, Richard Povinelli, Marquette University, United States

SP-P6.9: TRAPPING CONVERSATIONAL SPEECH: EXTENDING TRAP/TANDEM APPROACHES TO CONVERSATIONAL TELEPHONE SPEECH RECOGNITION

SP-P6.10: ON USE OF TASK INDEPENDENT TRAINING DATA IN TANDEM FEATURE EXTRACTION
Sunil Sivadas, Oregon Health & Science University, United States / IDIAP, Switzerland; Hynek Hermansky, IDIAP, Switzerland

SP-P6.11: FEATURE GENERATION BASED ON MAXIMUM-normalized ACOUSTIC LIKELIHOOD FOR IMPROVED SPEECH RECOGNITION
Xiang Li, Richard Stern, Carnegie Mellon University, United States

SP-P6.12: ENTROPY-BASED VARIABLE FRAME RATE ANALYSIS OF SPEECH SIGNALS AND ITS APPLICATION TO ASR
Hong You, University of California, Los Angeles, United States; Qifeng Zhu, ICSI, United States; Abeer Alwan, University of California, Los Angeles, United States
SP-P7: TOPICS IN SPEECH ANALYSIS

SP-P7.1: BAYESIAN MODELLING OF THE SPEECH SPECTRUM USING MIXTURE OF 5 GAUSSIANS
Parham Zolfaghari, Shinji Watanabe, Atsushi Nakamura, Shigeru Katagiri, NTT Corporation, Japan

SP-P7.2: A STRUCTURED SPEECH MODEL WITH CONTINUOUS HIDDEN DYNAMICS AND PREDICTION-RESIDUAL TRAINING FOR TRACKING VOCAL TRACT RESONANCES
Li Deng, Leo Lee, Hagai Attias, Alex Acero, Microsoft, United States

SP-P7.3: AN ESTIMATE OF PHYSICAL SCALE FROM SPEECH
Lawrence Smith, National Institutes of Health (NIH), United States; Douglas Nelson, United States Department of Defense, United States

SP-P7.4: FORMANT TRACKING BY MIXTURE STATE PARTICLE FILTER
Yanli Zheng, Mark Hasegawa-Johnson, University of Illinois at Urbana-Champaign, United States

SP-P7.5: ACOUSTIC ANALYSIS OF FRIENDLY SPEECH
Fangxin Chen, IBM China Research Laboratory, China; Aijun Li, Haibo Wang, Tianqing Wang, Qiang Fang, Chinese Academy of Social Science, China

SP-P7.6: IMPORTANCE OF WINDOW SHAPE FOR PHASE-ONLY RECONSTRUCTION OF SPEECH
Leigh Alsteris, Kuldip. KPaliwal, Griffith University, Australia

SP-P7.7: SPEECH EMOTION RECOGNITION COMBINING ACOUSTIC FEATURES AND LINGUISTIC INFORMATION IN A HYBRID SUPPORT VECTOR MACHINE - BELIEF NETWORK ARCHITECTURE
Björn Schuller, Gerhard Rigoll, Manfred Lang, Technische Universität München, Germany

SP-P7.8: FORMANT FREQUENCY ESTIMATION IN NOISE
Bin Chen, Philipos Loizou, University of Texas, Dallas, United States

SP-P7.9: YET ANOTHER ACOUSTIC REPRESENTATION OF SPEECH SOUNDS
Nobuaki Minematsu, University of Tokyo, Japan

SP-P7.10: ESTIMATING VOCAL-TRACT AREA FUNCTIONS FROM VOWEL SOUND SIGNALS OVER CLOSED GLOTTAL PHASES
Huiqun Deng, Rabab K. Ward, Michael Beddoes, Murray Hodgson, University of British Columbia, Canada

SP-P8: VOICE ACTIVITY DETECTION AND SPEECH SEGMENTATION

SP-P8.1: A DIFFERENTIAL SPECTRAL VOICE ACTIVITY DETECTOR
Philip Garner, Toshiaki Fukada, Yasuhiro Komori, Canon, Inc., Japan

SP-P8.2: SPEECH DISCRIMINATION BASED ON MULTISCALE SPECTRO-TEMPORAL MODULATIONS
Nima Mesgarani, Shihab Shamma, University of Maryland, College Park, United States; Malcolm Slaney, IBM Almaden Research Center, United States

SP-P8.3: CLUSTERING AND SEGMENTING SPEAKERS AND THEIR LOCATIONS IN MEETINGS
Jitendra Ajmera, Guillaume Lathoud, Iain McCowan, IDIAP, Switzerland

SP-P8.4: VOICE ACTIVITY DETECTION USING VISUAL INFORMATION
Peng Liu, Zuoying Wang, Tsinghua University, China
SP-P8.5: SPEECH MODELING AND VOICED/UNVOICED/MIXED/SILENCE SPEECH
SEGMENTATION WITH FRACTIONALLY GAUSSIAN NOISE BASED MODELS
Shahab Oveisgharan, Mohammad Bagher Shamsollahi, Sharif University of Technology, Iran

SP-P8.6: SOUND FEATURE DETECTION USING LEAKY INTEGRATE-AND-FIRE NEURONS
Leslie Smith, Dagmar Fraser, University of Stirling, United Kingdom

SP-P8.7: CLOSED-FORM ESTIMATION OF THE AMPLITUDE COMMANDS IN THE AUTOMATIC EXTRACTION OF FUSISAKI’S MODEL
Solimar Silva, Sergio Netto, Federal University of Rio de Janeiro, Brazil

SP-P8.8: A VOICE ACTIVITY DETECTOR USING THE CHI-SQUARE TEST
Beena Ahmed, RMIT University, Australia; W. Harvey Holmes, University of New South Wales, Australia

SP-P9: TOPICS IN SPEECH SYNTHESIS

SP-P9.1: MINIMUM SEGMENTATION ERROR BASED DISCRIMINATIVE TRAINING FOR SPEECH SYNTHESIS APPLICATION
Yi-Jian Wu, University of Science and Technology of China, China; Hisashi Kawai, Jinfu Ni, ATR, Spoken Language Translation Laboratories, Japan; Ren-Hua Wang, University of Science and Technology of China, China

SP-P9.2: WATERMARKING OF SPEECH SIGNALS USING THE SINUSOIDAL MODEL AND FREQUENCY MODULATION OF THE PARTIALS
Laurent Girin, ICP/INPG, France; Sylvain Marchand, SCRIME/LaBRI, France

SP-P9.3: ANALYSIS BY SYNTHESIS OF ACOUSTIC CORRELATES OF BRITISH, AUSTRALIAN AND AMERICAN ACCENTS
Qin Yan, Saeed Vaseghi, Dimitrios Rentzos, Brunel University, United Kingdom; Ching-Hsiang Ho, Fortune Institute of Technology, Taiwan

SP-P9.4: REFINING SEGMENTAL BOUNDARIES FOR TTS DATABASE USING FINE CONTEXTUAL-DEPENDENT BOUNDARY MODELS
Lijuan Wang, Tsinghua University, China; Yong Zhao, Min Chu, Jian-Lai Zhou, Microsoft Research Asia, China; Zhigang Cao, Tsinghua University, China

SP-P9.5: A LOW-BAND SPECTRUM ENVELOPE MODELING FOR HIGH QUALITY PITCH MODIFICATION
Ryo Mochizuki, Waseda University / Matsushita Electric Industrial Co., Ltd., Japan; Tetsunori Kobayashi, Waseda University, Japan

SP-P9.6: PROBABILITY BASED PROSODY MODEL FOR UNIT SELECTION
Xi Jun Ma, Wei Zhang, Wei Bin Zhu, Qin Shi, Ling Jin, IBM China Research Laboratory, China

SP-P9.7: A REAL-TIME CANTONESE TEXT-TO-AUDIOVISUAL SPEECH SYNTHESIZER
Jian-Qing Wang, Ka-Ho Wong, Pheng-Ann Heng, Helen Mei-Ling Meng, Tien-Tsin Wong, Chinese University of Hong Kong, Hong Kong SAR of China

SP-P9.8: OPTIMIZING SUB-COST FUNCTIONS FOR SEGMENT SELECTION BASED ON PERCEPTUAL EVALUATIONS IN CONCATENATIVE SPEECH SYNTHESIS
Tomoki Toda, Nagoya Institute of Technology / ATR, Japan; Hisashi Kawai, Minoru Tsuzaki, ATR, Spoken Language Translation Laboratories, Japan

SP-P9.9: EVALUATION OF THE EFFECT OF STRESS ON FORMANTS IN Farsi VOWELS
Davood Gharavian, Mohammad Ahadi, Amirkabir University of Technology, Iran

SP-P9.10: A STRATEGY TO SOLVE DATA SCARCITY PROBLEMS IN CORPUS BASED INTONATION MODELLING
Valentin Cardeñoso, David Escudero, University of Valladolid, Spain
SP-P9.11: AN IMPROVED CORRECTION FORMULA FOR THE ESTIMATION OF HARMONIC MAGNITUDES AND ITS APPLICATION TO OPEN QUOTIENT ESTIMATION
Markus Iseli, Abeer Alwan, University of California, Los Angeles, United States

SP-P9.12: MODELING PRONUNCIATION VARIATION FOR SPONTANEOUS SPEECH SYNTHESIS
Steffen Werner, Matthias Wolff, Matthias Eichner, Rüdiger Hoffmann, Dresden University of Technology, Germany

SP-P9.13: AN EVALUATION OF AUTOMATIC PHONE SEGMENTATION FOR CONCATENATIVE SPEECH SYNTHESIS
Hisashi Kawai, ATR, Spoken Language Translation Laboratories, Japan; Tomoki Toda, Nagoya Institute of Technology, Japan

SP-P9.14: SCALING OF WAVEFORM SEGMENTS ALONG THE TIME AXIS FOR CONCATENATIVE SPEECH SYNTHESIS
Nobuyuki Nishizawa, Hisashi Kawai, ATR, Spoken Language Translation Laboratories, Japan

SP-P9.15: SPEECH SYNTHESIS FROM REAL TIME ULTRASOUND IMAGES OF THE TONGUE
Bruce Denby, Université Pierre et Marie Curie, France; Maureen Stone, University of Maryland Dental School, United States

SP-P10: TOPICS IN SPEECH ENHANCEMENT

SP-P10.1: SPHERICAL HARMONIC ANALYSIS OF EQUALIZATION IN A REVERBERANT ROOM
Terence Betlehem, Thushara Abhayapala, Australian National University, Australia

SP-P10.2: SPEECH ENHANCEMENT BY PERCEPTUAL FILTER WITH SEQUENTIAL NOISE PARAMETER ESTIMATION
Te-Won Lee, Kaisheng Yao, University of California, San Diego, United States

SP-P10.3: FEATURE SELECTION FOR IMPROVED BANDWIDTH EXTENSION OF SPEECH SIGNALS
Peter Jax, Peter Vary, Aachen University (RWTH), Germany

SP-P10.4: AUTOMATED LIP-READING FOR IMPROVED SPEECH INTELLIGIBILITY
Matthew McClain, University of Illinois, United States; Kevin Brady, Michael Brandstein, Thomas Quatieri, MIT Lincoln Laboratory, United States

SP-P10.5: ESTIMATION OF SHORT-TERM PREDICTOR PARAMETERS FOR CODING AND ENHANCEMENT OF NOISY SPEECH
Sriram Srinivasan, Jonas Samuelsson, W. Bastiaan Kleijn, Royal Institute of Technology (KTH), Sweden

SP-P10.6: HMM-BASED FREQUENCY BANDWIDTH EXTENSION FOR SPEECH ENHANCEMENT USING LINE SPECTRAL FREQUENCIES
Guo Chen, Vijay Parsa, National Centre for Audiology, Canada

SP-P10.7: COMBINING EQUALIZATION AND ESTIMATION FOR BANDWIDTH EXTENSION OF NARROWBAND SPEECH
Yasheng Qian, Peter Kabal, McGill University, Canada

SP-P10.8: PERCEPTUAL KALMAN FILTERING FOR SPEECH ENHANCEMENT IN COLORED NOISE
Ning Ma, Martin Bouchard, University of Ottawa, Canada; Rafik Goubran, Carleton University, Canada

SP-P10.9: SPEECH ENHANCEMENT USING ROBUST WEIGHTING FACTORS FOR CRITICAL-BAND-WAVELET-PACKET TRANSFORM
Ching-Ta Lu, Chin-Min College, Taiwan; Hsiao-Chuan Wang, National Tsing Hua University, Taiwan
SP-P10.10: AN MMSE SPEECH ENHANCEMENT APPROACH INCORPORATING MASKING PROPERTIES
Chang huai You, Institute for Infocomm Research, Singapore; Soo ngee Koh, Nanyang Technological University, Singapore; Susanto Rahardja, Institute for Infocomm Research, Singapore

SP-P10.11: NEW SPEECH HARMONIC STRUCTURE MEASURE AND IT APPLICATION TO POST SPEECH ENHANCEMENT
An-Tze Yu, Hsiao-Chuan Wang, National Tsing Hua University, Taiwan

SP-P10.12: SPEECH ENHANCEMENT WITH MISSING DATA TECHNIQUES USING RECURRENT NEURAL NETWORKS
Shahla Parveen, Phil Green, University of Sheffield, United Kingdom

SP-P11: TOPICS IN LARGE VOCABULARY CONTINUOUS SPEECH RECOGNITION

SP-P11.1: IMPROVING BROADCAST NEWS TRANSCRIPTION BY LIGHTLY SUPERVISED DISCRIMINATIVE TRAINING
H. Y. Chan, Phil Woodland, Cambridge University, United Kingdom

SP-P11.2: ADVANCES IN UNSUPERVISED AUDIO SEGMENTATION FOR THE BROADCAST NEWS AND NGSW CORPORA
Rongqing Huang, John H. L. Hansen, University of Colorado, Boulder, United States

SP-P11.3: HYBRID LANGUAGE MODELS FOR OUT OF VOCABULARY WORD DETECTION IN LARGE VOCABULARY CONVERSATIONAL SPEECH RECOGNITION
Ali Yazgan, Johns Hopkins University, United States; Murat Saraclar, AT&T Labs - Research, United States

SP-P11.4: CORRECTIVE LANGUAGE MODELING FOR LARGE VOCABULARY ASR WITH THE PERCEPTRON ALGORITHM
Brian Roark, Murat Saraclar, AT&T Labs - Research, United States; Michael Collins, MIT Artificial Intelligence Laboratory, United States

SP-P11.5: GENERATING AND EVALUATING SEGMENTATIONS FOR AUTOMATIC SPEECH RECOGNITION OF CONVERSATIONAL TELEPHONE SPEECH
Sue Tranter, Kai Yu, Gunnar Evermann, Phil Woodland, Cambridge University, United Kingdom

SP-P11.6: OUT-OF-DOMAIN DETECTION BASED ON CONFIDENCE MEASURES FROM MULTIPLE TOPIC CLASSIFICATION
Ian Lane, Tatsuya Kawahara, Kyoto University, Japan; Tomoko Matsui, The Institute of Statistical Mathematics, Japan; Satoshi Nakamura, ATR, Spoken Language Translation Laboratories, Japan

SP-P11.7: A GENERALIZED CONSTRUCTION OF INTEGRATED SPEECH RECOGNITION TRANSDUCERS
Cyril Allauzen, Mehryar Mohri, Michael Riley, Brian Roark, AT&T Labs - Research, United States

SP-P11.8: CROSS-DIALECTAL ACOUSTIC DATA SHARING FOR ARABIC SPEECH RECOGNITION
Katrin Kirchhoff, University of Washington, United States; Dimitra Vergyri, SRI International, United States

SP-P11.9: ADVANCES IN THE AUTOMATIC TRANSCRIPTION OF LECTURES
Mauro Cettolo, Fabio Brugnara, Marcello Federico, ITC-irst, Italy

SP-P11.10: THE 2003 ISL RICH TRANSCRIPTION SYSTEM FOR CONVERSATIONAL TELEPHONY SPEECH
Hagen Soltau, Hua Yu, Florian Metze, Christian Fügen, Qin Jin, Szu-Chen Jou, Interactive Systems Labs, Germany

SP-P11.11: LIGHTLY SUPERVISED AND DATA-DRIVEN APPROACHES TO MANDARIN BROADCAST NEWS TRANSCRIPTION
Berlin Chen, Jen-Wei Kuo, Wen-Hung Tsai, National Taiwan Normal University, Taiwan
SP-P11.12: FILLER MODEL BASED CONFIDENCE MEASURES FOR SPOKEN DIALOGUE .................................. 1 - 781
SYSTEMS: A CASE STUDY FOR TURKISH
Aydin Akyol, Hakan Erdogan, Sabanci University, Turkey

SP-P11.13: AN EVALUATION OF A NONLINEAR FEATURE TRANSFORMATION FOR CONVERSATIONAL SPEECH RECOGNITION
Mohamed Omar, University of Illinois at Urbana-Champaign, United States; Brian Kingsbury, IBM T. J. Watson Research Center, United States

SP-P11.14: IMPROVED NAME RECOGNITION WITH META-DATA DEPENDENT NAME NETWORKS
Sameer Maskey, Columbia University, United States; Michiel Bacchiani, Brian Roark, AT&T Labs - Research, United States; Richard Sproat, University of Illinois at Urbana-Champaign, United States

SP-P11.15: REAL-TIME WORD CONFIDENCE SCORING USING LOCAL POSTERIOR PROBABILITIES ON TREE TRELLIS SEARCH
Akinobu Lee, Kiyohiro Shikano, Nara Institute of Science and Technology, Japan; Tatsuya Kawahara, Kyoto University, Japan

SP-P12: ACOUSTIC MODELING: MODEL COMPLEXITY, GENERAL TOPICS

SP-P12.1: MODEL COMPLEXITY CONTROL AND COMPRESSION USING DISCRIMINATIVE GROWTH FUNCTIONS
Xunying Liu, Mark J. F. Gales, Cambridge University, United Kingdom

SP-P12.2: BASIS SUPERPOSITION PRECISION MATRIX MODELLING FOR LARGE VOCABULARY CONTINUOUS SPEECH RECOGNITION
Khe Chai Sim, Mark J. F. Gales, Cambridge University, United Kingdom

SP-P12.3: AUTOMATIC GENERATION OF NON-UNIFORM HMM STRUCTURES BASED ON VARIATIONAL BAYESIAN APPROACH
Takatoshi Jitsuhiro, Satoshi Nakamura, ATR, Spoken Language Translation Laboratories, Japan

SP-P12.4: RAO-BLACKWELLISED GIBBS SAMPLING FOR SWITCHING LINEAR DYNAMICAL SYSTEMS
Antti-Veikko Rosti, Mark J. F. Gales, Cambridge University, United Kingdom

SP-P12.5: AUTOMATIC DETERMINATION OF ACOUSTIC MODEL TOPOLOGY USING VARIATIONAL BAYESIAN ESTIMATION AND CLUSTERING
Shinji Watanabe, NTT Corporation, Japan; Atsushi Sako, Ryukoku University, Japan; Atsushi Nakamura, NTT Corporation, Japan

SP-P12.6: OPTIMIZING ACOUSTIC MODELS FOR COMMERCIAL SPEECH RECOGNITION USING FOREGROUND SCORES AND DATA WEIGHTING
Daniel Boies, Brian Strope, Mitchel Weintraub, Su-Lin Wu, Nuance Communications, United States

SP-P12.7: EXTENDED BAUM TRANSFORMATIONS FOR GENERAL FUNCTIONS
Dimitri Kanevsky, IBM T. J. Watson Research Center, United States

SP-P12.8: STUDIES IN MASSIVELY SPEAKER-SPECIFIC SPEECH RECOGNITION
Yu Shi, Eric Chang, Microsoft Research Asia, China

SP-P12.9: PHONE DURATION MODELING FOR LVCSR
Daniel Povey, IBM T. J. Watson Research Center, United States

SP-P12.10: SEQUENTIAL CLUSTERING ALGORITHM FOR GAUSSIAN MIXTURE INITIALIZATION
Ronaldo Messina, Denis Jouvet, France Télécom R&D, France
SP-P12.11: A VITERBI ALGORITHM FOR A TRAJECTORY MODEL DERIVED FROM HMM ................................. I - 837
WITH EXPLICIT RELATIONSHIP BETWEEN STATIC AND DYNAMIC FEATURES
Heiga Zen, Keitchi Tokuda, Tadashi Kitamura, Nagoya Institute of Technology, Japan

SP-P12.12: TRAINING FOR POLYNOMIAL SEGMENT MODEL USING THE EXPECTATION MAXIMIZATION ALGORITHM ................................. I - 841
Chak-Fai Li, Man-Hung Siu, Hong Kong University of Science and Technology, Hong Kong SAR of China

SP-P13: GENERAL TOPICS IN ROBUST SPEECH RECOGNITION

SP-P13.1: CODEBOOK DESIGN FOR ASR SYSTEMS USING CUSTOM ARITHMETIC UNITS................................. I - 845
Xiao Li, Jonathan Malkin, Jeff Bilmes, University of Washington, United States

SP-P13.2: A NEW VOICE ACTIVITY DETECTOR USING SUBBAND ORDER-STATISTICS ................................. I - 849
FILTERS FOR ROBUST SPEECH RECOGNITION
Javier Ramirez, José C. Segura, Carmen Bentéz, Ángel de la Torre, Antonio J. Rubio, Universidad de Granada, Spain

SP-P13.3: AN ANALYSIS OF INTERLEAVERS FOR ROBUST SPEECH RECOGNITION IN ................................. I - 853
BURST-LIKE PACKET LOSS
Alastair James, Ben Milner, University of East Anglia, United Kingdom

SP-P13.4: A STREAM-WEIGHT OPTIMIZATION METHOD FOR AUDIO-VISUAL SPEECH USING MULTI-STREAM HMMS
Satoshi Tamura, Koji Iwano, Sadaoki Furui, Tokyo Institute of Technology, Japan

SP-P13.5: A FACTORIAL HMM APPROACH TO SIMULTANEOUS RECOGNITION OF ISOLATED DIGITS SPOKEN BY MULTIPLE TALKERS ON ONE AUDIO CHANNEL
Ameya Deoras, Mark Hasegawa-Johnson, University of Illinois at Urbana-Champaign, United States

SP-P13.6: INVESTIGATIONS INTO THE RELATIONSHIP BETWEEN MEASURABLE SPEECH QUALITY AND SPEECH RECOGNITION RATE FOR TELEPHONY SPEECH
Hanwu Sun, Louis Shue, Jianfeng Chen, Institute for Infocomm Research, Singapore

SP-P13.7: ACOUSTIC MODEL ADAPTATION USING FIRST ORDER PREDICTION FOR REVERBERANT SPEECH
Tetsuya Takiguchi, Masafumi Nishimura, IBM Research, Japan

SP-P13.8: EXPERIMENTS IN KEYPAD-AIDED SPELLING RECOGNITION ................................................. I - 873
Sarangarajan Parthasarathy, AT&T Labs - Research, United States

SP-P13.9: SPEECH ENHANCEMENT BASED ON MULTIPLE DIRECTIVITY PATTERNS USING A MICROPHONE ARRAY
Toshiyuki Sekiya, Tetsunori Kobayashi, Waseda University, Japan

SP-P13.10: PARAMETER SHARING IN SUBBAND LIKELIHOOD-MAXIMIZING BEAMFORMING FOR SPEECH RECOGNITION USING MICROPHONE ARRAYS
Michael Selizer, Microsoft Research, United States; Richard Stern, Carnegie Mellon University, United States

SP-P13.11: FUSION BASED SPEECH SEGMENTATION IN DARPA SPINE2 TASK ........................................... I - 885
Chengyi Zheng, Yonghong Yan, OGI School of Science & Engineering, United States

SP-P13.12: EXTENDED CLUSTER INFORMATION VECTOR QUANTIZATION (ECI-VQ) FOR ROBUST CLASSIFICATION
Jon Arrowood, Nexidia, Inc., United States; Mark Clements, Georgia Institute of Technology, United States
SP-P14: ACOUSTIC MODELING: TONE, PROSODY, AND FEATURES

SP-P14.1: INTEGRATING THUMBNAIL FEATURES FOR SPEECH RECOGNITION USING CONDITIONAL EXPONENTIAL MODELS
Hua Yu, Alex Waibel, Carnegie Mellon University, United States

SP-P14.2: DISCRiminative Feature Transformation by guided Discriminative training
Roger Hsiao, Brian Mak, Hong Kong University of Science and Technology, Hong Kong SAR of China

SP-P14.3: segmentAL TONAL MODELING FOR PHONE SET DESIGN IN MANDARIN LVCSR
Chao Huang, Yu Shi, Jian-Lai Zhou, Min Chu, Terry Wang, Eric Chang, Microsoft Research Asia, China

SP-P14.4: Decision Tree Based Tone Modeling for Chinese Speech Recognition
Pui-Fung Wong, Man-Hung Siu, Hong Kong University of Science and Technology, Hong Kong SAR of China

SP-P14.5: Hidden Spectral Peak Trajectory Model for Phone Classification
Yiu-Pong Lai, Man-Hung Siu, Hong Kong University of Science and Technology, Hong Kong SAR of China

SP-P14.6: A Study on Robust Segmentation and Location of Tone Nuclei in Chinese Continuous Speech
Jinsong Zhang, ATR, Spoken Language Translation Laboratories, Japan; Keikichi Hirose, University of Tokyo, Japan

SP-P14.7: Chinese-English Bilingual Phone Modeling for Cross-Language Speech Recognition
Shengmin Yu, Shuwu Zhang, Bo Xu, Chinese Academy of Sciences, China

SP-P14.8: Voicing Feature Integration in SRI's Decipher LVCSR System
Martin Graciarena, Horacio Franco, Jing Zheng, Dimitra Vergyri, Andreas Stolcke, SRI International, United States

SP-P14.9: Parsing Speech into Articulatory Events
Kadri Hacioglu, Bryan Pellom, Wayne Ward, University of Colorado, Boulder, United States

SP-P14.10: Prosody-Based Recognition of Spoken German Varieties
Vedran Dizdarevic, Martin Hagmüller, Gernot Kubin, Franz Pernkopf, Graz University of Technology, Austria; Micha Baum, SPEX, Netherlands

SP-P14.11: Tone Variation Modeling for Fluent Mandarin Tone Recognition Based on Clustering
Wan-Yi Lin, National Taiwan University, Taiwan

SP-P14.12: Minimum Classification Error Training of Landmark Models for Real-Time Continuous Speech Recognition
Erik McDermott, NTT Corporation, Japan; Timothy Hazen, Massachusetts Institute of Technology, United States

SP-P15: Robustness in Noisy Environments

SP-P15.1: Robust Speech Recognition in Additive and Channel Noise Environments Using GMM and EM Algorithm
Masakiyo Fujimoto, Ryukoku University, Japan; Yasuo Ariki, Kobe University, Japan

SP-P15.2: Assessment of Signal Subspace Based Speech Enhancement for Noise Robust Speech Recognition
Kris Hermus, Patrick Wambacq, Katholieke Universiteit Leuven, Belgium
SP-P15.3: JOINT REMOVAL OF ADDITIVE AND CONVOLUTIONAL NOISE WITH MODEL-BASED FEATURE ENHANCEMENT
Veronique Stouten, Hugo Van hamme, Patrick Wambacq, Katholieke Universiteit Leuven, Belgium

SP-P15.4: NOISE ROBUST SPEECH RECOGNITION WITH A SWITCHING LINEAR DYNAMIC MODEL
Jasha Droppo, Alex Acero, Microsoft Research, United States

SP-P15.5: A MODIFIED EPHRAIM-MALAH NOISE SUPPRESSION RULE FOR AUTOMATIC SPEECH RECOGNITION
Roberto Gemello, Franco Mana, Loquendo, Italy; Renato De Mori, University of Avignon, France

SP-P15.6: UNIVERSAL COMPENSATION - AN APPROACH TO NOISY SPEECH RECOGNITION ASSUMING NO KNOWLEDGE OF NOISE
Ji Ming, Queen's University Belfast, United Kingdom

SP-P15.7: ON TRACKING NOISE WITH LINEAR DYNAMICAL SYSTEM MODELS
Bhiksha Raj, Mitsubishi Electric Research Labs, United States; Rita Singh, Richard Stern, Carnegie Mellon University, United States

SP-P15.8: COMBINING FEATURE COMPENSATION AND WEIGHTED VITERBI DECODING FOR NOISE ROBUST SPEECH RECOGNITION WITH LIMITED ADAPTATION DATA
Xiaodong Cui, Abeer Alwan, University of California, Los Angeles, United States

SP-P15.9: SNR-DEPENDENT NON-UNIFORM SPECTRAL COMPRESSION FOR NOISY SPEECH RECOGNITION
Kam-keung Chu, Shu Hung Leung, City University of Hong Kong, China

SP-P15.10: MINIMUM MEAN SQUARE ERROR FILTERING OF NOISY CEPSTRAL COEFFICIENTS WITH APPLICATIONS TO ASR
Tor André Myrvoll, Norges Teknisk Naturvitenskaplige Universitet, Norway; Satoshi Nakamura, ATR, Spoken Language Translation Laboratories, Japan

SP-P15.11: A TREE-STRUCTURED CLUSTERING METHOD INTEGRATING NOISE AND SNR FOR PIECEWISE LINEAR-TRANSFORMATION-BASED NOISE ADAPTATION
Zhipeng Zhang, Toshiaki Sugimura, NTT DoCoMo, Japan; Sadaoki Furui, Tokyo Institute of Technology, Japan

SP-P15.12: NONLINEAR NOISE COMPENSATION IN FEATURE DOMAIN FOR SPEECH RECOGNITION WITH NUMERICAL METHODS
Hui Jiang, Qi Wang, York University, Canada

SP-P15.13: PCMM-BASED FEATURE COMPENSATION SCHEMES USING MODEL INTERPOLATION AND MIXTURE SHARING
Wooil Kim, Korea University, Republic of Korea; Ohil Kwon, Hyundai Autonet Co. Ltd., Republic of Korea; Hanseok Ko, Korea University, Republic of Korea

SP-P16: SPEECH MODELING FOR ROBUST SPEECH RECOGNITION

SP-P16.1: DBN BASED MULTI-STREAM MODELS FOR AUDIO-VISUAL SPEECH RECOGNITION
John Gowdy, Amarnag Subramanya, Clemson University, United States; Chris Bartels, Jeff Bilmes, University of Washington, United States

SP-P16.2: TONE ARTICULATION MODELING FOR MANDARIN SPONTANEOUS SPEECH RECOGNITION
Jian-Lai Zhou, Ye Tian, Yu Shi, Chao Huang, Eric Chang, Microsoft Research Asia, China

SP-P16.3: SPATIO-TEMPORAL PROCESSING FOR DISTANT SPEECH RECOGNITION
Siow Yong Low, Western Australian Telecommunications Research Institute, Australia; Roberto Togneri, University of Western Australia, Australia; Sven Nordholm, Western Australian Telecommunications Research Institute, Australia
Volume II

SAM-L1: MIMO SYSTEMS AND SPACE-TIME CODING

SAM-L1.1: A MAXIMIN APPROACH FOR ROBUST MIMO DESIGN: COMBINING OSTBC AND BEAMFORMING WITH MINIMUM TRANSMIT POWER REQUIREMENTS
Antonio Pascual-Iserte, Ana I. Pérez-Neira, Polytechnic University of Catalonia (UPC), Spain; Miguel Ángel Lagunas, Telecommunications Technological Center of Catalonia (CTTC), Spain

SAM-L1.2: OUTAGE PROBABILITY OF MULTI-CELLULAR MIMO SYSTEMS IN RAYLEIGH FADING
Yeliz Tokgoz, Bhaskar Rao, University of California, San Diego, United States

SAM-L1.3: ROBUST LINEAR RECEIVERS FOR SPACE-TIME BLOCK CODED MULTIPLE-ACCESS MIMO WIRELESS SYSTEMS
Yue Rong, University of Duisburg-Essen, Germany; Shahram Shahbazpanahi, Alex Gershman, McMaster University, Canada

SAM-L1.4: TRANSMIT/RECEIVE MIMO ANTENNA SUBSET SELECTION
Alexei Gorokhov, Manel Collados, Philips Research Laboratories, Netherlands; Dhananjay Gore, Qualcomm, Inc., United States; Arogyaswami Paulraj, Stanford University, United States