BMVC 2017 LONDON

Pocket Programme Guide

British Machine Vision Conference 2017

4th-7th September 2017
Imperial College London

Platinum Sponsors

Microsoft

facebook

oculus

amazon

SCAPE

Google

Gold Sponsors

TOSHIBA
Leading Innovation

QUALCOMM

SLAMCORE

Silver Sponsors

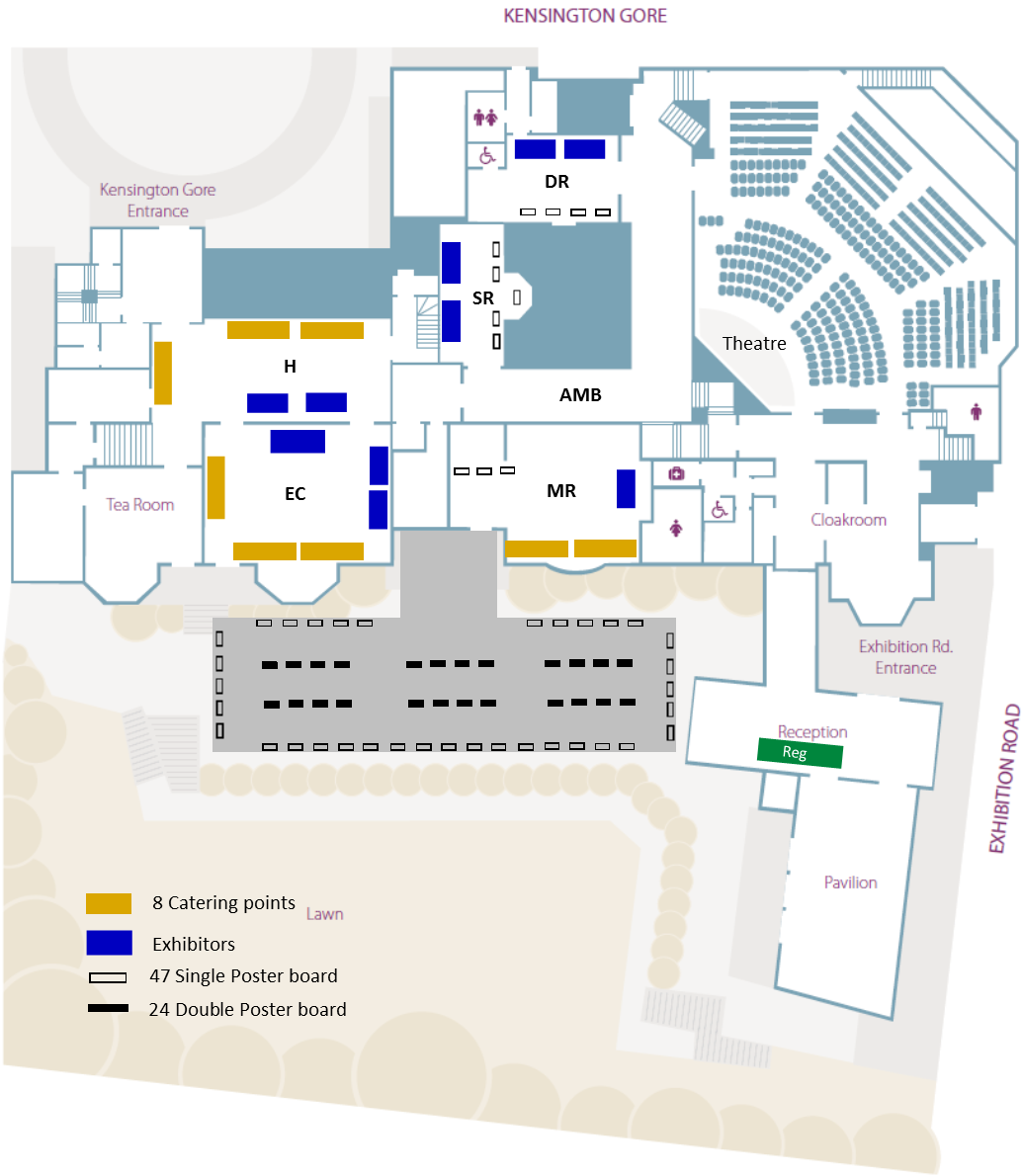
OSRAM

sengled

Points of Interest:

1. Royal Geographical Society (RGS)
2. Imperial College London (ICL)
3. Huxley Building
4. Millennium Gloucester Hotel
5. Underground

RGS Floorplan



Programme overview

	Mon 4 Sep	Tue 5 Sep	Wed 6 Sep	Thur 7 Sep
7:30		Registration RGS	Registration RGS	Registration RGS
8:00		Coffee & Pastry	Coffee & Pastry	Coffee & Pastry
8:15				
8:30		Welcome	Welcome	Welcome
8:45		Spotlights	Spotlights	Orals
9:00				
9:15				
9:30		Orals	Orals	
9:45				
10:00				
10:15				Break
10:30		Break	Break	
10:45				
11:00				Orals
11:15		Keynote	Keynote	
11:30		R. Szeliski	P. Perona	
11:45				
12:00				
12:15		Lunch	Lunch	
12:30				Lunch
12:45				
13:00		Posters	Posters	
13:15				RGS venue closed
13:30	Registration RGS			Keynote
13:45				J. Shotton
14:00				ICL Huxley Bld.
14:15				Workshops
14:30	Invited Tutorial 1			ICL Huxley Bld.
14:45	A. Geiger			
15:00		Tea break	Tea break	
15:15				
15:30	Invited Tutorial 2	Orals	Orals	
15:45	S. Whiteson			
16:00				Tea break
16:15				
16:30	Tea break			Workshops
16:45		Break		ICL Huxley Bld.
17:00	Invited Tutorial 3			
17:15	V. Lepetit	Orals		
17:30			RGS venue closed	
17:45				
18:00	Invited Tutorial 4			
18:15	L. Agapito			
18:30				
18:45				
19:00	Reception		Banquet	
19:15	RGS		Millennium Gloucester	
19:30			Hotel	
19:45				
20:00				
20:15				
20:30				
20:45				

Welcome Message from The General and Programme Chairs

It is our great pleasure to welcome you to London for the British Machine Vision Conference (BMVC)! This is the 28th BMVC since its inception in 1990, and it's the second time in London. The conference is hosted at the Royal Geographical Society of London (RGS). RGS was founded in 1830, overlooking Imperial College London, Kensington Gardens, and the Royal Albert Hall. It has been a key associate and supporter of many explorers including Darwin, Livingstone, Stanley, Scott, Shackleton, Hunt and Hillary.

BMVC is one of the top events in the Computer Vision conference calendar, and a truly international event, with the majority of papers coming from outside the UK. This year, BMVC attracted a total of 635 full paper submissions, which is the highest number in the history of BMVC.
This year, we introduced a few changes to the reviewing process. Despite a very tight schedule bounded by ICCV deadlines from one side and UK immigration visas from the other, we ran a full reviewing process that included bidding for papers by ACs and reviewers, as well as reviewer suggestions by ACs, author rebuttals, and reviewer and AC discussions with the help of CMT and TPMS systems. There were 550 reviewers and 41 area chairs involved in this process, generously donating their time. The Program Chairs made an effort to assign papers on similar topics to the same set of Area Chairs (ACs), so that related papers could be compared directly. To achieve this, 70 papers were pre-assigned to each AC with the help of the Toronto Paper Matching System. ACs then went over the paper titles and abstracts and bid on those they would like to handle. Using all of this information as input, ACs received a final assignment of around 30 papers, for which they provided ranked lists of candidate reviewers. ACs were asked to suggest candidate reviewers who were experts in the topic, had a broad view of the field, were reliable, and could provide high quality reviews. Their choices were additionally guided by the reviewers' bids (the process followed for reviewer bidding was similar to that used for ACs). Oral papers were selected based on the reviews, AC consolidation reports, ranking in the ACs' batches, and suitability of the content for a general audience. We would like to thank all the reviewers and area chairs for their hard work and prompt responses!

Of the 635 submissions, just 188 were accepted for presentation in BMVC 2017, which is a 30% acceptance rate. Of the accepted papers, 36 were accepted as oral for a podium spot. We also introduced spotlight sessions to promote 20 further outstanding papers. The accepted papers represent a truly international research community, with 15% of the papers from the UK, 24% from Europe excluding the UK, 34% from Asia, 27% from North America, and 1 paper from Egypt. As is now standard for many top conferences, the proceedings are published entirely online.

We have put together an interesting programme and are delighted to welcome Richard Szeliski and Pietro Perona to the conference as keynote speakers, Lourdes Agapito, Shimon Whiteson, Andreas Geiger and Vincent Lepetit as tutorial speakers, as well as Jamie Shotton as workshops keynote speaker.
BMVC has always had strong links with industry, and again we are very grateful to our industrial sponsors for supporting the event. Platinum Sponsors: Microsoft Research, Facebook Oculus, Amazon, SCAPE, Snap and Google. Gold Sponsors: Toshiba,

Qualcomm, Intel, Twitter, and Slamcore. Silver Sponsors: Cosmonio, Osram, Ocado, Sengled, Toyota Research Institute, IET The Institution of Engineering and Technology, and Adobe.

Last but not least, we wish to thank all members of the Organizing Committee, the Area Chairs, reviewers, emergency reviewers, authors, and the CMT and TPMS teams for the immense amount of hard work and professionalism that has gone into making BMVC 2017 a first rate conference.
We hope you find BMVC 2017 in London both an enjoyable and productive experience.

Tae-Kyun Kim, Stefanos Zafeiriou, Gabriel Brostow and Krystian Mikolajczyk
BMVC 2017 organising committee

General Chair	Tae-Kyun Kim
	Stefanos Zafeiriou
Programme Chair	Gabriel Brostow
	Krystian Mikolajczyk
Advisory Board	Andrew Davison
	Maja Pantic
	Yannis Demiris
Tutorials/Workshops	Yannis Panagakis
	Tania Stathaki
Sponsorship/Demo	Ben Glocker
	Stefan Leutenegger
Local Arrangements	Guillermo Garcia-Hernando
	Athanasios Papaioannou
	Rigas Kouskouridas
	Anastasios Roussos
	Vassileios Balntas
Publicity	Giovanni Maria Farinella
Website	Athanasios Papaioannou
	James Booth
	Patrick Snape

Area Chairs

Lourdes Agapito	University College London
Antonis Argyros	University of Crete
Richard Bowden	University of Surrey
Edmond Boyer	INRIA Rhône Alpes
Toby Breckon	Durham University
Neill Campbell	University of Bath
Ondrej Chum	Czech Technical University
John Collomosse	University of Surrey
Tim Cootes	University of Manchester
Daniel Cremers	Technical University of Munich
Dima Damen	University of Bristol
Alessio Del Bue	Istituto Italiano di Tecnologia
Mario Fritz	Max-Planck-Institut
Shaogang Gong	Queen Mary University of London
Peter Hall	University of Bath
Edwin Hancock	University of York
Adrian Hilton	University of Surrey
Timothy Hospedales	University of Edinburgh
Frederic Jurie	University of Caen
Joni Kamarainen	Tampere University of Technology
Iasonas Kokkinos	Facebook Inc.
Nikos Komodakis	Ecole des Ponts ParisTech
M. Pawan Kumar	University of Oxford
Kyoung Mu Lee	Seoul National University
Bastian Leibe	RWTH Aachen University
Ales Leonardis	University of Birmingham
Vincent Lepetit	University of Bordeaux
Majid Mirmehdi	Bristol University
Francesc Moreno-Noguer	Polytechnic University of Catalonia
Patrick Perez	Technicolor
Yoichi Sato	University of Tokyo
Ling Shao	University of East Anglia
William Smith	University of York
Cees Snoek	University of Amsterdam
Tinne Tuytelaars	Katholieke Universiteit Leuven
Andrea Vedaldi	University of Oxford
Richard Wilson	University of York
Tony Xiang	Queen Mary University of London
Xianghua Xie	Swansea University
Fei Yan	University of Surrey
Jianguo Zhang	University of Dundee

Outstanding Reviewers

High reviewing load, useful reviews, delivered by the deadline.

Alexander Andreopoulos
Ognjen Arandjelovic
Chetan Arora
Artem Babenko
Andrew Bagdanov
Vassilis Balntas
Christos Bampis
Adrian Barbu
Ardhendu Behera
Vasileios Belagiannis
Guillaume-Alexandre Bilodeau
Horst Bischof
Terrance Boult
Michael Breuss
Michael Brown
Andrei Bursuc
Joao Carreira
David Chan
Hyung Jin Chang
HyungJinÊ ChangÊ
Xinlei Chen
Ke Chen
Tat-Jun Chin
Wongun Choi
Salil Deena
Jia Deng
Joachim Denzler
Francisco Escolano Ruiz
Ryan Farrell
Andrew French
Efstratios Gavves
David Geronimo
Andrea Giachetti
Ioannis Gkioulekas
Guy Godin
Boqing Gong
Philippe-Henri Gosselin
Saurabh Gupta
Ankush Gupta
Bohyung Han
Tatsuya Harada
Lei He
Shengfeng He
Seunghoon Hong
Heikki Huttunen
Go Irie
Suyog Jain
Hao Jiang
Michael Jones
Margret Keuper
Akisato Kimura
Josip Krapac
Shiro Kumano
Kaustav Kundu
Junseok Kwon
Adrian Lopez Rodriguez
Kevin Lai
Jochen Lang
Oswald Lanz
Hongyang Li
Stephen Lin
James Little
Stephan Liwicki
Roberto Lopez-Sastre
Chen-Change Loy
Le Lu
Jiwen Lu
Subhransu Maji
Yasushi Makihara
Clement Mallet
Ioannis Marras
Kevin Matzen
Frank Michel
Ishan Misra
Damien Muselet
Magnus Oskarsson
Mustafa Ozuysal
Federico Pernici
Andrea Prati
Gang Qian
Peter Roth
Christian Rupprecht
Chris Russell
Mathieu Salzmann
Jorge Sanchez
Conrad Sanderson
Shin'ichi Satoh
Alexander Schwing
Gaurav Sharma
Pramod Sharma
Richa Singh
Greg Slabaugh
Yi-zhe Song
Ran Song
Michael Stark
Justin Strait
Hao Su
Akihiro Sugimoto
Jean-Philippe Tarel
Diego Thomas
Radu Timofte
Giorgos Tolias
Mercedes Torres Torres
Du Tran
Stavros Tsogkas
Ernest Valveny
Jan van Gemert
Shanmukha Vedantam
Catherine Wah
Yang Wang
Lei Wang
Longyin Wen
Tianfu Wu
Jiajun Wu
Song Wu
Yu Xiang
Junchi Yan
Angela Yao
Runze Zhang
Bin Zhao
Zihan Zhou
Jun-yan Zhu
Siyu Zhu
Zeeshan Zia

13:30 – Registration, RGS

14:30 – 16:30 Tutorials, RGS Theatre

Chair: Stefan Leutenegger and Stefanos Zafeiriou

14^{30} Probabilistic and Deep Models for 3D Reconstruction

Andreas Geiger (Max Planck Institute)

15^{30} Intro to Reinforcement Learning

Shimon Whiteson (University of Oxford)

16:30 – 17:00 Break, RGS

17:00 – 19:00 Tutorials, RGS Theatre

Chair: Andrew Davison and Tae-Kyun Kim

17^{00} Deep Learning for 3D Localization

Vincent Lepetit (University of Bordeaux)

18^{00} 3D Reconstruction of Dynamic Scenes from Monocular Video

Lourdes Agapito (University College London)

19:00 – 21:00 Reception, RGS

7:30 – Registration, RGS

8:00 – 8:30 Coffee & pastry, RGS

8:30 – 8:45 Welcome, RGS Theatre

8:45 – 9:30 Spotlights, RGS Theatre

8^{45} 1. Semantics-Preserving Locality Embedding for Zero-Shot Learning

Shih-Yen Tao (Academia Sinica), Yi-Ren Yeh and Yu-Chiang Frank Wang

8^{49} 2. Hyperspectral CNN Classification with Limited Training Samples

Lloyd Windrim (University of Sydney), Rishi Ramakrishnan Arman Melkumyan and Richard Murphy

8^{53} 3. Which is the better inpainted image? Learning without subjective annotation

Mariko Isogawa (NTT), Dan Mikami Kosuke Takahashi and Hideaki Kimata

8^{57} 4. Spatio-Temporal Consistency to Detect and Segment Carried Objects

Farnoosh Ghadiri (Laval University), Robert Bergevin and Guillaume-Alexandre Bilodeau

9^{01} 5. Improved Image Segmentation via Cost Minimization of Multiple Hypotheses

Marc Bosch Ruiz (The Johns Hopkins University), Christopher Gifford Austin Dress Clare Lau and Jeffrey Skibo

9^{05} 6. Face Parsing via Recurrent Propagation

Sifei Liu (UCMERCED), Jianping Shi Liang Ji and Ming-Hsuan Yang

9^{09} 7. Accurate Camera Registration in Urban Environments Using High-Level Feature Matching

Anil Armagan (TU GRAZ), Martin Hirzer Peter Roth and Vincent Lepetit

9^{13} 8. The Devil is in the Decoder

Zbigniew Wojna (University College London), Jasper Uijlings Sergio Guadarrama Nathan Silberman Liang-Chieh Chen Alireza Fathi and Vittorio Ferrari

9^{17} 9. Virtual to Real Reinforcement Learning for Autonomous Driving

Xinlei Pan (UC Berkeley), Yurong You Ziyan Wang and Cewu Lu

9^{21} 10. Real-Time Salient Closed Boundary Tracking using Perceptual Grouping and Shape Priors

Xuebin Qin (University of Alberta), Shida He Zichen Zhang Masood Dehghan and Martin Jagersand

9:30 – 10:30 Orals: Pose Estimation, RGS Theatre

Chair: Roberto Cipolla

9^{30} 11. Kinematic-Layout-aware Random Forests for Depth-based Action Recognition

Seungryul Baek (Imperial College London), Zhiyuan Shi Masato Kawade and Tae-Kyun Kim

9^{45} 12. Total Capture: 3D Human Pose Estimation Fusing Video and Inertial Sensors

Matthew Trumble (University of Surrey), Andrew Gilbert Charles Malleson Adrian Hilton and John Collomosse

10^{00} 13. Indirect deep structured learning for 3D human body shape and pose prediction

Vince Tan (University of Cambridge), Ignas Budvytis and Roberto Cipolla

10^{15} 14. Real-time Visual-Inertial Odometry for Event Cameras using Keyframe-based Nonlinear Optimization

Henri Rebecq (University of Zurich), Timo Horstschaefer and Davide Scaramuzza

10:30 – 11:15 Break, RGS

11:15 – 12:15 Keynote, RGS Theatre

Chair: Krystian Mikolajczyk

11^{15} **Visual Reconstruction and Image-Based Rendering**

Richard Szeliski (Facebook)

12:15 – 13:00 Lunch, RGS

13:00 – 15:00 Posters, RGS

15. Weakly Supervised Semantic Segmentation Based on Co-segmentation

Tong Shen (The University of Adelaide), Guosheng Lin Lingqiao Liu Chunhua Shen and Ian Reid

16. Semantic Segmentation with Reverse Attention

Qin Huang (University of Southern Califor), Chihao Wu Chunyang Xia Ye Wang and C.-C. Jay Kuo

17. Double Expansion for Optimization of Multilabel Energies

Yelena Gorelick (Western University), Zhengqin Li and Olga Veksler

18. Discovering Class-Specific Pixels for Weakly-Supervised Semantic Segmentation

Arslan Chaudhry (University of Oxford), Puneet Kumar Dokania and Philip Torr

19. Adapting Models to Signal Degradation using Distillation

Jong-Chyi Su (UMass-Amherst), and Subhransu Maji

20. Localizing Actions from Video Labels and Pseudo-Annotations

Pascal Mettes (University of Amsterdam), Cees Snoek and Shih-Fu Chang

21. Introduction to Coherent Depth Fields for Dense Monocular Surface Recovery

Vladislav Golyanik (DFKI), Torben Fetzer and Didier Stricker

22. Real-time Factored ConvNets: Extracting the X Factor in Human Parsing

James Charles (University of Cambridge), Ignas Budvytis and Roberto Cipolla

23. Holistic, Instance-level Human Parsing

Qizhu Li (University of Oxford), Anurag Arnab and Philip Torr

24. CT from Motion: Volumetric Capture of Moving Shapes using X-rays

Julien Pansiot (Inria), and Edmond Boyer

25. Concise Radiometric Calibration Using The Power of Ranking

Han Gong (University of East Anglia), Graham Finlayson and Maryam Darrodi

26. Scale Exploiting Minimal Solvers for Relative Pose with Calibrated Cameras

Stephan Liwicki (Toshiba Research Europe Ltd), and Christopher Zach

27. Multiple Instance Curriculum Learning for Weakly Supervised Object Detection

Siyang Li (USC), Xiangxin Zhu Qin Huang Hao Xu and C.-C. Jay Kuo

28. Fast Feature Fool: A data independent approach to universal adversarial perturbations

Konda Reddy Mopuri (Indian Institute of Science), Utsav Garg and Venkatesh Babu Radhakrishnan

29. 3D color charts for camera spectral sensitivity estimation

Rada Deeb () Damien Muselet Mathieu Hebert Alain Tremeau and Joost van de Weijer

30. Guided Robust Matte-Model Fitting for Accelerating Multi-light Reflectance Processing Techniques

Ruggero Pintus (CRS4), Andrea Giachetti Gianni Pintore and Enrico Gobbetti

31. Fast Event-based Corner Detection

Elias Mueggler (University of Zurich), Chiara Bartolozzi and Davide Scaramuzza

32. PCN: Part and Context Information for Pedestrian Detection with CNNs

Shiguang Wang, UESTC

33. Adaptive Local Contrast Normalization for Robust Object Detection and Pose Estimation

Mahdi Rad (TUG), Vincent Lepetit and Peter Roth

34. Object-Extent Pooling for Weakly Supervised Single-Shot Localization

Amogh Gudi (TU Delft), Nicolai van Rosmalen Marco Loog and Jan van Gemert

35. Transformed Anti-Sparse Hashing

Zhangyang Wang (TAMU), Ji Liu Shuai Huang Xinchao Wang and Shiyu Chang

36. Deeply Supervised 3D Recurrent FCN for Salient Object Detection in Videos

Trung-Nghia Le (NII), and Akihiro Sugimoto

37. Robust Dense Depth Maps Generations from Sparse DVS Stereos

Dongqing Zou (Samsung Electronic), Feng Shi Weiheng Liu Jia Li Qiang Wang Paul-K.J. Park and Hyunsurk Eric Ryu

38. Adaptive Temporal Pooling for Object Detection using Dynamic Vision Sensor

Jia Li (Samsung - SAIT China Lab), Feng Shi Weiheng Liu Dongqing Zou Qiang Wang Paul-K.J. Park and Hyunsurk Eric Ryu

39. High-Resolution Stereo Matching based on Sampled Photoconsistency Computation

Chloe Legendre (usc), Konstantinos Batsos and Philippos Mordohai

40. Generative OpenMax for Multi-Class Open Set Classification

Zongyuan Ge (IBM), Sergey Demyanov and Rahil Garnavi

41. ANSAC: Adaptive Non-Minimal Sample and Consensus

Victor Fragoso (West Virginia University), Christopher Sweeney Pradeep Sen and Matthew Turk

42. Color Restoration of Underwater Images

Dana Menaker (Tel Aviv University), Tali Treibitz and Shai Avidan

43. Fine-Grained Image Retrieval: the Text/Sketch Input Dilemma

Jifei Song (Queen Mary), Yi-zhe Song Tony Xiang and Timothy Hospedales

44. Cross-domain Generative Learning for Fine-Grained Sketch-Based Image Retrieval
Kaiyue Pang (QMUL), Yi-zhe Song Tony Xiang and Timothy Hospedales

45. Epipolar Plane Diffusion: An Efficient Approach for Light Field Editing

Oriel Frigo, INRIA

46. Group Cost-sensitive Boosting with Multi-scale Decorrelated Filters for Pedestrian Detection
Chengju Zhou (NTU), Meiqing Wu and SiewKei Lam

47. Moving Object Segmentation in Jittery Videos by Stabilizing Trajectories Modeled in Kendall's Shape Space

Geethu Jacob (Indian Institute of Technology), and Sukhendu Das

48. Urban Image Stitching using Planar Perspective Guidance

Joo Ho Lee (KAIST), Seung-Hwan Baek and Min H. Kim

49. Detecting Parts for Action Localization

Nicolas Chesneau (Inria), Gregory Rogez Karteek Alahari and Cordelia Schmid

50. Cascaded Boundary Regression for Temporal Action Detection

Jiyang Gao (USC), Zhenheng Yang and Ram Nevatia

51. Confidence and Diversity for Active Selection of Feedback in Image Retrieval
Bhavin Modi and Adriana Kovashka

52. Shape Generation using Spatially Partitioned Point Clouds

Matheus Gadelha (University of Massachusetts Am), Subhransu Maji and Rui Wang

53. Person Re-Identification by Localizing Discriminative Regions

Tanzila Rahman (University of Manitoba), Mrigank Rochan and Yang Wang

54. Adapting Object Detectors from Images to Weakly Labeled Videos

Omit Chanda (University of Manitoba), Eu Wern Teh Mrigank Rochan Zhenyu Guo and Yang Wang

55. Bayesian SegNet: Model Uncertainty in Deep Convolutional Encoder-Decoder Architectures for Scene Understanding

Alex Kendall (University of Cambridge), Vijay Badrinarayanan and Roberto Cipolla

56. DepthComp: Real-time Depth Image Completion Based on Prior Semantic Scene Segmentation

Amir Atapour Abarghouei (Durham University), and Toby Breckon

57. DEEP eye contact detector: Robust eye contact bid detection using convolutional neural network

Yu Mitsuzumi (Kyoto University), Atsushi Nakazawa and Toyoaki Nishida

58. BiSeg: Simultaneous Instance Segmentation and Semantic Segmentation with Fully Convolutional Networks

Viet Pham (Toshiba), Satoshi Ito and Tatsuo Kozakaya

59. Salient Object Detection using a Context-Aware Refinement Network

Md Amirul Islam (University of Manitoba), Mahmoud Kalash Mrigank Rochan Neil Bruce and Yang Wang

60. Visual Odometry with Drift-Free Rotation Estimation Using Indoor Scene Regularities

Pyojin Kim (Seoul National University), Brian Coltin and Hyounjin Kim

61. One For All: Adaptive Learning-based Temporal Tracker for 3D Head Shape Models

David Joseph Tan (Technische Universität München), Federico Tombari and Nassir Navab

62. Dominant Set Clustering and Pooling for Multi-View 3D Object Recognition.

Chu Wang (McGill University), Marcello Pelillo and Kaleem Siddiqi

63. A Generalised Formulation for Collaborative Representation of Image Patches (GP-CRC)

Tapabrata Chakraborti (University of Otago), Brendan McCane Steven Mills and Umapada Pal

64. DeepShoe: A Multi-Task View-Invariant CNN for Street-to-Shop Shoe Retrieval

Huijing Zhan (NTU EEE), boxin shi and ALEX KOT

65. Weakly Supervised Saliency Detection with A Category-Driven Map Generator

Kuang-Jui Hsu (National Taiwan University), Yen-Yu Lin and Yung-Yu Chuang

66. Saliency Detection by Compactness Diffusion

Qi Zheng (Huazhong Univ. of Sci. & Tech.), Peng Zhang and Xinge You

67. Weakly-supervised Learning of Mid-level Features for Pedestrian Attribute Recognition and Localization

Yang Zhou (Beijing University of Chemical), Kai Yu Biao Leng zhang Zhang Dangwei Li and Kaiqi Huang

68. AST-Net: An Attribute-based Siamese Temporal Network for Real-Time Emotion Recognition

Shu-Hui Wang (National Tsing Hua University), and Chiou-Ting Hsu

69. Combined Internal and External Category-Specific Image Denoising

Saeed Anwar (ANU), Cong Huynh and Fatih Porikli

70. Temporal Perceptive Network for Skeleton-Based Action Recognition

Yueyu Hu (Peking University), Chunhui Liu Yanghao Li and Jiaying Liu

71. Detecting Semantic Parts on Partially Occluded Objects

Jianyu Wang Zhishuai Zhang Cihang Xie Jun Zhu Lingxi Xie and Alan Yuille

72. Place Recognition in Semi-Dense Maps

Yawei Ye (ETH Zurich), Titus Cieslewski Antonio Loquercio and Davide Scaramuzza

73. Image Completion with Intrinsic Reflectance Guidance

Soomin Kim (KAIST), Taeyoung Kim Min H. Kim and Sung-Eui Yoon

74. Enhancement of SSD by concatenating feature maps for object detection

Jisoo Jeong (Seoul National University), Hyojin Park and Nojun Kwak

75. Colour Constancy: Biologically-inspired Contrast Variant Pooling Mechanism

Arash Akbarinia (Centre de Visió per Computador), Raquel Gil Rodríguez and C. Alejandro Parraga

76. Superpixel-based semantic segmentation trained by statistical process control

Hyojin Park (Seoul National University), Jisoo Jeong Youngjoon Yoo and Nojun Kwak

77. Robust Pixel-wise Dehazing Algorithm based on Advanced Haze-Relevant Features

Guisik Kim (CAU), and Junseok Kwon

78. General Deep Image Completion with Lightweight Conditional Generative Adversarial Networks

Ching-Wei Tseng (National Tsing Hua University), Hung Jin Lin and Shang-Hong Lai

79. Augmented Reality meets Deep Learning

Hassan Abu Alhaija (TU Dresden), Siva Karthik Mustikovela Lars Mescheder Andreas Geiger and Carsten Rother

80. Improved IR-Colorization using Adversarial Training and Estuary Networks

Matthias Limmer (Daimler AG), and Hendrik Lensch

81. Realtime Novel View Synthesis with Eigen-Texture Regression

Yuta Nakashima (Osaka University), Fumio Okura Norihiko Kawai ryosuke Kimura Hiroshi Kawasaki Katsushi Ikeuchi and Ambrosio Blanco

82. Human Action Recognition Using A Multi-Modal Hybrid Deep Learning Model

Hany El-Ghaish (Egypt-Japan University(EJUST)), Mohamed Hussein and Amin Shoukry

83. Probabilistic Image Colorization

Amelie Royer Alexander Kolesnikov and Christoph Lampert

84. Iterated Lifting for Robust Cost Optimization

Christopher Zach (Toshiba Cambridge Research Laboratory), and Guillaume Bourmaud

15:00 – 15:30 Tea break, RGS

15:30 – 16:45 Orals: Scene Understanding, RGS Theatre

Chair: Pushmeet Kohli

15^{30} 85. Slanted Stixels: Representing San Francisco's Steepest Streets

Daniel Hernandez-Juarez (UAB), Lukas Schneider Antonio Espinosa Juan Moure David Vazquez Antonio López Uwe Franke and Marc Pollefeys

15^{45} 86. Deformable Part-based Fully Convolutional Network for Object Detection

Taylor Mordan (LIP6), Nicolas Thome Gilles Henaff and Matthieu Cord

16^{00} 87. Multiple Instance Visual-Semantic Embedding

Zhou Ren (Snap Inc.), Hailin Jin Zhe Lin Chen Fang and Alan Yuille

16^{15} 88. Thinking Outside the Box: Spatial Anticipation of Semantic Categories

Martin Garbade (Uni Bonn), and Juergen Gall

16^{30} 89. Visually Aligned Word Embeddings for Improving Zero-shot Learning

Ruizhi Qiao (The University of Adelaide), Lingqiao Liu Chunhua Shen and Anton van den Hengel

16:45 – 17:15 Break, RGS

17:15 – 18:30 Orals: Action Recognition, RGS Theatre

Chair: Rahul Sukthankar

17^{15} 90. Real-Time Temporal Action Localization in Untrimmed Videos by Sub-Action Discovery

Rui Hou (University of Central Florida), Rahul Sukthankar and Mubarak Shah

17^{30} 91. RED: Reinforced Encoder-Decoder Networks for Action Anticipation

Jiyang Gao (USC), Zhenheng Yang and Ram Nevatia

17^{45} 92. End-to-End, Single-Stream Temporal Action Detection in Untrimmed Videos

Shyamal Buch (Stanford University), Victor Escorcia Bernard Ghanem and Juan Carlos Niebles

18^{00} 93. Recognizing and Curating Photo Albums via Event-Specific Image Importance

Yufei Wang (UCSD), Zhe Lin Xiaohui Shen Radomir Mech Gavin Miller and Garrison Cottrell

18^{15} 94. Spatio-Temporal Action Detection with Cascade Proposal and Location Anticipation

Zhenheng Yang (USC), Jiyang Gao and Ram Nevatia

7:30 – Registration, RGS

8:00 – 8:30 Coffee & pastry, RGS

8:30 – 8:45 Welcome, RGS Theatre

8:45 – 9:30 Spotlights, RGS Theatre

8^{45} 1. Video Segmentation with Background Motion Models

Scott Wehrwein (Cornell University), and Richard Szeliski

8^{49} 2. Orientation-boosted Voxel Nets for 3D Object Recognition

Nima Sedaghat (University of Freiburg), Mohammadreza Zolfaghari Ehsan Amiri and Thomas Brox

8^{53} 3. Semantic 3D Reconstruction with Finite Element Bases

Audrey Richard (ETH Zurich), Christoph Vogel Maros Blaha Thomas Pock and Konrad Schindler

8^{57} 4. SilNet : Single- and Multi-View Reconstruction by Learning from Silhouettes

Olivia Wiles (Oxford University), and Andrew Zisserman

9^{01} 5. A Differential Approach to Shape from Polarization

Roberto Mecca (University of Cambridge), Fotios Logothetis and Roberto Cipolla

9^{05} 6. Fast dense feature extraction with convolutional neural networks that have pooling or striding layers

Christian Bailer (DFKI), tewodros Habtegebrial Kiran Varanasi and Didier Stricker

9^{09} 7. Local Visual Microphones: Improved Sound Extraction from Silent Video

Mohammad Amin Sadeghi (University of Tehran), Mohammad Amin Shabani and Laleh Samadfam

9^{13} 8. Unsupervised Deep Generative Hashing

Yuming Shen (University of East Anglia), li Liu and Ling Shao

9^{17} 9. Sampled Image Tagging and Retrieval Methods on User Generated Content

Karl Ni (In-Q-Tel), Kyle Zaragoza Alexander Gude Yonas Tesfaye Carmen Carrano Charles Foster and Barry Chen

9^{21} 10. Criteria Sliders: Learning Continuous Database Criteria via Interactive Ranking

James Tompkin (Brown University), Kwang In Kim Hanspeter Pfister and Christian Theobalt

9:30 – 10:30 Orals: Face Analysis, RGS Theatre

Chair: Josef Kittler

9^{30} 11. Now You See Me: Deep Face Hallucination for Unviewed Sketches

Conghui Hu (Queen Mary Uni. of Londo), Da Li Yi-zhe Song and Timothy Hospedales

9^{45} 12. From Benedict Cumberbatch to Sherlock Holmes: Character Identification in TV series without a Script

Arsha Nagrani (Oxford University), and Andrew Zisserman

10^{00} 13. Quantifying Facial Age by Posterior of Age Comparisons

Yunxuan Zhang (Sensetime), Li Liu Cheng Li and Chen-Change Loy

10^{15} 14. You said that?

Joon Son Chung (University of Oxford), Amir Jamaludin and Andrew Zisserman

10:30 – 11:15 Break, RGS

11:15 – 12:15 Keynote, RGS Theatre

Chair: Gabriel Brostow

11^{15} **Visipedia - A universal visual expert**

Pietro Perona (California Institute of Technology)

12:15 – 13:00 Lunch, RGS

13:00 – 15:00 Posters, RGS

15. Reflectance and Shape Estimation with a Light Field Camera under Natural Illumination

Trung Ngo (Osaka University), Hajime Nagahara Ko Nishino Rin-ichiro Taniguchi and Yasushi Yagi

16. GeneGAN: Learning Object Transfiguration and Object Subspace from Unpaired Data

Shuchang Zhou (Megvii Inc.), Taihong Xiao Yi Yang Dieqiao Feng Qinyao He and Weiran He

17. PixColor: Pixel Recursive Colorization

Sergio Guadarrama (Google), Ryan Dahl David Bieber Jonathon Shlens Mohammad Norouzi and Kevin Murphy

18. Depth Estimation and Blur Removal from a Single Out-of-focus Image

Saeed Anwar (ANU), Zeeshan Hayder and Fatih Porikli

19. Exploring the structure of a real-time, arbitrary neural artistic stylization network

Golnaz Ghiasi (Google Inc.), Honglak Lee Manjunath Kudlur Vincent Dumoulin and Jonathon Shlens

20. Fine-Pruning: Joint Fine-Tuning and Compression of a Convolutional Network with Bayesian Optimization

Frederick Tung (Simon Fraser University), Srikanth Muralidharan and Greg Mori

21. Online Adaptation of Convolutional Neural Networks for Video Object Segmentation

Paul Voigtlaender (RWTH Aachen University), and Bastian Leibe

22. Improved Bilinear Pooling with CNNs

Tsung-Yu Lin (UMass Amherst), and Subhransu Maji

23. Video to Text Summary: Joint Video Summarization and Captioning with Recurrent Neural Networks

Bor-Chun Chen (University of Maryland), Yan-Ying Chen and Francine Chen

24. A Recurrent Variational Autoencoder for Human Motion Synthesis

Ikhsanul Habibie (University of Edinburgh), Daniel Holden Jonathan Schwarz Joe Yearsley and Taku Komura

25. A deep learning pipeline for semantic facade segmentation

Radwa Fathalla (AAST), and George Vogiatzis

26. Deep Reinforcement Learning Attention Selection For Person Re-Identification

XU LAN (Queen Mary University of Londo), HangXiao Wang Shaogang Gong and Xiatian Zhu

27. Case-Based Histopathological Malignancy Diagnosis using Convolutional Neural Networks

Qicheng Lao (Concordia University), and Thomas Fevens

28. Labelless Scene Classification with Semantic Matching

Meng Ye (Temple University), and Yuhong Guo

29. Patch-based Interferometric Phase Estimation via Mixture of Gaussian Density Modelling & Non-local Averaging in the Complex Domain

Joshin Krishnan (Instituto de Telecomunicações), and José Bioucas-Dias

30. Feature Sequence Representation Via Slow Feature Analysis For Action Classification

Takumi Kobayashi, National Institute of AIST

31. A Convolutional Temporal Encoder for Video Caption Generation

Qingle Huang (Zhejiang University), and Zicheng Liao

32. Marginalized CNN: Learning Deep Invariant Representations

Jian ZHAO (NUS), Jianshu Li Fang Zhao Xuecheng Nie Yunpeng Chen Shuicheng Yan and Jiashi Feng

33. Correlation Hashing Network for Efficient Cross-Modal Retrieval

Yue Cao (Tsinghua University), Mingsheng Long and Jianmin Wang

34. Sunrise or Sunset: Selective Comparison Learning for Subtle Attribute Recognition

Hong-Yu Zhou (Nanjing University), Bin-Bin Gao and Jianxin Wu

35. BIM: Ball Intersection Multi Template Matching

Bat El Shlomo, Tel Aviv University

36. Visual Textbook Network: Watch Carefully before Answering Visual Questions

Difei Gao (ICT), Ruiping Wang Shiguang Shan and Xilin Chen

37. Key Person Aided Re-identification in Partially Ordered Pedestrian Set

Chen Chen (Chinese academy of sciences), Min Cao and Silong Peng

38. Learning confidence measures in the wild

Fabio Tosi (University of Bologna), Matteo Poggi Stefano Mattoccia Alessio Tonioni and Luigi Di Stefano

39. Deep View-Sensitive Pedestrian Attribute Inference in an end-to-end Model

M. Saquib Sarfraz (KIT), Arne Schumann Yan Wang and Rainer Stiefelhagen

40. Divide and Fuse: A Re-ranking Approach for Person Re-identification

Rui Yu (Huazhong University of Science and Technology), Zhichao Zhou Song Bai and Xiang Bai

41. Joint Optimization of Coded Illumination and Grayscale Conversion for One-Shot Raw Material Classification

Chao Wang (Kyushu Institute of Technoloigy), and Takahiro Okabe

42. Hand Pose Learning: Combining Deep Learning and Hierarchical Refinement for 3D Hand Pose Estimation

Min-Yu Wu (National Taiwan University), Ya Hui Tang, Pai-Wei Ting and Li-Chen Fu

43. Learning Cross-scale Correspondence and Patch-based Synthesis for Reference-based Super-Resolution

Haitian Zheng (HKUST), Mengqi Ji Lei Han Ziwei Xu Haoqian Wang Yebin Liu and Lu Fang

44. Anomaly Detection using a Convolutional Winner-Take-All Autoencoder

Hanh Tran (University of Leeds), and David Hogg

45. A Parallel Architecture for High Frame Rate Stereo using Semi-Global Matching

Akshay Jain (IIIT Delhi), Alexander Fell and Saket Anand

46. Accelerating Computation of Exemplar-SVM by Binary Approximation based on Matrix Decomposition

Takato Kurokawa (Chubu University), Yuji Yamauchi Mitsuru Ambai Takayoshi YAMASHITA and Hironobu Fujiyoshi

47. Sparse and Noisy to Dense Depth Map Upsampling Based on Mesh and Colour Consistency

Hanshin Lim (Electronics and Telecommunications Research Institute), and Junseok Lee

48. Efficient 3D Tracking in Urban Environments with Semantic Segmentation

Martin Hirzer (Graz University of Technology), Peter Roth and Vincent Lepetit

49. Class Weighted Convolutional Features for Visual Instance Search

Albert Jimenez Jose Alvarez and Xavier Giro-i-Nieto

50. Dynamic Steerable Blocks in Deep Residual Networks

Jörn-Henrik Jacobsen (University of Amsterdam), Bert De Brabandere and Arnold Smeulders

51. Towards Complete Scene Reconstruction from Single-View Depth and Human Motion

Sam Fowler (University of Surrey), Hansung Kim and Adrian Hilton

52. Flow Based Video Super-Resolution with Spatio-temporal Patch Similarity

Joan Duran (Universitat Illes Balears), and Antoni Buades

53. BV-CNNs: Binary Volumetric Convolutional Networks for 3D Object Recognition

Chao Ma (NUDT), Wei An Yinjie Lei and Yulan Guo

54. Combining Edge Images and Depth Maps for Robust Visual Odometry

Fabian Schenk (TU Graz), and Friedrich Fraundorfer

55. Order-Adaptive and Illumination-Aware Variational Optical Flow Refinement

Daniel Maurer (University of Stuttgart), Michael Stoll and Andres Bruhn

56. Large-scale Continual Road Inspection: Visual Infrastructure Assessment in the Wild

Ke Ma (Stony Brook University), Minh Hoai and Dimitris Samaras

57. Correlation Filter Tracking: Beyond an Open-loop System

Qingyong Hu (National University of Defense), Yulan Guo Yunjin Chen Jingjing Xiao and Wei An

58. Photorealistic Style Transfer with Screened Poisson Equation

Roey Mechrez (Technion), Eli Shechtman and Lihi Zelnik-Manor

59. Probabilistic Spatial Regression using a Deep Fully Convolutional Neural Network

S M Masudur Rahman Al Arif (City University of London), Karen Knapp and Greg Slabaugh

60. Lip Reading in Profile

Joon Son Chung (University of Oxford), and Andrew Zisserman

61. Adversarial Robustness: Softmax versus Openmax

Andras Rozsa (UCCS), Manuel Gunther and Terrance Boult

62. Generative 3D Hand Tracking with Spatially Constrained Pose Sampling

Konstantinos Roditakis (FORTH), Alexandros Makris and Antonis Argyros

63. Efficient Online Surface Correction for Real-time Large-Scale 3D Reconstruction

Robert Maier (Technische Universität München), Raphael Schaller and Daniel Cremers

64. Long Range Stereo from Synchronized Monocular Optical Flow Streamss

Darius Burschka, Technical University of Munich

65. Learning temporal structures for human activity recognition

Tiantian Xu (New York University), and Edward Wong

66. End-to-end multi-view lipreading

Stavros Petridis (Imperial College), Yujiang Wang Zuwei Li and Maja Pantic

67. Supervised Scale-Regularized Linear Convolutionary Filters

Marco Loog and Francois Lauze

68. Sparse Deep Feature Representation for Object Detection from Wearable Cameras

Quanfu Fan (IBM), and Richard Chen

69. Human Action Segmentation using 3D Fully Convolutional Network

Pei Yu (Northwestern University), Jiang Wang and Ying Wu

70. Deep fisher faces

Harald Hanselmann (HLTPR Group at RWTH Aachen), Shen Yan and Hermann Ney

71. Improving target tracking robustness with Bayesian data fusion

Yevgeniy Reznichenko (Marquette University), and Henry Medeiros

72. One-Shot Learning for Semantic Segmentation

Amirreza Shaban (Georgia Institute of Tech.), Shray Bansal Zhen Liu Irfan Essa and Byron Boots

73. Efficent Traffic-Sign Recognition with Scale-aware CNN

Yuchen Yang (BJUT), Shuo Liu Wei Ma Qiuyuan Wang and Zheng Liu

74. Plane-Aided Visual-Inertial Odometry for Pose Estimation of a 3D Camera based Indoor Blind Navigation System

He Zhang (Univ. of Arkansas at LR (UALR)), and cang Ye

75. Dissecting scale from pose estimation in visual odometry

Rong yuan (Brown University), Hongyi Fan and Benjamin Kimia

76. Automatic Image transformation for inducing affect

MOhsen Ali (Information Technology Uni.), and Afsheen Rafaqat

77. Subpixel Semantic Flow

Berk Sevilmis (Brown University), and Benjamin Kimia

78. Light Cascaded Convolutional Neural Networks for Accurate Player Detection

Keyu Lu (National University of Defense Technology), Jianhui Chen James Little and Hangen He

79. Exploiting Protrusion Cues for Fast and Effective Shape Modeling via Ellipses

Alex Wong (UCLA), Alan Yuille and Brian Taylor

80. A CNN-Based Approach for Automatic License Plate Recognition in the Wild

Meng Dong (University of Science and Technology of China), Dongliang He Chong Luo Dong Liu and Wenjun Zeng

81. Multi-Region Ensemble Convolutional Neural Network for High Accuracy Age Estimation

Yiliang Chen (Macau University of Science and Technology), zichang Tan Alex Po Leung Jun Wan and Jianguo Zhang

82. Primitive-based Surface Regularization for Urban 3D Reconstruction

Thomas Holzmann (Graz University of Technology), Martin Oswald Marc Pollefeys Friedrich Fraundorfer and Horst Bischof

83. Cross-modal Retrieval via Memory Network

Ge Song (Nanjing University of Aeronaut), and Xiaoyang Tan

84. A Compact Parametric Solution to Depth Sensor Calibration

Andrew Spek (Monash University), and Tom Drummond

85. Pansharpening via Locality-Constrained Sparse Representation

Songze Tang (Nanjing Forest Police College), Nan zhou and Liang Xiao

86. Residual Conv-Deconv Grid Network for Semantic Segmentation

Damien Fourure (Université Jean Monnet), Rémi Emonet Elisa Fromont Damien Muselet Alain Tremeau and Christian Wolf

87. Deep GrabCut for Object Selection

Ning Xu (University of illinois), Brian Price Scott Cohen Jimei Yang and Thomas Huang

88. Solar Power Plant Detection on Multi-Spectral Satellite Imagery using Weakly-Supervised CNN with Feedback Features and m-PCNN Fusion

Nevrez Imamoglu (AIST), Motoki Kimura Hiroki Miyamoto Aito Fujita and Ryosuke Nakamura

15:00 – 15:30 Tea break, RGS

15:30 – 17:00 Orals: Matching, RGS Theatre

Chair: Tom Drummond

15^{30} 89. Multiple-Kernel Local-Patch Descriptor

Arun Mukundan (CMP), Giorgos Tolias and Ondrej Chum

15^{45} 90. AutoScaler: Scale-Attention Networks for Visual Correspondence

Shenlong Wang (University of Toronto), Linjie Luo Ning Zhang and Jia Li

16^{00} 91. Cross-View GAN Based Vehicle Generation for Re-identification

Yi Zhou (Northumbria University), and Ling Shao

16^{15} 92. Efficient Video Summarization Using Principal Person Appearance for Video-Based Person Re-Identification

Seongro Yoon (INRIA), Furqan Khan and Francois Bremond

16^{30} 93. Fine-Grained Forensic Matching

Bailey Kong (UC Irvine), James Supancic Deva Ramanan and Charless Fowlkes

16^{45} 94. Learning Accurate Low-Bit Deep Neural Networks with Stochastic Quantization

Yinpeng Dong (Tsinghua University), Jianguo Li and Renkun Ni

19:00 – 21:00 Banquet, Millennium Gloucester Hotel

7:30 – Registration, RGS

8:00 – 8:30 Coffee & pastry, RGS

8:30 – 8:45 Welcome, RGS Theatre

8:45 – 10:15 Orals: Segmentation, RGS Theatre

Chair: Alan Yuille

8^{45} Weakly Supervised Semantic Segmentation Based on Co-segmentation

Tong Shen (The University of Adelaide), Guosheng Lin Lingqiao Liu Chunhua Shen and Ian Reid

9^{00} Semantic Segmentation with Reverse Attention

Qin Huang (University of Southern Califor), Chihao Wu Chunyang Xia Ye Wang and C.-C. Jay Kuo

9^{15} Double Expansion for Optimization of Multilabel Energies

Yelena Gorelick (Western University), Zhengqin Li and Olga Veksler

9^{30} Discovering Class-Specific Pixels for Weakly-Supervised Semantic Segmentation

Arslan Chaudhry (University of Oxford), Puneet Kumar Dokania and Philip Torr

9^{45} Online Adaptation of Convolutional Neural Networks for Video Object Segmentation

Paul Voigtlaender (RWTH Aachen University), and Bastian Leibe

10^{00} Improved Bilinear Pooling with CNNs

Tsung-Yu Lin (UMass Amherst), and Subhransu Maji

10:15 – 11:00 Break, RGS

11:00 – 12:30 Orals: Enhancement, RGS Theatre

Chair: Lihi Zelnik-Manor

11^{15} Reflectance and Shape Estimation with a Light Field Camera under Natural Illumination

Trung Ngo (Osaka University), Hajime Nagahara Ko Nishino Rin-ichiro Taniguchi and Yasushi Yagi

11^{30} GeneGAN: Learning Object Transfiguration and Object Subspace from Unpaired Data

Shuchang Zhou (Megvii Inc.), Taihong Xiao Yi Yang Dieqiao Feng Qinyao He and Weiran He

11^{45} PixColor: Pixel Recursive Colorization

Sergio Guadarrama (Google), Ryan Dahl David Bieber Jonathon Shlens Mohammad Norouzi and Kevin Murphy

12^{00} Depth Estimation and Blur Removal from a Single Out-of-focus Image

Saeed Anwar (ANU), Zeeshan Hayder and Fatih Porikli

12^{15} Exploring the structure of a real-time, arbitrary neural artistic stylization network

Golnaz Ghiasi (Google Inc.), Honglak Lee Manjunath Kudlur Vincent Dumoulin and Jonathon Shlens

12^{30} Fine-Pruning: Joint Fine-Tuning and Compression of a Convolutional Network with Bayesian Optimization

Frederick Tung (Simon Fraser University), Srikanth Muralidharan and Greg Mori

12:30 – 13:15 Lunch, RGS

13:30 – 14:15 Workshops Keynote, ICL Huxley Bld.

Chair: Yannis Panagakis

13^{30} **HoloLens: Computer Vision meets Mixed Reality**

Jamie Shotton (Microsoft Research)

14:15 – 16:00 5th Activity Monitoring by Multiple Distributed Sensing (AMMDS 2017), ICL Huxley Bld.

Chair: Pier Luigi Mazzeo, Paolo Spagnolo (National Research Council Italy)

14^{15} Welcome

14^{20} Superpixels based Manifold Structured Sparse RPCA for Moving Object Detection. *Sajid Javed, Arif Mahmood, Thierry Bouwmans, Soon Ki Jung*

14^{45} ATLAS: Adaptive Single Object Tracking using Off line Learned Motion and Visual Patterns. *Ruxandra G. Tapu, Bogdan Mocanu, Titus Zaharia*

15^{10} Vehicle Re-Identification by Fine-Grained Cross-Level Deep Learning. *Aytaç Kanacı, Xiatian Z hu, Shaogang Gong*

15^{35} Geometry-Based Multiple Camera Head Detection in Dense Crowds. *Nicola Pellicano, Emanuel Aldea, Sylvie Le Hegarat-Mascle*

14:15 – 16:00 Deep Learning on Irregular Domains (DLID 2017), ICL Huxley Bld.

Chair: Xianghua Xie, Michael Edward (Swansea University), Pierre Vandergheynst, Michaël Defferrard (École Polytechnique Fédérale de Lausanne)

14^{15} Welcome

14^{20} Invited Speaker – *Michael Defferrard (EPFL)*

14^{50} Invited Speaker – *Joost Bastings (University of Amsterdam)*

15^{20} Invited Speaker – *Michael Edwards (Swansea University)*

15^{50} Speaker – *Gilles Puy (Technicolor)*

14:15 – 16:00 Lip-Reading using deep learning methods (LRDLM 2017), ICL Huxley Bld.

Chair: Themos Stafylakis, Georgios Tzimiropoulos (University of Nottingham), Stavros Petridis, Maja Pantic (Imperial College London)

14^{25} Welcome

14^{30} Keynote – *Andrew Zisserman and Joon Son Chung (University of Oxford & Google DeepMind)*

15^{30} Invited Speaker – *Helen L. Bear (University of East London)*

14:15 – 16:00 Automatic Face Analytics for Human Behavior Understanding (FaceHUB 2017), ICL Huxley Bld.

Chair: Xiaohua Huang, Guoying Zhao (University of Oulu)

14^{15} Welcome

14^{20} Oral Session

15^{00} Keynote – *Hatice Gunes (University of Cambridge)*

16:00 – 16:30 Tea break, ICL Huxley Bld.

16:30 – 18:30 5th Activity Monitoring by Multiple Distributed Sensing (AMMDS 2017), ICL Huxley Bld.

Chair: Pier Luigi Mazzeo, Paolo Spagnolo (National Research Council Italy)

16^{30} Knowledge Driven Activity Recognition from Patterns of Object Use. *Isibor Kennedy Ihianle, Usman Naeem, Syed Islam*

16^{55} Extraction of Spatio temporal Descriptors for Maritime Vessel Detection using Attentive Sensing. *Roman Palenychka, Rafael Falcon, Rami Abielmona, Emil Petriu*

17^{20} Dense Net with pre-activated deconvolution for estimating depth map from single image *Saurav Sharma, Ram P Padhy, Suman Choudhury, Nabarun Goswami, Pankaj Kumar Sa*

17^{45} Fine segmentation for Activity of Daily Living analysis in a wide-angle multi-camera set-up. *Philipe Ambrozio Dias, Henry Medeiros, Francesca Odone*

16:30 – 18:30 Deep Learning on Irregular Domains (DLID 2017), ICL Huxley Bld.

Chair: Xianghua Xie, Michael Edward (Swansea University), Pierre Vandergheynst, Michaël Defferrard (École Polytechnique Fédérale de Lausanne)

16^{30} Invited Speaker – *Federico Monti*

17^{00} Speaker – *Sei-ichiro Kamata*

17^{15} Speaker – *Jan P Siebert*

17^{30} Speaker – *Abhinav Agarwalla*

17^{45} Speaker – *Aditya Nigam*

16:30 – 18:30 Lip-Reading using deep learning methods (LRDLM 2017), ICL Huxley Bld.

Chair: Themos Stafylakis, Georgios Tzimiropoulos (University of Nottingham), Stavros Petridis, Maja Pantic (ICL Huxley Bld.)

16^{30} Oral Session

17^{30} Poster Session

16:30 – 18:30 Automatic Face Analytics for Human Behavior Understanding (FaceHUB 2017), ICL Huxley Bld.

Chair: Xiaohua Huang, Guoying Zhao (University of Oulu)

16^{30} Keynote – *Javier Orozo (Realeyes)*

17^{30} Oral Session

Intro to Reinforcement Learning

Shimon Whiteson, *University of Oxford*

This tutorial will give a brief introduction to the fundamental concepts in reinforcement learning. These include the exploration/exploitation dilemma, the credit assignment problem, bandit algorithms, planning and learning in Markov decision processes, and deep reinforcement learning.

Shimon Whiteson is an associate professor in the Department of Computer Science at the University of Oxford, and a tutorial fellow at St. Catherine's College. His research focuses on artificial intelligence, with a particular focus on machine learning and decision-theoretic planning. In addition to theoretical work on these topics, he has in recent years also focused on applying them to practical problems in robotics and search engine optimisation. He studied English and Computer Science at Rice University before completing a doctorate in Computer Science at the University of Texas at Austin in 2007. He then spent eight years as an Assistant and then an Associate Professor at the University of Amsterdam before joining Oxford as an Associate Professor in 2015. He was awarded an ERC Starting Grant from the European Research Council in 2014 and a Google Faculty Research Award in 2017.

3D Reconstruction of Dynamic Scenes from Monocular Video

Lourdes Agapito, *University College London*

As humans we take the ability to perceive the dynamic world around us in three dimensions for granted. From an early age we can grasp an object by adapting our fingers to its 3D shape; we can understand our mother's feelings by interpreting her facial expressions; or we can effortlessly navigate through a busy street. All of these tasks require some internal 3D representation of shape, deformations and motion.

Building algorithms that can emulate this level of human 3D perception has proved to be an extremely challenging task. In this tutorial I will show progress from early systems which captured sparse 3D models with primitive representations of deformation towards the most recent algorithms which can capture every fold and detail of hands or faces in 3D using as input video sequences taken with a single consumer camera. There is now great short-term potential for commercial uptake of this technology.

Professor Lourdes Agapito obtained her BSc, MSc and PhD (1996) degrees from the Universidad Complutense de Madrid (Spain). She held an EU Marie Curie Postdoctoral Fellowship at The University of Oxford's Robotics Research Group before being appointed as a Lecturer at Queen Mary, University of London in 2001. In 2008 she was awarded an ERC Starting Grant to carry out research on the estimation of 3D models of dynamic scenes from monocular video sequences. In July 2013 she joined the Department of Computer Science at University College London (UCL) where she leads a research team that focuses on developing algorithms for 3D understanding of the real world from video.

Lourdes was Program Chair for CVPR 2016, the top annual conference in computer vision; in addition she was Programme Chair for 3DV'14 and has served as Area Chair for the top Computer Vision conferences. Lourdes is Associate Editor for the International Journal of Computer Vision (IJCV) and IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), a member of the Executive Committee of the British Machine Vision Association and a member of the EPSRC Peer Review College.

Probabilistic and Deep Models for 3D Reconstruction

Andreas Geiger, *Max Planck Institut*

3D reconstruction from multiple 2D images is an inherently ill-posed problem. Prior knowledge is required to resolve ambiguities and probabilistic models are desirable to capture the ambiguities in the reconstructed model. In this talk, I will present two recent results tackling these two aspects. First, I will introduce a probabilistic framework for volumetric 3D reconstruction where the reconstruction problem is cast as inference in a Markov random field using ray potentials. Our main contribution is a discrete-continuous inference algorithm which computes marginal distributions of each voxel's occupancy and appearance. I will show that the proposed algorithm allows for Bayes optimal predictions with respect to a natural reconstruction loss. I will further demonstrate several extensions which integrate non-local CAD priors into the reconstruction process. In the second part of my talk, I will present a novel framework for deep learning with 3D data called OctNet which enables 3D CNNs on high-dimensional inputs. I will demonstrate the utility of the OctNet representation on several 3D tasks including classification, orientation estimation and point cloud labeling. Finally, I will present an extension of OctNet called OctNetFusion which jointly predicts the space partitioning function with the output representation, resulting in an end-to-end trainable model for volumetric depth map fusion.

Andreas Geiger is a Max Planck Research Group Leader at the MPI for Intelligent Systems in Tübingen heading the Autonomous Vision Group (AVG), and a Visiting Professor at ETH Zürich. Prior to this, he was a research scientist in the Perceiving Systems department at MPI Tübingen. He studied at KIT, EPFL and MIT and received his PhD degree in 2013 from the Karlsruhe Institute of Technology. His research interests are at the intersection of 3D reconstruction and visual scene understanding with a particular focus on rich semantic and geometric priors for bridging the gap between low-level and high-level vision. He is particularly interested in autonomous driving applications. His work has received several prices, including the Heinz Maier Leibnitz Prize, the Ernst-Schoemperlen Award, as well as best paper awards at CVPR, GCPR and 3DV. He is an associate member of the Max Planck ETH Center for Learning Systems and serves as area chair and associate editor in computer vision (CVPR, ECCV, PAMI).

Deep Learning for 3D Localization

Vincent Lepetit, *University of Bordeaux*

The first part of the talk will describe a novel method for 3D object detection and pose estimation from color images only. We introduce a "holistic" approach that relies on a representation of a 3D pose suitable to Deep Networks and on a feedback loop. This approach, like many previous ones is however not sufficient for handling objects with an axis of rotational symmetry, as the pose of these objects is in fact ambiguous. We show how to relax this ambiguity with a combination of classification and regression. The second part will describe an approach bridging the gap between learning-based approaches and geometric approaches, for accurate and robust camera pose estimation in urban environments from single images and simple 2D maps.

Dr. Vincent Lepetit is a Full Professor at the LaBRI, University of Bordeaux, and an associate member of the Inria Manao team. He also supervizes a research group in Computer Vision for Augmented Reality at the Institute for Computer Graphics and Vision, TU Graz. He received the PhD degree in Computer Vision in 2001 from the University of Nancy, France, after working in the ISA INRIA team. He then joined the Virtual Reality Lab at EPFL as a post-doctoral fellow and became a founding member of the Computer Vision Laboratory. He became a Professor at TU Graz in February 2014, and at University of Bordeaux in January 2017. His research interests include computer vision and machine learning, and their application to 3D hand pose estimation, feature point detection and description, and 3D registration from images.

Visual Reconstruction and Image-Based Rendering

Richard Szeliski, *Facebook*

The reconstruction of 3D scenes and their appearance from imagery is one of the longest-standing problems in computer vision. Originally developed to support robotics and artificial intelligence applications, it has found some of its most widespread use in the support of interactive 3D scene visualization. One of the keys to this success has been the melding of 3D geometric and photometric reconstruction with a heavy re-use of the original imagery, which produces more realistic rendering than a pure 3D model-driven approach. In this talk, I give a retrospective of two decades of research in this area, touching on topics such as sparse and dense 3D reconstruction, the fundamental concepts in image-based rendering and computational photography, applications to virtual reality, as well as ongoing research in the areas of layered decompositions and 3D-enabled video stabilization.

Richard Szeliski is a Research Scientist in the Computational Photography group at Facebook, which he founded in 2015. He is also an Affiliate Professor at the University of Washington, and is member of the NAE and a Fellow of the ACM and IEEE. Dr. Szeliski has done pioneering research in the fields of Bayesian methods for computer vision, image-based modeling, image-based rendering, and computational photography, which lie at the intersection of computer vision and computer graphics. His research on Photo Tourism, Photosynth, and Hyperlapse are exciting examples of the promise of large-scale image and video-based rendering.

Dr. Szeliski received his Ph.D. degree in Computer Science from Carnegie Mellon University, Pittsburgh, in 1988 and joined Facebook as founding Director of the Computational Photography group in 2015. Prior to Facebook, he worked at Microsoft Research for twenty years, the Cambridge Research Lab of Digital Equipment Corporation for six years, and several other industrial research labs. He has published over 150 research papers in computer vision, computer graphics, neural nets, and numerical analysis, as well as the books Computer Vision: Algorithms and Applications and Bayesian Modeling of Uncertainty in Low-Level Vision. He was a Program Committee Chair for CVPR'2013 and ICCV'2003, served as an Associate Editor of the IEEE Transactions on Pattern Analysis and Machine Intelligence and on the Editorial Board of the International Journal of Computer Vision, and as Founding Editor of Foundations and Trends in Computer Graphics and Vision.

Visipedia - A universal visual expert

Pietro Perona, *California Institute of Technology*

Each day we are faced with visual puzzles: What is the species of that bird? What script was used to write on this old stone and what does it say? Who painted that picture? In most cases, an expert would be able to quickly inform us, but we do not know whom to ask. I will discuss the challenge of of building a universal visual expert – a network of people, data and machines designed to harvest and organize visual information and make it accessible to anyone anywhere. I will explore the technical challenges arising from Visipedia and discuss their implications for computer vision, machine learning, artificial intelligence, human-machine systems and visual psychology. I will present data from large-scale experiments carried out by building systems that people use: iNaturalist, eBird and regisTree. I will conclude by discussing open challenges for Computer Vision and Machine Learning researchers.

Pietro Perona received his PhD from UC Berkeley, was a post-doctoral fellow at MIT and is now a Professor at the California Institute of Technology in Pasadena. He is currently interested in visual categorization and in the analysis of behavior. He has worked on partial differential equations for image processing, on modeling visual perception, on visual search and attention and on the role of visual mechanisms in art.

HoloLens: Computer Vision meets Mixed Reality

Jamie Shotton, *Microsoft Research*

Microsoft HoloLens is the world's first fully-untethered, self-contained Holographic computer, and has been made possible by recent advances in computer vision. In this talk, we'll explore the remarkable combination of hardware and software innovation involved in building a device that can track your head motion, recognize your hand gestures, and reconstruct the room in 3D, all in real-time on a low-power embedded processor without external sensors. We'll also talk through some of the exciting opportunities for HoloLens, for both research and real-world applications.

Jamie Shotton is a Partner Scientist and leads the HoloLens Science team at Microsoft in Cambridge, UK, where his team focuses on the visual understanding of people to improve interaction and communication in mixed reality. He studied Computer Science at the University of Cambridge, where he remained for his PhD in computer vision and machine learning. He joined Microsoft Research in 2008 where he was a research scientist and head of the Machine Intelligence & Perception group, before founding the HoloLens Science Cambridge team in 2016. His research focuses at the intersection of computer vision, AI, machine learning, and graphics, with particular emphasis on systems that allow people to interact naturally with computers. He has received multiple Best Paper and Best Demo awards at top-tier academic conferences. His work on machine learning for body part recognition for Kinect was awarded the Royal Academy of Engineering's MacRobert Award 2011, and he shares Microsoft's Outstanding Technical Achievement Award for 2012 with the Kinect engineering team. In 2014 he received the PAMI Young Researcher Award, and in 2015 the MIT Technology Review Innovator Under 35 Award ("TR35").

.

.

.

Platinum Sponsors

Microsoft

facebook

oculus

amazon

SCAPE

Google

Gold Sponsors

TOSHIBA
Leading Innovation

QUALCOMM

SLAMCORE

Silver Sponsors

OSRAM

sengled

IET The Institution of Engineering and Technology