

2022 IEEE International Symposium on Multimedia (ISM 2022)

**Naples, Italy
5-7 December 2022**



**IEEE Catalog Number: CFP22197-POD
ISBN: 978-1-6654-7173-2**

**Copyright © 2022 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:	CFP22197-POD
ISBN (Print-On-Demand):	978-1-6654-7173-2
ISBN (Online):	978-1-6654-7172-5

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
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings
.com

2022 IEEE International Symposium on Multimedia (ISM) **ISM 2022**

Table of Contents

Message from the General Co-Chairs	xii
Message from the Program Co-Chairs	xiii

Session 1: ISM-1

Complete Cross-Triplet Loss in Label Space for Audio-Visual Cross-Modal Retrieval	1
<i>Donghuo Zeng (KDDI Research, Inc., Japan), Yanan Wang (KDDI Research, Inc., Japan), Jianming Wu (KDDI Research, Inc., Japan), and Kazushi Ikeda (KDDI Research, Inc., Japan)</i>	
Interactive RGB Image Segmentation via Depth-Modified Click Encoding and Estimated Depth	10
<i>Furkan Kaynar (Technical University of Munich, Germany), Adrian Michl (Technical University of Munich, Germany), and Eckehard Steinbach (Technical University of Munich, Germany)</i>	
Efficient Multi-Threading Strategies in VVenC, an Open and Optimized VVC Encoder Implementation	18
<i>Valeri George (Fraunhofer HHI, Germany), Jens Brandenburg (Fraunhofer HHI, Germany), Gabriel Hege (Fraunhofer HHI, Germany), Tobias Hinz (Fraunhofer HHI, Germany), Adam Wieckowski (Fraunhofer HHI, Germany), Benjamin Bross (Fraunhofer HHI, Germany), and Detlev Marpe (Fraunhofer HHI, Germany)</i>	
Effects of Color Stain Normalization in Histopathology Image Retrieval Using Deep Learning	26
<i>Antonio M. Rinaldi (University of Napoli Federico II, Italy), Cristiano Russo (University of Napoli Federico II, Italy), and Cristian Tommasino (University of Napoli Federico II, Italy)</i>	

Session 3: ISM-2 - Best Paper Session (I)

Robust Depth Estimation in Foggy Environments Combining RGB Images and mmWave Radar	34
<i>Mengchen Xiong (Technical University of Munich, Germany), Xiao Xu (Technical University of Munich, Germany), Dong Yang (Technical University of Munich, Germany), and Eckehard Steinbach (Technical University of Munich, Germany)</i>	

Segmentation Consistency Training: Out-of-Distribution Generalization for Medical Image Segmentation	42
<i>Birk Torpmann-Hagen (University of Tromsø, Norway; SimulaMet, Norway; University of Oslo, Norway), Vajira Thambawita (SimulaMet, Norway), Michael A. Riegler (University of Tromsø, Norway; SimulaMet, Norway), Pål Halvorsen (SimulaMet, Norway; University of Oslo, Norway), and Kyrre Glette (University of Oslo, Norway)</i>	
Aggregated Bidirectional Local Binary Pattern for Robust Perceptual Image Hashing	50
<i>Qasim Abbas (La Trobe University, Australia), S. Jannat Shirazi (Riphah International University, Pakistan), and Yi-Ping Phoebe Chen (La Trobe University, Australia)</i>	

Session 5: ISM-3 - Best Paper Session (II)

TAME: Attention Mechanism Based Feature Fusion for Generating Explanation Maps of Convolutional Neural Networks	58
<i>Mariano Ntroungkas (CERTH-ITI, Greece), Nikolaos Gkalelis (CERTH-ITI, Greece), and Vasileios Mezaris (CERTH-ITI, Greece)</i>	
Detecting Happiness from 14-Channel Binary-Valued EEG Charts via Deep Learning	66
<i>Michele Baldassini (University of Pisa, Italy), Francesco Pistolesi (University of Pisa, Italy), and Beatrice Lazzerini (University of Pisa, Italy)</i>	

Session 7: ISM-4

Impact of Conventional and Deep Learning-Based Point Cloud Geometry Coding on Deep Learning-Based Classification Performance	74
<i>Abdelrahman Seleem (Instituto Superior Técnico - Universidade de Lisboa and Instituto de Telecomunicações, Portugal; South Valley University, Egypt), André F. R. Guarda (Instituto de Telecomunicações, Portugal), Nuno M. M. Rodrigues (ESTG – Politécnico de Leiria and Instituto de Telecomunicações, Portugal), and Fernando Pereira (Instituto Superior Técnico - Universidade de Lisboa and Instituto de Telecomunicações, Portugal)</i>	
Teardrop Magnification: A Hybrid Linear-Fisheye Magnifier for the Border and Corner of the Screen	82
<i>Florian Schmiederjann (HSW University of Applied Sciences, Germany), Darius Rausch (HSW University of Applied Sciences, Germany), Jens Wiggenbrock (HSW University of Applied Sciences, Germany), and Robert Mertens (HSW University of Applied Sciences, Germany; BHH University of Applied Sciences, Germany)</i>	
Multimodal 2D/3D Registration for Open Augmented Reality Applications	84
<i>Madjid Maida (ESME Engineering School, France) and Samir Otmane (Evry University, France)</i>	
On Requirements for Field of Light Displays to Pass the Visual Turing Test	86
<i>Matthew Hamilton (Memorial University of Newfoundland, Canada), Nicholas Wells (Memorial University of Newfoundland, Canada), and Amilcar Soares (Memorial University of Newfoundland, Canada)</i>	

Session 11: ISM-5

To Sparsify or not to Sparsify: Simplifying Visual Feature Maps for Mobile Agents	88
<i>Sebastian Eger (Technical University of Munich, Germany), Rastin Pries (Nokia Bell Labs, Hungary), Gábor Sörös (Nokia Bell Labs, Hungary), Michael G. Adam (Technical University of Munich, Germany), Martin Piccolrovazzi (Technical University of Munich, Germany), and Eckehard Steinbach (Technical University of Munich, Germany)</i>	
Analysis of Heart Sound Signals Using Sparse Modeling with Gabor Dictionary	92
<i>Mahmoud Fakhry (Aswan University, Egypt; Carlos III University of Madrid, Spain), Abeer FathAllah Brery (Aswan University, Egypt), and Ascensión Gallardo-Antolín (Carlos III University of Madrid, Spain)</i>	

Session 12: ISM-6

Improving Multimodal Object Detection with Individual Sensor Monitoring	97
<i>Christopher B. Kuhn (BMW Group, Germany), Goran Petrovic (BMW Group, Germany), Markus Hofbauer (Technical University Munich, Germany), Ma Bowen (Technical University Munich, Germany), and Eckehard Steinbach (Technical University Munich, Germany)</i>	
Self-Supervised Object Recognition Based on Repeated Re-Capturing of Dynamic Indoor Environments	105
<i>Martin Piccolrovazzi (Technical University of Munich, Germany), Michael G. Adam (Technical University of Munich, Germany), Sebastian Eger (Technical University of Munich, Germany), Marsil Zakour (Technical University of Munich, Germany), and Eckehard Steinbach (Technical University of Munich, Germany)</i>	
Gated-ViGAT: Efficient Bottom-Up Event Recognition and Explanation Using a New Frame Selection Policy and Gating Mechanism	113
<i>Nikolaos Gkalelis (CERTH-ITI, Greece), Dimitrios Daskalakis (CERTH-ITI, Greece), and Vasileios Mezaris (CERTH-ITI, Greece)</i>	
Semantic-Aware View Prediction for 360-Degree Videos at the 5G Edge	121
<i>Shivi Vats (University of Klagenfurt, Austria), Jounsup Park (California Baptist University, USA), Klara Nahrstedt (University of Illinois, USA), Michael Zink (University of Massachusetts, USA), Ramesh Sitaraman (University of Massachusetts, USA), and Hermann Hellwagner (University of Klagenfurt, Austria)</i>	
S2CMF: Multi-Method Assessment Fusion for Scan-to-CAD Methods	129
<i>Driton Salihu (Technical University of Munich), Adam Misik (Technical University of Munich; Siemens Technology, Germany), Markus Hofbauer (Technical University of Munich), and Eckehard Steinbach (Technical University of Munich)</i>	

Session 14: ISM-7

Low-Precision Post-Filtering in Video Coding	137
<i>Ruiying Yang (Nokia Technologies, Finland), Maria Santamaria (Nokia Technologies, Finland), Francesco Cricri (Nokia Technologies, Finland), Honglei Zhang (Nokia Technologies, Finland), Jani Lainema (Nokia Technologies, Finland), Ramin G. Youvalari (Nokia Technologies, Finland), and Miska M. Hannuksela (Nokia Technologies, Finland)</i>	
The Lottery Ticket Adaptation for Neural Video Coding	141
<i>Nannan Zou (Tampere University, Finland), Francesco Cricri (Nokia Technologies, Finland), Honglei Zhang (Nokia Technologies, Finland), Hamed R. Tavakoli (Nokia Technologies, Finland), Miska M. Hannuksela (Nokia Technologies, Finland), and Esa Rahtu (Tampere University, Finland)</i>	
Explaining Video Summarization Based on the Focus of Attention	146
<i>Evlampios Apostolidis (CERTH-ITI & Queen Mary University of London, Greece), Georgios Balaouras (CERTH-ITI, Greece), Vasileios Mezaris (CERTH-ITI, Greece), and Ioannis Patras (Queen Mary University of London, UK)</i>	
Measuring the Influence of Image Preprocessing on the Rate-Distortion Performance of Video Encoding	151
<i>Markus Hofbauer (Technical University of Munich, Germany), Christopher B. Kuhn (Technical University of Munich, Germany), Goran Petrovic (BMW Group, Germany), and Eckehard Steinbach (Technical University of Munich, Germany)</i>	

Session 18: ISM-8

Towards Efficient Video Super Resolution for Faster Streaming	153
<i>Himanshu Gupta (University of Massachusetts at Amherst, USA), Sowmya Vasuki Jallepalli (University of Massachusetts at Amherst, USA), Pratik Mulchandani (University of Massachusetts at Amherst, USA), Chirag Trasikar (University of Massachusetts at Amherst, USA), Chetan Manjesh (University of Massachusetts at Amherst, USA), Vishy Swaminathan (Adobe Research, USA), and Stefano Petrangeli (Adobe Research, USA)</i>	
Detection of Commercial Fishing-Related Slipping Events Using Multimodal Data	155
<i>Tor-Arne S. Nordmo (UiT: The Arctic University of Norway, Norway), Martine M. Espeseth (KSAT, Norway), Bjørn A. Juliussen (UiT: The Arctic University of Norway, Norway), Michael A. Riegler (UiT: The Arctic University of Norway, Norway), and Dag Johansen (UiT: The Arctic University of Norway, Norway)</i>	
Synthesizing Ultraviolet Skin Images via GAN with Gaussian Weighted Patch Blending	157
<i>Ruowei Jiang (ModiFace, Canada), Brendan Duke (University of Toronto, Canada), Frederic Flament (L'Oreal, Research & Innovation, France), and Parham Aarabi (University of Toronto, Canada)</i>	

A Smartphone app to Collect Emotion-Labeled Signals in the wild Using a Body Sensor Network	159
<i>Francesco Pistolesi (University of Pisa, Italy), Michele Baldassini (University of Pisa, Italy), and Beatrice Lazzerini (University of Pisa, Italy)</i>	
Experiences and Lessons Learned from a Crowdsourced-Remote Hybrid User Survey Framework	161
<i>Cise Midoglu (SimulaMet, Norway), Andrea Storås (SimulaMet, Norway; Oslo Metropolitan University, Norway), Saeed Shafiee Sabet (SimulaMet, Norway), Malek Hammou (SimulaMet, Norway), Steven Alexander Hicks (SimulaMet, Norway), Inga Strømke (SimulaMet, Norway; The Norwegian University of Science and Technology, Norway), Michael Alexander Riegler (SimulaMet, Norway; The Arctic University of Norway, Norway), Carsten Griwodz (University of Oslo, Norway), and Pål Halvorsen (SimulaMet, Norway; Oslo Metropolitan University, Norway)</i>	

Session 21: ISM-9

An Artificial Neural Network-Based System for Detecting Machine Failures Using a Tiny Sound Dataset: A Case Study	163
<i>Thanh Tran (Mid Sweden University; Group Technology Team, Sweden), Sebastian Bader (Mid Sweden University), and Jan Lundgren (Mid Sweden University)</i>	
Contextualized Styling of Images for Web Interfaces Using Reinforcement Learning	169
<i>Pooja Guhan (University of Maryland College Park, USA; Adobe Research, San Jose, USA), Saayan Mitra (Adobe Research, San Jose, USA), Somdeb Sarkhel (Adobe Research, San Jose, USA), Stefano Petrangeli (Adobe Research, San Jose, USA), Ritwik Sinha (Adobe Research, San Jose, USA), Viswanathan Swaminathan (Adobe Research, San Jose, USA), Aniket Bera (Purdue University, USA), and Dinesh Manocha (University of Maryland College Park, USA)</i>	
Benchmarking the Second Edition of the Omnidirectional Media Format Standard	173
<i>Burak Kara (Ozyegin University, Turkey), Mehmet N. Akcay (Ozyegin University, Turkey), Ali C. Begen (Ozyegin University, Turkey), Saba Ahsan (Nokia Technologies, Finland), Igor D.D. Curcio (Nokia Technologies, Finland), Kashyap Kammachi-Sreedhar (Nokia Technologies, Finland), and Emre Aksu (Nokia Technologies, Finland)</i>	
Optimizing Storage and Delivery of Omnidirectional Videos in Viewport-Dependent Streaming ...	177
<i>Kashyap Kammachi-Sreedhar (Nokia Technologies, Finland), Miska M. Hannuksela (Nokia Technologies, Finland), Emre Aksu (Nokia Technologies, Finland), Lauri Ilola (Nokia Technologies, Finland), and Lukasz Condrad (Nokia Technologies, Finland)</i>	

Session 23: ISM-10

Using Web 3D and WebXR Game to Enhance Engagement in Primary School Learning	181
<i>Xi Guo (Birmingham City University, UK) and Imran Mogra (Birmingham City University, UK)</i>	

Effectiveness of Machine Learning on Heartprint Biometric Recognition with Multisession ECG Signals	N/A
<i>Md. Saiful Islam (King Saud University, Saudi Arabia), Haikel Alhichri (King Saud University, Saudi Arabia), Yakoub Bazi (King Saud University, Saudi Arabia), Nassim Ammour (King Saud University, Saudi Arabia), Naif Alajlan (King Saud University, Saudi Arabia), and Rami M. Jomaa (University of Prince Mugrin, Saudi Arabia)</i>	
A Topological Approach for Facial Region Segmentation in Thermal Images	189
<i>Michael Lilley (University of Michigan-Dearborn, USA), Kapotaksha Das (University of Michigan-Dearborn, USA), Kais Riani (University of Michigan-Dearborn, USA), and Mohamed Abouelenien (University of Michigan-Dearborn, USA)</i>	

Session A: ISM-V1- Augmented Reality/ Virtual Reality

A Comparative Study of Interactive Environments for Investigative Interview of A Virtual Child Avatar	194
<i>Syed Zohaib Hassan (SimulaMet, Norway), Saeed Shafiee Sabet (SimulaMet, Norway), Pegah Salehi (SimulaMet, Norway), Hayley Ko (OsloMet, Norway), Ingvild Riiser (OsloMet, Norway), Miriam Johnson (OsloMet, Norway), Gunn Astrid Baugerud (OsloMet, Norway), Michael Alexander Riegler (SimulaMet, Norway), and Pål Halvorsen (SimulaMet, Norway)</i>	
On Selection of Efficient Sequential Pattern Mining Algorithm Based on Characteristics of Data	202
<i>Jakub Peschel (Masaryk University, Czechia), Michal Batko (Masaryk University, Czechia), and Pavel Zezula (Masaryk University, Czechia)</i>	
Hand Tracking vs Motion Controllers: The Effects on Immersive Virtual Reality Game Experience	206
<i>Sophia Neamoniti (Univeristy of the Aegean, Greece) and Vlasios Kasapakis (Univeristy of the Aegean, Greece)</i>	
On The Exploration of Vision Transformers in Remote Sensing Building Extraction	208
<i>G.-F. Angelis (Information and Technologies Institute, Centre for Research and Technology Hellas, Greece), A. Domi (Information and Technologies Institute, Centre for Research and Technology Hellas, Greece), A. Zamichos (Information and Technologies Institute, Centre for Research and Technology Hellas, Greece), M. Tsourma (Information and Technologies Institute, Centre for Research and Technology Hellas, Greece), A. Drosou (Information and Technologies Institute, Centre for Research and Technology Hellas, Greece), I. Manakos (Information and Technologies Institute, Centre for Research and Technology Hellas, Greece), and D. Tzovaras (Information and Technologies Institute, Centre for Research and Technology Hellas, Greece)</i>	
Contextually Guided Convolutional Neural Networks for Learning Most Transferable Representations	210
<i>Olcay Kursun (Auburn University at Montgomery, USA), Semih Dinc (EagleView Technologies, USA), and Oleg V. Favorov (University of North Carolina at Chapel Hill, USA)</i>	

Towards Accurate Positioning in Multiuser Augmented Reality on Mobile Devices	214
<i>Na Wang (George Mason University), Haoliang Wang (Adobe Research), Stefano Petrangeli (Adobe Research), Viswanathan Swaminathan (Adobe Research), Fei Li (George Mason University), and Songqing Chen (George Mason University)</i>	
Emotionally Expressive Motion Controller for Virtual Character Locomotion Animations	218
<i>Diogo Silva (Universidade de Lisboa, Portugal), Pedro A. Santos (Universidade de Lisboa, Portugal), and João Dias (University of Algarve, Portugal)</i>	

Session D: ISM-V2-Melody Generation and Retrieval

Singing Melody Extraction Based on Combined Frequency-Temporal Attention and Attentional Feature Fusion with Self-Attention	220
<i>Xi Qi (Xi'an Jiaotong University, China), Lihua Tian (Xi'an Jiaotong University, China), Chen Li (Xi'an Jiaotong University, China), Hui Song (University of PAP, China), and Jiahui Yan (Xi'an Jiaotong University, China)</i>	
Retaining Semantics in Image to Music Conversion	228
<i>Zeyu Xiong (The Hong Kong University of Science and Technology (Guangzhou), China), Pei-Chun Lin (Feng Chia University, Taiwan), and Amin Farjudian (University of Nottingham Ningbo China, China)</i>	
Deep Attention-Based Alignment Network for Melody Generation from Incomplete Lyrics	236
<i>Gurunath Reddy M (National Institute of Informatics, SOKENDAI, Japan), Zhe Zhang (National Institute of Informatics, SOKENDAI, Japan), Yi Yu (National Institute of Informatics, SOKENDAI, Japan), Florian Harscoet (National Institute of Informatics, SOKENDAI, Japan), Simon Canales (National Institute of Informatics, SOKENDAI, Japan), and Suhua Tang (National Institute of Informatics, SOKENDAI, Japan; The University of Electro-Communications)</i>	
Performance Analysis of Deep Learning Based Speech Quality Model with Mixture of Features	240
<i>Rahul Jaiswal (University of Agder, Norway)</i>	
HSD: A Hierarchical Singing Annotation Dataset	245
<i>Xiao Fu (Waseda University, Japan), Xin Yuan (Waseda University, Japan), and Jinglu Hu (Waseda University, Japan)</i>	
Comparison of Multi-Scale Speaker Vectors and S-Vectors for Zero-Shot Speech Synthesis	247
<i>Tristin Cory (Missouri State University, USA) and Razib Iqbal (Missouri State University, USA)</i>	

Session E: ISM-V3 - Image Processing (I)

Roundwood Tracking from the Forest to the Sawmill Using Filter Approaches to Highlight the Annual Ring Pattern	249
<i>Georg Wimmer (University of Salzburg, Austria), Rudolf Schraml (University of Salzburg, Austria), Andreas Uhl (University of Salzburg, Austria), and Alexander Petuschnigg (University of Applied Sciences Salzburg, Austria)</i>	

Task-Oriented Near-Lossless Burst Compression	257
<i>Weixin Jiang (Northwestern University, USA), Gang Wu (Adobe Research, Adobe Inc., USA), Viswanathan Swaminathan (Adobe Research, Adobe Inc., USA), Stefano Petrangeli (Adobe Research, Adobe Inc., USA), Haoliang Wang (Adobe Research, Adobe Inc., USA), Ryan Rossi (Adobe Research, Adobe Inc., USA), and Nedim Lipka (Adobe Research, Adobe Inc., USA)</i>	
Cloud Region Segmentation from All Sky Images Using Double K-Means Clustering	261
<i>Semih Dinc (EagleView Technologies, USA), Randy Russell (Auburn University Montgomery, USA), and Luis Alberto Cueva Parra (University of North Georgia, USA)</i>	
Toward Energy Efficient Curvature in Range Images	263
<i>Jacob D. Hauenstein (The University of Alabama in Huntsville) and Timothy S. Newman (The University of Alabama in Huntsville)</i>	

Session F: ISM-V4 - Image Processing (II)

Attention-Driven RetinaNet for Parasitic Egg Detection	265
<i>Khoa Pho (Japan Advanced Institute of Science and Technology, Japan), Han Lam (Japan Advanced Institute of Science and Technology, Japan), Tung Le (University of Science, Vietnam; Vietnam National University, Vietnam), Huy Tien Nguyen (University of Science, Vietnam; Vietnam National University, Vietnam), and Atsuo Yoshitaka (Japan Advanced Institute of Science and Technology, Japan)</i>	
Actor-Critic Bilateral Filter for Noise-Robust Image Smoothing	273
<i>Yi-Jie Chen (Yuan Ze University), Yen-Chiao Wang (Yuan Ze University), Bo-Hao Chen (Yuan Ze University), Hsiang-Yin Cheng (Yuan Ze University), and Jia-Li Yin (Fuzhou University)</i>	
Conditional GAN for Small Datasets	278
<i>Komei Hiruta (Keio University, Japan), Ryusuke Saito (Keio University, Japan), Taro Hatakeyama (Keio University, Japan), Atsushi Hashimoto (Keio University, Japan), and Satoshi Kurihara (Keio University, Japan)</i>	
Author Index	283