

- 8.30-10.00** **Session 0: Opening Session**
(Napalai Room)
Chair: Thanarat Chalidabhongse
- Welcome Speech**
Assoc. Prof. Dr. Supot Techavorasinskun
Dean, Faculty of Engineering Chulalongkorn University
- Conference Report**
Assoc. Prof. Dr. Supavadee Aramvith
General Co-Chairs, MMM 2018
Faculty of Engineering Chulalongkorn University
- Keynote (1)**
- Multimedia and Art in the Age of Creative AI**
Prof. Ahmed Elgammal
Department of Computer Science
Rutgers University, USA
- 10.00-10.30** **Refreshment Break**
(Foyer)
- 10.30-12.00** **Session 1: Best Paper Session**
(Napalai Room)
Chairs: Chong-wah Ngo and Thanarat Chalidabhongse
- 1 . **[18]** LOCO: Local Context based Faster R-CNN for Small Traffic Sign Detection
 - 2 . **[86]** Crowd Distribution Estimation with Multi-scale Recursive Convolutional Neural Network
 - 3 . **[110]** A Markov Network Based Passage Retrieval Method for Multimodal Question Answering in the Cultural Heritage Domain
 - 4 . **[170]** Frame-Based Classification of Operation Phases in Cataract Surgery Videos
- 12.00-14.00** **Lunch**
(Vimarn Suriya Room)
VBS preparation
(Dusit Thani Hall)
- 14.00-15.30** **Session 2: Demo/Poster Fast Forward Session**
(Napalai Room)
Chairs: Wolfgang Huerst and Suvit Nakpeerayuth
- 15.30-16.00** **Refreshment Break**
(Foyer)
- 16.00-19.00** **Session 3: Video Browser Showdown (VBS) Competition**
(Dusit Thani Hall)
Chairs: Cathal Gurrin, Werner Bailer, Klaus Schoeffmann, Jakub Lokoc, Kunwadee Sripanidkulchai

1. **[195]** Enhanced VIREO KIS at VBS 2018
2. **[184]** Competitive Video Retrieval with vitrivr
3. **[202]** The ITEC Collaborative Video Search System at the Video Browser Showdown 2018
4. **[192]** Sketch-based Similarity Search for Collaborative Feature Maps
5. **[200]** Video Search Based on Semantic Extraction and Locally Regional Object Proposal
6. **[189]** Revisiting SIRET Video Retrieval Tool
7. **[201]** Sloth Search System
8. **[188]** VERGE IN VBS 2018
9. **[196]** Fusing keyword search and visual exploration for untagged videos

17.00-20.00 **Welcome reception [International Cocktails]**
(Dusit Thani Hall)

9.00-10.00 **Session 4: Keynote (2)** (Napalai Room)
Chair: Yo-Song Ho

Perception of Visual Sentiment: From Experimental Psychology to Computational Modeling

Prof. Mohan Kankanhalli
School of Computing,
National University of Singapore
Singapore

10.00-10.30 **Refreshment Break**
(Foyer)

10.30-12.00 **Session 5A: Music, Audio, Art**
(Napalai Room)
Chair: Nisachon Tangsangiumvisai

- 1 . **[21]** Stitch-Based Image Stylization for Thread Art Using Sparse Modeling
- 2 . **[141]** Towards Demographic-based Photographic Aesthetics Prediction for Portraits
- 3 . **[48]** An RNN-Based Speech-Music Discrimination Used for Hybrid Audio Coder
- 4 . **[5]** Triplet Convolutional Network for Music Version Identification

Session 5B: Multimedia Application
(Vimarn Suriya Room)
Chair: Cathal Gurrin

- 1 . **[171]** Multimodal Augmented Reality - Augmenting Auditory-Tactile Feedback to Change the Perception of Thickness
- 2 . **[173]** Spatiotemporal 3D Models of Aging Fruit from Multi-View Time-Lapse Videos
- 3 . **[91]** Two-level Segment-based Bitrate Control for Live ABR Streaming
- 4 . **[9]** Multi-Hypothesis-based Error Concealment for Whole Frame Loss in HEVC

12.00-13.30 **Lunch**
(The Cellar)

13.30-15.00 **Session 6A: Quality Enhancement**
(Napalai Room)
Chair: Thanarat Chalidabhongse

- 1 . **[118]** Parameter Selection for Denoising Algorithms Using NR-IQA with CNN
- 2 . **[100]** Co-occurrent Structural Edge Detection for Color-Guided Depth Map Super-resolution

Session 6B: SS - Multimedia Analytics: Perspectives, Techniques and Applications
(Vimarn Suriya Room)
Chair: Björn Þór Jónsson

- 1 . **[158]** Rethinking Summarization and Storytelling for Modern Social Multimedia
- 2 . **[63]** On the Traceability of Results from Deep Learning-based Cloud Services

- 3 . [19] Shallow-water Image Enhancement Using Relative Global Histogram Stretching Based on Adaptive Parameter Acquisition
- 4 . [97] The CAMETRON lecture recording system: High quality video recording and editing with minimal human supervision
- 3 . [64] Category Specific Post Popularity Prediction
- 4 . [167] Image Aesthetics and Content in Selecting Memorable Keyframes from Lifelogs
- 5 . [46] Approaches for Event Segmentation of Visual Lifelog Data

15.00-16.30

Session 7A: Face and Text Processing

(Napalai Room)

Chair: Takao Onoye

- 1 . [150] Teacher and Student Joint Learning for Compact Facial Landmark Detection Network
- 2 . [34] Font Recognition in Natural Images via Transfer Learning
- 3 . [147] Uyghur Text Localization with Fast Component Detection
- 4 . [94] Text Image Deblurring via Intensity Extremums Prior

Session 7B: Indexing and Classification

(Vimarn Suriya Room)

Chair: Bernd Münzer

- 1 . [84] Learning to Index in Large-Scale Datasets
- 2 . [10] Iterative Active Classification of Large Image Collection
- 3 . [36] Cost-sensitive Deep Metric Learning for Fine-Grained Image Classification

16.30-17.00

Refreshment Break

(Foyer)

17.00-18.30

Session 8A: Video Search and Recognition

(Napalai Room)

Chair: Charnchai Pluempitiwiriyaewej

- 1 . [180] Evaluation of Visual Content Descriptors for Supporting Ad-Hoc Video Search Tasks at the Video Browser Showdown
- 2 . [104] A Novel 3D Human Action Recognition Framework for Video Content Analysis
- 3 . [102] Multi-Stream Fusion Model for Social Relation Recognition from Videos

Session 8B: Video and Image Processing

(Vimarn Suriya Room)

Chairs: Kanokphan Lertniphonphan and Manfred Jürgen Primus

- 1 . [132] A motion-driven approach for fine-grained temporal segmentation of user-generated videos
- 2 . [164] Real-Time Polyps Segmentation for Colonoscopy Video Frames Using Compressed Fully Convolutional Network
- 3 . [124] Find me a sky: a data-driven method for color-consistent sky search & replacement
- 4 . [39] Remote Sensing Image Fusion Based on Two-stream Fusion Network

19.00-21.00

MMM 2018 Banquet

(Dusit Thani Hall)

MC: Proadpran Punyabukkana

9.00-10.00 **Session 9: Keynote (3)** (Napalai Room)
Chair: Supavadee Aramvith

Multimodal Social Signals Analysis
Prof. Nicu Sebe
Department of Information Engineering and
Computer Science
University of Trento
Italy

10.00-10.30 **Refreshment Break**
 (Foyer)

10.30-12.00 **Session 10A: Cross-modal Retrieval and**
Processing
 (The Cellar)
Chair: Klaus Schoeffmann

- 1 . **[159]** Efficient and Interactive Spatial-Semantic Image Retrieval
- 2 . **[35]** Recursive Pyramid Network with Joint Attention for Cross-media Retrieval
- 3 . **[172]** Improving the quality of video-to-language models by optimizing annotation of the training material
- 4 . **[52]** Deep Convolutional Neural Network for Correlating Images and Sentences

Session 10B: Visual Parsing and Tracking
 (Vimarn Suriya Room)
Chair: Kai Barthel

- 1 . **[82]** Reinforcing Pedestrian Parsing on Small Scale Dataset
- 2 . **[69]** Deep Pedestrian Detection using Contextual Information and Multi-level Features
- 3 . **[65]** REVT: Robust and Efficient Visual Tracking by Region-Convolutional Regression Network
- 4 . **[25]** Collision-Free LSTM for Human Trajectory Prediction

12.00-13.30 **Lunch**
 (22 Kitchen & Bar)

13.30-15.00 **Session 11A: Vision & Graphics**
 (The Cellar)
Chair: Suree Pumrin

- 1 . **[50]** Light Field Foreground Matting Based on Defocus and Correspondence
- 2 . **[3]** Dual-way Guided Depth Image Inpainting With RGBD Image Pairs
- 3 . **[15]** High-Precision 3D Coarse Registration Using RANSAC and Randomly-Picked Rejections

Session 11B: Deep Learning and Its Applications
 (Vimarn Suriya Room)
Chair: Peerapon Vateekul

- 1 . **[149]** Convolution with Logarithmic Filter Groups for Efficient Shallow CNN
- 2 . **[14]** Image Aesthetic Distribution Prediction with Fully Convolutional Network

4 . **[22]** AGO: Accelerating Global Optimization for Accurate Stereo Matching

- 3 . **[140]** A Method of weather radar echo extrapolation based on convolutional neural networks
- 4 . **[162]** Adaptive Image Representation using Information Gain and Saliency: Application to Cultural Heritage Datasets

15.00-17.00

Refreshment Break

(Foyer)

Session 12A: Poster Session

(Vimarn Suriya Room)

Chairs: Chong-Wah Ngo, Thanarat

Chalidabhongse, Klaus Schoeffmann

- 1 . **[2]** Scalable Bag of Selected Deep Features for Visual Instance Retrieval
- 2 . **[27]** Efficient Two-layer Model towards Cover Song Identification
- 3 . **[32]** Source Distortion Estimation for Wyner-Ziv Distributed Video Coding
- 4 . **[37]** A Novel Frontal Facial Synthesis Algorithm based on Individual Residual Faces
- 5 . **[43]** A Text Recognition and Retrieval System for e-Business Image Management
- 6 . **[53]** ShapeCreator: 3D Shape Generation from Isomorphic Datasets Based on Autoencoder
- 7 . **[57]** k-Labelsets for Multimedia Classification with Global and Local Label Correlation
- 8 . **[58]** Domain Invariant Subspace Learning for Cross-modal Retrieval
- 9 . **[72]** Coarse-to-Fine Image Super-Resolution Using Convolutional Neural Networks
- 10 . **[76]** CNN-Based DCT-Like Transform for Image Compression
- 11 . **[83]** Data Augmentation for EEG-Based Emotion Recognition with Deep Convolutional Neural Networks
- 12 . **[88]** A New Accurate Image Denoising Method Based on Sparse Coding Coefficients
- 13 . **[90]** Fusion Networks for Air-writing Recognition

Session 12B: Demonstration Session and VBS Showcase

(Vimarn Suriya Room)

Chairs: Wolfgang Huerst, Suvit Nakpeerayuth

- 1 . **[126]** Lifelog Exploration Prototype in Virtual Reality
- 2 . **[181]** Auto accessory segmentation and interactive try-on system
- 3 . **[185]** ActionVis: An Explorative Tool to Visualize Surgical Actions in Gynecologic Laparoscopy
- 4 . **[186]** Automatic Smoke Classification in Endoscopic Video
- 5 . **[187]** Depth Representation of LiDAR Point Cloud with Adaptive Surface Patching for Object Classification
- 6 . **[190]** Ontlus: 3D Content Collaborative Creation via Virtual Reality
- 7 . **[191]** Programmatic 3D Printing of a Revolving Camera Track to Automatically Capture Dense Images for 3D Scanning of Objects
- 8 . **[193]** Video Browsing on a Circular Timeline
- 9 . **[194]** Multi-Camera Microenvironment to Capture Multi-view Time-lapse Videos for 3D Analysis of Aging Objects
- 10 . **[197]** A Virtual Reality Interface for Interactions with Spatiotemporal 3D Data
- 11 . **[198]** AR DeepCalorieCam: An iOS App for Food Calorie Estimation with Augmented Reality
- 12 . **[199]** ImageX – Explore and search local/private images

- 14 . **[96]** Cloud of Line Distribution and Random Forest Text Detection from Natural/Video Scene Images
- 15 . **[98]** Room Floor Plan Generation on a Project Tango Device
- 16 . **[101]** Person Re-identification by Incorporating PCA Loss in CNN
- 17 . **[105]** Ensemble System for First Person Interactive Action Recognition
- 18 . **[108]** SRN: The Movie Character Relationship Analysis via Social Network
- 19 . **[111]** Effective Action Detection using Temporal Context and Posterior Probability of Length
- 20 . **[115]** Vehicle Semantics Extraction and Retrieval for Long-term Carpark Video Surveillance
- 21 . **[117]** Implicit Affective Video Tagging using Pupillary Response
- 22 . **[122]** Accurate Detection for Scene Texts with a Cascaded CNN Networks
- 23 . **[127]** Robust and Real-Time Visual Tracking based on Complementary Learners
- 24 . **[135]** The Long Tail of Web Video
- 25 . **[151]** Venue Prediction for Social Images by Exploiting Rich Temporal Patterns in LBSNs
- 26 . **[155]** LVFS: a Lightweight Video Storage File System for IP Camera-based Surveillance Applications
- 27 . **[160]** Food Photo Recognition for Dietary Tracking: System and Experiment
- 28 . **[169]** SeqSense: video recommendation using topic sequence mining

17.00-17.30

Closing Session

(Vimarn Suriya Room)

Chair: Supavadee Aramvith