Aloha!

Welcome to Honolulu, Hawaii and the 30th IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). In addition to the main four-day program of presentations, interactive sessions, plenary talks, demos, exhibitions, and social functions, CVPR 2017 has a number of co-located events, including 44 workshops and 20 tutorials. As the fields of computer vision, pattern recognition, machine learning and artificial intelligence continue to break new ground and scale new heights, so does our conference. This year CVPR 2017 received a record 2680 valid submissions to the main conference, of which 2620 were fully reviewed (the others were administratively rejected for technical or ethical reasons or were withdrawn before review).

The number of papers reviewed for CVPR 2017 was 40% larger than the number reviewed for last year’s edition of the conference, a growth rate that posed significant organizational challenges.

To select papers for the program from these submissions, we invited 85 researchers to act as Area Chairs (ACs). ACs were selected to provide a broad range of expertise, to balance junior and senior members, and to represent a variety of geographical locations. Additionally, we recruited a record number of experienced reviewers from the broader computer vision and pattern recognition community. The original list of reviewers was augmented with reviewers recommended by the ACs to add expertise for papers where appropriate reviewers were not initially available.

The reviewing process accepted 783 papers (29% of valid submissions). 71 of these were accepted as oral presentations (2.65% of valid submissions) and 144 were accepted as spotlight oral presentations, for a total of 8% of valid submissions with live presentations. Continuing the successful innovation from CVPR 2016, the inclusion of spotlight oral presentations has allowed us to increase the number of works presented from the podium.

All accepted papers will appear in the interactive poster sessions where we hope that lively discussions will ensue. The total number of papers presented at CVPR 2017 is 22% larger than the number presented at last year’s edition.

The review process was similar to previous years. Each paper was reviewed by at least three reviewers and considered by at least three ACs before a decision was made. Borderline papers and candidate orals and spotlights were discussed in groups of three non-conflicted ACs with common areas of expertise. Oral and spotlight recommendations were made by panels of ACs after extensive discussion.

The Program Chairs did not submit any papers to CVPR 2017, allowing them to avoid direct conflicts throughout the review process. This year, General Chairs who were allowed to submit papers, did not have any software access to the CMT system beyond that of an author. The double-blind nature of the CVPR review process was maintained throughout.

This year we have expanded the format of the technical program in two important ways. First, there are three parallel oral sessions for the first two days of the conference. While CVPR has grown immensely over the years, it has not increased the number of parallel oral tracks for more than two decades. Expanding from two to three tracks has enabled us to have a higher combined percentage of long and spotlight oral presentations. Second, we have continued the 2016 innovation of four days for the main conference, instead of three, but reduced the third day to a half-day to allow time for relaxation and mental regrouping. In recent post-conference surveys, the community has voted overwhelmingly for a four-day main conference.

Continuing the tradition established in CVPR 2016, we are providing an exciting, "trade-show" like atmosphere to foster maximal visibility and exposure for each onsite exhibitor from promising startups and creative standouts to the biggest industry leaders. Over 110 companies are showcasing their technologies at CVPR 2017 and demonstrating the impact that their hardware and software products are
having on a wide range of industries. Many of these companies are sponsoring the conference through a wide range of promotional mechanisms, resulting in a sponsorship funding increase of more than 50% from 2016. The conference would not be possible in its current form without the generosity of our corporate partners and their support is deeply appreciated.

Last but not least, we wish to thank all members of the Organizing Committee, the Area Chairs, reviewers, emergency reviewers and authors for the immense amount of hard work and professionalism that has gone into making CVPR 17 one of the most important venues in Computer Science. Our thanks also go to the organizers of previous CVPRs, many of whom provided helpful advice and guidance. The organizers are particularly indebted to Jana Košecká and René Vidal for patiently answering many questions. Critical aspects of the paper review process were handled using Microsoft’s CMT system, the Toronto Paper Matching System and Researcher.cc and we would like to thank everyone who works on those project teams. We are particularly grateful to Laurent Charlin from TPMS and Ari Kobren from Researcher.cc for their role in helping us to handle the increased scale of our conference. Once again Eric Mortensen has done the seemingly impossible in pulling together the many aspects of the publication process and we are very grateful for all of his hard work. We also want to thank Nicole Finn and her staff at C to C Events for the crucial organizational support that allowed us to stage this conference.

Finally, we wish all the attendees a highly stimulating, informative, and enjoyable conference.

Enjoy CVPR 2017 and Hawaiian hospitality!

Program Chairs: Yanxi Liu, James M. Rehg, Camillo J. Taylor, Ying Wu
General Chairs: Rama Chellappa, Anthony Hoogs, Zhengyou Zhang

QR Codes for the IEEE CPS Mobile App

Apple iTunes App Store (iOS)

Google Play (Android)
# CVPR 2017 Organizing Committee

**General Chairs:**
- Rama Chellappa
- Zhengyou Zhang
- Anthony Hoogs

**Program Chairs:**
- Jim Rehg
- Yanxi Liu
- Ying Wu
- Camillo Taylor

**Workshops Chairs:**
- Jason Corso
- Mei Chen

**Tutorials Chairs:**
- Robert Pless
- David Crandall

**Finance Chair:**
- Walter Scheirer

**Demos Chair:**
- Le Lu

**Publications Chair:**
- Eric Mortensen

**Website Chair:**
- Junsong Yuan

**Corporate Relations Chairs:**
- Matt Turek
- Shuicheng Yan
- Mei Han

**Doctoral Consortium Chairs:**
- Adriana Kovashka
- Nathan Jacobs

**Technology Chair:**
- Terry Boult

**Publicity Chair:**
- Yihong Gong

**Student Volunteers Chair:**
- Amit Roy-Chowdhury

**Logistics Advisor:**
- Ginger Boult

**Program Coordination Lead:**
- Christopher Funk

# CVPR 2017 Area Chairs

<table>
<thead>
<tr>
<th>Area Chairs</th>
<th>Area Chairs</th>
<th>Area Chairs</th>
<th>Area Chairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shai Avidan</td>
<td>Andreas Geiger</td>
<td>Simon Lucey</td>
<td>Qi Tian</td>
</tr>
<tr>
<td>Dhruv Batra</td>
<td>Ross Girshick</td>
<td>Jiebo Luo</td>
<td>Sinisa Todorovic</td>
</tr>
<tr>
<td>Margrit Betke</td>
<td>Bohyung Han</td>
<td>Dimitris Metaxas</td>
<td>Carlo Tomasi</td>
</tr>
<tr>
<td>Gabriel Brostow</td>
<td>Martial Hebert</td>
<td>Greg Mori</td>
<td>Lorenzo Torresani</td>
</tr>
<tr>
<td>Octavia Camps</td>
<td>Gang Hua</td>
<td>PJ Narayanan</td>
<td>Alexander Toshev</td>
</tr>
<tr>
<td>Peter Carr</td>
<td>David Jacobs</td>
<td>Ko Nishino</td>
<td>Matthew Turk</td>
</tr>
<tr>
<td>Xilin Chen</td>
<td>Jiaya Jia</td>
<td>Nikos Paragios</td>
<td>René Vidal</td>
</tr>
<tr>
<td>Bob Collins</td>
<td>Neel Joshi</td>
<td>Devi Parikh</td>
<td>Jingdong Wang</td>
</tr>
<tr>
<td>Jason Corso</td>
<td>Sing Bing Kang</td>
<td>Pietro Perona</td>
<td>Xiaogang Wang</td>
</tr>
<tr>
<td>David Crandall</td>
<td>Kris Kitani</td>
<td>Hamed Pirsiavash</td>
<td>Daphna Weinshall</td>
</tr>
<tr>
<td>Kristin Dana</td>
<td>Pushmeet Kohli</td>
<td>Robert Pless</td>
<td>Lior Wolf</td>
</tr>
<tr>
<td>Kostas Daniilidis</td>
<td>Nikos Komodakis</td>
<td>Marc Pollefeys</td>
<td>Jianxin Wu</td>
</tr>
<tr>
<td>Rogerio Feris</td>
<td>Pawan Kumar</td>
<td>Amit Roy-Chowdhury</td>
<td>Ruigang Yang</td>
</tr>
<tr>
<td>Cornelia Fermuller</td>
<td>Kyros Kutulakos</td>
<td>Imari Sato</td>
<td>Zhaozheng Yin</td>
</tr>
<tr>
<td>Vittorio Ferrari</td>
<td>Erik Learned-Miller</td>
<td>Yoichi Sato</td>
<td>Jingyi Yu</td>
</tr>
<tr>
<td>Sanja Fidler</td>
<td>Honglak Lee</td>
<td>Daniel Scharstein</td>
<td>Junsong Yuan</td>
</tr>
<tr>
<td>David Forsyth</td>
<td>Fuxin Li</td>
<td>Stan Sclaroff</td>
<td>Cha Zhang</td>
</tr>
<tr>
<td>Charless Fowlkes</td>
<td>Ce Liu</td>
<td>Nicu Sebe</td>
<td>S. Kevin Zhou</td>
</tr>
<tr>
<td>Dieter Fox</td>
<td>Xiaoming Liu</td>
<td>Greg Shakharovich</td>
<td>Song-Chun Zhu</td>
</tr>
<tr>
<td>William Freeman</td>
<td>Zicheng Liu</td>
<td>Eli Shechtman</td>
<td></td>
</tr>
</tbody>
</table>
We are pleased to recognize the following researchers as "CVPR 2017 Outstanding Reviewers". These reviewers were identified by one or more of the CVPR Area Chairs for their hard work in providing high quality and detailed reviews for their assigned papers.

Radhakrishna Achanta
Aishwarya Agrawal
Xavier Alameda-Pineda
Jose M. Alvarez
Yannis Avrithis
Xiang Bai
Ohad Ben-Shahar
Katie Bouman
Stefan Breuers
Michael Brown
Thomas Brox
Dylan Campbell
Ayan Chakrabarti
Anoop Cherian
Sunghyun Cho
Yung-Yu Chuang
Michael Cogswell
Daniel Cremers
Marco Cristani
Abhishek Das
Luca del Pero
Sergio Escalera
Victor Escorcia
Francisco Estrada
Boris Flach
Wolfgang Foerstner
Victor Fragoso
Orazio Gallo
Ioannis Gkioulekas
Georgia Gkioxari
Michael Goesele
Boqing Gong
Yahong Han
Christian Häne
Christian Hoffmann
Kaiming He
Jia-Bin Huang
Go Irie
Nathan Jacobs
Varun Jampani
C.V. Jawahar
Laszlo Jeni
Kevin Karsch
Laurent Kneip
Piotr Koniusz
Adriana Kovashka
Abhijit Kundu
Jean-Francois Lalonde
Diane Larlus
Stefan Lee
Yin Li
Chia-Kai Liang
Yen-Yu Lin
Jonathan Long
Chen Change Loy
Mohammed Mahoor
Michael Maire
Yasushi Makihara
Mateusz Malinowski
Renaud Marlet
Stefan Mathe
Mason McGill
Thomas Mensink
Phillippos Mordohai
Roozbeh Mottaghi
Nikhil Naik
Martin R. Oswald
Tomas Pajdla
Vishal Patel
Ioannis Patras
Bojan Popović
Rene Ranftl
Michal Reinstein
Jerome Revaud
Elisa Ricci
Gernot Riegler
Ergys Ristani
Emanuele Rodola
Gregory Rogez
Marcus Rohrbach
Matteo Ronchi
Samuel Rota Bulò
Stefan Roth
Fereshteh Sadeghi
Torsten Sattler
Johannes Schoenberger
Alex Schwing
Amir Shahroury
Chunhua Shen
Boxin Shi
Arnold Smeulders
Yale Song
Alexander Sorkine-Hornung
Hang Su
Shuochen Su
Yusuke Sugano
David Suter
Chris Sweeney
Yuichi Taguchi
Gabriel Taubin
Joseph Tighe
Federico Tombari
Akihiko Torii
Georgios Tzimiropoulos
Osman Ulusoy
Michel Valstar
Jan van Gemert
Grant van Horn
Ramakrishna Vedantam
Christoph Vogel
Christian Vogler
Catherine Wah
Limin Wang
Wei Wang
Xiu-Shen Wei
Kyle Wilson
David Wipf
Tianfu Wu
Lingxi Xie
Kota Yamaguchi
Jinwei Ye
Ryo Yonetani
Lu Yuan
Stefanos Zafeiriou
Jianming Zhang
Lei Zhang
Wei-Shi Zheng
Jun-yan Zhu
Yixin Zhu
Silva Zuffi
Saturday, July 22

0700–1700 Registration (Main Lobby)

0730–0830 Breakfast (Kamehameha II)

0830–0845 Opening Remarks & Paper Awards (Kamehameha III)

0900–1030 Session 1-1A: Machine Learning 1 (Kamehameha III)

Papers in this session are also in Poster Session P1-1.

Chairs: Sanja Fidler (Univ. of Toronto) 
Kris Kitani (Carnegie Mellon Univ.)

0900 Spotlights (S1-1A)
Format (4 min. for presentation; no questions)
1. [0900] Exclusivity-Consistency Regularized Multi-View Subspace Clustering, Xiaobo Wang, Xiaojie Guo, Zhen Lei, Changlee Zhang, Stan Z. Li


4. [0912] Dynamic Edge-Conditioned Filters in Convolutional Neural Networks on Graphs, Martin Simonovsky, Nikos Komodakis

5. [0916] Convolutional Neural Network Architecture for Geometric Matching, Ignacio Rocco, Relja Arandjelović, Josef Sivic

6. [0920] Deep Affordance-Grounded Sensorimotor Object Recognition, Spyridon Thermos, Georgios Th. Papadopoulos, Petros Daras, Gerasimos Potamianos


8. [0928] On Compressing Deep Models by Low Rank and Sparse Decomposition, Xiuyu Yu, Tongliang Liu, Xinchao Wang, Dacheng Tao

0933 Orals (O1-1A)
Format (12 min. for presentation + 2 min. for questions)


10. [0947] Universal Adversarial Perturbations, Seyed-Mohsen Moosavi-Dezfooli, Alhussein Fawzi, Omar Fawzi, Pascal Frossard

11. [1001] Unsupervised Pixel-Level Domain Adaptation With Generative Adversarial Networks, Konstantinos Bousmalis, David Dohan, Dumitru Erhan, Dilip Krishnan


0900–1030 Session 1-1B: 3D Vision 1 (Kalākaua Ballroom A-B)

Papers in this session are also in Poster Session P1-1.

Chairs: Andreas Geiger (MPI Tübingen) 
Ruigang Yang (Univ. of Kentucky)

0900 Spotlights (S1-1B)
Format (4 min. for presentation; no questions)


16. [0912] Elastic Shape-From-Template With Spatially Sparse Deforming Forces, Abed Malti, Cédric Herzet
17. [0916] Distinguishing the Indistinguishable: Exploring Structural Ambiguities via Geodesic Context, Qingan Yan, Long Yang, Ling Zhang, Chunxia Xiao
18. [0920] Multi-Scale Continuous CRFs as Sequential Deep Networks for Monocular Depth Estimation, Dan Xu, Elisa Ricci, Wanli Ouyang, Xiaogang Wang, Nicu Sebe
19. [0924] Dynamic Time-Of-Flight, Michael Schober, Amit Adam, Omer Yair, Shai Mazor, Sebastian Nowozin
20. [0928] Training Object Class Detectors With Click Supervision, Dim P. Papadopoulos, Jasper R. R. Uijlings, Frank Keller, Vittorio Ferrari

0933 Orals (O1-1B)
Format (12 min. for presentation + 2 min. for questions)

0900–1030 Session 1-1C: Low- & Mid-Level Vision (Kalākaua Ballroom C)
Papers in this session are also in Poster Session P1-1.

Chairs: Octavia Camps (Northeastern Univ.)
David Jacobs (Univ. of Maryland)

0900 Spotlights (S1-1C)
Format (4 min. for presentation; no questions)
25. [0900] Designing Effective Inter-Pixel Information Flow for Natural Image Matting, Yaşğ Aksoy, Tunç Özan Aydin, Marc Pollefeys

27. [0908] Instance-Level Salient Object Segmentation, Guanbin Li, Yuan Xie, Liang Lin, Yizhou Yu
28. [0912] Deep Multi-Scale Convolutional Neural Network for Dynamic Scene Deblurring, Seungjun Nah, Tae Hyun Kim, Kyoung Mu Lee
29. [0916] Diversified Texture Synthesis With Feed-Forward Networks, Yijun Li, Chen Fang, Jimei Yang, Zhaowen Wang, Xin Lu, Ming-Hsuan Yang
30. [0920] Radiometric Calibration for Internet Photo Collections, Zihipeng Mo, Boxin Shi, Sai-Kit Yeung, Yasuyuki Matsushita
31. [0924] Deeply Aggregated Alternating Minimization for Image Restoration, Youngjung Kim, Hyungjoo Jung, Dongbo Min, Kwanghoon Sohn
32. [0928] End-To-End Instance Segmentation With Recurrent Attention, Mengye Ren, Richard S. Zemel

0933 Orals (O1-1C)
Format (12 min. for presentation + 2 min. for questions)
33. [0933] SRN: Side-output Residual Network for Object Symmetry Detection in the Wild, Wei Ke, Jie Chen, Jianbin Jiao, Guoying Zhao, Qixiang Ye
34. [0947] Deep Image Matting, Ning Xu, Brian Price, Scott Cohen, Thomas Huang
36. [1015] FC4: Fully Convolutional Color Constancy With Confidence-Weighted Pooling, Yuanming Hu, Baoyuan Wang, Stephen Lin

1030–1115 Break (Kamehameha II)
Saturday, July 22 (Morning)

Program

1030–1230 Poster Session P1-1 (Kamehameha I)

3D Computer Vision
37. Face Normals “In-The-Wild” Using Fully Convolutional Networks, George Trigeorgis, Patrick Snape, Iasonas Kokkinos, Stefanos Zafeiriou
38. A Non-Convex Variational Approach to Photometric Stereo Under Inaccurate Lighting, Yvain Quéau, Tao Wu, François Lauze, Jean-Denis Durou, Daniel Cremers
39. A Linear Extrinsic Calibration of Kaleidoscopic Imaging System From Single 3D Point, Kosuke Takahashi, Akihiro Miyata, Shohei Nobuhara, Takashi Matsuyama
40. Polarimetric Multi-View Stereo, Zhaopeng Cui, Jinwei Gu, Boxin Shi, Ping Tan, Jan Kautz
41. An Exact Penalty Method for Locally Convergent Maximum Consensus, Huu Le, Tat-Jun Chin, David Suter
42. Deep Supervision With Shape Concepts for Occlusion-Aware 3D Object Parsing, Chi Li, M. Zeeshan Zia, Quoc-Huy Tran, Xiang Yu, Gregory D. Hager, Manmohan Chandraker
43. Amodal Detection of 3D Objects: Inferring 3D Bounding Boxes From 2D Ones in RGB-Depth Images, Zhuo Deng, Longin Jan Latecki

Analyzing Humans in Images
44. Transition Forests: Learning Discriminative Temporal Transitions for Action Recognition and Detection, Guillermo García-Hernando, Tae-Kyun Kim
45. Scene Flow to Action Map: A New Representation for RGB-D Based Action Recognition With Convolutional Neural Networks, Pichao Wang, Wanqing Li, Zhimin Gao, Yuyao Zhang, Chang Tang, Philip Oggunbona
46. Detecting Masked Faces in the Wild With LLE-CNNs, Shiming Ge, Jia Li, Qiting Ye, Zhao Luo
47. A Domain Based Approach to Social Relation Recognition, Qianru Sun, Bernt Schiele, Mario Fritz
48. Spatio-Temporal Naive-Bayes Nearest-Neighbor (ST-NBNN) for Skeleton-Based Action Recognition, Junwu Weng, Chaoinq Weng, Junsong Yuan
49. Personalizing Gesture Recognition Using Hierarchical Bayesian Neural Networks, Ajjen Joshi, Soumya Ghosh, Margrit Betke, Stan Sclaroff, Hanspeter Pfister

Applications
50. Real-Time 3D Model Tracking in Color and Depth on a Single CPU Core, Wadim Kehl, Federico Tombari, Slobodan Ilic, Nassir Navab
51. Multi-Scale FCN With Cascaded Instance Aware Segmentation for Arbitrary Oriented Word Spotting in the Wild, Dafang He, Xiao Yang, Chen Liang, Zihan Zhou, Alexander G. Ororbi II, Daniel Kifer, C. Lee Giles
52. Viraliency: Pooling Local Virality, Xavier Alameda-Pineda, Andrea Pilzer, Dan Xu, Nicu Sebe, Elisa Ricci

Biomedical Image/Video Analysis
53. A Non-Local Low-Rank Framework for Ultrasound Speckle Reduction, Lei Zhu, Chi-Wing Fu, Michael S. Brown, Pheng-Ann Heng

Image Motion & Tracking
54. Video Acceleration Magnification, Yichao Zhang, Silvia L. Pintea, Jan C. van Gemert
55. Superpixel-Based Tracking-By-Segmentation Using Markov Chains, Donghun Yeo, Jeany Son, Bohyung Han, Joon Hee Han
56. BranchOut: Regularization for Online Ensemble Tracking With Convolutional Neural Networks, Bohyung Han, Jack Sim, Hartwig Adam
57. Learning Motion Patterns in Videos, Pavel Tokmakov, Karteeb Alahari, Cordelia Schmid

Low- & Mid-Level Vision
58. Deep Level Sets for Salient Object Detection, Ping Hu, Bing Shuai, Jun Liu, Gang Wang
59. Binary Constraint Preserving Graph Matching, Bo Jiang, Jun Tang, Chris Ding, Bin Luo
60. From Local to Global: Edge Profiles to Camera Motion in Blurred Images, Subeesh Vasu, A. N. Rajagopalan
61. What Is the Space of Attenuation Coefficients in Underwater Computer Vision? Derya Akkaynak, Tali Treibitz, Tom Shlesinger, Yossi Loya, Raz Tamir, David Iluz
62. Robust Energy Minimization for BRDF-Invariant Shape From Light Fields, Zhengqin Li, Zexiang Xu, Ravi Ramamoorthi, Manmohan Chandraker
63. Boundary-Aware Instance Segmentation, Zeeshan Hayder, Xuming He, Mathieu Salzmann
Spatially-Varying Blur Detection Based on Multiscale Fused and Sorted Transform Coefficients of Gradient Magnitudes, S. Alireza Golestaneh, Lina J. Karam

Model-Based Iterative Restoration for Binary Document Image Compression With Dictionary Learning, Yandong Guo, Cheng Lu, Jan P. Allebach, Charles A. Bouman

FCSS: Fully Convolutional Self-Similarity for Dense Semantic Correspondence, Seungryong Kim, Dongbo Min, Bumsub Ham, Sangryul Jeon, Stephen Lin, Kwanghoon Sohn

Machine Learning

Learning by Association — A Versatile Semi-Supervised Training Method for Neural Networks, Philip Haeusser, Alexander Mordvintsev, Daniel Cremers

Dilated Residual Networks, Fisher Yu, Vladlen Koltun, Thomas Funkhouser

Split-Brain Autoencoders: Unsupervised Learning by Cross-Channel Prediction, Richard Zhang, Phillip Isola, Alexei A. Efros

Nonnegative Matrix Underapproximation for Robust Multiple Model Fitting, Mariano Tepper, Guillermo Sapiro

Truncated Max-Of-Convex Models, Pankaj Parsani, M. Pawan Kumar

Additive Component Analysis, Calvin Murdock, Fernando De la Torre

Subspace Clustering via Variance Regularized Ridge Regression, Chong Peng, Zhao Kang, Qiang Cheng

The Incremental Multiresolution Matrix Factorization Algorithm, Vamsi K. Ithapu, Risi Kondor, Sterling C. Johnson, Vikas Singh


A-Lamp: Adaptive Layout-Aware Multi-Patch Deep Convolutional Neural Network for Photo Aesthetic Assessment, Shuang Ma, Jing Liu, Chang Wen Chen

Teaching Compositionality to CNNs, Austin Stone, Huayan Wang, Michael Stark, Yi Liu, D. Scott Phoenix, Dileep George

Object Recognition & Scene Understanding

Physics Inspired Optimization on Semantic Transfer Features: An Alternative Method for Room Layout Estimation, Hao Zhao, Ming Lu, Anbang Yao, Yiwen Guo, Yurong Chen, Li Zhang
93. Pixelwise Instance Segmentation With a Dynamically Instantiated Network, Anurag Arnab, Philip H. S. Torr

94. Object Detection in Videos With Tubelet Proposal Networks, Kai Kang, Hongsheng Li, Tong Xiao, Wanli Ouyang, Junjie Yan, Xihui Liu, Xiaogang Wang

95. AMVH: Asymmetric Multi-Valued Hashing, Cheng Da, Shibiao Xu, Kun Ding, Gaofeng Meng, Shiming Xiang, Chunhong Pan

96. Spindle Net: Person Re-Identification With Human Body Region Guided Feature Decomposition and Fusion, Haiyu Zhao, Maoqing Tian, Shuyang Sun, Jing Shao, Junjie Yan, Shuai Yi, Xiaogang Wang, Xiaou Tang

97. Deep Visual-Semantic Quantization for Efficient Image Retrieval, Yue Cao, Mingsheng Long, Jianmin Wang, Shichen Liu

98. Efficient Diffusion on Region Manifolds: Recovering Small Objects With Compact CNN Representations, Ahmet Iscen, Giorgos Tolias, Yannis Avrithis, Teddy Furon, Ondřej Chum

99. Feature Pyramid Networks for Object Detection, Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, Serge Belongie

100. Mind the Class Weight Bias: Weighted Maximum Mean Discrepancy for Unsupervised Domain Adaptation, Hongliang Yan, Yukang Ding, Peihua Li, Qilong Wang, Yong Xu, Wangmeng Zuo

101. StyleNet: Generating Attractive Visual Captions With Styles, Chuang Gan, Zhe Gan, Xiaodong He, Jianfeng Gao, Li Deng


103. Improving Interpretability of Deep Neural Networks With Semantic Information, Yinpeng Dong, Hang Su, Jun Zhu, Bo Zhang

104. Video Captioning With Transferred Semantic Attributes, Yingwei Pan, Ting Yao, Houqiang Li, Tao Mei

105. Fast Boosting Based Detection Using Scale Invariant Multimodal Multiresolution Filtered Features, Arthur Daniel Costea, Robert Varga, Sergiu Nedevschi

106. Temporal Convolutional Networks for Action Segmentation and Detection, Colin Lea, Michael D. Flynn, René Vidal, Austin Reiter, Gregory D. Hager

107. Surveillance Video Parsing With Single Frame Supervision, Si Liu, Changhu Wang, Ruihe Qian, Han Yu, Renda Bao, Yao Sun

108. Weakly Supervised Actor-Action Segmentation via Robust Multi-Task Ranking, Yan Yan, Chenliang Xu, Dawen Cai, Jason J. Corso

109. Unsupervised Visual-Linguistic Reference Resolution in Instructional Videos, De-An Huang, Joseph J. Lim, Li Fei-Fei, Juan Carlos Niebles

110. Zero-Shot Action Recognition With Error-Correcting Output Codes, Jie Qin, Li Liu, Ling Shao, Fumin Shen, Bingbing Ni, Jiaxin Chen, Yunhong Wang

111. Enhancing Video Summarization via Vision-Language Embedding, Bryan A. Plummer, Matthew Brown, Svetlana Lazebnik

112. Synthesizing Dynamic Patterns by Spatial-Temporal Generative ConvNet, Jianwen Xie, Song-Chun Zhu, Ying Nian Wu

1030–1230 Demos (Kamehameha I)

- Tensorflow Object Detection API, Jonathan Huang, Vivek Rathod, Chen Sun, Sergio Guadarrama, Tyler Zhu, George Papandreou, Menglong Zhu, Alienza Fathi, Derek Chow, Kevin Murphy (Google)
- DenseReg: Fully Convolutional Dense Shape Regression In-The-Wild, Riza Alp Güler, George Trigeorgis, Epameinondas Antonakos, Patrick Snape, Stefanos Zafeiriou, Iasonas Kokkinos (INRIA/CentraleSupélec, Imperial College, Univ. College London/Facebook AI Research)
- UnrealCV: Connecting Computer Vision to Unreal Engine, Weichao Qiu, Yi Zhang, Fangwei Zhong, Siyuan Qiao, Yizhou Wang, Alan Yuille (Johns Hopkins Univ., Tsinghua Univ., Peking Univ.)
- Minimum Delay Moving Object Detection, Samim Zahoor, Dong Lao, Ganesh Sundaramoorthi (King Abdullah Univ. of Science and Technology (KAUST))
### 1030–1230 Exhibitors (Kamehameha I)

Booth # in parenthesis—see map on next page (pg. 11).

**Platinum Level**
- A9 (305)
- Alibaba (116)
- Amazon (112)
- Apple (117)
- DeepGlint & UISEE (405)
- Didi Chuxing (605)
- Facebook (205)
- Google (106)
- Intel Nervana (635)

**Gold Level**
- Adobe Research (737)
- Argo AI (705)
- Bosch (856)
- Cognex (727)
- CVTE (222)
- DJI (707)
- HERE (135)
- HiScene (647)
- IBM (230)
- iRobot (127)
- Malong (226)

**Silver Level**
- AIMATTER (749)
- Datatang (246)
- Disney Research (147)
- Honda Research Inst. (853)
- Kitware (242)
- Pinterest (143)

**Bronze Level**
- 3dMD (857)
- AiCure (759)
- AltumView Systems (157)
- Axon (750)
- Body Labs (238)
- CloudSight (649)
- CrowdFlower (849)
- Cruise Automation (254)
- DeepVision (725)
- Elsevier (159)
- FLIR Systems (250)
- GumGum (653)

**Start-Up Level**
- Anantak (746)
- Augmented Pixels (751)
- AutoX (561)
- DeepCognAI (661)
- EgoVid (743)
- Eyenuk (846)
- Eyeris (651)
- FeatureX (740)
- iniLabs (659)
- ISEE AI (761)
- Markable (461)
- Morgan & Claypool (742)
- Morpx (760)
- MUKH (719)

**Non-Profit Level**
- NextAI (723)

### 1230–1330 Lunch (Kamehameha II)

**Notes:**

- Horizon Robotics (149)
- KAUST (258)
- Linkface (748)
- Lunit (240)
- Mapillary (137)
- MathWorks (153)
- MatterPort (141)
- NetPosa (847)
- Synaptics (755)
- Tesla (729)
- TouTiao/ByteDance (722)
- Vipshop US (655)
- Yandex (709)
- Zillow (151)
- Zoox (715)
- Octi (717)
- Perceptive Automata (724)
- Pixm (731)
- Saikou Optics (754)
- Shopagon (758)
- Speechocean (851)
- Spotscale (739)
- Surfing Tech (845)
- TuringVideo (563)
- Umbo CV (657)
- VanGogh Imaging (753)
- Watrix Technology (660)
- WRNCH (641)
- Samasource (735)
1330–1500 Session 1-2A: Object Recognition & Scene Understanding - Computer Vision & Language (Kamehameha III)

Papers in this session are also in Poster Session P1-2.

Chairs: Devi Parikh (Georgia Tech)
Nicu Sebe (Univ. of Trento)

1330 Spotlights (S1-2A)
Format (4 min. for presentation; no questions)
1. [1330] Context-Aware Captions From Context-Agnostic Supervision, Ramakrishna Vedantam, Samy Bengio, Kevin Murphy, Devi Parikh, Gal Chechik
2. [1334] Visual Dialog, Abhishek Das, Satwik Kottur, Khushi Gupta, Avi Singh, Deshraj Yadav, José M. F. Moura, Devi Parikh, Dhruv Batra
3. [1338] Discriminative Bimodal Networks for Visual Localization and Detection With Natural Language Queries, Yuting Zhang, Luyao Yuan, Yijie Guo, Zhiyuan He, I-An Huang, Honglak Lee
4. [1342] Automatic Understanding of Image and Video Advertisements, Zaeem Hussain, Mingda Zhang, Xiaozhong Zhang, Keren Ye, Christopher Thomas, Zuha Agha, Nathan Ong, Adriana Kovashka
5. [1346] Discover and Learn New Objects From Documentaries, Kai Chen, Hang Song, Chen Change Loy, Dahua Lin
7. [1354] Fully-Adaptive Feature Sharing in Multi-Task Networks With Applications in Person Attribute Classification, Yongxi Lu, Abhishek Kumar, Shuangfei Zhai, Yu Cheng, Tara Javidi, Rogerio Feris
8. [1358] Semantic Compositional Networks for Visual Captioning, Zhe Gan, Chuang Gan, Xiaodong He, Yunchen Pu, Kenneth Tran, Jianfeng Gao, Lawrence Carin, Li Deng

1403 Orals (O1-2A)
Format (12 min. for presentation + 2 min. for questions)
9. [1403] Deep Reinforcement Learning-Based Image Captioning With Embedding Reward, Zhou Ren, Xiaoyu Wang, Ning Zhang, Xutao Lv, Li-Jia Li

11. [1431] Captioning Images With Diverse Objects, Subhashini Venugopalan, Lisa Anne Hendricks, Marcus Rohrbach, Raymond Mooney, Trevor Darrell, Kate Saenko
12. [1445] Self-Critical Sequence Training for Image Captioning, Steven J. Rennie, Etienne Marcheret, Youssef Mroueh, Jerret Ross, Vaibhava Goel

1330–1500 Session 1-2B: Analyzing Humans 1
(Kalākaua Ballroom A-B)

Papers in this session are also in Poster Session P1-2.

Chairs: Greg Mori (Simon Fraser Univ.)
Amit Roy-Chowdhury (UC Riverside)

1330 Spotlights (S1-2B)
Format (4 min. for presentation; no questions)
13. [1330] Crossing Nets: Combining GANs and VAEs With a Shared Latent Space for Hand Pose Estimation, Chengde Wan, Thomas Probst, Luc Van Gool, Angela Yao
14. [1334] Predicting Behaviors of Basketball Players From First Person Videos, Shan Su, Jung Pyo Hong, Jianbo Shi, Hyun Soo Park
15. [1338] LCR-Net: Localization-Classification-Regression for Human Pose, Grégory Rogez, Philippe Weinzaepfel, Cordelia Schmid
16. [1342] Learning Residual Images for Face Attribute Manipulation, Wei Shen, Rujie Liu
1403 Orals (O1-2B)

Format (12 min. for presentation + 2 min. for questions)
22. [1417] Disentangled Representation Learning GAN for Pose-Invariant Face Recognition, Luan Tran, Xi Yin, Xiaoming Liu
24. [1445] Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields, Zhe Cao, Tomas Simon, Shih-En Wei, Yaser Sheikh

1330–1500 Session 1-2C: Image Motion & Tracking; Video Analysis
(Kalākaua Ballroom C)

Papers in this session are also in Poster Session P1-2.

Chairs: Bob Collins (Pennsylvania State Univ.)
René Vidal (Johns Hopkins Univ.)

1330 Spotlights (S1-2C)

Format (4 min. for presentation; no questions)
25. [1330] Template Matching With Deformable Diversity Similarity, Itamar Talmi, Roey Mechrez, Lihi Zelnik-Manor
26. [1334] Beyond Triplet Loss: A Deep Quadruplet Network for Person Re-Identification, Weihua Chen, Xiaotang Chen, Jianguo Zhang, Kaiqi Huang
27. [1338] Agent-Centric Risk Assessment: Accident Anticipation and Risky Region Localization, Kuo-Hao Zeng, Shih-Han Chou, Fu-Hsiang Chan, Juan Carlos Niebles, Min Sun
29. [1346] Action-Decision Networks for Visual Tracking With Deep Reinforcement Learning, Sangdoo Yun, Jongwon Choi, Youngjoon Yoo, Kimin Yun, Jin Young Choi
31. [1354] Making 360° Video Watchable in 2D: Learning Videography for Click Free Viewing, Yu-Chuan Su, Kristen Grauman

1403 Orals (O1-2C)

Format (12 min. for presentation + 2 min. for questions)
33. [1403] Context-Aware Correlation Filter Tracking, Matthias Mueller, Neil Smith, Bernhard Ghanem
34. [1417] Deep 360 Pilot: Learning a Deep Agent for Piloting Through 360° Sports Videos, Hou-Ning Hu, Yan-Chen Lin, Ming-Yu Liu, Hsien-Tzu Cheng, Yung-Ju Chang, Min Sun

1500–1545 Break (Kamehameha II)

1500–1700 Poster Session P1-2 (Kamehameha I)

3D Computer Vision

37. Exploiting 2D Floorplan for Building-Scale Panorama RGBD Alignment, Erik Wijmans, Yasutaka Furukawa
38. A Combinatorial Solution to Non-Rigid 3D Shape-To-Image Matching, Florian Bernard, Frank R. Schmidt, Johan Thunberg, Daniel Cremers
39. NID-SLAM: Robust Monocular SLAM Using Normalised Information Distance, Geoffrey Pascoe, Will Maddern, Michael Tanner, Pedro Piniés, Paul Newman
41. Learning Shape Abstractions by Assembling Volumetric Primitives, Shubham Tulsiani, Hao Su, Leonidas J. Guibas, Alexei A. Efros, Jitendra Malik
42. Locality-Sensitive Deconvolution Networks With Gated Fusion for RGB-D Indoor Semantic Segmentation, Yanhua Cheng, Rui Cai, Zhiwei Li, Xin Zhao, Kaiqi Huang

43. Acquiring Axially-Symmetric Transparent Objects Using Single-View Transmission Imaging, Jaewon Kim, Ilya Reshetouski, Abhijeet Ghosh

44. Regressing Robust and Discriminative 3D Morhable Models With a Very Deep Neural Network, Anh Tuấn Trần, Tal Hassner, Iacopo Masi, Gérard Medioni

45. End-To-End 3D Face Reconstruction With Deep Neural Networks, Pengfei Dou, Shishir K. Shah, Ioannis A. Kakadiaris

46. DUST: Dual Union of Spatio-Temporal Subspaces for Monocular Multiple Object 3D Reconstruction, Antonio Agudo, Francesc Moreno-Noguer

Analyzing Humans in Images

47. Finding Tiny Faces, Peiyun Hu, Deva Ramanan

48. Dynamic Facial Analysis: From Bayesian Filtering to Recurrent Neural Network, Jinwei Gu, Xiaodong Yang, Shalini De Mello, Jan Kautz


50. Joint Registration and Representation Learning for Unconstrained Face Identification, Munawar Hayat, Salman H. Khan, Naoufel Werghi, Roland Goecke

51. 3D Human Pose Estimation From a Single Image via Distance Matrix Regression, Francesc Moreno-Noguer

52. One-Shot Metric Learning for Person Re-Identification, Slawomir Bąk, Peter Carr

53. Generalized Rank Pooling for Activity Recognition, Anoop Cherian, Basura Fernando, Mehrtash Harandi, Stephen Gould

54. Deep Representation Learning for Human Motion Prediction and Classification, Judith Bütepage, Michael J. Black, Danica Kragic, Hedvig Kjellström

55. Interspecies Knowledge Transfer for Facial Keypoint Detection, Maheen Rashid, Xiuye Gu, Yong Jae Lee

56. Recurrent Convolutional Neural Networks for Continuous Sign Language Recognition by Staged Optimization, Runpeng Cui, Hu Liu, Changshui Zhang

Applications

57. Modeling Sub-Event Dynamics in First-Person Action Recognition, Hasan F. M. Zaki, Faisal Shafait, Ajmal Mian

Computational Photography

58. Turning an Urban Scene Video Into a Cinemagraph, Hang Yan, Yebin Liu, Yasutaka Furukawa

59. Light Field Reconstruction Using Deep Convolutional Network on EPI, Gaochang Wu, Mandan Zhao, Liangyong Wang, Qionghai Dai, Tianyou Chai, Yebin Liu

Image Motion & Tracking


Low- & Mid-Level Vision

61. Attention-Aware Face Hallucination via Deep Reinforcement Learning, Qingxing Cao, Liang Lin, Yukai Shi, Xiaodan Liang, Guanbin Li

62. Simple Does It: Weakly Supervised Instance and Semantic Segmentation, Anna Khoreva, Rodrigo Benenson, Jan Hosang, Matthias Hein, Bernt Schiele

63. Anti-Glare: Tightly Constrained Optimization for Eyeglass Reflection Removal, Tushar Sandhan, Jin Young Choi

64. Deep Joint Rain Detection and Removal From a Single Image, Wenhan Yang, Robby T. Tan, Jiashi Feng, Jiaying Liu, Zongming Guo, Shuicheng Yan

65. Radiometric Calibration From Faces in Images, Chen Li, Kun Zhou, Katsushi Ikeuchi

66. Weblly Supervised Semantic Segmentation, Bin Jin, Maria V. Ortiz Segovia, Sabine Süsstrunk

67. Removing Rain From Single Images via a Deep Detail Network, Xueyang Fu, Jiabin Huang, Delu Zeng, Yue Huang, Xinghao Ding, John Paisley

68. Deep Crisp Boundaries, Yupei Wang, Xin Zhao, Kaiqi Huang

69. Coarse-To-Fine Segmentation With Shape-Tailored Continuum Scale Spaces, Naemullah Khan, Byung-Woo Hong, Anthony Yezzi, Ganesh Sundaramoorthi

70. Large Kernel Matters — Improve Semantic Segmentation by Global Convolutional Network, Chao Peng, Xiangyu Zhang, Gang Yu, Guiming Luo, Jian Sun
<table>
<thead>
<tr>
<th>Program</th>
<th>Saturday, July 22 (Afternoon)</th>
</tr>
</thead>
<tbody>
<tr>
<td>71.</td>
<td>Single Image Reflection Suppression, <em>Nikolaos Arvanitopoulos, Radhakrishna Achanta, Sabine Süsstrunk</em></td>
</tr>
<tr>
<td>72.</td>
<td>CASENet: Deep Category-Aware Semantic Edge Detection, <em>Zhiding Yu, Chen Feng, Ming-Yu Liu, Srikrumar Ramalingam</em></td>
</tr>
<tr>
<td>73.</td>
<td>Reflectance Adaptive Filtering Improves Intrinsic Image Estimation, <em>Thomas Nestmeyer, Peter V. Gehler</em></td>
</tr>
<tr>
<td><strong>Machine Learning</strong></td>
<td></td>
</tr>
<tr>
<td>74.</td>
<td>Conditional Similarity Networks, <em>Andreas Veit, Serge Belongie, Theofanis Karaletsos</em></td>
</tr>
<tr>
<td>76.</td>
<td>Xception: Deep Learning With Depthwise Separable Convolutions, <em>François Chollet</em></td>
</tr>
<tr>
<td>78.</td>
<td>Online Summarization via Submodular and Convex Optimization, <em>Ehsan Elhamifar, M. Clara De Paolis Kaluza</em></td>
</tr>
<tr>
<td>79.</td>
<td>Deep MANTA: A Coarse-To-Fine Many-Task Network for Joint 2D and 3D Vehicle Analysis From Monocular Image, <em>Florian Chabot, Mohamed Chaouch, Jaonary Rabarisoa, Céline Teulière, Thierry Chateau</em></td>
</tr>
<tr>
<td>80.</td>
<td>Improving Pairwise Ranking for Multi-Label Image Classification, <em>Yuncheng Li, Yale Song, Jiebo Luo</em></td>
</tr>
<tr>
<td>81.</td>
<td>Active Convolution: Learning the Shape of Convolution for Image Classification, <em>Yunho Jeon, Junmo Kim</em></td>
</tr>
<tr>
<td>82.</td>
<td>Linking Image and Text With 2-Way Nets, <em>Aviv Eisenschtat, Lior Wolf</em></td>
</tr>
<tr>
<td>83.</td>
<td>Stacked Generative Adversarial Networks, <em>Xun Huang, Yixuan Li, Omid Poursaeed, John Hopcroft, Serge Belongie</em></td>
</tr>
<tr>
<td>84.</td>
<td>Image Splicing Detection via Camera Response Function Analysis, <em>Can Chen, Scott McClosey, Jingyi Yu</em></td>
</tr>
<tr>
<td>85.</td>
<td>Building a Regular Decision Boundary With Deep Networks, <em>Edouard Oyallon</em></td>
</tr>
<tr>
<td>88.</td>
<td>Scale-Aware Face Detection, <em>Zekun Hao, Yu Liu, Hongwei Qin, Junjie Yan, Xiu Li, Xiaolin Hu</em></td>
</tr>
<tr>
<td>90.</td>
<td>Generative Hierarchical Learning of Sparse FRAME Models, <em>Jianwen Xie, Yifei Xu, Erik Nijkamp, Ying Nian Wu, Song-Chun Zhu</em></td>
</tr>
<tr>
<td><strong>Object Recognition &amp; Scene Understanding</strong></td>
<td></td>
</tr>
<tr>
<td>91.</td>
<td>Generating Holistic 3D Scene Abstractions for Text-Based Image Retrieval, <em>Ang Li, Jin Sun, Joe Yue-Hei Ng, Ruichi Yu, Vlad I. Morariu, Larry S. Davis</em></td>
</tr>
<tr>
<td>92.</td>
<td>Perceptual Generative Adversarial Networks for Small Object Detection, <em>Jianan Li, Xiaodan Liang, Yunchao Wei, Tingfa Xu, Jiashi Feng, Shuicheng Yan</em></td>
</tr>
<tr>
<td>93.</td>
<td>Emotion Recognition in Context, <em>Ronak Kosti, Jose M. Alvarez, Adria Recasens, Agata Lapedriza</em></td>
</tr>
<tr>
<td>95.</td>
<td>Dense Captioning With Joint Inference and Visual Context, <em>Linjie Yang, Kevin Tang, Jianchao Yang, Li-Jia Li</em></td>
</tr>
<tr>
<td>96.</td>
<td>CLEVR: A Diagnostic Dataset for Compositional Language and Elementary Visual Reasoning, <em>Justin Johnson, Bharath Hariharan, Laurens van der Maaten, Li Fei-Fei, C. Lawrence Zitnick, Ross Girshick</em></td>
</tr>
<tr>
<td>98.</td>
<td>Matrix Tri-Factorization With Manifold Regularizations for Zero-Shot Learning, <em>Xing Xu, Fumin Shen, Yang Yang, Dongxiang Zhang, Heng Tao Shen, Jingkuan Song</em></td>
</tr>
</tbody>
</table>
Saturday, July 22 (Afternoon)

100. Learning Spatial Regularization With Image-Level Supervisions for Multi-Label Image Classification, Feng Zhu, Hongsheng Li, Wanli Ouyang, Nenghai Yu, Xiaogang Wang

101. Semantically Consistent Regularization for Zero-Shot Recognition, Pedro Morgado, Nuno Vasconcelos

102. Can Walking and Measuring Along Chord Bunches Better Describe Leaf Shapes? Bin Wang, Yongsheng Gao, Changming Sun, Michael Blumenstein, John La Salle

**Video Analytics**

103. Self-Learning Scene-Specific Pedestrian Detectors Using a Progressive Latent Model, Qixiang Ye, Tianliang Zhang, Wei Ke, Qiang Qiu, Jie Chen, Guillermo Sapito, Baochang Zhang

104. Predictive-Corrective Networks for Action Detection, Achal Dave, Olga Russakovsky, Deva Ramanan

105. Budget-Aware Deep Semantic Video Segmentation, Behrooz Mahasseni, Sinisa Todorovic, Alan Fern


107. Spatiotemporal Pyramid Network for Video Action Recognition, Yunbo Wang, Mingsheng Long, Jianmin Wang, Philip S. Yu


109. FusionSeg: Learning to Combine Motion and Appearance for Fully Automatic Segmentation of Generic Objects in Videos, Suyog Dutt Jain, Bo Xiong, Kristen Grauman

110. Query-Focused Video Summarization: Dataset, Evaluation, and a Memory Network Based Approach, Aidean Sharghi, Jacob S. Laurel, Boqing Gong

111. Flexible Spatio-Temporal Networks for Video Prediction, Chaochao Lu, Michael Hirsch, Bernhard Schölkopf

112. Temporal Action Co-Segmentation in 3D Motion Capture Data and Videos, Konstantinos Papoutsakis, Costas Panagiotakis, Antonis A. Argyros

---

**Program**

1500–1700 Demos (Kamehameha I)
- Same as Saturday morning Demos (see pg. 9)

1500–1700 Exhibits (Kamehameha I)
- Same as Saturday morning Exhibits (see pg. 10)

1715–1830 Plenary Session (Kamehameha III)
- **Keynote Talk:** Extracting Social Meaning From Language, Dan Jurafsky (Stanford Univ.)
  
  **Abstract:** I describe research in our lab on computationally extracting social meaning from language, meaning that takes into account social relationships between people. I describe our study of interactions between police and community members in traffic stops recorded in body-worn camera footage. We automatically measure the quality of the interaction from language, study the role of race in the interaction, and draw suggestions for going forward in this fraught area. In another we computationally model the language of scientific papers together with the network formed by scientists and their research areas to better understand scientific innovation, how it progresses, and the role of interdisciplinarity. I discuss implications for the history of science and specifically of artificial intelligence. Both studies highlight the importance of social context and social models for interpreting the latent meanings behind the words we use.

1840–2040 PAMI Technical Committee Meeting (Kamehameha III)

---

**Notes:**

---
Sunday, July 23

0730–1700 Registration (Main Lobby)

0730–0830 Breakfast (Kamehameha II)

0830–1000 Session 2-1A: Machine Learning 2
(Kamehameha III)

Papers in this session are also in Poster Session P2-1.

Chairs: Lorenzo Torresani (Dartmouth)
Alexander Toshev (Google)

0830 Spotlights (S2-1A)

Format (4 min. for presentation; no questions)

1. [0830] Dual Attention Networks for Multimodal Reasoning and Matching, Hyeonseob Nam, Jung-Woo Ha, Jeonghee Kim

2. [0834] DESIRE: Distant Future Prediction in Dynamic Scenes With Interacting Agents, Namhoon Lee, Wongun Choi, Paul Vernaza, Christopher B. Choy, Philip H. S. Torr, Manmohan Chandraker

3. [0838] Interpretable Structure-Evolving LSTM, Xiaodan Liang, Liang Lin, Xiaohui Shen, Jiashi Feng, Shuicheng Yan, Eric P. Xing

4. [0842] ShapeOdds: Variational Bayesian Learning of Generative Shape Models, Shireen Elhabian, Ross Whitaker

5. [0846] Fast Video Classification via Adaptive Cascading of Deep Models, Haichen Shen, Seungyeop Han, Matthai Philipose, Arvind Krishnamurthy

6. [0850] Deep Metric Learning via Facility Location, Hyun Oh Song, Stefanie Jegelka, Vivek Rathod, Kevin Murphy

7. [0854] Semi-Supervised Deep Learning for Monocular Depth Map Prediction, Yevhen Kuznietsov, Jörg Stückler, Bastian Leibe

8. [0858] Weakly Supervised Semantic Segmentation Using Web-Crawled Videos, Seunghoon Hong, Donghun Yeo, Suha Kwak, Honglak Lee, Bohyung Han

0903 Orals (O2-1A)

Format (12 min. for presentation + 2 min. for questions)


11. [0931] Inverse Compositional Spatial Transformer Networks, Chen-Hsuan Lin, Simon Lucey

12. [0945] Densely Connected Convolutional Networks, Gao Huang, Zhuang Liu, Laurens van der Maaten, Kilian Q. Weinberger

0830–1000 Session 2-1B: Computational Photography
(Kalākaua Ballroom A-B)

Papers in this session are also in Poster Session P2-1.

Chairs: Imari Sato (National Inst. of Informatics)
Eli Shechtman (Adobe Research)

0830 Spotlights (S2-1B)

Format (4 min. for presentation; no questions)


14. [0834] FastMask: Segment Multi-Scale Object Candidates in One Shot, Hexiang Hu, Shiyi Lan, Yuning Jiang, Zhimin Cao, Fei Sha

15. [0838] Reconstructing Transient Images From Single-Photon Sensors, Matthew O’Toole, Felix Heide, David B. Lindell, Kai Zang, Steven Diamond, Gordon Wetzstein


18. [0850] Illuminant-Camera Communication to Observe Moving Objects Under Strong External Light by Spread Spectrum Modulation, Ryusuke Sagawa, Yutaka Satoh
19. [0854] Photorealistic Facial Texture Inference Using Deep Neural Networks, Shunsuke Saito, Lingyu Wei, Liwen Hu, Koki Nagano, Hao Li

0903 Orals (O2-1B)

Format (12 min. for presentation + 2 min. for questions)
21. [0903] Unrolling the Shutter: CNN to Correct Motion Distortions, Vijay Rengarajan, Yogesh Balaji, A. N. Rajagopalan
22. [0917] Light Field Blind Motion Deblurring, Pratul P. Srinivasan, Ren Ng, Ravi Ramamoorthi
23. [0931] Computational Imaging on the Electric Grid, Mark Sheinin, Yoav Y. Schechner, Kiriakos N. Kutulakos
24. [0945] Deep Outdoor Illumination Estimation, Yannick Hold-Geoffroy, Kalyan Sunkavalli, Sunil Hadap, Emiliano Gambaretto, Jean-François Lalonde

0830–1000 Session 2-1C: 3D Vision 2 (Kalâkaua Ballroom C)

Papers in this session are also in Poster Session P2-1.

Chairs: Yasutaka Furukawa (Washington Univ. St. Louis) Jingyi Yu (Univ. of Delaware)

0830 Spotlights (S2-1C)

Format (4 min. for presentation; no questions)
25. [0830] Efficient Solvers for Minimal Problems by Syzygy-Based Reduction, Viktor Larsson, Kalle Åström, Magnus Oskarsson
26. [0834] HSfM: Hybrid Structure-from-Motion, Hainan Cui, Xiang Gao, Shuhan Shen, Zhanyi Hu
27. [0838] Efficient Global Point Cloud Alignment Using Bayesian Nonparametric Mixtures, Julian Straub, Trevor Campbell, Jonathan P. How, John W. Fisher III
29. [0846] IM2CAD, Hamid Izadinia, Qi Shan, Steven M. Seitz
30. [0850] ScanNet: Richly-Annotated 3D Reconstructions of Indoor Scenes, Angela Dai, Angel X. Chang, Manolis Savva, Maciej Halber, Thomas Funkhouser, Matthias Nießner
31. [0854] Noise Robust Depth From Focus Using a Ring Difference Filter, Jaeheung Surh, Hae-Gon Jeon, Yunwon Park, Sunghoon Im, Hyowon Ha, In So Kweon
32. [0858] Group-Wise Point-Set Registration Based on Rényi’s Second Order Entropy, Luis G. Sanchez Giraldo, Eron Hasanbelliu, Murali Rao, Jose C. Principe

0903 Orals (O2-1C)

Format (12 min. for presentation + 2 min. for questions)
34. [0917] 3D Point Cloud Registration for Localization Using a Deep Neural Network Auto-Encoder, Gil Elbaz, Tamar Avraham, Anath Fischer
35. [0931] Flight Dynamics-Based Recovery of a UAV Trajectory Using Ground Cameras, Artem Rozantsev, Sudipta N. Sinha, Debadepta Dey, Pascal Fua
36. [0945] DSAC - Differentiable RANSAC for Camera Localization, Eric Brachmann, Alexander Krull, Sebastian Nowozin, Jamie Shotton, Frank Michel, Stefan Gumhold, Carsten Rother

1000–1045 Break (Kamehameha II)

1000–1200 Poster Session P2-1 (Kamehameha I)

3D Computer Vision

37. Scalable Surface Reconstruction From Point Clouds With Extreme Scale and Density Diversity, Christian Mostegel, Rudolf Pretenthaler, Friedrich Fraundorfer, Horst Bischof
39. General Models for Rational Cameras and the Case of Two-Slit Projections, Matthew Trager, Bernd Sturmels, John Canny, Martial Hebert, Jean Ponce
40. Accurate Depth and Normal Maps From Occlusion-Aware Focal Stack Symmetry, Michael Strecke, Anna Alperovich, Bastian Goldluecke
41. A Multi-View Stereo Benchmark With High-Resolution Images and Multi-Camera Videos, Thomas Schöps, Johannes L. Schönberger, Silvano Galliari, Torsten Sattler, Konrad Schindler, Marc Pollefeys, Andreas Geiger
42. Non-Contact Full Field Vibration Measurement Based on Phase-Shifting, Hiroyuki Kayaba, Yuji Kokumai
43. A Minimal Solution for Two-View Focal-Length Estimation Using Two Affine Correspondences, Daniel Barath, Tekla Toth, Levente Hajder
44. PoseAgent: Budget-Constrained 6D Object Pose Estimation via Reinforcement Learning, Alexander Krull, Eric Brachmann, Sebastian Nowozin, Frank Michel, Jamie Shotton, Carsten Rother
45. An Efficient Background Term for 3D Reconstruction and Tracking With Smooth Surface Models, Mariano Jaimez, Thomas J. Cashman, Andrew Fitzgibbon, Javier Gonzalez-Jimenez, Daniel Cremers

Analyzing Humans in Images
46. Reliable Crowdsourcing and Deep Locality-Preserving Learning for Expression Recognition in the Wild, Shan Li, Weihong Deng, JunPing Du
47. Procedural Generation of Videos to Train Deep Action Recognition Networks, César Roberto de Souza, Adrien Gaidon, Yohann Cabon, Antonio Manuel López
48. BigHand 2.2M Benchmark: Hand Pose Dataset and State of the Art Analysis, Shanxin Yuan, Qi Ye, Björn Stenger, Siddhant Jain, Tae-Kyun Kim
49. DenseReg: Fully Convolutional Dense Shape Regression In-The-Wild, Rıza Alp Güler, George Trigeorgis, Epameinondas Antonakos, Patrick Snape, Stefanos Zafeiriou, Iasonas Kokkinos
50. Adaptive Class Preserving Representation for Image Classification, Jian-Xun Mi, Qiankun Fu, Weisheng Li

Applications
52. EAST: An Efficient and Accurate Scene Text Detector, Xinyu Zhou, Cong Yao, He Wen, Yuzhi Wang, Shuchang Zhou, Weiran He, Jiajun Liang
53. VidLoc: A Deep Spatio-Temporal Model for 6-DoF Video-Clip Relocalization, Ronald Clark, Sen Wang, Andrew Markham, Niki Trigoni, Hongkai Wen

Biomedical Image/Video Analysis
54. Improving RANSAC-Based Segmentation Through CNN Encapsulation, Dustin Morley, Hassan Foroosh

Computational Photography
55. Position Tracking for Virtual Reality Using Commodity WiFi, Manikanta Kotaru, Sachin Katti
56. Designing Illuminant Spectral Power Distributions for Surface Classification, Henryk Blasinski, Joyce Farrell, Brian Wandell
57. One-Shot Hyperspectral Imaging Using Faced Reflectors, Tsuyoshi Takatani, Takahito Aoto, Yasuhiro Mukaigawa

Image Motion & Tracking
58. Direct Photometric Alignment by Mesh Deformation, Kaimo Lin, Nianjuan Jiang, Shuaicheng Liu, Loong-Fah Cheong, Minh Do, Jiangbo Lu
59. CNN-Based Patch Matching for Optical Flow With Thresholded Hinge Embedding Loss, Christian Bailer, Kiran Varanasi, Didier Stricker
60. Optical Flow Estimation Using a Spatial Pyramid Network, Anurag Ranjan, Michael J. Black
61. Deep Network Flow for Multi-Object Tracking, Samuel Schulter, Paul Vernaza, Wongun Choi, Manmohan Chandraker

Low- & Mid-Level Vision
62. Material Classification Using Frequency- and Depth-Dependent Time-Of-Flight Distortion, Kenichiro Tanaka, Yasuhiro Mukaigawa, Takuya Funatomi, Hiroyuki Kubo, Yasuyuki Matsushita, Yasushi Yagi
63. Benchmarking Denoising Algorithms With Real Photographs, Tobias Plötz, Stefan Roth
64. A Unified Approach of Multi-Scale Deep and Hand-Crafted Features for Defocus Estimation, Jinsun Park, Yu-Wing Tai, Donghyeon Cho, In So Kweon
65. StyleBank: An Explicit Representation for Neural Image Style Transfer, Dongdong Chen, Lu Yuan, Jing Liao, Nenghai Yu, Gang Hua
66. Specular Highlight Removal in Facial Images, Chen Li, Stephen Lin, Kun Zhou, Katsushi Ikeuchi
67. Image Super-Resolution via Deep Recursive Residual Network, Ying Tai, Jian Yang, Xiaoming Liu
68. Deep Image Harmonization, Yi-Hsuan Tsai, Xiaohui Shen, Zhe Lin, Kalyan Sunkavalli, Xin Lu, Ming-Hsuan Yang
70. A Novel Tensor-Based Video Rain Streaks Removal Approach via Utilizing Discriminatively Intrinsic Priors, Tai-Xiang Jiang, Ting-Zhu Huang, Xi-Le Zhao, Liang-Jian Deng, Yao Wang
71. GMS: Grid-based Motion Statistics for Fast, Ultra-Robust Feature Correspondence, JiaWang Bian, Wen-Yan Lin, Yasuyuki Matsushita, Sai-Kit Yeung, Tan-Dat Nguyen, Ming-Ming Cheng
72. Video Desnowing and Deraining Based on Matrix Decomposition, Weihong Ren, Jiandong Tian, Zhi Han, Antoni Chan, Yandong Tang
73. Real-Time Video Super-Resolution With Spatio-Temporal Networks and Motion Compensation, Jose Caballero, Christian Ledig, Andrew Aitken, Alejandro Acosta, Johannes Totz, Zehan Wang, Wenzhe Shi
74. Deep Watershed Transform for Instance Segmentation, Min Bai, Raquel Urtasun
75. AnchorNet: A Weakly Supervised Network to Learn Geometry-Sensitive Features for Semantic Matching, David Novotny, Diane Larlus, Andrea Vedaldi
76. Learning Diverse Image Colorization, Aditya Deshpande, Jiajun Lu, Mao-Chuang Yeh, Min Jin Chong, David Forsyth
77. Awesome Typography: Statistics-Based Text Effects Transfer, Shuai Yang, Jiaying Liu, Zhouhui Lian, Zongming Guo

**Machine Learning**

78. Unsupervised Video Summarization With Adversarial LSTM Networks, Behrooz Mahasseni, Michael Lam, Sinisa Todorovic
79. Deep TEN: Texture Encoding Network, Hang Zhang, Jia Xue, Kristin Dana
80. Order-Preserving Wasserstein Distance for Sequence Matching, Bing Su, Gang Hua
82. Hierarchical Multimodal Metric Learning for Multimodal Classification, Heng Zhang, Vishal M. Patel, Rama Chellappa
83. Efficient Linear Programming for Dense CRFs, Thalaiyasingam Ajanthan, Alban Desmaison, Rudy Bunel, Mathieu Salzmann, Philip H. S. Torr, M. Pawan Kumar
84. Variational Autoencoded Regression: High Dimensional Regression of Visual Data on Complex Manifold, YoungJoon Yoo, Sangdoo Yun, Hyung Jin Chang, Yiannis Demiris, Jin Young Choi
85. Learning Random-Walk Label Propagation for Weakly-Supervised Semantic Segmentation, Paul Vernaza, Manmohan Chandraker
86. Adversarial Discriminative Domain Adaptation, Eric Tzeng, Judy Hoffman, Kate Saenko, Trevor Darrell
87. Low-Rank-Sparse Subspace Representation for Robust Regression, Yongqiang Zhang, Daming Shi, Junbin Gao, Dansong Cheng

**Object Recognition & Scene Understanding**

88. Generating the Future With Adversarial Transformers, Carl Vondrick, Antonio Torralba
89. Semantic Amodal Segmentation, Yan Zhu, Yuandong Tian, Dimitris Metaxas, Piotr Dollár
90. Learning a Deep Embedding Model for Zero-Shot Learning, Li Zhang, Tao Xiang, Shaogang Gong
91. BIND: Binary Integrated Net Descriptors for Texture-Less Object Recognition, Jacob Chan, Jimmy Addison Lee, Qian Kemao
92. Growing a Brain: Fine-Tuning by Increasing Model Capacity, Yu-Xiong Wang, Deva Ramanan, Martial Hebert
94. Multiple Instance Detection Network With Online Instance Classifier Refinement, Peng Tang, Xinggang Wang, Xiang Bai, Wenyu Liu
95. Kernel Pooling for Convolutional Neural Networks, Yin Cui, Feng Zhou, Jiang Wang, Xiao Liu, Yuanqing Lin, Serge Belongie
97. Zero-Shot Learning - the Good, the Bad and the Ugly, Yongqin Xian, Bernt Schiele, Zeynep Akata
98. DeepNav: Learning to Navigate Large Cities, Samarth Brahmbhatt, James Hays
99. Scene Graph Generation by Iterative Message Passing, Danfei Xu, Yuke Zhu, Christopher B. Choy, Li Fei-Fei
100. Visual Translation Embedding Network for Visual Relation Detection, Hanwang Zhang, Zawlin Kyaw, Shih-Fu Chang, Tat-Seng Chua
101. Unsupervised Part Learning for Visual Recognition, Ronan Sicre, Yannis Avrithis, Ewa Kijak, Frédéric Jurie
102. Comprehension-Guided Referring Expressions, Ruotian Luo, Gregory Shakhnarovich
103. Top-Down Visual Saliency Guided by Captions, Vasili Ramanishka, Abir Das, Jianming Zhang, Kate Saenko

Theory
104. Grassmannian Manifold Optimization Assisted Sparse Spectral Clustering, Qiong Wang, Junbin Gao, Hong Li

Video Analytics
105. Video Propagation Networks, Varun Jampani, Raghudeep Gadde, Peter V. Gehler
108. Hierarchical Boundary-Aware Neural Encoder for Video Captioning, Lorenzo Baraldi, Costantino Grana, Rita Cucchiara
109. HOPE: Hierarchical Object Prototype Encoding for Efficient Object Instance Search in Videos, Tan Yu, Yuwei Wu, Junsong Yuan

1000–1200 Demos (Kamehameha I)
- FlowNet 2.0: Evolution of Optical Flow Estimation With Deep Networks, Eddy Ilg, Nikolaus Mayer, Tonmoy Saikia, Margret Keuper, Alexey Dosovitskiy, Thomas Brox (Univ. of Freiburg)
- A Low Power, Fully Event-Based Gesture Recognition System, Arnon Amir, Brian Tabata, 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, Dharmendra Modha (IBM Almaden Research Center)
- Fast Moving Objects – Detection, Recognition, Description, De-Blurring, Denys Rozumnyi, Aleš Hrabalík, Jan Koteru, Filip Šroubek, Jiří Matas (Czech Technical Univ. in Prague, Czech Academy of Sciences)
- OpenPose: A Real-Time Multi-Person Keypoint Detection Library, Gines Hidalgo, Zhe Cao, Tomas Simon, Shih-En Wei, Hanbyul Joo, Yaser Sheikh (Carnegie Mellon Univ.)

1000–1200 Exhibits (Kamehameha I)
- Same as Saturday morning Exhibits (see pg. 10)

1200–1300 Lunch (Kamehameha II)
1300–1430 Session 2-2A: Object Recognition & Scene Understanding 1  
(Kamehameha III)

Papers in this session are also in Poster Session P2-2.

Chairs: Vittorio Ferrari (Univ. of Edinburgh)  
Cha Zhang (Microsoft Research)

1300 Spotlights (S2-2A)  
Format (4 min. for presentation; no questions)
1. [1300] Graph-Structured Representations for Visual Question Answering, Damien Teney, Lingqiao Liu, Anton van den Hengel
3. [1308] Learned Contextual Feature Reweighting for Image Geo-Localization, Hyo Jin Kim, Enrique Dunn, Jan-Michael Frahm
4. [1312] End-To-End Concept Word Detection for Video Captioning, Retrieval, and Question Answering, Youngjae Yu, Hyungjin Ko, Jongwook Choi, Gunhee Kim
5. [1316] Deep Cross-Modal Hashing, Qing-Yuan Jiang, Wujun Li
6. [1320] Unambiguous Text Localization and Retrieval for Cluttered Scenes, Xuejian Rong, Chucai Yi, Yingli Tian
7. [1324] Bayesian Supervised Hashing, Zihao Hu, Junxuan Chen, Hongtao Lu, Tongzhen Zhang
8. [1328] Speed/Accuracy Trade-Offs for Modern Convolutional Object Detectors, Jonathan Huang, Vivek Rathod, Chen Sun, Menglong Zhu, Anoop Korattikara, Alireza Fathi, Ian Fischer, Zbigniew Wojna, Yang Song, Sergio Guadarrama, Kevin Murphy

1333 Orals (O2-2A)  
Format (12 min. for presentation + 2 min. for questions)
9. [1333] Detecting Visual Relationships With Deep Relational Networks, Bo Dai, Yuqi Zhang, Dahua Lin
12. [1415] AGA: Attribute-Guided Augmentation, Mandar Dixit, Roland Kwitt, Marc Niethammer, Nuno Vasconcelos

1300–1430 Session 2-2B: Analyzing Humans 2  
(Kalākaua Ballroom A-B)

Papers in this session are also in Poster Session P2-2.

Chairs: Jingdong Wang (Microsoft Research Asia)  
Daphna Weinshall (Hebrew Univ. of Jerusalem)

1300 Spotlights (S2-2B)  
Format (4 min. for presentation; no questions)
14. [1304] Person Re-Identification in the Wild, Liang Zheng, Hengheng Zhang, Shaoyan Sun, Manmohan Chandraker, Yi Yang, Qi Tian
15. [1308] Scalable Person Re-Identification on Supervised Smoothed Manifold, Song Bai, Xiang Bai, Qi Tian
17. [1316] Joint Detection and Identification Feature Learning for Person Search, Tong Xiao, Shuang Li, Bochao Wang, Liang Lin, Xiaogang Wang
18. [1320] Synthesizing Normalized Faces From Facial Identity Features, Forrester Cole, David Belanger, Dilip Krishnan, Aaron Sarna, Inbar Mosseri, William T. Freeman
19. [1324] Consistent-Aware Deep Learning for Person Re-Identification in a Camera Network, Ji Lin, Liangliang Ren, Jiwen Lu, Jianjiang Feng, Jie Zhou
20. [1328] Level Playing Field for Million Scale Face Recognition, Aaron Nech, Ira Kemelmacher-Shlizerman

1333 Orals (O2-2B)  
Format (12 min. for presentation + 2 min. for questions)
22. [1347] Social Scene Understanding: End-To-End Multi-Person Action Localization and Collective Activity Recognition, Timur Bagautdinov, Alexandre Alahi, François Fleuret, Pascal Fua, Silvio Savarese
Detangling People: Individuating Multiple Close People and Their Body Parts via Region Assembly, Hao Jiang, Kristen Grauman

Lip Reading Sentences in the Wild, Joon Son Chung, Andrew Senior, Oriol Vinyals, Andrew Zisserman

**1300–1430 Session 2-2C: Applications** (Kalākaua Ballroom C)

Papers in this session are also in Poster Session P2-2.

**Chairs:** Zicheng Liu (Microsoft Research)
S. Kevin Zhou (Siemens Research)

**1300 Spotlights (S2-2C)**
Format (4 min. for presentation; no questions)


28. [1312] Detecting Oriented Text in Natural Images by Linking Segments, Baoguang Shi, Xiang Bai, Serge Belongie

29. [1316] Learning Video Object Segmentation From Static Images, Federico Perazzi, Anna Khoreva, Rodrigo Benenson, Bernt Schiele, Alexander Sorkine-Hornung


**1333 Orals (O2-2C)**
Format (12 min. for presentation + 2 min. for questions)

33. [1333] End-To-End Learning of Driving Models From Large-Scale Video Datasets, Huazhe Xu, Yang Gao, Fisher Yu, Trevor Darrell

34. [1347] Deep Future Gaze: Gaze Anticipation on Egocentric Videos Using Adversarial Networks, Mengmi Zhang, Keng Teck Ma, Joo Hwee Lim, Qi Zhao, Jiashi Feng

35. [1401] MDNet: A Semantically and Visually Interpretable Medical Image Diagnosis Network, Zizhao Zhang, Yuanpu Xie, Fuyong Xing, Mason McGough, Lin Yang

**1430–1515 Break** (Kamehameha II)

**1430–1630 Poster Session P2-2** (Kamehameha I)

**3D Computer Vision**

36. Surface Motion Capture Transfer With Gaussian Process Regression, Adnane Boukhayma, Jean-Sébastien Franco, Edmond Boyer

37. Visual-Inertial-Semantic Scene Representation for 3D Object Detection, Jingming Dong, Xiaohan Fei, Stefano Soatto

38. Template-Based Monocular 3D Recovery of Elastic Shapes Using Lagrangian Multipliers, Nazim Haouchine, Stephane Cotin

39. Learning Category-Specific 3D Shape Models From Weakly Labeled 2D Images, Dingwen Zhang, Junwei Han, Yang Yang, Dong Huang

40. Simultaneous Geometric and Radiometric Calibration of a Projector-Camera Pair, Marjan Shahpaski, Luis Ricardo Sapaico, Gaspard Chevassus, Sabine Süsstrunk

41. A Clever Elimination Strategy for Efficient Minimal Solvers, Zuzana Kukelova, Joe Kileel, Bernd Sturmfsels, Tomas Pajdla

42. Learning Barycentric Representations of 3D Shapes for Sketch-Based 3D Shape Retrieval, Jin Xie, Guoxian Dai, Fan Zhu, Yi Fang

43. Geodesic Distance Descriptors, Gil Shamai, Ron Kimmel
Analyzing Humans in Images
44. Modeling Temporal Dynamics and Spatial Configurations of Actions Using Two-Stream Recurrent Neural Networks, Hongsong Wang, Liang Wang
45. Forecasting Human Dynamics From Static Images, Yu-Wei Chao, Jimei Yang, Brian Price, Scott Cohen, Jia Deng
46. Re-Ranking Person Re-Identification With $k$-Reciprocal Encoding, Zhun Zhong, Liang Zheng, Donglin Cao, Shaozi Li
47. Deep Sequential Context Networks for Action Prediction, Yu Kong, Zhiqiang Tao, Yun Fu
48. Global Context-Aware Attention LSTM Networks for 3D Action Recognition, Jun Liu, Gang Wang, Ping Hu, Ling-Yu Duan, Alex C. Kot
49. Dynamic Attention-Controlled Cascaded Shape Regression Exploiting Training Data Augmentation and Fuzzy-Set Sample Weighting, Zhen-Hua Feng, Josef Kittler, William Christmas, Patrik Huber, Xiao-Jun Wu
50. A Deep Regression Architecture With Two-Stage Re-Initialization for High Performance Facial Landmark Detection, Jiangjing Lv, Xiaohu Shao, Junliang Xing, Cheng Cheng, Xi Zhou
51. Multiple People Tracking by Lifted Multicut and Person Re-Identification, Siyu Tang, Mykhaylo Andriluka, Bjoern Andres, Bernt Schiele
52. Towards Accurate Multi-Person Pose Estimation in the Wild, George Papandreou, Tyler Zhu, Nori Kanazawa, Alexander Toshev, Jonathan Tompson, Chris Bregler, Kevin Murphy

Applications
53. Towards a Quality Metric for Dense Light Fields, Vamsi Kiran Adhikarla, Marek Vinkler, Denis Sumin, Rafał K. Mantiuk, Karol Myszkowski, Hans-Peter Seidel, Piotr Didyk
54. Controlling Perceptual Factors in Neural Style Transfer, Leon A. Gatys, Alexander S. Ecker, Matthias Bethge, Aaron Hertzmann, Eli Shechtman

Biomedical Image/Video Analysis
55. Joint Sequence Learning and Cross-Modality Convolution for 3D Biomedical Segmentation, Kuan-Lun Tseng, Yen-Liang Lin, Winston Hsu, Chung-Yang Huang
56. LSTM Self-Supervision for Detailed Behavior Analysis, Biagio Brattoli, Uta Büchner, Anna-Sophia Wahl, Martin E. Schwab, Björn Ommer

Computational Photography
57. A Wide-Field-Of-View Monocentric Light Field Camera, Donald G. Dansereau, Glenn Schuster, Joseph Ford, Gordon Wetzstein

Image Motion & Tracking
58. S2F: Slow-To-Fast Interpolator Flow, Yanchao Yang, Stefano Soatto
59. CLKN: Cascaded Lucas-Kanade Networks for Image Alignment, Che-Han Chang, Chun-Nan Chou, Edward Y. Chang
60. Multi-Object Tracking With Quadruplet Convolutional Neural Networks, Jeany Son, Mooyeol Baek, Minsu Cho, Bohyung Han

Low- & Mid-Level Vision
61. Learning to Detect Salient Objects With Image-Level Supervision, Lijun Wang, Huchuan Lu, Yifan Wang, Mengyang Feng, Dong Wang, Baocai Yin, Xiang Ruan
62. From Motion Blur to Motion Flow: A Deep Learning Solution for Removing Heterogeneous Motion Blur, Dong Gong, Jie Yang, Lingqiao Liu, Yanning Zhang, Ian Reid, Chunhua Shen, Anton van den Hengel, Qinfeng Shi
63. Co-Occurrence Filter, Roy J. Jevnisek, Shai Avidan
64. Fractal Dimension Invariant Filtering and Its CNN-Based Implementation, Hongteng Xu, Junchi Yan, Nils Persson, Weiyao Lin, Hongyuan Zha
65. Noise-Blind Image Deblurring, Meiguang Jin, Stefan Roth, Paolo Favaro
66. Simultaneous Visual Data Completion and Denoising Based on Tensor Rank and Total Variation Minimization and Its Primal-Dual Splitting Algorithm, Tatsuya Yokota, Hidekata Hontani
67. HPatches: A Benchmark and Evaluation of Handcrafted and Learned Local Descriptors, Vassileios Balntas, Karel Lenc, Andrea Vedaldi, Krystian Mikolajczyk
68. Hyperspectral Image Super-Resolution via Non-Local Sparse Tensor Factorization, Renwei Dian, Leyuan Fang, Shutao Li
69. Reflection Removal Using Low-Rank Matrix Completion, Byeong-Ju Han, Jae-Young Sim
70. Object Co-Skeletonization With Co-Segmentation, Koteswar Rao Jerrypothula, Jianfei Cai, Jiangbo Lu, Junsong Yuan
Sunday, July 23 (Afternoon)

**Program**

**Machine Learning**

71. Mining Object Parts From CNNs via Active Question-Answering, Quanshi Zhang, Ruiming Cao, Ying Nian Wu, Song-Chun Zhu
72. PolyNet: A Pursuit of Structural Diversity in Very Deep Networks, Xingcheng Zhang, Zhizhong Li, Chen Change Loy, Dahua Lin
73. The VQA-Machine: Learning How to Use Existing Vision Algorithms to Answer New Questions, Peng Wang, Qi Wu, Chunhua Shen, Anton van den Hengel
74. Joint Discriminative Bayesian Dictionary and Classifier Learning, Naveed Akhtar, Ajmal Mian, Fatih Porikli
75. Quad-Networks: Unsupervised Learning to Rank for Interest Point Detection, Nikolay Savinov, Akihito Seki, Lubor Ladický, Torsten Sattler, Marc Pollefeys
76. Outlier-Robust Tensor PCA, Pan Zhou, Jiashi Feng
77. Learning Adaptive Receptive Fields for Deep Image Parsing Network, Zhen Wei, Yao Sun, Jinqiao Wang, Hanjiang Lai, Si Liu
78. Learning an Invariant Hilbert Space for Domain Adaptation, Samitha Herath, Mehrtash Harandi, Fatih Porikli
79. Fixed-Point Factorized Networks, Peisong Wang, Jian Cheng
80. Discriminative Optimization: Theory and Applications to Point Cloud Registration, Jayakorn Vongkulbhisal, Fernando De la Torre, João P. Costeira
81. Online Asymmetric Similarity Learning for Cross-Modal Retrieval, Yiling Wu, Shuhui Wang, Qingming Huang
82. Improving Training of Deep Neural Networks via Singular Value Bounding, Kui Jia, Dacheng Tao, Shenghua Gao, Xiangmin Xu
83. S3Pool: Pooling With Stochastic Spatial Sampling, Shuangfei Zhai, Hui Wu, Abhishek Kumar, Yu Cheng, Yongxi Lu, Zhongfei Zhang, Rogerio Feris
84. Sports Field Localization via Deep Structured Models, Namdar Homayounfar, Sanja Fidler, Raquel Urtasun
85. Noisy Softmax: Improving the Generalization Ability of DCNN via Postponing the Early Softmax Saturation, Binghui Chen, Weihong Deng, Junping Du
86. Switching Convolutional Neural Network for Crowd Counting, Deepak Babu Sam, Shiv Surya, R. Venkatesh Babu
87. Network Sketching: Exploiting Binary Structure in Deep CNNs, Yiwen Guo, Anbang Yao, Hao Zhao, Yurong Chen
88. Multi-Task Clustering of Human Actions by Sharing Information, Xiaqiang Yan, Shizhe Hu, Yangdong Ye
89. Soft-Margin Mixture of Regressions, Dong Huang, Longfei Han, Fernando De la Torre
90. Multigrid Neural Architectures, Tsung-Wei Ke, Michael Maire, Stella X. Yu
91. High-Resolution Image Inpainting Using Multi-Scale Neural Patch Synthesis, Chao Yang, Xin Lu, Zhe Lin, Eli Shechtman, Oliver Wang, Hao Li
92. Deep Quantization: Encoding Convolutional Activations With Deep Generative Model, Zhaofan Qiu, Ting Yao, Tao Mei
93. DOPE: Distributed Optimization for Pairwise Energies, Jose Dolz, Ismail Ben Ayed, Christian Desrosiers
94. Improved Texture Networks: Maximizing Quality and Diversity in Feed-Forward Stylization and Texture Synthesis, Dmitry Ulyanov, Andrea Vedaldi, Victor Lempitsky

**Object Recognition & Scene Understanding**

95. Polyhedral Conic Classifiers for Visual Object Detection and Classification, Hakan Cevikalp, Bill Triggs
96. Incremental Kernel Null Space Discriminant Analysis for Novelty Detection, Juncheng Liu, Zhouhui Lian, Yi Wang, Jianguo Xiao
97. Predicting Ground-Level Scene Layout From Aerial Imagery, Menghua Zhai, Zachary Bessinger, Scott Workman, Nathan Jacobs
98. Deep Feature Flow for Video Recognition, Xizhou Zhu, Yuwen Xiong, Jifeng Dai, Lu Yuan, Yichen Wei
99. Object-Aware Dense Semantic Correspondence, Fan Yang, Xin Li, Hong Cheng, Jianping Li, Leiting Chen
100. Semantic Regularisation for Recurrent Image Annotation, Feng Liu, Tao Xiang, Timothy M. Hospedales, Wankou Yang, Changyin Sun
101. Video2Shop: Exact Matching Clothes in Videos to Online Shopping Images, Zhi-Qi Cheng, Xiao Wu, Yang Liu, Xiansheng Hua
103. Multi-Level Attention Networks for Visual Question Answering, Dongfei Yu, Jianlong Fu, Tao Mei, Yong Rui
104. Generating Descriptions With Grounded and Co-Referenced People, Anna Rohrbach, Marcus Rohrbach, Siyu Tang, Seong Joon Oh, Bernt Schiele
106. Simultaneous Feature Aggregating and Hashing for Large-Scale Image Search, Thanh-Toan Do, Dang-Khoa Le Tran, Trung T. Pham, Ngai-Man Cheung
107. Improving Facial Attribute Prediction Using Semantic Segmentation, Mahdi M. Kalayeh, Boqing Gong, Mubarak Shah

Video Analytics
110. CERN: Confidence-Energy Recurrent Network for Group Activity Recognition, Tianmin Shu, Sinisa Todorovic, Song-Chun Zhu
111. Understanding Traffic Density From Large-Scale Web Camera Data, Shanghang Zhang, Guanhang Wu, João P. Costeira, José M. F. Moura
112. Collaborative Summarization of Topic-Related Videos, Rameswar Panda, Amit K. Roy-Chowdhury

1430–1630 Demos (Kamehameha I)
• Same as Sunday morning Demos (see pg. 21)

1430–1630 Exhibits (Kamehameha I)
• Same as Saturday morning Exhibits (see pg. 10)

1645–1800 Plenary Session (Kamehameha III)
• Keynote Talk: Commercializing Computer Vision: Success Stories and Lessons Learned, Harry Shum (Microsoft Research)
• Abstract: It is an exciting time for all of us computer vision researchers and practitioners. We have seen an unprecedented growth in the conversion of years of progress into marketable technologies. Microsoft has long been committed to developing new computer vision technologies, making them available to developers, and incorporating them into many products. In this talk, I will first briefly review 25 years of computer vision research at Microsoft Research (MSR), highlighting MSR’s contributions to the vision community and emphasizing the importance of long-term commitment to funding successful industrial research labs. I will also describe some of our latest research work in computational photography, image understanding, and vision and language before detailing our commercialization successes. In particular, I will share our experiences in developing three products: Microsoft Pix, HoloLens, and Cognitive Services, which leverage computer vision systems and technologies in different ways.

Pix is an AI-powered camera app that makes taking great pictures easy and fun: “point, shoot, perfect!” It has incorporated technologies from more than a dozen CVPR, ICCV, and SIGGRAPH papers from MSR. HoloLens is the first commercially available mixed reality system in the market. Cognitive Services allow you to build useful AI-based apps using just a few lines of code, across different devices and platforms. I will show IRIS, which is an interactive visual learning service for developers to create image recognition applications. I will also show the latest cool demos using HoloLens, including real-time environment understanding. There are challenges in accelerating the cycle from research to product, and I will discuss the lessons learned in productizing Pix, HoloLens, and Cognitive Services.

1900–2200 Luau Reception (The Museum Quad, Hale Koa Hotel beachfront)

Directions: On the beachfront between the Hale Koa Hotel (2055 Kalia Rd.) and the Fort DeRussy US Army Museum (2161 Kalia Rd.).
Monday, July 24

0730–1200 Registration (Main Lobby)

0730–0830 Breakfast (Kamehameha II)

0830–1000 Session 3-1A: Machine Learning 3 (Kamehameha III)
Papers in this session are also in Poster Session P3-1.

Chairs: Simon Lucey (Carnegie Mellon Univ.)
         Gabriel Brostow (Univ. College London)

0830 Spotlights (S3-1A)
Format (4 min. for presentation; no questions)
1. [0830] Local Binary Convolutional Neural Networks, Felix Juefei-Xu, Vishnu Naresh Boddeti, Marios Savvides
2. [0834] Deep Self-Taught Learning for Weakly Supervised Object Localization, Zequn Jie, Yunchao Wei, Xiaojie Jin, Jiashi Feng, Wei Liu
3. [0838] Multi-Modal Mean-Fields via Cardinality-Based Clamping, Pierre Baqué, François Fleuret, Pascal Fua
4. [0842] Probabilistic Temporal Subspace Clustering, Behnam Gholami, Vladimir Pavlovic
5. [0846] Provable Self-Representation Based Outlier Detection in a Union of Subspaces, Chong You, Daniel P. Robinson, René Vidal
6. [0850] Latent Multi-View Subspace Clustering, Changqing Zhang, Qinghua Hu, Huazhu Fu, Pengfei Zhu, Xiaochun Cao
7. [0854] Learning to Extract Semantic Structure From Documents Using Multimodal Fully Convolutional Neural Networks, Xiao Yang, Ersin Yumer, Paul Asente, Mike Kraley, Daniel Kifer, C. Lee Giles
8. [0858] Age Progression/Regression by Conditional Adversarial Autoencoder, Zhifei Zhang, Yang Song, Hairong Qi

0903 Orals (O3-1A)
Format (12 min. for presentation + 2 min. for questions)
9. [0903] Compact Matrix Factorization With Dependent Subspaces, Viktor Larsson, Carl Olsson

10. [0917] FFTLasso: Large-Scale LASSO in the Fourier Domain, Adel Bibi, Hani Itani, Bernard Ghanem
12. [0945] Global Optimality in Neural Network Training, Benjamin D. Haeffele, René Vidal

0830–1000 Session 3-1B: Object Recognition & Scene Understanding 2 (Kalākaua Ballroom)
Papers in this session are also in Poster Session P3-1.

Chairs: Gang Hua (Microsoft Research Asia)
         Hamed Pirsiavash (Univ. of Maryland Baltimore County)

0830 Spotlights (S3-1B)
Format (4 min. for presentation; no questions)
15. [0838] Modeling Relationships in Referential Expressions With Compositional Modular Networks, Ronghang Hu, Marcus Rohrbach, Jacob Andreas, Trevor Darrell, Kate Saenko
16. [0842] Counting Everyday Objects in Everyday Scenes, Prithvijit Chattopadhyay, Ramakrishna Vedantam, Ramprasaath R. Selvaraju, Dhruv Batra, Devi Parikh
17. [0846] Fully Convolutional Instance-Aware Semantic Segmentation, Yi Li, Haozhi Qi, Jifeng Dai, Xiangyang Ji, Yichen Wei
18. [0850] Semantic Autoencoder for Zero-Shot Learning, Elyor Kodirov, Tao Xiang, Shaogang Gong
19. [0854] CityPersons: A Diverse Dataset for Pedestrian Detection, Shanshan Zhang, Rodrigo Benenson, Bernt Schiele
0903 Orals (O3-1B)

Format (12 min. for presentation + 2 min. for questions)


22. [0917] Annotating Object Instances With a Polygon-RNN, Lluís Castrejón, Kaustav Kundu, Raquel Urtasun, Sanja Fidler


1000–1045 Break (Kamehameha II)

1000–1200 Poster Session P3-1 (Kamehameha I)

3D Computer Vision

25. Self-Calibration-Based Approach to Critical Motion Sequences of Rolling-Shutter Structure From Motion, Eisuke Ito, Takayuki Okatani

26. Semi-Calibrated Near Field Photometric Stereo, Fotios Logothetis, Roberto Mecca, Roberto Cipolla


28. Learning to Predict Stereo Reliability Enforcing Local Consistency of Confidence Maps, Matteo Poggi, Stefano Mattoccia

29. The Misty Three Point Algorithm for Relative Pose, Tobias Palmér, Kalle Åström, Jan-Michael Frahm

30. The Surfacing of Multiview 3D Drawings via Lofting and Occlusion Reasoning, Anil Usumezbas, Ricardo Fabbri, Benjamin B. Kimia

31. A New Representation of Skeleton Sequences for 3D Action Recognition, Qiuhong Ke, Mohammed Bennamoun, Senjian An, Ferdous Sohel, Farid Boussaid

32. A General Framework for Curve and Surface Comparison and Registration With Oriented Varifolds, Irène Kaltenmark, Benjamin Charlier, Nicolas Charon

33. Learning to Align Semantic Segmentation and 2.5D Maps for Geolocation, Anil Armagan, Martin Hirzer, Peter M. Roth, Vincent Lepetit

34. A Generative Model for Depth-Based Robust 3D Facial Pose Tracking, Lu Sheng, Jianfei Cai, Tat-Jen Cham, Vladimir Pavlovic, King Ngï Ngan

35. Fast 3D Reconstruction of Faces With Glasses, Fabio Maninchedda, Martin R. Oswald, Marc Pollefeys

36. An Efficient Algebraic Solution to the Perspective-Three-Point Problem, Tong Ke, Stergios I. Roumeliotis

Analyzing Humans in Images

37. Learning From Synthetic Humans, Gül Varol, Javier Romero, Xavier Martin, Naureen Mahmood, Michael J. Black, Ivan Laptev, Cordelia Schmid

38. Forecasting Interactive Dynamics of Pedestrians With Fictitious Play, Wei-Chiu Ma, De-An Huang, Namhoon Lee, Kris M. Kitani


40. PoseTrack: Joint Multi-Person Pose Estimation and Tracking, Umar Iqbal, Anton Milan, Juergen Gall

41. Expecting the Unexpected: Training Detectors for Unusual Pedestrians With Adversarial Imposters, Shiyu Huang, Deva Ramanan

42. On Human Motion Prediction Using Recurrent Neural Networks, Julieta Martinez, Michael J. Black, Javier Romero

43. Learning and Refining of Privileged Information-Based RNNs for Action Recognition From Depth Sequences, Zhiyuan Shi, Tae-Kyun Kim

44. Quality Aware Network for Set to Set Recognition, Yu Liu, Junjie Yan, Wanli Ouyang

45. Unite the People: Closing the Loop Between 3D and 2D Human Representations, Christoph Lassner, Javier Romero, Martin Kiefel, Federica Bogo, Michael J. Black, Peter V. Gehler


47. Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset, João Carreira, Andrew Zisserman
Monday, July 24 (Morning)

Applications
48. Identifying First-Person Camera Wearers in Third-Person Videos, Chenyou Fan, Jangwon Lee, Mingze Xu, Krishna Kumar Singh, Yong Jae Lee, David J. Crandall, Michael S. Ryoo
49. Learning to Rank Retargeted Images, Yang Chen, Yong-Jin Liu, Yu-Kun Lai

Biomedical Image/Video Analysis
51. Fine-Tuning Convolutional Neural Networks for Biomedical Image Analysis: Actively and Incrementally, Zongwei Zhou, Jae Shin, Lei Zhang, Suryakanth Gurudu, Michael Gotway, Jianming Liang

Computational Photography
52. Depth From Defocus in the Wild, Huixuan Tang, Scott Cohen, Brian Price, Stephen Schiller, Kiriakos N. Kutulakos
53. Matting and Depth Recovery of Thin Structures Using a Focal Stack, Chao Liu, Srinivasa G. Narasimhan, Artur W. Dubrawski

Image Motion & Tracking
54. Robust Interpolation of Correspondences for Large Displacement Optical Flow, Yinlin Hu, Yunsong Li, Rui Song
55. Large Margin Object Tracking With Circulant Feature Maps, Mengmeng Wang, Yong Liu, Zeyi Huang
56. Minimum Delay Moving Object Detection, Dong Lao, Ganesh Sundaramoorthi
57. Multi-Task Correlation Particle Filter for Robust Object Tracking, Tianzhu Zhang, Changsheng Xu, Ming-Hsuan Yang
59. The World of Fast Moving Objects, Denys Rozumnyi, Jan Kotera, Filip Šroubek, Lukáš Novotný, Jiří Matas
60. Discriminative Correlation Filter With Channel and Spatial Reliability, Alan Lukežič, Tomáš Vojíř, Luka Čehovin Zajc, Jiří Matas, Matej Kristan

Low- & Mid-Level Vision
61. Learning Deep Binary Descriptor With Multi-Quantization, Yueqi Duan, Jiwen Lu, Ziwei Wang, Jianjiang Feng, Jie Zhou
62. One-To-Many Network for Visually Pleasing Compression Artifacts Reduction, Jun Guo, Hongyang Chao
64. BRISKS: Binary Features for Spherical Images on a Geodesic Grid, Hao Guan, William A. P. Smith
65. Superpixels and Polygons Using Simple Non-Iterative Clustering, Radhakrishna Achanta, Sabine Süsstrunk
66. Hardware-Efficient Guided Image Filtering for Multi-Label Problem, Longquan Dai, Mengke Yuan, Zecho Li, Xiaopeng Zhang, Jinhui Tang
67. Alternating Direction Graph Matching, D. Khuê Lê-Huu, Nikos Paragios
68. Learning Discriminative and Transformation Covariant Local Feature Detectors, Xu Zhang, Felix X. Yu, Svebor Karaman, Shih-Fu Chang

Machine Learning
69. Correlational Gaussian Processes for Cross-Domain Visual Recognition, Chengjiang Long, Gang Hua
70. DeLiGAN : Generative Adversarial Networks for Diverse and Limited Data, Swaminathan Gurumurthy, Ravi Kiran Sarvadevabhatla, R. Venkatesh Babu
71. A Dual Ascent Framework for Lagrangean Decomposition of Combinatorial Problems, Paul Swoboda, Jan Kuske, Bogdan Savchynsky
72. Oriented Response Networks, Yanzhao Zhou, Qixiang Ye, Qiang Qiu, Jianbin Jiao
73. Missing Modalities Imputation via Cascaded Residual Autoencoder, Luan Tran, Xiaoming Liu, Jiayu Zhou, Rong Jin
74. Efficient Optimization for Hierarchically-structured Interacting Segments (HINTS), Hossam Isack, Olga Veksler, Ipek Oguz, Milan Sonka, Yuri Boykov
75. A Message Passing Algorithm for the Minimum Cost Multicut Problem, Paul Swoboda, Bjoern Andres
76. End-To-End Representation Learning for Correlation Filter Based Tracking, Jack Valmadre, Luca Bertinetto, João Henriques, Andrea Vedaldi, Philip H. S. Torr

77. Filter Flow Made Practical: Massively Parallel and Lock-Free, Sathya N. Ravi, Yunyang Xiong, Lopamudra Mukherjee, Vikas Singh

78. Online Graph Completion: Multivariate Signal Recovery in Computer Vision, Won Hwa Kim, Mona Jalal, Seongjae Hwang, Sterling C. Johnson, Vikas Singh

79. Point to Set Similarity Based Deep Feature Learning for Person Re-Identification, Sanping Zhou, Jinjun Wang, Jiayun Wang, Yihong Gong, Nanning Zheng

80. Exploiting Saliency for Object Segmentation From Image Level Labels, Seong Joon Oh, Rodrigo Benenson, Anna Khoreva, Zeynep Akata, Mario Fritz, Bernt Schiele

81. Consensus Maximization With Linear Matrix Inequality Constraints, Pablo Speciale, Danda Pani Paudel, Martin R. Oswald, Till Kroeger, Luc Van Gool, Marc Pollefeys

82. Physically-Based Rendering for Indoor Scene Understanding Using Convolutional Neural Networks, Yinda Zhang, Shuran Song, Ersin Yumer, Manolis Savva, Joon-Young Lee, Hailin Jin, Thomas Funkhouser

83. Deep Multimodal Representation Learning From Temporal Data, Xitong Yang, Palghat Ramesh, Radha Chitta, Sriganesh Madhvanath, Edgar A. Bernal, Jiebo Luo

84. All You Need Is Beyond a Good Init: Exploring Better Solution for Training Extremely Deep Convolutional Neural Networks With Orthonormality and Modulation, Di Xie, Jiang Xiong, Shiliang Pu

85. Hard Mixtures of Experts for Large Scale Weakly Supervised Vision, Sam Gross, Marc’Aurelio Ranzato, Arthur Szlam

86. A Reinforcement Learning Approach to the View Planning Problem, Mustafa Devrim Kaba, Mustafa Gokhan Uzunbas, Ser Nam Lim

87. Zero-Shot Classification With Discriminative Semantic Representation Learning, Meng Ye, Yuhong Guo

**Object Recognition & Scene Understanding**

88. Automatic Discovery, Association Estimation and Learning of Semantic Attributes for a Thousand Categories, Ziad Al-Halah, Rainer Stiefelhagen

89. Scene Parsing Through ADE20K Dataset, Bolei Zhou, Hang Zhao, Xavier Puig, Sanja Fidler, Adela Barriuso, Antonio Torralba


91. Discretely Coding Semantic Rank Orders for Supervised Image Hashing, Li Liu, Ling Shao, Fumin Shen, Mengyang Yu

92. Joint Geometrical and Statistical Alignment for Visual Domain Adaptation, Jing Zhang, Wanqing Li, Philip Ogurbona

93. Weakly Supervised Dense Video Captioning, Zhiqiang Shen, Jianguo Li, Zhou Su, Minjun Li, Yurong Chen, Yu-Gang Jiang, Xiangyang Xue

94. RefineNet: Multi-Path Refinement Networks for High-Resolution Semantic Segmentation, Guosheng Lin, Anton Milan, Chunjua Shen, Ian Reid

95. Semantic Segmentation via Structured Patch Prediction, Context CRF and Guidance CRF, Falong Shen, Rui Gan, Shuicheng Yan, Gang Zeng

96. Person Search With Natural Language Description, Shuang Li, Tong Xiao, Hongsheng Li, Bolei Zhou, Dayu Yue, Xiaogang Wang

97. Weakly Supervised Affordance Detection, Johann Sawatzky, Abhilash Srikanta, Juergen Gall


99. Neural Aggregation Network for Video Face Recognition, Jiaolong Yang, Peiran Ren, Dongqing Zhang, Dong Chen, Fang Wen, Hongdong Li, Gang Hua

100. Relationship Proposal Networks, Ji Zhang, Mohamed Elhoseiny, Scott Cohen, Walter Chang, Ahmed Elgammal

101. Learning Object Interactions and Descriptions for Semantic Image Segmentation, Guangrun Wang, Ping Luo, Liang Lin, Xiaogang Wang

102. RON: Reverse Connection With Objectness Prior Networks for Object Detection, Tao Kong, Fuchun Sun, Anbang Yao, Huaping Liu, Ming Lu, Yurong Chen
Monday, July 24 (Morning)

103. Weakly-Supervised Visual Grounding of Phrases With Linguistic Structures, Fanyi Xiao, Leonid Sigal, Yong Jae Lee

104. Incorporating Copying Mechanism in Image Captioning for Learning Novel Objects, Ting Yao, Yingwei Pan, Yehao Li, Tao Mei


106. MuCaLe-Net: Multi Categorical-Level Networks to Generate More Discriminating Features, Youssef Tamaazousti, Hervé Le Borgne, Céline Hudelot

107. Zero Shot Learning via Multi-Scale Manifold Regularization, Shay Deutsch, Soheil Kolouri, Kyungnam Kim, Yuri Owechko, Stefano Soatto

**Theory**

108. Deeply Supervised Salient Object Detection With Short Connections, Qibin Hou, Ming-Ming Cheng, Xiaowei Hu, Ali Borji, Zhuowen Tu, Philip H. S. Torr


**Video Analytics**

110. One-Shot Video Object Segmentation, Sergi Caelles, Kevis-Kokitsi Maninis, Jordi Pont-Tuset, Laura Leal-Taixé, Daniel Cremers, Luc Van Gool

111. Fast Person Re-Identification via Cross-Camera Semantic Binary Transformation, Jiaxin Chen, Yinhong Wang, Jie Qin, Li Liu, Ling Shao

112. SPFTN: A Self-Paced Fine-Tuning Network for Segmenting Objects in Weakly Labelled Videos, Dingwen Zhang, Le Yang, Deyu Meng, Dong Xu, Junwei Han

**1000–1200 Demos (Kamehameha I)**


- ShapeSearch: A Generic Search Engine for 3D Models, Images and Sketches, Flora Ponjou Tasse (Univ. of Cambridge/Selerio)

- Online Social Media Image Processing Using AIDR 2.0: Artificial Intelligence for Digital Response, Firoj Alam, Muhammad Imran, Ferda Ofli (Qatar Computing Research Inst., Hamad Bin Khalifa Univ.)

**1000–1200 Exhibits (Kamehameha I)**

- Same as Saturday morning Exhibits (see pg. 10)

1200 Lunch (on your own)

Notes:
Monday, July 24 (Morning)

1200–1400 Doctoral Consortium (Kamehameha II) (by invitation only)

Supported by:

- Shervin Ardeshir Univ. of Central Florida
- Xiao Chu CUHK
- Zhengming Ding Northeastern Univ.
- Christoph Feichtenhofer Graz Univ. of Technology
- Je Hyeong Hong Univ. of Cambridge
- Ahmet Iscen INRIA/Renne
- Dinesh Jayaraman Univ. of Texas at Austin
- Dinghuang Ji UNC Chapel Hill
- Mahdi Kalayeh Univ. of Central Florida
- Kuldeep Kulkarni Arizona State Univ.
- Ang Li Univ. of Maryland at College Park
- Dong Li Tsinghua Univ.
- Yang Long Univ. of Sheffield
- Ondrej Miksik Univ. of Oxford
- Saeid Motiian West Virginia Univ.
- Tae-Hyun Oh KAIST
- Eshed Ohn-Bar Univ. of California San Diego
- Jinshan Pan Dalian Univ. of Technology
- Wenjie Pei Delft Univ. of Technology
- Wenqi Ren Tianjin Univ.
- Christos Sagonas Imperial College London
- Sergey Tulyakov Univ. of Trento
- Subhashini Venugopalan Univ. of Texas at Austin
- Scott Workman Univ. of Kentucky
- Zhongwen Xu Univ. of Technology Sydney
- Qingan Yan Wuhan Univ.
- Fisher Yu Princeton Univ.
- Hang Zhang Rutgers Univ.
Tuesday, July 25

0730–1700 Registration (Main Lobby)

0730–0830 Breakfast (Kamehameha II)

0830–1000 Session 4-1A: Machine Learning 4
(Kamehameha III)

Papers in this session are also in Poster Session P4-1.

Chairs: Greg Shakhnarovich (TTI Chicago)
Lior Wolf (Tel Aviv Univ.)

0830 Spotlights (S4-1A)
Format (4 min. for presentation; no questions)
1. [0830] Hidden Layers in Perceptual Learning, Gad Cohen, Daphna Weinshall
3. [0838] Hallucinating Very Low-Resolution Unaligned and Noisy Face Images by Transformative Discriminative Autoencoders, Xin Yu, Fatih Porikli
5. [0846] Deep Hashing Network for Unsupervised Domain Adaptation, Hemanth Venkateswara, Jose Eusebio, Shayok Chakraborty, Sethuraman Panchanathan
7. [0854] Deep Learning With Low Precision by Half-Wave Gaussian Quantization, Zhaowei Cai, Xiaodong He, Jian Sun, Nuno Vasconcelos

0903 Orals (O4-1A)
Format (12 min. for presentation + 2 min. for questions)
10. [0917] Full Resolution Image Compression With Recurrent Neural Networks, George Toderici, Damien Vincent, Nick Johnston, Sung Jin Hwang, David Minnen, Joel Shor, Michele Covell
11. [0931] Neural Face Editing With Intrinsic Image Disentangling, Zhixin Shu, Ersin Yumer, Sunil Hadap, Kalyan Sunkavalli, Eli Shechtman, Dimitris Samaras
12. [0945] Ubernet: Training a Universal Convolutional Neural Network for Low-, Mid-, and High-Level Vision Using Diverse Datasets and Limited Memory, Iasonas Kokkinos

0830–1000 Session 4-1B: Analyzing Humans with 3D Vision (Kalākaua Ballroom)

Papers in this session are also in Poster Session P4-1.

Chairs: Xilin Chen (Chinese Academy of Sciences)
Matthew Turk (UC Santa Barbara)

0830 Spotlights (S4-1B)
Format (4 min. for presentation; no questions)
15. [0838] Detailed, Accurate, Human Shape Estimation From Clothed 3D Scan Sequences, Chao Zhang, Sergi Pujades, Michael J. Black, Gerard Pons-Moll
16. [0842] POSEidon: Face-From-Depth for Driver Pose Estimation, Guido Borghi, Marco Venturelli, Roberto Vezzani, Rita Cucchiara
17. [0846] Human Shape From Silhouettes Using Generative HKS Descriptors and Cross-Modal Neural Networks, Endri Dibra, Himanshu Jain, Cengiz Öztireli, Remo Ziegler, Markus Gross
18. [0850] Parametric T-Spline Face Morphable Model for Detailed Fitting in Shape Subspace, Weilong Peng, Zhiyong Feng, Chao Xu, Yong Su
19. [0854] 3D Menagerie: Modeling the 3D Shape and Pose of Animals, Silvia Zuffi, Angjoo Kanazawa, David W. Jacobs, Michael J. Black
20. [0858] iCaRL: Incremental Classifier and Representation Learning, Sylvestre-Alvise Rebuffi, Alexander Kolesnikov, Georg Sperl, Christoph H. Lampert

0903 Orals (O4-1B)
Format (12 min. for presentation + 2 min. for questions)
21. [0903] Recurrent 3D Pose Sequence Machines, Mude Lin, Liang Lin, Xiaodan Liang, Keze Wang, Hui Cheng
22. [0917] Learning Detailed Face Reconstruction From a Single Image, Elad Richardson, Matan Sela, Roy Or-El, Ron Kimmel

1000–1200 Poster Session P4-1 (Kamehameha I)

3D Computer Vision
25. Semantically Coherent Co-Segmentation and Reconstruction of Dynamic Scenes, Armin Mustafa, Adrian Hilton
26. On the Two-View Geometry of Unsynchronized Cameras, Cenek Albl, Zuzana Kukelova, Andrew Fitzgibbon, Jan Heller, Matej Smid, Tomas Pajdla
27. Using Locally Corresponding CAD Models for Dense 3D Reconstructions From a Single Image, Chen Kong, Chen-Hsuan Lin, Simon Lucey
28. Convex Global 3D Registration With Lagrangian Duality, Jesus Biales, Javier Gonzalez-Jimenez

Program
29. DeMoN: Depth and Motion Network for Learning Monocular Stereo, Benjamin Ummenhofer, Huizhong Zhou, Jonas Uhrig, Niko laus Mayer, Eddy Ilg, Alexey Dosovitskiy, Thomas Brox
30. 3D Bounding Box Estimation Using Deep Learning and Geometry, Arsalan Mousavian, Dragomir Anguelov, John Flynn, Jana Košecká
31. A Dataset for Benchmarking Image-Based Localization, Xun Sun, Yuanfan Xie, Pei Luo, Liang Wang

Analyzing Humans in Images
32. Asynchronous Temporal Fields for Action Recognition, Gunnar A. Sigurdsson, Santosh Divvala, Ali Farhadi, Abhinav Gupta
33. Sequential Person Recognition in Photo Albums With a Recurrent Network, Yao Li, Guosheng Lin, Bohan Zhuang, Lingqiao Liu, Chunjua Shen, Anton van den Hengel
34. Multi-Context Attention for Human Pose Estimation, Xiao Chu, Wei Yang, Wanli Ouyang, Cheng Ma, Alan L. Yuille, Xiaogang Wang
35. 3D Convolutional Neural Networks for Efficient and Robust Hand Pose Estimation From Single Depth Images, Liu Hao Ge, Hui Liang, Junsong Yuan, Daniel Thalmann
36. Lifting From the Deep: Convolutional 3D Pose Estimation From a Single Image, Denis Tome, Chris Russell, Lourdes Agapito
38. Deep Structured Learning for Facial Action Unit Intensity Estimation, Robert Walecki, Ognjen (Oggi) Rudovic, Vladimir Pavlovic, Björn Schuller, Maja Pantic
39. Simultaneous Facial Landmark Detection, Pose and Deformation Estimation Under Facial Occlusion, Yue Wu, Chao Gou, Qiang Ji
40. Self-Supervised Video Representation Learning With Odd-One-Out Networks, Basura Fernando, Hakan Bilen, Efstratios Gavves, Stephen Gould
41. Robust Joint and Individual Variance Explained, Christos Sagonas, Yannis Panagakis, Alina Leidinger, Stefanos Zafeiriou
| 42. | Discriminative Covariance Oriented Representation Learning for Face Recognition With Image Sets, Wen Wang, Ruiping Wang, Shiguang Shan, Xilin Chen |
| 43. | 3D Human Pose Estimation = 2D Pose Estimation + Matching, Ching-Hang Chen, Deva Ramanan |

**Applications**

| 44. | Joint Gap Detection and Inpainting of Line Drawings, Kazuma Sasaki, Satoshi Iizuka, Edgar Simo-Serra, Hiroshi Ishikawa |

**Biomedical Image/Video Analysis**


**Computational Photography**

| 47. | Multiple-Scattering Microphysics Tomography, Aviad Levis, Yoav Y. Schechner, Anthony B. Davis |

**Image Motion & Tracking**

| 48. | Accurate Optical Flow via Direct Cost Volume Processing, Jia Xu, René Ranftl, Vladlen Koltun |
| 49. | Event-Based Visual Inertial Odometry, Alex Zihao Zhu, Nikolay Atanasov, Kostas Daniilidis |
| 50. | Robust Visual Tracking Using Oblique Random Forests, Le Zhang, Jagannadan Varadarajan, Ponnuthurai Nagaratnam Suganthan, Narendra Ahuja, Pierre Moulin |

**Low- & Mid-Level Vision**

| 51. | Deep Laplacian Pyramid Networks for Fast and Accurate Super-Resolution, Wei-Sheng Lai, Jia-Bin Huang, Narendra Ahuja, Ming-Hsuan Yang |
| 52. | Learning Non-Lambertian Object Intrinsics Across ShapeNet Categories, Jian Shi, Yue Dong, Hao Su, Stella X. Yu |
| 53. | MCMLSD: A Dynamic Programming Approach to Line Segment Detection, Emilio J. Almazàn, Ron Tal, Yiming Qian, James H. Elder |

| 55. | Richer Convolutional Features for Edge Detection, Yun Liu, Ming-Ming Cheng, Xiaowei Hu, Kai Wang, Xiang Bai |
| 56. | Non-Local Color Image Denoising With Convolutional Neural Networks, Stamatios Lefkimmiatis |
| 57. | Generative Face Completion, Yijun Li, Sifei Liu, Jimei Yang, Ming-Hsuan Yang |
| 58. | Hyper-Laplacian Regularized Unidirectional Low-Rank Tensor Recovery for Multispectral Image Denoising, Yi Chang, Luxin Yan, Sheng Zhong |
| 59. | Unsupervised Semantic Scene Labeling for Streaming Data, Maggie Wigness, John G. Rogers III |
| 60. | Why You Should Forget Luminance Conversion and Do Something Better, Rang M. H. Nguyen, Michael S. Brown |
| 61. | Deep Semantic Feature Matching, Nikolai Ufer, Björn Ommer |

**Machine Learning**

| 62. | Revisiting the Variable Projection Method for Separable Nonlinear Least Squares Problems, Je Hyeong Hong, Christopher Zach, Andrew Fitzgibbon |
| 63. | Efficient Multiple Instance Metric Learning Using Weakly Supervised Data, Marc T. Law, Yaoiliang Yu, Raquel Urtasun, Richard S. Zemel, Eric P. Xing |
| 64. | WILDCAT: Weakly Supervised Learning of Deep ConvNets for Image Classification, Pointwise Localization and Segmentation, Thibaut Durand, Taylor Mordan, Nicolas Thome, Matthieu Cord |
| 66. | Deep Roots: Improving CNN Efficiency With Hierarchical Filter Groups, Yani Ioannou, Duncan Robertson, Roberto Cipolla, Antonio Criminisi |
| 67. | Aggregated Residual Transformations for Deep Neural Networks, Saining Xie, Ross Girshick, Piotr Dollár, Zhuowen Tu, Kaiming He |
| 68. | MIML-FCN+: Multi-Instance Multi-Label Learning via Fully Convolutional Networks With Privileged Information, Hao Yang, Joey Tianyi Zhou, Jianfei Cai, Yew Soon Ong |
| 69. | Low-Rank Embedded Ensemble Semantic Dictionary for Zero-Shot Learning, Zhengming Ding, Ming Shao, Yun Fu |
Factorized Variational Autoencoders for Modeling Audience Reactions to Movies, Zhiwei Deng, Rajitha Navarathna, Peter Carr, Stephan Mandt, Yisong Yue, Iain Matthews, Greg Mori

Learning Features by Watching Objects Move, Deepak Pathak, Ross Girshick, Piotr Dollár, Trevor Darrell, Bharath Hariharan

What Can Help Pedestrian Detection? Jiayuan Mao, Tete Xiao, Yuning Jiang, Zhimin Cao

DeepPermNet: Visual Permutation Learning, Rodrigo Santa Cruz, Basura Fernando, Anoop Cherian, Stephen Gould

Learning the Multilinear Structure of Visual Data, Mengjiao Wang, Yannis Panagakis, Patrick Snape, Stefanos Zafeiriou

Adaptive and Move Making Auxiliary Cuts for Binary Pairwise Energies, Lena Gorelick, Yuri Boykov, Olga Veksler

Joint Multi-Person Pose Estimation and Semantic Part Segmentation, Fangting Xia, Peng Wang, Xianjie Chen, Alan L. Yuille

Deep Feature Interpolation for Image Content Changes, Paul Upchurch, Jacob Gardner, Geoff Pleiss, Robert Pless, Noah Snavely, Kavita Bala, Vivienne Sze

FASON: First and Second Order Information Fusion Network for Texture Recognition, Xiyang Dai, Joe Yue-Hei Ng, Larry S. Davis

Lean Crowdsourcing: Combining Humans and Machines in an Online System, Steve Branson, Grant Van Horn, Pietro Perona

Object Recognition & Scene Understanding

Supervising Neural Attention Models for Video Captioning by Human Gaze Data, Youngjae Yu, Jongwook Choi, Yeonhwa Kim, Kyung Yoo, Sang-Hun Lee, Gunhee Kim

L2-Net: Deep Learning of Discriminative Patch Descriptor in Euclidean Space, Yurun Tian, Bin Fan, Fuchao Wu

Convolutional Random Walk Networks for Semantic Image Segmentation, Gedas Bertasius, Lorenzo Torresani, Stella X. Yu, Jianbo Shi

Knowledge Acquisition for Visual Question Answering via Iterative Querying, Yuke Zhu, Joseph J. Lim, Li Fei-Fei

Memory-Augmented Attribute Manipulation Networks for Interactive Fashion Search, Bo Zhao, Jiashi Feng, Xiao Wu, Shuicheng Yan

From Zero-Shot Learning to Conventional Supervised Classification: Unseen Visual Data Synthesis, Yang Long, Li Liu, Ling Shao, Fumin Shen, Guiguang Ding, Jungong Han

Are Large-Scale 3D Models Really Necessary for Accurate Visual Localization? Torsten Sattler, Akihiko Torii, Josef Sivic, Marc Pollefeys, Hajime Taira, Masatoshi Okutomi, Tomas Pajdla

Asymmetric Feature Maps With Application to Sketch Based Retrieval, Giorgos Tolias, Ondřej Chum

Diverse Image Annotation, Baoyuan Wu, Fan Jia, Wei Liu, Bernard Ghanem

AMC: Attention guided Multi-modal Correlation Learning for Image Search, Kan Chen, Trung Bui, Chen Fang, Zhaowen Wang, Ram Nevatia

Multi-Attention Network for One Shot Learning, Peng Wang, Lingqiao Liu, Chunhua Shen, Zi Huang, Anton van den Hengel, Heng Tao Shen

Fried Binary Embedding for High-Dimensional Visual Features, Weixiang Hong, Junsong Yuan, Sreyasee Das Bhattacharjee

Pyramid Scene Parsing Network, Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, Jiaya Jia

Learning Deep Match Kernels for Image-Set Classification, Haoliang Sun, Xiantong Zhen, Yuanjie Zheng, Gongping Yang, Yilong Yin, Shuo Li

Task-Driven Dynamic Fusion: Reducing Ambiguity in Video Description, Xishan Zhang, Ke Gao, Yongdong Zhang, Dongming Zhang, Jintao Li, Qi Tian

Learning Multifunctional Binary Codes for Both Category and Attribute Oriented Retrieval Tasks, Haomiao Liu, Ruiping Wang, Shiguang Shan, Xilin Chen

Indoor Scene Parsing With Instance Segmentation, Semantic Labeling and Support Relationship Inference, Wei Zhuo, Mathieu Salzmann, Xuming He, Miaomiao Liu
98. Episodic CAMN: Contextual Attention-Based Memory Networks With Iterative Feedback for Scene Labeling, Abrar H. Abdulnabi, Bing Shuai, Stefan Winkler, Gang Wang

99. Link the Head to the “Beak“: Zero Shot Learning From Noisy Text Description at Part Precision, Mohamed Elhoseiny, Yizhe Zhu, Han Zhang, Ahmed Elgammal

100. SCA-CNN: Spatial and Channel-Wise Attention in Convolutional Networks for Image Captioning, Long Chen, Hanwang Zhang, Jun Xiao, Liqiang Nie, Jian Shao, Wei Liu, Tat-Seng Chua

101. Deep Pyramidal Residual Networks, Dongyoon Han, Jiwhan Kim, Junmo Kim

102. Product Split Trees, Artem Babenko, Victor Lempitsky

103. Making the v in VQA Matter: Elevating the Role of Image Understanding in Visual Question Answering, Yash Goyal, Tejas Khot, Douglas Summers-Stay, Dhruv Batra, Devi Parikh


105. Cross-Modality Binary Code Learning via Fusion Similarity Hashing, Hong Liu, Rongrong Ji, Yongjian Wu, Feiyue Huang, Baochang Zhang

**Theory**


107. InterpoNet, a Brain Inspired Neural Network for Optical Flow Dense Interpolation, Shay Zweig, Lior Wolf

**Video Analytics**


109. Video Segmentation via Multiple Granularity Analysis, Rui Yang, Bingbing Ni, Chao Ma, Yi Xu, Xiaokang Yang

110. Spatio-Temporal Alignment of Non-Overlapping Sequences From Independently Panning Cameras, Seyed Morteza Safdarnajad, Xiaoming Liu

111. UntrimmedNets for Weakly Supervised Action Recognition and Detection, Limin Wang, Yuanjun Xiong, Dahua Lin, Luc Van Gool

---

**1000–1200 Demos** (Kamehameha I)

- Open-Source Simulator in UE4: Photorealistic Fully Annotated Datasets, Real-Time Object Tracking With VR, Deep Learning Interface, Matthias Mueller, Neil Smith, Bernard Ghanem (King Abdullah Univ. of Science and Technology)
- CloudCV, Deshraj Yadav, Viraj Prabhu, Prithvijit Chattopadhyay, Abhishek Das, Devi Parikh, Dhruv Batra (Virginia Tech, Georgia Tech)
- Procedural Human Action Videos, César De Souza, Yohann Cabon, Adrien Gaidon, Antonio M. Lopez (Xerox Research Centre Europe)
- Hierarchical 3D Fully Convolutional Networks for Multi-Organ and Vessel Segmentation Used in Surgical Navigation, Holger R. Roth, Yusuke Tatemura, Hirohisa Oda, Yuichiro Hayashi, Masahiro Oda, Natsuki Shimizu, Michitaka Fujiwara, Kazunari Misawa, Kensaku Mori (Nagoya Univ., Nagoya Univ. Graduate School of Medicine, Aichi Cancer Center)

---

**1000–1200 Exhibits** (Kamehameha I)

- Same as Saturday morning Exhibits (see pg. 10)

---

**1200–1300 Lunch** (Kamehameha II)

---

**Notes:**

---

37
1300–1430 Session 4-2A: Object Recognition & Scene Understanding 3
(Kamehameha III)
Papers in this session are also in Poster Session P4-2.

Chairs: Kostas Daniilidis (Univ. of Pennsylvania)
Junsong Yuan (Nanyang Technological Univ)

1300 Spotlights (S4-2A)
Format (4 min. for presentation; no questions)
1. [1300] Gaze Embeddings for Zero-Shot Image Classification, Nour Karessli, Zeynep Akata, Bernt Schiele, Andreas Bulling
3. [1308] Attend to You: Personalized Image Captioning With Context Sequence Memory Networks, Cesc Chunseong Park, Byeongchang Kim, Gunhee Kim
4. [1312] Adversarially Tuned Scene Generation, VSR Veeravasarapu, Constantin Rothkopf, Ramesh Visvanathan
5. [1316] Residual Attention Network for Image Classification, Fei Wang, Mengqing Jiao, Chen Qian, Shuo Yang, Cheng Li, Honggang Zhang, Xiaogang Wang, Xiaowu Tang
7. [1324] Learning Non-Maximum Suppression, Jan Hosang, Rodrigo Benenson, Bernt Schiele

1333 Orals (O4-2A)
Format (12 min. for presentation + 2 min. for questions)
9. [1333] Object Region Mining With Adversarial Erasing: A Simple Classification to Semantic Segmentation Approach, Yunchao Wei, Jiashi Feng, Xiaodan Liang, Ming-Ming Cheng, Yao Zhao, Shuicheng Yan
10. [1347] Fine-Grained Recognition as HSnet Search for Informative Image Parts, Michael Lam, Behrooz Mahassani, Sinisa Todorovic
11. [1401] G²DeNet: Global Gaussian Distribution Embedding Network and Its Application to Visual Recognition, Qilong Wang, Peihua Li, Lei Zhang

1300–1430 Session 4-2B: Machine Learning for 3D Vision (Kalākaua Ballroom)
Papers in this session are also in Poster Session P4-2.

Chairs: Dieter Fox (Univ. of Washington)
Marc Pollefeys (ETH Zurich)

1300 Spotlights (S4-2B)
Format (4 min. for presentation; no questions)
13. [1300] Multi-View 3D Object Detection Network for Autonomous Driving, Xiaozhi Chen, Huimin Ma, Ji Wan, Bo Li, Tian Xia
14. [1304] UltraStereo: Efficient Learning-Based Matching for Active Stereo Systems, Sean Ryan Fanello, Julien Valentin, Christoph Rheemann, Adarsh Kowdle, Vladimir Tankovich, Philip Davidson, Shahram Izadi
15. [1308] Shape Completion Using 3D-Encoder-Predictor CNNs and Shape Synthesis, Angela Dai, Charles Ruizhongtai Qi, Matthias Nießer
16. [1312] Geometric Loss Functions for Camera Pose Regression With Deep Learning, Alex Kendall, Roberto Cipolla
17. [1316] CNN-SLAM: Real-Time Dense Monocular SLAM With Learned Depth Prediction, Keisuke Tateno, Federico Tombari, Iro Laina, Nassir Navab
18. [1320] Learning From Noisy Large-Scale Datasets With Minimal Supervision, Andreas Veit, Neil Alldrin, Gal Chechik, Ivan Krasin, Abhinav Gupta, Serge Belongie
20. [1328] Non-Local Deep Features for Salient Object Detection, Zhiming Luo, Akshaya Mishra, Andrew Achkar, Justin Eichel, Shaohui Li, Pierre-Marc Jodoin
1333 Orals (O4-2B)

Format (12 min. for presentation + 2 min. for questions)

21. [1333] Unsupervised Monocular Depth Estimation With Left-Right Consistency, Clément Godard, Oisin Mac Aodha, Gabriel J. Brostow

22. [1347] Unsupervised Learning of Depth and Ego-Motion From Video, Tinghui Zhou, Matthew Brown, Noah Snavely, David G. Lowe


24. [1415] 3D Shape Segmentation With Projective Convolutional Networks, Evangelos Kalogerakis, Melinos Averkiou, Subhransu Maji, Siddhartha Chaudhuri

1430–1515 Break (Kamehameha II)

1430–1630 Poster Session P4-2 (Kamehameha I)

3D Computer Vision

25. SGM-Nets: Semi-Global Matching With Neural Networks, Akihito Seki, Marc Pollefeys

26. Stereo-Based 3D Reconstruction of Dynamic Fluid Surfaces by Global Optimization, Yiming Qian, Minglun Gong, Yee-Hong Yang

27. Fine-To-Coarse Global Registration of RGB-D Scans, Maciej Halber, Thomas Funkhouser

28. Analyzing Computer Vision Data - The Good, the Bad and the Ugly, Oliver Zendel, Katrin Honauer, Markus Murschitz, Martin Humenberger, Gustavo Fernández Domínguez

29. Product Manifold Filter: Non-Rigid Shape Correspondence via Kernel Density Estimation in the Product Space, Matthias Vestner, Roei Litman, Emanuele Rodolà, Alex Bronstein, Daniel Cremers


31. Toroidal Constraints for Two-Point Localization Under High Outlier Ratios, Federico Camposeco, Torsten Sattler, Andrea Cohen, Andreas Geiger, Marc Pollefeys

32. 4D Light Field Superpixel and Segmentation, Hao Zhu, Qi Zhang, Qing Wang

33. Exploiting Symmetry and/or Manhattan Properties for 3D Object Structure Estimation From Single and Multiple Images, Yuan Gao, Alan L. Yuille

Analyzing Humans in Images

34. Binary Coding for Partial Action Analysis With Limited Observation Ratios, Jie Qin, Li Liu, Ling Shao, Bingbing Ni, Chen Chen, Fumin Shen, Yunhong Wang

35. SphereFace: Deep Hypersphere Embedding for Face Recognition, Weiyang Liu, Yandong Wen, Zhiding Yu, Ming Li, Bhiksha Raj, Le Song

36. IRINA: Iris Recognition (Even) in Inaccurately Segmented Data, Hugo Proença, João C. Neves

37. Look Into Person: Self-Supervised Structure-Sensitive Learning and a New Benchmark for Human Parsing, Ke Gong, Xiaodan Liang, Dongyu Zhang, Xiaohui Shen, Liang Lin

38. Action Unit Detection With Region Adaptation, Multi-Labeling Learning and Optimal Temporal Fusing, Wei Li, Farnaz Abtahi, Zhigang Zhu

39. See the Forest for the Trees: Joint Spatial and Temporal Recurrent Neural Networks for Video-Based Person Re-identification, Zhen Zhou, Yan Huang, Wei Wang, Liang Wang, Tieniu Tan

40. Joint Intensity and Spatial Metric Learning for Robust Gait Recognition, Yasushi Makihara, Atsuyuki Suzuki, Daigo Muramatsu, Xiang Li, Yasushi Yagi

41. Pose-Aware Person Recognition, Vijay Kumar, Anoop Namboodiri, Manohar Paluri, C. V. Jawahar

42. Not Afraid of the Dark: NIR-VIS Face Recognition via Cross-Spectral Hallucination and Low-Rank Embedding, José Lezama, Qiang Qiu, Guillermo Sapiro

Applications

43. Jointly Learning Energy Expenditures and Activities Using Egocentric Multimodal Signals, Katsuyuki Nakamura, Serena Yeung, Alexandre Alahi, Li Fei-Fei

44. Binarized Mode Seeking for Scalable Visual Pattern Discovery, Wei Zhang, Xiaochun Cao, Rui Wang, Yuanfang Guo, Zhineng Chen
45. Scribbler: Controlling Deep Image Synthesis With Sketch and Color, Patsorn Sangkloy, Jingwan Lu, Chen Fang, Fisher Yu, James Hays

**Biomedical Image/Video Analysis**

46. Multi-Way Multi-Level Kernel Modeling for Neuroimaging Classification, Lifang He, Chun-Ta Lu, Hao Ding, Shen Wang, Linlin Shen, Philip S. Yu, Ann B. Ragin

47. WSISA: Making Survival Prediction From Whole Slide Histopathological Images, Xinliang Zhu, Jiawen Yao, Feiyun Zhu, Junzhou Huang

**Computational Photography**

48. On the Effectiveness of Visible Watermarks, Tali Dekel, Michael Rubinstein, Ce Liu, William T. Freeman

49. Snapshot Hyperspectral Light Field Imaging, Zhiwei Xiong, Lizhi Wang, Huijun Li, Dong Liu, Feng Wu


**Image Motion & Tracking**

51. Fast Multi-Frame Stereo Scene Flow With Motion Segmentation, Tatsunori Taniai, Sudipta N. Sinha, Yoichi Sato

52. Improved Stereo Matching With Constant Highway Networks and Reflective Confidence Learning, Amit Shaked, Lior Wolf

53. Optical Flow in Mostly Rigid Scenes, Jonas Wulff, Laura Sevilla-Lara, Michael J. Black

54. Optical Flow Requires Multiple Strategies (but Only One Network), Tal Schuster, Lior Wolf, David Gadot

55. ECO: Efficient Convolution Operators for Tracking, Martin Danelljan, Goutam Bhat, Fahad Shahbaz Khan, Michael Felsberg

**Low- & Mid-Level Vision**

56. Differential Angular Imaging for Material Recognition, Jia Xue, Hang Zhang, Kristin Dana, Ko Nishino

57. Fast Fourier Color Constancy, Jonathan T. Barron, Yun-Ta Tsai

58. Comparative Evaluation of Hand-Crafted and Learned Local Features, Johannes L. Schönberger, Hans Hardmeier, Torsten Sattler, Marc Pollefeys

59. Learning Fully Convolutional Networks for Iterative Non-Blind Deconvolution, Jiawei Zhang, Jinshan Pan, Weisheng Lai, Rynson W. H. Lau, Ming-Hsuan Yang

60. Image Deblurring via Extreme Channels Prior, Yanyang Yan, Wengi Ren, Yuanfang Guo, Rui Wang, Xiaochun Cao

61. Simultaneous Stereo Video Deblurring and Scene Flow Estimation, Liyuan Pan, Yuchao Dai, Miaomiao Liu, Fatih Porikli

62. Deep Photo Style Transfer, Fujun Luan, Sylvain Paris, Eli Shechtman, Kavita Bala

63. Generative Attribute Controller With Conditional Filtered Generative Adversarial Networks, Takuhiro Kaneko, Kaoru Hiramatsu, Kunio Kashino

64. Fast Haze Removal for Nighttime Image Using Maximum Reflectance Prior, Jing Zhang, Yang Cao, Shuai Fang, Yu Kang, Chang Wen Chen

**Machine Learning**

65. Low-Rank Bilinear Pooling for Fine-Grained Classification, Shu Kong, Charless Fowlkes

66. Neural Scene De-Rendering, Jiajun Wu, Joshua B. Tenenbaum, Pushmeet Kohli

67. Real-Time Neural Style Transfer for Videos, Haozhi Huang, Hao Wang, Wenhan Luo, Lin Ma, Wenhao Jiang, Xiaolong Zhu, Zhifeng Li, Wei Liu

68. A Graph Regularized Deep Neural Network for Unsupervised Image Representation Learning, Shijie Yang, Liang Li, Shuhui Wang, Weigang Zhang, Qingming Huang

69. A Study of Lagrangean Decompositions and Dual Ascent Solvers for Graph Matching, Paul Swoboda, Carsten Rother, Hassan Abu Alhaija, Dagmar Kainmüller, Bogdan Savchynskyy

70. Collaborative Deep Reinforcement Learning for Joint Object Search, Xiangyu Kong, Bo Xin, Yizhou Wang, Gang Hua

71. Loss Max-Pooling for Semantic Image Segmentation, Samuel Rota Bulò, Gerhard Neuhold, Peter Kontschieder

72. Deep View Morphing, Dinghuang Ji, Junghyun Kwon, Max McFarland, Silvio Savarese

73. Unsupervised Learning of Long-Term Motion Dynamics for Videos, Zelun Luo, Boya Peng, De-An Huang, Alexandre Alahi, Li Fei-Fei
74. Revisiting Metric Learning for SPD Matrix Based Visual Representation, Luping Zhou, Lei Wang, Jianjia Zhang, Yinghuan Shi, Yang Gao
75. Expert Gate: Lifelong Learning With a Network of Experts, Rahaf Aljundi, Punarjay Chakravarty, Tinne Tuytelaars
76. A Gift From Knowledge Distillation: Fast Optimization, Network Minimization and Transfer Learning, Junho Yim, Donggyu Joo, Jihoon Bae, Junmo Kim
77. Domain Adaptation by Mixture of Alignments of Second- or Higher-Order Scatter Tensors, Piotr Koniusz, Yusuf Tas, Fatih Porikli
78. Deep Mixture of Linear Inverse Regressions Applied to Head-Pose Estimation, Stéphane Lathuilière, Rémi Juge, Pablo Mesejo, Rafael Muñoz-Salinas, Radu Horaud
79. STD2P: RRGB Semantic Segmentation Using Spatio-Temporal Data-Driven Pooling, Yang He, Wei-Chen Chiu, Margret Keuper, Mario Fritz
80. Harmonic Networks: Deep Translation and Rotation Equivariance, Daniel E. Worrall, Stephan J. Garbin, Daniyar Turmukhambetov, Gabriel J. Brostow
82. Detect, Replace, Refine: Deep Structured Prediction for Pixel Wise Labeling, Spyros Gidaris, Nikos Komodakis
83. Weighted-Entropy-Based Quantization for Deep Neural Networks, Eunhyeok Park, Junwhan Ahn, Sungjoo Yoo
84. Residual Expansion Algorithm: Fast and Effective Optimization for Nonconvex Least Squares Problems, Daiki Ikami, Toshihiko Yamasaki, Kiyoharu Aizawa
85. Bidirectional Beam Search: Forward-Backward Inference in Neural Sequence Models for Fill-In-The-Blank Image Captioning, Qing Sun, Stefan Lee, Dhruv Batra
86. Newton-Type Methods for Inference in Higher-Order Markov Random Fields, Hariprasad Kannan, Nikos Komodakis, Nikos Paragios

Object Recognition & Scene Understanding
88. ViP-CNN: Visual Phrase Guided Convolutional Neural Network, Yikang Li, Wanli Ouyang, Xiaogang Wang, Xiao'ou Tang
89. Instance-Aware Image and Sentence Matching With Selective Multimodal LSTM, Yan Huang, Wei Wang, Liang Wang
90. Kernel Square-Loss Exemplar Machines for Image Retrieval, Rafael S. Rezende, Joaquin Zepeda, Jean Ponce, Francis Bach, Patrick Pérez
92. Combining Bottom-Up, Top-Down, and Smoothness Cues for Weakly Supervised Image Segmentation, Anirban Roy, Sinisa Todorovic
93. Seeing Into Darkness: Scotopic Visual Recognition, Bo Chen, Pietro Perona
94. Deep Co-Occurrence Feature Learning for Visual Object Recognition, Ya-Fang Shih, Yang-Ming Yeh, Yen-Yu Lin, Ming-Fang Weng, Yi-Chang Lu, Yung-Yu Chuang
96. InstanceCut: From Edges to Instances With MultiCut, Alexander Kirillov, Evgeny Levinkov, Bjoern Andres, Bogdan Savchynskyy, Carsten Rother
97. Fine-Grained Image Classification via Combining Vision and Language, Xiangteng He, Xinhua Huang, Junjie Yan
98. Mimicking Very Efficient Network for Object Detection, Quanquan Li, Shengyin Jin, Junjie Yan
100. A Dataset and Exploration of Models for Understanding Video Data Through Fill-In-The-Blank Question-Answering, Tegan Maharaj, Nicolas Ballas, Anna Rohrbach, Aaron Courville, Christopher Pal
101. Learning Detection With Diverse Proposals, Samaneh Azadi, Jiashi Feng, Trevor Darrell
Tuesday, July 25 (Afternoon)

102. Skeleton Key: Image Captioning by Skeleton-Attribute Decomposition, Yufei Wang, Zhe Lin, Xiaohui Shen, Scott Cohen, Garrison W. Cottrell

Theory

103. A Low Power, Fully Event-Based Gesture Recognition System, 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, Dharmendra Modha

Video Analytics

104. Learning Deep Context-Aware Features Over Body and Latent Parts for Person Re-Identification, Dangwei Li, Xiaotang Chen, Zhang Zhang, Kaiqi Huang

105. Recurrent Modeling of Interaction Context for Collective Activity Recognition, Minsi Wang, Bingbing Ni, Xiaokang Yang

106. Primary Object Segmentation in Videos Based on Region Augmentation and Reduction, Yeong Jun Koh, Chang-Su Kim

107. ROAM: A Rich Object Appearance Model With Application to Rotoscopying, Ondrej Miksik, Juan-Manuel Pérez-Rúa, Philip H. S. Torr, Patrick Pérez

108. Temporal Residual Networks for Dynamic Scene Recognition, Christoph Feichtenhofer, Axel Pinz, Richard P. Wildes

109. Spatiotemporal Multiplier Networks for Video Action Recognition, Christoph Feichtenhofer, Axel Pinz, Richard P. Wildes

110. Learning to Learn From Noisy Web Videos, Serena Yeung, Vignesh Ramanathan, Olga Russakovsky, Liyue Shen, Greg Mori, Li Fei-Fei

111. YouTube-Bounding Boxes: A Large High-Precision Human-Annotated Data Set for Object Detection in Video, Esteban Real, Jonathon Shlens, Stefano Mazzocchi, Xin Pan, Vincent Vanhoucke

112. Online Video Object Segmentation via Convolutional Trident Network, Won-Dong Jang, Chang-Su Kim

1430–1630 Demos (Kamehameha I)
- Same as Tuesday morning Demos (see pg. 37)

1430–1630 Exhibits (Kamehameha I)
- Same as Saturday morning Exhibits (see pg. 10)

1645–1800 Plenary Session (Kamehameha III)


Abstract: The fields of neuroscience and cognitive science are hard at work on one of our last great scientific quests — to reverse engineer the human mind. In comparison to other areas of science, these sciences are still in their infancy. Not surprisingly, forward engineering approaches that aim to emulate human intelligence in artificial systems (AI) are also still in their infancy. Yet the intelligence and cognitive flexibility apparent in human behavior are an existence proof that machines can be constructed to emulate and work alongside the human mind. In this talk, I will argue that these challenges of reverse engineering the mind will be solved by tightly combining the efforts of brain and cognitive scientists (hypothesis generation and data acquisition), and forward engineering aiming to emulate the mind (hypothesis instantiation and data prediction). To support that thesis, I will focus on one aspect of perceptual intelligence — object categorization and detection — and I will tell the story of how work in brain science, cognitive science and computer science converged to create deep neural networks that can support such tasks. These networks not only reach human performance for many images, but their internal workings are modeled after — and largely emulate — the internal workings of the primate visual system. Yet, the primate visual system (NI) still outperforms current generation artificial deep neural networks (AI), and I will show some new clues that neuroscience can offer. More broadly, this is just the beginning of the last great human science quest — to understand natural intelligence — and I hope to motivate others to engage that frontier alongside us.

END OF MAIN CVPR 2017 CONFERENCE