23rd Annual Conference of the International Speech Communication Association (INTERSPEECH 2022)

Human and Humanizing Speech Technology

Incheon, South Korea
18-22 September 2022

Volume 1 of 7

ISBN: 978-1-7138-8879-6

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571

Some format issues inherent in the e-media version may also appear in this print version.

International Speech Communication Association
c/o Mme Emmanuelle FOXONET
4 Rue des Fauvettes - Lous Tourils
F-66390 Baixas, France

Phone: 49 228 735 643
Fax: 33 468 385 827

secretariat@isca-speech.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

TABLE OF CONTENTS

VOLUME 1

ACOUSTIC PHONETICS AND PROSODY

SPOKEN MACHINE TRANSLATION

(MULTIMODAL) SPEECH EMOTION RECOGNITION I

DEREVERBERATION, NOISE REDUCTION, AND SPEAKER EXTRACTION

SOURCE SEPARATION II

EMBEDDING AND NETWORK ARCHITECTURE FOR SPEAKER RECOGNITION

SPEECH REPRESENTATION II

SPEECH SYNTHESIS: LINGUISTIC PROCESSING, PARADIGMS AND OTHER TOPICS II

OTHER TOPICS IN SPEECH RECOGNITION

AUDIO DEEP PLC (PACKET LOSS CONCEALMENT) CHALLENGE

ROBUST SPEAKER RECOGNITION

SPEECH PRODUCTION

SPEECH QUALITY ASSESSMENT

SPEECH PROCESSING & MEASUREMENT

SPEECH SYNTHESIS: ACOUSTIC MODELING AND NEURAL WAVEFORM GENERATION I

SHOW AND TELL I

SPATIAL AUDIO

SINGLE-CHANNEL SPEECH ENHANCEMENT II

NOVEL MODELS AND TRAINING METHODS FOR ASR II

SPOKEN DIALOGUE SYSTEMS AND MULTIMODALITY

SHOW AND TELL I(VR)

SPEECH EMOTION RECOGNITION I

SINGLE-CHANNEL SPEECH ENHANCEMENT I

SPEECH SYNTHESIS: NEW APPLICATIONS

SPOKEN LANGUAGE UNDERSTANDING I

INCLUSIVE AND FAIR SPEECH TECHNOLOGIES I

INCLUSIVE AND FAIR SPEECH TECHNOLOGIES II

PHONETICS I

MULTI-, CROSS-LINGUAL AND OTHER TOPICS IN ASR I

ZERO, LOW-RESOURCE AND MULTI-MODAL SPEECH RECOGNITION I

SPEAKER EMBEDDING AND DIARIZATION

ACOUSTIC EVENT DETECTION AND CLASSIFICATION

VOLUME 3

SPEECH SYNTHESIS: ACOUSTIC MODELING AND NEURAL WAVEFORM GENERATION II

ASR: ARCHITECTURE AND SEARCH

SPOKEN LANGUAGE PROCESSING II

SOURCE SEPARATION I

ASR TECHNOLOGIES AND SYSTEMS

SPEECH PERCEPTION

SPOKEN TERM DETECTION AND VOICE SEARCH

SPEECH AND LANGUAGE IN HEALTH: FROM REMOTE MONITORING TO MEDICAL CONVERSATIONS I

SPEECH SYNTHESIS: LINGUISTIC PROCESSING, PARADIGMS AND OTHER TOPICS I

SHOW AND TELL II

MULTIMODAL SPEECH EMOTION RECOGNITION AND PARALINGUISTICS

NEURAL TRANSDUCERS, STREAMING ASR AND NOVEL ASR MODELS

ZERO, LOW-RESOURCE AND MULTI-MODAL SPEECH RECOGNITION II

ATYPICAL SPEECH ANALYSIS AND DETECTION

ADAPTATION, TRANSFER LEARNING, AND DISTILLATION FOR ASR

SPEAKER AND LANGUAGE RECOGNITION I

PATHOLOGICAL SPEECH ANALYSIS

CROSS/MULTI-LINGUAL ASR

SPEAKING STYLES AND INTERACTION STYLES I

VOLUME 4

DEREVERBERATION AND ECHO CANCELLATION

VOICE CONVERSION AND ADAPTATION III

NOVEL MODELS AND TRAINING METHODS FOR ASR III

SPOKEN LANGUAGE MODELING AND UNDERSTANDING

ACOUSTIC SIGNAL REPRESENTATION AND ANALYSIS I

PRIVACY AND SECURITY IN SPEECH COMMUNICATION

MULTIMODAL SYSTEMS

ATYPICAL SPEECH DETECTION

SPOOFING-AWARE AUTOMATIC SPEAKER VERIFICATION (SASV) I

SINGLE-CHANNEL AND MULTI-CHANNEL SPEECH ENHANCEMENT

VOICE CONVERSION AND ADAPTATION II

RESOURCE-CONSTRAINED ASR

SPEECH PRODUCTION, PERCEPTION AND MULTIMODALITY

VOLUME 5

MULTI-, CROSS-LINGUAL AND OTHER TOPICS IN ASR II

SPOKEN LANGUAGE PROCESSING III

NON-INTRUSIVE OBJECTIVE SPEECH QUALITY ASSESSMENT (NISQA) CHALLENGE FOR ONLINE CONFERENCING APPLICATIONS

SPEECH AND LANGUAGE IN HEALTH: FROM REMOTE MONITORING TO MEDICAL CONVERSATIONS III

SPEECH SYNTHESIS: PROSODY MODELING

SELF-SUPERVISED, SEMI-SUPERVISED, ADAPTATION AND DATA AUGMENTATION FOR ASR

PHONETICS AND PHONOLOGY

SPOKEN LANGUAGE UNDERSTANDING II

SPEECH INTELLIGIBILITY PREDICTION FOR HEARING-IMPAIRED LISTENERS I

LOW-RESOURCE ASR DEVELOPMENT I

SPEECH REPRESENTATION I

PATHOLOGICAL SPEECH ASSESSMENT

ASR: LINGUISTIC COMPONENTS

VOLUME 6

SPEECH INTELLIGIBILITY PREDICTION FOR HEARING-IMPAIRED LISTENERS II

SHOW AND TELL III(VR)

SUMMARIZATION, ENTITY EXTRACTION, EVALUATION AND OTHERS

AUTOMATIC ANALYSIS OF PARALINGUISTICS

SELF SUPERVISION AND ANTI-SPOOFING

SPEECH ARTICULATION & NEURAL PROCESSING

ACOUSTIC SCENE ANALYSIS

SPEECH CODING AND PRIVACY

SPEECH SYNTHESIS: SINGING, MULTIMODAL, CROSSLINGUAL SYNTHESIS

APPLICATIONS IN TRANSCRIPTION, EDUCATION AND LEARNING II

STREAMING ASR

APPLICATIONS IN TRANSCRIPTION, EDUCATION AND LEARNING I

SPOKEN DIALOGUE SYSTEMS

THE VOICEMOS CHALLENGE

SPEECH SYNTHESIS: SPEAKING STYLE, EMOTION AND ACCENTS I

SPEECH SEGMENTATION I

HUMAN SPEECH & SIGNAL PROCESSING

SPEECH EMOTION RECOGNITION II

VOLUME 7

SPEAKER RECOGNITION AND ANTI-SPOOFING

MISCELLANEOUS TOPICS IN SPEECH, VOICE AND HEARING DISORDERS

LOW-RESOURCE ASR DEVELOPMENT II

VOICE CONVERSION AND ADAPTATION I

SEARCH/DECODING ALGORITHMS FOR ASR

EMOTIONAL SPEECH PRODUCTION AND PERCEPTION

SPEECH ANALYSIS

TRUSTWORTHY SPEECH PROCESSING

SPEAKER RECOGNITION AND DIARIZATION

SELF-SUPERVISED, SEMI-SUPERVISED, ADAPTATION AND DATA AUGMENTATION FOR ASR II

SPOKEN LANGUAGE PROCESSING I

SOURCE SEPARATION III

SPEECH ENHANCEMENT AND INTELLIGIBILITY

SPEECH SYNTHESIS: SPEAKING STYLE, EMOTION AND ACCENTS II

SHOW & TELL IV(VR)