

Thematic Workshop Pitches - I (Monday, Oct 23) 09:00 AM - 10:30 AM

1. An evaluation of large-scale methods for image instance and class discovery
2. Towards a Tangible Storytelling Kit for Exploring Emotions with Children
3. Intelligent Portrait Composition Assistance - Integrating Deep-learned Models and Photography Idea Retrieval
4. Beyond Saliency: Assessing Visual Balance with High-level Cues
5. Multispectral Object Detection for Autonomous Vehicles
6. Aggregated Deep Activation Clusters for Particular Object Retrieval
7. Local Deep Descriptors in Bag-of-Words for Image Retrieval
8. Beyond Sum and Weighted Aggregation: An Efficient Mixed Aggregation Method with Multiple Weights for Image Search
9. Detecting culture-specific tags for news videos through multimodal embedding
10. Hybrid Indexes for Spatial-Visual Search
11. Unsupervised Triplet Hashing for Fast Image Retrieval
12. Relational Variational Autoencoder for Link Prediction with Multimedia Data
13. First Mile in Crowdsourced Live Streaming: A Content Harvest Network Approach
14. efficient communications in training large scale neural networks
15. Augmenting Cognitive Processes and Behavior of Intelligent Virtual Agents by Synthetic Perception
16. Mobile Multispectral Video Streaming
17. Impact of three-dimensional video scalability on multi-view activity recognition using deep learning
18. CloudHide: Towards the Latency Hiding Techniques for Thin-client Cloud Games
19. Mobile Instant Video Sharing: Does More Information Help?
20. QoE-fair Adaptive Streaming of Free-viewpoint Videos over LTE Networks
21. Adaptive Multicast Streaming of Virtual Reality Content to Mobile Users
22. Self-Gated Recurrent Neural Networks for Human Activity Recognition on Wearable Devices
23. SeeNav: Seamless and Energy-Efficient Indoor Navigation using Augmented Reality
24. CloudAR: A Cloud-based Framework for Mobile Augmented Reality
25. A Stereoscopic Vision System with Delay Compensation for 360° Remote Reality
26. Recognizing the presence of hidden visual markers in digital images
27. Generating Virtual Avatars with Personalized Walking Gaits using Commodity Hardware
28. Towards the Security of Motion Detection-based Video Surveillance on IoT Devices
29. Multimodal Context-Aware Recommenders for Post Popularity Prediction in Social Media
30. Multimodal Classification of Violent Online Political Extremism Content with Graph Convolutional Networks
31. Compressed-domain Video Synopsis via Blank Frame Deletion
32. A Collaborative Representation Approach to Detecting Error-Related Potentials in SSVEP-BCIs

**MMHealth 2017: Workshop on Multimedia for Personal Health and Health Care
(Monday, Oct 23) 09:00 AM - 06:00 PM**

Session 1: 9:00-10:30

9:00 Welcome, Introduction

9:15 Keynote

Prof. Alan F. Smeaton Insight Centre for Data Analytics Dublin City University

Insights from Data Analytics Into Our Personal Sensor Data

10:10 Paper presentations 1: Cross-topic issues

(15min presentation, 5min discussion)

Research Challenges of Emerging Technologies Supporting Life-Long Health and Wellbeing

(Jochen Meyer and Parisa Eslambolchilar)

Session 2: 11:00-12:30:

Paper presentations 2: Data Analytics

11:00 Combining the benefits of CCA and SVMs for SSVEP-based BCIs in real-world conditions (Elisavet Chatzilari, Georgios Liaros, Kostas Georgiadis, Spiros Nikolopoulos and Yiannis Kompatsiaris)

11:20 Empirical Mode Decomposition of Throat Microphone Recordings for Intake Classification (Mehmet Ali Tugtekin Turan and Engin Erzin)

11:40 Wearable Emotion Recognition System based on GSR and PPG Signals (Goran Udovičić, Jurica Đerek, Mladen Russo and Marjan Sikora)

12:00 Denoising of Joint Tracking Data by Kinect Sensors using clustered Gaussian Process Regression (An-Ti Chiang and Yao Wang)

12:20 Discussion

Session 3: 14:00-16:00

14:00 Paper presentations 3: Systems and applications

14:00 Artificial Intelligence in XPRIZE DeepQ Tricorder (Edward Chang, Meng-Hsi Wu, Kai-Fu Tang, Hao-Cheng Kao and Chun-Nan Chou)

14:20 Live Personalized Nutrition Recommendation Engine (Nitish Nag, Vaibhav Pandey and Ramesh Jain)

14:40 Gamification of Heel Raise Plantarflexion Physiotherapy (Jeannie Lee, Darren Goh and Alfred Tan)

15:00 Discussion

15:10 Poster and demo presentations

(7 x 5min presentation (possibly a "rapid fire" format to be defined), 2min discussion)

- DeepQ Arrhythmia Database: A Large-Scale Dataset for Arrhythmia Detector Evaluation (Meng-Hsi Wu and Edward Chang)
- Toward Personalized Treatment of Chronic Diseases - the CKD Case Study (Chih-Yang Chen, Chun-Nan Chou and I-Wen Wu)
- Elderly Living Alone: Detecting Home Visits with Ambient and Wearable Sensing (Rui Hu, Hieu Pham, Philipp Buluscsek and Daniel Gatica-Perez)
- Few-shot Learning Using a Small-Sized Dataset of High-Resolution FUNDUS Images for Glaucoma Diagnosis (Mijung Kim, Jasper Zuallaert and Wesley De Neve)
- Managing Family Healthcare with Multimedia Chat Apps: A Survey on What is Missing (Britta Meixner, Matthew Lee and Scott Carter)
- Exploring Challenges in Automated Just-In-Time Adaptive Food Choice Interventions (Nađa Terzimehić, Christina Schneegass and Heinrich Hussmann)
- MedFit: a Mobile Application for Recovering CVD Patients (Jogile Kuklyte, Leonardo Gualano, Ghanashyama Prabhu, Kaushik Venkataraman, Deirdre Walsh, Catherine Woods, Kieran Moran and Noel E. O'Connor)

Session 4: 16:30-18:00

16:30-17:30 Poster and demo session

17:30-18:00 Wrap-up, closing.

***Summary for AVEC 2017 – Real-life Depression and Affect Challenge and Workshop
(Monday, Oct 23) 09:00 AM – 06:00 PM***

Session 0 – Keynote

Chair: Fabien Ringeval

9:00 – 10:00: AVEC'17 Keynote Talk – “Body Language Without Body: Social Signals in Technology Mediated Communication”, Alessandro Vinciarelli

Session 1 – Introduction

Chair: Alessandro Vinciarelli

10:00 – 10:30: “Challenge Introduction: AVEC 2017 – Real-life Depression, and Affect Recognition Workshop and Challenge”, Fabien Ringeval, Björn Schuller, Michel Valstar, Jonathan Gratch, Roddy Cowie, Stefan Scherer, Sharon Mozgai, Nicholas Cummins, Maximilian Schmitt, Maja Pantic

Session 2 – AVEC 2017 Part 1

Chair: Yue Zhang

11:00 – 11:25: “Continuous Multimodal Emotion Prediction Based on Long Short Term Memory Recurrent Neural Network”, Jian Huang, Ya Li, Jianhua Tao, Zheng Lian, Zhengqi Wen, Minghao Yang and Jiangyan Yi

11:25 – 11:50: “Multimodal Multi-task Learning for Dimensional and Continuous Emotion Recognition”, Shizhe Chen, Qin Jin, Jinming Zhao and Shuai Wang

11:50 – 12:15: “Investigating Word Affect Features and Fusion of Probabilistic Predictions Incorporating Uncertainty in AVEC 2017”, Ting Dang, Brian Stasak, Zhaocheng Huang, Sadari Jayawardena, Mia Atcheson, Munawar Hayat, Phu Le, Vidhyasaharan Sethu, Roland Goecke and Julien Epps

Session 3 – AVEC 2017 Part 2

Chair: Fabien Ringeval

14:00 – 14:25: “Depression Severity Prediction Based on Biomarkers of Psychomotor Retardation”, Zafi Sherhan Syed, Kirill Sidorov and David Marshall

14:25 – 14:50: “Hybrid Depression Classification and Estimation from Audio Video and Text Information”, Le Yang, Hichem Sahli, Xiaohan Xia, Ercheng Pei, Meshia Cédric Oveneke and Dongmei Jiang

14:50 – 15:15: “Multimodal Measurement of Depression Using Deep Learning Models”, Le Yang, Dongmei Jiang, Xiaohan Xia, Ercheng Pei, Meshia Cédric Oveneke and Hichem Sahli

15:15 – 15:40: “A Random Forest Regression Method With Selected-Text Feature For Depression Assessment”, Bo Sun, Yinghui Zhang and Jun He

15:40 – 16:05: “Topic Modeling Based Multi-modal Depression Detection”, Yuan Gong and Christian Poellabauer

16:05 – 16:15: Challenge Results

Session 4 – Panel Session

Chairs: Yue Zhang, Fabien Ringeval

17:00 – 17:30: Panel Session

SAWACMMM'17: The 1st Workshop on Multi Media Applications within the South African Context (Monday, Oct 23) 09:00 AM – 12:30 PM

Human-like Visual Learning and Reasoning [Peng Cui (Tsinghua University); Wenwu Zhu (Tsinghua University)] (Monday, Oct 23) 11:00 AM - 12:30 PM

Social Multimedia Sentiment Analysis [Jiebo Luo (University of Rochester); Damian Borth (DFKI); Quanzeng You (University of Rochester)] (Monday, Oct 23) 02:00 PM - 03:30 PM

LTA 2017: The Second Workshop on Lifelogging Tools and Applications (Monday, Oct 23) 02:00 PM - 06:00 PM

Welcome & Introduction (14:00 - 14:10)
LTA 2017 Organising Committee

Keynote (14:10 to 15:00)
LIAM - A Lifelog-driven Life Management System. Tahl Milburn - CEO LifeState.io

Full Papers (15:00 - 15:40)
Semantic Summarization of Egocentric Photo Stream Events. Aniol Lidon (Universitat Politecnica de Catalunya); Marc Bolaños (Universitat de Barcelona); Mariella Dimiccoli (Universitat de Barcelona); Petia Radeva (Universitat de Barcelona); Maite Garolera (Consorti Sanitari de Terrassa); Xavier Giro-I-Nieto (Universitat Politecnica de Catalunya)

Are conversation logs useful sources for generating memory cues for recalling past memories? Seyed Ali Bahrainian (University of Lugano - USI); Fabio Crestani (University of Lugano - USI).

Short Papers (15:40 - 16:00)
A Baseline Search Engine for Personal Life Archives. Liting Zhou (Insight Centre for Data Analytics at Dublin City University); Duc-Tien Dang-Nguyen (Insight Centre for Data Analytics at Dublin City University); Cathal Gurrin (Insight Centre for Data Analytics at Dublin City University)

Big Five Personality Measurement Based On Lifelogs
Yewen Wang (Tsinghua University); Min Zhang (Tsinghua University); Pouneh Soleimaninejadian (Tsinghua University); Haoyue Tong (Tsinghua University); Zehui Feng (Tsinghua University).

Pilot Study to Investigate Feasibility of Visual Lifelog Exploration in Virtual Reality
Aaron Duane (Insight Centre for Data Analytics at Dublin City University); Cathal Gurrin (Insight Centre for Data Analytics at Dublin City University)

Coffee Break (16:00 - 16:30)

Poster Session (16:30 - 17:20)
All five papers will present posters.

Panel (17:20 - 17:50)
A short panel discussion will be chaired by Cathal Gurrin and include three leaders in the field of lifelogging.

Closing remarks (17:50 - 18:00)
Closing remarks by the organizers and suggestions for how to organizers LTA2018.

Deep Learning for Intelligent Video Analysis [Ting Yao (Microsoft Research); Tao Mei (Microsoft Research); Cha Zhang (Microsoft AI & Research)] (Monday, Oct 23) 04:00 PM - 06:00 PM

Fast Forward 1 (Tuesday, Oct 24) 08:25 AM - 09:00 AM

Attention Transfer from Web Images for Video Recognition

SketchParse: Towards Rich Descriptions for Poorly Drawn Sketches using Multi-Task Hierarchical Deep Networks

Place-centric Visual Urban Perception with Deep Multi-instance Regression

Future-Supervised Retrieval of Unseen Queries for Live Video

Learning to Compose with Professional Photographs on the Web

StructCap: Structured Semantic Embedding for Image Captioning

Is Foveated Rendering Perceivable in Virtual Reality? Exploring the Efficiency and Consistency of Quality Assessment Methods

FaceCollage: A Rapidly Deployable System for Real-time Head Reconstruction for On-The-Go 3D Telepresence

LiveJack: Integrating CDNs and Edge Clouds for Live Content Broadcasting

Face Aging with Contextual Generative Adversarial Nets

Fashion World Map: Understanding Cities Through Streetwear Fashion

Automatic Adjustment of Stereoscopic Content for Long-Range Projections in Outdoor Areas

Multiview and Multimodal Pervasive Indoor Localization

Searching Personal Photos on the Phone with Instant Visual Query Suggestion and Joint Text-Image Hashing

A Unified Personalized Video Recommendation via Dynamic Recurrent Neural Networks

Keynote Address 1 (Tuesday, Oct 24) 09:00 AM - 10:00 AM

Enhancing and Augmenting Human Perception with Artificial Intelligence
[Achin Bhowmik: CTO & EVP, Starkey, USA]

Best Paper Presentation (Tuesday, Oct 24) 10:30 AM - 12:00 PM

H-TIME: Haptic-enabled Tele-Immersive Musculoskeletal Examination

Catching the Temporal Regions-of-Interest for Video Captioning

Adversarial Cross-Modal Retrieval
Deep Low-rank Sparse Collective Factorization for Cross-Domain Recommendation
<i>Fast Forward 2 (Tuesday, Oct 24) 01:30 PM - 02:00 PM</i>
Unconstrained Fashion Landmark Detection via Hierarchical Recurrent Transformer Networks
Deep Attribute-preserving Metric Learning for Natural Language Object Retrieval
Understanding Fashion Trends from Street Photos via Neighbor-Constrained Embedding Learning
Skeleton-Aided Articulated Motion Generation
Deep Progressive Hashing for Image Retrieval
The Role of Visual Attention in Sentiment Prediction
Robust Visual Object Tracking with Top-down Reasoning
Pedestrian Path Forecasting in Crowd: A Deep Spatio-Temporal Perspective
Stylized Adversarial AutoEncoder for Image Generation
ReGLe: Spatially Regularized Graph Learning for Visual Tracking
Deep Unsupervised Convolutional Domain Adaptation
Improving Event Extraction via Multimodal Integration
A Dual-Network Progressive Approach to Weakly Supervised Object Detection
Multimodal Learning for Web Information Extraction
Fast Deep Matting for Portrait Animation on Mobile Phone
An HTTP/2-Based Adaptive Streaming Framework for 360° Virtual Reality Videos
360ProbDASH: Improving QoE of 360 Video Streaming Using Tile-based HTTP Adaptive Streaming
ShareRender: Bypassing GPU Virtualization to Enable Fine-grained Resource Sharing for Cloud Gaming
Temporal Binary Coding for Large-Scale Video Search
One-Shot Fine-Grained Instance Retrieval

Modeling the Intransitive Pairwise Image Preference from Multiple Angles
PD-Survey - Supporting Audience-Centric Research through Surveys on Pervasive Display Networks
Learning Visual Emotion Distributions via Multi-Modal Features Fusion
Exploiting High-Level Semantics for No-Reference Image Quality Assessment of Realistic Blur Images
A Paralinguistic Approach To Speaker Diarisation
Wheel: Accelerating CNNs with Distributed GPUs via Hybrid Parallelism and Alternate Strategy
A Delicious Recipe Analysis Framework for Exploring Multi-Modal Recipes with Various Attributes
Multi-Modal Knowledge Representation Learning via Webly-Supervised Relationships Mining
GLAD: Global-Local-Alignment Descriptor for Pedestrian Retrieval
<i>Keynote Address 2 (Tuesday, Oct 24) 02:00 PM - 03:00 PM</i>
Efficient Methods and Hardware for Deep Learning [Bill Dally: Senior Vice President and Chief Scientist, NVidia, USA]
<i>Understanding 1 - Deep Learning for MM (1) (Tuesday, Oct 24) 03:30 PM - 05:00 PM</i>
Magic-wall: Visualizing Room Decoration
Multi-Scale Cascade Network for Salient Object Detection
Sketch Recognition with Deep Visual-Sequential Fusion Model
<i>Panel 1 (Tuesday, Oct 24) 03:30 PM - 05:00 PM</i>
Privacy Protection in Online Multimedia [Panel: Yung-Hsiang Lu (Purdue University), Andrea Cavallaro (Queen Mary University of London); Catherine Crump (University of California Berkeley); Gerald Friedland (University of California Berkeley); Keith Winstein (Stanford University)]
<i>Experience 1 - Social and Affective Multimedia (Tuesday, Oct 24) 03:30 PM - 05:00 PM</i>
What your Facebook Profile Picture Reveals about your Personality
Capturing Spatial and Temporal Patterns for Distinguishing between Posed and Spontaneous Expressions
An Image-based Deep Spectrum Feature Representation for the Recognition of Emotional Speech
Automatic Generation of Lyrics Parodies

Systems 1 - Systems and Applications (Tuesday, Oct 24) 05:15 PM - 06:45 PM

On Server Provisioning for Cloud Gaming

FastShrinkage: Perceptually-aware Retargeting Toward Mobile Platforms

Real-time Monocular Dense Mapping for Augmented Reality

Engagement 1 - Multimedia Search and Recommendation (Tuesday, Oct 24) 05:15 PM - 06:45 PM

Automatic Music Video Generation Based on Simultaneous Soundtrack Recommendation and Video Editing

Region-based Image Retrieval Revisited by Semantic Region Specification and Spatial Relationship Recommendation

Learning Multimodal Attention LSTM Networks for Video Captioning

Business Idea Venture (Tuesday, Oct 24) 05:15 PM - 06:45 PM

Profilio: Psychometric profiling to boost social media advertising

Pl@ntNet My Business

Interactive Art (Tuesday, Oct 24) 09:00 AM - 06:45 PM

Drag A Star 3.0

PPROBABLY/POSSIBLY?

Touch Me Here

Filters

Split Consideration for Foreground and Background Painting Using Artificial Neural Networks

Spatial Magnetic Field Visualization

A Quatre Mains

Las Barricadas Misteriosas

Empowering Clerks Network

Presently Untitled: Data Mapping of 2016 U.S. Presidential Election Twitter Activity, Phase III

Fast Forward 3 (Wednesday, Oct 25) 08:25 AM - 09:00 AM

Query-adaptive Video Summarization via Quality-aware Relevance Estimation

Predicting Human Intentions from Motion Cues Only: A 2D+3D Fusion Approach

RGB-D Scene Recognition with Object-to-Object Relation

Data Generation for Improving Person Re-identification

Salient Object Detection with Chained Multi-Scale Fully Convolutional Network

Fine-grained Discriminative Localization via Saliency-guided Faster R-CNN

Learning to Recognise Unseen Classes by A Few Similes

Deep Cross-Modality Alignment for Multi-Shot Person Re-IDentification

Improved Multimodal Representation Learning with Skip Connections

Modeling Image Virality with Pairwise Spatial Transformer Networks

Metric-based Generative Adversarial Network

More Than An Answer: Neural Pivot Network for Visual Question Answering

Aristo: An Augmented Reality Platform for Immersion and Interactivity

Sports VR Content Generation from Regular Camera Feeds

OpTile: Toward Optimal Tiling in 360-degree Video Streaming

Too Many Pixels to Perceive: Subpixel Shutoff for Display Energy Reduction on OLED Smartphones

Exploring Consistent Preferences: Discrete Hashing with Pair-Exemplar for Scalable Landmark Search

Fast and Accurate Pedestrian Detection using Dual-Stage Group Cost-Sensitive RealBoost with Vector Form Filters

Online Cross-Modal Scene Retrieval by Binary Representation and Semantic Graph

NeuroStylist: Neural Compatibility Modeling for Clothing Matching

It's All Around You: Exploring 360° Video Viewing Experiences on Mobile Devices

Exploring Domain Knowledge for Affective Video Content Analyses

Occlusion-aware Video Temporal Consistency
Protest Activity Detection and Perceived Violence Estimation from Social Media Images
Multimodal Fusion with Recurrent Neural Networks for Rumor Detection on Microblogs
Keynote Address 3 (Wednesday, Oct 25) 09:00 AM - 10:00 AM
Building Multi-Modal Interfaces for Smartphones [Injong Rhee: CTO & EVP, Samsung Electronics, Korea]
SIGMM Award session (Wednesday, Oct 25) 10:30 AM - 12:00 PM
ACM SIGMM Technical Achievement Award 2017 [Recipient: Prof. Dr. Arnold Smeulders]
ACM SIGMM Rising Star Award 2017 [Recipient: Dr. Liangliang Cao]
ACM SIGMM Outstanding Ph.D. Thesis Award 2017 [Recipient: Dr. Chien-Nan Chen]
Doctoral Symposium (Wednesday, Oct 25) 12:00 PM - 01:30 PM
Using DASH Assisting Network Elements for Optimizing Video Streaming Quality
Who Composes the Music? Musicality Evaluation for Algorithmic Composition via Electroencephalography
Cross-media Relevance Computation for Multimedia Retrieval
Towards Global Optimization in Display Advertising by Integrating Multimedia Metrics to Real-Time Bidding
On Job Training: Automated Interpersonal Behavior Assessment & Real-Time Feedback
Fast Forward 4 (Wednesday, Oct 25) 01:30 PM - 02:00 PM
Indefinite Kernel Logistic Regression
Positive and Unlabeled Learning for Anomaly Detection with Multi-features
Hierarchical Recurrent Neural Network for Video Summarization
Learning a Target Sample Re-Generator for Cross-Database Micro-Expression Recognition
From Multimedia Logs to Personal Chronicles
From Hard to Soft: Towards more Human-like Emotion Recognition by Modelling the Perception Uncertainty

Two Birds One Stone
Multi-Networks Joint Learning for Large-Scale Cross-Modal Retrieval
Photo2Trip: Exploiting Visual Contents in Geo-tagged Photos for Personalized Tour Recommendation
Rethinking HTTP Adaptive Streaming with the Mobile User Perception
REQUEST: Seamless Dynamic Adaptive Streaming over HTTP for Multi-Homed Smartphone under Resource Constraints
Optimal Set of 360-Degree Videos for Viewport-Adaptive Streaming
Deep Active Learning Through Cognitive Information Parcels
3DensiNet: A Robust Neural Network Architecture towards 3D Volumetric Object Prediction from 2D Image
Towards Micro-video Understanding by Joint Sequential-Sparse Modeling
LEAF: Latent Extended Attribute Features Discovery for Visual Classification
Single Shot Temporal Action Detection
Finding the Secret of CNN Parameter Layout under Strict Size Constraint
Deep Temporal Models using Identity Skip-Connections for Speech Emotion Recognition
Video Description with Spatial-Temporal Attention
Pedestrian Detection via Bi-directional Multi-scale Analysis
Fine-Grained Recognition via Attribute-Guided Attentive Feature Aggregation
NormFace: L2 Hypersphere Embedding for Face Verification
Video Question Answering via Hierarchical Dual-Level Attention Network Learning
Region-based Activity Recognition Using Conditional GAN
<i>Keynote Address 4 (Wednesday, Oct 25) 02:00 PM - 03:00 PM</i>
DeepQ: Advancing Healthcare Through AI and VR [Edward Y. Chang: President, HTC, Taiwan]
<i>Understanding 2 - Deep Learning for MM (2) (Wednesday, Oct 25) 03:30 PM - 05:00 PM</i>
Detecting Temporal Proposal for Action Localization with Tree-structured Search Policy

Learning Fashion Compatibility with Bidirectional LSTMs
3D CNNs on Distance Matrices for Human Action Recognition
Sync-DRAW: Automatic Video Generation using Deep Recurrent Attentive Architectures
<i>Systems 2 - Video Streaming (Wednesday, Oct 25) 03:30 PM - 05:00 PM</i>
16K Cinematic VR Streaming
Where are the Sweet Spots? A Systematic Approach to Reproducible DASH Player Comparisons
Towards Forward-looking Online Bitrate Adaptation for DASH
QUETRA: A Queuing Theory Approach to DASH Rate Adaptation
<i>Experience 2 - Perceptual, Affect and Interaction (Wednesday, Oct 25) 03:30 PM - 05:00 PM</i>
Vocktail: A Virtual Cocktail for Pairing Digital Taste, Smell, and Color Sensations
Affect Recognition in Ads with Application to Computational Advertising
Image Quality Assessment for DIBR Synthesized Views using Elastic Metric
ElasticPlay: Interactive Video Summarization with Dynamic Time Budgets
<i>Panel 2 (Wednesday, Oct 25) 05:15 PM - 06:45 PM</i>
Cross-media Intelligence [Panel: Yueting Zhuang (Zhejiang University); Ramesh Jain (University of California, Irvine); Wen Gao (Peking University); Liu Ren (Robert Bosch Research)]
<i>Engagement 2 - Digital Society & Multimedia Art, Entertainment and Culture (Wednesday, Oct 25) 05:15 PM - 06:45 PM</i>
From Part to Whole: Who is Behind the Painting?
DeepArt: Learning Joint Representations of Visual Arts
Enhancing Micro-video Understanding by Harnessing External Sounds
<i>Open Source Software Competition (Wednesday, Oct 25) 05:15 PM - 06:45 PM</i>
TensorLayer: A Versatile Library for Efficient Deep Learning Development
NUBOMEDIA: The First Open Source WebRTC PaaS

BMXNet: An Open-Source Binary Neural Network Implementation Based on MXNet
WebDNN: Fastest DNN Execution Framework on Web Browser
ChainerCV: a Library for Deep Learning in Computer Vision
UnrealCV: Virtual Worlds for Computer Vision
<i>Demo (Wednesday, Oct 25) 01:30 PM - 06:45 PM</i>
A System for Spatiotemporal Anomaly Localization in Surveillance Videos
A Tag Recommendation System for Popularity Boosting
DeepCADx: Automated Prostate Cancer Detection and Diagnosis in mp-MRI based on Multimodal Convolutional Neural Networks
MatPlanner: Plan Your Days in Conferences by Resolving Conflicting Events
Natural Experiences in Museums through Virtual Reality and Voice Commands
FaceCloud: Heterogeneous Cloud Visualization of Multiplex Networks for Multimedia Archive Exploration
Real-Time Dense Monocular SLAM for Augmented Reality
Enhancing Music Events Using Physiological Sensor Data
Teleconsultant: Communication and analysis of wearable videos in Emergency Medical Environments
mIDoT-key: A Smart Key Instantly Generated on Your Item
Time Traveler: a real-time face aging system
Outdoor Object Recognition for Smart Audio Guides
IBM High-Five: Highlights From Intelligent Video Engine
Shadow Puppetry with Robotic Arms
Smart Mirror: Intelligent Makeup Recommendation and Synthesis
Real-Time Deep Video SpaTial Resolution UpConversion SysTem (STRUCT++ Demo)
RSVP: A Real-Time Surveillance Video Parsing System with Single Frame Supervision
NexGenTV: Providing Real-Time Insight during Political Debates in a Second Screen Application

Interactive Visualization for Travel Route Recommendations
A Hybrid P2P/Multi-Server Quality-Adaptive Live-Streaming Solution Enhancing End-User's QoE
Diversified and Summarized Video Search System
A Deep Cross-Platform Video Recommendation System with Novel Interface
Sketch-based Image Retrieval using Generative Adversarial Networks
<i>Video (Wednesday, Oct 25) 01:30 PM - 06:45 PM</i>
PIC2DISH: A Customized Cooking Assistant System
<i>Fast Forward 5 (Thursday, Oct 26) 08:25 AM - 09:00 AM</i>
Visual Sentiment Analysis for Review Images with Item-Oriented and User-Oriented CNN
Mutually Guided Image Filtering
Learning Semantic Feature Map for Visual Content Recognition
Video Visual Relation Detection
Deep Location-Specific Tracking
A Multi-Task Framework for Weather Recognition
Discriminative Training of Complex-valued Deep Recurrent Neural Network for Singing Voice Separation
Adaptive Low-Rank Multi-Label Active Learning for Image Classification
Adaptively Attending to Visual Attributes and Linguistic Knowledge for Captioning
Efficient Binary Coding for Subspace-based Query-by-Image Video Retrieval
FRACTaL: FEC-based Rate Control for RTP
When Cloud Meets Uncertain Crowd: An Auction Approach for Crowdsourced Livecast Transcoding
Multicamera Summarization of Rehabilitation Sessions in Home Environment
Visualization of Stone Trajectories in Live Curling Broadcasts using Online Machine Learning
Deep Binary Reconstruction for Cross-modal Hashing

Semi-Dense Depth Interpolation using Deep Convolutional Neural Networks
Venues in Social Media: Examining Ambiance Perception Through Scene Semantics
Moving as a Leader: Detecting Emergent Leadership in Small Groups using Body Pose
#VisualHashtags: Visual Summarization of Social Media Events Using Mid-Level Visual Elements
Multi-scale Context Based Attention for Dynamic Music Emotion Prediction
A Simplified Topological Representation of Text for Local and Global Context
Experimental Analysis of Bandwidth Allocation in Automated Video Surveillance Systems
Multimedia Semantic Integrity Assessment Using Joint Embedding Of Images And Text
Real-Time False-Contours Removal for Inverse Tone Mapped HDR Content
Deep Matching and Validation Network
Keynote Address 5 (Thursday, Oct 26) 09:00 AM - 10:00 AM
Bringing a Billion Hours to Life [Scott Silver: Vice President, Google, USA]
Understanding 3 - Deep Learning for MM (3) (Thursday, Oct 26) 10:30 AM - 12:00 PM
Learning Object-Centric Transformation for Video Prediction
Two-Stream Attentive CNNs for Image Retrieval
Deep Asymmetric Pairwise Hashing
Integrated Face Analytics Networks through Cross-Dataset Hybrid Training
Novel 1 (Thursday, Oct 26) 10:30 AM - 12:00 PM
Exploring Outliers in Crowdsourced Ranking for QoE
Fluency-Guided Cross-Lingual Image Captioning
Mr.MAPP: Mixed Reality for MAnaging Phantom Pain
Anti-camera LED Lighting

Fast Forward 6 (Thursday, Oct 26) 01:30 PM - 02:00 PM

Incremental Accelerated Kernel Discriminant Analysis

Pseudo Label based Unsupervised Deep Discriminative Hashing for Image Retrieval

Multi-Modal Localization and Enhancement of Multiple Sound Sources from a Micro Aerial Vehicle

Selective Deep Convolutional Features for Image Retrieval

Statistical Inference of Gaussian-Laplace Distribution for Person Verification

Beyond Human-level License Plate Super-resolution with Progressive Vehicle Search and Domain Prior GAN

Learning to Generate and Edit Hairstyles

Adaptively Weighted Multi-task Deep Network for Person Attribute Classification

Video Question Answering via Gradually Refined Attention over Appearance and Motion

Cross-Domain Image Retrieval with Attention Modeling

Modeling the Resource Requirements of Convolutional Neural Networks on Mobile Devices

Adaptive Audio Classification for Smartphone in Noisy Car Environment

A Novel System for Visual Navigation of Educational Videos Using Multimodal Cues

Adaptive 360-Degree Video Streaming using Scalable Video Coding

Cross-media Retrieval by Learning Rich Semantic Embeddings of Multimedia

Deep Supervised Quantization by Self-Organizing Map

Laplacian-Steered Neural Style Transfer

PQk-means: Billion-scale Clustering for Product-quantized Codes

Outlining Objects for Interactive Segmentation on Touch Devices

Temporally Selective Attention Model for Social and Affective State Recognition in Multimedia Content

Quality-of-Experience of Adaptive Video Streaming: Exploring the Space of Adaptations

Keynote Address 6 (Thursday, Oct 26) 02:00 PM - 03:00 PM

Bringing Gaming, VR, and AR to Life with Deep Learning
[Danny Lange: Vice President, Unity Technologies, USA]

Understanding 4 - Multimodal/Multisensor Analysis and Description (Thursday, Oct 26) 03:30 PM - 05:00 PM

Semi-Relaxation Supervised Hashing for Cross-Modal Retrieval

Cross-modal Recipe Retrieval with Rich Food Attributes

Exploring the use of Time-Dependent Cross-Network Information for Personalized Recommendations

Brave New Ideas (Thursday, Oct 26) 03:30 PM - 05:00 PM

To Create What You Tell: Generating Videos from Captions

Harnessing A.I. for Augmenting Creativity: Application to Movie Trailer Creation

Brain2Image: Converting Brain Signals into Images

Do Individuals Smile More in Diverse Social Company? Studying Smiles and Diversity Via Social Media Photos

How Personality Affects our Likes: Towards a Better Understanding of Actionable Images

Grand Challenge (Thursday, Oct 26) 04:15 PM - 06:47 PM

Knowing Yourself: Improving Video Caption via In-depth Recap

Multirate Multimodal Video Captioning

MANet: A Modal Attention Network for Describing Videos

Richer Semantic Visual and Language Representation for Video Captioning

View Count Prediction Base on Residual Learning and Random Forest for Social Media

Combining Multiple Features for Image Popularity Prediction in Social Media

Towards SMP challenge: Stacking of Diverse Models for Social Image Popularity Prediction

BUPTMM Submission to Social Media Prediction 2017

A Hybrid Model Combining Convolutional Neural Network with XGBoost for Predicting Social Media Popularity

Popularity Meter: An influence- and aesthetics-aware social media popularity predictor

Understanding 5 (Thursday, Oct 26) 05:15 PM - 06:45 PM

Video Captioning with Guidance of Multimodal Latent Topics

Learning Non-local Image Diffusion for Image Denoising

Weighted Sparse Representation Regularized Graph Learning for RGB-T Object Tracking

Novel 2 (Thursday, Oct 26) 05:15 PM - 06:45 PM

Hashtag-centric Immersive Search on Social Media

Spatio-Temporal AutoEncoder for Video Anomaly Detection

Deep Siamese Network with Multi-level Similarity Perception for Person Re-identification

Thematic Workshop Pitches - II (Friday, Oct 26) 09:00 AM - 10:30 PM

1. Learning Deep Contextual Attention Network for Narrative Photo Stream Captioning
2. Robust and Real-Time Visual Tracking with Triplet Convolutional Neural Network
3. Spatiotemporal Multi-Task Network for Human Activity Understanding
4. Real-Time Image-based Smoke Detection in Endoscopic Videos
5. Watch What You Just Said: Image Captioning with Text-Conditional Semantic Attention
6. Sensing Urban with Wi-Fi and Satellite: Functional Region Discovery across Cities
7. Reconstructable and Interpretable Representations for Time Series with Time-skip Sparse Dictionary Learning
8. Towards improving canonical correlation analysis for cross-modal retrieval
9. Variation Robust Cross-Modal Metric Learning for Caricature Recognition
10. Deep Cross-Modal Audio-Visual Generation
11. Video Imagination from a Single Image with Transformation Generation
12. Image-Based Food Calorie Estimation Using Knowledge on Food Categories, Ingredients and Cooking Directions
13. Tag Prediction at Flickr: A View from the Darkroom
14. Cross-View Gait Identification with Embedded Learning
15. CTC Network with Statistical Language Modeling for Action Sequence Recognition in Videos
16. Evolution of Trajectories: A Novel Representation for Action Recognition
17. Deep Face Recognition with Center Invariant Loss
18. Generative Attention Model with Adversarial Self-learning for Visual Question Answering
19. Learning Feature Embedding with Strong Neural Activations for Fine-Grained Retrieval
20. Image Caption with Synchronous Cross-Attention
21. Efficient Depth-aware Image Deformation Adaptation for Curved Screen Displays

22. Layout Style Modeling for Automating Banner Design
23. Learning Social Image Embedding with Deep Multimodal Attention Networks
24. Liquid Jets as Logic-Computing Fluid-User-Interfaces
25. An Experiential Model of the Atmosphere: Rich State Transitions in a Media Choreography Framework
26. Comparing user QoE of AR and VR applications using physiological and interaction measurements
27. Measuring and Improving the Viewing Experience of First-person Videos
28. An Experimental Study of Markerless Image Registration Methods on Varying Quality of Images for Augmented Reality Applications
29. Comparison of Subjective Quality Evaluation for HEVC Encoded Omnidirectional Videos at Different Bit-rates for UHD and FHD Resolution
30. Surveillance Video Quality Assessment Based on Face Recognition
31. Learning Transferable Features for Speech Emotion Recognition
32. Hierarchical Representation based on Bayesian Nonparametric Tree-Structured Mixture Model for Playing Technique Classification
33. A Smart Mirror for Music Conducting Exercises

**AltMM 2017 - 2nd International Workshop on Multimedia Alternate Realities
(Friday, Oct 26) 09:00 AM - 12:30 PM**

MUSA2 - First ACM Workshop on Multimodal Understanding of Social, Affective and Subjective Attributes (Friday, Oct 26) 09:00 AM - 12:30 PM

9h00 — Opening

9h05 — Keynote by David A. Shamma and Saide Bakhshi

9h50 — Oral Session 1: Mining

User Group based Viewpoint Recommendation using User Attributes for Multiview Videos

Beyond Concept Detection: The Potential of User Intent for Image Retrieval

10h25 — Spotlights

Image Captioning in the Wild: How people caption images on Flickr

A Deep Multi-Modal Fusion Approach for Semantic Place Prediction in Social Media

Movie Genre Classification based on Poster Images with Deep Neural Networks

Robust Multi-Modal Cues for Dyadic Human Interaction Recognition

[Coffee break ~10h30]

11h00 — Oral Session 2: Modeling

Head Pose Recommendation for Taking Good Selfies

More cat than cute? Interpretable Prediction of Adjective-Noun Pairs

11h35 — Conclusion

11h40 — Poster session (all ORAL and SPOTLIGHT papers)

MultiEdTech 2017: 1st International Workshop on Multimedia-based Educational and Knowledge Technologies for Personalized and Social Online Training (Friday, Oct 26) 09:00 AM - 12:30 PM

09:00-09:15 Opening remarks

09:15-10:30 Keynote 1, Pablo Cesar: "Sensing Engagement: Helping Performers to Evaluate their Impact"

10:30-11:00 Coffee break

11:00-12:30 Oral session

11:00-11:20 Train in Virtual Court: Basketball Tactic Training via Virtual Reality

11:20-11:40 Surgical Action Retrieval for Assisting Video Review of Laparoscopic Skills

11:40-12:00 Automatic MOOC video classification using transcript features and convolutional neural networks

12:00-12:20 Chat2Doc: From Chats to How-to Instructions, FAQ, and Report

12:20-12:30 Closing remarks and closing of the workshop

Deep reinforce learning with applications in computer games such as AlphaGo, StarCraft [Yuandong Tian (Facebook)] (Friday, Oct 26) 11:00 AM - 12:30 PM

Deep learning/CNN models [Francois Chollet (Google)] (Friday, Oct 26) 02:00 PM - 03:30 PM

[RFIW 2017: 1st Large-Scale Kinship Recognition Challenge](#) (Friday, Oct 26) 02:00 PM - 06:00 PM

[Large-Scale Video Classification Challenge](#) (LSVC 2017) (Friday, Oct 26) 02:00 PM
- 06:00 PM

MuVer'17 - First International Workshop on Multimedia Verification (Friday, Oct 26) 02:00 PM - 06:00 PM

14:00-14:05 Opening remarks

14:05-15:00 Keynote 1, Giulia Boato: "Multimedia Forensics: An Overview of Recent Advances and Open Issues"

15:00-16:00 Oral session

15:00-15:20 Collaborative Networks for Person Verification

15:20-15:40 Video Retrieval for Multimedia Verification of Breaking News on Social Networks

15:40-16:00 The InVID Plug-in: Web Video Verification on the Browser

16:00-16:30 Coffee break

16:30-17:30 Keynote 2, Gerald Friedland: "Privacy vs Multimedia Verification: A Conundrum"

17:30-17:45 Closing remarks and closing of the workshop

Medical Multimedia Information Systems [Klaus Schoeffmann (Klagenfurt University); Bernd Münzer (Klagenfurt University); Michael Riegler (Simula Research and University of Oslo); Pål Halvorsen (Simula Research and University of Oslo)] (Friday, Oct 26) 04:00 PM - 06:00 PM