

# **2024 IEEE 5th International Symposium on the Internet of Sounds (IS2 2024)**

**Erlangen, Germany  
30 September - 2 October 2024**



**IEEE Catalog Number: CFP24UB1-POD  
ISBN: 979-8-3503-6653-2**

**Copyright © 2024 by the Institute of Electrical and Electronics Engineers, Inc.  
All Rights Reserved**

*Copyright and Reprint Permissions:* Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

***\*\*\* This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

|                         |                   |
|-------------------------|-------------------|
| IEEE Catalog Number:    | CFP24UB1-POD      |
| ISBN (Print-On-Demand): | 979-8-3503-6653-2 |
| ISBN (Online):          | 979-8-3503-6652-5 |

**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  
E-mail: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

CURRAN ASSOCIATES INC.  
**proceedings**  
.com

## Table of Contents

|                                                                                                                                            |     |
|--------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Multi-Signal Informed Attention for Beat and Downbeat Detection .....                                                                      | 1   |
| <i>James Bolt, Johan Pauwels and George Fazekas</i>                                                                                        |     |
| Can You DAW it Online? The Challenges of a Web-Based Open Source Workstation .....                                                         | 8   |
| <i>Michel Buffa and Samuel Demont</i>                                                                                                      |     |
| SoundSignature: What Type of Music Do You Like? .....                                                                                      | 16  |
| <i>Brandon Carone and Pablo Ripollés</i>                                                                                                   |     |
| tinyVocos: Neural Vcoders on MCUs .....                                                                                                    | 26  |
| <i>Stefano Ciapponi, Francesco Paissan, Alberto Ancilotto and Elisabetta Farella</i>                                                       |     |
| Ethernet Real-time Audio Transmission to FPGA .....                                                                                        | 36  |
| <i>Pierre Cochard, Jurek Weber, Romain Michon, Tanguy Risset and Stéphane Letz</i>                                                         |     |
| DECT-2020 NR for Professional Live Audio: Design Space Exploration and Practical Evaluation .....                                          | 43  |
| <i>Jan Dürre and Norbert Werner</i>                                                                                                        |     |
| Evaluation of Compressed Higher-Order Ambisonics at Off-Center Listening Positions ....                                                    | 50  |
| <i>Matthias Frank and Sonja Lutz</i>                                                                                                       |     |
| FLCrypt – Secure Federated Learning for Audio Event Classification using Homomorphic Encryption .....                                      | 57  |
| <i>Kay Fuhrmeister, Hao Cui, Artem Yaroshchok and Thomas Köllmer</i>                                                                       |     |
| The OCON model: an old but green solution for distributable supervised classification for acoustic monitoring in smart cities .....        | 64  |
| <i>Stefano Giacomelli, Marco Giordano and Claudia Rinaldi</i>                                                                              |     |
| TinyChirp: Bird Song Recognition Using TinyML Models on Low-power Wireless Acoustic Sensors .....                                          | 74  |
| <i>Zhaolan Huang, Adrien Tousnakhoff, Polina Kozyr, Roman Rehausen, Felix Bießmann, Robert Lachlan, Cedric Adjih and Emmanuel Baccelli</i> |     |
| Enhancing Video Music Recommendation with Transformer-Driven Audio-Visual Embeddings .....                                                 | 84  |
| <i>Shimiao Liu and Alexander Lerch</i>                                                                                                     |     |
| Rebound Effects in Digital Audio: Past and Present .....                                                                                   | 90  |
| <i>Vincent LOSTANLEN and Lucie Bouchet</i>                                                                                                 |     |
| Dice Cubes: Quantum-Based Networked Performance with CubeHarmonic .....                                                                    | 97  |
| <i>Maria Mannone and Luca Turchet</i>                                                                                                      |     |
| Large-Scale Room Impulse Response Dataset Compression With Neural Audio Codecs ....                                                        | 102 |
| <i>Alessandro Ilic Mezza, Alberto Bernardini and Fabio Antonacci</i>                                                                       |     |
| Benchmarking Music Demixing Models for Deep Drum Source Separation .....                                                                   | 110 |
| <i>Alessandro Ilic Mezza, Riccardo Giampiccolo, Alberto Bernardini and Augusto Sarti</i>                                                   |     |
| Noise-to-mask Ratio Loss for Deep Neural Network based Audio Watermarking .....                                                            | 116 |
| <i>Martin Moritz, Toni Olán and Tuomas Virtanen</i>                                                                                        |     |

|                                                                                                                                    |     |
|------------------------------------------------------------------------------------------------------------------------------------|-----|
| Telemersive audio systems for online jamming .....                                                                                 | 122 |
| <i>Patrick Müller, Roman Haefeli, Johannes Schütt and Matthias Ziegler</i>                                                         |     |
| Towards Measuring and Forecasting Noise Exposure at the VELTINS-Arena in Gelsenkirchen, Germany .....                              | 137 |
| <i>Pitchapa Ngamthipwatthana, Marco Götze, András Kátai and Jakob Abeßer</i>                                                       |     |
| GENPIA: A Genre-Conditioned Piano Music Generation System .....                                                                    | 145 |
| <i>Quoc-Viet Nguyen, Hao-Wei Lai, Khanh-Duy Nguyen, Min-Te Sun, Wu-Yuin Hwang, Kazuya Sakai and Wei-Shinn Ku</i>                   |     |
| Internet of Musical Things for Children .....                                                                                      | 155 |
| <i>Luc Nijs and Luca Turchet</i>                                                                                                   |     |
| Extending Networked Mapping Research Middleware Into the Browser Sandbox .....                                                     | 164 |
| <i>Matthew Peachey, Kyle Smith and Joseph Malloch</i>                                                                              |     |
| Deepfake Audio Detection Using Spectrogram-based Feature and Ensemble of Deep Learning Models .....                                | 170 |
| <i>Lam Pham, Phat Lam, Truong Nguyen, Huyen Nguyen and Alexander Schindler</i>                                                     |     |
| Supporting the Identification of Natural Sounds in Ecoacoustic Soundscapes through Spectrogram-based Analysis Methods .....        | 175 |
| <i>Jana Popp, Bernd Edler and Cornelius Poepel</i>                                                                                 |     |
| BCHJam: a Brain-Computer Music Interface for Live Music Performance in Shared Mixed Reality Environments .....                     | 184 |
| <i>Michele Romani, Gregorio Andrea Giudici, Domenico Stefani, Devis Zanoni, Alberto Boem and Luca Turchet</i>                      |     |
| ANIRA: An Architecture for Neural Network Inference in Real-Time Audio Applications .                                              | 193 |
| <i>Fares Schulz and Valentin Ackva</i>                                                                                             |     |
| Introducing DUST: a Dataset of real-time UDP Sound packet Traces .....                                                             | 203 |
| <i>Leonardo Severi, Matteo Sacchetto, Cristina Rottondi and Andrea Bianco</i>                                                      |     |
| Prototyping a sustainable GUI for an IoMusT project: A reflection-on-action on the adoption of sustainable practices .....         | 207 |
| <i>Elisa Silene Panozzo, Mela Bettega, Antonio Rodà and Nicolo' Merendino</i>                                                      |     |
| AudioInsight: Online exploration of Large Audio Datasets for Musical Acoustics .....                                               | 217 |
| <i>Michael Starakis and Chrisoula Alexandraki</i>                                                                                  |     |
| RBS-based Time Synchronization Approach with Autonomous Acoustic Sensors – a Simulative Proof of Concept .....                     | 227 |
| <i>Alexander Tessmer, Leonhard Brüggemann and Nils Aschenbruck</i>                                                                 |     |
| The Internet of Musical Things Meets Satellites: Evaluating Starlink Support for Networked Music Performances in Rural Areas ..... | 235 |
| <i>Luca Turchet and Paolo Casari</i>                                                                                               |     |
| MusicoNet: a social network for musicians based on the Internet of Musical Things and People paradigm .....                        | 243 |
| <i>Luca Turchet, Jacopo Tomelleri and Alessandro Ruffo</i>                                                                         |     |

|                                                                                                                                     |     |
|-------------------------------------------------------------------------------------------------------------------------------------|-----|
| Remastering Divide and Remaster: A Cinematic Audio Source Separation Dataset with Multilingual Support .....                        | 252 |
| <i>Karn Watcharasupat, Chih-Wei Wu and Iroko Orife</i>                                                                              |     |
| Real-Time Automatic Drum Transcription Using Dynamic Few-Shot Learning .....                                                        | 262 |
| <i>Philipp Weber, Christian Uhle, Meinard Müller and Matthias Lang</i>                                                              |     |
| "It Takes Two" - Shared and Collaborative Virtual Musical Instruments in the Musical Metaverse .....                                | 270 |
| <i>Alberto Boem, Damian Dziwis, Matteo Tomasetti, Sascha Etezazi and Luca Turchet</i>                                               |     |
| Presence and Flow in Virtual and Mixed Realities for Music-Related Educational Settings                                             | 280 |
| <i>Leonard Bruns, Benedict Saurbier, Tray Minh Voong and Michael Oehler</i>                                                         |     |
| Using Web Audio Modules for Immersive Audio Collaboration in the Musical Metaverse ..                                               | 287 |
| <i>Michel Buffa, Dorian Girard and Ayoub Hofr</i>                                                                                   |     |
| Holodeck: A Research Framework for Distributed Multimedia Concert Performances .....                                                | 297 |
| <i>Andrea Genovese, Zack Nguyen, Marta Gospodarek, Robert Pahle, Corinne Brenner and Agnieszka Roginska</i>                         |     |
| Key Considerations for Duo Singing in Virtual Reality and Videoconferencing: An Exploratory Study with Bigscreen and Zoom .....     | 307 |
| <i>Ben Loveridge</i>                                                                                                                |     |
| The Musical Metaverse: Advancements and Applications in Networked Immersive Audio .                                                 | 314 |
| <i>Claudia Rinaldi and Carlo Centofanti</i>                                                                                         |     |
| Remote Orchestral Conduction via a Virtual Reality System .....                                                                     | 321 |
| <i>Leonardo Severi, Matteo Sacchetto, Andrea Bianco, Cristina Rottondi, Gabriele Abbate, Antonio Paolillo and Alessandro Giusti</i> |     |
| Immersive Io3MT Environments: Design Guidelines, Use Cases and Future Directions ....                                               | 327 |
| <i>Rômulo Vieira, Shu Wei, Thomas Röggl, Débora C. Muchaluat-Saade and Pablo César</i>                                              |     |
| Ethical Considerations in the Production and Consumption of Music in the Metaverse ....                                             | 337 |
| <i>Gareth W. Young</i>                                                                                                              |     |
| Investigation of Server-Based Spatial Audio for Metaverse Concert Distribution .....                                                | 346 |
| <i>Patrick Cairns, Helena Daffern and Gavin Kearney</i>                                                                             |     |
| On the Use of a Hybrid Linear-ERB Frequency Scale for Lightweight Simulation of HRTFs                                               | 354 |
| <i>Maurício Do Vale Madeira da Costa, Luiz Wagner Pereira Biscainho and Michael Oehler</i>                                          |     |
| Use-Cases of the new 3GPP Immersive Voice and Audio Services (IVAS) Codec and a Web Demo Implementation .....                       | 362 |
| <i>Eleni Fotopoulou, Kacper Sagnowski, Karin Prebeck, Moumita Chakraborty, Suhas Medicherla and Stefan Döhla</i>                    |     |
| Locally Adapted Immersive Environments for Distributed Music Performance in Mixed Reality .....                                     | 368 |
| <i>Andrea Genovese, Marta Gospodarek, Zack Nguyen, Robert Pahle and Agnieszka Roginska</i>                                          |     |

|                                                                                                                             |     |
|-----------------------------------------------------------------------------------------------------------------------------|-----|
| Remote Immersive Audio Production: State of the Art Implementation, Challenges, and Improvements .....                      | 378 |
| <i>Stefano Giacomelli, Carlo Centofanti, José Santos, Mauro Galbiati, Tiziano Salvi, Fabio Graziosi and Claudia Rinaldi</i> |     |
| EEG Hyperscanning in the Internet of Sounds: Low-Delay Real-time multi-modal Transmission using the OVBOX.....              | 388 |
| <i>Giso Grimm, Mareike Daeglau, Volker Hohmann and Stefan Debener</i>                                                       |     |
| Immersive Networked Music Performance: Impact of Extended Reality on the Quality of Experience .....                        | 396 |
| <i>Robert Hupke, Stephan Preihs and Juergen Peissig</i>                                                                     |     |
| Metadata-Assisted Spatial Audio (MASA) - An Overview .....                                                                  | 405 |
| <i>Jouni Paulus, Lasse Laaksonen, Tapani Pihlajakujä, Mikko-Ville Laitinen, Juha Vilkamo and Adriana Vasilache</i>          |     |