

2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017)

**Honolulu, Hawaii, USA
21-26 July 2017**

Pages 1-751



**IEEE Catalog Number: CFP17003-POD
ISBN: 978-1-5386-0458-8**

**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:	CFP17003-POD
ISBN (Print-On-Demand):	978-1-5386-0458-8
ISBN (Online):	978-1-5386-0457-1
ISSN:	1063-6919

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
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings
.com

2017 IEEE Conference on Computer Vision and Pattern Recognition

CVPR 2017

Table of Contents

Message from the General and Program Chairs.....	liv
Committees.....	lvi
Reviewers.....	lvii

Oral and Spotlight 1-1A

S1-1A: Machine Learning 1

Exclusivity-Consistency Regularized Multi-view Subspace Clustering	1
<i>Xiaobo Wang, Xiaojie Guo, Zhen Lei, Changqing Zhang, and Stan Z. Li</i>	
Borrowing Treasures from the Wealthy: Deep Transfer Learning through Selective Joint Fine-Tuning	10
<i>Weifeng Ge and Yizhou Yu</i>	
The More You Know: Using Knowledge Graphs for Image Classification	20
<i>Kenneth Marino, Ruslan Salakhutdinov, and Abhinav Gupta</i>	
Dynamic Edge-Conditioned Filters in Convolutional Neural Networks on Graphs	29
<i>Martin Simonovsky and Nikos Komodakis</i>	
Convolutional Neural Network Architecture for Geometric Matching	39
<i>Ignacio Rocco, Relja Arandjelovic, and Josef Sivic</i>	
Deep Affordance-Grounded Sensorimotor Object Recognition	49
<i>Spyridon Thermos, Georgios Th. Papadopoulos, Petros Daras, and Gerasimos Potamianos</i>	
Discovering Causal Signals in Images	58
<i>David Lopez-Paz, Robert Nishihara, Soumith Chintala, Bernhard Schölkopf, and Léon Bottou</i>	
On Compressing Deep Models by Low Rank and Sparse Decomposition	67
<i>Xiyu Yu, Tongliang Liu, Xinchao Wang, and Dacheng Tao</i>	

O1-1A

PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation	77
<i>R. Qi Charles, Hao Su, Mo Kaichun, and Leonidas J. Guibas</i>	
Universal Adversarial Perturbations	86
<i>Seyed-Mohsen Moosavi-Dezfooli, Alhussein Fawzi, Omar Fawzi, and Pascal Frossard</i>	
Unsupervised Pixel-Level Domain Adaptation with Generative Adversarial Networks	95
<i>Konstantinos Bousmalis, Nathan Silberman, David Dohan, Dumitru Erhan, and Dilip Krishnan</i>	
Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network	105
<i>Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, and Wenzhe Shi</i>	

Oral and Spotlight 1-1B

S1-1B: 3D Vision 1

Global Hypothesis Generation for 6D Object Pose Estimation	115
<i>Frank Michel, Alexander Kirillov, Eric Brachmann, Alexander Krull, Stefan Gumhold, Bogdan Savchynskyy, and Carsten Rother</i>	
A Practical Method for Fully Automatic Intrinsic Camera Calibration Using Directionally Encoded Light	125
<i>Mahdi Abbaspour Tehrani, Thabo Beeler, and Anselm Grundhöfer</i>	
CATS: A Color and Thermal Stereo Benchmark	134
<i>Wayne Treible, Philip Saponaro, Scott Sorensen, Abhishek Kolagunda, Michael O’Neal, Brian Phelan, Kelly Sherbondy, and Chandra Kambhamettu</i>	
Elastic Shape-from-Template with Spatially Sparse Deforming Forces	143
<i>Abed Malti and Cédric Herzet</i>	
Distinguishing the Indistinguishable: Exploring Structural Ambiguities via Geodesic Context	152
<i>Qingan Yan, Long Yang, Ling Zhang, and Chunxia Xiao</i>	
Multi-scale Continuous CRFs as Sequential Deep Networks for Monocular Depth Estimation	161
<i>Dan Xu, Elisa Ricci, Wanli Ouyang, Xiaogang Wang, and Nicu Sebe</i>	
Dynamic Time-of-Flight	170
<i>Michael Schober, Amit Adam, Omer Yair, Shai Mazor, and Sebastian Nowozin</i>	
Training Object Class Detectors with Click Supervision	180
<i>Dim P. Papadopoulos, Jasper R. R. Uijlings, Frank Keller, and Vittorio Ferrari</i>	

O1-1B

Semantic Scene Completion from a Single Depth Image	190
<i>Shuran Song, Fisher Yu, Andy Zeng, Angel X. Chang, Manolis Savva, and Thomas Funkhouser</i>	
3DMatch: Learning Local Geometric Descriptors from RGB-D Reconstructions	199
<i>Andy Zeng, Shuran Song, Matthias Nießner, Matthew Fisher, Jianxiong Xiao, and Thomas Funkhouser</i>	
Multi-view Supervision for Single-View Reconstruction via Differentiable Ray Consistency	209
<i>Shubham Tulsiani, Tinghui Zhou, Alexei A. Efros, and Jitendra Malik</i>	
On-the-Fly Adaptation of Regression Forests for Online Camera Relocalisation	218
<i>Tommaso Cavallari, Stuart Golodetz, Nicholas A. Lord, Julien Valentin, Luigi Di Stefano, and Philip H. S. Torr</i>	

Oral and Spotlight 1-1C

S1-1C: Low- & Mid-Level Vision

Designing Effective Inter-Pixel Information Flow for Natural Image Matting	228
<i>Yagiz Aksoy, Tunç Ozan Aydin, and Marc Pollefeys</i>	
Deep Video Deblurring for Hand-Held Cameras	237
<i>Shuo Chen Su, Mauricio Delbracio, Jue Wang, Guillermo Sapiro, Wolfgang Heidrich, and Oliver Wang</i>	
Instance-Level Salient Object Segmentation	247
<i>Guanbin Li, Yuan Xie, Liang Lin, and Yizhou Yu</i>	
Deep Multi-scale Convolutional Neural Network for Dynamic Scene Deblurring	257
<i>Seungjun Nah, Tae Hyun Kim, and Kyoung Mu Lee</i>	
Diversified Texture Synthesis with Feed-Forward Networks	266
<i>Yijun Li, Chen Fang, Jimei Yang, Zhaowen Wang, Xin Lu, and Ming-Hsuan Yang</i>	
Radiometric Calibration for Internet Photo Collections	275
<i>Zhipeng Mo, Boxin Shi, Sai-Kit Yeung, and Yasuyuki Matsushita</i>	
Deeply Aggregated Alternating Minimization for Image Restoration	284
<i>Youngjung Kim, Hyungjoo Jung, Dongbo Min, and Kwanghoon Sohn</i>	
End-to-End Instance Segmentation with Recurrent Attention	293
<i>Mengye Ren and Richard S. Zemel</i>	

O1-1C

SRN: Side-Output Residual Network for Object Symmetry Detection in the Wild	302
<i>Wei Ke, Jie Chen, Jianbin Jiao, Guoying Zhao, and Qixiang Ye</i>	
Deep Image Matting	311
<i>Ning Xu, Brian Price, Scott Cohen, and Thomas Huang</i>	

Witness and Color from a Single Multispectral Image	321
<i>Mihoko Shimano, Hiroki Okawa, Yuta Asano, Ryoma Bise, Ko Nishino, and Imari Sato</i>	
FC ⁴ : Fully Convolutional Color Constancy with Confidence-Weighted Pooling	330
<i>Yuanming Hu, Baoyuan Wang, and Stephen Lin</i>	
Poster Session P1-1	
Face Normals "In-the-Wild" Using Fully Convolutional Networks	340
<i>George Trigeorgis, Patrick Snape, Iasonas Kokkinos, and Stefanos Zafeiriou</i>	
A Non-convex Variational Approach to Photometric Stereo under Inaccurate Lighting	350
<i>Yvain Quéau, Tao Wu, François Lauze, Jean-Denis Durou, and Daniel Cremers</i>	
A Linear Extrinsic Calibration of Kaleidoscopic Imaging System from Single 3D Point	360
<i>Kosuke Takahashi, Akihiro Miyata, Shohei Nobuhara, and Takashi Matsuyama</i>	
Polarimetric Multi-view Stereo	369
<i>Zhaopeng Cui, Jinwei Gu, Boxin Shi, Ping Tan, and Jan Kautz</i>	
An Exact Penalty Method for Locally Convergent Maximum Consensus	379
<i>Huu Le, Tat-Jun Chin, and David Suter</i>	
Deep Supervision with Shape Concepts for Occlusion-Aware 3D Object Parsing	388
<i>Chi Li, M. Zeeshan Zia, Quoc-Huy Tran, Xiang Yu, Gregory D. Hager, and Manmohan Chandraker</i>	
Amodal Detection of 3D Objects: Inferring 3D Bounding Boxes from 2D Ones in RGB-Depth Images	398
<i>Zhuo Deng and Longin Jan Latecki</i>	
Transition Forests: Learning Discriminative Temporal Transitions for Action Recognition and Detection	407
<i>Guillermo Garcia-Hernando and Tae-Kyun Kim</i>	
Scene Flow to Action Map: A New Representation for RGB-D Based Action Recognition with Convolutional Neural Networks	416
<i>Pichao Wang, Wanqing Li, Zhimin Gao, Yuyao Zhang, Chang Tang, and Philip Ogunbona</i>	
Detecting Masked Faces in the Wild with LLE-CNNs	426
<i>Shiming Ge, Jia Li, Qiting Ye, and Zhao Luo</i>	
A Domain Based Approach to Social Relation Recognition	435
<i>Qianru Sun, Bernt Schiele, and Mario Fritz</i>	
Spatio-Temporal Naive-Bayes Nearest-Neighbor (ST-NBNN) for Skeleton-Based Action Recognition	445
<i>Junwu Weng, Chaoqun Weng, and Junsong Yuan</i>	
Personalizing Gesture Recognition Using Hierarchical Bayesian Neural Networks	455
<i>Ajjen Joshi, Soumya Ghosh, Margrit Betke, Stan Sclaroff, and Hanspeter Pfister</i>	
Real-Time 3D Model Tracking in Color and Depth on a Single CPU Core	465
<i>Wadim Kehl, Federico Tombari, Slobodan Ilic, and Nassir Navab</i>	

Multi-scale FCN with Cascaded Instance Aware Segmentation for Arbitrary Oriented Word Spotting in the Wild	474
<i>Dafang He, Xiao Yang, Chen Liang, Zihan Zhou, Alexander G. Ororbia, Daniel Kifer, and C. Lee Giles</i>	
Viraliency: Pooling Local Virality	484
<i>Xavier Alameda-Pineda, Andrea Pilzer, Dan Xu, Nicu Sebe, and Elisa Ricci</i>	
A Non-local Low-Rank Framework for Ultrasound Speckle Reduction	493
<i>Lei Zhu, Chi-Wing Fu, Michael S. Brown, and Pheng-Ann Heng</i>	
Video Acceleration Magnification	502
<i>Yichao Zhang, Silvia L. Pinteá, and Jan C. van Gemert</i>	
Superpixel-Based Tracking-by-Segmentation Using Markov Chains	511
<i>Donghun Yeo, Jeany Son, Bohyung Han, and Joon Hee Han</i>	
BranchOut: Regularization for Online Ensemble Tracking with Convolutional Neural Networks	521
<i>Bohyung Han, Jack Sim, and Hartwig Adam</i>	
Learning Motion Patterns in Videos	531
<i>Pavel Tokmakov, Karteek Alahari, and Cordelia Schmid</i>	
Deep Level Sets for Salient Object Detection	540
<i>Ping Hu, Bing Shuai, Jun Liu, and Gang Wang</i>	
Binary Constraint Preserving Graph Matching	550
<i>Bo Jiang, Jin Tang, Chris Ding, and Bin Luo</i>	
From Local to Global: Edge Profiles to Camera Motion in Blurred Images	558
<i>Subeesh Vasu and A. N. Rajagopalan</i>	
What is the Space of Attenuation Coefficients in Underwater Computer Vision?	568
<i>Derya Akkaynak, Tali Treibitz, Tom Shlesinger, Yossi Loya, Raz Tamir, and David Iluz</i>	
Robust Energy Minimization for BRDF-Invariant Shape from Light Fields	578
<i>Zhengqin Li, Zexiang Xu, Ravi Ramamoorthi, and Manmohan Chandraker</i>	
Boundary-Aware Instance Segmentation	587
<i>Zeeshan Hayder, Xuming He, and Mathieu Salzmann</i>	
Spatially-Varying Blur Detection Based on Multiscale Fused and Sorted Transform Coefficients of Gradient Magnitudes	596
<i>S. Alireza Golestaneh and Lina J. Karam</i>	
Model-Based Iterative Restoration for Binary Document Image Compression with Dictionary Learning	606
<i>Yandong Guo, Cheng Lu, Jan P. Allebach, and Charles A. Bouman</i>	
FCSS: Fully Convolutional Self-Similarity for Dense Semantic Correspondence	616
<i>Seungryoung Kim, Dongbo Min, Bumsub Ham, Sangryul Jeon, Stephen Lin, and Kwanghoon Sohn</i>	
Learning by Association — A Versatile Semi-Supervised Training Method for Neural Networks	626
<i>Philip Haeusser, Alexander Mordvintsev, and Daniel Cremers</i>	

Dilated Residual Networks	636
<i>Fisher Yu, Vladlen Koltun, and Thomas Funkhouser</i>	
Split-Brain Autoencoders: Unsupervised Learning by Cross-Channel Prediction	645
<i>Richard Zhang, Phillip Isola, and Alexei A. Efros</i>	
Nonnegative Matrix Underapproximation for Robust Multiple Model Fitting	655
<i>Mariano Tepper and Guillermo Sapiro</i>	
Truncated Max-of-Convex Models	664
<i>Pankaj Pansari and M. Pawan Kumar</i>	
Additive Component Analysis	673
<i>Calvin Murdock and Fernando De la Torre</i>	
Subspace Clustering via Variance Regularized Ridge Regression	682
<i>Chong Peng, Zhao Kang, and Qiang Cheng</i>	
The Incremental Multiresolution Matrix Factorization Algorithm	692
<i>Vamsi K. Ithapu, Risi Kondor, Sterling C. Johnson, and Vikas Singh</i>	
Transformation-Grounded Image Generation Network for Novel 3D View Synthesis	702
<i>Eunbyung Park, Jimei Yang, Ersin Yumer, Duygu Ceylan, and Alexander C. Berg</i>	
Learning Dynamic Guidance for Depth Image Enhancement	712
<i>Shuhang Gu, Wangmeng Zuo, Shi Guo, Yunjin Chen, Chongyu Chen, and Lei Zhang</i>	
A-Lamp: Adaptive Layout-Aware Multi-patch Deep Convolutional Neural Network for Photo Aesthetic Assessment	722
<i>Shuang Ma, Jing Liu, and Chang Wen Chen</i>	
Teaching Compositionality to CNNs	732
<i>Austin Stone, Huayan Wang, Michael Stark, Yi Liu, D. Scott Phoenix, and Dileep George</i>	
Using Ranking-CNN for Age Estimation	742
<i>Shixing Chen, Caojin Zhang, Ming Dong, Jialiang Le, and Mike Rao</i>	
Accurate Single Stage Detector Using Recurrent Rolling Convolution	752
<i>Jimmy Ren, Xiaohao Chen, Jianbo Liu, Wenxiu Sun, Jiahao Pang, Qiong Yan, Yu-Wing Tai, and Li Xu</i>	
A Compact DNN: Approaching GoogLeNet-Level Accuracy of Classification and Domain Adaptation	761
<i>Chunpeng Wu, Wei Wen, Tariq Afzal, Yongmei Zhang, Yiran Chen, and Hai (Helen) Li</i>	
The Impact of Typicality for Informative Representative Selection	771
<i>Jawadul H. Bappy, Sujoy Paul, Ertem Tuncel, and Amit K. Roy-Chowdhury</i>	
Infinite Variational Autoencoder for Semi-Supervised Learning	781
<i>M. Ehsan Abbasnejad, Anthony Dick, and Anton van den Hengel</i>	
SurfNet: Generating 3D Shape Surfaces Using Deep Residual Networks	791
<i>Ayan Sinha, Asim Unmesh, Qixing Huang, and Karthik Ramani</i>	
Intrinsic Grassmann Averages for Online Linear and Robust Subspace Learning	801
<i>Rudrasis Chakraborty, Søren Hauberg, and Baba C. Vemuri</i>	
Variational Bayesian Multiple Instance Learning with Gaussian Processes	810
<i>Manuel Haußmann, Fred A. Hamprecht, and Melih Kandemir</i>	

Temporal Attention-Gated Model for Robust Sequence Classification	820
<i>Wenjie Pei, Tadas Baltrušaitis, David M. J. Tax, and Louis-Philippe Morency</i>	
Non-uniform Subset Selection for Active Learning in Structured Data	830
<i>Sujoy Paul, Jawadul H. Bappy, and Amit K. Roy-Chowdhury</i>	
Colorization as a Proxy Task for Visual Understanding	840
<i>Gustav Larsson, Michael Maire, and Gregory Shakhnarovich</i>	
Shading Annotations in the Wild	850
<i>Balazs Kovacs, Sean Bell, Noah Snavely, and Kavita Bala</i>	
LCNN: Lookup-Based Convolutional Neural Network	860
<i>Hessam Bagherinezhad, Mohammad Rastegari, and Ali Farhadi</i>	
Physics Inspired Optimization on Semantic Transfer Features: An Alternative Method for Room Layout Estimation	870
<i>Hao Zhao, Ming Lu, Anbang Yao, Yiwen Guo, Yurong Chen, and Li Zhang</i>	
Pixelwise Instance Segmentation with a Dynamically Instantiated Network	879
<i>Anurag Arnab and Philip H. S. Torr</i>	
Object Detection in Videos with Tubelet Proposal Networks	889
<i>Kai Kang, Hongsheng Li, Tong Xiao, Wanli Ouyang, Junjie Yan, Xihui Liu, and Xiaogang Wang</i>	
AMVH: Asymmetric Multi-Valued hashing	898
<i>Cheng Da, Shibiao Xu, Kun Ding, Gaofeng Meng, Shiming Xiang, and Chunhong Pan</i>	
Spindle Net: Person Re-identification with Human Body Region Guided Feature Decomposition and Fusion	907
<i>Haiyu Zhao, Maoqing Tian, Shuyang Sun, Jing Shao, Junjie Yan, Shuai Yi, Xiaogang Wang, and Xiaoou Tang</i>	
Deep Visual-Semantic Quantization for Efficient Image Retrieval	916
<i>Yue Cao, Mingsheng Long, Jianmin Wang, and Shichen Liu</i>	
Efficient Diffusion on Region Manifolds: Recovering Small Objects with Compact CNN Representations	926
<i>Ahmet Iscen, Giorgos Tolias, Yannis Avrithis, Teddy Furon, and Ondrej Chum</i>	
Feature Pyramid Networks for Object Detection	936
<i>Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie</i>	
Mind the Class Weight Bias: Weighted Maximum Mean Discrepancy for Unsupervised Domain Adaptation	945
<i>Hongliang Yan, Yukang Ding, Peihua Li, Qilong Wang, Yong Xu, and Wangmeng Zuo</i>	
StyleNet: Generating Attractive Visual Captions with Styles	955
<i>Chuang Gan, Zhe Gan, Xiaodong He, Jianfeng Gao, and Li Deng</i>	
Fine-Grained Recognition of Thousands of Object Categories with Single-Example Training	965
<i>Leonid Karlinsky, Joseph Shtok, Yochay Tzur, and Asaf Tzadok</i>	

Improving Interpretability of Deep Neural Networks with Semantic Information	975
<i>Yinpeng Dong, Hang Su, Jun Zhu, and Bo Zhang</i>	
Video Captioning with Transferred Semantic Attributes	984
<i>Yingwei Pan, Ting Yao, Houqiang Li, and Tao Mei</i>	
Fast Boosting Based Detection Using Scale Invariant Multimodal Multiresolution Filtered Features	993
<i>Arthur Daniel Costea, Robert Varga, and Sergiu Nedevschi</i>	
Temporal Convolutional Networks for Action Segmentation and Detection	1003
<i>Colin Lea, Michael D. Flynn, René Vidal, Austin Reiter, and Gregory D. Hager</i>	
Surveillance Video Parsing with Single Frame Supervision	1013
<i>Si Liu, Changhu Wang, Ruihe Qian, Han Yu, Renda Bao, and Yao Sun</i>	
Weakly Supervised Actor-Action Segmentation via Robust Multi-task Ranking	1022
<i>Yan Yan, Chenliang Xu, Dawen Cai, and Jason J. Corso</i>	
Unsupervised Visual-Linguistic Reference Resolution in Instructional Videos	1032
<i>De-An Huang, Joseph J. Lim, Li Fei-Fei, and Juan Carlos Niebles</i>	
Zero-Shot Action Recognition with Error-Correcting Output Codes	1042
<i>Jie Qin, Li Liu, Ling Shao, Fumin Shen, Bingbing Ni, Jiaxin Chen, and Yunhong Wang</i>	
Enhancing Video Summarization via Vision-Language Embedding	1052
<i>Bryan A. Plummer, Matthew Brown, and Svetlana Lazebnik</i>	
Synthesizing Dynamic Patterns by Spatial-Temporal Generative ConvNet	1061
<i>Jianwen Xie, Song-Chun Zhu, and Ying Nian Wu</i>	

Oral and Spotlight 1-2A

S1-2A: Object Recognition & Scene Understanding — Computer Vision & Language

Context-Aware Captions from Context-Agnostic Supervision	1070
<i>Ramakrishna Vedantam, Samy Bengio, Kevin Murphy, Devi Parikh, and Gal Chechik</i>	
Visual Dialog	1080
<i>Abhishek Das, Satwik Kottur, Khushi Gupta, Avi Singh, Deshraj Yadav, José M. F. Moura, Devi Parikh, and Dhruv Batra</i>	
Discriminative Bimodal Networks for Visual Localization and Detection with Natural Language Queries	1090
<i>Yuting Zhang, Luyao Yuan, Yijie Guo, Zhiyuan He, I-An Huang, and Honglak Lee</i>	
Automatic Understanding of Image and Video Advertisements	1100
<i>Zaeem Hussain, Mingda Zhang, Xiaozhong Zhang, Keren Ye, Christopher Thomas, Zuha Agha, Nathan Ong, and Adriana Kovashka</i>	
Discover and Learn New Objects from Documentaries	1111
<i>Kai Chen, Hang Song, Chen Change Loy, and Dahua Lin</i>	
Spatial-Semantic Image Search by Visual Feature Synthesis	1121
<i>Long Mai, Hailin Jin, Zhe Lin, Chen Fang, Jonathan Brandt, and Feng Liu</i>	

Fully-Adaptive Feature Sharing in Multi-Task Networks with Applications in Person Attribute Classification	1131
<i>Yongxi Lu, Abhishek Kumar, Shuangfei Zhai, Yu Cheng, Tara Javidi, and Rogerio Feris</i>	
Semantic Compositional Networks for Visual Captioning	1141
<i>Zhe Gan, Chuang Gan, Xiaodong He, Yunchen Pu, Kenneth Tran, Jianfeng Gao, Lawrence Carin, and Li Deng</i>	

O1-2A

Deep Reinforcement Learning-Based Image Captioning with Embedding Reward	1151
<i>Zhou Ren, Xiaoyu Wang, Ning Zhang, Xutao Lv, and Li-Jia Li</i>	
From Red Wine to Red Tomato: Composition with Context	1160
<i>Ishan Misra, Abhinav Gupta, and Martial Hebert</i>	
Captioning Images with Diverse Objects	1170
<i>Subhashini Venugopalan, Lisa Anne Hendricks, Marcus Rohrbach, Raymond Mooney, Trevor Darrell, and Kate Saenko</i>	
Self-Critical Sequence Training for Image Captioning	1179
<i>Steven J. Rennie, Etienne Marcheret, Youssef Mroueh, Jerret Ross, and Vaibhava Goel</i>	

Oral and Spotlight 1-2B

S1-2B: Analyzing Humans 1

Crossing Nets: Combining GANs and VAEs with a Shared Latent Space for Hand Pose Estimation	1196
<i>Chengde Wan, Thomas Probst, Luc Van Gool, and Angela Yao</i>	
Predicting Behaviors of Basketball Players from First Person Videos	1206
<i>Shan Su, Jung Pyo Hong, Jianbo Shi, and Hyun Soo Park</i>	
LCR-Net: Localization-Classification-Regression for Human Pose	1216
<i>Grégory Rogez, Philippe Weinzaepfel, and Cordelia Schmid</i>	
Learning Residual Images for Face Attribute Manipulation	1225
<i>Wei Shen and Rujie Liu</i>	
Seeing What is Not There: Learning Context to Determine Where Objects are Missing	1234
<i>Jin Sun and David W. Jacobs</i>	
Deep Learning on Lie Groups for Skeleton-Based Action Recognition	1243
<i>Zhiwu Huang, Chengde Wan, Thomas Probst, and Luc Van Gool</i>	
Harvesting Multiple Views for Marker-Less 3D Human Pose Annotations	1253
<i>Georgios Pavlakos, Xiaowei Zhou, Konstantinos G. Derpanis, and Kostas Daniilidis</i>	
Coarse-to-Fine Volumetric Prediction for Single-Image 3D Human Pose	1263
<i>Georgios Pavlakos, Xiaowei Zhou, Konstantinos G. Derpanis, and Kostas Daniilidis</i>	

O1-2B

Weakly Supervised Action Learning with RNN Based Fine-to-Coarse Modeling	1273
<i>Alexander Richard, Hilde Kuehne, and Juergen Gall</i>	
Disentangled Representation Learning GAN for Pose-Invariant Face Recognition	1283
<i>Luan Tran, Xi Yin, and Xiaoming Liu</i>	
ArtTrack: Articulated Multi-Person Tracking in the Wild	1293
<i>Eldar Insafutdinov, Mykhaylo Andriluka, Leonid Pishchulin, Siyu Tang, Evgeny Levinkov, Bjoern Andres, and Bernt Schiele</i>	
Realtime Multi-person 2D Pose Estimation Using Part Affinity Fields	1302
<i>Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh</i>	

Oral and Spotlight 1-2C

S1-2C: Image Motion & Tracking; Video Analysis

Template Matching with Deformable Diversity Similarity	1311
<i>Itamar Talmi, Roey Mechrez, and Lihi Zelnik-Manor</i>	
Beyond Triplet Loss: A Deep Quadruplet Network for Person Re-identification	1320
<i>Weihua Chen, Xiaotang Chen, Jianguo Zhang, and Kaiqi Huang</i>	
Agent-Centric Risk Assessment: Accident Anticipation and Risky Region Localization	1330
<i>Kuo-Hao Zeng, Shih-Han Chou, Fu-Hsiang Chan, Juan Carlos Niebles, and Min Sun</i>	
Bidirectional Multirate Reconstruction for Temporal Modeling in Videos	1339
<i>Linchao Zhu, Zhongwen Xu, and Yi Yang</i>	
Action-Decision Networks for Visual Tracking with Deep Reinforcement Learning	1349
<i>Sangdoon Yun, Jongwon Choi, Youngjoon Yoo, Kimin Yun, and Jin Young Choi</i>	
TGIF-QA: Toward Spatio-Temporal Reasoning in Visual Question Answering	1359
<i>Yunseok Jang, Yale Song, Youngjae Yu, Youngjin Kim, and Gunhee Kim</i>	
Making 360° Video Watchable in 2D: Learning Videography for Click Free Viewing	1368
<i>Yu-Chuan Su and Kristen Grauman</i>	
Unsupervised Adaptive Re-identification in Open World Dynamic Camera Networks	1377
<i>Rameswar Panda, Amran Bhuiyan, Vittorio Murino, and Amit K. Roy-Chowdhury</i>	

O1-2C

Context-Aware Correlation Filter Tracking	1387
<i>Matthias Mueller, Neil Smith, and Bernard Ghanem</i>	
Deep 360 Pilot: Learning a Deep Agent for Piloting through 360° Sports Videos	1396
<i>Hou-Ning Hu, Yen-Chen Lin, Ming-Yu Liu, Hsien-Tzu Cheng, Yung-Ju Chang, and Min Sun</i>	
Slow Flow: Exploiting High-Speed Cameras for Accurate and Diverse Optical Flow Reference Data	1406
<i>Joel Janai, Fatma Güney, Jonas Wulff, Michael J. Black, and Andreas Geiger</i>	

CDC: Convolutional-De-Convolutional Networks for Precise Temporal Action Localization in Untrimmed Videos	1417
<i>Zheng Shou, Jonathan Chan, Alireza Zareian, Kazuyuki Miyazawa, and Shih-Fu Chang</i>	

Poster Session P1-2

Exploiting 2D Floorplan for Building-Scale Panorama RGBD Alignment	1427
<i>Erik Wijmans and Yasutaka Furukawa</i>	
A Combinatorial Solution to Non-Rigid 3D Shape-to-Image Matching	1436
<i>Florian Bernard, Frank R. Schmidt, Johan Thunberg, and Daniel Cremers</i>	
NID-SLAM: Robust Monocular SLAM Using Normalised Information Distance	1446
<i>Geoffrey Pascoe, Will Maddern, Michael Tanner, Pedro Piniés, and Paul Newman</i>	
End-to-End Training of Hybrid CNN-CRF Models for Stereo	1456
<i>Patrick Knöbelreiter, Christian Reinbacher, Alexander Shekhovtsov, and Thomas Pock</i>	
Learning Shape Abstractions by Assembling Volumetric Primitives	1466
<i>Shubham Tulsiani, Hao Su, Leonidas J. Guibas, Alexei A. Efros, and Jitendra Malik</i>	
Locality-Sensitive Deconvolution Networks with Gated Fusion for RGB-D Indoor Semantic Segmentation	1475
<i>Yanhua Cheng, Rui Cai, Zhiwei Li, Xin Zhao, and Kaiqi Huang</i>	
Acquiring Axially-Symmetric Transparent Objects Using Single-View Transmission Imaging	1484
<i>Jaewon Kim, Ilya Reshetouski, and Abhijeet Ghosh</i>	
Regressing Robust and Discriminative 3D Morphable Models with a Very Deep Neural Network	1493
<i>Anh Tuan Tran, Tal Hassner, Iacopo Masi, and Gérard Medioni</i>	
End-to-End 3D Face Reconstruction with Deep Neural Networks	1503
<i>Pengfei Dou, Shishir K. Shah, and Ioannis A. Kakadiaris</i>	
DUST: Dual Union of Spatio-Temporal Subspaces for Monocular Multiple Object 3D Reconstruction	1513
<i>Antonio Agudo and Francesc Moreno-Noguer</i>	
Finding Tiny Faces	1522
<i>Peiyun Hu and Deva Ramanan</i>	
Dynamic Facial Analysis: From Bayesian Filtering to Recurrent Neural Network	1531
<i>Jinwei Gu, Xiaodong Yang, Shalini De Mello, and Jan Kautz</i>	
Deep Temporal Linear Encoding Networks	1541
<i>Ali Diba, Vivek Sharma, and Luc Van Gool</i>	
Joint Registration and Representation Learning for Unconstrained Face Identification	1551
<i>Munawar Hayat, Salman H. Khan, Naoufel Werghi, and Roland Goecke</i>	
3D Human Pose Estimation from a Single Image via Distance Matrix Regression	1561
<i>Francesc Moreno-Noguer</i>	
One-Shot Metric Learning for Person Re-identification	1571
<i>Sławomir Bak and Peter Carr</i>	

Generalized Rank Pooling for Activity Recognition	1581
<i>Anoop Cherian, Basura Fernando, Mehrtash Harandi, and Stephen Gould</i>	
Deep Representation Learning for Human Motion Prediction and Classification	1591
<i>Judith Bütepage, Michael J. Black, Danica Kragic, and Hedvig Kjellström</i>	
Interspecies Knowledge Transfer for Facial Keypoint Detection	1600
<i>Maheen Rashid, Xiuye Gu, and Yong Jae Lee</i>	
Recurrent Convolutional Neural Networks for Continuous Sign Language Recognition by Staged Optimization	1610
<i>Runpeng Cui, Hu Liu, and Changshui Zhang</i>	
Modeling Sub-Event Dynamics in First-Person Action Recognition	1619
<i>Hasan F. M. Zaki, Faisal Shafait, and Ajmal Mian</i>	
Turning an Urban Scene Video into a Cinemagraph	1629
<i>Hang Yan, Yebin Liu, and Yasutaka Furukawa</i>	
Light Field Reconstruction Using Deep Convolutional Network on EPI	1638
<i>Gaochang Wu, Mandan Zhao, Liangyong Wang, Qionghai Dai, Tianyou Chai, and Yebin Liu</i>	
FlowNet 2.0: Evolution of Optical Flow Estimation with Deep Networks	1647
<i>Eddy Ilg, Nikolaus Mayer, Tonmoy Saikia, Margret Keuper, Alexey Dosovitskiy, and Thomas Brox</i>	
Attention-Aware Face Hallucination via Deep Reinforcement Learning	1656
<i>Qingxing Cao, Liang Lin, Yukai Shi, Xiaodan Liang, and Guanbin Li</i>	
Simple Does It: Weakly Supervised Instance and Semantic Segmentation	1665
<i>Anna Khoreva, Rodrigo Benenson, Jan Hosang, Matthias Hein, and Bernt Schiele</i>	
Anti-Glare: Tightly Constrained Optimization for Eyeglass Reflection Removal	1675
<i>Tushar Sandhan and Jin Young Choi</i>	
Deep Joint Rain Detection and Removal from a Single Image	1685
<i>Wenhan Yang, Robby T. Tan, Jiashi Feng, Jiaying Liu, Zongming Guo, and Shuicheng Yan</i>	
Radiometric Calibration from Faces in Images	1695
<i>Chen Li, Stephen Lin, Kun Zhou, and Katsushi Ikeuchi</i>	
Webly Supervised Semantic Segmentation	1705
<i>Bin Jin, Maria V. Ortiz Segovia, and Sabine Sússtrunk</i>	
Removing Rain from Single Images via a Deep Detail Network	1715
<i>Xueyang Fu, Jiabin Huang, Delu Zeng, Yue Huang, Xinghao Ding, and John Paisley</i>	
Deep Crisp Boundaries	1724
<i>Yupei Wang, Xin Zhao, and Kaiqi Huang</i>	
Coarse-to-Fine Segmentation with Shape-Tailored Continuum Scale Spaces	1733
<i>Naeemullah Khan, Byung-Woo Hong, Anthony Yezzi, and Ganesh Sundaramoorthi</i>	
Large Kernel Matters — Improve Semantic Segmentation by Global Convolutional Network	1743
<i>Chao Peng, Xiangyu Zhang, Gang Yu, Guiming Luo, and Jian Sun</i>	
Single Image Reflection Suppression	1752
<i>Nikolaos Arvanitopoulos, Radhakrishna Achanta, and Sabine Sússtrunk</i>	

CASENet: Deep Category-Aware Semantic Edge Detection	1761
<i>Zhiding Yu, Chen Feng, Ming-Yu Liu, and Srikumar Ramalingam</i>	
Reflectance Adaptive Filtering Improves Intrinsic Image Estimation	1771
<i>Thomas Nestmeyer and Peter V. Gehler</i>	
Conditional Similarity Networks	1781
<i>Andreas Veit, Serge Belongie, and Theofanis Karaletsos</i>	
Spatially Adaptive Computation Time for Residual Networks	1790
<i>Michael Figurnov, Maxwell D. Collins, Yukun Zhu, Li Zhang, Jonathan Huang, Dmitry Vetrov, and Ruslan Salakhutdinov</i>	
Xception: Deep Learning with Depthwise Separable Convolutions	1800
<i>François Chollet</i>	
Feedback Networks	1808
<i>Amir R. Zamir, Te-Lin Wu, Lin Sun, William B. Shen, Bertram E. Shi, Jitendra Malik, and Silvio Savarese</i>	
Online Summarization via Submodular and Convex Optimization	1818
<i>Ehsan Elhamifar and M. Clara De Paolis Kaluza</i>	
Deep MANTA: A Coarse-to-Fine Many-Task Network for Joint 2D and 3D Vehicle Analysis from Monocular Image	1827
<i>Florian Chabot, Mohamed Chaouch, Jaonary Rabarisoa, Céline Teulière, and Thierry Chateau</i>	
Improving Pairwise Ranking for Multi-label Image Classification	1837
<i>Yuncheng Li, Yale Song, and Jiebo Luo</i>	
Active Convolution: Learning the Shape of Convolution for Image Classification	1846
<i>Yunho Jeon and Junmo Kim</i>	
Linking Image and Text with 2-Way Nets	1855
<i>Aviv Eisenschtat and Lior Wolf</i>	
Stacked Generative Adversarial Networks	1866
<i>Xun Huang, Yixuan Li, Omid Poursaeed, John Hopcroft, and Serge Belongie</i>	
Image Splicing Detection via Camera Response Function Analysis	1876
<i>Can Chen, Scott McCloskey, and Jingyi Yu</i>	
Building a Regular Decision Boundary with Deep Networks	1886
<i>Edouard Oyallon</i>	
More is Less: A More Complicated Network with Less Inference Complexity	1895
<i>Xuanyi Dong, Junshi Huang, Yi Yang, and Shuicheng Yan</i>	
Joint Graph Decomposition & Node Labeling: Problem, Algorithms, Applications	1904
<i>Evgeny Levinkov, Jonas Uhrig, Siyu Tang, Mohamed Omran, Eldar Insafutdinov, Alexander Kirillov, Carsten Rother, Thomas Brox, Bernt Schiele, and Bjoern Andres</i>	
Scale-Aware Face Detection	1913
<i>Zekun Hao, Yu Liu, Hongwei Qin, Junjie Yan, Xiu Li, and Xiaolin Hu</i>	
Deep Unsupervised Similarity Learning Using Partially Ordered Sets	1923
<i>Miguel A. Bautista, Artsiom Sanakoyeu, and Björn Ommer</i>	

Generative Hierarchical Learning of Sparse FRAME Models	1933
<i>Jianwen Xie, Yifei Xu, Erik Nijkamp, Ying Nian Wu, and Song-Chun Zhu</i>	
Generating Holistic 3D Scene Abstractions for Text-Based Image Retrieval	1942
<i>Ang Li, Jin Sun, Joe Yue-Hei Ng, Ruichi Yu, Vlad I. Morariu, and Larry S. Davis</i>	
Perceptual Generative Adversarial Networks for Small Object Detection	1951
<i>Jianan Li, Xiaodan Liang, Yunchao Wei, Tingfa Xu, Jiashi Feng, and Shuicheng Yan</i>	
Emotion Recognition in Context	1960
<i>Ronak Kostli, Jose M. Alvarez, Adria Recasens, and Agata Lapedriza</i>	
Deep Learning of Human Visual Sensitivity in Image Quality Assessment Framework	1969
<i>Jongyoo Kim and Sanghoon Lee</i>	
Dense Captioning with Joint Inference and Visual Context	1978
<i>Linjie Yang, Kevin Tang, Jianchao Yang, and Li-Jia Li</i>	
CLEVR: A Diagnostic Dataset for Compositional Language and Elementary Visual Reasoning	1988
<i>Justin Johnson, Bharath Hariharan, Laurens van der Maaten, Li Fei-Fei, C. Lawrence Zitnick, and Ross Girshick</i>	
Cross-View Image Matching for Geo-Localization in Urban Environments	1998
<i>Yicong Tian, Chen Chen, and Mubarak Shah</i>	
Matrix Tri-Factorization with Manifold Regularizations for Zero-Shot Learning	2007
<i>Xing Xu, Fumin Shen, Yang Yang, Dongxiang Zhang, Heng Tao Shen, and Jingkuan Song</i>	
Self-Supervised Learning of Visual Features through Embedding Images into Text Topic Spaces	2017
<i>Lluís Gomez, Yash Patel, Marçal Rusiñol, Dimosthenis Karatzas, and C. V. Jawahar</i>	
Learning Spatial Regularization with Image-Level Supervisions for Multi-label Image Classification	2027
<i>Feng Zhu, Hongsheng Li, Wanli Ouyang, Nenghai Yu, and Xiaogang Wang</i>	
Semantically Consistent Regularization for Zero-Shot Recognition	2037
<i>Pedro Morgado and Nuno Vasconcelos</i>	
Can Walking and Measuring Along Chord Bunches Better Describe Leaf Shapes?	2047
<i>Bin Wang, Yongsheng Gao, Changming Sun, Michael Blumenstein, and John La Salle</i>	
Self-Learning Scene-Specific Pedestrian Detectors Using a Progressive Latent Model	2057
<i>Qixiang Ye, Tianliang Zhang, Wei Ke, Qiang Qiu, Jie Chen, Guillermo Sapiro, and Baochang Zhang</i>	
Predictive-Corrective Networks for Action Detection	2067
<i>Achal Dave, Olga Russakovsky, and Deva Ramanan</i>	
Budget-Aware Deep Semantic Video Segmentation	2077
<i>Behrooz Mahasseni, Sinisa Todorovic, and Alan Fern</i>	
Unified Embedding and Metric Learning for Zero-Exemplar Event Detection	2087
<i>Noureddien Hussein, Efstratios Gavves, and Arnold W. M. Smeulders</i>	
Spatiotemporal Pyramid Network for Video Action Recognition	2097
<i>Yunbo Wang, Mingsheng Long, Jianmin Wang, and Philip S. Yu</i>	

ER3: A Unified Framework for Event Retrieval, Recognition and Recounting	2107
<i>Zhanning Gao, Gang Hua, Dongqing Zhang, Nebojsa Jojic, Le Wang, Jianru Xue, and Nanning Zheng</i>	
FusionSeg: Learning to Combine Motion and Appearance for Fully Automatic Segmentation of Generic Objects in Videos	2117
<i>Suyog Dutt Jain, Bo Xiong, and Kristen Grauman</i>	
Query-Focused Video Summarization: Dataset, Evaluation, and a Memory Network Based Approach	2127
<i>Aidean Sharghi, Jacob S. Laurel, and Boqing Gong</i>	
Flexible Spatio-Temporal Networks for Video Prediction	2137
<i>Chaochao Lu, Michael Hirsch, and Bernhard Schölkopf</i>	
Temporal Action Co-Segmentation in 3D Motion Capture Data and Videos	2146
<i>Konstantinos Papoutsakis, Costas Panagiotakis, and Antonis A. Argyros</i>	

Oral and Spotlight 2-1A

S2-1A: Machine Learning 2

Dual Attention Networks for Multimodal Reasoning and Matching	2156
<i>Hyeonseob Nam, Jung-Woo Ha, and Jeonghee Kim</i>	
DESIRE: Distant Future Prediction in Dynamic Scenes with Interacting Agents	2165
<i>Namhoon Lee, Wongun Choi, Paul Vernaza, Christopher B. Choy, Philip H. S. Torr, and Manmohan Chandraker</i>	
Interpretable Structure-Evolving LSTM	2175
<i>Xiaodan Liang, Liang Lin, Xiaohui Shen, Jiashi Feng, Shuicheng Yan, and Eric P. Xing</i>	
ShapeOdds: Variational Bayesian Learning of Generative Shape Models	2185
<i>Shireen Elhabian and Ross Whitaker</i>	
Fast Video Classification via Adaptive Cascading of Deep Models	2197
<i>Haichen Shen, Seungyeop Han, Matthai Philipose, and Arvind Krishnamurthy</i>	
Deep Metric Learning via Facility Location	2206
<i>Hyun Oh Song, Stefanie Jegelka, Vivek Rathod, and Kevin Murphy</i>	
Semi-Supervised Deep Learning for Monocular Depth Map Prediction	2215
<i>Yevhen Kuznietsov, Jörg Stückler, and Bastian Leibe</i>	
Weakly Supervised Semantic Segmentation Using Web-Crawled Videos	2224
<i>Seunghoon Hong, Donghun Yeo, Suha Kwak, Honglak Lee, and Bohyung Han</i>	

O2-1A

Making Deep Neural Networks Robust to Label Noise: A Loss Correction Approach	2233
<i>Giorgio Patrini, Alessandro Rozza, Aditya Krishna Menon, Richard Nock, and Lizhen Qu</i>	
Learning from Simulated and Unsupervised Images through Adversarial Training	2242
<i>Ashish Shrivastava, Tomas Pfister, Oncel Tuzel, Joshua Susskind, Wenda Wang, and Russell Webb</i>	

Inverse Compositional Spatial Transformer Networks	2252
<i>Chen-Hsuan Lin and Simon Lucey</i>	
Densely Connected Convolutional Networks	2261
<i>Gao Huang, Zhuang Liu, Laurens van der Maaten, and Kilian Q. Weinberger</i>	

Oral and Spotlight 2-1B

S2-1B: Computational Photography

Video Frame Interpolation via Adaptive Convolution	2270
<i>Simon Niklaus, Long Mai, and Feng Liu</i>	
FastMask: Segment Multi-scale Object Candidates in One Shot	2280
<i>Hexiang Hu, Shiyi Lan, Yuning Jiang, Zhimin Cao, and Fei Sha</i>	
Reconstructing Transient Images from Single-Photon Sensors	2289
<i>Matthew O'Toole, Felix Heide, David B. Lindell, Kai Zang, Steven Diamond, and Gordon Wetzstein</i>	
Deep Sketch Hashing: Fast Free-Hand Sketch-Based Image Retrieval	2298
<i>Li Liu, Fumin Shen, Yuming Shen, Xianglong Liu, and Ling Shao</i>	
DeshadowNet: A Multi-context Embedding Deep Network for Shadow Removal	2308
<i>Liangqiong Qu, Jiandong Tian, Shengfeng He, Yandong Tang, and Rynson W. H. Lau</i>	
Illuminant-Camera Communication to Observe Moving Objects under Strong External Light by Spread Spectrum Modulation	2317
<i>Ryusuke Sagawa and Yutaka Satoh</i>	
Photorealistic Facial Texture Inference Using Deep Neural Networks	2326
<i>Shunsuke Saito, Lingyu Wei, Liwen Hu, Koki Nagano, and Hao Li</i>	
The Geometry of First-Returning Photons for Non-Line-of-Sight Imaging	2336
<i>Chia-Yin Tsai, Kiriakos N. Kutulakos, Srinivasa G. Narasimhan, and Aswin C. Sankaranarayanan</i>	

O2-1B

Unrolling the Shutter: CNN to Correct Motion Distortions	2345
<i>Vijay Rengarajan, Yogesh Balaji, and A. N. Rajagopalan</i>	
Light Field Blind Motion Deblurring	2354
<i>Pratul P. Srinivasan, Ren Ng, and Ravi Ramamoorthi</i>	
Computational Imaging on the Electric Grid	2363
<i>Mark Sheinin, Yoav Y. Schechner, and Kiriakos N. Kutulakos</i>	
Deep Outdoor Illumination Estimation	2373
<i>Yannick Hold-Geoffroy, Kalyan Sunkavalli, Sunil Hadap, Emiliano Gambaretto, and Jean-François Lalonde</i>	

Oral and Spotlight 2-1C

S2-1C: 3D Vision 2

Efficient Solvers for Minimal Problems by Syzygy-Based Reduction	2383
<i>Viktor Larsson, Kalle Åström, and Magnus Oskarsson</i>	
HSfM: Hybrid Structure-from-Motion	2393
<i>Hainan Cui, Xiang Gao, Shuhan Shen, and Zhanyi Hu</i>	
Efficient Global Point Cloud Alignment Using Bayesian Nonparametric Mixtures	2403
<i>Julian Straub, Trevor Campbell, Jonathan P. How, and John W. Fisher III</i>	
A New Rank Constraint on Multi-view Fundamental Matrices, and Its Application to Camera Location Recovery	2413
<i>Soumyadip Sengupta, Tal Amir, Meirav Galun, Tom Goldstein, David W. Jacobs, Amit Singer, and Ronen Basri</i>	
IM2CAD	2422
<i>Hamid Izadinia, Qi Shan, and Steven M. Seitz</i>	
ScanNet: Richly-Annotated 3D Reconstructions of Indoor Scenes	2432
<i>Angela Dai, Angel X. Chang, Manolis Savva, Maciej Halber, Thomas Funkhouser, and Matthias Nießner</i>	
Noise Robust Depth from Focus Using a Ring Difference Filter	2444
<i>Jaeheung Surh, Hae-Gon Jeon, Yunwon Park, Sunghoon Im, Hyowon Ha, and In So Kweon</i>	
Group-Wise Point-Set Registration Based on Rényi's Second Order Entropy	2454
<i>Luis G. Sanchez Giraldo, Erion Hasanbelliu, Murali Rao, and Jose C. Principe</i>	

O2-1C

A Point Set Generation Network for 3D Object Reconstruction from a Single Image	2463
<i>Haoqiang Fan, Hao Su, and Leonidas Guibas</i>	
3D Point Cloud Registration for Localization Using a Deep Neural Network Auto-Encoder	2472
<i>Gil Elbaz, Tamar Avraham, and Anath Fischer</i>	
Flight Dynamics-Based Recovery of a UAV Trajectory Using Ground Cameras	2482
<i>Artem Rozantsev, Sudipta N. Sinha, Debadeepta Dey, and Pascal Fua</i>	
DSAC — Differentiable RANSAC for Camera Localization	2492
<i>Eric Brachmann, Alexander Krull, Sebastian Nowozin, Jamie Shotton, Frank Michel, Stefan Gumhold, and Carsten Rother</i>	

Poster Session P2-1

Scalable Surface Reconstruction from Point Clouds with Extreme Scale and Density Diversity	2501
<i>Christian Mostegel, Rudolf Prettenthaler, Friedrich Fraundorfer, and Horst Bischof</i>	

Synthesizing 3D Shapes via Modeling Multi-view Depth Maps and Silhouettes with Deep Generative Networks	2511
<i>Amir Arsalan Soltani, Haibin Huang, Jiajun Wu, Tejas D. Kulkarni, and Joshua B. Tenenbaum</i>	
General Models for Rational Cameras and the Case of Two-Slit Projections	2520
<i>Matthew Trager, Bernd Sturmfels, John Canny, Martial Hebert, and Jean Ponce</i>	
Accurate Depth and Normal Maps from Occlusion-Aware Focal Stack Symmetry	2529
<i>Michael Strecke, Anna Alperovich, and Bastian Goldluecke</i>	
A Multi-view Stereo Benchmark with High-Resolution Images and Multi-camera Videos	2538
<i>Thomas Schöps, Johannes L. Schönberger, Silvano Galliani, Torsten Sattler, Konrad Schindler, Marc Pollefeys, and Andreas Geiger</i>	
Non-contact Full Field Vibration Measurement Based on Phase-Shifting	2548
<i>HiroYuki Kayaba and Yuji Kokumai</i>	
A Minimal Solution for Two-View Focal-Length Estimation Using Two Affine Correspondences	2557
<i>Daniel Barath, Tekla Toth, and Levente Hajder</i>	
PoseAgent: Budget-Constrained 6D Object Pose Estimation via Reinforcement Learning	2566
<i>Alexander Krull, Eric Brachmann, Sebastian Nowozin, Frank Michel, Jamie Shotton, and Carsten Rother</i>	
An Efficient Background Term for 3D Reconstruction and Tracking with Smooth Surface Models	2575
<i>Mariano Jaimez, Thomas J. Cashman, Andrew Fitzgibbon, Javier Gonzalez-Jimenez, and Daniel Cremers</i>	
Reliable Crowdsourcing and Deep Locality-Preserving Learning for Expression Recognition in the Wild	2584
<i>Shan Li, Weihong Deng, and JunPing Du</i>	
Procedural Generation of Videos to Train Deep Action Recognition Networks	2594
<i>César Roberto de Souza, Adrien Gaidon, Yohann Cabon, and Antonio Manuel López</i>	
BigHand2.2M Benchmark: Hand Pose Dataset and State of the Art Analysis	2605
<i>Shanxin Yuan, Qi Ye, Björn Stenger, Siddhant Jain, and Tae-Kyun Kim</i>	
DenseReg: Fully Convolutional Dense Shape Regression In-the-Wild	2614
<i>Riza Alp Güler, George Trigeorgis, Epameinondas Antonakos, Patrick Snape, Stefanos Zafeiriou, and Iasonas Kokkinos</i>	
Adaptive Class Preserving Representation for Image Classification	2624
<i>Jian-Xun Mi, Qiankun Fu, and Weisheng Li</i>	
Generalized Semantic Preserving Hashing for N-Label Cross-Modal Retrieval	2633
<i>Devraj Mandal, Kunal N. Chaudhury, and Soma Biswas</i>	
EAST: An Efficient and Accurate Scene Text Detector	2642
<i>Xinyu Zhou, Cong Yao, He Wen, Yuzhi Wang, Shuchang Zhou, Weiran He, and Jiajun Liang</i>	
VidLoc: A Deep Spatio-Temporal Model for 6-DoF Video-Clip Relocalization	2652
<i>Ronald Clark, Sen Wang, Andrew Markham, Niki Trigoni, and Hongkai Wen</i>	

Improving RANSAC-Based Segmentation through CNN Encapsulation	2661
<i>Dustin Morley and Hassan Foroosh</i>	
Position Tracking for Virtual Reality Using Commodity WiFi	2671
<i>Manikanta Kotaru and Sachin Katti</i>	
Designing Illuminant Spectral Power Distributions for Surface Classification	2682
<i>Henryk Blasinski, Joyce Farrell, and Brian Wandell</i>	
One-Shot Hyperspectral Imaging Using Faced Reflectors	2692
<i>Tsuyoshi Takatani, Takahito Aoto, and Yasuhiro Mukaigawa</i>	
Direct Photometric Alignment by Mesh Deformation	2701
<i>Kaimo Lin, Nianjuan Jiang, Shuaicheng Liu, Loong-Fah Cheong, Minh Do, and Jiangbo Lu</i>	
CNN-Based Patch Matching for Optical Flow with Thresholded Hinge Embedding Loss	2710
<i>Christian Bailer, Kiran Varanasi, and Didier Stricker</i>	
Optical Flow Estimation Using a Spatial Pyramid Network	2720
<i>Anurag Ranjan and Michael J. Black</i>	
Deep Network Flow for Multi-object Tracking	2730
<i>Samuel Schuster, Paul Vernaza, Wongun Choi, and Manmohan Chandraker</i>	
Material Classification Using Frequency- and Depth-Dependent Time-of-Flight Distortion	2740
<i>Kenichiro Tanaka, Yasuhiro Mukaigawa, Takuya Funatomi, Hiroyuki Kubo, Yasuyuki Matsushita, and Yasushi Yagi</i>	
Benchmarking Denoising Algorithms with Real Photographs	2750
<i>Tobias Plötz and Stefan Roth</i>	
A Unified Approach of Multi-scale Deep and Hand-Crafted Features for Defocus Estimation	2760
<i>Jinsun Park, Yu-Wing Tai, Donghyeon Cho, and In So Kweon</i>	
StyleBank: An Explicit Representation for Neural Image Style Transfer	2770
<i>Dongdong Chen, Lu Yuan, Jing Liao, Nenghai Yu, and Gang Hua</i>	
Specular Highlight Removal in Facial Images	2780
<i>Chen Li, Stephen Lin, Kun Zhou, and Katsushi Ikeuchi</i>	
Image Super-Resolution via Deep Recursive Residual Network	2790
<i>Ying Tai, Jian Yang, and Xiaoming Liu</i>	
Deep Image Harmonization	2799
<i>Yi-Hsuan Tsai, Xiaohui Shen, Zhe Lin, Kalyan Sunkavalli, Xin Lu, and Ming-Hsuan Yang</i>	
Learning Deep CNN Denoiser Prior for Image Restoration	2808
<i>Kai Zhang, Wangmeng Zuo, Shuhang Gu, and Lei Zhang</i>	
A Novel Tensor-Based Video Rain Streaks Removal Approach via Utilizing Discriminatively Intrinsic Priors	2818
<i>Tai-Xiang Jiang, Ting-Zhu Huang, Xi-Le Zhao, Liang-Jian Deng, and Yao Wang</i>	
GMS: Grid-Based Motion Statistics for Fast, Ultra-Robust Feature Correspondence	2828
<i>Jiawang Bian, Wen-Yan Lin, Yasuyuki Matsushita, Sai-Kit Yeung, Tan-Dat Nguyen, and Ming-Ming Cheng</i>	

Video Desnowing and Deraining Based on Matrix Decomposition	2838
<i>Weihong Ren, Jiandong Tian, Zhi Han, Antoni Chan, and Yandong Tang</i>	
Real-Time Video Super-Resolution with Spatio-Temporal Networks and Motion Compensation	2848
<i>Jose Caballero, Christian Ledig, Andrew Aitken, Alejandro Acosta, Johannes Totz, Zehan Wang, and Wenzhe Shi</i>	
Deep Watershed Transform for Instance Segmentation	2858
<i>Min Bai and Raquel Urtasun</i>	
AnchorNet: A Weakly Supervised Network to Learn Geometry-Sensitive Features for Semantic Matching	2867
<i>David Novotny, Diane Larlus, and Andrea Vedaldi</i>	
Learning Diverse Image Colorization	2877
<i>Aditya Deshpande, Jiajun Lu, Mao-Chuang Yeh, Min Jin Chong, and David Forsyth</i>	
Awesome Typography: Statistics-Based Text Effects Transfer	2886
<i>Shuai Yang, Jiaying Liu, Zhouhui Lian, and Zongming Guo</i>	
Deep TEN: Texture Encoding Network	2896
<i>Hang Zhang, Jia Xue, and Kristin Dana</i>	
Order-Preserving Wasserstein Distance for Sequence Matching	2906
<i>Bing Su and Gang Hua</i>	
Attend in Groups: A Weakly-Supervised Deep Learning Framework for Learning from Web Data	2915
<i>Bohan Zhuang, Lingqiao Liu, Yao Li, Chunhua Shen, and Ian Reid</i>	
Hierarchical Multimodal Metric Learning for Multimodal Classification	2925
<i>Heng Zhang, Vishal M. Patel, and Rama Chellappa</i>	
Efficient Linear Programming for Dense CRFs	2934
<i>Thalaiyasingam Ajanthan, Alban Desmaison, Rudy Bunel, Mathieu Salzmann, Philip H. S. Torr, and M. Pawan Kumar</i>	
Variational Autoencoded Regression: High Dimensional Regression of Visual Data on Complex Manifold	2943
<i>Youngjoon Yoo, Sangdoon Yun, Hyung Jin Chang, Yiannis Demiris, and Jin Young Choi</i>	
Learning Random-Walk Label Propagation for Weakly-Supervised Semantic Segmentation	2953
<i>Paul Vernaza and Manmohan Chandraker</i>	
Adversarial Discriminative Domain Adaptation	2962
<i>Eric Tzeng, Judy Hoffman, Kate Saenko, and Trevor Darrell</i>	
Low-Rank-Sparse Subspace Representation for Robust Regression	2972
<i>Yongqiang Zhang, Daming Shi, Junbin Gao, and Dansong Cheng</i>	
Unsupervised Video Summarization with Adversarial LSTM Networks	2982
<i>Behrooz Mahasseni, Michael Lam, and Sinisa Todorovic</i>	
Generating the Future with Adversarial Transformers	2992
<i>Carl Vondrick and Antonio Torralba</i>	

Semantic Amodal Segmentation	3001
<i>Yan Zhu, Yuandong Tian, Dimitris Metaxas, and Piotr Dollár</i>	
Learning a Deep Embedding Model for Zero-Shot Learning	3010
<i>Li Zhang, Tao Xiang, and Shaogang Gong</i>	
BIND: Binary Integrated Net Descriptors for Texture-Less Object Recognition	3020
<i>Jacob Chan, Jimmy Addison Lee, and Qian Kemao</i>	
Growing a Brain: Fine-Tuning by Increasing Model Capacity	3029
<i>Yu-Xiong Wang, Deva Ramanan, and Martial Hebert</i>	
A-Fast-RCNN: Hard Positive Generation via Adversary for Object Detection	3039
<i>Xiaolong Wang, Abhinav Shrivastava, and Abhinav Gupta</i>	
Kernel Pooling for Convolutional Neural Networks	3049
<i>Yin Cui, Feng Zhou, Jiang Wang, Xiao Liu, Yuanqing Lin, and Serge Belongie</i>	
Multiple Instance Detection Network with Online Instance Classifier Refinement	3059
<i>Peng Tang, Xinggang Wang, Xiang Bai, and Wenyu Liu</i>	
Learning Cross-Modal Embeddings for Cooking Recipes and Food Images	3068
<i>Amaia Salvador, Nicholas Hynes, Yusuf Aytar, Javier Marin, Ferda Ofli, Ingmar Weber, and Antonio Torralba</i>	
Zero-Shot Learning — The Good, the Bad and the Ugly	3077
<i>Yongqin Xian, Bernt Schiele, and Zeynep Akata</i>	
DeepNav: Learning to Navigate Large Cities	3087
<i>Samarth Brahmbhatt and James Hays</i>	
Scene Graph Generation by Iterative Message Passing	3097
<i>Danfei Xu, Yuke Zhu, Christopher B. Choy, and Li Fei-Fei</i>	
Visual Translation Embedding Network for Visual Relation Detection	3107
<i>Hanwang Zhang, Zawlin Kyaw, Shih-Fu Chang, and Tat-Seng Chua</i>	
Unsupervised Part Learning for Visual Recognition	3116
<i>Ronan Sicre, Yannis Avrithis, Ewa Kijak, and Frédéric Jurie</i>	
Comprehension-Guided Referring Expressions	3125
<i>Ruotian Luo and Gregory Shakhnarovich</i>	
Top-Down Visual Saliency Guided by Captions	3135
<i>Vasili Ramanishka, Abir Das, Jianming Zhang, and Kate Saenko</i>	
Grassmannian Manifold Optimization Assisted Sparse Spectral Clustering	3145
<i>Qiong Wang, Junbin Gao, and Hong Li</i>	
Video Propagation Networks	3154
<i>Varun Jampani, Raghudeep Gadde, and Peter V. Gehler</i>	
ActionVLAD: Learning Spatio-Temporal Aggregation for Action Classification	3165
<i>Rohit Girdhar, Deva Ramanan, Abhinav Gupta, Josef Sivic, and Bryan Russell</i>	
SCC: Semantic Context Cascade for Efficient Action Detection	3175
<i>Fabian Caba Heilbron, Wayner Barrios, Victor Escorcia, and Bernard Ghanem</i>	

Hierarchical Boundary-Aware Neural Encoder for Video Captioning	3185
<i>Lorenzo Baraldi, Costantino Grana, and Rita Cucchiara</i>	
HOPE: Hierarchical Object Prototype Encoding for Efficient Object Instance Search in Videos	3195
<i>Tan Yu, Yuwei Wu, and Junsong Yuan</i>	
Spatio-Temporal Vector of Locally Max Pooled Features for Action Recognition in Videos	3205
<i>Ionut Cosmin Duta, Bogdan Ionescu, Kiyoharu Aizawa, and Nicu Sebe</i>	
Temporal Action Localization by Structured Maximal Sums	3215
<i>Zehuan Yuan, Jonathan C. Stroud, Tong Lu, and Jia Deng</i>	
Predicting Salient Face in Multiple-Face Videos	3224
<i>Yufan Liu, Songyang Zhang, Mai Xu, and Xuming He</i>	

Oral and Spotlight 2-2A

S2-2A: Object Recognition & Scene Understanding 1

Graph-Structured Representations for Visual Question Answering	3233
<i>Damien Teney, Lingqiao Liu, and Anton van den Hengel</i>	
Knowing When to Look: Adaptive Attention via a Visual Sentinel for Image Captioning	3242
<i>Jiasen Lu, Caiming Xiong, Devi Parikh, and Richard Socher</i>	
Learned Contextual Feature Reweighting for Image Geo-Localization	3251
<i>Hyo Jin Kim, Enrique Dunn, and Jan-Michael Frahm</i>	
End-to-End Concept Word Detection for Video Captioning, Retrieval, and Question Answering	3261
<i>Youngjae Yu, Hyungjin Ko, Jongwook Choi, and Gunhee Kim</i>	
Deep Cross-Modal Hashing	3270
<i>Qing-Yuan Jiang and Wu-Jun Li</i>	
Unambiguous Text Localization and Retrieval for Cluttered Scenes	3279
<i>Xuejian Rong, Chucai Yi, and Yingli Tian</i>	
Bayesian Supervised Hashing	3288
<i>Zihao Hu, Junxuan Chen, Hongtao Lu, and Tongzhen Zhang</i>	
Speed/Accuracy Trade-Offs for Modern Convolutional Object Detectors	3296
<i>Jonathan Huang, Vivek Rathod, Chen Sun, Menglong Zhu, Anoop Korattikara, Alireza Fathi, Ian Fischer, Zbigniew Wojna, Yang Song, Sergio Guadarrama, and Kevin Murphy</i>	

O2-2A

Detecting Visual Relationships with Deep Relational Networks	3298
<i>Bo Dai, Yuqi Zhang, and Dahua Lin</i>	
Full-Resolution Residual Networks for Semantic Segmentation in Street Scenes	3309
<i>Tobias Pohlen, Alexander Hermans, Markus Mathias, and Bastian Leibe</i>	

Network Dissection: Quantifying Interpretability of Deep Visual Representations	3319
<i>David Bau, Bolei Zhou, Aditya Khosla, Aude Oliva, and Antonio Torralba</i>	
AGA: Attribute-Guided Augmentation	3328
<i>Mandar Dixit, Roland Kwitt, Marc Niethammer, and Nuno Vasconcelos</i>	

Oral and Spotlight 2-2B

S2-2B: Analyzing Humans 2

A Hierarchical Approach for Generating Descriptive Image Paragraphs	3337
<i>Jonathan Krause, Justin Johnson, Ranjay Krishna, and Li Fei-Fei</i>	
Person Re-identification in the Wild	3346
<i>Liang Zheng, Hengheng Zhang, Shaoyan Sun, Manmohan Chandraker, Yi Yang, and Qi Tian</i>	
Scalable Person Re-identification on Supervised Smoothed Manifold	3356
<i>Song Bai, Xiang Bai, and Qi Tian</i>	
Binge Watching: Scaling Affordance Learning from Sitcoms	3366
<i>Xiaolong Wang, Rohit Girdhar, and Abhinav Gupta</i>	
Joint Detection and Identification Feature Learning for Person Search	3376
<i>Tong Xiao, Shuang Li, Bochao Wang, Liang Lin, and Xiaogang Wang</i>	
Synthesizing Normalized Faces from Facial Identity Features	3386
<i>Forrester Cole, David Belanger, Dilip Krishnan, Aaron Sarna, Inbar Mosseri, and William T. Freeman</i>	
Consistent-Aware Deep Learning for Person Re-identification in a Camera Network	3396
<i>Ji Lin, Liangliang Ren, Jiwen Lu, Jianjiang Feng, and Jie Zhou</i>	
Level Playing Field for Million Scale Face Recognition	3406
<i>Aaron Nech and Ira Kemelmacher-Shlizerman</i>	

O2-2B

Re-Sign: Re-Aligned End-to-End Sequence Modelling with Deep Recurrent CNN-HMMs	3416
<i>Oscar Koller, Sepehr Zargaran, and Hermann Ney</i>	
Social Scene Understanding: End-to-End Multi-person Action Localization and Collective Activity Recognition	3425
<i>Timur Bagautdinov, Alexandre Alahi, François Fleuret, Pascal Fua, and Silvio Savarese</i>	
Detangling People: Individuating Multiple Close People and Their Body Parts via Region Assembly	3435
<i>Hao Jiang and Kristen Grauman</i>	
Lip Reading Sentences in the Wild	3444
<i>Joon Son Chung, Andrew Senior, Oriol Vinyals, and Andrew Zisserman</i>	

Oral and Spotlight 2-2C

S2-2C: Applications

Deep Matching Prior Network: Toward Tighter Multi-oriented Text Detection	3454
<i>Yuliang Liu and Lianwen Jin</i>	
ChestX-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases	3462
<i>Xiaosong Wang, Yifan Peng, Le Lu, Zhiyong Lu, Mohammadhadi Bagheri, and Ronald M. Summers</i>	
Attentional Push: A Deep Convolutional Network for Augmenting Image Saliency with Shared Attention Modeling in Social Scenes	3472
<i>Siavash Gorji and James J. Clark</i>	
Detecting Oriented Text in Natural Images by Linking Segments	3482
<i>Baoguang Shi, Xiang Bai, and Serge Belongie</i>	
Learning Video Object Segmentation from Static Images	3491
<i>Federico Perazzi, Anna Khoreva, Rodrigo Benenson, Bernt Schiele, and Alexander Sorkine-Hornung</i>	
Seeing Invisible Poses: Estimating 3D Body Pose from Egocentric Video	3501
<i>Hao Jiang and Kristen Grauman</i>	
Plug & Play Generative Networks: Conditional Iterative Generation of Images in Latent Space	3510
<i>Anh Nguyen, Jeff Clune, Yoshua Bengio, Alexey Dosovitskiy, and Jason Yosinski</i>	
A Joint Speaker-Listener-Reinforcer Model for Referring Expressions	3521
<i>Licheng Yu, Hao Tan, Mohit Bansal, and Tamara L. Berg</i>	

O2-2C

End-to-End Learning of Driving Models from Large-Scale Video Datasets	3530
<i>Huazhe Xu, Yang Gao, Fisher Yu, and Trevor Darrell</i>	
Deep Future Gaze: Gaze Anticipation on Egocentric Videos Using Adversarial Networks	3539
<i>Mengmi Zhang, Keng Teck Ma, Joo Hwee Lim, Qi Zhao, and Jiashi Feng</i>	
MDNet: A Semantically and Visually Interpretable Medical Image Diagnosis Network	3549
<i>Zizhao Zhang, Yuanpu Xie, Fuyong Xing, Mason McGough, and Lin Yang</i>	

Poster Session P2-2

Surface Motion Capture Transfer with Gaussian Process Regression	3558
<i>Adnane Boukhayma, Jean-Sébastien Franco, and Edmond Boyer</i>	
Visual-Inertial-Semantic Scene Representation for 3D Object Detection	3567
<i>Jingming Dong, Xiaohan Fei, and Stefano Soatto</i>	

Template-Based Monocular 3D Recovery of Elastic Shapes Using Lagrangian Multipliers	3578
<i>Nazim Haouchine and Stephane Cotin</i>	
Learning Category-Specific 3D Shape Models from Weakly Labeled 2D Images	3587
<i>Dingwen Zhang, Junwei Han, Yang Yang, and Dong Huang</i>	
Simultaneous Geometric and Radiometric Calibration of a Projector-Camera Pair	3596
<i>Marjan Shahpaski, Luis Ricardo Sapaico, Gaspard Chevassus, and Sabine Süssstrunk</i>	
A Clever Elimination Strategy for Efficient Minimal Solvers	3605
<i>Zuzana Kukelova, Joe Kileel, Bernd Sturmfels, and Tomas Pajdla</i>	
Learning Barycentric Representations of 3D Shapes for Sketch-Based 3D Shape Retrieval	3615
<i>Jin Xie, Guoxian Dai, Fan Zhu, and Yi Fang</i>	
Geodesic Distance Descriptors	3624
<i>Gil Shamaï and Ron Kimmel</i>	
Modeling Temporal Dynamics and Spatial Configurations of Actions Using Two-Stream Recurrent Neural Networks	3633
<i>Hongsong Wang and Liang Wang</i>	
Forecasting Human Dynamics from Static Images	3643
<i>Yu-Wei Chao, Jimei Yang, Brian Price, Scott Cohen, and Jia Deng</i>	
Re-ranking Person Re-identification with k-Reciprocal Encoding	3652
<i>Zhun Zhong, Liang Zheng, Donglin Cao, and Shaozi Li</i>	
Deep Sequential Context Networks for Action Prediction	3662
<i>Yu Kong, Zhiqiang Tao, and Yun Fu</i>	
Global Context-Aware Attention LSTM Networks for 3D Action Recognition	3671
<i>Jun Liu, Gang Wang, Ping Hu, Ling-Yu Duan, and Alex C. Kot</i>	
Dynamic Attention-Controlled Cascaded Shape Regression Exploiting Training Data Augmentation and Fuzzy-Set Sample Weighting	3681
<i>Zhen-Hua Feng, Josef Kittler, William Christmas, Patrik Huber, and Xiao-Jun Wu</i>	
A Deep Regression Architecture with Two-Stage Re-initialization for High Performance Facial Landmark Detection	3691
<i>Jiangjing Lv, Xiaohu Shao, Junliang Xing, Cheng Cheng, and Xi Zhou</i>	
Multiple People Tracking by Lifted Multicut and Person Re-identification	3701
<i>Siyu Tang, Mykhaylo Andriluka, Bjoern Andres, and Bernt Schiele</i>	
Towards Accurate Multi-person Pose Estimation in the Wild	3711
<i>George Papandreou, Tyler Zhu, Nori Kanazawa, Alexander Toshev, Jonathan Tompson, Chris Bregler, and Kevin Murphy</i>	
Towards a Quality Metric for Dense Light Fields	3720
<i>Vamsi Kiran Adhikarla, Marek Vinkler, Denis Sumin, Rafal K. Mantiuk, Karol Myszkowski, Hans-Peter Seidel, and Piotr Didyk</i>	
Controlling Perceptual Factors in Neural Style Transfer	3730
<i>Leon A. Gatys, Alexander S. Ecker, Matthias Bethge, Aaron Hertzmann, and Eli Shechtman</i>	

Joint Sequence Learning and Cross-Modality Convolution for 3D Biomedical Segmentation	3739
<i>Kuan-Lun Tseng, Yen-Liang Lin, Winston Hsu, and Chung-Yang Huang</i>	
LSTM Self-Supervision for Detailed Behavior Analysis	3747
<i>Biagio Brattoli, Uta Büchler, Anna-Sophia Wahl, Martin E. Schwab, and Björn Ommer</i>	
A Wide-Field-of-View Monocentric Light Field Camera	3757
<i>Donald G. Dansereau, Glenn Schuster, Joseph Ford, and Gordon Wetzstein</i>	
S2F: Slow-to-Fast Interpolator Flow	3767
<i>Yanchao Yang and Stefano Soatto</i>	
CLKN: Cascaded Lucas-Kanade Networks for Image Alignment	3777
<i>Che-Han Chang, Chun-Nan Chou, and Edward Y. Chang</i>	
Multi-object Tracking with Quadruplet Convolutional Neural Networks	3786
<i>Jeany Son, Mooyeol Baek, Minsu Cho, and Bohyung Han</i>	
Learning to Detect Salient Objects with Image-Level Supervision	3796
<i>Lijun Wang, Huchuan Lu, Yifan Wang, Mengyang Feng, Dong Wang, Baocai Yin, and Xiang Ruan</i>	
From Motion Blur to Motion Flow: A Deep Learning Solution for Removing Heterogeneous Motion Blur	3806
<i>Dong Gong, Jie Yang, Lingqiao Liu, Yanning Zhang, Ian Reid, Chunhua Shen, Anton Van Den Hengel, and Qinfeng Shi</i>	
Co-occurrence Filter	3816
<i>Roy J. Jevnisek and Shai Avidan</i>	
Fractal Dimension Invariant Filtering and Its CNN-Based Implementation	3825
<i>Hongteng Xu, Junchi Yan, Nils Persson, Weiyao Lin, and Hongyuan Zha</i>	
Noise-Blind Image Deblurring	3834
<i>Meiguang Jin, Stefan Roth, and Paolo Favaro</i>	
Simultaneous Visual Data Completion and Denoising Based on Tensor Rank and Total Variation Minimization and Its Primal-Dual Splitting Algorithm	3843
<i>Tatsuya Yokota and Hidekata Hontani</i>	
HPatches: A Benchmark and Evaluation of Handcrafted and Learned Local Descriptors	3852
<i>Vassileios Balntas, Karel Lenc, Andrea Vedaldi, and Krystian Mikolajczyk</i>	
Hyperspectral Image Super-Resolution via Non-local Sparse Tensor Factorization	3862
<i>Renwei Dian, Leyuan Fang, and Shutao Li</i>	
Reflection Removal Using Low-Rank Matrix Completion	3872
<i>Byeong-Ju Han and Jae-Young Sim</i>	
Object Co-skeletonization with Co-segmentation	3881
<i>Koteswar Rao Jerripothula, Jianfei Cai, Jiangbo Lu, and Junsong Yuan</i>	
Mining Object Parts from CNNs via Active Question-Answering	3890
<i>Quanshi Zhang, Ruiming Cao, Ying Nian Wu, and Song-Chun Zhu</i>	
PolyNet: A Pursuit of Structural Diversity in Very Deep Networks	3900
<i>Xingcheng Zhang, Zhizhong Li, Chen Change Loy, and Dahua Lin</i>	

The VQA-Machine: Learning How to Use Existing Vision Algorithms to Answer New Questions	3909
<i>Peng Wang, Qi Wu, Chunhua Shen, and Anton van den Hengel</i>	
Joint Discriminative Bayesian Dictionary and Classifier Learning	3919
<i>Naveed Akhtar, Ajmal Mian, and Fatih Porikli</i>	
Quad-Networks: Unsupervised Learning to Rank for Interest Point Detection	3929
<i>Nikolay Savinov, Akihito Seki, L'Ubor Ladický, Torsten Sattler, and Marc Pollefeys</i>	
Outlier-Robust Tensor PCA	3938
<i>Pan Zhou and Jiashi Feng</i>	
Learning Adaptive Receptive Fields for Deep Image Parsing Network	3947
<i>Zhen Wei, Yao Sun, Jinqiao Wang, Hanjiang Lai, and Si Liu</i>	
Learning an Invariant Hilbert Space for Domain Adaptation	3956
<i>Samitha Herath, Mehrtash Harandi, and Fatih Porikli</i>	
Fixed-Point Factorized Networks	3966
<i>Peisong Wang and Jian Cheng</i>	
Discriminative Optimization: Theory and Applications to Point Cloud Registration	3975
<i>Jayakorn Vongkulbhisal, Fernando De la Torre, and João P. Costeira</i>	
Online Asymmetric Similarity Learning for Cross-Modal Retrieval	3984
<i>Yiling Wu, Shuhui Wang, and Qingming Huang</i>	
Improving Training of Deep Neural Networks via Singular Value Bounding	3994
<i>Kui Jia, Dacheng Tao, Shenghua Gao, and Xiangmin Xu</i>	
S3Pool: Pooling with Stochastic Spatial Sampling	4003
<i>Shuangfei Zhai, Hui Wu, Abhishek Kumar, Yu Cheng, Yongxi Lu, Zhongfei Zhang, and Rogerio Feris</i>	
Sports Field Localization via Deep Structured Models	4012
<i>Namdar Homayounfar, Sanja Fidler, and Raquel Urtasun</i>	
Noisy Softmax: Improving the Generalization Ability of DCNN via Postponing the Early Softmax Saturation	4021
<i>Binghui Chen, Weihong Deng, and Junping Du</i>	
Switching Convolutional Neural Network for Crowd Counting	4031
<i>Deepak Babu Sam, Shiv Surya, and R. Venkatesh Babu</i>	
Network Sketching: Exploiting Binary Structure in Deep CNNs	4040
<i>Yiwen Guo, Anbang Yao, Hao Zhao, and Yurong Chen</i>	
Multi-task Clustering of Human Actions by Sharing Information	4049
<i>Xiaoqiang Yan, Shizhe Hu, and Yangdong Ye</i>	
Soft-Margin Mixture of Regressions	4058
<i>Dong Huang, Longfei Han, and Fernando De la Torre</i>	
Multigrid Neural Architectures	4067
<i>Tsung-Wei Ke, Michael Maire, and Stella X. Yu</i>	
High-Resolution Image Inpainting Using Multi-scale Neural Patch Synthesis	4076
<i>Chao Yang, Xin Lu, Zhe Lin, Eli Shechtman, Oliver Wang, and Hao Li</i>	

Deep Quantization: Encoding Convolutional Activations with Deep Generative Model	4085
<i>Zhaofan Qiu, Ting Yao, and Tao Mei</i>	
DOPE: Distributed Optimization for Pairwise Energies	4095
<i>Jose Dolz, Ismail Ben Ayed, and Christian Desrosiers</i>	
Improved Texture Networks: Maximizing Quality and Diversity in Feed-Forward Stylization and Texture Synthesis	4105
<i>Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky</i>	
Polyhedral Conic Classifiers for Visual Object Detection and Classification	4114
<i>Hakan Cevikalp and Bill Triggs</i>	
Incremental Kernel Null Space Discriminant Analysis for Novelty Detection	4123
<i>Juncheng Liu, Zhouhui Lian, Yi Wang, and Jianguo Xiao</i>	
Predicting Ground-Level Scene Layout from Aerial Imagery	4132
<i>Menghua Zhai, Zachary Bessinger, Scott Workman, and Nathan Jacobs</i>	
Deep Feature Flow for Video Recognition	4141
<i>Xizhou Zhu, Yuwen Xiong, Jifeng Dai, Lu Yuan, and Yichen Wei</i>	
Object-Aware Dense Semantic Correspondence	4151
<i>Fan Yang, Xin Li, Hong Cheng, Jianping Li, and Leiting Chen</i>	
Semantic Regularisation for Recurrent Image Annotation	4160
<i>Feng Liu, Tao Xiang, Timothy M. Hospedales, Wankou Yang, and Changyin Sun</i>	
Video2Shop: Exact Matching Clothes in Videos to Online Shopping Images	4169
<i>Zhi-Qi Cheng, Xiao Wu, Yang Liu, and Xian-Sheng Hua</i>	
Fast-At: Fast Automatic Thumbnail Generation Using Deep Neural Networks	4178
<i>Seyed A. Esmaeili, Bharat Singh, and Larry S. Davis</i>	
Multi-level Attention Networks for Visual Question Answering	4187
<i>Dongfei Yu, Jianlong Fu, Tao Mei, and Yong Rui</i>	
Generating Descriptions with Grounded and Co-referenced People	4196
<i>Anna Rohrbach, Marcus Rohrbach, Siyu Tang, Seong Joon Oh, and Bernt Schiele</i>	
Straight to Shapes: Real-Time Detection of Encoded Shapes	4207
<i>Saumya Jetley, Michael Sapienza, Stuart Golodetz, and Philip H. S. Torr</i>	
Simultaneous Feature Aggregating and Hashing for Large-Scale Image Search	4217
<i>Thanh-Toan Do, Dang-Khoa Le Tan, Trung T. Pham, and Ngai-Man Cheung</i>	
Improving Facial Attribute Prediction Using Semantic Segmentation	4227
<i>Mahdi M. Kalayeh, Boqing Gong, and Mubarak Shah</i>	
Learning Cross-Modal Deep Representations for Robust Pedestrian Detection	4236
<i>Dan Xu, Wanli Ouyang, Elisa Ricci, Xiaogang Wang, and Nicu Sebe</i>	
Spatio-Temporal Self-Organizing Map Deep Network for Dynamic Object Detection from Videos	4245
<i>Yang Du, Chunfeng Yuan, Bing Li, Weiming Hu, and Stephen Maybank</i>	
CERN: Confidence-Energy Recurrent Network for Group Activity Recognition	4255
<i>Tianmin Shu, Sinisa Todorovic, and Song-Chun Zhu</i>	

Understanding Traffic Density from Large-Scale Web Camera Data	4264
<i>Shanghang Zhang, Guanhang Wu, João P. Costeira, and José M. F. Moura</i>	
Collaborative Summarization of Topic-Related Videos	4274
<i>Rameswar Panda and Amit K. Roy-Chowdhury</i>	

Oral and Spotlight 3-1A

S3-1A: Machine Learning 3

Local Binary Convolutional Neural Networks	4284
<i>Felix Juefei-Xu, Vishnu Naresh Boddeti, and Marios Savvides</i>	
Deep Self-Taught Learning for Weakly Supervised Object Localization	4294
<i>Zequn Jie, Yunchao Wei, Xiaojie Jin, Jiashi Feng, and Wei Liu</i>	
Multi-modal Mean-Fields via Cardinality-Based Clamping	4303
<i>Pierre Baqué, François Fleuret, and Pascal Fua</i>	
Probabilistic Temporal Subspace Clustering	4313
<i>Behnam Gholami and Vladimir Pavlovic</i>	
Provable Self-Representation Based Outlier Detection in a Union of Subspaces	4323
<i>Chong You, Daniel P. Robinson, and René Vidal</i>	
Latent Multi-view Subspace Clustering	4333
<i>Changqing Zhang, Qinghua Hu, Huazhu Fu, Pengfei Zhu, and Xiaochun Cao</i>	
Learning to Extract Semantic Structure from Documents Using Multimodal Fully Convolutional Neural Networks	4342
<i>Xiao Yang, Ersin Yumer, Paul Asente, Mike Kraley, Daniel Kifer, and C. Lee Giles</i>	
Age Progression/Regression by Conditional Adversarial Autoencoder	4352
<i>Zhifei Zhang, Yang Song, and Hairong Qi</i>	

O3-1A

Compact Matrix Factorization with Dependent Subspaces	4361
<i>Viktor Larsson and Carl Olsson</i>	
FFTLasso: Large-Scale LASSO in the Fourier Domain	4371
<i>Adel Bibi, Hani Itani, and Bernard Ghanem</i>	
On the Global Geometry of Sphere-Constrained Sparse Blind Deconvolution	4381
<i>Yuqian Zhang, Yenson Lau, Han-Wen Kuo, Sky Cheung, Abhay Pasupathy, and John Wright</i>	
Global Optimality in Neural Network Training	4390
<i>Benjamin D. Haeffele and René Vidal</i>	

Oral and Spotlight 3-1B

S3-1B: Object Recognition & Scene Understanding 2

What is and What is Not a Salient Object? Learning Salient Object Detector by Ensembling Linear Exemplar Regressors	4399
<i>Changqun Xia, Jia Li, Xiaowu Chen, Anlin Zheng, and Yu Zhang</i>	
Deep Variation-Structured Reinforcement Learning for Visual Relationship and Attribute Detection	4408
<i>Xiaodan Liang, Lisa Lee, and Eric P. Xing</i>	
Modeling Relationships in Referential Expressions with Compositional Modular Networks	4418
<i>Ronghang Hu, Marcus Rohrbach, Jacob Andreas, Trevor Darrell, and Kate Saenko</i>	
Counting Everyday Objects in Everyday Scenes	4428
<i>Prithvijit Chattopadhyay, Ramakrishna Vedantam, Ramprasaath R. Selvaraju, Dhruv Batra, and Devi Parikh</i>	
Fully Convolutional Instance-Aware Semantic Segmentation	4438
<i>Yi Li, Haozhi Qi, Jifeng Dai, Xiangyang Ji, and Yichen Wei</i>	
Semantic Autoencoder for Zero-Shot Learning	4447
<i>Elyor Kodirov, Tao Xiang, and Shaogang Gong</i>	
CityPersons: A Diverse Dataset for Pedestrian Detection	4457
<i>Shanshan Zhang, Rodrigo Benenson, and Bernt Schiele</i>	
GuessWhat?! Visual Object Discovery through Multi-modal Dialogue	4466
<i>Harm de Vries, Florian Strub, Sarath Chandar, Olivier Pietquin, Hugo Larochelle, and Aaron Courville</i>	

O3-1B

Look Closer to See Better: Recurrent Attention Convolutional Neural Network for Fine-Grained Image Recognition	4476
<i>Jianlong Fu, Heliang Zheng, and Tao Mei</i>	
Annotating Object Instances with a Polygon-RNN	4485
<i>Lluís Castrejón, Kaustav Kundu, Raquel Urtasun, and Sanja Fidler</i>	
Connecting Look and Feel: Associating the Visual and Tactile Properties of Physical Materials	4494
<i>Wenzhen Yuan, Shaoxiong Wang, Siyuan Dong, and Edward Adelson</i>	
Deep Learning Human Mind for Automated Visual Classification	4503
<i>C. Spampinato, S. Palazzo, I. Kavasidis, D. Giordano, N. Souly, and M. Shah</i>	

Poster Session P3-1

Self-Calibration-Based Approach to Critical Motion Sequences of Rolling-Shutter Structure from Motion	4512
<i>Eisuke Ito and Takayuki Okatani</i>	
Semi-Calibrated Near Field Photometric Stereo	4521
<i>Fotios Logothetis, Roberto Mecca, and Roberto Cipolla</i>	
Semantic Multi-view Stereo: Jointly Estimating Objects and Voxels	4531
<i>Ali Osman Ulusoy, Michael J. Black, and Andreas Geiger</i>	
Learning to Predict Stereo Reliability Enforcing Local Consistency of Confidence Maps	4541
<i>Matteo Poggi and Stefano Mattoccia</i>	
The Misty Three Point Algorithm for Relative Pose	4551
<i>Tobias Palmér, Kalle Åström, and Jan-Michael Frahm</i>	
The Surfacing of Multiview 3D Drawings via Lofting and Occlusion Reasoning	4560
<i>Anil Usumezbas, Ricardo Fabbri, and Benjamin B. Kimia</i>	
A New Representation of Skeleton Sequences for 3D Action Recognition	4570
<i>Qihong Ke, Mohammed Bennamoun, Senjian An, Ferdous Sohel, and Farid Boussaid</i>	
A General Framework for Curve and Surface Comparison and Registration with Oriented Varifolds	4580
<i>Irène Kaltenmark, Benjamin Charlier, and Nicolas Charon</i>	
Learning to Align Semantic Segmentation and 2.5D Maps for Geolocalization	4590
<i>Anil Armagan, Martin Hirzer, Peter M. Roth, and Vincent Lepetit</i>	
A Generative Model for Depth-Based Robust 3D Facial Pose Tracking	4598
<i>Lu Sheng, Jianfei Cai, Tat-Jen Cham, Vladimir Pavlovic, and King Ngi Ngan</i>	
Fast 3D Reconstruction of Faces with Glasses	4608
<i>Fabio Maninchedda, Martin R. Oswald, and Marc Pollefeys</i>	
An Efficient Algebraic Solution to the Perspective-Three-Point Problem	4618
<i>Tong Ke and Stergios I. Roumeliotis</i>	
Learning from Synthetic Humans	4627
<i>Gül Varol, Javier Romero, Xavier Martin, Naureen Mahmood, Michael J. Black, Ivan Laptev, and Cordelia Schmid</i>	
Forecasting Interactive Dynamics of Pedestrians with Fictitious Play	4636
<i>Wei-Chiu Ma, De-An Huang, Namhoon Lee, and Kris M. Kitani</i>	
Hand Keypoint Detection in Single Images Using Multiview Bootstrapping	4645
<i>Tomas Simon, Hanbyul Joo, Iain Matthews, and Yaser Sheikh</i>	
PoseTrack: Joint Multi-person Pose Estimation and Tracking	4654
<i>Umar Iqbal, Anton Milan, and Juergen Gall</i>	
Expecting the Unexpected: Training Detectors for Unusual Pedestrians with Adversarial Imposters	4664
<i>Shiyu Huang and Deva Ramanan</i>	

On Human Motion Prediction Using Recurrent Neural Networks	4674
<i>Julieta Martinez, Michael J. Black, and Javier Romero</i>	
Learning and Refining of Privileged Information-Based RNNs for Action Recognition from Depth Sequences	4684
<i>Zhiyuan Shi and Tae-Kyun Kim</i>	
Quality Aware Network for Set to Set Recognition	4694
<i>Yu Liu, Junjie Yan, and Wanli Ouyang</i>	
Unite the People: Closing the Loop Between 3D and 2D Human Representations	4704
<i>Christoph Lassner, Javier Romero, Martin Kiefel, Federica Bogo, Michael J. Black, and Peter V. Gehler</i>	
Deep Multitask Architecture for Integrated 2D and 3D Human Sensing	4714
<i>Alin-Ionut Popa, Mihai Zanfir, and Cristian Sminchisescu</i>	
Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset	4724
<i>João Carreira and Andrew Zisserman</i>	
Identifying First-Person Camera Wearers in Third-Person Videos	4734
<i>Chenyou Fan, Jangwon Lee, Mingze Xu, Krishna Kumar Singh, Yong Jae Lee, David J. Crandall, and Michael S. Ryoo</i>	
Learning to Rank Retargeted Images	4743
<i>Yang Chen, Yong-Jin Liu, and Yu-Kun Lai</i>	
Parsing Images of Overlapping Organisms with Deep Singling-Out Networks	4752
<i>Victor Yurchenko and Victor Lempitsky</i>	
Fine-Tuning Convolutional Neural Networks for Biomedical Image Analysis: Actively and Incrementally	4761
<i>Zongwei Zhou, Jae Shin, Lei Zhang, Suryakanth Gurudu, Michael Gotway, and Jianming Liang</i>	
Depth from Defocus in the Wild	4773
<i>Huixuan Tang, Scott Cohen, Brian Price, Stephen Schiller, and Kiriakos N. Kutulakos</i>	
Matting and Depth Recovery of Thin Structures Using a Focal Stack	4782
<i>Chao Liu, Srinivasa G. Narasimhan, and Artur W. Dubrawski</i>	
Robust Interpolation of Correspondences for Large Displacement Optical Flow	4791
<i>Yinlin Hu, Yunsong Li, and Rui Song</i>	
Large Margin Object Tracking with Circulant Feature Maps	4800
<i>Mengmeng Wang, Yong Liu, and Zeyi Huang</i>	
Minimum Delay Moving Object Detection	4809
<i>Dong Lao and Ganesh Sundaramoorthi</i>	
Multi-task Correlation Particle Filter for Robust Object Tracking	4819
<i>Tianzhu Zhang, Changsheng Xu, and Ming-Hsuan Yang</i>	
Attentional Correlation Filter Network for Adaptive Visual Tracking	4828
<i>Jongwon Choi, Hyung Jin Chang, Sangdoon Yun, Tobias Fischer, Yiannis Demiris, and Jin Young Choi</i>	

The World of Fast Moving Objects	4838
<i>Denys Rozumnyi, Jan Kotera, Filip Šroubek, Lukáš Novotný, and Jirí Matas</i>	
Discriminative Correlation Filter with Channel and Spatial Reliability	4847
<i>Alan Lukežić, Tomáš Vojir, Luka Cehovin Zajc, Jirí Matas, and Matej Kristan</i>	
Learning Deep Binary Descriptor with Multi-quantization	4857
<i>Yueqi Duan, Jiwen Lu, Ziwei Wang, Jianjiang Feng, and Jie Zhou</i>	
One-To-Many Network for Visually Pleasing Compression Artifacts Reduction	4867
<i>Jun Guo and Hongyang Chao</i>	
Gated Feedback Refinement Network for Dense Image Labeling	4877
<i>Md Amirul Islam, Mrigank Rochan, Neil D. B. Bruce, and Yang Wang</i>	
BRISKS: Binary Features for Spherical Images on a Geodesic Grid	4886
<i>Hao Guan and William A. P. Smith</i>	
Superpixels and Polygons Using Simple Non-iterative Clustering	4895
<i>Radhakrishna Achanta and Sabine Süsstrunk</i>	
Hardware-Efficient Guided Image Filtering for Multi-label Problem	4905
<i>Longquan Dai, Mengke Yuan, Zechao Li, Xiaopeng Zhang, and Jinhui Tang</i>	
Alternating Direction Graph Matching	4914
<i>D. Khuê Lê-Huu and Nikos Paragios</i>	
Learning Discriminative and Transformation Covariant Local Feature Detectors	4923
<i>Xu Zhang, Felix X. Yu, Svebor Karaman, and Shih-Fu Chang</i>	
Correlational Gaussian Processes for Cross-Domain Visual Recognition	4932
<i>Chengjiang Long and Gang Hua</i>	
DeLiGAN: Generative Adversarial Networks for Diverse and Limited Data	4941
<i>Swaminathan Gurumurthy, Ravi Kiran Sarvadevabhatla, and R. Venkatesh Babu</i>	
A Dual Ascent Framework for Lagrangean Decomposition of Combinatorial Problems	4950
<i>Paul Swoboda, Jan Kuske, and Bogdan Savchynskyy</i>	
Oriented Response Networks	4961
<i>Yanzhao Zhou, Qixiang Ye, Qiang Qiu, and Jianbin Jiao</i>	
Missing Modalities Imputation via Cascaded Residual Autoencoder	4971
<i>Luan Tran, Xiaoming Liu, Jiayu Zhou, and Rong Jin</i>	
Efficient Optimization for Hierarchically-Structured Interacting Segments (HINTS)	4981
<i>Hossam Isack, Olga Veksler, Ipek Oguz, Milan Sonka, and Yuri Boykov</i>	
A Message Passing Algorithm for the Minimum Cost Multicut Problem	4990
<i>Paul Swoboda and Bjoern Andres</i>	
End-to-End Representation Learning for Correlation Filter Based Tracking	5000
<i>Jack Valmadre, Luca Bertinetto, João Henriques, Andrea Vedaldi, and Philip H. S. Torr</i>	
Filter Flow Made Practical: Massively Parallel and Lock-Free	5009
<i>Sathya N. Ravi, Yunyang Xiong, Lopamudra Mukherjee, and Vikas Singh</i>	
Online Graph Completion: Multivariate Signal Recovery in Computer Vision	5019
<i>Won Hwa Kim, Mona Jalal, Seongjae Hwang, Sterling C. Johnson, and Vikas Singh</i>	

Point to Set Similarity Based Deep Feature Learning for Person Re-Identification	5028
<i>Sanping Zhou, Jinjun Wang, Jiayun Wang, Yihong Gong, and Nanning Zheng</i>	
Exploiting Saliency for Object Segmentation from Image Level Labels	5038
<i>Seong Joon Oh, Rodrigo Benenson, Anna Khoreva, Zeynep Akata, Mario Fritz, and Bernt Schiele</i>	
Consensus Maximization with Linear Matrix Inequality Constraints	5048
<i>Pablo Speciale, Danda Pani Paudel, Martin R. Oswald, Till Kroeger, Luc Van Gool, and Marc Pollefeys</i>	
Physically-Based Rendering for Indoor Scene Understanding Using Convolutional Neural Networks	5057
<i>Yinda Zhang, Shuran Song, Ersin Yumer, Manolis Savva, Joon-Young Lee, Hailin Jin, and Thomas Funkhouser</i>	
Deep Multimodal Representation Learning from Temporal Data	5066
<i>Xitong Yang, Palghat Ramesh, Radha Chitta, Sriganesh Madhvanath, Edgar A. Bernal, and Jiebo Luo</i>	
All You Need is Beyond a Good Init: Exploring Better Solution for Training Extremely Deep Convolutional Neural Networks with Orthonormality and Modulation	5075
<i>Di Xie, Jiang Xiong, and Shiliang Pu</i>	
Hard Mixtures of Experts for Large Scale Weakly Supervised Vision	5085
<i>Sam Gross, Marc'Aurelio Ranzato, and Arthur Szlam</i>	
A Reinforcement Learning Approach to the View Planning Problem	5094
<i>Mustafa Devrim Kaba, Mustafa Gokhan Uzunbas, and Ser Nam Lim</i>	
Zero-Shot Classification with Discriminative Semantic Representation Learning	5103
<i>Meng Ye and Yuhong Guo</i>	
Automatic Discovery, Association Estimation and Learning of Semantic Attributes for a Thousand Categories	5112
<i>Ziad Al-Halah and Rainer Stiefelhagen</i>	
Scene Parsing through ADE20K Dataset	5122
<i>Bolei Zhou, Hang Zhao, Xavier Puig, Sanja Fidler, Adela Barriuso, and Antonio Torralba</i>	
Weakly Supervised Cascaded Convolutional Networks	5131
<i>Ali Diba, Vivek Sharma, Ali Pazandeh, Hamed Pirsiavash, and Luc Van Gool</i>	
Discretely Coding Semantic Rank Orders for Supervised Image Hashing	5140
<i>Li Liu, Ling Shao, Fumin Shen, and Mengyang Yu</i>	
Joint Geometrical and Statistical Alignment for Visual Domain Adaptation	5150
<i>Jing Zhang, Wanqing Li, and Philip Ogunbona</i>	
Weakly Supervised Dense Video Captioning	5159
<i>Zhiqiang Shen, Jianguo Li, Zhou Su, Minjun Li, Yurong Chen, Yu-Gang Jiang, and Xiangyang Xue</i>	
RefineNet: Multi-path Refinement Networks for High-Resolution Semantic Segmentation	5168
<i>Guosheng Lin, Anton Milan, Chunhua Shen, and Ian Reid</i>	

Semantic Segmentation via Structured Patch Prediction, Context CRF and Guidance CRF	5178
<i>Falong Shen, Rui Gan, Shuicheng Yan, and Gang Zeng</i>	
Person Search with Natural Language Description	5187
<i>Shuang Li, Tong Xiao, Hongsheng Li, Bolei Zhou, Dayu Yue, and Xiaogang Wang</i>	
Weakly Supervised Affordance Detection	5197
<i>Johann Sawatzky, Abhilash Srikantha, and Juergen Gall</i>	
Zero-Shot Recognition Using Dual Visual-Semantic Mapping Paths	5207
<i>Yanan Li, Donghui Wang, Huanhang Hu, Yuetan Lin, and Yueting Zhuang</i>	
Neural Aggregation Network for Video Face Recognition	5216
<i>Jiaolong Yang, Peiran Ren, Dongqing Zhang, Dong Chen, Fang Wen, Hongdong Li, and Gang Hua</i>	
Relationship Proposal Networks	5226
<i>Ji Zhang, Mohamed Elhoseiny, Scott Cohen, Walter Chang, and Ahmed Elgammal</i>	
Learning Object Interactions and Descriptions for Semantic Image Segmentation	5235
<i>Guangrun Wang, Ping Luo, Liang Lin, and Xiaogang Wang</i>	
RON: Reverse Connection with Objectness Prior Networks for Object Detection	5244
<i>Tao Kong, Fuchun Sun, Anbang Yao, Huaping Liu, Ming Lu, and Yurong Chen</i>	
Weakly-Supervised Visual Grounding of Phrases with Linguistic Structures	5253
<i>Fanyi Xiao, Leonid Sigal, and Yong Jae Lee</i>	
Incorporating Copying Mechanism in Image Captioning for Learning Novel Objects	5263
<i>Ting Yao, Yingwei Pan, Yehao Li, and Tao Mei</i>	
Beyond Instance-Level Image Retrieval: Leveraging Captions to Learn a Global Visual Representation for Semantic Retrieval	5272
<i>Albert Gordo and Diane Larlus</i>	
MuCaLe-Net: Multi Categorical-Level Networks to Generate More Discriminating Features	5282
<i>Youssef Tamaazousti, Hervé Le Borgne, and Céline Hudelot</i>	
Zero Shot Learning via Multi-scale Manifold Regularization	5292
<i>Shay Deutsch, Soheil Kolouri, Kyungham Kim, Yuri Owechko, and Stefano Soatto</i>	
Deeply Supervised Salient Object Detection with Short Connections	5300
<i>Qibin Hou, Ming-Ming Cheng, Xiaowei Hu, Ali Borji, Zhuowen Tu, and Philip Torr</i>	
A Matrix Splitting Method for Composite Function Minimization	5310
<i>Ganzhao Yuan, Wei-Shi Zheng, and Bernard Ghanem</i>	
One-Shot Video Object Segmentation	5320
<i>S. Caelles, K. -K. Maninis, J. Pont-Tuset, L. Leal-Taixé, D. Cremers, and L. Van Gool</i>	
Fast Person Re-identification via Cross-Camera Semantic Binary Transformation	5330
<i>Jiaxin Chen, Yunhong Wang, Jie Qin, Li Liu, and Ling Shao</i>	
SPFTN: A Self-Paced Fine-Tuning Network for Segmenting Objects in Weakly Labelled Videos	5340
<i>Dingwen Zhang, Le Yang, Deyu Meng, Dong Xu, and Junwei Han</i>	

Oral and Spotlight 4-1A

S4-1A: Machine Learning 4

Hidden Layers in Perceptual Learning	5349
<i>Gad Cohen and Daphna Weinshall</i>	
Few-Shot Object Recognition from Machine-Labeled Web Images	5358
<i>Zhongwen Xu, Linchao Zhu, and Yi Yang</i>	
Hallucinating Very Low-Resolution Unaligned and Noisy Face Images by Transformative Discriminative Autoencoders	5367
<i>Xin Yu and Fatih Porikli</i>	
Are You Smarter Than a Sixth Grader? Textbook Question Answering for Multimodal Machine Comprehension	5376
<i>Aniruddha Kembhavi, Minjoon Seo, Dustin Schwenk, Jonghyun Choi, Ali Farhadi, and Hannaneh Hajishirzi</i>	
Deep Hashing Network for Unsupervised Domain Adaptation	5385
<i>Hemanth Venkateswara, Jose Eusebio, Shayok Chakraborty, and Sethuraman Panchanathan</i>	
Generalized Deep Image to Image Regression	5395
<i>Venkataraman Santhanam, Vlad I. Morariu, and Larry S. Davis</i>	
Deep Learning with Low Precision by Half-Wave Gaussian Quantization	5406
<i>Zhaowei Cai, Xiaodong He, Jian Sun, and Nuno Vasconcelos</i>	
Creativity: Generating Diverse Questions Using Variational Autoencoders	5415
<i>Unnat Jain, Ziyu Zhang, and Alexander Schwing</i>	

O4-1A

Geometric Deep Learning on Graphs and Manifolds Using Mixture Model CNNs	5425
<i>Federico Monti, Davide Boscaini, Jonathan Masci, Emanuele Rodolà, Jan Svoboda, and Michael M. Bronstein</i>	
Full Resolution Image Compression with Recurrent Neural Networks	5435
<i>George Toderici, Damien Vincent, Nick Johnston, Sung Jin Hwang, David Minnen, Joel Shor, and Michele Covell</i>	
Neural Face Editing with Intrinsic Image Disentangling	5444
<i>Zhixin Shu, Ersin Yumer, Sunil Hadap, Kalyan Sunkavalli, Eli Shechtman, and Dimitris Samaras</i>	
UberNet: Training a Universal Convolutional Neural Network for Low-, Mid-, and High-Level Vision Using Diverse Datasets and Limited Memory	5454
<i>Iasonas Kokkinos</i>	

Oral and Spotlight 4-1B

S4-1B: Analyzing Humans with 3D Vision

3D Face Morphable Models "In-the-Wild"	5464
<i>James Booth, Epameinondas Antonakos, Stylianos Ploumpis, George Trigeorgis, Yannis Panagakis, and Stefanos Zafeiriou</i>	
KillingFusion: Non-rigid 3D Reconstruction without Correspondences	5474
<i>Miroslava Slavcheva, Maximilian Baust, Daniel Cremers, and Slobodan Ilic</i>	
Detailed, Accurate, Human Shape Estimation from Clothed 3D Scan Sequences	5484
<i>Chao Zhang, Sergi Pujades, Michael Black, and Gerard Pons-Moll</i>	
POSEidon: Face-from-Depth for Driver Pose Estimation	5494
<i>Guido Borghi, Marco Venturelli, Roberto Vezzani, and Rita Cucchiara</i>	
Human Shape from Silhouettes Using Generative HKS Descriptors and Cross-Modal Neural Networks	5504
<i>Endri Dibra, Himanshu Jain, Cengiz Öztireli, Remo Ziegler, and Markus Gross</i>	
Parametric T-Spline Face Morphable Model for Detailed Fitting in Shape Subspace	5515
<i>Weilong Peng, Zhiyong Feng, Chao Xu, and Yong Su</i>	
3D Menagerie: Modeling the 3D Shape and Pose of Animals	5524
<i>Silvia Zuffi, Angjoo Kanazawa, David W. Jacobs, and Michael J. Black</i>	
iCaRL: Incremental Classifier and Representation Learning	5533
<i>Sylvestre-Alvise Rebuffi, Alexander Kolesnikov, Georg Sperl, and Christoph H. Lampert</i>	

O4-1B

Recurrent 3D Pose Sequence Machines	5543
<i>Mude Lin, Liang Lin, Xiaodan Liang, Keze Wang, and Hui Cheng</i>	
Learning Detailed Face Reconstruction from a Single Image	5553
<i>Elad Richardson, Matan Sela, Roy Or-El, and Ron Kimmel</i>	
Thin-Slicing Network: A Deep Structured Model for Pose Estimation in Videos	5563
<i>Jie Song, Limin Wang, Luc Van Gool, and Otmar Hilliges</i>	
Dynamic FAUST: Registering Human Bodies in Motion	5573
<i>Federica Bogo, Javier Romero, Gerard Pons-Moll, and Michael J. Black</i>	

Poster Session P4-1

Semantically Coherent Co-Segmentation and Reconstruction of Dynamic Scenes	5583
<i>Armin Mustafa and Adrian Hilton</i>	
On the Two-View Geometry of Unsynchronized Cameras	5593
<i>Cenek Albl, Zuzana Kukelova, Andrew Fitzgibbon, Jan Heller, Matej Smid, and Tomas Pajdla</i>	

Using Locally Corresponding CAD Models for Dense 3D Reconstructions from a Single Image	5603
<i>Chen Kong, Chen-Hsuan Lin, and Simon Lucey</i>	
Convex Global 3D Registration with Lagrangian Duality	5612
<i>Jesus Briales and Javier Gonzalez-Jimenez</i>	
DeMoN: Depth and Motion Network for Learning Monocular Stereo	5622
<i>Benjamin Ummerhofer, Huizhong Zhou, Jonas Uhrig, Nikolaus Mayer, Eddy Ilg, Alexey Dosovitskiy, and Thomas Brox</i>	
3D Bounding Box Estimation Using Deep Learning and Geometry	5632
<i>Arsalan Mousavian, Dragomir Anguelov, John Flynn, and Jana Košecká</i>	
A Dataset for Benchmarking Image-Based Localization	5641
<i>Xun Sun, Yuanfan Xie, Pei Luo, and Liang Wang</i>	
Asynchronous Temporal Fields for Action Recognition	5650
<i>Gunnar A. Sigurdsson, Santosh Divvala, Ali Farhadi, and Abhinav Gupta</i>	
Sequential Person Recognition in Photo Albums with a Recurrent Network	5660
<i>Yao Li, Guosheng Lin, Bohan Zhuang, Lingqiao Liu, Chunhua Shen, and Anton van den Hengel</i>	
Multi-context Attention for Human Pose Estimation	5669
<i>Xiao Chu, Wei Yang, Wanli Ouyang, Cheng Ma, Alan L. Yuille, and Xiaogang Wang</i>	
3D Convolutional Neural Networks for Efficient and Robust Hand Pose Estimation from Single Depth Images	5679
<i>Liu Hao Ge, Hui Liang, Junsong Yuan, and Daniel Thalmann</i>	
Lifting from the Deep: Convolutional 3D Pose Estimation from a Single Image	5689
<i>Denis Tome, Chris Russell, and Lourdes Agapito</i>	
AdaScan: Adaptive Scan Pooling in Deep Convolutional Neural Networks for Human Action Recognition in Videos	5699
<i>Amlan Kar, Nishant Rai, Karan Sikka, and Gaurav Sharma</i>	
Deep Structured Learning for Facial Action Unit Intensity Estimation	5709
<i>Robert Walecki, Ognjen (Oggi) Rudovic, Vladimir Pavlovic, Björn Schuller, and Maja Pantic</i>	
Simultaneous Facial Landmark Detection, Pose and Deformation Estimation Under Facial Occlusion	5719
<i>Yue Wu, Chao Gou, and Qiang Ji</i>	
Self-Supervised Video Representation Learning with Odd-One-Out Networks	5729
<i>Basura Fernando, Hakan Bilen, Efstratios Gavves, and Stephen Gould</i>	
Robust Joint and Individual Variance Explained	5739
<i>Christos Sagonas, Yannis Panagakis, Alina Leidinger, and Stefanos Zafeiriou</i>	
Discriminative Covariance Oriented Representation Learning for Face Recognition with Image Sets	5749
<i>Wen Wang, Ruiping Wang, Shiguang Shan, and Xilin Chen</i>	
3D Human Pose Estimation = 2D Pose Estimation + Matching	5759
<i>Ching-Hang Chen and Deva Ramanan</i>	

Joint Gap Detection and Inpainting of Line Drawings	5768
<i>Kazuma Sasaki, Satoshi Iizuka, Edgar Simo-Serra, and Hiroshi Ishikawa</i>	
Riemannian Nonlinear Mixed Effects Models: Analyzing Longitudinal Deformations in Neuroimaging	5777
<i>Hyunwoo J. Kim, Nagesh Adluru, Heemanshu Suri, Baba C. Vemuri, Sterling C. Johnson, and Vikas Singh</i>	
Simultaneous Super-Resolution and Cross-Modality Synthesis of 3D Medical Images Using Weakly-Supervised Joint Convolutional Sparse Coding	5787
<i>Yawen Huang, Ling Shao, and Alejandro F. Frangi</i>	
Multiple-Scattering Microphysics Tomography	5797
<i>Aviad Levis, Yoav Y. Schechner, and Anthony B. Davis</i>	
Accurate Optical Flow via Direct Cost Volume Processing	5807
<i>Jia Xu, René Ranftl, and Vladlen Koltun</i>	
Event-Based Visual Inertial Odometry	5816
<i>Alex Zihao Zhu, Nikolay Atanasov, and Kostas Daniilidis</i>	
Robust Visual Tracking Using Oblique Random Forests	5825
<i>Le Zhang, Jagannadan Varadarajan, Ponnuthurai Nagaratnam Suganthan, Narendra Ahuja, and Pierre Moulin</i>	
Deep Laplacian Pyramid Networks for Fast and Accurate Super-Resolution	5835
<i>Wei-Sheng Lai, Jia-Bin Huang, Narendra Ahuja, and Ming-Hsuan Yang</i>	
Learning Non-Lambertian Object Intrinsic Across ShapeNet Categories	5844
<i>Jian Shi, Yue Dong, Hao Su, and Stella X. Yu</i>	
MCMLSD: A Dynamic Programming Approach to Line Segment Detection	5854
<i>Emilio J. Almazán, Ron Tal, Yiming Qian, and James H. Elder</i>	
Contour-Constrained Superpixels for Image and Video Processing	5863
<i>Se-Ho Lee, Won-Dong Jang, and Chang-Su Kim</i>	
Richer Convolutional Features for Edge Detection	5872
<i>Yun Liu, Ming-Ming Cheng, Xiaowei Hu, Kai Wang, and Xiang Bai</i>	
Non-local Color Image Denoising with Convolutional Neural Networks	5882
<i>Stamatios Lefkimmiatis</i>	
Generative Face Completion	5892
<i>Yijun Li, Sifei Liu, Jimei Yang, and Ming-Hsuan Yang</i>	
Hyper-Laplacian Regularized Unidirectional Low-Rank Tensor Recovery for Multispectral Image Denoising	5901
<i>Yi Chang, Luxin Yan, and Sheng Zhong</i>	
Unsupervised Semantic Scene Labeling for Streaming Data	5910
<i>Maggie Wigness and John G. Rogers III</i>	
Why You Should Forget Luminance Conversion and Do Something Better	5920
<i>Rang M. H. Nguyen and Michael S. Brown</i>	
Deep Semantic Feature Matching	5929
<i>Nikolai Ufer and Björn Ommer</i>	

Revisiting the Variable Projection Method for Separable Nonlinear Least Squares Problems	5939
<i>Je Hyeong Hong, Christopher Zach, and Andrew Fitzgibbon</i>	
Efficient Multiple Instance Metric Learning Using Weakly Supervised Data	5948
<i>Marc T. Law, Yaoliang Yu, Raquel Urtasun, Richard S. Zemel, and Eric P. Xing</i>	
WILDCAT: Weakly Supervised Learning of Deep ConvNets for Image Classification, Pointwise Localization and Segmentation	5957
<i>Thibaut Durand, Taylor Mordan, Nicolas Thome, and Matthieu Cord</i>	
Image-to-Image Translation with Conditional Adversarial Networks	5967
<i>Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros</i>	
Deep Roots: Improving CNN Efficiency with Hierarchical Filter Groups	5977
<i>Yani Ioannou, Duncan Robertson, Roberto Cipolla, and Antonio Criminisi</i>	
Aggregated Residual Transformations for Deep Neural Networks	5987
<i>Saining Xie, Ross Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He</i>	
MIML-FCN+: Multi-Instance Multi-Label Learning via Fully Convolutional Networks with Privileged Information	5996
<i>Hao Yang, Joey Tianyi Zhou, Jianfei Cai, and Yew Soon Ong</i>	
Low-Rank Embedded Ensemble Semantic Dictionary for Zero-Shot Learning	6005
<i>Zhengming Ding, Ming Shao, and Yun Fu</i>	
Factorized Variational Autoencoders for Modeling Audience Reactions to Movies	6014
<i>Zhiwei Deng, Rajitha Navarathna, Peter Carr, Stephan Mandt, Yisong Yue, Iain Matthews, and Greg Mori</i>	
Learning Features by Watching Objects Move	6024
<i>Deepak Pathak, Ross Girshick, Piotr Dollár, Trevor Darrell, and Bharath Hariharan</i>	
What Can Help Pedestrian Detection?	6034
<i>Jiayuan Mao, Tete Xiao, Yuning Jiang, and Zhimin Cao</i>	
DeepPermNet: Visual Permutation Learning	6044
<i>Rodrigo Santa Cruz, Basura Fernando, Anoop Cherian, and Stephen Gould</i>	
Learning the Multilinear Structure of Visual Data	6053
<i>Mengjiao Wang, Yannis Panagakis, Patrick Snape, and Stefanos Zafeiriou</i>	
Adaptive and Move Making Auxiliary Cuts for Binary Pairwise Energies	6062
<i>Lena Gorelick, Yuri Boykov, and Olga Veksler</i>	
Designing Energy-Efficient Convolutional Neural Networks Using Energy-Aware Pruning	6071
<i>Tien-Ju Yang, Yu-Hsin Chen, and Vivienne Sze</i>	
Joint Multi-person Pose Estimation and Semantic Part Segmentation	6080
<i>Fangting Xia, Peng Wang, Xianjie Chen, and Alan L. Yuille</i>	
Deep Feature Interpolation for Image Content Changes	6090
<i>Paul Upchurch, Jacob Gardner, Geoff Pleiss, Robert Pless, Noah Snavely, Kavita Bala, and Kilian Weinberger</i>	

FASON: First and Second Order Information Fusion Network for Texture Recognition	6100
<i>Xiyang Dai, Joe Yue-Hei Ng, and Larry S. Davis</i>	
Lean Crowdsourcing: Combining Humans and Machines in an Online System	6109
<i>Steve Branson, Grant Van Horn, and Pietro Perona</i>	
Supervising Neural Attention Models for Video Captioning by Human Gaze Data	6119
<i>Youngjae Yu, Jongwook Choi, Yeonhwa Kim, Kyung Yoo, Sang-Hun Lee, and Gunhee Kim</i>	
L2-Net: Deep Learning of Discriminative Patch Descriptor in Euclidean Space	6128
<i>Yurun Tian, Bin Fan, and Fuchao Wu</i>	
Convolutional Random Walk Networks for Semantic Image Segmentation	6137
<i>Gedas Bertasius, Lorenzo Torresani, Stella X. Yu, and Jianbo Shi</i>	
Knowledge Acquisition for Visual Question Answering via Iterative Querying	6146
<i>Yuke Zhu, Joseph J. Lim, and Li Fei-Fei</i>	
Memory-Augmented Attribute Manipulation Networks for Interactive Fashion Search	6156
<i>Bo Zhao, Jiashi Feng, Xiao Wu, and Shuicheng Yan</i>	
From Zero-Shot Learning to Conventional Supervised Classification: Unseen Visual Data Synthesis	6165
<i>Yang Long, Li Liu, Ling Shao, Fumin Shen, Guiguang Ding, and Jungong Han</i>	
Are Large-Scale 3D Models Really Necessary for Accurate Visual Localization?	6175
<i>Torsten Sattler, Akihiko Torii, Josef Sivic, Marc Pollefeys, Hajime Taira, Masatoshi Okutomi, and Tomas Pajdla</i>	
Asymmetric Feature Maps with Application to Sketch Based Retrieval	6185
<i>Giorgos Tolias and Ondrej Chum</i>	
Diverse Image Annotation	6194
<i>Baoyuan Wu, Fan Jia, Wei Liu, and Bernard Ghanem</i>	
AMC: Attention Guided Multi-modal Correlation Learning for Image Search	6203
<i>Kan Chen, Trung Bui, Chen Fang, Zhaowen Wang, and Ram Nevatia</i>	
Multi-attention Network for One Shot Learning	6212
<i>Peng Wang, Lingqiao Liu, Chunhua Shen, Zi Huang, Anton van den Hengel, and Heng Tao Shen</i>	
Fried Binary Embedding for High-Dimensional Visual Features	6221
<i>Weixiang Hong, Junsong Yuan, and Sreyasee Das Bhattacharjee</i>	
Pyramid Scene Parsing Network	6230
<i>Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia</i>	
Learning Deep Match Kernels for Image-Set Classification	6240
<i>Haoliang Sun, Xiantong Zhen, Yuanjie Zheng, Gongping Yang, Yilong Yin, and Shuo Li</i>	
Task-Driven Dynamic Fusion: Reducing Ambiguity in Video Description	6250
<i>Xishan Zhang, Ke Gao, Yongdong Zhang, Dongming Zhang, Jintao Li, and Qi Tian</i>	
Learning Multifunctional Binary Codes for Both Category and Attribute Oriented Retrieval Tasks	6259
<i>Haomiao Liu, Ruiping Wang, Shiguang Shan, and Xilin Chen</i>	

Indoor Scene Parsing with Instance Segmentation, Semantic Labeling and Support Relationship Inference	6269
<i>Wei Zhuo, Mathieu Salzmann, Xuming He, and Miaomiao Liu</i>	
Episodic CAMN: Contextual Attention-Based Memory Networks with Iterative Feedback for Scene Labeling	6278
<i>Abrar H. Abdulnabi, Bing Shuai, Stefan Winkler, and Gang Wang</i>	
Link the Head to the "Beak": Zero Shot Learning from Noisy Text Description at Part Precision	6288
<i>Mohamed Elhoseiny, Yizhe Zhu, Han Zhang, and Ahmed Elgammal</i>	
SCA-CNN: Spatial and Channel-Wise Attention in Convolutional Networks for Image Captioning	6298
<i>Long Chen, Hanwang Zhang, Jun Xiao, Liqiang Nie, Jian Shao, Wei Liu, and Tat-Seng Chua</i>	
Deep Pyramidal Residual Networks	6307
<i>Dongyoon Han, Jiwhan Kim, and Junmo Kim</i>	
Product Split Trees	6316
<i>Artem Babenko and Victor Lempitsky</i>	
Making the V in VQA Matter: Elevating the Role of Image Understanding in Visual Question Answering	6325
<i>Yash Goyal, Tejas Khot, Douglas Summers-Stay, Dhruv Batra, and Devi Parikh</i>	
Commonly Uncommon: Semantic Sparsity in Situation Recognition	6335
<i>Mark Yatskar, Vicente Ordonez, Luke Zettlemoyer, and Ali Farhadi</i>	
Cross-Modality Binary Code Learning via Fusion Similarity Hashing	6345
<i>Hong Liu, Rongrong Ji, Yongjian Wu, Feiyue Huang, and Baochang Zhang</i>	
Saliency Revisited: Analysis of Mouse Movements Versus Fixations	6354
<i>Hamed R. Tavakoli, Fawad Ahmed, Ali Borji, and Jorma Laaksonen</i>	
InterpoNet, a Brain Inspired Neural Network for Optical Flow Dense Interpolation	6363
<i>Shay Zweig and Lior Wolf</i>	
SST: Single-Stream Temporal Action Proposals	6373
<i>Shyamal Buch, Victor Escorcia, Chuanqi Shen, Bernard Ghanem, and Juan Carlos Niebles</i>	
Video Segmentation via Multiple Granularity Analysis	6383
<i>Rui Yang, Bingbing Ni, Chao Ma, Yi Xu, and Xiaokang Yang</i>	
Spatio-Temporal Alignment of Non-overlapping Sequences from Independently Panning Cameras	6393
<i>S. Morteza Safdarnejad and Xiaoming Liu</i>	
UntrimmedNets for Weakly Supervised Action Recognition and Detection	6402
<i>Limin Wang, Yuanjun Xiong, Dahua Lin, and Luc Van Gool</i>	

Oral and Spotlight 4-2A

S4-2A: Object Recognition & Scene Understanding 3

Gaze Embeddings for Zero-Shot Image Classification	6412
<i>Nour Kaessli, Zeynep Akata, Bernt Schiele, and Andreas Bulling</i>	
What's in a Question: Using Visual Questions as a Form of Supervision	6422
<i>Siddha Ganju, Olga Russakovsky, and Abhinav Gupta</i>	
Attend to You: Personalized Image Captioning with Context Sequence Memory Networks	6432
<i>Cesc Chunseong Park, Byeongchang Kim, and Gunhee Kim</i>	
Adversarially Tuned Scene Generation	6441
<i>Vsr Veeravasrapu, Constantin Rothkopf, and Ramesh Visvanathan</i>	
Residual Attention Network for Image Classification	6450
<i>Fei Wang, Mengqing Jiang, Chen Qian, Shuo Yang, Cheng Li, Honggang Zhang, Xiaogang Wang, and Xiaoou Tang</i>	
Not All Pixels Are Equal: Difficulty-Aware Semantic Segmentation via Deep Layer Cascade	6459
<i>Xiaoxiao Li, Ziwei Liu, Ping Luo, Chen Change Loy, and Xiaoou Tang</i>	
Learning Non-maximum Suppression	6469
<i>Jan Hosang, Rodrigo Benenson, and Bernt Schiele</i>	
The Amazing Mysteries of the Gutter: Drawing Inferences Between Panels in Comic Book Narratives	6478
<i>Mohit Iyyer, Varun Manjunatha, Anupam Guha, Yogarshi Vyas, Jordan Boyd-Graber, Hal Daumé III, and Larry Davis</i>	

O4-2A

Object Region Mining with Adversarial Erasing: A Simple Classification to Semantic Segmentation Approach	6488
<i>Yunchao Wei, Jiashi Feng, Xiaodan Liang, Ming-Ming Cheng, Yao Zhao, and Shuicheng Yan</i>	
Fine-Grained Recognition as HSnet Search for Informative Image Parts	6497
<i>Michael Lam, Behrooz Mahasseni, and Sinisa Todorovic</i>	
G2DeNet: Global Gaussian Distribution Embedding Network and Its Application to Visual Recognition	6507
<i>Qilong Wang, Peihua Li, and Lei Zhang</i>	
YOLO9000: Better, Faster, Stronger	6517
<i>Joseph Redmon and Ali Farhadi</i>	

Oral and Spotlight 4-2B

S4-2B: Machine Learning for 3D Vision

Multi-view 3D Object Detection Network for Autonomous Driving	6526
<i>Xiaozhi Chen, Huimin Ma, Ji Wan, Bo Li, and Tian Xia</i>	
UltraStereo: Efficient Learning-Based Matching for Active Stereo Systems	6535
<i>Sean Ryan Fanello, Julien Valentin, Christoph Rhemann, Adarsh Kowdle, Vladimir Tankovich, Philip Davidson, and Shahram Izadi</i>	
Shape Completion Using 3D-Encoder-Predictor CNNs and Shape Synthesis	6545
<i>Angela Dai, Charles Ruizhongtai Qi, and Matthias Nießner</i>	
Geometric Loss Functions for Camera Pose Regression with Deep Learning	6555
<i>Alex Kendall and Roberto Cipolla</i>	
CNN-SLAM: Real-Time Dense Monocular SLAM with Learned Depth Prediction	6565
<i>Keisuke Tateno, Federico Tombari, Iro Laina, and Nassir Navab</i>	
Learning from Noisy Large-Scale Datasets with Minimal Supervision	6575
<i>Andreas Veit, Neil Alldrin, Gal Chechik, Ivan Krasin, Abhinav Gupta, and Serge Belongie</i>	
SyncSpecCNN: Synchronized Spectral CNN for 3D Shape Segmentation	6584
<i>Li Yi, Hao Su, Xingwen Guo, and Leonidas Guibas</i>	
Non-local Deep Features for Salient Object Detection	6593
<i>Zhiming Luo, Akshaya Mishra, Andrew Achkar, Justin Eichel, Shaozi Li, and Pierre-Marc Jodoin</i>	

O4-2B

Unsupervised Monocular Depth Estimation with Left-Right Consistency	6602
<i>Clément Godard, Oisín Mac Aodha, and Gabriel J. Brostow</i>	
Unsupervised Learning of Depth and Ego-Motion from Video	6612
<i>Tinghui Zhou, Matthew Brown, Noah Snavely, and David G. Lowe</i>	
OctNet: Learning Deep 3D Representations at High Resolutions	6620
<i>Gernot Riegler, Ali Osman Ulusoy, and Andreas Geiger</i>	
3D Shape Segmentation with Projective Convolutional Networks	6630
<i>Evangelos Kalogerakis, Melinos Averkiou, Subhransu Maji, and Siddhartha Chaudhuri</i>	

Poster Session P4-2

SGM-Nets: Semi-Global Matching with Neural Networks	6640
<i>Akihito Seki and Marc Pollefeys</i>	
Stereo-Based 3D Reconstruction of Dynamic Fluid Surfaces by Global Optimization	6650
<i>Yiming Qian, Minglun Gong, and Yee-Hong Yang</i>	
Fine-to-Coarse Global Registration of RGB-D Scans	6660
<i>Maciej Halber and Thomas Funkhouser</i>	
Analyzing Computer Vision Data — The Good, the Bad and the Ugly	6670
<i>Oliver Zendel, Katrin Honauer, Markus Murschitz, Martin Humenberger, and Gustavo Fernández Domínguez</i>	

Product Manifold Filter: Non-rigid Shape Correspondence via Kernel Density Estimation in the Product Space	6681
<i>Matthias Vestner, Roee Litman, Emanuele Rodolà, Alex Bronstein, and Daniel Cremers</i>	
Unsupervised Vanishing Point Detection and Camera Calibration from a Single Manhattan Image with Radial Distortion	6691
<i>Michel Antunes, João P. Barreto, Djamila Aouada, and Björn Ottersten</i>	
Toroidal Constraints for Two-Point Localization Under High Outlier Ratios	6700
<i>Federico Camposeco, Torsten Sattler, Andrea Cohen, Andreas Geiger, and Marc Pollefeys</i>	
4D Light Field Superpixel and Segmentation	6709
<i>Hao Zhu, Qi Zhang, and Qing Wang</i>	
Exploiting Symmetry and/or Manhattan Properties for 3D Object Structure Estimation from Single and Multiple Images	6718
<i>Yuan Gao and Alan L. Yuille</i>	
Binary Coding for Partial Action Analysis with Limited Observation Ratios	6728
<i>Jie Qin, Li Liu, Ling Shao, Bingbing Ni, Chen Chen, Fumin Shen, and Yunhong Wang</i>	
SphereFace: Deep Hypersphere Embedding for Face Recognition	6738
<i>Weiyang Liu, Yandong Wen, Zhiding Yu, Ming Li, Bhiksha Raj, and Le Song</i>	
IRINA: Iris Recognition (Even) in Inaccurately Segmented Data	6747
<i>Hugo Proença and João C. Neves</i>	
Look into Person: Self-Supervised Structure-Sensitive Learning and a New Benchmark for Human Parsing	6757
<i>Ke Gong, Xiaodan Liang, Dongyu Zhang, Xiaohui Shen, and Liang Lin</i>	
Action Unit Detection with Region Adaptation, Multi-labeling Learning and Optimal Temporal Fusing	6766
<i>Wei Li, Farnaz Abtahi, and Zhigang Zhu</i>	
See the Forest for the Trees: Joint Spatial and Temporal Recurrent Neural Networks for Video-Based Person Re-identification	6776
<i>Zhen Zhou, Yan Huang, Wei Wang, Liang Wang, and Tieniu Tan</i>	
Joint Intensity and Spatial Metric Learning for Robust Gait Recognition	6786
<i>Yasushi Makihara, Atsuyuki Suzuki, Daigo Muramatsu, Xiang Li, and Yasushi Yagi</i>	
Pose-Aware Person Recognition	6797
<i>Vijay Kumar, Anoop Namboodiri, Manohar Paluri, and C. V. Jawahar</i>	
Not Afraid of the Dark: NIR-VIS Face Recognition via Cross-Spectral Hallucination and Low-Rank Embedding	6807
<i>José Lezama, Qiang Qiu, and Guillermo Sapiro</i>	
Jointly Learning Energy Expenditures and Activities Using Egocentric Multimodal Signals	6817
<i>Katsuyuki Nakamura, Serena Yeung, Alexandre Alahi, and Li Fei-Fei</i>	
Binarized Mode Seeking for Scalable Visual Pattern Discovery	6827
<i>Wei Zhang, Xiaochun Cao, Rui Wang, Yuanfang Guo, and Zhineng Chen</i>	

Scribbler: Controlling Deep Image Synthesis with Sketch and Color	6836
<i>Patsorn Sangkloy, Jingwan Lu, Chen Fang, Fisher Yu, and James Hays</i>	
Multi-way Multi-level Kernel Modeling for Neuroimaging Classification	6846
<i>Lifang He, Chun-Ta Lu, Hao Ding, Shen Wang, Linlin Shen, Philip S. Yu, and Ann B. Ragin</i>	
WSISA: Making Survival Prediction from Whole Slide Histopathological Images	6855
<i>Xinliang Zhu, Jiawen Yao, Feiyun Zhu, and Junzhou Huang</i>	
On the Effectiveness of Visible Watermarks	6864
<i>Tali Dekel, Michael Rubinstein, Ce Liu, and William T. Freeman</i>	
Snapshot Hyperspectral Light Field Imaging	6873
<i>Zhiwei Xiong, Lizhi Wang, Huiqun Li, Dong Liu, and Feng Wu</i>	
Semantic Image Inpainting with Deep Generative Models	6882
<i>Raymond A. Yeh, Chen Chen, Teck Yian Lim, Alexander G. Schwing, Mark Hasegawa-Johnson, and Minh N. Do</i>	
Fast Multi-frame Stereo Scene Flow with Motion Segmentation	6891
<i>Tatsunori Taniai, Sudipta N. Sinha, and Yoichi Sato</i>	
Improved Stereo Matching with Constant Highway Networks and Reflective Confidence Learning	6901
<i>Amit Shaked and Lior Wolf</i>	
Optical Flow in Mostly Rigid Scenes	6911
<i>Jonas Wulff, Laura Sevilla-Lara, and Michael J. Black</i>	
Optical Flow Requires Multiple Strategies (but Only One Network)	6921
<i>Tal Schuster, Lior Wolf, and David Gadot</i>	
ECO: Efficient Convolution Operators for Tracking	6931
<i>Martin Danelljan, Goutam Bhat, Fahad Shahbaz Khan, and Michael Felsberg</i>	
Differential Angular Imaging for Material Recognition	6940
<i>Jia Xue, Hang Zhang, Kristin Dana, and Ko Nishino</i>	
Fast Fourier Color Constancy	6950
<i>Jonathan T. Barron and Yun-Ta Tsai</i>	
Comparative Evaluation of Hand-Crafted and Learned Local Features	6959
<i>Johannes L. Schönberger, Hans Hardmeier, Torsten Sattler, and Marc Pollefeys</i>	
Learning Fully Convolutional Networks for Iterative Non-blind Deconvolution	6969
<i>Jiawei Zhang, Jinshan Pan, Wei-Sheng Lai, Rynson W. H. Lau, and Ming-Hsuan Yang</i>	
Image Deblurring via Extreme Channels Prior	6978
<i>Yanyang Yan, Wenqi Ren, Yuanfang Guo, Rui Wang, and Xiaochun Cao</i>	
Simultaneous Stereo Video Deblurring and Scene Flow Estimation	6987
<i>Liyuan Pan, Yuchao Dai, Miaomiao Liu, and Fatih Porikli</i>	
Deep Photo Style Transfer	6997
<i>Fujun Luan, Sylvain Paris, Eli Shechtman, and Kavita Bala</i>	
Generative Attribute Controller with Conditional Filtered Generative Adversarial Networks	7006
<i>Takuhiro Kaneko, Kaoru Hiramatsu, and Kunio Kashino</i>	

Fast Haze Removal for Nighttime Image Using Maximum Reflectance Prior	7016
<i>Jing Zhang, Yang Cao, Shuai Fang, Yu Kang, and Chang Wen Chen</i>	
Low-Rank Bilinear Pooling for Fine-Grained Classification	7025
<i>Shu Kong and Charless Fowlkes</i>	
Neural Scene De-rendering	7035
<i>Jiajun Wu, Joshua B. Tenenbaum, and Pushmeet Kohli</i>	
Real-Time Neural Style Transfer for Videos	7044
<i>Haozhi Huang, Hao Wang, Wenhan Luo, Lin Ma, Wenhao Jiang, Xiaolong Zhu, Zhifeng Li, and Wei Liu</i>	
A Graph Regularized Deep Neural Network for Unsupervised Image Representation Learning	7053
<i>Shijie Yang, Liang Li, Shuhui Wang, Weigang Zhang, and Qingming Huang</i>	
A Study of Lagrangean Decompositions and Dual Ascent Solvers for Graph Matching	7062
<i>Paul Swoboda, Carsten Rother, Hassan Abu Alhaja, Dagmar Kainmüller, and Bogdan Savchynskyy</i>	
Collaborative Deep Reinforcement Learning for Joint Object Search	7072
<i>Xiangyu Kong, Bo Xin, Yizhou Wang, and Gang Hua</i>	
Loss Max-Pooling for Semantic Image Segmentation	7082
<i>Samuel Rota Bulò, Gerhard Neuhold, and Peter Kotschieder</i>	
Deep View Morphing	7092
<i>Dinghuang Ji, Junghyun Kwon, Max McFarland, and Silvio Savarese</i>	
Unsupervised Learning of Long-Term Motion Dynamics for Videos	7101
<i>Zelun Luo, Boya Peng, De-An Huang, Alexandre Alahi, and Li Fei-Fei</i>	
Revisiting Metric Learning for SPD Matrix Based Visual Representation	7111
<i>Luping Zhou, Lei Wang, Jianjia Zhang, Yinghuan Shi, and Yang Gao</i>	
Expert Gate: Lifelong Learning with a Network of Experts	7120
<i>Rahaf Aljundi, Punarjay Chakravarty, and Tinne Tuytelaars</i>	
A Gift from Knowledge Distillation: Fast Optimization, Network Minimization and Transfer Learning	7130
<i>Junho Yim, Donggyu Joo, Jihoon Bae, and Junmo Kim</i>	
Domain Adaptation by Mixture of Alignments of Second-or Higher-Order Scatter Tensors	7139
<i>Piotr Koniusz, Yusuf Tas, and Fatih Porikli</i>	
Deep Mixture of Linear Inverse Regressions Applied to Head-Pose Estimation	7149
<i>Stéphane Lathuilière, Rémi Juge, Pablo Mesejo, Rafael Muñoz-Salinas, and Radu Horaud</i>	
STD2P: RGBD Semantic Segmentation Using Spatio-Temporal Data-Driven Pooling	7158
<i>Yang He, Wei-Chen Chiu, Margret Keuper, and Mario Fritz</i>	
Harmonic Networks: Deep Translation and Rotation Equivariance	7168
<i>Daniel E. Worrall, Stephan J. Garbin, Daniyar Turmukhambetov, and Gabriel J. Brostow</i>	

Multimodal Transfer: A Hierarchical Deep Convolutional Neural Network for Fast Artistic Style Transfer	7178
<i>Xin Wang, Geoffrey Oxholm, Da Zhang, and Yuan-Fang Wang</i>	
Detect, Replace, Refine: Deep Structured Prediction for Pixel Wise Labeling	7187
<i>Spyros Gidaris and Nikos Komodakis</i>	
Weighted-Entropy-Based Quantization for Deep Neural Networks	7197
<i>Eunhyeok Park, Junwhan Ahn, and Sungjoo Yoo</i>	
Residual Expansion Algorithm: Fast and Effective Optimization for Nonconvex Least Squares Problems	7206
<i>Daiki Ikami, Toshihiko Yamasaki, and Kiyoharu Aizawa</i>	
Bidirectional Beam Search: Forward-Backward Inference in Neural Sequence Models for Fill-in-the-Blank Image Captioning	7215
<i>Qing Sun, Stefan Lee, and Dhruv Batra</i>	
Newton-Type Methods for Inference in Higher-Order Markov Random Fields	7224
<i>Hariprasad Kannan, Nikos Komodakis, and Nikos Paragios</i>	
Adaptive Relaxed ADMM: Convergence Theory and Practical Implementation	7234
<i>Zheng Xu, Mário A. T. Figueiredo, Xiaoming Yuan, Christoph Studer, and Tom Goldstein</i>	
ViP-CNN: Visual Phrase Guided Convolutional Neural Network	7244
<i>Yikang Li, Wanli Ouyang, Xiaogang Wang, and Xiao'Ou Tang</i>	
Instance-Aware Image and Sentence Matching with Selective Multimodal LSTM	7254
<i>Yan Huang, Wei Wang, and Liang Wang</i>	
Kernel Square-Loss Exemplar Machines for Image Retrieval	7263
<i>Rafael S. Rezende, Joaquin Zepeda, Jean Ponce, Francis Bach, and Patrick Pérez</i>	
Cognitive Mapping and Planning for Visual Navigation	7272
<i>Saurabh Gupta, James Davidson, Sergey Levine, Rahul Sukthankar, and Jitendra Malik</i>	
Combining Bottom-Up, Top-Down, and Smoothness Cues for Weakly Supervised Image Segmentation	7282
<i>Anirban Roy and Sinisa Todorovic</i>	
Seeing into Darkness: Scotopic Visual Recognition	7292
<i>Bo Chen and Pietro Perona</i>	
Deep Co-occurrence Feature Learning for Visual Object Recognition	7302
<i>Ya-Fang Shih, Yang-Ming Yeh, Yen-Yu Lin, Ming-Fang Weng, Yi-Chang Lu, and Yung-Yu Chuang</i>	
An Empirical Evaluation of Visual Question Answering for Novel Objects	7312
<i>Santhosh K. Ramakrishnan, Ambar Pal, Gaurav Sharma, and Anurag Mittal</i>	
InstanceCut: From Edges to Instances with MultiCut	7322
<i>Alexander Kirillov, Evgeny Levinkov, Bjoern Andres, Bogdan Savchynskyy, and Carsten Rother</i>	
Fine-Grained Image Classification via Combining Vision and Language	7332
<i>Xiangteng He and Yuxin Peng</i>	

Mimicking Very Efficient Network for Object Detection	7341
<i>Quanquan Li, Shengying Jin, and Junjie Yan</i>	
Tracking by Natural Language Specification	7350
<i>Zhenyang Li, Ran Tao, Efstratios Gavves, Cees G. M. Snoek, and Arnold W. M. Smeulders</i>	
A Dataset and Exploration of Models for Understanding Video Data through Fill-in-the-Blank Question-Answering	7359
<i>Tegan Maharaj, Nicolas Ballas, Anna Rohrbach, Aaron Courville, and Christopher Pal</i>	
Learning Detection with Diverse Proposals	7369
<i>Samaneh Azadi, Jiashi Feng, and Trevor Darrell</i>	
Skeleton Key: Image Captioning by Skeleton-Attribute Decomposition	7378
<i>Yufei Wang, Zhe Lin, Xiaohui Shen, Scott Cohen, and Garrison W. Cottrell</i>	
A Low Power, Fully Event-Based Gesture Recognition System	7388
<i>Arnon Amir, Brian Taba, David Berg, Timothy Melano, Jeffrey McKinstry, Carmelo Di Nolfo, Tapan Nayak, Alexander Andreopoulos, Guillaume Garreau, Marcela Mendoza, Jeff Kusnitz, Michael Debole, Steve Esser, Tobi Delbruck, Myron Flickner, and Dharmendra Modha</i>	
Learning Deep Context-Aware Features over Body and Latent Parts for Person Re-identification	7398
<i>Dangwei Li, Xiaotang Chen, Zhang Zhang, and Kaiqi Huang</i>	
Recurrent Modeling of Interaction Context for Collective Activity Recognition	7408
<i>Minsi Wang, Bingbing Ni, and Xiaokang Yang</i>	
Primary Object Segmentation in Videos Based on Region Augmentation and Reduction	7417
<i>Yeong Jun Koh and Chang-Su Kim</i>	
ROAM: A Rich Object Appearance Model with Application to Rotoscoping	7426
<i>Ondrej Miksik, Juan-Manuel Pérez-Rúa, Philip H. S. Torr, and Patrick Pérez</i>	
Temporal Residual Networks for Dynamic Scene Recognition	7435
<i>Christoph Feichtenhofer, Axel Pinz, and Richard P. Wildes</i>	
Spatiotemporal Multiplier Networks for Video Action Recognition	7445
<i>Christoph Feichtenhofer, Axel Pinz, and Richard P. Wildes</i>	
Learning to Learn from Noisy Web Videos	7455
<i>Serena Yeung, Vignesh Ramanathan, Olga Russakovsky, Liyue Shen, Greg Mori, and Li Fei-Fei</i>	
YouTube-BoundingBoxes: A Large High-Precision Human-Annotated Data Set for Object Detection in Video	7464
<i>Esteban Real, Jonathon Shlens, Stefano Mazzocchi, Xin Pan, and Vincent Vanhoucke</i>	
Online Video Object Segmentation via Convolutional Trident Network	7474
<i>Won-Dong Jang and Chang-Su Kim</i>	

Author Index