

# **Tutorials                      Sunday, December 1**

<b><u>Date</u></b>	<b><u>Workshop Title / Organizers</u></b>
12/01/13 (AM only)	<b>Don't Relax: Why Non-Convex Algorithms are Often Needed for Sparse Estimation</b> <i>Organizers: David Wipf</i>
12/01/13 (AM only)	<b>Part-based Models for Recognition</b> <i>Organizers: Subhransu Maji, Lubomir Bourdev, Ross Girshick</i>
12/01/13 (AM only)	<b>Sparsity Estimation and Robust Learning: A Half-quadratic Minimization View</b> <i>Organizers: Ran He, Wei-Shi Zheng, Wang Liang</i>
12/01/13 (AM only)	<b>Spectral geometry methods in 3D data analysis</b> <i>Organizers: Alexander Bronstein, Michael Bronstein</i>
12/01/13 (PM only)	<b>Image and Video Matting</b> <i>Organizers: Ehsan Shahrian, Margrit Gelautz, Brian Price</i>
12/01/13 (PM only)	<b>Introduction to Statistical Optimization for Geometric Estimation</b> <i>Organizers: Kenichi Kanatani</i>
12/01/13 (PM only)	<b>Low-Dimensional Subspaces in Computer Vision</b> <i>Organizers: Roland Angst</i>
12/01/13 (PM only)	<b>Sparse and Low-Rank Representations in Computer Vision -- Theory, Algorithms, and Applications</b> <i>Organizers: Bernard Ghanem, John Wright, Allen Y. Yang</i>
12/02/13 (AM only)	<b>Decision Forests and Fields for Computer Vision</b> <i>Organizers: Jamie Shotton, Sebastian Nowozin</i>
12/02/13 (PM only)	<b>Dense Image Correspondences for Computer Vision</b> <i>Organizers: Ce Liu, Zhuowen Tu, Michael Rubinstein</i>

## **Main Conference At a Glance**

	<b>Dec.3rd Tuesday</b>
8:30-9:00	<i>Welcome &amp; Opening Remarks</i>
9:00-10:00	Oral (4 papers)
10:00-10:45	Poster Spotlights(52)
10:45-12:15	Poster Session
12:15-14:15	lunch break
14:15-15:00	oral (3 papers)
15:00-15:45	Poster Spotlights(52)
15:45-16:15	Break
16:15-17.15	Oral (4 papers)
17:15-18:45	Poster Session
19:00-finish	PAMI TC
<hr/>	
	<b>Dec.4<sup>th</sup> Wednesday</b>
9:00-10:00	Oral (4 papers)
10:00-10:45	Poster Spotlights(52)
10:45-12:15	Poster Session
12:15-14:15	lunch break
14:15-15:00	Oral (3 papers)
15:00-15:45	Poster Spotlights(52)
15:45-16:15	Break
16:15-17.15	Oral (4 papers)
17:15-18:45	Poster Session
18:45-19:00	Award Session
19:00-finish	Welcome Reception

## **Main Conference At a Glance**

	<b>Dec.5<sup>th</sup> Thursday</b>
9:00-10:00	Oral (4 papers)
10:00-10:45	Poster Spotlights(52)
10:45-12:15	Poster Session
12:15-14:15	lunch break
14:15-15:00	oral (3 papers)
15:00-15:45	Poster Spotlights(52)
15:45-16:15	Break
16:15-17:15	Oral (4 papers)
17:15-18:45	Poster Session

---

	<b>Dec.6<sup>th</sup> Friday</b>
9:00-10:00	Oral (4 papers)
10:00-10:45	Poster Spotlights(52)
10:45-12:15	Poster Session
12:15-13:45	lunch break
13:45-14:45	Keynote Speech + half hour break
14:45-15:15	
15:15-16:15	Oral (4 papers)
16:15-17:00	Poster Spotlights(50)
17:00-18:30	Poster Session

# **Main Conference Tuesday, December 3**

## **Day 1:**

**8:50 - 9:00 Welcome & Opening remarks (Aannounce Marr Prize and student best paper)**

**9:00 - 10:00 Oral session 1A (4 papers, 60min)**

### **Recognition**

**Orals 1A:01 – 1A:03**

#### **1 HOGgles: Visualizing Object Detection Features,**

Carl Vondrick\*, MIT; Aditya Khosla, ; Tomasz Malisiewicz, ; Antonio Torralba, MIT

#### **2 How Do You Tell a Blackbird from a Crow?**

Thomas Berg\*, Columbia University; Peter Belhumeur, "Columbia University, USA"

#### **3 Regionlets for Generic Object Detection,**

Xiaoyu Wang\*, NEC Labs America; Ming Yang, NEC Labs America; Shenghuo Zhu, ; Yuanqing Lin,

#### **4 Learning Graphs to Match,**

Minsu Cho\*, ; Karteek Alahari, ENS-Willow; Jean Ponce, "ENS, France"

**10:00 - 10:45 Spotlights (45min, 52 spotlights at 50secs each)**

**10:45 - 12:15 Poster session 1A: Posters 1A:01 - 1A:52**

**1 Shape Anchors for Data-driven Multi-view Reconstruction,**  
Andrew Owens, MIT; Jianxiong Xiao\*, MIT; Antonio Torralba,  
MIT; Bill Freeman, "MIT, USA"

**2 Deterministic Fitting of Multiple Structures using Iterative  
MaxFS with Inlier Scale Estimation,** Kwang Hee Lee\*, Sogang  
University; Sang Lee, "Sogang University, Korea"

**3 Constant Time Weighted Median Filtering for Stereo  
Matching and Beyond,** Ziyang Ma\*, Institute of Software, CAS;  
Kaiming He, Microsoft Research Asia; Yichen Wei, ; Jian Sun,  
"Microsoft Research, China"; Enhua Wu, Faculty of Science and  
Technology, University of Macau

**4 Refractive Structure-from-Motion on Underwater Images,**  
Anne Jordt-Sedlazeck\*, Kiel University; Reinhard Koch, Christian-  
Albrechts-Universitat Kiel

**5 Live Metric 3D Reconstruction on Mobile Phones,** Petri  
Tanskanen, ETH Zurich; Kalin Kolev\*, ETH Zurich; Lorenz Meier,  
ETH Zurich; Federico Camposeco Paulsen, ETH Zurich; Olivier  
Saurer, ETH Zurich; Marc Pollefeys, ETH

**6 Joint Subspace Stabilization for Stereoscopic Video,** Feng  
Liu\*, Portland State University; Yuzhen Niu, ; Hailin Jin,

**7 Video Synopsis by Heterogeneous Multi-Source Correlation,**  
Xiatian Zhu\*, Queen Mary, Univ. of London; Chen Change Loy,  
CUHK; Shaogang Gong, EECS, QMUL

**8 DCSH - Matching Patches in RGBD Images,** Yaron Eshet, ;  
Simon Korman\*, Tel-Aviv University; Eyal Ofek, Microsoft; Shai  
Avidan, Tel-Aviv University

**9 Scene Text Localization and Recognition with Oriented Stroke  
Detection,** Lukas Neumann\*; Jiri Matas, Czech Technical  
University

**10 Adapting Classification Cascades to New Domains,** vidit  
Jain\*, ; Sachin Farfade,

**11 Learning Identity-Preserving Features**, Zhenyao Zhu\*,  
CUHK; Ping Luo, CUHK

**12 A Cascaded Deep Learning Architecture for Pedestrian Detection**, Xingyu ZENG\*, The Chinese University of HK; Wanli Ouyang, The Chinese University of HK; Xiaogang Wang, "The Chinese University of Hong Kong, Hongkong"

**13 Unsupervised Random Forest Manifold Alignment for Lipreading**, Yuru PEI\*, Peking University; Tae-Kyun Kim, Imperial College London; Hongbin Zha, Peking University

**14 Calibration-free Gaze Estimation using Human Gaze Patterns**, Fares Alnajar\*, University of Amsterdam; Theo Gevers, University of Amsterdam; Roberto Valenti, ; Sennay Ghebreab, University of Amsterdam

**15 Partial Sum Minimization of Singular Values in RPCA for Low-Level Vision**, Tae-Hyun Oh, KAIST; Hyeonwoo Kim, KAIST; Yu-Wing Tai, "KAIST, Korea"; Jean-Charles Bazin, ETH-Z; In So Kweon\*, KAIST

**16 Saliency Detection: A Boolean Map Approach**, Jianming Zhang\*, Boston University; Stan Sclaroff, Boston University

**17 Topology-Constrained Layered Tracking with Latent Flow**, Jason Chang\*, CSAIL, MIT; John Fisher, MIT

**18 Stacked Predictive Sparse Coding for Classification of Distinct Regions in Tumor Histopathology**, Hang Chang\*, Lawrence Berkeley National Lab; Yin Zhou, Nandita Nayak, ; Paul Spellman, ; Bahram Parvin,

**19 Higher Order Matching for Consistent Multiple Target Tracking**, Chetan Arora\*, ; Amir Globerson,

**20 Image Matching: a General Framework Combining Direct and Feature-based Costs**, Jim Braux-Zin\*, CEA, LIST; Adrien Bartoli, Université d'Auvergne; Romain Dupont, CEA, LIST

**21 Revisiting Example Dependent Cost-Sensitive Learning with Decision Trees**, Oisin Mac Aodha\*, UCL; Gabriel Brostow,

**22 Modeling Self-Occlusions in Dynamic Shape and Appearance Tracking**, Yanchao Yang, KAUST; Ganesh Sundaramoorthi\*, KAUST

**23 A Convex Optimization Framework for Active Learning**, Ehsan Elhamifar\*, UC Berkeley; Guillermo Sapiro, Duke; Shankar Sastry, UC Berkeley

**24 A Generalized Iterated Shrinkage Algorithm for Non-convex Sparse Coding**, Wangmeng Zuo, Harbin Institute of Technology; Deyu Meng, Xi'an Jiaotong University; Lei Zhang\*, The Hong Kong Polytechnic University; Xiangchu Feng, School of Science, Xidian University; David Zhang, The Hong Kong Polytechnic University

**25 Latent Space Sparse Subspace Clustering**, Vishal Patel\*, UMIACS; Hien Nguyen , UMIACS; Rene Vidal, Johns Hopkins University

**26 Tracking Revisited using RGBD Camera: Benchmark and Baselines**, Shuran Song, HKUST; Jianxiong Xiao\*, MIT

**27 A Simple Model for Intrinsic Image Decomposition with Depth Cues**, Qifeng Chen, ; Vladlen Koltun\*, Stanford University

**28 Quadruplet-wise Image Similarity Learning**, Marc Law\*, LIP6; Nicolas Thome, LIP6; Matthieu Cord,

**29 Complementary Projection Hashing**, Zhongming Jin\*, Zhejiang University; Deng Cai, ; Yao Hu, Zhejiang university; Debing Zhang, Zhejiang university; Xuelong Li,

**30 Find the Best Path: an Efficient and Accurate Classifier for Image Hierarchies**, Min Sun\*, ; Wan Huang, University of Michigan at Ann Arbor; Silvio Savarese, University of Michigan at Ann Arbor

**31 Detecting Dynamic Objects with Multi-View Background Subtraction**, Raúl Díaz, University of California, Irvine; Sam Hallman, University of California, Irvine; Charless Fowlkes\*, University of California, Irvine

**32 Low-Rank Sparse Coding for Image Classification**, Tianzhu Zhang\*, ADSC of UIUC in Singapore; Bernard Ghanem, KAUST; Si Liu, National University of Singapore; Changsheng Xu, CASIA; Narendra Ahuja,

**33 Allocentric Pose Estimation**, José Oramas\*, KU Leuven - ESAT; Luc De Raedt, KU Leuven - CS; Tinne Tuytelaars, KU Leuven

**34 Attribute Pivots for Guiding Relevance Feedback in Image Search**, Adriana Kovashka\*, ; Kristen Grauman, University of Texas at Austin

**35 Decomposing Bag of Words Histograms**, Ankit Gandhi\*, IIIT Hyderabad; Karteek Alahari, ENS-Willow; c. v. Jawahar, IIIT Hyderabad

**36 SYM-FISH: A Symmetry-aware Flip Invariant Sketch Histogram Shape Descriptor**, Xiaochun Cao, Chinese Academy of Sciences; Hua Zhang\*, Tju; Si Liu, National University of Singapore; Xiaojie Guo, Tianjin University

**37 Symbiotic Segmentation and Part Localization for Fine-Grained Categorization**, Yuning Chai\*, University of Oxford; Victor Lempitsky, Skolkovo Institute of Science and Technology; Andrew Zisserman, University of Oxford

**38 Image Set Classification Using Holistic Multiple Order Statistics Features and Localized Multi-Kernel Metric Learning**, Jiwen Lu\*, Advanced Digital Sciences Center, Singapore; Gang Wang, NTU; Pierre Moulin, UIUC

**39 Learning a Dictionary of Shape Epitomes with Application to Semantic Labeling**, Liang-Chieh Chen\*, UCLA; George Papandreou, UCLA; Alan Yuille, UCLA



**40 Pyramid Coding for Functional Scene Element Recognition in Video Scenes**, Eran Swears\*, Kitware Inc.; Anthony Hoogs, Kitware, USA; Kim Boyer, RPI

**41 Box In the Box: Joint 3D Layout and Object Reasoning from Single Images**, Alexander Schwing\*, ETH Zurich; Sanja Fidler, TTIC; Marc Pollefeys, ETH; Raquel Urtasun, Toyota Technological Institute at Chicago

**42 Semantic Transform: Weakly Supervised Semantic Inference for Relating Visual Attributes**, Sukrit Shankar\*, Cambridge University; Joan Lasenby, University of Cambridge; Roberto Cipolla, Cambridge University

**43 From Subcategories to Visual Composites: A Multi-Level Framework for Object Detection**, Tian Lan\*, Simon Fraser University; Leonid Sigal, ; Michalis Raptis, Disney Research Pittsburgh; Greg Mori, Simon Fraser University

**44 Online Video Superpixels for Temporal Window Objectness**, Michael Van den Bergh, ETH; Gemma Roig, ETH; Xavier Boix\*, ETH; Santiago Manen, BIWI ETH Zurich; Luc Van Gool, ETH

**45 Temporally Consistent Superpixels**, Matthias Reso\*, TNT LUH Hannover; Joern Jachalsky, Technicolor; Joern Ostermann, Institut für Informationsverarbeitung / Universität Hannover; Bodo Rosenhahn,

**46 Semi-supervised learning for large scale image cosegmentation**, Zhengxiang Wang\*, Fujitsu R&D Center ; Rujie Liu, Fujitsu R&D Center

**47 Automatic Kronecker Product Model Based Detection of Repeated Patterns in 2D Urban Images**, Juan Liu, Graduate Center City University of New York; Emmanouil Psarakis, University of Patras; Ioannis Stamos\*, CUNY

**48 Group Norm for Learning Structured SVMs with Unstructured Latent Variables**, Daozheng Chen, UMD; Dhruv Batra\*, Virginia Tech; Bill Freeman, "MIT, USA"

**49 Alternating Regression Forests for Object Detection and Pose Estimation**, Samuel Schulter\*, TUGraz; Christian Leistner, Microsoft; Paul Wohlhart, TU Graz; Peter Roth, ; Horst Bischof, Graz University of Technology

**50 Dynamic Label Propagation for Semi-supervised Multi-class Multi-label Classification**, Bo Wang\*, Stanford University; Zhuowen Tu, University of California, San Diego, USA; John Tsotsos, "York University, Canada"

**51 Translating video into natural language descriptions**, Marcus Rohrbach\*, MPI Informatics; Wei Qiu, Coli.uni-saarland.de; Ivan Titov, Saarland University; Stefan Thater, ; Manfred Pinkal, Saarland University; Bernt Schiele, "MPI Informatics, Germany"

**52 Person Re-Identification Post-Rank Optimisation**, Chunxiao Liu, Tsinghua University; Chen Change Loy, CUHK; Shaogang Gong\*, EECs, QMUL; Guijin Wang, Tsinghua University

### **Lunch 12:15 - 14:15**

### **14:15 - 15:00 Oral session 1B (3 papers, 45mins)**

#### **Computational Photography**

#### **Orals 1B:01 – 1B:03**

#### **1 Scene Intensity and Depth Acquisition from One Detected Photon per Pixel**

Ahmed Kirmani\*, MIT; Donggeek Shin, MIT; Dheera Venkatraman, MIT; Franco Wong, MIT; Vivek Goyal, MIT

#### **2 Separating Reflective and Fluorescent Components using High Frequency Illumination in the Spectral Domain**

Ying Fu\*, The University of Tokyo; Antony Lam, National Institute of Informatics; Imari Sato, ; Takahiro Okabe, ; Yoichi Sato, The University of Tokyo, Japan"

#### **3 Rolling Shutter Stereo**

Olivier Saurer\*, ETH Zurich; Kevin Kooser, ; Jean-Yves Bouguet, Google ; Marc Pollefeys, ETH

**15:00 - 15:45 Spotlights (45mins)**

**15:45 - 16:15 Break (30 mins)**

**16:15 - 17:15 Oral session 1C (4 papers 60 mins) 3D**

**Vision**

**Orals 1C:01 – 1C:04**

**1 Elastic Fragments for Dense Scene Reconstruction,**

Qian-Yi Zhou\*, Stanford University; Stephen Miller, Stanford University; Vladlen Koltun, Stanford University

**2 A Global Linear Method for Camera Pose Registration,**

Nianjuan Jiang\*, ADSC; Zhaopeng Cui, NUS; Ping Tan

**3 A Rotational Stereo Model Based On XSlit Imaging,**

Jinwei Ye\*, University of Delaware; Yu Ji, University of Delaware; Jingyi Yu, University of Delaware

**4 Lifting 3D Manhattan Lines from a Single Image,**

Srikumar Ramalingam\*, MERL; Matthew Brand, MERL

**17:15 - 18:45 Poster session 1B : Posters 1B:01-1B:52**

**1 Dynamic Probabilistic Volumetric Models,** Ali Ulusoy\*, Brown University; Joseph Mundy, Brown University

**2 Network Principles for SfM: Disambiguating Repeated**

**Structures with Local Context,** Kyle Wilson\*, Cornell University; Noah Snavely, Cornell University

**3 Efficient and Robust Large-Scale Rotation Averaging**, Avishek Chatterjee, Indian Institute of Science; Venu Madhav Govindu\*, Indian Institute of Science

**4 Pose Estimation with Unknown Focal Length using Points**, Directions and Lines, Yubin Kuang\*, Lund University; Kalle Åström, Lund University

**5 Unsupervised intrinsic calibration from a single frame using a "plumb-line" approach**, Rui Melo\*, ISR-Coimbra; Michel Antunes, ; Joao Barreto, ; Gabriel Falco, ; Nuno Goncalves,

**6 Structured Light in Sunlight**, Mohit Gupta, Qi Yin\*, Columbia University; Shree Nayar, Columbia University

**7 Content-Aware Rotation**, Kaiming He\*, Microsoft Research Asia; Huiwen Chang, Tsinghua University; Jian Sun, "Microsoft Research, China"

**8 Fast Direct Super-Resolution by Simple Functions**, Chih-Yuan Yang\*, UC Merced; Ming-Hsuan Yang, "UC Merced, USA"

**9 Recognizing Text with Perspective Distortion in Natural Scenes**, Trung Quy Phan\*, National University of Singapore; Palaiahnakote Shivakumara, University of Malaya; Shangxuan Tian, National University of Singapore; Chew Lim Tan, National University of Singapore

**10 Rank Minimization across Appearance and Shape for AAM Ensemble Fitting**, Xin Cheng\*, Queensland U of Tech; Sridharan Sridha, Queensland U of Tech; Jason Saragih, Queensland U of Tech; Simon Lucey,

**11 Face Recognition via Archetype Hull Ranking**, Yuanjun Xiong\*, The Chinese University of Hong Kong; Wei Liu, IBM T. J. Watson Research; Deli Zhao, The Chinese University of Hong Kong; Xiaoou Tang, The Chinese University of Hong Kong

**12 Optimization problems for fast AAM fitting in-the-wild**, Georgios Tzimiropoulos\*, University of Lincoln/Imperial College London; Maja Pantic, Imperial College

**13 Robust Feature Set Matching for Partial Face Recognition,** RENLIANG WENG\*, Nanyang Technological University; Jiwen Lu, Advanced Digital Sciences Center, Singapore; Junlin Hu, NTU; Gao Yang, Nanyang Technological University; Yap-Peng Tan,

**14 Cross-view Action Recognition over Heterogeneous Feature Spaces,** Xinxiao Wu\*, ; Han Wang, Beijing Institute of Technology; Cuiwei Liu, ; Yunde Jia,

**15 Efficient Image Dehazing with Boundary Constraint and Contextual Regularization,** Gaofeng MENG\*, Chinese Academy of Sciences; Ying WANG, ; Jiangyong DUAN, ; Shiming Xiang, NLPR, CASIA; Chunhong Pan, NLPR, CASIA

**16 From Where and How to What We See,** Karthikeyan Shanmuga Vadivel\*, UCSB; Vignesh Jagadeesh, UCSB; Renuka Shenoy, UCSB; Miguel Eckstein, UCSB; B.S. Manjunath, UCSB

**17 Restoring An Image Taken Through a Window Covered with Dirt or Rain,** David Eigen\*, Courant Institute, NYU; Dilip Krishnan, NYU; Rob Fergus, New York University

**18 Uncertainty-driven Efficiently-Sampled Sparse Graphical Models for Concurrent Tumor Segmentation and Atlas Registration,** Sarah Parisot\*, Ecole Centrale Paris; William Wells, Surgical Planning Laboratory, Harvard Medical School; Stéphane Chemouny, Intrasure SA; Hugues Duffau, Hpital Gui de Chauliac; Nikos Paragios, Ecole Centrale de Paris

**19 Tracking via Robust Multi-Task Multi-View Joint Sparse Representation,** Zhibin Hong\*, University of Technology, Sydney; Xue Mei, Future Mobility Research department, Toyota Research Institute, North America; Danil Prokhorov, TTC; Dacheng Tao, University of Technology, Sydney

**20 Online Robust Non-negative Dictionary Learning for Visual Tracking,** Naiyan Wang\*, HKUST; Jingdong Wang, Microsoft Research Asia; Dit-Yan Yeung, HKUST

**21 Robust Object Tracking with Online Multi-lifespan Dictionary Learning**, Junliang Xing\*, Institute of Automation, Chinese Academy of Sciences; Jin Gao, Institute of Automation Chinese Academy of Sciences; Bing Li, NLPR, CASIA; Weiming Hu, Institute of Automation Chinese Academy of Sciences; Shuicheng Yan, "NUS, Singapore"

**22 Depth from Combining Defocus and Correspondence Using Light-Field Cameras**, Michael Tao\*, ; Sunil Hadap, Adobe Inc.; Jitendra Malik, UC Berkeley; Ravi Ramamoorthi, U.C. Berkeley

**23 Affine Constraint Group Sparse Coding**, Yu-Tseh Chi\*, University of Florida; Mohsen Ali, ; Muhammad Rushdi, ; Jeffrey Ho, University of Florida

**24 Sparse Variation Dictionary Learning for Face Recognition with A Single Training Sample Per Person**, Meng YANG\*, ETH Zurich; Luc Van Gool, ETH; Lei Zhang, The Hong Kong Polytechnic University

**25 On the Mean Curvature Flow on Graphs with Applications in Image and Manifold Processing**, El chakik Abdallah\*, Greyc Laboratory; abderrahim ELmoataz, ; ahcen Sadi,

**26 Perceptual Fidelity Aware Mean Squared Error**, Wufeng Xue, Xi'an Jiaotong University; Xuanqin Mou\*, Xi'an Jiaotong University; Lei Zhang, The Hong Kong Polytechnic University; Xiangchu Feng, School of Science, Xidian University

**27 Real-World Normal Map Capture for Nearly Flat Reflective Surfaces**, Bastien Jacquet\*, ETH Zurich; Christian Hne, ETH Zurich; Kevin Koeser, ; Marc Pollefeys, ETH

**28 Human Attribute Recognition By Rich Appearance Dictionary**, Jungseock Joo\*, UCLA; Shuo Wang, ; Song Chun Zhu, UCLA

**29 Deformable Part Descriptors for Fine-grained Recognition and Attribute Prediction**, Ning Zhang\*, EECS, UC Berkeley; Ryan Farrell, ICSI, UC Berkeley; Trevor Darrell,

**30 Handling Uncertain Tags in Visual Recognition**, Arash Vahdat\*, Simon Fraser University; Greg Mori, Simon Fraser University

**31 Implied Feedback: Learning Nuances of User Behavior in Image Search**, Devi Parikh\*, Virginia Tech; Kristen Grauman, University of Texas at Austin

**32 Learning Near-Optimal Cost-Sensitive Decision Policy for Object Detection**, Tianfu Wu\*, UCLA; Song Chun Zhu, UCLA

**33 NYC3DCars: A Dataset of 3D Vehicles in Geographic Context**, Kevin Matzen\*, Cornell University; Noah Snavely, "Cornell, USA"

**34 Unsupervised Domain Adaptation by Domain Invariant Projection**, Mahsa Baktashmotlagh\*, University of Queensland; Mehrtash Harandi, NICTA; Brian Lovell, ; Mathieu Salzmann, NICTA

**35 SIFTpack: a compact representation for efficient SIFT matching**, Alexandra Gilinsky\*, Technion; Lihi Zelnik-Manor, "Technion, Israel"

**36 PhotoOCR: Reading Text in Uncontrolled Conditions**, Mark Cummins\*, Google; Alessandro Bissacco, Google Inc.; Yuval Netzer, Google; Hartmut Neven, Google

**37 Probabilistic Elastic Part Model for Unsupervised Face Detector Adaptation**, Haoxiang Li, Stevens Institute of Technology; Gang Hua\*, Stevens Institute of Technology; Zhe Lin, Adobe Research; Jonathan Brandt, Adobe; Jianchao Yang, Adobe Systems Inc.

**38 Multi-Label Image Annotations Using New Graph Structured Sparsity Model**, xiao Cai, University of Texas at Arlington; Feiping Nie, University of Texas at Arlington; Heng Huang\*, UTA

**39 Heterogeneous Auto-Similarities of Characteristics (HASC): exploiting relational information for classification**, Marco San

Biagio\*, IIT; Marco Crocco, IIT; Marco Cristani, IIT; Samuele Martelli, IIT; Vittorio Murino, Istituto Italiano di Tecnologia

**40 A Fully Hierarchical Approach for Finding Correspondences in Non-rigid Shapes**, Ivan Sipiran\*, Department of Computer and Information Science - University of Konstanz; Benjamin Bustos, University of Chile

**41 Learning to Rank Using Privileged Information**, Viktoriia Sharmanska\*, IST Austria; Novi Quadrianto, University of Cambridge; Christoph Lampert, Institute of Science and Technology Austria

**42 Joint Segmentation and Pose Tracking of Human in Natural Videos**, Taegyu Lim\*, Samsung; Seunghoon Hong, POSTECH; Bohyung Han, POSTECH; JoonHee Han, POSTECH

**43 Characterizing Layouts of Outdoor Scenes Using Spatial Topic Processes**, Dahua Lin\*, TTIC; Jianxiong Xiao, MIT

**44 Image Co-Segmentation via Consistent Functional Maps**, Fan Wang\*, Stanford University; Qixing Huang, Stanford University; Leonidas Guibas, Stanford University

**45 Exemplar Cut**, Jimei Yang\*, UC Merced; Yi-Hsuan Tsai, UC Merced; Ming-Hsuan Yang, "UC Merced, USA"

**46 Parallel Transport of Deformations in Shape Space of Elastic Surfaces**, QIAN XIE\*, Florida State University; Sebastian Kurtek, Ohio State University; Huiling Le, ; Anuj Srivastava,

**47 A Method of Perceptual-based Shape Decomposition**, Chang Ma, Peking University; Zhongqian Dong, Peking University; Tingting Jiang, Peking University; yizhou Wang\*, Peking University

**48 Curvature-aware Regularization on Riemannian Submanifolds**, Kwang In Kim\*, MPI for Informatics; James Tompkin, MPI Informatik; Christian Theobalt, MPI fuer Informatik



**49 Linear Sequence Discriminant Analysis: A Model-Based Dimensionality Reduction Method for Vector Sequences**, Bing Su\*, Tsinghua University; Xiaoqing Ding, Tsinghua University

**50 Frustratingly Easy NBNN Domain Adaptation**, Tatiana Tommasi\*, Tatiana Tommasi KU Leuven, Belgium; Barbara Caputo, University of Rome La Sapienza, Italy

**51 Video Event Understanding using Natural Language Descriptions**, Vignesh Ramanathan\*, Stanford University; Percy Liang, Stanford University; Fei-Fei Li, Stanford University

**52 ACTIVE: Activity Concept Transitions in Video Event Classification**, Chen Sun\*, University of Southern Califor; Ram Nevatia,

**19:00 IEEE PAMI Technical Committee Meeting**

# **Main Conference Wednesday, December 4**

## **Day 2:**

**9:00 -10:00 Oral session 2A ( 4 papers 60 min) Low level vision**

**Orals 2A:01 – 2A:03**

**1 Benchmarking Computational Model of Visual Saliency**,  
Ali Borji\*, ; Dicky Sihite, University of Southern California (USC);  
Hamed Rezazadegan Tavakoli, University of Oulu; Laurent Itti,  
University of Southern California (USC)

**2 A Color Constancy Model with Double-Opponency Mechanisms**,  
Shaobing Gao, UESTC; Kaifu Yang, UESTC; Yongjie Li\*, UESTC

**3 Towards Guaranteed Illumination Models for Non-Convex Objects**,  
Yuqian Zhang\*, Columbia University; Cun Mu, Columbia  
University; Han-wen Kuo, Columbia University; John Wright,

**4 Nonparametric Blind Super-Resolution**,  
Tomer Michaeli\*, Weizmann Institute of Science; Michal Irani,  
Weizmann Institute, Israel

**10:00 - 10:45 Spotlights (45min, 52 spotlights at 50secs each)**

**10:45 - 12:15 Poster session 2A: Posters 2A:01 – 2A:52**

**1 Street View Structure-from-Motion**, Bryan Klingner\*, Google;  
David Martin, Google; James Roseborough, Google

**2 A Robust Analytical Solution to Isometric Non-Rigid Pose with Focal Length Calibration**, Adrien Bartoli\*, Université d'Auvergne; Daniel Pizarro, ALCoV-ISIT; Toby Collins, ALCoV-ISIT

**3 Point-Based 3D Reconstruction of Thin Objects**, Benjamin Ummenhofer\*, University of Freiburg; Thomas Brox,

**4 Space-Time Tradeoffs in Photo Sequencing**, Tali Basha Dekel\*, TAU; Shai Avidan, Tel-Aviv University; Yael Moses,

**5 An Enhanced Structure-from-Motion Paradigm based on the Absolute Dual Quadric and Images of Circular Points**, Lilian Calvet\*, University of Toulouse; Pierre Gurdjos, IRIT

**6 Image Guided Depth Upsampling using Anisotropic Total Generalized Variation**, David Ferstl\*, Graz University of Technology; Christian Reinbacher, ; Matthias Rother, ; Horst Bischof, Graz University of Technology

**7 Fluttering Pattern Generation using Modified Legendre Sequence for Coded Exposure Imaging**, Hae-Gon Jeon, KAIST; Joon-Young Lee, KAIST; Yudeog Han, KAIST; Seon Joo Kim, Yonsei University; In So Kweon\*, "KAIST, Korea"

**8 Towards Motion-Aware Light Field Video for Dynamic Scenes**, Salil Tambe, Rice University; Ashok Veeraraghavan\*, Rice University; Amit Agrawal, MERL

**9 Handwritten Word Spotting with Corrected Attributes**, Jon Almazn\*, Computer Vision Center; Albert Gordo, INRIA; Alicia Forns, Computer Vision Center; Ernest Valveny, Computer Vision Center

**10 Exemplar-based Graph Matching for Robust Facial Landmark Localization**, Feng Zhou\*, Carnegie Mellon University; Jonathan Brandt, Adobe; Zhe Lin, Adobe Research

**11 Cascaded Shape Space Pruning for Robust Facial Landmark Detection**, Xiaowei Zhao\*, ICT,CAS; Shiguang Shan, "Chinese Academy of Sciences, China"; Xiujuan Chai, jdl; Xilin Chen,

**12 Two-Point Gait: Decoupling Gait from Body Shape**, Stephen Lombardi\*, Drexel University; Ko Nishino, "Drexel University, USA"; Yasushi Makihara, Osaka university; Yasushi Yagi,

**13 Learning People Detectors for Tracking in Crowded Scenes**, Siyu Tang\*, Max planck institute ; Mykhaylo Andriluka, Max Planck Institute for Informatics; Anton Milan, TU Darmstadt; Konrad Schindler, ETH Zurich; Stefan Roth, TU Darmstadt, Bernt Schiele, MPI Informatics, Germany

**14 Efficient pedestrian detection by directly optimizing the partial area under the ROC curve**, Sakrapee Paisitkriangkrai\*, The University of Adelaide; Chunhua Shen, The University of Adelaide; Anton Van den Hengel, University of Adelaide

**15 Example-based Facade Texture Synthesis**, Dengxin Dai\*, CVL, ETH Zurich; Hayko Riemenschneider, CVL, ETH Zurich; Luc Van Gool, ETH; gerhard Schmitt, ETH Zurich

**16 Single-patch low-rank prior for non-pointwise impulse noise removal**, Ruixuan Wang\*, University of Dundee; Emanuele Trucco, University of Dundee

**17 SGTD: Structure Gradient and Texture Decorrelating Regularization for Image Decomposition**, Qiegen Liu, ; Jianbo Liu, ; Pei Dong, ; Dong Liang\*, Shenzhen Institutes of Advance

**18 Drosophila Embryo Stage Annotation using Label Propagation**, Tomas Kazmar\*, IMP/ISTA; Evgeny Kvon, IMP; Alexander Stark, IMP; Christoph Lampert, Institute of Science and Technology Austria

**19 Measuring Flow Complexity in Videos**, Saad Ali\*,

**20 Real-time Body Tracking with One Depth Camera and Inertial Sensors**, Thomas Helten\*, MPI Informatik; Meinard Mller, International Audio Laboratories Erlangen; Hans-Peter Seidel, MPI Informatik; Christian Theobalt, MPI fuer Informatik

**21 Constructing Adaptive Complex Cells for Robust Visual Tracking**, Dapeng Chen\*, Xi'an Jiaotong University; Zejian Yuan, Xi'an jiaotong University; Yang Wu, Kyoto University; Geng Zhang, Xi'an Jiaotong University; Nanning Zheng, Xi'an Jiaotong University

**22 Camera Alignment using Trajectory Intersections in Unsynchronized Videos**, Thomas Kuo\*, UC Santa Barbara; Santhoshkumar Sunderrajan, UCSB; B.S. Manjunath, UCSB

**23 Slice Sampling Particle Belief Propagation**, Oliver Mueller\*, Leibniz University Hannover; Michael Ying Yang, Leibniz University Hannover; Bodo Rosenhahn, Leibniz University Hannover

**24 Multi-Attributed Dictionary Learning for Sparse Coding**, Chen-Kuo Chiang\*, National Tsing Hua University; Te-Feng Su, ; Yen Chih, ; Shang-Hong Lai, NTHU

**25 Semi-Supervised Robust Dictionary Learning via Efficient  $L_{2,0}+L_1$ -Norms Minimization**, Hua Wang\*, Colorado School of Mines; Feiping Nie, University of Texas at Arlington; Heng Huang, UTA

**26 Saliency and Human Fixations: State-of-the-art and Study of Comparison Metrics**, Nicolas Riche\*, UMONS; Matthieu Duvinage, UMONS; Matei Mancas, UMONS; Bernard Gosselin, UMONS; Thierry Dutoit, UMONS

**27 Multiview Photometric Stereo using Planar Mesh Parameterization**, Jaesik Park, KAIST; Sudipta Sinha, ; Yasuyuki Matsushita, Microsoft Research Asia; Yu-Wing Tai, "KAIST, Korea"; In So Kweon\*, KAIST

**28 Elastic Net Constraints for Shape Matching**, Emanuele Rodola\*, The University of Tokyo; Andrea Torsello, ; Tatsuya Harada, University of Tokyo; Yasuo Kuniyoshi, The University of Tokyo; Daniel Cremers, Technical University Munich

**29 No Matter Where You Are: Flexible Graph-guided Multi-task Learning for Multi-view Head Pose Classification Under Target Motion**, Yan Yan\*, University of Trento; Elisa Ricci, University of Perugia; Ramanathan Subramanian, Advanced Digital Sciences Center in Singapore; Oswald Lanz, FBK Fondazione Bruno Kessler; nicu Sebe, University of Trento

**30 Compositional Models for Video Event Detection: A Multiple Kernel Learning Latent Variable Approach**, Arash Vahdat\*, Simon Fraser University; Kevin Cannons, Simon Fraser University; Greg Mori, Simon Fraser University; Sangmin Oh, Kitware Inc.; Ilseo Kim,

**31 Event Recognition in Photo Collections with a Stopwatch HMM**, Lukas Bossard\*, ETH Zurich; Matthieu Guillaumin, ETH Zurich; Luc Van Gool, ETH Zurich

**32 Nested Shape Descriptors**, Jeffrey Byrne\*, University of Pennsylvania

**33 Collaborative Active Learning of a Kernel Machine Ensemble for Recognition**, Gang Hua\*, Stevens Institute of Technology; Chengjiang Long, Stevens Institute of Technology; Ming Yang, NEC Labs America; Yan Gao, Northwestern University

**34 Look Into Sparse Representation-based Classification: A Margin-based Perspective**, Zhaowen Wang\*, UIUC; Jianchao Yang, Adobe Systems Inc.; Nasser Nasrabadi, US Army Research Lab; Thomas Huang, University of Illinois at Urbana-Champaign

**35 Attribute Dominance: What Pops Out?**, Naman Turakhia, ; Devi Parikh\*, Virginia Tech

**36 Neighbor-To-Neighbor Search for Fast Coding of Feature Vectors**, Nakamasa Inoue\*, Tokyo Institute of Technology; Koichi Shinoda, Tokyo Institute of Technology

**37 Text Localization in Natural Images using Stroke Feature Transform and Text Covariance Descriptors**, Weilin Huang, Adobe Research; Zhe Lin, Adobe Research; Jianchao Yang, Adobe Systems Inc.; Jue Wang\*,

**38 A Framework for Shape Analysis via Hilbert Space Embedding**, Sadeep Jayasumana\*, ANU; Mathieu Salzmann, NICTA; Hongdong Li, Australia National University; Mehrtash Harandi, NICTA

**39 Offline Mobile Instance Retrieval with a Small Memory Footprint**, Jayaguru Panda\*, IIIT Hyderabad; Michael Brown, National University of Singapore; c. v. Jawahar, IIIT Hyderabad

**40 BOLD features to detect texture-less objects**, Federico Tombari\*, University of Bologna; Alessandro Franchi, ; Luigi Di Stefano,

**41 Estimating the 3D Layout of Indoor Scenes and its Clutter from Depth Sensors**, Jian Zhang, ; Kan Chen, ; Alexander Schwing\*, ETH Zurich; Raquel Urtasun, Toyota Technological Institute at Chicago

**42 A Non-parametric Bayesian Network Prior of Human Pose**, Andreas Lehrmann\*, MPI for Intelligent Systems; Peter Gehler, Max Planck; Sebastian Nowozin, Microsoft Research Cambridge

**43 Pictorial Human Spaces: How Well do Humans Perceive a 3D Articulated Pose?**, Elisabeta Marinoiu\*, Romanian Academy of Science; Dragos Papava, Institute of Mathematics of the Romanian Academy; Cristian Sminchisescu, Lund University

**44 Co-Segmentation by Composition**, Alon Faktor\*, Weizmann Institute of Science; Michal Irani, Weizmann Institute, Israel

**45 Cosegmentation and Cosegmentation by Unsupervised Learning**, Jifeng Dai\*, Tsinghua University; UCLA; Ying Nian Wu, UCLA; Jie Zhou, Tsinghua University; Song Chun Zhu, UCLA

**46 Predicting Sufficient Annotation Strength for Interactive Foreground Segmentation**, Suyog Jain\*, ; Kristen Grauman, University of Texas at Austin

**47 Sequential Bayesian Model Update under Structured Scene Prior for Semantic Road Scenes Labeling**, Evgeny Levinkov, MPII; Mario Fritz\*, MPI Informatics

**48 Learning Graph Matching for Category Modeling from Large Scenes**, Quanshi Zhang\*, University of Tokyo; Xuan Song, University of Tokyo; Xiaowei Shao, University of Tokyo; Ryosuke Shibasaki, University of Tokyo; Huijing Zhao, Peking University

**49 Robust Matrix Factorization with Unknown Noise**, Deyu Meng\*, Xi'an Jiaotong University; Fernando de la Torre, Carnegie Mellon University

**50 Correlation Adaptive Subspace Segmentation by Trace Lasso**, Canyi Lu\*, National University of Singapore; Jiashi Feng, NUS; Zhouchen Lin, Peking University ; Shuicheng Yan, "NUS, Singapore"

**51 Monte Carlo Tree Search for Scheduling Activity Recognition**, Mohamed Amer\*, OREGON STATE UNIVERSITY; Sinisa Todorovic, "Oregon State University, USA"; Alan Fern, Oregon State University; Song Chun Zhu, UCLA

**52 Manipulation Pattern Discovery: A Nonparametric Bayesian Approach**, Bingbing NI\*, Advanced Digital Sciences Center (ADSC), Singapor; Pierre Moulin, UIUC

### **12:15 - 14:15 lunch**

### **14:15 - 15:00 Oral session 2B (3 papers 45 mins)**

#### **Motion and Tracking**

#### **Orals 2B:01 - 2B:03**

**1 Perspective Motion Segmentation via Collaborative Clustering**, Zhuwen Li\*, NUS; Jiaming Guo, NUS; Loong-Fah Cheong, NUS; Zhiying Zhou, NUS

#### **2 Piecewise Rigid Scene Flow,**

Christoph Vogel\*, ETH Zurich; Konrad Schindler, ETH Zurich; Stefan Roth, TU Darmstadt

#### **3 Large displacement optical flow with deep matching,**

Philippe Weinzaepfel\*, INRIA; Jerome Revaud, ; Zaid Harchaoui, INRIA; Cordelia Schmid, "INRIA, France



**15:00 - 15:45 Spotlights (45mins)**

**15:45 - 16:15 Break (30 min)**

**16:15 – 17:15 Oral session 2C (4 papers 60 mins)**

**Recognition**

**Orals 2C:01 – 2C:04**

**1 Shufflets: shared mid-level parts for fast multi-category detection,**

Iasonas Kokkinos\*, "Ecole Centrale Paris, France"

**2 To aggregate or not to aggregate: Selective match kernels for image search,**

Giorgos Tolias, ; Yannis Avrithis, NTUA; Herv Jgou\*, INRIA

**3 NEIL: Extracting Visual Knowledge from Web Data,**

Xinlei Chen, CMU; Abhinav Shrivastava, Carnegie Mellon University; Abhinav Gupta\*,

**4 Holistic Scene Understanding for 3D Object Detection with RGBD cameras,**

Dahua Lin\*, TTIC; Sanja Fidler, TTI Chicago; Raquel Urtasun, Toyota Technological Institute at Chicago

**17:15 – 18:45 Poster session 2B: Posters 2B:01 – 2B:52**

**1 3D Scene Understanding by Voxel-CRF, Byung-soo Kim\*, ;**

Pushmeet Kohli, "Microsoft Research, UK"; Silvio Savarese, "University of Michigan, USA"

**2 Complex 3D General Object Reconstruction from Line Drawings, Linjie Yang\*, CUHK; Jianzhuang Liu, CUHK; Xiaouu**

Tang, Chinese University of Hong Kong

**3 Subpixel scanning invariant to indirect lighting using quadratic code length**, Nicolas Martin\*, Université de Montréal; Vincent Couture, Université de Montréal; Sebastien Roy, Université de Montréal

**4 Semi-Dense Visual Odometry for a Monocular Camera**, Jakob Engel\*, Technical University Munich; Juergen Sturm, TU Munich; Daniel Cremers, Technische Universität München

**5 Globally-Optimal ICP: Solving the 3D Registration Problem Efficiently and Globally Optimally**, Jiaolong Yang, Beijing Inst. of Tech.; Hongdong Li\*, Australia National University; Yunde Jia,

**6 Forward Motion Deblurring**, Shicheng Zheng\*, CUHK; Li Xu, CUHK; Jiaya Jia, Chinese University of Hong Kong

**7 Fibonacci Exposure Bracketing for High Dynamic Range Imaging**, Mohit Gupta\*, Daisuke Iso, Columbia University; Shree Nayar, Columbia University

**8 Compensating for Motion During Direct-Global Separation**, Supreeth Achar\*, Carnegie Mellon University; Stephen Nuske, Carnegie Mellon University; Srinivasa G. Narasimhan, Carnegie Mellon University

**9 Hybrid Deep Learning for Computing Face Similarities**, Yi Sun\*, CUHK; Xiaogang Wang, "The Chinese University of Hong Kong, Hongkong"

**10 Like Father, Like Son: Facial Expression Dynamics for Kinship Verification**, Hamdi Dibeklioglu\*, University of Amsterdam; Albert Salah, Bogazici University; Theo Gevers, University of Amsterdam

**11 Handling Occlusions with Franken-classifiers**, Markus Mathias\*, KU Leuven; Rodrigo Benenson, MPI-Inf; Radu Timofte, KU Leuven; Luc Van Gool, KU Leuven

**12 Robust face landmark estimation under occlusion**, Xavier Burgos Artizzu\*, Caltech; Pietro Perona, "Caltech, USA"; Piotr Dollar,

**13 Fingerspelling recognition with semi-Markov conditional random fields**, Taehwan Kim, TTIC; Gregory Shakhnarovich, TTIC; Karen Livescu\*, TTIC

**14 Efficient Salient Region Detection with Soft Image Abstraction**, Ming-Ming Cheng\*, Oxford Brookes University; Shuai Zheng, Oxford Brookes University; Jonathan Warrell, Oxford Brookes University; Vibhav Vineet, Oxford Brookes University; Wenyan Lin, Oxford Brookes University

**15 Cross-Field Joint Image Restoration via Scale Map**, Qiong Yan\*, CUHK; Xiaoyong Shen, CUHK; Li Xu, CUHK; Shaojie Zhuo, qualcomm.com; Xiaopeng Zhang, ; Liang Shen, ; Jiaya Jia, Chinese University of Hong Kong

**16 A Joint Intensity and Depth Co-Sparse Analysis Model for Depth Map Super-Resolution**, Martin Kiechle\*, Technische Universitt Mnchen; Simon Hawe, Technische Universitt Mnchen; Martin Kleinsteuber, Technische Universitt Mnchen

**17 Detecting Irregular Curvilinear Structures in Gray Scale and Color Imagery using Multi-Directional Oriented Flux**, Engin Turetken\*, EPFL; Carlos Becker, EPFL; Przemyslaw Glowacki, EPFL; fethallah Benmansour, ; Pascal Fua, "EPFL, Switzerland"

**18 STAR3D: Simultaneous Tracking And Reconstruction of 3D Objects Using RGB-D Data**, Yuheng (Carl) Ren\*, Oxford University; Victor Prisacariu, Oxford; David Murray, Oxford; Ian Reid, University of Adelaide

**19 Discriminant Tracking Using Tensor Representation with Semi-supervised Improvement**, Jin Gao\*, Institute of Automation Chinese Academy of Sciences; Junliang Xing, Institute of Automation, Chinese Academy of Sciences ; Weiming Hu, Institute of Automation Chinese Academy of Sciences; Steve Maybank, Birkbeck College, London

**20 Coherent motion segmentation in moving camera videos using optical flow orientations**, Manjunath Narayana\*, ; Erik Learned-Miller, University of Massachusetts at Amherst; Allen Hanson, University of Massachusetts Amherst

**21 Minimal Basis Facility Location for Subspace Segmentation**, Choon Meng Lee\*, NUS; Loong-Fah Cheong, NUS

**22 Non-Convex P-norm Projection for Robust Sparsity**, Mithun Das Gupta\*, Ricoh Innovations Pvt. Ltd.; Sanjeev Kumar, Qualcomm

**23 Log-Euclidean Kernels for Sparse Representation and Dictionary Learning**, Peihua Li\*, Dalian University of Technology; Qilong Wang, ; Wangmeng Zuo, Harbin Institute of Technology; Lei Zhang, The Hong Kong Polytechnic University

**24 Large-scale Image Annotation by Efficient and Robust Kernel Metric Learning**, Zheyun Feng\*, Michigan State University; Rong Jin, Michigan State University; Anil Jain, Michigan State University

**25 High Quality Shape from a single RGB-D Image under Uncalibrated Natural Illumination**, Yudeog Han, KAIST Joon-Young Lee, KAIST; In So Kweon\*, "KAIST, Korea" ,

**26 Semantic RGB-D Bundle Adjustment with Human in the Loop**, Jianxiong Xiao\*, MIT; Andrew Owens, MIT; Antonio Torralba, MIT

**27 The Interestingness of Images**, Michael Gygli\*, ETH Zurich; Helmut Grabner, ETH Zurich; Hayko Riemenschneider, CVL, ETH Zurich; Fabian Nater, upicto GmbH; Luc Van Gool, ETH

**28 Hierarchical Part Matching for Fine-Grained Visual Categorization**, Lingxi Xie\*, Tsinghua University; Qi Tian,

University of Texas at San Antonio; Shuicheng Yan, "NUS, Singapore"; Bo Zhang, Tsinghua University

**29 Joint optimization for consistent multiple graph matching,** Junchi Yan\*, Shanghai Jiao Tong University; Yu Tian, Shanghai Jiao Tong University; Hongyuan Zha, Georgia Tech; Xiaokang Yang, ; Ya Zhang, Shanghai Jiao Tong University

**30 Unbiased Metric Learning: On the Utilization of Multiple Datasets and Web Images for Softening Bias,** Chen Fang\*, Dartmouth College; Ye Xu, Dartmouth College; Daniel Rockmore, Dartmouth College

**31 Saliency Detection via Absorbing Markov Chain,** Bowen Jiang\*, DUT; Lihe Zhang, DUT, China; huchuan Lu, DUT,China; Ming-Hsuan Yang, "UC Merced, USA"; Chuan Yang,

**32 Semantic-aware Co-indexing for Near-duplicate Image Retrieval,** Shiliang Zhang, UTSA; Ming Yang\*, NEC Labs America; Xiaoyu Wang, NEC Labs America; Yuanqing Lin, ; Qi Tian, University of Texas at San Antonio

**33 Learning the Visual Interpretation of Sentences,** Larry Zitnick\*, "Microsoft Research, USA"; Devi Parikh, Virginia Tech; Lucy Vanderwende, Microsoft Research

**34 A Novel Earth Mover's Distance Methodology for Image Matching with Gaussian Mixture Models,** Peihua Li\*, Dalian University of Technology; Qilong Wang, ; Lei Zhang, The Hong Kong Polytechnic University

**35 Hierarchical Joint Max-Margin Learning of Mid and Top Level Representations for Visual Recognition,** Hans Lbel\*, Universidad Catlica de Chile; Rene Vidal, Johns Hopkins University; Ivaro Soto, Universidad Catlica de Chile

**36 Query-adaptive asymmetrical dissimilarities for visual object retrieval,** Cai-Zhi Zhu\*, National Institute of Informatics; Herv Jgou, INRIA; shin'ichi satoh, NII

**37 Fine-Grained Categorization by Alignments**, Efstratios Gavves\*, University of Amsterdam; Basura Fernando, KU Leuven; Cees Snoek, University of Amsterdam; Arnold Smeulders, ; Tinne Tuytelaars, KU Leuven

**38 Locally Affine Sparse-to-Dense Matching for Motion and Occlusion Estimation**, Marius Leordeanu\*, Institute of Mathematics of the Romanian Academy; Andrei Zanfir, ; Cristian Sminchisescu, Lund University

**39 Predicting an Object Location using a Global Image Representation**, Jose Rodriguez-Serrano\*, ; Diane Larlus,

**40 Heterogeneous Image Feature Integration via Multi-Modal Semi-Supervised Learning for Image Categorization**, xiao Cai\*, University of Texas at Arlington; Feiping Nie, University of Texas at Arlington; Heng Huang, UTA

**41 Building Parts-based Object Detectors via 3D Geometry**, Abhinav Shrivastava, Carnegie Mellon University; Abhinav Gupta\*,

**42 Detecting Curved Symmetric Parts using a Deformable Disc Model**, Tom Lee\*, University of Toronto; Sanja Fidler, TTI Chicago; Sven Dickinson, University of Toronto

**43 Category-Independent Object-level Saliency Detection**, Yangqing Jia\*, UC Berkeley; Mei Han, Google Research

**44 GrabCut in One Cut**, Meng Tang, University of Western Ontario; Lena Gorelick\*, University of Western Ontario; Olga Veksler, University of Western Ontario; Yuri Boykov, "University of Western Ontario, Canada"

**45 Fast object segmentation in unconstrained video**, Anestis Papazoglou\*, University of Edinburgh; Vittorio Ferrari, University of Edinburgh

**46 Bayesian Robust Matrix Factorization for Image and Video Processing**, Naiyan Wang\*, HKUST; Dit-Yan Yeung, HKUST

**47 Recursive Estimation of the Stein Center of SPD Matrices and its Applications**, Hesamoddin Salehian\*, University of Florida; Guang Cheng, ; Baba Vemuri, "University of Florida, USA"; Jeffrey Ho, University of Florida

**48 Correntropy Induced L2 Graph for Robust Subspace Clustering**, Canyi Lu\*, National University of Singapo; Zhouchen Lin, Peking University; Shuicheng Yan, "NUS, Singapore"

**49 Group Sparsity and Geometry Constrained Dictionary Learning for Action Recognition from Depth Maps**, Jiajia Luo\*, The University of Tennessee; Wei Wang; Hairong Qi,

**50 Action and Event Recognition with Fisher vectors on a Compact Feature Set**, Dan Oneata\*, INRIA; Jakob Verbeek, "INRIA, France"; Cordelia Schmid, "INRIA, France"

**51 Stable hyper-pooling and query expansion for event detection**, Matthijs Douze, INRIA; Jerome Revaud, ; Cordelia Schmid, "INRIA, France"; Herv Jgou\*, INRIA

**52 Active Learning of an Action Detector from Untrimmed Videos**, Sunil Bandla, University of Texas at Austin; Kristen Grauman\*, University of Texas at Austin

**18:45-19:00 Award Session**

**19:00: Welcome Reception**

# **Main Conference    Thursday, December 5**

## **Day 3:**

**9:00 - 10:00 Oral session 3A (4 papers, 60min)**

**Recognition + Segmentation**

**Orals 3A:01 - 3A:04**

**1 Structured Forest for Fast Edge Detection,**

Piotr Dollar\*, Larry Zitnick, "Microsoft Research, USA"

**2 Weakly supervised learning of image partitioning using decision trees with structured split criteria,**

Christoph Straehle\*, HCI, University of Heidelberg; Ullrich Koethe,  
; Fred Hamprecht, HCI, University of Heidelberg

**3 Style-aware Mid-level Representation for Discovering Visual Connections in Space and Time,**

Yong Jae Lee\*, Robotics Institute, Carnegie Mellon University;  
Alexei (Alyosha) Efros, CMU; Martial Hebert, "CMU, USA"

**4 Coarse-to-fine Semantic Video Segmentation using Supervoxel Trees,**

Rene Vidal\*, Johns Hopkins University; Aastha Jain, LinkedIn

**10:00 - 10:45 Spotlights (45min, 52 spotlights at 50secs each)**

**10:45 - 12:15 Poster session 3A: Posters 3A:01 - 3A:52**

**1 3DNN: Viewpoint Invariant 3D Geometry Matching for Scene Understanding,**  
Scott Satkin\*, google, Carnegie Mellon University

**2 Local Signal Equalization for Correspondence Matching,**

Derek Bradley\*, Disney Research Zurich; Thabo Beeler, Disney Research Zurich



**3 Monocular Image 3D Human Pose Estimation under Self-Occlusion**, Ibrahim Radwan\*, University of Canberra; Abhinav Dhall, Australian National University; Roland Goecke,

**4 Multi-View 3D Reconstruction from Uncalibrated Radially-Symmetric Cameras**, Jae-Hak Kim, Australian National University; Yuchao Dai, Australian National University; Hongdong Li, Australian National University and NICTA; Xin Du, Zhejiang University; Jonghyuk Kim, Australian National University

**5 Corrected-Moment Illuminant Estimation**, Graham Finlayson\*, University of East Anglia

**6 Target-Driven Moire Pattern Synthesis by Phase Modulation**, Pei-Hen Tsai, National Taiwan University; Yung-Yu Chuang\*, National Taiwan University

**7 Anchored Neighborhood Regression for Fast Example-Based Super-Resolution**, Radu Timofte\*, KU Leuven; Vincent De Smet, KU Leuven; Luc Van Gool, KU Leuven

**8 Multi-Scale Topological Features for Hand Posture Representation and Analysis**, Kaoning Hu\*, SUNY - Binghamton; Lijun Yin, Binghamton University State University of New York

**9 Sieving Regression Forest Votes for Facial Feature Detection in the Wild**, Heng Yang\*, Queen Mary Uni. of London; Ioannis Patras,

**10 Pose-free Facial Landmark Fitting via Optimized Part Mixtures and Cascaded Deformable Shape Model**, Xiang Yu\*, Rutgers University; Junzhou Huang, University of Texas at Arlington; Shaoting Zhang, Rutgers University; Wang Yan, Rutgers University; Dimitris Metaxas, Rutgers University

**11 Markov Network-based Unified Classifier for Face Identification**, Wonjun Hwang\*, Samsung AIT; Kyungshik Roh, ; Junmo Kim, KAIST

**12 Fast High Dimensional Vector Multiplication Based Face Recognition**, Oren Barkan\*, Tel Aviv University; Jonathan Weill, Tel Aviv University; Lior Wolf, "Tel Aviv University, Israel"; Hagai Aronowitz, IBM Research;

**13 A Generalized Low-Rank Appearance Model for Spatio-Temporally Correlated Rain Streaks**, Yi-Lei Chen\*, NTHU, Taiwan; Chiou-Ting Hsu, NTHU, Taiwan

**14 Salient Region Detection by UFO: Uniqueness, Focusness and Objectness**, Peng Jiang\*, Shandong University; Jingliang Peng, cs.sdu.edu.cn; Haibin Ling,

**15 Estimating the Material Properties of Fabric from Video**, Katherine L. Bouman, MIT; Bei Xiao, MIT; Peter Battaglia, MIT; Bill Freeman MIT

**16 Multiple Non-Rigid Surface Detection and Registration**, Yi Wu\*, UC Merced; Yoshihisa Ijiri, OMRON; Ming-Hsuan Yang, "UC Merced, USA"

**17 Discriminative Label Propagation for Multi-object Tracking with Sporadic Appearance Features**, Amit K.C.\*, Universit catholique de Louva; Christophe De Vleeschouwer, UCL

**18 Online Motion Segmentation using Dynamic Label Propagation**, Ali Elqursh\*, Rutgers University; Ahmed Elgammal,

**19 A Unified Rolling Shutter and Motion Blur Model for Dense 3D Visual Tracking**, Maxime Meilland\*, I3S-CNRS-UNS; Andrew Comport, CNRS-I3S/UNS; Tom Drummond, Monash University

**20 Shortest Paths with Curvature and Torsion**, Petter Strandmark\*, Lund University; Johannes Ulen, Lund University; Fredrik Kahl, Lund University; Leo Grady, HeartFlow Inc.

**21 Bounded Labeling Function for Global Segmentation of Multi-Part Objects with Geometric Constraints**, Masoud S. Nosrati\*, Simon Fraser University; Shawn Andrews, SFU; Ghassan Hamarneh, SFU

**22 Randomized Ensemble Tracking**, Qinxun Bai\*, Boston University; Zheng Wu, ; Stan Sclaroff, Boston University; Margrit Betke, Boston University; Camille Monnier,

**23 Discriminatively Trained Templates for 3D Object Detection: A Real Time Scalable Approach**, Reyes Rios-Cabrera\*, KULeuven; Tinne Tuytelaars, KU Leuven

**24 Joint Deep Learning for Pedestrian Detection**, Wanli Ouyang\*, The Chinese University of HK; Xiaogang Wang, "The Chinese University of Hong Kong, Hongkong"

**25 Detecting avocados to zucchinis: what have we done, and where are we going?**, Olga Russakovsky\*, ; Jia Deng, Stanford University; Zhiheng Huang, Stanford University; Alexander Berg, Stony Brook University; Fei-Fei Li, Stanford University

**26 Ensemble Projections for Semi-supervised Image Classification**, Dengxin Dai\*, CVL, ETH Zurich; Luc Van Gool, ETH

**27 Latent Task Adaptation with Large-scale Hierarchies**, Yangqing Jia\*, UC Berkeley; Trevor Darrell,

**28 Learning Coupled Feature Spaces for Cross-modal Matching**, Kaiye Wang\*, NLPR; Ran He, NLPR, CASIA; Wei Wang, NLPR; Liang Wang, unknown; Tieniu Tan, "NLPR, China"

**29 CoDeL: An Efficient Human Co-detection and Labeling Framework**, Jianping Shi\*, CUHK; Renjie Liao, CUHK; Jiaya Jia, Chinese University of Hong Kong

**30 How Related Exemplars Help Complex Event Detection in Web Videos?**, Yi Yang\*, cmu.edu; Zhigang Ma, The University of Trento; Zhongwen Xu, Carnegie Mellon University; Shuicheng Yan, "NUS, Singapore"; Alexander Hauptmann, Carnegie Mellon University

**31 Pose Estimation and Segmentation of People in 3D Movies**, Karteek Alahari\*, ENS-Willow; Guillaume Seguin; Josef Sivic, Ecole Normale Suprieure; Ivan Laptev, "INRIA, France"

**32 A Unified Probabilistic Approach Modeling Relationships between Attributes and Objects**, Xiaoyang Wang\*, RPI; Qiang Ji,

**33 Fast Neighborhood Graph Search using Cartesian Concatenation**, Jing Wang, Peking University; Jingdong Wang\*, Microsoft Research Asia; Gang Zeng, Peking University; Shipeng Li,

**34 Codemaps Segment, Classify and Search Objects Locally**, Zhenyang Li\*, University of Amsterdam; Efstratios Gavves, University of Amsterdam; Koen van de Sande, ; Cees Snoek, University of Amsterdam; Arnold Smeulders,

**35 Support surface prediction in indoor scenes**, Ruiqi Guo\*, UIUC; Derek Hoiem, University of Illinois at Urbana-Champaign

**36 Take Your Favorite Cloud with You**, Kuan-Chuan Peng\*, Cornell University; Tsuhan Chen,

**37 Spoken Attributes: Mixing Binary and Relative Attributes to Say the Right Thing**, Amir Sadovnik\*, Cornell University; Andrew Gallagher, ; Devi Parikh, Virginia Tech; Tsuhan Chen,

**38 Image Segmentation with Cascaded Hierarchical Models and Logistic Disjunctive Normal Networks**, Mojtaba Seyedhosseini\*, SCI; Mehdi Sajjadi, SCI; Tolga Tasdizen, SCI

**39 Semantic Segmentation without Annotating Segments**, Wei Xia\*, NUS; Csaba Domokos, NUS; Jian Dong, NUS; Loong-Fah Cheong, NUS; Shuicheng Yan, "NUS, Singapore"

**40 Progressive Multigrid Eigensolvers for Multiscale Spectral Segmentation**, Michael Maire\*, California Institute of Technology; Stella Yu, University of California, Berkeley / ICSI

**41 Video Segmentation by Tracking Many Figure-Ground Segments**, Fuxin Li\*, Georgia Inst. of Tech.; Taeyoung Kim, Georgia Institute of Technology; Ahmad Humayun, ; David Tsai, Georgia Institute of Technology; James Rehg, Georgia Institute of Technology

**42 Transfer Feature Learning with Joint Distribution**

**Adaptation**, Mingsheng Long\*, Tsinghua University; Jianmin Wang, Tsinghua University; Guiguang Ding, Tsinghua University; Philip Yu, University of Illinois at Chicago

**43 Manifold based Image Synthesis from Sparse Samples**

Hongteng Xu\*, Georgia Tech; Hongyuan Zha, Georgia Tech

**44 Robust Dictionary Learning by Error Source Decomposition**

Zhuoyuan Chen\*, Northwestern University; Ying Wu, Northwestern University

**45 Inferring "Dark Matter" and "Dark Energy" from Videos**

Dan Xie, University of California Los Angeles; Sinisa Todorovic\*, "Oregon State University, USA"; Song Chun Zhu, UCLA

**46 Video Co-segmentation for Meaningful Action Extraction**

Jiaming Guo\*, NUS; Zhuwen Li, NUS; Loong-Fah Cheong, NUS; Zhiying Zhou, NUS

**47 Flattening Supervoxel Hierarchies by the Uniform Entropy Slice**

Chenliang Xu\*, SUNY at Buffalo; Spencer Whitt, SUNY at Buffalo; Jason Corso, "SUNY Buffalo, USA"

**48 From Actemes to Action: A Strongly-supervised**

**Representation for Detailed Action Understanding**, Weiyu

Zhang\*, University of Pennsylvania; Konstantinos Derpanis, University of Pennsylvania; Menglong Zhu, University of Pennsylvania

**49 From Semi-Supervised to Transfer Counting of Crowds**

Chen Change Loy\*, CUHK; Shaogang Gong, EECS, QMUL; Tao Xiang, EECS, QMUL

**50 Learning to Share Latent Tasks for Action Recognition**

Qiang Zhou\*, NUS, Singapore; Gang Wang, NTU; Qi Zhao, National Univ. of Singapore

**51 Large Scale Video Hashing via Structure Learning**

Guangnan Ye\*, Columbia University; Dong Liu, Columbia University; Jun

Wang, IBM T. J. Watson Research; Shih-Fu Chang, Columbia University

**52 Finding Actors and Actions in Movies**, Piotr Bojanowski\*, INRIA; Francis Bach, "ENS and INRIA, France"; Ivan Laptev, "INRIA, France"; Jean Ponce, "ENS, France"; Cordelia Schmid, "INRIA, France"; Josef Sivic, Ecole Normale Suprieure

**12:15 - 14:15 Lunch**

**14:15 - 15:00 Oral session 3B (3 papers, 45mins)**  
**Motion and Tracking**

**Orals 3B:01 – 3B:03**

**1 Hierarchical Data-driven Descent for Efficient Optimal Deformation Estimation,**

Yuandong Tian\*, Carnegie Mellon University; Srinivasa G. Narasimhan, Carnegie Mellon University

**2 Orderless Tracking through Model-Averaged Posterior Estimation,**

Seunghoon Hong\*, POSTECH; Suha Kwak, POSTECH; Bohyung Han, POSTECH

**3 The Way They Move: Tracking Multiple Targets with Similar Appearance,**

Caglayan Dicle, neu.edu; Octavia Camps\*, ; Mario Sznaiier, Northeastern University

**15:00 – 15:45 Spotlights (45mins)**

**15:45 – 16:15 Break (30mins)**

**16:15 – 17:15 Oral session 3C (4 papers, 60mins)**

**Optimization**

**Orals 3C:01 – 3C:04**

**1 Active MAP Inference in CRFs for Efficient Semantic Segmentation,**

Gemma Roig, ETH; Xavier Boix\*, ETH; Roderick De Nijs, TUM; Sebastian Ramos, Computer Vision Center (CVC); Luc Van Gool, ETH

**2 Potts model, parametric maxflow and k-submodular functions,**  
Igor Gridchyn, IST Austria; Vladimir Kolmogorov\*, "IST, Austria"

**3 Proportion Priors for Image Sequence Segmentation,**

Claudia Nieuwenhuis\*, ; Evgeny Strekalovskiy, TU Munich; Daniel Cremers, Technische Universität München

**4 Tree Shape Priors with Connectivity Constraints using Convex Relaxation on General Graphs,**

Jan Stühmer\*, Technische Universität München; Peter Schröder, Caltech; Daniel Cremers, Technische Universität München

**17:15 – 18:45 Poster session 3B: Posters 3B:01 – 3B:52**

**1 Revisiting the  $P(n)P$  Problem: A Fast, General and Optimal**

**Solution,** Yinqiang Zheng\*, Tokyo Institute of Technology; Yubin Kuang, Lund University; shigeki Sugimoto, Tokyo Institute of Technology; Kalle Åström, Lund University ; Masatoshi Okutomi, Tokyo Institute of Technology

**2 Direct Optimization of Frame-to-Frame Rotation,** Laurent Kneip\*, ETH Zurich; Simon Lynen, ETH Zurich

**3 PM-Huber: PatchMatch with Huber Regularization for Stereo Matching,** Philipp Heise\*, TU München; Sebastian Klose, TU München; Brian Jensen, TU München; Alois Knoll, TU München

**4 Extrinsic Camera Calibration Without A Direct View Using Spherical Mirror,** Amit Agrawal\*, MERL

**5 Robust non-parametric Correspondence Fitting**, Wen-Yan Lin\*, Oxford Brookes; Ming Ming Chen, ; Kyle Zheng, ; Jiangbo Lu, Advanced Digital Sciences Cent; Crook Nigel,

**6 Deblurring by Example using Dense Correspondence**, Yoav HaCohen\*, Hebrew University; Eli Shechtman, Adobe Research; Dani Lischinski, Hebrew University

**7 On One-Shot Kernels: explicit feature maps and properties**, Stefanos Zafeiriou\*, Imperial; Irene Kotsia, Imperial College London/Middlesex University

**8 Facial Action Unit Event Detection by Cascade of Tasks**, Xiaoyu Ding\*, , Southeast University; Wen-Sheng Chu, Carnegie Mellon University; Fernando De la Torre, Carnegie Mellon University ; Jeffrey Cohn, University of Pittsburgh; Qiao Wang, Southeast University

**9 Similarity Metric Learning for Face Recognition**, Qiong Cao, University of Exeter; Yiming Ying\*, University of Exeter; Peng Li, University of Bristol

**10 Random Faces Guided Sparse Many-to-One Encoder for Pose-Invariant Face Recognition**, Yizhe Zhang\*, ; Ming Shao, Northeastern University; Edward Wong, Polytechnic Institute of NYU; Yun Fu, Northeastern University

**11 Breaking the chain: liberation from the temporal Markov assumption for tracking human poses**, ryan Tokola\*, ; Wongun Choi, University of Michigan; Silvio Savarese, "University of Michigan, USA"

**12 Exploiting Reflection Change for Automatic Reflection Removal**, Yu Li\*, NUS; Michael Brown, National University of Singapore

**13 Shape Index Descriptors Applied to Texture-Based Galaxy Analysis**, Kim Pedersen\*, diku.dk; Kristoffer Stensbo-Smidt, DIKU; Andrew Zirm, NBI KU; Christian Igel, DIKU



**14 Robust Tucker Tensor Decomposition for Efficient Image Storage**, Miao Zhang\*, University of Texas at Arlington; Chris Ding,

**15 Interactive Markerless Articulated Hand Motion Tracking Using RGB and Depth Data**, Srinath Sridhar\*, MPI Informatik and Saarland University; Antti Oulasvirta, MPI Informatik and Saarland University; Christian Theobalt, MPI Informatik and Saarland University

**16 Video Motion for Every Visible Point**, Susanna Ricco\*, Duke University; Carlo Tomasi, Duke University

**17 EVSAC: Accelerating Hypotheses Generation by Modeling Matching Scores with Extreme Value Theory**, Victor Fragoso\*, UCSB; Pradeep Sen, UCSB; Sergio Rodriguez, UCSB; Matthew Turk, "UC Santa Barbara, USA"

**18 PixelTrack: a fast adaptive algorithm for tracking non-rigid objects**, Stefan Duffner\*, LIRIS, INSA de Lyon; Christophe Garcia, LIRIS, INSA de Lyon

**19 Unifying Nuclear Norm and Bilinear Factorization Approaches for Low-rank Matrix Decomposition**, Ricardo Cabral\*, Carnegie Mellon University; Fernando de la Torre, Carnegie Mellon University ; Joao Costeira, Instituto de Sistemas e Robotica; Alexandre Bernardino, Instituto de Sistemas e Robotica

**20 Coupled Dictionary and Feature Space Learning with Applications to Cross-Domain Image Synthesis and Recognition**, De-An Huang, Academia Sinica; Yu-Chiang Frank Wang\*, Academia Sinica

**21 Multi-View Normal Field Integration for 3D Reconstruction of Mirroring Objects**, Michael Weinmann\*, University of Bonn; Aljosa Osep, University of Bonn; Roland Ruiters, University of Bonn; Reinhard Klein, University of Bonn

**22 Discovering Object Functionality**, Bangpeng Yao\*, Stanford University; Jiayuan Ma, ; Fei-Fei Li, Stanford University

**23 Bird Part Localization Using Exemplar-Based Models with Enforced Pose and Subcategory Consistency**, Jiongxin Liu\*, COLUMBIA UNIVERSITY; Peter Belhumeur, "Columbia University, USA"

**24 Person Re-identification by Saliency Matching**, Rui Zhao\*, CUHK; Wanli Ouyang, The Chinese University of HK; Xiaogang Wang, "The Chinese University of Hong Kong, Hongkong"

**25 Prime Object Proposals with Randomized Prim's Algorithm**, Santiago Manen\*, BIWI ETH Zurich; Matthieu Guillaumin, ETH Zurich; Luc Van Gool, ETH

**26 Mining Multiple Queries for Image Retrieval: On-the-fly learning of an Object-specific Mid-level Representation**, Basura Fernando\*, KU Leuven; Tinne Tuytelaars, KU Leuven

**27 A General Two-step Approach to Learning-Based Hashing**, Guosheng Lin\*, The University of Adelaide; Chunhua Shen, The University of Adelaide; Anton Van den Hengel, University of Adelaide ; David Suter,

**28 Modeling Occlusion by Discriminative AND-OR Structures**, Bo Li\*, Beijing Institute of Tech.; Wenzhe Hu, ; Tianfu Wu, UCLA; Song Chun Zhu, UCLA

**29 An Adaptive Descriptor Design for Object Recognition in the Wild**, Zhenyu Guo\*, University of British Columbia; Z.Jane Wang, University of British Columbia

**30 Coherent Object Detection with 3D Geometric Context from a Single Image**, Jiyan Pan\*, Carnegie Mellon University; Takeo Kanade,

**31 Write a Classifier: Zero Shot Learning Using Purely Textual Descriptions**, Mohamed Elhoseiny\*, Rutgers University; Ahmed Elgammal, ; Babak Saleh, Rutgers University

**32 Random Forests of Local Experts for Pedestrian Detection**, Javier Marin\*, Computer Vision Center; David Vazquez, Computer Vision Center, UAB; Jaume Amorós, Computer Vision Center,

UAB; Antonio Lopez, Computer Vision Center, UAB; Bastian Leibe, "RWTH Aachen University, Germany"

**33 Visual Reranking through Weakly Supervised Multi-Graph Learning**, Cheng Deng\*, Xidian University; Rongrong Ji, Xiamen University; Wei Liu, IBM T. J. Watson Research; Dacheng Tao, University of Technology, Sydney; xinbo Gao

**34 Domain Adaptive Classification**, Fatemeh Mirrashed\*, University of Maryland; Mohammad Rastegari, University of Maryland

**35 Supervise Binary Hash Code Learning With Jensen Shannon Divergence**, Lixin Fan\*, Nokia Research Center

**36 Model Recommendation with Virtual Probes for Ego-Centric Hand Detection**, Cheng Li, Tsinghua University; Kris Kitani\*, Carnegie Mellon University

**37 Geometric Registration Based on Distortion Estimation**, Wei Zeng\*, Florida International Univ.; Mayank Goswami, Stony Brook University; Feng Luo, Rutgers University; xianfeng Gu, Stony Brook University

**38 Multi-View Object Segmentation in Space and Time**, Abdelaziz Djelouah\*, Technicolor; Edmond Boyer, ; Jean-Sebastien Franco, Grenoble Universities; Patrick Perez, "Technicolor, France"; Francois Le Clerc,

**39 Pedestrian Parsing via Deep Compositional Network**, Ping Luo\*, CUHK

**40 Dynamic Structured Model Selection**, David Weiss\*, University of Pennsylvania; Benjamin Sapp, Google; Ben Taskar, University of Washington

**41 From Point to Set: Extend the Learning of Distance Metrics**, Pengfei Zhu, The Hong Kong Polytechnic University; Lei Zhang\*, The Hong Kong Polytechnic University; Wangmeng Zuo, Harbin Institute of Technology; David Zhang, The Hong Kong Polytechnic University

**42 Class-Specific Simplex-Latent Dirichlet Allocation for Image Classification**, Mandar Dixit\*, UC San Diego; Nikhil Rasiwasia, Yahoo Research; Nuno Vasconcelos, "UC San Diego, USA"

**43 Mining Motion Atoms and Phrases for Complex Action Recognition**, LiMin Wang\*, CUHK; Yu Qiao, SIAT

**44 Learning Maximum Margin Temporal Warping for Action Recognition**, Jiang Wang\*, Northwestern University; Ying Wu, Northwestern University

**45 Combining the Right Features for Complex Event Recognition**, Kevin Tang\*, Stanford U.; Bangpeng Yao, Stanford University; Fei-Fei Li, Stanford University; Daphne Koller,

**46 Space-Time Robust Video Representation for Action Recognition**, Nicolas Ballas\*, CEA/Mines-ParisTech; Yi Yang, cmu.edu; Lan Zshzsh, CMU; Betrand Delezoide, CEA; Franoise Preteux, Mines ParisTech; Alexander Hauptmann, Carnegie Mellon University

**47 YouTube2Text: Recognizing and Describing Arbitrary Activities Using Semantic Hierarchies and Zero-Shoot Recognition**, Sergio Guadarrama\*, University of California, Berk; Niveda Krishnamoorthy, UT Austin; Girish Malkarnenkar, UT Austin; Raymond Mooney, ; Trevor Darrell, ; Kate Saenko, UMass Lowell

**48 Very Fast Abnormal Event Detection**, Cewu Lu\*, The Chinese University of Hong; Jianping Shi, CUHK ; Jiaya Jia, Chinese University of Hong Kong

**49 Dynamic Pooling for Complex Event Recognition**, Weixin Li\*, UC San Diego; Qian Yu, Sarnoff; Ajay Divakaran, Nuno Vasconcelos, UC San Diego

**50 Relative Attributes For Large-scale Abandoned Object Detection**, Quanfu Fan\*, IBM; Prasad Gabbur, ; Sharath Pankanti,

**51 Action Recognition and Localization by Hierarchical Space-Time Segments**, Shugao Ma\*, Boston University; Stan Sclaroff, Boston University; Jianming Zhang, Boston University; Nazli Ikizler-Cinbis, Department of Computer Engineering, Hacettepe University

**52 The "Moving Pose": An Efficient 3D Kinematics Descriptor for Low-Latency Action Recognition and Detection**, Mihai Zanfir, ; Marius Leordeanu\*, Institute of Mathematics of the Romanian Academy; Cristian Sminchisescu, Lund University

# **Main Conference**      **Friday, December 6**

## **Day 4:**

**9:00 – 10:00 Oral session 4A (4 papers, 60mins)**

### **Recognition**

**Orals 4A:01 – 4A:04**

#### **1 Beyond Hard Negative Mining: Efficient Detector Learning via Block-Circulant Decomposition,**

João F. Henriques\*, Institute of Systems and Robotics - University of Coimbra; João Carreira, Institute of Systems and Robotics - University of Coimbra; Rui Caseiro, Institute of Systems and Robotics - University of Coimbra; Jorge Batista, Institute of Systems and Robotics - University of Coimbra

#### **2 From Large Scale Image Categorization to Entry-Level**

**Categories,** Vicente Ordonez\*, Stony Brook University; Jia Deng, Stanford University; Yejin Choi, Stony Brook University; Alexander Berg, Stony Brook University; Tamara Berg, "Stony Brook University, USA"  
Carl Vondrick\*, MIT; Aditya Khosla, ; Tomasz Malisiewicz, ; Antonio Torralba, MIT

#### **3 Fast Subspace Search via Grassmannian Based Hashing,** Xu

Wang\*, University of Minnesota; Stefan Atef, ; John Wright, ; Gilad Lerman, University of Minnesota

#### **4 Finding the Best from the Second Bests -- Inhibiting Subjective Bias in Evaluation of Visual Tracking Algorithms,**

YU PANG\*, Temple univeristy; Haibin Ling,

**10:00 – 10:45 Spotlights (45min, 52 spotlights at 50secs each)**

**10:45 – 12:15 Poster session 4A: Posters 4A:01 – 4A:52**

**1 Line Assisted Light Field Triangulation and Stereo Matching,**  
Zhan Yu\*, University of Delaware; Xinqing Guo, University of  
Delaware; Jingyi Yu, University of Delaware

**2 A Flexible Scene Representation for 3D Reconstruction Using  
an RGB-D Camera,** Diego Thomas\*, National,  
Institute of Informatics, Akihiro Sugimoto

**3 Automatic Registration of RGB-D Scans via Salient  
Directions,** Bernhard Zeisl\*, ETH; Kevin Koeser, ; Marc Pollefeys,  
ETH

**4 Real-time solution to the absolute pose problem with unknown  
radial distortion and focal length,** Zuzana Kukelova\*, ; Martin  
Bujnak, Bzovicka, Bratislava, Slovakia; Tomas Pajdla, Czech  
Technical University

**5 Rectangling Stereographic Projection for Wide-Angle Image  
Visualization,** Che-Han Chang\*, National Taiwan University; Min-  
Chun Hu, Wen-Huang Cheng, Yung-Yu Chuang, National Taiwan  
University

**6 Accurate Blur Models vs. Image Priors in Single Image Super-  
Resolution,** Netalee Efrat\*, Weizmann institute; Daniel Glasner,  
Weizmann Institute; sasha Apartsin, weizmann institute; Boaz  
Nadler, weizmann institute; Anat Levin, Weizmann Institute, Israel

**7 Learning Slow Features for Behavior Analysis,** Lazaros  
Zafeiriou, Imperial College London; Mihalís Nicolaou, Imperial  
College London; Stefanos Zafeiriou\*, Imperial; Symeon Nikitidis,  
Imperial College London; Maja Pantic, Imperial College

**8 Fast Face Detector Training Using Tailored Views,** Kristina  
Scherbaum\*, MMCI; Rogerio Feris, ; James Petterson, ; Volker  
Blanz , University of Siegen; Hans-Peter Seidel, MPI Saarland

**9 Simultaneous Clustering and Tracklet Linking for Multi-Face  
Tracking in Videos,** Baoyuan Wu\*, CASIA & RPI; Siwei Lyu,  
SUNY Albany; Baogang Hu, CASIA; Qiang Ji,

**10 A Deep Sum-Product Architecture for Robust Facial Attributes Analysis**, Ping Luo\*, CUHK

**11 Hidden Factor Analysis for Age Invariant Face Recognition**, Dihong GONG, SIAT; Zhifeng LI\*, SIAT; Dahua Lin, TTIC; Jianzhuang Liu, CUHK; Xiaoou Tang, Chinese University of Hong Kong

**12 A Learning-Based Approach to Reduce JPEG Artifacts in Image Matting**, Inchang Choi, ; Sunyeong Kim, KAIST; Michael Brown, National University of Singapore; Yu-Wing Tai\*, "KAIST, Korea"

**13 A New Image Quality Metric for Image Auto-Denoising**, Xiangfei Kong\*, City University of Hong Kong; Kuan Li, NUDT; Qingxiong Yang, City University of Hong Kong; Ming-Hsuan Yang, "UC Merced, USA"; Liu Wenyin, Shanghai University of Electronic Power

**14 Joint Noise Level Estimation from Personal Photo Collections**, YiChang Shih\*, M.I.T.; Vivek Kwatra, ; Sergey Ioffe, ; Hui Fang, ; Troy Chinen,

**15 Latent Data Association: Bayesian Model Selection for Multitarget Tracking**, Aleksandr Segal\*, Oxford University Robotics Lab; Ian Reid, University of Adelaide

**16 Initialization-Insensitive Visual Tracking Through Voting with Salient Local Features**, Kwang Yi\*, Seoul National University; Hawook Jeong, Seoul National University; Byeongho Heo, Seoul National University; Hyung Jin Chang, Imperial College London; Jin Young Choi, Seoul National University

**17 Pose-Configurable Generic Tracking of Elongated Objects**, Daniel Wesierski\*, University; Patrick Horain,

**18 Conservation Tracking**, Martin Schiegg\*, University of Heidelberg, HCI; Philipp Hanslovsky, Universitt Heidelberg, HCI;



Bernhard Kausler, University of Heidelberg, HCI; Lars Hufnagel, EMBL, Heidelberg; Fred Hamprecht, HCI, University of Heidelberg

**19 Curvature Regularization using Partial Enumeration**, Carl Olsson\*, ; Johannes Ulén, Lund University; Yuri Boykov, "University of Western Ontario, Canada"; Vladimir Kolmogorov, "IST, Austria"

**20 Total Variation Regularization for Functions with Values in a Manifold**, Jan Lellmann\*, University of Cambridge; Evgeny Strelakovsky, TU Munich; Sabrina Koetter, TU Munich; Daniel Cremers, Technische Universität München

**21 Matching Dry to Wet Materials**, Yaser Yacoob\*, Univ of Maryland

**22 Unsupervised Visual Domain Adaptation Using Subspace Alignment**, Basura Fernando\*, KU Leuven; Amaury Habrard, University Jean Monnet of Saint-Etienne; Marc Sebban, University of Jean Monnet in Saint-Etienne, Hubert Curien Laboratory; Tinne Tuytelaars, KU Leuven

**23 Segmentation Driven Object Detection with Fisher Vectors**, Ramazan Gokberk CINBIS\*, INRIA Grenoble; Jakob Verbeek, "INRIA, France"; Cordelia Schmid, "INRIA, France"

**24 Saliency Detection via Dense and Sparse Reconstruction**, Xiaohui Li\*, DUT, China; Huchuan Lu, DUT, China; Ming-Hsuan Yang, "UC Merced, USA"; Lihe Zhang, DUT, China; Xiang Ruan,

**25 Bayesian Joint Topic Modelling for Weakly Supervised Object Localisation**, Zhiyuan Shi\*, Queen Mary Univ. of London; Timothy Hospedales, EECS, QMUL; Tao Xiang, EECS, QMUL

**26 Parsing IKEA Objects: Fine Pose Estimation**, Joseph Lim\*, MIT; Hamed Pirsiavash, MIT; Antonio Torralba, MIT

**27 Active Visual Recognition with Expertise Estimation in Crowdsourcing**, Chengjiang Long, Stevens Institute of

Technology; Gang Hua\*, Stevens Institute of Technology; Ashish Kapoor, Microsoft Research

**28 A Scalable Unsupervised Feature Merging Approach to Efficient Dimensionality Reduction of High-dimensional Visual Data**, Lingqiao Liu\*, Australian National University; Lei Wang, University of Wollongong

**29 Training deformable part models with decorrelated features**, Ross Girshick\*, UC Berkeley; Jitendra Malik, UC Berkeley

**30 Quantize and Conquer: A dimensionality-recursive solution to nearest neighbor search, clustering, and image retrieval**, Yannis Avrithis\*, NTUA

**31 Learning Hash Codes with Listwise Supervision**, Jun Wang\*, IBM T. J. Watson Research; Wei Liu, IBM T. J. Watson Research; Andy Sun, IBM ; Yugang Jiang, Fudan University

**32 Image Retrieval using Textual Cues**, Anand Mishra\*, IIIT Hyderabad; Karteek Alahari, ENS-Willow; c. v. Jawahar, IIIT Hyderabad

**33 Scene Collaging: Analysis and Synthesis of Natural Images with Semantic Layers**, Phillip Isola\*, MIT; Ce Liu, Microsoft Research New England

**34 Understanding High-Level Semantics by Modeling Traffic Patterns**, Hongyi Zhang, Peking University; Andreas Geiger\*, KIT ; Raquel Urtasun, Toyota Technological Institute at Chicago

**35 Efficient 3D Scene Labeling Using Fields of Trees**, Olaf Kähler\*, University of Oxford; Ian Reid, University of Adelaide

**36 Multi-Channel Correlation Filters**, Hamed Kiani galoogahi\*, NUS; Terence Sim, NUS; Simon Lucey,

**37 Learning CRFs for Image Parsing with Adaptive Subgradient Descent**, Honghui Zhang\*, HKUST; Long Quan, "The

Hong Kong University of Science and Technology, China"; Ping Tan, ; Jingdong Wang, Microsoft Research Asia

**38 Robust Trajectory Clustering for Motion Segmentation**, Feng Shi, Beihang University; zhong Zhou\*, ; Jiangjian Xiao, ; Wei Wu,

**39 Robust Subspace Clustering via Half-Quadratic Minimization**, Yingya Zhang\*, NLPR, CASIA; Zhenan Sun, NLPR, CASIA; Ran He, NLPR, CASIA; Tieniu Tan, "NLPR, China"

**40 Structured learning of sum-of-submodular higher order energy functions**, Alex Fix, Cornell; Thorsten Joachims, Cornell; Sam Park, Cornell; Ramin Zabih\*, Cornell University

**41 Joint Learning of Discriminative Prototypes and Large Margin Nearest Neighbor Classifiers**, Martin Koestinger\*, Graz University of Technology; Peter Roth, Graz University of Technology; Horst Bischof, Graz University of Technology

**42 Dictionary Learning and Sparse Coding on Grassmann Manifolds**, An Extrinsic Solution, Mehrtash Harandi, NICTA and CECS, ANU, Conrad Sanderson, CECS, ANU, Chunhua Shen, University of Adelaide, Brian C. Lovell, University of Queensland

**43 Latent Multitask Learning for View-Invariant Action Recognition**, Behrooz Mahasseni\*, Oregon State University; Sinisa Todorovic, Oregon State University

**44 Concurrent Action Detection with Structural Prediction**, Ping Wei\*, Xi'an Jiaotong University, UCLA; Nanning Zheng, Xi'an Jiaotong University ; Yibiao Zhao, UCLA; Song Chun Zhu, UCLA

**45 Understanding Human-Object Interaction via Exemplar based Modelling**, Jian-Fang Hu\*, Sun yat-sen university; Wei-Shi Zheng, ; Jian-Huang Lai, Sun Yat-sen University; Shaogang Gong, EECS, QMUL; Tao Xiang, EECS, QMUL

**46 Human Re-identification by Matching Compositional Template with Cluster Sampling**, Yuanlu Xu, Sun Yat-Sen University; Liang Lin\*, Sun Yat-Sen University; Weishi Zheng, ; Xiaobai Liu, Ucla.edu

**47 Dynamic Scene Deblurring**, Tae Hyun Kim, Seoul National Univ.; Byeongjoo Ahn, Seoul National University; Kyoung Mu Lee\*, Seoul National University

**48 Directed Acyclic Graph Kernels for Action Recognition**, Ling WANG\*, Telecom ParisTech; Hichem SAHBI, LTCI, CNRS, (TELECOM ParisTech)

**49 Learning View-invariant Sparse Representations for Cross-view Action Recognition**, Jingjing Zheng\*, University of Maryland; Zhuolin Jiang, Huawei Noah's Ark Lab

**50 Event Detection in Complex Scene Using Interval Temporal Constraints**, Yifan Zhang\*, CASIA; Qiang Ji, ; Hanqing Lu,

**51 Towards understanding action recognition**, Hueihan Jhuang\*, Max Planck Institute for Intelligent Systems; Juergen Gall, University of Bonn, Germany; Silvia Zuffi, Michael Black, "Max Planck Institute for Intelligent Systems, Germany"; Cordelia Schmid, INRIA

**52 Modifying the Memorability of Face Photographs**, Aditya Khosla\*, MIT; Wilma Bainbridge, MIT; Antonio Torralba, MIT; Aude Oliva,

**12:15 – 13:45 Lunch**

**13:45 – 14:45 Keynote Speech**

**14:45 – 15:15 Break**

**15:15 – 16:15 Oral session 4B (4 papers, 60mins)**

**Orals 4B:01 – 4B:04**

**1 A Practical Transfer Learning Algorithm for Face Verification,**

Xudong Cao\*, Microsoft Research Asia; David Wipf,

**2 Learning to predict gaze in egocentric video,** Yin Li\*, Georgia Institute of Technolog; Alireza Fathi, Georgia Institute of Technology; James Rehg, Georgia Institute of Technology

**3 Real-time Articulated Hand Pose Estimation using Semi-supervised Transductive Regression Forests,**

Danhang Tang\*, Imperial College London; Tsz-Ho Yu, University of Cambridge; Tae-Kyun Kim, Imperial College London

**4 Estimating Human Scanpath Using Hidden Markov Model,** Huiying Liu\*, Institute of Computing Technology; Dong Xu, "NTU, Singapore"; Qingming Huang, Graduate Univ of Chinese Academy of Sciences; Wen LI, NTU; Stephen Lin, Microsoft Research Asia

**16:15 – 17:00 Spotlights (45mins)**

**17:00 – 18:30 Poster session 4B: Posters 4B:01 – 4B:50**

**1 Enhanced Continuous Tabu Search for Parameter Estimation in Multiview Geometry,** Guoqing Zhou, Northwestern Polytechnical University; Qing Wang\*, Northwestern Polytechnical Uni

**2 Global Fusion of Relative Motions for Robust, Accurate and Scalable Structure from Motion,** Pierre Moulon, Université Paris-Est, LIGM (UMR CNRS), ENPC, Mikros Image; Pascal Monasse, Université Paris-Est, LIGM (UMR CNRS), ENPC ; Renaud Marlet, Université Paris-Est, LIGM (UMR CNRS), ENPC

**3 Internet-based morphable model,** Ira Kemelmacher\*, University of Washington

**4 Large-Scale Multi-Resolution Surface Reconstruction from RGB-D Sequences**, Frank Steinbrcker\*, Technical University of Munich; Christian Kerl, in.tum.de; Juergen Sturm, TU Munich; Daniel Cremers, in.tum.de

**5 Modeling 4D Human-Object Interactions for Event and Object Recognition**, Ping Wei\*, Xi'an Jiaotong University, UCLA; Yibiao Zhao, UCLA; Nanning Zheng, Xi'an Jiaotong University ; Song Chun Zhu, UCLA

**6 Modeling the calibration pipeline of the Lytro camera and its application in high quality light-field image reconstruction**, donghyeon Cho\*, KAIST; Minhaeng Lee, KAIST; Sunyeong Kim, KAIST; Yu-Wing Tai, "KAIST, Korea"

**7 Face Recognition via Random Path Measure Over Face Patch Network**, Chaochao Lu\*, CUHK; Deli Zhao, The Chinese University of Hong Kong

**8 Coupling Alignments with Recognition for Still-to-Video Face Recognition**, Zhiwu Huang\*, ICT, CAS; Xiaowei Zhao, ICT, CAS; Shiguang Shan, "Chinese Academy of Sciences, China"; Ruiping Wang, Institute of Computing Technology, Chinese Academy of Sciences; Xilin Chen,

**9 Capturing Global Semantic Relationships for Facial Action Unit Recognition**, Ziheng Wang\*, RPI; Yongqiang Li, Harbin Institute of Technology; SHANGFEI WANG, USTC; Qiang Ji,

**10 Estimating Human Pose with Flowing Puppets**, Silvia Zuffi\*, ; Javier Romero, MPI PS; Cordelia Schmid, "INRIA, France"; Michael Black, "Max Planck Institute for Intelligent Systems, Germany"

**11 Illuminant Chromaticity from Image Sequences**, Veronique Prinet\*, Hebrew University of Jerusalem; Dani Lischinski, ; Michael Werman,

**12 Contextual Hypergraph Modeling for Salient Object Detection**, Xi Li\*, University of Adelaide; Yao Li, University of Adelaide; Chunhua Shen, The University of Adelaide; Anthony Dick, University of Adelaide ; Anton Van den Hengel, University of Adelaide

**13 Super-resolution via Transform-invariant Group-sparse Regularization**, Carlos Fernandez-Granda\*, Stanford University, Emmanuel J. Candès, Stanford University

**14 Optical Flow via Locally Adaptive Fusion of Complementary Data Costs**, Tae Hyun Kim, Seoul National Univ.; Heeseok Lee, Seoul National University; Kyoung Mu Lee\*, Seoul National University

**15 Optimal Orthogonal Basis and Image Assimilation: Motion Modeling**, Isabelle Herlin\*, ; Etienne Huot, Inria; Giuseppe Papari, Lithicon

**16 A Generic Deformation Model for Dense Non-Rigid Surface Registration: a Higher-Order MRF-based Approach**, Yun Zeng\*, Harvard; Chaohui Wang, MPI ; David Gu, ; Dimitris Samaras, Stony Brook Univ.; Nikos Paragios, Ecole Centrale de Paris

**17 Bayesian 3D tracking from monocular video**, Ernesto Brau\*, ; Kobus Barnard, "University of Arizona, USA"; Jinyan Guan, University of Arizona; Kyle Simek, University of Arizona; Luca del Pero, ; Colin Dawson, University of Arizona

**18 GOSUS: Grassmannian Online Subspace Updates with Structured-sparsity**, Jia Xu\*, UW-Madison; Vamsi Ithapu, ; Lopamudra Mukherjee, University of Wisc Whitewater; James Rehg, Georgia Institute of Technology ; Vikas Singh,

**19 Fast online orthogonal dictionary learning and image restoration**, Chenglong Bao, nus.edu.sg; Jianfeng Cai, uiowa.edu; Hui Ji\*, "NUS, Singapore"

**20 Data-Driven 3D Primitives for Single Image Understanding,**  
David Fouhey\*, Carnegie Mellon University; Abhinav Gupta, ;  
Martial Hebert, "CMU, USA"

**21 Learning Discriminative Part Detectors for Image  
Classification and Cosegmentation,** Jian Sun\*, Xi'an Jiaotong  
University; Jean Ponce, "ENS, France"

**22 A Deformable Mixture Parsing Model with Parselets,** Jian  
Dong\*, NUS; Qiang Chen, ; Wei Xia, NUS; Zhongyang Huang, ;  
Shuicheng Yan, "NUS, Singapore"

**23 Joint Inverted Indexing,** Yan Xia\*, USTC; Kaiming He,  
Microsoft Research Asia; Fang Wen, ; Jian Sun, "Microsoft  
Research, China"

**24 Improving Graph Matching via Density Maximization,** Chao  
Wang\*, University of Wollongong; Lei Wang, University of  
Wollongong

**25 Attribute Adaptation for Personalized Image Search,** Adriana  
Kovashka\*; Kristen Grauman, University of Texas at Austin

**26 Feature Weighting via Optimal Thresholding for Video  
Analysis,** Zhongwen Xu\*, Zhejiang University; Yi Yang, cmu.edu;  
Ivor Tsang, ; Nicu Sebe, University of Trento; Alexander  
Hauptmann, Carnegie Mellon University

**27 Volumetric Semantic Segmentation using Pyramid Context  
Features,** Jonathan Barron\*, UC Berkeley; Pablo Arbelaez, ; Soile  
Keranen, LBL; David Knowles, LBL; Mark Biggin, LBL; Jitendra  
Malik, UC Berkeley

**28 Efficient Hand Pose Estimation from a Single Depth Image,**  
Chi Xu, Bioinformatics Institute; Li Cheng\*, Bioinformatics  
Institute



**29 Synergistic Clustering of Image and Segment Descriptors for Unsupervised Scene Understanding**, Daniel Steinberg\*, ACFR; Oscar Pizarro, ACFR; Stefan Williams, ACFR

**30 Random Grids: Fast Approximate Nearest Neighbors and Range Searching for Image Search**, Dror Aiger\*, Google; Effrosyni Kokiopoulou, Google; Ehud Rivlin, Google Research

**31 Discovering Details and Scene Structure with Hierarchical Iconoid Shift**, Tobias Weyand\*, RWTH Aachen; Bastian Leibe, RWTH Aachen

**32 Strong Appearance and Expressive Spatial Models for Human Pose Estimation**, Leonid Pishchulin\*, Max Planck Institute for Infor; Mykhaylo Andriluka, Max Planck Institute for Informatics; Peter Gehler, Max Planck; Bernt Schiele, "MPI Informatics, Germany"

**33 3D Sub-Query Expansion for Improving Sketch-based Multi-View Image Retrieval**, Yenliang Lin\*, National Taiwan University; Cheng Yu Huang, ; Hao Jeng Wang, ; Winston Hsu,

**34 Predicting Primary Gaze Behavior using Social Saliency Fields**, Hyun Soo Park\*, CMU; Eakta Jain, TI; Yaser Sheikh,

**35 Efficient Higher-Order Clustering on the Grassmann Manifold**, Suraj Jain, Indian Institute of Science; Venu Madhav Govindu\*, Indian Institute of Science

**36 Paper Doll Parsing: Retrieving Similar Styles to Parse Clothing Items**, Kota Yamaguchi\*, Stony Brook University; M. Hadi Kiapour, UNC at Chapel Hill; Tamara Berg, UNC at Chapel Hill

**37 A Unified Video Segmentation Benchmark: Annotation, Metrics and Analysis**, Fabio Galasso\*, MPI Informatics; Naveen Nagaraja, University of Freiburg; Tatiana Jimenez Cardenas, University of Freiburg; Thomas Brox, ; Bernt Schiele, "MPI Informatics, Germany"

**38 What Is the Most Efficient Way to Select Nearest Neighbor Candidates for Fast Approximate Nearest Neighbor Search?**, Masakazu Iwamura\*, Osaka Prefecture University; Tomokazu Sato, Osaka Prefecture University; Koichi Kise, Osaka Prefecture University

**39 Distributed Low-rank Subspace Segmentation**, Ameet Talwalkar\*, UC Berkeley; Lester Mackey, Stanford University; Yadong MU, Columbia University; Shih-Fu Chang, Columbia University; Michael Jordan, Berkeley

**40 Action recognition with improved trajectories**, Heng Wang\*; Cordelia Schmid, INRIA, France

**41 Action Recognition with Actons**, Jun Zhu\*, Shanghai Jiao Tong University; Baoyuan Wang, ; Xiaokang Yang, ; Wenjun Zhang, Zhuowen Tu, UCSD;

**42 Domain Transfer Support Vector Ranking for Person Re-Identification without Target Camera Label Information**, Andy Jinhua Ma, Hong Kong Baptist University; Pong Chi Yuen\*, Hong Kong Baptist University; Jiawei Li,

**43 Finding Causal Interactions in Video Sequences**, Mustafa Ayazoglu, Northeastern University; Burak Yilmaz, Northeastern University; Mario Sznaier\*, Northeastern University; Octavia Camps,

**44 A New Adaptive Segmental Matching Measure for Human Activity Recognition**, Shahriar Shariat\*, Rutgers; Vladimir Pavlovic,

**45 Saliency Detection in Large Point Sets**, Elizabeth Shtrom\*, Technion; George Leifman, Technion; Ayellet Tal, Technion

**46 Motion-Aware KNN Laplacian for Video Matting**, Dingzeyu Li\*, Columbia University; Qifeng Chen, ; Chi-Keung Tang, "Hong Kong University of Science and Technology, Hongkong"

**47 Viewing Real-World Faces in 3D**, Tal Hassner\*,

**48 Accurate and Robust 3D Facial Performance Capture using Single RGBD Images**, Yen-Lin Chen, Texas A&M University; Muscle Wu, Microsoft Research Asia; Fuhao Shi, Texas A&M University; Xin Tong, Microsoft Asia; Jinxiang Chai\*, TAMU

**49 Visual Semantic Complex Network for Web Images**, Shi Qiu\*, CUHK; Xiaogang Wang, "The Chinese University of Hong Kong, Hongkong"; Xiaoou Tang, Chinese University of Hong Kong

**50 What Do You Do? Recognize Occupations in a Photo via Social Context**, Ming Shao\*, Northeastern University; Liangyue Li, Northeastern University; Yun Fu, Northeastern University

## Workshops At a Glance

Dec-2nd							
08:30-10:00 W1 Room 103	08:50-10:00 W2: CDC4V Room 102	09:15-10:15 W3 Room 101	08:30-10:00 W4 Room 104	08:45-10:00 W5 Room 106	9:00-10:30 W6 Room 201	09:10-10:00 W7 Room 105	
Coffee Break							
10:30-12:30 W1 Room 103	10:30-12:40 W2: CDC4V Room 102	10:30-12:50 W3 Room 101	10:30-12:40 W4 Room 104	10:30-12:30 W5 Room 106	10:30-12:10 W6 Room 201	10:30-12:40 W7 Room 105	
Lunch							
13:00-15:55 W8: PGMs Room 103	14:00-15:40 W2: CDC4V Room 102	14:40-15:40 W3 Room 101	14:30-15:40 W4 Room 104	14:30-15:40 W5 Room 106	13:30-15:00 W6 Room 201	14:30-15:40 W7 Room 105	
Afternoon Coffee							
15:55-17:00 W8: PGMs Room 103	16:10-17:50 W2: CDC4V Room 102	16:00-17:40 W3 Room 101	16:10-18:20 W4 Room 104	16:10-17:40 W5 Room 106	15:30-17:00 W6 Room 201	16:10-17:15 W7 Room 105	

## Workshops At a Glance

<b>Dec-7th</b>							
08:30-10:00 W9:4D MOD Room 101	09:00-10:00 W10:LS VSM Room 201	09:30-10:15 W11:30 O-W Room 102	08:30-10:00 W12:ILS VRC Room 104		09:00-10:05 W14 Room 103	08:30-10:00 W15 Room 106	09:00-10:00 W16:WCV S Room 202
Coffee Break							
10:30-12:30 W9:4D MOD Room 101	10:30-12:10 W10:LS VSM Room 201	10:30-12:30 W11:30 O-W Room 102	10:30-12:40 W12:ILS VRC Room 104	10:30-12:15 W13 Room 105	10:30-12:50 W14 Room 103	10:30-12:30 W15 Room 106	10:30-12:30 W16:WCV S Room 202
Lunch							
14:00-16:30 W9:4D MOD Room 101	14:00-15:40 W10:LS VSM Room 201		14:00-15:40 W12:ILS VRC Room 104	13:45-15:40 W13 Room 105	14:00-15:40 W14 Room 103	14:20-15:40 W15 Room 106	14:00-15:30 W16:WCV S Room 202
Afternoon Coffee							
	16:10-17:50 W10:LS VSM Room 201		16:10-17:30 W12:ILS VRC Room 104	16:10-17:30 W13 Room 105	16:10-17:00 W14 Room 103	16:10-17:30 W15 Room 106	16:00-17:30 W16:WC VS Room 202

## Workshops At a Glance

Dec-8th							
08:30-10:00 W17 Room 101	08:45-10:00 W18:HA CI Room 102	09:20-10:20 W19:3d RR Room 104	08:30-10:00 W20 Room 202	09:00-10:00 W21 Room 201	09:00-10:00 W22 Room 103	08:30-10:10 W23:VE CTaR Room 105	08:30-10:30 W24 Room 106
Coffee Break							
10:30-12:50 W17 Room 101	10:30-12:30 W18:HA CI Room 102	10:40-13:00 W19:3d RR Room 104	10:30-12:40 W20 Room 202	10:30-11:50 W21 Room 201	10:30-12:20 W22 Room 103	10:30-12:10 W23:VE CTaR Room 105	11:00-12:20 W24 Room 106
Lunch							
14:00-15:40 W25:C VPVRom 101	13:20-15:40W 18:HACI Room 102	14:30-16:10W 19:3dRR Room 104	14:30-15:40W 20Room 202	14:00-15:40W21 Room 201	14:30-15:40W22 Room 103		13:00-14:30W 24Room 106
Afternoon Coffee							
16:00-18:20 W25:C VPV Room 101	16:10-17:00 W18:HA CI Room 102	16:30-17:25 W19:3d RR Room 104	16:10-17:20 W20 Room 202	16:10-17:10 W21 Room 201	16:10-18:10 W22 Room 103		

## **Workshops**

**Monday, December 2**

### **W1: Graphical models for scene understanding: challenges and perspectives**

**Date:** December 2

**Location:** Room 103

**Overview:** The aim of this workshop is to bring together experts in the areas of graphical models and scene understanding to provide an overview of the next big challenges that need to be addressed (in terms of either modeling, inference or learning) for further advancing the state-of-the-art in natural scene understanding through the proper use of graphical models.

We acknowledge the support of the Computer Vision Foundation and the IEEE Computer Society, who have been sponsoring this workshop.

**08:30      Welcome**

08:35      **Invited talk:** Beyond MAP: Hedging Against Uncertainty via Multiple Diverse Predictions, *Dhruv Batra (Virginia Tech)*

09:10      **Invited talk:** Consistency Potentials: from Pairwise to Higher-order, *Stephen Gould (Australian National University)*

09:45      Multi-instance Object Segmentation with Exemplars, *Xuming He, Stephen Gould*

**10:00      Coffee break**

10:30      **Invited talk:** Modeling Complex Dependencies through Sequential Prediction, *Derek Hoiem (University of Illinois at Urbana-Champaign)*

11:05      **Invited talk:** Inference machines for scene understanding and recognition, *Martial Hebert (Carnegie Mellon University)*

11:40      **Invited talk:** Graphical models for interpreting shape, *Bill Freeman (MIT, CSAIL)*

12:15      Discussion

12:30      Poster session

### **W2: 3rd Workshop on Consumer Depth Cameras for Computer Vision (CDC4CV)**

**Date:** December 2

**Location:** Room 102

**Overview:** The aim of this workshop is to explore recent progress in computer vision with depth cameras. It is also time to take stock of the past five years of work in this field, by evaluating different algorithms. This workshop will help the community to understand the benefits and challenges of depth cameras, and to be prepared for the next generation of these devices.

**08:50 Introduction**

09:00 **Invited Talk** by *Derek Hoiem (University of Illinois at Urbana-Champaign)*

**10:00 Coffee break**

S1: Segmentation (10:30-11:30)

10:30 Saliency-aware Stereo Segmentation using High-Order Energy Optimization - *Jianbing Shen (Beijing Institute of Technology)*, *JianTeng Peng (Beijing Institute of Technology)*

10:50 Depth Interpolation via Smooth Surface Segmentation using Tangent Planes Based on the Superpixels of a Color Image - *Kiyoshi Matsuo (Keio University)*, *Aoki Yoshimitsu (Keio University)*

11:10 External mask based depth and light field camera - *Dikpal Reddy (Nvidia Research)*, *Jiamin Bai (UC Berkeley)*, *Ravi Ramamoorthi (UC Berkeley)*

**11:30 Short break**

S2: Tracking and Events (11:40-12:40)

11:40 Asynchronous 3D Reconstruction Using Event-Driven Dynamic Stereo Vision and an Adaptive Cooperative Approach - *Ahmed Nabil Belbachir (Austrian Institute of Technology)*

12:00 Tracking an RGB-D Camera Using Points and Planes - *Esra Ataer-Cansizoglu (Northeastern University)*, *Yuichi Taguchi (MERL)*, *Srikumar Ramalingam (MERL)*, *Tyler Garaas (MERL)*

12:20 Reliable Left Luggage Detection Using Stereo Depth and Intensity Cues - *Csaba Beleznai (Austrian Institute of Technology)*, *Peter Gemeiner (Austrian Institute of Technology)*, *Christian Zinner (Austrian Institute of Technology)*

**12:40 Lunch**

14:00 **Invited Talk** by *Daniel Cremers (Technische Universität München)*



- S3: Face Analysis (15:00-15:40)
- 15:00 Compact and Accurate 3-D Face Modeling Using a RGB-D Camera: Let's Open the Door to 3-D Video Conference - *Pavankumar Anasosalu (USC)*, *Diego Thomas (National Institute of Informatics)*, *Akihiro Sugimoto (National Institute of Informatics)*
- 15:20 Automatic Detection of Emotion Valence on Faces Using Consumer Depth Cameras - *Arman Savran (University of Pennsylvania)*, *Ragini Verma (University of Pennsylvania)*, *Ruben Gur (University of Pennsylvania)*
- 15:40 Coffee break**
- 16:10
- 17:10 Real-time sign language recognition using a consumer depth camera - *Alina Kuznetsova (Leibniz University Hannover)*, *Laura Leal-Taixé (Leibniz University Hannover)*, *Bodo Rosenhahn (Leibniz University Hannover)*
- 17:30 Fusion of Skeletal and Silhouette-based Features for Human Action Recognition with RGB-D Devices - *Alexandros Chaaraoui (University of Alicante)*, *José Padilla-López (University of Alicante)*, *Francisco Flórez-Revuelta (Kingston University)*
- 17:50 Concluding Remarks

### **W3: IEEE Workshop on the VOT2013 Visual Object Tracking Challenge**

**Date:** December 2

**Location:** Room 101

**Organizers:** Roman Pflugfelder, Matej Kristan, Ales Leonardis, Jiri Matas, Fatih Porikli

**Overview:** The large numbers of publications on tracking-related problems have made it impossible to follow all developments. Without standardised methodology it is very difficult to distinguish significant progress. This workshop will present, compare and rank the results of more than 20 short-term trackers on the VOT2013 challenge comprising 16 well-known image sequences and a standardised SW evaluation kit. The organizers invite interested researchers to participate in the VOT initiative ([votchallenge.net](http://votchallenge.net)) and to discuss methodologies for tracker comparison.

S1: VOT2013 challenge results (09:15 – 10:15)

- 09:15 The Visual Object Tracking VOT2013 challenge results, *Matej Kristan (University of Ljubljana), Roman Pflugfelder (Austrian Institute of Technology), Ales Leonardis (University of Birmingham), Jiri Matas (Czech Technical University in Prague), Fatih Porikli (Mitsubishi Electric Research Laboratories), Luka Cehovin (University of Ljubljana), Georg Nebehay (Austrian Institute of Technology), Gustavo Fernandez (Austrian Institute of Technology), Tomas Vojir (Czech Technical University in Prague) et al.*

**10:15 Morning Break**

S2: Tracker Presentations (10:30 – 11:45)

- 10:30 VOT2013 Winner: PLT - Single scale pixel based LUT tracker, *Cher Keng Heng (Panasonic R&D Center Singapore), Samantha Yue Ying Lim (Panasonic R&D Center Singapore), Zhi Heng Niu (Panasonic R&D Center Singapore), Bo Li (Panasonic R&D Center Singapore)*
- 10:55 Robust Real-Time Tracking with Diverse Ensembles and Random Projections, *Ahmed Salaheldin (Nile University), Mohamed ELHelw (Nile University), Sara Maher (Nile University)*
- 11:20 Enhanced Distribution Field Tracking using Channel Representations, *Michael Felsberg (Linköping University)*

**11:45 Break**

S3: Tracker Presentations (12:00 – 12:50)

- 12:00 An Adaptive Combination of Multiple Features for Robust Tracking in Real Scene, *Weihua Chen (Chinese Academy of Sciences), Lijun Cao (Chinese Academy of Sciences), Junge Zhang (Chinese Academy of Sciences), Kaiqi Huang (Chinese Academy of Sciences)*
- 12:25 An enhanced adaptive coupled-layer LGTracker++, *Jingjing Xiao (University of Birmingham), Rustam Stolkin (University of Birmingham), Aleš Leonardis (University of Birmingham)*

**12:50 Lunch**

S4: Keynote Talk (14:40 – 15:40)

- 14:40 **Invited Speaker:** *Mubarak Shah (University of Central Florida)*

**15:40 Afternoon Break**

S5: Tracker Presentations, Discussion (16:00 – 17:40)

- 16:00 Graph Embedding Based Semi-Supervised Discriminative Tracker, *Jin Gao (Chinese Academy of Sciences), Junliang Xing (Chinese Academy of Sciences), Weiming Hu (Chinese Academy of Sciences), Xiaoqin Zhang (Chinese Academy of Sciences)*
- 16:25 Long-Term Tracking Through Failure Cases, *Karel Lebeda (University of Surrey), Simon Hadfield, Jiri Matas (Czech Technical University), Richard Bowden (University of Surrey),*
- 16:50 Panel Discussion:  
How to design a standardised evaluation kit?  
Which criterias are important? speed vs. accuracy and reliability  
What is the next VOT2014 challenge?
- 17:40 Closing Remarks

## W4: Reconstruction meets Recognition Challenge

**Date:** December 2

**Location:** Room 104

**Organizers:** Raquel Urtasun (TTI Chicago), Rob Fergus (NYU), Derek Hoiem (UIUC), Antonio Torralba (MIT), Andreas Geiger (MPI Tubingen), Philip Lenz (KIT), Nathan Silberman (NYU), Jianxiong Xiao (Princeton)

**Overview:** Understanding the 3D world is one of the fundamental challenges in computer vision. A wide variety of approaches have been developed to either reconstruct the 3D world or recognize it. However, until very recently the interactions between these two tasks were mostly ignored.

In this workshop, we propose a set of challenges to study how reconstruction and recognition algorithms can jointly be exploited to push forward the state-of-the-art in visual perception tasks. Towards this goal, we propose a set of benchmarks that cover both outdoor scenarios in the context of autonomous driving, as well as indoor scenes for personal robotics.

- 8:30 Session 1: Reconstruction I  
- Conclusions from stereo and optical flow  
- Participant talks
- 10:00 Coffee break**
- 10:30 Session 2: Reconstruction II  
- Conclusions from reconstruction from RGB-D and

- visual odometry
- Participant talks
- 12:40 Lunch break**
- 14:30 Session 3: Recognition I**
  - Conclusions from detection and tracking in outdoors
  - Participant talks
- 15:40 Coffee break**
- 16:10 Session 4: Recognition II**
  - Conclusions from 3D detection and semantic segmentation in indoor scenes
  - Participant talks
- 18:00 Discussion (20 minutes)**

## **W5: Vision-Based Sports Analytics**

**Date:** December 2

**Location:** Room 106

**Website:** <http://www.visionbasedsportsanalytics.net>

**Organizers:** Patrick Lucey (Disney Research, Pittsburgh), Peter Carr (Disney Research, Pittsburgh), Stuart Morgan (Australian Institute of Sport), Iain Matthews (Disney Research, Pittsburgh)

**Overview:** Over the past 10 years, computer vision has played a central role in transforming how sports are watched, played, coached, officiated, broadcasted and organized. Even though tremendous progress has been made, there are still many problems to solve, such as automatic player and ball tracking in team sports, team tactic analysis and prediction; marker-less motion capture of athletes and bio-mechanical analysis; and automatic broadcast solutions. This workshop brings together top researchers in academia and industry together to talk about these problems and foster potential collaborations, as well as releasing a database to the research community to promote research into these areas.

- 08:45 Opening Remarks**
  - S1: Game Analysis (09:00-10:00)
- 09:00 Invited Talk:** Vision-Based Sports Analytics in the AFL - A Coaches View of Constraints, Questions and Opportunities, *David Rath (Hawthorn Football Club, Australia)*
- 09:30 Invited Talk:** From Numbers to Insight: Making Sense of Sports Tracking Data, *Rajiv Maheswaran (University of Southern California, USA)*

- 10:00 Coffee Break**  
S2: Individual Athlete Analysis (10:30-12:30)
- 10:30 Invited Talk:** Busted Bowlers - Fact and Fiction  
*Peter Blanch (Cricket Australia, Australia)*
- 11:00 Invited Talk:** Capturing Olympic Divers and Swimmers  
Above and Below the Water for London 2012, *Chris Bregler (New York University, USA)*
- 11:30 Invited Talk:** Advances in Sport Performance Analytics:  
Enriching the Decision Making Environment or Adding to  
Confusion? *David Martin (Australian Institute of Sport, Australia)*
- 12:00 Invited Talk:** Capturing Intent in Sports, *Yaser Sheikh (Carnegie Mellon University, USA)*
- 12:30 Lunch Break**  
S3: Future Game-Day and Viewer Experiences (14:30-15:40)
- 14:30 Invited Talk:** The Future of Connected Fan Experiences,  
*Stuart Taggart (Relevant Innovation, Australia)*
- 15:00 Invited Talk:** Vision Methods for Sports Video Analysis:  
Tracking Players, Plays, and Cameras, *Irfan Essa (Georgia Tech., USA)*
- 15:40 Coffee Break**  
S4: Broadcasting (16:10-17:40)
- 16:10 Invited Talk:** New Production and Distribution Options  
for Sport in the New Media Landscape, *Gus Seebeck (ESPN Australia, Australia)*
- 16:40 Invited Talk:** Automated Sport Production, *Christophe De Vleeschouwer (UCL, Belgium)*
- 17:10 Description of CMU Basketball Dataset, Peter Carr (Disney Research Pittsburgh, USA)**

## **W6: Workshop on Computer Vision for Accelerated Bioscience**

**Date:** 02 December 2013

**Location:** Room 201

**Organizers:** David Lovell, Matt Adcock, Shahram Izadi, Chuong Nguyen and Hongdong Li.

**Overview:** This workshop aims to combine biological and computer vision research to enhance the scientific understanding of life. This workshop will explore and showcase research efforts that apply

novel Computer Vision techniques to better our understanding of natural organisms.

**Workshop website:** <http://australianbioinformatics.net/cvab-2013>

- 9:00 Workshop introduction
- 9:10 **Invited Talk:** Advancing Biodiversity  
Discovery with Computer Vision, *John La Salle*  
(*Atlas of Living Australia and CSIRO*)
- 9:40 **Invited Talk:** What there is to see: Imaging  
Spectroscopy for Scene Analysis, *Antonio*  
*Robles-Kelly* (*National ICT Australia*)
- 10:10 Virtual 3D Models of Insects for Accelerated  
Quarantine Control, *Chuong Nguyen, David*  
*Lovell, Rolf Oberprieler, Debbie Jennings, Matt*  
*Adcock, Eleanor Gates-Stuart, John LaSalle*  
(*CSIRO*)
- 10:30 Coffee Break**
- 11:00 **Invited Talk:** Computer Vision: Can it help us  
digest Insect Soup?, *Paul Flemons* (*Australian*  
*Museum*)
- 11:15 Insect Soup Challenge - Segmentation,  
Counting, and Simple Classification, *Katarina*  
*Mele* (*CSIRO*)
- 11:30 3D Plant Modelling via Hyperspectral Imaging,  
*Jie Liang, Ali Zia, Jun Zhou, Xavier Sirault*  
(*Australian National University, Griffith*  
*University and CSIRO*)
- 11:50 Super-resolution 3D Reconstruction of thick  
biological samples: a computer vision  
perspective, *Alessio Del Bue, Francesca Cella*  
*Zanacchi, Alberto Diaspro* (*Italian Institute of*  
*Technology*)
- 12:10 Morning session discussion
- 12:30 Lunch**
- 13:30 **Invited Talk:** Imaging Less than Meets the  
Eye, *Kyros Kutulakos* (*University of Toronto*)
- 14:00 Extended Gaussian-filtered Local Binary  
Patterns for colonoscopy image classification,  
*Siyamalan Manivannan, Ruixuan Wang,*

- Emanuele Trucco, Adrian Hood (University of Dundee and University of Leeds)*
- 14:20 Learning to Detect Basal Tubules of Nematocysts in SEM Images, *Michael Lam, Janardhan Rao Doppa, Hu Xu, Abigail Reft, Sinisa Todorovic, Thomas Dietterich, Marymegan Daly (Oregon State University)*
- 14:40 Dirichlet Process Mixtures of Multinomials for Data Mining in Mice Behaviour Analysis, *Matteo Zanotto, Diego Sona, Francesco Papaleo, Vittorio Murino (Italian Institute of Technology)*
- 15:00 Coffee Break**
- 15:30 **Invited Talk:** A Framework for Ultra-High Resolution 3D Imaging, *Michael S. Brown (National University of Singapore)*
- 16:00 Zero-Shot Learning and Detection of Teeth in Images of Bat Skulls, *Hu Xu, Michael Lam, Sinisa Todorovic, Thomas Dietterich, Andrea Cirranello, Paul Velazco, Nancy Simmons, Maureen O'Leary (Oregon State University and American Museum of Natural History)*
- 16:20 High Precision Localization of bacterium and Scientific Visualization, *Mohammadreza Hosseini, Arcot Sowmya, Pascal Vallotton, Tomasz Bednarz (University of New South Wales and CSIRO)*
- 16:40 Afternoon session discussion and review
- 17:00 Workshop concludes**

## **W7: Computer Vision for Autonomous Driving**

**Date:** December 2

**Location:** Room 105

**Organizers:** Bart Nabbe, Yaser Sheikh

**Overview:** The goal of this workshop is to bring together leaders from both academia and industry to determine how well Computer Vision is aiding Autonomous Driving applications. We are going to identify the most relevant aspects of computer vision problems to solve, and to learn from others about proposed avenues and

solutions. Within the scope of the workshop will be core computer vision tasks such as dynamic 3D reconstruction, pedestrian and vehicle detection, and predictive scene understanding.

- 09:10 Opening notes from Workshop Organizers
- 09:20 **Invited Talk:** *Making Bertha See*, Uwe Franke  
(Daimler AG)
- 10:00 Coffee Break**
- 10:30 Visual Odometry by Multi-frame Feature Integration,  
*Akihiro Yamamoto, Hernan Badino, Takeo Kanade*
- 11:00 Integrated Pedestrian and Direction Classification  
using a Random Decision Forest, *Junli Tao, Reinhard  
Klette*
- 11:30 **Invited Talk:** Programmable Headlights: Smart and  
Safe Lighting Solutions for the Road Ahead,  
*Srinivasa Narasimhan (Carnegie Mellon University)*
- 12:10 Priors for Stereo Vision under Adverse Weather  
Conditions, *Stefan Gehrig, Maxim Reznitskii, Nicolai  
Schneider, Uwe Franke*
- 12:40 Lunch**
- 14:30 Spatio-Temporal Good Features to Track, *Christoph  
Feichtenhofer, Axel Pinz*
- 15:00 **Invited Talk:** Is the self-driving car around the  
corner? Mobileye's work on Computer Vision centric  
approach to self-driving at consumer level cost,  
*Amnon Shashua (Mobileye)*
- 15:40 Coffee Break**
- 16:10 Panel Discussion: Are computer vision algorithms  
ready to take the wheel? Leaders of industry and  
academia will discuss avenues of research to pursue.
- 17:10 Best paper award (Sponsored by Tandent) and  
raffle (Sponsored by Texas Instruments)

## **W8: Inference for probabilistic graphical models (PGMs)**

Date: December 2

Location: Room 103

**Organizers:** Qinfeng (Javen) Shi, Chunhua Shen, Stephen Gould,  
Jason L. Williams, Tiberio Caetano

**Overview:** With advances in inference techniques of Probabilistic  
graphical models (PGMs), new insights are emerging, as are new



problems that are motivated by these applications. The purpose of this workshop is to bring together an examination of theoretical advances in inference techniques with emerging problem formulations motivated by applications. We will also discuss novel inference methods, new views or understandings, novel inference problems and/or methods for solving them.

- 14: 00     **welcome and overview by organizers**
- 14: 40     Invited Talk 1: Supervised Hierarchical Dirichlet Process  
with Variational Inference  
*Cheng Zhang, Carl Henrik Ek, Xavi Gratal, Florian  
Pokorny, and Hedvig Kjellström*
- 15: 20     Oral presentation 1: Supervised Hierarchical Dirichlet  
Process with Variational Inference, *Cheng Zhang, Carl  
Henrik Ek, Xavi Gratal, Florian Pokorny, and Hedvig  
Kjellström*
- 15:40     **Coffee break**
- 16:20     Invited talk 2: Title: Likelihood, Inference, and other  
hassles: Who needs graphical models? by *Sebastian  
Nowozin*
- 17: 00     Oral presentation 2: Infinite Latent Conditional Random  
Fields, *Yun Jiang and Ashutosh Saxena*
- 17: 20     Spotlight 1: Getting Feasible Variable Estimates From  
Infeasible Ones: MRF Local Polytope Study, *Bogdan  
Savchynskyy and Stefan Schmidt*
- 17: 25     Spotlight 2: Superpixel Coherency and Uncertainty  
Models for Semantic Segmentation, *SeungRyul Baek,  
Taegyu Lim, Yongseok Heo, Sungbum Park, Hantak  
Kwak, and Woosung Shim*
- 17: 30     Spotlight 3: Video Object Segmentation by Salient  
Segment Chain Composition, *Dan Banica, Alexandru  
Agape, Adrian Ion, and Cristian Sminchisescu*
- 17: 35     Poster session

## **Workshops**

**Saturday, December 7**

### **W9: 2nd International Workshop on Dynamic Shape Capture and Analysis (4DMOD)**

Date: December 7

Location: Room 101

**Organizers:** Slobodan Ilic (TUM), Edmond Boyer (INRIA), Adriana Hilton (University of Surrey)

**Overview:** 4DMOD is the workshop on the modeling of dynamic scenes. Modeling shapes that evolve over time, analyzing and interpreting their motion is a subject of increasing interest of many research communities including the computer vision, the computer graphics and the medical imaging community. Following the 1st edition in 2011, the purpose of this workshop is to provide a venue for researchers, from various communities, working in the field of dynamic scene modeling from various modalities to present their work, exchange ideas and identify challenging issues in this domain

- 8:30      **Invited 1, *L. Sigal***
- 9:30      Paper 1 : A Convex Relaxation Approach to Space Time  
Multi-view 3D Reconstruction, *Oswald Martin, Daniel  
Cremers*
- 10:00      Break**
- 10:30      Invited 2, *M. Pollefeys*
- 11:30      Paper 2 : One-shot Entire Shape Scanning by Utilizing  
Multiple Projector-Camera Constraints of Grid Patterns,  
*Kasuya Nozomu, Ryusuke Sagawa, Ryo Furukawa*
- 12:00      Paper 3 : Single-view RGBD-based Reconstruction of  
Dynamic Human Geometry, *Charles Malleon, martin  
Klaudiny, Adrian Hilton, Jean-Yves Guillemaut*
- 12:30      Lunch**
- 14:00      Invited 3, *D. Samaras*
- 15:00      paper 4 : Robust Model-based 3D Torso Pose Estimation  
in RGB-D sequences, *Markos Sigalas, Maria Pateraki,  
Iason Oikonomidis, Panos Trahanias*
- 15:30      Invited 4 5, *M. Salzmann*(1 hr)

### **W10: 2<sup>nd</sup> International Workshop on Large-Scale Video Search and Mining (LSVSM'13)**

**Date:** December 7

**Location:** Room 201

**Co-chairs:** Junsong Yuan, John Smith, and Shih-Fu Chang

- 09:00 *Junsong Yuan (Nanyang Tech., Singapore)*  
09:10 *Josef Sivic (École Normale Supérieure, France)*  
10:00 **Tea / Coffee Break**  
10:30 **Invited talk:** TBD  
11:20 *Junsong Yuan (Nanyang Tech., Singapore)*  
12:10 **Lunch Break**  
14:00 Jason Corso (SUNY Buffalo, US)  
14:50 **Invited Talk:** TBD  
15:40 **Tea / Coffee Break**  
16:10 Learning Non-linear Calibration for Score Fusion with Applications to Image and Video Classification, *Tianyang Ma (Temple University), Sangmin Oh, Kitware, Amitha Perera, Kitware, Longin Jan Latecki (Temple University)*  
16:30 Pedestrian Attribute Classification in Surveillance: Database and Evaluation, *Jianqing Zhu, Shengcai Liao, Dong Yi, Stan Z. Li (Chinese Academy of Science, China)*  
17:10 An Adaptive Query Prototype Modeling Method for Image Search Reranking, *Hong Lu, Guobao Jiang, Xiangyang Xue (Fudan University, China)*  
17:30 Thematic Saliency Detection using Spatial-Temporal Context, *Ye Luo, Gangqiang Zhao, Junsong Yuan (Nanyang Technological University, Singapore)*

## **W11:300 Faces in-the-Wild Challenge (300-W)**

**Date:** December 7

**Location:** Room 102

**Overview:**

- 09:30 Welcome (5 min)  
9:40 Oral Presentations: Constrained Local Neural Fields for robust facial landmark detection in the wild, *Tadas Baltrusaitis, Louis-Philippe Morency and Peter Robinson*  
**10:00 Morning Break**  
10:30 Oral Presentations: (10:30 - 12:30)  
1. Localizing Facial Keypoints with Global Descriptor Search, Neighbour Alignment and Locally Linear Models, *Md. Kamrul Hasan, Sharon Moalem and Chris Pal*  
2. Guided Unsupervised Learning of Mode Specific

Models for Facial Point Detection in the Wild, *Shashank Jaiswal, Timur Almaev and Michel Valstar*

3. Multiview Active Shape Models with SIFT Descriptors for the 300-W Face Landmark Challenge, *Stephen Milborrow, Tom Bishop and Fred Nicolls*

4. Facial Landmark Localization with Coarse-to-fine Convolutional Network Cascade, *Erjin Zhou, Haoqiang Fan, Zhimin Cao, Yuning Jiang and Qi Yin*

5. Learn to Combine Multiple Hypotheses for Face Alignment, Junjie Yan, Zhen Lei, Dong Yi and Stan Z. Li

6. 300 Faces in-the-Wild Challenge: The First Facial Landmark Localization Challenge, *Christos Sagonas, Georgios Tzimiropoulos, Stefanos Zafeiriou and Maja Pantic*

## **W12: ImageNet Large Scale Visual Recognition Challenge 2013 (ILSVRC2013) Workshop**

**Date:** December 7

**Location:** Room 104

**Organizers:** Olga Russakovsky (Stanford University), Jonathan Krause (Stanford University), Jia Deng (University of Michigan), Alexander C. Berg (UNC Chapel Hill), Fei-Fei Li (Stanford University)

**Overview:** The purpose of the workshop is to present the methods and results of the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2013 and the new Fine-Grained Challenge 2013. Challenge participants with the most successful and innovative entries are invited to present. The ILSVRC is sponsored in part by Google.

Detailed schedule is available at

<http://imagenet.org/challenges/LSVRC/2013/iccv2013.php>

8:30 ILSVRC classification competition: introduction, results and invited talks

**10:00 Coffee break**

10:30 ILSVRC detection competition: introduction, results and invited talks

**12:40 Lunch**

14:00 **Invited talk:** *Vittorio Ferrari, University of Edinburgh*

14:40 Fine-grained competition: introduction, results and invited talks

- 15:40**     **Coffee break**
- 16:10     Fine-grained competition: invited talks
- 17:30**     **Close**

## **W13: Large Scale Visual Commerce**

**Date:** December 7

**Location:** Room 105

**Organizers:** Anurag Bhardwaj (*eBay Research Labs*), Robinson Piramuthu (*eBay Research Labs*), Serge Belongie (*University of California, San Diego*), Tsuhan Chen (*Cornell University*)

**Overview:** Past few years have seen a paradigm shift in commerce experiences moving closer to purely visual forms of interaction. These developments offer exciting opportunities for state-of-the-art vision research to enable the next generation of commerce experiences. This workshop aims to foster growth in this emerging area by bring together researchers working on the cutting-edge of computer vision for visual commerce.

**Sponsors:** IEEE computer society, CVF and eBay Research Labs

- 10:30     Introducing LSVisCom
- 10:45     **Invited Talk:** Real-time object class recognition and search in large databases, *Lorenzo Torresani (Dartmouth College)*
- 11:30     **Invited Talk:** Privacy and Security in Large Scale Visual Commerce, *Venu Govindaraju (SUNY, Buffalo)*
- 12:15**     **Lunch Break**
- 13:45     **Invited Talk:** You are what you wear: Discovering Attributes, Parsing Clothing, and Recognizing Socio-Identity, *Tamara Berg (University of North Carolina, Chapel Hill)*
- 14:30     **Invited Talk:** Perceiving 3D objects and space from images, *Silvio Savarese (Stanford University)*
- 15:15     Discovering Pictorial Brand Associations from Large-Scale Online Image Data, *Gunhee Kim (CMU), Eric Xing (CMU)*
- 15:35     Distribution of Best Paper Award
- 15:40**     **Coffee Break**
- 16:10     Panel Discussions moderated by Serge Belongie  
Panelists: TBD
- 17:00     Social event (30 min)

## **W14: THUMOS Challenge: Action Recognition with a Large Number of Classes**

**Date:** December 7

**Location:** Room 103

**Organizers:** Yu-Gang Jiang, Jingen Liu, Amir Roshan Zamir

, Ivan Laptev, Massimo Piccardi, Mubarak Shah, Rahul Sukthankar

**Overview:** The goal of THUMOS challenge is to encourage researchers to develop methods for action recognition that scale to a large numbers of action categories captured in natural settings. The results of this competition, in both detection and classification tasks, will be presented and discussed in this workshop.

### ***Opening Session***

- 09:00 Opening Remarks
- 09:05 Invited Talk, Cordelia Schmid (INRIA).
- 09:35 Invited Talk, Jason Corso (SUNY at Buffalo).
- 10:05 Coffee Break
- Action Classification Session
- 10:30 Overview and the Results of the Classification Challenge
- 10:50 Classification Challenge Winner Presentation
- 11:10 Classification Challenge Runner-up Presentation
- 11:30 Invited Talk: Silvio Savarese (Stanford).
- 12:00 “Ordered Trajectories for Large Scale Human Action Recognition”, Ramana Oruganti, Roland Goecke, University of Canberra, Canberra, Australia
- 12:20 “A Spatio-Temporal Feature based on Triangulation of Dense SURF”, Keiji Yanai, Do Nga, UEC, Tokyo, Japan.
- 12:40 Lunch
- Action Detection Session
- 14:00 Overview and the Results of the Detection Challenge
- 14:20 Detection Challenge Winner Presentation
- 14:40 Invited Talk: Stan Sclaroff (Boston University)
- 15:10 Action Detection Invited Talk, TBA.
- 15:40 Coffee Break
- Conclusion Session
- 16:10 Invited Talk: Tal Hassner (Open University, Israel)
- 16:40 Summarization of the Challenge Results and Closing Remarks
- 17:00 End

# W15: The First International Workshop on Visual Domain Adaptation and Dataset Bias

**Date:** December 7

**Location:** Room 106

**Organizers:** Brian Kulis, Ruonan Li, Kate Saenko, Fei Sha

- 08:30 Welcome
- 08:40 **Invited Talk:** Dataset bias and how to live with it, *Alyosha Efros (University of California, Berkeley)*
- 09:20 Spotlight Talks
- 09:20 Datasets, Sample Selection, and Test Performance, *Omid Aghazadeh, Stefan Carlsson*
- 09:25 Heterogeneous Domain Adaptation: Learning Visual Classifiers from Textual Description, *Mohamed Elhoseiny, Babak Saleh, Ahmed Elgammal*
- 09:30 Sampling For Unsupervised Domain Adaptive Object Recognition, *Fatemeh Mirrashed, Vlad Morariu, Larry Davis*
- 09:35 Domain Adaptive Classification, *Fatemeh Mirrashed, Mohammad Rastegari*
- 09:40 Does Evolution cause a Domain Shift? *Konstantinos Rematas, Basura Fernando, Tatiana Tommasi, Tinne Tuytelaars*
- 09:45 Scalable Transform-based Domain Adaptation, *Erik Rodner, Judy Hoffman, Jeff Donahue, Trevor Darrell, Kate Saenko*
- 09:50 Adapting Pedestrian Detection from Synthetic to Far Infrared Images, *Yainuvis Socarras, Sebastian Ramos, David Vazquez, Antonio Lopez, Theo Gevers*
- 09:55 Sparse Embedding-based Domain Adaptation for Object Recognition, *Jingjing Zheng, Rama Chellappa, Jonathon Phillips*
- 10:00 Coffee Break and Poster Session**
- 10:30 **Invited Talk:** Recent progress towards deep domain adaptation, *Trevor Darrell (University of California, Berkeley)*
- 11:10 **Invited Talk:** SVM based approaches for domain adaptation, *Dong Xu (Nanyang Technological University)*
- 11:50 **Invited Talk:** Domain Adaptation using Manifolds and Dictionary-based Methods for Object/Face/Event Recognition, *Rama Chellappa (University of Maryland)*
- 12:30 Lunch Break**

- 14:20 **Invited Talk:** Adaptation for Objects and Attributes, *Kristen Grauman (University of Texas, Austin)*
- 15:00 **Invited Talk:** Virtual and Real World Adaptation for Pedestrian Detection, *Antonio Lopez (Universitat Autònoma de Barcelona)*
- 15:40 Coffee Break and Poster Session**
- 16:10 **Invited Talk:** Towards life-long visual learning: from practice to theory and back, *Christoph Lampert (Institute of Science and Technology Austria)*
- 16:50 **Invited Talk:** Domain Adaptation for Large-Scale Visual Recognition, *Jia Deng (University of Michigan)*

## **W16: Wearable Computer Vision Systems (WCVS)**

**Date:** December 7

**Location:** Room 202

**Organizers:** Ana C. Murillo, Javier Civera, Mohammad Moghimi, Serge Belongie

**Overview:** This workshop addresses the topic of wearable computer vision systems, understanding by that cameras that are worn by a person, somehow mounted or attached to a human body, but not necessarily with the same point of view than the user. Advances on vision technologies, miniaturization of the camera optics and electronics and growing computing capabilities, either locally or in the cloud, are pushing even further the potential for this technology. This workshop is generously sponsored by Facebook, Thanks!

- 9:00 Introduction by the organizers
- 9:15 **Invited Talk:** TBA, *James Rehg ( Georgia Institute of Technology)*
- 10:00 Coffee break**
- 10:30 **Invited Talk:** Going out in public: Humans using wearable computers, *Bruce Thomas (University of South Australia)*
- 11:15 **Invited Talk:** Summarizing egocentric video, *Kristen Grauman (University of Texas, USA)*
- 12:00 Short talk: Live Metric 3D Reconstruction on Mobile Phones, *Marc Pollefeys (ETH Zurich, Switzerland)*
- 12:15 Short talk: TBA, *Mohammad Moghimi (UCSD, USA)*
- 12:30 Lunch break**
- 14:00 **Invited Talk:** *Takeo Kanade (Carnegie Mellon University, USA)*
- 14:45 Short talk: Wearable Computer Vision Systems for a Cortical Visual Prosthesis, *Wai Ho Li (Monash*



- University, Victoria, Australia)*
- 15:00 Short talk: Wearable smartphone hybrid framework for hand and foot gesture interaction, *Zhihan Lv (Umea University, Sweden)*
- 15:15 Short talk: A Smartphone-Based Obstacle Detection and Classification System for Assisting Visually Impaired People, *Ruxandra Tapu, Bogdan Mocanu, Titus Zaharia, Andrei Bursuc (Telecom-SudParis, France)*
- 15:30 Coffee break**
- 16:00 Invited Talk: PhotoOCR: Reading Text in Uncontrolled Conditions, *Mark Cummins (Google Australia)*
- 16:45 Demos, Concluding remarks and discussion (45 min)

## **Workshops**

**Sunday, December 8**

### **W17: The First IEEE International Workshop on Computer Vision for Converging Perspectives**

**Date:** December 8

**Location:** Room 101

**Overview:** The First IEEE International Workshop on Computer Vision for Converging Perspectives brings together interested researchers from academia, government, and industry working in the field of computer vision, machine learning, pattern recognition, and remote sensing to address the challenges involved in developing vision systems capable of assimilating image and video data from heterogeneous, multi-scale and multi-perspective visual sensing platforms for actionable intelligence and scientific discoveries. The workshop is sponsored by Google and Oak Ridge National Laboratory.

- 08:30      Opening Remarks
- 08:40      **Keynote Talk:** Prof. Robert Pless (*Washington University*)
- 09:30      Paper Presentation: Observing the natural world through Flickr, D. Crandall, J. Wang, M. Korayem
- 10:00      Morning Break**
- 10:30      **Keynote Talk:** Luc Vincent, Director, Google StreetView
- 11:20      **Keynote Talk:** Prof. Anton van den Hengel (*Australian Centre of Vision Technologies*)
- 12:10      Paper Presentation: Processing Geotagged Image Sets for Collaborative Compositing and View Construction, L. Kovacs (*MTA SZTAKI*)
- 12:30      Paper Presentation: Probabilistic Semantic Segmentation from Label Strokes Applied to StreetView Scenes, S. Zhu, Y. Yang, Li Zhang ( *University of Wisconsin - Madison*)
- 12:50      Best Paper Award and Concluding Remarks

### **W18: 1st Workshop on Understanding Human Activities: Context and Interactions (HACI 2013)**

**Date:** December 8

**Location:** Room 102,

**Overview:** Motivated by the rich and complex temporal, spatial, and social structure of human activities, activity recognition today features several new challenges, including modeling group activities, complex temporal reasoning, activity hierarchies, human-object interactions and human-scene interactions. This workshop aims to bring together researchers in computer vision and machine learning to share ideas and propose solutions to address the many aspects of this field.

- 8:45      Opening Remarks
- 9:00      **Keynote 1:** TBA, *Abhinav Gupta (Carnegie Mellon University)*
- 9:40      Oral 1: Iterative Action and Pose Recognition using Global-and-Pose Features and Action-specific Models, *Norimichi Ukita (NAIST)*
- 10:00      Morning Break**
- 10:30      **Keynote 2:** TBA, *Ivan Laptev (INRIA / Ecole Normale Supérieure)*
- 11:10      **Keynote 3:** TBA, *Amitha Perera (Kitware Inc)*
- 11:50      Oral 2: A multi-scale approach to gesture detection and recognition, *Natalia Neverova (INSA-Lyon)*
- 12:10      Oral 3: Context-sensitive Conditional Ordinal Random Fields for Facial Action Intensity Estimation, *Ognjen Rudovic (Imperial College London)*
- 12:30      Lunch**
- 13:20      **Keynote 4:** TBA, *Michael S. Ryoo (NASA / Jet Propulsion Laboratory)*
- 14:00      **Keynote 5:** TBA, *Ashutosh Saxena (Cornell University)*
- 14:40      Oral 4: Temporal Poselets for Collective Activity Detection and Recognition, *Moin Nabi (Istituto Italiano di Tecnologia)*
- 15:00      Oral 5: Spatio-Temporal Human-Object Interactions for Action Recognition in Videos, *Victor Escorcia (Universidad del Norte)*
- 15:20      Oral 6: Less is More: Video Trimming for Action Recognition, *Borislav Antic (University of Heidelberg)*
- 15:40      Afternoon Break**
- 16:10      Panel members: Sameh Khamis, Mohamed Amer, Wongun Choi, Tian Lan  
Panel Discussion (20 min)

## W19: 3D Representation and Recognition (3dRR-13)

**Date:** December 8

**Location:** Room 104

**Overview:** Object categorization and scene understanding have long been a central goal of computer vision research. Changes in lighting, viewpoint, and pose, as well as intra-class differences, lead to enormous appearance variation, making the problem highly challenging. While advances in machine learning and image feature representations have led to great progress in 2D pattern recognition approaches, recent work suggests that large gains can be made by acknowledging that objects live in a physical, three-dimensional world. When modeling scenes, objects and their relations in 3D, we must answer several fundamental questions. How can we effectively learn 3D object representations from images or video? What level of supervision is required? How can we infer spatial knowledge of the scene and use it to aid in recognition? How can both depth sensors and RGB data be used to enable more descriptive representations for scenes and objects?

Specific questions we aim to address include: Object Representation, Kinect: Combining Depth and RGB Sensors, Reconstruction and Recognition, Spatial Inference, Spatial constraints and contextual recognition, and Human vision.

We especially thank Microsoft Research for sponsoring the workshop.

- 9:20      Opening: Min Sun (University of Washington)
- 9:40      **Keynote speak**, *Raquel Urtasun (TTI Chicago)*
- 10:20      Coffee break**
- 10:40      Oral session I – Modeling Object Shape and Pose  
Chair: *Michael Stark (Stanford University and Max Planck Institute for Informatics)*  
[1] Cubistic Representation for Realtime 3D Shape and Pose Estimation of Unknown Rigid Object, *Hiromasa Yoshimoto and Yuichi Nakamura*  
[2] Object Detection by 3D Aspectlets and Occlusion Reasoning, *Yu Xiang and Silvio Savarese*  
[3] A Computationally Efficient Approach to 3D Model Matching for 3D Object Localization and Fine-Grained 3D Pose Estimation from a Single 2D Image, *Erdem Yoruk and Rene Vidal*
- 11:40      **Keynote speak**, *Marc Pollefeys (ETH Zurich)*

- 12:20 Oral session II – 3D Scenes and Fine-Grained 3D Object Categorization  
Chair: Min Sun (University of Washington)  
[1] Behind the Scenes: What Moving Targets Reveal About Static Scene Geometry, *Geoffrey Taylor and Fei Mai*  
[2] 3D Object Representations for Fine-Grained Categorization, *Jonathan Krause, Michael Stark, Jia Deng, and Fei-Fei Li*
- 13:00 Lunch Break**
- 14:30 **Keynote speak**, *Ashutosh Saxena (Cornell)*
- 15:10 Oral session III – Features and Cues for Recognition  
Chair: *Silvio Savarese (Stanford University)*  
[1] The mesh-LBP: computing Local Binary Patterns on Discrete Manifolds, *Naoufel Werghi, Stefano Berretti, Pietro Pala, and Alberto Del Bimbo*  
[2] Multiscale TILT Feature Detection with Application to Geometric Image Segmentation, *Chi-Pang Lam, Allen Yang, Ehsan Elhamifar, and Shankar Sastry*  
[3] Which Edges Matter?, *Aayush Bansal, Adarsh Kowdle, Devi Parikh, Larry Zitnick, and Andrew Gallagher*
- 16:10 Coffee break**
- 16:30 **Keynote speak**: *Jitendra Malik (UC Berkeley)*
- Prize presentation –Sponsored by Microsoft Research.**
- 17:20 Conclusions: Min Sun, Michael Stark, and Silvio Savarese (5 min)

## **W20: Computer Vision in Vehicle Technology: From Earth To Mars**

**Date:** December 8

**Location:** Room 202

**Organizers:** David Geronimo, Atsushi Imiya, Antonio M. López, Tomas Pajdla, Dariu Gavrila, Urbano Nunes, Steven Beauchemin, Theo Gevers

**Overview:** CVVT: E2M aims to get together researchers to promote development and spreading of new ideas and results on the use of computer vision in vehicle technology. Some examples are advanced driver assistance systems, exploratory and service robotics, unmanned aerial vehicles and underwater robots.

- 08:30 Welcome Message
- 08:40 **Invited talk:** RatSLAM: Using Models of Rodent Hippocampus for Vision-based Robot Navigation and Beyond, *Michael Milford (Queensland Univ. of Technology, Australia)*
- 09:30 Oral 1: Enhanced Target Tracking in Aerial Imagery with P-N Learning and Structural Constraints, *Mennatullah Siam (Nile University, Egypt), Mohamed El-Helw (Nile University, Egypt)*
- 10:00 Coffee Break**
- 10:30 Oral 2: Evaluating Color Representations for Online Road Detection, *Jose Alvarez (NICTA, Australia), Theo Gevers (University of Amsterdam, The Netherlands), Antonio López (CVC and Univ. Autònoma de Barcelona, Spain)*
- 10:50 Oral 3: Direct Generation of Regular-Grid Ground Surface Map From In-Vehicle Stereo Image Sequences, *Shigeki Sugimoto (Tokyo Institute of Technology, Japan), Kouma Motooka (Tokyo Institute of Technology, Japan), Masatoshi Okutomi (Tokyo Institute of Technology, Japan)*
- 11:10 Oral 4: From Video Matching to Video Grounding, *Georgios Evangelidis (INRIA, France), Ferran Diego Andilla (HCI, Germany), Radu Horaud (INRIA, France)*
- 11:30 **Invited Talk:** Live and Semantic 3D Maps for Autonomous Driving, *Antonio M. López, Germán Ros, Sebastián Ramos, Jiaolong Xu (CVC, Spain)*
- 12:40 Lunch**
- 14:30 **Invited Talk:** Autonomous systems for environmental monitoring: from the air and underwater, *Mitch Bryson (University of Sydney, Australia)*
- 15:20 Oral 5: Visual Approaches for Driver and Driving Behavior Monitoring: A Review, *Hang-Bong Kang (The Catholic University of Korea, Republic of Korea)*
- 15:40 Coffee break**
- 16:10 Oral 6: Evaluation of the Capabilities of Confidence Measures for Assessing Optical Flow Quality, *Patricia Márquez-Valle (CVC, Spain), Debora Gil (CVC, Spain), Aura Hernández (CVC and Univ. Autònoma de Barcelona, Spain), Debora Gil (CVC, Spain)*
- 16:30 Oral 7: Exploiting Sparsity for Real Time Video Labelling, *Lachlan Horne (NICTA, Australia), Jose Alvarez (NICTA, Australia), Nick Barnes (NICTA, Australia)*

- 16:50     **Invited Talk:** PROViDE Planetary Robotics, *Tomas Pajdla (Czech Technical University in Prague, Czech Republic)*
- 17:20     Best Paper Announcement and Closing

## **W21: Big Data in 3D Computer Vision (BigData3DCV)**

**Date:** December 8

**Location:** Room: 201

**Organizers:** Jian Zhang, Mohammed Bennamoun, Dan Schonfeld, Zhengyou Zhang, Fatih Porikli, Dong Xu, Hongdong Li, Lixin Fan, Qiang Wu.

**Overview:** The main goal of this workshop is to explore scientific research on Big Data in 3D computer vision. It will be the venue for papers to highlight the recent advanced research from academic and industry labs through the connection of big data in 3D computer vision. This workshop appreciates the generous supports from NOKIA Research Centre for the sponsor of the best paper award.

- 9:00     Workshop introduction
- 9:10     **Invited Talk:** Topics in 2D & 3D Computer Vision, *Mohammed Bennamoun (The University of Western Australia)*
- 10:00     Coffee Break**
- 10:30     A Novel Local Surface Description for Automatic 3D Object Recognition in Low Resolution Cluttered Scenes, *Syed Afaq Ali Shah, Mohammed Bennamoun, Farid Boussaid, Amar El-Sallam (The University of Western Australia)*
- 10:50     \*A Scalable Collaborative Online System for City Reconstruction, *Ole Untzelmann, Torsten Sattler, Sven Middelberg, Leif Kobbelt (RWTH Aachen University)*  
**\*Note: Candidate of best paper**
- 11:10     Fury of the Swarm: Efficient and Very Accurate Triangulation for Multi-View Scene Reconstruction, *Shawn Recker, Mauricio Hess-Flores, Kenneth Joy, (University of California Davis)*
- 11:30     Sparse Approximations of 3D Mesh Geometry Using Frames as Overcomplete Dictionaries, *Maja Krivokuca, Waleed Abdulla, Burkhard Wuensche (The University of Auckland)*

- 11:50 Lunch**
- 14:00 Targetless Calibration of a Lidar - Perspective Camera Pair, *Levente Tamas (Technical University of Cluj-Napoca)*, *Zoltan Kato (University of Szeged)*
- 14:20 Optimal Reduction of Large Image Databases for Location Recognition, *Michal Havlena, Wilfried Hartmann, Konrad Schindler (ETH Zürich)*
- 14:40 \* Three Dimensional Motion Trail Model for Gesture Recognition, *Bin Liang, Lihong Zheng (Charles Sturt University)*  
**\*Note: Candidate of best paper**
- 15:00 3D Surface Extraction using Incremental Tetrahedra Carving, *Takayuki Sugiura, Akihiko Torii, Masatoshi Okutomi (Tokyo Institute of Technology)*
- 15:20 Fast and Accurate Large-scale Stereo Reconstruction using Variational Methods, *Georg Kusch (Martin-Luther University)*, *Daniel Cremers (Technische Universität München)*
- 15:40 Coffee Break**
- 16:10 Kinect Shadow Detection and Classification, *Teng Deng, Hui Li, Jianfei Cai, Tat Jen Cham (Nanyang Technological University)*
- 16:30 Semantic Parsing of Street Scene Images Using 3D LiDAR Point Cloud, *Pouria Babahajiani, Lixin Fan and Moncef Gabbouj (Tampere University of Technology and Nokia Research Center)*
- 16:50 Memory Efficient 3D Integral Volumes, *Martin Urschler, Alexander Bornik (Graz University of Technology)*

## W22: Decoding Subtle Cues from Social Interactions

**Date:** December 8

**Location:** Room 103

**Overview:** A great deal of meaning in social interactions is subtly encoded in the posture, movement, expressions, tone of voice, etc. of the participants. Interpreting these cues requires novel analysis methods that go beyond standard techniques for the classification of actions and activities; this workshop will define and explore new challenges and approaches for decoding human behavior in social interactions.

09:00 **Invited talk,** *Wendy Stone*



- 10:00      Coffee**
- 10:30      **Invited talk, Jim Rehg**
- 11:20      Home Alone: Social Robots for Digital Ethnography of Toddler Behavior, *Mohsen Malmir (University of California, San Diego); Deborah Forster (University of California, San Diego); Javier Movellan (Emotient.com); Kendall Youngstrom (University of California, San Diego)*
- 11:40-      Self-Stimulatory Behaviours in the Wild for Autism Diagnosis, *Shyam Sundar Rajagopalan (University of Canberra); Abhinav Dhall (Australian National University), Roland Goecke (University of Canberra)*
- 12:00      Ivan Laptev
- 12:20      Lunch and posters**
- 14:30      Jitendra Malik
- 15:20      Young One 1
- 15:40      Coffee**
- 16:10      Young One 2
- 16:40      Yaser Sheikh
- 17:10      Posters+discussion (1 hr)

Posters:

Joint Alignment and Modeling of Correlated Behavior Streams, *Liliana Lo Presti (Boston University), Stan Sclaroff (Boston University), Agata Rozga (Georgia Institute of Technology)*

Markov Random Field Structures for Facial Action Unit Intensity Estimation, *Georgia Sandbach (Imperial College London), Stefanos Zafieriou (Imperial College London), Maja Pantic (Imperial College London)*

Video based Children's Social Behavior Classification in Peer-play Scenarios, *Lu Tian (Peking University), Dingrui Duan (Peking University), Jinshi Cui (Peking University), Li Wang (Peking University), Hongbin Zha (Peking University), Hamid Aghajan (Stanford University)*

Trusting Skype: Learning the Way People Chat for Fast User Recognition and Verification, *Giorgio Roffo (University of Verona) Marco Cristani (University of Verona), Loris Bazzani (Italian Institute of Technology), Vittorio Murino (Italian Institute of Technology)*

Self-Stimulatory Behaviours in the Wild for Autism Diagnosis, *Shyam Sundar Rajagopalan (University of Canberra), Abhinav Dhall (Australian National University), Roland Goecke (University of Canberra)*

Home Alone: Social Robots for Digital Ethnography of Toddler Behavior, *Mohsen Malmir (University of California, San Diego), Deborah Forster (University of California, San Diego), Javier Movellan (Emotient.com), Kendall Youngstrom (University of California, San Diego)*

Hand Gestures for Intelligent Tutoring Systems: Dataset, Techniques & Evaluation, *Suchitra Sathyanarayana (University of California, San Diego), Gwen Littlewort (University of California, San Diego), Marian Bartlett (University of California, San Diego)*

Human Body-parts Tracking for Fine-grained Behavior Classification, *Norimichi Ukita (Nara Institute of Science and Technology), Atsushi Nakazawa (Kyoto University)*

## **W23: The 5th International Workshop on Video Event Categorization, Tagging and Retrieval (VECTaR2013)**

**Date:** December 8

**Location:** Room 105

**Organizers:** Tieniu Tan, Thomas S. Huang, Liang Wang, Ling Shao, Jianguo Zhang, Yun Fu

08:30      Welcome

08:35      **Keynote Talk:** Dr. Zhengyou Zhang, IEEE Fellow, Microsoft Corp., One Microsoft Way, Redmond WA 98052-6399, US.

### **S1: Oral Session 1 (09:20-10:10)**

09:20      Spatio-Temporal Context Modeling for BoW-Based Video Classification, *Saehoon Yi (Rutgers University), Vladimir Pavlovic (Rutgers University)*

09:45      Semantic Video-to-Video Search using Sub-Graph

Grouping and Matching, *Tae Eun Choe (ObjectVideo Inc.)*, *Hongli Deng (ObjectVideo Inc.)*, *Feng Guo (Google Inc.)*, *Mun Wai Lee (Intelligent Automation, Inc.)*, *Niels Haering (ObjectVideo Inc.)*

**10:10 Morning Coffee Break**

**S2: Oral Session 2 (10:30-12:10)**

- 10:30 NSH: Normality Sensitive Hashing for Anomaly Detection, *Hiroataka Hachiya (Canon INC.)*, *Masakazu Matsugu (Canon INC.)*
- 10:55 Deeply-Learned Slow Feature Analysis for Action Recognition, *Lin Sun (NUS)*
- 11:20 Dynamic Scene Classification using Spatial and Temporal Cues, *Arun Balajee Vasudevan (IIT Jodhpur)*, *Srikanth Muralidharan (IIT Jodhpur)*, *Pratheek Chintapalli (IIT Jodhpur)*, *Shanmuganathan Raman (IIT Gandhinagar)*
- 11:45 VGRAPH: A Novel Approach for Generating Static Video Summaries using Nearest Neighbor Graph, *Karim Mahmoud (IBM,Alexandria University)*
- 12:10 Concluding Remarks

## **W24: Underwater Vision Workshop**

**Date:** December 8

**Location:** Room 106

- 8:30 Arrival
- 9:00 Introduction to ACFR and the Marine Robotics Group
- 9:30 **Keynote** (Ecology)
- 10:10 1x Oral presentation (15+5)
- 10:30 Morning Tea Break**
- 11:00 2x Oral Presentations (15+5)
- 11:40 **Keynote** (Computer Vision/Robotics)
- 12:20 1x Oral Presentations (15+5)
- 12:40 Lunch Break and Travel to the ACFR (Sydney Univesity)**
- 14:00 Lab tour @ the ACFR
- 14:30 Discussion @ The Rose Pub
- 16:30 End of workshop

## **W25: Color and Photometry in Computer Vision (CVPV)**

**Date:** December 8

**Location:** Room 101

**Organizers:** Theo Gevers, Joost van de Weijer, Jose Alvarez, Todd Zickler

**Overview:** This workshop focuses on new insights for the understanding of color and photometry in computer vision. As color and photometry are shared among various research fields, this workshop places them at the junctions of different areas, including color science, applied optics, computational photography, computer vision, computer graphics, and machine learning.

### **S1: Keynote Session 1 (14:00-14:40)**

- 14:00    **Keynote:** Robust Photometric Stereo, *Yasuyuki Matsushita (Microsoft Research Asia)*

### **S2: Oral Session 1 (14:40-15:40)**

- 14:40    Quick Approximation of Camera's Spectral Response From Casual Lighting, *Dilip Prasad, Rang Nguyen, and Michael Brown (National Univ. of Singapore)*
- 15:00    Polarization-based Dehazing using Two Reference Objects, *Daisuke Miyazaki, Daisuke Akiyama, Masashi Baba, Ryo Furukawa, Shinsaku Hiura, Naoki Asada, (Hiroshima City University)*
- 15:20    Approximate Cross Channel Color Mapping from Sparse Color Correspondences, *Hasan Sheikh Faridul, Jurgen Stauder, and Alain Tremeau (Technicolor R&D, Université Jean Monnet)*

### **15:40    Afternoon Break**

### **S3: Oral session 2 (16:00–18:20)**

- 16:00    Multiplex Image Projection using Multi-Band Projectors, *Makoto Nonoyama, Fumihiko Sakaue, and Jun Sato (Nagoya Institute of Technology)*
- 16:20    Mixing Paints for Generating Metamerism Art under 2 Lights and 3 Object Colors, *Daisuke Miyazaki, Kanami Takahashi, Masashi Baba, Hirooki Aoki, Ryo Furukawa, Masahito Aoyama, and Shinsaku Hiura (Hiroshima City University)*
- 16:40    Visual Material Traits: Recognizing Per-Pixel Material Context, *Gabriel Schwartz and Ko Nishino (Drexel University)*

- 17:00 Separating Specular and Diffuse Reflection Components in the HSI Color Space, Jianwei Yang, Lixing Liu, Zhen Lei, and Stan Z. Li (CBSR&NLPR, USC)
- 17:20 Colour Constancy from Both Sides of the Shadow Edge, *Stuart Lynch, Mark Drew, and Graham Finlayson (University of East Anglia, Simon Fraser University)*
- 17:40 Verification of Sky Models for Image Calibration, Rishi Ramakrishnan, *Juan Nieto, and Steve Scheduling (University of Sydney)*
- 18:00 BRDF Estimation of the Structural Color Object by Using Hyper Spectral Image, *Yoshie Kobayashi, Tetsuro Morimoto, Imari Sato, Yasuhiro Mukaigawa, Katsushi Ikeuchi (University of Tokyo, Toppan Printing Co., Ltd, National Institute of Informatics, Osaka University)*

**Day 1:****Demo: 1**

**Date:** December 3 (Morning)

**Name:** Li CHENG, Bioinformatics Institute, A\*STAR, Singapore, chengli@bii.a-star.edu.sg

**Associated paper:** Our Paper with title "Efficient Hand Pose Estimation from a Single Depth Image" and id (1078) has been accepted into the main conference as a poster paper. In the paper, we tackle the practical problem of hand pose estimation from a single noisy depth image. Our approach is able to work with Kinect type noisy depth images, and reliably produces pose estimations of general motions efficiently. There are some further improvement we have made after the paper submission (which leads to improved results and faster rate). We are keen to demo to the ICCV audience our up-to-date system.

**Full list of authors:** Chi XU, and Li CHENG, both from Bioinformatics Institute, A\*STAR, Singapore

**Abstract:** We tackle the practical problem of hand pose estimation from a single noisy depth image. A dedicated three-step pipeline is proposed: Initial estimation step provides an initial estimation of the hand in-plane orientation and 3D location; Candidate generation step produces a set of 3D pose candidate from the Hough voting space with the help of the rotational invariant depth features; Verification step delivers the final 3D hand pose as the solution to an optimization problem. We analyze the depth noises, and suggest tips to minimize their negative impacts on the overall performance. Our approach is able to work with Kinect-type noisy depth images, and reliably produces pose estimations of general motions efficiently (12 frames per second). Extensive experiments are conducted to qualitatively and quantitatively

evaluate the performance with respect to the state-of-the-art methods that have access to additional RGB images. Our approach is shown to deliver on par or even better results.

**Demo: 2**

**Date:** December 3 (Morning)

**Name:** Lorenz Meier, ETH Zurich, lm@inf.ethz.ch

**Associated paper:** ICCV Paper Submission #1725: Live Metric 3D Reconstruction on Mobile Phones

**Full list of authors:** Petri Tanskanen, Kalin Kolev, Lorenz Meier, Federico Camposeco, Olivier Saurer and Marc Pollefeys, all authors are affiliated to: Computer Vision and Geometry Group, ETH Zurich, Switzerland

**Abstract:** In this paper, we propose a complete on-device 3D reconstruction pipeline for mobile monocular hand-held devices, which generates dense 3D models with absolute scale on-site while simultaneously supplying the user with real-time interactive feedback. The method fills a gap in current cloud-based mobile reconstruction services as it ensures at capture time that the acquired image set fulfills desired quality and completeness criteria. In contrast to existing systems, the developed framework offers multiple innovative solutions. In particular, we investigate the usability of the available on-device inertial sensors to make the tracking and mapping process more resilient to rapid motions and to estimate the metric scale of the captured scene. Moreover, we propose an efficient and accurate scheme for dense stereo matching which allows to reduce the processing time to interactive speed. We demonstrate the performance of the reconstruction pipeline on multiple challenging indoor and outdoor scenes of different size and depth variability.

**Demo: 3**

**Date:** December 3 (Morning)

**Name:** Jürgen Sturm, Technical University of Munich, sturmju@in.tum.de

**Associated paper:** This work has been presented recently at the German Conference on Pattern Recognition (GCPR, formerly DAGM): CopyMe3D: Scanning and Printing Persons in 3D (J. Sturm, E. Bylow, F. Kahl, D. Cremers), In German Conference on Pattern Recognition (GCPR), 2013.

Paper:

[http://vision.in.tum.de/\\_media/spezial/bib/sturm\\_et\\_al\\_2013gcpr.pdf](http://vision.in.tum.de/_media/spezial/bib/sturm_et_al_2013gcpr.pdf)

Video: <http://youtu.be/9B2IdD2jHGw>

**Full list of authors:** Jürgen Sturm, Technical University of Munich; Erik Bylow, Lund University; Fredrik Kahl, Lund University; Daniel Cremers, Technical University of Munich

**Abstract:** <http://youtu.be/9B2IdD2jHGw>. We developed an approach to scan persons in 3D using a Kinect while they are rotating on a swivel chair. We post-process the scanned models automatically to close holes and to make it hollow, so that it can be printed cost-efficiently on a color 3D printer. Scanning a person takes around 10 seconds, the reconstruction runs in real-time so that the resulting 3D model is immediately visible on the computer. We can save these models to a VRML file so that the ICCV participants can take their 3D models home. We recently scanned persons live during a talk at the GCPR conference. We believe that this is a nice demo, as it showcases novel applications and challenges that arise from advances in 3D sensor and printer technology. We will also bring some of the 3D models along to demonstrate how the final result looks like.

#### **Demo: 4**

**Date:** December 3 (Morning)

**Name:** Jakob Engel, Technical University Munich, [jajuengel@gmail.com](mailto:jajuengel@gmail.com)

**Associated paper:** Paper#: 1567 Title: Real-Time Semi-Dense Monocular SLAM: A Statistical Approach. Accepted as Poster.

**Full list of authors:** Jakob Engel, Jürgen Sturm, Daniel Cremers all from Technical University Munich



**Abstract:** We propose a statistical approach to perform real-time, semi-dense monocular SLAM on a single CPU. The key idea is to continuously estimate a semi-dense inverse depth map associated with each camera frame and to in turn use this depth map to track the motion of the camera. More specifically, we estimate for all pixels which have a non-negligible image gradient a Gaussian-distributed hypothesis for the inverse depth at that location. We propagate this inverse

depth information over time and update it with stereo observations from images which provide the optimal baseline for the given pixel, allowing to accurately estimate the depth for both, far-away and close-by regions. In terms of tracking accuracy and computational speed, the proposed method compares favorably to both state-of-the-art dense and feature-based SLAM algorithms. We believe that this real-time capable CPU-based mono-SLAM approach is of major practical value for numerous applications such as the autonomous navigation of flying robots.

**Demo: 5**

**Date:** December 3 (Afternoon)

**Name:** Shicheng Zheng, The Chinese University of Hong Kong, zsc.leigh@gmail.com

**Associated paper:**

[1] Unnatural L0 Sparse Representation For Natural Image Deblurring, CVPR 2013

[2] Two-Phase Kernel Estimation for Robust Motion Deblurring, ECCV 2010

**Full list of authors:** Shicheng Zheng, The Chinese University of Hong Kong; Li Xu, The Chinese University of Hong Kong

**Abstract:** We build a fast and robust deblurring software incorporating latest powerful blind deconvolution technologies to remove blur. It is capable of handling large blur kernel and recovering subtle structures and fine details with fast cpu implementation. GPU is also enabled.

**Demo: 6**

**Date:** December 3 (Afternoon)

**Name:** Guofeng Zhang, State Key Lab of CAD&CG, Zhejiang University, zhangguofeng@cad.zju.edu.cn

**Associated paper:**

[1] Wei Tan, Haomin Liu, Zilong Dong, Guofeng Zhang and Hujun Bao. Robust Monocular SLAM in Dynamic Environments. International Symposium on Mixed and Augmented Reality (ISMAR), 2013.

[2] Zilong Dong, Guofeng Zhang, Jiaya Jia, and Hujun Bao. Efficient Keyframe-Based Real-Time Camera Tracking. Computer Vision and Image Understanding, accepted.

[3] Guofeng Zhang, Zilong Dong, Jiaya Jia, Tien-Tsin Wong, and Hujun Bao. Efficient Non-Consecutive Feature Tracking for Structure-from-Motion. European Conference on Computer Vision (ECCV), 2010.

**Full list of authors:** Guofeng Zhang, State Key Lab of CAD&CG, Zhejiang University; Haomin Liu, State Key Lab of CAD&CG, Zhejiang University; Zilong Dong, State Key Lab of CAD&CG, Zhejiang University; Wei Tan, State Key Lab of CAD&CG, Zhejiang University; Jiaya Jia, The Chinese University of Hong Kong; Hujun Bao, State Key Lab of CAD&CG, Zhejiang University

**Abstract:** We propose a robust monocular SLAM system which can work in challenging environments. A novel keyframe-based global localization method is proposed, which requires an offline module to select features from a group of reference images and constructs an optimal set of keyframes to approximately cover the entire space. In order to alleviate the effort of offline preprocessing, an online reference map extension module is proposed to realtime reconstruct new 3D features and select online keyframes to extend the keyframe set. Our system also can handle the dynamic scenes. Our online keyframe representation and updating method can adaptively model the dynamic environments, where the appearance or structure changes can

be effectively detected and handled. The keyframes with large changed areas will be replaced by newly selected frames. In addition, a novel prior-based adaptive RANSAC algorithm (PARSAC) is proposed to efficiently remove outliers even when the inlier ratio is rather low. Experimental results demonstrate that the proposed system can robustly work in a large-scale scene with dynamic objects, and outperforms the state-of-the-art SLAM systems (e.g. PTAM).

**Demo: 7**

**Date:** December 3 (Afternoon)

**Name:** Tae-Kyun Kim, Imperial College London,  
tk.kim@imperial.ac.uk

**Associated paper:** ICCV 2013 oral paper: "Real-time Articulated Hand Pose Estimation using Semi-supervised Transductive Regression Forests"

**Full list of authors:** Danhang Tang (Imperial College London, London, UK), Tsz-Ho Yu (University of Cambridge, Cambridge, UK), Tae-Kyun Kim (Imperial College London, London, UK)

**Abstract:** This paper presents the first semi-supervised transductive algorithm for real-time articulated hand pose estimation. Noisy data and occlusions are the major challenges of articulated hand pose estimation. In addition, the discrepancies among realistic and synthetic pose data undermine the performances of existing approaches that use synthetic data extensively in training. We therefore propose the Semi-supervised Transductive Regression (STR) forest which learns the relationship between a small, sparsely labelled realistic dataset and a large synthetic dataset. We also design a novel data-driven, pseudo-kinematic technique to refine noisy or occluded joints. Our contributions include: (i) capturing the benefits of both realistic and synthetic data via transductive learning; (ii) showing accuracies can be improved by considering unlabelled data; and (iii) introducing a pseudo-kinematic technique to refine articulations

efficiently. Experimental results show not only the promising performance of our method with respect to noise and occlusions, but also its superiority over state-of-the-arts in accuracy, robustness and speed.

## **Demo Session      Wednesday, December 4**

### **Day 2:**

#### **Demo: 8**

**Date:** December 4 (Morning)

**Name:** David Ferstl, Institute for Computer Graphics and Vision, Graz University of Technology , ferstl@icg.tugraz.at

**Associated paper:** In this demo we try to show the practical application of our accepted ICCV paper: Image Guided Depth Upsampling using Anisotropic Total Generalized Variation.

**Full list of authors:** David Ferstl - Institute for Computer Graphics and Vision / Graz University of Technology; Christian Reinbacher - Institute for Computer Graphics and Vision / Graz University of Technology; Rene Ranftl - Institute for Computer Graphics and Vision / Graz University of Technology; Matthias Rüther - Institute for Computer Graphics and Vision / Graz University of Technology; Horst Bischof - Institute for Computer Graphics and Vision / Graz University of Technology

**Abstract:** In this demo we present a novel method to calculate accurate scene flow from depth and intensity image pair in a convex energy minimization approach. While the flow estimation in 2D is a well-studied problem, the calculation of the 3D movement over time is hard to solve. We make use calculate scene flow from modern consumer systems such as Microsoft Kinect, Intel Gesture Camera or Time-of-Flight sensors. Each of these systems simultaneously delivers a depth and an intensity image over time. Through this time sequenced depth and intensity image pair we estimate the metric 3D movement between consecutive acquisitions. We formulate the 3D flow as convex optimization problem with anisotropic higher order regularization. We derive a numerical algorithm based on the classical primal-dual formulation that is efficiently

parallelized. We show that this novel scene flow calculation clearly outperforms existing approaches in terms of accuracy. We further show the practical applicability in a wide range of applications such as super-resolution, tracking, segmentation and camera pose estimation.

**Demo: 9**

**Date:** December 4 (Morning)

**Name:** Ralf Dragon, ETH Zurich, dragon@vision.ee.ethz.ch

**Associated paper:**

[1] Ralf Dragon, Jörn Ostermann, Luc Van Gool, "Robust Realtime Motion-Split-And-Merge for Motion Segmentation", GCPR (DAGM), 2013

[2] Ralf Dragon, Bodo Rosenhahn and Jörn Ostermann, "Multi-Scale Clustering of Frame-to-Frame Correspondences for Motion Segmentation", ECCV 2012

**Full list of authors:** Ralf Dragon, Luc Van Gool (Computer Vision Lab, ETH Zurich)

**Abstract:** Real-Time Motion Segmentation Motion segmentation is the grouping of trajectories according to their motion over time. In the past years, it has been receiving increasing attention since it can be used as a strong prior in dense object segmentation, for the unsupervised learning of object detectors, or for tracking. However, although more than a decade has passed since the first methods were proposed, most approaches are still not suitable for real-world applications as they are either limited by slow processing, they need the number of motions as prior knowledge, or they are not able to handle incomplete or erroneous trajectories. In our ECCV 2012 paper, we presented multi-scale motion clustering as a two-step approach which overcomes these problems and which performs very good on standard benchmarks. It consists of many frame-to-frame motion segmentations which are combined in one final clustering step. After a concise analysis of the frame-to-frame step in our

DAGM 2013 paper, we were able to reduce the runtime and the error a lot which enables robust and realtime motion segmentation. In our live demo, we carry out motion segmentation on consecutive blocks. The result is displayed immediately such that an interactive feedback is provided.

**Demo: 10**

**Date:** December 4 (Morning)

**Name:** Akshay Asthana, Imperial College London,  
a.asthana@imperial.ac.uk

**Associated paper:** Robust Discriminative Response Map Fitting with Constrained Local Models. Akshay Asthana, Stefanos Zafeiriou, Shiyang Cheng and Maja Pantic. In Proc. of 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2013), Portland, Oregon, USA, June 2013.

[http://www.doc.ic.ac.uk/~aasthana/CVPR2013/AAsthana\\_CVPR2013.pdf](http://www.doc.ic.ac.uk/~aasthana/CVPR2013/AAsthana_CVPR2013.pdf)

Demo videos available at:  
<https://sites.google.com/site/akshayasthana/clm-wild-code>

**Full list of authors:** Akshay Asthana (Imperial College London); Stefanos Zafeiriou (Imperial College London); Shiyang Cheng (Imperial College London); Jie Shen (Imperial College London); Maja Pantic (Imperial College London and University of Twente)

**Abstract:** We present a fully-automatic and real-time facial landmark point tracking system that can robustly handle uncontrolled imaging conditions. The system can detect 66 facial landmark points and estimate the 3D head-pose at over 30fps on Intel Xeon 3.80 GHz processor with NVIDIA GeForce GTX 660 (960 CUDA cores). The system is based on the novel discriminative fitting of response map framework, proposed by Asthana et al (CVPR 2013), which has shown state-of-the-art performance for the task of facial landmark detection on multiple uncontrolled databases. The

novelty lies in the use of texture model based on the response maps generated via discriminantly trained filters (i.e. patch-experts). Another benefit of this framework is that it allows for the direct use of 3D shape model during the alignment procedure and therefore, can be easily used for applications that require 3D facial information (for example, pose normalization for face recognition or facial expression recognition, facial animation, HCI interfaces etc). Therefore, using the current facial landmark point tracking system, we also show that the applications such as facial expression recognition and facial performance transfer can be performed in real-time under uncontrolled imaging conditions. The software will be made public at <http://ibug.doc.ic.ac.uk/resources>.

### **Demo: 11**

**Date:** December 4 (Afternoon)

**Name:** Markos Sigalas, Institute of Computer Science, Foundation for Research and Technology – Hellas (FORTH), [msigalas@ics.forth.gr](mailto:msigalas@ics.forth.gr)

### **Associated paper:**

[1] Robust Model-based 3D Torso Pose Estimation in RGB-D sequences. Sigalas M., Pateraki M., Oikonomidis I., Trahanias P. (submitted to 2nd IEEE Workshop on Dynamic Shape Capture and Analysis, to be held within ICCV 2013)

[2] Model-based 3D Torso Pose Estimation from RGB-D data. Sigalas M., Pateraki M., Trahanias P. Demo Session in Computer Vision Pattern Recognition (CVPR), June 25-27, 2013.

[3] Visual tracking of hands, faces and facial features of multiple persons. Baltzakis H., Pateraki M., Trahanias P., 2012. Machine Vision and Applications. [doi: <http://dx.doi.org/10.1007/s00138-012-0409-5>]

[4] Using Dempster's rule of combination to robustly estimate pointed targets. Pateraki M., Baltzakis H., Trahanias P., 2012. In Proc. of the IEEE International Conference on Robotics and Automation (ICRA), 14-18 May, St. Paul, Minnesota, USA. [doi: <http://dx.doi.org/10.1109/ICRA.2012.6224870>]



- [5] Visual estimation of pointed targets for robot guidance via fusion of face pose and hand orientation. Pateraki M., Baltzakis H., Trahanias P., 2011. In Proc. of the 1st IEEE Workshop on Challenges and Opportunities in Robot Perception, held within the 13th International Conference on Computer Vision (ICCV), 6-13 November, Barcelona, Spain. [doi: <http://dx.doi.org/10.1109/ICCVW.2011.6130368>]
- [6] Gesture recognition based on arm tracking for human-robot interaction. Sigalas M., Baltzakis H., Trahanias P. In Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Taipei, Taiwan, October 2010. [doi: <http://dx.doi.org/10.1109/IROS.2010.5648870>]

**Full list of authors:** Markos Sigalas(1,2), Maria Pateraki(1), Panos Trahanias(1,2)

1 Institute of Computer Science (ICS) - Foundation for Research and Technology - Hellas (FORTH), Greece

2 Department of Computer Science, University of Crete, Greece

**Abstract:** The extraction of pose related information from the human body is a challenging computer vision task and of utmost importance in free-form Human Robot Interaction in naturalistic environments. In this context the proposed demo presents our implemented system for extracting upper body pose, utilizing information from commercial RGB-D Sensors. A model-based method for the estimation of the upper body configuration, including torso and arms, has been developed and implemented under the assumption that no initialization phase is possible in order to commence interaction, and that pose recovery and tracking should remain unaffected from partial occlusions. The segmentation of the body of each user relies on our previous work on face identification. Then the formulation of the “minimum projection criterion” is employed to estimate the specific user top-view that is aligned with the main axis of the human torso, thus giving rise to robust 3D torso orientation. Interestingly, the mentioned top-view re-projection of each segmented user can also effectively cope with partial occlusions. The extracted torso pose

provides estimations of the 3D shoulder locations. The latter are used, along with detected hypotheses of hands, in order to extract arm configurations for each user via Particle Filtering.

**Demo: 12**

**Date:** December 4 (Afternoon)

**Name:** Yin Li, Georgia Institute of Technology, yli440@gatech.edu

**Associated paper:**

Yin Li, Alireza Fathi, James M. Rehg. Learning to Predict Gaze in Egocentric Video, ICCV 2013

Alireza Fathi, Yin Li, James M. Rehg. Learning to Recognize Daily Actions using Gaze, ECCV 2012

Zhefan Ye, Yin Li, Alireza Fathi, Yi Han, Agata Rozga, Gregory D. Abowd, James M. Rehg. Detecting Eye Contact using Wearable Eye-Tracking Glasses, 2nd International Workshop on Pervasive Eye Tracking and Mobile Eye-Based Interaction (PETMEI) in conjunction with UbiComp 2012

**Full list of authors:** Yin Li, Zhefan Ye, Alireza Fathi, James Mathew Rehg (School of Interactive Computing, College of Computing, Georgia Institute of Technology)

**Abstract:** With the recent advent of wearable cameras, there has been an increasing interest in egocentric vision. Egocentric vision aims at the automatic analysis of video captured from a first-person perspective. We plan to demonstrate a real-time system for detecting eye contact, an important aspect of face-to-face interactions. With a single pair of video-recording glasses, our method detects moments of eye contact between the camera wearer and a second person during naturalistic social interactions. Our system highlights a set of key functions in egocentric vision, including gaze tracking, ego-motion estimation, face detection/tracking and facial analysis. Our real-time implementation utilizes gaze tracking technology from SMI eye-tracking glasses to determine the camera wearer's point of

gaze. We combine the egocentric video from the glasses and the face analysis technology by Omron OKAO face SDK to estimate the gaze direction of others. We identify moments of eye contact as the event of simultaneous, mutual looking at faces. Our demo suggests the proposed egocentric vision platform as a promising vehicle for understanding the camera wearer's behavior.

**Day 3:****Demo: 13****Date:** December 4 (Afternoon)**Name:** Federico Pernici, MICC University Of Florence,  
pernici@dsi.unifi.it**Associated paper:**[http://www.micc.unifi.it/pernici/index\\_files/eccv2012demo.pdf](http://www.micc.unifi.it/pernici/index_files/eccv2012demo.pdf)**Full list of authors:** Federico Pernici and Alberto Del Bimbo**Abstract:** Back to Back Comparison of Long Term Tracking Systems

Recently several comparisons of tracking systems have been presented in [1] [2] and others are soon to be published in the workshop VOT2013 (in conjunction with the ICCV2013) and in the website: [www.alov300.org](http://www.alov300.org) which consists of more than 300 sequences. In particular the website reports several experiments showing that none of the current methods is able to track objects in the long term. This demonstrates an evident bias on how performance are so far computed. Indeed most of the works on object tracking focused on nuisance factors such: occlusion, lighting, pose, blur, excluding time. According to this, we are interested in showing some of the best long term tracking methods (as for example [3], ([4] demo authors), [5]) simultaneously compared in real time on a laptop (2 or 3 systems will run in parallel on the same machine). This will help participants understanding the pros and cons of the various approaches.

[1] S. Salti, A. Cavallaro, and L. Di Stefano. Adaptive

appearance modeling for video tracking: Survey and evaluation. TIP2012.

[2] Y. Wu, J. Lim, , and M.-H. Yang. Online object tracking: A benchmark. CVPR2013.

[3] S. Hare, A. Saffari, and P. Torr. Struck: Structured output tracking with kernels. ICCV2011.

[4] FaceHugger: The ALIEN Tracker Applied to Faces. ECCV2012 Demo Session.

[5] Kalal, Zdenek, Krystian Mikolajczyk, and Jiri Matas. "Tracking-learning-detection." Pattern Analysis and Machine Intelligence, IEEE Transactions on 34.7 (2012): 1409-1422.

### **Demo: 14**

**Date:** December 5 (Morning)

**Name:** William Nguatem, Bundeswehr University Munich, william.nguatem@unibw.de

**Associated paper:** Nguatem, W., Drauschke, M. and Mayer, H., ROOF RECONSTRUCTION FROM POINT CLOUDS USING IMPORTANCE SAMPLING., Workshop (CMRT13) - City Models, Roads and Traffic 2013, Antalya

**Full list of authors:** Nguatem, William., Drauschke, Martin. and Mayer, Helmut (Bundeswehr University Munich, Institute of Applied Computer Science)

**Abstract:** We propose a novel fully automatic technique for roof fitting in 3D point clouds based on sequential importance sampling (SIS). Our approach makes no assumption of the nature (sparse, dense) or origin (LIDAR, image matching) of the point clouds and further distinguishes, automatically, between different basic roof types based on model selection. The algorithm comprises an inherent data parallelism, the lack of which has been a major drawback of most Monte Carlo schemes. A further speedup is achieved by applying a coarse to fine search within all probable roof configurations in the sample space of roofs. The robustness and effectiveness of our roof reconstruction algorithm is illustrated for point clouds of

varying nature.

**Demo: 15**

**Date:** December 5 (Morning)

**Name:** Andrew Comport, CNRS-I3S, University of Nice  
Sophia-Antipolis, Andrew.Comport@cnrs.fr

**Associated paper:**

[1] M. Meilland, T. Drummond. and A. I. Comport. "A Unified Rolling Shutter and Motion Model for Dense 3D Visual Tracking". International Conference on Computer Vision. 2013.

[2] M. Meilland and A. I. Comport. "Super-resolution 3D Tracking and Mapping". IEEE International Conference on Robotics and Automation. 2013.

[3] M. Meilland, C. Barat and A. I. Comport. "3D High Dynamic Range Dense Visual SLAM and Its Application to Real-time Object Re-lighting". International Symposium on Mixed and Augmented Reality. 2013.

**Full list of authors:** M. Meilland, I3S/CNRS, University of Nice Sophia Antipolis France; T. Drummond, Electrical and Computer Systems Engineering, Monash University, Australia; A.I. Comport, I3S/CNRS, University of Nice Sophia Antipolis France.

**Abstract:** We will demonstrate our live real-time dense localisation and mapping approach that unifies two different representations commonly used to define dense models. On one hand, much research has looked at 3D dense model representations using voxel grids in 3D. On the other hand, image-based key-frame representations for dense environment mapping have been developed. The proposed demonstration will show a new model which unifies various concepts and exhibits the main advantages of each approach within a common multi-keyframe framework. One of the main results of the proposed approach is its ability to perform large scale reconstruction accurately at the scale of mapping a building in

real-time. The demonstrator also includes our work on super-resolution localisation and mapping, high dynamic range 3D mapping at large scales along with fast visual SLAM while modelling rolling shutter and motion blur deformations.

**Demo: 16**

**Date:** December 5 (Morning)

**Name:** Frank Steinbrücker, Technical University of Munich, frank.steinbruecker@in.tum.de

**Associated paper:** Paper No. 1688: "Large-Scale Multi-Resolution Surface Reconstruction from RGB-D Sequences".

**Full list of authors:** Frank Steinbrücker, Jürgen Sturm, Daniel Cremers (Technical University of Munich)

**Abstract:** We demonstrate a system for online reconstruction of large environments. We track the camera by aligning the dense images of a standard RGB-D sensor and perform online loop-closures by pose graph optimization. For the dense 3D reconstruction of the geometry and texture, we combine the RGB-D images in a volumetric signed distance function. The system runs in real-time on a standard laptop computer and it is capable of preserving fine details in large reconstruction volumes by embedding the signed distance function in a multi-scale octree data structure.

**Demo: 17**

**Date:** December 5 (Afternoon)

**Name:** Johannes Jordan, University of Erlangen-Nuremberg, johannes.jordan@cs.fau.de

**Associated paper:**

[1] J. Jordan, E. Angelopoulou. "Mean-shift Clustering for Interactive Multispectral Image Analysis", IEEE International Conference on Image Processing, September 2013

[2] J. Jordan, E. Angelopoulou. "Hyperspectral Image

Visualization With a 3-D Self-organizing Map”, IEEE 5th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, June 2013

[3] J. Jordan, E. Angelopoulou. “Supervised Multispectral Image Segmentation with Power Watersheds”, IEEE International Conference on Image Processing, September 2012, pp. 1585-1588.

[4] J. Jordan, E. Angelopoulou. “Gerbil - A Novel Software Framework for Visualization and Analysis in the Multispectral Domain”, VMV 2010: Vision, Modeling and Visualization, November 2010, pp. 259-266.

**Full list of authors:** Johannes Jordan, University of Erlangen-Nuremberg; Elli Angelopoulou, University of Erlangen-Nuremberg

**Abstract:** Multispectral imaging is an important tool for better understanding of image formation and reflectance phenomena. Wideband RGB data is not sufficient to draw meaningful interpretations from the captured data; instead, a significant amount of filter bands needs to be available. Research on computer vision methods that interpret or rely on the reflectance in the scene often relies on analyzing multispectral images. However, due to the high dimensionality of the data, both human observers as well as computers have difficulty interpreting this wealth of information.

In this demonstration, we show how analysis of a multispectral image can be conducted interactively by the researcher with a powerful visual framework. Our software 'Gerbil' incorporates a new paradigm for visual assistance of multispectral analysis that specifically addresses the lack of a seamless integration of spectral distribution and topology. It puts emphasis on the spectral gradient, which is shown to provide enhanced information for many reflectance analysis tasks. It also includes a rich toolbox for evaluation of image segmentation, global clustering and false-color display algorithms in the multispectral domain.

We demonstrate how several specific research interests in



scene reflectance can be tackled in a simple workflow on captured image data. The software is open-source (<http://gerbilvis.org/>) and researchers are encouraged to apply it on their own analysis tasks.

**Demo: 18**

**Date:** December 5 (Afternoon)

**Name:** Nabil Belbachir, AIT Austrian Institute of Technology, [nabil.belbachir@ait.ac.at](mailto:nabil.belbachir@ait.ac.at)

**Associated paper:** Papers published on the topic:

[1] A.N. Belbachir, R. Pflugfelder and R. Gmeiner, “A Neuromorphic Smart Camera for Real-time 360° Distortion-free Panoramas,” IEEE Conference on Distributed Smart Cameras, USA, 2010.

[2] A.N. Belbachir, M. Mayerhofer, D. Matolin and J. Colineau, “Real-time Panoramic Views Using BiCa360, the fast Rotating Dynamic Vision Sensor to up to 10 Rotations per Sec,” IEEE ISCAS2012 International Symposium on Circuits and Systems, South Korea, 2012.

Patents

[3] A.N. Belbachir and R. Pflugfelder, “Mechanism and approach for the acquisition of panoramic views (“360° Panorama visual scanner”), Austrian patent AT507.543. (15.10.2010); European patent application EP 2182720A2

[4] A.N. Belbachir and M. Mayerhofer, “Mechanism and approach for Real-time 360° 3D Panoramic Views (“Rotating Stereo Dynamic Line Vision Sensor”), German patent application DE 102011120718.3 (12. 12.2011)

**Full list of authors:** Manfred Mayerhofer, AIT Austrian Institute of Technology

**Abstract:** This live demonstration aims to show a new smart HDR camera TUCO-3D providing real-time 360° panoramic stereo views by scanning the surroundings horizontally at a speed up to 10 rotations per seconds. The demonstration will show event data generated from the rotating dynamic vision

sensor (Contrast sensitive) and rendered in 360° panoramic views (10 panoramas per sec). Furthermore, we will show the corresponding intensity images, which are reconstructed by an on board software in real-time for up to 10 panoramas /sec. In a further demonstration anaglyph representations of the panorama will be displayed, which show the 3D views by using specific glasses (anaglyph), which will be made available to the visitors. Panoramic 3D vision can be very fascinating to the visitors of our demo desk as it is not frequently used in consumer technologies. However this technology is gaining importance for several applications like wide area surveillance and robotics. The visitors will have the opportunity to learn about the technology behind the camera and interact for new experiments.

**Demo: 19**

**Date:** December 5 (Afternoon)

**Name:** Lixin Fan, Nokia Research Center, fanlixin@ieee.org

**Associated paper:** No associated papers.

**Full list of authors:** Lixin Fan, Junsheng Fu, Yu You, Kimmo Roimela, Petri Piippo and Ville-Veikko Mattila

**Abstract:** Mobile phone industry has witnessed tremendous growth of embedded sensory units including GPS, accelerometer, gyros and, most importantly, multi-megapixel camera. Huge amount of photographs & video clips are captured by nowadays mobile phone cameras, and consequently, it remains a challenging task to present these media data in an intuitive and user-friendly manner. By exploiting 3D reconstructed scene as well as camera pose information, “Photo tourism” system provided an intriguing user experience for browsing large collections of photographs. One of the limitations of the system, however, is its incapability to register photos that belong to parts of the scenes disconnected from the reconstructed ones. This demo presents our research endeavor to improve 3D photo browsing user experiences by combining user created photographs with

large scale Nokia HERE 3D map data. Using the map street-view images, our system is able to register disconnected photos that cannot be handled by PhotoToursim system. Furthermore, since reconstructed scenes and camera poses are now placed in a global map context, rendering photographs together with 3D map data e.g. Terrain and textured 3D building models leads to a novel photo navigation experience on mobile devices.

**Demo: 20**

**Date:** December 5 (Afternoon)

**Name:** Lixin Fan, Nokia Research Center, fanlixin@ieee.org

**Associated paper:** No associated papers.

**Full list of authors:** Yu You, Lixin Fan, Alain Boyer, Tuomas Kantonen and Ville-Veikko Mattila

**Abstract:** Augmented reality applications are gaining popularity due to increased capabilities of modern mobile devices. Creating AR content however is tedious and traditionally done on desktop environments by professionals, with extensive knowledge and/or even programming skills required. In this demo, we demonstrate a complete mobile approach for creating vision-based AR in both indoor and outdoor environment. Using hyperlinks, Web mashups are built to dynamically augment the physical world by normal users without programing skills.