PROCEEDINGS OF SPIE

# *International Workshop on Advanced Imaging Technology (IWAIT) 2021*

**Masayuki Nakajima**
**Jae-Gon Kim**
**Wen-Nung Lie**
**Qian Kemao**
*Editors*

**5–6 January 2021**
**Kagoshima (Virtual), Japan**

*Organized by*
Kagoshima University (Japan)
Korean Institute of Broadcast and Communication Engineers (Korea, Republic of)
IEICE—Institute of Electronics, Information and Communication Engineers (Japan)
Institute of Image Information and Television Engineers (Japan)
Japanese Society of Precision Engineering (JSPE_IAPI) (Japan)
Image Processing and Pattern Recognition Society (Taiwan China)

*Published by*
SPIE

**Volume 11766**

Proceedings of SPIE 0277-786X, V. 11766

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 SPIEDigitalLibrary.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) 2021*, edited by Masayuki Nakajima, Jae-Gon Kim, Wen-Nung Lie, Qian Kemao, Proceedings of SPIE Vol. 11766 (SPIE, Bellingham, WA, 2021) Seven-digit Article CID Number.

ISSN: 0277-786X
ISSN: 1996-756X (electronic)

ISBN: 9781510643642
ISBN: 9781510643659 (electronic)

Published by
**SPIE**
P.O. Box 10, Bellingham, Washington 98227-0010 USA
Telephone +1 360 676 3290 (Pacific Time)· Fax +1 360 647 1445
SPIE.org
Copyright © 2021, Society of Photo-Optical Instrumentation Engineers.

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 copying fees. The Transactional Reporting Service base fee for this volume is $21.00 per article (or portion thereof), which should be paid directly to the Copyright Clearance Center (CCC), 222 Rosewood Drive, Danvers, MA 01923. Payment may also be made electronically through CCC Online 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. The CCC fee code is 0277-786X/21/$21.00.

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.

SPIEDigitalLibrary.org

---

**Paper Numbering:** *Proceedings of SPIE* follow an e-First publication model. A unique citation identifier (CID) number is assigned to each article 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


**ADVANCED IMAGE TECHNOLOGY I**

11766 02 **Comparing pixel predictors for lossless image coding** [11766-59]

11766 03 **Attention-based cross-modality interaction for multispectral pedestrian detection** [11766-22]

11766 04 **Multi-frame super-resolution for versatile video coding** [11766-72]

11766 05 **Probability model adjustment for the CNN-based lossless image coding method** [11766-87]

11766 06 **CAVRoom: scalable online multiplayer and multiplatform sandbox** [11766-14]

11766 07 **DSBN: detail-structure blending network for video super-resolution** [11766-21]

11766 08 **Tool to aid view of PC display for USN observers: magnifier with KUI** [11766-82]

**ADVANCED IMAGE TECHNOLOGY II**

11766 09 **CycleGAN without checkerboard artifacts for counter-forensics of fake-image detection** [11766-83]

11766 0A **Depth estimation from 4D light field videos** [11766-105]

11766 0B **A hardware-oriented algorithm of GMM-MRCoHOG for high-performance human detection by an FPGA** [11766-117]

11766 0C **Multi-channel image deblurring using coded flashes** [11766-86]

**ADVANCED IMAGE TECHNOLOGY III**

11766 0D **Ball flight path determination in VR table tennis** [11766-3]

11766 0E **Automatic generation of background computer graphics by two types of deep learning according to user's preference** [11766-33]

11766 0F **A contour drawing method for 3D animal models' fur** [11766-125]

11766 0G **Sound field visualization for primary reflection using equivalent sources and image source methods** [11766-113]

11766 0H **A support system for tracheal suction skill acquisition based on augmented reality technology** [11766-66]

## ADVANCED IMAGE TECHNOLOGY IV

11766 0I **Interest estimation method based on 2D pose features on security camera** [11766-28]

11766 0J **Learning to synthesize dense focal stack and all-in-focus image from sparse focal stack** [11766-56]

11766 0K **Hidden surface removal method using object point based ray tracing in CGH** [11766-55]

11766 0L **Improvement of texture reconstruction based on underlying pattern modification** [11766-107]

11766 0M **2D filter design for coding artifacts reduction using structural similarity as a metric** [11766-53]

11766 0N **An extraction attack on image recognition model using VAE-kdtree model** [11766-62]

## ADVANCED IMAGE TECHNOLOGY V

11766 0O **A study of fabric estimation using CNN from clothes images** [11766-80]

11766 0P **A note on detection of sports action based on temporal cycle consistency learning** [11766-90]

11766 0Q **Data compression for DNN weighting coefficients using layer adaptive quantization** [11766-23]

11766 0R **3D deformable convolution for action classification in videos** [11766-132]

11766 0S **Controlling agents using recurrent neural network in crowd simulation** [11766-118]

11766 0T **Deep neural network for handcrafted cost-based multi-view stereo** [11766-101]

## VIDEO CODING

11766 0U **A watermarked image search method using compressed representation by a convolutional autoencoder** [11766-5]

11766 0V **Two-layer image coding compatible with JPEG XS** [11766-25]

11766 0W **EMOS-based spatio-temporal rate control technique for 2K240P videos** [11766-135]

11766 0X **Most probable direct mode for chroma intra mode coding** [11766-85]

**IMAGE PROCESSING I**

11766 0Y **Scale adaptive side window bilateral filter** [11766-12]

11766 0Z **Ensemble of illumination estimation methods using support vector regression** [11766-77]

11766 10 **Person re-identification by shifting-subclip** [11766-19]

11766 11 **Raster-based mathematical morphology for cutting and packing problems** [11766-94]

11766 12 **The synthesis of weathering using multi-label map** [11766-68]

**IMAGE PROCESSING II**

11766 13 **Wettability method based on surface free energy between solid and liquid** [11766-31]

11766 14 **Illumination robust template matching based on nearest neighbor** [11766-43]

11766 15 **Super-resolution on remote sensing images** [11766-13]

11766 16 **Interior coordination image retrieval with object-detection-based and color features** [11766-27]

11766 17 **Wafer defect pattern classification robust for rotation** [11766-98]

11766 18 **A contactless training support system of holding chopsticks by posture estimation of chopsticks, hand, and fingers** [11766-126]

**IMAGE PROCESSING III**

11766 19 **Formulation of optical properties of cloth for CG curtain simulation** [11766-89]

11766 1A **Object tracking for tangible projection mapping by using multiple contours** [11766-120]

11766 1B **Sketch-based normal map generation with geometric sampling** [11766-36]

11766 1C **Affine subspace clustering with nearest subspace neighbor** [11766-38]

11766 1D **Single image shadow removal via detection-free spatially adaptive denormalization** [11766-34]

**MULTIMEDIA SYSTEMS AND APPLICATIONS I**

11766 1E **Interaction between high resolution and high dynamic range in psychological effects of ultra-high-definition television** [11766-6]

11766 1F **A real-time feedback system of strength training with motion capture and head mounted display** [11766-127]

11766 1G **Climbing pathfinding with the holds and a decision method of the difficulty level of the holds** [11766-130]

11766 1H **Localizing fish in highly turbid underwater images** [11766-93]

11766 1I **Composition presentation based on compositional features for photography assistance** [11766-35]

11766 1J **Immersive audio system based on 2.5D local sound field synthesis using high-speed 1-bit signal** [11766-115]

**MULTIMEDIA SYSTEMS AND APPLICATIONS II**

11766 1K **Shot detection using skeleton position in badminton videos** [11766-17]

11766 1L **Influence of audiovisual feedback on player behavior and performance in response to video game failure** [11766-84]

11766 1M **A study of a rally scene estimation method in a table tennis video** [11766-78]

11766 1N **Abrupt and gradual change analyses of moving objects in stereoscopic viewing** [11766-99]

**THREE-DIMENSIONAL**

11766 1O **Projection mapping using a rotating volumetric 3D display** [11766-122]

11766 1P **Study on three-dimensional extraction and classification of neonatal murine cardiomyocytes** [11766-54]

11766 1Q **Estimation of 3D Information of multiple objects in a monocular Image** [11766-110]

11766 1R **KUI-based modeling : a study for real world oriented modeling using real shadow as an interface** [11766-104]

11766 1S **Video denoising by BM3D technique with an improved cube construction and SURE shrinkage techniques** [11766-134]

**VIRTUAL REALITY I**

11766 1T **Steering autonomous animals in VR hunting** [11766-2]

11766 1U **A VR-based learning system of tuna dismantling for fisheries high school students** [11766-76]

11766 1V **Smart modeling of bobsleigh track surface VR** [11766-71]

11766 1W **Usability comparison of screen-based and virtual-reality-based virtual museums** [11766-88]

11766 1X **A new verification approach for subjective evaluation of actions in HMD-VR with EEG** [11766-52]

11766 1Y **A comparative study on MSSQ and VR sickness caused by swimming motion as VR space transfer** [11766-67]

**VIRTUAL REALITY II**

11766 1Z **A virtual keyboard with multimodal feedback using physical tabletop** [11766-106]

11766 20 **A mixed reality-based fire evacuation drill system** [11766-108]

11766 21 **Development of a virtual try-on system for eyeglasses considering lens distortion** [11766-42]

11766 22 **A real camera interface enabling to shoot objects in virtual space** [11766-97]

11766 23 **A VR-based learning system for operating gas burner in secondary education** [11766-103]

11766 24 **Interaction with virtual objects through eye-tracking** [11766-92]

**DEEP LEARNING I**

11766 25 **An approach with CNN to unsafe manhole classification for visually impaired people** [11766-81]

11766 26 **Transfer learning for small-scale data classification using CNN filter replacement** [11766-75]

11766 27 **A weight-wise watermarking technique for DNN models and its robustness against overwriting attack** [11766-45]

11766 28 **Detection of main persons in snapshots using deep neural networks** [11766-95]

**DEEP LEARNING II**

11766 29 **No-reference video quality assessment metric using spatiotemporal features through LSTM** [11766-16]

11766 2A **Neural-network-based optimal mode estimation for adaptive affine motion compensation** [11766-109]

11766 2B **DMNAS: Differentiable Multi-modal Neural Architecture Search** [11766-47]

11766 2C **Training digital hologram watermarking deep neural network considering hologram distributions** [11766-79]

**POINT CLOUD IMAGE PROCESSING**

11766 2D **Data transmission based on fractal compression of object light data in computer-generated hologram** [11766-15]

11766 2E **A study on stone tool identification based on depth image and 3D point cloud** [11766-44]

11766 2F **An examination of stone chamber extraction based on 3D-measured point cloud for evaluation of stone chamber similarity** [11766-40]

11766 2G **Light field display using stacked imaginary image display** [11766-116]

11766 2H **Pre-adjustment of positions of point-clouds for the ICP algorithm using IMU and parallel projection** [11766-58]

**IMMERSIVE APPLICATIONS**

11766 2I **A VR-based piano self-training portable system on standalone HMD** [11766-124]

11766 2J **An examination of displaying reassembly procedure by recording disassembly procedure with AR markers** [11766-41]

11766 2K **A wearable augmented reality system with anamorphosis projection** [11766-57]

11766 2L **Evaluation of effectiveness of position/orientation measurement method for indoor AR display guidance** [11766-112]

11766 2M **An AR-based conversion system of classical Chinese text with return mark to Japanese reading** [11766-123]

**IMAGE GENERATION AND DISPLAY**

11766 2N **Interaction support for virtual studio by vibration feedback** [11766-11]

11766 2O **Automation of parts segmentation in 2D character animation production** [11766-20]

11766 2P **Speckleless reconstructed image display using convergent light in CGH** [11766-37]

11766 2Q **A basic study of holographic image reconstruction in colors using blue-violet laser light** [11766-32]

11766 2R **Scratch detection based on image reconstruction using ideal-image-pair dataset of just before/after defect occurrence** [11766-102]

11766 2S **Improvement of GrabCut form Kumamoto Castle Ishigaki region extraction method** [11766-128]

11766 2T **Development of a method for animating water flow in an illustration** [11766-65]

**IMAGE RECOGNITION**

11766 2U **Facial feature exaggeration system for caricature generation** [11766-96]

11766 2V **Performance evaluation of liquid pouch inspection using conventional deep learning models** [11766-10]

11766 2W **Graph generation using point cloud data for path planning of autonomous mobile robots** [11766-131]

11766 2X **Time complexity reduction on light-field super-resolution with graph-based regularization** [11766-29]

11766 2Y **A method for assigning appropriate elevation to water portion in the DEM data** [11766-60]

11766 2Z **Development of a method for calculating overground-openness and underground-openness with high precision based on the digital elevation model data** [11766-63]

11766 30 **Intra CTU depth decision for HEVC by using neural networks** [11766-8]

**APPLICATION AND THEORY**

11766 31 **Vectorized implementation of K-means** [11766-61]

11766 32 **Aspect-based sentiment analysis: movie and television series reviews** [11766-46]

11766 33 **Study of minor dry-eye detection using smartphone camera based on deep learning** [11766-18]

11766 34 **Analysis of generative adversarial network targeting anti-forensic in JPEG compressed domain** [11766-73]

11766 35 **Cross-domain recommendation based on multilayer graph analysis using subgraph representation** [11766-91]

11766 36 **An autonomous control method for UAV using the relation between flight command and 3D moving vector** [11766-129]

**AUTONOMOUS VEHICLES**

11766 37 **Information display design on turn-by-turn navigation for visually impaired people** [11766-50]

11766 38 **Real-time gesture control for automotive infotainment system** [11766-9]

11766 39 **A method for visualizing train delay trends on railway lines using open data** [11766-114]

11766 3A **Nighttime image dehazing based on improved erosion dark channel and multi-scale clipping limit histogram equalization** [11766-137]

11766 3B **Relationship between covert attention and eye movements during observation of natural scenes** [11766-111]

11766 3C **Video dividing method for high-resolution video regardless of people position** [11766-51]