



IEEE Computer Society  
Conference on  
**Computer Vision and  
Pattern Recognition**

Pocket Guide

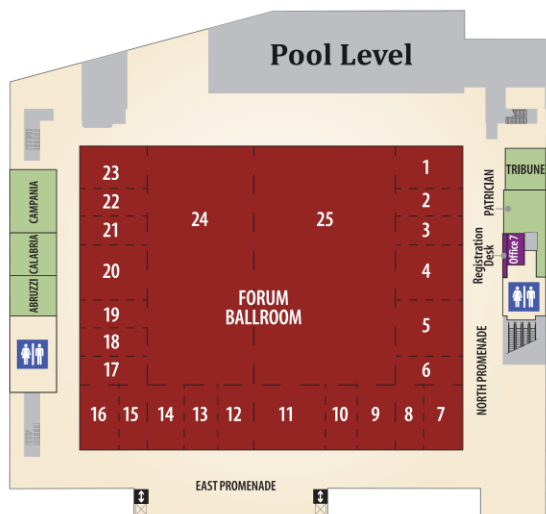
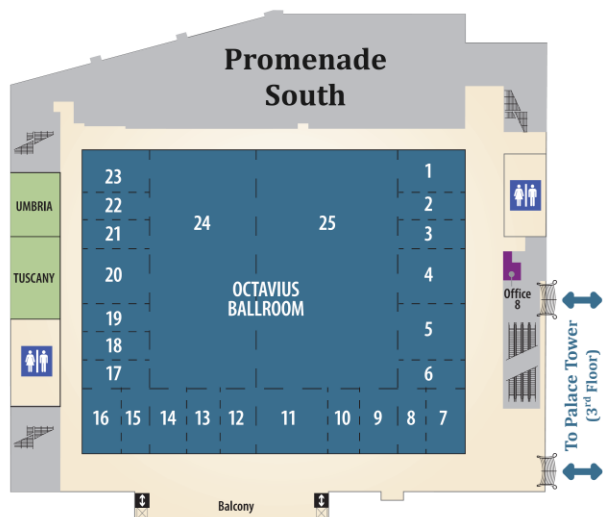
**CVPR**

June 26 – July 1, 2016

**Las Vegas, NV**

# Caesars Palace Convention Center

# CVPR 2016



**CVPR 2016 Sponsors:**

IEEE  
computer  
society

CVF The  
Computer  
Vision  
Foundation

# Message from the General and Program Chairs

Welcome to Las Vegas and the 29th IEEE 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 2016 has a number of colocated events, including 29 workshops and 22 tutorials. As the field of artificial intelligence has become a major player in the technology world, this year's CVPR has made history in a number of exciting ways.

First and foremost, we received a record 2145 valid submissions to the main conference, of which 1865 were fully reviewed (the others were administratively rejected for technical or ethical reasons or were withdrawn before review). To select papers from these submissions, we invited 72 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 team of 1141 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 643 papers (29.9% of valid submissions). 83 of these were accepted as oral presentations (3.9% of valid submissions) and 123 were accepted as spotlight presentations (for a total of 9.7% of valid submissions with live presentations). All papers will also appear in the interactive poster sessions.

CVPR 2016 introduces a new form of presentation — “Spotlights”. Spotlights are short oral presentations aimed at highlighting the main contributions, novelties and results of papers. Our rationale behind adding spotlight presentations to CVPR 2016 was to increase the number of works that get presented to a large audience thus gaining maximum exposure and visibility. We believe that this new presentation format (oral/spotlight/poster) will make CVPR 2016 a more exciting conference.

The review process was similar to previous years. The Program Chairs made an effort to assign papers on similar topics to the same set of Area Chairs (ACs), so that related papers could be compared directly. To achieve this, between 70 and 130 papers were pre-assigned to each AC with the help of the Toronto Paper Matching System while accounting for author provided suggestions. ACs then went over the paper titles and abstracts and bid on those they would like to handle. Using all of this information as input, ACs received a final assignment of around 30 papers for which they provided ranked lists of candidate reviewers. ACs were asked to suggest candidate reviewers who were experts in the topic, had a broad view of the field, were reliable, and could provide high quality reviews. Their choices were additionally guided by the reviewers' bids (the process followed for reviewer bidding was similar to that used for ACs). For 99.9% of the papers all 3 of the assigned reviewers were amongst those listed by ACs as candidate reviewers. For 70% of the papers at least 2 of the assigned reviewers were ranked by ACs amongst their top 3 candidates for that paper, and for 95% of the papers at least 1 reviewer was ranked amongst the top 3 candidates. 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 12 ACs after extensive discussion.

The Program Chairs did not submit any papers to CVPR 2016, allowing them to avoid direct conflicts throughout the review process. This year, General Chairs were allowed to submit papers, and therefore did not have any software access to the CMT system beyond that of an author. Additionally, CVPR 2016 is the first CVPR to use Researcher.cc to identify conflicts between reviewers and authors. Reviewers and ACs were excluded from any access to papers from their research groups, affiliated institutions, or collaborators. The double-blind nature of the CVPR review process was thus strictly maintained throughout.

## Message from the General and Program Chairs

Since its inception in the 1960s, Computer Vision has grown into one of the most important areas of research and development in modern artificial intelligence and computer science. As we observe the blossom of computer vision technologies applied in industry, we take this opportunity at CVPR 2016 to celebrate the coming of age of computer vision, and to facilitate an even closer tie between our academic world and the industry world. As a result, for the first time in CVPR's history, we are providing an exciting, "trade-show" like atmosphere to foster maximal visibility and exposure for each on-site exhibitor from promising startups and creative standouts to the biggest industry leaders. The technologies developed by these companies range from software to hardware, and demonstrate the great impact of computer vision on different industry areas. As of now, more than 70 companies have signed up to participate in our Expo. In addition to on site booths, we also offer a unique web-presence video this year, enabling recruiting opportunities and technology highlights. So far, over 20 companies are committed to participate in this special format of presence. Furthermore, companies in the computer vision industry are showing strong interest in sponsoring the CVPR community. Together they have contributed a historical amount of sponsorship funding. Additionally, more than 10 companies showed interest in sponsoring promotional items, resulting in a significant amount of additional fundings for the conference.

CVPR 2016 has also made history in its diverse organizing committee. A record number of women are involved, from General Chairs to Area Chairs. The field of computer vision has historically been male-dominated. We are pleased to see a growing number of women and under-represented minority researchers in our community. CVPR 2016 has made an explicit statement to celebrate their contributions to our field. We hope that at the conference, you get to know some of the women leaders of our field, and encourage more women and under-represented minority students to study and work in this exciting field. We truly hope that this is only the beginning!

Last but not least, we wish to thank all members of the Organizing Committee, the Area Chairs, reviewers, emergency reviewers, authors, and the CMT and TPMS teams for the immense amount of hard work and professionalism that has gone into making CVPR 2016 a first rate conference. Our thanks also go to the organizers of previous CVPRs, in particular Rene Vidal and Andrew Zisserman, for their helpful advice and support, and to Ari Kobren for a huge effort maintaining Researcher.cc. A few people working behind the scenes deserve special thanks. Yipin Zhou and Arsalan Mousavian provided critical technical assistance with CMT and managing the largest CVPR ever. It could not have been a success without them. Nicole Finn and Liz Ryan provided critical support to the general conference organization and financial well being of CVPR. Tali Treibitz and Dana Berman have provided high quality and rapid turnaround of the CVPR web pages. Eric Mortensen and Kate Saenko have again lent expertise developed over many years to shepherd the papers through the publication process. A special thank-you should also go to the Corporate Relations team - Jia Li, Kristin Dana, Brendan Morris and Andrea Frome. They have worked tirelessly to put together the inaugural CVPR Expo and to solicit industry support. Greg Mori worked with a number of teams and contributed tremendously to local arrangements, the AC meeting, CVPR website, and industry Expo. Last but not the least, this CVPR would not be possible without the record amount of support from our corporate sponsors and participants.

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

Program Chairs: **Lourdes Agapito, Tamara Berg, Jana Kosecka, Lihi Zelnik-Manor**

General Chairs: **Tinne Tuytelaars, Fei-Fei Li, Ruzena Bajcsy**

## CVPR 2016 Organizing Committee

|                         |                                                                     |                                     |                                         |
|-------------------------|---------------------------------------------------------------------|-------------------------------------|-----------------------------------------|
| <b>General Chairs:</b>  | Ruzena Bajcsy<br>Fei-Fei Li<br>Tinne Tuytelaars                     | <b>Program Coordination Chairs:</b> | Ira Kemelmacher-Shlizerman<br>Greg Mori |
| <b>Program Chairs:</b>  | Lourdes Agapito<br>Tamara Berg<br>Jana Kosecka<br>Lihi Zelnik-Manor | <b>Publications Chairs:</b>         | Eric Mortensen<br>Kate Saenko           |
| <b>Workshops Chair:</b> | Svetlana Lazebnik                                                   | <b>Corporate Relations Chairs:</b>  | Andrea Frome<br>Jia Li<br>Kristin Dana  |
| <b>Tutorials Chair:</b> | Sanja Fidler<br>Raquel Urtasun                                      | <b>Doctoral Consortium Chair:</b>   | Adriana Kovashka                        |
| <b>Finance Chairs:</b>  | Octavia Camps<br>Bryan Morse                                        | <b>Technology Chair:</b>            | Terry Boulton                           |
| <b>Website Chairs:</b>  | Tali Treibitz<br>Dana Berman                                        | <b>Local Arrangements Chair:</b>    | Bredan Morris                           |
| <b>Demos Chair:</b>     | Andrew Davison                                                      | <b>Publicity &amp; Press Chair:</b> | Olga Russakovsky                        |
| <b>Posters Chair:</b>   | Devi Parikh                                                         | <b>Student Volunteers Chair:</b>    | Tatiana Tommasi                         |
|                         |                                                                     | <b>Logistics Advisor:</b>           | Ginger Boulton                          |

## CVPR 2016 Area Chairs

|                      |                   |                   |                    |
|----------------------|-------------------|-------------------|--------------------|
| Tal Arbel            | Alyosha Efros     | Frederic Jurie    | Chris Pal          |
| Ronen Basri          | Irfan Essa        | Ira Kemelmacher   | Sylvain Paris      |
| Dhruv Batra          | Ryan Farrell      | Hedvig Kjellström | Florent Perronnin  |
| Serge Belongie       | Paolo Favaro      | Iasonas Kokkinos  | Thomas Pock        |
| Alex Berg            | Rogério Feris     | Vladlen Koltun    | Marc Ranzato       |
| Michal Black         | Sanja Fidler      | Nikos Komodakis   | Bryan Russell      |
| Lubomir Bourdev      | Andrew Fitzgibbon | Pawan Kumar       | Kate Saenko        |
| Edmond Boyer         | Charles Fowlkes   | Kyros Kutulakos   | Yoav Schechner     |
| Yuri Boykov          | Dieter Fox        | Ivan Laptev       | Greg Shakhnarovich |
| Gabriel Brostow      | Jan-Michael Frahm | Bastian Leibe     | Jianbo Shi         |
| Octavia Camps        | Pascal Fua        | Jongwoo Lim       | Richard Socher     |
| Manmohan Chandraker  | Ross Girshick     | Steve Lin         | Jian Sun           |
| Rama Chellapa        | Stephen Gould     | Zhouchen Lin      | Sabine Susstrunk   |
| Dorin Comaniciu      | Tal Hassner       | Haibin Ling       | Ayellet Tal        |
| Rita Cucchiara       | Kaiming He        | Yi Ma             | Jianxiong Xiao     |
| Kostas Daniilidis    | Adrian Hilton     | Julien Mairal     | Todd Zickler       |
| Fernando de la Torre | Michal Irani      | Michael Maire     | Harpreet Sawhney   |
| Jia Deng             | Jiayia Jia        | Subhransu Maji    | Nuno Vasconcelos   |

## Sunday, June 26

---

**NOTE:** Tutorial locations are in the online proceedings and the mobile app

**0730–1700 Registration** (Octavius Prefunction)

**0730–0830 Breakfast** (Forum Ballroom)

**1230–1345 Lunch** (Forum Ballroom)

### State of the Art and Recent Trends in Social Signal Processing

**Organizers:** Marco Cristani  
Vittorio Murino  
Alessandro Vinciarelli  
Bruno Lepri

**Time:** 0830-1800 (Full Day)

**Description:** Social Signal Processing is the domain aimed at studying social behaviour, in particular its nonverbal aspects, comprehensively covering the aspects of analysis, synthesis and modeling. To date, the field has focused on face-to-face interactions where it is possible to use the whole range of nonverbal cues that people utilizes to communicate or, more in general, conveying information during interplays. This scenario triggered several advancements in computer vision in order to capture subtle signals coming from gestures, facial expressions, vocalizations, and other explicit and implicit communication means, so as novel machine learning and pattern recognition strategies to embed those signals into behavioral models.

However, increasingly more interactions take place in unconstrained scenarios. In particular, groups and crowd, that have been the focus of the many surveillance studies so far, can be analyzed following a social signal processing perspective. This amounts to exploit proxemic and dynamic cues, which could indicate group/crowd membership, intention to aggregate in/leave a group and the tendency of being social/antisocial or suspect, bringing new tools to the surveillance field for example. But social signals are not only present in “real” social situation, but also in “virtual” settings, i.e., Internet. Nowadays, people can interact with virtually anybody at virtually every moment: According to the latest statistics, internet users worldwide are

now three billions (roughly 40% of the population), for a total of over two billions of active social media accounts (29% of the world population). These “virtual” social interactions take place through communication technologies that limit the use of nonverbal cues (e.g., phones do not allow one to display facial expressions) or require the adoption of artificial cues that do not belong to the natural repertoire of human beings (e.g., “likes” and emoticons on social media). This opens a new frontier for Social Signal Processing where the main questions are whether people still exchange social signals and, if so, what pattern recognition technologies are effective in this novel domain. The tutorial will investigate these two novel fields of the social signal processing field. In particular, it will illustrate how SSP can deal with open, “wild” actual scenarios and how social cues can be retrieved and utilised in virtual contexts.

### Low-Rank and Sparse Modeling for Visual Analytics

**Organizers:** René Vidal  
Ehsan Elhamifar  
Zhouchen Lin  
Jiashi Feng  
Sheng Li  
Yun Fu

**Time:** 0830-1800 (Full Day)

**Description:** Low-rank and sparse modeling are emerging mathematical tools dealing with uncertainties of real-world visual data. Leveraging on the underlying structure of data, low-rank and sparse modeling approaches have achieved impressive performance in many visual learning tasks. In this tutorial, we will introduce in detail the latest subspace clustering techniques. For both sparsity based and low-rankness based approaches, we will present some representative subspace clustering models, analyze their theoretical properties, such as exact recovery and closed-form solutions, and present applications in various real problems in computer vision. We will also discuss convex and nonconvex algorithms and scalable methods that can effectively address the problem of sparse and low-rank recovery for visual data.

## Fitting Surface Models to Data: Accuracy, Speed, Robustness

**Organizers:** Andrew Fitzgibbon  
Jonathan Taylor

**Time:** 0830-1800 (Full Day)

**Description:** In vision and machine learning, almost everything we do may be considered to be a form of model fitting. Whether estimating the parameters of a convolutional neural network, computing structure and motion from image collections, tracking objects in video, computing low-dimensional representations of datasets, estimating parameters for an inference model such as Markov random fields, or extracting shape spaces such as active appearance models, it almost always boils down to minimizing an objective containing some parameters of interest as well as some latent or nuisance parameters. This tutorial will describe several tools and techniques for solving such optimization problems, with a focus on fitting 3D smooth-surface models, such as subdivision surfaces, to 2D and 3D data.

## Advances in Geometry and Reflectance Acquisition Vision

**Organizer:** Michael Weinmann  
Fabian Langguth  
Michael Goesele  
Reinhard Klein

**Time:** 0830-1230 (Half Day — Morning)

**Description:** This tutorial is focused on acquisition methods for geometry and reflectance as well as strategies towards an efficient acquisition pipeline to fulfill the demands of industry with respect to mass digitization of 3D contents. We provide a thorough overview of the standard methods for the acquisition of both geometry and reflectance of surfaces with different types of reflectance behavior ranging from diffuse over opaque to specular surfaces or even translucent and transparent surfaces as well as the necessary preliminaries of material appearance and setup calibration. As standard acquisition techniques are only well-suited for a limited range of surface materials, we will also discuss strategies on how an efficient, fully automatic acquisition can still be achieved when no prior information with respect to the surface reflectance behavior is available. In addition, a discussion of strategies regarding an acquisition in the wild (i.e., under uncontrolled conditions) is provided.

## Image Tag Assignment, Refinement and Retrieval

**Organizer:** Xirong Li  
Tiberio Uricchio  
Lamberto Ballan  
Marco Bertini  
Ceess Snoek  
Alberto Del Bimbo

**Time:** 0830-1230 (Half Day — Morning)

**Description:** This tutorial focuses on challenges and solutions for content-based image annotation and retrieval in the context of online image sharing and tagging. We present a unified review on three closely linked problems, i.e., tag assignment, tag refinement, and tag-based image retrieval. We introduce a taxonomy to structure the growing literature, understand the ingredients of the main works, clarify their connections and difference, and recognize their merits and limitations. Moreover, we present an open-source testbed, with training sets of varying sizes and three test datasets, to evaluate methods of varied learning complexity. A selected set of eleven representative works have been implemented and evaluated. During the tutorial we provide a practice session for hands on experience with the methods, software and datasets. For repeatable experiments all data and code are online at <http://www.micc.unifi.it/tagsurvey>.

## Understanding the In-Camera Image Processing Pipeline for Computer Vision

**Organizer:** Michael S. Brown

**Time:** 0830-1230 (Half Day — Morning)

**Description:** Computer vision algorithms generally assume a simple camera imaging model, where output pixel values are considered to correspond to physical scene values. The reality, however, is that there is a wide range of processing steps (e.g. Bayer pattern demosaicing, flat-field correction, white-balance, de-noising, color space transforms, color manipulation, tone manipulation, etc) applied on-board a camera to obtain the final RGB output. These collective processing steps make up the "in-camera image processing pipeline". This tutorial provides a comprehensive overview of the in-camera processing pipeline and describes how it is used to produce camera images. Recent work on modeling camera pipelines for physics-based computer vision research will also be discussed.

## Feature Learning for Image Data: From Dictionary Learning to Deep Learning

**Organizer:** Reza Borhani

Jeremy Watt

Aggelos K. Katsaggelos

**Time:** 0830-1230 (Half Day — Morning)

**Description:** Extracting proper features from image data is one of the holy grails of computer vision. Dictionary and deep learning schemes are among our recent successful attempts at incorporating our understanding of the human brain and the way it processes visual information in order to teach a computer how to 'see' images. In this tutorial, which is based on our new machine learning textbook (see [mlrefined.com](http://mlrefined.com)), we go beyond this biological perspective and study these two frameworks side by side at the mathematical modeling and optimization level. Specifically, we aim to provide a user-friendly introduction to the basic tools of dictionary and deep learning, describe their many applications, discuss how they relate to each other and more traditional ideas in machine learning, and provide an introduction to most useful techniques from numerical optimization crucial to their implementation. To make full use of this tutorial one only needs a basic understanding of linear algebra and vector calculus. No prior knowledge of numerical optimization or machine learning is expected.

## First-Person Visual Sensing: Theory, Models, and Application

**Organizer:** Hyun Soo Park

Kris Kitani

Jianbo Shi

**Time:** 0830-1230 (Half Day — Morning)

**Description:** Still photos are often taken intentionally, capturing specific moments in life, often reflecting individual and relational preferences. In contrast, a first person camera is always on, continually capturing unscripted and often candid moments in life. It allows re-experiencing other's visual sensation and perception via "putting yourself in his/her shoes"; it tells us not only what the person sees but also his/her future intention and momentary visual sensorimotor behaviors. Such unique in-situ characteristics of first person videos are leveraged to achieve numerous tasks in computer vision, graphics, and robotics, e.g., first person activity recognition, life-log summarization, social footage editing, and human-robotic collaboration planning.

In this tutorial, we will address two fundamental questions to learn human perceptual behaviors from first person videos: (1) "What does a first person video tell us about surroundings?" We will address this question by characterizing first person visual signals in terms of anthropometric constraints, geometric relationship with people, objects, and scenes, and its associated visual semantics. For instance, we can reach our hands up to the fixed arm length, which constrain the pose of an interacting object, and therefore, the object visually looks similar if interacted. We will study how such first person characteristics are beneficial to learn computational representations. (2) "What does a first person video tell us about the wearer?" First person visual signals provide feedback control signals for our motor behaviors, and vice versa. For example, when biking, we control the bike handle to balance our posture by incorporating our visual sensation. We will explore computational models for this tight relationship between first person visual signals and motor behaviors.

## 3D Deep Learning with Marvin

**Organizer:** Jianxiong Xiao

Shuran Song

Daniel Suo

Fisher Yu

**Time:** 0830-1230 (Half Day — Morning)

**Description:** Deep learning has made unprecedented progress for one-dimensional data (e.g. text) and two-dimensional data (e.g. image). But it is much less explored for three dimensions, at which our physical world naturally is. Therefore, it is critical for the computer vision, robotics, and graphics communities to study deep learning with 3D input and 3D output. This tutorial will focus on teaching the basic concepts for three-dimensional deep learning. We will start by teaching the basic concepts illustrated using our 3D deep learning software framework <http://marvin.is>. Then, we will give concrete examples about how to use 3D deep learning technique for object recognition, including amodal object detection, pose estimation, and view planning for active recognition, shape completion. Going beyond objects, we will further provide case studies to use 3D deep learning to learn local geometry descriptors, as well as global holistic scene understanding with 3D context. We will also mention about the datasets that are available to train these 3D deep models.

## Clinical Imaging of Human Body: For Health, Visualization and Predictive Analytics

**Organizer:** Ramesh Raskar  
Pratik Shah  
Mrinal Mohit  
Tristan Swedish

**Time:** 0830-1230 (Half Day — Morning)

**Description:** There is an increasing interest in research at the intersection of computer vision, machine learning and health. A multidisciplinary approach is necessary to identify the key challenges to expand computer vision applications into clinical settings and healthcare. This tutorial will cover the major topics and challenges in using image acquisition to model the human body and combine them with vision and learning techniques to perform predictive analytics for clinical conditions. Attendees will receive an overview of the anatomy, physiology and biochemistry that create the variability in appearance of the eye, ear, skin, mouth, and hair among individuals. Topics include state of the art methods used to create more accurate models by incorporating tools used in biomedical imaging. Finally, we will explore health applications of image acquisition and computer vision, along with machine learning to analyze large clinical datasets. The goal of this amalgamated session is to benefit a wider cross-section of audiences who are looking to expand applications into clinical settings and healthcare.

## Compact Features for Visual Search

**Organizer:** Rongrong Ji  
Wei Liu  
Yue Gao

**Time:** 0830-1230 (Half Day — Morning)

**Description:** Handheld mobile devices, such as smart camera phones, have great potential for emerging mobile visual search and augmented reality applications, such as location recognition, scene retrieval, product search, or CD/book cover search. In real-world mobile visual search applications, the reference database typically has millions of images and can only be stored on the remote server(s). Therefore, online querying involves transferring the query image from the mobile device to the remote server. This query delivery is often over a relatively slow wireless link. The quality of the user experience heavily depends on how much information has to be transferred. This issue

becomes even more crucial when we move towards streaming augmented reality applications. This time consuming query delivery is often unnecessary, since the server only performs similarity search rather than query image reconstruction. With the ever growing computational power on mobile devices, recent works have proposed to directly extract compact visual descriptor(s) from the query image on the mobile end, and then send this descriptor over the wireless link with a low bit rate. This descriptor is expected to be compact, discriminative, and meanwhile efficient in extraction to reduce overall query delivery latency. In particular, the research, development, and standardization of compact descriptors for visual search are involving big industry efforts from Nokia, Qualcomm, Aptina, NEC, etc. In large-scale retrieval, it is also well-known that hashing is among the most commonly used scheme to perform approximate nearest neighbor search. Indeed, such methods have been significantly improved and adaptive for the search application. In this tutorial, we will review and introduce some very recent trends and discuss the importance and significance of handling the feature compactness in hashing and visual attributes.

## Recent Image Search Techniques

**Organizer:** Sung-Eui Yoon  
Zhe Lin

**Time:** 1400-1800 (Half Day — Afternoon)

**Description:** Image search has been gaining more and more importance, since many people use image search as well as text search on mobile and desktop platforms. On top of that, image search has been considered as one of the fundamental tools for various computer vision and image processing applications such as object recognition/localization, scene completion, duplicate or near duplicate detection, intelligent personal photo management, image tagging, keyword-based image retrieval, artwork and asset recommendation, data mining, image grouping, product recognition, image editing and stylization, etc.

There have been dramatic advances on image search recently, starting from adopting the bag-of-visual-word model. Some of recent advances include various hashing techniques for designing compact image representation, indexing structures for large-scale image search, etc. We will cover those recent developments.

## Mathematics of Deep Learning

**Organizer:** René Vidal  
Guillermo Sapiro  
Joan Bruna  
Benjamin Haeffele  
Raja Giryes

**Time:** 1400-1800 (Half Day — Afternoon)

**Description:** The past five years have seen a dramatic increase in the performance of recognition systems due to the introduction of deep architectures for feature learning and classification. However, the mathematical reasons for this success remain elusive. This tutorial will review recent work that aims to provide a mathematical justification for properties of special classes of deep networks, such as global optimality, invariance, & stability of the learned representations.

## Optimization Algorithms for Subset Selection and Summarization in Large Data Sets

**Organizer:** Ehsan Elhamifar  
Jeff Bilmes  
Alex Kulesza  
Michael Gyli

**Time:** 1400-1800 (Half Day — Afternoon)

**Description:** The increasing amounts of data in computer vision and other science and engineering fields requires robust tools and techniques to extract important relevant information from massive data collections. Subset selection addresses this challenge by systematically casting the problem as optimizations searching for a subset of high quality diverse items from a large and possibly exponential set of items. However, this in general leads to combinatorial NP-hard problems, which are extremely difficult to solve. This tutorial will present techniques and recent advances in submodular optimization, convex and nonconvex optimization, sparse modeling and determinantal point processes to effectively address the problem of information selection and summarization. The presentation of the formulations, algorithms and theoretical foundations will be complemented with applications in computer vision including video summarization, multiple pose estimation, learning nonlinear dynamic models, human activity segmentation, active learning and feature selection for visual data and more.

## Multimodal Machine Learning

**Organizer:** Louis-Philippe Morency  
Tadas Baltrušaitis

**Time:** 1400-1800 (Half Day — Afternoon)

**Description:** Multimodal machine learning is a vibrant multi-disciplinary research field which addresses some of the original goals of artificial intelligence by integrating and modeling multiple communicative modalities, including linguistic, acoustic and visual messages. With the initial research on audio-visual speech recognition and more recently with image and video captioning projects, this research field brings some unique challenges for multimodal researchers given the heterogeneity of the data and the contingency often found between modalities. This course will explore the fundamental concepts related to multimodal machine learning including modality alignment, heterogeneous representation learning, multimodal fusion and multi-stream sequential modeling. We will also review recent state-of-the-art probabilistic models and computational algorithms for multimodal machine learning and discuss the current and upcoming challenges.

## Computer Vision and Applied Deep Learning with Mathematica

**Organizer:** Sebastian Bodenstein  
Matthias Odisio

**Time:** 1400-1800 (Half Day — Afternoon)

**Description:** This half-day tutorial is a hands-on introduction to using Mathematica and the Wolfram Language for developing and deploying applications in computer vision and machine learning. Mathematica is a fully integrated system that spans all areas of technical computing. We will give a comprehensive overview of Mathematica's state of the art capabilities in computer vision and machine learning, including its completely new and integrated deep learning framework. You will be practicing hands-on with the Wolfram Language through various application examples such as image retrieval, inpainting, feature extraction, or deep learning-based face detection.

## Computer Vision for Automotive/DAS Market: Challenges and Embedded Vision Solutions

**Organizer:** Mihir Mody  
Kedar Chitnis  
Pramod Swami  
Prashant Vishwanath

**Time:** 1400-1800 (Half Day — Afternoon)

**Description:** Computer vision is key in automotive market especially various ADAS applications e.g. front camera, rear camera, surround view, parking assistance. This tutorial discusses various constraints on algorithms quality, computation complexity, cost, power dissipation and functional safety in automotive market. The tutorial explains typical embedded hardware solution with TI's Driver Assist (TDA) product line as an example. The tutorial goes in details of mix of typical programmable processors (e.g. CPU and DSP) as well as Domain specific hardware accelerators for computer vision (e.g. vector engines and image pipe) to address these constraints. The tutorial covers intelligent system partitioning as well as software infrastructure (open compute frameworks and vision development frameworks) to enable simplified software programmer view. The last part of tutorial teaches optimal mapping of various algorithms (e.g. Forward collision warning, pedestrian detection, lane departure warning etc.) in front camera systems on TDA product line. The tutorial will be immensely help students, researchers as well as industry folks to understand ADAS market and embedded solutions.

## Statistical Methods for Open Set Recognition

**Organizer:** Walter J. Scheirer  
Terrance Boulton

**Time:** 1400-1800 (Half Day — Afternoon)

**Description:** The purpose of this tutorial is to introduce the difficult problem of open set recognition in statistical learning specifically within the context of important vision applications. A number of different topics will be introduced, including: recent formalizations of the open set recognition problem, the statistical extreme value theory for visual recognition, which facilitates generalization in probabilistic decision models, and new supervised learning algorithms that minimize the risk of the unknown. Original case studies will be covered for applications related to the analysis of faces, objects and scenes.

## Power Efficient Recognition Systems for Embedded Applications

**Organizer:** Michelle Mao  
Samer Hijazi  
Chris Rowen

**Time:** 1400-1800 (Half Day — Afternoon)

**Description:** Deep learning is becoming most widely used techniques for computer vision and pattern recognition. This is primarily driven by it is outstanding recognition rate. However, its high computational requirements drives the power requirements to levels that are not reasonable for most embedded systems. In this tutorial, we will address this problem from two aspects. First, from problem definition and algorithm optimization point of view. Second form target deployment platform selection and optimization perspective.

In the first part, we will touch on the importance of defining the problem to address only the need we are targeting the embedded device for. Then we will discuss how to choose an algorithm and optimize it according to problem definition dimensions. We will focus on complexity reduction vs performance, i.e., how to strike a balance between performance and complexity for deep learning algorithms. In the second part of this tutorial, we will discuss how to further reduce the power consumption on the target embedded platform. By examining the characteristics of the embedded system and understanding the challenge of power optimization, we discuss how to select and design the right embedded vision target. Finally, specific optimization with vision DSP will be addressed.

## Semantic 3D Modeling

**Organizer:** Christian Häne  
Lubor Ladický  
Marc Pollefeys

**Time:** 1400-1800 (Half Day — Afternoon)

**Description:** In semantic 3D modeling the goal is to find a dense geometric model from images and at the same time also infer the semantic classes of the individual parts of the reconstructed model. Having a semantically annotated dense 3D model gives a much richer representation of the scene than just the geometry. By solving the problem of dense 3D reconstruction and class segmentation jointly, prior knowledge of surface orientations can be included. In this tutorial, we use a volumetric representation of the scene. Traditionally, each voxel gets label into either being in the free space or in the occupied space. For

our semantic formulation we extend the representation to a multi-label formulation. A voxel either belongs to the free space or to one out of multiple semantic classes. While such a formulation is quite memory intensive it allows for a very rich description of the scene.

First we will introduce the underlying continuously inspired multi-label formulation. This formulation is then used for our joint formulation of dense 3D reconstruction and class segmentation. Next, we will present how 3D object shape priors can be modeled by looking at surface normal statistics. In the case of moving objects physical constraints such as solid bodies should not intersect are included. The input that we consider, is given in image space in terms of depth maps and semantic class likelihoods. This naturally induces a data term which is formulated as a potential over a viewing ray. We will present methods which facilitate the usage of a ray potential data term within our semantic 3D modeling framework.

## Diversity Meets Deep Networks — Inference, Ensemble Learning, and Applications

**Organizer:** Dhruv Batra  
Alexander Kirillov  
Stefan Lee  
Carsten Rother  
Bogdan Savchynsky

**Time:** 1400-1800 (Half Day — Afternoon)

**Description:** Perception systems must operate under significant levels of ambiguity. For instance, out of context, a patch from an image may seem like a face, but may simply be an incidental arrangement of tree branches and shadows, causing a face detection system to produce nonsensical results, such as hallucinating faces floating on tree branches and building walls. Similarly, when a human reads the sentence “I eat sushi with tuna”, it is clear that the preposition phrase “with tuna” modifies “sushi” and not the act of eating, but this may be ambiguous to a machine. In addition to implicit task ambiguity, uncertainty can arise from incomplete modelling, inaccurate sensors, and difficulties with optimizations. One paradigm that has recently emerged across a range of AI subfields in response to this ambiguity is the notion of producing a set of diverse plausible hypotheses. These multiple hypotheses may be extracted by finding a set of diverse M-Best solutions from a single model or by training an ensemble of models under a diversity inducing

loss. This tutorial will present an overview of the diverse predictions paradigm, review classical techniques, and present new methods for making diverse predictions and training diverse ensembles across a range of applications in Computer Vision and at the intersection of Vision and Language.

## 3D Keypoint Detection and Feature Description

**Organizer:** Mohammed Bannamoun  
Yulan Guo  
Federico Tombari  
Ferdous Sohel

**Time:** 1400-1800 (Half Day — Afternoon)

**Description:** With the rapid development of point-cloud acquisition techniques (e.g., Microsoft Kinect) and computing devices, 3D data (pointcloud, depth image, mesh) processing has become a rapidly growing research area in computer vision, pattern recognition and computer graphics. Extensive investigations have been conducted on 3D related topics, such as 3D modeling, 3D scene reconstruction, 3D shape retrieval, 3D object recognition, 3D face recognition, and 3D animation. It is considered that feature extraction forms the core of many 3D computer vision systems.

Existing 3D feature extraction methods can broadly be divided into two categories, i.e., global feature and local feature based methods. Intensive research has been done on local feature based methods due to their robustness to occlusion, clutter and missing points which are frequently present in real-world scenarios. A typical 3D local feature extraction algorithm consists of two main phases: 3D keypoint detection and local feature description. A large number of 3D keypoint detection and local feature description have been proposed in the literature. It is however unclear which keypoint detectors and descriptors are more appropriate for a particular application. This tutorial will therefore present a comprehensive review and analysis of the state-of-the-art 3D keypoint detection and local feature description algorithms. The tutorial will also provide extensive performance evaluation results of the state-of-the-art algorithms on several benchmark datasets, along with insightful discussions and analyses. Moreover, a number of interesting 3D related applications will be introduced in the tutorial. Finally, several directions for future work will also be presented.

**END OF SUNDAY CVPR 2016 TUTORIALS**

## Sunday, June 26

**NOTE:** Workshop locations are in the online proceedings and the mobile app

**0730–1700 Registration** (Octavius Prefunction)

**0730–0830 Breakfast** (Forum Ballroom)

**1230–1345 Lunch** (Forum Ballroom)

## Visual Question Answering Challenge

**Organizers:** Aishwarya Agrawal

Jiasen Lu

Yash Goyal

Peng Zhang

Larry Zitnick

Dhruv Batra

Devi Parikh

**Schedule:** Full Day

0900 **Welcome**

0910 **Invited Talk:** *Yuandong Tian (Facebook AI Research)*

0935 **Invited Talk:** *Margaret Mitchell (Microsoft Research)*

1000 **Invited Talk:** *Alex Berg (UNC Chapel Hill)*

1025 **Morning Break** (Forum Ballroom)

1045 **Invited Talk:** *Jitendra Malik (UC Berkeley)*

1110 **Overview of Challenge, Winner Announcements, Analysis of Results**

1140 **Challenge Winner Talk**

1155 **Challenge Winner Talk**

1210 **Challenge Winner Talk**

1225 **Lunch** (Forum Ballroom)

1400 **Invited Talk:** *Kevin Murphy (Google Research)*

1425 **Invited Talk:** *Trevor Darrell (UC Berkeley)*

1450 **Afternoon Break** (Forum Ballroom) & **Poster Session**

1600 **Invited Talk:** *Ali Farhadi (Univ. of Washington)*

1625 **Invited Talk:** *Mario Fritz (Max-Planck-Institut für Informatik)*

1650 **Panel:** Future Directions

1800 **Closing Remarks**

## Computer Vision in Vehicle Technology

**Organizers:** José M. Álvarez

David Vázquez Bermúdez

Tomas Pajdla

Antonio M. López

**Schedule:** Full Day

0800 **Welcome**

0815 **Invited Talk:** *Jianxiong Xiao (Princeton)*

0915 **Contributed Talks**

1000 **Posters**

- UAV-Based Autonomous Image Acquisition With Multi-View Stereo Quality Assurance by Confidence Prediction, *Christian Mostegel, Markus Rumpfer, Friedrich Fraundorfer, Horst Bischof*
- Mobile Device Based Outdoor Navigation With On-Line Learning Neural Network: A Comparison With Convolutional Neural Network, *Zejia Zheng, Juyang Weng*
- The HCI Benchmark Suite: Stereo and Flow Ground Truth With Uncertainties for Urban Autonomous Driving, *Daniel Kondermann, Rahul Nair, Katrin Honauer, Karsten Krispin, Jonas Andriulis, Alexander Brock, Burkhard Giese, Mohsen Rahimimoghaddam, Sabine Hofmann, Claus Brenner, Bernd Jähne*
- Monocular Long-Term Target Following on UAVs, *Rui Li, Minjian Pang, Cong Zhao, Guyue Zhou, Lu Fang*
- DeepLanes: End-To-End Lane Position Estimation Using Deep Neural Networks, *Alexandru Gurghian, Tejaswi Koduri, Smriti V. Baidur, Kyle J. Carey, Vidya N. Murali*
- Multiple Scale Faster-RCNN Approach to Driver's Cell-Phone Usage and Hands on Steering Wheel Detection, *T. Hoang Ngan Le, Yutong Zheng, Chenchen Zhu, Khoa Luu, Marios Savvides*
- DR(Eye)Ve: A Dataset for Attention-Based Tasks With Applications to Autonomous and Assisted Driving, *Stefano Alletto, Andrea Palazzi, Francesco Solera, Simone Calderara, Rita Cucchiara*
- A Low-Cost Mirror-Based Active Perception System for Effective Collision Free Underwater Robotic Navigation, *Noel Cortés Pérez, Luz Abril Torres Méndez*

1000 **Morning Break** (Forum Ballroom)

1030 **Invited Talk:** *Tomas Pajdla (CTU)*

1130 **Contributed Talks**

1235 **Lunch** (Forum Ballroom)

1435 **Invited Talk:** *Kai Yu (Baidu)*

1500 **Invited Talk:** *Marc Pollefeys (ETH)*

- 1525 **Afternoon Break** (Forum Ballroom)
- 1600 **Invited Talk:** *Urs Muller (NVIDIA)*
- 1700 **Closing Remarks**

## Biometrics

**Organizers:** Bir Bhanu  
Ajay Kumar

**Schedule:** Full Day  
0815 **Welcome**

### S1: Biometrics Security and Protection (0820-0920)

- 0820 What Do You Do When You Know That You Don't Know?  
*Abhijit Bendale, Terrance E. Boul*
- 0840 Deep Secure Encoding for Face Template Protection,  
*Rohit Kumar Pandey, Yingbo Zhou, Bhargava Urala Kota, Venu Govindaraju*
- 0900 Feature Vector Compression Based on Least Error Quantization, *Tomokazu Kawahara, Osamu Yamaguchi*

### S2: Deep Learning in Biometrics I (0820-1000)

- 0920 Weakly Supervised Facial Analysis With Dense Hyper-Column Features, *Chenchen Zhu, Yutong Zheng, Khoa Luu, T. Hoang Ngan Le, Chandrasekhar Bhagavatula, Marios Savvides*
- 0930 A Comprehensive Analysis of Deep Learning Based Representation for Face Recognition, *Mostafa Mehdipour Ghazi, Hazim Kemal Ekenel*
- 0940 Two-Stream CNNs for Gesture-Based Verification and Identification: Learning User Style, *Jonathan Wu, Prakash Ishwar, Janusz Konrad*
- 0950 Deep Tattoo Recognition, *Xing Di, Vishal M. Patel*

### **1000 Morning Break** (Forum Ballroom)

### S3: Deep Learning in Biometrics II (1030-1130)

- 1030 Pooling Faces: Template Based Face Recognition With Pooled Face Images, *Tal Hassner, Iacopo Masi, Jungyeon Kim, Jongmoo Choi, Shai Harel, Prem Natarajan, Gérard Medioni*
- 1050 DeepGender: Occlusion and Low Resolution Robust Facial Gender Classification via Progressively Trained Convolutional Neural Networks With Attention, *Felix Juefei-Xu, Eshan Verma, Parag Goel, Anisha Cherodan, Marios Savvides*
- 1110 Real-Time Face Identification via CNN and Boosted Hashing Forest, *Yuri Vizilter, Vladimir Gorbatshevich, Andrey Vorotnikov, Nikita Kostromov*

### S4: Face Recognition I (1130-1230)

- 1130 Gaussian Conditional Random Fields for Face Recognition, *Jonathon M. Smerka, B.V.K. Vijaya Kumar, Andres Rodriguez*
- 1140 Grouper: Optimizing Crowdsourced Face Annotations, *Jocelyn C. Adams, Kristen C. Allen, Timothy Miller, Nathan D. Kalka, Anil K. Jain*
- 1150 PARAPH: Presentation Attack Rejection by Analyzing Polarization Hypotheses, *Ethan M. Rudd, Manuel Günther, Terrance E. Boul*
- 1200 Heterogeneous Face Recognition Using Inter-Session Variability Modelling, *Tiago de Freitas Pereira, Sébastien Marcel*
- 1210 A Polarimetric Thermal Database for Face Recognition Research, *Shuowen Hu, Nathaniel J. Short, Benjamin S. Riggan, Christopher Gordon, Kristan P. Gurton, Matthew Thielke, Prudhvi Gurram, Alex L. Chan*
- 1220 CALIPER: Continuous Authentication Layered With Integrated PKI Encoding Recognition, *Ethan M. Rudd, Terrance E. Boul*

### S5: Fingerprint Identification (1230-1300)

- 1230 Frequency Map by Structure Tensor in Logarithmic Scale Space and Forensic Fingerprints, *Josef Bigun, Anna Mikaelyan*
- 1240 Latent Fingerprint Image Segmentation Using Fractal Dimension Features and Weighted Extreme Learning Machine Ensemble, *Jude Ezeobiesi, Bir Bhanu*
- 1250 GMM-SVM Fingerprint Verification Based on Minutiae Only, *Berkay Topcu, Yusuf Ziya Isik, Hakan Erdogan*

### **1300 Lunch** (Forum Ballroom)

### S5: Face Recognition II (1400-1500)

- 1400 A Comparison of Human and Automated Face Verification Accuracy on Unconstrained Image Sets, *Austin Blanton, Kristen C. Allen, Timothy Miller, Nathan D. Kalka, Anil K. Jain*
- 1420 Soft-Margin Learning for Multiple Feature-Kernel Combinations With Domain Adaptation, for Recognition in Surveillance Face Dataset, *Samik Banerjee, Sukhendu Das*
- 1440 Simultaneous Semi-Coupled Dictionary Learning for Matching RGBD Data, *Nilotpal Das, Devraj Mandal, Soma Biswas*

## S7: Secured Signature Verification (1500-1520)

- 1500 Offline Signature Verification Based on Bag-Of-Visual Words Model Using KAZE Features and Weighting Schemes, *Manabu Okawa*
- 1510 Implementation of Fixed-Length Template Protection Based on Homomorphic Encryption With Application to Signature Biometrics, *Marta Gomez-Barrero, Julian Fierrez, Javier Galbally, Emanuele Maiorana, Patrizio Campisi*

1520 **Afternoon Break** (Forum Ballroom)

## S8: Panel Session (1600-1720)

- 1600 **Panel:** Challenging Questions for Biometrics in its Adulthood, *Chair: Josef Bigun*
- 1720 **Awards, Valedictory, and Closing Remarks**, *Bir Bhanu*

## Perceptual Organization in Computer Vision: The Role of Feedback in Recognition and Reorganization

**Organizers:** Katerina Fragkiadaki  
Phillip Isola  
João Carreira

**Schedule:** Full Day

- 0915 **Welcome**
- 0930 **Invited Talk:** *Piotr Dollar (Facebook)*
- 1000 **Invited Talk:** *Bruno Olshausen (UC Berkeley)*
- 1030 **Morning Break** (Forum Ballroom) & **Posters**
- 1100 **Invited Talk:** *Deva Ramanan (CMU)*
- 1130 **Invited Talk:** *Aude Oliva (MIT)*
- 1200 **Posters**
- 1235 **Lunch** (Forum Ballroom) & **Posters**
- 1330 **Invited Talk:** *Thomas Brox (Univ. of Freiburg)*
- 1400 **Poster Orals**
- 1430 **Invited Talk:** *Viren Jain (Google)*
- 1500 **Invited Talk:** *Daniel Kersten (Univ. of Minnesota)*
- 1530 **Afternoon Break** (Forum Ballroom)
- 1600 **Invited Talk:** *Jianbo Shi (Univ. of Pennsylvania)*
- 1630 **Invited Talk:** *Jitendra Malik (UC Berkeley)*
- 1700 **Discussion**

## Vision Industry and Entrepreneur Workshop

**Organizers:** Arnab Dhua  
Himanshu Arora  
Ziyan Wu  
Samson Timoner

**Schedule:** Half Day — Morning

0815 **Poster and Demo Setup**

### S1: Presentations (0830-1015)

- 0830 **Welcome**
- 0845 **Venture Pitch Contest:** *Moderator — Samson Timoner*
- 1000 **Industry Session Spotlights:** *Moderator — Ziyan Wu*
- 1015 **Morning Break** (Forum Ballroom)

### S2: Poster Session (1030-1145)

- 1030 **Venture Pitch Feedback:** *Moderator — Samson Timoner*
- 1030 **Industry Session: Demos, Posters, and Recruiting:** *Moderator — Ziyan Wu*
- Indian Traffic Sign Detection and Classification Using Neural Networks, *Arun Nandewal, Satyam Chandra*
  - Mapping the World Using Drones, *Tianwei Shen, Tian Fang, Long Quan*
  - Collaborative Computer Vision R&D at Kitware, *Anthony Hoogs, Matt Turek, Keith Fieldhouse*
  - Efficient Video Similarity Using Convolutional Neural Networks, *Pegah Massoudifar, Ana Murillo, Hardik Shah*
  - Visual Search technologies at Ag, *Himanshu Arora, Arnab Dhua, Sunil Ramesh*
  - SRI's Center of Vision Technology, *Sek Chai*
  - Vision Technologies and Solutions at Siemens, *Jan Ernst, Ziyan Wu*
  - Efficient Annotation Platform for Vision Applications, *Yu Wang, Raymon Ng*
  - Reimagining Retail: Fellow Robots, *Sivapriya Kaza, Thavidu Ranatunga*
  - Comcast Labs: Enhancing the Entertainment Experiences of the Future, *Hongcheng Wang*

### S3: Awards and Closing Remarks (1145-1230)

- 1145 **Venture Pitch Results:** *Moderator — Samson Timoner*
- 1220 **Closing Remarks**

## Scene Understanding

**Organizers:** James Hays  
Aditya Khosla  
Silvio Savarese  
Jianxiong Xiao

**Schedule:** Half Day — Morning

0755 **Welcome**

0800 **Invited Talk:** Alex Berg (UNC, Chapel Hill)

0830 **Invited Talk:** Trevor Darrell (UC Berkeley)

0900 **Invited Talk:** Kristen Grauman (Univ. of Texas, Austin)

0930 **Invited Talk:** Honglak Lee (Univ. of Michigan)

**1000 Morning Break** (Forum Ballroom) & **Poster session**

1100 **Invited Talk:** Irving Biederman (Univ. of Southern California)

1130 **Invited Talk:** Svetlana Lazebnik (Univ. of Illinois, Urbana Champaign)

1200 **Invited Talk:** Derek Hoiem (Univ. of Illinois, Urbana Champaign)

## Computational Models for Learning Systems and Educational Assessment

**Organizers:** Saad M. Khan  
Yuchi Huang  
Alina von Davier  
Sidney D'Mello  
Demitris Metaxas

**Schedule:** Half Day — Morning

0815 **Welcome & Opening Remarks**

0830 **Invited Talk:** Ajay Divakaran (SRI International)

0900 **Invited Talk:** Robert Mislevy, ETS)

0930 **Short Talks & Discussion**

- Improving Teaching and Learning Through Video Summaries of Student Engagement, Jonathan Ventura, Steve Cruz, Terrance Boulton
- Development of an Audiovisual Database of Human-Machine Conversations for Educational Learning and Assessment Applications, Vikram Ramanarayanan, David Suendermann-Oeft, Patrick Lange, Robert Mundkowsky, Alexei V. Ivanov, Zhou Yu, Yao Qian, Keelan Evanini

**1000 Morning Break** (Forum Ballroom)

**1015 Short Talks & Discussion**

- Understanding & Predicting Engagement in Lecture Videos, Rachit Dubey, Zachary Pardos
- Assessing Adaptive Learning Gain and Its Confluence With Task Difficulty, Habiba Khan, Gahangir Hossain

**1045 Invited Talk:** Louis-Philippe Morency (CMU)

**1115 Invited Talk:** Jonathan Huang (Google)

**1145 Open Discussion:** Attendees and CMLA Organizing Committee

**1200 Closing Remarks**

## Women in Computer Vision

**Organizers:** Samaneh Azadi  
Lisa Anne Hendricks  
Brittany Morago  
Brigit Schroeder  
Lingling Tao  
Serena Yeung

**Schedule:** Half Day — Afternoon

**1330 Introduction**

**1340 Invited Talk:** Interactive Image and Video Segmentation, Kristen Grauman (Univ. of Texas at Austin)

**1400 Oral Session 1**

- Exploring Joint Language-Visual Models for Visual Content Search in Movies, Atousa Torabi, Leonid Sigal
- Ambiguity Helps: Classification With Disagreements in Crowdsourced Annotations, Viktoriia Sharmanska, Daniel Hernandez-Lobato, Jose Miguel Hernandez-Lobato, Novi Quadrianto

**1430 Invited Talk:** Computer Vision Research at Google, Caroline Pantofaru (Google)

**1450 Afternoon Break** (Forum Ballroom) & **Poster Session**

**1620 Invited Talk:** TBA, Raquel Urtasun (Univ. of Toronto)

**1640 Oral Session 2**

- Real-Time 3D Reconstruction via Implicit-To-Implicit Registration, Miroslava Slavcheva, Wadim Kehl, Nassir Navab, Slobodan Ilic
- Clockwork Convnets for Semantic Segmentation, Kate Rakelly, Evan Shelhamer, Judy Hoffman, Trevor Darrell

**1710 Closing Remarks**

## Perception Beyond the Visible Spectrum

**Organizers:** Riad I. Hammoud

Guoliang Fan

Erik Blasch

Firooz Sadjadi

Yi Ding

**Schedule:** Half Day — Afternoon

- 1330 **Keynote:** GPU-Accelerated Deep Learning Toolbox, *Jon Barker (NVIDIA)*
- 1400 Learning Cross-Spectral Similarity Measures With Deep Convolutional Neural Networks, *Cristhian A. Aguilera, Francisco J. Aguilera, Angel D. Sappa, Cristhian Aguilera, Ricardo Toledo*
- 1415 Distinguishing Weather Phenomena From Bird Migration Patterns in Radar Imagery, *Aruni RoyChowdhury, Daniel Sheldon, Subhansu Maji, Erik Learned-Miller*
- 1430 A Modular NMF Matching Algorithm for Radiation Spectra, *Melissa Koudelka, Daniel J. Dorsey*
- 1445 Evaluation of Feature Channels for Correlation-Filter-Based Visual Object Tracking in Infrared Spectrum, *Erhan Gundogdu, Aykut Koc, Berkan Solmaz, Riad I. Hammoud, A. Aydin Alatan*
- 1500 **Keynote:** A Decade of Progress in Computer Vision, *Behzad Kamgar-Parsi (Office of Naval Research)*
- 1545 Adaptive Object Classification Using Complex SAR Signatures, *Firooz Sadjadi*
- 1600 Scale Invariant Human Action Detection From Depth Cameras Using Class Templates, *Kartik Gupta, Arnav Bhavsar*
- 1615 Real-Time Physiological Measurement and Visualization Using a Synchronized Multi-Camera System, *Otkrist Gupta, Dan McDuff, Ramesh Raskar*
- 1630 Seeing the Forest From the Trees: A Holistic Approach to Near-Infrared Heterogeneous Face Recognition, *Christopher Reale, Nasser M. Nasrabadi, Heesung Kwon, Rama Chellappa*
- 1645 Non-Planar Infrared-Visible Registration for Uncalibrated Stereo Pairs, *Dinh-Luan Nguyen, Pierre-Luc St-Charles, Guillaume-Alexandre Bilodeau*
- 1700 A Novel Visualization Tool for Evaluating the Accuracy of 3D Sensing and Reconstruction Algorithms for Automatic Dormant Pruning Applications, *Fangda Li, Somrita Chattopadhyay, Shayan A. Akbar, Noha M. Elfiky, Avinash Kak*

- 1715 A Novel Benchmark RGBD Dataset for Dormant Apple Trees and Its Application to Automatic Pruning, *Shayan A. Akbar, Somrita Chattopadhyay, Noha M. Elfiky, Avinash Kak*

1730 **Concluding Remarks**

## Large-Scale Scene Understanding Challenge

**Organizers:** Jianxiong Xiao

Yinda Zhang

Fisher Yu

Shuran Song

Pingmei Xu

Ari Seff

**Schedule:** Half Day — Afternoon

- 1330 **Welcome**
- 1335 **Keynote:** *Jitendra Malik (Univ. of California at Berkeley)*
- 1405 **Keynote:** *Yann LeCun (Facebook & New York Univ.)*
- 1435 **Award Session**
- 1445 **Afternoon Break** (Forum Ballroom)
- 1505 **Introduction to LSUN Dataset & Classification Task**
- 1520 **Classification Winner Talk**
- 1535 **Introduction to Saliency Prediction Task**
- 1540 **Saliency Winner Talk**
- 1555 **Afternoon Break** (Forum Ballroom)
- 1615 **Introduction to Room Layout Task**
- 1620 **Room Layout Winner Talk**
- 1635 **Introduction to RGBD 3D Object Detection Task**
- 1640 **Detection Winner Talk**
- 1655 **Closing Remarks**

## END OF SUNDAY CVPR 2016 WORKSHOPS

# Monday, June 27

**0730–1700 Registration** (Octavius Prefunction)

**0730–0830 Breakfast** (Forum Ballroom)

**0845–0900 Welcome by the General & Program Chairs** (Augustus Ballroom)

**0900–1030 Oral & Spotlight Session 1-1A**  
(Augustus III–VI)

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

**Chairs:** Jongwoo Lim (*Hanyang Univ.*)  
Subhransu Maji (*UMass Amherst*)

**0900 O1-1A: Image Captioning and Question Answering**  
Format (12 min. for presentation + 1 min. for questions)

1. Deep Compositional Captioning: Describing Novel Object Categories Without Paired Training Data, *Lisa Anne Hendricks, Subhashini Venugopalan, Marcus Rohrbach, Raymond Mooney, Kate Saenko, Trevor Darrell*
2. Generation and Comprehension of Unambiguous Object Descriptions, *Junhua Mao, Jonathan Huang, Alexander Toshev, Oana Camburu, Alan L. Yuille, Kevin Murphy*
3. Stacked Attention Networks for Image Question Answering, *Zichao Yang, Xiaodong He, Jianfeng Gao, Li Deng, Alex Smola*
4. Image Question Answering Using Convolutional Neural Network With Dynamic Parameter Prediction, *Hyeonwoo Noh, Paul Hongsuck Seo, Bohyung Han*
5. Neural Module Networks, *Jacob Andreas, Marcus Rohrbach, Trevor Darrell, Dan Klein*

**1005 S1-1A: Language and Vision**

Format (4 min. for presentation; no questions)

6. Learning Deep Representations of Fine-Grained Visual Descriptions, *Scott Reed, Zeynep Akata, Honglak Lee, Bernt Schiele*
7. Multi-Cue Zero-Shot Learning With Strong Supervision, *Zeynep Akata, Mateusz Malinowski, Mario Fritz, Bernt Schiele*
8. Latent Embeddings for Zero-Shot Classification, *Yongqin Xian, Zeynep Akata, Gaurav Sharma, Quynh Nguyen, Matthias Hein, Bernt Schiele*

9. One-Shot Learning of Scene Locations via Feature Trajectory Transfer, *Roland Kwitt, Sebastian Hegenbart, Marc Niethammer*
10. Learning Attributes Equals Multi-Source Domain Generalization, *Chuang Gan, Tianbao Yang, Boqing Gong*
11. Anticipating Visual Representations From Unlabeled Video, *Carl Vondrick, Hamed Pirsiavash, Antonio Torralba*

**0900–1030 Oral & Spotlight Session 1-1B**  
(Augustus I–II; Emperors Ballroom)

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

**Chairs:** Jianxiong Xiao (*Princeton Univ.*)  
Tal Hassner (*USC*)

**0900 O1-1B: Matching and Alignment**

Format (12 min. for presentation + 1 min. for questions)

12. Learning to Assign Orientations to Feature Points, *Kwang Moo Yi, Yannick Verdie, Pascal Fua, Vincent Lepetit*
13. Learning Dense Correspondence via 3D-Guided Cycle Consistency, *Tinghui Zhou, Philipp Krähenbühl, Mathieu Aubry, Qixing Huang, Alexei A. Efros*
14. The Global Patch Collider, *Shenlong Wang, Sean Ryan Fanello, Christoph Riemann, Shahram Izadi, Pushmeet Kohli*
15. Joint Probabilistic Matching Using  $m$ -Best Solutions, *Seyed Hamid Reza Tofighi, Anton Milan, Zhen Zhang, Qinfeng Shi, Anthony Dick, Ian Reid*
16. Face Alignment Across Large Poses: A 3D Solution, *Xiangyu Zhu, Zhen Lei, Xiaoming Liu, Hailin Shi, Stan Z. Li*

**1005 S1-1B: Segmentation and Contour Detection**

Format (4 min. for presentation; no questions)

17. Interactive Segmentation on RGBD Images via Cue Selection, *Jie Feng, Brian Price, Scott Cohen, Shih-Fu Chang*
18. Layered Scene Decomposition via the Occlusion-CRF, *Chen Liu, Pushmeet Kohli, Yasutaka Furukawa*
19. Affinity CNN: Learning Pixel-Centric Pairwise Relations for Figure/Ground Embedding, *Michael Maire, Takuya Narihira, Stella X. Yu*
20. Weakly Supervised Object Boundaries, *Anna Khoreva, Rodrigo Benenson, Mohamed Omran, Matthias Hein, Bernt Schiele*
21. Object Contour Detection With a Fully Convolutional Encoder-Decoder Network, *Jimei Yang, Brian Price, Scott Cohen, Honglak Lee, Ming-Hsuan Yang*

**1030–1100 Break** (Octavius Ballroom)

**1030–1230 Exhibits (Octavius Ballroom)**

|              |                                                                                                                                                                                                                                                                                                                                                                                                                       |         |                                                                                                                                                                                                                                                                                                                                                                                             |
|--------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Platinum     | <ul style="list-style-type: none"> <li>• Ag</li> <li>• Adobe Systems</li> <li>• Amazon</li> <li>• Apple</li> <li>• Athena Eyes (a-eye)</li> <li>• DiDi</li> <li>• DJI</li> <li>• Facebook</li> <li>• Google</li> <li>• Intel</li> <li>• Megvii/Face++</li> <li>• MERL</li> <li>• Microsoft Research</li> <li>• NVIDIA</li> <li>• Prism</li> <li>• Toyota Research Inst.</li> <li>• TuSimple</li> </ul>                | Bronze  | <ul style="list-style-type: none"> <li>• AiCure</li> <li>• Avigilon</li> <li>• Blippar</li> <li>• CEVA</li> <li>• Datatang</li> <li>• IntelliVision</li> <li>• Kitware</li> <li>• Magic Leap</li> <li>• Movidius</li> <li>• Omron</li> <li>• Pinterest</li> <li>• Point Grey</li> <li>• Rethink Robotics</li> <li>• Second Spectrum</li> <li>• uSens</li> <li>• Wolfram Research</li> </ul> |
|              | <ul style="list-style-type: none"> <li>• Baidu</li> <li>• Bosch</li> <li>• Cognex</li> <li>• DAQRI</li> <li>• Ebay</li> <li>• HERE</li> <li>• IBM Research</li> <li>• iRobot</li> <li>• Linkface</li> <li>• Qihoo</li> <li>• SenseTime</li> <li>• Sighthound</li> <li>• Skydio</li> <li>• Spare5</li> <li>• Twitter</li> <li>• Uber ATC</li> <li>• DeepGlint</li> <li>• Disney Research</li> <li>• HiScene</li> </ul> |         | <ul style="list-style-type: none"> <li>• Dextro</li> <li>• FiveFocal</li> <li>• Markable</li> <li>• Nervana</li> <li>• SportLogiq</li> <li>• Spotscale</li> <li>• Umbo CV</li> </ul>                                                                                                                                                                                                        |
|              | <ul style="list-style-type: none"> <li>• MathWorks</li> <li>• Panasonic</li> <li>• Snapchat</li> <li>• SRI International</li> <li>• Tencent</li> <li>• UISEE</li> </ul>                                                                                                                                                                                                                                               |         | <ul style="list-style-type: none"> <li>• wrnch</li> <li>• 3dMD</li> <li>• CSIRO DATA61</li> <li>• Elsevier</li> <li>• KAUST</li> <li>• Matterport</li> <li>• Morgan &amp; Claypool</li> <li>• NREC</li> <li>• Nuctech</li> <li>• Occam Vision Group</li> <li>• Samasource</li> <li>• Sony Electronic</li> <li>• Springer</li> <li>• Van Gogh Imaging</li> </ul>                             |
| Gold         |                                                                                                                                                                                                                                                                                                                                                                                                                       | Startup | <ul style="list-style-type: none"> <li>• Cizr</li> <li>• GemHunt.com</li> <li>• iniLabs Ltd.</li> <li>• Morpx Inc.</li> <li>• QR Hub Ltd.</li> </ul>                                                                                                                                                                                                                                        |
|              |                                                                                                                                                                                                                                                                                                                                                                                                                       |         |                                                                                                                                                                                                                                                                                                                                                                                             |
| Silver       |                                                                                                                                                                                                                                                                                                                                                                                                                       | Basic   |                                                                                                                                                                                                                                                                                                                                                                                             |
|              |                                                                                                                                                                                                                                                                                                                                                                                                                       |         |                                                                                                                                                                                                                                                                                                                                                                                             |
| Startup Hive |                                                                                                                                                                                                                                                                                                                                                                                                                       |         |                                                                                                                                                                                                                                                                                                                                                                                             |
|              |                                                                                                                                                                                                                                                                                                                                                                                                                       |         |                                                                                                                                                                                                                                                                                                                                                                                             |

**1030–1230 Demos (Octavius Ballroom)**

- Modulated Video Camera – Light Separation for Dynamic Scenes, Amir Kolaman, Rami Hagege, Hugo Guterman (Ben-Gurion Univ. of the Negev)

- FANNG: Fast Approximate Nearest Neighbour Graphs, Ben Harwood and Tom Drummond (Monash Univ.)
- DeepDriving: Learning Affordance for Direct Perception in Autonomous Driving, Chenyi Chen, Ari Seff, Alain L. Kornhauser, Jianxiong Xiao (Princeton Univ.)
- VL-SLAM: Real-Time Visual-Inertial Navigation and Semantic Mapping, Xiaohan Fei, Yanchao Yang, Jingming Dong, Nikolaos Karianakis, Konstantine Tsotsos, Virginia Estellers, Stefano Soatto (UCLA)

**1030–1230 Poster Session P1-1 (Octavius Ballroom)****Images and Language**

22. What Value Do Explicit High Level Concepts Have in Vision to Language Problems? Qi Wu, Chunhua Shen, Lingqiao Liu, Anthony Dick, Anton van den Hengel

**Edge Contour Detection**

23. Fast Detection of Curved Edges at Low SNR, Nati Ofir, Meirav Galun, Boaz Nadler, Ronen Basri
24. Object Skeleton Extraction in Natural Images by Fusing Scale-Associated Deep Side Outputs, Wei Shen, Kai Zhao, Yuan Jiang, Yan Wang, Zhijiang Zhang, Xiang Bai
25. Learning Relaxed Deep Supervision for Better Edge Detection, Yu Liu, Michael S. Lew
26. Occlusion Boundary Detection via Deep Exploration of Context, Huan Fu, Chaohui Wang, Dacheng Tao, Michael J. Black
27. SemiContour: A Semi-Supervised Learning Approach for Contour Detection, Zizhao Zhang, Fuyong Xing, Xiaoshuang Shi, Lin Yang

**Feature Extraction and Description**

28. Learning to Localize Little Landmarks, Saurabh Singh, Derek Hoiem, David Forsyth
29. InterActive: Inter-Layer Activeness Propagation, Lingxi Xie, Liang Zheng, Jingdong Wang, Alan L. Yuille, Qi Tian
30. Exploit Bounding Box Annotations for Multi-Label Object Recognition, Hao Yang, Joey Tianyi Zhou, Yu Zhang, Bin-Bin Gao, Jianxin Wu, Jianfei Cai
31. TI-POOLING: Transformation-Invariant Pooling for Feature Learning in Convolutional Neural Networks, Dmitry Laptev, Nikolay Savinov, Joachim M. Buhmann, Marc Pollefeys
32. Fashion Style in 128 Floats: Joint Ranking and Classification Using Weak Data for Feature Extraction, Edgar Simo-Serra, Hiroshi Ishikawa

- 33. Equiangular Kernel Dictionary Learning With Applications to Dynamic Texture Analysis, *Yuhui Quan, Chenglong Bao, Hui Ji*
- 34. Compact Bilinear Pooling, *Yang Gao, Oscar Beijbom, Ning Zhang, Trevor Darrell*

#### Feature Extraction and Matching

- 35. Accumulated Stability Voting: A Robust Descriptor From Descriptors of Multiple Scales, *Tsun-Yi Yang, Yen-Yu Lin, Yung-Yu Chuang*
- 36. CoMaL: Good Features to Match on Object Boundaries, *Swarna K. Ravindran, Anurag Mittal*
- 37. Progressive Feature Matching With Alternate Descriptor Selection and Correspondence Enrichment, *Yuan-Ting Hu, Yen-Yu Lin*

#### Image Segmentation

- 38. A New Finsler Minimal Path Model With Curvature Penalization for Image Segmentation and Closed Contour Detection, *Da Chen, Jean-Marie Mirebeau, Laurent D. Cohen*
- 39. Scale-Aware Alignment of Hierarchical Image Segmentation, *Yuhua Chen, Dengxin Dai, Jordi Pont-Tuset, Luc Van Gool*
- 40. Deep Interactive Object Selection, *Ning Xu, Brian Price, Scott Cohen, Jimei Yang, Thomas S. Huang*
- 41. Pull the Plug? Predicting If Computers or Humans Should Segment Images, *Danna Gurari, Suyog Jain, Margrit Betke, Kristen Grauman*
- 42. In the Shadows, Shape Priors Shine: Using Occlusion to Improve Multi-Region Segmentation, *Yuka Kihara, Matvey Soloviev, Tsuhan Chen*
- 43. Convexity Shape Constraints for Image Segmentation, *Loic A. Royer, David L. Richmond, Carsten Rother, Bjoern Andres, Dagmar Kainmueller*
- 44. MCMC Shape Sampling for Image Segmentation With Nonparametric Shape Priors, *Ertunc Erdil, Sinan Yildirim, Müjdat Cetin, Tolga Tasdizen*

#### Low-Level Vision

- 45. From Noise Modeling to Blind Image Denoising, *Fengyuan Zhu, Guangyong Chen, Pheng-Ann Heng*
- 46. Efficient and Robust Color Consistency for Community Photo Collections, *Jaesik Park, Yu-Wing Tai, Sudipta N. Sinha, In So Kweon*
- 47. Needle-Match: Reliable Patch Matching Under High Uncertainty, *Or Lotan, Michal Irani*
- 48. ReconNet: Non-Iterative Reconstruction of Images From Compressively Sensed Measurements, *Kuldeep Kulkarni, Suhas Lohit, Pavan Turaga, Ronan Kerviche, Amit Ashok*

- 49. Soft-Segmentation Guided Object Motion Deblurring, *Jinshan Pan, Zhe Hu, Zhixun Su, Hsin-Ying Lee, Ming-Hsuan Yang*
- 50. Two Illuminant Estimation and User Correction Preference, *Dongliang Cheng, Abdelrahman Abdelhamed, Brian Price, Scott Cohen, Michael S. Brown*
- 51. Deep Contrast Learning for Salient Object Detection, *Guanbin Li, Yizhou Yu*
- 52. Multiview Image Completion With Space Structure Propagation, *Seung-Hwan Baek, Inchang Choi, Min H. Kim*
- 53. Composition-Preserving Deep Photo Aesthetics Assessment, *Long Mai, Hailin Jin, Feng Liu*
- 54. Automatic Image Cropping: A Computational Complexity Study, *Jiansheng Chen, Gaocheng Bai, Shaoheng Liang, Zhengqin Li*
- 55. A Deeper Look at Saliency: Feature Contrast, Semantics, and Beyond, *Neil D. B. Bruce, Christopher Catton, Sasa Janjic*
- 56. Spatially Binned ROC: A Comprehensive Saliency Metric, *Calden Wloka, John Tsotsos*
- 57. GraB: Visual Saliency via Novel Graph Model and Background Priors, *Qiaosong Wang, Wen Zheng, Robinson Piramuthu*
- 58. Predicting When Saliency Maps Are Accurate and Eye Fixations Consistent, *Anna Volokitin, Michael Gygli, Xavier Boix*
- 59. Split and Match: Example-Based Adaptive Patch Sampling for Unsupervised Style Transfer, *Oriel Frigo, Neus Sabater, Julie Delon, Pierre Hellier*
- 60. Detection and Accurate Localization of Circular Fiducials Under Highly Challenging Conditions, *Lilian Calvet, Pierre Gurdjos, Carsten Griwodz, Simone Gasparini*

#### Scene Understanding

- 61. Scene Recognition With CNNs: Objects, Scales and Dataset Bias, *Luis Herranz, Shuqiang Jiang, Xiangyang Li*
- 62. Learning Action Maps of Large Environments via First-Person Vision, *Nicholas Rhinehart, Kris M. Kitani*
- 63. Single-Image Crowd Counting via Multi-Column Convolutional Neural Network, *Yingying Zhang, Desen Zhou, Siqin Chen, Shenghua Gao, Yi Ma*
- 64. Shallow and Deep Convolutional Networks for Saliency Prediction, *Junting Pan, Elisa Sayrol, Xavier Giro-i-Nieto, Kevin McGuinness, Noel E. O'Connor*
- 65. Sample and Filter: Nonparametric Scene Parsing via Efficient Filtering, *Mohammad Najafi, Sarah Taghavi Namin, Mathieu Salzmann, Lars Petersson*
- 66. DeLay: Robust Spatial Layout Estimation for Cluttered Indoor Scenes, *Saumitro Dasgupta, Kuan Fang, Kevin Chen, Silvio Savarese*

67. A Text Detection System for Natural Scenes With Convolutional Feature Learning and Cascaded Classification, *Siyu Zhu, Richard Zanibbi*

### Segmentation and Saliency

68. Reversible Recursive Instance-Level Object Segmentation, *Xiaodan Liang, Yunchao Wei, Xiaohui Shen, Zequn Jie, Jiashi Feng, Liang Lin, Shuicheng Yan*
69. Coherent Parametric Contours for Interactive Video Object Segmentation, *Yao Lu, Xue Bai, Linda Shapiro, Jue Wang*
70. Manifold SLIC: A Fast Method to Compute Content-Sensitive Superpixels, *Yong-Jin Liu, Cheng-Chi Yu, Min-Jing Yu, Ying He*
71. Deep Saliency With Encoded Low Level Distance Map and High Level Features, *Gayoung Lee, Yu-Wing Tai, Junmo Kim*
72. Instance-Level Segmentation for Autonomous Driving With Deep Densely Connected MRFs, *Ziyu Zhang, Sanja Fidler, Raquel Urtasun*
73. DHSNet: Deep Hierarchical Saliency Network for Salient Object Detection, *Nian Liu, Junwei Han*
74. Object Co-Segmentation via Graph Optimized-Flexible Manifold Ranking, *Rong Qian, Junwei Han, Dingwen Zhang, Feipeng Nie*

## Video Segmentation

75. Primary Object Segmentation in Videos via Alternate Convex Optimization of Foreground and Background Distributions, *Won-Dong Jang, Chulwoo Lee, Chang-Su Kim*
76. Automatic Fence Segmentation in Videos of Dynamic Scenes, *Renjiao Yi, Jue Wang, Ping Tan*
77. Discovering the Physical Parts of an Articulated Object Class From Multiple Videos, *Luca Del Pero, Susanna Ricco, Rahul Sukthankar, Vittorio Ferrari*
78. A Benchmark Dataset and Evaluation Methodology for Video Object Segmentation, *Federico Perazzi, Jordi Pont-Tuset, Brian McWilliams, Luc Van Gool, Markus Gross, Alexander Sorkine-Hornung*
79. Learning Temporal Regularity in Video Sequences, *Mahmudul Hasan, Jonghyun Choi, Jan Neumann, Amit K. Roy-Chowdhury, Larry S. Davis*
80. Bilateral Space Video Segmentation, *Nicolas Maerki, Federico Perazzi, Oliver Wang, Alexander Sorkine-Hornung*
81. ReD-SFA: Relation Discovery Based Slow Feature Analysis for Trajectory Clustering, *Zhang Zhang, Kaiqi Huang, Tieniu Tan, Peipei Yang, Jun Li*  
ReD-SFA: Relation Discovery Based Slow Feature Analysis for Trajectory Clustering, *Zhang Zhang, Kaiqi Huang, Tieniu Tan, Peipei Yang, Jun Li*

**1230-1345 Lunch** (Forum Ballroom)

**Notes:**

[illegible]

**1345–1520 Oral & Spotlight Session 1-2A**

(Augustus III–VI)

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

**Chairs:** Olga Russakovsky (*Carnegie Mellon Univ.*)  
Katerina Fragkiadaki (*Google Research*)

**1345 O1-2A: Object Recognition and Detection**Format (12 min. for presentation + 1 min. for questions)

1. Training Region-Based Object Detectors With Online Hard Example Mining, *Abhinav Shrivastava, Abhinav Gupta, Ross Girshick*
2. Deep Residual Learning for Image Recognition, *Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun*
3. You Only Look Once: Unified, Real-Time Object Detection, *Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi*
4. LocNet: Improving Localization Accuracy for Object Detection, *Spyros Gidaris, Nikos Komodakis*
5. Sketch Me That Shoe, *Qian Yu, Feng Liu, Yi-Zhe Song, Tao Xiang, Timothy M. Hospedales, Chen-Change Loy*

**1450 S1-2A: Object Detection 1**Format (4 min. for presentation; no questions)

6. Deep Sliding Shapes for Amodal 3D Object Detection in RGB-D Images, *Shuran Song, Jianxiong Xiao*
7. Object Detection From Video Tubelets With Convolutional Neural Networks, *Kai Kang, Wanli Ouyang, Hongsheng Li, Xiaogang Wang*
8. Learning With Side Information Through Modality Hallucination, *Judy Hoffman, Saurabh Gupta, Trevor Darrell*
9. Object-Proposal Evaluation Protocol is 'Gameable', *Neelima Chavali, Harsh Agrawal, Aroma Mahendru, Dhruv Batra*
10. HyperNet: Towards Accurate Region Proposal Generation and Joint Object Detection, *Tao Kong, Anbang Yao, Yurong Chen, Fuchun Sun*
11. We Don't Need No Bounding-Boxes: Training Object Class Detectors Using Only Human Verification, *Dim P. Papadopoulos, Jasper R. R. Uijlings, Frank Keller, Vittorio Ferrari*
12. Factors in Finetuning Deep Model for Object Detection With Long-Tail Distribution, *Wanli Ouyang, Xiaogang Wang, Cong Zhang, Xiaokang Yang*

**1345–1520 Oral & Spotlight Session 1-2B**

(Augustus I–II; Emperors Ballroom)

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

**Chairs:** Karteek Alahari (*INRIA*)  
Neill Campbell (*Univ. of Bath*)

**1345 O1-2B: Vision with Alternative Sensors**Format (12 min. for presentation + 1 min. for questions)

13. Information-Driven Adaptive Structured-Light Scanners, *Guy Rosman, Daniela Rus, John W. Fisher III*
14. Simultaneous Optical Flow and Intensity Estimation From an Event Camera, *Patrick Bardow, Andrew J. Davison, Stefan Leutenegger*
15. Macroscopic Interferometry: Rethinking Depth Estimation With Frequency-Domain Time-Of-Flight, *Achuta Kadambi, Jamie Schiel, Ramesh Raskar*
16. ASP Vision: Optically Computing the First Layer of Convolutional Neural Networks Using Angle Sensitive Pixels, *Huaijin G. Chen, Suren Jayasuriya, Jiyue Yang, Judy Stephen, Sriram Sivaramakrishnan, Ashok Veeraraghavan, Alyosha Molnar*
17. Computational Imaging for VLBi Image Reconstruction, *Katherine L. Bouman, Michael D. Johnson, Daniel Zoran, Vincent L. Fish, Sheperd S. Doleman, William T. Freeman*

**1450 S1-2B: Video Analysis 1**Format (4 min. for presentation; no questions)

18. You Lead, We Exceed: Labor-Free Video Concept Learning by Jointly Exploiting Web Videos and Images, *Chuang Gan, Ting Yao, Kuiyuan Yang, Yi Yang, Tao Mei*
19. Track and Segment: An Iterative Unsupervised Approach for Video Object Proposals, *Fanyi Xiao, Yong Jae Lee*
20. Beyond Local Search: Tracking Objects Everywhere With Instance-Specific Proposals, *Gao Zhu, Fatih Porikli, Hongdong Li*
21. Groupwise Tracking of Crowded Similar-Appearance Targets From Low-Continuity Image Sequences, *Hongkai Yu, Youjie Zhou, Jeff Simmons, Craig P. Przybyla, Yuewei Lin, Xiaochuan Fan, Yang Mi, Song Wang*
22. Social LSTM: Human Trajectory Prediction in Crowded Spaces, *Alexandre Alahi, Kratarth Goel, Vignesh Ramanathan, Alexandre Robicquet, Li Fei-Fei, Silvio Savarese*
23. What Players Do With the Ball: A Physically Constrained Interaction Modeling, *Andrii Maksai, Xinchao Wang, Pascal Fua*
24. Highlight Detection With Pairwise Deep Ranking for First-Person Video Summarization, *Ting Yao, Tao Mei, Yong Rui*

**1520–1545 Break** (Forum Ballroom)**1545–1645 Plenary Session** (Augustus Ballroom)

- **Plenary Talk:** Autonomous Driving, Computer Vision and Machine Learning: Disrupting Transportation Is Just Around the Corner, *Amnon Shashua (The Hebrew Univ. of Jerusalem)*

**1645–1845 Exhibits** (Octavius Ballroom)

- Same as Monday morning Exhibits (see pg. 17)

**1645–1845 Demos** (Octavius Ballroom)

- Live Prediction of Age, Gender and Facial Attractiveness, *Rasmus Rothe, Radu Timofte, Luc van Gool (ETH Zurich)*
- A Real-Time RGB-NIR Imaging System Using a Single Image Sensor, *Kazunori Yoshizaki, Munenori Fukunishi, Yasuhiro Komiya, Yusuke Monno, Masayuki Tanaka, Masatoshi Okutomi, Steven Lansel (Olympus R&D Group, Tokyo Institute of Technology, Olympus Communication Technology of America)*
- The Menpo Project, *Epameinondas Antonakos, James Booth, Patrick Snape, George Trigeorgis (Imperial College London)*
- Real-Time 3D Object Detection and Tracking, *David Joseph Tan, Federico Tombari, Nassir Navab (TU Munich, Univ. of Bologna, Johns Hopkins Univ.)*

**1645–1845 Poster Session P1-2** (Octavius Ballroom)**Events, Activities, and Surveillance**

25. Direct Prediction of 3D Body Poses From Motion Compensated Sequences, *Bugra Tekin, Artem Rozantsev, Vincent Lepetit, Pascal Fua*
26. Video2GIF: Automatic Generation of Animated GIFs From Video, *Michael Gygli, Yale Song, Liangliang Cao*
27. NTU RGB+D: A Large Scale Dataset for 3D Human Activity Analysis, *Amir Shahroudy, Jun Liu, Tian-Tsong Ng, Gang Wang*
28. Progressively Parsing Interactional Objects for Fine Grained Action Detection, *Bingbing Ni, Xiaokang Yang, Shenghua Gao*
29. Hierarchical Recurrent Neural Encoder for Video Representation With Application to Captioning, *Pingbo Pan, Zhongwen Xu, Yi Yang, Fei Wu, Yueting Zhuang*
30. From Keyframes to Key Objects: Video Summarization by Representative Object Proposal Selection, *Jingjing Meng, Hongxing Wang, Junsong Yuan, Yap-Peng Tan*
31. Temporal Action Localization in Untrimmed Videos via Multi-Stage CNNs, *Zheng Shou, Dongang Wang, Shih-Fu Chang*

32. Summary Transfer: Exemplar-Based Subset Selection for Video Summarization, *Ke Zhang, Wei-Lun Chao, Fei Sha, Kristen Grauman*
33. POD: Discovering Primary Objects in Videos Based on Evolutionary Refinement of Object Recurrence, Background, and Primary Object Models, *Yeong Jun Koh, Won-Dong Jang, Chang-Su Kim*
34. What If We Do Not Have Multiple Videos of the Same Action? — Video Action Localization Using Web Images, *Waqas Sultani, Mubarak Shah*
35. Beyond F-Formations: Determining Social Involvement in Free Standing Conversing Groups From Static Images, *Lu Zhang, Hayley Hung*

**Fine-Grained Categorization**

36. DeepFashion: Powering Robust Clothes Recognition and Retrieval With Rich Annotations, *Ziwei Liu, Ping Luo, Shi Qiu, Xiaogang Wang, Xiaoou Tang*
37. SketchNet: Sketch Classification With Web Images, *Hua Zhang, Si Liu, Changqing Zhang, Wenqi Ren, Rui Wang, Xiaochun Cao*
38. Embedding Label Structures for Fine-Grained Feature Representation, *Xiaofan Zhang, Feng Zhou, Yuanqing Lin, Shaoting Zhang*
39. Fine-Grained Image Classification by Exploring Bipartite-Graph Labels, *Feng Zhou, Yuanqing Lin*
40. Picking Deep Filter Responses for Fine-Grained Image Recognition, *Xiaopeng Zhang, Hongkai Xiong, Wengang Zhou, Weiyao Lin, Qi Tian*
41. SPDA-CNN: Unifying Semantic Part Detection and Abstraction for Fine-Grained Recognition, *Han Zhang, Tao Xu, Mohamed Elhoseiny, Xiaolei Huang, Shaoting Zhang, Ahmed Elgammal, Dimitris Metaxas*
42. Fine-Grained Categorization and Dataset Bootstrapping Using Deep Metric Learning With Humans in the Loop, *Yin Cui, Feng Zhou, Yuanqing Lin, Serge Belongie*
43. Mining Discriminative Triplets of Patches for Fine-Grained Classification, *Yaming Wang, Jonghyun Choi, Vlad Morariu, Larry S. Davis*
44. Part-Stacked CNN for Fine-Grained Visual Categorization, *Shaoli Huang, Zhe Xu, Dacheng Tao, Ya Zhang*

**Feature Matching and Indexing**

45. Learning Compact Binary Descriptors With Unsupervised Deep Neural Networks, *Kevin Lin, Jiwen Lu, Chu-Song Chen, Jie Zhou*

46. Solving Small-Piece Jigsaw Puzzles by Growing Consensus, *Kilho Son, daniel Moreno, James Hays, David B. Cooper*
47. Pairwise Matching Through Max-Weight Bipartite Belief Propagation, *Zhen Zhang, Qinfeng Shi, Julian McAuley, Wei Wei, Yanning Zhang, Anton van den Hengel*
48. Structured Feature Similarity With Explicit Feature Map, *Takumi Kobayashi*

49. Temporal Epipolar Regions, *Mor Dar, Yael Moses*

#### **Human ID**

50. Recurrent Attention Models for Depth-Based Person Identification, *Albert Haque, Alexandre Alahi, Li Fei-Fei*
51. Learning a Discriminative Null Space for Person Re-Identification, *Li Zhang, Tao Xiang, Shaogang Gong*
52. Learning Deep Feature Representations With Domain Guided Dropout for Person Re-Identification, *Tong Xiao, Hongsheng Li, Wanli Ouyang, Xiaogang Wang*
53. How Far Are We From Solving Pedestrian Detection? *Shanshan Zhang, Rodrigo Benenson, Mohamed Omran, Jan Hosang, Bernt Schiele*
54. Similarity Learning With Spatial Constraints for Person Re-Identification, *Dapeng Chen, Zejian Yuan, Badong Chen, Nanning Zheng*
55. Sample-Specific SVM Learning for Person Re-Identification, *Ying Zhang, Baohua Li, Huchuan Lu, Atshushi Irie, Xiang Ruan*
56. Joint Learning of Single-Image and Cross-Image Representations for Person Re-Identification, *Faqiang Wang, Wangmeng Zuo, Liang Lin, David Zhang, Lei Zhang*
57. A Multi-Level Contextual Model For Person Recognition in Photo Albums, *Haoxiang Li, Jonathan Brandt, Zhe Lin, Xiaohui Shen, Gang Hua*
58. Unsupervised Cross-Dataset Transfer Learning for Person Re-Identification, *Peixi Peng, Tao Xiang, Yaowei Wang, Massimiliano Pontil, Shaogang Gong, Tiejun Huang, Yonghong Tian*
59. Pedestrian Detection Inspired by Appearance Constancy and Shape Symmetry, *Jiale Cao, Yanwei Pang, Xuelong Li*
60. Recurrent Convolutional Network for Video-Based Person Re-Identification, *Niall McLaughlin, Jesus Martinez del Rincon, Paul Miller*
61. Person Re-Identification by Multi-Channel Parts-Based CNN With Improved Triplet Loss Function, *De Cheng, Yihong Gong, Sanping Zhou, Jinjun Wang, Nanning Zheng*
62. Top-Push Video-Based Person Re-Identification, *Jinjie You, Ancong Wu, Xiang Li, Wei-Shi Zheng*
63. Improving Person Re-Identification via Pose-Aware Multi-Shot Matching, *Yeong-Jun Cho, Kuk-Jin Yoon*

64. Hierarchical Gaussian Descriptor for Person Re-Identification, *Tetsu Matsukawa, Takahiro Okabe, Einoshin Suzuki, Yoichi Sato*

#### **Motion and Tracking**

65. STCT: Sequentially Training Convolutional Networks for Visual Tracking, *Lijun Wang, Wanli Ouyang, Xiaogang Wang, Huchuan Lu*
66. Determining Occlusions From Space and Time Image Reconstructions, *Juan-Manuel Pérez-Rúa, Tomas Crivelli, Patrick Bouthemy, Patrick Pérez*
67. Online Multi-Object Tracking via Structural Constraint Event Aggregation, *Ju Hong Yoon, Chang-Ryeol Lee, Ming-Hsuan Yang, Kuk-Jin Yoon*
68. Staple: Complementary Learners for Real-Time Tracking, *Luca Bertinetto, Jack Valmadre, Stuart Golodetz, Ondrej Miksik, Philip H. S. Torr*
69. Robust Optical Flow Estimation of Double-Layer Images Under Transparency or Reflection, *Jiaolong Yang, Hongdong Li, Yuchao Dai, Robby T. Tan*
70. Siamese Instance Search for Tracking, *Ran Tao, Efstratios Gavves, Arnold W.M. Smeulders*
71. Adaptive Decontamination of the Training Set: A Unified Formulation for Discriminative Visual Tracking, *Martin Danelljan, Gustav Häger, Fahad Shahbaz Khan, Michael Felsberg*
72. 3D Part-Based Sparse Tracker With Automatic Synchronization and Registration, *Adel Bibi, Tianzhu Zhang, Bernard Ghanem*
73. Recurrently Target-Attending Tracking, *Zhen Cui, Shengtao Xiao, Jiashi Feng, Shuicheng Yan*

#### **Supervised Learning**

74. Structured Regression Gradient Boosting, *Ferran Diego, Fred A. Hamprecht*
75. Loss Functions for Top-k Error: Analysis and Insights, *Maksim Lapin, Matthias Hein, Bernt Schiele*
76. Metric Learning as Convex Combinations of Local Models With Generalization Guarantees, *Valentina Zantedeschi, Rémi Emonet, Marc Sebban*
77. Efficient Training of Very Deep Neural Networks for Supervised Hashing, *Ziming Zhang, Yuting Chen, Venkatesh Saligrama*
78. Information Bottleneck Learning Using Privileged Information for Visual Recognition, *Saeid Motiian, Marco Piccirilli, Donald A. Adjeroh, Gianfranco Doretto*

### 1900-1915 Awards Ceremony (Augustus Ballroom)

### 1915-2100 Reception (Garden of the God's Pool Complex)

**Directions:** Between Palace Tower and Octavius Tower.

**2100-2200 PAMI Technical Committee Meeting**  
(Augustus III-VI)

## CVPR 2016 Outstanding Reviewers:

We are pleased to recognize the following researchers as "CVPR 2016 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.

- Zeynep Akata
- Relja Arandjelovic
- Antonis Argyros
- Tamar Avraham
- Steve Branson
- Jongmoo Choi
- Andrew Davison
- Alexander Fix
- Andreas Geiger
- Ioannis Gkioulekas
- Kristen Grauman
- Abhinav Gupta
- C.V. Jawahar
- Dinesh Jayaraman
- Philipp Krahenbuhl
- Josip Krapac
- Jonathan Long
- Philippos Mordohai
- Alison Noble
- Peter Ochs
- Matthew O'Toole
- Emanuele Rodola
- Marcus Rohrbach
- Javier Romero
- Miki Rubinstein
- Olga Russakovsky
- Bogdan Savchynskyy
- Stan Sclaroff
- Joseph Sivic
- Cees Snoek
- Deqing Sun
- Akihiko Torii
- Tinne Tuytelaars
- Anton van den Hengel
- Jan van Gemert
- Andrea Vedaldi
- René Vidal
- David Wipf
- Lior Wolf
- Christian Wolf
- John Wright
- Tao Xiang
- Stefanos Zafeiriou
- Maria Zontak

**Notes:**

[illegible]

## Tuesday, June 28

### 0730–1700 Registration (Octavius Prefunction)

### 0730–0830 Breakfast (Forum Ballroom)

### 0900–1030 Oral & Spotlight Session 2-1A (Augustus III–VI)

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

**Chairs:** Walter Scheirer (*Univ. of Notre Dame*)  
Laura Sevilla-Lara (*MPI*)

#### 0900 O2-1A: Recognition and Parsing In 3D

Format (12 min. for presentation + 1 min. for questions)

1. 3D Action Recognition From Novel Viewpoints, *Hossein Rahmani, Ajmal Mian*
2. 3D Shape Attributes, *David F. Fouhey, Abhinav Gupta, Andrew Zisserman*
3. Three-Dimensional Object Detection and Layout Prediction Using Clouds of Oriented Gradients, *Zhile Ren, Erik B. Sudderth*
4. 3D Semantic Parsing of Large-Scale Indoor Spaces, *Iro Armeni, Ozan Sener, Amir R. Zamir, Helen Jiang, Ioannis Brilakis, Martin Fischer, Silvio Savarese*
5. Dense Human Body Correspondences Using Convolutional Networks, *Lingyu Wei, Qixing Huang, Duygu Ceylan, Etienne Vouga, Hao Li*

#### 1005 S2-1A: Recognition Beyond Objects

Format (4 min. for presentation; no questions)

6. Geometry-Informed Material Recognition, *Joseph DeGol, Mani Golparvar-Fard, Derek Hoiem*
7. Towards Open Set Deep Networks, *Abhijit Bendale, Terrance E. Boult*
8. What's Wrong With That Object? Identifying Images of Unusual Objects by Modelling the Detection Score Distribution, *Peng Wang, Lingqiao Liu, Chunhua Shen, Zi Huang, Anton van den Hengel, Heng Tao Shen*
9. Large-Scale Location Recognition and the Geometric Burstiness Problem, *Torsten Sattler, Michal Havlena, Konrad Schindler, Marc Pollefeys*
10. Regularity-Driven Facade Matching Between Aerial and Street Views, *Mark Wolff, Robert T. Collins, Yanxi Liu*
11. Do Computational Models Differ Systematically From Human Object Perception? *R. T. Pramod, S. P. Arun*

### 0900–1030 Oral & Spotlight Session 2-1B (Augustus I–II; Emperors Ballroom)

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

**Chairs:** Phillip Isola (*UC Berkeley*)  
Matthias Nießner (*Stanford Univ.*)

#### 0900 O2-1B: Image Processing and Restoration

Format (12 min. for presentation + 1 min. for questions)

12. Contour Detection in Unstructured 3D Point Clouds, *Timo Hackel, Jan D. Wegner, Konrad Schindler*
13. Unsupervised Learning of Edges, *Yin Li, Manohar Paluri, James M. Rehg, Piotr Dollár*
14. Blind Image Deblurring Using Dark Channel Prior, *Jinshan Pan, Deqing Sun, Hanspeter Pfister, Ming-Hsuan Yang*
15. Deeply-Recursive Convolutional Network for Image Super-Resolution, *Jiwon Kim, Jung Kwon Lee, Kyoung Mu Lee*
16. Accurate Image Super-Resolution Using Very Deep Convolutional Networks, *Jiwon Kim, Jung Kwon Lee, Kyoung Mu Lee*

#### 1005 S2-1B: Image Processing and Restoration

Format (4 min. for presentation; no questions)

17. RAW Image Reconstruction Using a Self-Contained sRGB-JPEG Image With Only 64 KB Overhead, *Rang M. H. Nguyen, Michael S. Brown*
18. Group MAD Competition - A New Methodology to Compare Objective Image Quality Models, *Kede Ma, Qingbo Wu, Zhou Wang, Zhengfang Duanmu, Hongwei Yong, Hongliang Li, Lei Zhang*
19. Non-Local Image Dehazing, *Dana Berman, Tali Treibitz, Shai Avidan*
20. A Holistic Approach to Cross-Channel Image Noise Modeling and Its Application to Image Denoising, *Seonghyeon Nam, Youngbae Hwang, Yasuyuki Matsushita, Seon Joo Kim*
21. Multispectral Images Denoising by Intrinsic Tensor Sparsity Regularization, *Qi Xie, Qian Zhao, Deyu Meng, Zongben Xu, Shuhang Gu, Wangmeng Zuo, Lei Zhang*
22. A Comparative Study for Single Image Blind Deblurring, *Wei-Sheng Lai, Jia-Bin Huang, Zhe Hu, Narendra Ahuja, Ming-Hsuan Yang*

### 1030–1100 Break (Octavius Ballroom)

### 1030–1230 Exhibits (Octavius Ballroom)

- Same as Monday morning Exhibits (see pg. 17)

## 1030–1230 Demos (Octavius Ballroom)

- Live Demo for Semantic Image Segmentation, *Shuai Zheng, Sadeep Jayasumana, Bernardina Romera-Paredes, Philip H. S. Torr (Univ. of Oxford)*
- Big Data Analytics for Intelligent Farming: i-EKBase, *Ritaban Dutta (CSIRO)*
- Photo-Realistic Virtual Worlds to Boost Video Understanding and Autonomous Driving: Introducing SYNTHIA & Virtual KITTI, *Adrien Gaidon, German Ros, Antonio M. Lopez, Eleonora Vig, David Vázquez, Qiao Wang (Xerox Research Center Europe, Computer Vision Center)*
- Monocular Real-Time Surface Reconstruction using Dynamic Level of Detail, *Jacek Zienkiewicz, Stefan Leutenegger (Imperial College London)*

## 1030–1230 Poster Session P2-1 (Octavius Ballroom)

### 3D Vision

23. Spatiotemporal Bundle Adjustment for Dynamic 3D Reconstruction, *Minh Vo, Srinivasa G. Narasimhan, Yaser Sheikh*
24. Inextensible Non-Rigid Shape-From-Motion by Second-Order Cone Programming, *Ajad Chhatkuli, Daniel Pizarro, Toby Collins, Adrian Bartoli*
25. Optimal Relative Pose With Unknown Correspondences, *Johan Fredriksson, Viktor Larsson, Carl Olsson, Fredrik Kahl*
26. Homography Estimation From the Common Self-Polar Triangle of Separate Ellipses, *Haifei Huang, Hui Zhang, Yiuming Cheung*
27. Heterogeneous Light Fields, *Maximilian Diebold, Bernd Jähne, Alexander Gatto*
28. A Consensus-Based Framework for Distributed Bundle Adjustment, *Anders Eriksson, John Bastian, Tat-Jun Chin, Mats Isaksson*
29. Globally Optimal Manhattan Frame Estimation in Real-Time, *Kyungdon Joo, Tae-Hyun Oh, Junsik Kim, In So Kweon*
30. Mirror Surface Reconstruction Under an Uncalibrated Camera, *Kai Han, Kwan-Yee K. Wong, Dirk Schnieders, Miaomiao Liu*
31. A Hole Filling Approach Based on Background Reconstruction for View Synthesis in 3D Video, *Guibo Luo, Yuesheng Zhu, Zhaotian Li, Liming Zhang*
32. A Direct Least-Squares Solution to the PnP Problem With Unknown Focal Length, *Yinqiang Zheng, Laurent Kneip*

33. Efficient Intersection of Three Quadrics and Applications in Computer Vision, *Zuzana Kukelova, Jan Heller, Andrew Fitzgibbon*
34. Using Spatial Order to Boost the Elimination of Incorrect Feature Matches, *Lior Talker, Yael Moses, Ilan Shimshoni*
35. A Probabilistic Framework for Color-Based Point Set Registration, *Martin Danelljan, Giulia Meneghetti, Fahad Shahbaz Khan, Michael Felsberg*

### Deblurring and Super-Resolution

36. Blind Image Deconvolution by Automatic Gradient Activation, *Dong Gong, Mingkui Tan, Yanning Zhang, Anton van den Hengel, Qinfeng Shi*
37. PSyCo: Manifold Span Reduction for Super Resolution, *Eduardo Pérez-Pellitero, Jordi Salvador, Javier Ruiz-Hidalgo, Bodo Rosenhahn*
38. Parametric Object Motion From Blur, *Jochen Gast, Anita Sellent, Stefan Roth*
39. Image Deblurring Using Smartphone Inertial Sensors, *Zhe Hu, Lu Yuan, Stephen Lin, Ming-Hsuan Yang*
40. Seven Ways to Improve Example-Based Single Image Super Resolution, *Radu Timofte, Rasmus Rothe, Luc Van Gool*
41. Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network, *Wenzhe Shi, Jose Caballero, Ferenc Huszar, Johannes Totz, Andrew P. Aitken, Rob Bishop, Daniel Rueckert, Zehan Wang*

### Events, Actions, and Activity Recognition

42. They Are Not Equally Reliable: Semantic Event Search Using Differentiated Concept Classifiers, *Xiaojun Chang, Yao-Liang Yu, Yi Yang, Eric P. Xing*
43. Going Deeper into First-Person Activity Recognition, *Minghuang Ma, Haoqi Fan, Kris M. Kitani*
44. Cascaded Interactional Targeting Network for Egocentric Video Analysis, *Yang Zhou, Bingbing Ni, Richang Hong, Xiaokang Yang, Qi Tian*
45. Fast Temporal Activity Proposals for Efficient Detection of Human Actions in Untrimmed Videos, *Fabian Caba Heilbron, Juan Carlos Nieves, Bernard Ghanem*
46. Discriminative Hierarchical Rank Pooling for Activity Recognition, *Basura Fernando, Peter Anderson, Marcus Hutter, Stephen Gould*
47. Convolutional Two-Stream Network Fusion for Video Action Recognition, *Christoph Feichtenhofer, Axel Pinz, Andrew Zisserman*

48. Learning Activity Progression in LSTMs for Activity Detection and Early Detection, *Shugao Ma, Leonid Sigal, Stan Sclaroff*
49. VLAD3: Encoding Dynamics of Deep Features for Action Recognition, *Yingwei Li, Weixin Li, Vijay Mahadevan, Nuno Vasconcelos*
50. A Multi-Stream Bi-Directional Recurrent Neural Network for Fine-Grained Action Detection, *Bharat Singh, Tim K. Marks, Michael Jones, Oncel Tuzel, Ming Shao*
51. A Hierarchical Deep Temporal Model for Group Activity Recognition, *Mostafa S. Ibrahim, Srikanth Muralidharan, Zhiwei Deng, Arash Vahdat, Greg Mori*
52. A Hierarchical Pose-Based Approach to Complex Action Understanding Using Dictionaries of Actionlets and Motion Poselets, *Ivan Lillo, Juan Carlos Niebles, Alvaro Soto*
53. A Key Volume Mining Deep Framework for Action Recognition, *Wangjiang Zhu, Jie Hu, Gang Sun, Xudong Cao, Yu Qiao*

## Image Indexing and Retrieval

54. Improved Hamming Distance Search Using Variable Length Substrings, *Eng-Jon Ong, Miroslaw Bobier*
55. Shortlist Selection With Residual-Aware Distance Estimator for K-Nearest Neighbor Search, *Jae-Pil Heo, Zhe Lin, Xiaohui Shen, Jonathan Brandt, Sung-eui Yoon*
56. Supervised Quantization for Similarity Search, *Xiaojuan Wang, Ting Zhang, Guo-Jun Qi, Jinhui Tang, Jingdong Wang*
57. Efficient Large-Scale Approximate Nearest Neighbor Search on the GPU, *Patrick Wieschollek, Oliver Wang, Alexander Sorkine-Hornung, Hendrik P. A. Lensch*
58. Collaborative Quantization for Cross-Modal Similarity Search, *Ting Zhang, Jingdong Wang*
59. Aggregating Image and Text Quantized Correlated Components, *Thi Quynh Nhi Tran, Hervé Le Borgne, Michel Crucianu*
60. Efficient Indexing of Billion-Scale Datasets of Deep Descriptors, *Artem Babenko, Victor Lempitsky*
61. Deep Supervised Hashing for Fast Image Retrieval, *Haomiao Liu, Ruiping Wang, Shiguang Shan, Xilin Chen*
62. Efficient Large-Scale Similarity Search Using Matrix Factorization, *Ahmet Iscen, Michael Rabbat, Teddy Furon*
63. Incremental Object Discovery in Time-Varying Image Collections, *Theodora Kontogianni, Markus Mathias, Bastian Leibe*

## Motion and Tracking

64. Detecting Migrating Birds at Night, *Jia-Bin Huang, Rich Caruana, Andrew Farnsworth, Steve Kelling, Narendra Ahuja*

## Object Class Detection and Recognition

65. When Naïve Bayes Nearest Neighbors Meet Convolutional Neural Networks, *Ilya Kuzborskiy, Fabio Maria Carlucci, Barbara Caputo*
66. Traffic-Sign Detection and Classification in the Wild, *Zhe Zhu, Dun Liang, Songhai Zhang, Xiaolei Huang, Baoli Li, Shimin Hu*
67. Large Scale Semi-Supervised Object Detection Using Visual and Semantic Knowledge Transfer, *Yuxing Tang, Josiah Wang, Boyang Gao, Emmanuel Dellandréa, Robert Gaizauskas, Liming Chen*
68. Exploit All the Layers: Fast and Accurate CNN Object Detector With Scale Dependent Pooling and Cascaded Rejection Classifiers, *Fan Yang, Wongun Choi, Yuanqing Lin*
69. Dictionary Pair Classifier Driven Convolutional Neural Networks for Object Detection, *Keze Wang, Liang Lin, Wangmeng Zuo, Shuhang Gu, Lei Zhang*
70. Monocular 3D Object Detection for Autonomous Driving, *Xiaozhi Chen, Kaustav Kundu, Ziyu Zhang, Huimin Ma, Sanja Fidler, Raquel Urtasun*
71. How Hard Can It Be? Estimating the Difficulty of Visual Search in an Image, *Radu Tudor Ionescu, Bogdan Alexe, Marius Leordeanu, Marius Popescu, Dim P. Papadopoulos, Vittorio Ferrari*
72. Deep Relative Distance Learning: Tell the Difference Between Similar Vehicles, *Hongye Liu, Yonghong Tian, Yaowei Yang, Lu Pang, Tiejun Huang*

## Recognition and Detection

73. Eye Tracking for Everyone, *Kyle Krafka, Aditya Khosla, Petr Kellnhofer, Suchendra Bhandarkar, Wojciech Matusik, Antonio Torralba*
74. Efficient Globally Optimal 2D-To-3D Deformable Shape Matching, *Zorah Löhner, Emanuele Rodolà, Frank R. Schmidt, Michael M. Bronstein, Daniel Cremers*
75. Ambiguity Helps: Classification With Disagreements in Crowdsourced Annotations, *Viktoria Sharmanska, Daniel Hernández-Lobato, José Miguel Hernández-Lobato, Novi Quadrianto*
76. A Task-Oriented Approach for Cost-Sensitive Recognition, *Roozbeh Mottaghi, Hannaneh Hajishirzi, Ali Farhadi*
77. Refining Architectures of Deep Convolutional Neural Networks, *Sukrit Shankar, Duncan Robertson, Yoni Ioannou, Antonio Criminisi, Roberto Cipolla*
78. iLab-20M: A Large-Scale Controlled Object Dataset to Investigate Deep Learning, *Ali Borji, Saeed Izadi, Laurent Itti*
79. Recursive Recurrent Nets With Attention Modeling for OCR in the Wild, *Chen-Yu Lee, Simon Osindero*



## 1345-1520 Oral & Spotlight Session 2-2A (Augustus III-VI)

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

**Chairs:** Oisín Mac Aodha (*Univ. College London*)  
Ross Girshick (*Facebook AI Research*)

### 1345 O2-2A: Recognition and Labeling

Format (12 min. for presentation + 1 min. for questions)

1. Hierarchically Gated Deep Networks for Semantic Segmentation, *Guo-Jun Qi*
2. Deep Structured Scene Parsing by Learning With Image Descriptions, *Liang Lin, Guangrun Wang, Rui Zhang, Ruimao Zhang, Xiaodan Liang, Wangmeng Zuo*
3. CNN-RNN: A Unified Framework for Multi-Label Image Classification, *Jiang Wang, Yi Yang, Junhua Mao, Zhiheng Huang, Chang Huang, Wei Xu*
4. Walk and Learn: Facial Attribute Representation Learning From Egocentric Video and Contextual Data, *Jing Wang, Yu Cheng, Rogerio Schmidt Feris*
5. CNN-N-Gram for Handwriting Word Recognition, *Arik Poznanski, Lior Wolf*

### 1450 S2-2A: Object Detection 2

Format (4 min. for presentation; no questions)

6. Synthetic Data for Text Localisation in Natural Images, *Ankush Gupta, Andrea Vedaldi, Andrew Zisserman*
7. End-To-End People Detection in Crowded Scenes, *Russell Stewart, Mykhaylo Andriluka, Andrew Y. Ng*
8. Real-Time Salient Object Detection With a Minimum Spanning Tree, *Wei-Chih Tu, Shengfeng He, Qingxiong Yang, Shao-Yi Chien*
9. Local Background Enclosure for RGB-D Salient Object Detection, *David Feng, Nick Barnes, Shaodi You, Chris McCarthy*
10. Adaptive Object Detection Using Adjacency and Zoom Prediction, *Yongxi Lu, Tara Javidi, Svetlana Lazebnik*
11. Semantic Channels for Fast Pedestrian Detection, *Arthur Daniel Costea, Sergiu Nedevshi*
12. G-CNN: An Iterative Grid Based Object Detector, *Mahyar Najibi, Mohammad Rastegari, Larry S. Davis*

## 1345-1520 Oral & Spotlight Session 2-2B (Augustus I-II; Emperors Ballroom)

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

**Chairs:** Tali Treibitz (*Univ. of Haifa*)  
Guy Rosman (*MIT/CSAIL*)

### 1345 O2-2B: Computational Photography and Faces

Format (12 min. for presentation + 1 min. for questions)

13. Recurrent Face Aging, *Wei Wang, Zhen Cui, Yan Yan, Jiashi Feng, Shuicheng Yan, Xiangbo Shu, Nicu Sebe*
14. Face2Face: Real-Time Face Capture and Reenactment of RGB Videos, *Justus Thies, Michael Zollhöfer, Marc Stamminger, Christian Theobalt, Matthias Nießner*
15. Self-Adaptive Matrix Completion for Heart Rate Estimation From Face Videos Under Realistic Conditions, *Sergey Tulyakov, Xavier Alameda-Pineda, Elisa Ricci, Lijun Yin, Jeffrey F. Cohn, Nicu Sebe*
16. Visually Indicated Sounds, *Andrew Owens, Phillip Isola, Josh McDermott, Antonio Torralba, Edward H. Adelson, William T. Freeman*
17. Image Style Transfer Using Convolutional Neural Networks, *Leon A. Gatys, Alexander S. Ecker, Matthias Bethge*

### 1450 S2-2B: Computational Photography and Biomedical Applications

Format (4 min. for presentation; no questions)

18. Patch-Based Convolutional Neural Network for Whole Slide Tissue Image Classification, *Le Hou, Dimitris Samaras, Tahsin M. Kurc, Yi Gao, James E. Davis, Joel H. Saltz*
19. Hedgehog Shape Priors for Multi-Object Segmentation, *Hossam Isack, Olga Veksler, Milan Sonka, Yuri Boykov*
20. Latent Variable Graphical Model Selection Using Harmonic Analysis: Applications to the Human Connectome Project (HCP), *Won Hwa Kim, Hyunwoo J. Kim, Nagesh Adluru, Vikas Singh*
21. Simultaneous Estimation of Near IR BRDF and Fine-Scale Surface Geometry, *Gyeongmin Choe, Srinivasa G. Narasimhan, In So Kweon*
22. Do It Yourself Hyperspectral Imaging With Everyday Digital Cameras, *Seoung Wug Oh, Michael S. Brown, Marc Pollefeys, Seon Joo Kim*
23. Automatic Content-Aware Color and Tone Stylization, *Joon-Young Lee, Kalyan Sunkavalli, Zhe Lin, Xiaohui Shen, In So Kweon*

24. Combining Markov Random Fields and Convolutional Neural Networks for Image Synthesis, *Chuan Li, Michael Wand*

## 1520–1545 Break (Forum Ballroom)

## 1545–1645 Plenary Session (Augustus Ballroom)

- **Plenary Talk:** The Conversation on Long-term AI Impacts, *Nick Bostrom (Univ. of Oxford)*

## 1645–1845 Exhibits (Octavius Ballroom)

- Same as Monday morning Exhibits (see pg. 17)

## 1645–1845 Demos (Octavius Ballroom)

- Don't Think, Just Sketch - Deep Sketch-Based Fine-Grained Retrieval, *Yi-Zhe Song, Timothy Hospedales, Tao Xiang (Queen Mary, Univ. of London)*
- Staple: Complementary Learners for Real-Time Tracking, *Luca Bertinetto, Jack Valmadre, Stuart Golodetz, Ondrej Miksik, Philip H.S. Torr (Univ. of Oxford)*
- EmotioNet: An Accurate, Real-Time Algorithm for the Automatic Annotation of Half a Million Facial Expressions in the Wild, *Aleix M. Martinez, Fabian Benitez-Quiroz (The Ohio State Univ.)*
- Grounding Natural Language Phrases in Images With Bounding Boxes and Segmentation Masks, *Anna Rohrbach, Ronghang Hu, Marcus Rohrbach, Bernt Schiele, Trevor Darrell (Max Planck Institute for Informatics, UC Berkeley, ICSI)*

## 1645–1845 Poster Session P2-2 (Octavius Ballroom)

### Biomedical Image Analysis

25. DCAN: Deep Contour-Aware Networks for Accurate Gland Segmentation, *Hao Chen, Xiaojuan Qi, Lequan Yu, Pheng-Ann Heng*
26. Learning to Read Chest X-Rays: Recurrent Neural Cascade Model for Automated Image Annotation, *Hoo-Chang Shin, Kirk Roberts, Le Lu, Dina Demner-Fushman, Jianhua Yao, Ronald M. Summers*
27. Conformal Surface Alignment With Optimal Möbius Search, *Huu Le, Tat-Jun Chin, David Suter*
28. Coupled Harmonic Bases for Longitudinal Characterization of Brain Networks, *Seong Jae Hwang, Nagesh Adluru, Maxwell D. Collins, Sathya N. Ravi, Barbara B. Bendlin, Sterling C. Johnson, Vikas Singh*

29. Automating Carotid Intima-Media Thickness Video Interpretation With Convolutional Neural Networks, *Jae Shin, Nima Tajbakhsh, R. Todd Hurst, Christopher B. Kendall, Jianming Liang*

### Deep Learning and CNNs

30. Context Encoders: Feature Learning by Inpainting, *Deepak Pathak, Philipp Krähenbuhl, Jeff Donahue, Trevor Darrell, Alexei A. Efros*
31. Comparative Deep Learning of Hybrid Representations for Image Recommendations, *Chenyi Lei, Dong Liu, Weiping Li, Zheng-Jun Zha, Houqiang Li*
32. Fast ConvNets Using Group-Wise Brain Damage, *Vadim Lebedev, Victor Lempitsky*
33. Learning to Co-Generate Object Proposals With a Deep Structured Network, *Zeeshan Hayder, Xuming He, Mathieu Salzmann*
34. DeepFool: A Simple and Accurate Method to Fool Deep Neural Networks, *Seyed-Mohsen Moosavi-Dezfooli, Alhussein Fawzi, Pascal Frossard*
35. Blockout: Dynamic Model Selection for Hierarchical Deep Networks, *Calvin Murdock, Zhen Li, Howard Zhou, Tom Duerig*
36. FireCaffe: Near-Linear Acceleration of Deep Neural Network Training on Compute Clusters, *Forrest N. Iandola, Matthew W. Moskewicz, Khalid Ashraf, Kurt Keutzer*
37. MDL-CW: A Multimodal Deep Learning Framework With Cross Weights, *Sarah Rastegar, Mahdieh Soleymani, Hamid R. Rabiee, Seyed Mohsen Shojaei*
38. Structured Receptive Fields in CNNs, *Jörn-Henrik Jacobsen, Jan van Gemert, Zhongyu Lou, Arnold W. M. Smeulders*

### Events, Actions, and Activity Recognition

39. First Person Action Recognition Using Deep Learned Descriptors, *Suriya Singh, Chetan Arora, C. V. Jawahar*
40. Recognizing Micro-Actions and Reactions From Paired Egocentric Videos, *Ryo Yonetani, Kris M. Kitani, Yoichi Sato*
41. Mining 3D Key-Pose-Motifs for Action Recognition, *Chunyu Wang, Yizhou Wang, Alan L. Yuille*
42. Predicting the Where and What of Actors and Actions Through Online Action Localization, *Khurram Soomro, Haroon Idrees, Mubarak Shah*
43. Actions ~ Transformations, *Xiaolong Wang, Ali Farhadi, Abhinav Gupta*
44. Visual Path Prediction in Complex Scenes With Crowded Moving Objects, *Youngjoon Yoo, Kimin Yun, Sangdo Yun, JongHee Hong, Hawoock Jeong, Jin Young Choi*

45. End-To-End Learning of Action Detection From Frame Glimpses in Videos, *Serena Yeung, Olga Russakovsky, Greg Mori, Li Fei-Fei*
46. Action Recognition in Video Using Sparse Coding and Relative Features, *Anali Alfaro, Domingo Mery, Alvaro Soto*
47. Improving Human Action Recognition by Non-Action Classification, *Yang Wang, Minh Hoai*
48. Actionness Estimation Using Hybrid Fully Convolutional Networks, *Limin Wang, Yu Qiao, Xiaoou Tang, Luc Van Gool*
49. Real-Time Action Recognition With Enhanced Motion Vector CNNs, *Bowen Zhang, Limin Wang, Zhe Wang, Yu Qiao, Hanli Wang*

#### **Image Enhancement, Restoration, and Texture**

50. Laplacian Patch-Based Image Synthesis, *Joo Ho Lee, Inchang Choi, Min H. Kim*
51. Rain Streak Removal Using Layer Priors, *Yu Li, Robby T. Tan, Xiaojie Guo, Jiangbo Lu, Michael S. Brown*
52. Gradient-Domain Image Reconstruction Framework With Intensity-Range and Base-Structure Constraints, *Takashi Shibata, Masayuki Tanaka, Masatoshi Okutomi*
53. Removing Clouds and Recovering Ground Observations in Satellite Image Sequences via Temporally Contiguous Robust Matrix Completion, *Jialei Wang, Peder A. Olsen, Andrew R. Conn, Aurélie C. Lozano*
54. D<sup>3</sup>: Deep Dual-Domain Based Fast Restoration of JPEG-Compressed Images, *Zhangyang Wang, Ding Liu, Shiyu Chang, Qing Ling, Yingzhen Yang, Thomas S. Huang*
55. From Bows to Arrows: Rolling Shutter Rectification of Urban Scenes, *Vijay Rengarajan, Ambasamudram N. Rajagopalan, Rangarajan Aravind*
56. A Weighted Variational Model for Simultaneous Reflectance and Illumination Estimation, *Xueyang Fu, Delu Zeng, Yue Huang, Xiao-Ping Zhang, Xinghao Ding*
57. Visualizing and Understanding Deep Texture Representations, *Tsung-Yu Lin, Subhransu Maji*

#### **Low-Level Vision**

58. Robust Kernel Estimation With Outliers Handling for Image Deblurring, *Jinshan Pan, Zhouchen Lin, Zhixun Su, Ming-Hsuan Yang*

#### **Large Scale Visual Recognition**

59. Online Collaborative Learning for Open-Vocabulary Visual Classifiers, *Hanwang Zhang, Xindi Shang, Wenzhuo Yang, Huan Xu, Huanbo Luan, Tat-Seng Chua*

60. Rethinking the Inception Architecture for Computer Vision, *Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, Zbigniew Wojna*

#### **Object Class Detection and Recognition**

61. Cross Modal Distillation for Supervision Transfer, *Saurabh Gupta, Judy Hoffman, Jitendra Malik*
62. Efficient Point Process Inference for Large-Scale Object Detection, *Trung T. Pham, Seyed Hamid Rezatofighi, Ian Reid, Tat-Jun Chin*
63. Weakly Supervised Deep Detection Networks, *Hakan Bilen, Andrea Vedaldi*
64. BORDER: An Oriented Rectangles Approach to Texture-Less Object Recognition, *Jacob Chan, Jimmy Addison Lee, Qian Kemao*
65. Active Image Segmentation Propagation, *Suyog Dutt Jain, Kristen Grauman*
66. Inside-Outside Net: Detecting Objects in Context With Skip Pooling and Recurrent Neural Networks, *Sean Bell, C. Lawrence Zitnick, Kavita Bala, Ross Girshick*
67. RIFD-CNN: Rotation-Invariant and Fisher Discriminative Convolutional Neural Networks for Object Detection, *Gong Cheng, Peicheng Zhou, Junwei Han*
68. Reinforcement Learning for Visual Object Detection, *Stefan Mathe, Aleksis Pirinen, Cristian Sminchisescu*
69. Detecting Repeating Objects Using Patch Correlation Analysis, *Inbar Huberman, Raanan Fattal*
70. Analyzing Classifiers: Fisher Vectors and Deep Neural Networks, *Sebastian Bach, Alexander Binder, Grégoire Montavon, Klaus-Robert Müller, Wojciech Samek*

#### **Scene and Image Classification**

71. Learning Deep Features for Discriminative Localization, *Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva, Antonio Torralba*
72. Seeing Through the Human Reporting Bias: Visual Classifiers From Noisy Human-Centric Labels, *Ishan Misra, C. Lawrence Zitnick, Margaret Mitchell, Ross Girshick*
73. Learning Aligned Cross-Modal Representations From Weakly Aligned Data, *Lluís Castrejón, Yusuf Aytar, Carl Vondrick, Hamed Pirsiavash, Antonio Torralba*
74. A Probabilistic Collaborative Representation Based Approach for Pattern Classification, *Sijia Cai, Lei Zhang, Wangmeng Zuo, Xiangchu Feng*
75. Learning Structured Inference Neural Networks With Label Relations, *Hexiang Hu, Guang-Tong Zhou, Zhiwei Deng, Zicheng Liao, Greg Mori*

76. Discriminative Multi-Modal Feature Fusion for RGBD Indoor Scene Recognition, *Hongyuan Zhu, Jean-Baptiste Weibel, Shijian Lu*
77. Conditional Graphical Lasso for Multi-Label Image Classification, *Qiang Li, Maoying Qiao, Wei Bian, Dacheng Tao*
78. Region Ranking SVM for Image Classification, *Zijun Wei, Minh Hoai*

### Scene Understanding

79. Predicting Motivations of Actions by Leveraging Text, *Carl Vondrick, Deniz Oktay, Hamed Pirsiavash, Antonio Torralba*

### Video Surveillance

80. BoxCars: 3D Boxes as CNN Input for Improved Fine-Grained Vehicle Recognition, *Jakub Sochor, Adam Herout, Jiří Havel*
81. Highway Vehicle Counting in Compressed Domain, *Xu Liu, Zilei Wang, Jiashi Feng, Hongsheng Xi*
82. Camera Calibration From Periodic Motion of a Pedestrian, *Shiyao Huang, Xianghua Ying, Jiangpeng Rong, Zeyu Shang, Hongbin Zha*

**Notes:**

[illegible]

## Wednesday, June 29

### 0730–1700 Registration (Octavius Prefunction)

### 0730–0830 Breakfast (Forum Ballroom)

### 0900–1030 Oral & Spotlight Session 3-1A (Augustus III–VI)

Papers in this session are also in Poster Session P3-1.

**Chairs:** Hamed Pirsiavash (*Univ. of Maryland*)  
Grégory Rogez (*INRIA*)

#### 0900 O3-1A: Actions and Human Pose

Format (12 min. for presentation + 1 min. for questions)

1. Dynamic Image Networks for Action Recognition, *Hakan Bilen, Basura Fernando, Efstratios Gavves, Andrea Vedaldi, Stephen Gould*
2. Detecting Events and Key Actors in Multi-Person Videos, *Vignesh Ramanathan, Jonathan Huang, Sami Abu-El-Hajja, Alexander Gorban, Kevin Murphy, Li Fei-Fei*
3. Regularizing Long Short Term Memory With 3D Human-Skeleton Sequences for Action Recognition, *Behrooz Mahasseni, Sinisa Todorovic*
4. Personalizing Human Video Pose Estimation, *James Charles, Tomas Pfister, Derek Magee, David Hogg, Andrew Zisserman*
5. End-To-End Learning of Deformable Mixture of Parts and Deep Convolutional Neural Networks for Human Pose Estimation, *Wei Yang, Wanli Ouyang, Hongsheng Li, Xiaogang Wang*

#### 1005 S3-1A: Activity Recognition

Format (4 min. for presentation; no questions)

6. Actor-Action Semantic Segmentation With Grouping Process Models, *Chenliang Xu, Jason J. Corso*
7. Temporal Action Localization With Pyramid of Score Distribution Features, *Jun Yuan, Bingbing Ni, Xiaokang Yang, Ashraf A. Kassim*
8. Recognizing Activities of Daily Living With a Wrist-Mounted Camera, *Katsunori Ohnishi, Atsushi Kanehira, Asako Kanezaki, Tatsuya Harada*
9. Harnessing Object and Scene Semantics for Large-Scale Video Understanding, *Zuxuan Wu, Yanwei Fu, Yu-Gang Jiang, Leonid Sigal*

10. Video-Story Composition via Plot Analysis, *Jinsoo Choi, Tae-Hyun Oh, In So Kweon*
11. Temporal Action Detection Using a Statistical Language Model, *Alexander Richard, Juergen Gall*

### 0900–1030 Oral & Spotlight Session 3-1B (Augustus I–II; Emperors Ballroom)

Papers in this session are also in Poster Session P3-1.

**Chairs:** João Carreira (*Google DeepMind*)  
Hyun Oh Song (*Stanford Univ.*)

#### 0900 O3-1B: Semantic Segmentation

Format (12 min. for presentation + 1 min. for questions)

12. Multi-Scale Patch Aggregation (MPA) for Simultaneous Detection and Segmentation, *Shu Liu, Xiaojuan Qi, Jianping Shi, Hong Zhang, Jiaya Jia*
13. Instance-Aware Semantic Segmentation via Multi-Task Network Cascades, *Jifeng Dai, Kaiming He, Jian Sun*
14. ScribbleSup: Scribble-Supervised Convolutional Networks for Semantic Segmentation, *Di Lin, Jifeng Dai, Jiaya Jia, Kaiming He, Jian Sun*
15. Feature Space Optimization for Semantic Video Segmentation, *Abhijit Kundu, Vibhav Vineet, Vladlen Koltun*
16. Large-Scale Semantic 3D Reconstruction: An Adaptive Multi-Resolution Model for Multi-Class Volumetric Labeling, *Maroš Bláha, Christoph Vogel, Audrey Richard, Jan D. Wegner, Thomas Pock, Konrad Schindler*

#### 1005 S3-1B: Semantic Parsing and Segmentation

Format (4 min. for presentation; no questions)

17. Semantic Object Parsing With Local-Global Long Short-Term Memory, *Xiaodan Liang, Xiaohui Shen, Donglai Xiang, Jiashi Feng, Liang Lin, Shuicheng Yan*
18. Efficient Piecewise Training of Deep Structured Models for Semantic Segmentation, *Guosheng Lin, Chunhua Shen, Anton van den Hengel, Ian Reid*
19. Learning Transferrable Knowledge for Semantic Segmentation With Deep Convolutional Neural Network, *Seunghoon Hong, Junhyuk Oh, Honglak Lee, Bohyung Han*
20. The Cityscapes Dataset for Semantic Urban Scene Understanding, *Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, Bernt Schiele*
21. Gaussian Conditional Random Field Network for Semantic Segmentation, *Raviteja Vemulapalli, Oncel Tuzel, Ming-Yu Liu, Rama Chellapa*

22. The SYNTHIA Dataset: A Large Collection of Synthetic Images for Semantic Segmentation of Urban Scenes, *German Ros, Laura Sellart, Joanna Materzynska, David Vazquez, Antonio M. López*

## 1030–1100 Break (Octavius Ballroom)

## 1030–1230 Exhibits (Octavius Ballroom)

- Same as Monday morning Exhibits (see pg. 17)

## 1030–1230 Demos (Octavius Ballroom)

- Face2Face, *Justus Thies, Michael Zollhoefer, Christian Theobalt, Marc Stamminger, Matthias Niessner (Univ. of Erlangen-Nuremberg, Max Planck Institute for Informatics, Stanford Univ.)*
- AntiCrime: A System for Suspect Retrieval and Loitering Discovery in Large Scale Surveillance Videos, *Jianquan Liu, Shoji Nishimura, and Takuya Araki (NEC Corporation)*
- 3D Semantic Parsing of Large-Scale Indoor Spaces, *Iro Armeni, Ozan Sener, Amir R. Zamir (Stanford Univ., Cornell Univ.)*
- Robust Spatial Layout Estimation for Cluttered Indoor Scenes on Mobile Devices, *Kuan Fang, Kevin Chen, Silvio Savarese (Stanford Univ.)*

## 1030–1230 Poster Session P3-1 (Octavius Ballroom)

### 3D Vision

23. Progressive Prioritized Multi-View Stereo, *Alex Locher, Michal Perdoch, Luc Van Gool*
24. WarpNet: Weakly Supervised Matching for Single-View Reconstruction, *Angjoo Kanazawa, David W. Jacobs, Manmohan Chandraker*
25. What Sparse Light Field Coding Reveals About Scene Structure, *Ole Johannsen, Antonin Sulc, Bastian Goldluecke*
26. Online Reconstruction of Indoor Scenes From RGB-D Streams, *Hao Wang, Jun Wang, Wang Liang*
27. Patches, Planes and Probabilities: A Non-Local Prior for Volumetric 3D Reconstruction, *Ali Osman Ulusoy, Michael J. Black, Andreas Geiger*
28. Single Image Camera Calibration With Lenticular Arrays for Augmented Reality, *Ian Schillebeecx, Robert Pless*
29. Augmented Blendshapes for Real-Time Simultaneous 3D Head Modeling and Facial Motion Capture, *Diego Thomas, Rin-ichiro Taniguchi*

30. Learned Binary Spectral Shape Descriptor for 3D Shape Correspondence, *Jin Xie, Meng Wang, Yi Fang*
31. Multiple Model Fitting as a Set Coverage Problem, *Luca Magri, Andrea Fusiello*
32. Piecewise-Planar 3D Approximation From Wide-Baseline Stereo, *Cédric Verleysen, Christophe De Vleeschouwer*
33. Sparse to Dense 3D Reconstruction From Rolling Shutter Images, *Olivier Saurer, Marc Pollefeys, Gim Hee Lee*
34. Consistency of Silhouettes and Their Duals, *Matthew Trager, Martial Hebert, Jean Ponce*
35. Rolling Shutter Absolute Pose Problem With Known Vertical Direction, *Cenek Albl, Zuzana Kukelova, Tomas Pajdla*
36. Uncertainty-Driven 6D Pose Estimation of Objects and Scenes From a Single RGB Image, *Eric Brachmann, Frank Michel, Alexander Krull, Michael Ying Yang, Stefan Gumhold, carsten Rother*
37. Multicamera Calibration From Visible and Mirrored Epipoles, *Andrey Bushnevskiy, Lorenzo Sorgi, Bodo Rosenhahn*

### Face and Gesture

38. Joint Unsupervised Deformable Spatio-Temporal Alignment of Sequences, *Lazaros Zafeiriou, Epameinondas Antonakos, Stefanos Zafeiriou, Maja Pantic*
39. Deep Region and Multi-Label Learning for Facial Action Unit Detection, *Kaili Zhao, Wen-Sheng Chu, Honggang Zhang*
40. Constrained Joint Cascade Regression Framework for Simultaneous Facial Action Unit Recognition and Facial Landmark Detection, *Yue Wu, Qiang Ji*
41. Unconstrained Face Alignment via Cascaded Compositional Learning, *Shizhan Zhu, Cheng Li, Chen-Chang Loy, Xiaoou Tang*
42. Automated 3D Face Reconstruction From Multiple Images Using Quality Measures, *Marcel Pietraschke, Volker Blanz*
43. Occlusion-Free Face Alignment: Deep Regression Networks Coupled With De-Corrupt AutoEncoders, *Jie Zhang, Meina Kan, Shiguang Shan, Xilin Chen*
44. Multimodal Spontaneous Emotion Corpus for Human Behavior Analysis, *Zheng Zhang, Jeff M. Girard, Yue Wu, Xing Zhang, Peng Liu, Umur Ciftci, Shaun Canavan, Michael Reale, Andy Horowitz, Huiyuan Yang, Jeffrey F. Cohn, Qiang Ji, Lijun Yin*
45. Learning Reconstruction-Based Remote Gaze Estimation, *Pei Yu, Jiahuan Zhou, Ying Wu*
46. Joint Training of Cascaded CNN for Face Detection, *Hongwei Qin, Junjie Yan, Xiu Li, Xiaolin Hu*
47. Facial Expression Intensity Estimation Using Ordinal Information, *Rui Zhao, Quan Gan, Shangfei Wang, Qiang Ji*

***Recognition and Detection***

48. Proposal Flow, *Bumsu Ham, Minsu Cho, Cordelia Schmid, Jean Ponce*
49. ProNet: Learning to Propose Object-Specific Boxes for Cascaded Neural Networks, *Chen Sun, Manohar Paluri, Ronan Collobert, Ram Nevatia, Lubomir Bourdev*
50. Seeing Behind the Camera: Identifying the Authorship of a Photograph, *Christopher Thomas, Adriana Kovashka*
51. Material Classification Using Raw Time-Of-Flight Measurements, *Shuochen Su, Felix Heide, Robin Swanson, Jonathan Klein, Clara Callenberg, Matthias Hullin, Wolfgang Heidrich*
52. Weakly Supervised Object Localization With Progressive Domain Adaptation, *Dong Li, Jia-Bin Huang, Yali Li, Shengjin Wang, Ming-Hsuan Yang*
53. Newtonian Scene Understanding: Unfolding the Dynamics of Objects in Static Images, *Roozbeh Mottaghi, Hessam Bagherinezhad, Mohammad Rastegari, Ali Farhadi*
54. Identifying Good Training Data for Self-Supervised Free Space Estimation, *Ali Harakeh, Daniel Asmar, Elie Shammass*
55. Learning to Match Aerial Images With Deep Attentive Architectures, *Hani Altwaijry, Eduard Trulls, James Hays, Pascal Fua, Serge Belongie*
56. Track and Transfer: Watching Videos to Simulate Strong Human Supervision for Weakly-Supervised Object Detection, *Krishna Kumar Singh, Fanyi Xiao, Yong Jae Lee*
57. DeepCAMP: Deep Convolutional Action & Attribute Mid-Level Patterns, *Ali Diba, Ali Mohammad Pazandeh, Hamed Pirsiavash, Luc Van Gool*
58. Canny Text Detector: Fast and Robust Scene Text Localization Algorithm, *Hojin Cho, Myungchul Sung, Bongjin Jun*
59. Temporal Multimodal Learning in Audiovisual Speech Recognition, *Di Hu, Xuelong Li, Xiaoqiang lu*
60. Recovering 6D Object Pose and Predicting Next-Best-View in the Crowd, *Andreas Doumanoglou, Rigas Kouskouridas, Sotiris Malassiotis, Tae-Kyun Kim*
61. Robust 3D Hand Pose Estimation in Single Depth Images: From Single-View CNN to Multi-View CNNs, *Liuha Ge, Hui Liang, Junsong Yuan, Daniel Thalmann*

***Semantic Image Segmentation***

62. Semantic Segmentation With Boundary Neural Fields, *Gedas Bertasius, Jianbo Shi, Lorenzo Torresani*
63. HD Maps: Fine-Grained Road Segmentation by Parsing Ground and Aerial Images, *Gellért Mátyus, Shenlong Wang, Sanja Fidler, Raquel Urtasun*

64. DAG-Recurrent Neural Networks For Scene Labeling, *Bing Shuai, Zhen Zuo, Bing Wang, Gang Wang*
65. Saliency Guided Dictionary Learning for Weakly-Supervised Image Parsing, *Baisheng Lai, Xiaojin Gong*
66. Attention to Scale: Scale-Aware Semantic Image Segmentation, *Liang-Chieh Chen, Yi Yang, Jiang Wang, Wei Xu, Alan L. Yuille*
67. Scene Labeling Using Sparse Precision Matrix, *Nasim Souly, Mubarak Shah*
68. Iterative Instance Segmentation, *Ke Li, Bharath Hariharan, Jitendra Malik*
69. Recurrent Attentional Networks for Saliency Detection, *Jason Kuen, Zhenhua Wang, Gang Wang*

***Semantic Video Segmentation***

70. Instance-Level Video Segmentation From Object Tracks, *Guillaume Seguin, Piotr Bojanowski, Rémi Lajugie, Ivan Laptev*
71. Semantic Instance Annotation of Street Scenes by 3D to 2D Label Transfer, *Jun Xie, Martin Kiefel, Ming-Ting Sun, Andreas Geiger*

***Shape From X***

72. Amplitude Modulated Video Camera - Light Separation in Dynamic Scenes, *Amir Kolaman, Maxim Lvov, Rami Hagege, Hugo Guterman*
73. A Benchmark Dataset and Evaluation for Non-Lambertian and Uncalibrated Photometric Stereo, *Boxin Shi, Zhe Wu, Zhipeng Mo, Dinglong Duan, Sai-Kit Yeung, Ping Tan*
74. Depth From Semi-Calibrated Stereo and Defocus, *Ting-Chun Wang, Manohar Srikanth, Ravi Ramamoorthi*
75. Exploiting Spectral-Spatial Correlation for Coded Hyperspectral Image Restoration, *Ying Fu, Yinqiang Zheng, Imari Sato, Yoichi Sato*
76. Variable Aperture Light Field Photography: Overcoming the Diffraction-Limited Spatio-Angular Resolution Tradeoff, *Julie Chang, Isaac Kauvar, Xuemei Hu, Gordon Wetzstein*
77. Convolutional Networks for Shape From Light Field, *Stefan Heber, Thomas Pock*
78. Panoramic Stereo Videos With a Single Camera, *Rajat Aggarwal, Amrisha Vohra, Anoop M. Nambodiri*
79. The Next Best Underwater View, *Mark Sheinin, Yoav Y. Schechner*
80. Reconstructing Shapes and Appearances of Thin Film Objects Using RGB Images, *Yoshie Kobayashi, Tetsuro Morimoto, Imari Sato, Yasuhiro Mukaigawa, Takao Tomono, Katsushi Ikeuchi*
81. Noisy Label Recovery for Shadow Detection in Unfamiliar Domains, *Tomás F. Yago Vicente, Minh Hoai, Dimitris Samaras*

**1230-1345 Lunch** (Forum Ballroom)

**Notes:**

[illegible]

## 1345–1520 Oral & Spotlight Session 3-2A

(Augustus III–VI)

Papers in this session are also in Poster Session P3-2.

**Chairs:** Marcus Rohrbach (*MPI*)  
Jifeng Dai (*Microsoft Research Asia*)

### 1345 O3-2A: Video Understanding

Format (12 min. for presentation + 1 min. for questions)

1. Deep Hand: How to Train a CNN on 1 Million Hand Images When Your Data Is Continuous and Weakly Labelled, *Oscar Koller, Hermann Ney, Richard Bowden*
2. Recognizing Car Fluenta From Video, *Bo Li, Tianfu Wu, Caiming Xiong, Song-Chun Zhu*
3. Pairwise Decomposition of Image Sequences for Active Multi-View Recognition, *Edward Johns, Stefan Leutenegger, Andrew J. Davison*
4. Inferring Forces and Learning Human Utilities From Videos, *Yixin Zhu, Chenfanfu Jiang, Yibiao Zhao, Demetri Terzopoulos, Song-Chun Zhu*
5. Force From Motion: Decoding Physical Sensation in a First Person Video, *Hyun Soo Park, jyh-Jing Hwang, Jianbo Shi*

### 1450 S3-2A: Video Analysis 2

Format (4 min. for presentation; no questions)

6. Robust Multi-Body Feature Tracker: A Segmentation-Free Approach, *Pan Ji, Hongdong Li, Mathieu Salzmann, Yiran Zhong*
7. Slow and Steady Feature Analysis: Higher Order Temporal Coherence in Video, *Dinesh Jayaraman, Kristen Grauman*
8. Volumetric 3D Tracking by Detection, *Chun-Hao Huang, Benjamin Allain, Jean-Sébastien Franco, Nassir Navab, Slobodan Ilic, Edmond Boyer*
9. The Solution Path Algorithm for Identity-Aware Multi-Object Tracking, *Shoou-I Yu, Deyu Meng, Wangmeng Zuo, Alexander Hauptmann*
10. In Defense of Sparse Tracking: Circulant Sparse Tracker, *Tianzhu Zhang, Adel Bibi, Bernard Ghanem*
11. Optical Flow With Semantic Segmentation and Localized Layers, *Laura Sevilla-Lara, Deqing Sun, Varun Jampani, Michael J. Black*
12. Video Segmentation via Object Flow, *Yi-Hsuan Tsai, Ming-Hsuan Yang, Michael J. Black*

## 1345–1520 Oral & Spotlight Session 3-2B

(Augustus I–II; Emperors Ballroom)

Papers in this session are also in Poster Session P3-2.

**Chairs:** Tatiana Tommasi (*UNC Chapel Hill*)  
Kate Saenko (*UMass Lowell*)

### 1345 O3-2B: Grouping and Optimization Methods

Format (12 min. for presentation + 1 min. for questions)

13. Closed-Form Training of Mahalanobis Distance for Supervised Clustering, *Marc T. Law, YaoLiang Yu, Matthieu Cord, Eric P. Xing*
14. Scalable Sparse Subspace Clustering by Orthogonal Matching Pursuit, *Chong You, Daniel Robinson, René Vidal*
15. Oracle Based Active Set Algorithm for Scalable Elastic Net Subspace Clustering, *Chong You, Chun-Guang Li, Daniel P. Robinson, René Vidal*
16. Sparse Coding and Dictionary Learning With Linear Dynamical Systems, *Wenbing Huang, Fuchun Sun, Lele Cao, Deli Zhao, Huaping Liu, Mehrtaash Harandi*
17. Sublabel-Accurate Relaxation of Nonconvex Energies, *Thomas Möllenhoff, Emanuel Laude, Michael Moeller, Jan Lellmann, Daniel Cremers*

### 1450 S3-2B: Statistical Methods and Transfer Learning

Format (4 min. for presentation; no questions)

18. The Multiverse Loss for Robust Transfer Learning, *Etai Littwin, Lior Wolf*
19. Learning From the Mistakes of Others: Matching Errors in Cross-Dataset Learning, *Viktorii Sharmanska, Novi Quadrianto*
20. An Efficient Exact-PGA Algorithm for Constant Curvature Manifolds, *Rudrasis Chakraborty, Dohyung Seo, Baba C. Vemuri*
21. Online Learning With Bayesian Classification Trees, *Samuel Rota Bulò, Peter Kotschieder*
22. Cross-Stitch Networks for Multi-Task Learning, *Ishan Misra, Abhinav Shrivastava, Abhinav Gupta, Martial Hebert*
23. Deep Metric Learning via Lifted Structured Feature Embedding, *Hyun Oh Song, Yu Xiang, Stefanie Jegelka, Silvio Savarese*
24. Fast Algorithms for Convolutional Neural Networks, *Andrew Lavin, Scott Gray*

## 1520–1545 Break (Forum Ballroom)

## 1545–1645 Plenary Session (Augustus Ballroom)

- **Plenary Talk:** What Babies See, *Elizabeth Spelke (Harvard Univ.)*

## 1645–1845 Exhibits (Octavius Ballroom)

- Same as Monday morning Exhibits (see pg. 17)

## 1645–1845 Demos (Octavius Ballroom)

- 3D Object Tracking from Monocular Images Using Stable Parts, *Alberto Crivellaro, Mahdi Rad, Yannick Verdie, Kwang Moo Yi, Pascal Fua, Vincent Lepetit (EPFL, Graz Univ. of Technology, NCam-Tech)*
- Real-Time Image and Video Super-Resolution on Mobile, Desktop and in the Browser, *Wenzhe Shi, Jose Caballero, Andrew Aitken, Zehan Wang (Magic Pony Technology)*
- CloudCV Visual Question Answering (VQA), *Harsh Agrawal, Aishwarya Agrawal, Jiasen Lu, Deshraj Yadav, Akrit Mohapatra, Devi Parikh, Dhruv Batra (Virginia Tech, JSSATE Noida)*
- Convolutional Pose Machines, *Shih-En Wei, Varun Ramakrishna, Yaser Sheikh (CMU, Uber Advanced Technologies Center)*

## 1645–1845 Poster Session P3-2 (Octavius Ballroom)

### 3D Vision

- Coordinating Multiple Disparity Proposals for Stereo Computation, *Ang Li, Dapeng Chen, Yuanliu Liu, Zejian Yuan*
- Joint Multiview Segmentation and Localization of RGB-D Images Using Depth-Induced Silhouette Consistency, *Chi Zhang, Zhiwei Li, Rui Cai, Hongyang Chao, Yong Rui*
- A Large Dataset to Train Convolutional Networks for Disparity, Optical Flow, and Scene Flow Estimation, *Nikolaus Mayer, Eddy Ilg, Philip Häusser, Philipp Fischer, Daniel Cremers, Alexey Dosovitskiy, Thomas Brox*
- 6D Dynamic Camera Relocalization From Single Reference Image, *Wei Feng, Fei-Peng Tian, Qian Zhang, Jizhou Sun*
- Dense Monocular Depth Estimation in Complex Dynamic Scenes, *René Ranftl, Vibhav Vineet, Qifeng Chen, Vladlen Koltun*
- Using Self-Contradiction to Learn Confidence Measures in Stereo Vision, *Christian Mostegel, Markus Rumpel, Friedrich Fraundorfer, Horst Bischof*

- Understanding Real World Indoor Scenes With Synthetic Data, *Ankur Handa, Viorica Pătrăucean, Vijay Badrinarayanan, Simon Stent, Roberto Cipolla*
- Stereo Matching With Color and Monochrome Cameras in Low-Light Conditions, *Hae-Gon Jeon, Joan-Young Lee, Sunghoon Im, Hyowon Ha, In So Kweon*
- Camera Calibration From Dynamic Silhouettes Using Motion Barcodes, *Gil Ben-Artzi, Yoni Kasten, Shmuel Peleg, Michael Werman*
- Structure-From-Motion Revisited, *Johannes L. Schönberger, Jan-Michael Frahm*
- Constructing Canonical Regions for Fast and Effective View Selection, *Wencheng Wang, Tianhao Gao*
- Prior-Less Compressible Structure From Motion, *Chen Kong, Simon Lucey*
- Rolling Shutter Camera Relative Pose: Generalized Epipolar Geometry, *Yuchao Dai, Hongdong Li, Laurent Kneip*
- Structure From Motion With Objects, *Marco Crocco, Cosimo Rubino, Alessio Del Bue*
- DeepHand: Robust Hand Pose Estimation by Completing a Matrix Imputed With Deep Features, *Ayan Sinha, Chihoi Choi, Karthik Ramani*

### Document Analysis

- Multi-Oriented Text Detection With Fully Convolutional Networks, *Zheng Zhang, Chengquan Zhang, Wei Shen, Cong Yao, Wenyu Liu, Xiang Bai*
- Robust Scene Text Recognition With Automatic Rectification, *Baoguang Shi, Xinggang Wang, Pengyuan Lyu, Cong Yao, Xiang Bai*

### Face and Gesture

- Mnemonic Descent Method: A Recurrent Process Applied for End-To-End Face Alignment, *George Trigeorgis, Patrick Snape, Mihalis A. Nicolaou, Epameinondas Antonakos, Stefanos Zafeiriou*
- Large-Pose Face Alignment via CNN-Based Dense 3D Model Fitting, *Amin Jourabloo, Xiaoming Liu*
- Adaptive 3D Face Reconstruction From Unconstrained Photo Collections, *Joseph Roth, Yiyi Tong, Xiaoming Liu*
- Online Detection and Classification of Dynamic Hand Gestures With Recurrent 3D Convolutional Neural Network, *Pavlo Molchanov, Xiaodong Yang, Shalini Gupta, Kihwan Kim, Stephen Tyree, Jan Kautz*
- Kinematic Structure Correspondences via Hypergraph Matching, *Hyung Jin Chang, Tobias Fischer, Maxime Petit, Martina Zambelli, Yiannis Demiris*

47. CP-mtML: Coupled Projection multi-task Metric Learning for Large Scale Face Retrieval, *Binod Bhattarai, Gaurav Sharma, Frederic Jurie*

## Motion and Tracking

48. PatchBatch: A Batch Augmented Loss for Optical Flow, *David Gadot, Lior Wolf*
49. Joint Recovery of Dense Correspondence and Cosegmentation in Two Images, *Tatsunori Tanai, Sudipta N. Sinha, Yoichi Sato*
50. Multi-View People Tracking via Hierarchical Trajectory Composition, *Yuanlu Xu, Xiaobai Liu, Yang Liu, Song-Chun Zhu*
51. Object Tracking via Dual Linear Structured SVM and Explicit Feature Map, *Jifeng Ning, Jimei Yang, Shaojie Jiang, Lei Zhang, Ming-Hsuan Yang*
52. Robust, Real-Time 3D Tracking of Multiple Objects With Similar Appearances, *Taiki Sekii*
53. An Egocentric Look at Video Photographer Identity, *Yedid Hoshen, Shmuel Peleg*
54. Learning Multi-Domain Convolutional Neural Networks for Visual Tracking, *Hyeonseob Nam, Bohyung Han*
55. Hedged Deep Tracking, *Yuankai Qi, Shengping Zhang, Lei Qin, Hongxun Yao, Qingming Huang, Jongwoo Lim, Ming-Hsuan Yang*
56. Structural Correlation Filter for Robust Visual Tracking, *Si Liu, Tianzhu Zhang, Xiaochun Cao, Changsheng Xu*
57. Visual Tracking Using Attention-Modulated Disintegration and Integration, *Jongwon Choi, Hyung Jin Chang, Jiyeoup Jeong, Yiannis Demiris, Jin Young Choi*
58. A Continuous Occlusion Model for Road Scene Understanding, *Vikas Dhiman, Quoc-Huy Tran, Jason J. Corso, Manmohan Chandraker*
59. Virtual Worlds as Proxy for Multi-Object Tracking Analysis, *Adrien Gaidon, Qiao Wang, Yohann Cabon, Eleonora Vig*

## Shape From X

60. Uncalibrated Photometric Stereo by Stepwise Optimization Using Principal Components of Isotropic BRDFs, *Keisuke Midorikawa, Toshihiko Yamasaki, Kiyoharu Aizawa*
61. Unbiased Photometric Stereo for Colored Surfaces: A Variational Approach, *Yvain Quéau, Roberto Mecca, Jean-Denis Durou*
62. 3D Reconstruction of Transparent Objects With Position-Normal Consistency, *Yiming Qian, Minglun Gong, Yee Hong Yang*

63. Real-Time Depth Refinement for Specular Objects, *Roy Or-El, Rom Hershkovitz, Aaron Wetzler, Guy Rosman, Alfred M. Bruckstein, Ron Kimmel*
64. Recovering Transparent Shape From Time-Of-Flight Distortion, *Kenichiro Tanaka, Yasuhiro Mukaigawa, Hiroyuki Kubo, Yasuyuki Matsushita, Yasushi Yagi*
65. Robust Light Field Depth Estimation for Noisy Scene With Occlusion, *Williem, In Kyu Park*
66. Rotational Crossed-Slit Light Field, *Nianyi Li, Haiting Lin, Bilin Sun, Mingyuan Zhou, Jingyi Yu*
67. Single Image Object Modeling Based on BRDF and R-Surfaces Learning, *Fabrizio Natola, Valsamis Ntouskos, Fiora Pirri, Marta Sanzari*

## Statistical Methods and Learning

68. A Nonlinear Regression Technique for Manifold Valued Data With Applications to Medical Image Analysis, *Monami Banerjee, Rudrasis Chakraborty, Edward Ofori, Michael S. Okun, David E. Viallancourt, Baba C. Vemuri*
69. RAID-G: Robust Estimation of Approximate Infinite Dimensional Gaussian With Application to Material Recognition, *Qilong Wang, Peihua Li, Wangmeng Zuo, Lei Zhang*
70. An Empirical Evaluation of Current Convolutional Architectures' Ability to Manage Nuisance Location and Scale Variability, *Nikolaos Karianakis, Jingming Dong, Stefano Soatto*
71. Learning Sparse High Dimensional Filters: Image Filtering, Dense CRFs and Bilateral Neural Networks, *Varun Jampani, Martin Kiefel, Peter V. Gehler*
72. Mixture of Bilateral-Projection Two-Dimensional Probabilistic Principal Component Analysis, *Fujiao Ju, Yanfeng Sun, Junbin Gao, Simeng Liu, Yongli Hu, Baocai Yin*
73. Rolling Rotations for Recognizing Human Actions From 3D Skeletal Data, *Raviteja Vemulapalli, Rama Chellapa*
74. Improving the Robustness of Deep Neural Networks via Stability Training, *Stephan Zheng, Yang Song, Thomas Leung, Ian Goodfellow*
75. Logistic Boosting Regression for Label Distribution Learning, *Chao Xing, Xin Geng, Hui Xue*
76. Efficient Temporal Sequence Comparison and Classification Using Gram Matrix Embeddings on a Riemannian Manifold, *Xikang Zhang, Yin Wang, Mengran Gou, Mario Sznaier, Octavia Camps*

### Vision For Graphics

77. Deep Reflectance Maps, *Konstantinos Rematas, Tobias Ritschel, Mario Fritz, Efstratios Gavves, Tinne Tuytelaars*
78. Semantic Filtering, *Qingxiong Yang*

## Vision For Robotics

79. UAV Sensor Fusion With Latent-Dynamic Conditional Random Fields in Coronal Plane Estimation, *Amir M. Rahimi, Raphael Ruschel, B.S. Manjunath*
80. Robust Visual Place Recognition with Graph Kernels, *Elena Stumm, Christopher Mei, Simon Lacroix, Juan Nieto, Marco Hutter, Roland Siegwart*

### Semantic Image Segmentation

81. Semantic Image Segmentation With Task-Specific Edge Detection Using CNNs and a Discriminatively Trained Domain Transform, *Liang-Chieh Chen, Jonathan T. Barron, George Papandreou, Kevin Murphy, Alan L. Yuille*

**1900-2100 Social Event** (Forum Ballroom)

**Notes:**

[illegible]

# Thursday, June 30

## 0730-1700 Registration (Octavius Prefunction)

## 0730-0830 Breakfast (Forum Ballroom)

## 0900-1030 Oral & Spotlight Session 4-1A (Augustus III-VI)

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

**Chairs:** Alexander Schwing (*Univ. of Toronto*)  
Adriana Kovashka (*Univ. of Pittsburgh*)

### 0900 O4-1A: Image & Video Captioning and Descriptions

Format (12 min. for presentation + 1 min. for questions)

1. Natural Language Object Retrieval, *Ronghang Hu, Huazhe Xu, Marcus Rohrbach, Jiashi Feng, Kate Saenko, Trevor Darrell*
2. DenseCap: Fully Convolutional Localization Networks for Dense Captioning, *Justin Johnson, Andrej Karpathy, Li Fei-Fei*
3. Unsupervised Learning From Narrated Instruction Videos, *Jean-Baptiste Alayrac, Piotr Bojanowski, Nishant Agrawal, Josef Sivic, Ivan Laptev, Simon Lacoste-Julien*
4. Video Paragraph Captioning Using Hierarchical Recurrent Neural Networks, *Haonan Yu, Jiang Wang, Zhiheng Huang, Yi Yang, Wei Xu*
5. Jointly Modeling Embedding and Translation to Bridge Video and Language, *Yingwei Pan, Tao Mei, Ting Yao, Houqiang Li, Yong Rui*

### 1005 S4-1A: High Level Semantics

Format (4 min. for presentation: no questions)

6. We Are Humor Beings: Understanding and Predicting Visual Humor, *Arjun Chandrasekaran, Ashwin K. Vijayakumar, Stanislaw Antol, Mohit Bansal, Dhruv Batra, C. Lawrence Zitnick, Devi Parikh*
7. Where to Look: Focus Regions for Visual Question Answering, *Kevin J. Shih, Saurabh Singh, Derek Hoiem*
8. Ask Me Anything: Free-Form Visual Question Answering Based on Knowledge From External Sources, *Qi Wu, Peng Wang, Chunhua Shen, Anthony Dick, Anton van den Hengel*
9. MovieQA: Understanding Stories in Movies Through Question-Answering, *Makarand Tapaswi, Yukun Zhu, Rainer Stiefelhagen, Antonio Torralba, Raquel Urtasun, Sanja Fidler*

10. TGIF: A New Dataset and Benchmark on Animated GIF Description, *Yuncheng Li, Yale Song, Liangliang Cao, Joel Tetreault, Larry Goldberg, Alejandro Jaimes, Jiebo Luo*
11. Image Captioning With Semantic Attention, *Quanzeng You, Hailin Jin, Zhaoen Wang, Chen Fang, Jiebo Luo*

## 0900-1030 Oral & Spotlight Session 4-1B

(Augustus I-II; Emperors Ballroom)

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

**Chairs:** Jonathan Taylor (*Microsoft Research Cambridge*)  
Victor Prisacariu (*Univ. of Oxford*)

### 0900 O4-1B: Non-Rigid Reconstruction and Motion Analysis Format (12 min. for presentation + 1 min. for questions)

12. Temporally Coherent 4D Reconstruction of Complex Dynamic Scenes, *Armin Mustafa, Hansung Kim, Jean-Yves Guillemaut, Adrian Hilton*
13. Consensus of Non-Rigid Reconstructions, *Minsik Lee, Jungchan Cho, Songhwai Oh*
14. Isometric Non-Rigid Shape-From-Motion in Linear Time, *Shaifali Parashar, Daniel Pizarro, Adrien Bartoli*
15. Learning Online Smooth Predictors for Realtime Camera Planning Using Recurrent Decision Trees, *Jianhui Chen, Hoang M. Le, Peter Carr, Yisong Yue, James J. Little*
16. Egocentric Future Localization, *Hyun Soo Park, Jyh-Jing Hwang, Yedong Niu, Jianbo Shi*
17. Full Flow: Optical Flow Estimation By Global Optimization Over Regular Grids, *Qifeng Chen, Vladlen Koltun*

### 1018 S4-1B: Human Pose Estimation

Format (4 min. for presentation: no questions)

18. Structured Feature Learning for Pose Estimation, *Xiao Chu, Wanli Ouyang, Hongsheng Li, Xiaogang Wang*
19. Convolutional Pose Machines, *Shih-En Wei, Varun Ramakrishna, Takeo Kanade, Yaser Sheikh*
20. Human Pose Estimation With Iterative Error Feedback, *João Carreira, Pulkit Agrawal, Katerina Fragkiadaki, Jitendra Malik*

## 1030-1100 Break (Octavius Ballroom)

## 1030-1230 Exhibits (Octavius Ballroom)

- Same as Monday morning Exhibits (see pg. 17)

**1030–1230 Demos (Octavius Ballroom)**

- SemanticPaint: Mobile Interactive Segmentation of 3D Worlds, *Michael Sapienza, Stuart Golodetz, Olaf Kahler, Victor A. Prisacariu, Philip H. S. Torr (Univ. of Oxford)*
- Real-Time Human Segmentation Using Deep Fully Convolutional Encoder-Decoder Network, *Ping-Lin Chang, Tingfan Wu and Shawn Guan (Umbo CV)*
- Glass-Enabled Contextual Interactions Through Video-Based Guidance, *Dima Damen, Davide Moltisanti (Univ. of Bristol)*
- Mobile 3D Scanner, *Oliver Saurer, Ioannis Mariggis, Petri Tanskanen, Marc Pollefeys (ETH Zurich)*

**1030–1230 Poster Session P4-1 (Octavius Ballroom)****Deep Learning and CNNs**

21. WELDON: Weakly Supervised Learning of Deep Convolutional Neural Networks, *Thibaut Durand, Nicolas Thome, Matthieu Cord*
22. DisturbLabel: Regularizing CNN on the Loss Layer, *Lingxi Xie, Jingdong Wang, Zhen Wei, Meng Wang, Qi Tian*
23. Gradual DropIn of Layers to Train Very Deep Neural Networks, *Leslie N. Smith, Emily M. Hand, Timothy Doster*
24. Structure Inference Machines: Recurrent Neural Networks for Analyzing Relations in Group Activity Recognition, *Zhiwei Deng, Arash Vahdat, Hexiang Hu, Greg Mori*
25. Deep SimNets, *Nadav Cohen, Or Sharir, Amnon Shashua*
26. Studying Very Low Resolution Recognition Using Deep Networks, *Zhangyang Wang, Shiyu Chang, Yingzhen Yang, Ding Liu, Thomas S. Huang*
27. Deep Gaussian Conditional Random Field Network: A Model-Based Deep Network for Discriminative Denoising, *Raviteja Vemulapalli, Oncel Tuzel, Ming-Yu Liu*
28. Event-Specific Image Importance, *Yufei Wang, Zhe Lin, Xiaohui Shen, Radomir Měch, Gavin Miller, Garrison W. Cottrell*
29. Quantized Convolutional Neural Networks for Mobile Devices, *Jiaxiang Wu, Cong Leng, Yuhang Wang, Qinghao Hu, Jian Cheng*
30. Inverting Visual Representations With Convolutional Networks, *Alexey Dosovitskiy, Thomas Brox*

**Face Recognition**

31. Pose-Aware Face Recognition in the Wild, *Iacopo Masi, Stephen Rawls, Gérard Medioni, Prem Natarajan*
32. Multi-View Deep Network for Cross-View Classification, *Meina Kan, Shiguang Shan, Xilin Chen*

33. Sparsifying Neural Network Connections for Face Recognition, *Yi Sun, Xiaogang Wang, Xiaoou Tang*
34. Pairwise Linear Regression Classification for Image Set Retrieval, *Qingxiang Feng, Yicong Zhou, Rushi Lan*
35. The MegaFace Benchmark: 1 Million Faces for Recognition at Scale, *Ira Kemelmacher-Shlizerman, Steven M. Seitz, Daniel Miller, Evan Brossard*
36. Learnt Quasi-Transitive Similarity for Retrieval From Large Collections of Faces, *Ognjen Arandjelović*
37. Latent Factor Guided Convolutional Neural Networks for Age-Invariant Face Recognition, *Yandong Wen, Zhifeng Li, Yu Qiao*
38. Copula Ordinal Regression for Joint Estimation of Facial Action Unit Intensity, *Robert Walecki, Ognjen Rudovic, Vladimir Pavlovic, Maja Pantic*

**Face and Gesture**

39. A Robust Multilinear Model Learning Framework for 3D Faces, *Timo Bolkart, Stefanie Wuhrer*
40. Ordinal Regression With Multiple Output CNN for Age Estimation, *Zhenxing Niu, Mo Zhou, Le Wang, Xinbo Gao, Gang Hua*

**Human Pose Estimation**

41. DeepCut: Joint Subset Partition and Labeling for Multi Person Pose Estimation, *Leonid Pishchulin, Eldar Insafutdinov, Siyu Tang, Bjørn Andres, Mykhaylo Andriluka, Peter V. Gehler, Bernt Schiele*
42. Thin-Slicing for Pose: Learning to Understand Pose Without Explicit Pose Estimation, *Suha Kwak, Minsu Cho, Ivan Laptev*
43. A Dual-Source Approach for 3D Pose Estimation From a Single Image, *Hashim Yasin, Umar Iqbal, Björn Krüger, Andreas Weber, Juergen Gall*
44. Efficiently Creating 3D Training Data for Fine Hand Pose Estimation, *Markus Oberweger, Gernot Riegler, Paul Wohlhart, Vincent Lepetit*
45. Sparseness Meets Deepness: 3D Human Pose Estimation From Monocular Video, *Xiaowei Zhou, Menglong Zhu, Spyridon Leonardos, Konstantinos G. Derpanis, Kostas Daniilidis*

**Images and Language**

46. Answer-Type Prediction for Visual Question Answering, *Kushal Kafle, Christopher Kanan*
47. Visual Word2Vec (vis-w2v): Learning Visually Grounded Word Embeddings Using Abstract Scenes, *Satwik Kottur, Ramakrishna Vedantam, José M. F. Moura, Devi Parikh*
48. VisualyW: Grounded Question Answering in Images, *Yuke Zhu, Oliver Groth, Michael Bernstein, Li Fei-Fei*

- 49. Learning Deep Structure-Preserving Image-Text Embeddings, *Liwei Wang, Yin Li, Svetlana Lazebnik*
- 50. Yin and Yang: Balancing and Answering Binary Visual Questions, *Peng Zhang, Yash Goyal, Douglas Summers-Stay, Dhruv Batra, Devi Parikh*

## Shape Representations and Matching

- 51. GIFT: A Real-Time and Scalable 3D Shape Search Engine, *Song Bai, Xiang Bai, Zhichao Zhou, Zhaoxiang Zhang, Longin Jan Latecki*
- 52. Functional Faces: Groupwise Dense Correspondence Using Functional Maps, *Chao Zhang, William A. P. Smith, Arnaud Dessein, Nick Pears, Hang Dai*
- 53. Similarity Metric For Curved Shapes In Euclidean Space, *Girum G. Demisse, Djamila Aouada, Björn Ottersten*
- 54. Shape Analysis With Hyperbolic Wasserstein Distance, *Jie Shi, Wen Zhang, Yalin Wang*
- 55. Tensor Power Iteration for Multi-Graph Matching, *Xinchu Shi, Haibin Ling, Weiming Hu, Junliang Xing, Yanning Zhang*

## Transfer Learning

- 56. Multivariate Regression on the Grassmannian for Predicting Novel Domains, *Yongxin Yang, Timothy M. Hospedales*
- 57. Learning Cross-Domain Landmarks for Heterogeneous Domain Adaptation, *Yao-Hung Hubert Tsai, Yi-Ren Yeh, Yu-Chiang Frank Wang*
- 58. Geospatial Correspondences for Multimodal Registration, *Diego Marcos, Raffay Hamid, Devis Tuia*
- 59. Constrained Deep Transfer Feature Learning and Its Applications, *Yue Wu, Qiang Ji*

## Unsupervised, Semi-Supervised and Interactive Learning

- 60. Deep Canonical Time Warping, *George Trigeorgis, Mihalís A. Nicolaou, Stefanos Zafeiriou, Björn W. Schuller*
- 61. Multilinear Hyperplane Hashing, *Xianglong Liu, Xinjie Fan, Cheng Deng, Zhujin Li, Hao Su, Dacheng Tao*
- 62. Large Scale Hard Sample Mining With Monte Carlo Tree Search, *Olivier Canévet, François Fleuret*
- 63. Multi-Label Ranking From Positive and Unlabeled Data, *Atsushi Kanehira, Tatsuya Harada*
- 64. Joint Unsupervised Learning of Deep Representations and Image Clusters, *Jianwei Yang, Devi Parikh, Dhruv Batra*
- 65. Kernel Sparse Subspace Clustering on Symmetric Positive Definite Manifolds, *Ming Yin, Yi Guo, Junbin Gao, Zhaoshui He, Shengli Xie*
- 66. Symmetry reCAPTCHA, *Chris Funk, Yanxi Liu*
- 67. Unsupervised Learning of Discriminative Attributes and Visual Representations, *Chen Huang, Chen Change Loy, Xiaoou Tang*

- 68. When VLAD Met Hilbert, *Mehrtash Harandi, Mathieu Salzmann, Fatih Porikli*
- 69. Approximate Log-Hilbert-Schmidt Distances Between Covariance Operators for Image Classification, *Hà Quang Minh, Marco San Biagio, Loris Bazzani, Vittorio Murino*
- 70. Subspace Clustering With Priors via Sparse Quadratically Constrained Quadratic Programming, *Yongfang Cheng, Yin Wang, Mario Sznai, Octavia Camps*
- 71. Robust Tensor Factorization With Unknown Noise, *Xi'ai Chen, Zhi Han, Yao Wang, Qian Zhao, Deyu Meng, Yandong Tang*
- 72. Kernel Approximation via Empirical Orthogonal Decomposition for Unsupervised Feature Learning, *Yusuke Mukuta, Tatsuya Harada*
- 73. Active Learning for Delineation of Curvilinear Structures, *Agata Mosinska-Domanska, Raphael Sznitman, Przemyslaw Glowacki, Pascal Fua*
- 74. Recognizing Emotions From Abstract Paintings Using Non-Linear Matrix Completion, *Xavier Alameda-Pineda, Elisa Ricci, Yan Yan, Nicu Sebe*
- 75. Tensor Robust Principal Component Analysis: Exact Recovery of Corrupted Low-Rank Tensors via Convex Optimization, *Canyi Lu, Jiashi Feng, Yudong Chen, Wei Liu, Zhouchen Lin, Shuicheng Yan*
- 76. Sliced Wasserstein Kernels for Probability Distributions, *Soheil Kolouri, Yang Zou, Gustavo K. Rohde*
- 77. Trace Quotient Meets Sparsity: A Method for Learning Low Dimensional Image Representations, *Xian Wei, Hao Shen, Martin Kleinsteuber*
- 78. Backtracking ScSPM Image Classifier for Weakly Supervised Top-Down Saliency, *Hisham Cholakkal, Jubin Johnson, Deepu Rajan*

## Video and Language

- 79. MSR-VTT: A Large Video Description Dataset for Bridging Video and Language, *Jun Xu, Tao Mei, Ting Yao, Yong Rui*

## 1230-1345 Lunch (Forum Ballroom)

## 1345–1520 Oral & Spotlight Session 4-2A

(Augustus III–VI)

Papers in this session are also in Poster Session P4-2.

**Chairs:** Basura Fernando (*Australian National Univ.*)  
Alexandre Alahi (*Stanford Univ.*)

### 1345 O4-2A: Learning and CNN Architectures

Format (12 min. for presentation + 1 min. for questions)

1. NetVLAD: CNN Architecture for Weakly Supervised Place Recognition, *Relja Arandjelović, Petr Gronat, Akihiko Torii, Tomas Pajdla, Josef Sivic*
2. Structural-RNN: Deep Learning on Spatio-Temporal Graphs, *Ashesh Jain, Amir R. Zamir, Silvio Savarese, Ashutosh Saxena*
3. Learning to Select Pre-Trained Deep Representations With Bayesian Evidence Framework, *Yong-Deok Kim, Taewoong Jang, Bohyung Han, Seungjin Choi*
4. Synthesized Classifiers for Zero-Shot Learning, *Soravit Changpinyo, Wei-Lun Chao, Boqing Gong, Fei Sha*
5. Semi-Supervised Vocabulary-Informed Learning, *Yanwei Fu, Leonid Sigal*

### 1450 S4-2A: Learning and Optimization

Format (4 min. for presentation; no questions)

6. Simultaneous Clustering and Model Selection for Tensor Affinities, *Zhuwen Li, Shuoguang Yang, Loong-Fah Cheong, Kim-Chuan Toh*
7. Discriminatively Embedded K-Means for Multi-View Clustering, *Jinglin Xu, Junwei Han, Feiping Nie*
8. Min Norm Point Algorithm for Higher Order MRF-MAP Inference, *Ishant Shanu, Chetan Arora, Parag Singla*
9. Learning Deep Representation for Imbalanced Classification, *Chen Huang, Yining Li, Chen Change Loy, Xiaoou Tang*
10. Learning Local Image Descriptors With Deep Siamese and Triplet Convolutional Networks by Minimising Global Loss Functions, *Vijay Kumar B G, Gustavo Carneiro, Ian Reid*
11. Sparse Coding for Third-Order Super-Symmetric Tensor Descriptors With Application to Texture Recognition, *Piotr Koniusz, Anoop Cherian*
12. Random Features for Sparse Signal Classification, *Jen-Hao Rick Chang, Aswin C. Sankaranarayanan, B. V. K. Vijaya Kumar*

## 1345–1520 Oral & Spotlight Session 4-2B

(Augustus I–II; Emperors Ballroom)

Papers in this session are also in Poster Session P4-2.

**Chairs:** Enrique Dunn (*UNC Chapel Hill*)  
Daniel Zoran (*MIT/CSAIL*)

### 1345 O4-2B: 3D Shape Reconstruction

Format (12 min. for presentation + 1 min. for questions)

13. High-Quality Depth From Uncalibrated Small Motion Clip, *Hyowon Ha, Sunghoon Im, Jaesik Park, Hae-Gon Jeon, In So Kweon*
14. Efficient 3D Room Shape Recovery From a Single Panorama, *Hao Yang, Hui Zhang*
15. Structured Prediction of Unobserved Voxels From a Single Depth Image, *Michael Firman, Oisin Mac Aodha, Simon Julier, Gabriel J. Brostow*
16. HyperDepth: Learning Depth From Structured Light Without Matching, *Sean Ryan Fanello, Christoph Rhemann, Vladimir Tankovich, Adarsh Kowdle, Sergio Orts Escolano, David Kim, Shahram Izadi*
17. SVBRDF-Invariant Shape and Reflectance Estimation From Light-Field Cameras, *Ting-Chun Wang, Manmohan Chandraker, Alexei A. Efros, Ravi Ramamoorthi*

### 1450 S4-2B: 3D Reconstruction

Format (4 min. for presentation; no questions)

18. Semantic 3D Reconstruction With Continuous Regularization and Ray Potentials Using a Visibility Consistency Constraint, *Nikolay Savinov, Christian Häne, Lubor Ladický, Marc Pollefeys*
19. Theory and Practice of Structure-From-Motion Using Affine Correspondences, *Carolina Raposo, João P. Barreto*
20. Just Look at the Image: Viewpoint-Specific Surface Normal Prediction for Improved Multi-View Reconstruction, *Silvano Galliani, Konrad Schindler*
21. From Dusk Till Dawn: Modeling in the Dark, *Filip Radenović, Johannes L. Schönberger, Dinghuang Ji, Jan-Michael Frahm, Ondřej Chum, Jiří Matas*
22. Accelerated Generative Models for 3D Point Cloud Data, *Benjamin Eckart, Kihwan Kim, Alejandro Troccoli, Alonzo Kelly, Jan Kautz*
23. Monocular Depth Estimation Using Neural Regression Forest, *Anirban Roy, Sinisa Todorovic*
24. DeepStereo: Learning to Predict New Views From the World's Imagery, *John Flynn, Ivan Neulander, James Philbin, Noah Snavely*

**1520–1545 Break** (Forum Ballroom)

**1545–1645 Oral & Spotlight Session 4-3A**  
(Augustus III–VI)

Papers in this session are also in Poster Session P4-2.

**Chairs:** Ira Kemelmacher-Schlizerman (*Univ. of Washington*)  
Michael Zollhoefer (*MPI*)

**1545 O4-3A: Face, Gesture, & Situation Recognition: Algorithms and Datasets**

Format (12 min. for presentation + 1 min. for questions)

25. WIDER FACE: A Face Detection Benchmark, *Shuo Yang, Ping Luo, Chen-Change Loy, Xiaoou Tang*
26. Situation Recognition: Visual Semantic Role Labeling for Image Understanding, *Mark Yatskar, Luke Zettlemoyer, Ali Farhadi*

**1610 S4-3A: People and Faces**

Format (4 min. for presentation; no questions)

27. A 3D Morphable Model Learnt From 10,000 Faces, *James Booth, Anastasios Roussos, Stefanos Zafeiriou, Allan Ponniah, David Dunaway*
28. Some Like It Hot - Visual Guidance for Preference Prediction, *Rasmus Rothe, Radu Timofte, Luc Van Gool*
29. EmotiNet: An Accurate, Real-Time Algorithm for the Automatic Annotation of a Million Facial Expressions in the Wild, *C. Fabian Benitez-Quiroz, Ramprakash Srinivasan, Aleix M. Martinez*
30. ForgetMeNot: Memory-Aware Forensic Facial Sketch Matching, *Shuxin Ouyang, Timothy M. Hospedales, Yi-Zhe Song, Xueming Li*
31. LOMo: Latent Ordinal Model for Facial Analysis in Videos, *Karan Sikka, Gaurav Sharma, Marian Bartlett*
32. Discriminative Invariant Kernel Features: A Bells-and-Whistles-Free Approach to Unsupervised Face Recognition and Pose Estimation, *Dipan K. Pal, Felix Juefei-Xu, Marios Savvides*
33. Bottom-Up and Top-Down Reasoning With Hierarchical Rectified Gaussians, *Peiyun Hu, Deva Ramanan*
34. Fits Like a Glove: Rapid and Reliable Hand Shape Personalization, *David Joseph Tan, Thomas Cashman, Jonathan Taylor, Andrew Fitzgibbon, Daniel Tarlow, Sameh Khamis, Shahram Izadi, Jamie Shotton*
35. Slicing Convolutional Neural Network for Crowd Video Understanding, *Jing Shao, Chen-Change Loy, Kai Kang, Xiaogang Wang*

**1545–1645 Spotlight Session 4-3B** (Augustus I–II; Emperors Ballroom)

Papers in this session are also in Poster Session P4-2.

**Chairs:** Stefan Leutenegger (*Imperial College London*)  
Derek Hoiem (*UIUC*)

**1545 S4-3B: 3D, Stereo, Matching, and Saliency Estimation**  
Format (4 min. for presentation; no questions)

36. Linear Shape Deformation Models With Local Support Using Graph-Based Structured Matrix Factorisation, *Florian Bernard, Peter Gemmar, Frank Hertel, Jorge Goncalves, Johan Thunberg*
37. Motion From Structure (MfS): Searching for 3D Objects in Cluttered Point Trajectories, *Jayakorn Vongkulbhisal, Ricardo Cabral, Fernando De la Torre, João P. Costeira*
38. Volumetric and Multi-View CNNs for Object Classification on 3D Data, *Charles R. Qi, Hao Su, Matthias Niessner, Angela Dai, Mengyuan Yan, Leonidas J. Guibas*
39. Detecting Vanishing Points Using Global Image Context in a Non-Manhattan World, *Menghua Zhai, Scott Workman, Nathan Jacobs*
40. Learning Weight Uncertainty With Stochastic Gradient MCMC for Shape Classification, *Chunyuan Li, Andrew Stevens, Changyou Chen, Yunchen Pu, Zhe Gan, Lawrence Carin*
41. A Field Model for Repairing 3D Shapes, *Duc Thanh Nguyen, Binh-Son Hua, Khoi Tran, Quang-Hieu Pham, Sai-Kit Yeung*
42. GOGMA: Globally-Optimal Gaussian Mixture Alignment, *Dylan Campbell, Lars Petersson*
43. Efficient Deep Learning for Stereo Matching, *Wenjie Luo, Alexander G. Schwing, Raquel Urtasun*
44. Efficient Coarse-To-Fine PatchMatch for Large Displacement Optical Flow, *Yinlin Hu, Rui Song, Yunsong Li*
45. FANNG: Fast Approximate Nearest Neighbour Graphs, *Ben Harwood, Tom Drummond*
46. Exemplar-Driven Top-Down Saliency Detection via Deep Association, *Shengfeng He, Rynson W.H. Lau, Qingxiong Yang*
47. Unconstrained Salient Object Detection via Proposal Subset Optimization, *Jianming Zhang, Stan Sclaroff, Zhe Lin, Xiaohui Shen, Brian Price, Radomir Mech*
48. Recombinator Networks: Learning Coarse-To-Fine Feature Aggregation, *Sina Honari, Jason Yosinski, Pascal Vincent, Christopher Pal*
49. End-To-End Saliency Mapping via Probability Distribution Prediction, *Saumya Jetley, Naila Murray, Eleonora Vig*

**1645–1845 Exhibits** (Octavius Ballroom)

- Same as Monday morning Exhibits (see pg. 17)

**1645–1845 Demos (Octavius Ballroom)**

- Pairwise Conditional Random Forests for Facial Expression Recognition, *Arnaud Dapogny (Universite Pierre et Marie Curie)*
- A Storyline-Based Query-Answering Benchmark for Understanding Scenes and Events from a Network of Cameras, *Qi Hang, Tianfu Wu, Mun-Wai Lee, Song-Chun Zhu (UCLA)*
- Vision-Aided Inertial Navigation for Virtual Reality Applications, *Ryan DuToit, Georgios Georgiou, Kejian Wu, Chao Guo, Stergios Roumeliotis (Univ. of Minnesota)*
- Robust Keyframe-Based Monocular SLAM for Augmented Reality, *Haomin Liu, Guofeng Zhang, Hujun Bao (Zhejiang University)*

**1645–1845 Poster Session P4-2 (Octavius Ballroom)****Biologically Inspired Vision**

50. A Paradigm for Building Generalized Models of Human Image Perception Through Data Fusion, *Shaoping Fan, Tian-Tsong Ng, Bryan L. Koenig, Ming Jiang, Qi Zhao*
51. Longitudinal Face Modeling via Temporal Deep Restricted Boltzmann Machines, *Chi Nhan Duong, Khoa Luu, Kha Gia Quach, Tien D. Bui*
52. Saliency Unified: A Deep Architecture for Simultaneous Eye Fixation Prediction and Salient Object Segmentation, *Srinivas S. S. Kruthiventi, Vennela Gudisa, Jaley H. Dholakiya, R. Venkatesh Babu*

**Image Alignment and Registration**

53. Estimating Correspondences of Deformable Objects "In-The-Wild", *Yuxiang Zhou, Epameinondas Antonakos, Joan Alabort-i-Medina, Anastasios Roussos, Stefanos Zafeiriou*
54. Gravitational Approach for Point Set Registration, *Vladislav Golyanik, Sk Aziz Ali, Didier Stricker*
55. Context-Aware Gaussian Fields for Non-Rigid Point Set Registration, *Gang Wang, Zhicheng Wang, Yupei Chen, Qiangqiang Zhou, Weidong Zhao*

**Optimization**

56. Trust No One: Low Rank Matrix Factorization Using Hierarchical RANSAC, *Magnus Oskarsson, Kenneth Batstone, Kalle Åström*
57. Relaxation-Based Preprocessing Techniques for Markov Random Field Inference, *Chen Wang, Ramin Zabih*
58. Sparse Coding for Classification via Discrimination Ensemble, *Yuhui Quan, Yong Xu, Yuping Sun, Yan Huang, Hui Ji*

59. Principled Parallel Mean-Field Inference for Discrete Random Fields, *Pierre Baqué, Timur Bagautdinov, François Fleuret, Pascal Fua*
60. Guaranteed Outlier Removal With Mixed Integer Linear Programs, *Tat-Jun Chin, Yang Heng Kee, Anders Eriksson, Frank Neumann*
61. Memory Efficient Max Flow for Multi-Label Submodular MRFs, *Thalaiyasingam Ajanthan, Richard Hartley, Mathieu Salzmann*
62. Proximal Riemannian Pursuit for Large-Scale Trace-Norm Minimization, *Mingkui Tan, Shijie Xiao, Junbin Gao, Dong Xu, Anton van den Hengel, Qinfeng Shi*
63. Minimizing the Maximal Rank, *Erik Bylow, Carl Olsson, Fredrik Kahl, Mikael Nilsson*
64. Solving Temporal Puzzles, *Caglayan Dicle, Burak Yilmaz, Octavia Camps, Mario Sznaier*
65. Estimating Sparse Signals With Smooth Support via Convex Programming and Block Sparsity, *Sohil Shah, Tom Goldstein, Christoph Studer*
66. TenSR: Multi-Dimensional Tensor Sparse Representation, *Na Qi, Yunhui Shi, Xiaoyan Sun, Baocai Yin*
67. Moral Lineage Tracing, *Florian Jug, Evgeny Levinkov, Corinna Blasse, Eugene W. Myers, Bjørn Andres*
68. Globally Optimal Rigid Intensity Based Registration: A Fast Fourier Domain Approach, *Behrooz Nasihatkan, Frida Fejné, Fredrik Kahl*
69. On Benefits of Selection Diversity via Bilevel Exclusive Sparsity, *Haichuan Yang, Yijun Huang, Lam Tran, Ji Liu, Shuai Huang*

**Recognition and Detection**

70. Fast Training of Triplet-Based Deep Binary Embedding Networks, *Bohan Zhuang, Guosheng Lin, Chunhua Shen, Ian Reid*
71. Marr Revisited: 2D-3D Alignment via Surface Normal Prediction, *Aayush Bansal, Bryan Russell, Abhinav Gupta*
72. Recovering the Missing Link: Predicting Class-Attribute Associations for Unsupervised Zero-Shot Learning, *Ziad Al-Halah, Makarand Tapaswi, Rainer Stiefelhofen*
73. Fast Zero-Shot Image Tagging, *Yang Zhang, Boqing Gong, Mubarak Shah*
74. Modality and Component Aware Feature Fusion For RGB-D Scene Classification, *Anran Wang, Jianfei Cai, Jiwen Lu, Tat-Jen Cham*
75. PPP: Joint Pointwise and Pairwise Image Label Prediction, *Yilin Wang, Suhang Wang, Jiliang Tang, Huan Liu, Baoxin Li*

76. Cataloging Public Objects Using Aerial and Street-Level Images – Urban Trees, *Jan D. Wegner, Steven Branson, David Hall, Konrad Schindler, Pietro Perona*
77. Deep Exemplar 2D-3D Detection by Adapting From Real to Rendered Views, *Francisco Massa, Bryan C. Russell, Mathieu Aubry*
78. Zero-Shot Learning via Joint Latent Similarity Embedding, *Ziming Zhang, Venkatesh Saligrama*
79. CRAFT Objects From Images, *Bin Yang, Junjie Yan, Zhen Lei, Stan Z. Li*

**END OF MAIN CVPR 2016 CONFERENCE**

**Notes:**

[illegible]

## Friday, July 1

**NOTE:** Workshop locations are in the online proceedings and the mobile app.

**0730–1700 Registration** (Octavius Prefunction)

**0730–0830 Breakfast** (Forum Ballroom)

## Egocentric Vision

**Organizers:** Michael Ryoo

Kris Kitani

Yin Li

Yong Jae Lee

**Schedule:** Full Day

0920 **Welcome**

0930 **Keynote Talk:** *Kristen Grauman (Univ. of Texas at Austin)*

1000 **Keynote Talk:** *Dov Katz (Oculus VR)*

1030 **Morning Break** (Forum Ballroom)

1045 **Spotlight Presentation 1**

- Body Part Based Re-Identification From an Egocentric Perspective, *Federica Fergnani, Stefano Alletto, Giuseppe Serra, Joaquim De Mira, Rita Cucchiara*
- Discovering Objects of Joint Attention via First-Person Sensing, *Hiroshi Kera, Ryo Yonetani, Keita Higuchi, Yoichi Sato*
- KrishnaCam: Using a Longitudinal, Single-Person, Egocentric Dataset for Scene Understanding Tasks, *Krishna Kumar Singh, Kayvon Fatahalian, Alexei A. Efros*
- Egocentric Place Recognition for AR Wayfinding Assistive Technology, *John S. Zelek, Charbel Azzi*
- Towards Social Interaction Detection in Egocentric Photo-Streams, *Maedeh Aghaei*
- Embedding of Egocentric Action Videos in Semantic-Visual Graph, *Davide Moltisanti, Michael Wray, Walterio Mayol-Cuevas, Dima Damen*
- Where Did I Leave My Phone? *Cristian Reyes, Eva Mohamedano, Kevin McGuinness, Noel E. O'Connor, Xavier Giro-i-Nieto*
- Segmenting Egocentric Videos to Highlight Personal Locations of Interest, *Antonino Furnari, Giovanni Maria Farinella, Sebastiano Battiato*

- Finding Egocentric Image Topics Through Convolutional Neural Network Based Representations, *Kai Zhen, David Crandall*
- Text Detection in Stores From a Wearable Camera Using a Repetition Prior, *Bo Xiong, Kristen Grauman*

1115 **Poster Session 1**

1200 **Lunch** (Forum Ballroom)

1330 **Keynote Talk:** *David Crandall (Indiana Univ.)*

1400 **Spotlight Presentation 2**

- A Pointing Gesture Based Egocentric Interaction System: Dataset, Approach and Application, *Yichao Huang, Xiaorui Liu, Xin Zhang, Lianwen Jin*
- Multimodal Multi-Stream Deep Learning for Egocentric Activity Recognition, *Sibo Song, Vijay Chandrasekhar, Bappaditya Mandal, Liyuan Li, Joo-Hwee Lim, Giduthuri Sateesh Babu, Phyo Phyo San, Ngai-Man Cheung*
- Generating Notifications for Missing Actions: Don't Forget to Turn the Lights Off!, *Bilge Soran, Ali Farhadi, Linda Shapiro*
- Egocentric Multi-Modal Dataset With Visual and Physiological Signals, *Katsuyuki Nakamura, Alexandre Alahi, Serena Yeung, Li Fei-Fei*
- Head Gesture Recognition in Spontaneous Human Conversations: A Benchmark, *Wu Yang, Kai Akiyama, Kris Kitani, Laszlo Jeni, Yasuhiro Mukaigawa*
- Attention-Driven Egocentric Computer Vision for Robotic Wheelchair Navigation, *Haoxiang Liy, Philippos Mordohaiy, Gang Hua*
- EgoMemNet: Visual Memorability Adaptation to Egocentric Images, *Marc Carne', Xavier Giro-i-Nieto, Petia Radeva, Cathal Gurrin*
- Efficiently Creating 3D Training Data for Fine Hand Pose Estimation, *Markus Oberweger, Gernot Riegler, Paul Wohlhart, Vincent Lepetit*
- Recovering 6D Object Pose and Predicting Next-Best-View in the Crowd, *Andreas Domanoglou, Rigas Kouskouridas, Sotiris Malassiotis, Tae-Kyun Kim*

1425 **Poster Session 2**

1530 **Keynote Talk:** *Wenyi Zhao (DAQRI)*

1540 **Afternoon Break** (Forum Ballroom)

1600 **Keynote Talk:** *Shmuel Peleg (The Hebrew Univ. of Jerusalem)*

1630 **Keynote Talk:** *James Rehg (Georgia Tech)*

1700 **Closing Remarks**

## Large Scale Visual Recognition and Retrieval

**Organizers:** Vicente Ordóñez Román

Olga Russakovsky

Jia Deng

Alex Berg

Yuanqing Lin

Jason Corso

Fei-fei Li

Samy Bengio

**Schedule:** Full Day

0900 **Introduction**

0915 **Invited Talk:** Towards Recognition in the Open-World,  
*Deva Ramanan (Carnegie Mellon University)*

1000 **Morning Break** (Forum Ballroom)

1015 **Invited Talk:** Guaranteed Machine Learning Using Tensor  
Methods, *Anima Anandkumar (University of California, Irvine)*

1100 **Invited Talk:** Mining Relationships From Large Scale  
Photos, *Jia Li (Snapchat)*

1145 **Spotlight Presentations**

- VisFlow: A Declarative Platform for Parallelizing Large-Scale Vision Programs, *Yao Lu, Aakanksha Chowdhery, Srikanth Kandula*
- The Unreasonable Effectiveness of Noisy Data for Fine-Grained Recognition, *Jonathan Krause, Benjamin Sapp, Andrew Howard, Howard Zhou, Alexander Toshev, Tom Duerig, James Philbin, Li Fei-Fei*

1215 **Lunch** (Forum Ballroom)

1330 **Invited Talk:** What Does It Mean to Scale Visual  
Recognition? *Ali Farhadi (University of Washington)*

1415 **Invited Talk:** Visual Question Answering, *Devi Parikh (Virginia Tech)*

1500 **Spotlight Presentations**

- The Open World of Micro-Videos, *Phuc X. Nguyen, Gregory Rogez, Charles Fowlkes, Deva Ramanan*
- Discovering Unsupervised Binary Codes for Learning From Small Sample Sets, *Yu-Xiong Wang, Martial Hebert*

1530 **Afternoon Break** (Forum Ballroom) & **Poster Session**

1630 **Invited Talk:** How a Driverless Car Sees the World, *Jiajun Zhu (Google)*

1715 **Closing Remarks**

## DeepVision: Deep Learning in Computer Vision

**Organizers:** José M. Álvarez

Yann LeCun

Fatih Porikli

Yi Li

**Schedule:** Full Day

0800 **Welcome**

0810 **Invited Talk:** *Tomaso Poggio (MIT)*

0855 **Invited Talk:** *Miguel A. Carreria Perpignan (Univ. of California, Merced)*

0940 **Posters / Spotlights**

- Joint Learning of Convolutional Neural Networks and Temporally Constrained Metrics for Tracklet Association, *Bing Wang, Li Wang, Bing Shuai, Zhen Zuo, Ting Liu, Kap Luk Chan, Gang Wang*
- Faster R-CNN Features for Instance Search, *Amaia Salvador, Xavier Giró-i-Nieto, Ferran Marqués, Shin'ichi Satoh*
- Deep End2End Voxel2Voxel Prediction, *Du Tran, Lubomir Bourdev, Rob Fergus, Lorenzo Torresani, Manohar Paluri*
- Adversarial Diversity and Hard Positive Generation, *Andras Rozsa, Ethan M. Rudd, Terrance E. Boult*
- Learning by Tracking: Siamese CNN for Robust Target Association, *Laura Leal-Taixé, Cristian Canton-Ferrer, Konrad Schindler*
- ReSeg: A Recurrent Neural Network-Based Model for Semantic Segmentation, *Francesco Visin, Marco Ciccone, Adriana Romero, Kyle Kastner, Kyunghyun Cho, Yoshua Bengio, Matteo Matteucci, Aaron Courville*
- Rich Image Captioning in the Wild, *Kenneth Tran, Xiaodong He, Lei Zhang, Jian Sun, Cornelia Carapcea, Chris Thrasher, Chris Buehler, Chris Sienkiewicz*

1000 **Morning Break** (Forum Ballroom)

1030 **Invited Talk:** *Graham Taylor (Univ. of Guelph)*

1115 **Invited Talk:** *Cees Snoek (Univ. of Amsterdam / Qualcomm)*

1200 **Invited Talk:** *Andrej Karpathy (Stanford Univ.)*

1245 **Lunch** (Forum Ballroom)

1435 **Invited Talk:** *Anton van den Hengel (Univ. of Adelaide)*

1525 **Afternoon Break** (Forum Ballroom)

1600 **Invited Talk:** *Matthew Zeiler (Clarifai)*

1700 **Closing Remarks**

## Biomedical Image Registration

**Organizers:** Julia Schnabel  
Kensaku Mori  
Ben Glocker  
Mattias Heinrich

**Schedule:** Full Day

0830 **Welcome**

### S1: Advances in Regularisation & Optimisation (0840-1000)

- 0840 Discrete Optimisation for Group-Wise Cortical Surface Atlas, *Emma C. Robinson, Ben Glocker, Martin Rajchl, Daniel Rueckert*
- 0900 Sparse Kernel Machines for Discontinuous Registration and Nonstationary Regularization, *Christoph Jud, Nadia Möri, Philippe C. Cattin*
- 0920 Accurate Small Deformation Exponential Approximant to Integrate Large Velocity Fields: Application to Image Registration, *Sebastiano Ferraris, Marco Lorenzi, Pankaj Daga, Marc Modat, Tom Vercauteren*
- 0940 Fast Deformable Image Registration With Non-Smooth Dual Optimization, *Martin Rajchl, John S.H Baxter, Wu Qiu, Ali R. Khan, Aaron Fenster, Terry M. Peters, Daniel Rueckert, Jing Yuan*

### **1000 Morning Break (Forum Ballroom) & Poster Session 1**

- Image Registration for Placenta Reconstruction, *Floris Gaisser, Pieter P. Jonker, Toshio Chiba*
  - Tissue-Volume Preserving Deformable Image Registration for 4DCT Pulmonary Images, *Bowen Zhao, Gary E. Christensen, Joo Hyun Song, Yue Pan, Sarah E. Gerard, Joseph M. Reinhardt, Kaifang Du, Taylor Patton, John M. Bayouth, Geoffrey D. Hugo*
  - Registering Retinal Vessel Images From Local to Global via Multiscale and Multicycle Features, *Haiyong Zheng, Lin Chang, Tengda Wei, Xinxin Qiu, Ping Lin, Yangfan Wang*
  - The Design of SuperElastix — A Unifying Framework for a Wide Range of Image Registration Methodologies, *Floris F. Berendsen, Kasper Marstal, Stefan Klein, Marius Staring*
  - Tumor Growth Estimation via Registration of DCE-MRI Derived Tumor Specific Descriptors, *Thaïs Roque, Bartłomiej W. Papiez, Veerle Kersemans, Sean Smart, Danny Allen, Michael Chappell, Julia A. Schnabel*
  - Graph-Constrained Surface Registration Based on Tutte Embedding, *Wei Zeng, Yi-Jun Yang, Muhammad Razib*
- 1045 **Keynote:** Non-Rigid 2D and 3D Registration, *Lourdes Agapito (Univ. College London)*

### **S2: Mosaicing With Temporal Stability (1145-1225)**

- 1145 A Combined EM and Visual Tracking Probabilistic Model for Robust Mosaicking: Application to Fetoscopy, *Marcel Tella-Amo, Pankaj Daga, François Chadebecq, Stephen Thompson, Dzhoshkun I. Shakir, George Dwyer, Ruwan Wimalasundera, Jan Deprest, Danail Stoyanov, Tom Vercauteren, Sebastien Ourselin*
- 1205 Reducing Drift in Mosaicing Slit-Lamp Retinal Images, *Kristina Prokopetc, Adrien Bartoli*

### **1230 Lunch (Forum Ballroom)**

- 1330 **Keynote:** VarPro, ICP, Lifting, and All That, *Andrew Fitzgibbon (Microsoft Cambridge)*

### **S3: How to Deal with Missing Data (1430-1530)**

- 1430 How to Build an Average Model When Samples Are Variably Incomplete? Application to Fossil Data, *Jean Dumoncel, Gérard Subsol, Stanley Durlleman, Jean-Pierre Jessel, Amélie Beaudet, José Braga*
- 1450 Population Shape Collapse in Large Deformation Registration of MR Brain Images, *Wei Shao, Gary E. Christensen, Hans J. Johnson, Joo Hyun Song, Oguz C. Durumeric, Casey P. Johnson, Joseph J. Shaffer, Vincent A. Magnotta, Jess G. Fiedorowicz, John A. Wemmie*
- 1510 Registration of Developmental Image Sequences With Missing Data, *Istvan Csapo, Yundi Shi, Mar Sanchez, Martin Styner, Marc Niethammer*

### **1530 Afternoon Break (Forum Ballroom) & Poster Session 2**

- Current- and Varifold-Based Registration of Lung Vessel and Airway Trees, *Yue Pan, Gary E. Christensen, Oguz C. Durumeric, Sarah E. Gerard, Joseph M. Reinhardt, Geoffrey D. Hugo*
- SimpleElastix: A User-Friendly, Multi-Lingual Library for Medical Image Registration, *Kasper Marstal, Floris Berendsen, Marius Staring, Stefan Klein*
- Effects of Resolution and Registration Algorithm on the Accuracy of EPI vNavs for Real Time Head Motion Correction in MRI, *Yingzhuo Zhang, Iman Aganj, André J. van der Kouwe, M. Dylan Tisdall*
- Graph Cuts-Based Registration Revisited: A Novel Approach for Lung Image Registration Using Supervoxels and Image-Guided Filtering, *Adam Szmul, Bartłomiej W. Papiez, Russell Bates, Andre Hallack, Julia A. Schnabel, Vicente Grau*
- Multi-Atlas Based Pseudo-CT Synthesis Using Multimodal Image Registration and Local Atlas Fusion Strategies, *Johanna Degen, Mattias P. Heinrich*

- A Fast DRR Generation Scheme for 3D-2D Image Registration Based on the Block Projection Method, *Zhiping Mu*

## **S4: Registration for Advanced Magnetic Resonance Imaging (1615-1715)**

- 1615 Optimal Estimation of Diffusion in DW-MRI by High-Order MRF-Based Joint Deformable Registration and Diffusion Modeling, *Evgenios N. Kornaropoulos, Evangelia I. Zacharakis, Pierre Zerbib, Chieh Lin, Alain Rahmouni, Nikos Paragios*
- 1635 Total Correlation-Based Groupwise Image Registration for Quantitative MRI, *Jean-Marie Guyader, Wyke Huizinga, Valerio Fortunati, Dirk H. J. Poot, Matthijs van Kranenburg, Jifke F. Veenland, Margarethus M. Paulides, Wiro J. Niessen, Stefan Klein*
- 1655 Multimodal Whole Brain Registration: MRI and High Resolution Histology, *Maryana Alegro, Edson Amaro-Jr, Buren Loring, Helmut Heinsen, Eduardo Alho, Lilla Zöllei, Daniela Ushizima, Lea T. Grinberg*
- 1715 **Closing Remarks**

## **Large Scale 3D Data: Acquisition, Modelling and Analysis**

**Organizers:** Jian Zhang  
Myron Brown  
Charalambos Poullis  
Konrad Schindler

**Schedule:** Full Day

- 0845 **Registration & Welcome**
- 0900 **Invited Talk:** 3D Models From Satellite Imagery: History and Future, *Joseph Mundy (Brown Univ.)*
- 0945 **Invited Talk:** Fusing 100's of 3D Point Clouds of Objects, *Robert B. Fisher (Univ. of Edinburgh)*
- 1030 **Morning Break** (Forum Ballroom)
- 1045 **Invited Talk:** Efficient Interactive Labeling of Small Objects in Urban LIDAR Scans, *Tom Funkhouser (Princeton Univ.)*
- 1130 **Overview:** 3D Point Cloud Semantic Segmentation Challenge, *Marc Pollefeys, Timo Hackel, Nikolay Savinov*
- 1215 **Lunch** (Forum Ballroom) **& Poster session**
- Fast and Accurate Registration of Structured Point Clouds With Small Overlaps, *Yanxin Ma, Yulan Guo, Jian Zhao, Min Lu, Jun Zhang, Jianwei Wan*

- Comprehensive Automated 3D Urban Environment Modelling Using Terrestrial Laser Scanning Point Cloud, *Pouria Babahajani, Lixin Fan, Joni-Kristian Kämäräinen, Moncef Gabbouj*
  - RGBD Datasets: Past, Present and Future, *Michael Firman*
  - Real Time Complete Dense Depth Reconstruction for a Monocular Camera, *Xiaoshui Huang, Lixin Fan, Jian Zhang, Qiang Wu, Chun Yuan*
- 1345 **Invited Talk:** Shape Reconstruction and Approximation: Robustness and Guarantees, *Pierre Alliez (INRIA Sophia Antipolis)*
- 1430 **Invited Talk:** 3D Reconstruction of Dynamic Cityscapes, *Ming C. Lin (Univ. of North Carolina at Chapel Hill)*
- 1515 **Overview:** IARPA Multi-View Stereo 3D Mapping Challenge, *HakJae Kim, Myron Brown*
- 1600 **Closing Remarks**

## **Visual Analysis of Satellite to Street Imagery**

**Organizers:** Asaad Hakeem  
Jiangye Yuan

**Schedule:** Full Day

- 0800 **Welcome**
- 0815 **Invited Talk:** *Torsten Sattler (ETH Zurich)*
- 0900 **Invited Talk:** *Nathan Jacobs (Univ. of Kentucky)*
- 1000 **Morning Break** (Forum Ballroom)
- 1030 Semantic Segmentation of Small Objects and Modeling of Uncertainty in Urban Remote Sensing Images Using Deep Convolutional Neural Networks, *Michael Kampffmeyer, Arnt-Børre Salberg, Robert Jenssen*
- 1055 Automatic Alignment of Indoor and Outdoor Building Models Using 3D Line Segments, *Tobias Koch, Marco Körner, Friedrich Fraundorfer*
- 1120 The TUM-DLR Multimodal Earth Observation Evaluation Benchmark, *Tobias Koch, Pablo d'Angelo, Franz Kurz, Friedrich Fraundorfer, Peter Reinartz, Marco Körner*
- 1150 **Invited Talk:** *Raquel Urtasun (Univ. of Toronto)*
- 1235 **Lunch** (Forum Ballroom)
- 1435 **Invited Talk:** *Himaanshu Gupta (Nokia Here)*
- 1510 **Invited Talk:** *John Leonard (MIT)*
- 1525 **Afternoon Break** (Forum Ballroom)
- 1600 **Panel Discussion**
- 1700 **Closing Remarks & Awards**

## ChaLearn Looking at People and Faces of the World: Face Analysis Workshop and Challenge

**Organizers:** Sergio Escalera

Michel Valstar

Xavier Baró

Brais Martinez

Hugo Jair Escalante

Mercedes Torres Torres

Georgios Tzimiropoulos

Isabelle Guyon

**Schedule:** Full Day

845 **Welcome:** *Sergio Escalera, Michel Valstar*

900 **Invited Talk:** DCNN: The Gift That Keeps on Giving, *Rama Chellapa (Univ. of Maryland)*

### S1: Challenge Results & Awards Ceremony (945-1000)

945 ChaLearn Looking at People and Faces of the World: Face Analysis Workshop and Challenge 2016, *Sergio Escalera, Mercedes Torres Torres, Brais Martínez, Xavier Baró, Hugo Jair Escalante, Isabelle Guyon, Georgios Tzimiropoulos, Ciprian Corneou, Marc Oliu, Mohammad Ali Bagheri, Michel Valstar*

**1000 Morning Break** (Forum Ballroom)

### S2: Winners — Apparent Age Estimation (1030-1115)

1030 Apparent Age Estimation Using Ensemble of Deep Learning Models, *Refik Can Malli, Mehmet Aygün, Hazim Kemal Ekenel*

1045 Deep Age Distribution Learning for Apparent Age Estimation, *Zengwei Huo, Xu Yang, Chao Xing, Ying Zhou, Peng Hou, Jiaqi Lv, Xin Geng*

1100 Structured Output SVM Prediction of Apparent Age, Gender and Smile From Deep Features, *Michal Uříčar, Radu Timofte, Rasmus Rothe, Jiří Matas, Luc Van Gool*

### S3: Winners — Smile & Gender Recognition (1030-1200)

1115 Gender and Smile Classification Using Deep Convolutional Neural Networks, *Kaipeng Zhang, Lianzhi Tan, Zhifeng Li, Yu Qiao*

1130 DeepBE: Learning Deep Binary Encoding for Multi-Label Classification, *Chenghua Li, Qi Kang, Guojing Ge, Qiang Song, Hanqing Lu, Jian Cheng*

1145 Facial Attributes Classification Using Multi-Task Representation Learning, *Max Ehrlich, Timothy J. Shields, Timur Almaev, Mohamed R. Amer*

1200 **Invited Talk:** TBA, *Jeffrey Cohn (Univ. of Pittsburgh; CMU)*

**1235 Lunch** (Forum Ballroom)

### S4: Looking at Faces I (1435-1525)

1435 ChaLearn Looking at People RGB-D Isolated and Continuous Datasets for Gesture Recognition, *Jun Wan, Yibing Zhao, Shuai Zhou, Isabelle Guyon, Sergio Escalera, Stan Z. Li*

1450 Dominant Codewords Selection With Topic Model for Action Recognition, *Hirokatsu Kataokai, Kenji Iwata, Yutaka Satoh, Masaki Hayashi, Yoshimitsu Aoki, Slobodan Ilic*

1505 Inferring Visual Persuasion via Body Language, Setting, and Deep Features, *Xinyue Huang, Adriana Kovashka*

**1525 Afternoon Break** (Forum Ballroom)

1600 **Invited Talk:** TBA, *Florian Schroff (Google)*

### S5: Looking at Faces II (1640-1740)

1640 Kernel ELM and CNN Based Facial Age Estimation, *Furkan Gürpınar, Heysem Kaya, Hamdi Dibekioğlu, Ali Salah*

1655 Person-Independent 3D Gaze Estimation Using Face Frontalization, *László A. Jeni, Jeffrey F. Cohn*

1710 Apparent Age Estimation From Face Images Combining General and Children-Specialized Deep Learning Models, *Grigory Antipov, Moez Baccouche, Sid-Ahmed Berrani, Jean-Luc Dugelay*

1725 Identifying Same Persons From Temporally Synchronized Videos Taken by Multiple Wearable Cameras, *Kang Zheng, Hao Guo, Xiaochuan Fan, Hongkai Yu, Song Wang*

1740 **Closing Ceremony & Announcements:** *Sergio Escalera, Michel Valstar*

## Embedded Vision

**Organizers:** Stefano Mattoccia

Jagadeesh Sankaran

Martin Humenberger

Swarup Medasani

**Schedule:** Full Day

0800 **Welcome**

0810 **Keynote:** Vision-Centric ADAS Silicon Architectures for Platform Efficiency, Intelligence and Flexibility, *Chris Rowen (Cadence IP)*

0855 A Diverse Low Cost High Performance Platform for Advanced Driver Assistance System (ADAS) Applications, *Prashanth Viswanath, Kedar Chitnis, Pramod Swami, Mihir Mody, Sujith Shivalingappa, Soyeb Nagori, Manu Mathew,*

*Kumar Desappan, Shyam Jagannathan, Deepak Poddar, Anshu Jain, Hrushikesh Garud, Vikram Appia, Mayank Mangla, Shashank Dabral*

- 0915 A Visual Attention Algorithm Designed for Coupled Oscillator Acceleration, *Christopher Thomas, Adriana Kovashka, Donald Chiarulli, Steven Levitan*
- 0935 Embedded Motion Detection via Neural Response Mixture Background Modeling, *Mohammad Javad Shafiee, Parthipan Siva, Paul Fieguth, Alexander Wong*

**1000 Morning Break** (Forum Ballroom)

- 1030 **Keynote:** The Challenges and Fun of Building Mu: A Tiny Smart Eye for Toys, *Tianli Yu (Morpx Inc.)*
- 1115 A Scalable High-Performance Hardware Architecture for Real-Time Stereo Vision by Semi-Global Matching, *Jaco Hofmann, Jens Korinth, Andreas Koch*
- 1135 Vision Based Autonomous Orientational Control for Aerial Manipulation via On-Board FPGA, *Leewiwatwong Suphachart, Shouhei Shimahara, Robert Ladig, Kazuhiro Shimonomura*
- 1155 Embedded Vision System for Atmospheric Turbulence Mitigation, *Ajinkya Deshmukh, Gaurav Bhosale, Swarup Shanti, Karthik Reddy, Hemanthkumar P., Chandrasekhar A., Kirankumar P., Vijaysagar K.*

**1235 Lunch** (Forum Ballroom)

- 1330 **Keynote:** Energy-Efficient Hardware for Embedded Vision and Deep Convolutional Neural Networks, *Vivienne Sze (MIT)*
- 1415 **Spotlights Posters and Demos**
- 1435 **Posters**
- Approximated Prediction Strategy for Reducing Power Consumption of Convolutional Neural Network Processor, *Takayuki Ujii, Masayuki Hiromoto, Takashi Sato*
  - Visual Monocular Obstacle Avoidance for Small Unmanned Vehicles, *Levente Kovács*
  - Real-Time, Embedded Scene Invariant Crowd Counting Using Scale-Normalized Histogram of Moving Gradients (HoMG), *Parthipan Siva, Mohammad Javad Shafiee, Michael Jamieson, Alexander Wong*
  - 3DCapture: 3D Reconstruction for a Smartphone, *Oleg Muratov, Yury Slynko, Vitaly Chernov, Maria Lyubimtseva, Artem Shamsuarov, Victor Bucha*
  - Embedded Computing Framework for Vision-Based Real-Time Surround Threat Analysis and Driver Assistance, *Frankie Lu, Sean Lee, Ravi Kumar Satzoda, Mohan Trivedi*

**1525 Afternoon Break** (Forum Ballroom)

- 1600 **Keynote:** Multi-Sensor Fusion for Self-Healing Vision Based Navigation, *Stephan Weiss (Alpen-Adria-Universität)*

**1645 Demos**

- Aquifi SIR: Extremely Low Latency Depth and Color Camera With Computation on FPGA, *Paolo Di Febbo, Alessandro Mattei, Alessandro Temil, Aryan Hazeghi, Stefano Mattoccia, Carlo Dal Mutto, Abbas Rafii, Jason Trachewsky*
- Embedded Computing Framework for Vision-Based Real-Time Surround Threat Analysis and Driver Assistance, *Frankie Lu, Sean Lee, Ravi Satzoda, Mohan Trivedi*
- Embedding Convolutional Neural Networks for Real-Time Vehicle Localization Into Consumer Grade Mobile Phones, *Nauto, Inc.*

**1730 Closing Remarks and Awards**

## Computational Cameras and Displays

**Organizers:** Matthew O'Toole

Jean-Charles Bazin

Aswin C. Sankaranarayanan

El Mustapha Mouaddib

**Schedule:** Full Day

**0845 Welcome**

- 0900 **Keynote Talk:** Computational Microscopy for Gigapixel-Scale Imaging, *Laura Waller (UC Berkeley)*

**1000 Fast-Forward Presentations for Posters**

**1020 Morning Break** (Forum Ballroom) & **Poster Session**

- 1100 Depth Camera Based on Color-Coded Aperture, *Vladimir Paramonov, Ivan Panchenko, Victor Bucha, Andrey Drogolyub, Sergey Zagoruyko*
- 1120 SparkleGeometry: Glitter Imaging for 3D Point Tracking, *Abigail Stylianou, Robert Pless*
- 1140 Time-Offset Conversations on a Life-Sized Automultiscopic Projector Array, *Andrew Jones, Jonas Unger, Koki Nagano, Jay Busch, Xueming Yu, Hsuan-Yueh Peng, Joseph Barreto, Oleg Alexander, Mark Bolas, Paul Debevec*
- 1200 Lunch** (Forum Ballroom)
- 1300 **Keynote Talk:** A Systems Approach to Designing Computational Cameras, *Rajiv Loria (Light)*
- 1400 Avoiding the Deconvolution: Framework Oriented Color Transfer for Enhancing Low-Light Images, *Laura Florea, Corneliu Florea, Ciprian Ionaşcu*

- 1420 Power-Efficient Cameras Using Natural Image Statistics,  
*Roni Feldman, Yair Weiss, Yonina C. Eldar*
- 1440 Strategies for Resolving Camera Metamers Using 3+1  
Channel, *Dilip K. Prasad*
- 1500 **Afternoon Break** (Forum Ballroom) & **Poster Session**
- 1600 **Keynote Talk:** Computational Near-Eye Displays, *Gordon Wetzstein (Stanford Univ.)*
- 1700 **Closing Remarks and Best Paper Award**

## Medical Computer Vision

**Organizers:** Le Lu

Gustavo Carneiro  
Bjoern Menze  
Georg Langs  
Leo Grady

**Schedule:** Half Day — Morning

- 0810 **Invited Talk:** Deep Learning for Histopathology Image Analysis, *Pheng Ann Heng (Chinese Univ. of Hong Kong)*
- 0835 **Invited Talk:** Multi-Instance Deep Learning: Bridging Local Pattern Recognition to Holistic Image Understanding, *Yiqiang Zhan (Siemens Healthcare)*
- 0900 **Invited Talk:** Big Imaging-Genomics Data Analytics for Clinical Outcome Prediction, *Junzhou Huang (Univ. of Texas at Arlington)*
- 0925 **Invited Talk:** Applications of Discrete Optimization to Medical Imaging, *Ramin Zabih (Cornell Univ.)*
- 0950 **Invited Talk:** Clinician Detection and Pose Estimation in the Operating Room, *Nicolas Padoy (Univ. of Strasbourg)*
- 1015 **Morning Break** (Forum Ballroom)
- 1030 **Invited Talk:** Convolutional Neural Networks in Biomedical Imaging for Diagnosis, Therapy and Surgery, *Jianming Liang (Arizona State Univ.)*
- 1055 **Invited Talk:** Deep CNN in Radiology: Preventative and Precision Medicine Perspectives, *Le Lu (NIH)*
- 1120 **Invited Talk:** Classifying MRIs Based on Group Cardinality Constrained Solutions, *Killian Pohl (SRI)*
- 1145 **Invited Talk:** Robotics and Augmented Reality for Patient and Process Specific Imaging and Visualization, *Nassir Navab (Technische Universität München & Johns Hopkins Univ.)*
- 1210 **Invited Talk:** Discrete Methods for Fast Registration of Multi-Modal and Dynamic Imaging, *Julia Schnabel (King's College London)*

## Differential Geometry in Computer Vision and Machine Learning

**Organizers:** Hassen Drira

Sebastian Kurtek  
Anuj Srivastava  
Pavan Turaga

**Schedule:** Half Day — Morning

- 0810 **Welcome**
- 0815 **Keynote Talk:** *Baba C. Vemuri (Univ. of Florida)*

### S1: Oral Presentations (900-1000)

- 0900 The Assignment Manifold: A Smooth Model for Image Labeling, *Freddie Åström, Stefania Petra, Bernhard Schmitzer, Christoph Schnörr*
- 0915 A Statistical Framework for Elastic Shape Analysis of Spatio-Temporal Evolutions of Planar Closed Curves, *Chafik Samir, Sebastian Kurtek, Justin Strait, Shantanu H. Joshi*
- 0930 Testing Stationarity of Brain Functional Connectivity Using Change-Point Detection in fMRI Data, *Mengyu Dai, Zhengwu Zhang, Anuj Srivastava*
- 0945 Partial Matchings and Growth Mapped Evolutions in Shape Spaces, *Irène Kaltenmark, Alain Trouvé*
- 1000 **Morning Break** (Forum Ballroom)
- 1015 **Keynote Talk:** *Rama Chellappa (Univ. of Maryland)*

### 1045 Short Break

### S2: Poster Highlights (1050-1130)

- 1050 Human Object Interaction Recognition Using Rate-Invariant Shape Analysis of Inter Joint Distances Trajectories, *Meng Meng, Hassen Drira, Mohamed Daoudi, Jacques Boonaert*
- 1054 Riemannian Geometric Approaches for Measuring Movement Quality, *Anirudh Som, Rushil Anirudh, Qiao Wang, Pavan Turaga*
- 1058 Differential Geometry Boosts Convolutional Neural Networks for Object Detection, *Chu Wang, Kaleem Siddiqi*
- 1102 On Time-Series Topological Data Analysis: New Data and Opportunities, *Lee M. Seversky, Shelby Davis, Matthew Berger*
- 1106 A Riemannian Framework for Statistical Analysis of Topological Persistence Diagrams, *Rushil Anirudh, Vinay Venkataraman, Karthikeyan Natesan Ramamurthy, Pavan Turaga*

- 1110 A Survey on Rotation Optimization in Structure From Motion, *Roberto Tron, Xiaowei Zhou, Kostas Daniilidis*
- 1114 Bayesian Model-Based Automatic Landmark Detection for Planar Curves, *Justin Strait, Sebastian Kurtek*
- 1118 Consensus-Based Image Segmentation via Topological Persistence, *Qian Ge, Edgar Lobaton*
- 1122 Robust Domain Adaptation on the  $L_1$ -Grassmannian Manifold, *Sriram Kumar, Andreas Savakis*
- 1126 Fast Dynamic Programming for Elastic Registration of Curves, *Javier Bernal, Günay Doğan, Charles R. Hagwood*
- 1130 **Poster Session**
- 1225 **Best Paper Award and Closing Remarks**

## Robust Features for Computer Vision

**Organizers:** Jie Chen

Zhen Lei

Li Liu

Guoying Zhao

Matti Pietikäinen

**Schedule:** Half Day — Morning

0800 **Welcome**

### S1: Oral Session (810-950)

- 0810 **Invited Talk:** Convolutional Patch Representations for Image Retrieval, *Cordelia Schmid (INRIA)*
- 0850 Fast Image Gradients Using Binary Feature Convolutions, *Pierre-Luc St-Charles, Guillaume-Alexandre Bilodeau, Robert Bergevin*
- 0910 Texture Complexity Based Redundant Regions Ranking for Object Proposal, *Wei Ke, Tianliang Zhang, Jie Chen, Fang Wan, Qixiang Ye, Zhenjun Han*
- 0930 Deeply Exploit Depth Information for Object Detection, *Saihui Hou, Zilei Wang, Feng Wu*

### **0950 Morning Break (Forum Ballroom) & Poster Session**

- Efficient Deep Feature Learning and Extraction via StochasticNets, *Mohammad Javad Shafiee, Parthipan Siva, Paul Fieguth, Alexander Wong*
- Embedding Sequential Information Into Spatiotemporal Features for Action Recognition, *Yuancheng Ye, YingLi Tian*
- Learning Discriminative Features With Class Encoder, *Hailin Shi, Xiangyu Zhu, Zhen Lei, Shengcai Liao, Stan Z. Li*
- Do We Need Binary Features for 3D Reconstruction? *Bin Fan, Qingqun Kong, Wei Sui, Zhiheng Wang, Xinchao Wang, Shiming Xiang, Chunhong Pan, Pascal Fua*

- Deep Features or Not: Temperature and Time Prediction in Outdoor Scenes, *Anna Volokitin, Radu Timofte, Luc Van Gool*
- Euclidean and Hamming Embedding for Image Patch Description With Convolutional Networks, *Zishun Liu, Zhenxi Li, Juyong Zhang, Ligang Liu*
- Robust zDPCA and Its Application, *Qianqian Wang, Quanxue Gao*
- Background Subtraction Using Local SVD Binary Pattern, *Lili Guo, Dan Xu, Zhenping Qiang*
- Generating Discriminative Object Proposals via Submodular Ranking, *Yangmuzi Zhang, Zhuolin Jiang, Xi Chen, Larry S. Davis*
- Time Series Representation and Similarity Based on Local Autopatterns, *Mustafa Gökçe Baydoğan, George Runger*

### S2: Oral Session II (1050-1230)

- 1050 **Invited Talk:** Deep Features for Face Recognition and Related Problems, *Rama Chellappa (Univ. of Maryland)*
- 1130 Improving Gradient Histogram Based Descriptors for Pedestrian Detection in Datasets With Large Variations, *Prashanth Balasubramanian, Sathak Pathak, Anurag Mittal*
- 1150 Unsupervised Robust Feature-Based Partition Ensembling to Discover Categories, *Roberto J. López-Sastre*
- 1210 The Best of Both Worlds: Combining Data-Independent and Data-Driven Approaches for Action Recognition, *Zhenzhong Lan, Shou-I Yu, Dezhong Yao, Ming Lin, Bhiksha Raj, Alexander Hauptmann*
- 1230 **Concluding Remarks**

## Observing and Understanding Hands in Action

**Organizers:** Gregory Rogez  
Tae-Kyun Kim

**Schedule:** Half Day — Morning

0820 **Welcome**

0830 **Invited Talk:** Antonis Argyros (*ICS-FORTH*)

0900 **Invited Talk:** Jamie Shotton (*Microsoft Research Cambridge*)

0930 **Invited Talk:** Mohan Trivedi (*Univ. of California, San Diego*)

**1000 Morning Break** (Forum Ballroom)

1030 **Posters**

- Skeleton-Based Dynamic Hand Gesture Recognition, *Quentin De Smedt, Hazem Wannous, Jean-Philippe Vandeboor*
- Learning Marginalization Through Regression for Hand Orientation Inference, *Muhammad Asad, Gregory Slabaugh*
- Hidden Hands: Tracking Hands With an Occlusion Aware Tracker, *Akshay Rangesh, Eshed Ohn-Bar, Mohan M. Trivedi*
- Effectiveness of Grasp Attributes and Motion-Constraints for Fine-Grained Recognition of Object Manipulation Actions, *Kartik Gupta, Darius Burschka, Arnav Bhavsar*
- Understanding Everyday Hands in Action From RGB-D Images, *Gregory Rogez, James Steven Supancic III, Deva Ramanan*
- Opening the Black Box: Hierarchical Sampling Optimization for Estimating Human Hand Pose, *Danhang Tang, Jonathan Taylor, Pushmeet Kohli, Cem Keskin, Tae-Kyun Kim, Jamie Shotton*
- Deep Learning for Gesture Recognition in Automotive User Interfaces, *Pavlo Molchanov, Shalini Gupta, Xiaodong Yang, Kihwan Kim, Stephen Tyree, and Jan Kautz*
- Efficiently Creating 3D Training Data for Fine Hand Pose Estimation, *Markus Oberweger, Gernot Riegler, Paul Wohlhart, Vincent Lepetit*
- Hand Posture Recognition Based on Bottom-Up Structured Convolutional Neural Network With Curriculum Learning, *Takayoshi Yamashita, Hiroshi Fukui, Yuji Yamauchi, Hironobu Fujiyoshi*
- Analyzing Hands to Recognize Social Interactions With a Large-Scale Egocentric Hands Dataset, *Sven Bambach, Stefan Lee, David J. Crandall, Chen Yu*
- Robust 3D Hand Pose Estimation in Single Depth Images: From Single-View CNN to Multi-View CNNs, *Liuhao Ge, Hui Liang, Junsong Yuan, and Daniel Thalmann*
- Multi-Layered Random Forest-Based Metaphoric Hand Gesture Interface in Vr, *Youngkyoon Jang, Ikbeom Jeon, Tae-Kyun Kim, Woontack Woo*

- On- and Above-Skin Input Sensing Through a Wearable Depth Sensor, *Srinath Sridhar, Anders Markussen, Antti Oulasvirta, Christian Theobalt, Sebastian Boring*
- Deep Hand: How to Train a CNN on 1 Million Hand Images When Your Data Is Continuous and Weakly Labelled, *Oscar Koller, Hermann Ney, Richard Bowden*

1130 **Invited Talk:** Vincent Lepetit (*Graz Univ. of Technology*)

1200 **Invited Talk:** Deva Ramanan (*Carnegie Mellon Univ.*)

1230 **Closing & Prizes**

## Performance Evaluation of Tracking and Surveillance

**Organizers:** Luis Patino  
James Ferryman  
James L. Crowley

**Schedule:** Half Day — Morning

0830 **Welcome**

0840 **PETS 2016: Dataset and Challenge**, *Luis Patino, Tom Cane, Alain Vallee, James Ferryman*

0855 **Channel Coded Distribution Field Tracking for Thermal Infrared Imagery**, *Amanda Berg, Jörgen Ahlberg, Michael Felsberg*

0935 **Saliency-Based Detection for Maritime Object Tracking**, *Tom Cane, James Ferryman*

0935 **Robust Visual Tracking With Deep Convolutional Neural Network Based Object Proposals on PETS**, *Gao Zhu, Fatih Porikli, Hongdong Li*

0955 **Online Multi-Object Tracking Based on Hierarchical Association Framework**, *Jaeyong Ju, Daehun Kim, Bonhwa Ku, David K. Han, Hanseok Ko*

**1015 Morning Break** (Forum Ballroom)

1040 **Invited Talk:** Video Analytics in Pubic Safety (VAPS), *Jason Corso (Univ. of Michigan)*

1110 **Semantic Modelling for Behaviour Characterisation and Threat Detection**, *Luis Patino, James Ferryman*

1130 **Realtime Anomaly Detection Using Trajectory-Level Crowd Behavior Learning**, *Aniket Bera, Sujeong Kim, Dinesh Manocha*

1150 **Abnormal Event Recognition: A Hybrid Approach Using Semantic Web**, *Luca Greco, Pierluigi Ritrovato, Alessia Saggese, Mario Vento*

1210 **Evaluation on Submitted PETS2016 Results & Conclusions: Workshop Chairs**

## Computer Vision for Microscopy Image Analysis

**Organizers:** Mei Chen

Margrit Betke  
Weidong Cai  
An-An Liu  
Jens Rittscher  
Zhaozheng Yin  
Shaoting Zhang

**Schedule:** Half Day — Afternoon

1300 **Welcome**

1310 **Keynote:** Dimitri Metaxas (*Rutgers Univ.*)

1400 **Spotlight Presentations**

- Spatially Aware Dictionary Learning and Coding for Fossil Pollen Identification, *Shu Kong, Surangi Punyasena, Charless Fowlkes*
- Multi-View Multi-Modal Feature Embedding for Endomicroscopy Mosaic Classification, *Yun Gu, Jie Yang, Guang-Zhong Yang*
- Neuron Segmentation Based on CNN With Semi-Supervised Regularization, *Kun Xu, Hang Su, Jun Zhu, Ji-Song Guan, Bo Zhang*
- 3D Structure Modeling of Dense Capillaries by Multi-Objects Tracking, *Ryoma Bise, Imari Sato, Kentaro Kajiya, Toyonobu Yamashita*
- Analysing the Structure of Collagen Fibres in SBFSEM Images, *Yassar Almutairi, Timothy Cootes, Karl Kadler*
- Cluster Sensing Superpixel and Grouping, *Rui Li, Lu Fang*
- 3D Convolutional Networks-Based Mitotic Event Detection in Time-Lapse Phase Contrast Microscopy Image Sequences of Stem Cell Populations, *Wei-Zhi Nie, Wei-Hui Li, An-An Liu, Tong Hao, Yu-Ting Su*
- Segmentation of Overlapping Cervical Cells in Microscopic Images With Superpixel Partitioning and Cell-Wise Contour Refinement, *Hansang Lee, Junmo Kim*
- Unsupervised Segmentation of Cervical Cell Images Using Gaussian Mixture Model, *Srikanth Ragothaman, Sridharakumar Narasimhan, Madivala G. Basavaraj, Rajan Dewar*
- ICORD: Intelligent Collection of Redundant Data - A Dynamic System for Crowdsourcing Cell Segmentations Accurately and Efficiently, *Sameki Mehrnoosh, Danna Gurari, Margrit Betke*
- Four Dimensional Image Registration For Intravital Microscopy, *Chichen Fu, Neeraj Gadgil, Khalid K. Tahboub, Paul Salama, Kenneth W. Dunn, Edward J. Delp*

- Methodology for Increasing the Measurement Accuracy of Image Features, *Michael Majurski, Joe Chalfoun, Steven P. Lund, Peter Bajcsy, Mary Brady*

1500 **Afternoon Break (Forum Ballroom) & Poster Session**

1600 **Postdoc Colloquium**

- Danna Gurari (*Univ. of Texas Austin*)
- Hang Su (*Tsinghua Univ.*)

1700 **Invited Talks**

- Dan Hoepfner (*Lieber Inst. for Brain Development*)
- Jens Rittscher (*Oxford Univ.*)

1800 **Closing Remarks**

## Moving Cameras Meet Video Surveillance: From Body-Borne Cameras to Drones

**Organizers:** Rogerio Feris  
Quanfu Fan  
Steve Mann

**Schedule:** Half Day — Afternoon

1330 **Welcome**

1340 **Keynote:** Surveillance (Oversight), Sousveillance (Undersight), and Metaveillance (Seeing Sight Itself), *Steve Mann (Univ. of Toronto)*

1430 Covert Video Classification by Codebook Growing Pattern, *Liang Du, Haitao Lang, Ying-Li Tian, Chiu C. Tan, Jie Wu, Haibin Ling*

1440 Detecting Anomalous Objects on Mobile Platforms, *Wallace Lawson, Laura Hiatt, Keith Sullivan*

1450 **Keynote:** The Role of Computer Vision in Public Safety and Vice Versa, *Jason Corso (Univ. of Michigan)*

1530 **Afternoon Break (Forum Ballroom)**

1600 **Keynote:** Summarizing Long First-Person Videos, *Kristen Grauman (Univ. of Texas at Austin)*

1640 Robust Detection of Moving Vehicles in Wide Area Motion Imagery, *Michael Teutsch, Michael Grinberg*

1650 Real-Time Vehicle Tracking in Aerial Video Using Hyperspectral Features, *Burak Uzkent, Matthew J. Hoffman, Anthony Vodacek*

1700 **Concluding Remarks**

## Context-Based Affect Recognition and Affective Face In-the-Wild

**Organizers:** Zakia Hammal

Merlin Teodosia Suarez

Ognjen Rudovic

Irene Kotsia

Mihalis Nicolaou

Guoying Zhao

Stefanos Zafeiriou

**Schedule:** Half Day — Afternoon

1330 **Welcome**

### S1: CBAR (1330-1515)

**Chairs:** Zakia Hammal, Merlin Teodosia Suarez, Ognjen Rudovic

1330 **Invited Talk:** Consensus Bayesian Models for Analysis of Distributed Spatio-Temporal Processes: Human Affect, Crowds, and Beyond, *Vladimir Pavlovic (Rutgers Univ.)*

1415 **Extended DISFA Dataset:** Investigating Posed and Spontaneous Facial Expressions, *Mohammad Mavadati, Peyton Sanger, Mohammad H. Mahoor*

1430 **A Framework for Joint Estimation and Guided Annotation of Facial Action Unit Intensity,** *Robert Walecki, Ognjen Rudovic, Maja Pantic, Vladimir Pavlovic, Jeffrey F. Cohn*

1445 **Gaussian Process Domain Experts for Model Adaptation in Facial Behavior Analysis,** *Stefanos Eleftheriadis, Ognjen Rudovic, Marc P. Deisenroth, Maja Pantic*

1500 **Automatic Recognition of Emotions and Membership in Group Videos,** *Wenxuan Mou, Hatice Gunes, Ioannis Patras*

### S2: WILD (1515-1725) (Papers in this session are also in the Poster Session)

**Chairs:** Irene Kotsia, Mihalis Nicolaou, Guoying Zhao, Stefanos Zafeiriou

1515 **Facial Affect "In-The-Wild": A Survey and a New Database,** *Stefanos Zafeiriou, Athanasios Papaioannou, Irene Kotsia, Mihalis Nicolaou, Guoying Zhao*

1530 **Afternoon Break** (Forum Ballroom)

1555 **Invited Talk:** The Face of Emotion: From Human Perception to Computer Vision and Back, *Alex Martinez (Ohio State Univ.)*

1640 **Fusing Aligned and Non-Aligned Face Information for Automatic Affect Recognition in the Wild: A Deep Learning Approach,** *Bo-Kyeong Kim, Suh-Yeon Dong, Jihyeon Roh, Geonmin Kim, Soo-Young Lee*

1655 **Facial Expression Recognition from World Wild Web,** *Ali Mollahosseini, Behzad Hasani, Michelle J. Salvador, Hojjat Abdollahi, David Chan, Mohammad H. Mahoor*

1710 **Facial Expression Recognition in the Wild Using Improved Dense Trajectories and Fisher Vector Encoding,** *Sadaf Afshar, Albert Ali Salah*

### 1725 **Poster Session**

- **Towards An "In-The-Wild" Emotion Dataset Using a Game-Based Framework,** *Wei Li, Farnaz Abtahi, Christina Tsangouri, Zhigang Zhu*

- **Recurrent Convolutional Neural Network Regression for Continuous Pain Intensity Estimation in Video,** *Jing Zhou, Xiaopeng Hong, Fei Su, Guoying Zhao*

- **Towards Facial Expression Recognition in the Wild: A New Database and Deep Recognition System,** *Xianlin Peng, Zhaoqiang Xia, Lei Li, Xiaoyi Peng*

- **A 3D Mask Face Anti-Spoofing Database With Real World Variations,** *Siqi Liu, Baoyao Yang, Pong C. Yuen, Guoying Zhao*

- **Sequential Face Alignment via Person-Specific Modeling in the Wild,** *Xi Peng, Junzhou Huang, Dimitris N. Metaxas*

1825 **Closing Remarks**

## ActivityNet Large Scale Activity Recognition Challenge

**Organizers:** Cees Snoek

Bernard Ghanem

Juan Carlos Niebles

Fabian Caba Heilbron

Wayner Barrios

**Schedule:** Half Day — Afternoon

1330 **Welcome**

1340 **Invited Talk:** Grouping Process Models in Actor-Action Segmentation, *Jason Corso, (Univ. of Michigan)*

1410 **Invited Talk:** Human Pose Estimation and Activity Recognition, *Bernt Schiele (Max-Planck-Institut für Informatik)*

1440 **ActivityNet: Dataset, Challenge Tasks and Results**

1500 **Afternoon Break** (Forum Ballroom)

1515 **Classification Task - Challenge Participant Talks**

1615 **Detection Task - Challenge Participant Talk**

1715 **Closing Remarks**



## Silver Donors

---



## Bronze Donors

---



## Startup Donors

---



## Platinum Donors



## Gold Donors

