

AES International Conference on Artificial Intelligence and Machine Learning for Audio 2025

London, United Kingdom
8 - 10 September 2025

ISBN: 979-8-3313-3434-5

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.

Copyright© (2025) by the Audio Engineering Society
All rights reserved.

Printed with permission by Curran Associates, Inc. (2026)

For permission requests, please contact the Audio Engineering Society
at the address below.

Audio Engineering Society
International Headquarters
551 Fifth Ave., Suite 1225
New York, NY 10176
USA

Phone: +1 212 661 8528

www.aes.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

PAPERS

Motor2Synth: Leveraging Differentiable Digital Signal Processing for Generating Combustion Engine Sounds Compatible with Active Sound Design Frameworks.....	3
<i>T.H.G. Lobato, S. Hank, Z. Hanyi, L. Haofu</i>	
Sound Matching an Analogue Levelling Amplifier Using the Newton-Raphson Method	12
<i>C.Y. Yu, G. Fazekas</i>	
Compression of Scene-Based Higher Order Ambisonics with Multichannel RVQGAN	22
<i>T. Hirvonen, M. Namazi</i>	
Extraction and Neural Synthesis of Low-Order Spectral Components for Head-Related Transfer Functions	32
<i>M. Pilataki, C.G. Buchanan, C. Armstrong</i>	
Hybrid Learning-Based Active Noise Control in Encapsulated Structures	41
<i>A. Aboutiman, H.R. Karimi, F. Ripamonti</i>	
Flute Tone Quality Classification: A Machine-Learning-Based Evaluation Tool.....	50
<i>N. Sane, J. Abel</i>	
Automatic Audio Equalization with Semantic Embeddings.....	60
<i>E. Moliner, V. Valimaki, K. Drossos, M.S. Hamalainen</i>	
Transfer Learning for Neural Modelling of Nonlinear Distortion Effects.....	70
<i>T. Vanhatalo, P. Legrand, M. Desainte-Catherine, P. Hanna, G. Pille, A. Brusco, J.D. Reiss</i>	
Perceptions of an Artificial Intelligence Musical Collaborator	80
<i>B. Allen, R. Mo</i>	
Multiple Loudspeaker Localization with Simultaneous Deconvolution.....	90
<i>S. Bharitkar, A. Celestinos</i>	
Complex-Valued Physics-Informed Neural Networks for Sound Field Estimation	98
<i>V.S. Paul, N. Hahn, P.A. Nelson</i>	
Compressing Neural Network Models of Audio Distortion Effects Using Knowledge Distillation Techniques.....	108
<i>R. Simionato, A. Tidemann</i>	
A Scalable AI Architecture for Audio and Multimodal Analysis on Mobile Devices: A Case of Environmental Monitoring	118
<i>M.E. Stamatiadou, A. Mpesmerti, N. Vryzas, L. Vrysis, C. Dimoulas</i>	
Challenges in Predicting the Lyric Intelligibility of Musical Segments for Older Individuals with Hearing Loss.....	128
Unstable Audio: Code Bending Text-To-Music Generation.....	135
<i>N. Collins</i>	

Simulating 3D Acoustic Radiation and Scattering in the Frequency Domain with Fourier Neural Operators (FNOs).....	142
<i>J. Hipperson, T. Cox, J. Hargreaves</i>	
AudioGAN: A Compact and Efficient Framework for Real-Time High-Fidelity Text-To-Audio Generation	152
<i>H.C. Chung</i>	
Predicting Binaural Colouration Using VGGish Embeddings	162
<i>T. McKenzie, A. Wright, D. Turner, P. Llado</i>	
A Dataset of Amateur Karaoke Singing and Audiobook Narration	172
<i>E. Georgieva, P. Ripolles, B. McFee</i>	
It's All About Speed: AI's Impact on Workflow in Music Production	182
<i>F. McClellan, F. Morreale</i>	
From CNN to Reservoir Computing: A New Perspective on Acoustic Scene Classification	192
<i>Y. He, A.M.A. Hosseini, J. Abeber, L. Jaurigue, A. Raake, K. Ludge</i>	
Improvement and Cross-Domain Evaluation of SlowFast Networks	202
<i>R. Kumar, S. Grollmisch, J. Abeber</i>	
Improved Singing Voice Conversion with Frame-Level Content and Melody-Informed Speaker Embeddings Using Cross-Attention	211
<i>J.W. Yeh, E.M. Liu, Y.W. Liu</i>	
Procedural Music Generation Systems in Games	220
<i>S. Luo, J. Reiss</i>	
Establishing a Virtual Listener Panel for Audio Characterisation	230
<i>M. Herlufsen, N.A. Fuglsang, B. Pedersen, S.V. Legarth</i>	
A Machine Learning Approach to Modal Control in Small Rooms	238
<i>C.M. Bolla, T.J. Cox, B.M. Fazenda</i>	
Perceiving AI in Music: Human Evaluation of AI-Generated Melodies and AI Detection Sensitivity	247
<i>M. Oehler, J. Oldach, F. Zwibler</i>	
Faust Autodiff: Towards Audio Domain-Specific Machine Learning	256
<i>T. Rushton, Y. Orlarey, R. Michon, T. Risset, S. Letz</i>	
Integrating IP Broadcasting with Audio Tags: Workflow and Challenges.....	266
<i>R. Burchett-Vass, A. Singh, G. Bibbo, M.D. Plumbley</i>	
Perceptual Evaluation of Machine Learning and Non-ML Emulations of the Vox AC30 Amplifier.....	274
<i>M.A.V. Reyes, G. Kearney, M. McLoughlin</i>	
Broadcast-Quality Synthetic Narration: A Workflow for Fine-Grained Text-To-Speech Intonation and Emotion Control	284
<i>L.F. Kruszielski, P.H.L. Leite, M.P. Fernandes, A. Pereira, L.W.P. Biscaíno</i>	
Neutone SDK: An Open Source Framework for Neural Audio Processing.....	293
<i>C. Mitcheltree, B. Teleaga, A. Fyfe, N. Masuda, M. Schafer, A. Bradic, N. Tokui</i>	

Adaptive Neural Audio Mixing with Human-In-The-Loop Feedback: A Reinforcement Learning Approach	303
<i>S. Zhu, M. Nasim</i>	

NabIAFx: A Framework for Differentiable Black-Box and Gray-Box Modeling of Audio Effects	313
<i>M. Comunita, C.J. Steinmetz, J.D. Reiss</i>	

LATE BREAKING DEMO PAPERS

OBSIDIAN-Neural: Real-Time AI Music Generation Plugin for Live Performance.....	325
<i>A. Charretier</i>	

Bias Research on Generated Music from Around the World.....	327
<i>A. Solak, F. Grotschla, L.A. Lanzendorfer, R. Wattenhofer</i>	

Real-Time Low Latency Audio Features with Resonate	329
<i>A.R.J. Francois</i>	

Real-Time Audio Monitoring Pipeline with Edge Inference and IoT Supervision Via ThingsBoard.....	331
<i>J. Naranjo-Alcazar, J. Grau-Haro, R. Ribes-Serrano, P. Zuccarello</i>	

Class-Aware Hybrid Ensemble for Query-By-Vocal Imitation	333
<i>R. Peter, V. Mohan</i>	

Investigating the Use of Deep Convolutional Neural Networks for Direction-Of-Arrival Estimation on Raw Stereo Audio.....	335
<i>S. Hobern, A. Archer-Boyd, D. Murphy</i>	

Beat-Based Rhythm Quantization of MIDI Performances	337
<i>M. Wachter, S. Murgul, M. Heizmann</i>	

Decoding Melodic Acoustic Features from Neural Data.....	339
<i>Z. Bozilovic, I.R. Roman</i>	

Perceptual Probing of Deep Audio Representations.....	341
<i>E. Bost, R. Kronland-Martinet, M. Aramaki, S. Ystad, T. Artieres, T. Schatz</i>	

Black-Box Modeling of Immersive Reverberation with Convolutional Wavenet Audio Effects Modeling Network	343
<i>Y. Buruk, T. Ozdemir</i>	

Transformer-Based Sustain Pedal Reconstruction for Expressive Piano Performance MIDI	345
<i>W. Liu, G. Fazekas, J. Tang</i>	

Latent Rhythm Transformation of Drum Recordings.....	347
<i>J. Hockman, J. Drysdale</i>	

FLARE: An Open-Source Library for RIR Synthesis and Analysis in PyTorch	349
<i>G. Dal Santo, K. Prawda, S.J. Schlecht, V. Valimaki</i>	

Towards Leveraging Machine Learning for Multichannel Audio Upmix	351
<i>J.M. Jot, G. Wichern, T. Paez, S. Bharitkar</i>	

Deep Localization Refinement for Bat Abundance Estimation.....	353
<i>R. Alessandri, L. Gilmour, J.J. Tan, A. Osses, D. Stowell</i>	

Denoising Marine Recordings of Humpback Whales	355
<i>T. Gifford</i>	
MULTIVOX: Spatial & Multimodal Data of Singing Groups	357
<i>G.M. Gomez, M. Sepulveda, A. Roman, R.S. Sefchovich, I. Roman</i>	
Towards Intelligent Music Education: Score-Informed Transcription and Performance Assessment	359
<i>J. Loth, A.M. Primenta, J. Tang, X. Riley, S. Dixon, E. Benetos</i>	
Real-Time Voice Conversion with Online Speaker Adaptation for Creative Audio Processing	361
<i>A.R. Bargum, S. Serafin, C. Erkut</i>	

Author Index