

PROCEEDINGS OF SPIE

International Workshop on Advanced Imaging Technology (IWAIT) 2026

**Masayuki Nakajima
Chuan-Yu Chang
Chien-Chou Lin
Shogo Tokai
Kwang-Deok Seo
Chia-Hung Yeh
Budianto Tandianus**
Editors

**12–14 January 2026
Kaohsiung, Taiwan**

Organized by
National Yunlin University of Science and Technology (Taiwan)
National Cheng Kung University (Taiwan)

Co-organized by
The Chinese Image Processing and Pattern Recognition Society (Taiwan)
Institute of Electronics, Information and Communication Engineers (IEICE), (Japan)
Institute of Image Information and Television Engineers (ITE), (Japan)
Korean Institute of Broadcast and Media Engineers (KIBME), (Korea)

Published by
SPIE

Volume 14072

Proceedings of SPIE 0277-786X, V. 14072

SPIE is an international society advancing an interdisciplinary approach to the science and application of light.

The papers in this volume were part of the technical conference cited on the cover and title page. Papers were selected and subject to review by the editors and conference program committee. Some conference presentations may not be available for publication. Additional papers and presentation recordings may be available online in the SPIE Digital Library at SPIDigitalLibrary.org.

The papers reflect the work and thoughts of the authors and are published herein as submitted. The publisher is not responsible for the validity of the information or for any outcomes resulting from reliance thereon.

Please use the following format to cite material from these proceedings:

Author(s), "Title of Paper," in *International Workshop on Advanced Imaging Technology (IWAIT) 2026*, edited by Masayuki Nakajima, Chuan-Yu Chang, Chien-Chou Lin, Shogo Tokai, Kwang-Deok Seo, Chia-Hung Yeh, Budianto Tandianus, Proc. of SPIE 14072, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X

ISSN: 1996-756X (electronic)

ISBN: 9798902321033

ISBN: 9798902321040 (electronic)

Published by

SPIE

P.O. Box 10, Bellingham, Washington 98227-0010 USA

Telephone +1 360 676 3290 (Pacific Time)

SPIE.org

Copyright © 2026 Society of Photo-Optical Instrumentation Engineers (SPIE).

Copying of material in this book for internal or personal use, or for the internal or personal use of specific clients, beyond the fair use provisions granted by the U.S. Copyright Law is authorized by SPIE subject to payment of fees. To obtain permission to use and share articles in this volume, visit Copyright Clearance Center at copyright.com. Other copying for republication, resale, advertising or promotion, or any form of systematic or multiple reproduction of any material in this book is prohibited except with permission in writing from the publisher.

Printed in the United States of America by Curran Associates, Inc., under license from SPIE.

Publication of record for individual papers is online in the SPIE Digital Library.

**SPIE. DIGITAL
LIBRARY**

SPIDigitalLibrary.org

Paper Numbering: A unique citation identifier (CID) number is assigned to each article in the Proceedings of SPIE at the time of publication. Utilization of CIDs allows articles to be fully citable as soon as they are published online, and connects the same identifier to all online and print versions of the publication. SPIE uses a seven-digit CID article numbering system structured as follows:

- The first five digits correspond to the SPIE volume number.
- The last two digits indicate publication order within the volume using a Base 36 numbering system employing both numerals and letters. These two-number sets start with 00, 01, 02, 03, 04, 05, 06, 07, 08, 09, 0A, 0B ... 0Z, followed by 10-1Z, 20-2Z, etc. The CID Number appears on each page of the manuscript.

Contents

xv Conference Committee

3D MODEL & POINT CLOUD

- 14072 02 **Difference visualization of 3D point sets from trench investigations** [14072-9]
- 14072 03 **A study on boundary segmentation and assembly of fractured bone fragments using measured point clouds** [14072-10]
- 14072 04 **A study on 3D model construction of Morioka Castle stone walls using photogrammetry to estimate camera parameters for old photograph of Morioka Castle** [14072-44]
- 14072 05 **Zero-shot text-to-3D-object generation with visual-geometry-grounded transformer** [14072-56]
- 14072 06 **Performance analysis of 3D semantic segmentation using 3D Gaussian splatting** [14072-58]
- 14072 07 **Curriculum learning for 3D vehicle geometry editing** [14072-147]
- 14072 08 **Quaternion-based deformation of cloth point clouds and its verification** [14072-54]

AI FOR MEDICAL IMAGE I

- 14072 09 **Estimating the difficulty of laparoscopic surgery for renal tumor using perinephric fat radiomics features** [14072-46]
- 14072 0A **Multi-scale vessel and lesion feature fusion for early diabetic retinopathy detection using fundus images** [14072-196]
- 14072 0B **Proposal for cuticle damage classification using deep learning in hair care** [14072-92]
- 14072 0C **Class-controlled histopathology image synthesis for robust tumor-infiltrating lymphocyte segmentation** [14072-91]

AI FOR COMPUTER VISION I

- 14072 0D **Compressed-image recognition for a feature extractable CMOS image sensor** [14072-60]
- 14072 0E **Enhancing story visualization via subject-aware inpainting with SOEDiff** [14072-140]

- 14072 OF **Phase component compression for computer-generated holograms using modified JPEG-LS prediction** [14072-181]
- 14072 OG **Fast implementation of SSIMULACRA2 for image quality assessment** [14072-25]

AI FOR MEDICAL APPLICATION AND OTHERS

- 14072 OH **Deep-learning-based diagnosis for nasal morphological abnormalities** [14072-38]
- 14072 OI **Joint modeling of perspective, lens distortion, and rolling shutter for accurate monocular 3D pose/velocity estimation** [14072-55]
- 14072 OJ **Estimation of quality of life using sublingual vein images based on deep learning and traditional Chinese medicine** [14072-19]
- 14072 OK **Enhancing image-text alignment for medical report generation via low-rank random-mask contrastive learning** [14072-30]

AUTONOMOUS MOBILITY

- 14072 OL **Interactive system for real-time detection and GIS registration of damaged houses using mobile devices to streamline damage assessment** [14072-134]
- 14072 OM **SkyPose: real-time camera pose estimation via skyline matching in mountainous terrain** [14072-36]
- 14072 ON **Signal-duration-based drone classification using YOLO detection on time-frequency spectrograms** [14072-127]
- 14072 OO **Reinforcement-learning-based indoor drone for path finding and collision avoidance** [14072-151]
- 14072 OP **A YOLOv7-based system for traffic object and road sign recognition** [14072-163]

LLM APPLICATION I

- 14072 OQ **Performance analysis of face generation using an integrated GAN and deep reinforcement learning framework** [14072-164]
- 14072 OR **A study of video compression and its impact on video LLMs in open-ended VideoQA** [14072-69]

- 14072 0S **Proposal for a real-world cooking robot based on the integration of LLM and visual information** [14072-106]
- 14072 0T **Digital strategies and implementation for small and medium-sized enterprises in Japan** [14072-184]

AI FOR MEDICAL IMAGE II

- 14072 0U **Risk stratification of asymptomatic carotid plaque: a comparison of symptom onset and treatment eligibility prediction models** [14072-169]
- 14072 0V **Impact of image processing pipeline on video-based heart rate estimation** [14072-119]
- 14072 0W **Similar video retrieval for laparoscopic cholecystectomy videos based on image and text feature representations with temporal context integration** [14072-72]

IMAGE PROCESSING AND PATTERN RECOGNITION

- 14072 0X **Effects of differences in tasks on the degree of color constancy** [14072-78]
- 14072 0Y **Optimization-based reflection removal from glass surfaces using focal length variation** [14072-180]
- 14072 0Z **Semantic and stability-aware grasp pose estimation for unseen objects using deformable shape models** [14072-158]
- 14072 10 **Transform selection for joint coding of chroma residuals** [14072-142]
- 14072 11 **Spatially adaptive and flexible deep network for blind compressed image reconstruction** [14072-182]

3D IMAGE AND VR

- 14072 12 **Safe route estimation from bicycle-mounted camera in pedestrian-active environments** [14072-81]
- 14072 13 **Pruning-guided point cloud enhancement via 3D Gaussian splatting** [14072-101]
- 14072 14 **Analysis of object detection performance according to implicit and explicit representations in novel view synthesis** [14072-59]
- 14072 15 **VNL: a scalable and immersive 3D sandbox virtual lab for computer network education** [14072-176]

LLM APPLICATION II

- 14072 16 **Development of a plant modeling interface to enhance creativity for generating design ideas** [14072-186]
- 14072 17 **De-identification in game chargeback fraud detection** [14072-175]
- 14072 18 **Structure-aware knowledge distillation using centered kernel alignment** [14072-167]
- 14072 19 **Intelligent federated deep learning for credit-card fraud: client-side rebalancing and update-norm analytics** [14072-177]
- 14072 1A **Selecting collaborative learning paradigms for edge IoT: an E-M-O-based comparative simulation with design guidelines** [14072-191]
- 14072 1B **BCF: a blockchain-enabled function for enhancing data consistency and trust in 5GC** [14072-190]

3D DISPLAY

- 14072 1C **Composition transformation of a single landscape image using segmentation and 3D reconstruction** [14072-121]
- 14072 1D **A handheld transparent 3D display using a light-guide and holographic optical element** [14072-84]
- 14072 1E **Generating textured 3D dairy cow models from a few images** [14072-149]
- 14072 1F **HOE generating virtual images viewable from specific viewing zones for VAC-free transparent 3D display** [14072-35]
- 14072 1G **Large-scale time-division 360-degree light field display** [14072-178]
- 14072 1H **Attribute optimization method for geometry compression of 3D Gaussian splat data** [14072-194]
- 14072 1I **Light field representation with constrained viewpoint Gaussian splatting** [14072-144]
- 14072 1J **Toward multi-view texture consistency in novel view synthesis from sparse-view 3DGS** [14072-139]

SIGNAL PROCESSING AND PATTERN RECOGNITION

- 14072 1K **Pedestrians' gaze object estimation in traffic scene considering relationships among multiple pedestrians** [14072-87]
- 14072 1L **Automatic lighting control based on person localization using Pseudo-CAM on edge devices** [14072-145]
- 14072 1M **Sketch-based optimization of residential layouts with fine-grained urban heat island prediction** [14072-123]
- 14072 1N **Image-to-image scene matching by representing semantic structures using scene graphs** [14072-34]
- 14072 1O **Enhancement of TIMD in ECM with skipped non-angular fusion** [14072-131]
- 14072 1P **AI-assisted mmWave beamforming for QoS enhancement in networked multimedia systems** [14072-33]

MODELS OF MACHINE LEARNING

- 14072 1Q **Is YES/NO question answering by vision-language models effective for video retrieval?** [14072-23]
- 14072 1R **Logical anomaly detection based on relative similarity analysis of region segments** [14072-98]

POSTURE RECOGNITION

- 14072 1S **3D human pose estimation from pseudo-two-viewpoints videos** [14072-37]
- 14072 1T **Analysis of bowling posture using center-of-gravity-based motion features** [14072-67]
- 14072 1U **Recognition of sumo attack styles based on the features between sumo wrestlers** [14072-100]
- 14072 1V **Development of a pedestrian floor-projection method considering walking position and direction** [14072-133]
- 14072 1W **Unsupervised estimation of tennis players' fatigue level assuming temporal monotonicity in their fatigue accumulation** [14072-99]
- 14072 1X **Spatio-temporal evaluation of white cane maneuvering based on a single video** [14072-113]
- 14072 1Y **Controllable motion generation for diverse skeletal structures using phase manifolds** [14072-128]

AR AND VR I

- 14072 1Z **A study of method for rapid foreground extraction of natural objects for the construction of AR system for the Morioka Castle site** [14072-8]
- 14072 20 **Development of a guitar practice support system based on chord estimation** [14072-61]
- 14072 21 **A VR-based evacuation experience system considering crowd psychology during disasters** [14072-188]
- 14072 22 **Visual perception support of sound movement using peripheral displays in VR space** [14072-42]
- 14072 23 **GomuGomu Arm: kinetic-energy-based extensible virtual arm** [14072-136]

AR AND VR II

- 14072 24 **Interactive maze generation from vector-based outlines with scalar fields** [14072-150]
- 14072 25 **MagicLiveRoom: a livestream system for VTubers to seamlessly interact with real-world objects** [14072-85]
- 14072 26 **Negative–positive inversion effects for puzzle games using cellophane-attached polarizer sheets** [14072-148]
- 14072 27 **Stereoscopic imaging beyond the boundaries of displays and projectors through perceptual completion based on transparency illusion** [14072-74]
- 14072 28 **The influence of jerk asymmetry in asymmetric vibration on the perceived intensity of virtual force sensation** [14072-111]
- 14072 29 **Emerflux: proposal for a water surface display using a two-layer liquid** [14072-130]

AI FOR AOI

- 14072 2A **A high-fidelity industrial image defect generation method based on DDPM and real-world defect simulation** [14072-66]
- 14072 2B **Transformer-based placement and harmonized composition for synthetic data generation in PCB component detection** [14072-6]
- 14072 2C **Design and performance evaluation of a USRP-based VDES transceiver** [14072-122]

- 14072 2D **Enhanced YOLOv7 for flame and smoke detection in industrial environments** [14072-24]
- 14072 2E **Fractional delay modeling in multipath fading channels with time-varying propagation delays** [14072-157]

AR AND VR III

- 14072 2F **Annular arrangement of multi-view aerial images using a conical frustum half-mirror in an AIRR system** [14072-109]
- 14072 2G **A training system for Jugote routine in Japanese sword fighting choreography** [14072-102]
- 14072 2H **Manga-style rendering of clapping using animated onomatopoeia with high visibility** [14072-71]
- 14072 2I **On the improvement of temporal stability of mist screen system employing air curtain structure derived from spray nozzle and its effect on holographic image projection** [14072-1]
- 14072 2J **Construction of an asymmetric-information board game field using triangular-prism lenses and interlaced images** [14072-76]
- 14072 2K **Dynamic projection mapping system using event-based camera** [14072-96]
- 14072 2L **A fundamental study on optimal duration of screen shake for enhancing affective feedback in action video games** [14072-110]

IMAGE PROCESSING I

- 14072 2M **A study on image acquisition methods to improve classification accuracy of excavated ceramics** [14072-7]
- 14072 2N **Pencil style transfer of digital illustrations using LIC and depth estimation** [14072-28]
- 14072 2O **Generation of semantic segmentation masks based on conditional discrete flow matching** [14072-152]
- 14072 2P **Region-of-interest image coding with content-aware resizing based on an elastic model** [14072-143]
- 14072 2Q **Discussion on introduction of quantum information and computational science in the image data compression procedure of 360 degrees camera** [14072-49]

AI FOR COMPUTER VISION IMAGE PROCESSING II

- 14072 2R **Verification of fixation for UHD images** [14072-155]
- 14072 2S **Enhancement of intra-prediction fusion in ECM** [14072-132]
- 14072 2T **Video deblurring using transformer-based deep neural networks** [14072-32]
- 14072 2U **Study on disparity estimation for monochrome sensor images using extended feature extraction via tone mapping** [14072-39]
- 14072 2V **Lightweight U-Net-based image reconstruction for inverse halftoning** [14072-22]
- 14072 2W **Image watermarking based on Hessenberg decomposition and q-logarithmic transformation** [14072-5]
- 14072 2X **Extensions of bilateral filter for intra-prediction improvement in ECM** [14072-129]

AI FOR COMPUTER VISION II

- 14072 2Y **A selfie system adapted to background environment using an omnidirectional camera** [14072-70]
- 14072 2Z **Movement-specific vehicle counting robust to occlusion at intersections** [14072-125]
- 14072 30 **Joint markerless estimation of ball spin axis and rotation speed from real pitch videos** [14072-115]
- 14072 31 **Attention-guided frequency and locality learning GAN for limited-angle CT reconstruction** [14072-50]
- 14072 32 **Construction of the Sim2Real object grasping pipeline and effect analysis of images on grasp point estimation** [14072-118]
- 14072 33 **Drift adaptation via training data sampling for object detection in surveillance system** [14072-146]
- 14072 34 **Tiny object detection on low-cost edge AI systems** [14072-20]
- 14072 35 **Comparison of deep-learning-based face recognition characteristics across multiple animal species** [14072-48]

AI FOR COMPUTER VISION III

- 14072 36 **A lightweight channel bias for small object detection in aerial images** [14072-29]
- 14072 37 **Real-time vehicle tracking and traffic metric estimation using YOLOv11 with CUDA optical flow** [14072-183]
- 14072 38 **Bus boarding-to-alighting section estimation via integrated passenger re-identification and tracking** [14072-162]
- 14072 39 **Computer-vision-based approaches for monitoring honeybee behavior using AI image recognition** [14072-17]

POSTER SESSION PART I

- 14072 3A **Color image enhancement method based on convex combination coefficients transformation using CIELAB color space** [14072-43]
- 14072 3B **Visibility improvement of sand-dust image using S-shaped functions and constrained histogram specification** [14072-15]
- 14072 3C **Validation of watermark embedding in diffusion models** [14072-103]
- 14072 3D **Front parallel alignment of circular patterns for improvement of OpenCV camera calibration** [14072-79]
- 14072 3E **Lightweight dynamic-aware SLAM via object detection and depth-guided feature filtering** [14072-47]
- 14072 3F **Low-light image enhancement based on inverse S-shaped function and histogram equalization** [14072-11]

POSTER SESSION PART II

- 14072 3G **Redundant flexible representation of quantum images for efficient reconstruction** [14072-80]
- 14072 3H **On-device stereo matching using NPU acceleration for real-time depth estimation** [14072-138]
- 14072 3I **Altering the visual impression of 3D Gaussian splatting via style transfer** [14072-126]
- 14072 3J **Improvement of a video image generation method for a frailty-prevention walking exercise system** [14072-165]
- 14072 3K **Self-attention in vision transformers using ridge regression and the kernel method** [14072-170]

- 14072 3L **An automatic photo selection and trimming for graduation albums** [14072-31]
- 14072 3M **Autonomous oil palm harvesting robot with tree-trunk-image-based visual servoing and navigation framework** [14072-117]
- 14072 3N **Accuracy improvement of multi-speaker speech recognition by introducing speaker diarization** [14072-166]

POSTER SESSION PART III

- 14072 3O **Shadow removal from document image using normalized convolution smoothing and plane approximation** [14072-12]
- 14072 3P **Temporal window analysis for 1D-CNN in skeleton-based continuous sign language recognition** [14072-51]
- 14072 3Q **Image analysis-based support for skateboard training** [14072-4]
- 14072 3R **Proposal for an acne diagnosis using image processing of smartphone photographs and analysis by generative AI** [14072-18]
- 14072 3S **Improvement of image quality in wavefront-coding-reconstructed images under different apodization conditions using deep learning** [14072-168]
- 14072 3T **SAGE: segmentation-aware 3D object extraction from single images** [14072-82]
- 14072 3U **Measuring emotional changes caused by pet robots from facial images** [14072-88]
- 14072 3V **Distributed compressed video sensing with a pre-learned consensus convolutional dictionary** [14072-174]
- 14072 3W **A classification method of the static and dynamic states of continuous fingerspelling** [14072-137]

POSTER SESSION PART IV

- 14072 3X **An efficient real-time pipeline for flare removal on mobile devices** [14072-3]
- 14072 3Y **M3DDM+: an improved video outpainting by a modified masking strategy** [14072-62]
- 14072 3Z **A performance evaluation of sequential inference for missing pixel restoration method** [14072-112]
- 14072 40 **Distributed compressed video sensing with multiple key frames** [14072-153]

- 14072 41 **Linear-color-transformation-based hue-preserving image enhancement for dichromats**
[14072-13]
- 14072 42 **A manga speech balloon segmentation method using T2I-generated synthetic data**
[14072-57]
- 14072 43 **Study on estimation of light reflection for AR try-on systems for accessories** [14072-65]
- 14072 44 **Analysis of bowling throwing posture using body-part-specific features** [14072-89]
- 14072 45 **Extraction of goal-scoring segments and player positions in a panoramic ice hockey video**
[14072-107]
- 14072 46 **L1-L1 norm-based convolutional sparse coding via Anderson-accelerated Douglas-Rachford splitting** [14072-179]
- 14072 47 **A real-time image generation method for head-mounted displays to correct the discrepancy in perceived object size between reality and virtual reality** [14072-171]