

# **19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2024)**

Volume 4: VISAPP

Rome, Italy  
27-29 February 2024

## **Editors:**

**Petia Radeva  
Antonino Furnari  
Kadi Bouatouch  
A. Augusto Sousa**

ISBN: 978-1-7138-9758-3

**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© (2024) by SCITEPRESS – Science and Technology Publications, Lda.  
All rights reserved.

Printed with permission by Curran Associates, Inc. (2025)

For permission requests, please contact SCITEPRESS – Science and Technology Publications, Lda.  
at the address below.

SCITEPRESS – Science and Technology Publications, Lda.  
Avenida de S. Francisco Xavier, Lote 7 Cv. C,  
2900-616 Setúbal, Portugal

Phone: +351 265 520 185

Fax: +351 265520 186

[info@scitepress.org](mailto:info@scitepress.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](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

# CONTENTS

---

## INVITED SPEAKERS

### KEYNOTE SPEAKERS

- Towards Detailed Understanding of the Visual World 5  
*Fahad Khan*
- Virtual Reality in Mental Health: A Self-Counselling Approach 7  
*Mel Slater*
- Multi-Modal Human-Machine Interaction: Joint Optimization of Single Modalities and Automatic Learning of Communication Channel Fusion 9  
*Gerhard Rigoll*
- The Dance of Logic and Unpredictability: Examining the Predictability of User Behavior on Visual Analytics Tasks 11  
*Alvitta Ottley*

## MOTION, TRACKING AND STEREO VISION

### FULL PAPERS

- BEVSeg2TP: Surround View Camera Bird's-Eye-View Based Joint Vehicle Segmentation and Ego Vehicle Trajectory Prediction 25  
*Sushil Sharma, Arindam Das, Ganesh Sistu, Mark Halton and Ciarán Eising*
- Are Semi-Dense Detector-Free Methods Good at Matching Local Features? 35  
*Matthieu Vilain, Rémi Giraud, Hugo Germain and Guillaume Bourmaud*
- ImgAdaPoinTr: Improving Point Cloud Completion via Images and Segmentation 47  
*Vaagn Chopuryan, Mikhail Kuznetsov, Vasilii Latonov and Natalia Semenova*
- Evaluation of 3D Point Cloud Distances: A Comparative Study in Multi-Point Cloud Fusion Environments 59  
*Ulugbek Alibekov, Vanessa Staderini, Geetha Ramachandran, Philipp Schneider and Doris Antensteiner*
- Toward Multi-Plane Image Reconstruction from a Casually Captured Focal Stack 72  
*Shiori Ueda, Hideo Saito and Shohei Mori*

## SHORT PAPERS

Anisotropic Diffusion for Depth Estimation in Shape from Focus Systems <i>Bilal Ahmad, Ivar Farup and Pål Anders Floor</i>	85
Attacking the Loop: Adversarial Attacks on Graph-Based Loop Closure Detection <i>Jonathan J. Y. Kim, Martin Urschler, Patricia J. Riddle and Jörg S. Wicker</i>	90
PanoTherm: Panoramic Thermal Imaging for Object Detection and Tracking <i>Thomas Kernbauer, Philipp Fleck and Clemens Arth</i>	98
BASE: Probably a Better Approach to Visual Multi-Object Tracking <i>Martin Vonheim Larsen, Sigmund Rolfsjord, Daniel Gusland, Jörgen Ahlberg and Kim Mathiassen</i>	110
Learning 3D Human UV with Loose Clothing from Monocular Video <i>Meng-Yu Jennifer Kuo, Jingfan Guo and Ryo Kawahara</i>	122
A Multilevel Strategy to Improve People Tracking in a Real-World Scenario <i>Cristiano B. de Oliveira, Joao C. Neves, Rafael O. Ribeiro and David Menotti</i>	130
A Tool for 3D Representation of the 2D Thermographic Breast Acquisitions <i>Eudoxia L. S. Moura, Gleidson M. Costa, Tiago B. Borchardt and Aura Conci</i>	138
Self-Mounted Motion Capture System Using Mutual Projection of Asynchronous Cameras <i>Kazusa Ozaki, Fumihiko Sakaue and Jun Sato</i>	147

## MOBILE AND EGOCENTRIC VISION FOR HUMANS AND ROBOTS

### FULL PAPERS

Event-Based Semantic-Aided Motion Segmentation <i>Chenao Jiang, Julien Moreau and Franck Davoine</i>	159
Semantic State Estimation in Robot Cloth Manipulations Using Domain Adaptation from Human Demonstrations <i>Georgies Tzelepis, Eren Erdal Aksoy, Júlia Borràs and Guillem Alenyà</i>	172
Hand Mesh and Object Pose Reconstruction Using Cross Model Autoencoder <i>Chaitanya Bandi and Ulrike Thomas</i>	183
Multi-View Inversion for 3D-aware Generative Adversarial Networks <i>Florian Barthel, Anna Hilsmann and Peter Eisert</i>	194
HD-VoxelFlex: Flexible High-Definition Voxel Grid Representation <i>Igor Vozniak, Pavel Astreika, Philipp Müller, Nils Lipp, Christian Müller and Philipp Slusallek</i>	204
BEVFastLine: Single Shot Fast BEV Line Detection for Automated Parking Applications <i>Praveen Narasappareddyari, Venkatesh Moorthi Karunamoorthy, Shubham Sonarghare, Ganesh Sistu and Prasad Deshpande</i>	220
Fooling Neural Networks for Motion Forecasting via Adversarial Attacks <i>Edgar Medina and Leyong Loh</i>	232

## SHORT PAPERS

Finding and Navigating to Humans in Complex Environments for Assistive Tasks <i>Asfand Yaar, Antonino Furnari, Marco Rosano, Aki Härmä and Giovanni Maria Farinella</i>	245
Informative Rays Selection for Few-Shot Neural Radiance Fields <i>Marco Orsingher, Anthony Dell'Eva, Paolo Zani, Paolo Medici and Massimo Bertozzi</i>	253
Augmenting Human-Robot Collaboration Task by Human Hand Position Forecasting <i>Shyngyskhan Abilkassov, Michael Gentner and Mirela Popa</i>	262
Enhancing Object Detection Accuracy with Variational Autoencoders as a Filter in YOLO <i>Shubham Kumar Dubey, J. V. Satyanarayana and C. Krishna Mohan</i>	270
Analysis of Point Cloud Domain Gap Effects for 3D Object Detection Evaluation <i>Aitor Iglesias, Mikel García, Nerea Aranjuelo, Ignacio Arganda-Carreras and Marcos Nieto</i>	278
Incorporating Temporal Information into 3D Hand Pose Estimation Using Scene Flow <i>Niklas Hermes, Alexander Bigalke and Mattias P. Heinrich</i>	286
PIRO: Permutation-Invariant Relational Network for Multi-Person 3D Pose Estimation <i>Nicolas Ugrinovic, Adria Ruiz, Antonio Agudo, Alberto Sanfeliu and Francesc Moreno-Noguer</i>	295
Social Distancing Monitoring by Human Detection Through Bird's-Eye View Technique <i>Gona Rozhbayani, Amel Tuama and Fadwa Al-Azzo</i>	306
Automatic Error Correction of GPT-Based Robot Motion Generation by Partial Affordance of Tool <i>Takahiro Suzuki, Yuta Ando and Manabu Hashimoto</i>	314
World-Map Misalignment Detection for Visual Navigation Systems <i>Rosario Forte, Michele Mazzamuto, Francesco Ragusa, Giovanni Maria Farinella and Antonino Furnari</i>	324
Combining Progressive Hierarchical Image Encoding and YOLO to Detect Fish in Their Natural Habitat <i>Antoni Burguera</i>	333
Region-Transformer: Self-Attention Region Based Class-Agnostic Point Cloud Segmentation <i>Dipesh Gyawali, Jian Zhang and Bijaya B. Karki</i>	341
Multidimensional Compressed Sensing for Spectral Light Field Imaging <i>Wen Cao, Ehsan Miandji and Jonas Unger</i>	349
Visual Perception of Obstacles: Do Humans and Machines Focus on the Same Image Features? <i>Constantinos A. Kyriakides, Marios Thoma, Zenonas Theodosiou, Harris Partaourides, Loizos Michael and Andreas Lanitis</i>	357
Reliability and Stability of Mean Opinion Score for Image Aesthetic Quality Assessment Obtained Through Crowdsourcing <i>Egor Ershov, Artyom Panshin, Ivan Ermakov, Nikola Banić, Alex Savchik and Simone Bianco</i>	365
Detecting Anomalous 3D Point Clouds Using Pre-Trained Feature Extractors <i>Dario Mantegazza and Alessandro Giusti</i>	373

Decoding Visual Stimuli and Visual Imagery Information from EEG Signals Utilizing Multi-Perspective 3D-CNN Based Hierarchical Deep-Fusion Learning Network <i>Fatma Yusranur Emanet and Kazim Sekeroglu</i>	381
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----

## APPLICATIONS AND SERVICES

### FULL PAPERS

End-to-End Chess Recognition <i>Athanasios Masouris and Jan C. van Gemert</i>	393
Animating NeRFs from Texture Space: A Framework for Pose-Dependent Rendering of Human Performances <i>Paul Knoll, Wieland Morgenstern, Anna Hilsmann and Peter Eisert</i>	404
Symmetry Completion Test: A Novel Approach for Visual Distortion Mapping and Correction Using Symmetry Constraints <i>Ye Ling, David M. Frohlich, Tom H. Williamson and Jean-Yves Guillemaut</i>	414
Defying Limits: Super-Resolution Refinement with Diffusion Guidance <i>Marcelo Dos Santos, João C. R. Neves, Hugo Proença and David Menotti</i>	426
On Feasibility of Transferring Watermarks from Training Data to GAN-Generated Fingerprint Images <i>Venkata Srinath Mannam, Andrey Makrushin and Jana Dittmann</i>	435
Towards the Detection of Diffusion Model Deepfakes <i>Jonas Ricker, Simon Damm, Thorsten Holz and Asja Fischer</i>	446

### SHORT PAPERS

Enhanced Multimodal Timely Prediction of Pulmonary Fibrosis Progression with Uncertainty Estimation from Chest CT Images and Clinical Metadata <i>Mohamed Dahmane</i>	461
Multi-View 3D Reconstruction for Construction Site Monitoring <i>Guangan Chen, Michiel Vlamincx, Wilfried Philips and Hiep Luong</i>	469
A Study of Real World Information Mapping for Information Sharing Using Edge Devices <i>Takenori Hara and Hideo Saito</i>	477
Real-Time Detection and Mapping of Crowd Panic Emergencies <i>Ilias Lazarou, Anastasios L. Kesidis and Andreas Tsatsaris</i>	485
Attentive-YOLO: On-Site Water Pipeline Inspection Using Efficient Channel Attention and Reduced ELAN-Based YOLOv7 <i>Atanu Shuvam Roy and Priyanka Bagade</i>	492
Deep Learning-Powered Assembly Step Classification for Intricate Machines <i>Luca Rodiga, Eva Eggeling, Ulrich Krispel and Torsten Ullrich</i>	500
Teacher-Student Models for AI Vision at the Edge: A Car Parking Case Study <i>Mbasa Joaquim Molo, Emanuele Carlini, Luca Ciampi, Claudio Gennaro and Lucia Vadicamo</i>	508
Robotics and Computer Vision in the Brazilian Electoral Context: A Case Study <i>Jairton da Silva Falcão Filho, Matheus Hopper Jansen Costa, Jonas Ferreira Silva, Cecília Virgínia Santos da Silva, Felipe Augusto da Silva Mendonca, Jefferson Medeiros Norberto, Riei Joaquim Matos Rodrigues and Marcondes Ricarte da Silva Júnior</i>	516

Direct 3D Body Measurement Estimation from Sparse Landmarks <i>David Bojanić, Kristijan Bartol, Tomislav Petković and Tomislav Pribanić</i>	524
Diffusion-Inspired Dynamic Models for Enhanced Fake Face Detection <i>Mamadou Dian Bah, Rajjeshwar Ganguly and Mohamed Dahmane</i>	532
Physical Ergonomics Anticipation with Human Motion Prediction <i>Mattias Billast, Jonas De Bruyne, Klaas Bombeke, Tom De Schepper and Kevin Mets</i>	539
Coreset Based Medical Image Anomaly Detection and Segmentation <i>Ciprian-Mihai Ceaușescu, Bogdan Alexe and Riccardo Volpi</i>	549
A Mobile-Phone Pose Estimation for Gym-Exercise Form Correction <i>Matthew Turner, Kofi Appiah and Sze Chai Kwok</i>	559
Deep Discriminative Feature Learning for Document Image Manipulation Detection <i>Kais Rouis, Petra Gomez-Krämer, Mickaël Coustaty, Saddock Kébaïri and Vincent Poulain d'Andecy</i>	567
Deep Learning Model Compression for Resource Efficient Activity Recognition on Edge Devices: A Case Study <i>Dieter Balemans, Benjamin Vandersmissen, Jan Steckel, Siegfried Mercelis, Phil Reiter and José Oramas</i>	575
Explainability and Interpretability for Media Forensic Methods: Illustrated on the Example of the Steganalysis Tool Stegdetect <i>Christian Kraetzer and Mario Hildebrandt</i>	585
Recognizing Actions in High-Resolution Low-Framerate Videos: A Feasibility Study in the Construction Sector <i>Benjamin Vandersmissen, Arian Sabaghi, Phil Reiter and Jose Oramas</i>	593
An Assistive Technology Based on Object Detection for Automated Task List Generation <i>Frédéric Rayar</i>	601
Microplankton Discrimination in FlowCAM Images Using Deep Learning <i>Francisco Bonin-Font, Gorka Buenvaron, Mary K. Kane and Idan Tuval</i>	606
Fact-Checked Claim Detection in Videos Using a Multimodal Approach <i>Frédéric Rayar</i>	614
Evaluating the Usability of a 3D Map Visualizer Augmented Reality Application <i>Franklin L. N. Fracaro, Fabiana F. F. Peres, Claudio R. M. Mauricio and João Marcelo X. N. Teixeira</i>	621
Automatic Computation of the Posterior Nipple Line from Mammographies <i>Quynh Nhu Jennifer Tran, Tina Santner and Antonio Rodríguez-Sánchez</i>	629
Automatic Assessment of Skill and Performance in Fencing Footwork <i>Filip Malawski, Marek Krupa and Ksawery Kapela</i>	637
Semantic Image Synthesis for Realistic Image Generation in Robotic Assisted Partial Nephrectomy <i>Stefano Mazzocchetti, Laura Cercenelli, Lorenzo Bianchi, Riccardo Schiavina and Emanuela Marcelli</i>	645
AUTHOR INDEX	653