TU Delft

Delft University of Technology

# Preface

Rudinac, Stevan; Hanjalic, Alan; Liem, Cynthia; Worring, Marcel; Jónsson, Björn Þór; Liu, Bei; Yamakata, Yoko

**DOI**
10.1007/978-3-031-53305-1

**Publication date**
2024

**Document Version**
Final published version

**Published in**
Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)

**Citation (APA)**
Rudinac, S., Hanjalic, A., Liem, C., Worring, M., Jónsson, B. Þ., Liu, B., & Yamakata, Y. (2024). Preface. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *14554 LNCS*. https://doi.org/10.1007/978-3-031-53305-1

**Important note**
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

**Copyright**
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

*This work is downloaded from Delft University of Technology.*
*For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.*

*Green Open Access added to TU Delft Institutional Repository*

*'You share, we take care!' - Taverne project*

*https://www.openaccess.nl/en/you-share-we-take-care*

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.

**Lecture Notes in Computer Science** **14565**

Founding Editors

Gerhard Goos
Juris Hartmanis

Editorial Board Members

Elisa Bertino, *Purdue University, West Lafayette, IN, USA*
Wen Gao, *Peking University, Beijing, China*
Bernhard Steffen, *TU Dortmund University, Dortmund, Germany*
Moti Yung, *Columbia University, New York, NY, USA*

The series Lecture Notes in Computer Science (LNCS), including its subseries Lecture Notes in Artificial Intelligence (LNAI) and Lecture Notes in Bioinformatics (LNBI), has established itself as a medium for the publication of new developments in computer science and information technology research, teaching, and education.

LNCS enjoys close cooperation with the computer science R & D community, the series counts many renowned academics among its volume editors and paper authors, and collaborates with prestigious societies. Its mission is to serve this international community by providing an invaluable service, mainly focused on the publication of conference and workshop proceedings and postproceedings. LNCS commenced publication in 1973.

Stevan Rudinac · Alan Hanjalic · Cynthia Liem ·
Marcel Worring · Björn Þór Jónsson · Bei Liu ·
Yoko Yamakata
Editors

# MultiMedia Modeling

30th International Conference, MMM 2024
Amsterdam, The Netherlands, January 29 – February 2, 2024
Proceedings, Part V

Springer

*Editors*
Stevan Rudinac
University of Amsterdam
Amsterdam, The Netherlands

Alan Hanjalic
Delft University of Technology
Delft, The Netherlands

Cynthia Liem
Delft University of Technology
Delft, The Netherlands

Marcel Worring
University of Amsterdam
Amsterdam, The Netherlands

Björn Þór Jónsson
Reykjavik University
Reykjavik, Iceland

Bei Liu
Microsoft Research Lab – Asia
Beijing, China

Yoko Yamakata
The University of Tokyo
Tokyo, Japan

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Computer Science
ISBN 978-3-031-56434-5 ISBN 978-3-031-56435-2 (eBook)
https://doi.org/10.1007/978-3-031-56435-2

© The Editor(s) (if applicable) and The Author(s), under exclusive license
to Springer Nature Switzerland AG 2024

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Paper in this product is recyclable.

# Preface

These five-volume proceedings contain the papers presented at MMM 2024, the International Conference on Multimedia Modeling. This 30th anniversary edition of the conference was held in Amsterdam, The Netherlands, from 29 January to 2 February 2024. The event showcased recent research developments in a broad spectrum of topics related to multimedia modelling, particularly: audio, image, video processing, coding and compression, multimodal analysis for retrieval applications, and multimedia fusion methods.

We received 297 regular, special session, Brave New Ideas, demonstration and Video Browser Showdown paper submissions. Out of 238 submitted regular papers, 27 were selected for oral and 86 for poster presentation through a double-blind review process in which, on average, each paper was judged by at least three program committee members and reviewers. In addition, the conference featured 23 special session papers, 2 Brave New Ideas and 8 demonstrations. The following four special sessions were part of the MMM 2024 program:

- FMM: Special Session on Foundation Models for Multimedia
- MDRE: Special Session on Multimedia Datasets for Repeatable Experimentation
- ICDAR: Special Session on Intelligent Cross-Data Analysis and Retrieval
- XR-MACCI: Special Session on eXtended Reality and Multimedia - Advancing Content Creation and Interaction

The program further included four inspiring keynote talks by Anna Vilanova from the Eindhoven University of Technology, Cees Snoek from the University of Amsterdam, Fleur Zeldenrust from the Radboud University and Ioannis Kompatsiaris from CERTH-ITI.

In addition, the annual MediaEval workshop was organised in conjunction with the conference. The attractive and high-quality program was completed by the Video Browser Showdown, an annual live video retrieval competition, in which 13 teams participated.

We would like to thank the members of the organizing committee, special session and VBS organisers, steering and technical program committee members, reviewers, keynote speakers and authors for making MMM 2024 a success.

December 2023

Stevan Rudinac
Alan Hanjalic
Cynthia Liem
Marcel Worring
Björn Þór Jónsson
Bei Liu
Yoko Yamakata

# Organization

## General Chairs

| | |
|---|---|
| Stevan Rudinac | University of Amsterdam, The Netherlands |
| Alan Hanjalic | Delft University of Technology, The Netherlands |
| Cynthia Liem | Delft University of Technology, The Netherlands |
| Marcel Worring | University of Amsterdam, The Netherlands |

## Technical Program Chairs

| | |
|---|---|
| Björn Þór Jónsson | Reykjavik University, Iceland |
| Bei Liu | Microsoft Research, China |
| Yoko Yamakata | University of Tokyo, Japan |

## Community Direction Chairs

| | |
|---|---|
| Lucia Vadicamo | ISTI-CNR, Italy |
| Ichiro Ide | Nagoya University, Japan |
| Vasileios Mezaris | Information Technologies Institute, Greece |

## Demo Chairs

| | |
|---|---|
| Liting Zhou | Dublin City University, Ireland |
| Binh Nguyen | University of Science, Vietnam National University Ho Chi Minh City, Vietnam |

## Web Chairs

| | |
|---|---|
| Nanne van Noord | University of Amsterdam, The Netherlands |
| Yen-Chia Hsu | University of Amsterdam, The Netherlands |

## Video Browser Showdown Organization Committee

| | |
|---|---|
| Klaus Schoeffmann | Klagenfurt University, Austria |
| Werner Bailer | Joanneum Research, Austria |
| Jakub Lokoc | Charles University in Prague, Czech Republic |
| Cathal Gurrin | Dublin City University, Ireland |
| Luca Rossetto | University of Zurich, Switzerland |

## MediaEval Liaison

| | |
|---|---|
| Martha Larson | Radboud University, The Netherlands |

## MMM Conference Liaison

| | |
|---|---|
| Cathal Gurrin | Dublin City University, Ireland |

## Local Arrangements

| | |
|---|---|
| Emily Gale | University of Amsterdam, The Netherlands |

## Steering Committee

| | |
|---|---|
| Phoebe Chen | La Trobe University, Australia |
| Tat-Seng Chua | National University of Singapore, Singapore |
| Kiyoharu Aizawa | University of Tokyo, Japan |
| Cathal Gurrin | Dublin City University, Ireland |
| Benoit Huet | Eurecom, France |
| Klaus Schoeffmann | Klagenfurt University, Austria |
| Richang Hong | Hefei University of Technology, China |
| Björn Þór Jónsson | Reykjavik University, Iceland |
| Guo-Jun Qi | University of Central Florida, USA |
| Wen-Huang Cheng | National Chiao Tung University, Taiwan |
| Peng Cui | Tsinghua University, China |
| Duc-Tien Dang-Nguyen | University of Bergen, Norway |

## Special Session Organizers

### FMM: Special Session on Foundation Models for Multimedia

| | |
|---|---|
| Xirong Li | Renmin University of China, China |
| Zhineng Chen | Fudan University, China |
| Xing Xu | University of Electronic Science and Technology of China, China |
| Symeon (Akis) Papadopoulos | Centre for Research and Technology Hellas, Greece |
| Jing Liu | Chinese Academy of Sciences, China |

### MDRE: Special Session on Multimedia Datasets for Repeatable Experimentation

| | |
|---|---|
| Klaus Schöffmann | Klagenfurt University, Austria |
| Björn Þór Jónsson | Reykjavik University, Iceland |
| Cathal Gurrin | Dublin City University, Ireland |
| Duc-Tien Dang-Nguyen | University of Bergen, Norway |
| Liting Zhou | Dublin City University, Ireland |

### ICDAR: Special Session on Intelligent Cross-Data Analysis and Retrieval

| | |
|---|---|
| Minh-Son Dao | National Institute of Information and Communications Technology, Japan |
| Michael Alexander Riegler | Simula Metropolitan Center for Digital Engineering, Norway |
| Duc-Tien Dang-Nguyen | University of Bergen, Norway |
| Binh Nguyen | University of Science, Vietnam National University Ho Chi Minh City, Vietnam |

### XR-MACCI: Special Session on eXtended Reality and Multimedia - Advancing Content Creation and Interaction

| | |
|---|---|
| Claudio Gennaro | Information Science and Technologies Institute, National Research Council, Italy |
| Sotiris Diplaris | Information Technologies Institute, Centre for Research and Technology Hellas, Greece |
| Stefanos Vrochidis | Information Technologies Institute, Centre for Research and Technology Hellas, Greece |
| Heiko Schuldt | University of Basel, Switzerland |
| Werner Bailer | Joanneum Research, Austria |

## Program Committee

| | |
|---|---|
| Alan Smeaton | Dublin City University, Ireland |
| Anh-Khoa Tran | National Institute of Information and Communications Technology, Japan |
| Chih-Wei Lin | Fujian Agriculture and Forestry University, China |
| Chutisant Kerdvibulvech | National Institute of Development Administration, Thailand |
| Cong-Thang Truong | Aizu University, Japan |
| Fan Zhang | Macau University of Science and Technology/Communication University of Zhejiang, China |
| Hilmil Pradana | Sepuluh Nopember Institute of Technology, Indonesia |
| Huy Quang Ung | KDDI Research, Inc., Japan |
| Jakub Lokoc | Charles University, Czech Republic |
| Jiyi Li | University of Yamanashi, Japan |
| Koichi Shinoda | Tokyo Institute of Technology, Japan |
| Konstantinos Ioannidis | Centre for Research & Technology Hellas/Information Technologies Institute, Greece |
| Kyoung-Sook Kim | National Institute of Advanced Industrial Science and Technology, Japan |
| Ladislav Peska | Charles University, Czech Republic |
| Li Yu | Huazhong University of Science and Technology, China |
| Linlin Shen | Shenzhen University, China |
| Luca Rossetto | University of Zurich, Switzerland |
| Maarten Michiel Sukel | University of Amsterdam, The Netherlands |
| Martin Winter | Joanneum Research, Austria |
| Naoko Nitta | Mukogawa Women's University, Japan |
| Naye Ji | Communication University of Zhejiang, China |
| Nhat-Minh Pham-Quang | Aimesoft JSC, Vietnam |
| Pierre-Etienne Martin | Max Planck Institute for Evolutionary Anthropology, Germany |
| Shaodong Li | Guangxi University, China |
| Sheng Li | National Institute of Information and Communications Technology, Japan |
| Stefanie Onsori-Wechtitsch | Joanneum Research, Austria |
| Takayuki Nakatsuka | National Institute of Advanced Industrial Science and Technology, Japan |
| Tao Peng | UT Southwestern Medical Center, USA |

| | |
|---|---|
| Thitirat Siriborvornratanakul | National Institute of Development Administration, Thailand |
| Vajira Thambawita | SimulaMet, Norway |
| Wei-Ta Chu | National Cheng Kung University, Taiwan |
| Wenbin Gan | National Institute of Information and Communications Technology, Japan |
| Xiangling Ding | Hunan University of Science and Technology, China |
| Xiao Luo | University of California, Los Angeles, USA |
| Xiaoshan Yang | Institute of Automation, Chinese Academy of Sciences, China |
| Xiaozhou Ye | AsiaInfo, China |
| Xu Wang | Shanghai Institute of Microsystem and Information Technology, China |
| Yasutomo Kawanishi | RIKEN, Japan |
| Yijia Zhang | Dalian Maritime University, China |
| Yuan Lin | Kristiania University College, Norway |
| Zhenhua Yu | Ningxia University, China |
| Weifeng Liu | China University of Petroleum, China |

## Additional Reviewers

Alberto Valese
Alexander Shvets
Ali Abdari
Bei Liu
Ben Liang
Benno Weck
Bo Wang
Bowen Wang
Carlo Bretti
Carlos Cancino-Chacón
Chen-Hsiu Huang
Chengjie Bai
Chenlin Zhao
Chenyang Lyu
Chi-Yu Chen
Chinmaya Laxmikant Kaundanya
Christos Koutlis
Chunyin Sheng
Dennis Hoppe
Dexu Yao
Die Yu
Dimitris Karageorgiou
Dong Zhang
Duy Dong Le
Evlampios Apostolidis
Fahong Wang
Fang Yang
Fanran Sun
Fazhi He
Feng Chen
Fengfa Li
Florian Spiess
Fuyang Yu
Gang Yang
Gopi Krishna Erabati
Graham Healy
Guangjie Yang
Guangrui Liu
Guangyu Gao
Guanming Liu
Guohua Lv
Guowei Wang

Gylfi Þór Guðmundsson
Hai Yang Zhang
Hannes Fassold
Hao Li
Hao-Yuan Ma
Haochen He
Haotian Wu
Haoyang Ma
Haozheng Zhang
Herng-Hua Chang
Honglei Zhang
Honglei Zheng
Hu Lu
Hua Chen
Hua Li Du
Huang Lipeng
Huanyu Mei
Huishan Yang
Ilias Koulalis
Ioannis Paraskevopoulos
Ioannis Sarridis
Javier Huertas-Tato
Jiacheng Zhang
Jiahuan Wang
Jianbo Xiong
Jiancheng Huang
Jiang Deng
Jiaqi Qiu
Jiashuang Zhou
Jiaxin Bai
Jiaxin Li
Jiayu Bao
Jie Lei
Jing Zhang
Jingjing Xie
Jixuan Hong
Jun Li
Jun Sang
Jun Wu
Jun-Cheng Chen
Juntao Huang
Junzhou Chen
Kai Wang
Kai-Uwe Barthel
Kang Yi
Kangkang Feng
Katashi Nagao
Kedi Qiu
Kha-Luan Pham
Khawla Ben Salah
Konstantin Schall
Konstantinos Apostolidis
Konstantinos Triaridis
Kun Zhang
Lantao Wang
Lei Wang
Li Yan
Liang Zhu
Ling Shengrong
Ling Xiao
Linyi Qian
Linzi Xing
Liting Zhou
Liu Junpeng
Liyun Xu
Loris Sauter
Lu Zhang
Luca Ciampi
Luca Rossetto
Luotao Zhang
Ly-Duyen Tran
Mario Taschwer
Marta Micheli
Masatoshi Hamanaka
Meiling Ning
Meng Jie Zhang
Meng Lin
Mengying Xu
Minh-Van Nguyen
Muyuan Liu
Naomi Ubina
Naushad Alam
Nicola Messina
Nima Yazdani
Omar Shahbaz Khan
Panagiotis Kasnesis
Pantid Chantangphol
Peide Zhu
Pingping Cai
Qian Cao

Qian Qiao
Qiang Chen
Qiulin Li
Qiuxian Li
Quoc-Huy Trinh
Rahel Arnold
Ralph Gasser
Ricardo Rios M. Do Carmo
Rim Afdhal
Ruichen Li
Ruilin Yao
Sahar Nasirihaghighi
Sanyi Zhang
Shahram Ghandeharizadeh
Shan Cao
Shaomin Xie
Shengbin Meng
Shengjia Zhang
Shihichi Ka
Shilong Yu
Shize Wang
Shuai Wang
Shuaiwei Wang
Shukai Liu
Shuo Wang
Shuxiang Song
Sizheng Guo
Song-Lu Chen
Songkang Dai
Songwei Pei
Stefanos Iordanis Papadopoulos
Stuart James
Su Chang Quan
Sze An Peter Tan
Takafumi Nakanishi
Tanya Koohpayeh Araghi
Tao Zhang
Theodor Clemens Wulff
Thu Nguyen
Tianxiang Zhao
Tianyou Chang
Tiaobo Ji
Ting Liu
Ting Peng
Tongwei Ma
Trung-Nghia Le
Ujjwal Sharma
Van-Tien Nguyen
Van-Tu Ninh
Vasilis Sitokonstantinou
Viet-Tham Huynh
Wang Sicheng
Wang Zhou
Wei Liu
Weilong Zhang
Wenjie Deng
Wenjie Wu
Wenjie Xing
Wenjun Gan
Wenlong Lu
Wenzhu Yang
Xi Xiao
Xiang Li
Xiangzheng Li
Xiaochen Yuan
Xiaohai Zhang
Xiaohui Liang
Xiaoming Mao
Xiaopei Hu
Xiaopeng Hu
Xiaoting Li
Xiaotong Bu
Xin Chen
Xin Dong
Xin Zhi
Xinyu Li
Xiran Zhang
Xitie Zhang
Xu Chen
Xuan-Nam Cao
Xueyang Qin
Xutong Cui
Xuyang Luo
Yan Gao
Yan Ke
Yanyan Jiao
Yao Zhang
Yaoqin Luo
Yehong Pan
Yi Jiang

Yi Rong
Yi Zhang
Yihang Zhou
Yinqing Cheng
Yinzhou Zhang
Yiru Zhang
Yizhi Luo
Yonghao Wan
Yongkang Ding
Yongliang Xu
Yosuke Tsuchiya
Youkai Wang
Yu Boan
Yuan Zhou
Yuanjian He
Yuanyuan Liu
Yuanyuan Xu
Yufeng Chen
Yuhang Yang
Yulong Wang
Yunzhou Jiang
Yuqi Li
Yuxuan Zhang
Zebin Li
Zhangziyi Zhou
Zhanjie Jin
Zhao Liu
Zhe Kong
Zhen Wang
Zheng Zhong
Zhengye Shen
Zhenlei Cui
Zhibin Zhang
Zhongjie Hu
Zhongliang Wang
Zijian Lin
Zimi Lv
Zituo Li
Zixuan Hong

# Contents – Part V

RESET: Relational Similarity Extension for V3C1 Video Dataset . . . . . 1
*Patrik Veselý and Ladislav Peška*

A New Benchmark and OCR-Free Method for Document Image Topic Classification . . . . . 15
*Zhen Wang, Peide Zhu, Fuyang Yu, and Manabu Okumura*

The Rach3 Dataset: Towards Data-Driven Analysis of Piano Performance Rehearsal . . . . . 28
*Carlos Eduardo Cancino-Chacón and Ivan Pilkov*

WikiMuTe: A Web-Sourced Dataset of Semantic Descriptions for Music Audio . . . . . 42
*Benno Weck, Holger Kirchhoff, Peter Grosche, and Xavier Serra*

PDTW150K: A Dataset for Patent Drawing Retrieval . . . . . 57
*Chan-Ming Hsu, Tse-Hung Lin, Yu-Hsien Chen, and Chih-Yi Chiu*

Interactive Question Answering for Multimodal Lifelog Retrieval . . . . . 68
*Ly-Duyen Tran, Liting Zhou, Binh Nguyen, and Cathal Gurrin*

Event Recognition in Laparoscopic Gynecology Videos with Hybrid Transformers . . . . . 82
*Sahar Nasirihaghighi, Negin Ghamsarian, Heinrich Husslein, and Klaus Schoeffmann*

GreenScreen: A Multimodal Dataset for Detecting Corporate Greenwashing in the Wild . . . . . 96
*Ujjwal Sharma, Stevan Rudinac, Joris Demmers, Willemijn van Dolen, and Marcel Worring*

**Author Index** . . . . . 111