

# **2017 IEEE International Conference on Computer Vision (ICCV 2017)**

**Venice, Italy  
22 - 29 October 2017**

**Pages 1-744**



**IEEE Catalog Number: CFP17198-POD  
ISBN: 978-1-5386-1033-6**

**Copyright © 2017 by the Institute of Electrical and Electronics Engineers, Inc.  
All Rights Reserved**

*Copyright and Reprint Permissions:* Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

***\*\*\* This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP17198-POD
ISBN (Print-On-Demand):	978-1-5386-1033-6
ISBN (Online):	978-1-5386-1032-9
ISSN:	1550-5499

**Additional Copies of This Publication Are Available From:**

Curran Associates, Inc  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: (845) 758-0400  
Fax: (845) 758-2633  
E-mail: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

CURRAN ASSOCIATES INC.  
**proceedings**  
.com

# 2017 IEEE International Conference on Computer Vision

## ICCV 2017

### Table of Contents

Message from the General Chairs .....	xliii
Message from the Program Chairs.....	xliv
Organizing Committee .....	xlviii
Area Chairs .....	xliv
2017 Outstanding and Emergency Reviewers.....	l

---

### 3D Vision & Video Analysis

#### Oral Session 1

Globally-Optimal Inlier Set Maximisation for Simultaneous Camera Pose and Feature Correspondence .....	1
<i>Dylan Campbell, Lars Petersson, Laurent Kneip, and Hongdong Li</i>	
Robust Pseudo Random Fields for Light-Field Stereo Matching .....	11
<i>Chao-Tsung Huang</i>	
A Lightweight Approach for On-the-Fly Reflectance Estimation .....	20
<i>Kihwan Kim, Jinwei Gu, Stephen Tyree, Pavlo Molchanov, Matthias Nießner, and Jan Kautz</i>	
Distributed Very Large Scale Bundle Adjustment by Global Camera Consensus .....	29
<i>Runze Zhang, Siyu Zhu, Tian Fang, and Long Quan</i>	
Practical Projective Structure from Motion (P2SfM) .....	39
<i>Ludovic Magerand and Alessio Del Bue</i>	

#### Spotlight Session 1

Anticipating Daily Intention Using On-wrist Motion Triggered Sensing .....	48
<i>Tz-Ying Wu, Ting-An Chien, Cheng-Sheng Chan, Chan-Wei Hu, and Min Sun</i>	
Rethinking Reprojection: Closing the Loop for Pose-Aware Shape Reconstruction from a Single Image .....	57
<i>Rui Zhu, Hamed Kiani Galoogahi, Chaoyang Wang, and Simon Lucey</i>	
End-to-End Learning of Geometry and Context for Deep Stereo Regression .....	66
<i>Alex Kendall, Hayk Martirosyan, Saumitro Dasgupta, and Peter Henry</i>	

Using Sparse Elimination for Solving Minimal Problems in Computer Vision .....	76
<i>Janne Heikkilä</i>	
High-Resolution Shape Completion Using Deep Neural Networks for Global Structure and Local Geometry Inference .....	85
<i>Xiaoguang Han, Zhen Li, Haibin Huang, Evangelos Kalogerakis, and Yizhou Yu</i>	
Temporal Tessellation: A Unified Approach for Video Analysis .....	94
<i>Dotan Kaufman, Gil Levi, Tal Hassner, and Lior Wolf</i>	
Learning Policies for Adaptive Tracking with Deep Feature Cascades .....	105
<i>Chen Huang, Simon Lucey, and Deva Ramanan</i>	
Temporal Shape Super-Resolution by Intra-frame Motion Encoding Using High-fps Structured Light .....	115
<i>Yuki Shiba, Satoshi Ono, Ryo Furukawa, Shinsaku Hiura, and Hiroshi Kawasaki</i>	

## Poster 1

Real-Time Monocular Pose Estimation of 3D Objects Using Temporally Consistent Local Color Histograms .....	124
<i>Henning Tjaden, Ulrich Schwanecke, and Elmar Schömer</i>	
CAD Priors for Accurate and Flexible Instance Reconstruction .....	133
<i>Tolga Birdal and Slobodan Ilic</i>	
Colored Point Cloud Registration Revisited .....	143
<i>Jaesik Park, Qian-Yi Zhou, and Vladlen Koltun</i>	
Learning Compact Geometric Features .....	153
<i>Marc Houry, Qian-Yi Zhou, and Vladlen Koltun</i>	
Joint Layout Estimation and Global Multi-view Registration for Indoor Reconstruction .....	162
<i>Jeong-Kyun Lee, Jaewon Yea, Min-Gyu Park, and Kuk-Jin Yoon</i>	
A Geometric Framework for Statistical Analysis of Trajectories with Distinct Temporal Spans .....	172
<i>Rudrasis Chakraborty, Vikas Singh, Nagesh Adluru, and Baba C. Vemuri</i>	
An Optimal Transportation Based Univariate Neuroimaging Index .....	182
<i>Liang Mi, Wen Zhang, Junwei Zhang, Yonghui Fan, Dhruvan Goradia, Kewei Chen, Eric M. Reiman, Xianfeng Gu, and Yalin Wang</i>	
S <sup>3</sup> FD: Single Shot Scale-Invariant Face Detector .....	192
<i>Shifeng Zhang, Xiangyu Zhu, Zhen Lei, Hailin Shi, Xiaobo Wang, and Stan Z. Li</i>	
Amulet: Aggregating Multi-level Convolutional Features for Salient Object Detection .....	202
<i>Pingping Zhang, Dong Wang, Huchuan Lu, Hongyu Wang, and Xiang Ruan</i>	
Learning Uncertain Convolutional Features for Accurate Saliency Detection .....	212
<i>Pingping Zhang, Dong Wang, Huchuan Lu, Hongyu Wang, and Baocai Yin</i>	
Zero-Order Reverse Filtering .....	222
<i>Xin Tao, Chao Zhou, Xiaoyong Shen, Jue Wang, and Jiaya Jia</i>	
Learning Blind Motion Deblurring .....	231
<i>Patrick Wieschollek, Michael Hirsch, Bernhard Schölkopf, and Hendrik P.A. Lensch</i>	

Joint Adaptive Sparsity and Low-Rankness on the Fly: An Online Tensor Reconstruction Scheme for Video Denoising .....	241
<i>Bihan Wen, Yanjun Li, Luke Pfister, and Yoram Bresler</i>	
Learning to Super-Resolve Blurry Face and Text Images .....	251
<i>Xiangyu Xu, Deqing Sun, Jinshan Pan, Yujin Zhang, Hanspeter Pfister, and Ming-Hsuan Yang</i>	
Video Frame Interpolation via Adaptive Separable Convolution .....	261
<i>Simon Niklaus, Long Mai, and Feng Liu</i>	
Deep Occlusion Reasoning for Multi-camera Multi-target Detection .....	271
<i>Pierre Baqué, Francois Fleuret, and Pascal Fua</i>	
Encouraging LSTMs to Anticipate Actions Very Early .....	280
<i>Mohammad Sadegh Aliakbarian, Fatemeh Sadat Saleh, Mathieu Salzmann, Basura Fernando, Lars Petersson, and Lars Andersson</i>	
PathTrack: Fast Trajectory Annotation with Path Supervision .....	290
<i>Santiago Manen, Michael Gygli, Dengxin Dai, and Luc Van Gool</i>	
Tracking the Untrackable: Learning to Track Multiple Cues with Long-Term Dependencies .....	300
<i>Amir Sadeghian, Alexandre Alahi, and Silvio Savarese</i>	
MirrorFlow: Exploiting Symmetries in Joint Optical Flow and Occlusion Estimation .....	312
<i>Junhwa Hur and Stefan Roth</i>	
Tracking as Online Decision-Making: Learning a Policy from Streaming Videos with Reinforcement Learning .....	322
<i>James Supancic III and Deva Ramanan</i>	
Non-convex Rank/Sparsity Regularization and Local Minima .....	332
<i>Carl Olsson, Marcus Carlsson, Fredrik Andersson, and Viktor Larsson</i>	
A Revisit of Sparse Coding Based Anomaly Detection in Stacked RNN Framework .....	341
<i>Weixin Luo, Wen Liu, and Shenghua Gao</i>	
HydraPlus-Net: Attentive Deep Features for Pedestrian Analysis .....	350
<i>Xihui Liu, Haiyu Zhao, Maoqing Tian, Lu Sheng, Jing Shao, Shuai Yi, Junjie Yan, and Xiaogang Wang</i>	
No Fuss Distance Metric Learning Using Proxies .....	360
<i>Yair Movshovitz-Attias, Alexander Toshev, Thomas K. Leung, Sergey Ioffe, and Saurabh Singh</i>	
Benchmarking and Error Diagnosis in Multi-instance Pose Estimation .....	369
<i>Matteo Ruggero Ronchi and Pietro Perona</i>	
Orientation Invariant Feature Embedding and Spatial Temporal Regularization for Vehicle Re-identification .....	379
<i>Zhongdao Wang, Luming Tang, Xihui Liu, Zhuliang Yao, Shuai Yi, Jing Shao, Junjie Yan, Shengjin Wang, Hongsheng Li, and Xiaogang Wang</i>	
Fashion Forward: Forecasting Visual Style in Fashion .....	388
<i>Ziad Al-Halah, Rainer Stiefelhagen, and Kristen Grauman</i>	

Towards 3D Human Pose Estimation in the Wild: A Weakly-Supervised Approach .....	398
<i>Xingyi Zhou, Qixing Huang, Xiao Sun, Xiangyang Xue, and Yichen Wei</i>	
Flow-Guided Feature Aggregation for Video Object Detection .....	408
<i>Xizhou Zhu, Yujie Wang, Jifeng Dai, Lu Yuan, and Yichen Wei</i>	
Reasoning About Fine-Grained Attribute Phrases Using Reference Games .....	418
<i>Jong-Chyi Su, Chenyun Wu, Huaizu Jiang, and Subhansu Maji</i>	
DeNet: Scalable Real-Time Object Detection with Directed Sparse Sampling .....	428
<i>Lachlan Tychsen-Smith and Lars Petersson</i>	
MIHash: Online Hashing with Mutual Information .....	437
<i>Fatih Cakir, Kun He, Sarah Adel Bargal, and Stan Sclaroff</i>	
SafetyNet: Detecting and Rejecting Adversarial Examples Robustly .....	446
<i>Jiajun Lu, Theerasit Issaranon, and David Forsyth</i>	
Recurrent Models for Situation Recognition .....	455
<i>Arun Mallya and Svetlana Lazebnik</i>	
Multi-label Image Recognition by Recurrently Discovering Attentional Regions .....	464
<i>Zhouxia Wang, Tianshui Chen, Guanbin Li, Ruijia Xu, and Liang Lin</i>	
Deep Determinantal Point Process for Large-Scale Multi-label Classification .....	473
<i>Pengtao Xie, Ruslan Salakhutdinov, Luntian Mou, and Eric P. Xing</i>	
Visual Semantic Planning Using Deep Successor Representations .....	483
<i>Yuke Zhu, Daniel Gordon, Eric Kolve, Dieter Fox, Li Fei-Fei, Abhinav Gupta, Roozbeh Mottaghi, and Ali Farhadi</i>	
Neural Person Search Machines .....	493
<i>Hao Liu, Jiashi Feng, Zequn Jie, Karlekar Jayashree, Bo Zhao, Meibin Qi, Jianguo Jiang, and Shuicheng Yan</i>	
DualNet: Learn Complementary Features for Image Recognition .....	502
<i>Saihui Hou, Xu Liu, and Zilei Wang</i>	
Higher-Order Integration of Hierarchical Convolutional Activations for Fine-Grained Visual Categorization .....	511
<i>Sijia Cai, Wangmeng Zuo, and Lei Zhang</i>	
Show, Adapt and Tell: Adversarial Training of Cross-Domain Image Captioner .....	521
<i>Tseng-Hung Chen, Yuan-Hong Liao, Ching-Yao Chuang, Wan-Ting Hsu, Jianlong Fu, and Min Sun</i>	
Attribute Recognition by Joint Recurrent Learning of Context and Correlation .....	531
<i>Jingya Wang, Xiatian Zhu, Shaogang Gong, and Wei Li</i>	
VegFru: A Domain-Specific Dataset for Fine-Grained Visual Categorization .....	541
<i>Saihui Hou, Yushan Feng, and Zilei Wang</i>	
Increasing CNN Robustness to Occlusions by Reducing Filter Support .....	550
<i>Elad Osherov and Michael Lindenbaum</i>	
Exploiting Multi-grain Ranking Constraints for Precisely Searching Visually-similar Vehicles .....	562
<i>Ke Yan, Yonghong Tian, Yaowei Wang, Wei Zeng, and Tiejun Huang</i>	

Recurrent Scale Approximation for Object Detection in CNN .....	571
<i>Yu Liu, Hongyang Li, Junjie Yan, Fangyin Wei, Xiaogang Wang, and Xiaoou Tang</i>	
Embedding 3D Geometric Features for Rigid Object Part Segmentation .....	580
<i>Yafei Song, Xiaowu Chen, Jia Li, and Qinqing Zhao</i>	
Towards Context-Aware Interaction Recognition for Visual Relationship Detection .....	589
<i>Bohan Zhuang, Lingqiao Liu, Chunhua Shen, and Ian Reid</i>	
When Unsupervised Domain Adaptation Meets Tensor Representations .....	599
<i>Hao Lu, Lei Zhang, Zhiguo Cao, Wei Wei, Ke Xian, Chunhua Shen, and Anton van den Hengel</i>	
Look, Listen and Learn .....	609
<i>Relja Arandjelovic and Andrew Zisserman</i>	
Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization .....	618
<i>Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra</i>	
Image-Based Localization Using LSTMs for Structured Feature Correlation .....	627
<i>F. Walch, C. Hazirbas, L. Leal-Taixé, T. Sattler, S. Hilsenbeck, and D. Cremers</i>	
Personalized Image Aesthetics .....	638
<i>Jian Ren, Xiaohui Shen, Zhe Lin, Radomír Mech, and David J. Foran</i>	
Predicting Deeper into the Future of Semantic Segmentation .....	648
<i>Pauline Luc, Natalia Neverova, Camille Couprie, Jakob Verbeek, and Yann LeCun</i>	
Coordinating Filters for Faster Deep Neural Networks .....	658
<i>Wei Wen, Cong Xu, Chunpeng Wu, Yandan Wang, Yiran Chen, and Hai Li</i>	
Unsupervised Representation Learning by Sorting Sequences .....	667
<i>Hsin-Ying Lee, Jia-Bin Huang, Maneesh Singh, and Ming-Hsuan Yang</i>	
A Read-Write Memory Network for Movie Story Understanding .....	677
<i>Seil Na, Sangho Lee, Jisung Kim, and Gunhee Kim</i>	
SegFlow: Joint Learning for Video Object Segmentation and Optical Flow .....	686
<i>Jingchun Cheng, Yi-Hsuan Tsai, Shengjin Wang, and Ming-Hsuan Yang</i>	
Unsupervised Action Discovery and Localization in Videos .....	696
<i>Khurram Soomro and Mubarak Shah</i>	
Dense-Captioning Events in Videos .....	706
<i>Ranjay Krishna, Kenji Hata, Frederic Ren, Li Fei-Fei, and Juan Carlos Niebles</i>	
Learning Long-Term Dependencies for Action Recognition with a Biologically-Inspired Deep Network .....	716
<i>Yemin Shi, Yonghong Tian, Yaowei Wang, Wei Zeng, and Tiejun Huang</i>	
Compressive Quantization for Fast Object Instance Search in Videos .....	726
<i>Tan Yu, Zhenzhen Wang, and Junsong Yuan</i>	
Complex Event Detection by Identifying Reliable Shots from Untrimmed Videos .....	736
<i>Hehe Fan, Xiaojun Chang, De Cheng, Yi Yang, Dong Xu, and Alexander G. Hauptmann</i>	
Deep Direct Regression for Multi-oriented Scene Text Detection .....	745
<i>Wenhao He, Xu-Yao Zhang, Fei Yin, and Cheng-Lin Liu</i>	

## Recognition I

### Oral Session 2

Open Set Domain Adaptation .....	754
<i>Pau Panareda Busto and Juergen Gall</i>	
Deformable Convolutional Networks .....	764
<i>Jifeng Dai, Haozhi Qi, Yuwen Xiong, Yi Li, Guodong Zhang, Han Hu, and Yichen Wei</i>	
Ensemble Diffusion for Retrieval .....	774
<i>Song Bai, Zhichao Zhou, Jingdong Wang, Xiang Bai, Longin Jan Latecki, and Qi Tian</i>	
FoveaNet: Perspective-Aware Urban Scene Parsing .....	784
<i>Xin Li, Zequn Jie, Wei Wang, Changsong Liu, Jimei Yang, Xiaohui Shen, Zhe Lin, Qiang Chen, Shuicheng Yan, and Jiashi Feng</i>	
Beyond Planar Symmetry: Modeling Human Perception of Reflection and Rotation Symmetries in the Wild .....	793
<i>Christopher Funk and Yanxi Liu</i>	

### Spotlight Session 2

Learning to Reason: End-to-End Module Networks for Visual Question Answering .....	804
<i>Ronghang Hu, Jacob Andreas, Marcus Rohrbach, Trevor Darrell, and Kate Saenko</i>	
Hard-Aware Deeply Cascaded Embedding .....	814
<i>Yuhui Yuan, Kuiyuan Yang, and Chao Zhang</i>	
Query-Guided Regression Network with Context Policy for Phrase Grounding .....	824
<i>Kan Chen, Rama Kovvuri, and Ram Nevatia</i>	
SuBiC: A Supervised, Structured Binary Code for Image Search .....	833
<i>Himalaya Jain, Joaquin Zepeda, Patrick Pérez, and Rémi Gribonval</i>	
Revisiting Unreasonable Effectiveness of Data in Deep Learning Era .....	843
<i>Chen Sun, Abhinav Shrivastava, Saurabh Singh, and Abhinav Gupta</i>	
A Generative Model of People in Clothing .....	853
<i>Christoph Lassner, Gerard Pons-Moll, and Peter V. Gehler</i>	
Escape from Cells: Deep Kd-Networks for the Recognition of 3D Point Cloud Models .....	863
<i>Roman Klokov and Victor Lempitsky</i>	
Improved Image Captioning via Policy Gradient optimization of SPIDER .....	873
<i>Siqi Liu, Zhenhai Zhu, Ning Ye, Sergio Guadarrama, and Kevin Murphy</i>	

### Poster Session 2

Rolling Shutter Correction in Manhattan World .....	882
<i>Pulak Purkait, Christopher Zach, and Ales Leonardis</i>	
Local-to-Global Point Cloud Registration Using a Dictionary of Viewpoint Descriptors .....	891
<i>David Avidar, David Malah, and Meir Barzohar</i>	



3D-PRNN: Generating Shape Primitives with Recurrent Neural Networks .....	900
<i>Chuhang Zou, Ersin Yumer, Jimei Yang, Duygu Ceylan, and Derek Hoiem</i>	
BodyFusion: Real-Time Capture of Human Motion and Surface Geometry Using a Single Depth Camera .....	910
<i>Tao Yu, Kaiwen Guo, Feng Xu, Yuan Dong, Zhaoqi Su, Jianhui Zhao, Jianguo Li, Qionghai Dai, and Yebin Liu</i>	
Quasiconvex Plane Sweep for Triangulation with Outliers .....	920
<i>Qiangong Zhang, Tat-Jun Chin, and David Suter</i>	
“Maximizing Rigidity” Revisited: A Convex Programming Approach for Generic 3D Shape Reconstruction from Multiple Perspective Views .....	929
<i>Pan Ji, Hongdong Li, Yuchao Dai, and Ian Reid</i>	
Surface Registration via Foliation .....	938
<i>Xiaopeng Zheng, Chengfeng Wen, Na Lei, Ming Ma, and Xianfeng Gu</i>	
Rolling-Shutter-Aware Differential SfM and Image Rectification .....	948
<i>Bingbing Zhuang, Loong-Fah Cheong, and Gim Hee Lee</i>	
Corner-Based Geometric Calibration of Multi-focus Plenoptic Cameras .....	957
<i>Sotiris Nousias, Francois Chadebecq, Jonas Pichat, Pearse Keane, Sebastien Ourselin, and Christos Bergeles</i>	
Focal Track: Depth and Accommodation with Oscillating Lens Deformation .....	966
<i>Qi Guo, Emma Alexander, and Todd Zickler</i>	
Reconfiguring the Imaging Pipeline for Computer Vision .....	975
<i>Mark Buckler, Suren Jayasuriya, and Adrian Sampson</i>	
Catadioptric HyperSpectral Light Field Imaging .....	985
<i>Yujia Xue, Kang Zhu, Qiang Fu, Xilin Chen, and Jingyi Yu</i>	
Cross-View Asymmetric Metric Learning for Unsupervised Person Re-Identification .....	994
<i>Hong-Xing Yu, Ancong Wu, and Wei-Shi Zheng</i>	
Real Time Eye Gaze Tracking with 3D Deformable Eye-Face Model .....	1003
<i>Kang Wang and Qiang Ji</i>	
Ensemble Deep Learning for Skeleton-Based Action Recognition Using Temporal Sliding LSTM Networks .....	1012
<i>Inwoong Lee, Doyoung Kim, Seoungyoon Kang, and Sanghoon Lee</i>	
How Far are We from Solving the 2D & 3D Face Alignment Problem? (and a Dataset of 230,000 3D Facial Landmarks) .....	1021
<i>Adrian Bulat and Georgios Tzimiropoulos</i>	
Large Pose 3D Face Reconstruction from a Single Image via Direct Volumetric CNN Regression .....	1031
<i>Aaron S. Jackson, Adrian Bulat, Vasileios Argyriou, and Georgios Tzimiropoulos</i>	
RankIQA: Learning from Rankings for No-Reference Image Quality Assessment .....	1040
<i>Xialei Liu, Joost van de Weijer, and Andrew D. Bagdanov</i>	

Look, Perceive and Segment: Finding the Salient Objects in Images via Two-stream Fixation-Semantic CNNs .....	1050
<i>Xiaowu Chen, Anlin Zheng, Jia Li, and Feng Lu</i>	
Delving into Salient Object Subitizing and Detection .....	1059
<i>Shengfeng He, Jianbo Jiao, Xiaodan Zhang, Guoqiang Han, and Rynson W.H. Lau</i>	
Visual Relationship Detection with Internal and External Linguistic Knowledge Distillation .....	1068
<i>Ruichi Yu, Ang Li, Vlad I. Morariu, and Larry S. Davis</i>	
Learning Discriminative Data Fitting Functions for Blind Image Deblurring .....	1077
<i>Jinshan Pan, Jiangxin Dong, Yu-Wing Tai, Zhixun Su, and Ming-Hsuan Yang</i>	
Video Deblurring via Semantic Segmentation and Pixel-Wise Non-linear Kernel .....	1086
<i>Wenqi Ren, Jinshan Pan, Xiaochun Cao, and Ming-Hsuan Yang</i>	
On-demand Learning for Deep Image Restoration .....	1095
<i>Ruohan Gao and Kristen Grauman</i>	
Multi-channel Weighted Nuclear Norm Minimization for Real Color Image Denoising .....	1105
<i>Jun Xu, Lei Zhang, David Zhang, and Xiangchu Feng</i>	
Coherent Online Video Style Transfer .....	1114
<i>Dongdong Chen, Jing Liao, Lu Yuan, Nenghai Yu, and Gang Hua</i>	
SHaPE: A Novel Graph Theoretic Algorithm for Making Consensus-Based Decisions in Person Re-identification Systems .....	1124
<i>Arko Barman and Shishir K. Shah</i>	
Need for Speed: A Benchmark for Higher Frame Rate Object Tracking .....	1134
<i>Hamed Kiani Galoogahi, Ashton Fagg, Chen Huang, Deva Ramanan, and Simon Lucey</i>	
Learning Background-Aware Correlation Filters for Visual Tracking .....	1144
<i>Hamed Kiani Galoogahi, Ashton Fagg, and Simon Lucey</i>	
Robust Object Tracking Based on Temporal and Spatial Deep Networks .....	1153
<i>Zhu Teng, Junliang Xing, Qiang Wang, Congyan Lang, Songhe Feng, and Yi Jin</i>	
Real-Time Hand Tracking under Occlusion from an Egocentric RGB-D Sensor .....	1163
<i>Franziska Mueller, Dushyant Mehta, Oleksandr Sotnychenko, Srinath Sridhar, Dan Casas, and Christian Theobalt</i>	
Predicting Human Activities Using Stochastic Grammar .....	1173
<i>Siyuan Qi, Siyuan Huang, Ping Wei, and Song-Chun Zhu</i>	
ProbFlow: Joint Optical Flow and Uncertainty Estimation .....	1182
<i>Anne S. Wannenwetsch, Margret Keuper, and Stefan Roth</i>	
Sublabel-Accurate Discretization of Nonconvex Free-Discontinuity Problems .....	1192
<i>Thomas Möllenhoff and Daniel Cremers</i>	
DeepContext: Context-Encoding Neural Pathways for 3D Holistic Scene Understanding .....	1201
<i>Yinda Zhang, Mingru Bai, Pushmeet Kohli, Shahram Izadi, and Jianxiong Xiao</i>	
BAM! The Behance Artistic Media Dataset for Recognition Beyond Photography .....	1211
<i>Michael J. Wilber, Chen Fang, Hailin Jin, Aaron Hertzmann, John Collomosse, and Serge Belongie</i>	

Adversarial PoseNet: A Structure-Aware Convolutional Network for Human Pose Estimation .....	1221
<i>Yu Chen, Chunhua Shen, Xiu-Shen Wei, Lingqiao Liu, and Jian Yang</i>	
An Empirical Study of Language CNN for Image Captioning .....	1231
<i>Jiuxiang Gu, Gang Wang, Jianfei Cai, and Tsuhan Chen</i>	
Attributes2Classname: A Discriminative Model for Attribute-Based Unsupervised Zero-Shot Learning .....	1241
<i>Berkan Demirel, Ramazan Gokberk Cinbis, and Nazli Ikizler-Cinbis</i>	
Areas of Attention for Image Captioning .....	1251
<i>Marco Pedersoli, Thomas Lucas, Cordelia Schmid, and Jakob Verbeek</i>	
Generative Modeling of Audible Shapes for Object Perception .....	1260
<i>Zhoutong Zhang, Jiajun Wu, Qiuqia Li, Zhengjia Huang, James Traer, Josh H. McDermott, Joshua B. Tenenbaum, and William T. Freeman</i>	
Scene Graph Generation from Objects, Phrases and Region Captions .....	1270
<i>Yikang Li, Wanli Ouyang, Bolei Zhou, Kun Wang, and Xiaogang Wang</i>	
Recurrent Multimodal Interaction for Referring Image Segmentation .....	1280
<i>Chenxi Liu, Zhe Lin, Xiaohui Shen, Jimei Yang, Xin Lu, and Alan Yuille</i>	
Learning Feature Pyramids for Human Pose Estimation .....	1290
<i>Wei Yang, Shuang Li, Wanli Ouyang, Hongsheng Li, and Xiaogang Wang</i>	
Structured Attentions for Visual Question Answering .....	1300
<i>Chen Zhu, Yanpeng Zhao, Shuaiyi Huang, Kewei Tu, and Yi Ma</i>	
Cut, Paste and Learn: Surprisingly Easy Synthesis for Instance Detection .....	1310
<i>Debidatta Dwibedi, Ishan Misra, and Martial Hebert</i>	
Cascaded Feature Network for Semantic Segmentation of RGB-D Images .....	1320
<i>Di Lin, Guangyong Chen, Daniel Cohen-Or, Pheng-Ann Heng, and Hui Huang</i>	
Encoder Based Lifelong Learning .....	1329
<i>Amal Rannen, Rahaf Aljundi, Matthew B. Blaschko, and Tinne Tuytelaars</i>	
Transitive Invariance for Self-Supervised Visual Representation Learning .....	1338
<i>Xiaolong Wang, Kaiming He, and Abhinav Gupta</i>	
Weakly Supervised Learning of Deep Metrics for Stereo Reconstruction .....	1348
<i>Stepan Tulyakov, Anton Ivanov, and Francois Fleuret</i>	
Fine-Grained Recognition in the Wild: A Multi-task Domain Adaptation Approach .....	1358
<i>Timnit Gebru, Judy Hoffman, and Li Fei-Fei</i>	
SORT: Second-Order Response Transform for Visual Recognition .....	1368
<i>Yan Wang, Lingxi Xie, Chenxi Liu, Siyuan Qiao, Ya Zhang, Wenjun Zhang, Qi Tian, and Alan Yuille</i>	
Adversarial Examples for Semantic Segmentation and Object Detection .....	1378
<i>Cihang Xie, Jianyu Wang, Zhishuai Zhang, Yuyin Zhou, Lingxi Xie, and Alan Yuille</i>	
Genetic CNN .....	1388
<i>Lingxi Xie and Alan Yuille</i>	

Channel Pruning for Accelerating Very Deep Neural Networks .....	1398
<i>Yihui He, Xiangyu Zhang, and Jian Sun</i>	
Infinite Latent Feature Selection: A Probabilistic Latent Graph-Based Ranking Approach .....	1407
<i>Giorgio Roffo, Simone Melzi, Umberto Castellani, and Alessandro Vinciarelli</i>	
Video Fill In the Blank Using LR/RL LSTMs with Spatial-Temporal Attentions .....	1416
<i>Amir Mazaheri, Dong Zhang, and Mubarak Shah</i>	
Primary Video Object Segmentation via Complementary CNNs and Neighborhood Reversible Flow .....	1426
<i>Jia Li, Anlin Zheng, Xiaowu Chen, and Bin Zhou</i>	
Attentive Semantic Video Generation Using Captions .....	1435
<i>Tanya Marwah, Gaurav Mittal, and Vineeth N. Balasubramanian</i>	
Following Gaze in Video .....	1444
<i>Adrià Recasens, Carl Vondrick, Aditya Khosla, and Antonio Torralba</i>	
Adaptive RNN Tree for Large-Scale Human Action Recognition .....	1453
<i>Wenbo Li, Longyin Wen, Ming-Ching Chang, Ser Nam Lim, and Siwei Lyu</i>	
Spatio-Temporal Person Retrieval via Natural Language Queries .....	1462
<i>Masataka Yamaguchi, Kuniaki Saito, Yoshitaka Ushiku, and Tatsuya Harada</i>	
Automatic Spatially-Aware Fashion Concept Discovery .....	1472
<i>Xintong Han, Zuxuan Wu, Phoenix X. Huang, Xiao Zhang, Menglong Zhu, Yuan Li, Yang Zhao, and Larry S. Davis</i>	
ChromaTag: A Colored Marker and Fast Detection Algorithm .....	1481
<i>Joseph DeGol, Timothy Bretl, and Derek Hoiem</i>	
Adversarial Image Perturbation for Privacy Protection A Game Theory Perspective .....	1491
<i>Seong Joon Oh, Mario Fritz, and Bernt Schiele</i>	
WeText: Scene Text Detection under Weak Supervision .....	1501
<i>Shangxuan Tian, Shijian Lu, and Chongshou Li</i>	

## **Vision for X**

### **Oral Session 3**

Arbitrary Style Transfer in Real-Time with Adaptive Instance Normalization .....	1510
<i>Xun Huang and Serge Belongie</i>	
Photographic Image Synthesis with Cascaded Refinement Networks .....	1520
<i>Qifeng Chen and Vladlen Koltun</i>	
SSD-6D: Making RGB-Based 3D Detection and 6D Pose Estimation Great Again .....	1530
<i>Wadim Kehl, Fabian Manhardt, Federico Tombari, Slobodan Ilic, and Nassir Navab</i>	
Unsupervised Creation of Parameterized Avatars .....	1539
<i>Lior Wolf, Yaniv Taigman, and Adam Polyak</i>	
Learning for Active 3D Mapping .....	1548
<i>Karel Zimmermann, Tomáš Petříček, Vojtěch Šalanský, and Tomáš Svoboda</i>	

### Poster Session 3

Toward Perceptually-Consistent Stereo: A Scanline Study .....	1557
<i>Jialiang Wang, Daniel Glasner, and Todd Zickler</i>	
Surface Normals in the Wild .....	1566
<i>Weifeng Chen, Donglai Xiang, and Jia Deng</i>	
Unsupervised Learning of Stereo Matching .....	1576
<i>Chao Zhou, Hong Zhang, Xiaoyong Shen, and Jiaya Jia</i>	
Unrestricted Facial Geometry Reconstruction Using Image-to-Image Translation .....	1585
<i>Matan Sela, Elad Richardson, and Ron Kimmel</i>	
Learned Multi-patch Similarity .....	1595
<i>Wilfried Hartmann, Silvano Galliani, Michal Havlena, Luc Van Gool, and Konrad Schindler</i>	
Click Here: Human-Localized Keypoints as Guidance for Viewpoint Estimation .....	1604
<i>Ryan Szeto and Jason J. Corso</i>	
Unsupervised Adaptation for Deep Stereo .....	1614
<i>Alessio Tonioni, Matteo Poggi, Stefano Mattoccia, and Luigi Di Stefano</i>	
Composite Focus Measure for High Quality Depth Maps .....	1623
<i>Parikshit Sakurikar and P. J. Narayanan</i>	
Reconstruction-Based Disentanglement for Pose-Invariant Face Recognition .....	1632
<i>Xi Peng, Xiang Yu, Kihyuk Sohn, Dimitris N. Metaxas, and Manmohan Chandraker</i>	
Recurrent 3D-2D Dual Learning for Large-Pose Facial Landmark Detection .....	1642
<i>Shengtao Xiao, Jiashi Feng, Luoqi Liu, Xuecheng Nie, Wei Wang, Shuicheng Yan, and Ashraf Kassim</i>	
Anchored Regression Networks Applied to Age Estimation and Super Resolution .....	1652
<i>Eirikur Agustsson, Radu Timofte, and Luc Van Gool</i>	
Infant Footprint Recognition .....	1662
<i>Eryun Liu</i>	
Self-Paced Kernel Estimation for Robust Blind Image Deblurring .....	1670
<i>Dong Gong, Mingkui Tan, Yanning Zhang, Anton van den Hengel, and Qinfeng Shi</i>	
Super-Trajectory for Video Segmentation .....	1680
<i>Wenguan Wang, Jianbing Shen, Jianwen Xie, and Fatih Porikli</i>	
Be Your Own Prada: Fashion Synthesis with Structural Coherence .....	1689
<i>Shizhan Zhu, Sanja Fidler, Raquel Urtasun, Dahua Lin, and Chen Change Loy</i>	
Wavelet-SRNet: A Wavelet-Based CNN for Multi-scale Face Super Resolution .....	1698
<i>Huaibo Huang, Ran He, Zhenan Sun, and Tieniu Tan</i>	
Learning Gaze Transitions from Depth to Improve Video Saliency Estimation .....	1707
<i>George Leifman, Dmitry Rudoy, Tristan Swedish, Eduardo Bayro-Corrochano, and Ramesh Raskar</i>	
Joint Convolutional Analysis and Synthesis Sparse Representation for Single Image Layer Separation .....	1717
<i>Shuhang Gu, Deyu Meng, Wangmeng Zuo, and Lei Zhang</i>	

Modelling the Scene Dependent Imaging in Cameras with a Deep Neural Network .....	1726
<i>Seonghyeon Nam and Seon Joo Kim</i>	
Transformed Low-Rank Model for Line Pattern Noise Removal .....	1735
<i>Yi Chang, Luxin Yan, and Sheng Zhong</i>	
Weakly Supervised Manifold Learning for Dense Semantic Object Correspondence .....	1744
<i>Utkarsh Gaur and B. S. Manjunath</i>	
PanNet: A Deep Network Architecture for Pan-Sharpener .....	1753
<i>Junfeng Yang, Xueyang Fu, Yuwen Hu, Yue Huang, Xinghao Ding, and John Paisley</i>	
Dual Motion GAN for Future-Flow Embedded Video Prediction .....	1762
<i>Xiaodan Liang, Lisa Lee, Wei Dai, and Eric P. Xing</i>	
Online Robust Image Alignment via Subspace Learning from Gradient Orientations .....	1771
<i>Qingqing Zheng, Yi Wang, and Pheng Ann Heng</i>	
Learning Dynamic Siamese Network for Visual Object Tracking .....	1781
<i>Qing Guo, Wei Feng, Ce Zhou, Rui Huang, Liang Wan, and Song Wang</i>	
High Order Tensor Formulation for Convolutional Sparse Coding .....	1790
<i>Adel Bibi and Bernard Ghanem</i>	
Learning Proximal Operators: Using Denoising Networks for Regularizing Inverse Imaging Problems .....	1799
<i>Tim Meinhardt, Michael Moeller, Caner Hazirbas, and Daniel Cremers</i>	
ScaleNet: Guiding Object Proposal Generation in Supermarkets and Beyond .....	1809
<i>Siyuan Qiao, Wei Shen, Weichao Qiu, Chenxi Liu, and Alan Yuille</i>	
Temporal Dynamic Graph LSTM for Action-Driven Video Object Detection .....	1819
<i>Yuan Yuan, Xiaodan Liang, Xiaolong Wang, Dit-Yan Yeung, and Abhinav Gupta</i>	
VQS: Linking Segmentations to Questions and Answers for Supervised Attention in VQA and Question-Focused Semantic Segmentation .....	1829
<i>Chuang Gan, Yandong Li, Haoxiang Li, Chen Sun, and Boqing Gong</i>	
Multi-modal Factorized Bilinear Pooling with Co-attention Learning for Visual Question Answering .....	1839
<i>Zhou Yu, Jun Yu, Jianping Fan, and Dacheng Tao</i>	
SCNet: Learning Semantic Correspondence .....	1849
<i>Kai Han, Rafael S. Rezende, Bumsub Ham, Kwan-Yee K. Wong, Minsu Cho, Cordelia Schmid, and Jean Ponce</i>	
Soft Proposal Networks for Weakly Supervised Object Localization .....	1859
<i>Yi Zhu, Yanzhao Zhou, Qixiang Ye, Qiang Qiu, and Jianbin Jiao</i>	
Class Rectification Hard Mining for Imbalanced Deep Learning .....	1869
<i>Qi Dong, Shaogang Gong, and Xiatian Zhu</i>	
Generating High-Quality Crowd Density Maps Using Contextual Pyramid CNNs .....	1879
<i>Vishwanath A. Sindagi and Vishal M. Patel</i>	
See the Glass Half Full: Reasoning About Liquid Containers, Their Volume and Content .....	1889
<i>Roosbeh Mottaghi, Connor Schenck, Dieter Fox, and Ali Farhadi</i>	

Hierarchical Multimodal LSTM for Dense Visual-Semantic Embedding .....	1899
<i>Zhenxing Niu, Mo Zhou, Le Wang, Xinbo Gao, and Gang Hua</i>	
Identity-Aware Textual-Visual Matching with Latent Co-attention .....	1908
<i>Shuang Li, Tong Xiao, Hongsheng Li, Wei Yang, and Xiaogang Wang</i>	
Learning Deep Neural Networks for Vehicle Re-ID with Visual-spatio-Temporal Path Proposals .....	1918
<i>Yantao Shen, Tong Xiao, Hongsheng Li, Shuai Yi, and Xiaogang Wang</i>	
Learning from Noisy Labels with Distillation .....	1928
<i>Yuncheng Li, Jianchao Yang, Yale Song, Liangliang Cao, Jiebo Luo, and Li-Jia Li</i>	
DSOD: Learning Deeply Supervised Object Detectors from Scratch .....	1937
<i>Zhiqiang Shen, Zhuang Liu, Jianguo Li, Yu-Gang Jiang, Yurong Chen, and Xiangyang Xue</i>	
Phrase Localization and Visual Relationship Detection with Comprehensive Image-Language Cues .....	1946
<i>Bryan A. Plummer, Arun Mallya, Christopher M. Cervantes, Julia Hockenmaier, and Svetlana Lazebnik</i>	
Chained Cascade Network for Object Detection .....	1956
<i>Wanli Ouyang, Kun Wang, Xin Zhu, and Xiaogang Wang</i>	
VPGNet: Vanishing Point Guided Network for Lane and Road Marking Detection and Recognition .....	1965
<i>Seokju Lee, Junsik Kim, Jae Shin Yoon, Seunghak Shin, Oleksandr Bailo, Namil Kim, Tae-Hee Lee, Hyun Seok Hong, Seung-Hoon Han, and In So Kweon</i>	
Unsupervised Learning of Important Objects from First-Person Videos .....	1974
<i>Gedas Bertasius, Hyun Soo Park, Stella X. Yu, and Jianbo Shi</i>	
An Analysis of Visual Question Answering Algorithms .....	1983
<i>Kushal Kafle and Christopher Kanan</i>	
A Two Stream Siamese Convolutional Neural Network for Person Re-identification .....	1992
<i>Dahjung Chung, Khalid Tahboub, and Edward J. Delp</i>	
Joint Learning of Object and Action Detectors .....	2001
<i>Vicky Kalogeiton, Philippe Weinzaepfel, Vittorio Ferrari, and Cordelia Schmid</i>	
No More Discrimination: Cross City Adaptation of Road Scene Segmenters .....	2011
<i>Yi-Hsin Chen, Wei-Yu Chen, Yu-Ting Chen, Bo-Cheng Tsai, Yu-Chiang Frank Wang, and Min Sun</i>	
Open Vocabulary Scene Parsing .....	2021
<i>Hang Zhao, Xavier Puig, Bolei Zhou, Sanja Fidler, and Antonio Torralba</i>	
Learned Watershed: End-to-End Learning of Seeded Segmentation .....	2030
<i>Steffen Wolf, Lukas Schott, Ullrich Köthe, and Fred Hamprecht</i>	
Curriculum Domain Adaptation for Semantic Segmentation of Urban Scenes .....	2039
<i>Yang Zhang, Philip David, and Boqing Gong</i>	
Scale-Adaptive Convolutions for Scene Parsing .....	2050
<i>Rui Zhang, Sheng Tang, Yongdong Zhang, Jintao Li, and Shuicheng Yan</i>	

Privacy-Preserving Visual Learning Using Doubly Permuted Homomorphic Encryption .....	2059
<i>Ryo Yonetani, Vishnu Naresh Boddeti, Kris M. Kitani, and Yoichi Sato</i>	
Multi-task Self-Supervised Visual Learning .....	2070
<i>Carl Doersch and Andrew Zisserman</i>	
A Self-Balanced Min-Cut Algorithm for Image Clustering .....	2080
<i>Xiaojun Chen, Joshua Zhexue Haung, Feiping Nie, Renjie Chen, and Qingyao Wu</i>	
Is Second-Order Information Helpful for Large-Scale Visual Recognition? .....	2089
<i>Peihua Li, Jiangtao Xie, Qilong Wang, and Wangmeng Zuo</i>	
Factorized Bilinear Models for Image Recognition .....	2098
<i>Yanghao Li, Naiyan Wang, Jiaying Liu, and Xiaodi Hou</i>	
Octree Generating Networks: Efficient Convolutional Architectures for High-resolution 3D Outputs .....	2107
<i>Maxim Tatarchenko, Alexey Dosovitskiy, and Thomas Brox</i>	
Truncating Wide Networks Using Binary Tree Architectures .....	2116
<i>Yan Zhangy, Mete Ozayy, Shuohao Li, and Takayuki Okatani</i>	
Bringing Background into the Foreground: Making All Classes Equal in Weakly-Supervised Video Semantic Segmentation .....	2125
<i>Fatemeh Sadat Saleh, Mohammad Sadegh Aliakbarian, Mathieu Salzmann, Lars Petersson, and Jose M. Alvarez</i>	
View Adaptive Recurrent Neural Networks for High Performance Human Action Recognition from Skeleton Data .....	2136
<i>Pengfei Zhang, Cuiling Lan, Junliang Xing, Wenjun Zeng, Jianru Xue, and Nanning Zheng</i>	
Joint Discovery of Object States and Manipulation Actions .....	2146
<i>Jean-Baptiste Alayrac, Josef Sivic, Ivan Laptev, and Simon Lacoste-Julien</i>	
What Actions are Needed for Understanding Human Actions in Videos? .....	2156
<i>Gunnar A. Sigurdsson, Olga Russakovsky, and Abhinav Gupta</i>	
Lattice Long Short-Term Memory for Human Action Recognition .....	2166
<i>Lin Sun, Kui Jia, Kevin Chen, Dit Yan Yeung, Bertram E. Shi, and Silvio Savarese</i>	
Common Action Discovery and Localization in Unconstrained Videos .....	2176
<i>Jiong Yang and Junsong Yuan</i>	
Pixel-Level Matching for Video Object Segmentation Using Convolutional Neural Networks .....	2186
<i>Jae Shin Yoon, Francois Rameau, Junsik Kim, Seokju Lee, Seunghak Shin, and In So Kweon</i>	
Am I a Baller? Basketball Performance Assessment from First-Person Videos .....	2196
<i>Gedas Bertasius, Hyun Soo Park, Stella X. Yu, and Jianbo Shi</i>	
Deep Cropping via Attention Box Prediction and Aesthetics Assessment .....	2205
<i>Wenguan Wang and Jianbing Shen</i>	
Raster-to-Vector: Revisiting Floorplan Transformation .....	2214
<i>Chen Liu, Jiajun Wu, Pushmeet Kohli, and Yasutaka Furukawa</i>	



Deep TextSpotter: An End-to-End Trainable Scene Text Localization and Recognition Framework .....	2223
<i>Michal Buřta, Lukáš Neumann, and Jiri Matas</i>	

## Vision for X & Computational Photography

### Spotlight Session 3

Playing for Benchmarks .....	2232
<i>Stephan R. Richter, Zeeshan Hayder, and Vladlen Koltun</i>	
Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks .....	2242
<i>Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros</i>	
GANs for Biological Image Synthesis .....	2252
<i>Anton Osokin, Anatole Chessel, Rafael E. Carazo Salas, and Federico Vaggi</i>	
Learning to Synthesize a 4D RGBD Light Field from a Single Image .....	2262
<i>Pratul P. Srinivasan, Tongzhou Wang, Ashwin Sreelal, Ravi Ramamoorthi, and Ren Ng</i>	
Neural EPI-Volume Networks for Shape from Light Field .....	2271
<i>Stefan Heber, Wei Yu, and Thomas Pock</i>	
Material Editing Using a Physically Based Rendering Network .....	2280
<i>Guilin Liu, Duygu Ceylan, Ersin Yumer, Jimei Yang, and Jyh-Ming Lien</i>	
Turning Corners into Cameras: Principles and Methods .....	2289
<i>Katherine L. Bouman, Vickie Ye, Adam B. Yedidia, Frédo Durand, Gregory W. Wornell, Antonio Torralba, and William T. Freeman</i>	
Linear Differential Constraints for Photo-Polarimetric Height Estimation .....	2298
<i>Silvia Tozza, William A. P. Smith, Dizhong Zhu, Ravi Ramamoorthi, and Edwin R. Hancock</i>	

### Poster Session 4

Polynomial Solvers for Saturated Ideals .....	2307
<i>Viktor Larsson, Kalle Åström, and Magnus Oskarsson</i>	
Shape Inpainting Using 3D Generative Adversarial Network and Recurrent Convolutional Networks .....	2317
<i>Weiyue Wang, Qiangui Huang, Suyu You, Chao Yang, and Ulrich Neumann</i>	
SurfaceNet: An End-to-End 3D Neural Network for Multiview Stereopsis .....	2326
<i>Mengqi Ji, Juergen Gall, Haitian Zheng, Yebin Liu, and Lu Fang</i>	
Making Minimal Solvers for Absolute Pose Estimation Compact and Robust .....	2335
<i>Viktor Larsson, Zuzana Kukelova, and Yinqiang Zheng</i>	
3D Surface Detail Enhancement from a Single Normal Map .....	2344
<i>Wuyuan Xie, Miaohui Wang, Xianbiao Qi, and Lei Zhang</i>	
RMPE: Regional Multi-person Pose Estimation .....	2353
<i>Hao-Shu Fang, Shuqin Xie, Yu-Wing Tai, and Cewu Lu</i>	
Online Video Object Detection Using Association LSTM .....	2363
<i>Yongyi Lu, Cewu Lu, and Chi-Keung Tang</i>	

PolyFit: Polygonal Surface Reconstruction from Point Clouds .....	2372
<i>Liangliang Nan and Peter Wonka</i>	
Progressive Large Scale-Invariant Image Matching in Scale Space .....	2381
<i>Lei Zhou, Siyu Zhu, Tianwei Shen, Jinglu Wang, Tian Fang, and Long Quan</i>	
Efficient Global 2D-3D Matching for Camera Localization in a Large-Scale 3D Map .....	2391
<i>Liu Liu, Hongdong Li, and Yuchao Dai</i>	
Multi-view Non-rigid Refinement and Normal Selection for High Quality 3D Reconstruction .....	2401
<i>Sk. Mohammadul Haque and Venu Madhav Govindu</i>	
Multi-stage Multi-recursive-input Fully Convolutional Networks for Neuronal Boundary Detection .....	2410
<i>Wei Shen, Bin Wang, Yuan Jiang, Yan Wang, and Alan Yuille</i>	
Depth and Image Restoration from Light Field in a Scattering Medium .....	2420
<i>Jiandong Tian, Zak Murez, Tong Cui, Zhen Zhang, David Kriegman, and Ravi Ramamoorthi</i>	
Video Reflection Removal Through Spatio-Temporal Optimization .....	2430
<i>Ajay Nandoriya, Mohamed Elgharib, Changil Kim, Mohamed Hefeeda, and Wojciech Matusik</i>	
Efficient Online Local Metric Adaptation via Negative Samples for Person Re-identification .....	2439
<i>Jiahuan Zhou, Pei Yu, Wei Tang, and Ying Wu</i>	
Stepwise Metric Promotion for Unsupervised Video Person Re-identification .....	2448
<i>Zimo Liu, Dong Wang, and Huchuan Lu</i>	
Beyond Face Rotation: Global and Local Perception GAN for Photorealistic and Identity Preserving Frontal View Synthesis .....	2458
<i>Rui Huang, Shu Zhang, Tianyu Li, and Ran He</i>	
Group Re-identification via Unsupervised Transfer of Sparse Features Encoding .....	2468
<i>Giuseppe Lisanti, Niki Martinel, Alberto Del Bimbo, and Gian Luca Foresti</i>	
Visual Transformation Aided Contrastive Learning for Video-Based Kinship Verification .....	2478
<i>Hamdi Dibeklioglu</i>	
Decoder Network over Lightweight Reconstructed Feature for Fast Semantic Style Transfer .....	2488
<i>Ming Lu, Hao Zhao, Anbang Yao, Feng Xu, Yurong Chen, and Li Zhang</i>	
Blind Image Deblurring with Outlier Handling .....	2497
<i>Jiangxin Dong, Jinshan Pan, Zhixun Su, and Ming-Hsuan Yang</i>	
Paying Attention to Descriptions Generated by Image Captioning Models .....	2506
<i>Hamed R. Tavakoliy, Rakshith Shetty, Ali Borji, and Jorma Laaksonen</i>	
Fast Image Processing with Fully-Convolutional Networks .....	2516
<i>Qifeng Chen, Jia Xu, Vladlen Koltun, and Intel Labs</i>	
Robust Video Super-Resolution with Learned Temporal Dynamics .....	2526
<i>Ding Liu, Zhaowen Wang, Yuchen Fan, Xianming Liu, Zhangyang Wang, Shiyu Chang, and Thomas Huang</i>	

Should We Encode Rain Streaks in Video as Deterministic or Stochastic? .....	2535
<i>Wei Wei, Lixuan Yi, Qi Xie, Qian Zhao, Deyu Meng, and Zongben Xu</i>	
Joint Bi-layer Optimization for Single-Image Rain Streak Removal .....	2545
<i>Lei Zhu, Chi-Wing Fu, Dani Lischinski, and Pheng-Ann Heng</i>	
Low-Dimensionality Calibration through Local Anisotropic Scaling for Robust Hand Model Personalization .....	2554
<i>Edoardo Remelli, Anastasia Tkach, Andrea Tagliasacchi, and Mark Pauly</i>	
Non-Markovian Globally Consistent Multi-object Tracking .....	2563
<i>Andrii Maksai, Xinchao Wang, François Fleuret, and Pascal Fua</i>	
CREST: Convolutional Residual Learning for Visual Tracking .....	2574
<i>Yibing Song, Chao Ma, Lijun Gong, Jiawei Zhang, Rynson W.H. Lau, and Ming-Hsuan Yang</i>	
Volumetric Flow Estimation for Incompressible Fluids Using the Stationary Stokes Equations .....	2584
<i>Katrin Lasinger, Christoph Vogel, and Konrad Schindler</i>	
Bounding Boxes, Segmentations and Object Coordinates: How Important is Recognition for 3D Scene Flow Estimation in Autonomous Driving Scenarios? .....	2593
<i>Aseem Behl, Omid Hosseini Jafari, Siva Karthik Mustikovela, Hassan Abu Alhajja, Carsten Rother, and Andreas Geiger</i>	
Performance Guaranteed Network Acceleration via High-Order Residual Quantization .....	2603
<i>Zefan Li, Bingbing Ni, Wenjun Zhang, Xiaokang Yang, and Wen Gao</i>	
Deep Metric Learning with Angular Loss .....	2612
<i>Jian Wang, Feng Zhou, Shilei Wen, Xiao Liu, and Yuanqing Lin</i>	
Compositional Human Pose Regression .....	2621
<i>Xiao Sun, Jiayang Shang, Shuang Liang, and Yichen Wei</i>	
MUTAN: Multimodal Tucker Fusion for Visual Question Answering .....	2631
<i>Hedi Ben-younes, Remi Cadene, Matthieu Cord, and Nicolas Thome</i>	
Revisiting IM2GPS in the Deep Learning Era .....	2640
<i>Nam Vo, Nathan Jacobs, and James Hays</i>	
Scene Parsing with Global Context Embedding .....	2650
<i>Wei-Chih Hung, Yi-Hsuan Tsai, Xiaohui Shen, Zhe Lin, Kalyan Sunkavalli, Xin Lu, and Ming-Hsuan Yang</i>	
A Simple Yet Effective Baseline for 3d Human Pose Estimation .....	2659
<i>Julieta Martinez, Rayat Hossain, Javier Romero, and James J. Little</i>	
Dual-Glance Model for Deciphering Social Relationships .....	2669
<i>Junnan Li, Yongkang Wong, Qi Zhao, and Mohan S. Kankanhalli</i>	
Sketching with Style: Visual Search with Sketches and Aesthetic Context .....	2679
<i>John Collomosse, Tu Bui, Michael Wilber, Chen Fang, and Hailin Jin</i>	
Point Set Registration with Global-Local Correspondence and Transformation Estimation .....	2688
<i>Su Zhang, Yang Yang, Kun Yang, Yi Luo, and Sim Heng Ong</i>	

SceneNet RGB-D: Can 5M Synthetic Images Beat Generic ImageNet Pre-training on Indoor Segmentation? .....	2697
<i>John McCormac, Ankur Handa, Stefan Leutenegger, and Andrew J. Davison</i>	
A Unified Model for Near and Remote Sensing .....	2707
<i>Scott Workman, Menghua Zhai, David J. Crandall, and Nathan Jacobs</i>	
Directionally Convolutional Networks for 3D Shape Segmentation .....	2717
<i>Haotian Xu, Ming Dong, and Zichun Zhong</i>	
AMAT: Medial Axis Transform for Natural Images .....	2727
<i>Stavros Tsogkas and Sven Dickinson</i>	
Deep Dual Learning for Semantic Image Segmentation .....	2737
<i>Ping Luo, Guangrun Wang, Liang Lin, and Xiaogang Wang</i>	
Regional Interactive Image Segmentation Networks .....	2746
<i>JunHao Liew, Yunchao Wei, Wei Xiong, Sim-Heng Ong, and Jiashi Feng</i>	
Learning Efficient Convolutional Networks through Network Slimming .....	2755
<i>Zhuang Liu, Jianguo Li, Zhiqiang Shen, Gao Huang, Shoumeng Yan, and Changshui Zhang</i>	
CVAE-GAN: Fine-Grained Image Generation through Asymmetric Training .....	2764
<i>Jianmin Bao, Dong Chen, Fang Wen, Houqiang Li, and Gang Hua</i>	
Universal Adversarial Perturbations Against Semantic Image Segmentation .....	2774
<i>Jan Hendrik Metzen, Mummadi Chaithanya Kumar, Thomas Brox, and Volker Fischer</i>	
Associative Domain Adaptation .....	2784
<i>Philip Haeusser, Thomas Frerix, Alexander Mordvintsev, and Daniel Cremers</i>	
Introspective Neural Networks for Generative Modeling .....	2793
<i>Justin Lazarow, Long Jin, and Zhuowen Tu</i>	
Towards a Unified Compositional Model for Visual Pattern Modeling .....	2803
<i>Wei Tang, Pei Yu, Jiahuan Zhou, and Ying Wu</i>	
Least Squares Generative Adversarial Networks .....	2813
<i>Xudong Mao, Qing Li, Haoran Xie, Raymond Y.K. Lau, Zhen Wang, and Stephen Paul Smolley</i>	
Centered Weight Normalization in Accelerating Training of Deep Neural Networks .....	2822
<i>Lei Huang, Xianglong Liu, Yang Liu, Bo Lang, and Dacheng Tao</i>	
Deep Growing Learning .....	2831
<i>Guangcong Wang, Xiaohua Xie, Jianhuang Lai, and Jiaxuan Zhuo</i>	
Smart Mining for Deep Metric Learning .....	2840
<i>Ben Harwood, Vijay Kumar B. G, Gustavo Carneiro, Ian Reid, and Tom Drummond</i>	
Temporal Generative Adversarial Nets with Singular Value Clipping .....	2849
<i>Masaki Saito, Eiichi Matsumoto, and Shunta Saito</i>	
Sampling Matters in Deep Embedding Learning .....	2859
<i>R. Manmatha, Chao-Yuan Wu, Alexander J. Smola, and Philipp Krähenbühl</i>	
DualGAN: Unsupervised Dual Learning for Image-to-Image Translation .....	2868
<i>Zili Yi, Hao Zhang, Ping Tan, and Minglun Gong</i>	

Learning View-Invariant Features for Person Identification in Temporally Synchronized Videos Taken by Wearable Cameras .....	2877
<i>Kang Zheng, Xiaochuan Fan, Yuewei Lin, Hao Guo, Hongkai Yu, Dazhou Guo, and Song Wang</i>	
MarioQA: Answering Questions by Watching Gameplay Videos .....	2886
<i>Jonghwan Mun, Paul Hongsuck Seo, Ilchae Jung, and Bohyung Han</i>	
SBGAR: Semantics Based Group Activity Recognition .....	2895
<i>Xin Li and Mooi Choo Chuah</i>	
Trespassing the Boundaries: Labeling Temporal Bounds for Object Interactions in Egocentric Video .....	2905
<i>Davide Moltisanti, Michael Wray, Walterio Mayol-Cuevas, and Dima Damen</i>	
Unmasking the Abnormal Events in Video .....	2914
<i>Radu Tudor Ionescu, Sorina Smeureanu, Bogdan Alexe, and Marius Popescu</i>	
Chained Multi-stream Networks Exploiting Pose, Motion, and Appearance for Action Classification and Detection .....	2923
<i>Mohammadreza Zolfaghari, Gabriel L. Oliveira, Nima Sedaghat, and Thomas Brox</i>	
Temporal Action Detection with Structured Segment Networks .....	2933
<i>Yue Zhao, Yuanjun Xiong, Limin Wang, Zhirong Wu, Xiaoou Tang, and Dahua Lin</i>	
Jointly Recognizing Object Fluents and Tasks in Egocentric Videos .....	2943
<i>Yang Liu, Ping Wei, and Song-Chun Zhu</i>	
Transferring Objects: Joint Inference of Container and Human Pose .....	2952
<i>Hanqing Wang, Wei Liang, and Lap-Fai Yu</i>	
Interpretable Learning for Self-Driving Cars by Visualizing Causal Attention .....	2961
<i>Jinkyu Kim and John Canny</i>	

## Recognition 2

### Oral Session 4

Learning Cooperative Visual Dialog Agents with Deep Reinforcement Learning .....	2970
<i>Abhishek Das, Satwik Kottur, José M. F. Moura, Stefan Lee, and Dhruv Batra</i>	
Mask R-CNN .....	2980
<i>Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick</i>	
Towards Diverse and Natural Image Descriptions via a Conditional GAN .....	2989
<i>Bo Dai, Sanja Fidler, Raquel Urtasun, and Dahua Lin</i>	
Focal Loss for Dense Object Detection .....	2999
<i>Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár</i>	
Inferring and Executing Programs for Visual Reasoning .....	3008
<i>Justin Johnson, Bharath Hariharan, Laurens van der Maaten, Judy Hoffman, Li Fei-Fei, C. Lawrence Zitnick, and Ross Girshick</i>	

## Spotlight Session 4

Visual Forecasting by Imitating Dynamics in Natural Sequences .....	3018
<i>Kuo-Hao Zeng, William B. Shen, De-An Huang, Min Sun, and Juan Carlos Niebles</i>	
TorontoCity: Seeing the World with a Million Eyes .....	3028
<i>Shenlong Wang, Min Bai, Gellert Mattyus, Hang Chu, Wenjie Luo, Bin Yang, Justin Liang, Joel Cheverie, Sanja Fidler, and Raquel Urtasun</i>	
Low-Shot Visual Recognition by Shrinking and Hallucinating Features .....	3037
<i>Bharath Hariharan and Ross Girshick</i>	
A Coarse-Fine Network for Keypoint Localization .....	3047
<i>Shaoli Huang, Mingming Gong, and Dacheng Tao</i>	
Detect to Track and Track to Detect .....	3057
<i>Christoph Feichtenhofer, Axel Pinz, and Andrew Zisserman</i>	
Single Shot Text Detector with Regional Attention .....	3066
<i>Pan He, Weilin Huang, Tong He, Qile Zhu, Yu Qiao, and Xiaolin Li</i>	
SubUNets: End-to-End Hand Shape and Continuous Sign Language Recognition .....	3075
<i>Necati Cihan Camgoz, Simon Hadfield, Oscar Koller, and Richard Bowden</i>	
A Spatiotemporal Oriented Energy Network for Dynamic Texture Recognition .....	3085
<i>Isma Hadji and Richard P. Wildes</i>	

## Poster Session 5

Probabilistic Structure from Motion with Objects (PSfMO) .....	3094
<i>Paul Gay, Vaibhav Bansal, Cosimo Rubino, and Alessio Del Bue</i>	
A 3D Morphable Model of Craniofacial Shape and Texture Variation .....	3104
<i>Hang Dai, Nick Pears, William Smith, and Christian Duncan</i>	
Multi-view Dynamic Shape Refinement Using Local Temporal Integration .....	3113
<i>Vincent Leroy, Jean-Sebastien Franco, and Edmond Boyer</i>	
Learning Hand Articulations by Hallucinating Heat Distribution .....	3123
<i>Chiho Choi, Sangpil Kim, and Karthik Ramani</i>	
Intrinsic3D: High-Quality 3D Reconstruction by Joint Appearance and Geometry Optimization with Spatially-Varying Lighting .....	3133
<i>Robert Maier, Kihwan Kim, Daniel Cremers, Jan Kautz, and Matthias Nießner</i>	
Robust Hand Pose Estimation during the Interaction with an Unknown Object .....	3142
<i>Chiho Choi, Sang Ho Yoon, Chin-Ning Chen, and Karthik Ramani</i>	
Detailed Surface Geometry and Albedo Recovery from RGB-D Video under Natural Illumination .....	3152
<i>Xinxin Zuo, Sen Wang, Jiangbin Zheng, and Ruigang Yang</i>	
Monocular Free-Head 3D Gaze Tracking with Deep Learning and Geometry Constraints .....	3162
<i>Haoping Deng and Wangjiang Zhu</i>	

Filter Selection for Hyperspectral Estimation .....	3172
<i>Boaz Arad and Ohad Ben-Shahar</i>	
A Microfacet-Based Reflectance Model for Photometric Stereo with Highly Specular Surfaces .....	3181
<i>Lixiong Chen, Yinqiang Zheng, Boxin Shi, Art Subpa-Asa, and Imari Sato</i>	
Detecting Faces Using Inside Cascaded Contextual CNN .....	3190
<i>Kaipeng Zhang, Zhanpeng Zhang, Hao Wang, Zhifeng Li, Yu Qiao, and Wei Liu</i>	
A Novel Space-Time Representation on the Positive Semidefinite Cone for Facial Expression Recognition .....	3199
<i>Anis Kacem, Mohamed Daoudi, Boulbaba Ben Amor, and Juan Carlos Alvarez-Paiva</i>	
DeepCoder: Semi-Parametric Variational Autoencoders for Automatic Facial Action Coding .....	3209
<i>Dieu Linh Tran, Robert Walecki, Ognjen (Oggi) Rudovic, Stefanos Eleftheriadis, Bjorn Schuller, and Maja Pantic</i>	
Pose-Invariant Face Alignment with a Single CNN .....	3219
<i>Amin Jourabloo, Mao Ye, Xiaoming Liu, and Liu Ren</i>	
Unsupervised Learning of Object Landmarks by Factorized Spatial Embeddings .....	3229
<i>James Thewlis, Hakan Bilen, and Andrea Vedaldi</i>	
Deeply-Learned Part-Aligned Representations for Person Re-identification .....	3239
<i>Liming Zhao, Xi Li, Yueting Zhuang, and Jingdong Wang</i>	
Semantic Line Detection and Its Applications .....	3249
<i>Jun-Tae Lee, Han-UI Kim, Chul Lee, and Chang-Su Kim</i>	
A Generic Deep Architecture for Single Image Reflection Removal and Image Smoothing .....	3258
<i>Qingnan Fan, Jialong Yang, Gang Hua, Baoquan Chen, and David Wipf</i>	
Revisiting Cross-Channel Information Transfer for Chromatic Aberration Correction .....	3268
<i>Tiancheng Sun, Yifan Peng, and Wolfgang Heidrich</i>	
High-Quality Correspondence and Segmentation Estimation for Dual-Lens Smart-Phone Portraits .....	3277
<i>Xiaoyong Shen, Hongyun Gao, Xin Tao, Chao Zhou, and Jiaya Jia</i>	
Learning Visual Attention to Identify People with Autism Spectrum Disorder .....	3287
<i>Ming Jiang and Qi Zhao</i>	
DSLR-Quality Photos on Mobile Devices with Deep Convolutional Networks .....	3297
<i>Andrey Ignatov, Nikolay Kobyshev, Radu Timofte, Kenneth Vanhoey, and Luc Van Gool</i>	
Non-uniform Blind Deblurring by Reblurring .....	3306
<i>Yuval Bahat, Netalee Efrat, and Michal Irani</i>	
Misalignment-Robust Joint Filter for Cross-Modal Image Pairs .....	3315
<i>Takashi Shibata, Masayuki Tanaka, and Masatoshi Okutomi</i>	
Low-Rank Tensor Completion: A Pseudo-Bayesian Learning Approach .....	3325
<i>Wei Chen and Nan Song</i>	

DeepCD: Learning Deep Complementary Descriptors for Patch Representations .....	3334
<i>Tsun-Yi Yang, Jo-Han Hsu, Yen-Yu Lin, and Yung-Yu Chuang</i>	
Beyond Standard Benchmarks: Parameterizing Performance Evaluation in Visual Object Tracking .....	3343
<i>Luka Cehovin Zajc, Alan Lukežič, Aleš Leonardis, and Matej Kristan</i>	
The Pose Knows: Video Forecasting by Generating Pose Futures .....	3352
<i>Jacob Walker, Kenneth Marino, Abhinav Gupta, and Martial Hebert</i>	
What will Happen Next? Forecasting Player Moves in Sports Videos .....	3362
<i>Panna Felsen, Pulkit Agrawal, and Jitendra Malik</i>	
Robust Kronecker-Decomposable Component Analysis for Low-Rank Modeling .....	3372
<i>Mehdi Bahri, Yannis Panagakis, and Stefanos Zafeiriou</i>	
Recurrent Topic-Transition GAN for Visual Paragraph Generation .....	3382
<i>Xiaodan Liang, Zhiting Hu, Hao Zhang, Chuang Gan, and Eric P. Xing</i>	
A Two-Streamed Network for Estimating Fine-Scaled Depth Maps from Single RGB Images .....	3392
<i>Jun Li, Reinhard Klein, and Angela Yao</i>	
Weakly Supervised Object Localization Using Things and Stuff Transfer .....	3401
<i>Miaoqing Shi, Holger Caesar, and Vittorio Ferrari</i>	
Single Image Action Recognition Using Semantic Body Part Actions .....	3411
<i>Zhichen Zhao, Huimin Ma, and Shaodi You</i>	
Incremental Learning of Object Detectors without Catastrophic Forgetting .....	3420
<i>Konstantin Shmelkov, Cordelia Schmid, and Karteek Alahari</i>	
Generative Adversarial Networks Conditioned by Brain Signals .....	3430
<i>S. Palazzo, C. Spampinato, I. Kavasidis, D. Giordano, and M. Shah</i>	
Learning to Disambiguate by Asking Discriminative Questions .....	3439
<i>Yining Li, Chen Huang, Xiaoou Tang, and Chen Change Loy</i>	
Interpretable Explanations of Black Boxes by Meaningful Perturbation .....	3449
<i>Ruth C. Fong and Andrea Vedaldi</i>	
DeepRoadMapper: Extracting Road Topology from Aerial Images .....	3458
<i>Gellért Mátyus, Wenjie Luo, and Raquel Urtasun</i>	
Monocular 3D Human Pose Estimation by Predicting Depth on Joints .....	3467
<i>Bruce Xiaohan Nie, Ping Wei, and Song-Chun Zhu</i>	
Large-Scale Image Retrieval with Attentive Deep Local Features .....	3476
<i>Hyeonwoo Noh, Andre Araujo, Jack Sim, Tobias Weyand, and Bohyung Han</i>	
Deep Globally Constrained MRFs for Human Pose Estimation .....	3486
<i>Ioannis Marras, Petar Palasek, and Ioannis Patras</i>	
Predicting Visual Exemplars of Unseen Classes for Zero-Shot Learning .....	3496
<i>Soravit Changpinyo, Wei-Lun Chao, and Fei Sha</i>	
Multi-label Learning of Part Detectors for Heavily Occluded Pedestrian Detection .....	3506
<i>Chunluan Zhou and Junsong Yuan</i>	



SGN: Sequential Grouping Networks for Instance Segmentation .....	3516
<i>Shu Liu, Jiaya Jia, Sanja Fidler, and Raquel Urtasun</i>	
Adaptive Feeding: Achieving Fast and Accurate Detections by Adaptively Combining Object Detectors .....	3525
<i>Hong-Yu Zhou, Bin-Bin Gao, and Jianxin Wu</i>	
Aesthetic Critiques Generation for Photos .....	3534
<i>Kuang-Yu Chang, Kung-Hung Lu, and Chu-Song Chen</i>	
Hide-and-Seek: Forcing a Network to be Meticulous for Weakly-Supervised Object and Action Localization .....	3544
<i>Krishna Kumar Singh and Yong Jae Lee</i>	
Two-Phase Learning for Weakly Supervised Object Localization .....	3554
<i>Dahun Kim, Donghyeon Cho, and Donggeun Yoo</i>	
Curriculum Dropout .....	3564
<i>Pietro Morerio, Jacopo Cavazza, Riccardo Volpi, René Vidal, and Vittorio Murino</i>	
Predictor Combination at Test Time .....	3573
<i>Kwang In Kim, James Tompkin, and Christian Richardt</i>	
Guided Perturbations: Self-Corrective Behavior in Convolutional Neural Networks .....	3582
<i>Swami Sankaranarayanan, Arpit Jain, and Ser Nam Lim</i>	
Learning Robust Visual-Semantic Embeddings .....	3591
<i>Yao-Hung Hubert Tsai, Liang-Kang Huang, and Ruslan Salakhutdinov</i>	
PUNDA: Probabilistic Unsupervised Domain Adaptation for Knowledge Transfer Across Visual Categories .....	3601
<i>Behnam Gholami, Ognjen (Oggi) Rudovic, and Vladimir Pavlovic</i>	
Learning in an Uncertain World: Representing Ambiguity Through Multiple Hypotheses .....	3611
<i>Christian Rupprecht, Iro Laina, Robert DiPietro, and Maximilian Baust</i>	
CDTS: Collaborative Detection, Tracking, and Segmentation for Online Multiple Object Segmentation in Videos .....	3621
<i>Yeong Jun Koh and Chang-Su Kim</i>	
Temporal Superpixels Based on Proximity-Weighted Patch Matching .....	3630
<i>Se-Ho Lee, Won-Dong Jang, and Chang-Su Kim</i>	
Joint Detection and Recounting of Abnormal Events by Learning Deep Generic Knowledge .....	3639
<i>Ryota Hinami, Tao Mei, and Shin'ichi Satoh</i>	
TURN TAP: Temporal Unit Regression Network for Temporal Action Proposals .....	3648
<i>Jiyang Gao, Zhenheng Yang, Chen Sun, Kan Chen, and Ram Nevatia</i>	
Online Real-Time Multiple Spatiotemporal Action Localisation and Prediction .....	3657
<i>Gurkirt Singh, Suman Saha, Michael Sapienza, Philip Torr, and Fabio Cuzzolin</i>	
Leveraging Weak Semantic Relevance for Complex Video Event Classification .....	3667
<i>Heng Tao Shen, Chao Li, Jiwei Cao, Zi Huang, and Lei Zhu</i>	
Weakly Supervised Summarization of Web Videos .....	3677
<i>Rameswar Panda, Abir Das, Ziyang Wu, Jan Ernst, and Amit K. Roy-Chowdhury</i>	

FCN-rLSTM: Deep Spatio-Temporal Neural Networks for Vehicle Counting in City Cameras .....	3687
<i>Shanghang Zhang, Guanhang Wu, João P. Costeira, and José M. F. Moura</i>	
Fast Face-Swap Using Convolutional Neural Networks .....	3697
<i>Iryna Korshunova, Wenzhe Shi, Joni Dambre, and Lucas Theis</i>	
Towards a Visual Privacy Advisor: Understanding and Predicting Privacy Risks in Images .....	3706
<i>Tribhuvanesh Orekondy, Bernt Schiele, and Mario Fritz</i>	

## Face and Human Behaviour Analysis

### Oral Session 5

First-Person Activity Forecasting with Online Inverse Reinforcement Learning .....	3716
<i>Nicholas Rhinehart and Kris M. Kitani</i>	
Binarized Convolutional Landmark Localizers for Human Pose Estimation and Face Alignment with Limited Resources .....	3726
<i>Adrian Bulat and Georgios Tzimiropoulos</i>	
MoFA: Model-Based Deep Convolutional Face Autoencoder for Unsupervised Monocular Reconstruction .....	3735
<i>Ayush Tewari, Michael Zollhöfer, Hyeonwoo Kim, Pablo Garrido, Florian Bernard, Patrick Pérez, and Christian Theobalt</i>	
RPAN: An End-to-End Recurrent Pose-Attention Network for Action Recognition in Videos .....	3745
<i>Wenbin Du, Yali Wang, and Yu Qiao</i>	
Temporal Non-volume Preserving Approach to Facial Age-Progression and Age-Invariant Face Recognition .....	3755
<i>Chi Nhan Duong, Kha Gia Quach, Khoa Luu, T. Hoang Ngan Le, and Marios Savvides</i>	

### Spotlight Session 5

Attribute-Enhanced Face Recognition with Neural Tensor Fusion Networks .....	3764
<i>Guosheng Hu, Yang Hua, Yang Yuan, Zhihong Zhang, Zheng Lu, Sankha S. Mukherjee, Timothy M. Hospedales, Neil M. Robertson, and Yongxin Yang</i>	
Unlabeled Samples Generated by GAN Improve the Person Re-identification Baseline in Vitro .....	3774
<i>Zhedong Zheng, Liang Zheng, and Yi Yang</i>	
Egocentric Gesture Recognition Using Recurrent 3D Convolutional Neural Networks with Spatiotemporal Transformer Modules .....	3783
<i>Congqi Cao, Yifan Zhang, Yi Wu, Hanqing Lu, and Jian Cheng</i>	
Recursive Spatial Transformer (ReST) for Alignment-Free Face Recognition .....	3792
<i>Wanglong Wu, Meina Kan, Xin Liu, Yi Yang, Shiguang Shan, and Xilin Chen</i>	
Learning Discriminative Aggregation Network for Video-Based Face Recognition .....	3801
<i>Yongming Rao, Ji Lin, Jiwen Lu, and Jie Zhou</i>	

Synergy between Face Alignment and Tracking via Discriminative Global Consensus Optimization .....	3811
<i>Muhammad Haris Khan, John McDonagh, and Georgios Tzimiropoulos</i>	
SVDNet for Pedestrian Retrieval .....	3820
<i>Yifan Sun, Liang Zheng, Weijian Deng, and Shengjin Wang</i>	
Towards More Accurate Iris Recognition Using Deeply Learned Spatially Corresponding Features .....	3829
<i>Zijing Zhao and Ajay Kumar</i>	

## Poster Session 6

Semantically Informed Multiview Surface Refinement .....	3839
<i>Maroš Bláha, Mathias Rothermel, Martin R. Oswald, Torsten Sattler, Audrey Richard, Jan D. Wegner, Marc Pollefeys, and Konrad Schindler</i>	
BB8: A Scalable, Accurate, Robust to Partial Occlusion Method for Predicting the 3D Poses of Challenging Objects without Using Depth .....	3848
<i>Mahdi Rad and Vincent Lepetit</i>	
Modeling Urban Scenes from Pointclouds .....	3857
<i>William Nguatem and Helmut Mayer</i>	
Parameter-Free Lens Distortion Calibration of Central Cameras .....	3867
<i>Filippo Bergamasco, Luca Cosmo, Andrea Gasparetto, Andrea Albarelli, and Andrea Torsello</i>	
Pose Guided RGBD Feature Learning for 3D Object Pose Estimation .....	3876
<i>Vassileios Balntas, Andreas Doumanoglou, Caner Sahin, Juil Sock, Rigas Kouskouridas, and Tae-Kyun Kim</i>	
Efficient Global Illumination for Morphable Models .....	3885
<i>Andreas Schneider, Sandro Schönborn, Bernhard Egger, Lavrenti Frobeen, and Thomas Vetter</i>	
Low Compute and Fully Parallel Computer Vision with HashMatch .....	3894
<i>Sean Ryan Fanello, Julien Valentin, Adarsh Kowdle, Christoph Rhemann, Vladimir Tankovich, Carlo Ciliberto, Philip Davidson, and Shahram Izadi</i>	
Dense Non-rigid Structure-from-Motion and Shading with Unknown Albedos .....	3904
<i>Mathias Gallardo, Toby Collins, and Adrien Bartoli</i>	
From Point Clouds to Mesh Using Regression .....	3913
<i>L'ubor Ladický, Olivier Saurer, SoHyeon Jeong, Fabio Maninchedda, and Marc Pollefeys</i>	
Stereo DSO: Large-Scale Direct Sparse Visual Odometry with Stereo Cameras .....	3923
<i>Rui Wang, Martin Schwörer, and Daniel Cremers</i>	
Space-Time Localization and Mapping .....	3932
<i>Minhaeng Lee and Charless C. Fowlkes</i>	
Benchmarking Single-Image Reflection Removal Algorithms .....	3942
<i>Renjie Wan, Boxin Shi, Ling-Yu Duan, Ah-Hwee Tan, and Alex C. Kot</i>	

Attention-Aware Deep Reinforcement Learning for Video Face Recognition .....	3951
<i>Yongming Rao, Jiwen Lu, and Jie Zhou</i>	
Learning to Fuse 2D and 3D Image Cues for Monocular Body Pose Estimation .....	3961
<i>Bugra Tekin, Pablo Márquez-Neila, Mathieu Salzmann, and Pascal Fua</i>	
Deep Facial Action Unit Recognition from Partially Labeled Data .....	3971
<i>Shan Wu, Shangfei Wang, Bowen Pan, and Qiang Ji</i>	
Pose-Driven Deep Convolutional Model for Person Re-identification .....	3980
<i>Chi Su, Jianing Li, Shiliang Zhang, Junliang Xing, Wen Gao, and Qi Tian</i>	
Recognition of Action Units in the Wild with Deep Nets and a New Global-Local Loss .....	3990
<i>C. Fabian Benitez-Quiroz, Yan Wang, and Aleix M. Martinez</i>	
Faster than Real-Time Facial Alignment: A 3D Spatial Transformer Network Approach in Unconstrained Poses .....	4000
<i>Chandrasekhar Bhagavatula, Chenchen Zhu, Khoa Luu, and Marios Savvides</i>	
Towards Large-Pose Face Frontalization in the Wild .....	4010
<i>Xi Yin, Xiang Yu, Kihyuk Sohn, Xiaoming Liu, and Manmohan Chandraker</i>	
A Joint Intrinsic-Extrinsic Prior Model for Retinex .....	4020
<i>Bolun Cai, Xianming Xu, Kailing Guo, Kui Jia, Bin Hu, and Dacheng Tao</i>	
Going Unconstrained with Rolling Shutter Deblurring .....	4030
<i>Mahesh Mohan M. R. and A. N. Rajagopalan</i>	
A Stagewise Refinement Model for Detecting Salient Objects in Images .....	4039
<i>Tiantian Wang, Ali Borji, Lihe Zhang, Pingping Zhang, and Huchuan Lu</i>	
From Square Pieces to Brick Walls: The Next Challenge in Solving Jigsaw Puzzles .....	4049
<i>Shir Gur and Ohad Ben-Shahar</i>	
Online Video Deblurring via Dynamic Temporal Blending Network .....	4058
<i>Tae Hyun Kim, Kyoung Mu Lee, Bernhard Schölkopf, and Michael Hirsch</i>	
Supervision by Fusion: Towards Unsupervised Learning of Deep Salient Object Detector .....	4068
<i>Dingwen Zhang, Junwei Han, and Yu Zhang</i>	
Fast Multi-image Matching via Density-Based Clustering .....	4077
<i>Roberto Tron, Xiaowei Zhou, Carlos Esteves, and Kostas Daniilidis</i>	
Characterizing and Improving Stability in Neural Style Transfer .....	4087
<i>Agrim Gupta, Justin Johnson, Alexandre Alahi, and Li Fei-Fei</i>	
Cross-Modal Deep Variational Hashing .....	4097
<i>Venice Erin Liong, Jiwen Lu, Yap-Peng Tan, and Jie Zhou</i>	
Spatial Memory for Context Reasoning in Object Detection .....	4106
<i>Xinlei Chen and Abhinav Gupta</i>	
Deep Binaries: Encoding Semantic-Rich Cues for Efficient Textual-Visual Cross Retrieval .....	4117
<i>Yuming Shen, Li Liu, Ling Shao, and Jingkuan Song</i>	
Learning a Recurrent Residual Fusion Network for Multimodal Matching .....	4127
<i>Yu Liu, Yanming Guo, Erwin M. Bakker, and Michael S. Lew</i>	

Rotational Subgroup Voting and Pose Clustering for Robust 3D Object Recognition .....	4137
<i>Anders Glent Buch, Liliya Kiforenko, and Dirk Kraft</i>	
CoupleNet: Coupling Global Structure with Local Parts for Object Detection .....	4146
<i>Yousong Zhu, Chaoyang Zhao, Jinqiao Wang, Xu Zhao, Yi Wu, and Hanqing Lu</i>	
Speaking the Same Language: Matching Machine to Human Captions by Adversarial Training .....	4155
<i>Rakshith Shetty, Marcus Rohrbach, Lisa Anne Hendricks, Mario Fritz, and Bernt Schiele</i>	
Drone-Based Object Counting by Spatially Regularized Regional Proposal Network .....	4165
<i>Meng-Ru Hsieh, Yen-Liang Lin, and Winston H. Hsu</i>	
BlitzNet: A Real-Time Deep Network for Scene Understanding .....	4174
<i>Nikita Dvornik, Konstantin Shmelkov, Julien Mairal, and Cordelia Schmid</i>	
Situation Recognition with Graph Neural Networks .....	4183
<i>Ruiyu Li, Makarand Tapaswi, Renjie Liao, Jiaya Jia, Raquel Urtasun, and Sanja Fidler</i>	
Learning Visual N-Grams from Web Data .....	4193
<i>Ang Li, Allan Jabri, Armand Joulin, and Laurens van der Maaten</i>	
Attention-Based Multimodal Fusion for Video Description .....	4203
<i>Chiori Hori, Takaaki Hori, Teng-Yok Lee, Ziming Zhang, Bret Harsham, John R. Hershey, Tim K. Marks, and Kazuhiko Sumi</i>	
Learning the Latent “Look”: Unsupervised Discovery of a Style-Coherent Embedding from Fashion Images .....	4213
<i>Wei-Lin Hsiao and Kristen Grauman</i>	
Aligned Image-Word Representations Improve Inductive Transfer Across Vision-Language Tasks .....	4223
<i>Tanmay Gupta, Kevin Shih, Saurabh Singh, and Derek Hoiem</i>	
Learning Discriminative Latent Attributes for Zero-Shot Classification .....	4233
<i>Huajie Jiang, Ruiping Wang, Shiguang Shan, Yi Yang, and Xilin Chen</i>	
PPR-FCN: Weakly Supervised Visual Relation Detection via Parallel Pairwise R-FCN .....	4243
<i>Hanwang Zhang, Zawlin Kyaw, Jinyang Yu, and Shih-Fu Chang</i>	
Higher-Order Minimum Cost Lifted Multicuts for Motion Segmentation .....	4252
<i>Margret Keuper</i>	
Deep Free-Form Deformation Network for Object-Mask Registration .....	4261
<i>Haoyang Zhang and Xuming He</i>	
Region-Based Correspondence Between 3D Shapes via Spatially Smooth Biclustering .....	4270
<i>Matteo Denitto, Simone Melzi, Manuele Bicego, Umberto Castellani, Alessandro Farinelli, Mario A. T. Figueiredo, Yanir Kleiman, and Maks Ovsjanikov</i>	
Learning Discriminative $\alpha\beta$ -Divergences for Positive Definite Matrices .....	4280
<i>A. Cherian, P. Stanitsas, M. Harandi, V. Morellas, and N. Papanikolopoulos</i>	
Consensus Convolutional Sparse Coding .....	4290
<i>Biswarup Choudhury, Robin Swanson, Felix Heide, Gordon Wetzstein, and Wolfgang Heidrich</i>	

Domain-Adaptive Deep Network Compression .....	4299
<i>Marc Masana, Joost van de Weijer, Luis Herranz, Andrew D. Bagdanov, and Jose M. Álvarez</i>	
Self-Supervised Learning of Pose Embeddings from Spatiotemporal Relations in Videos .....	4308
<i>Ömer Sümer, Tobias Dencker, and Björn Ommer</i>	
Approximate Grassmannian Intersections: Subspace-Valued Subspace Learning .....	4318
<i>Calvin Murdock and Fernando De la Torre</i>	
Side Information in Robust Principal Component Analysis: Algorithms and Applications .....	4327
<i>Niannan Xue, Yannis Panagakis, and Stefanos Zafeiriou</i>	
Summarization and Classification of Wearable Camera Streams by Learning the Distributions over Deep Features of Out-of-Sample Image Sequences .....	4336
<i>Alessandro Penna, Sadegh Mohammadi, Nebojsa Jojic, and Vittorio Murino</i>	
Unsupervised Learning from Video to Detect Foreground Objects in Single Images .....	4345
<i>Ioana Croitoru, Simion-Vlad Bogolin, and Marius Leordeanu</i>	
Supplementary Meta-Learning: Towards a Dynamic Model for Deep Neural Networks .....	4354
<i>Feihu Zhang and Benjamin W. Wah</i>	
Adversarial Inverse Graphics Networks: Learning 2D-to-3D Lifting and Image-to-Image Translation from Unpaired Supervision .....	4364
<i>Hsiao-Yu Fish Tung, Adam W. Harley, William Seto, and Katerina Fragkiadaki</i>	
Active Learning for Human Pose Estimation .....	4373
<i>Buyu Liu and Vittorio Ferrari</i>	
Interleaved Group Convolutions .....	4383
<i>Ting Zhang, Guo-Jun Qi, Bin Xiao, and Jingdong Wang</i>	
Learning-Based Cloth Material Recovery from Video .....	4393
<i>Shan Yang, Junbang Liang, and Ming C. Lin</i>	
Unsupervised Video Understanding by Reconciliation of Posture Similarities .....	4404
<i>Timo Milbich, Miguel Bautista, Ekaterina Sutter, and Björn Ommer</i>	
Action Tubelet Detector for Spatio-Temporal Action Localization .....	4415
<i>Vicky Kalogeiton, Philippe Weinzaepfel, Vittorio Ferrari, and Cordelia Schmid</i>	
AMTnet: Action-Micro-Tube Regression by End-to-end Trainable Deep Architecture .....	4424
<i>Suman Saha, Gurkirt Singh, and Fabio Cuzzolin</i>	
Constrained Convolutional Sparse Coding for Parametric Based Reconstruction of Line Drawings .....	4434
<i>Sara Shaheen, Lama Affara, and Bernard Ghanem</i>	
Neural Ctrl-F: Segmentation-Free Query-by-String Word Spotting in Handwritten Manuscript Collections .....	4443
<i>Tomas Wilkinson, Jonas Lindström, and Anders Brun</i>	

## Video Analysis

### Oral Session 6

Spatial-Aware Object Embeddings for Zero-Shot Localization and Classification of Actions .....	4453
<i>Pascal Mettes and Cees G. M. Snoek</i>	
Semantic Video CNNs Through Representation Warping .....	4463
<i>Raghudeep Gadde, Varun Jampani, and Peter V. Gehler</i>	
Video Frame Synthesis Using Deep Voxel Flow .....	4473
<i>Ziwei Liu, Raymond A. Yeh, Xiaoou Tang, Yiming Liu, and Aseem Agarwala</i>	
Detail-Revealing Deep Video Super-Resolution .....	4482
<i>Xin Tao, Hongyun Gao, Renjie Liao, Jue Wang, and Jiaya Jia</i>	
Learning Video Object Segmentation with Visual Memory .....	4491
<i>Pavel Tokmakov, Karteek Alahari, and Cordelia Schmid</i>	

## Low-Level Vision

### Oral Session 7

EnhanceNet: Single Image Super-Resolution Through Automated Texture Synthesis .....	4501
<i>Mehdi S. M. Sajjadi, Bernhard Schölkopf, and Michael Hirsch</i>	
Makeup-Go: Blind Reversion of Portrait Edit .....	4511
<i>Ying-Cong Chen, Xiaoyong Shen, and Jiaya Jia</i>	
Shadow Detection with Conditional Generative Adversarial Networks .....	4520
<i>Vu Nguyen, Tomas F. Yago Vicente, Maozheng Zhao, Minh Hoai, and Dimitris Samaras</i>	
Learning High Dynamic Range from Outdoor Panoramas .....	4529
<i>Jinsong Zhang and Jean-François Lalonde</i>	
DCTM: Discrete-Continuous Transformation Matching for Semantic Flow .....	4539
<i>Seungryong Kim, Dongbo Min, Stephen Lin, and Kwanghoon Sohn</i>	

### Spotlight Session 6

MemNet: A Persistent Memory Network for Image Restoration .....	4549
<i>Ying Tai, Jian Yang, Xiaoming Liu, and Chunyan Xu</i>	
Structure-Measure: A New Way to Evaluate Foreground Maps .....	4558
<i>Deng-Ping Fan, Ming-Ming Cheng, Yun Liu, Tao Li, and Ali Borji</i>	
Weakly- and Self-Supervised Learning for Content-Aware Deep Image Retargeting .....	4568
<i>Donghyeon Cho, Jinsun Park, Tae-Hyun Oh, Yu-Wing Tai, and In So Kweon</i>	
Practical and Efficient Multi-view Matching .....	4578
<i>Eleonora Maset, Federica Arrigoni, and Andrea Fusiello</i>	

Unrolled Memory Inner-Products: An Abstract GPU Operator for Efficient Vision-Related Computations .....	4587
<i>Yu-Sheng Lin, Wei-Chao Chen, and Shao-Yi Chien</i>	
Learning to Push the Limits of Efficient FFT-Based Image Deconvolution .....	4596
<i>Jakob Kruse, Carsten Rother, and Uwe Schmidt</i>	
Learning Spread-Out Local Feature Descriptors .....	4605
<i>Xu Zhang, Felix X. Yu, Sanjiv Kumar, and Shih-Fu Chang</i>	
Visual Odometry for Pixel Processor Arrays .....	4614
<i>Laurie Bose, Jianing Chen, Stephen J. Carey, Piotr Dudek, and Walterio Mayol-Cuevas</i>	

## Poster Session 7

Joint Estimation of Camera Pose, Depth, Deblurring, and Super-Resolution from a Blurred Image Sequence .....	4623
<i>Haesol Park and Kyoung Mu Lee</i>	
2D-Driven 3D Object Detection in RGB-D Images .....	4632
<i>Jean Lahoud and Bernard Ghanem</i>	
Ray Space Features for Plenoptic Structure-from-Motion .....	4641
<i>Yingliang Zhang, Peihong Yu, Wei Yang, Yuanxi Ma, and Jingyi Yu</i>	
Depth Estimation Using Structured Light Flow — Analysis of Projected Pattern Flow on an Object's Surface .....	4650
<i>Ryo Furukawa, Ryusuke Sagawa, and Hiroshi Kawasaki</i>	
Monocular Dense 3D Reconstruction of a Complex Dynamic Scene from Two Perspective Frames .....	4659
<i>Suryansh Kumar, Yuchao Dai, and Hongdong Li</i>	
Optimal Transformation Estimation with Semantic Cues .....	4668
<i>Luc Van Gool, Danda Pani Paudel, and Adlane Haded</i>	
Dynamics Enhanced Multi-camera Motion Segmentation from Unsynchronized Videos .....	4678
<i>Xikang Zhang, Bengisu Ozbay, Mario Sznaiier, and Octavia Camps</i>	
Taking the Scenic Route to 3D: Optimising Reconstruction from Moving Cameras .....	4687
<i>Oscar Mendez, Simon Hadfield, Nicolas Pugeault, and Richard Bowden</i>	
FLaME: Fast Lightweight Mesh Estimation Using Variational Smoothing on Delaunay Graphs .....	4696
<i>W. Nicholas Greene and Nicholas Roy</i>	
Efficient Algorithms for Moral Lineage Tracing .....	4705
<i>Markus Rempfler, Jan-Hendrik Lange, Florian Jug, Corinna Blasse, Eugene W. Myers, Bjoern H. Menze, and Bjoern Andres</i>	
From RGB to Spectrum for Natural Scenes via Manifold-Based Mapping .....	4715
<i>Yan Jia, Yinqiang Zheng, Lin Gu, Art Subpa-Asa, Antony Lam, Yoichi Sato, and Imari Sato</i>	
DeepFuse: A Deep Unsupervised Approach for Exposure Fusion with Extreme Exposure Image Pairs .....	4724
<i>K. Ram Prabhakar, V Sai Srikar, and R. Venkatesh Babu</i>	



Learning Dense Facial Correspondences in Unconstrained Images .....	4733
<i>Ronald Yu, Shunsuke Saito, Haoxiang Li, Duygu Ceylan, and Hao Li</i>	
Jointly Attentive Spatial-Temporal Pooling Networks for Video-Based Person Re-identification .....	4743
<i>Shuangjie Xu, Yu Cheng, Kang Gu, Yang Yang, Shiyu Chang, and Pan Zhou</i>	
Automatic Content-Aware Projection for 360° Videos .....	4753
<i>Yeong Won Kim, Chang-Ryeol Lee, Dae-Yong Cho, Yong Hoon Kwon, Hyeok-Jae Choi, and Kuk-Jin Yoon</i>	
Blur-Invariant Deep Learning for Blind-Deblurring .....	4762
<i>T M Nimisha, Akash Kumar Singh, and A N Rajagopalan</i>	
Non-linear Convolution Filters for CNN-Based Learning .....	4771
<i>Georgios Zoumpourlis, Alexandros Doumanoglou, Nicholas Vretos, and Petros Daras</i>	
AOD-Net: All-in-One Dehazing Network .....	4780
<i>Boyi Li, Xiulian Peng, Zhangyang Wang, Jizheng Xu, and Dan Feng</i>	
Simultaneous Detection and Removal of High Altitude Clouds from an Image .....	4789
<i>Tushar Sandhan and Jin Young Choi</i>	
Understanding Low- and High-Level Contributions to Fixation Prediction .....	4799
<i>Matthias Kümmerer, Thomas S.A. Wallis, Leon A. Gatys, and Matthias Bethge</i>	
Image Super-Resolution Using Dense Skip Connections .....	4809
<i>Tong Tong, Gen Li, Xiejie Liu, and Qinquan Gao</i>	
Convergence Analysis of MAP Based Blur Kernel Estimation .....	4818
<i>Sunghyun Cho and Seungyong Lee</i>	
Blob Reconstruction Using Unilateral Second Order Gaussian Kernels with Application to High-ISO Long-Exposure Image Denoising .....	4827
<i>Gang Wang, Carlos Lopez-Molina, and Bernard De Baets</i>	
Deep Generative Adversarial Compression Artifact Removal .....	4836
<i>Leonardo Galteri, Lorenzo Seidenari, Marco Bertini, and Alberto Del Bimbo</i>	
Online Multi-object Tracking Using CNN-Based Single Object Tracker with Spatial-Temporal Attention Mechanism .....	4846
<i>Qi Chu, Wanli Ouyang, Hongsheng Li, Xiaogang Wang, Bin Liu, and Nenghai Yu</i>	
Mutual Enhancement for Detection of Multiple Logos in Sports Videos .....	4856
<i>Yuan Liao, Xiaoqing Lu, Chengcui Zhang, Yongtao Wang, and Zhi Tang</i>	
Referring Expression Generation and Comprehension via Attributes .....	4866
<i>Jingyu Liu, Liang Wang, and Ming-Hsuan Yang</i>	
RoomNet: End-to-End Room Layout Estimation .....	4875
<i>Chen-Yu Lee, Vijay Badrinarayanan, Tomasz Malisiewicz, and Andrew Rabinovich</i>	
SSH: Single Stage Headless Face Detector .....	4885
<i>Mahyar Najibi, Pouya Samangouei, Rama Chellappa, and Larry S. Davis</i>	
AnnArbor: Approximate Nearest Neighbors Using Arborescence Coding .....	4895
<i>Artem Babenko Yandex and Victor Lempitsky</i>	

Boosting Image Captioning with Attributes .....	4904
<i>Ting Yao, Yingwei Pan, Yehao Li, Zhaofan Qiu, and Tao Mei</i>	
Learning to Estimate 3D Hand Pose from Single RGB Images .....	4913
<i>Christian Zimmermann and Thomas Brox</i>	
Locally-Transferred Fisher Vectors for Texture Classification .....	4922
<i>Yang Song, Fan Zhang, Qing Li, Heng Huang, Lauren J. O'Donnell, and Weidong Cai</i>	
Object-Level Proposals .....	4931
<i>Jianxiang Ma, Anlong Ming, Zilong Huang, Xinggong Wang, and Yu Zhou</i>	
Extreme Clicking for Efficient Object Annotation .....	4940
<i>Dim P. Papadopoulos, Jasper R. R. Uijlings, Frank Keller, and Vittorio Ferrari</i>	
WordSup: Exploiting Word Annotations for Character Based Text Detection .....	4950
<i>Han Hu, Chengquan Zhang, Yuxuan Luo, Yuzhuo Wang, Junyu Han, and Errui Ding</i>	
Illuminating Pedestrians via Simultaneous Detection and Segmentation .....	4960
<i>Garrick Brazil, Xi Yin, and Xiaoming Liu</i>	
Generalized Orderless Pooling Performs Implicit Salient Matching .....	4970
<i>Marcel Simon, Yang Gao, Trevor Darrell, Joachim Denzler, and Erik Rodner</i>	
Exploiting Spatial Structure for Localizing Manipulated Image Regions .....	4980
<i>Jawadul H. Bappy, Amit K. Roy-Chowdhury, Jason Bunk, Lakshmanan Nataraj, and B.S. Manjunath</i>	
RDFNet: RGB-D Multi-level Residual Feature Fusion for Indoor Semantic Segmentation .....	4990
<i>Seungyong Lee, Seong-Jin Park, and Ki-Sang Hong</i>	
Self-Organized Text Detection with Minimal Post-processing via Border Learning .....	5010
<i>Yue Wu and Prem Natarajan</i>	
Sparse Exact PGA on Riemannian Manifolds .....	5020
<i>Monami Banerjee, Rudrasis Chakraborty, and Baba C. Vemuri</i>	
Tensor RPCA by Bayesian CP Factorization with Complex Noise .....	5029
<i>Qiong Luo, Zhi Han, Xi'ai Chen, Yao Wang, Deyu Meng, Dong Liang, and Yandong Tang</i>	
Multimodal Gaussian Process Latent Variable Models with Harmonization .....	5039
<i>Guoli Song, Shuhui Wang, Qingming Huang, and Qi Tian</i>	
Segmentation-Aware Convolutional Networks Using Local Attention Masks .....	5048
<i>Adam W. Harley, Konstantinos G. Derpanis, and Iasonas Kokkinos</i>	
Rotation Equivariant Vector Field Networks .....	5058
<i>Diego Marcos, Michele Volpi, Nikos Komodakis, and Devis Tuia</i>	
ThiNet: A Filter Level Pruning Method for Deep Neural Network Compression .....	5068
<i>Jian-Hao Luo, Jianxin Wu, and Weiyao Lin</i>	
AutoDIAL: Automatic Domain Alignment Layers .....	5077
<i>Fabio Maria Cariucci, Lorenzo Porzi, Barbara Caputo, Elisa Ricci, and Samuel Rota Bulò</i>	
Focusing Attention: Towards Accurate Text Recognition in Natural Images .....	5086
<i>Zhanzhan Cheng, Fan Bai, Yunlu Xu, Gang Zheng, Shiliang Pu, and Shuigeng Zhou</i>	

Unsupervised Object Segmentation in Video by Efficient Selection of Highly Probable Positive Features .....	5095
<i>Emanuela Haller and Marius Leordeanu</i>	
Nonparametric Variational Auto-Encoders for Hierarchical Representation Learning .....	5104
<i>Prasoon Goyal, Zhiting Hu, Xiaodan Liang, Chenyu Wang, Eric P. Xing, and Carnegie Mellon</i>	
Dense and Low-Rank Gaussian CRFs Using Deep Embeddings .....	5113
<i>Siddhartha Chandra, Nicolas Usunier, and Iasonas Kokkinos</i>	
A Multimodal Deep Regression Bayesian Network for Affective Video Content Analyses .....	5123
<i>Quan Gan, Shangfei Wang, Longfei Hao, and Qiang Ji</i>	
Moving Object Detection in Time-Lapse or Motion Trigger Image Sequences Using Low-Rank and Invariant Sparse Decomposition .....	5133
<i>Moein Shakeri and Hong Zhang</i>	
A Multilayer-Based Framework for Online Background Subtraction with Freely Moving Cameras .....	5142
<i>Yizhe Zhu and Ahmed Elgammal</i>	
Dynamic Label Graph Matching for Unsupervised Video Re-identification .....	5152
<i>Mang Ye, Andy J. Ma, Liang Zheng, Jiawei Li, and Pong C. Yuen</i>	
Spatiotemporal Modeling for Crowd Counting in Videos .....	5161
<i>Feng Xiong, Xingjian Shi, and Dit-Yan Yeung</i>	
Personalized Cinemagraphs Using Semantic Understanding and Collaborative Learning .....	5170
<i>Tae-Hyun Oh, Kyungdon Joo, Neel Joshi, Baoyuan Wang, In So Kweon, and Sing Bing Kang</i>	
What is Around the Camera? .....	5180
<i>Stamatios Georgoulis, Konstantinos Rematas, Tobias Ritschel, Mario Fritz, Tinne Tuytelaars, and Luc Van Gool</i>	

### **Recognition 3**

#### **Oral Session 8**

Weakly-Supervised Learning of Visual Relations .....	5189
<i>Julia Peyre, Ivan Laptev, Cordelia Schmid, and Josef Sivic</i>	
BIER — Boosting Independent Embeddings Robustly .....	5199
<i>Michael Opitz, Georg Waltner, Horst Possegger, and Horst Bischof</i>	
3D Graph Neural Networks for RGBD Semantic Segmentation .....	5209
<i>Xiaojuan Qi, Renjie Liao, Jiaya Jia, Sanja Fidler, and Raquel Urtasun</i>	
Learning Multi-attention Convolutional Neural Network for Fine-Grained Image Recognition .....	5219
<i>Heliang Zheng, Jianlong Fu, Tao Mei, and Jiebo Luo</i>	
Learning 3D Object Categories by Looking Around Them .....	5228
<i>David Novotny, Diane Larlus, and Andrea Vedaldi</i>	

## Spotlight Session 7

Quantitative Evaluation of Confidence Measures in a Machine Learning World .....	5238
<i>Matteo Poggi, Fabio Tosi, and Stefano Mattoccia</i>	
Towards End-to-End Text Spotting with Convolutional Recurrent Neural Networks .....	5248
<i>Hui Li, Peng Wang, and Chunhua Shen</i>	
DeepSetNet: Predicting Sets with Deep Neural Networks .....	5257
<i>S. Hamid Rezatofighi, Vijay Kumar B. G, Anton Milan, Ehsan Abbasnejad, Anthony Dick, and Ian Reid</i>	
Learning from Video and Text via Large-Scale Discriminative Clustering .....	5267
<i>Antoine Miech, Jean-Baptiste Alayrac, Piotr Bojanowski, Ivan Laptev, and Josef Sivic</i>	
TALL: Temporal Activity Localization via Language Query .....	5277
<i>Jiyang Gao, Chen Sun, Zhenheng Yang, and Ram Nevatia</i>	
End-to-End Face Detection and Cast Grouping in Movies Using Erdős-Rényi Clustering .....	5286
<i>SouYoung Jin, Hang Su, Chris Stauffer, and Erik Learned-Miller</i>	
Active Decision Boundary Annotation with Deep Generative Models .....	5296
<i>Miriam Huijser and Jan C. van Gemert</i>	
Convolutional Dictionary Learning via Local Processing .....	5306
<i>Vardan Papyan, Yaniv Romano, Michael Elad, and Jeremias Sulam</i>	

## Poster Session 8

Editable Parametric Dense Foliage from 3D Capture .....	5315
<i>Paul Beardsley and Gaurav Chaurasia</i>	
Refractive Structure-from-Motion Through a Flat Refractive Interface .....	5325
<i>François Chadebecq, Francisco Vasconcelos, George Dwyer, René Lacher, Sébastien Ourselin, Tom Vercauteren, and Danail Stoyanov</i>	
Submodular Trajectory Optimization for Aerial 3D Scanning .....	5334
<i>Mike Roberts, Shital Shah, Debadeepta Dey, Anh Truong, Sudipta Sinha, Ashish Kapoor, Pat Hanrahan, and Neel Joshi</i>	
Camera Calibration by Global Constraints on the Motion of Silhouettes .....	5344
<i>Gil Ben-Artzi</i>	
Deltille Grids for Geometric Camera Calibration .....	5354
<i>Hyowon Ha, Michal Perdoch, Hatem Alismail, In So Kweon, and Yaser Sheikh</i>	
A Lightweight Single-Camera Polarization Compass with Covariance Estimation .....	5363
<i>Wolfgang Stürzl</i>	
Reflectance Capture Using Univariate Sampling of BRDFs .....	5372
<i>Zhuo Hui, Kalyan Sunkavalli, Joon-Young Lee, Sunil Hadap, Jian Wang, and Aswin C. Sankaranarayanan</i>	
Estimating Defocus Blur via Rank of Local Patches .....	5381
<i>Guodong Xu, Yuhui Quan, and Hui Ji</i>	

RGB-Infrared Cross-Modality Person Re-identification .....	5390
<i>Ancong Wu, Wei-Shi Zheng, Hong-Xing Yu, Shaogang Gong, and Jianhuang Lai</i>	
Intrinsic 3D Dynamic Surface Tracking based on Dynamic Ricci Flow and Teichmüller Map .....	5400
<i>Xiaokang Yu, Na Lei, Yalin Wang, and Xianfeng Gu</i>	
Multi-scale Deep Learning Architectures for Person Re-identification .....	5409
<i>Xuelin Qian, Yanwei Fu, Yu-Gang Jiang, Tao Xiang, and Xiangyang Xue</i>	
Range Loss for Deep Face Recognition with Long-Tailed Training Data .....	5419
<i>Xiao Zhang, Zhiyuan Fang, Yandong Wen, Zhifeng Li, and Yu Qiao</i>	
Face Sketch Matching via Coupled Deep Transform Learning .....	5429
<i>Shruti Nagpal, Maneet Singh, Richa Singh, Mayank Vatsa, Afzel Noore, and Angshul Majumdar</i>	
Realistic Dynamic Facial Textures from a Single Image Using GANs .....	5439
<i>Kyle Olszewski, Zimo Li, Chao Yang, Yi Zhou, Ronald Yu, Zeng Huang, Sitao Xiang, Shunsuke Saito, Pushmeet Kohli, and Hao Li</i>	
Pixel Recursive Super Resolution .....	5449
<i>Ryan Dahl, Mohammad Norouzi, and Jonathon Shlens</i>	
Recurrent Color Constancy .....	5459
<i>Yanlin Qian, Ke Chen, Jarno Nikkanen, Joni-Kristian Kämäräinen, and Jiri Matas</i>	
Saliency Pattern Detection by Ranking Structured Trees .....	5468
<i>Lei Zhu, Haibin Ling, Jin Wu, Huiping Deng, and Jin Liu</i>	
Monocular Video-Based Trailer Coupler Detection Using Multiplexer Convolutional Neural Network .....	5478
<i>Yousef Atoum, Joseph Roth, Michael Bliss, Wende Zhang, and Xiaoming Liu</i>	
Parallel Tracking and Verifying: A Framework for Real-Time and High Accuracy Visual Tracking .....	5487
<i>Heng Fan and Haibin Ling</i>	
Non-rigid Object Tracking via Deformable Patches Using Shape-Preserved KCF and Level Sets .....	5496
<i>Xin Sun, Ngai-Man Cheung, Hongxun Yao, and Yiluan Guo</i>	
A Discriminative View of MRF Pre-processing Algorithms .....	5505
<i>Chen Wang, Charles Herrmann, and Ramin Zabih</i>	
Offline Handwritten Signature Modeling and Verification Based on Archetypal Analysis .....	5515
<i>Elias N. Zois, Ilias Theodorakopoulos, and George Economou</i>	
Long Short-Term Memory Kalman Filters: Recurrent Neural Estimators for Pose Regularization .....	5525
<i>Huseyin Coskun, Felix Achilles, Robert DiPietro, Nassir Navab, and Federico Tombari</i>	
Learning Spatio-Temporal Representation with Pseudo-3D Residual Networks .....	5534
<i>Zhaofan Qiu, Ting Yao, and Tao Mei</i>	
Deeper, Broader and Artier Domain Generalization .....	5543
<i>Da Li, Yongxin Yang, Yi-Zhe Song, and Timothy M. Hospedales</i>	

Deep Spatial-Semantic Attention for Fine-Grained Sketch-Based Image Retrieval .....	5552
<i>Jifei Song, Qian Yu, Yi-Zhe Song, Tao Xiang, and Timothy M. Hospedales</i>	
Soft-NMS — Improving Object Detection with One Line of Code .....	5562
<i>Navaneeth Bodla, Bharat Singh, Rama Chellappa, and Larry S. Davis</i>	
Semantic Jitter: Dense Supervision for Visual Comparisons via Synthetic Images .....	5571
<i>Aron Yu and Kristen Grauman</i>	
Video Scene Parsing with Predictive Feature Learning .....	5581
<i>Xiaojie Jin, Xin Li, Huaxin Xiao, Xiaohui Shen, Zhe Lin, Jimei Yang, Yunpeng Chen, Jian Dong, Luoqi Liu, Zequn Jie, Jiashi Feng, and Shuicheng Yan</i>	
Understanding and Mapping Natural Beauty .....	5590
<i>Scott Workman, Richard Souvenir, and Nathan Jacobs</i>	
Human Pose Estimation Using Global and Local Normalization .....	5600
<i>Ke Sun, Cuiling Lan, Junliang Xing, Wenjun Zeng, Dong Liu, and Jingdong Wang</i>	
HashNet: Deep Learning to Hash by Continuation .....	5609
<i>Zhangjie Cao, Mingsheng Long, Jianmin Wang, and Philip S. Yu</i>	
Scaling the Scattering Transform: Deep Hybrid Networks .....	5619
<i>Edouard Oyallon, Eugene Belilovsky, and Sergey Zagoruyko</i>	
Flip-Invariant Motion Representation .....	5629
<i>Takumi Kobayashi</i>	
Scene Categorization with Spectral Features .....	5639
<i>Salman H. Khan, Munawar Hayat, and Fatih Porikli</i>	
Image2song: Song Retrieval via Bridging Image Content and Lyric Words .....	5650
<i>Xuelong Li, Di Hu, and Xiaoqiang Lu</i>	
Deep Functional Maps: Structured Prediction for Dense Shape Correspondence .....	5660
<i>Or Litany, Tal Remez, Emanuele Rodolà, Alex Bronstein, and Michael Bronstein</i>	
Training Deep Networks to be Spatially Sensitive .....	5669
<i>Nicholas Kolkin, Gregory Shakhnarovich, and Eli Shechtman</i>	
3DCNN-DQN-RNN: A Deep Reinforcement Learning Framework for Semantic Parsing of Large-Scale 3D Point Clouds .....	5679
<i>Fangyu Liu, Shuaipeng Li, Liqiang Zhang, Chenghu Zhou, Rongtian Ye, Yuebin Wang, and Jiwen Lu</i>	
Semi Supervised Semantic Segmentation Using Generative Adversarial Network .....	5689
<i>Nasim Souly, Concetto Spampinato, and Mubarak Shah</i>	
Efficient Low Rank Tensor Ring Completion .....	5698
<i>Wenqi Wang, Vaneet Aggarwal, and Shuchin Aeron</i>	
Semantic Image Synthesis via Adversarial Learning .....	5707
<i>Hao Dong, Simiao Yu, Chao Wu, and Yike Guo</i>	
Unified Deep Supervised Domain Adaptation and Generalization .....	5716
<i>Saeid Motiian, Marco Piccirilli, Donald A. Adjeroh, and Gianfranco Doretto</i>	
Temporal Context Network for Activity Localization in Videos .....	5727
<i>Xiyang Dai, Bharat Singh, Guyue Zhang, Larry S. Davis, and Yan Qiu Chen</i>	

Interpretable Transformations with Encoder-Decoder Networks .....	5737
<i>Daniel E. Worrall, Stephan J. Garbin, Daniyar Turmukhambetov, and Gabriel J. Brostow</i>	
Deep Clustering via Joint Convolutional Autoencoder Embedding and Relative Entropy Minimization .....	5747
<i>Kamran Ghasedi Dizaji, Amirhossein Herandi, Cheng Deng, Weidong Cai, and Heng Huang</i>	
Deep Scene Image Classification with the MFAFVNet .....	5757
<i>Yunsheng Li, Mandar Dixit, and Nuno Vasconcelos</i>	
Learning Bag-of-Features Pooling for Deep Convolutional Neural Networks .....	5766
<i>Nikolaos Passalis and Anastasios Tefas</i>	
Adversarial Examples Detection in Deep Networks with Convolutional Filter Statistics .....	5775
<i>Xin Li and Fuxin Li</i>	
Joint Prediction of Activity Labels and Starting Times in Untrimmed Videos .....	5784
<i>Tahmida Mahmud, Mahmudul Hasan, and Amit K. Roy-Chowdhury</i>	
R-C3D: Region Convolutional 3D Network for Temporal Activity Detection .....	5794
<i>Huijuan Xu, Abir Das, and Kate Saenko</i>	
Localizing Moments in Video with Natural Language .....	5804
<i>Lisa Anne Hendricks, Oliver Wang, Eli Shechtman, Josef Sivic, Trevor Darrell, and Bryan Russell</i>	
TORNADO: A Spatio-Temporal Convolutional Regression Network for Video Action Proposal .....	5814
<i>Hongyuan Zhu, Romain Vial, and Shijian Lu</i>	
Tube Convolutional Neural Network (T-CNN) for Action Detection in Videos .....	5823
<i>Rui Hou, Chen Chen, and Mubarak Shah</i>	
Learning Action Recognition Model from Depth and Skeleton Videos .....	5833
<i>Hossein Rahmani and Mohammed Bennamoun</i>	
The “Something Something” Video Database for Learning and Evaluating Visual Common Sense .....	5843
<i>Raghav Goyal, Samira Ebrahimi Kahou, Vincent Michalski, Joanna Materzynska, Susanne Westphal, Heuna Kim, Valentin Haenel, Ingo Freund, Peter Yianilos, Moritz Mueller-Freitag, Florian Hoppe, Christian Thureau, Ingo Bax, and Roland Memisevic</i>	
GPLAC: Generalizing Vision-Based Robotic Skills Using Weakly Labeled Images .....	5852
<i>Avi Singh, Larry Yang, and Sergey Levine</i>	
Semi-Global Weighted Least Squares in Image Filtering .....	5862
<i>Wei Liu, Xiaogang Chen, Chuanhua Shen, Zhi Liu, and Jie Yang</i>	
Scale Recovery for Monocular Visual Odometry Using Depth Estimated with Deep Convolutional Neural Fields .....	5871
<i>Xiaochuan Yin, Xiangwei Wang, Xiaoguo Du, and Qijun Chen</i>	

## Machine Learning

### Oral Session 9

Deep Adaptive Image Clustering .....	5880
<i>Jianlong Chang, Lingfeng Wang, Gaofeng Meng, Shiming Xiang, and Chunhong Pan</i>	
One Network to Solve Them All — Solving Linear Inverse Problems Using Deep Projection Models .....	5889
<i>J. H. Rick Chang, Chun-Liang Li, Barnabás Póczos, and B. V. K. Vijaya Kumar</i>	
Representation Learning by Learning to Count .....	5899
<i>Mehdi Noroozi, Hamed Pirsiavash, and Paolo Favaro</i>	
StackGAN: Text to Photo-Realistic Image Synthesis with Stacked Generative Adversarial Networks .....	5908
<i>Han Zhang, Tao Xu, and Hongsheng Li</i>	
Unsupervised Domain Adaptation for Face Recognition in Unlabeled Videos .....	5917
<i>Kihyuk Sohn, Sifei Liu, Guangyu Zhong, Xiang Yu, Ming-Hsuan Yang, and Manmohan Chandraker</i>	

### Additional Paper

The Mapillary Vistas Dataset for Semantic Understanding of Street Scenes .....	5000
<i>Gerhard Neuhold, Tobias Ollmann, Samuel Rota Bulo and Peter Kotschieder</i>	

### Author Index