Welcome to the 2016 edition of the European Conference on Computer Vision in Amsterdam!

It is safe to say that the European Conference on Computer Vision is one of the top conferences in computer vision. It is good to reiterate the history of the conference to see the broad base the conference has built in the thirteen editions. First held in 1990 in Antibes (France), it was followed with subsequent conferences in Santa Margherita Ligure (Italy) in 1992, Stockholm (Sweden) in 1994, Cambridge (UK) in 1996, Freiburg (Germany) in 1998, Dublin (Ireland) in 2000, Copenhagen (Denmark) in 2002, Prague (Czech Republic) in 2004, Graz (Austria) in 2006, Marseille (France) in 2008, Heraklion (Greece) in 2010, Florence (Italy) in 2012, and Zürich (Switzerland) in 2014.

The program chairs, Jiri Matas, Nicu Sebe, Bastian Leibe and Max Welling, are delighted to present a strong and exciting program, the result of an extensive review process. In total, they received 1,561 paper submissions. Of these, 81 violated the ECCV submission guidelines or did not pass the iThenticate-plagiarism test and were rejected without review. Of the remaining papers, 415 were accepted (26.6 %): 342 as posters (22.6 %), 45 as spotlights (2.9 %), and 28 as oral presentations (1.8 %). The spotlights are novel to ECCV and were introduced after their success at the CVPR 2016 conference. All orals and spotlights are presented as posters as well. The selection process was a combined effort of four program co-chairs (PCs), 74 area chairs (ACs), 1,086 Program Committee members, and 77 additional reviewers. We thank the program chairs, area chairs and reviewers.

For the fourteenth edition, many people have worked hard to provide you with a most warm welcome while enjoying the best science. The program committee, Bastian Leibe RWTH Aachen, Jiri Matas TUPrague, Nicu Sebe UTrento, and Max Welling UvA, has done an excellent job. They provide a preface in the next pages. We are most grateful to them. In addition, Albert Ali Salah UBogazici and Robby Tan USingapore were handling all accepted papers into the proceedings. Apart from the scientific program, the workshops have been selected and handled by Herve Jegou Facebook and Gang Hua Microsoft, and the tutorials by Jacob Verbeek INRIA and Rita Cucchiara UModena. Thanks for the persistent job done. The coordination with ACM Multimedia 2016 Amsterdam offered an opportunity to expand the tutorials with an additional invited session, offered by the University of Amsterdam.
Of the many people who have worked hard as local organizers, we would like to single out Martine de Wit, Niels Klein, Melanie Venverloo and others at the UvA Conference Office, who have delicately, as well as efficiently, organised the main body of work. Also local organizers Hamdi Dibeklioglu TUDelft, Efstratios Gavves UvA, Jan van Gemert TUDelft, Thomas Mensink UvA and Mihir Jain Qualcomm NL had their hands full. As a main venue, we have the Royal Theatre Carre located on the canals of the Amstel river downtown Amsterdam. Space in Amsterdam is a sparse good so it will be a little tighter than usual. The university has lent us their downtown campuses for the tutorials and the workshops, supplemented with a downtown theater in art deco style and a hotel for some of the workshops. A relatively new thing is the industry and the sponsors for which Ronald Poppe and Peter de With have done a great job, while Andy Bagdanov U Florence and John Schavemaker TNO arranged the demos. Michael Wilkinson has taken care to make Yom Kippur as comfortable as possible for those for whom it is an important day. We thank Marc Pollefeys ETH, Alberto del Bimbo U Florence and Virginie Mes UvA for their advice and help behind the scenes. We thank all the anonymous volunteers for their hard and precise work.

We also thank our generous sponsors. Their support is an essential part of the program. It is good to see such a level of industrial interest in what our community is doing!

Amsterdam does not need any introduction. Please emerge yourself but do not drown in it, have a nice time!

October 2016

Also on behalf of
Jiri Matas, Nicu Sebe, Bastian Leibe and Max Welling, program chairs,
Theo Gevers and Arnold Smeulders, general chairs
6 Time Schedule
ECCV’16
Sunday, October 9
Monday, October 10
Tuesday, October 11
Wednesday, October 12
Thursday, October 13
Friday, October 14
Saturday, October 15
Sunday, October 16

13 Practical Information
ECCV’16
15 Venue Information
18 Practical Matters
19 Presenter Instructions
19 Oral Sessions
20 Spotlight Sessions
22 Poster Sessions
23 Demo Sessions
24 Proceedings
26 Restaurants
26 Cheap / on the go / coffee
26 Midprice / seated lunch
27 Luxurious / extensive

28 Saturday, October 8 | Oudemanhuispoort
Workshops
28 W1 Datasets and Performance Analysis in Early Vision
29 W2 Visual Analysis of Sketches
30 W3 Biological and Artificial Vision
31 W4 Brave New Ideas For Motion Representations in Videos

Regular Tutorials
33 Geometric deep learning
33 Zero-shot learning for vision and multimedia
34 New directions in saliency research: Developments in architecture, datasets and evaluation
34 Deep learning meets model optimization and statistical inference
35 Computational geometry tools and applications in computer vision
35 Distance metric learning for computer vision
36 **Sunday, October 9** | Oudemanhuispoort / NH Hotel Krasnapolsky  
Workshops  
36 W5 Joint Imagenet and MS Coco Visual Recognition Challenge Workshop  
37 W6 Geometry Meets Deep Learning  
39 W7 Action and Anticipation for Visual Learning  
40 W8 Computer Vision for Road Scene Understanding and Autonomous Driving  
41 W9 Challenge on Automatic Personality Analysis  
43 W10 BioImage Computing  
44 W11 Benchmarking Multi-Target Tracking: MOTChallenge  
45 W12 Assistive Computer Vision and Robotics  
46 W13 Transferring and Adapting Source Knowledge in Computer Vision  
47 W14 Recovering 6D Object Pose  
50 W15 Robust Reading  
51 W16 3D Face Alignment in the Wild & Challenge  
52 W17 Egocentric Perception, Interaction and Computing  

54 **Monday, October 10** | Theatre Tuschinski  
Workshops  
54 W18 Local Features: State of the art, open problems and performance evaluation  
55 W19 Crowd Understanding  
56 W20 Video Segmentation  
57 W21 The Visual Object Tracking Challenge Workshop  
58 W22 Web-scale Vision and Social Media  

60 **Tuesday, October 11** | Theatre Carré  
Main conference + Welcome drinks  

68 **Wednesday, October 12** | Theatre Carré  
Main conference  

75 **Thursday, October 13** | Theatre Carré  
Main conference + Conference dinner  

85 **Friday, October 14** | Theatre Carré  
Main conference  

93 **Saturday, October 15** | Roeterseiland  
Invited Tutorials  
93 Segmentation and Tracking  
93 Variational wrapping deformations for image processing applications  
93 Developments in deep learning (tentative)  
93 Multimodal human-robot interaction  
93 Output embedding for large-scale visual recognition  
93 Video event detection, localization and cross-media linking  
93 Human body shape modelling  
93 A guide to creating and managing lifelog collections  
93 Deep visual understanding from deep learning  
93 Bridging video and language with deep learning  
93 Color and spectral information in modern computer vision
94 **Sunday, October 16**

**Workshops**

94 W23 Computer Vision for Audio-visual Media
95 W24 Computer Vision for ART Analysis
96 W25 Virtual/Augmented Reality for Visual Artificial Intelligence
97 W26 Joint Workshop on Storytelling with Images and Videos and Large Scale Movie Description and Understanding Challenge

98 **Workshops and Tutorials**

ACM Multimedia 2016
### Saturday, October 8 / Oudemanhuispoort

<table>
<thead>
<tr>
<th>Time</th>
<th>Event Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:00</td>
<td>Registration</td>
</tr>
<tr>
<td>08:30</td>
<td>Workshop 1 (D0.08) - Datasets and Performance Analysis in Early Vision (p.28)</td>
</tr>
<tr>
<td>09:00</td>
<td>Workshop 2 (D0.09) - Visual Analysis of Sketches (p.29)</td>
</tr>
<tr>
<td>09:30</td>
<td>Workshop 3 (F0.01) - Biological and Artificial Vision (p.30)</td>
</tr>
<tr>
<td>10:00</td>
<td>Workshop 4 (D1.09) - Brave New Ideas For Motion Representations in videos (p.31)</td>
</tr>
<tr>
<td>10:30</td>
<td>Regular tutorial (D1.08) - Geometric deep learning (p.33)</td>
</tr>
<tr>
<td>11:00</td>
<td>Regular tutorial (C1.17) - Zero-shot learning for vision and multimedia (p.33)</td>
</tr>
<tr>
<td>11:30</td>
<td>Regular tutorial (A0.08) - New objectives in early vision research: Developments in architectures, datasets and evaluation (p.33)</td>
</tr>
</tbody>
</table>

### Sunday, October 9 / Oudemanhuispoort / NH Hotel Krasnapolsky

<table>
<thead>
<tr>
<th>Time</th>
<th>Event Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:00</td>
<td>Registration</td>
</tr>
<tr>
<td>08:30</td>
<td>Workshop 5 (Ruby, NH Kras.) - Joint Imagenet and MS Coco Visual Recognition Challenge Workshop (p.36)</td>
</tr>
<tr>
<td>09:00</td>
<td>Workshop 6 (Diamant, NH Kras.) - Geometry Meets Deep Learning (p.37)</td>
</tr>
<tr>
<td>09:30</td>
<td>Workshop 7 (D0.09) - Action and Anticipation for Visual Learning (p.39)</td>
</tr>
<tr>
<td>10:00</td>
<td>Workshop 8 (D0.08) - Computer Vision for Road Scene Understanding and Autonomous Driving (p.40)</td>
</tr>
<tr>
<td>10:30</td>
<td>Workshop 9 (C0.23) - Challenge on Automatic Personality Analysis (p.41)</td>
</tr>
<tr>
<td>11:00</td>
<td>Workshop 10 (C1.17) - BiImage Computing (p.43)</td>
</tr>
<tr>
<td>11:30</td>
<td>Workshop 11 (F0.01) - Benchmarking Multi-Target Tracking: MOTChallenge (p.44)</td>
</tr>
<tr>
<td>12:00</td>
<td>Workshop 12 (C0.17) - Assistive Computer Vision and Robotics (p.45)</td>
</tr>
<tr>
<td>12:30</td>
<td>Workshop 13 (D1.09) - Transferring and Adapting Source Knowledge in Computer Vision (p.46)</td>
</tr>
<tr>
<td>13:00</td>
<td>Workshop 14 (A0.08) - Recovering 3D Object Pose (p.47)</td>
</tr>
<tr>
<td>13:30</td>
<td>Workshop 15 (C2.23) - Robust Reading (p.50)</td>
</tr>
<tr>
<td>14:00</td>
<td>Workshop 16 (C2.17) - 3D Face Alignment in the Wild &amp; Challenge (p.51)</td>
</tr>
<tr>
<td>14:30</td>
<td>Workshop 17 (C3.17) - Egocentric Perception, Interaction and Computing (p.52)</td>
</tr>
</tbody>
</table>
### Monday, October 10 / Oudemanhuispoort / Theatre Tuschinski

<table>
<thead>
<tr>
<th>Hall</th>
<th>Registration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Workshop 18 Room 1</td>
<td>Local Features: State of the art, open problems and performance evaluation (p.54)</td>
</tr>
<tr>
<td>Workshop 19 Room 5</td>
<td>Crowd Understanding (p.55)</td>
</tr>
<tr>
<td>Workshop 20 Room 2</td>
<td>Video Segmentation (p.56)</td>
</tr>
<tr>
<td>Workshop 21 Room 3</td>
<td>The Visual Object Tracking Challenge Workshop (p.57)</td>
</tr>
<tr>
<td>Workshop 22 Room 4</td>
<td>Webscale Vision and Social Media (p.58)</td>
</tr>
</tbody>
</table>

### Tuesday, October 11 / Theatre Carré

<table>
<thead>
<tr>
<th>Theatre</th>
<th>Oral Session 1A – Detection, Recognition &amp; Retrieval (p.60)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foyer</td>
<td>Registration</td>
</tr>
<tr>
<td>Roof Garden</td>
<td>Welcome Reception</td>
</tr>
<tr>
<td></td>
<td>Coffee Break</td>
</tr>
<tr>
<td></td>
<td>Lunch Break</td>
</tr>
<tr>
<td></td>
<td>Coffee Break</td>
</tr>
<tr>
<td></td>
<td>Coffee Break</td>
</tr>
<tr>
<td></td>
<td>Welcome Reception</td>
</tr>
</tbody>
</table>

### Wednesday, October 12 / Theatre Carré

<table>
<thead>
<tr>
<th>Theatre</th>
<th>Oral Session 2A – Optimization (p.68)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foyer</td>
<td>Registration</td>
</tr>
<tr>
<td>Roof Garden</td>
<td>Oral Session 2B – Image &amp; Video Processing (p.71)</td>
</tr>
<tr>
<td></td>
<td>Lunch Break</td>
</tr>
<tr>
<td></td>
<td>Coffee Break</td>
</tr>
<tr>
<td></td>
<td>Coffee Break</td>
</tr>
<tr>
<td></td>
<td>Coffee Break</td>
</tr>
</tbody>
</table>

### Thursday, October 13 / Theatre Carré & Ocean Diva

<table>
<thead>
<tr>
<th>Theatre</th>
<th>Oral Session 3A – Learning (p.75)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foyer</td>
<td>Registration</td>
</tr>
<tr>
<td>Roof Garden</td>
<td>Spotlight Session 3A &amp; Structural Motion (p.78)</td>
</tr>
<tr>
<td>Ocean Diva (PTA)</td>
<td>Demos 48-49 (p.81)</td>
</tr>
<tr>
<td></td>
<td>Coffee Break</td>
</tr>
</tbody>
</table>
### Friday, October 14 / Theatre Carré

<table>
<thead>
<tr>
<th>Time</th>
<th>Hall</th>
<th>Activity</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:00</td>
<td>REC E</td>
<td>Registration</td>
<td>Theatre Carré</td>
</tr>
<tr>
<td>08:30</td>
<td>REC E</td>
<td>Coffee Break</td>
<td>Theatre Carré</td>
</tr>
<tr>
<td>09:00</td>
<td>REC E</td>
<td>Oral Session 4A, 1B, (p.85)</td>
<td>Theatre Carré</td>
</tr>
<tr>
<td>09:30</td>
<td>REC E</td>
<td>Spotlight Session 4A, Feature Extraction, &amp; Matching (p.85)</td>
<td>Theatre Carré</td>
</tr>
<tr>
<td>10:00</td>
<td>REC E</td>
<td>Poster Session 4A – poster 1-47 (p.85)</td>
<td>Theatre Carré</td>
</tr>
<tr>
<td>10:30</td>
<td>REC E</td>
<td>Lunch Break</td>
<td>Theatre Carré</td>
</tr>
<tr>
<td>11:00</td>
<td>REC E</td>
<td>Oral Session 4B, Action, Activity &amp; Tracking (p.85)</td>
<td>Theatre Carré</td>
</tr>
<tr>
<td>11:30</td>
<td>REC E</td>
<td>Spotlight Session 4B, Human &amp; Face (p.85)</td>
<td>Theatre Carré</td>
</tr>
<tr>
<td>12:00</td>
<td>REC E</td>
<td>Poster Session 4B – poster 1-4 (p.85)</td>
<td>Theatre Carré</td>
</tr>
<tr>
<td>12:30</td>
<td>REC E</td>
<td>Coffee Break</td>
<td>Theatre Carré</td>
</tr>
<tr>
<td>13:00</td>
<td>REC E</td>
<td>Invited tutorials C1.03</td>
<td>Theatre Carré</td>
</tr>
<tr>
<td>13:30</td>
<td>REC E</td>
<td>Segmentation and Tracking (p.93)</td>
<td>Theatre Carré</td>
</tr>
<tr>
<td>14:00</td>
<td>REC E</td>
<td>Developments in deep learning (p.93)</td>
<td>Theatre Carré</td>
</tr>
<tr>
<td>14:30</td>
<td>REC E</td>
<td>Output embedding for large-scale visual recognition (p.93)</td>
<td>Theatre Carré</td>
</tr>
<tr>
<td>15:00</td>
<td>REC E</td>
<td>Human body shape modelling (p.93)</td>
<td>Theatre Carré</td>
</tr>
<tr>
<td>15:30</td>
<td>REC E</td>
<td>Deep visual understanding from deep learning (p.93)</td>
<td>Theatre Carré</td>
</tr>
<tr>
<td>16:00</td>
<td>REC E</td>
<td>Color and spectral information in modern computer vision (p.93)</td>
<td>Theatre Carré</td>
</tr>
<tr>
<td>16:30</td>
<td>REC E</td>
<td>Variational swapping deformations for image processing applications (p.93)</td>
<td>Theatre Carré</td>
</tr>
<tr>
<td>17:00</td>
<td>REC E</td>
<td>Multimodal human-robot interaction (p.93)</td>
<td>Theatre Carré</td>
</tr>
<tr>
<td>17:30</td>
<td>REC E</td>
<td>Video event detection, localization and cross-media linking (p.93)</td>
<td>Theatre Carré</td>
</tr>
<tr>
<td>18:00</td>
<td>REC E</td>
<td>A guide to creating and managing lifelog collections (p.93)</td>
<td>Theatre Carré</td>
</tr>
<tr>
<td>18:30</td>
<td>REC E</td>
<td>Bridging video and language with deep learning (p.93)</td>
<td>Theatre Carré</td>
</tr>
</tbody>
</table>

### Saturday, October 15 / Roeterseiland

<table>
<thead>
<tr>
<th>Time</th>
<th>Hall</th>
<th>Activity</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:00</td>
<td>REC E</td>
<td>Registration</td>
<td>Roeterseiland</td>
</tr>
<tr>
<td>08:30</td>
<td>REC E</td>
<td>Coffee and drinks are served all day</td>
<td>Roeterseiland</td>
</tr>
<tr>
<td>09:00</td>
<td>REC E</td>
<td>Invited tutorials C1.03</td>
<td>Roeterseiland</td>
</tr>
<tr>
<td>09:30</td>
<td>REC E</td>
<td>Segmentation and Tracking (p.93)</td>
<td>Roeterseiland</td>
</tr>
<tr>
<td>10:00</td>
<td>REC E</td>
<td>Developments in deep learning (p.93)</td>
<td>Roeterseiland</td>
</tr>
<tr>
<td>10:30</td>
<td>REC E</td>
<td>Output embedding for large-scale visual recognition (p.93)</td>
<td>Roeterseiland</td>
</tr>
<tr>
<td>11:00</td>
<td>REC E</td>
<td>Human body shape modelling (p.93)</td>
<td>Roeterseiland</td>
</tr>
<tr>
<td>11:30</td>
<td>REC E</td>
<td>Deep visual understanding from deep learning (p.93)</td>
<td>Roeterseiland</td>
</tr>
<tr>
<td>12:00</td>
<td>REC E</td>
<td>Color and spectral information in modern computer vision (p.93)</td>
<td>Roeterseiland</td>
</tr>
<tr>
<td>12:30</td>
<td>REC E</td>
<td>Variational swapping deformations for image processing applications (p.93)</td>
<td>Roeterseiland</td>
</tr>
<tr>
<td>13:00</td>
<td>REC E</td>
<td>Multimodal human-robot interaction (p.93)</td>
<td>Roeterseiland</td>
</tr>
<tr>
<td>13:30</td>
<td>REC E</td>
<td>Video event detection, localization and cross-media linking (p.93)</td>
<td>Roeterseiland</td>
</tr>
<tr>
<td>14:00</td>
<td>REC E</td>
<td>A guide to creating and managing lifelog collections (p.93)</td>
<td>Roeterseiland</td>
</tr>
<tr>
<td>14:30</td>
<td>REC E</td>
<td>Bridging video and language with deep learning (p.93)</td>
<td>Roeterseiland</td>
</tr>
<tr>
<td>15:00</td>
<td>REC E</td>
<td>Invited tutorials C1.03</td>
<td>Roeterseiland</td>
</tr>
<tr>
<td>15:30</td>
<td>REC E</td>
<td>Coffee and drinks are served all day</td>
<td>Roeterseiland</td>
</tr>
<tr>
<td>16:00</td>
<td>REC E</td>
<td>Invited tutorials C1.03</td>
<td>Roeterseiland</td>
</tr>
<tr>
<td>16:30</td>
<td>REC E</td>
<td>Segmentation and Tracking (p.93)</td>
<td>Roeterseiland</td>
</tr>
<tr>
<td>17:00</td>
<td>REC E</td>
<td>Developments in deep learning (p.93)</td>
<td>Roeterseiland</td>
</tr>
<tr>
<td>17:30</td>
<td>REC E</td>
<td>Output embedding for large-scale visual recognition (p.93)</td>
<td>Roeterseiland</td>
</tr>
<tr>
<td>18:00</td>
<td>REC E</td>
<td>Human body shape modelling (p.93)</td>
<td>Roeterseiland</td>
</tr>
<tr>
<td>18:30</td>
<td>REC E</td>
<td>Deep visual understanding from deep learning (p.93)</td>
<td>Roeterseiland</td>
</tr>
</tbody>
</table>
Practical Information

ECCV’16

About Amsterdam
Amsterdam is one of the greatest small cities in the world. From Amsterdam canals to world-famous Amsterdam museums and historical Amsterdam sights, it is one of the most romantic and beautiful cities in Europe. Canal cruises are a popular way to see the city from the perspective of its canals.
Amsterdam is also a city of tolerance and diversity. It has all the advantages of a big city: rich culture, lively Amsterdam nightlife, international restaurants, good transport - but is quiet, and largely thanks to its extensive canals, has a little road traffic. In this city your destination is never far away, but get a bike for an authentic local experience.

Important.
ECCV’16 is sold out.

Your badges serves as your entrance ticket.

No badge is no entrance (If you have a full registration you also need your ticket for the conference dinner).
Practical Information

Venue Information

The ECCV’16 conference will take place at the locations described below. A map showing all locations is provided.

A

**Oudemanhuispoort**

*(Faculty of Law, University of Amsterdam)*

**For**

Workshops and Regular Tutorials (no lunch)

Conference Registration (Saturday 8th and Sunday 9th October, from 8am)

**Address**

Oudemanhuispoort 4-6  
1012 CN Amsterdam

**Directions**

To get to the Oudemanhuispoort from the Amsterdam Central Station, you can use tramlines 4, 9, 16, 24, or 25.

Get off at the second stop (*Spui*). The street you are on now is called the *Rokin*. Take a left and cross the bridge to enter the alley *Langebrugsteeg*. Follow it through to *Grimburgwal*. At the end of *Grimburgwal*, take a left onto the canal named *Oudezijds Achterburgwal*. Immediately on your right-hand side you will see the gate to the *Oudemanhuispoort*.

Go through the gate and go halfway down the passage. On your left hand side, you will find the courtyard with the main entrance to the Faculty of Law.

B

**NH Hotel Krasnapolsky**

**For**

Workshops W5 + W6

Conference Registration (Monday 10th October, from 8am)

**Address**

Dam 9  
1012 JS Amsterdam

**Directions**

To get to the NH Hotel Krasnapolsky from the Amsterdam Central Station, you can use tramlines 4, 9, 16, 24, or 25.

Get off at the first stop (*Dam*). The square you are at is called *Dam*.

C

**Theatre Tuschinski**

*(Pathé Tuschinski is a movie theater in the Netherlands in Amsterdam commissioned by Abraham Icek Tuschinski in 1921 at a cost of 4 million guilders. The interior and exterior are a spectacular mix of styles, as designed by Hijman Louis de Jong; Amsterdam School, Jugendstil, Art Nouveau and Art Deco. The main auditorium hosts many premieres of Dutch films. It is considered to be one of the most beautiful cinemas in the world.)*
**Workshops**

*Conference Registration (Monday 10th October, from 8am)*

**Address**
Reguliersbreestraat 26-34
1017 CN Amsterdam

**Directions**
To get to the Theatre Tuschinski from the Amsterdam Central Station, you can use tramlines 4, 9, 16, 24, or 25.

Get off at the second stop (Spui). The street you are on now is called the Rokin. Take a right and cross the 'Muntplein’ to enter the alley Reguliersbreestraat.

---

**Theatre Carré**
*(Carré opened its doors for the first time on 3 December 1887. The majestic building on the banks of the Amstel, built as a circus theatre, grew to become the country’s most famous theatre and venue for cabaret, theatre, music, ballet, musicals and, happily, still for the circus as well. After 126 years, the history of Carré is still very palpable while, at the same time, there is, behind the scenes, an ultramodern theatre.)*

---

**Main conference**

*Conference Registration (Tuesday 11th to Friday 14th October, from 8am)*

**Address**
Amstel 115 / 125
1018 EM Amsterdam

**Directions**
To get to the Theatre Carre from the Amsterdam Central Station, you can use all the metro's.

Head north on Rhijnspoorplein toward Sarphatistraat, Continue onto Weesperplein, Turn left onto Voormalige Stadstimmertuin and Turn right onto Amstel.

---

**Roeterseiland**
*(Faculty of Economics and Business / Faculty of Social and Behavioural Sciences)*

**Workshops and invited tutorials**

*Conference Registration (Saturday 15th and Sunday 16th October, from 8am)*

**Address**
Roeterstraat 11
1018 WB Amsterdam

**Directions**
To get to the Roeterseiland from the Amsterdam Central Station, you can use all the metro’s.

Get off at the third stop (Weesperplein).
Head north on Rhijnspoorplein toward Weesperplein, Turn right at Weesperplein,
Turn left toward Sarphatistraat, Turn right onto Sarphatistraat and Turn left onto Roetersstraat
Destination will be on the left.
**Ocean Diva (Conference Dinner)**
(Venue: Passenger Terminal Amsterdam (PTA))

**For**

**Conference Dinner (Thursday 13th October, from 6pm.**
The ship sails between 8 and 9pm)

Your badge is your entrance ticket

**Address**
Piet Heinkade 27
1019 BR Amsterdam

**Directions**
There are several ways to reach the conference dinner, we highlight 3 of them:

1. Walk, it is a pleasant 2.5 km (30 minutes) walk between the conference venue and the PTA.
2. Public Transport:
   - First get to Central Station, take the metro from *Weesperplein*.
   - Then take tram 26, and exit on the first stop (*Muziekgebouw Bimhuis*) or walk from Central Station (about 1 km)
3. Taxi - Use the Uber app, call +31 20 777 77 77, or ask (wie) to help to get you a cab.
More information about Amsterdam can be found at: www.iamsterdam.nl.

Safety & Crime
Amsterdam is considered to be a relatively safe city. In general, violent crimes are as rare as in other European Cities. However, as in most crowded cities, beware of pick-pockets; they are mainly active in crowded areas, and on the train connecting Amsterdam Central Station and Schiphol Airport. Watch your belongings carefully.

National Emergency numbers / Police
Police, fire brigade, ambulance: National: +31 112
Police, no emergency: +31 900 8844

Electricity
Within the Netherlands 230 Volts/50 Hertz electrical supply is used. Visitors may need converters and adapters.

Cloak room/luggage deposit
Is available at all venues.

Lost and found
There is a lost and found service at the registration desk in the hall of all the locations.

Lunch
At some days we are offering lunch bags. Please check the program.

Questions and requests
Volunteer assistants – wearing an orange ECCV-shirt – will be present at the registration desk conference rooms. They will provide assistance to speakers and other participants with practical answers.

Wifi
A limited capacity is available at all conference venues.
Oral Sessions

**Time**
Each paper in an oral session is allocated 13 minutes. Additional 2 minutes are allocated for questions by the session chairs, switching between speakers and introducing the next speaker. You have to leave the podium once your time is up. **Do not exceed the time limit.**

**Testing laptop**
Oral speakers use their own laptop. They must testrun the laptop in combination with the audio equipment 20 minutes before the session starts at last.

**Arrival time**
Oral presenters should be present at the podium at least 20 minutes before the start of the session.

**Recording**
Oral presentations will be filmed and posted on the web by videolectures.net. If you have not signed the consent form with the video chair yet please do so before September 26th. We cannot broadcast your presentation without a signed consent form.

**Poster presentation**
All oral presentations have also been allocated a poster presentation. The poster presentation will be scheduled in the subsequent poster session after the talk. Each poster session will have 47 poster boards numbered 1-49. Oral papers have been assigned to poster boards 1-4. Poster boards 1-4 together with the demos (poster boards 48-49) are located in the beautiful roof garden.
Spotlight Sessions

Time
Presentations in the spotlight session are limited to **5 minutes and 30 seconds**. Additional **30 seconds** are allocated for switching between speakers. There will be no time allotted for “questions” after the spotlight. The next speaker is starting at the **6 minute** mark.

Presentation format
Presentations should be prepared as a video. (The slides are on the video, we do not mean you should make a video recording of yourself presenting.) All presentations should be finalized and submitted by October 4th. This will enable us to load all talks onto the same laptop without any configuration issues, while allowing presenters to use whatever graphics/video tools they choose to generate the presentation. The session chairs will not be introducing each spotlight therefore please make sure that you start with a title slide and that you introduce your talk.

Authors will be talking in front of their video as it plays projected on the main screen. Note that the audio channel from your submitted video will NOT be used. There will be a speaker’s monitor on the stage, so you can see what is being projected without turning to look at the screen. **The video should be at most 5 min and 30 sec long**, and should start with a title slide. It should be .MP4 (Mpeg4), HD resolution (1080p preferred, 720p accepted), with H264 encoding (the default for almost all MP4 encoders). It should NOT be an AVI file or a MOV but a proper MP4 (and don’t just rename it).

If you prepare your presentation using PowerPoint you can time your slides and save the presentation as a video directly from PowerPoint. Some instructions on how to do this can be found here (https://support.office.com/en-us/article/Turn-your-presentation-into-a-video-c140551f-cb37-4818-b5d4-3e30815c3e83). You will then need to convert the WMV video into MP4 (e.g. via Youtube or FFMPEG).

Alternatively, there are many free screen capture programs that directly produce proper MP4:

- VLC (http://www.videolan.org/vlc/index.html) which works on all platforms
- OBS (https://obsproject.com/) for windows, mac or linux
- TinyTake (http://tinytake.com) for windows or mac
- For windows 10 there is the built-in recorder (http://betanews.com/2015/07/19/windows-10-secret-screen-recording-tool/)
For Mac there is the built in QuickTime (which will need to be exported as MP4 as the default is MOV).

If you only have recorders producing other formats, g. avi, Youtube may be able to convert for you. If not try FFMPEG (ffmpeg i source_video.avi input acodec aac ab 128kb vcodec mpeg4 b 8000kb mbd 2 flags +4mv +trell aic 2 cmp 2 subcmp 2 s 1920×1080 title X final_video.mp4)

Submitting your spotlight presentation (mp4): Please name your “.mp4” file as “X-XX-XX.mp4”, where X-XX-XX is the presentation ID assigned to your paper (e.g. P-8A-55) in the conference program. In CMT, go to the Author console under column “Camera-Ready|Presentations”:

- Click on “Edit” corresponding to
- “Browse” to locate your “.mp4”
- Click on “Upload” to upload the “.mp4”

Note that the deadline for this is October 4th. When absent beyond this point, we cannot ascertain the video is presented and beyond this point in time no changes can be made.

Arrival time
Presenters should be present at the podium before the start of the session.

Recording
Spotlight presentations will be filmed and posted on the web by videolectures.net. If you have not signed the consent form with the video chair yet please do so before September 26th.

Poster presentation
All spotlight papers have also been allocated a poster presentation. The poster presentation will be scheduled in the subsequent poster session after the talk. Each poster session will have 49 poster boards numbered 1-49. Spotlight papers have been assigned to poster boards 5-9. Spotlight presenters are asked to install the posters during the coffee break prior to the poster session. Posters are to be removed from the poster boards by the presenters at the end of the session. Volunteers will collect the remaining ones.
Poster Sessions

Time
Each poster session lasts 1 hour 30 minutes. Poster papers have been assigned to poster boards 5-47 as indicated in the program. Check with the volunteers or the registration desk if you cannot find your poster board.

- Poster numbers 1-4 are allocated to oral posters. They are shown in the roof garden.
- Posters 5-9 are spotlight presentations. They are shown in the regular poster space.
- Posters 10-47 are regular posters.
- Posters 48-49 are demos. They are shown in the roof garden as well.

Presentation format
The poster format is A0 portrait. Poster boards are 841mm wide x 1189mm tall (equivalent to 33.1 inch wide x 46.8 inch tall). Adhesive material and/or pins will be provided for mounting the posters to the boards. If you have special requirements, please contact poster chairs as soon as possible. We will try to accommodate your requests as much as possible.

Arrival time
Poster presenters are asked to install their posters during the coffee break prior to the poster session. Remove posters promptly at the conclusion of each poster session.

Arrival and take off time
Poster presenters are asked to install their posters during the coffee break prior to the poster session. Posters are to be removed from the poster boards by the presenters at the end of the session. If not removed, volunteers will collect the remaining ones.

Poster printing
In case you misplaced or need to reprint your poster, there is an external poster printing facility close to the conference venue. You contact them via pr@printerette.nl.

Poster posted on ECCV website
As a new service to the community, all authors are offered the facility to upload posters using the CMT facility, up to October 4th. This is a service on trial, there is no obligation to do so. This will allow conference attendees to preview your poster through the ECCV website. Please upload your file as a single pdf file and keep
the file size of your upload to be **under 5MB** (Other formats and larger files will not be accepted).

**Submitting your poster presentation (pdf):**
Please name your “.pdf” file as “X-XX-XX.pdf”, where X-XX-XX is the presentation ID assigned to your paper (e.g. P-8A-55) in the conference program. In CMT, go to the Author console under column “Camera-Ready|Presentations”:

- Click on “Edit” corresponding to
- “Browse” to locate your “.pdf”
- Click on “Upload” to upload the “.pdf”
- Note that the deadline for this is October 4th. When absent beyond this point, we cannot ascertain the posters are posted and beyond this point in time no changes can be made.

**Demo Sessions**

**Time**
Demonstrations will take place in parallel with poster sessions (**1 hour 30 minutes**). Demos have been assigned to tables, electricity and poster boards (48-49). Check with the volunteers or the registration desk if you cannot find your table.

**Format**
Tables and power connections are provided for demonstrations. We encourage you to prepare your demonstrations to be independent of internet access.

**Arrival and take off time**
Demonstrators are asked to prepare their setup during the coffee break prior to the session. Remove your setup promptly at the conclusion of your demo session. Poster presenters are asked to install their posters during the coffee break prior to the poster session. Remove posters promptly at the conclusion of each poster session.
Practical Information

Proceedings

The proceedings will be freely available as part of the Lecture Notes on Computer Science from Springer until November 1, 2016. To access the papers, please visit http://www.eccv2016.org/proceedings/ and click on the volume you wish to access.

LNCS 9905 (ISBN: 978-3-319-46447-3)
Computer Vision – ECCV 2016
14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I
Editors: Bastian Leibe, Jiri Matas, Nicu Sebe, Max Welling

LNCS 9906 (ISBN: 978-3-319-46474-9)
Computer Vision – ECCV 2016
14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part II
Editors: Bastian Leibe, Jiri Matas, Nicu Sebe, Max Welling

LNCS 9907 (ISBN: 978-3-319-46486-2)
Computer Vision – ECCV 2016
14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part III
Editors: Bastian Leibe, Jiri Matas, Nicu Sebe, Max Welling

LNCS 9908 (ISBN: 978-3-319-46492-3)
Computer Vision – ECCV 2016
14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV
Editors: Bastian Leibe, Jiri Matas, Nicu Sebe, Max Welling

LNCS 9909 (ISBN: 978-3-319-46453-4)
Computer Vision – ECCV 2016
14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part V
Editors: Bastian Leibe, Jiri Matas, Nicu Sebe, Max Welling

LNCS 9910 (ISBN: 978-3-319-46465-7)
Computer Vision – ECCV 2016
14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part VI
Editors: Bastian Leibe, Jiri Matas, Nicu Sebe, Max Welling
LNCS 9911 (ISBN: 978-3-319-46477-0)
Computer Vision – ECCV 2016
14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part VII
Editors: Bastian Leibe, Jiri Matas, Nicu Sebe, Max Welling

LNCS 9912 (ISBN: 978-3-319-46483-1)
Computer Vision – ECCV 2016
14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part VIII
Editors: Bastian Leibe, Jiri Matas, Nicu Sebe, Max Welling

LNCS 9913 (ISBN: 978-3-319-46603-3)
Computer Vision – ECCV 2016 Workshops
Amsterdam, The Netherlands, October 8-10 and 15-16, 2016, Proceedings, Part I
Editors: Gang Hua, Hervé Jégou
Practical Information
Restaurants

Cheap / on the go / coffee

**Bakhuys Amsterdam**
Sarphatistraat 61
1018 EX Amsterdam
+31 20 370 4861
www.bakhuys-amsterdam.nl

**Cantarell**
Weesperplein 13-15
1018 WZ Amsterdam
+31 20 33 07 973
www.cantarell.nl

**Brezel**
Sarphatistraat 42
1018 GM Amsterdam
+31 6 25 588 235
www.brezel.nl

**Convenience to go**
Inside subwaystation **Weesperplein**
1018 XA Amsterdam

Midprice / seated lunch

**Amstelhaven**
Mauritskade 1
1091 EW Amsterdam
+31 20 66 52 672
www.amstelhaven.nl

**Café Noir**
Weesperplein 19-21
1018 WZ Amsterdam
+31 6 230 901
www.cafenoiramsterdam.nl

**Bar Lempicka**
Sarphatistraat 23
1018 EV Amsterdam
+31 20 622 0209
www.barlempicka.com

**Café De Magere Brug**
Amstel 81
1018 EK Amsterdam
+31 20 221 3400
www.demagerebrug.nl
Luxurious / extensive

De Ysbreker
Weesperzijde 23
1091 EC Amsterdam
+31 20 468 1808
www.deysbreker.nl

Dick & Cunningham
Kerkstraat 377
1017 HW Amsterdam
+31 20 422 2766
www.restaurantdenc.nl

La Riva
Professor Tulpplein 1
1018 GX Amsterdam
+31 20 520 3264
www.restaurantlarive.nl

Neva
Amstel 51
1018 DR Amsterdam
+31 20 530 7483
www.neva.nl
Workshops
Saturday, October 8

W1
Datasets and Performance Analysis in Early Vision
Organizers: M. Goesele, B. Jähne, K. Honauer, M. Waechter
Location: Oudemanhuispoort / Room D0.08

Please visit the website for the latest information on this workshop:
www.eccv2016.org/workshops
W2

Visual Analysis of Sketches
Location: Oudemanhuispoort / Room D0.09

Please visit the website for the latest information on this workshop:
www.eccv2016.org/workshops

08:30 – 08:45 Opening Remarks
08:45 – 09:30 Keynote 1
    Changhu Wang, MRSA
09:30 – 09:50 Oral 1
09:50 – 10:10 Oral 2
10:10 – 11:40 Posters + Coffee Break
11:40 – 12:00 Oral 3
12:00 – 12:20 Oral 4
12:20 – 13:05 Keynote 2
    James Hays, Georgia Tech
W3

Biological and Artificial Vision

Organizers: K. Ramakrishnan, R. M. Cichy, S. Ghebreab, H. S. Scholte, A. Smeulders

Location: Oudemanhuispoort / Room F0.01

Please visit the website for the latest information on this workshop:
www.eccv2016.org/workshops
W4

Brave New Ideas For Motion Representations in Videos

Organizers: E. Gavves, B. Fernando, J. van Gemert

Location: Oudemanhuispoort / Room D1.09

Please visit the website for the latest information on this workshop:
www.eccv2016.org/workshops

08:45 – 09:00 Welcome
09:00 – 09:45 Machine learning keynote
   Max Welling
09:45 – 10:00 Best paper 1
10:00 – 10:15 Break
10:15 – 11:00 Neuroscience keynote
   Aman Saleem
11:00 – 11:45 Computer Vision keynote
   Ivan Laptev
11:45 – 13:30 Break
13:30 – 14:15 Computer Vision keynote
   Andrew Zisserman
14:15 – 15:00 Visualization/Graphics keynote
   Elmar Eisemann
15:00 – 15:00 Best paper 2
15:15 – 15:30 Conclusion
15:30 – 17:30 Posters

P-1  Signs in time: Encoding human motion as a temporal image, Joon Son Chung, Andrew Zisserman
P-2  Back to Basics: Unsupervised Learning of Optical Flow via Brightness Constancy and Motion Smoothness, Jason J. Yu, Adam W. Harley, Konstantinos G. Derpanis
P-3  Human Action Recognition without Human, Yun He, Soma Shirakabe, Yutaka Satoh, Hirokatsu Kataoka
P-4  Motion Representation with Acceleration Images, Hirokatsu Kataoka, Yun He, Soma Shirakabe, Yutaka Satoh
P-5  Efficient Two-Stream Motion and Appearance 3D CNNs for Video Classification, Ali Diba, Ali Mohammad Pazandeh, Luc Van Gool
P-6  Segmentation Free Object Discovery in Video, Giovanni Cuffaro, Federico Becattini, Claudio Baecchi, Lorenzo Seidenari, Alberto Del Bimbo
P-7  Human Pose Estimation in Space and Time using 3D CNN, Agne Grinciunaite, Amogh Gudi, Emrah Tasli, Marten den Uyl
P-8  Autonomous driving challenge: To Infer the property of a dynamic object based on its motion pattern, Mona Fathollahi, Rangachar Kasturi

P-9  Integrated Variational And Nearest Neighbor field (IVANN) for Optical Flow, Zhuoyuan Chen, Ying Wu, Hailin Jin, Zhe Lin, Scott Cohen

P-10 Mining Spatial and Spatio-Temporal ROIs for Action Recognition, Xiaochen Lian, Zhuoyuan Chen, Yi Yang, Jiang Wang, Alan Yuille

P-11 Temporal Convolutional Networks: A Unified Approach to Action Segmentation, Colin Lea, Rene Vidal, Austin Reiter, Gregory D. Hager

P-12 Do Motion Boundaries Improve Semantic Segmentation?, Yu-Hui Huang, Jose Oramas, Tinne Tuytelaars, Luc Van Gool

P-13 Making a Case for Learning Motion Representations with Phase, S. L. Pintea, J. C. van Gemert
Regular Tutorials
Saturday, October 8

Morning Session

09:00 – 12:00  Geometric deep learning
Rodola, Masci, Bronstein
Location: Oudemanhuispoort / Room D1.08
https://sites.google.com/site/eccvgdl/

Dr. Emanuele Rodolà
USI Lugano, Switzerland
emanuele.rodola@usi.ch
https://sites.google.com/site/erodola/

Dr. Jonathan Masci
USI Lugano, Switzerland
jonathan.masci@usi.ch
http://www.people.usi.ch/mascij/

Prof. Michael Bronstein
USI Lugano, Switzerland / Intel Perceptual Computing, Israel
michael.bronstein@usi.ch
http://www.inf.usi.ch/bronstein/

09:00 – 12:00  Zero-shot learning for vision and multimedia
Mensink, Gavves, Snoek
Location: Oudemanhuispoort / Room C0.17
https://staff.fnwi.uva.nl/t.e.j.mensink/zsl2016/

Thomas Mensink
University of Amsterdam, The Netherlands
thomas.mensink@uva.nl

Efstratios Gavves
University of Amsterdam, The Netherlands
egavves@uva.nl

Cees G.M. Snoek
University of Amsterdam, The Netherlands
cgmsnoek@uva.nl
09:00 – 12:00  New directions in saliency research: Developments in architecture, datasets and evaluation
Borji, Bylinskii, Judd
Location: Oudemanhuispoort / Room A0.08

Ali Borji
Assistant Professor
Center for Research in Computer Vision, University of Central Florida
aborji@crcv.ucf.edu

Zoya Bylinskii
Graduate Student
Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology
zoya@mit.edu

Tilke Judd
Product Manager
Google
tilke.judd@gmail.com

Afternoon Session

14:00 – 17:00  Deep learning meets model optimization and statistical inference
Wang, Zuo, Liu, Chen, Peng
Location: Oudemanhuispoort / Room D1.08

Zhangyang Wang
Assistant Professor
Department of Computer Science and Engineering, Texas A&M University, USA
masterwant@gmail.com

Wangmeng Zuo
Professor
School of Computer Science and Technology, Harbin Institute of Technology, China
wmzuo@hit.edu.cn

Risheng Liu
Associate Professor
School of Software, Dalian University of Technology, China
rsliu@dlut.edu.cn

Yunjin Chen
Lecturer
National Laboratory for Parallel and Distributed Processing,
National University of Defense Technology, China.
chenyunjin_nudt@hotmail.com
14:00 – 17:00  Computational geometry tools and applications in computer vision
Lafarge, Alliez, Fabri
Location: Oudemanhuispoort / Room C0.17

Dr. Florent Lafarge
Inria Sophia Antipolis
florent.lafarge@inria.fr
http://www-sop.inria.fr/members/Florent.Lafarge/

Dr. Pierre Alliez
Inria Sophia Antipolis
pierre.alliez@inria.fr
https://team.inria.fr/titane/pierre-alliez/

Dr. Andreas Fabri
GeometryFactory
andreas.fabri@geometryfactory.com
http://geometryfactory.com/who-we-are/

14:00 – 17:00  Distance metric learning for computer vision
Lu, Wang
Location: Oudemanhuispoort / Room A0.08
http://ivg.au.tsinghua.edu.cn/ECCV16_tutorial/

Dr. Jiwen Lu
Associate Professor
Department of Automation
Tsinghua University, China
ljiwen@tsinghua.edu.cn
http://ivg.au.tsinghua.edu.cn/Jiwen_Lu/

Dr. Ruiping Wang
Associate Professor
Institute of Computing Technology
Chinese Academy of Sciences, China
ruiping.wang@vipl.ict.ac.cn
http://vipl.ict.ac.cn/homepage/rpwang
Workshops
Sunday, October 9

W5
Joint Imagenet and MS Coco Visual Recognition Challenge Workshop
Organizers: W. Liu, G. Patterson, M. Ronchi, Y. Cui, T.Y. Lin, L. Zitnick, P. Dollar, O. Russakovsky, J. Deng, F-F. Li, A. Berg
Location: NH Hotel Krasnapolsky / Room Ruby

Please visit the website for the latest information on this workshop:
www.eccv2016.org/workshops

08:30 – 08:50  Opening remarks
08:50 – 09:00  ILSVRC Localization / Classification Challenge
09:00 – 09:30  Invited teams 1
09:30 – 10:00  Coffee + Posters
10:00 – 10:30  Invited Speaker 1
                 Sanja Fidler
10:30 – 10:40  ILSVRC Video Detection Challenge
10:40 – 11:10  Invited teams 2
11:10 – 11:30  Places2 Challenge: Scene Classification and Parsing
11:30 – 13:00  Break
13:00 – 13:30  Invited Speaker 2
                 Vittorio Ferrari
13:30 – 14:00  COCO & ILSVRC Object Detection and Segmentation Challenges
14:00 – 15:00  Invited teams 3
15:00 – 15:30  Coffee + Posters
15:30 – 16:00  COCO Person Keypoint Detection Challenge
16:00 – 16:30  Invited teams 4
16:30 – 16:45  Awards
16:45 – 17:00  Future Plans & Discussion
17:00 – 18:00  Posters
**W6**

**Geometry Meets Deep Learning**


**Location:** NH Hotel Krasnapolsky / Room Diamant

Please visit the website for the latest information on this workshop: www.eccv2016.org/workshops

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:20 – 08:30</td>
<td>Opening</td>
</tr>
<tr>
<td>08:30 – 09:00</td>
<td>Invited speaker 1</td>
</tr>
<tr>
<td></td>
<td>Joan Bruna</td>
</tr>
<tr>
<td>09:00 – 09:30</td>
<td>Invited speaker 2</td>
</tr>
<tr>
<td></td>
<td>Robert Cipolla</td>
</tr>
<tr>
<td>09:30 – 10:00</td>
<td>Invited speaker 3</td>
</tr>
<tr>
<td></td>
<td>Jitendra Malik</td>
</tr>
<tr>
<td>10:00 – 10:15</td>
<td>Oral Session 1</td>
</tr>
<tr>
<td></td>
<td>Learning Covariant Feature Detectors, Karel Lenc, Andrea Vedaldi</td>
</tr>
<tr>
<td>10:15 – 10:30</td>
<td>Oral Session 2</td>
</tr>
<tr>
<td></td>
<td>Overcoming Occlusion with Inverse Graphics, Pol Moreno, Chris</td>
</tr>
<tr>
<td></td>
<td>Williams, Charlie Nash, Pushmeet Kohli</td>
</tr>
<tr>
<td>10:30 – 11:00</td>
<td>Coffee + Posters</td>
</tr>
<tr>
<td>11:00 – 11:30</td>
<td>Invited speaker 4</td>
</tr>
<tr>
<td></td>
<td>Honglak Lee</td>
</tr>
<tr>
<td>11:30 – 12:00</td>
<td>Invited speaker 5</td>
</tr>
<tr>
<td></td>
<td>Vladlen Koltun</td>
</tr>
<tr>
<td>12:00 – 12:30</td>
<td>Invited speaker 6</td>
</tr>
<tr>
<td></td>
<td>Raquel Urtasun</td>
</tr>
<tr>
<td>12:30 – 14:00</td>
<td>Break</td>
</tr>
<tr>
<td>14:00 – 14:30</td>
<td>Invited speaker 7</td>
</tr>
<tr>
<td></td>
<td>Michael Black</td>
</tr>
<tr>
<td>14:30 – 15:00</td>
<td>Invited speaker 8</td>
</tr>
<tr>
<td></td>
<td>Hao Li</td>
</tr>
<tr>
<td>15:00 – 15:30</td>
<td>Invited speaker 9</td>
</tr>
<tr>
<td></td>
<td>Michael Bronstein</td>
</tr>
<tr>
<td>15:30 – 15:45</td>
<td>Oral Session 3</td>
</tr>
<tr>
<td></td>
<td>Learning the structure of objects from Web supervision, David</td>
</tr>
<tr>
<td></td>
<td>Novotny, Diane Larlus, Andrea Vedaldi</td>
</tr>
</tbody>
</table>

**Workshops** | Sunday, October 9 | NH Hotel Krasnapolsky
15:45 – 16:00  Oral Session 4  
- Deep Shape from A Low Number of Silhouettes, Xinhan Di, Rozenn Dahyot, Mukta Prasad

16:00 – 16:15  Oral Session 5  
- How useful is photo-realistic rendering for visual learning?, Yair Movshovitz-Attias, Yaser Sheikh, Takeo Prof. Kanade

16:15 – 17:00  Coffee + Posters

17:00 – 17:30  Invited speaker 10  
Jana Kosecka

17:30 – 18:00  Invited speaker 11  
Jianxiong Xiao
W7

Action and Anticipation for Visual Learning
Organizers: D. Jayaraman, K. Grauman, S. Levine
Location: Oudemanhuispoort / Room D0.09

Please visit the website for the latest information on this workshop:
www.eccv2016.org/workshops

09:00 – 09:15 Introduction
09:15 – 09:45 Invited speaker 1
   Jitendra Malik (tentative, see goo.gl/Fdtnoj)
09:45 – 10:15 Invited speaker 2
   Nando de Freitas (tentative, see goo.gl/Fdtnoj)
10:15 – 10:45 Invited speaker 3
   Jeanette Bohg (tentative, see goo.gl/Fdtnoj)
10:45 – 11:00 Coffee break
11:00 – 11:45 Poster spotlights
11:45 – 13:00 Break
13:00 – 14:00 Poster session
14:00 – 14:30 Invited speaker 4
   Abhinav Gupta (tentative, see goo.gl/Fdtnoj)
14.30 – 15.00 Invited speaker 5
   Andrea Censi (tentative, see goo.gl/Fdtnoj)
15:00 – 15:30 Invited speaker 6
   Ali Farhadi (tentative, see goo.gl/Fdtnoj)
15:30 – 15:45 Coffee break
15:45 – 16:30 Panel discussion with all invited speakers
   Jitendra Malik, Nando de Freitas, Jeanette Bohg, Andrea Censi, Abhinav Gupta, Ali Farhadi
W8

Computer Vision for Road Scene Understanding and Autonomous Driving

Location: Oudemanhuispoort / Room D0.08

Please visit the website for the latest information on this workshop:
www.eccv2016.org/workshops

08:50 – 09:00 Opening remarks
09:00 – 09:45 Invited speaker 1
Branislav Kisacanin, NVIDIA
09:45 – 10:30 Invited speaker 2
Dariu Gavrila, TU Delft
10:30 – 11:45 Poster session / coffee
11:45 – 12:30 Invited speaker 3
Petros Kapsalas, Panasonic
12:30 – 14:00 Break
14:00 – 14:45 Invited speaker 4
James Philbin, Zoox
14:45 – 15:30 Invited speaker 5
Shanti Swarup, Uurmi Systems
15:30 – 16:45 Poster session / coffee
16:45 – 17:30 Invited speaker 6
Joakim Lin-Sorstedt, Volvo Cars
W9
Challenge on Automatic Personality Analysis
Organizers: S. Escalera, X. Baró, I. Guyon, H. J. Escalante
Location: Oudemanhuispoort / Room C0.23

Please visit the website for the latest information on this workshop:
www.eccv2016.org/workshops

08:30 – 08:45  Presentation (Sergio Escalera)
08:45 – 09:30  Invited speaker 1
Björn W. Schuller
09:30 – 09:45  Challenge results and award ceremony
Contributed paper:
V. Ponce Lopez, B. Chen, A. Places, M. Oliu, C. Corneanu, X. Baro, H. J. Escalante,
I. Guyon, S. Escalera
ChaLearn LAP 2016: First Round
Challenge on First Impressions - Dataset and Results.
ChaLearn Looking at People Workshop on Apparent Personality Analysis.

09:45 – 10:15  Coffee Break
10:15 – 11:00  Invited speaker 2
Maja Pantic
11:00 – 11:20  Deep Bimodal Regression for Apparent Personality Analysis
1st place of the Challenge
C. L. Zhang, H. Zhang, X. S. Wei and J. Wu
11:20 – 11:50  Bi-modal First Impressions Recognition using Temporally Ordered Deep
Audio and Stochastic Visual Features
2nd place of the challenge
A. Subramaniam, V. Patel, A. Mishra, P. Balasubramanian, A. Mittal
11:50 – 12:10  Deep Impression: Audiovisual Deep Residual Networks for Multimodal
Apparent Personality Trait Recognition
3rd place of the challenge
Y. Güçlütürk, U. Güçlü, M. van Gerven, R. van Lier
12:10 – 13:30  Break
13:30 – 13:50  SASE: RGB-Depth Database for Human Head Pose Estimation
I. Lüsi, S. Escalera, G. Anbarjafari
13:50 – 14:10  Deep Impression: Audiovisual Deep Residual Networks for Multimodal
Apparent Personality Trait Recognition
14:10 – 14:30  Deep Learning for Facial Action Unit Detection under Large Head Poses
Zoltan Toser, Laszlo Jeni, Andras Lorincz, Jeffrey Cohn
14:30 – 14:50  Combining Deep Facial and Ambient Features for First Impression Estimation
F. Gürpınar, H. Kaya, A. Salah

14:50 – 15:30  Invited Speaker 3
Daniel Gatica-Perez

15:30 – 15:50  Coffe break

15:50 – 16:30  Invited Speaker 4
Roland Goecke

16:30 – 16:50  The static MultiModal Dyadic Behavior Dataset for Engagement Prediction
P. Tsatsoulis, P. Kordas, M. Marshall, D. Forsyth, A. Rozga

16:50 – 17:10  Overcoming Calibration Problems in Pattern Labeling with Pairwise Ratings: Application to Personality Traits

17:10 – 17:45  Invited Speaker 5
Michael Valstar

17:45 – 18:00  Closing ceremony (Sergio Escalera)
### W10

**BioImage Computing**  
Organizers: P. Bouthemy, F. Hamprecht, E. Meijering, T. Pécot, P. Perona, C. Rother  
**Location:** Oudemanhuispoort / Room C1.17  

Please visit the website for the latest information on this workshop:  
[www.eccv2016.org/workshops](http://www.eccv2016.org/workshops)

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:45 – 09:00</td>
<td><strong>Welcome and Introduction</strong></td>
</tr>
</tbody>
</table>
| 09:00 – 09:40| **Invited Speaker 1**  
Charless Fowlkes (UC Irvine)  
Learning to detect, segment, and classify biological structures |
| 09:40 – 10:20| **Invited Speaker 2**  
Pascal Fua (EPFL)  
Comparing synaptic structure in the cortex of adult and old mice |
| 10:20 – 10:40| **Contributed talk 1**  
| 10:40 – 10:50| **Coffee break**                                                                           |
| 10:50 – 12:00| **Poster Session 1 (4 posters)**                                                           |
| 12:00 – 13:30| **Break**                                                                                 |
| 13:30 – 14:10| **Invited Speaker 3**  
Jens Rittscher (University Oxford)  
Monitoring complex biological environments |
| 14:10 – 14:50| **Invited Speaker 4**  
Jeroen van der Laak (Radboud University Medical Center)  
Challenges of using computer vision for evaluating whole-slide pathology images |
| 14:50 – 15:10| **Contributed talk 2**  
Poisson point processes for solving stochastic inverse problems in fluorescence microscopy, Ihor Smal and Erik Meijering (Erasmus MC, University Medical Center Rotterdam) |
| 15:10 – 15:20| **Coffee break**                                                                           |
| 15:20 – 16:30| **Poster Session 2 (4 posters)**                                                           |
| 16:30 – 17:10| **Invited Speaker 5**  
Thomas Walter (Institut Curie, Ecole des Mines Paris)  
Bioimage Informatics for Phenomics – Exploring morphological phenotypes and spatial organization by large-scale screening approaches |
| 17:10 – 17:50| **Invited Speaker 6**  
Gonzalo de Polavieja (Instituto Cajal, Madrid)  
Tracking animals in groups from video |
| 17:50 – 18:00| **Closing**                                                                               |
W11
Benchmarking Multi-Target Tracking: MOTChallenge
Organizers: L. Leal-Taixé, A. Milan, K. Schindler, I. Reid, S. Roth
Location: Oudemanhuispoort / Room F0.01

Please visit the website for the latest information on this workshop:
www.eccv2016.org/workshops

08:50 – 09:00 Opening remarks
09:00 – 09:30 Invited Speaker 1
Horesh Ben Shitrit
09:30 – 09:50 Oral Session 1
- Multi-Person Tracking by Multicut and Deep Matching, Siyu Tang, Max Planck Institute for Informatics
09:50 – 10:10 Oral Session 2
- Tracking Multiple Persons Based on a Variational Bayesian Model, Yutong Ban, INRIA
10:10 – 10:40 Invited Speaker 2
Bastian Leibe
10:40 – 11:00 Coffee break
11:00 – 11:30 Invited Speaker 3
Alexandre Alahi
11:30 – 12:00 Invited Speaker 4
Wongun Choi
12:00 – 13:30 Break
13:30 – 15:00 Poster Session
15:00 – 15:30 Invited Speaker 5
Andreas Geiger
15:30 – 15:50 Oral 3
- Online multi-target tracking with strong and weak detections, Ricardo Sanchez Matilla, Queen Mary University of London
15:50 – 16:20 Invited Speaker 6
Bernt Schiele
16:20 – 17:00 Challenge, Results, Awards, Discussion and Closing
W12
Assistive Computer Vision and Robotics
Organizers: G. Medioni, G. M. Farinella, M. Leo, M. Trivedi
Location: Oudemanhuispoort / Room C0.17

Please visit the website for the latest information on this workshop:
www.eccv2016.org/workshops

| 08:15 – 08:30 | Opening Remarks |
| 08:30 – 09:15 | Invited Speaker 1  
Devi Parikh, Virginia Tech, US |
| 09:15 – 10:00 | Oral Session 1 |
| 10:00 – 11:45 | Poster Session 1 |
| 10:30 – 11:00 | Coffee Break |
| 11:45 – 12:30 | Oral Session 2 |
| 12:30 – 14:00 | Break |
| 14:00 – 14:45 | Invited Speaker 2  
Barbara Caputo, University of Rome La Sapienza, IT |
| 14:45 – 15:30 | Oral Session 3 |
| 15:30 – 16:00 | Coffee Break |
| 16:00 – 17:15 | Poster Session 2 |
| 17:15 – 18:00 | Oral Session 4 |
| 18:00 – 18:15 | Closing Remarks |
W13
Transferring and Adapting Source Knowledge in Computer Vision
Organizers: W. Li, T. Tommasi, F. Orabona, D. Vázquez, M. López, J. Xu, H. Larochelle
Location: Oudemanhuispoort / Room D1.09

Please visit the website for the latest information on this workshop; www.eccv2016.org/workshops
W14

Recovering 6D Object Pose


Location: Oudemanhuispoort / Room A0.08

Please visit the website for the latest information on this workshop; www.eccv2016.org/workshops

08:50 – 09:00 Opening
09:00 – 09:30 Invited talk 1
Silvio Savarese
09:30 – 10:00 Invited talk 2
Carsten Steger
10:00 – 10:30 Coffee break
10:30 – 11:30 Oral session
paper 1 Direct-from-Video: Unsupervised NRSfM, Karel Lebeda, Simon Hadfield, Richard Bowden
paper 2 Towards Categorization and Pose Estimation of Sets of Occluded Objects in Cluttered Scenes from Depth Data and Generic Object Models Using Joint Parsing, Hector Basevi, Ales Leonardis
paper 3 Reconstructing Articulated Rigged Models from RGB-D Videos, Dimitrios Tzionas, Juergen Gall
paper 4 On Evaluation of 6D Object Pose Estimation, Tomas Hodan, Jiri Matas, Stepan Obdrzalek

11:30 – 12:00 Invited talk 3
Jianxiong Xiao
12:00 – 13:30 Break
13:30 – 14:00 Spotlights
14:00 – 15:00 Poster session
P-1 RobotFusion: Grasping with a Robotic Manipulator via Multi-view Reconstruction, Daniele De Gregorio, Federico Tombari, Luigi Stefano
P-2 A Direct Method for Robust Model-Based 3D Object Tracking from a Single RGB Image, Byung-Kuk Seo, Wuest Harald
P-3 A Radial Search Method for fast Nearest Neighbor Search on Range Images, Federico Tombari, Samuele Salti, Luca Puglia, Luigi Stefano, Giancarlo Raiconi
P-4 Physical Reasoning for 3D Object Recognition using Global Hypothesis Verification, Shuichi Akizuki, Manabu Hashimoto
P-5 SemanticFusion: Joint Labeling, Tracking and Mapping, Tommaso Cavallari, Luigi Stefano
P-6 Affordance-focused Features for Generic Object Recognition, Shuichi Akizuki, Masaki Iizuka, Manabu Hashimoto

P-7 Fast Feature-Less Quaternion-based Particle Swarm Optimization for Rigid and Articulated Pose Estimation From RGB-D Images, Giorgio Toscana, Stefano Rosa

P-8 Object Detection for Random Bin Picking using Point Pair Features, Wim Abbeloos, Toon Goedeme

P-9 Fast 6D Pose Estimation from a Monocular Image using Hierarchical Pose Trees, Yoshinori Konishi, Yuki Hanzawa, Masato Kawade, Manabu Hashimoto

P-10 ObjectNet3D: A Large Scale Database for 3D Object Recognition, Yu Xiang, Wonhui Kim, Wei Chen, Jingwei Ji, Christopher Choy, Hao Su, Roozbeh Mottaghi, Leonidas J. Guibas, Silvio Savarese

P-11 SDF-2-SDF: Highly Accurate 3D Object Reconstruction, Miroslava Slavcheva, Wadim Kehl, Nassir Navab, Slobodan Ilic

P-12 Bayesian Image based 3D Pose Estimation, Valºamis Ntouskos, Fiera Pirri, Marta Sanzari

P-13 Learning a Predictable and Generative Vector Representation for Objects, Rohit Girdhar, David Fouhey, Mikel Rodriguez, Abhinav Gupta

P-14 Localizing and Orienting Street Views Using Overhead Imagery (location and orientation of street images), Nam Vo, James Hays

P-15 Weakly Supervised Object Localization Using Size Estimates, Miaojing Shi, Vittorio Ferrari

P-16 Real-time Joint Tracking of a Hand Manipulating an Object from RGB-D Input, Srinath Sridhar, Franziska Mueller, Michael Zollhoefer, Dan Casas, Antti Oulasvirta, Christian Theobalt

P-17 Global Registration of 3D Point Sets via LRS decomposition, Federica Arrigoni, Beatrice Rossi, Andrea Fusiello

P-18 4D Match Trees for Non-rigid Surface Alignment, Armin Mustafa, Hansung Kim, Adrian Hilton

P-19 Rigid 3D Shape Retrieval via Large Margin Nearest Neighbor Embedding, Ioannis Chiotellis, Rudolph Treibel, Thomas Windheuser, Daniel Cremers

P-20 Deep learning 3D shape surfaces using geometry images, Ayan Sinha, Jing Bai, Karthik Ramani

P-21 Single Image 3D Interpreter Network, Jiajun Wu, Tianfan Xue, Joseph Lim, Yuandong Tian, Joshua Tenenbaum, Antonio Torralba, William Freeman

P-22 3D-R2N2: A unified approach for single and multi-view 3D object reconstruction, Christopher Choy, Danfei Xu, JunYoung Gwak, Kevin Chen, Silvio Savarese

P-23 Deep Deformation Network for Object Landmark Localization, Xiang Yu, Feng Zhou, Manmohan Chandraker

P-24 DOC: Deep OCclusion Estimation From a Single Image, Peng Wang, Alan Yuille

P-25 Facilitating and Exploring Planar Homogeneous Texture for Indoor Scene Understanding, Shahzor Ahmad, Loong-Fah Cheong

P-26 Hand pose estimation from local surface normals, Chengde Wan, Angela Yao, Luc Van Gool
Deep Learning of Local RGB-D Patches for 3D Object Detection and 6D Pose Estimation, Wadim Kehl, Fausto Milletari, Federico Tombari, Slobodan Ilic, Nassir Navab

Multi-view 3D Models from Single Images With a Convolutional Network, Maxim Tatarchenko, Alexey Dosovitskiy, Thomas Brox

15:00 – 15:30 Coffee break
15:30 – 16:00 Invited talk 4
   Tinne Tuytelaars
16:00 – 16:30 Invited talk 5
   Zoltan Marton
16:30 – 17:00 Invited talk 6
   Carsten Rother
17:00 – 17:20 Awards and Closing
**W15**

**Robust Reading**

Location: Oudemanhuispoort / Room C2.23

Please visit the website for the latest information on this workshop;  
www.eccv2016.org/workshops

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:50 – 09:00</td>
<td>Welcome</td>
</tr>
<tr>
<td>09:00 – 10:00</td>
<td>Invited Talk 1</td>
</tr>
<tr>
<td></td>
<td>- Deep Learning and Synthetic Data for Text Spotting, Dr Max Jaderberg (Google Deep Mind, UK)</td>
</tr>
<tr>
<td>10:00 – 10:30</td>
<td>Coffee break</td>
</tr>
<tr>
<td>10:30 – 11:30</td>
<td>Oral Session 1</td>
</tr>
<tr>
<td></td>
<td>- Robust Text Detection with Vertically-Regressed Proposal Network, Donglai Xiang, Qiang Guo, Yan Xia</td>
</tr>
<tr>
<td></td>
<td>- Scene Text Detection with Adaptive Line Clustering, Xinxu Qiao, He Zhu, Weiping Li</td>
</tr>
<tr>
<td></td>
<td>- From text detection to text segmentation: a unified evaluation scheme, Stefania Calarasanu, Jonathan Fabrizio, Séverine Dubuisson</td>
</tr>
<tr>
<td>11:30 – 12:30</td>
<td>Invited Talk 2</td>
</tr>
<tr>
<td></td>
<td>- COCO-Text: Dataset and Benchmark for Text Detection and Recognition in Natural Images, Dr Serge Belongie (Cornell Tech, USA)</td>
</tr>
<tr>
<td>12:30 – 13:30</td>
<td>Break</td>
</tr>
<tr>
<td>13:30 – 14:50</td>
<td>Oral Session 2</td>
</tr>
<tr>
<td></td>
<td>- Efficient exploration of text regions in natural scene images using, adaptive image sampling, Ismet Zeki Yalniz, Douglas Gray, R. Manmatha</td>
</tr>
<tr>
<td></td>
<td>- Dynamic Lexicon Generation for Natural Scene Images, Yash Patel, Lluis Gomez, Maral Rusinol, Dimosthenis Karatzas</td>
</tr>
<tr>
<td></td>
<td>- End-to-End Interpretation of the French Street Name Signs Dataset, Raymond Smith, Chunhui Gu, Dar-Shyang Lee, Huiyi Hu, Ranjith Unnikrishnan, Julian Ibarz, Sacha Arnould, Sophia Lin</td>
</tr>
<tr>
<td></td>
<td>- Downtown Osaka Scene Text Dataset, Masakazu Iwamura, Takahiro Matsuda, Naoyuki Morimoto, Hitomi Sato, Yuki Ikeda, Koichi Kise”</td>
</tr>
<tr>
<td>14:50 – 15:20</td>
<td>Coffee Break</td>
</tr>
<tr>
<td>15:20 – 16:20</td>
<td>Invited Talk 3</td>
</tr>
<tr>
<td></td>
<td>- Robust Reading in Mobile Apps, Prof R. Manmatha (A9/Amazon, Palo Alto, CA)</td>
</tr>
<tr>
<td>16:20 – 17:20</td>
<td>Panel Discussion</td>
</tr>
<tr>
<td>17:20 – 17:30</td>
<td>Closing Remarks</td>
</tr>
</tbody>
</table>
W16
3D Face Alignment in the Wild & Challenge

Organizers: J. Cohn, L. Jeni, N. Sebe, S. Tulyakov, L. Yin
Location: Oudemanhuispoort / Room C2.17

Please visit the website for the latest information on this workshop;
www.eccv2016.org/workshops
W17

Egocentric Perception, Interaction and Computing

Organizers: G. Serra, R. Cucchiara, W. Mayol-Cuevas, A. Bulling, D. Damen

Location: Oudemanhuispoort / Room C3.17

Please visit the website for the latest information on this workshop;
www.eccv2016.org/workshops

08.30 – 08:40 Welcome and opening
08:40 – 09:25 Invited Talk 1
- Understanding group activities from multiple first-person perspectives, Yoichi Sato
09:25 – 10:10 Oral Session 1
- DeepDiary: Automatically Captioning Lifelogging Image Streams, Chenyou Fan, David Crandall
- Temporal Segmentation of Egocentric Videos to Highlight Personal Locations of Interest, Antonino Furnari, Giovanni Maria Farinella, Sebastiano Battiato
- Face-off: a Face, Reconstruction Technique for Virtual Reality (VR) Scenarios, Muhammad Sikandar Lal Khan, Shafiq Ur Rehman, Ulrik Söderström, Alaa Halawani, Haibo Li

10:10 – 10:40 Coffee break
10:40 – 11:15 Invited Talk 2
- Glancing for help: integrating attention models and visual appearance for multi user egocentric assistance, Walterio Mayol-Cuevas and Dima Damen
11:15 – 12:15 Oral Session 2
- A3D: A device for studying gaze in 3D, Mahmoud Qodseya, Valsamis Ntouskos, Marta Sanzari, Fiora Pirri
- Interactive feature growing for accurate object detection in megapixel images, Julius Schöning, Patrick Faion, Gunther Heidemann
- Egocentric Vision for Visual Market Basket Analysis, Vito Santarcangelo, Giovanni Maria Farinella, Sebastiano Battiato
- SEMBED: Semantic Embedding of Egocentric Action Videos, Michael Wray, Davide Moltisanti, Walterio Mayol-Cuevas, Dima Damen

12:15 – 13:45 Break
13:45 – 14:30 Invited Talk 3
- Activity and Photographer Recognition from Visual Motion, Shmuel Peleg
14:30 – 15:15 Oral Session 3
- GPU Accelerated Left/Right Hand-segmentation in First Person Vision, Alejandro Betancourt, Lucio Marcenaro, Emilia Barakova, Matthias Rauterberg, Carlo,Regazzoni
- Towards Semantic Fast-Forward and Stabilized Egocentric Videos, Michel Silva, Washington Ramos, João Ferreira, Mario Campos, Erickson Nascimento
- Context Change Detection for an Ultra-Low Power Low-Resolution Ego-Vision Imager, Francesco Paci, Lorenzo Baraldi, Giuseppe Serra, Rita,Cucchiara, Luca Benini
15:15 – 15:45 Coffee break
15:45 – 16:15 Spotlight session

- Automatic analysis of in-the-wild mobile eye-tracking experiment, Stijn De Beugher, Geert Brône, Toon Goedemé
- Optimizing Deep CNNs for Object Recognition in Egocentric Video for Vision-based Neuro-Prostheses Design, Philippe Pérez De San Roman, Jenny Benois-Pineau, Daniel Cattaert, Florent Paclet, Aymar De Rugy
- Detecting and Segmenting Hands to Recognize Social Interactions in Egocentric Video, Sven Bambach, Stefan Lee, David Crandall, Chen Yu
- Detecting Engagement in Egocentric Video, Yu-Chuan Su, Kristen Grauman
- Spatio-temporal Localization of Egocentric Viewers in Top-view Videos, Shervin Ardeshir, Ali Borji
- EgoCap: Egocentric Marker-less Motion Capture with Two Fisheye Cameras, Helge Rhodin, Christian Richardt, Dan Casas, Eldar Insafutdinov, Mohammad Shafiei, Hans-Peter Seidel, Bernt Schiele, Christian Theobalt
- Development of a Tiny, Low-cost and Wireless Motion Sensor for Interacting with Virtual Reality Games, Mahdi Kazempour, Hakan Urey, Erdem Ulusoy
- Solfège hand sign recognition with smart glasses, Gábor Sörös, Julia Giger, Jie Song
- Human Tracking and Re-identification in Egocentric Videos, Jyoti Nigam, Renu M. Rameshan
- A Software Retina for Egocentric & Robotic Vision Applications on Mobile Platforms, Jan Siebert, Adam Schmidt, Gerardo Aragon-Camarasa, Nick Hockings, Xiaomeng Wang, W. Paul Cockshott

16:15 – 17:20 Poster session
17:20 – 17:30 Closing remarks
Workshops
Monday, October 10

W18
Local Features: State of the art, open problems and performance evaluation
Organizers: J. Matas, K. Mikolajczyk, T. Tuytelaars, A. Vedaldi
Location: Theatre Tuschinski / Room 1

Please visit the website for the latest information on this workshop;
www.eccv2016.org/workshops
W19
Crowd Understanding
Organizers: F. Bremond, V. Libal, A. Cavallaro, T. Pajdla
Location: Theatre Tuschinski / Room 5

Please visit the website for the latest information on this workshop;
www.eccv2016.org/workshops

09.00 – 09:15 Welcome and opening
09:15 – 10:00 Invited Talk 1
  Pascal Fua
  Identity Preserving Multi-People Tracking through Linear programming
10:00 – 10:40 Spotlight session
10:40 – 11:00 Cofee break
11:00 – 11:45 Invited Talk 2
  Mubarak Shah
  Counting, Detecting And Tracking of People in Crowded Scenes
11:45 – 12:45 Poster session
12:45 – 13:00 Cofee break
13:00 – 13:45 Invited Talk 3
  Gennadiy Reznicenko
  Turning Big Data into Rich data: Proactive Defensive Strategie in an Age of Modern Surveillance
13:45 – 14:00 Closing remarks
W20 Video Segmentation

Organizers: T. Brox, K. Fragkiadaki, F. Galasso, F. Li, J. M. Rehg, B. Schiele, M. Ying Yang
Location: Theatre Tuschinski / Room 2

Please visit the website for the latest information on this workshop; www.eccv2016.org/workshops

08.30 – 08:40 Welcome and Opening
08:40 – 09:10 Invited Talk 1
   Jitendra Malik
09:10 – 09:35 Talk
   - Improved Image Boundaries for Better Video Segmentation
   - Pointwise mutual information-based video segmentation with high temporal consistency
09:35 – 10:05 Invited Talk 2
   Michael Black (MPI Intelligent Systems)
   Towards semantic optical flow: Jointly estimating image motion, object segmentation and scene structure
10:05 – 10:30 Talk
   - Causal Motion Segmentation in Moving Camera Videos, P. Bideau, E. Learned-Miller
   - Can Ground Truth Label Propagation from Video Help Semantic Segmentation?
     S. Mustikovela, M.Y. Yang, C. Rother
10:30 – 11:00 Invited Talk 3
   TBD, Stefano Soatto
11:00 – 11:25 Talk
   - A Benchmark Dataset and Evaluation Methodology for Video Object Segmentation
   - 3D Point Cloud Video Segmentation based on Interaction Analysis
11:25 – 11:45 Coffee break
11:45 – 12:15 Invited Talk 4
   Michel Irani (Weizmann Institute)
   Video Segmentation and Inference by Composition
12:15 – 12:40 Talk
   - Deep learning based fence segmentation and removal from an image using a video sequence, S. Jonna, K.K. Nakka, R.R. Sahay
   - Clockwork Convnets for Video Semantic Segmentation?, E. Shelhamer, K. Rakelly, J. Hoffman, T. Darrell
12:40 – 13:10 Invited Talk 5
   Vittorio Ferrari
   Learning articulated object classes from video
13:10 – 14:00 Panel Discussion
The Visual Object Tracking Challenge Workshop

Organizers: M. Kristan, A. Leonardis, J. Matas, M. Felsberg, R. Pflugfelder
Location: Theatre Tuschinski / Room 3

Please visit the website for the latest information on this workshop; www.eccv2016.org/workshops

09:10 – 09:20 Welcome and Opening

   Session 1

09:20 – 10:00 The Visual Object Tracking VOT2016 challenge results
10:00 – 10:20 Talk
   Winner of VOT2016
10:20 – 10:40 Talk
   Best tracker VOT2016
10:40 – 11:10 Coffee break
11:10 – 11:55 Talk
   Invited speaker
11:55 – 12:15 Talk
   Contributed paper
12:15 – 12:45 Coffee break

   Session 2

12:45 – 13:05 The Visual Object Tracking VOT-TIR2016 challenge results
13:05 – 13:25 Talk
   Winner of the VOT-TIR challenge
13:25 – 14:00 Panel discussion
14:00 Closing, concluding remarks
W22
Web-scale Vision and Social Media
Organizers: L. Ballan, M. Bertini, T. Mensink
Location: Theatre Tuschinski / Room 4

Please visit the website for the latest information on this workshop; www.eccv2016.org/workshops

08.30 – 08:40 Welcome and opening
08:40 – 09:25 Invited speaker 1
Barbara Caputo
Robots learning about objects from the Web: where are we?

09:25 – 10:10 Oral session
OP1 Label-based Automatic Alignment of Video with Narrative Sentences, P. Dogan, M. Gross, J.-C. Bazin
OP2 Learning Joint Representations of Videos and Sentences with Web ImageSearch, M. Otani, Y. Nakashima, E. Rahtu, J. Heikkila, N. Yokoya
OP3 Cross-dimensional Weighting for Aggregated Deep Convolutional Features, Y. Kalantidis, C. Mellina, S. Osindero

10:10 – 10:40 Coffee break
10:40 – 11:00 Spotlight session
SP1 Towards Category Based Large-Scale Image Retrieval Using Transductive Support Vector Machines, H. Cevikalp, M. Elmas, S. Ozkan
SP2 Depth2Action: Exploring Embedded Depth for Large-Scale Action Recognition, Y. Zhu and S. Newsam
SP3 LOH and Behold: Web-scale visual search, recommendation and clustering using Locally Optimized Hashing, Y. Kalantidis, L. Kennedy, H. Nguyen, C. Mellina, D.A. Shamma
SP4 Solving multi-codebook quantization in the GPU, J. Martinez, H.H. Hoos, J.J. Little

11:00 – 11:45 Invited speaker 2
Vittorio Ferrari
Fun with annotators for object class detection

11:45 – 12:45 Poster session (ext. abstracts, spotlights and orals)
EA1 Faceless Person Recognition; Privacy Implications in Social Media, S J Oh, R Benenson, M Fritz, B Schiele
EA2 Estimating Inter-Class Visual Compatibility through Mid-level Elements, J Oramas, T Tuytelaars
EA3 Deep Convolutional Networks for Modeling Image Virality, A Dubey, S Agarwal
EA4 Tracking Natural Events through Social Media and Computer Vision, J Wang, M Korayem, S Blanco, D Crandall
EA5 Fast Zero-Shot Image Tagging, Y Zhang, B Gong, M Shah
Improving Multi-label Learning with Missing Labels by Structured Semantic Correlations, H Yang, J T Zhou, J Cai
Learning Visual Features from Large Weakly Supervised Data, A Joulin, L van der Maaten, A Jabri, N Vasilache
Learning global representations for image retrieval, A Gordo, J Almazan, J Revaud, D Larlus
Webly-supervised Video Recognition by Mutually Voting for Relevant Web Images and Web Video Frames, C Gan, C Sun, L Duan, B Gong
Video Stream Retrieval of Unseen Queries using Semantic Memory, S Cappallo, T Mensink, C Snoek
Learning the semantic structure of objects from Web supervision, D Novotny, D Larlus, A Vedaldi

12:45 – 13:30 Invited speaker 3
Program
Tuesday, October 11
Location: Theatre Carré / Room Theatre

08:45 – 09:00 Welcome

09:00 – 10:00 Oral Session 1A (Detection, Recognition & Retrieval)
Chairs: Barbara Caputo (University of Rome La Sapienza), Vincent Lepetit (TU Graz)

O-1A-1. **CNN Image Retrieval Learns from BoW: Unsupervised Fine-Tuning with Hard Examples**, Filip Radenovic, CMP, CVUT; Giorgos Tolias, CMP, CVUT; Ondra Chum, CMP, CVUT

O-1A-2. **SSD: Single Shot MultiBox Detector**, Wei Liu, UNC Chapel Hill; Dragomir Anguelov, Zoox; dumitru Erhan, Google; Christian Szegedy, Google; Scott Reed, University of Michigan, Ann-Arbor; Cheng-Yang Fu, UNC Chapel Hill; Alex Berg, UNC Chapel Hill

O-1A-3. **A Recurrent Encoder-Decoder Network for Sequential Face Alignment**, Xi Peng, Rutgers University; Rogerio Feris, IBM Research Center, USA; Xiaoyu Wang, Snapchat Research; Dimitris Metaxas, Rutgers University

O-1A-4. **Robust Facial Landmark Detection via Recurrent Attentive-Refinement Networks**, Shengtao Xiao, National University of Singapore; Jiashi Feng, NUS; Junliang Xing, Chinese Academy of Sciences; Hanjiang Lai, SUN YAT-SEN UNIVERSITY; Shuicheng Yan, National University of Singapore; Ashraf Kassim, National University of Singapore

10:00 – 10:30 Spotlight Session 1A (Segmentation)
Chairs: Vittorio Ferrari (University of Edinburgh), Cristian Sminchisescu (Lund University)

S-1A-5. **Learning to Refine Object Segments**, Pedro Pinheiro, EPFL; Tsung-Yi Lin, Cornell; Ronan Collobert, Facebook; Piotr Dollar, Facebook

S-1A-6. **Deep Automatic Portrait Matting**, Xiaoyong Shen, CUHK; Xin Tao, CUHK; Hongyun Gao, CUHK; Chao Zhou, ; Jiaya Jia, Chinese University of Hong Kong

S-1A-7. **Segmentation from Natural Language Expressions**, Ronghang Hu, UC Berkeley; Marcus Rohrbach, UC Berkeley; Trevor Darrell

S-1A-8. **Semantic Object Parsing with Graph LSTM**, Xiaodan Liang, Sun Yat-sen University; Xiaohui Shen, Adobe; Jiashi Feng, NUS; Liang Lin, Sun Yat-sen University; Shuicheng Yan, NUS

S-1A-9. **SSHMT: Semi-supervised Hierarchical Merge Tree for Electron Microscopy Image Segmentation**, Ting Liu, University of Utah; Miaomiao Zhang, MIT; Mehran Javanmardi, University of Utah; Nisha Ramesh, University of Utah; Tolga Tasdizen, University of Utah

10:30 – 10:35 Opening

10:35 – 11:00 Coffee Break

11:00 – 12:30 Poster Session 1A & Demos
Chairs: Jasper Uijlings (University of Edinburgh), Roberto Valenti (Sightcorp)

O-1A-1. **CNN Image Retrieval Learns from BoW: Unsupervised Fine-Tuning with Hard Examples**, Filip Radenovic, CMP, CVUT; Giorgos Tolias, CMP, CVUT; Ondra Chum, CMP, CVUT (shown in roof garden)
O-1A-2. **SSD: Single Shot MultiBox Detector**, Wei Liu, UNC Chapel Hill; Dragomir Anguelov, Zoox; Dumitru Erhan, Google; Christian Szegedy, Google; Scott Reed, University of Michigan; Ann Arbor; Cheng-Yang Fu, UNC Chapel Hill; Alex Berg, UNC Chapel Hill (shown in roof garden)

O-1A-3. **A Recurrent Encoder-Decoder Network for Sequential Face Alignment**, Xi Peng, Rutgers University; Rogerio Feris, IBM Research Center, USA; Xiaoyu Wang, Snapchat Research; Dimitris Metaxas, Rutgers University (shown in roof garden)

O-1A-4. **Robust Facial Landmark Detection via Recurrent Attentive-Refinement Networks**, Shengtao Xiao, National University of Singapore; Jiashi Feng, NUS; Junliang Xing, Chinese Academy of Sciences; Hanjiang Lai, SUN YAT-SEN UNIVERSITY; Shuicheng Yan, National University of Singapore; Ashraf Kassim, National University of Singapore (shown in roof garden)

S-1A-5. **Learning to Refine Object Segments**, Pedro Pinheiro, EPFL; Tsung-Yi Lin, Cornell; Ronan Collobert, Facebook; Piotr Dollar, Facebook

S-1A-6. **Deep Automatic Portrait Matting**, Xiaoyong Shen, CUHK; Xin Tao, CUHK; Hongyun Gao, CUHK; Chao Zhou, ; Jiaya Jia, Chinese University of Hong Kong

S-1A-7. **Segmentation from Natural Language Expressions**, Ronghang Hu, UC Berkeley; Marcus Rohrbach, UC Berkeley; Trevor Darrell

S-1A-8. **Semantic Object Parsing with Graph LSTM**, Xiaodan Liang, SUN YAT-SEN UNIVERSITY; Xiaohui Shen, Adobe; Jiashi Feng, NUS; Liang Lin, SUN YAT-SEN UNIVERSITY; Shuicheng Yan, NUS

S-1A-9. **SSHMT: Semi-supervised Hierarchical Merge Tree for Electron Microscopy Image Segmentation**, Ting Liu, University of Utah; Miaomiao Zhang, MIT; Mehran Javanmardi, University of Utah; Nisha Ramesh, University of Utah; Tolga Tasdizen, University of Utah

P-1A-10. **Fundamental Matrices from Moving Objects Using Line Motion Barcodes**, Yoni Kasten, Hebrew University; Gil Ben Artzi, Hebrew University of Jerusalem; Shmuel Peleg, Hebrew University; Michael Werman


P-1A-12. **Deep Learning the City: Quantifying Urban Perception At A Global Scale**, Abhimanyu Dubey, ; Nikhil Naik, MIT ; Devi Parikh, Virginia Tech; Ramesh Raskar, MIT Media Lab; Cesar Hidalgo, MIT

P-1A-13. **4D Match Trees for Non-rigid Surface Alignment**, Armin Mustafa, University of Surrey; Hansung Kim, U. Surrey; Adrian Hilton, University of Surrey

P-1A-14. **Eigen Appearance Maps of Dynamic Shapes**, Adnane Boukhayma, Inria; vogia Tsiminaki, Inria; Jean-Sebastien Franco, LJK, Université Grenoble Alpes, Inria; Edmond Boyer, Inria

P-1A-15. **Learnable Histogram: Statistical Context Features for Deep Neural Networks**, Zhe Wang, CUHK; Hongsheng Li, CUHK; Wanli Ouyang, Chinese University of Hong Kong; Xiaoqiang Wang, The Chinese University of Hong Kong

P-1A-16. **Pedestrian Behavior Understanding and Prediction with Deep Neural Networks**, Shuai Yi, CUHK; Hongsheng Li, CUHK; Xiaoqiang Wang, The Chinese University of Hong Kong

P-1A-17. **Real-time RGB-D Activity Prediction by Soft Regression**, Jian-Fang Hu, SUN YAT-SEN UNIVERSITY; Wei-Shi Zheng, ; Lianyang Ma, ; Gang Wang, ; Jianhuang Lai

P-1A-18. **A 3D Morphable Eye Region Model for Gaze Estimation**, Erroll Wood, University of Cambridge; Tadas Baltrušaitis, Carnegie Mellon University; Louis-Philippe Morency, Carnegie Mellon University; Peter Robinson, University of Cambridge; Andreas Bulling, Max Planck Institute for Informatics

Program | Tuesday, October 11 | Theatre Carré
P-1A-19. Foreground Segmentation via Dynamic Tree-Structured Sparse RPCA, Salehe Erfanian Ebadi, Queen Mary University of London; Ebroul Izquierdo, Queen Mary University of London

P-1A-20. Contextual Priming & Feedback for Faster R-CNN, Abhinav Shrivastava, Carnegie Mellon University; Abhinav Gupta

P-1A-21. Efficient Multi-view Surface Refinement with Adaptive Resolution Control, Shiwei Li, HKUST; Sing Yu Siu, HKUST; Tian Fang, HKUST; Long Quan, The Hong Kong University of Science and Technology, Hong Kong

P-1A-22. Gaussian Process Density Counting from Weak Supervision, Matthias von Borstel, Heidelberg University; Melih Kandemir, Heidelberg University; Philip Schmidt, Heidelberg University; Madhavi Kachur Rao, Robert Bosch; Kumar Rajamani, Bosch India; Fred Hamprecht, Heidelberg University

P-1A-23. Region-based semantic segmentation with end-to-end training, Holger Caesar, University of Edinburgh; Jasper Uijlings, Univ. of Edinburgh; Vittorio Ferrari

P-1A-24. Fast 6D Pose Estimation from a Monocular Image using Hierarchical Pose Trees, Yoshinori Konishi, OMRON Corporation; Yuki Hanzawa, OMRON Corporation; Masato Kawade, OMRON Corporation; Manabu Hashimoto, Chukyo University

P-1A-25. Learning Models for Actions and Person-Object Interactions with Transfer to Question Answering, Arun Mallya, UIUC; Svetlana Lazebnik

P-1A-26. A Software Platform for Manipulating the Camera Imaging Pipeline, Hakki Karaimer, NUS; Michael Brown, NUS

P-1A-27. A Benchmark and Simulator for UAV Tracking, Matthias Mueller, KAUST; Bernard Ghanem, KAUST; Neil Smith, KAUST

P-1A-28. Scene depth profiling using Helmholtz Stereopsis, Hironori Mori, Sony Deutschland GmbH; Roderick Koehle, ; Markus Kamm,

P-1A-29. Projective Bundle Adjustment from Arbitrary Initialization using the Variable Projection Method, Je Hyeong Hong, University of Washington; Christopher Zach, Toshiba Research Europe; Andrew Fitzgibbon, Microsoft Research; Roberto Cipolla

P-1A-30. Localizing and Orienting Street Views Using Overhead Imagery, Nam Vo, Georgia Institute of Technology; James Hays, Georgia Institute of Technology

P-1A-31. Hollywood in Homes: Crowdsourcing Data Collection for Activity Understanding, Gunnar Sigurdsson, Carnegie Mellon University; Gul Varol, INRIA; Xiaolong Wang, CMU; Ali Farhadi, University of Washington; Ivan Laptev, ; Abhinav Gupta

P-1A-32. Shuffle and Learn: Unsupervised Learning using Temporal Order Verification, Ishan Misra, CMU; Larry Zitnick, ; Martial Hebert, Carnegie Mellon University

P-1A-33. DOC: Deep OCclusion Estimation From a Single Image, Peng Wang, UCLA; Alan Yuille, UCLA

P-1A-34. RepMatch: Robust Feature Matching and Pose for Reconstructing Modern Cities, Wen-Yan Lin, ADSC; Siying Liu, UIUC; Nianjuan Jiang, Advanced Digital Sciences Center (ADSC); Minh Do, University of Illinois at Urbana-Champaign; ping tan, ; Jiangbo Lu, ADSC Singapore

P-1A-35. Convolutional Oriented Boundaries, Keviss-Kokitsi Maninis, ETH Zürich; Jordi Pont-Tuset, ETHZ; Pablo Arbelaez, Universidad de los Andes, Colombia; Luc Van Gool, ETH Zurich

P-1A-36. Superpixel Convolutional Networks using Bilateral Inceptions, Raghudeep Gadde, Ecole des Ponts Paris Tech; Varun Jampani, MPI-IS; Martin Kiefel, MPI for Intelligent Systems; Daniel Kappler, MPI Intelligent Systems; Peter Gehler

P-1A-37. Sublabel-Accurate Convex Relaxation of Vectorial Multilabel Energies, Emanuel Laude, Technical University of Munich; Thomas Möllenhoff, ; Michael Moeller, TUM; Jan Lellmann, ; Daniel Cremers
P-1A-38. **Building Dual-Domain Representations for Compression Artifacts Reduction**, Jun Guo, Sun Yat-Sen University; Hongyang Chao, Sun Yat-sen University

P-1A-39. **Geometric Neural Phrase Pooling: Modeling the Spatial Co-occurrence of Neurons**, Lingxi Xie, UCLA; Qi Tian, ; John Flynn, UCLA; Jingdong Wang, Microsoft Research; Alan Yuille, UCLA

P-1A-40. **Photo Aesthetics Ranking Network with Attributes and Content Adaptation**, Shu Kong, UCI; Xiaohui Shen, Adobe; Zhe Lin, Adobe Systems, Inc.; Radomir Mech, Adobe; Charless Fowlkes

P-1A-41. **SDF-2-SDF: Highly Accurate 3D Object Reconstruction**, Miroslava Slavcheva, Siemens AG; Wadim Kehl, TU München; Nassir Navab, ; Slobodan Ilic, TUM

P-1A-42. **Knowledge transfer for scene-specific motion prediction**, Lamberto Ballan, Stanford University; Francesco Castaldo, Seconda Università di Napoli; Alexandre Alahi, Stanford University; Francesco Palmieri, Seconda Università di Napoli; Silvio Savarese


P-1A-44. **Embedding Deep Metric for Person Re-identification: A Study Against Large Variations**, Hailin Shi, NLPR; Yang Yang, Institute of Automation; Shengcai Liao, Institute of Automation, Chinese Academy of Sciences; Zhen Lei, ; Wei-Shi Zheng, ; Stan Li

P-1A-45. **Learning to Track at 100 FPS with Deep Regression Networks**, David Held, UC Berkeley; Sebastian Thrun, Stanford; Silvio Savarese

P-1A-46. **Matching Handwritten Document Images**, Praveen Krishnan, IIIT H; C.V. Jawahar, IIIT Hyderabad

P-1A-47. **Semantic Clustering for Robust Fine-Grained Scene Recognition**, Marian George, ETH Zurich; Dixit Mandar, University of California, San Diego; Gábor Zogg, ETH Zurich; Nuno Vasconcelos

**Demos**

D-1A-48. **Real-Time Monocular Segmentation and Pose Tracking of Multiple Objects**, Henning Tjaden, RheinMain University of Applied Sciences; Ulrich Schwanecke, RheinMain University of Applied Sciences; Elmar Schömer, Johannes Gutenberg Universität Mainz (shown in roof garden)

D-1A-49. **Live Template-based 3D tracking and 3D reconstruction of deformable objects in 2D videos**, Toby Collins, ALCoV, Université d’Auvergne; Adrien Bartoli, ALCoV, Université d’Auvergne (shown in roof garden)

**Program**

<table>
<thead>
<tr>
<th>Tuesday, October 11</th>
<th>Theatre Carré</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:30 – 14:00</td>
<td>Lunch Break</td>
</tr>
<tr>
<td>14:00 – 15:00</td>
<td>Oral Session 1B (Scene Understanding)</td>
</tr>
<tr>
<td></td>
<td>Chairs: Cordelia Schmid (INRIA), Abhinav Gupta (CMU)</td>
</tr>
</tbody>
</table>

O-1B-1. **Ambient sound provides supervision for visual learning**, Andrew Owens, MIT; Jiajun Wu, MIT; Josh McDermott, MIT; Antonio Torralba, MIT; William Freeman, MIT

O-1B-2. **Grounding of Textual Phrases in Images by Reconstruction**, Anna Rohrbach; Marcus Rohrbach, UC Berkeley; Ronghang Hu, UC Berkeley; Trevor Darrell, UC Berkeley; Bernt Schiele

O-1B-3. **Improving Multi-label Learning with Missing Labels by Structured Semantic Correlations**, Hao Yang, NTU; Joey Tianyi Zhou, IHPC; Jianfei Cai, NTU
O-1B-4. **Visual Relationship Detection with Language Priors**, Cewu Lu, Stanford University; Ranjay Krishna, Stanford University; Michael Bernstein, Stanford University; Fei-Fei Li, Stanford University

15:00 – 15:30 **Spotlight Session 1B (Recognition & Scene Understanding)**

**Chairs:** Svetlana Lazebnik (UIUC), Andreas Geiger (MPI)

- **S-1B-5.** **The Curious Robot: Learning Visual Representations via Physical Interactions**, Lerrel Pinto, Carnegie Mellon University; Dhiraj Gandhi; Yuanfeng Han; Yong-Lae Park; Abhinav Gupta
- **S-1B-6.** **Image Co-localization by Mimicking a Good Detector’s Confidence Score Distribution**, Yao Li, University of Adelaide; Lingqiao Liu, University of Adelaide; Chunhua Shen, University of Adelaide; Anton Van den Hengel, University of Adelaide
- **S-1B-7.** **Facilitating and Exploring Planar Homogeneous Texture for Indoor Scene Understanding**, Shahzor Ahmad, National University of Singapore; Loong-Fah Cheong, National University of Singapore
- **S-1B-8.** **An Empirical Study and Analysis of Generalized Zero-Shot Learning for Object Recognition in the Wild**, Wei-Lun Chao, USC; Soravit Changpinyo, U. of Southern California; Boqing Gong, University of Central Florida; Fei Sha, UCLA
- **S-1B-9.** **Modeling Context in Referring Expressions**, Licheng Yu, University of North Carolina; Patrick Poirson; Shang Yang; Alex Berg; Tamara Berg, University on North Carolina

15:30 – 16:00 **Coffee Break**

16:00 – 17:30 **Poster Session 1B & Demos**

**Chairs:** Jasper Uijlings (University of Edinburgh), Roberto Valenti (Sightcorp)

- **O-1B-1.** **Ambient sound provides supervision for visual learning**, Andrew Owens, MIT; Jiajun Wu, MIT; Josh McDermott, MIT; Antonio Torralba, MIT; William Freeman, MIT (shown in roof garden)
- **O-1B-2.** **Grounding of Textual Phrases in Images by Reconstruction**, Anna Rohrbach; Marcus Rohrbach, UC Berkeley; Ronghang Hu, UC Berkeley; Trevor Darrell, UC Berkeley; Bernt Schiele (shown in roof garden)
- **O-1B-3.** **Improving Multi-label Learning with Missing Labels by Structured Semantic Correlations**, Hao Yang, NTU; Joey Tianyi Zhou, IHPC; Jianfei Cai, NTU (shown in roof garden)
- **O-1B-4.** **Visual Relationship Detection with Language Priors**, Cewu Lu, Stanford University; Ranjay Krishna, Stanford University; Michael Bernstein, Stanford University; Fei-Fei Li, Stanford University (shown in roof garden)
- **S-1B-5.** **The Curious Robot: Learning Visual Representations via Physical Interactions**, Lerrel Pinto, Carnegie Mellon University; Dhiraj Gandhi; Yuanfeng Han; Yong-Lae Park; Abhinav Gupta
- **S-1B-6.** **Image Co-localization by Mimicking a Good Detector’s Confidence Score Distribution**, Yao Li, University of Adelaide; Lingqiao Liu, University of Adelaide; Chunhua Shen, University of Adelaide; Anton Van den Hengel, University of Adelaide
- **S-1B-7.** **Facilitating and Exploring Planar Homogeneous Texture for Indoor Scene Understanding**, Shahzor Ahmad, National University of Singapore; Loong-Fah Cheong, National University of Singapore
<table>
<thead>
<tr>
<th>Program</th>
<th>Tuesday, October 11</th>
<th>Theatre Carré</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-1B-8. An Empirical Study and Analysis of Generalized Zero-Shot Learning for Object Recognition in the Wild</td>
<td>Wei-Lun Chao, USC; Soravit Changpinoyo, U. of Southern California; Boqing Gong, University of Central Florida; Fei Sha, UCLA</td>
<td></td>
</tr>
<tr>
<td>S-1B-9. Modeling Context in Referring Expressions</td>
<td>Licheng Yu, University of North Carolina; Patrick Poirson; Shang Yang; Alex Berg; Tamara Berg, University on North Carolina</td>
<td></td>
</tr>
<tr>
<td>P-1B-10. Taxonomy-Regularized Semantic Deep Convolutional Neural Networks</td>
<td>Wonjoon Goo, Seoul National University; Juyong Kim, Seoul National University; Gunhee Kim, Seoul National University; Sung Ju Hwang, UNIST</td>
<td></td>
</tr>
<tr>
<td>P-1B-11. Training a semantic Segmentation Model by playing a computer game</td>
<td>Stephan Richter, TU Darmstadt; Vibhav Vineet, Intel Labs; Stefan Roth, TU Darmstadt; Vladlen Koltun, Intel Labs</td>
<td></td>
</tr>
<tr>
<td>P-1B-12. Human Re-identification in Crowd Videos using Personal, Social and Environmental Constraints</td>
<td>Shayan Modiri Assari, University of Central Florida; Haroon Idrees, University of Central Florida; Mubarak Shah, University of Central Florida</td>
<td></td>
</tr>
<tr>
<td>P-1B-13. Revisiting additive quantization</td>
<td>Julieta Martinez, University of British Columbia; Joris Clement, University of British Columbia; Holger Hoos, University of British Columbia; Jim Little, University of British Columbia</td>
<td></td>
</tr>
<tr>
<td>P-1B-14. Single Image Dehazing via Multi-Scale Convolutional Neural Networks</td>
<td>Wenheng Ren, Tianjin University; Si Liu, Chinese Academy of Sciences; Hua Zhang, iie.ac.cn; Jinshan Pan, UC Merced; Xiaochun Cao, Chinese Academy of Sciences; Ming-Hsuan Yang, UC Merced</td>
<td></td>
</tr>
<tr>
<td>P-1B-15. Photometric Stereo under Non-uniform Light Intensities and Exposures</td>
<td>Donghyeon Cho, KAIST; Yasuyuki Matsushita; Yu-Wing Tai, Korea Advanced Institute of Science and Technology; In So Kweon, KAIST</td>
<td></td>
</tr>
<tr>
<td>P-1B-16. Visual Motif Discovery via First-Person Vision</td>
<td>Ryo Yonetani, University of Tokyo; Kris Kitani, Carnegie Mellon University; Yoichi Sato</td>
<td></td>
</tr>
<tr>
<td>P-1B-17. A Cluster Sampling Method for Image Matting via Sparse Coding</td>
<td>Xiaoxue Feng, Beihang University; Xiaoquliang, State Key Lab. of Virtual Reality Technology and Systems; Zili Zhang, State Key Lab. of Virtual Reality Technology and Systems</td>
<td></td>
</tr>
<tr>
<td>P-1B-18. What's the Point: Semantic Segmentation with Point Supervision</td>
<td>Amy Bearman, Stanford University; Olga Russakovsky, Stanford University; Vittorio Ferrari, ; Fei-Fei Li, Stanford University</td>
<td></td>
</tr>
<tr>
<td>P-1B-19. Fashion Landmark Detection in the Wild</td>
<td>Ziwei Liu, The Chinese Univ. of Hong Kong; Sijie Yan, The Chinese Univ. of Hong Kong; Ping Luo, The Chinese University of Hong Kong; Xiaogang Wang, The Chinese University of Hong Kong; Xiaou Tang, Chinese University of Hong Kong</td>
<td></td>
</tr>
<tr>
<td>P-1B-20. Person Re-identification by Unsupervised \ell_1 Graph Learning</td>
<td>Elyor Kodirov, QMUL; Tao Xiang, Queen Mary University of London; Zhenyong Fu; Shaogang Gong, Queen Mary University</td>
<td></td>
</tr>
<tr>
<td>P-1B-21. Leveraging Visual Question Answering for Image-Caption Ranking</td>
<td>Xiao Lin, Virginia Tech; Devi Parikh, Virginia Tech</td>
<td></td>
</tr>
<tr>
<td>P-1B-22. DAVE: A Unified Framework for Fast Vehicle Detection and Annotation</td>
<td>Yi Zhou, Northumbria University; Li Liu, Northumbria University; Ling Shao, Northumbria University; Matt Mellor, Createc Ltd.</td>
<td></td>
</tr>
<tr>
<td>P-1B-23. Real-time Joint Tracking of a Hand Manipulating an Object from RGB-D Input</td>
<td>Srinath Sridhar, MPI for Informatics; Franziska Mueller, MPI Informatics; Michael Zollhoefer, MPI Informatics; Dan Casas, MPI; Antti Oulasvirta, Aalto University; Christian Theobalt, MPI Informatics</td>
<td></td>
</tr>
<tr>
<td>P-1B-24. DeepWarp: Photorealistic Image Resynthesis for Gaze Manipulation</td>
<td>Yaroslav Ganin, Skoltech; Daniil Kononenko, Skoltech; Diana Sungatullina, Skoltech; Victor Lempitsky, Skolkovo Institute of Science and Technology</td>
<td></td>
</tr>
</tbody>
</table>
P-1B-25. **Non-Rigid 3D Shape Retrieval via Large Margin Nearest Neighbor Embedding**, Ioannis Chiotellis, TUM; Rudolph Triebel, TU Munich; Thomas Windheuser, TU Muenchen; Daniel Cremers

P-1B-26. **Multi-Task Zero-Shot Action Recognition with Prioritised Data Augmentation**, Xun Xu, Queen Mary University of London; Timothy Hospedales, Queen Mary University of London; Shaogang Gong, Queen Mary University

P-1B-27. **Fine-grained material classification using micro-geometry and reflectance**, Christos Kampouris, Imperial College London; Stefanos Zafeiriou, Imperial College London; Abhijeet Ghosh, Imperial College London; Sotirios Malassiotis, Centre for Research and Technology Hellas

P-1B-28. **Support Discrimination Dictionary Learning for Image Classification**, Yang Liu, University of Cambridge; Wei Chen, University of Cambridge; Qingchao Chen, University College London; Ian Wassell, University of Cambridge

P-1B-29. **Accelerating the Super-Resolution Convolutional Neural Network**, CHAO DONG, The Chinese University of HK; Chen-Change Loy, the Chinese University of Hong Kong; Xiaoou Tang, Chinese University of Hong Kong

P-1B-30. **Symmetric Non-Rigid Structure from Motion for Category-Specific Object Structure Estimation**, Yuan Gao, City University of Hong Kong; Alan Yuille, UCLA

P-1B-31. **Peak-Piloted Deep Network for Facial Expression Recognition**, Xiangyun Zhao, University of California, San Diego; Xiaodan Liang, Sun Yat-sen University; Luoqi Liu, Qihoo/360; Teng Li, Anhui University; Yugang Han, 360 AI Institute; Nuno Vasconcelos, ; Shuicheng Yan

P-1B-32. **Is Faster R-CNN Doing Well for Pedestrian Detection?**, Liliang Zhang, Sun Yat-sen University; Liang Lin, Sun Yat-sen University; Xiaodan Liang, Sun Yat-sen University; Kaiming He, Microsoft Research Asia

P-1B-33. **Coarse-to-fine Planar Regularization for Dense Monocular Depth Estimation**, Stephan Liwicki, ; Christopher Zach, Toshiba Research Europe; Ondrej Miksik, University of Oxford; Philip Torr, Oxford University

P-1B-34. **Deep Attributes Driven Person Re-identification**, Chi Su, Peking University; Shiliang Zhang, Peking University; Junliang Xing, Chinese Academy of Sciences; Wen Gao, Peking University; Qi Tian

P-1B-35. **An Occlusion-Resistant Ellipse Detection Method by Joining Coelliptic Arcs**, Halil Cakir, Dumlupinar University; Cihan Topal, ; Cuneyt Akinlar

P-1B-36. **Branching Path Following for Graph Matching**, Tao Wang, Beijing Jiaotong University; Haibin Ling, Temple University; Congyan Lang, Beijing Jiaotong University; Jun Wu, Beijing Jiaotong University

P-1B-37. **Higher Order Conditional Random Fields in Deep Neural Networks**, Anurag Arnab, University of Oxford; Sadeep Jayasumana, University of Oxford; Shuai Zheng, University of Oxford; Philip Torr, Oxford University

P-1B-38. **LSTM-CF: Unifying Context Modeling and Fusion with LSTMs for RGB-D Scene Labeling**, Zhen Li, The University of Hong Kong; Yukang Gan, Sun Yat-sen University; Xiaodan Liang, Sun Yat-sen University; Yizhou Yu, The University of Hong Kong; Hui Cheng, ; Liang Lin, Sun Yat-sen University

P-1B-39. **Stereo Video Deblurring**, Anita Sellent, Technische Universität Darmstadt, Technische Universität Dresden; Carsten Rother, ; Stefan Roth, TU Darmstadt
P-1B-40. **Robust Image and Video Dehazing with Visual Artifact Suppression via Gradient Residual Minimization**, Chen Chen, UIUC; Minh Do, University of Illinois at Urbana-Champaign; Jue Wang, Adobe Research

P-1B-41. **Smooth Neighborhood Structure Mining on Multiple Affinity Graphs with Applications to Context-sensitive Similarity**, Song Bai, HUST; Shaoyan Sun, University of Science and Technology of China; Xiang Bai, Huazhong University of Science and Technology; Zhaoxiang Zhang, Institute of Automation, Chinese Academy of Sciences; Qi Tian

P-1B-42. **Title Generation for User Generated Videos**, Kuo-Hao Zeng, National Tsing Hua University; Tseng-Hung Chen, National Tsing Hua University; Juan Carlos Niebles, Stanford University; Min Sun, National Tsing Hua University

P-1B-43. **Natural Image Matting using Deep Convolutional Neural Networks**, Donghyeon Cho, KAIST; Yu-Wing Tai, Korea Advanced Institute of Science and Technology; In So Kweon, KAIST

P-1B-44. **Double-Opponent Vectorial Total Variation**, Freddie Astrom, Heidelberg University; Christoph Schnoer, Heidelberg University

P-1B-45. **Learning to Count with CNN Boosting**, Elad Walach, Tel Aviv University; Lior Wolf

P-1B-46. **Amodal Instance Segmentation**, Ke Li, UC Berkeley; Jitendra Malik, UC Berkeley

P-1B-47. **Perceptual Losses for Real-Time Style Transfer and Super-Resolution**, Justin Johnson, Stanford University; Alexandre Alahi, Stanford University; Fei-Fei Li, Stanford University

**Demos**

D-1B-48. **Object Detection and Retrieval using Natural Language**, Ronghang Hu, UC, Berkeley and ICSI; XingChao Peng, Boston University; Tao Zhou, Canon USA Innovation Center; Jie Yu, Canon USA Innovation Center; Sandra Skaff, Canon USA Innovation Center; Trevor Darrell, UC, Berkeley and ICSI; Kate Saenko, Boston University (shown in roof garden)

D-1B-49. **ZFace**, László A. Jeni, Carnegie Mellon University; Jeffrey F. Cohn, Carnegie Mellon University and University of Pittsburgh; Takeo Kanade, Carnegie Mellon University (shown in roof garden)

**18:00 – 19:30** **Welcome Reception in Theatre Carré**
Program
Wednesday, October 12
Location: Theatre Carré / Room Theatre

09:00 – 10:00 Oral Session 2A (Optimization)
Chairs: Carsten Rother (TU Dresden), Ramin Zabih (Cornell)

O-2A-01. An Efficient Fusion Move Algorithm for the Minimum Cost Lifted Multicut Problem, Thorsten Beier, University of Heidelberg; Bjoern Andres, Max-Planck Institute for Informatics; Ullrich Koethe, University of Heidelberg; Fred Hamprecht, Heidelberg University

O-2A-02. Sparse Subspace Clustering, Yingzhen Yang, UIUC; Jiashi Feng, NUS; Nebojsa Jojic, Microsoft Research; Jianchao Yang, Snapchat; Thomas Huang, UIUC

O-2A-03. Normalized Cut meets MRF, Meng Tang, UWO; Dmitrii Marin, UWO; Ismail Ben Ayed, École de technologie supérieure; Yuri Boykov, University of Western Ontario

O-2A-04. Fast Global Registration, Qian-Yi Zhou, Intel Labs; Jaesik Park, Intel Labs; Vladlen Koltun, Intel Labs

10:00 – 10:30 Spotlight Session 2A (Optimization & Mathematical Models)
Chairs: Frederik Kahl (Chalmers University of Technology), Christoph Lampert (IST Austria)

S-2A-05. Polysemous Codes, Matthijs Douze; Herve Jegou; Florent Perronnin, Facebook

S-2A-06. Binary Hashing with Semidefinite Relaxation and Augmented Lagrangian, Thanh-Toan Do, SUTD; Dung Doan, SUTD; Duc-Thanh Nguyen, University of Wollongong; Ngai-Man Cheung, SUTD

S-2A-07. Efficient Continuous Relaxations for Dense CRF, Rudy Bunel, University of Oxford; Alban Desmaison, University of Oxford; M. Pawan Kumar, University of Oxford; Philip Torr, Oxford University; Pushmeet Kohli, Microsoft Research Cambridge

S-2A-08. Complexity of Discrete Energy Minimization Problems, Mengtian Li, Carnegie Mellon University; Alexander Shekhovtsov, Graz University of Technology; Daniel Huber

S-2A-09. A Convex Solution to Spatially-Regularized Correspondence Problems, homas Windheuser, TU Muenchen; Daniel Cremers

10:30 – 11:00 Coffee Break

11:00 – 12:30 Poster Session 2A & Demos
Chairs: Jasper Uijlings (University of Edinburgh), Roberto Valenti (Sightcorp)

O-2A-01. An Efficient Fusion Move Algorithm for the Minimum Cost Lifted Multicut Problem, Thorsten Beier, University of Heidelberg; Bjoern Andres, Max-Planck Institute for Informatics; Ullrich Koethe, University of Heidelberg; Fred Hamprecht, Heidelberg University (shown in roof garden)

O-2A-02. Sparse Subspace Clustering, Yingzhen Yang, UIUC; Jiashi Feng, NUS; Nebojsa Jojic, Microsoft Research; Jianchao Yang, Snapchat; Thomas Huang, UIUC (shown in roof garden)

O-2A-03. Normalized Cut meets MRF, Meng Tang, UWO; Dmitrii Marin, UWO; Ismail Ben Ayed, École de technologie supérieure; Yuri Boykov, University of Western Ontario (shown in roof garden)
O-2A-04. **Fast Global Registration**, Qian-Yi Zhou, Intel Labs; Jaesik Park, Intel Labs; Vladlen Koltun, Intel Labs (shown in roof garden)

S-2A-05. **Polysemous Codes**, Matthijs Douze; Herve Jegou; Florent Perronnin, Facebook

S-2A-06. **Binary Hashing with Semidefinite Relaxation and Augmented Lagrangian**, Thanh-Toan Do, SUTD; Dung Doan, SUTD; Duc-Thanh Nguyen, University of Wollongong; Ngai-Man Cheung, SUTD

S-2A-07. **Efficient Continuous Relaxations for Dense CRF**, Rudy Bunel, University of Oxford; Alban Desmaison, University of Oxford; M. Pawan Kumar, University of Oxford; Philip Torr, Oxford University; Pushmeet Kohli, Microsoft Research Cambridge

S-2A-08. **Complexity of Discrete Energy Minimization Problems**, Mengtian Li, Carnegie Mellon University; Alexander Shekhovtsov, Graz University of Technology; Daniel Huber

S-2A-09. **A Convex Solution to Spatially-Regularized Correspondence Problems**, homas Windheuser, TU Muenchen; Daniel Cremers

P-2A-10. **Jensen Bregman LogDet Divergence Optimal Filtering in the Manifold of Positive Definite Matrices**, Yin Wang, northeastern University; Octavia Camps, Northeastern University; Mario Sznaier, Northeastern University; Biel Roig Solvas, Northeastern University


P-2A-12. **Faceless Person Recognition; Privacy Implications in Social Media**, Seong Joon Oh, MPI-INF; Rodrigo Benenson, MPI Informatics; Mario Fritz, ; Bernt Schiele

P-2A-13. **Segmental Spatiotemporal CNNs for Fine-grained Action Segmentation**, Colin Lea, Johns Hopkins University; Austin Reiter, Johns Hopkins University; Rene Vidal, Johns Hopkins University; Gregory Hager, Johns Hopkins University

P-2A-14. **Structure from motion on a sphere**, Jonathan Ventura, University of Colorado Colorado

P-2A-15. **Evaluation of LBP and Deep Texture Descriptors with A New Robustness Benchmark**, Li Liu, NUDT; Paul Fieguth, ; Xiaogang Wang, The Chinese University of Hong Kong; Matti Pietikainen, University of Oulu; Dewen Hu

P-2A-16. **MS-Celeb-1M: A Dataset and Benchmark for Large Scale Face Recognition**, Yandong Guo, Microsoft Corporation; Lei Zhang, Microsoft Corporation; Yuxiao Hu, Microsoft Corporation; Jianfeng Gao, Microsoft Corporation

P-2A-17. **Hierarchical Beta Process with Gaussian Process prior for Hyperspectral Image Super Resolution**, Naveed Akhtar, Uni. of Western Australia; Faisal Shafait, ; Ajmal Mian, UWA

P-2A-18. **A 4D Light-Field Dataset and CNN Architectures for Material Recognition**, Ting-Chun Wang, UC Berkeley; Jun-Yan Zhu, UC BERKELEY; Hiroaki Ebi, University of California, San ; Manmohan Chandraker, NEC Labs America; Alexei Efros, ; Ravi Ramamoorthi

P-2A-19. **Graph-Based Consistent Matching for Structure-from-Motion**, Tianwei Shen, HKUST; Siyu Zhu, HKUST; Tian Fang, HKUST; Runze Zhang, HKUST; Long Quan, The Hong Kong University of Science and Technology

P-2A-20. **All-around Depth from Small Motion with A Spherical Panoramic Camera**, Sunghoon Im, KAIST; Hyowon Ha, KAIST; Francois Rameau, ; Hae-Gon Jeon, KAIST; Gyeongmin Choe, KAIST; In So Kweon, KAIST

P-2A-21. **On Shape Reconstruction from Implicit Forms**, Li Wang, Inria; Franck Hétroy-Wheeler, University Grenoble Alpes; Edmond Boyer, Inria

P-2A-22. **Multi-attributed Graph Matching with Multi-layer Random Walks**, Han-Mu Park, GIST; Kuk-Jin Yoon, Gwangju Institute of Science and Technology

P-2A-23. **Deep Learning of Local RGB-D Patches for 3D Object Detection and 6D Pose Estimation**, Wadim Kehl, TU München; Fausto Milletari, Technische Universität München; Federico Tombari, University of Bologna; Slobodan Ilic, TUM; Nassir Navab

<table>
<thead>
<tr>
<th>Program</th>
<th>Wednesday, October 12</th>
<th>Theatre Carré</th>
</tr>
</thead>
</table>

P-2A-25. **Joint Face Representation Adaptation and Clustering in Videos**, Zhanpeng Zhang, The Chinese University of HK; Ping Luo, The Chinese University of Hong Kong; Chen-Change Loy, the Chinese University of Hong Kong; Xiaou Tang, Chinese University of Hong Kong


P-2A-27. **ATGV-Net: Accurate Depth Super-Resolution**, Gernot Riegler, Graz University of Technology; Matthias Rüther, Graz University of Technology; Horst Bischof, Graz University of Technology

P-2A-28. **Indoor-Outdoor 3D Reconstruction Alignment**, Andrea Cohen, ETHZ; Johannes Schönberger, ETH Zürich; Pablo Speciale, ETHZ; Torsten Sattler, ETH Zurich; Jan-Michael Frahm, ; Marc Pollefeys, ETH

P-2A-29. **The Unreasonable Effectiveness of Noisy Data for Fine-Grained Recognition**, Jonathan Krause, Stanford University; Benjamin Sapp, Google; Andrew Howard, Google; Howard Zhou, Google; Alexander Toshev, Google; Tom Duerig, Google; James Philbin, Google; Fei-Fei Li, Stanford University

P-2A-30. **A simple hierarchical pooling data structure for loop closure**, Xiaohan Fei, UCLA; Konstantine Tsotsos, ; Stefano Soatto, University of California, Los Angeles


P-2A-32. **Depth Map Super Resolution by Deep Multi-Scale Guidance**, Tak-Wai Hui, The Chinese University of HK; Chen-Change Loy, the Chinese University of Hong Kong; Xiaou Tang, Chinese University of Hong Kong

P-2A-33. **SEAGULL: Seam-guided Local Alignment for Parallax-tolerant Image Stitching**, Kaimo Lin, NUS; Nianjuan Jiang, Advanced Digital Sciences Center (ADSC); Loong-Fah Cheong, National University of Singapore; Minh Do, University of Illinois at Urbana-Champaign; Jiangbo Lu, ADSC Singapore

P-2A-34. **Grid Loss: Detecting Occluded Faces**, Michael Opitz, Graz University of Technology; Georg Waltner, TU Graz; Georg Poier, Graz University of Technology; Horst Possegger, Graz University of Technology; Horst Bischof, Graz University of Technology

P-2A-35. **Large-scale R-CNN with Classifier Adaptive Quantization**, Ryota Hinami, The University of Tokyo; Shin’ichi Satoh, National Institute of Informatics, Japan

P-2A-36. **Face Detection with End-to-End Integration of a ConvNet and a 3D Model**, Yunzhu Li, Peking University; Benyuan Sun, PKU; Tianfu Wu, UCLA; Yizhou Wang, Peking University

P-2A-37. **Large Scale Asset Extraction for Urban Images**, Lama Affara, KAUST; Liangliang Nan, KAUST; Bernard Ghanem, KAUST; Peter Wonka

P-2A-38. **Multi-label Active Learning Based on Maximum Correntropy Criterion: Towards Robust and Discriminative Labeling**, Zengmao Wang, Wuhan University; Bo Du, ; Lefei Zhang, Wuhan University; Liangpei Zhang, Wuhan University; Meng Fang, ; Dacheng Tao, University of Technology, Sydney


P-2A-40. **Fine-Scale Surface Normal Estimation using a Single NIR Image**, Youngjin Yoon, KAIST; Gyeongmin Choe, KAIST; Namil Kim, KAIST; Joon-Young Lee, Adobe ; In So Kweon, KAIST

P-2A-41. **Pixelwise View Selection for Unstructured Multi-View Stereo**, Johannes Schönberger, ETH Zürich; Enliang Zheng, UNC Chapel Hill; Marc Pollefeys, ETH; Jan-Michael Frahm

P-2A-42. **Laplacian Pyramid Reconstruction and Refinement for Semantic Segmentation**, Golnaz Ghiasi, University of California, Irvine; Charless Fowlkes

P-2A-43. **Generic 3D Representation via Pose Estimation and Matching**, Amir Zamir, Stanford University; Pulkit Agrawal, UC Berkeley; Silvio Savarese, ; Tilman Wekel, Stanford University; Colin Wei, Stanford University; Jitendra Malik, UC Berkeley
P-2A-44. Hand pose estimation from local surface normals, Chengde Wan, ETHZ; Angela Yao, University of Bonn; Luc Van Gool, ETH Zurich

P-2A-45. Abundant Inverse Regression using Sufficient Reduction and its applications, Hyunwoo Kim, UW-Madison; Brandon Smith, UW-Madison; Nagesh Adluru, UW-Madison; Charles Dyer, University of Wisconsin – Madison; Sterling Johnson, UW-Madison; Vikas Singh, University of Wisconsin-Madison

P-2A-46. Learning Diverse Models: The Coulomb Structured Support Vector Machine, Martin Schiegg, University of Heidelberg; Robert Bosch GmbH; Ferran Diego, University of Heidelberg, HCI; Fred Hamprecht, Heidelberg University

P-2A-47. Pose Hashing with Microlens Arrays, Ian Schillebeeckx, Washington University in St. L; Robert Pless, Washington University in St. Louis

Demos

D-2A-48. Multi-person Body Pose Estimation, Eldar Insafutdinov, Max Planck Institute for Informatics; Leonid Pishchulin, Max Planck Institute for Informatics; Evgeny Levinkov, Max Planck Institute for Informatics; Bjoern Andres, Max Planck Institute for Informatics; Mykhaylo Andriluka, Max Planck Institute for Informatics; Bernt Schiele, Max Planck Institute for Informatics (shown in roof garden)

D-2A-49. Realtime Multiperson Pose Estimation, Zhe Cao, Carnegie Mellon University; Shih En Wei, Carnegie Mellon University; Tomas Simon, Carnegie Mellon University; Yaser Sheikh, Carnegie Mellon University (shown in roof garden)

12:30 – 14:00 Lunch Break

14:00 – 15:00 Oral Session 2B (Image & Video Processing)

Chairs: Larry Zitnick (Facebook), Peter Gehler (MPI)

O-2B-01. The Fast Bilateral Solver, Jonathan Barron, UC Berkeley; Ben Poole, Stanford University

O-2B-02. Phase-based Modification Transfer for Video, Simone Meyer, ETH Zurich; Alexander Sorkine-Hornung, Disney Research Zurich; Markus Gross, ETH Zurich

O-2B-03. Colorful Image Colorization, Richard Zhang, UC Berkeley; Phillip Isola, MIT; Alexei Efros

O-2B-04. Focal flow: Measuring depth and velocity from defocus and differential motion, Emma Alexander, Harvard University; Qi Guo, Harvard University; Sanjeev Koppal, University of Florida; Steven Gortler; Todd Zickler

15:00 – 15:30 Spotlight Session 2B (Image Processing & Computational Photography)

Chairs: Stefan Roth (TU Darmstadt), Joost van de Weijer (CVC)

S-2B-05. An evaluation of computational imaging techniques for heterogeneous inverse scattering, Ioannis Gkioulekas, Harvard University; Todd Zickler, Anat Levin, Weizmann Institute of Science

S-2B-06. Precomputed Real-Time Texture Synthesis with Markovian Generative Adversarial Networks, Chuan Li, University of Mainz; Michael Wand, University of Mainz

S-2B-07. Fast Guided Global Interpolation for Depth and Motion, Yu Li, Advanced Digital Sci. Center; Dongbo Min, Chungnam National University; Minh Do, University of Illinois at Urbana-Champaign; Jiangbo Lu, ADSC Singapore

S-2B-08. Learning High-Order Filters for Efficient Blind Deconvolution of Document Photographs, Lei Xiao, University of British Columbia; Jue Wang, Adobe Research; Wolfgang Heidrich, KAUST; Michael Hirsch

Program | Wednesday, October 12 | Theatre Carré
S-2B-09. **Multi-View Inverse Rendering under Arbitrary Illumination and Albedo**, Kichang Kim, Tokyo Institute of Technology; Akihiko Torii, Tokyo Institute of Technology; Masatoshi Okutomi, Tokyo Institute of Technology

15:30 – 16:00 **Coffee Break**

16:00 – 17:30 **Poster Session 2B & Demos**  
Chairs: Jasper Uijlings (University of Edinburgh), Roberto Valenti (Sightcorp)

O-2B-01. **The Fast Bilateral Solver**, Jonathan Barron, UC Berkeley; Ben Poole, Stanford University (shown in roof garden)

O-2B-02. **Phase-based Modification Transfer for Video**, Simone Meyer, ETH Zurich; Alexander Sorkine-Hornung, Disney Research Zurich; Markus Gross, ETH Zurich (shown in roof garden)

O-2B-03. **Colorful Image Colorization**, Richard Zhang, UC Berkeley; Phillip Isola, MIT; Alexei Efros (shown in roof garden)

O-2B-04. **Focal flow: Measuring depth and velocity from defocus and differential motion**, Emma Alexander, Harvard University; Qi Guo, Harvard University; Sanjeev Koppal, University of Florida; Steven Gortler; Todd Zickler (shown in roof garden)

S-2B-05. **An evaluation of computational imaging techniques for heterogeneous inverse scattering**, Ioannis Gkioulekas, Harvard University; Todd Zickler; Anat Levin, Weizmann Institute of Science

S-2B-06. **Precomputed Real-Time Texture Synthesis with Markovian Generative Adversarial Networks**, Chuan Li, University of Mainz; Michael Wand, University of Mainz

S-2B-07. **Fast Guided Global Interpolation for Depth and Motion**, Yu Li, Advanced Digital Sci. Center; Dongbo Min, Chungnam National University; Minh Do, University of Illinois at Urbana-Champaign; Jiangbo Lu, ADSC Singapore

S-2B-08. **Learning High-Order Filters for Efficient Blind Deconvolution of Document Photographs**, Lei Xiao, University of British Columbia; Jue Wang, Adobe Research; Wolfgang Heidrich, KAUST; Michael Hirsch

S-2B-09. **Multi-View Inverse Rendering under Arbitrary Illumination and Albedo**, Kichang Kim, Tokyo Institute of Technology; Akihiko Torii, Tokyo Institute of Technology; Masatoshi Okutomi, Tokyo Institute of Technology

P-2B-10. **DAPs: Deep Action Proposals for Action Understanding**, Victor Escorcia, KAUST; FABIAN CABA HEILBRON, KAUST; Juan Carlos Niebles, Stanford University; Bernard Ghanem, KAUST

P-2B-11. **A Large Contextual Dataset for Classification**, Detection and Counting of Cars with Deep Learning, Terrell Mundhenk, Lawrence Livermore National L.; Wesam Sakla, LLNL; Goran Konjevod, LLNL; Kofi Boakye, LLNL

P-2B-12. **Reliable Attribute-Based Object Recognition Using High Predictive Value Classifiers**, Wentao Luan, University of Maryland; Yezhou Yang, University of Maryland; Cornelia Fermuller, University of Maryland; John Baras, University of Maryland

P-2B-13. **Spatio-Temporal LSTM with Trust Gates for 3D Human Action Recognition**, Jun Liu, Nanyang Technological University; Amir Shahroudy, Nanyang Technological University; Dong Xu; ; Gang Wang

P-2B-14. **Going Further with Point Pair Features**, Stefan Hinterstoisser, Google; Vincent Lepetit; ; Kurt Konolige, Google; Naresh Rajkumar, Google
P-2B-15. **Webly-supervised Video Recognition by Mutually Voting for Relevant Web Images and Web Video Frames**, Chuang Gan, Tsinghua University; Chen Sun, USC; Lixin Duan, Amazon; Boqing Gong, University of Central Florida

P-2B-16. **HFS: Hierarchical Feature Selection for Efficient Image Segmentation**, Ming-Ming Cheng, Nankai University; Yun Liu, Nankai University; Qibin Hou, Nankai University; Jiawang Bian, Nankai University; Philip Torr, Oxford University; Shiming Hu, Tsinghua University; Zhenwu Tu

P-2B-17. **Generating Visual Explanations**, Lisa Anne Hendricks, UC Berkeley; Zeynep Akata, ; Marcus Rohrbach, UC Berkeley; Jeff Donahue, UC Berkeley; Bernt Schiele, ; Trevor Darrell

P-2B-18. **Marker-less 3D Human Motion Capture with Monocular Image Sequence and Height-Maps**, Yu Du, ; Yongkang Wong, ; Yonghao Liu, ; Feilin Han, Zhejiang University; Yilin Gui, ; Zhen Wang, ; Mohan Kankanhalli, National University of Singapore; Weidong Geng

P-2B-19. **Tensor Representations via Kernel Linearization for Action Recognition from 3D Skeletons**, Piotr Koniusz, NICTA; Anoop Cherian, ANU; Fatih Porikli, ANU / NICTA

P-2B-20. **Manhattan-world Urban Reconstruction from Point Clouds**, Minglei Li, Nanjing University of Aeronautics and Astronautics; Peter Wonka, ; Liangliang Nan, KAUST

P-2B-21. **From Multiview Image Curves to 3D Drawings**, Anil Usmuzbas, Brown University; Ricardo Fabbri, State University of Rio de Janeiro; Benjamin Kimia, Brown University

P-2B-22. **Shape from Selfies: Human Body Shape Estimation using CCA Regression Forests**, Endri Dibra, ETH Zurich; Cengiz Oztireli, ETH Zurich; Remo Ziegler, Vizrt; Markus Gross, ETH Zurich

P-2B-23. **Can we Jointly Register and Reconstruct Creased Surfaces by Shape-from-Template Accurately?**, Mathias Gallardo, ISIT, UMR 6284 CNRS/UdA; Toby Collins, Universite d’Auzeragne; Adrien Bartoli, Universite d’Auzeragne

P-2B-24. **Distractor-supported single target tracking in extremely cluttered scenes**, Jingjing Xiao, University of Birmingham; LINBO QIAO, ; Rustam Stolkin, ; Aleš Leonardis

P-2B-25. **Connectionist Temporal Modeling for Weakly Supervised Action Labeling**, De-An Huang, Stanford University; Fei-Fei Li, Stanford University; Juan Carlos Niebles, Stanford University

P-2B-26. **Deep Joint Image Filter**, Yijun Li, UC Merced; Jia-Bin Huang, University of Illinois, Urbana-Champaign: Narendra Ahuja, University of Illinois at Urbana-Champaign; Ming-Hsuan Yang, UC Merced

P-2B-27. **Efficient Multi-Frequency Phase Unwrapping using Kernel Density Estimation**, Felix Jaremo-Lawin, Linköping University; Per-Erik Forssen, Linköping University; Hannes Ovren, Linköping university

P-2B-28. **A Multi-Scale CNN for Affordance Segmentation in RGB Images**, Anirban Roy, Oregon State University; Sinisa Todorovic, Oregon State University

P-2B-29. **Hierarchical Dynamic Parsing and Encoding for Action Recognition**, Bing Su, Chinese Academy of Sciences; JIAHUAN ZHOU, Northwestern University; Xiaoxiao Ding, Tsinghua University; Hao Wang, Chinese Academy of Sciences; Ying Wu, Northwestern University


P-2B-31. **A Diagram Is Worth A Dozen Images**, Aniruddha Kembhavi, AI2; Michael Salvato, Allen Institute for Artificial; Eric Kolve, Allen Institute for AI; Minjoon Seo, University of Washington; Hannaneh Hajishirzi, University of Washington; Ali Farhadi, University of Washington
P-2B-32. **Automatic Attribute Discovery with Neural Activations**, Sirion Vittayakorn, University of North Carolina at Chapel Hill; Takayuki Umeda, NTT; Kazuhiko Murasaki, NTT; Kyoko Sudo, NTT; Takayuki Okatani, Tohoku University; Kota Yamaguchi, Tohoku University

P-2B-33. **“What happens if…” Learning to predict the effect of forces in images**, Roozbeh Mottaghi, Allen Institute for AI; Mohammad Rastegari, AI2; Abhinav Gupta, ; Ali Farhadi, University of Washington

P-2B-34. **View synthesis by appearance flow**, Tinghui Zhou, UC Berkeley; Shubham Tulsiani, UC Berkeley; Weilun Sun, UC Berkeley; Jitendra Malik, UC Berkeley; Alexei Efros

P-2B-35. **Top-down Learning for Structured Labeling with Convolutional Pseudoprior**, Saining Xie, UCSD; Xun Huang, ; Zhuowen Tu

P-2B-36. **Generative Image Modeling using Style and Structure Adversarial Networks**, Xiaolong Wang, CMU; Abhinav Gupta

P-2B-37. **Joint learning of Semantic and Latent Attributes**, Peixi Peng, Peking University; Yonghong Tian, Peking University; Tao Xiang, Queen Mary University of London; Yaowei Wang, Beijing Institute of Technology; Tiejun Huang, Peking University

P-2B-38. **A Unified Multi-scale Deep Convolutional Neural Network for Fast Object Detection**, Zhaowei Cai, ucsd.edu; Quanfu Fan, IBM; Rogerio Feris, IBM Research Center; Nuno Vasconcelos

P-2B-39. **Deep Specialized Network for Illuminant Estimation**, Wu Shi, The Chinese University of HK; Chen-Change Loy, the Chinese University of Hong Kong; Xiaou Tang, Chinese University of Hong Kong

P-2B-40. **Weakly-Supervised Semantic Segmentation using Motion Cues**, Pavel Tokmakov, INRIA Grenoble Rhône-Alpes; Karteek Alahari, Inria; Cordelia Schmid

P-2B-41. **Human-In-The-Loop Person Re-Identification**, Hanxiao Wang, Queen Mary, Univ. of London; Shaogang Gong, Queen Mary University; Xiatian Zhu, ; Tao Xiang, Queen Mary University of London

P-2B-42. **Real-Time Monocular Segmentation and Pose Tracking of Multiple Objects**, Henning Tjaden, RheinMain University of Applied Sciences; Ulrich Schwanecke, RheinMain University of Applied Sciences; Elmar Schömer, Johannes Gutenberg University Mainz

P-2B-43. **Estimation of Human Body Shape in Motion with Wide Clothing**, Jinlong Yang, Inria; Jean-Sebastien Franco, LJK, Université Grenoble Alpes, Inria; Franck Hétroy-Wheeler, University Grenoble Alpes; Stefanie Wuhrer, INRIA

P-2B-44. **A Shape-based Approach for Salient Object Detection Using Deep Learning**, Jongpil Kim, Rutgers Univ.; Vladimir Pavlovic, Rutgers University

P-2B-45. **Fast Optical Flow using Dense Inverse Search**, Till Kroeger, ETH Zuerich; Radu Timofte, ETH Zurich; Dengxin Dai, ETH Zurich; Luc Van Gool, ETH Zurich

P-2B-46. **Global Registration of 3D Point Sets via LRS decomposition**, Federica Arrigoni, University of Udine; Beatrice Rossi, AST; Andrea Fusiello, University of Udine

P-2B-47. **Recognition from Hand Cameras: A Revisit with Deep Learning**, Cheng-Sheng Chan, ; Shou-Zhong CHEN, ; Pei-Xuan Xie, ; CHIUNG-CHIH CHANG, ; Min Sun, National Tsing Hua University

**Demos**

D-2B-48. **Mobile AR for Dentistry: Virtual Try-On in Live 3D**, Gábor Sörös, Kapanu AG and ETH Zurich; Marcel Lancelle, Kapanu AG and ETH Zurich; Nicolas Degen, Kapanu AG and ETH Zurich; Roland Mörzinger, Kapanu AG and ETH Zurich (shown in roof garden)

D-2B-49. **Facial Tracking**, Nial Stewart, ULSee Inc. (shown in roof garden)
Program

Thursday, October 13

Location: Theatre Carré / Room Theatre

09:00 – 10:00  Oral Session 3A (Learning)

Chairs: Raquel Urtasun (University of Toronto), Iasonas Kokkinos (ECP)

O-3A-01. XNOR-Net: ImageNet Classification Using Binary Convolutional Neural Networks, Mohammad Rastegari, AI2; Vicente Ordonez, Allen Institute for AI; Joe Redmon; Ali Farhadi, University of Washington

O-3A-02. Top-down Neural Attention by Excitation Backprop, Jianming Zhang; Zhe Lin, Adobe Systems, Inc.; Jonathan Brandt; Xiaohui Shen, Adobe; Stan Sclaroff, Boston University

O-3A-03. Learning Recursive Filters for Low-Level Vision via a Hybrid Neural Network, Sifei Liu, UC Merced; Jinshan Pan, UC Merced; Ming-Hsuan Yang, UC Merced

O-3A-04. Learning Representations for Automatic Colorization, Gustav Larsson, University of Chicago; Michael Maire, Toyota Technological Institute at Chicago; Greg Shakhnarovich, TTI Chicago, USA

10:00 – 10:30  Spotlight Session 3A (CNNs)

Chairs: Laurens van der Maaten (Facebook), Xiaogang Wang (CUHK)

S-3A-05. Deep Reconstruction-Classification Networks for Unsupervised Domain Adaptation, Muhammad Ghifary, VUW, Weta Digital; Bastiaan Kleijn, Victoria University of Wellington; Mengjie Zhang, Victoria University of Wellington; David Balduzzi, Victoria University of Wellington; Wen Li, ETH Zurich

S-3A-06. Learning without Forgetting, Zhizhong Li, UIUC; Derek Hoiem, UIUC

S-3A-07. Identity Mappings in Deep Residual Networks, Kaiming He, Microsoft Research Asia; Xiangyu Zhang, Xi’an Jiaotong University; Shaoqing Ren, University of Science & Technology of China; Jian Sun, Microsoft Research China

S-3A-08. Deep Networks with Stochastic Depth, Gao Huang, Cornell University; Yu Sun, Cornell University; Zhuang Liu, Tsinghua University; Daniel Sedra, Cornell University; Kilian Weinberger, Cornell University

S-3A-09. Less is More: Towards Compact CNNs, Hao Zhou, University of Maryland; Jose M. Alvarez, Data61 / CSIRO; Fatih Porikli, Australian National University

10:30 – 11:00  Coffee Break

11:00 – 12:30  Poster Session 3A & Demos

Chairs: Jasper Uijlings (University of Edinburgh), Roberto Valenti (Sightcorp)

O-3A-01. XNOR-Net: ImageNet Classification Using Binary Convolutional Neural Networks, Mohammad Rastegari, AI2; Vicente Ordonez, Allen Institute for AI; Joe Redmon; Ali Farhadi, University of Washington (shown in roof garden)

O-3A-02. Top-down Neural Attention by Excitation Backprop, Jianming Zhang; Zhe Lin, Adobe Systems, Inc.; Jonathan Brandt; Xiaohui Shen, Adobe; Stan Sclaroff, Boston University (shown in roof garden)
O-3A-03. Learning Recursive Filters for Low-Level Vision via a Hybrid Neural Network, Sifei Liu, UC Merced; Jinshan Pan, UC Merced; Ming-Hsuan Yang, UC Merced (shown in roof garden)

O-3A-04. Learning Representations for Automatic Colorization, Gustav Larsson, University of Chicago; Michael Maire, Toyota Technological Institute at Chicago; Greg Shakhnarovich, TTI Chicago, USA (shown in roof garden)

S-3A-05. Deep Reconstruction-Classification Networks for Unsupervised Domain Adaptation, Muhammad Ghifary, VUW, Weta Digital; Bastiaan Kleijn, Victoria University of Wellington; Mengjie Zhang, Victoria University of Wellington; David Balduzzi, Victoria University of Wellington; Wen Li, ETH Zurich

S-3A-06. Learning without Forgetting, Zhizhong Li, UIUC; Derek Hoiem, UIUC

S-3A-07. Identity Mappings in Deep Residual Networks, Kaiming He, Microsoft Research Asia; Xiangyu Zhang, Xi’an Jiaotong University; Shaoqing Ren, University of Science & Technology of China; Jian Sun, Microsoft Research China

S-3A-08. Deep Networks with Stochastic Depth, Gao Huang, Cornell University; Yu Sun, Cornell University; Zhuang Liu, Tsinghua University; Daniel Sedra, Cornell University; Kilian Weinberger, Cornell University

S-3A-09. Less is More: Towards Compact CNNs, Hao Zhou, University of Maryland; Jose M. Alvarez, Data61 / CSIRO; Fatih Porikli, Australian National University

P-3A-10. Unsupervised Visual Representation Learning by Graph-based Consistent Constraints, Dong Li, Tsinghua University; Wei-Chih Hung, USC; Jia-Bin Huang, University of Illinois at Urbana-Champaign; Shengjin Wang, wgsj1@tsinghua.edu.cn; Narendra Ahuja, University of Illinois at Urbana-Champaign; Ming-Hsuan Yang, UC Merced

P-3A-11. Seed, Expand and Constrain: Three Principles for Weakly-Supervised Image Segmentation, Alexander Kolesnikov, IST Austria; Christoph Lampert, IST Austria

P-3A-12. Patch-based low-rank matrix completion for learning of shape and motion models from few training samples, Jan Ehrhardt, ; Matthias Wilms, University of Lübeck; Heinz Handels

P-3A-13. Chained Predictions Using Convolutional Neural Networks, Georgia Gkioxari, ; Navdeep Jaitly, Google; Alexander Toshev, Google

P-3A-14. Multi-region two-stream R-CNN for action detection, Xiaojiang Peng, INRIA; Cordelia Schmid

P-3A-15. Semantic Co-segmentation in Videos, Yi-Hsuan Tsai, UC Merced; Guangyu Zhong, Dalian University of Technology; Ming-Hsuan Yang, UC Merced

P-3A-16. Attribute2Image: Conditional Image Generation from Visual Attributes, Xinchen Yan, University of Michigan; Jimei Yang, Abode Research; Kihyuk Sohn, NEC Lab; Honglak Lee

P-3A-17. Modeling Context Between Objects for Referring Expression Understanding, Varun Nagaraja, University of Maryland; Vlad Morariu, University of Maryland; Larry Davis, University of Maryland

P-3A-18. Friction from Reflectance: Deep Reflectance Codes for Predicting Physical Surface Properties from One-Shot In-Field Reflectance, Hang Zhang, Rutgers University; Kristin Dana, Rutgers University; Ko Nishino, Drexel University

P-3A-19. Saliency Detection with Recurrent Fully Convolutional Networks, Linzhao Wang; Lijun Wang, Dalian University of Technology; Huchuan Lu, ; Pingping Zhang, Dalian University of Technology; Xiang Ruan

P-3A-20. Deep3D: Fully Automatic 2D-to-3D Video Conversion with Deep Convolutional Neural Networks, Junyuan Xie, University of Washington; Ross Girshick, ; Ali Farhadi, University of Washington

P-3A-21. Temporal Model Adaptation for Person Re-Identification, Niki Martinel, University of Udine; Abir Das, University of Massachusetts Lowell; Christian Micheloni, University of Udine; Amit Roy-Chowdhury, UC Riverside
P-3A-22. **Image Quality Assessment Using Similar Scene as Reference**, Yudong Liang, Institute of Artificial Intelligence and Robotics, Xi’an Jiaotong University; Jinjun Wang; Xingyu Wan, xjtu.edu.cn; Yihong Gong; Nanning Zheng, xjtu.edu.cn; Xingyu Wan, xjtu.edu.cn; Yihong Gong; Nanning Zheng, xjtu.edu.cn

P-3A-23. **MOON: A Mixed Objective Optimization Network for the Recognition of Facial Attributes**, Ethan Rudd, UCCS; Manuel Günther, UCCS; Terrance Boult, University of Colorado Colorado Springs

P-3A-24. **Degeneracies in Rolling Shutter SfM**, Cenek Albl, Czech Technical University; Akihiro Sugimoto, NII; Tomas Pajdla

P-3A-25. **Deep Deformation Network for Object Landmark Localization**, Xiang Yu, NEC Labs; Feng Zhou, NEC Lab; Manmohan Chandraker, NEC Labs America

P-3A-26. **Learning Visual Storylines with Skipping Recurrent Neural Networks**, Gunnar Sigurdsson, Carnegie Mellon University; Xinlei Chen, CMU; Abhinav Gupta

P-3A-27. **Towards large-scale city reconstruction from satellites**, liuyun Duan; Florent Lafarge


P-3A-29. **Supervised Transformer Network for Efficient Face Detection**, Dong Chen, Microsoft Research; Gang Hua, Microsoft Research; Fang Wen, Microsoft Research; Jian Sun, Microsoft Research China

P-3A-30. **A Geometric Approach to Image Labeling**, Freddie Astrom, Heidelberg University; Stefania Petra; Bernhard Schmitzer, Université Paris-Dauphine; Christoph Schnoerr, Heidelberg University


P-3A-32. **A Minimal Solution for Non-Perspective Pose Estimation from Line Correspondences**, Gim Hee Lee, NUS

P-3A-33. **Natural Image Stitching with the Global Similarity Prior**, Yu-Sheng Chen, National Taiwan University; Yung-Yu Chuang, National Taiwan University

P-3A-34. **Minimal Solvers for Generalized Pose and Scale Estimation from Two Rays and One Point**, Federico Camposeco, ETHZ; Torsten Sattler, ETH Zurich; Marc Pollefeys, ETH

P-3A-35. **Learning to Hash with Binary Deep Neural Network**, Thanh-Toan Do, SUTD; Dung Doan, SUTD; Nga-can Cheung, SUTD

P-3A-36. **Automatically selecting inference algorithms for discrete energy minimisation**, Paul Henderson, University of Edinburgh; Vittorio Ferrari

P-3A-37. **Ego2Top: Matching Viewers in Egocentric and Top-view Videos**, Shervin Ardeshir, University of Central Florida; Ali Borji, University of Central Florida

P-3A-38. **Online Action Detection**, Roeland De Geest, KU Leuven; Stratis Gavves, University of Amsterdam; Amir Ghodrati; Zhenyang Li; Cees Snoek, University of Amsterdam; Tinne Tuytelaars, KU Leuven

P-3A-39. **Cross-modal Supervision for Learning Active Speaker Detection in Video**, Punarjay Chakravarty, KU Leuven; Tinne Tuytelaars, KU Leuven

P-3A-40. **Recurrent Temporal Deep Field for Semantic Video Labeling**, Peng Lei, Oregon State University; Sinisa Todorovic, Oregon State University

P-3A-41. **Ultra Resolution by Discriminative Generative Networks**, Xin Yu, Australian National University; Fatih Porikli, Australian National University


P-3A-44. **Network flow formulations for Learning Binary Hashing**, Lopamudra Mukherjee, University of Wisc Whitewater; Jiming Peng, University of Houston; Trevor Sigmund, University of Wisconsin Whitewater; Vikas Singh, University of Wisconsin-Madison

P-3A-45. **SPICE: Semantic Propositional Image Caption Evaluation**, Peter Anderson, Australian National University; Basura Fernando, ANU; Mark Johnson, Macquarie University; Stephen Gould, Australian National University

P-3A-46. **Transfer Neural Trees for Heterogeneous Domain Adaptation**, Wei Yu Chen, NTU; Academia Sinica; Tzu-Ming Hsu, Academia Sinica; Yu-Chiang Frank Wang, Academia Sinica; Ming-Syan Chen, National Taiwan University

P-3A-47. **Tracking Persons-of-Interest via Adaptive Discriminative Features**, Shun Zhang, Xi’an Jiaotong University; Yihong Gong, Jia-Bin Huang, University of Illinois at Urbana-Champaign; Jongwoo Lim, Hanyang University; Jinjun Wang, Narendra Ahuja, University of Illinois at Urbana-Champaign; Ming-Hsuan Yang, UC Merced

**Demos**


D-3A-49. **DeepWarp: Photorealistic Image Resynthesis for Gaze Manipulation**, Yaroslav Ganin, Skolkovo Institute of Science and Technology; Daniil Kononenko, Skolkovo Institute of Science and Technology; Diana Sungatullina, Skolkovo Institute of Science and Technology; Victor Lempitsky, Skolkovo Institute of Science and Technology (shown in roof garden)

**12:30 – 14:00 Lunch Break**

**14:00 – 15:00 Spotlight Session 3B-3C (3D & Structure from Motion)**

**Chairs:** Daniel Cremers (TU Munchen), Tomas Pajdla(CVUT)

S-3B-05. **When is Rotations Averaging Hard?**, Kyle Wilson, Washington College; David Bindel, Cornell University; Noah Snavely

S-3B-06. **Capturing Dynamic Textured Surfaces of Moving Targets**, Ruizhe Wang, USC; Lingyu Wei, USC; Etienne Vouga, UT Austin; Qixing Huang, TTI-Chicago; Duygu Ceylan, Adobe Research; Gerard Medioni, University of Southern California; Hao Li, USC

S-3B-07. **ShapeFit and ShapeKick for Robust, Scalable Structure from Motion**, Tom Goldstein, University of Maryland, College Park; Paul Hand, Vladislav Voroninski, MIT; Stefano Soatto, University of California, Los Angeles; Choongbum Lee

S-3B-08. **Heat Diffusion Long-Short Term Memory Learning for 3D Shape Analysis**, Fan Zhu, NYUAD; Jin Xie, NYUAD; Yi Fang, New York University Abu Dhabi

S-3B-09. **Multi-view 3D Models from Single Images With a Convolutional Network**, Maxim Tatarchenko, University of Freiburg; Alexey Dosovitskiy, University of Freiburg; Thomas Brox

S-3C-05. **Linear depth estimation from an uncalibrated**, monocular polarisation image, William Smith, University of York; Ravi Ramamoorthi, Silvia Tozza

S-3C-06. **Online Variational Bayesian Motion Averaging**, Guillaume Bourmaud, Toshiba CRL

S-3C-07. **Unified Depth Prediction and Intrinsic Image Decomposition from a Single Image via Joint Convolutional Neural Fields**, Seungryong Kim, Yonsei University; Kihong Park, Yonsei Univ.; Kwanghoon Sohn, Yonsei university; Stephen Lin, Microsoft Research Asia, China
S-3C-08. **ObjectNet3D: A Large Scale Database for 3D Object Recognition**, Yu Xiang, University of Michigan; Wonhui Kim, University of Michigan; Wei Chen, Stanford University; Jingwei Ji, Stanford University; Christopher Choy, Stanford University; Hao Su, Stanford; Roozbeh Mottaghi, Allen Institute for AI; Leonidas J. Guibas; Silvio Savarese

S-3C-09. **Branching Gaussian Processes with Applications to Spatiotemporal Reconstruction of 3D Trees**, Kyle Simek, University of Arizona; Ravishankar Palanivelu, University of Arizona; Kebus Barnard, University of Arizona

**15:00 – 16:30 Poster Session 3B & Demos**

*Chairs: Jasper Uijlings (University of Edinburgh), Roberto Valenti (Sightcorp)*

S-3B-05. **When is Rotations Averaging Hard?**, Kyle Wilson, Washington College; David Bindel, Cornell University; Noah Snavely

S-3B-06. **Capturing Dynamic Textured Surfaces of Moving Targets**, Ruizhe Wang, USC; Lingyu Wei, USC; Etienne Vouga, UT Austin; Qixing Huang, TTI-Chicago; Duygu Ceylan, Adobe Research; Gerard Medioni, University of Southern California; Hao Li, USC

S-3B-07. **ShapeFit and ShapeKick for Robust**, Scalable Structure from Motion, Tom Goldstein, University of Maryland, College Park; Paul Hand, ; Vladislav Voroninski, MIT; Stefano Soatto, University of California, Los Angeles; Choongbun Lee

S-3B-08. **Heat Diffusion Long-Short Term Memory Learning for 3D Shape Analysis**, Fan Zhu, NYUAD; Jin Xie, NYUAD; Yi Fang, New York University Abu Dhabi

S-3B-09. **Multi-view 3D Models from Single Images With a Convolutional Network**, Maxim Tatarchenko, University of Freiburg; Alexey Dosovitskiy, University of Freiburg; Thomas Brox

P-3B-10. **Extending Long Short-Term Memory for Multi-View Structured Learning**, Shyam Sundar Rajagopalan, University of Canberra; Louis-Philippe Morency, Carnegie Mellon University; Tadas Baltrušaitis, Carnegie Mellon University; Roland Goecke, University of Canberra

P-3B-11. **Gated Bi-directional CNN for Object Detection**, Xingyu Zeng, The Chinese University of Hong Kong; Wanli Ouyang, Chinese University of Hong Kong; Bin Yang, NLPR, CASIA; Junjie Yan, National Laboratory of Pattern Recognition, Chinese Academy of Sciences; Xiaogang Wang, The Chinese University of Hong Kong

P-3B-12. **Graph Based Skeleton Motion Representation and Similarity Measurement for Action Recognition**, Pei Wang, Chinese Academy of Sciences; Chunfeng Yuan, ; Weiming Hu, ; Bing Li, ; yanning Zhang, Northwestern Polytechnical University

P-3B-13. **Reliable Fusion of ToF and Stereo Depth Driven by Confidence Measures**, Giulio Marin, University of Padova; Pietro Zanuttigh, University of Padova; Stefano Mattoccia, University of Bologna

P-3B-14. **Fast, Exact and Multi-Scale Inference for Semantic Image Segmentation with Deep Gaussian CRFs**, Siddhartha Chandra, INRIA; Iasonas Kokkinos, INRIA

P-3B-15. **Kernel-Based Supervised Discrete Hashing for Image Retrieval**, Xiaoshuang Shi, University of Florida; Fuyong Xing, University of Florida; Jinzheng Cai, University of Florida; Zizhao Zhang, University of Florida; Yuanpu Xie, University of Florida; Lin Yang, University of Florida

P-3B-16. **Iterative Reference Driven Metric Learning for Signer Independent Isolated Sign Language Recognition**, Fang Yin, ICT,CAS; Xiujuan Chai, ICT, CAS; Xilin Chen, Institute of Computing Technology, Chinese Academy of Sciences

P-3B-17. **Ask, Attend and Answer: Exploring Question-Guided Spatial Attention for Visual Question Answering**, Huijuan Xu, UMass Lowell; Kate Saenko, University of Massachusetts Lowell
P-3B-18. **Relay Backpropagation for Effective Learning of Deep Convolutional Neural Networks**, Li Shen, UCAS; Zhouchen Lin, Peking University; Qingming Huang, University of Chinese Academy of Sciences

P-3B-19. **Counting in The Wild**, Carlos Arteta, Oxford University; Victor Lempitsky, Skolkovo Institute of Science and Technology; Andrew Zisserman, University of Oxford

P-3B-20. **A Discriminative Feature Learning Approach for Deep Face Recognition**, Yandong Wen, ; Kaipeng Zhang, ; Zhifeng Li, SIAT, CAS; Yu Qiao, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences

P-3B-21. **Network of Experts for Large-Scale Image Categorization**, Karim Ahmed, Dartmouth; Mohammad Haris Baig, Dartmouth College; Lorenzo Torresani, Dartmouth College

P-3B-22. **Zero-Shot Recognition via Structured Prediction**, Ziming Zhang, Boston University; Venkatesh Saligrama,

P-3B-23. **A Generalized Successive Shortest Paths Solver for Tracking Dividing Targets**, Carsten Haubold, University of Heidelberg; Janz Ales, ; Steffen Wolf, ; Fred Hamprecht, Heidelberg University

P-3B-24. **Accurate and Linear Time Pose Estimation from Points and Lines**, Alexander Vakhitov, SPbSU; Jan Funke, Institut de Robòtica i Informà; Francesc Moreno-Noguer, Institut de Robotica i Informatica Industrial (UPC/CSIC)

P-3B-25. **Pseudo-Geometric Formulation for Fitting Equidistant Parallel Lines**, FAISAL AZHAR, HP Labs; STEPHEN POLLARD, HP Labs.

P-3B-26. **Towards perspective-free object counting with deep learning**, Daniel Oñoro Rubio, University of Alcalá; Roberto Lopez-Sastre, University of Alcalá

P-3B-27. **Information Bottleneck Domain Adaptation with Privileged Information for Visual Recognition**, Saeid Motiian, West Virginia University; Gianfranco Doretto, West Virginia University

P-3B-28. **Template-free 3D Reconstruction of Poorly-textured Nonrigid Surfaces**, Xuan Wang, Xi’an Jiaotong University; Mathieu Salzmann, EPFL; Fei Wang, Xi’an Jiaotong University; Jizhong Zhao, Xi’an Jiaotong university

P-3B-29. **FigureSeer: Parsing Result-Figures in Research Papers**, Noah Siegel, ; Zachary Horvitz, ; Roie Levin, ; Santosh Kumar Divvala, Allen Institute for Artificial Intelligence; Ali Farhadi, University of Washington

P-3B-30. **Approximate search with quantized sparse representations**, Himalaya Jain, Inria; Patrick Perez, Technicolor, France; Rémi Gribonval, ; joaquin Zepeda, Technicolor; Hervé Jegou

P-3B-31. **Sympathy for the Details: Dense Trajectories and Hybrid Classification Architectures for Action Recognition**, Cesar De Souza, Xerox Research Center Europe; Adrien Gaidon, ; Eleonora Vig, German Aerospace Center; Antonio Lopez, Centro de Visio per Computador, UAB

P-3B-32. **Human pose estimation via Convolutional Part Heatmap Regression**, Adrian Bulat, University of Nottingham; Georgios Tzimiropoulos, Nottingham University

P-3B-33. **Collaborative Layer-wise Discriminative Learning in Deep Neural Networks**, XIAOJIE JIN, National Univ. of Singapore; YUNPENG CHEN, NUS; Jian Dong, Qihoo/360; Jiashi Feng, NUS; Shuicheng Yan,

P-3B-34. **Deep Decoupling of Defocus and Motion Blur for Dynamic Segmentation**, Abhijith Punnappurath, IIT Madras; YOGESH BALAJI, IIT MADRAS; Mahesh Mohan M R, IIT Madras; A N Rajagopalan, Indian Institute of Technology, Madras

P-3B-35. **Video Summarization with Long Short-term Memory**, Ke Zhang, USC; Wei-Lun Chao, USC; Fei Sha, UCLA; Kristen Grauman, University of Texas at Austin
P-3B-36. Leaving Some Stones Unturned: Dynamic Feature Prioritization for Activity Detection in Streaming Video, Yu-Chuan Su, University of Texas at Austin; Kristen Grauman, University of Texas at Austin

P-3B-37. Robust and Accurate Line- and/or Point-Based Pose Estimation without Manhattan Assumptions, Yohann Salaun, LIGM-Imagine; Renaud Marlet, ; Pascal Monasse, Universite Paris-Est

P-3B-38. MARLow: A Joint Multilinear Autoregressive and Low-Rank Approach for Image Completion, Mading Li, Peking University; Jiaying Liu, Peking University; Zhiwei Xiong, ; Xiaoyan Sun, ; Zongming Guo, Peking University


P-3B-40. Carried Object Detection based on an Ensemble of Contour Exemplars, FARNOOSH GHADIRI, Laval University; ROBERT BERGEVIN, Laval university; GUILLAUME-ALEXANDRE BILODEAU, École Polytechnique de Montréal

P-3B-41. Query-Focused Extractive Video Summarization, Aidean Sharghi Karganroodi, University of Central Florida; Boqing Gong, University of Central Florida; Mubarak Shah, University of Central Florida

P-3B-42. Temporal Segment Networks: Towards Good Practices for Deep Action Recognition, Limin Wang, ETHZ; Yuanjun Xiong, The Chinese University of HK; Zhe Wang, SIAT, CAS; Yu Qiao, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences; Dahua Lin, CUHK; Xiaou Tang, Chinese University of Hong Kong; Lu Van Gool, ETH Zurich

P-3B-43. PlaNet – Photo Geolocation with Convolutional Neural Networks, Tobias Weyand, Google; Ilya Kostrikov, RWTH Aachen University; James Philbin, Zoox

P-3B-44. Detecting Text in Natural Image with Connectionist Text Proposal Network, Zhi Tian, SIAT, Weilin Huang, The University of Oxford; Tong He, Wuhan University; Pan He, SIAT, MMLAB; Yu Qiao, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences

P-3B-45. Face recognition using a unified 3D morphable model, GUOSHENG HU, INRIA; fei Yan, ; chi-Ho Chan, CVSSP, University of Surrey; Weihong Deng, BUPT; william Christmas, CVSSP; Josef Kittler, University of Surrey; Neil Robertson, Queen’s University of Belfast

P-3B-46. Augmented Feedback in Semantic Segmentation under Image Level Supervision, Xiaojuan Qi, CUHK; Zhengzhe liu, ; Jianping Shi, SenseTime; Hengshuang Zhao, ; Jiaya Jia, Chinese University of Hong Kong

P-3B-47. Towards Viewpoint Invariant 3D Human Pose Estimation, Albert Haque, Stanford University; Boya Peng, Stanford University; Zelun Luo, Stanford University; Alexandre Alahi, Stanford University; Serena Yeung, Stanford University; Fei-Fei Li, Stanford University

Demos

D-3B-48. RGB-D SLAM with fast and robust relocalization and HDR texture mapping, Shuda Li, University of Bristol; Andrew Calway, University of Bristol (shown in roof garden)

D-3B-49. Extracting Driving Behavior: Global Metric Localization from Dashcam Videos in the Wild, Shao-Pin Chang, National Tsing Hua University; Jui-Ting Chien, National Tsing Hua University; Fu-En Wang, National Tsing Hua University; Shang-Da Yang, National Tsing Hua University; Hwann-Tzong Chen, National Tsing Hua University; Min Sun, National Tsing Hua University (shown in roof garden)
16:30 – 16:45  **Coffee Break**

16:45 – 18:15  **Poster Session 3C & Demos**  

Chairs: Jasper Uijlings (University of Edinburgh), Roberto Valenti (Sightcorp)

- **S-3C-05.** *Linear depth estimation from an uncalibrated* monocular polarisation image, William Smith, University of York; Ravi Ramamoorthi, ; Silvia Tozza
- **S-3C-06.** *Online Variational Bayesian Motion Averaging*, Guillaume Bourmaud, Toshiba CRL
- **S-3C-07.** *Unified Depth Prediction and Intrinsic Image Decomposition from a Single Image via Joint Convolutional Neural Fields*, Seungryong Kim, Yonsei University; Kihong Park, Yonsei Univ.; Kwanghoon Sohn, Yonsei university; Stephen Lin, Microsoft Research
- **S-3C-08.** *ObjectNet3D: A Large Scale Database for 3D Object Recognition*, Yu Xiang, University of Michigan; Wonhui Kim, University of Michigan; Wei Chen, Stanford University; Jingwei Ji, Stanford University; Christopher Choy, Stanford University; Hao Su, Stanford; Roozbeh Mottaghi, Allen Institute for AI; Leonidas J. Guibas, ; Silvio Savarese
- **S-3C-09.** *Branching Gaussian Processes with Applications to Spatiotemporal Reconstruction of 3D Trees*, Kyle Simek, University of Arizona; Ravishankar Palanivelu, University of Arizona; Kobus Barnard, University of Arizona
- **P-3C-10.** *Tracking Completion*, Yao Sui, University of Kansas; Guanghui Wang, University of Kansas; Yafei Tang, China Unicom Research; Li Zhang, Tsinghua University
- **P-3C-11.** *Inter-Battery Topic Representation Learning*, Cheng Zhang, KTH Royal Institute of Technology; Hedvig Kjellstrom, ; Carl Henrik Ek
- **P-3C-12.** *Online Adaptation for Joint Scene and Object Classification*, Md Jawadul Bappy, University of California, Riverside; Sujoy Paul, University of California, Riverside; Amit Roy-Chowdhury, UC Riverside
- **P-3C-13.** *Real-Time Facial Segmentation and Performance Capture from RGB Input*, Shunsuke Saito, USIC; Tianye Li, USC; Hao Li, USC
- **P-3C-14.** *Learning Temporal Transformations From Time-Lapse Videos*, Yipin Zhou, UNC; Tamara Berg, University on North Carolina
- **P-3C-15.** *Interactive Image Segmentation Using Constrained Dominant Sets*, Eyasu Zemene, Ca’Foscar University; Marcello Pelillo, University of Venice
- **P-3C-16.** *Deep Markov Random Field for Image Modeling*, Zhirong Wu, The Chinese University of Hong Kong; Dahua Lin, CUHK; Xiaou Tang, Chinese University of Hong Kong
- **P-3C-17.** *A Symmetry Prior for Convex Variational 3D Reconstruction*, Pablo Speciale, ETHZ; Martin R. Oswald, ETH Zurich; Andrea Cohen, ETHZ; Marc Pollefeys, ETH
- **P-3C-18.** *SPLeaP: Soft Pooling of Learned Parts for Image Classification*, Praveen Kulkarni, Technicolor; Frederic Jurie, University of Caen; joaquin Zepeda, Technicolor; Patrick Perez, Technicolor, France; Louis Chevallier, Technicolor
- **P-3C-19.** *Spatial Attention Deep Net with Partial PSO for Hierarchical Hybrid Hand Pose Estimation*, Qi Ye, ; Shaxnin Yuan, Imperial College London; Tae-Kyun Kim, Imperial College London
- **P-3C-20.** *VolumeDeform: Real-time Volumetric Non-rigid Reconstruction*, Matthias Innmann, University Erlangen-Nuremberg; Michael Zollhoefer, MPI Informatics; Matthias Niessner, Stanford University; Christian Theobalt, MPI Informatics; Marc Stamminger, University of Erlangen-Nuremberg
- **P-3C-21.** *piMatch: Monocular vSLAM and Piecewise Planar Reconstruction using Fast Plane Correspondences*, Carolina Raposo, University of Coimbra; Joao Barreto, Universidade de Coimbra
P-3C-22. **Peripheral Expansion of Depth Information via Layout Estimation with Fisheye Camera**, Alejandro Perez-Yus, Universidad de Zaragoza; Gonzalo Lopez-Nicolás, Universidad de Zaragoza; Josechu Guerrero, Universidad de Zaragoza

P-3C-23. **Built-in Foreground/Background Prior for Weakly-Supervised Semantic Segmentation**, Fatemehsadat Saleh, The Australian National University/CSIRO; Mohammad Sadegh Aliakbarian, National ICT Australia; Mathieu Salzmann, EPFL; Lars Petersson, NICTA; Stephen Gould, Australian National University; Jose M. Alvarez, Data61 / CSIRO

P-3C-24. **It's Moving! A Probabilistic Model for Causal Motion Segmentation in Moving Camera Videos**, Pia Bideau, Umass Amherst; Erik Miller

P-3C-25. **Kernelized Subspace Ranking for Saliency Detection**, Tianfeng Wang, Dalian University of Technology; Lihe Zhang, Dalian University of Technology; Huchuan Lu; Chong Sun, Dalian University of Technology

P-3C-26. **Depth-aware Motion Magnification**, Julian Kooij, Delft University of Technology

P-3C-27. **Stacked Hourglass Networks for Human Pose Estimation**, Alejandro Newell, University of Michigan; Kaiyu Yang, University of Michigan; Jia Deng, University of Michigan

P-3C-28. **Real-time Large-Scale Dense 3D Reconstruction with Loop Closure**, Olaf Kahler, University of Oxford; David Murray, Oxford; Victor Prisacariu, Oxford

P-3C-29. **Pixel-Level Domain Transfer**, Donggeun Yoo, KAIST; Namil Kim, KAIST; Sunggyun Park, KAIST; Anthony Paek, Lunit Inc.; In So Kweon, KAIST

P-3C-30. **Accelerating Convolutional Neural Networks with Dominant Convolutional Kernel and Knowledge Pre-regression**, Zhenyang Tang, Tsinghua University; Zhihong Deng, Tsinghua University; Shiyao Wang, Tsinghua University

P-3C-31. **Learning Social Etiquette: Human Trajectory Understanding in Crowded Scenes**, Alexandre Robicquet, Stanford University; Amir Sadeghian, Stanford University; Alexandre Alahi, Stanford University; Silvio Savarese

P-3C-32. **Bayesian Image-based 3D Pose Estimation**, Valsamis Ntouskos, Sapienza University of Rome; Fiore Pirri, University of Rome, Sapienza; Marta Sanzari, Sapienza

P-3C-33. **Efficient and Robust Semi-supervised Learning over a Sparse-regularized Graph**, Hang Su, Tsinghua University; Jun Zhu, Tsinghua University; Zhaoheng Yin, Missouri University of Science and Technology; Yinpeng Dong, Tsinghua University; Bo Zhang, Tsinghua University

P-3C-34. **Novel Coplanar Line-points Invariants for Robust Line Matching Across Views**, Qi Jia, Dalian university of technology; Xinkai Gao, Dalian university of technology; Xin Fan, Dalian University of Technology; Zhongxuan Luo, Dalian university of technology; Haojie Li, Dalian university of technology; Ziyao Chen, Dalian university of technology

P-3C-35. **Sparse Representation Based Complete Kernel Marginal Fisher Analysis Framework for Computational Art Painting Categorization**, Ajit Putthiputhussy, New Jersey Institute of Technology; Qingfeng Liu, New Jersey Institute of Techno; Chengjun Liu, New Jersey Institute of Technology

P-3C-36. **3D-R2N2: A unified approach for single and multi-view 3D object reconstruction**, Christopher Choy, Stanford University; Danfei Xu, Stanford University; GunYoung Gwak, Stanford; Kevin Chen, Stanford University; Silvio Savarese

P-3C-37. **Cascaded Continuous Regression for Real-time Incremental Face Tracking**, Enrique Sanchez-Lozano, University of Nottingham; Brais Martinez, University of Nottingham; Georgios Tzimiropoulos, Nottingham University; Michel Valstar, Nottingham University

P-3C-38. **Real-Time Visual Tracking: Promoting the Robustness of Correlation Filter Learning**, Yao Sui, University of Kansas; Ziming Zhang, Boston University; Guanghui Wang, University of Kansas; Yafei Tang, China Unicom Research; Li Zhang, Tsinghua University

---

**Program** | Thursday, October 13 | Theatre Carré
P-3C-39. **Deep Self-Correlation Descriptor for Dense Cross-Modal Correspondence**, Seungryong Kim, Yonsei University; Dongbo Min, Chungnam National University; Stephen Lin, Microsoft Research Asia, China; Kwanghoon Sohn, Yonsei University

P-3C-40. **Structured Matching for Phrase Localization**, Mingzhe Wang, University of Michigan; Mahmoud Azab, University of Michigan; Noriyuki Kojima, University of Michigan; Rada Mihalcea, University of Michigan; Jia Deng, University of Michigan

P-3C-41. **Crossing-line Crowd Counting with Two-phase Deep Neural Networks**, Zhuoyi Zhao, CUHK; Hongsheng Li, CUHK; Rui Zhao, The Chinese University of Hong Kong; Xiaogang Wang, The Chinese University of Hong Kong

P-3C-42. **Revisiting Visual Question Answering Baselines**, Allan Jabri, Facebook AI Research; Armand Joulin, Facebook AI Research; Laurens van der Maaten

P-3C-43. **Unsupervised CNN for Single View Depth Estimation: Geometry to the Rescue**, Ravi Garg, University of Adelaide; Vijay Kumar, University of Adelaide; Gustavo Carneiro, University of Adelaide; Ian Reid, University of Adelaide

P-3C-44. **A Continuous Optimization Approach for Efficient and Accurate Scene Flow**, Zhaoyang Lv, Georgia Tech; Chris Beall, Georgia Tech; Pablo Alcantarilla, iRobot Corporation; Fuxin Li, Oregon State University; Zsolt Kira, Georgia Tech Research Institute; Frank Dellaert

P-3C-45. **Improving Multi-Frame Data Association with Sparse Representations for Robust Near-Online Multi-Object Tracking**, Loïc Fagot-Bouquet, CEA LIST; Romaric Audigier, CEA LIST; Yoann Dhome, CEA LIST; Frédéric Lerasle, LAAS CNRS

P-3C-46. **Gated Siamese Convolutional Neural Network Architecture for Human Re-Identification**, Rahul Rama Varior, NTU Singapore; Mrinal Haloi, Nanyang technological Universi; Gang Wang

P-3C-47. **Saliency Detection via Combining Region-Level and Pixel-Level Predictions with CNNs**, Youbao Tang, Harbin Institute of Technology; Xiangqian Wu, Harbin Institute of Technology

**Demos**

D-3C-48. **Focal Flow Camera: a Prototype of Measuring Distance and Velocity with Defocus and Differential Motion**, Qi Guo, Harvard University; Emma Alexander, Harvard University; Avinash Uttamchandani, Harvard University; Sanjeev Koppal, Harvard University; Steven Gortler, Harvard University; Todd Zickler, Harvard University (shown in roof garden)

D-3C-49. **From Low- to High-Quality Depth Sensors for 3D Object Reconstruction with SDF-2-SDF**, Miroslava Slavcheva, Technische Universität München and Siemens AG; Wadim Kehl, Technische Universität München; Nassir Navab, Technische Universität München; Slobodan Ilic, Technische Universität München and Siemens AG (shown in roof garden)

18:00 – 22:00  Conference Dinner Ocean Diva (PTA)
Program
Friday, October 14
Location: Theatre Carré / Room Theatre

09:00 – 10:00 Oral Session 4A (3D)
Chairs: Jan-Michael Frahm (UNC), Hongdong Li (ANU)

O-4A-01. **Real-Time 3D Reconstruction and 6-DoF Tracking with an Event Camera**, Hanme Kim, Imperial College London; Stefan Leutenegger, Imperial College London; Andrew Davison

O-4A-02. **Single Image 3D Interpreter Network**, Jiajun Wu, MIT; Tianfan Xue, MIT; Joseph Lim, MIT; Yuandong Tian, FAIR; Joshua Tenenbaum, MIT; Antonio Torralba, ; William Freeman

O-4A-03. **Dual Structured Light 3D using a 1D Sensor**, Jian Wang, Carnegie Mellon University; Aswin Sankaranarayanan, Carnegie Mellon University; Mohit Gupta, ; Srinivasa Narasimhan, CMU

O-4A-04. **Shape acquisition and registration for 3D endoscope based on grid pattern projection**, Ryo Furukawa, Hiroshima City University; Hiroki Morinaga, Kagoshima University; Yoji Sanomura, Hiroshima University; Shinji Tanaka, Hiroshima University; Shigeto Yoshida, Hiroshima General Hospital of West Japan Railway Company; Hiroshi Kawasaki, Kagoshima University

10:00 – 10:30 Spotlight Session 4A (Feature Extraction & Matching)
Chairs: Thomas Brox (U Freiburg), Andrea Vedaldi (U Oxford)

S-4A-05. **Target Response Adaptation for Correlation Filter Tracking**, Adel Bibi, KAUST; Matthias Mueller, KAUST; Bernard Ghanem, KAUST

S-4A-06. **Learning Image Matching by Simply Watching Video**, Gucan Long, NUDT; Laurent Kneip, Australian National University; Jose M. Alvarez, Data61 / CSIRO; Hongdong Li, ; Xiaohu Zhang, NUDT; Qifeng Yu, NUDT

S-4A-07. **A Distance for HMMs based on Aggregated Wasserstein Metric and State Registration**, Yukun Chen, Penn State University; Jianbo Ye, College of Information Sciences and Technology, Penn State University; Jia Li, Department of Statistics, Penn State University

S-4A-08. **LIFT: Learned Invariant Feature Transform**, Kwang Yi, EPFL; Eduard Trulls, EPFL; Vincent Lepetit, ; Pascal Fua,

S-4A-09. **Learning a Predictable and Generative Vector Representation for Objects**, Rohit Girdhar, CMU; David Fouhey, Carnegie Mellon University; Mikel Rodriguez, MITRE INRIA; Abhinav Gupta

10:30 – 11:00 Coffee Break

11:00 – 12:30 Poster Session 4A
Chairs: Jasper Uijlings (University of Edinburgh), Roberto Valenti (Sightcorp)

O-4A-01. **Real-Time 3D Reconstruction and 6-DoF Tracking with an Event Camera**, Hanme Kim, Imperial College London; Stefan Leutenegger, Imperial College London; Andrew Davison
O-4A-02. **Single Image 3D Interpreter Network**, Jiajun Wu, MIT; Tianfan Xue, MIT; Joseph Lim, MIT; Yuandong Tian, FAIR; Joshua Tenenbaum, MIT; Antonio Torralba, ; William Freeman

O-4A-03. **Dual Structured Light 3D using a 1D Sensor**, Jian Wang, Carnegie Mellon University; Aswin Sankaranarayanan, Carnegie Mellon University; Mohit Gupta, ; Srinivasa Narasimhan, CMU

O-4A-04. **Shape acquisition and registration for 3D endoscope based on grid pattern projection**, Ryo Furukawa, Hiroshima City University; Hiroki Morinaga, Kagoshima University; Yoji Sanomura, Hiroshima University; Shinji Tanaka, Hiroshima University; Shigeto Yoshida, Hiroshima General Hospital of West Japan Railway Company; Hiroshi Kawasaki, Kagoshima University

S-4A-05. **Target Response Adaptation for Correlation Filter Tracking**, Adel Bibi, KAUST; Matthias Mueller, KAUST; Bernard Ghanem, KAUST

S-4A-06. **Learning Image Matching by Simply Watching Video**, Gucan Long, NUDT; Laurent Kneip, Australian National University; Jose M. Alvarez, Data61 / CSIRO; Hongdong Li, ; Xiaohu Zhang, NUDT; Qifeng Yu, NUDT

S-4A-07. **A Distance for HMMs based on Aggregated Wasserstein Metric and State Registration**, Yukun Chen, Penn State University; Jianbo Ye, College of Information Sciences and Technology, Penn State University; Jia Li, Department of Statistics, Penn State University

S-4A-08. **LIFT: Learned Invariant Feature Transform**, Kwang Yi, EPFL; Eduard Trulls, EPFL; Vincent Lepetit, ; Pascal Fua

S-4A-09. **Learning a Predictable and Generative Vector Representation for Objects**, Rohit Girdhar, CMU; David Fouhey, Carnegie Mellon University; Mikel Rodriguez, MITRE INRIA; Abhinav Gupta

P-4A-10. **HouseCraft: Building Houses from Rental Ads and Street Views**, Hang Chu, University of Toronto; Shenlong Wang, University of Toronto; Raquel Urtasun, ; Sanja Fidler, University of Toronto

P-4A-11. **Superpixel-based Two-view Deterministic Fitting for Multiple-structure Data**, Guobao Xiao, Xiamen University; Hanzi Wang, ; Yan Yan, ; David Suter

P-4A-12. **Instance-sensitive Fully Convolutional Networks**, Jifeng Dai, Microsoft Research Asia; Kaiming He, Microsoft Research Asia; Yi Li, Tsinghua University; Shaoqing Ren, University of Science & Technology of China; Jian Sun, Microsoft Research China

P-4A-13. **Domain Adaptive Fisher Vector for Visual Recognition**, Li Niu, NTU; Jianfei Cai, ; Dong Xu

P-4A-14. **Deep Robust Encoder through Locality Preserving Low-Rank Dictionary**, Zhengming Ding, Northeastern University; Ming Shao, Northeastern University; Yun Fu, Northeastern University

P-4A-15. **Pattern Mining Saliency**, Yuqiu Kong, Dalian University of Technology; Lijun Wang, Dalian University of Technology; Xiuping Liu, Dalian University of Technology; Huchuan Lu, ; Xiang Ruan

P-4A-16. **Streaming Video Segmentation via Short-Term Hierarchical Segmentation and Frame-by-Frame Markov Random Field Optimization**, Won-Dong Jang, Korea University; Chang-Su Kim, Korea University


P-4A-18. **Shape from Water: Bispectral Light Absorption for Depth Recovery**, Yuta Asano, ; Yinqiang Zheng, National Institute of Informatics, Japan; Ko Nishino, Drexel University; Imari Sato, National Institute of Informatics, Japan

P-4A-19. **Learning Dynamic Hierarchical Models for Anytime Scene Labeling**, Buyu Liu, NICTA; Xuming He, NICTA
P-4A-20. **Semantic 3D Reconstruction of Heads**, Fabio Maninchedda, ETH Zürich; Christian Haene, ETH Zurich; Bastien Jacquet, ETH Zürich; Amael Delaunoy, ; Marc Pollefeys, ETH

P-4A-21. **Human Attribute Recognition by Deep Hierarchical Contexts**, Yining Li, IE, CUHK; Chen Huang, The Chinese University of Hong Kong; Chen-Change Loy, the Chinese University of Hong Kong; Xiaou Tang, Chinese University of Hong Kong

P-4A-22. **Person Re-Identification via Recurrent Feature Aggregation**, Yichao Yan, Shanghai Jiaotong University; Bingbing Ni, SJTU; Zhichao Song, ; chao Ma, ; Yan Yan, ; xiaokang Yang, SJTU

P-4A-23. **Biconvex Relaxation for Semidefinite Programming in Computer Vision**, Sohil Shah, University of Maryland; Abhay Yadav, University Of Maryland; Carlos Castillo, University of Maryland, CP; David Jacobs, University of Maryland; Christoph Studer, Cornell University; Tom Goldstein, University of Maryland, College Park


P-4A-25. **End-to-End Localization and Ranking for Relative Attributes**, Krishna Kumar Singh, UC Davis; Yong Jae Lee, UC Davis

P-4A-26. **Efficient Large Scale Image Classification via Prediction Score Decomposition**, Duy-Dinh Le, National Institute of Informat; Tien-Dung Mai, University of Information Technology; Shin’ichi Satoh, National Institute of Informatics, Japan; Thanh Duc Ngo, University of Information Technology; Duc Anh Duong, University of Information Technology

P-4A-27. **Stochastic Dykstra Algorithms for Metric Learning with Positive Definite Covariance Descriptors**, Tsuyoshi Kato, Gunma University; Tomoki Matsuzawa, Gunma University; Jun Sese, ; Raissa Relator

P-4A-28. **MeshFlow: Minimum Latency Online Video Stabilization**, Shuaicheng Liu, Uestc.edu.cn; ping tan, ; Lu Yuan, Microsoft Research Asia; Jian Sun, Microsoft Research China; Bing Zeng, uestc.edu.cn

P-4A-29. **Large-scale training of shadow detectors with noisily-annotated shadow examples**, Tomas F Yago Vicente, Stony Brook University; Le Hou, Stony Brook University; Chen-Ping Yu, Stony Brook University; Minh Hoai, Stony Brook University; Dimitris Samaras, SUNY Stonybrook

P-4A-30. **RNN Fisher Vectors for Action Recognition and Image Annotation**, Guy Lev, ; Gil Sadeh, ; Benjamin Klein, Tel Aviv University; Lior Wolf

P-4A-31. **CDT: Cooperative Detection and Tracking for Tracing Multiple Objects in Video Sequences**, Hanul Kim, Korea university; Chang-Su Kim, Korea University

P-4A-32. **MARS: A Video Benchmark for Large-Scale Person Re-identification**, Liang Zheng, University of Texas at San Ant; Zhi Bie, Tsinghua University; Yifan Sun, Tsinghua University; Jingdong Wang, Microsoft Research; Chi Su, Peking University; Shengjin Wang, wgsj@tsinghua.edu.cn; Qi Tian

P-4A-33. **Angry Crowds: Detecting Violent Events in Videos**, Seyed Sadegh Mohammadi, Istituto Italiano di Tecnologi; Alessandro Perina, Microsoft; Hamed Kiani, ; Vittorio Murino, Istituto Italiano di Tecnologia

P-4A-34. **Sparse Recovery of Hyperspectral Signal from Natural RGB Images**, Boaz Arad, BGU; Ohad Ben-Shahar, Ben-Gurion University

P-4A-35. **Light Field Segmentation Using a Ray-Based Graph Structure**, Matthieu Hog, Technicolor R&I; Neus Sabater, Technicolor R&I; Christine Guillemot, INRIA

P-4A-36. **Design of Kernels in Convolutional Neural Networks for Image Classification**, Zhun Sun, Tohoku University; Mete Ozay, Tohoku University; Takayuki Okatani, Tohoku University, Japan
P-4A-37. **Learning Visual Features from Large Weakly Supervised Data**, Armand Joulin, Facebook AI Research; Laurens van der Maaten, ; Allan Jabri, Facebook AI Research; Nicolas Vasilache, Facebook AI Research

P-4A-38. **3D Mask Face Anti-spoofing with Remote Photoplethysmography**, Siqi Liu, Hong Kong Baptist University; Pong Chi YUEN, Hong Kong Baptist University; Shengping Zhang, Harbin Institute of Technology; Guoying Zhao, University of Oulu

P-4A-39. **Guided Matching based on Statistical Optical Flow for Fast and Robust Correspondence Analysis**, Josef Maier, AIT Austrian Institute of Technology; Martin Humenberger, AIT Austrian Institute of Technology; Markus Murschitz, AIT Austrian Institute of Technology; Oliver Zendel, AIT Austrian Institute of Technology; Markus Vincze, TU Wien

P-4A-40. **Pose Estimation Errors**, the Ultimate Diagnosis, Carolina Redondo-Cabrera, University of Alcalá; Roberto Lopez-Sastre, University of Alcalá; Yu Xiang, ; Tinne Tuytelaars, KU Leuven; Silvio Savarese

P-4A-41. **A Siamese Long Short-Term Memory Architecture for Human Re-identification**, Rahul Rama Varior, NTU Singapore; Bing Shuai, Nanyang Technological Univ; Jiwen Lu, Tsinghua University; Dong Xu, ; Gang Wang

P-4A-42. **Integration of Probabilistic Pose Estimates From Multiple Views**, Ozgur Erkent, Innsbruck University; Dadhichi Shukla, ; Justus Piater, University of Innsbruck

P-4A-43. **SurfCut: Free-Boundary Surface Extraction**, Marei Algarni, KAUST University; Ganesh Sundaramoorthi, KAUST

P-4A-44. **CATS: Co-saliency Activated Tracklet Selection for Video Co-localization**, Koteswar Jerripothula, Nanyang Technological University; Jianfei Cai, ; Junsong Yuan, Nanyang Technological University

P-4A-45. **Online Human Action Detection using Joint Classification-Regression Recurrent Neural Networks**, Yanghao Li, Peking University; Cuiling Lan, Microsoft Research; Junliang Xing, Chinese Academy of Sciences; Wenjun Zeng, Microsoft Research; Chunfeng Yuan, ; Jiaying Liu, Peking University

P-4A-46. **SyB3R: A Realistic Synthetic Benchmark for 3D Reconstruction from Images**, Andreas Ley, Technische Universität Berlin; Ronny Hänisch, Technische Universität Berlin; Olaf Hellwich, Technische Universität Berlin

P-4A-47. **A Deep Learning-based Approach to Progressive Vehicle Re-identification for Urban Surveillance**, Xinchen Liu, Beijing University of Posts and Telecommunication; Wu Liu, Beijing University of Posts and Telecommunication; Tao Mei, Microsoft Research Asia; Huadong Ma, Beijing University of Posts and Telecommunication

12:30 – 14:00  Lunch Break

14:00 – 15:00  Oral Session 4B (Action, Activity & Tracking)

**Chairs: Rita Cucchiara (U Modena), Juergen Gall (U Bonn)**

O-4B-01. **Spot On: Action Localization from Pointly-Supervised Proposals**, Pascal Mettes, University of Amsterdam; Jan van Gemert, Delft University of Technology; Cees Snoek, University of Amsterdam

O-4B-02. **Detecting Engagement in Egocentric Video**, Yu-Chuan Su, University of Texas at Austin; Kristen Grauman, University of Texas at Austin

O-4B-03. **Beyond Correlation Filters: Learning Continuous Convolution Operators for Visual Tracking**, Martin Danelljan, Linköping University; Andreas Robinson, Linköping University; Fahad Khan, Linköping University, Sweden; Michael Felsberg, Linköping University
**Spotlight Session 4B (Human & Face)**

Chairs: Maja Pantic (Imperial College), Bernt Schiele (MPI)

S-4B-05. **General Automatic Human Shape and Motion Capture Using Volumetric Contour Cues**, Helge Rhodin, MPI for Informatics; Nadia Robertini, Max-Planck-Institute for Informatics; Dan Casas, MPI; Christian Richardt, Intel VCI; Hans-Peter Seidel; Christian Theobalt, MPI Informatics

S-4B-06. **Globally Continuous and Non-Markovian Activity Analysis from Videos**, He Wang, Disney Research LA; Carol O’Sullivan, Trinity College Dublin

S-4B-07. **Joint Face Alignment and 3D Face Reconstruction**, Feng Liu, Sichuan University; Qijun Zhao, Sichuan University; Dan Zeng, Sichuan University; Xiaoming Liu, Michigan State University

S-4B-08. **Keep it SMPL: Automatic Estimation of 3D Human Pose and Shape from a Single Image**, Federica Bogo, Max Planck Institute Tuebingen; Angjoo Kanazawa, University of Maryland; Christoph Lassner, BCCN Tübingen; Peter Gehler; Javier Romero, MPI Intelligent Systems Tuebingen; Michael Black

S-4B-09. **Do We Really Need to Collect Millions of Faces for Effective Face Recognition?**, Iacopo Masi, USC; Anh Tran, USC; Tal Hassner, Open Univ Israel; Jatuporn Leksut, USC; Gerard Medioni, University of Southern California

**Coffee Break**

**Poster Session 4B**

Chairs: Jasper Uijlings (University of Edinburgh), Roberto Valenti (Sightcorp)

O-4B-01. **Spot On: Action Localization from Pointly-Supervised Proposals**, Pascal Mettes, University of Amsterdam; Jan van Gemert, Delft University of Technology; Cees Snoek, University of Amsterdam (shown in roof garden)

O-4B-02. **Detecting Engagement in Egocentric Video**, Yu-Chuan Su, University of Texas at Austin; Kristen Grauman, University of Texas at Austin (shown in roof garden)

O-4B-03. **Beyond Correlation Filters: Learning Continuous Convolution Operators for Visual Tracking**, Martin Danelljan, Linköping University; Andreas Robinson, Linköping University; Fahad Khan, Linköping University, Sweden; Michael Felsberg, Linköping University (shown in roof garden)

O-4B-04. **Look-ahead before you leap: end-to-end active recognition by forecasting the effect of motion**, Dinesh Jayaraman, UT Austin; Kristen Grauman, University of Texas at Austin (shown in roof garden)

S-4B-05. **General Automatic Human Shape and Motion Capture Using Volumetric Contour Cues**, Helge Rhodin, MPI for Informatics; Nadia Robertini, Max-Planck-Institute for Informatics; Dan Casas, MPI; Christian Richardt, Intel VCI; Hans-Peter Seidel; Christian Theobalt, MPI Informatics

S-4B-06. **Globally Continuous and Non-Markovian Activity Analysis from Videos**, He Wang, Disney Research LA; Carol O’Sullivan, Trinity College Dublin
S-4B-07. Joint Face Alignment and 3D Face Reconstruction, Feng Liu, Sichuan University; Qijun Zhao, Sichuan University; Dan Zeng, Sichuan University; Xiaoming Liu, Michigan State University

S-4B-08. Keep it SMPL: Automatic Estimation of 3D Human Pose and Shape from a Single Image, Federica Bogo, Max Planck Institute Tuebingen; Angjoo Kanazawa, University of Maryland; Christoph Lassner, BCCN Tübingen; Peter Gehler, ; Javier Romero, MPI Intelligent Systems Tuebingen; Michael Black

S-4B-09. Do We Really Need to Collect Millions of Faces for Effective Face Recognition?, Iacopo Masi, USC; Anh Tran, USC; Tal Hassner, Open Univ Israel; Jatuporn Leksut, USC; Gerard Medioni, University of Southern California

P-4B-10. Generative Visual Manipulation on the Natural Image Manifold, Jun-Yan Zhu, UC BERKELEY; Philipp Krahenbuhl, ; Eli Shechtman, Adobe; Alexei Efros,

P-4B-11. Deep Cascaded Bi-Network for Face Hallucination, Shizhan Zhu, Chinese University of HK; Sifei Liu, UC Merced; Chen-Change Loy, the Chinese University of Hong Kong; Xiaohou Tang, Chinese University of Hong Kong

P-4B-12. Cluster Sparsity Field for Hyperspectral Imagery Denoising, Lei Zhang, NPU; Wei Wei, Northwestern Polytechnical University; yanning Zhang, Northwestern Polytechnical University; Chunhua Shen, University of Adelaide; Anton Van den Hengel, University of Adelaide; Qinfeng Shi, The University of Adelaide


P-4B-14. Learning Common and Specific Features for RGB-D Semantic Segmentation with Deconvolutional Networks, Jinghua Wang, NTU; Zhenhua Wang, NTU, ROSE Lab; Dacheng Tao, University of Technology, Sydney; Simon See, ; Gang Wang

P-4B-15. MADMM: a generic algorithm for non-smooth optimization on manifolds, Artiom Kovnatsky, USI-Università della Svizzera ; Klaus Glashoff, USI-Università della Svizzera italiana; Michael Bronstein

P-4B-16. Interpreting the Ratio Criterion for Matching SIFT Descriptors, Avi Kaplan, Technion, Israel; Tamar Avraham, Technion; Michael Lindenbaum, Technion

P-4B-17. Semi-supervised learning based on joint diffusion of graph functions and Laplacians, Kwang In Kim, University of Bath

P-4B-18. Improving Semantic Embedding Consistency by Metric Learning for Zero-Shot Classification, Maxime Bucher, ONERA; Stephane Herbin, ONERA; Frederic Jurie, University of Caen

P-4B-19. A Sequential Approach to 3D Human Pose Estimation: Separation of Localization and Identification of Body Joints, Hoyub Jung, HUFS; Yumin Suh, Seoul National University; Gyeongsik Moon, Seoul National University; Kyoung Mu Lee

P-4B-20. A Novel Tiny Object Recognition Algorithm Based on Unit Statistical Curvature Feature, Yimei Kang, College of Software, Beihang University; Xiang Li, College of Software, Beihang University

P-4B-21. Head Reconstruction from Internet Photos, Shu Liang, ; Linda Shapiro, University of Washington; Ira Kemelmacher, University of Washington

P-4B-22. The Conditional Lucas & Kanade Algorithm, Chen-Hsuan Lin, Carnegie Mellon University; Rui Zhu, Carnegie Mellon University; Simon Lucey, CMU

P-4B-23. Where should saliency models look next?, Zoya Bylinskii, MIT; Adria Recasens, MIT; Ali Borji, University of Central Florida; Aude Oliva, MIT; Antonio Torralba, ; Fredo Durand

P-4B-24. Robust Face Alignment Using a Mixture of Invariant Experts, Oncel Tuzel, MERL; Tim Marks, MERL; Salil Tambe
P-4B-25. Partial Linearization based Optimization for Multi-class SVM, Pritish Mohapatra, IIIT, Hyderabad; Puneet Dokania, Ecole Centrale Paris; C.V. Jawahar, IIIT Hyderabad; M. Pawan Kumar, University of Oxford

P-4B-26. Search-based Depth Estimation via Coupled Dictionary Learning with Large-Margin Structure Inference, Yan Zhang, HIT; Rongrong Ji, Xiamen University; Xiaopeng Fan; Yan Wang, Microsoft; Feng Guo, Xiamen University; Yue Gao; debin Zhao

P-4B-27. Scalable Metric Learning via Weighted Approximate Rank Component Analysis, Cijo Jose, Idiap Research Institute; Francois Fleuret

P-4B-28. Spatio-Temporally Consistent Correspondence for Dense Dynamic Scene Modeling, Dinghuang Ji, UNC; Enrique Dunn, UNC Chapel Hill; Jan-Michael Frahm

P-4B-29. 3D Image Reconstruction from X-Ray Measurements with Overlap, Maria Klodt, University of Oxford; Raphael Hauser, University of Oxford

P-4B-30. DeeperCut: A Deeper, Stronger, and Faster Multi-Person Pose Estimation Model, Eldar Insafutdinov, Max-Planck Institute for Informatics; Bojan Andres, Max-Planck Institute for Informatics; Mykhaylo Andriluka, Max-Planck Institute for Informatics; Bernt Schiele

P-4B-31. Resonant Deformable Matching: Simultaneous Registration and Reconstruction, John Corring, University of Florida; Anand Rangarajan, University of Florida

P-4B-32. Unsupervised Learning of Visual Representations by Solving Jigsaw Puzzles, Mehdi Noroozi, University of Bern; Paolo Favaro

P-4B-33. COCO Attributes: Attributes for People, Animals, and Objects, Genevieve Patterson, Microsoft Research; James Hays, Georgia Institute of Technology

P-4B-34. Temporally Robust Global Motion Compensation by Keypoint-based Congealing, Seyed Morteza Safdarnejad, Michigan State University; Yousef Atoum, Michigan State University; Xiaoming Liu, Michigan State University

P-4B-35. Visualizing Image Priors, Tamar Shaham, Technion; Tomer Michaeli, Technion

P-4B-36. Exploiting Semantic Information and Deep Matching for Optical Flow, Min Bai, University of Toronto; Wenjie Luo, University of Toronto; Kaustav Kundu, University of Toronto; Raquel Urtasun

P-4B-37. Similarity Registration Problems for 2D/3D Ultrasound Calibration, Francisco Vasconcelos, UCL; Sebastien Ourselin, Donald Peebles, Danail Stoyanov

P-4B-38. Adaptive Signal Recovery on Graphs via Harmonic Analysis for Experiment Design in Neuroimaging, Won Hwa Kim, UW-Madison; Seong Jae Hwang, UW-Madison; Nagesh Adluru, UW-Madison; Sterling Johnson, UW-Madison; Vikas Singh, University of Wisconsin-Madison

P-4B-39. A Benchmark for Automatic Visual Classification of Clinical Skin Disease Images, SUN Xiao Xiao, NanKai University; Jufeng Yang, NanKai University; Ming Sun, NanKai University; Kai Wang, NanKai University

P-4B-40. Deep learning 3D shape surfaces using geometry images, Ayan Sinha, Purdue University; Jing Bai, Purdue University; Karthik Ramani, Purdue University


P-4B-42. Building Scene Models by Completing and Hallucinating Depth and Semantics, Miaomiao Liu, Data61,csiro, (NICTA); Xuming He, NICTA; Mathieu Salzmann, EPFL

P-4B-43. Weakly Supervised Learning of Heterogeneous Concepts in Videos, Sohil Shah, University of Maryland; Kuldeep Kulkarni, Arizona State University; Arijit Biswas, Amazon Development Center India; Ankit Gandhi, Xerox Research Centre India; Om Deshmukh, Xerox Research Centre India; Larry Davis, University of Maryland
P-4B-45. **Learning Semantic Deformation Flows with 3D Convolutional Networks**, M. Ersin Yumer, Adobe Research; Niloy Mitra, University College London

P-4B-46. **Recurrent Instance Segmentation**, Bernardino Romera-Paredes, University of Oxford; Philip Torr, Oxford University

P-4B-47. **Individualness and Determinantal Point Processes for Pedestrian Detection**, Donghoon Lee, Seoul National University; Ming-Hsuan Yang, UC Merced; Geonho Cha, Seoul National University; Songhwai Oh, Seoul National University
To celebrate the special occasion that both the European Conference on Computer Vision (8 – 16 October) and the ACM Multimedia Conference (15 – 19 October) meet in Amsterdam, the University of Amsterdam offers to registered attendees of either conference the opportunity to attend a maximum of two Invited tutorials by a renowned scientist in vision and/or multimedia on Saturday October 15th, free of charge.

The Invited tutorials are:

9:30 – 10:30  IT11 / Room C1.03  
Segmentation and Tracking, Mubarak Shah

9:30 – 10:30  IT12 / Room C0.01  
Variational wrapping deformations for image processing applications, Olga Sorkine-Hornung

10:45 – 11:45  IT21 / Room C1.03  
Developments in deep learning (tentative), Yoshua Bengio

10:45 – 11:45  IT22 / Room C0.01  
Multimodal human-robot interaction, Elisabeth Andre

11:45 – 13:00  Break

13:00 – 14:00  IT31 / Room C1.03  
Output embedding for large-scale visual recognition, Florent Perronnin

13:00 – 14:00  IT32 / Room C0.01  
Video event detection, localization and cross-media linking, Shih-Fu Chang

14:15 – 15:15  IT41 / Room C1.03  
Human body shape modelling, Michael Black

14:15 – 15:15  IT42 / Room C0.01  
A guide to creating and managing lifelog collections, Cathal Gurrin

15:30 – 16:30  IT51 / Room C1.03  
Deep visual understanding from deep learning, Jitendra Malik

15:30 – 16:30  IT52 / Room C0.01  
Bridging video and language with deep learning, Tao Mei

16:45 – 17:45  IT61 / Room C0.01  
Color and spectral information in modern computer vision, Sabine Süsstrunk
Workshops
Sunday, October 16

W23
Computer Vision for Audio-visual Media
Organizers: J-C. Bazin, Z. Zhang, W. Li
Location: Roeterseiland / Room C0.02

Please visit the website for the latest information on this workshop;
www.eccv2016.org/workshops

09:00 – 12:30 Workshops and Tutorials
ECCV Workshop on Computer Vision for Audio-visual Media
Room C0.02
W24

Computer VISION for ART Analysis

Organizers: J. P. Costeira, G. Carneiro, A. Del Bue, A. Elgammal, P. Hall, S. Lehmann, H. Brandhorst, E. Spratt

Location: Roeterseiland / Room M1.03

Please visit the website for the latest information on this workshop;
www.eccv2016.org/workshops

09:00 – 12:30  Workshops and Tutorials
ECCV Workshop on Computer VISION for ART Analysis
Room M1.03

12:30 – 14:00  Break

14:00 – 17:00  Workshops and Tutorials
ECCV Workshop on Computer VISION for ART Analysis
Room M1.03
W25
Virtual/Augmented Reality for Visual Artificial Intelligence

Organizers: A. López, A. Gaidon, G. Ros, E. Vig, D. Vázquez, H. Su, F. Perronnin
Location: Roeterseiland / Room C1.03

Please visit the website for the latest information on this workshop;
www.eccv2016.org/workshops

09:00 – 09:30 Welcome & Poster spotlights
09:30 – 10:00 Invited speaker 1
Andrea Vedaldi, Visual Geometry Group, Oxford
10:00 – 10:30 Invited speaker 2
Thomas Brox, University of Freiburg
10:30 – 11:00 Invited speaker 3
Ankur Handa, Dyson Lab, Imperial College
11:00 – 12:00 Coffee break + Posters
12:00 – 12:30 Invited speaker 4
Devi Parikh, Computer Vision Lab, Georgia Tech & Dpt. Electrical and Computer Engineering, Virginia Tech
12:30 – 13:00 Invited speaker 5
Bryan Russell, Creative Technologies Lab, Adobe Research
13:00 – 13:30 Invited speaker 6
Vladlen Koltun, Visual Computing Lab, Intel
13:30 – 14:00 Invited speaker 7
Adrien Treuille, VP of Simulation at Zoox & Computer Graphics Group, CMU

09:00 – 12:30 Workshops and Tutorials
ECCV Workshop on Virtual/Augmented Reality for Visual Artificial Intelligence
Room C1.03
W26

Joint Workshop on Storytelling with Images and Videos and Large Scale Movie Description and Understanding Challenge


Location: Roeterseiland / Room C1.04

Please visit the website for the latest information on this workshop; www.eccv2016.org/workshops

09:00 – 09:10 Introduction
09:10 – 09:40 Invited speaker 1
   Sanja Fidler, University of Toronto
09:40 – 10:10 Invited speaker 2
   Jason J. Corso, University of Michigan
10:10 – 10:30 Coffee Break
10:30 – 11:10 LSMDC Challenge Introduction
11:10 – 11:25 Winner 1 Spotlight Talk
11:25 – 11:40 Winner 2 Spotlight Talk
11:40 – 11:55 Winner 3 Spotlight Talk
12:00 – 13:30 Break + Poster (First 30 min)
13:30 – 14:00 Invited speaker 3
   Devi Parikh, Georgia Tech
   Learning Common Sense from Stories
14:00 – 15:00 Spotlight Talks for Accepted Papers
15:00 – 15:30 Invited speaker 4
   Cees Snoek, University of Amsterdam
   Recognizing events in videos without examples
15:30 – 15:40 Closing Remarks

09:00 – 12:30 Workshops and Tutorials
   ECCV Workshop on Storytelling with Images and Video and Large Scale Movie Description and Understanding Challenge
   Room C1.04

12:30 – 14:00 Break

14:00 – 17:00 Workshops and Tutorials
   ECCV Workshop on Storytelling with Images and Video and Large Scale Movie Description and Understanding Challenge
   Room C1.04
   Room M1.03
Workshops and Tutorials
ACM Multimedia 2016
Location: Roeterseiland

09:00 – 12:30  Workshops and Tutorials
  ECCV Workshop on Computer VISion for ART Analysis
  Room M1.03
  ECCV Workshop on Computer Vision for Audio-visual Media
  Room C0.02
  ECCV Workshop on Storytelling with Images and Video and Large Scale Movie Description
  and Understanding Challenge
  Room C1.04
  Workshop on Vision and Language Integration Meets Multimedia Fusion
  Room M1.02
  ECCV Workshop on Virtual/Augmented Reality for Visual Artificial Intelligence
  Room C1.03
  Workshop on Multimedia Alternate Realities
  Room C2.06
  Workshop on Human Behavior Understanding
  Room C0.01
  Workshop on Multimedia Assisted Dietary Management
  Room C2.05
  Workshop on Multimedia for personal health and health care
  Room C2.03
  Workshop on Lifelogging Tools and Applications
  Room C2.04
  Audio/Visual Emotion Challenge and Workshop
  Room C2.02
  Multimedia COMMONS Workshop 2016
  Room C2.01
  Tutorial on Social and Affective Robotics
  Room C3.01
  Tutorial on Technology and Art in Stimulating Creative Placemaking in Public-Use Spaces
  Room C2.07

12:30 – 14:00  Break

14:00 – 17:00  Workshops and Tutorials
  ECCV Workshop on Computer VISion for ART Analysis
  Room M1.03
  Tutorial on Emerging topics in learning from noisy and missing data
  Room C0.02
  ECCV Workshop on Storytelling with Images and Video and Large Scale Movie Description
  and Understanding Challenge
  Room C1.04
  Workshop on Vision and Language Integration Meets Multimedia Fusion
  Room M1.02
  Workshop on Multimedia Alternate Realities
  Room C2.06
Workshop on Human Behavior Understanding  
Room C0.01  

Workshop on Multimedia Assisted Dietary Management  
Room C2.05  

Workshop on Multimedia for personal health and health care  
Room C2.03  

Workshop on Lifelogging Tools and Applications  
Room C2.04  

Audio/Visual Emotion Challenge and Workshop  
Room C2.02  

Multimedia COMMONS Workshop 2016  
Room C2.01  

Tutorial on Social and Affective Robotics  
Room C3.01  

Tutorial on Multimedia Privacy  
Room C2.07  

Tutorial on Situation Recognition from Multimodal Data  
Room C3.02  

Tutorial on The Lifecycle of Geotagged Multimedia Data  
Room C3.03