Message from the General Chairs

Welcome to Venice, Italy, and to ICCV 2017, the 16th edition of the IEEE International Conference on Computer Vision.

Started in 1987, ICCV has soon established itself as the premier computer vision conference and a must-attend event for all those working in the field, both in academia and industry. Confirming the irresistible recent growth in computer vision and related areas, this year’s edition enjoyed a record level of participation. We received 2,143 valid submissions, with an increase of 26% from the previous edition. After a careful selection process coordinated by the Program Chairs and described in detail in their editorial, 621 papers were accepted for publication and presentation in the main program. Continuing a recently established tradition within our community, presentations are of three kinds: oral, spotlight and poster. All oral and spotlight presentations are video recorded and posted on the Computer Vision Foundation YouTube Channel.

As customary, in addition to the main conference sessions, this year’s program features a number of co-located events including a record number of 44 workshops (63% more than the previous edition), 9 tutorials, a doctoral consortium, industrial exhibitions, demos, and exciting social activities in a unique historical setting. Industrial participation is also very high this year. Over 60 companies, from promising startups to the biggest industry leaders, are showcasing their latest technologies at ICCV 2017 and are supporting the conference through different levels of sponsorship. In fact, we tripled the amount of industrial sponsorship from the previous edition. Needless to say, we are very pleased to see how our community is growing at such an astonishing pace and how our research work is attracting more and more interest on the part of industry, which is a clear sign of maturity of our field.

ICCV 2017 is held at the Venice Convention Center, located on Lido Island, in the prestigious area recognized worldwide as the setting for the annual Venice Film Festival, the oldest film festival in the world. The social program includes a welcome reception at the conference venue, in the “Palazzo del Casinò,” a monumental building in rationalist style built in the late 1930s and used for more than fifty years as the summer location of the Venice Casino, and in the “Palazzo del Cinema,” where the film festival takes place. We’ll also have a banquet at the Venetian Arsenal, one of the most impressive testimonies of the military power of the ancient Republic of Venice. This is a complex of former shipyards and armories which used to be the heart of the Venice naval industry from the XII century. With its thousands of operators employed, it is considered to be one of the world’s largest production centers in the pre-industrial age. Galileo, who served at the Arsenal as a consultant in 1593, used it as the setting of his last, epoch-making book *Two New Sciences*.

We wish to thank all members of the organization committee, area chairs, reviewers and authors for making this event possible. Special thanks go to Andrea Prati and Andrea Vedaldi, our local committee team, for their invaluable help in dealing with all practical aspects of the conference organization, and to Giovanni Farinella for setting up the conference webpage and keeping it up to date over the past three years. We are also grateful to Eric Mortensen for overseeing the whole publication process. Our thanks also go to the organizers of previous conferences for their helpful advice and support. The conference would not be possible in its current form without the generosity of our corporate partners and their support is deeply appreciated. We are also indebted to the technical staff of Ca’ Foscari University for its assistance. Last but not least, we are grateful to Cristiana Fiandra and her team at “The office” for managing the conference organization.

In closing, we wish all the attendees an inspiring and informative conference, and a memorable stay in one of the most beautiful and unique cities in the world.

General Chairs:  Katsushi Ikeuchi, Gérard Medioni, Marcello Pelillo
ICCV is the premier conference in Computer Vision, held bi-annually and attracting submissions worldwide in all areas of the field. The submission and review process gives us an opportunity to assess the state of a field experiencing rapid evolution. As evidenced by its growth and impact, the state of the field is strong. Submissions have grown in comparison to prior ICCV’s (2143 this year, 1698 for ICCV 2015). The overall acceptance rates have remained steady over the years (28.9% this year, 30.3% for ICCV 2015).

ICCV has traditionally been a single-track conference, and we have upheld this tradition in 2017. This choice presents challenges given the number of submissions and the “noise” intrinsic in the review process. Over the years, attendees and organizers have debated whether papers selected for oral presentation are distinguishably better than others. The tradition in Computer Vision conferences to give equal space in the proceedings to orals and posters reflects the belief that all accepted papers are equally worthy of the attention of the community. Recent trends have also seen decreased length allocated to oral presentations, and the birth of an intermediate category of “spotlight presentations”, or short orals, of increasing length. At ICCV 17, we have reinforced the trend, giving each oral presentation 12 minutes, each spotlight 3.5 minutes. At the end of the review process, 45 (2.09%) papers were accepted for oral presentation, 56 (2.61%) for spotlights, and 520 (24.26%) for poster, for a total of 621 accepted papers out of 2143 valid submissions (28.9% overall acceptance rate).

The review process was designed by the Program Chairs (PCs), in consultation with Chairs of previous conferences. In line with other conferences, PCs were aided in the design of the program by Area Chairs (ACs), who in turn oversaw the work of Peer Reviewers (PRs). The Program Committee thus consisted of 123 ACs, selected according to criteria including experience, rejuvenation, diversity, thematic and geographic distribution. We abided by the informal rule of excluding from consideration any individual who had served as Area Chair in more than 4 consecutive major conferences, and added 34 first-time Area Chairs. We hope future PCs also adopt these guidelines. PRs were selected considering reviewers from prior conferences, journals, as well as recommendations by ACs and PCs. A handful of self-nominations were also considered, a small subset of which deemed viable for inclusion in the reviewer pool. Geographic distribution and diversity remain a challenge. ACs included 44 researchers from European institutions, 22 from Asian, 4 from Australian and 53 from North American ones. 14 of them were female. 15 had a primary academic affiliation, 108 a primary industrial affiliation. We did not collect demographic or ethnic data. In a rapidly growing field with global reach, more attention should be devoted to ensuring that sufficient representation is achieved, especially from developing countries. Peer Reviewers were mostly researchers in Academia, Industry and Government. Having a large pool of Program Committee members allowed us to match expertise while bounding reviewer loads. Typically five papers, but never more than eight, were assigned to a Program Committee member. Graduate students had a maximum of four papers to review. A number of reviewers and Area Chairs were held back a “wild cards” to manage last-minute loads originating from others’ failure to complete their work by the due dates.

As with most recent conferences, there was considerable imbalance in seniority and experience in the reviewer pool as well as the Area Chair body. To avoid undue bias by senior researchers, the decision was made early on to conduct the review process electronically and anonymously, where Area Chairs were allowed to select and knew the identity of Reviewers, but Reviewers did not know the identity of Area Chairs. Similarly, Area Chairs did not know the identity of their peer ACs during discussions. There are pros and cons to this choice, evidenced by the fact that different conferences in the recent past have opted variously to conduct the Area Chair meeting in person, while others have done so remotely. Remote meetings allow for better management of conflicts, and since there is a recorded trail of all conversations and decisions, there is better accountability and transparency. The challenge is calibration as well as the ability to manage deadlines: A physical AC meeting forces all works to be completed in time, whereas
managing a remote process with a large volunteer workforce presents a significant challenge. In the case of ICCV 2017, the decision to conduct the AC process remotely was part of the conference bid in 2013, made in part on budget considerations since at that time the AC meeting was the heaviest line on the budget. Since then, the finances of Computer Vision conferences have improved significantly, so future chairs are not tied to a particular format.

Submission to ICCV was, nominally, double-blind: Authors did not know the identity of reviewers or Area Chairs, and papers are submitted anonymously. This is in line with tradition. However, recent trends have brought significant challenges to this model. It is now customary for researchers to post preliminary, non-peer reviewed work on public repositories, such as ArXiv, that are indexed by search engines (e.g., Google Scholar), and therefore are disseminated before they are vetted. This complicates the assessment of concurrent and prior art, content overlap, plagiarism and self-plagiarism. A few past incidents whereby non-peer reviewed papers were publicized broadly, challenging the foundations of the peer-reviewed system, prompted the PAMI TC to issue guidelines on this topic. Some conferences have chosen to disqualify submissions that were publicly disclosed, prompting backlash from authors. Dissemination prior to review effectively makes the double-blind process single-blind (reviewers can easily find the identity of the authors), and creates an accelerated cycle whereby papers become obsolete by the time the peer review process is over. Even more concerning is the “flag-planting” phenomenon, whereby authors rush to post preliminary work only to claim primacy later when more complete and peer-reviewed work appears in press. In this complex landscape, we have chosen to allow authors to post their work on ArXiv, but we have requested that such work not be discussed in the media, or promoted on social media or in wide-audience venues. We have requested that reviewers refrain from searching for the identity of the authors, but recognize that the interpretation of the honor code is inhomogeneous and generally ineffective. But while we have not disallowed dissemination before evaluation, we strongly believe that non-peer reviewed work should not be considered prior art, and authors should not be requested to compare their work to non-peer reviewed work, nor obliged to cite it. This creates at least a minimal cool-off period, and encourages publication of research work that aims at remaining relevant for at least the few months while it is being peer-reviewed. We feel this is a viable compromise between letting all flowers bloom, but not fostering dissemination of work that becomes obsolete by the time it is properly peer-reviewed. In the future, we hope that Program Chairs will make it a rule that authors should not be required to reference, cite, or compare the work to any disseminated work that has not been peer-reviewed, even though some do so spontaneously. Almost all controversies during the review process were related to arXiv submissions. Most of the reviewer inquiries were resolved by giving the benefit of the doubt to ICCV authors. We have also carefully checked other cases of double-submissions with the chairs of concurrent conferences: NIPS, BMVC, and SIGGRAPH Asia. 20 overlapping submissions were detected and rejected ex-officio. While a small number relative to the submitted pool, future Program Chairs should make it clear that such practices are unacceptable.

The initial reviewing period was five weeks long, after which reviewers provided reviews with preliminary recommendations. With the generous help of several last-minute reviewers, each paper received three reviews. Submissions with all three reviews suggesting rejection were independently checked by two ACs and if they agreed, the manuscript was rejected at this stage (“early rejects”). In total, 379 manuscripts (17.68%) were early-rejected, reducing the average AC load to about 27. Authors of the remaining submissions were then given the opportunity to rebut the reviews, primarily to identify factual errors. Following this, reviewers and ACs discussed papers at length, after which reviewers finalized their reviews and gave a final recommendation to the ACs. Each manuscript was evaluated independently by two ACs. In most of the cases, after extensive discussions, the two ACs reached consensus. In the very few borderline cases where an agreement was not reached, the PCs acted as tie-breakers, without entering the merit of any specific
Editorial Message from the Program Chairs

paper, but using broad rules to decide which were above threshold. While in a handful of cases worthwhile papers were rejected early, without giving authors the opportunity to rebut, lightening the load was overall beneficial to the conference.

In the rebuttal phase, we found that some reviewers and Area Chairs have come to interpret it as a second round of reviews, akin to an archival journal. We believe that, in such a fast-moving field, the role of the rebuttal should be to clarify the content of the paper, rather than to add additional material. For this reason we limited the rebuttal to text, and we have explicitly pushed back against reviewer or Area Chairs who demanded additional experiments or comparisons.

At least three Peer Reviewers and two Area Chairs evaluated every single paper. This is, in our opinion, the minimum amount of attention that a submission deserves. However, since there is no barrier to submission nor memory, there is a vast body of papers that are passed from one conference to the next, creating an enormous burden on the community that keeps re-reviewing the same (large number of) rejected papers. In effect, given the inherent stochasticity in the review process, substandard submissions function as lottery tickets, except that the lottery operates in reverse: Tickets are free (authors do not pay to submit), and winners pay (conference registration). There have been past proposals to reverse the process (pay to submit, registration is free if accepted) originally proposed by the Program Chairs of ICCV 2011 but never implemented. Future Program Chairs are encouraged to explore ways to align the incentives in the review process to make it function more efficiently.

There are other areas where incentives are misaligned: The best reviewer or Area Chair for a given paper is likely one with a similar or related paper in submission. Given the zero-sum nature of the conference, this creates a conflict, which is currently managed by relying on the good faith and professionalism of individuals. This is in part ameliorated by Best Reviewer Awards. As a community, we should give more weight to these awards, beyond their monetary value, as quality reviewing is vital to the successful progression of a scientific discipline. Senior members of the community should encourage giving weight to these awards in tenure and promotion proceedings. General Chairs should consider devoting a significant portion of the conference budget to paying reviewers, or to designate substantial amounts for the Best Reviewer Awards.

The process of Best Paper Awards is also a delicate tradeoff. Traditionally, Program Chairs are responsible for assigning Best Paper and Best Student Paper Awards, whereas the PAMI TC Meeting has been assigning career-level awards not tied to the conference program. The Program Chairs of CVPR 2015 first introduced a test-of-time award, for papers published up to 10 years prior to the current conference (in that case, all CVPR’s up to 2005). An external committee was put in place by the Program Chairs, independent of the conference Chairs or the PAMI TC. Since then, the PAMI TC has taken the assignment of such awards under its wing. Traditionally, Best Papers are chosen among candidate papers recommended by Area Chairs. Again, since the conference is a zero-sum game and Area Chairs represent some of the most successful researchers in the field, this creates a (real or perceived) conflict with no easy solution, since an Awards Committee cannot realistically look at dozens if not hundreds of candidates if one were to consider every oral, or even every paper, as an award candidate. We have chosen to elect an external committee to look at nominated papers, but given them leeway to consider other papers from the program if they deem them so relevant and if there is no conflict, real or perceived, with any member of the committee.

Particular attention was paid to handling conflicts. Conflicts of interest between ACs, Program Committee members, and papers were identified based on the authorship of ICCV 2017 submissions, on the home institutions and on previous collaborations of all researchers involved. To find institutional conflicts, all authors, Program Committee members, and ACs were asked to list the Internet domains of their current institutions. To find collaborators, the Researcher.cc database (http://researcher.cc/), funded by the
Editorial Message from the Program Chairs

Computer Vision Foundation (CVF), was used to find any co-authored papers in the period 2012-2017. We pre-assigned approximately 100 papers to each AC, based on affinity scores from the Toronto Paper Matching System. ACs then bid on these, indicating their level of expertise. Based on these bids, and conflicts of interest, approximately 35 papers were assigned to each AC. The ACs then suggested seven reviewers from the pool of Program Committee members for each paper, in ranked order, from which three were chosen automatically by CMT, taking load balancing and conflicts of interest into account. Authors were not allowed to suggest preferred reviewers or Area Chairs. Similarly, they were not allowed to name reviewers-non-grati. In a now hyper-competitive field, future Program Chairs may want to consider implementing this option.

General Chairs were not involved in any part of the decision process. Program Chairs did not weigh in on any individual decision except where conflicts or complaints arose. Program Chairs were not allowed to co-author any submission to the conference, and none of the PhD students, postdocs or co-workers of the PCs submitted any papers. This arrangement should become standard of all future conferences.

We want to thank everyone involved in making the ICCV 2017 possible. First and foremost, the quality of the program depends on the quality of papers submitted by authors, and on the very hard work of the ACs and the Program Committee members. We are particularly grateful to Rene Vidal for his continuous support and sharing experience from ICCV 2015, to Laurent Charlin for the use of the Toronto Paper Matching System, to Ari Kobren for the use of the Researcher.cc tools, to CVF for facilitating the use of the iThenticate plagiarism detection software and to Gloria Zen and Radu-Laurentiu Vieriu for setting up CMT, Microsoft’s Academic Conference Management Service.

Program Chairs: Rita Cucchiara, Yasuyuki Matsushita, Nicu Sebe, Stefano Soatto

ICCV 2017 Organizing Committee

General Chairs: Katsushi Ikeuchi
Gérard Medioni
Marcello Pelillo
Program Chairs: Rita Cucchiara
Yasuyuki Matsushita
Nicu Sebe
Stefano Soatto
Honorary Chair: Jan Koenderink
Workshops Chairs: Sing Bing Kang
Jingyi Yu
Tutorials Chairs: Tal Hassner
Vittorio Murino
Demos Chair: Andrew D. Bagdanov
Website Chair: Giovanni Maria Farinella
Publications Chair: Eric Mortensen
Industrial Chairs: Florent Perronnin
Walter Scheirer
Ning Zhang
Finance Chairs: Octavia Camps
Ramin Zabih
Local Arrangements Chairs: Andrea Prati
Andrea Vedaldi
Video Proceedings Chair: Roberto Vezzani
Publicity Chair: Ali Shokoufandeh
Doctoral Consortium Chair: Samuel Rota Bulò
Student Volunteers Chair: Lamberto Ballan
Organizing Secretariat: the office srl
ICCV 2017 Area Chairs

Xavier Alameda-Pineda  
Yiannis Aloimonos  
Michael Bronstein  
Michael S. Brown  
Barbara Caputo  
Antoni B. Chan  
Manmohan Chandraker  
Ramalingam Chellappa  
James Clark  
Daniel Cremers  
Kostas Daniilidis  
Andrew Davison  
Alberto Del Bimbo  
Luigi Di Stefano  
Piotr Dollar  
Gianfranco Doretto  
Giovanni Maria Farinella  
Paolo Favaro  
Vittorio Ferrari  
Sanja Fidler  
John W. Fisher III  
Charless Fowlkes  
Mario Fritz  
Yun Raymond Fu  
Yasutaka Furukawa  
Fabio Galasso  
Theo Gevers  
Michael Goesele  
Kristen Grauman  
Sergio Escalera Guerrero  
Leonidas Guibas  
Bohyung Han  
Kaiming He  
Derek Hoiem  
Byung-Woo Hong  
Timothy Hospedales  
Gang Hua  
Hiroshi Ishikawa  
C. V. Jawahar  
Hervé Jégou  
Jaïa Jia  
Hailin Jin  
Fredrik Kahl  
Seon Joo Kim  
Hedvig Kjellström  
Iasonas Kokkinos  
Vladlen Koltun  
Philipp Krähenbühl  
Kyros Kutulakos  
In So Kweon  
Shang-Hong Lai  
Svetlana Lazebnik  
Ales Leonardis  
Hongdong Li  
Jia Li  
Yi Ma  
Roberto Manduchi  
Jiri Matas  
Krystian Mikolajczyk  
Peyman Milanfar  
Yael Moses  
Mario Enrique Munich  
Vittorio Murino  
Hajime Nagahara  
Ko Nishino  
Tomas Pajdla  
Maja Pantic  
Devi Parikh  
Ioannis Patras  
Pietro Perona  
Robert Pless  
Marc Pollefeys  
Guo-Jun Qi  
Ian Reid  
Elisa Ricci  
Lorenzo Rosasco  
Amit Roy-Chowdhury  
Kate Saenko  
Yoichi Sato  
Shin'ichi Satoh  
Silvio Savarese  
Bernt Schiele  
Cordelia Schmid  
Stan Sclaroff  
Fei Sha  
Sudipta Sinha  
Josef Sivic  
Cristian Sminchisescu  
Cees Snoek  
Yusuke Sugano  
Ganesh Sundaramoorthi  
Yu-Wing Tai  
Chi Keung Tang  
Qi Tian  
Carlo Tomasi  
Antonio Torralba  
Lorenzo Torresani  
Sotirios A. Tsaftaris  
John K. Tsotsos  
Zhuowen Tu  
Raquel Urtasun  
Joost van de Weijer  
Nuno Vasconcelos  
Olga Veksler  
Roberto Vezzani  
Rene Vidal  
Xiaogang Wang  
Jingdong Wang  
Gang Wang  
David Wipf  
Lior Wolf  
John Wright  
Ying Wu  
Tao Xiang  
Dong Xu  
Ming-Hsuan Yang  
Ruigang Yang  
Yi Yang  
Allen Y. Yang  
Jingyi Yu  
Lei Zhang  
Song-Chun Zhu  
Larry Zitnick

QR Codes for the Conference Mobile App (by IEEE CPS)

Apple iTunes App Store (iOS)

Google Play (Android)
We are pleased to recognize the following researchers as “ICCV 2017 Outstanding Reviewers”. These reviewers were identified by one or more of the ICCV Area Chairs for their hard work in providing high quality and detailed reviews for their assigned papers.

Yannis Avrithis  Lisa Anne Hendricks  David Nilsson  Giorgos Tolias
Hakan Bilen  Go Irie  Deepak Pathak  Yi-Hsuan Tsai
Marten Bjoerkman  Yu-Gang Jiang  David Picard  Jasper Uijlings
Andrei Bursuc  Yannis Kalantidis  Leonid Pishchulin  Yashaswi Verma
Dapeng Chen  Margret Keuper  Alin Popa  Christoph Vogel
Dong Chen  Kris Kitani  Xiaojuan Qi  Liwei Wang
John Collomosse  Piotr Koniusz  Ergys Ristani  Michael Weinmann
Dima Damen  Yulia Kotseruba  Emanuele Rodola  Michael Wilber
Teofilo de Campos  Stefan Lee  Anna Rohrbach  Kyle Wilson
Santosh Kumar Divvala  Weixin Li  Javier Romero  Calden Wloka
Basura Fernando  Wen Li  Johannes Schoenberger  Angela Yao
Boris Flach  Haibin Ling  Ramprasaath Selvaraju  Sai Kit Yeung
Wolfgang Foerstner  Sifei Liu  Arnold Smuelders  Andrei Zanfir
Simone Frintrop  Elisabetta Marinoiu  Chris Sweeney  Quanshi Zhang
Diego Gutierrez  Pablo Mesejo-Santiago  Huixuan Tang  Guangming Zhu
Christian Haene  Pedro Morgado  Tatsunori Taniai
Wolfgang Heidrich  Varun Nagaraja  Damien Teney

We also want to recognize the following researchers as “ICCV 2017 Emergency Reviewers”. These reviewers were willing to provide an “emergency” review on short notice within a very short timeframe. Thank you for your service.

Sagie Benaim  Margret Keuper  Pablo Mesejo Santiago  Tatsunori Taniai
Marten Bjoerkman  Akisato Kimura  Pascal Mettes  Akihiko Torii
Andreas Bulling  Wei-Sheng Lai  Moin Nabi  Sergey Tulyakov
Shay Deutsch  Stephane Lathuiliere  Tae-Hyun Oh  Yoshitaka Ushiku
Abhinal Dhall  Joon-Young Lee  Jinshan Pan  Lior Uzan
William Freeman  Hsin-Ying Lee  Jaesik Park  Wei Wang
Amir Ghodrati  Gil Levi  Jordi Pont-Tuset  Jiajun Wu
Valerio Giuffrida  Yijun Li  Enver Sangineto  Tao Wu
Giann Gorospe  Xiao Lin  Boxin Shi  Takayoshi Yamashita
Luis Heranz  Sifei Liu  Gregory Slabaugh  Yan Yan
Jia-Bin Huang  Kevin Ma  Doron Sobol  Gloria Zen
Wei-Chih Hung  Yasushi Mekihara  Ramanathan Subramanian  Yinqiang Zheng
Asako Kamezaki  Yusuke Matsui  Chong Sun
Tuesday, October 24

0730–1730 Registration (Palazzo del Casinò Terrace)

0830–0900 Opening Remarks (Sala Grande, Sala Darsena, & Sala Perla)

0900–1030 Session O1-S1: 3D Vision & Video Analysis (Sala Grande, Sala Darsena, & Sala Perla)

Papers in this session are also in Poster Session P1.

Chairs: Raquel Urtasun (Uber ATG; Univ. of Toronto)
Daniel Cremers (Technical Univ. of Munchen)

0900 Orals (O1)

Format (9 min. for presentation + 3 min. for questions)

1. [0900] Globally-Optimal Inlier Set Maximisation for Simultaneous Camera Pose and Feature Correspondence, Dylan Campbell, Lars Petersson, Laurent Kneip, Hongdong Li
2. [0912] Robust Pseudo Random Fields for Light-Field Stereo Matching, Chao-Tsung Huang
4. [0936] Distributed Very Large Scale Bundle Adjustment by Global Camera Consensus, Runze Zhang, Siyu Zhu, Tian Fang, Long Quan
5. [0948] Practical Projective Structure From Motion (P²SfM), Ludovic Magerand, Alessio Del Bue

1000 Spotlights (S1)

Format (3 min. & 30 sec. for presentation; no questions)

1. [1000] Anticipating Daily Intention Using On-Wrist Motion Triggered Sensing, Tz-Ying Wu, Ting-An Chien, Cheng-Sheng Chan, Chan-Wei Hu, Min Sun
2. [1003] Rethinking Reprojection: Closing the Loop for Pose-Aware Shape Reconstruction From a Single Image, Rui Zhu, Hamed Kiani Galoogahi, Chaoyang Wang, Simon Lucey
4. [1011] Using Sparse Elimination for Solving Minimal Problems in Computer Vision, Janne Heikkilä
5. [1015] High-Resolution Shape Completion Using Deep Neural Networks for Global Structure and Local Geometry Inference, Xiaoguang Han, Zhen Li, Haibin Huang, Evangelos Kalogerakis, Yizhou Yu
7. [1022] Learning Policies for Adaptive Tracking With Deep Feature Cascades, Chen Huang, Simon Lucey, Deva Ramanan
8. [1026] Temporal Shape Super-Resolution by Intra-Frame Motion Encoding Using High-Fps Structured Light, Yuki Shiba, Satoshi Ono, Ryo Furukawa, Shinsaku Hiura, Hiroshi Kawasaki

1030–1115 Morning Break

1030–1230 Poster Session P1 (Sala Laguna, Salone Adriatico, & Sale Mosaici 1–2)

See poster locations map on pg. 1.

Oral O1 Posters

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

Spotlight S1 Posters

6. Anticipating Daily Intention Using On-Wrist Motion Triggered Sensing, Tz-Ying Wu, Ting-An Chien, Cheng-Sheng Chan, Chan-Wei Hu, Min Sun
7. Rethinking Reprojection: Closing the Loop for Pose-Aware Shape Reconstruction From a Single Image, Rui Zhu, Hamed Kiani Galoogahi, Chaoyang Wang, Simon Lucey
9. Using Sparse Elimination for Solving Minimal Problems in Computer Vision, Janne Heikkilä
10. High-Resolution Shape Completion Using Deep Neural Networks for Global Structure and Local Geometry Inference, Xiaoguang Han, Zhen Li, Haibin Huang, Evangelos Kalogerakis, Yizhou Yu
12. Learning Policies for Adaptive Tracking With Deep Feature Cascades, Chen Huang, Simon Lucey, Deva Ramanan
13. Temporal Shape Super-Resolution by Intra-Frame Motion Encoding Using High-Fps Structured Light, Yuki Shiba, Satoshi Ono, Ryo Furukawa, Shinsaku Hiura, Hiroshi Kawasaki

**3D Computer Vision**
14. Real-Time Monocular Pose Estimation of 3D Objects Using Temporally Consistent Local Color Histograms, Henning Tjaden, Ulrich Schwancke, Elmar Schömer
15. CAD Priors for Accurate and Flexible Instance Reconstruction, Tolga Birdal, Slobodan Ilic
16. Colored Point Cloud Registration Revisited, Jaesik Park, Qian-Yi Zhou, Vladlen Koltun
17. Learning Compact Geometric Features, Marc Khoury, Qian-Yi Zhou, Vladlen Koltun
18. Joint Layout Estimation and Global Multi-View Registration for Indoor Reconstruction, Jeong-Kyun Lee, Jaewon Yea, Min-Gyu Park, Kuk-Jin Yoon

**Biomedical Image Analysis**
20. An Optimal Transportation Based Univariate Neuroimaging Index, Liang Mi, Wen Zhang, Junwei Zhang, Yonghui Fan, Dhruman Goradia, Kewei Chen, Eric M. Reiman, Xianfeng Gu, Yalin Wang

**Face & Gesture**
21. S3FD: Single Shot Scale-Invariant Face Detector, Shifeng Zhang, Xiangyu Zhu, Zhen Lei, Hailin Shi, Xiaobo Wang, Stan Z. Li

---

**Low-Level Vision & Image Processing**
22. Amulet: Aggregating Multi-Level Convolutional Features for Salient Object Detection, Pingping Zhang, Dong Wang, Huchuan Lu, Hongyu Wang, Xiang Ruan
23. Learning Uncertain Convolutional Features for Accurate Saliency Detection, Pingping Zhang, Dong Wang, Huchuan Lu, Hongyu Wang, Baocai Yin
24. Zero-Order Reverse Filtering, Xin Tao, Chao Zhou, Xiaoyong Shen, Joe Wang, Jiaya Jia
25. Learning Blind Motion Deblurring, Patrick Wieschollek, Michael Hirsch, Bernhard Schölkopf, Hendrik P. A. Lenz
27. Learning to Super-Resolve Blurry Face and Text Images, Xiangyu Xu, Deqing Sun, Jinshan Pan, Yujin Zhang, Hanspeter Pfister, Ming-Hsuan Yang

**Motion & Tracking**
29. Deep Occlusion Reasoning for Multi-Camera Multi-Target Detection, Pierre Baqué, François Fleuret, Pascal Fua
30. Encouraging LSTMs to Anticipate Actions Very Early, Mohammad Sadegh Aliakbarian, Fatemeh Sadat Saleh, Mathieu Salzmann, Basura Fernando, Lars Petersson, Lars Andersson
31. PathTrack: Fast Trajectory Annotation With Path Supervision, Santiago Manen, Michael Gygli, Dengxin Dai, Luc Van Gool
32. Tracking the Untrackable: Learning to Track Multiple Cues With Long-Term Dependencies, Amir Sadeghian, Alexandre Alahi, Silvio Savarese
33. MirrorFlow: Exploiting Symmetries in Joint Optical Flow and Occlusion Estimation, Junhwa Hur, Stefan Roth
34. Tracking as Online Decision-Making: Learning a Policy From Streaming Videos With Reinforcement Learning, James Supančič, III, Deva Ramanan

**Optimization Methods**
35. Non-Convex Rank/Sparse Regularization and Local Minima, Carl Olsson, Marcus Carlsson, Fredrik Andersson, Viktor Larsson
36. A Revisit of Sparse Coding Based Anomaly Detection in Stacked RNN Framework, Weixin Luo, Wen Liu, Shenghua Gao

Recognition
37. HydraPlus-Net: Attentive Deep Features for Pedestrian Analysis, Xihui Liu, Haiyu Zhao, Maoqing Tian, Lu Sheng, Jing Shao, Shuai Yi, Junjie Yan, Xiaogang Wang
38. No Fuss Distance Metric Learning Using Proxies, Yair Movshovitz-Attias, Alexander Toshev, Thomas K. Leung, Sergey Ioffe, Saurabh Singh
39. Benchmarking and Error Diagnosis in Multi-Instance Pose Estimation, Matteo Ruggero Ronchi, Pietro Perona
40. Orientation Invariant Feature Embedding and Spatial Temporal Regularization for Vehicle Re-identification, Zhongdao Wang, Luming Tang, Xihui Liu, Zhuliang Yao, Shuai Yi, Jing Shao, Junjie Yan, Shengjin Wang, Hongsheng Li, Xiaogang Wang
41. Fashion Forward: Forecasting Visual Style in Fashion, Ziad Al-Halah, Rainer Stiefelhagen, Kristen Grauman
42. Towards 3D Human Pose Estimation in the Wild: A Weakly-Supervised Approach, Xingyi Zhou, Qixing Huang, Xiao Sun, Xiangyang Xue, Yichen Wei
43. Flow-Guided Feature Aggregation for Video Object Detection, Xizhou Zhu, Yujie Wang, Jifeng Dai, Lu Yuan, Yichen Wei
44. Reasoning About Fine-Grained Attribute Phrases Using Reference Games, Jong-Chyi Su, Chenyun Wu, Huaizu Jiang, Subhransu Maji
45. DeNet: Scalable Real-Time Object Detection With Directed Sparse Sampling, Lachsen Tychsen-Smith, Lars Petersson
46. MIHash: Online Hashing With Mutual Information, Fatih Cakir, Kun He, Sarah Adel Bargal, Stan Sclaroff
47. SafetyNet: Detecting and Rejecting Adversarial Examples Robustly, Jiajun Lu, Theerasit Issaranon, David Forsyth
48. Recurrent Models for Situation Recognition, Arun Mallya, Svetlana Lazebnik
49. Multi-Label Image Recognition by Recurrently Discovering Attentional Regions, Zhouxia Wang, Tianshui Chen, Guanbin Li, Ruijia Xu, Liang Lin
50. Deep Determinantal Point Process for Large-Scale Multi-Label Classification, Pengtao Xie, Ruslan Salakhutdinov, Luntian Mou, Eric P. Xing
52. Neural Person Search Machines, Hao Liu, Jiashi Feng, Zequn Jie, Karlekar Jayashree, Bo Zhao, Meibin Qi, Jianguo Jiang, Shuicheng Yan
53. DualNet: Learn Complementary Features for Image Recognition, Saihui Hou, Xu Liu, Zilei Wang
55. Show, Adapt and Tell: Adversarial Training of Cross-Domain Image Captioner, Tseng-Hung Chen, Yuan-Hong Liao, Ching-Yao Chuang, Wan-Ting Hsu, Jianlong Fu, Min Sun
56. Attribute Recognition by Joint Recurrent Learning of Context and Correlation, Jingya Wang, Xiatian Zhu, Shaogang Gong, Wei Li
57. VegFru: A Domain-Specific Dataset for Fine-Grained Visual Categorization, Saihui Hou, Yushan Feng, Zilei Wang
58. Increasing CNN Robustness to Occlusions by Reducing Filter Support, Elad Osherov, Michael Lindenbaum
59. Exploiting Multi-Grain Ranking Constraints for Precisely Searching Visually-Similar Vehicles, Ke Yan, Yonghong Tian, Yaowei Wang, Wei Zeng, Tiejun Huang
60. Recurrent Scale Approximation for Object Detection in CNN, Yu Liu, Hongyang Li, Junjie Yan, Fangyin Wei, Xiaogang Wang, Xiaou Tang

Segmentation, Grouping & Shape
61. Embedding 3D Geometric Features for Rigid Object Part Segmentation, Yafei Song, Xiaowu Chen, Jia Li, Qinping Zhao

Statistical Methods & Learning
63. When Unsupervised Domain Adaptation Meets Tensor Representations, Hao Lu, Lei Zhang, Zhiguo Cao, Wei Wei, Ke Xian, Chunhua Shen, Anton van den Hengel
64. Look, Listen and Learn, Relja Arandjelović, Andrew Zisserman
65. Grad-CAM: Visual Explanations From Deep Networks via Gradient-Based Localization, Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, Dhruv Batra
66. Image-Based Localization Using LSTMs for Structured Feature Correlation, Florian Walch, Caner Hazirbas, Laura Leal-Taixé, Torsten Sattler, Sebastian Hilsenbeck, Daniel Cremers
67. Personalized Image Aesthetics, Jian Ren, Xiaohui Shen, Zhe Lin, Radomír Měch, David J. Foran
68. Predicting Deeper Into the Future of Semantic Segmentation, Pauline Luc, Natalia Neverova, Camille Couprie, Jakob Verbeek, Yann LeCun
69. Coordinating Filters for Faster Deep Neural Networks, Wei Wen, Cong Xu, Chunpeng Wu, Yandan Wang, Yiran Chen, Hai Li
70. Unsupervised Representation Learning by Sorting Sequences, Hsin-Ying Lee, Jia-Bin Huang, Maneesh Singh, Ming-Hsuan Yang

**Video**

71. A Read-Write Memory Network for Movie Story Understanding, Seil Na, Sangho Lee, Jisung Kim, Gunhee Kim
72. SegFlow: Joint Learning for Video Object Segmentation and Optical Flow, Jingchun Cheng, Yi-Hsuan Tsai, Shengjin Wang, Ming-Hsuan Yang
73. Unsupervised Action Discovery and Localization in Videos, Khurram Soomro, Mubarak Shah
74. Dense-Captioning Events in Videos, Ranjay Krishna, Kenji Hata, Frederic Ren, Li Fei-Fei, Juan Carlos Niebles
75. Learning Long-Term Dependencies for Action Recognition With a Biologically-Inspired Deep Network, Yemin Shi, Yonghong Tian, Yaowei Wang, Wei Zeng, Tiejun Huang
76. Compressive Quantization for Fast Object Instance Search in Videos, Tan Yu, Zhenzhen Wang, Junsong Yuan
77. Complex Event Detection by Identifying Reliable Shots From Untrimmed Videos, Hehe Fan, Xiaojun Chang, De Cheng, Yi Yang, Dong Xu, Alexander G. Hauptmann

**Vision for X**

78. Deep Direct Regression for Multi-Oriented Scene Text Detection, Wenhao He, Xu-Yao Zhang, Fei Yin, Cheng-Lin Liu

1030–1230 **Exhibitors** (Various locations)

**Platinum Level**
- Adobe Research (CAS 11)
- Amazon (CAS 21)
- Baidu (CAS 5)
- CrowdFlower (CAS 3)
- Didi Chuxing (CAS 7)
- Facebook (CAS 130)
- Google (CAS 19)

**Gold Level**
- eBay (CAS 53)
- FiveAI (CAS 30)
- nuTonomy (CAS 31)
- KLA-Tencor (CAS 50)
- Lyft (CIN 134)
- Meitu (CIN 135)
- Microsoft (CIN 127)
- SenseTime (CAS 1)
- Tencent YouTu Lab (CAS 9)
- Uber ATG (CAS 29)
- Vuforia (CAS 32)

**Silver Level**
- Bosch Computer Vision Research (CIN 121)
- Magic Leap (CIN 110)
- Qihoo 360 (CAS 109)
- Samsung Electronics (CIN 107)
- Seetatech (CAS 108)
- Vion (CAS 23)

**Bronze Level**
- Datatang Technology Inc. (CAS 26)
- Frontiers (CAS 25)
- KAUST Visual Computing Center (CIN 122)
- Kina.ai (CAS 123)
- Linkface (CAS 106)
- MathWorks (CIN 105)
- Meta Company (CIN 111)
- Nexar (CIN 113)
- ObjectVideo Labs (CIN 112)
- OSRAM (CIN 132)
- Panasonic (CIN 123)
- SAP (CAS 24)
- Second Spectrum (CIN 118)
- Springer (CAS 18)
- Styria (CAS 52)
- Toshiba Corporation (CIN 125)
- Zalando (CAS 27)

**Start-Up Level**
- 8th Wall (CIN 103)
- AutoX (CAS 28)
- ATG (CIN 119)
- Fashwell (CIN 115)
- Finge Technologies (CIN 120)
- Good AI Lab (CIN 133)
- Imperial Vision Technology (CIN 114)
- iniLabs (CIN 126)
- Meitu (CIN 101)
- NextAI (CIN 119)
- nuTonomy (CAS 31)
- PaperFlower (CIN 5)
- Piaggio Technologies (CIN 129)
- SenseTime (CAS 1)
- Second Spectrum (CIN 118)
- Scape (CIN 120)
- SenseTime (CAS 1)
- SenceTime (CAS 1)
- SenceTime (CAS 1)
- SenseTime (CAS 1)
- SenseTime (CAS 1)

1230–1330 **Lunch**

**1030–1230 Demos** (Salone Adriatico)
**Tuesday, October 24 (Afternoon)**

**1330–1500 Session O2-S2: Recognition I (Sala Grande, Sala Darsena, & Sala Perla)**

Papers in this session are also in Poster Session P2.

**Chairs:** Barbara Caputo (Univ. of Roma La Sapienza)
Kate Saenko (Boston Univ.)

**1330 Orals (O2)**

Format (9 min. for presentation + 3 min. for questions)

1. [1330] Open Set Domain Adaptation, Pau Panareda Busto, Juergen Gall

2. [1342] Deformable Convolutional Networks, Jifeng Dai, Haozhi Qi, Yuwen Xiong, Yi Li, Guodong Zhang, Han Hu, Yichen Wei


4. [1406] FoveaNet: Perspective-Aware Urban Scene Parsing, Xin Li, Zequn Jie, Wei Wang, Changsong Liu, Jimei Yang, Xiaohui Shen, Zhe Lin, Qiang Chen, Shuicheng Yan, Jiashi Feng

5. [1418] Beyond Planar Symmetry: Modeling Human Perception of Reflection and Rotation Symmetries in the Wild, Christopher Funk, Yanxi Liu

**1430 Spotlights (S2)**

Format (3 min. & 30 sec. for presentation; no questions)

1. [1430] Learning to Reason: End-To-End Module Networks for Visual Question Answering, Ronghang Hu, Jacob Andreas, Marcus Rohrbach, Trevor Darrell, Kate Saenko

2. [1433] Hard-Aware Deeply Cascaded Embedding, Yuhui Yuan, Kuiyuan Yang, Chao Zhang

3. [1437] Query-Guided Regression Network With Context Policy for Phrase Grounding, Kan Chen, Rama Kovvuri, Ram Nevatia


5. [1445] Revisiting Unreasonable Effectiveness of Data in Deep Learning Era, Chen Sun, Abhinav Shrivastava, Saurabh Singh, Abhinav Gupta

6. [1448] A Generative Model of People in Clothing, Christoph Lassner, Gerard Pons-Moll, Peter V. Gehler


8. [1456] Improved Image Captioning via Policy Gradient Optimization of SPIDER, Siqi Liu, Zhenhai Zhu, Ning Ye, Sergio Guadarrama, Kevin Murphy

**1500–1700 Poster Session P2 (Sala Laguna, Salone Adriatico, & Sale Mosaici 1–2)**

See poster locations map on pg. 1.

**Oral O2 Posters**

1. Open Set Domain Adaptation, Pau Panareda Busto, Juergen Gall

2. Deformable Convolutional Networks, Jifeng Dai, Haozhi Qi, Yuwen Xiong, Yi Li, Guodong Zhang, Han Hu, Yichen Wei

3. Ensemble Diffusion for Retrieval, Song Bai, Zhichao Zhou, Jingdong Wang, Xiang Bai, Longin Jan Latecki, Qi Tian

4. FoveaNet: Perspective-Aware Urban Scene Parsing, Xin Li, Zequn Jie, Wei Wang, Changsong Liu, Jimei Yang, Xiaohui Shen, Zhe Lin, Qiang Chen, Shuicheng Yan, Jiashi Feng

5. Beyond Planar Symmetry: Modeling Human Perception of Reflection and Rotation Symmetries in the Wild, Christopher Funk, Yanxi Liu

**Spotlight S2 Posters**

6. Learning to Reason: End-To-End Module Networks for Visual Question Answering, Ronghang Hu, Jacob Andreas, Marcus Rohrbach, Trevor Darrell, Kate Saenko

7. Hard-Aware Deeply Cascaded Embedding, Yuhui Yuan, Kuiyuan Yang, Chao Zhang

8. Query-Guided Regression Network With Context Policy for Phrase Grounding, Kan Chen, Rama Kovvuri, Ram Nevatia


10. Revisiting Unreasonable Effectiveness of Data in Deep Learning Era, Chen Sun, Abhinav Shrivastava, Saurabh Singh, Abhinav Gupta

11. A Generative Model of People in Clothing, Christoph Lassner, Gerard Pons-Moll, Peter V. Gehler


13. Improved Image Captioning via Policy Gradient Optimization of SPIDER, Siqi Liu, Zhenhai Zhu, Ning Ye, Sergio Guadarrama, Kevin Murphy

---

13
3D Computer Vision
14. Rolling Shutter Correction in Manhattan World, Pulak Purkait, Christopher Zach, Aleš Leonardis
15. Local-To-Global Point Cloud Registration Using a Dictionary of Viewpoint Descriptors, David Avidar, David Malah, Meir Barzohar
16. 3D-PRNN: Generating Shape Primitives With Recurrent Neural Networks, Chuhang Zou, Ersin Yumer, Jimei Yang, Duygu Ceylan, Derek Hoiem
17. BodyFusion: Real-Time Capture of Human Motion and Surface Geometry Using a Single Depth Camera, Tao Yu, Kaiwen Guo, Feng Xu, Yuan Dong, Zhaqoi Su, Jianhui Zhao, Jianguo Li, Qionghai Dai, Yebin Liu
18. Quasiconvex Plane Sweep for Triangulation With Outliers, Qianggong Zhang, Tat-Jun Chin, David Suter
19. “Maximizing Rigidity“ Revisited: A Convex Programming Approach for Generic 3D Shape Reconstruction From Multiple Perspective Views, Pan Ji, Hongdong Li, Yuchao Dai, Ian Reid
20. Surface Registration via Foliation, Xiaopeng Zheng, Chengfeng Wen, Na Lei, Ming Ma, Xianfeng Gu
22. Corner-Based Geometric Calibration of Multi-Focus Plenoptic Cameras, Sotiris Nousias, François Chadebecq, Jonas Pichat, Pearse Keane, Sébastien Ourselin, Christos Bergeles
23. Focal Track: Depth and Accommodation With Oscillating Lens Deformation, Qi Guo, Emma Alexander, Todd Zickler
24. Reconfiguring the Imaging Pipeline for Computer Vision, Mark Buckler, Suren Jayasuriya, Adrian Sampson
25. Catadioptric HyperSpectral Light Field Imaging, Yujia Xue, Kang Zhu, Qiang Fu, Xilin Chen, Jingyi Yu

Computational Photography
26. Cross-View Asymmetric Metric Learning for Unsupervised Person Re-Identification, Hong-Xing Yu, Ancong Wu, Wei-Shi Zheng
27. Real Time Eye Gaze Tracking With 3D Deformable Eye-Face Model, Kang Wang, Qiang Ji
29. How Far Are We From Solving the 2D & 3D Face Alignment Problem? (And a Dataset of 230, 000 3D Facial Landmarks), Adrian Bulat, Georgios Tzimiropoulos
30. Large Pose 3D Face Reconstruction From a Single Image via Direct Volumetric CNN Regression, Aaron S. Jackson, Adrian Bulat, Vasileios Argyriou, Georgios Tzimiropoulos

Low-Level Vision & Image Processing
31. RankIQA: Learning From Rankings for No-Reference Image Quality Assessment, Xialei Liu, Joost van de Weijer, Andrew D. Bagdanov
32. Look, Perceive and Segment: Finding the Salient Objects in Images via Two-Stream Fixation-Semantic CNNs, Xiaowu Chen, Anlin Zheng, Jia Li, Feng Lu
33. Delving Into Salient Object Subitizing and Detection, Shengfeng He, Jianbo Jiao, Xiaodan Zhang, Guoqiang Han, Rynson W.H. Lau
34. Learning Discriminative Data Fitting Functions for Blind Image Deblurring, Jinshan Pan, Jiangxin Dong, Yu-Wing Tai, Zhixun Su, Ming-Hsuan Yang
35. Video Deblurring via Semantic Segmentation and Pixel-Wise Non-Linear Kernel, Wenqi Ren, Jinshan Pan, XiaoChun Cao, Ming-Hsuan Yang
36. On-Demand Learning for Deep Image Restoration, Ruohan Gao, Kristen Grauman
37. Multi-Channel Weighted Nuclear Norm Minimization for Real Color Image Denoising, Jun Xu, Lei Zhang, David Zhang, Xiangchu Feng
38. Coherent Online Video Style Transfer, Dongdong Chen, Jing Liao, Lu Yuan, Nenghai Yu, Gang Hua

Motion & Tracking
39. SHaPE: A Novel Graph Theoretic Algorithm for Making Consensus-Based Decisions in Person Re-Identification Systems, Arko Barman, Shishir K. Shah
40. Need for Speed: A Benchmark for Higher Frame Rate Object Tracking, Hamed Kiani Galoogahi, Ashton Fagg, Chen Huang, Deva Ramanan, Simon Lucey
41. Learning Background-Aware Correlation Filters for Visual Tracking, Hamed Kiani Galoogahi, Ashton Fagg, Simon Lucey
41. Robust Object Tracking Based on Temporal and Spatial Deep Networks, Zhu Teng, Junliang Xing, Qiang Wang, Congyan Lang, Songhe Feng, Yi Jin
42. Real-Time Hand Tracking Under Occlusion From an Egocentric RGB-D Sensor, Franziska Mueller, Dushyant Mehta, Oleksandr Sotnychenko, Srinath Sridhar, Dan Casas, Christian Theobalt
43. Predicting Human Activities Using Stochastic Grammar, Siyuan Qi, Siyuan Huang, Ping Wei, Song-Chun Zhu
44. ProbFlow: Joint Optical Flow and Uncertainty Estimation, Anne S. Wannenwetsch, Margret Keuper, Stefan Roth

**Optimization Methods**

45. Sublabel-Accurate Discretization of Nonconvex Free-Discontinuity Problems, Thomas Möllenhoff, Daniel Cremers

**Recognition**

46. DeepContext: Context-Encoding Neural Pathways for 3D Holistic Scene Understanding, Yinda Zhang, Mingru Bai, Pushmeet Kohli, Shahram Izadi, Jianxiong Xiao
47. BAM! The Behance Artistic Media Dataset for Recognition Beyond Photography, Michael J. Wilber, Chen Fang, Hailin Jin, Aaron Hertzmann, John Collomosse, Serge Belongie
49. An Empirical Study of Language CNN for Image Captioning, Jiuxiang Gu, Gang Wang, Jianfei Cai, Tsuhan Chen
50. Attributes2Classname: A Discriminative Model for Attribute-Based Unsupervised Zero-Shot Learning, Berkan Demirel, Ramazan Gokberk Cinbis, Nazli Ikizler-Cinbis
51. Areas of Attention for Image Captioning, Marco Pedersoli, Thomas Lucas, Cordelia Schmid, Jakob Verbeek
52. Generative Modeling of Audible Shapes for Object Perception, Zhoutong Zhang, Jiajun Wu, Qijia Li, Zhengjia Huang, James Traer, Josh H. McDermott, Joshua B. Tenenbaum, William T. Freeman
53. Scene Graph Generation From Objects, Phrases and Region Captions, Yikang Li, Wanli Ouyang, Bolei Zhou, Kun Wang, Xiaogang Wang
54. Recurrent Multimodal Interaction for Referring Image Segmentation, Chenxi Liu, Zhe Lin, Xiaohui Shen, Jiming Yang, Xin Lu, Alan Yuille
55. Learning Feature Pyramids for Human Pose Estimation, Wei Yang, Shuang Li, Wanli Ouyang, Hongsheng Li, Xiaogang Wang
56. Structured Attentions for Visual Question Answering, Chen Zhu, Yanpeng Zhao, Shuaiyi Huang, Kewei Tu, Yi Ma

**Segmentation, Grouping & Shape**


**Statistical Methods & Learning**

59. Encoder Based Lifelong Learning, Amal Rannen, Rahaf Aljundi, Matthew B. Blaschko, Tinne Tuytelaars
60. Transitive Invariance for Self-Supervised Visual Representation Learning, Xiaolong Wang, Kaiming He, Abhinav Gupta
61. Weakly Supervised Learning of Deep Metrics for Stereo Reconstruction, Stepan Tulyakov, Anton Ivanov, François Fleuret
63. SORT: Second-Order Response Transform for Visual Recognition, Yan Wang, Lingxi Xie, Chenxi Liu, Siyuan Qiao, Ya Zhang, Wenzun Zhang, Qi Tian, Alan Yuille
64. Adversarial Examples for Semantic Segmentation and Object Detection, Cihang Xie, Jianyu Wang, Zhishuai Zhang, Yuyin Zhou, Lingxi Xie, Alan Yuille
65. Genetic CNN, Lingxi Xie, Alan Yuille
66. Channel Pruning for Accelerating Very Deep Neural Networks, Yihui He, Xiangyu Zhang, Jian Sun
67. Infinite Latent Feature Selection: A Probabilistic Latent Graph-Based Ranking Approach, Giorgio Roffo, Simone Melzi, Umberto Castellani, Alessandro Vinciarelli

**Video**

68. Video Fill in the Blank Using LR/RL LSTMs With Spatial-Temporal Attentions, Amir Mazaheri, Dong Zhang, Mubarak Shah
69. Primary Video Object Segmentation via Complementary CNNs and Neighborhood Reversible Flow, Jia Li, Anlin Zheng, Xiaowu Chen, Bin Zhou
71. Attentive Semantic Video Generation Using Captions, 
   Tanya Marwah, Gaurav Mittal, Vineeth N. Balasubramanian
72. Following Gaze in Video, Adrià Recasens, Carl Vondrick, 
   Aditya Khosla, Antonio Torralba
73. Adaptive RNN Tree for Large-Scale Human Action 
   Recognition, Wenbo Li, Longyin Wen, Ming-Ching Chang, Ser Nam Lim, Siwei Lyu
74. Spatio-Temporal Person Retrieval via Natural Language 
   Queries, Masataka Yamaguchi, Kuniaki Saito, Yoshitaka Ushiki, Tatsuya Harada

Vision for X
75. Automatic Spatially-Aware Fashion Concept Discovery, 
   Xintong Han, Zuxuan Wu, Phoenix X. Huang, Xiao Zhang, 
   Menglong Zhu, Yuan Li, Yang Zhao, Larry S. Davis
76. ChromaTag: A Colored Marker and Fast Detection 
   Algorithm, Joseph DeGol, Timothy Bretl, Derek Hoiem
77. Adversarial Image Perturbation for Privacy Protection — A 
   Game Theory Perspective, Seong Joon Oh, Mario Fritz, Bernt Schiele
78. WeText: Scene Text Detection Under Weak Supervision, 
   Shangxuan Tian, Shijian Lu, Chongshou Li

1500–1700 Demos (Salone Adriatico)

1500–1700 Exhibitors (Various locations)
- See exhibitor list on page 12.

1545–1630 Afternoon Break

1700–1800 Oral Session O3: Vision for X (Sala Grande, Sala Darsena, & Sala Perla)

Papers in this session are also in Poster Session P3.

   Chairs: Elisa Ricci (Univ. of Perugia) 
          Stan Sclaroff (Boston Univ.)

   Format (9 min. for presentation + 3 min. for questions)
1. [1700] Arbitrary Style Transfer in Real-Time With Adaptive 
   Instance Normalization, Xun Huang, Serge Belongie
2. [1712] Photographic Image Synthesis With Cascaded 
   Refinement Networks, Qifeng Chen, Vladlen Koltun

1800–1900 PAMI Technical Committee Meeting 
   (Sala Darsena)

1900–2000 Reception (Palazzo del Cinema & 
   Palazzo del Casinò)

Notes:
Wednesday, October 25

0800–1400 Registration (Palazzo del Casinò Terrace)

0830–1030 Poster Session P3 (Sala Laguna, Salone Adriatico, & Sale Mosaici 1–2)

See poster locations map on pg. 1.

Oral O3 Posters
1. Arbitrary Style Transfer in Real-Time With Adaptive Instance Normalization, Xun Huang, Serge Belongie
2. Photographic Image Synthesis With Cascaded Refinement Networks, Qifeng Chen, Vladlen Koltun
3. SSD-6D: Making RGB-Based 3D Detection and 6D Pose Estimation Great Again, Wadim Kehl, Fabian Manhardt, Federico Tombari, Slobodan Ilic, Nassir Navab
4. Unsupervised Creation of Parameterized Avatars, Lior Wolf, Yaniv Taigman, Adam Polyak
5. Learning for Active 3D Mapping, Karel Zimmermann, Tomáš Petříček, Vojtěch Šalanský, Tomáš Svoboda

3D Computer Vision
7. Surface Normals in the Wild, Weifeng Chen, Donglai Xiang, Jia Deng
8. Unsupervised Learning of Stereo Matching, Chao Zhou, Hong Zhang, Xiaoyong Shen, Jiaya Jia
9. Unrestricted Facial Geometry Reconstruction Using Image-To-Image Translation, Matan Sela, Elad Richardson, Ron Kimmel
10. Learned Multi-Patch Similarity, Wilfried Hartmann, Silvano Galliani, Michal Havlena, Luc Van Gool, Konrad Schindler
11. Click Here: Human-Localized Keypoints as Guidance for Viewpoint Estimation, Ryan Szeto, Jason J. Corso
12. Unsupervised Adaptation for Deep Stereo, Alessio Tonioni, Matteo Poggi, Stefano Mattoccia, Luigi Di Stefano

Computational Photography
13. Composite Focus Measure for High Quality Depth Maps, Parikshit Sakurikar, P. J. Narayanan

Face & Gesture
14. Reconstruction-Based Disentanglement for Pose-Invariant Face Recognition, Xi Peng, Xiang Yu, Kihyuk Sohn, Dimitris N. Metaxas, Mannmohan Chandraker
15. Recurrent 3D-2D Dual Learning for Large-Pose Facial Landmark Detection, Shengtao Xiao, Jiashi Feng, Luoqi Liu, Xuecheng Nie, Wei Wang, Shuicheng Yan, Ashraf Kassim
16. Anchored Regression Networks Applied to Age Estimation and Super Resolution, Eirikur Agustsson, Radu Timofte, Luc Van Gool
17. Infant Footprint Recognition, Eryun Liu

Low-Level Vision & Image Processing
18. Self-Paced Kernel Estimation for Robust Blind Image Deblurring, Dong Gong, Mingkui Tan, Yanning Zhang, Anton van den Hengel, Qinfeng Shi
19. Super-Trajectory for Video Segmentation, Wenguan Wang, Jianbing Shen, Jianwen Xie, Fatih Porikli
20. Be Your Own Prada: Fashion Synthesis With Structural Coherence, Shizhan Zhu, Raquel Urtasun, Sanja Fidler, Dahua Lin, Chen Change Loy
21. Wavelet-SRNnet: A Wavelet-Based CNN for Multi-Scale Face Super Resolution, Huaibo Huang, Ran He, Zhenan Sun, Tieniu Tan
22. Learning Gaze Transitions From Depth to Improve Video Salience Estimation, George Leifman, Dmitry Rudoy, Tristan Swedish, Eduardo Bayro-Corrochano, Ramesh Raskar
24. Modelling the Scene Dependent Imaging in Cameras With a Deep Neural Network, Seonghyeon Nam, Seon Joo Kim
25. Transformed Low-Rank Model for Line Pattern Noise Removal, Yi Chang, Luxin Yan, Sheng Zhong
26. Weakly Supervised Manifold Learning for Dense Semantic Object Correspondence, Utkarsh Gaur, B. S. Manjunath
27. PanNet: A Deep Network Architecture for Pan-Sharpening, Junfeng Yang, Xueyang Fu, Yuwen Hu, Yue Huang, Xinghao Ding, John Paisley

Motion & Tracking
28. Dual Motion GAN for Future-Flow Embedded Video Prediction, Xiaodan Liang, Lisa Lee, Wei Dai, Eric P. Xing
29. Online Robust Image Alignment via Subspace Learning From Gradient Orientations, Qingqing Zheng, Yi Wang, Pheng-Ann Heng
30. Learning Dynamic Siamese Network for Visual Object Tracking, Qing Guo, Wei Feng, Ce Zhou, Rui Huang, Liang Wan, Song Wang

Optimization Methods
31. High Order Tensor Formulation for Convolutional Sparse Coding, Adel Bibi, Bernard Ghanem
32. Learning Proximal Operators: Using Denoising Networks for Regularizing Inverse Imaging Problems, Tim Meinhardt, Michael Möller, Caner Hazirbas, Daniel Cremers

Recognition
33. ScaleNet: Guiding Object Proposal Generation in Supermarkets and Beyond, Siyuan Qiao, Wei Shen, Weichao Qiu, Chenxi Liu, Alan Yuille
34. Temporal Dynamic Graph LSTM for Action-Driven Video Object Detection, Yuan Yuan, Xiaodan Liang, Xiaolong Wang, Dit-Yan Yeung, Abhinav Gupta
35. VQS: Linking Segmentations to Questions and Answers for Supervised Attention in VQA and Question-Focused Semantic Segmentation, Chuang Gan, Yandong Li, Haoxiang Li, Chen Sun, Boqing Gong
37. SCNet: Learning Semantic Correspondence, Kai Han, Rafael S. Rezende, Bumsub Ham, Kwan-Yee K. Wong, Minsu Cho, Cordelia Schmid, Jean Ponce
38. Soft Proposal Networks for Weakly Supervised Object Localization, Yi Zhu, Yanzhao Zhou, Qixiang Ye, Qiang Qiu, Jianbin Jiao
39. Class Rectification Hard Mining for Imbalanced Deep Learning, Qi Dong, Shaogang Gong, Xiatian Zhu
40. Generating High-Quality Crowd Density Maps Using Contextual Pyramid CNNs, Vishwanath A. Sindagi, Vishal M. Patel
41. See the Glass Half Full: Reasoning About Liquid Containers, Their Volume and Content, Roozbeh Mottaghi, Connor Schenck, Dieter Fox, Ali Farhadi
42. Hierarchical Multimodal LSTM for Dense Visual-Semantic Embedding, Zhenxing Niu, Mo Zhou, Le Wang, Xinbo Gao, Gang Hua
43. Identity-Aware Textual-Visual Matching With Latent Co-Attention, Shuang Li, Tong Xiao, Hongsheng Li, Wei Yang, Xiaogang Wang
44. Learning Deep Neural Networks for Vehicle Re-ID With Visual-Spatio-Temporal Path Proposals, Yantao Shen, Tong Xiao, Hongsheng Li, Shuai Yi, Xiaogang Wang
45. Learning From Noisy Labels With Distillation, Yuncheng Li, Jianchao Yang, Yafe Song, Liangliang Cao, Jiebo Luo, Li-Jia Li
46. DSOD: Learning Deeply Supervised Object Detectors From Scratch, Zhiqiang Shen, Zhuang Liu, Jianguo Li, Yu-Gang Jiang, Yurong Chen, Xiangyang Xue
47. Phrase Localization and Visual Relationship Detection With Comprehensive Image-Language Cues, Bryan A. Plummer, Arun Mallya, Christopher M. Cervantes, Julia Hockenmaier, Svetlana Lazebnik
49. VPGNet: Vanishing Point Guided Network for Lane and Road Marking Detection and Recognition, Seokju Lee, Junsik Kim, Jae Shin Yoon, Seunghak Shin, Oleksandr Bailo, Namil Kim, Tae-Hee Lee, Hyun Seok Hong, Seung-Hoon Han, In So Kweon
50. Unsupervised Learning of Important Objects From First-Person Videos, Gedas Bertasius, Hyun Soo Park, Stella X. Yu, Jianbo Shi
51. An Analysis of Visual Question Answering Algorithms, Kushal Kafle, Christopher Kanan
52. Visual Relationship Detection With Internal and External Linguistic Knowledge Distillation, Ruichi Yu, Ang Li, Vlad I. Morariu, Larry S. Davis
53. A Two Stream Siamese Convolutional Neural Network for Person Re-Identification, Dahjung Chung, Khalid Tahboub, Edward J. Delp
54. Joint Learning of Object and Action Detectors, Vicky Kalogeiton, Philippe Weinzaepfel, Vittorio Ferrari, Cordelia Schmid

Segmentation, Grouping & Shape
55. No More Discrimination: Cross City Adaptation of Road Scene Segmenters, Yi-Hsin Chen, Wei-Yu Chen, Yu-Ting Chen, Bo-Cheng Tsai, Yu-Chiang Frank Wang, Min Sun
56. Open Vocabulary Scene Parsing, Hang Zhao, Xavier Puig, Bolei Zhou, Sanja Fidler, Antonio Torralba
Abhinav Gupta

Video

Statistical Methods & Learning

Wednesday, October 25 (Morning)

0830–1030 Demos (Salone Adriatico)

0830–1030 Exhibitors (Various locations)

0945–1030 Morning Break

1030–1100 Spotlight Session S3: Vision for X & Computational Phytography (Sala Grande, Sala Darsena, & Sala Perla)

Papers in this session are also in Poster Session P4.

Format (3 min. & 30 sec. for presentation; no questions)

1. [1030] Playing for Benchmarks, Stephan R. Richter, Zeeshan Hayder, Vladlen Koltun
3. [1037] GANs for Biological Image Synthesis, Anton Osokin, Anatole Chessel, Rafael E. Carazo Salas, Federico Vaggi
<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:15</td>
<td>ICCV 2017 At-a-Glance (Main Conference)</td>
<td>Palazzo del Casino Terrace</td>
</tr>
<tr>
<td>08:30</td>
<td>Session O1-P1 (Orals &amp; Spotlights): 3D Vision &amp; Video Analysis</td>
<td>Sala Grande, Sala Darsena, &amp; Sala Perla</td>
</tr>
<tr>
<td>09:15</td>
<td>Poster Session P1 (Sala Laguna, Salone Adriatico, &amp; Sala Mosaic 1–2)</td>
<td>Sala Grande, Sala Darsena, &amp; Sala Perla</td>
</tr>
<tr>
<td>10:00</td>
<td>Lunch (On your own)</td>
<td></td>
</tr>
<tr>
<td>10:15</td>
<td>Session O2-P2 (Orals &amp; Spotlights): Recognition I</td>
<td>Sala Grande, Sala Darsena, &amp; Sala Perla</td>
</tr>
<tr>
<td>11:00</td>
<td>Poster Session P2 (Sala Laguna, Salone Adriatico, &amp; Sala Mosaic 1–2)</td>
<td>Sala Grande, Sala Darsena, &amp; Sala Perla</td>
</tr>
<tr>
<td>11:45</td>
<td>Lunch (On your own)</td>
<td></td>
</tr>
<tr>
<td>12:00</td>
<td>Session O4-P4 (Orals &amp; Spotlights): Recognition 2</td>
<td>Sala Grande, Sala Darsena, &amp; Sala Perla</td>
</tr>
<tr>
<td>12:45</td>
<td>Poster Session P4 (Sala Laguna, Salone Adriatico, &amp; Sala Mosaic 1–2)</td>
<td>Sala Grande, Sala Darsena, &amp; Sala Perla</td>
</tr>
<tr>
<td>13:30</td>
<td>Lunch (On your own)</td>
<td></td>
</tr>
<tr>
<td>13:45</td>
<td>Session O6-P6 (Orals &amp; Spotlights): Face and Human Behaviour Analysis</td>
<td>Sala Grande, Sala Darsena, &amp; Sala Perla</td>
</tr>
<tr>
<td>14:30</td>
<td>Poster Session P6 (Sala Laguna, Salone Adriatico, &amp; Sala Mosaic 1–2)</td>
<td>Sala Grande, Sala Darsena, &amp; Sala Perla</td>
</tr>
<tr>
<td>15:15</td>
<td>Lunch (On your own)</td>
<td></td>
</tr>
<tr>
<td>15:30</td>
<td>Session O7-P7 (Orals &amp; Spotlights): Low-Level Vision</td>
<td>Sala Grande, Sala Darsena, &amp; Sala Perla</td>
</tr>
<tr>
<td>16:15</td>
<td>Poster Session P7 (Sala Laguna, Salone Adriatico, &amp; Sala Mosaic 1–2)</td>
<td>Sala Grande, Sala Darsena, &amp; Sala Perla</td>
</tr>
<tr>
<td>17:00</td>
<td>Lunch (On your own)</td>
<td></td>
</tr>
<tr>
<td>17:15</td>
<td>Session O9-P9 (Orals &amp; Spotlights): Machine Learning</td>
<td>Sala Grande, Sala Darsena, &amp; Sala Perla</td>
</tr>
</tbody>
</table>

**Reception Desk (Palazzo del Casino Terrace)**

**Coffee Break**

**Demos (Salone Adriatico)**

**Exhibits (Various locations): pg. 12**
4. [1041] Learning to Synthesize a 4D RGBD Light Field From a Single Image, Pratul P. Srinivasan, Tongzhou Wang, Ashwin Sreelal, Ravi Ramamoorthi, Ren Ng
5. [1045] Neural EPI-Volume Networks for Shape From Light Field, Stefan Heber, Wei Yu, Thomas Pock
6. [1048] Material Editing Using a Physically Based Rendering Network, Guilin Liu, Duygu Ceylan, Ersin Yumer, Jimei Yang, Jyh-Ming Lien
8. [1056] Linear Differential Constraints for Photo-Polarimetric Height Estimation, Silvia Tozza, William A. P. Smith, Dizhong Zhu, Ravi Ramamoorthi, Edwin R. Hancock

1100–1300 Poster Session P4 (Sala Laguna, Salone Adriatico, & Sale Mosaici 1–2)

See poster locations map on pg. 1.

Spotlight S3 Posters
1. Playing for Benchmarks, Stephan R. Richter, Zeeshan Hayder, Vladlen Koltun
3. GANs for Biological Image Synthesis, Anton Osokin, Anatole Chessel, Rafael E. Carazo Salas, Federico Vaggi
4. Learning to Synthesize a 4D RGBD Light Field From a Single Image, Pratul P. Srinivasan, Tongzhou Wang, Ashwin Sreelal, Ravi Ramamoorthi, Ren Ng
5. Neural EPI-Volume Networks for Shape From Light Field, Stefan Heber, Wei Yu, Thomas Pock
6. Material Editing Using a Physically Based Rendering Network, Guilin Liu, Duygu Ceylan, Ersin Yumer, Jimei Yang, Jyh-Ming Lien
8. Linear Differential Constraints for Photo-Polarimetric Height Estimation, Silvia Tozza, William A. P. Smith, Dizhong Zhu, Ravi Ramamoorthi, Edwin R. Hancock

3D Computer Vision
9. Polynomial Solvers for Saturated Ideals, Viktor Larsson, Kalle Åström, Magnus Oskarsson
10. Shape Inpainting Using 3D Generative Adversarial Network and Recurrent Convolutional Networks, Weiyue Wang, Qiangui Huang, Suya You, Chao Yang, Ulrich Neumann
11. SurfaceNet: An End-To-End 3D Neural Network for Multiview Stereopsis, Mengqi Ji, Juergen Gall, Haitian Zheng, Yebin Liu, Lu Fang
12. Making Minimal Solvers for Absolute Pose Estimation Compact and Robust, Viktor Larsson, Zuzana Kukelova, Yinqiang Zheng
13. 3D Surface Detail Enhancement From a Single Normal Map, Wuyuan Xie, Miaohui Wang, Xianbiao Qi, Lei Zhang
14. RMPE: Regional Multi-Person Pose Estimation, Hao-Shu Fang, Shiqin Xie, Yu-Wing Tai, Cewu Lu
15. Online Video Object Detection Using Association LSTM, Yongyi Lu, Cewu Lu, Chi-Keung Tang
16. PolyFit: Polygonal Surface Reconstruction From Point Clouds, Liangliang Nan, Peter Wonka
17. Progressive Large Scale-Invariant Image Matching in Scale Space, Lei Zhou, Siyu Zhu, Tianwei Shen, Jinglu Wang, Tian Fang, Long Quan
18. Efficient Global 2D-3D Matching for Camera Localization in a Large-Scale 3D Map, Liu Liu, Hongdong Li, Yuchao Dai

Biomedical Image Analysis
20. Multi-Stage Multi-Recursive-Input Fully Convolutional Networks for Neuronal Boundary Detection, Wei Shen, Bin Wang, Yuan Jiang, Yan Wang, Alan Yuille

Computational Photography
21. Depth and Image Restoration From Light Field in a Scattering Medium, Jiadong Tian, Zachary Murez, Tong Cui, Zhen Zhang, David Kriegman, Ravi Ramamoorthi
22. Video Reflection Removal Through Spatio-Temporal Optimization, Ajay Nandoriya, Mohamed Elgharib, Changil Kim, Mohamed Hefeeda, Wojciech Matusik

Face & Gesture
23. Efficient Online Local Metric Adaptation via Negative Samples for Person Re-Identification, Jiahuan Zhou, Pei Yu, Wei Tang, Ying Wu
24. Stepwise Metric Promotion for Unsupervised Video Person Re-Identification, Zimo Liu, Dong Wang, Huichuan Lu
25. Beyond Face Rotation: Global and Local Perception GAN for Photorealistic and Identity Preserving Frontal View Synthesis, Rui Huang, Shu Zhang, Tianyu Li, Ran He
26. Group Re-Identification via Unsupervised Transfer of Sparse Features Encoding, Giuseppe Lisanti, Niki Martinel, Alberto Del Bimbo, Gian Luca Foresti
27. Visual Transformation Aided Contrastive Learning for Video-Based Kinship Verification, Hamdi Dibeklioğlu

**Low-Level Vision & Image Processing**
28. Decoder Network Over Lightweight Reconstructed Feature for Fast Semantic Style Transfer, Ming Lu, Hao Zhao, Anbang Yao, Feng Xu, Yurong Chen, Li Zhang
29. Blind Image Deblurring With Outlier Handling, Jiangxin Dong, Jinshan Pan, Zhiyun Su, Ming-Hsuan Yang
31. Fast Image Processing With Fully-Convolutional Networks, Qifeng Chen, Jia Xu, Vladlen Koltun
32. Robust Video Super-Resolution With Learned Temporal Dynamics, Ding Liu, Zhaowen Wang, Yuchen Fan, Xiaoming Liu, Zhangyang Wang, Shiyu Chang, Thomas Huang
33. Should We Encode Rain Streaks in Video as Deterministic or Stochastic?, Wei Wei, Lixuan Yi, Qi Xie, Qian Zhao, Deyu Meng, Zongben Xu

**Motion & Tracking**
35. Low-Dimensionality Calibration Through Local Anisotropic Scaling for Robust Hand Model Personalization, Edoardo Remelli, Anastasia Tkach, Andrea Tagliasacchi, Mark Pauly
36. Non-Markovian Globally Consistent Multi-Object Tracking, Andríi Maksai, Xinchao Wang, François Fleuret, Pascal Fua
37. CREST: Convolutional Residual Learning for Visual Tracking, Yibing Song, Chao Ma, Lijun Gong, Jiawei Zhang, Rynson W. H. Lau, Ming-Hsuan Yang
38. Volumetric Flow Estimation for Incompressible Fluids Using the Stationary Stokes Equations, Katrin Lasinger, Christoph Vogel, Konrad Schindler

**Optimization Methods**
40. Performance Guaranteed Network Acceleration via High-Order Residual Quantization, Zefan Li, Bingbing Ni, Wenjun Zhang, Xiaokang Yang, Wen Gao

**Recognition**
41. Deep Metric Learning With Angular Loss, Jian Wang, Feng Zhou, Shilei Wen, Xiao Liu, Yuanqing Lin
42. Compositional Human Pose Regression, Xiao Sun, Jiaxiang Shang, Shuang Liang, Yichen Wei
43. MUTAN: Multimodal Tucker Fusion for Visual Question Answering, Hedi Ben-younes, Remi Cadene, Matthieu Cord, Nicolas Thome
44. Revisiting IM2GPS in the Deep Learning Era, Nam Vo, Nathan Jacobs, James Hays
45. Scene Parsing With Global Context Embedding, Wei-Chih Hung, Yi-Hsuan Tsai, Xiaohui Shen, Zhe Lin, Kalyan Sunkavalli, Xin Lu, Ming-Hsuan Yang
46. A Simple yet Effective Baseline for 3D Human Pose Estimation, Julieta Martinez, Rayat Hossain, Javier Romero, James J. Little
47. Dual-Glance Model for Deciphering Social Relationships, Junnan Li, Yongkang Wong, Qi Zhao, Mohan S. Kankanhalli
48. Sketching With Style: Visual Search With Sketches and Aesthetic Context, John Collomosse, Tu Bui, Michael J. Wilber, Chen Fang, HaiLin Jin
49. Point Set Registration With Global-Local Correspondence and Transformation Estimation, Su Zhang, Yang Yang, Kun Yang, Yi Luo, Sim-Heng Ong

**Segmentation, Grouping & Shape**
52. Directionally Convolutional Networks for 3D Shape Segmentation, Haotian Xu, Ming Dong, Zichun Zhong
53. AMAT: Medial Axis Transform for Natural Images, Stavros Tsogkas, Sven Dickinson
55. Regional Interactive Image Segmentation Networks, Jun Hao Liew, Yunchao Wei, Wei Xiong, Sim-Heng Ong, Jiashi Feng

Statistical Methods & Learning
56. Learning Efficient Convolutional Networks Through Network Slimming, Zhuang Liu, Jianguo Li, Zhiqiang Shen, Gao Huang, Shoumeng Yan, Changshui Zhang
57. CVAE-GAN: Fine-Grained Image Generation Through Asymmetric Training, Jianmin Bao, Dong Chen, Fang Wen, Houqiang Li, Gang Hua
58. Universal Adversarial Perturbations Against Semantic Image Segmentation, Jan Hendrik Metzen, Mummadi Chaithanya Kumar, Thomas Brox, Volker Fischer
59. Associative Domain Adaptation, Philip Haeusser, Thomas Frerix, Alexander Mordvintsev, Daniel Cremers
60. Introspective Neural Networks for Generative Modeling, Justin Lazarow, Long Jin, Zhuowen Tu
61. Towards a Unified Compositional Model for Visual Pattern Modeling, Wei Tang, Pei Yu, Jiahuan Zhou, Ying Wu
62. Least Squares Generative Adversarial Networks, Xudong Mao, Qing Li, Haoran Xie, Raymond Y.K. Lau, Zhen Wang, Stephen Paul Smolley
63. Centered Weight Normalization in Accelerating Training of Deep Neural Networks, Lei Huang, Xianglong Liu, Yang Liu, Bo Lang, Dacheng Tao
64. Deep Growing Learning, Guangcong Wang, Xiaohua Xie, Jianhuang Lai, Jiaxuan Zhuo
65. Smart Mining for Deep Metric Learning, Ben Harwood, Vijay Kumar B G, Gustavo Carneiro, Ian Reid, Tom Drummond
66. Temporal Generative Adversarial Nets With Singular Value Clipping, Masaki Saito, Eiichi Matsumoto, Shunta Saito
67.Sampling Matters in Deep Embedding Learning, Chao-Yuan Wu, R. Manmatha, Alexander J. Smola, Philipp Krähenbühl
68. DualGAN: Unsupervised Dual Learning for Image-To-Image Translation, Zili Yi, Hao Zhang, Ping Tan, Minglun Gong

Video
69. Learning View-Invariant Features for Person Identification in Temporally Synchronized Videos Taken by Wearable Cameras, Kang Zheng, Xiaochuan Fan, Yuewei Lin, Hao Guo, Hongkai Yu, Dazhou Guo, Song Wang
70. MarioQA: Answering Questions by Watching Gameplay Videos, Jonghwan Mun, Paul Hongsuck Seo, Ilchae Jung, Bohyung Han
71. SBGAR: Semantics Based Group Activity Recognition, Xin Li, Mooi Choo Chuah
72. Trespassing the Boundaries: Labeling Temporal Bounds for Object Interactions in Egocentric Video, Davide Moltisanti, Michael Wray, Walterio Mayol-Cuevas, Dima Damen
73. Unmasking the Abnormal Events in Video, Radu Tudor Ionescu, Sorina Smeureanu, Bogdan Alexe, Marius Popescu
74. Chained Multi-Stream Networks Exploiting Pose, Motion, and Appearance for Action Classification and Detection, Mohammadreza Zolfaghari, Gabriel L. Oliveira, Nima Sedaghat, Thomas Brox
75. Temporal Action Detection With Structured Segment Networks, Yue Zhao, Yuanjun Xiong, Limin Wang, Zhirong Wu, Xiaou Tang, Dahua Lin
76. Jointly Recognizing Object Fluents and Tasks in Egocentric Videos, Yang Liu, Ping Wei, Song-Chun Zhu
77. Transferring Objects: Joint Inference of Container and Human Pose, Hanqing Wang, Wei Liang, Lap-Fai Yu

Vision for X
78. Interpretable Learning for Self-Driving Cars by Visualizing Causal Attention, Jinkyu Kim, John Canny

1100–1300 Demos (Salone Adriatico)

1100–1300 Exhibitors (Various locations)
- See exhibitor list on page 12.

1300 Lunch

Nothing Scheduled for the Afternoon
Thursday, October 26

0800–1800 Registration (Palazzo del Casinò Terrace)

0900–1030 Session O4-S4: Recognition 2 (Sala Grande, Sala Darsena, & Sala Perla)

Papers in this session are also in Poster Session P5.

Chairs: Bernt Schiele (MPI Informatics)
        Antonio Torralba (MIT)

O900 Orals (O4)

Format (9 min. for presentation + 3 min. for questions)

1. [0900] Learning Cooperative Visual Dialog Agents With Deep Reinforcement Learning, Abhishek Das, Satwik Kottur, José M. F. Moura, Stefan Lee, Dhruv Batra
2. [0912] Mask R-CNN, Kaiming He, Georgia Gkioxari, Piotr Dollár, Ross Girshick
3. [0924] Towards Diverse and Natural Image Descriptions via a Conditional GAN, Bo Dai, Sanja Fidler, Raquel Urtasun, Dahua Lin
4. [0936] Focal Loss for Dense Object Detection, Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, Piotr Dollár
5. [0948] Inferring and Executing Programs for Visual Reasoning, Justin Johnson, Bharath Hariharan, Laurens van der Maaten, Judy Hoffman, Li Fei-Fei, C. Lawrence Zitnick, Ross Girshick

1000 Spotlights (S4)

Format (3 min. & 30 sec. for presentation; no questions)

1. [1000] Visual Forecasting by Imitating Dynamics in Natural Sequences, Kuo-Hao Zeng, William B. Shen, De-An Huang, Min Sun, Juan Carlos Niebles
2. [1003] TorontoCity: Seeing the World With a Million Eyes, Shenlong Wang, Min Bai, Gellért Máttyus, Hang Chu, Wenjie Luo, Bin Yang, Justin Liang, Joel Cheverie, Sanja Fidler, Raquel Urtasun
3. [1007] Low-Shot Visual Recognition by Shrinking and Hallucinating Features, Bharath Hariharan, Ross Girshick
4. [1011] A Coarse-Fine Network for Keypoint Localization, Shaoli Huang, Mingming Gong, Dacheng Tao
5. [1015] Detect to Track and Track to Detect, Christoph Feichtenhofer, Axel Pinz, Andrew Zisserman

1030–1115 Morning Break

1030–1230 Poster Session P5 (Sala Laguna, Salone Adriatico, & Sale Mosaici 1–2)

See poster locations map on pg. 1.

Oral O4 Posters

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

Spotlight S4 Posters

6. Visual Forecasting by Imitating Dynamics in Natural Sequences, Kuo-Hao Zeng, William B. Shen, De-An Huang, Min Sun, Juan Carlos Niebles
7. TorontoCity: Seeing the World With a Million Eyes, Shenlong Wang, Min Bai, Gellért Máttyus, Hang Chu, Wenjie Luo, Bin Yang, Justin Liang, Joel Cheverie, Sanja Fidler, Raquel Urtasun
8. Low-Shot Visual Recognition by Shrinking and Hallucinating Features, Bharath Hariharan, Ross Girshick
9. A Coarse-Fine Network for Keypoint Localization, Shaoli Huang, Mingming Gong, Dacheng Tao
10. Detect to Track and Track to Detect, Christoph Feichtenhofer, Axel Pinz, Andrew Zisserman
Thursday, October 26 (Morning)

11. Single Shot Text Detector With Regional Attention, Pan He, Weilin Huang, Tong He, Qile Zhu, Yu Qiao, Xiaolin Li
12. SubUNets: End-To-End Hand Shape and Continuous Sign Language Recognition, Necati Cihan Camgoz, Simon Hadfield, Oscar Koller, Richard Bowden
13. A Spatiotemporal Oriented Energy Network for Dynamic Texture Recognition, Isma Hadji, Richard P. Wildes

3D Computer Vision
14. Probabilistic Structure From Motion With Objects (PSfMO), Paul Gay, Cosimo Rubino, Vaibhav Bansal, Alessio Del Bue
16. Multi-View Dynamic Shape Refinement Using Local Temporal Integration, Vincent Leroy, Jean-Sebastien Franco, Edmond Boyer
17. Learning Hand Articulations by Hallucinating Heat Distribution, Chihoo Choi, Sangpil Kim, Karthik Ramani
18. Intrinsic3D: High-Quality 3D Reconstruction by Joint Appearance and Geometry Optimization With Spatially-Varying Lighting, Robert Maier, Kihwan Kim, Daniel Cremers, Jan Kautz, Matthias Nießner
19. Robust Hand Pose Estimation During the Interaction With an Unknown Object, Chihoo Choi, Sang Ho Yoon, Chin-Ning Chen, Karthik Ramani

Computational Photography
22. Filter Selection for Hyperspectral Estimation, Boaz Arad, Ohad Ben-Shahar

Face & Gesture
24. Detecting Faces Using Inside Cascaded Contextual CNN, Kaipeng Zhang, Zhanpeng Zhang, Hao Wang, Zhifeng Li, Yu Qiao, Wei Liu
25. A Novel Space-Time Representation on the Positive Semidefinite Cone for Facial Expression Recognition, Anis Kacem, Mohamed Daoudi, Boulbaba Ben Amor, Juan Carlos Alvarez-Paiva
27. Pose-Invariant Face Alignment With a Single CNN, Amin Jourabloo, Mao Ye, Xiaoming Liu, Liu Ren
28. Unsupervised Domain Adaptation for Face Recognition in Unlabeled Videos, Kihyuk Sohn, Sifei Liu, Guangyu Zhong, Xiang Yu, Ming-Hsuan Yang, Manmohan Chandraker
29. Deeply-Learned Part-Aligned Representations for Person Re-Identification, Liming Zhao, Xi Li, Yueting Zhuang, Jingdong Wang

Low-Level Vision & Image Processing
32. Revisiting Cross-Channel Information Transfer for Chromatic Aberration Correction, Tiancheng Sun, Yifan Peng, Wolfgang Heidrich
33. High-Quality Correspondence and Segmentation Estimation for Dual-Lens Smart-Phone Portraits, Xiaoyong Shen, Hongyun Gao, Xin Tao, Chao Zhou, Jiaya Jia
34. Learning Visual Attention to Identify People With Autism Spectrum Disorder, Ming Jiang, Qi Zhao
35. DSLR-Quality Photos on Mobile Devices With Deep Convolutional Networks, Andrey Ignatov, Nikolay Kobyshev, Radu Timofte, Kenneth Vanhoey, Luc Van Gool
36. Non-Uniform Blind Deblurring by Reblurring, Yuval Bahat, Netalee Efrat, Michal Irani
37. Misalignment-Robust Joint Filter for Cross-Modal Image Pairs, Takashi Shibata, Masayuki Tanaka, Masatoshi Okutomi
38. Low-Rank Tensor Completion: A Pseudo-Bayesian Learning Approach, Wei Chen, Nan Song
Motion & Tracking
41. The Pose Knows: Video Forecasting by Generating Pose Futures, Jacob Walker, Kenneth Marino, Abhinav Gupta, Martial Hebert
42. What Will Happen Next? Forecasting Player Moves in Sports Videos, Panna Felsen, Pulkit Agrawal, Jitendra Malik

Optimization Methods
43. Robust Kronecker-Decomposable Component Analysis for Low-Rank Modeling, Mehdi Bahri, Yannis Panagakis, Stefanos Zafeiriou

Recognition
45. A Two-Streamed Network for Estimating Fine-Scaled Depth Maps From Single RGB Images, Jun Li, Reinhard Klein, Angela Yao
46. Weakly Supervised Object Localization Using Things and Stuff Transfer, Miaojing Shi, Holger Caesar, Vittorio Ferrari
47. Single Image Action Recognition Using Semantic Body Part Actions, Zhichen Zhao, Huimin Ma, Shaodi You
48. Incremental Learning of Object Detectors Without Catastrophic Forgetting, Konstantin Shmelkov, Cordelia Schmid, Karteek Alahari
49. Generative Adversarial Networks Conditioned by Brain Signals, Simone Palazzo, Concetto Spampinato, Isaak Kavasidis, Daniela Giordano, Mubarak Shah
50. Learning to Disambiguate by Asking Discriminative Questions, Yining Li, Chen Huang, Xiaohu Tang, Chen Change Loy
51. Interpretable Explanations of Black Boxes by Meaningful Perturbation, Ruth C. Fong, Andrea Vedaldi
52. DeepRoadMapper: Extracting Road Topology From Aerial Images, Gellért Mátyus, Wenjie Luo, Raquel Urtasun
53. Monocular 3D Human Pose Estimation by Predicting Depth on Joints, Bruce Xiaohan Nie, Ping Wei, Song-Chun Zhu
54. Large-Scale Image Retrieval With Attentive Deep Local Features, Hyeonwoo Noh, Andre Araujo, Jack Sim, Tobias Weyand, Bohyung Han
55. Deep Globally Constrained MRFs for Human Pose Estimation, Ioannis Marras, Petar Palasek, Ioannis Patras
56. Predicting Visual Exemplars of Unseen Classes for Zero-Shot Learning, Soravit Changpinyo, Wei-Lun Chao, Fei Sha
57. Multi-Label Learning of Part Detectors for Heavily Occluded Pedestrian Detection, Chunluan Zhou, Junsong Yuan
58. SGN: Sequential Grouping Networks for Instance Segmentation, Shuo Liu, Jiaya Jia, Sanja Fidler, Raquel Urtasun
60. Aesthetic Critiques Generation for Photos, Kuang-Yu Chang, Kung-Hung Lu, Chu-Song Chen
61. Hide-And-Seek: Forcing a Network to Be Meticulous for Weakly-Supervised Object and Action Localization, Krishna Kumar Singh, Yong Jae Lee

Segmentation, Grouping & Shape
62. Two-Phase Learning for Weakly Supervised Object Localization, Dahun Kim, Donghyeon Cho, Donggeun Yoo, In So Kweon

Statistical Methods & Learning
63. Curriculum Dropout, Pietro Morerio, Jacopo Cavazza, Riccardo Volpi, René Vidal, Vittorio Murino
64. Predictor Combination at Test Time, Kwang In Kim, James Tompkin, Christian Richardt
65. Guided Perturbations: Self-Corrective Behavior in Convolutional Neural Networks, Swami Sankaranarayanan, Aprit Jain, Ser Nam Lim
66. Learning Robust Visual-Semantic Embeddings, Yao-Hung Hubert Tsai, Liang-Kang Huang, Ruslan Salakhutdinov
67. PUnda: Probabilistic Unsupervised Domain Adaptation for Knowledge Transfer Across Visual Categories, Behnam Gholami, Ognjen (Oggi) Rudovic, Vladimir Pavlovic

Video
69. CDTS: Collaborative Detection, Tracking, and Segmentation for Online Multiple Object Segmentation in Videos, Yeong Jun Koh, Chang-Su Kim

25
70. Temporal Superpixels Based on Proximity-Weighted Patch Matching, Se-Ho Lee, Won-Dong Jang, Chang-Su Kim
71. Joint Detection and Recounting of Abnormal Events by Learning Deep Generic Knowledge, Ryota Hinami, Tao Mei, Shin‘ichi Satoh
72. TURN TAP: Temporal Unit Regression Network for Temporal Action Proposals, Jiayang Gao, Zhenheng Yang, Kan Chen, Chen Sun, Ram Nevatia
73. Online Real-Time Multiple Spatiotemporal Action Localisation and Prediction, Gurkirt Singh, Suman Saha, Michael Sapienza, Philip H. S. Torr, Fabio Cuzzolin
74. Leveraging Weak Semantic Relevance for Complex Video Event Classification, Chao Li, Jiewei Cao, Zi Huang, Lei Zhu, Heng Tao Shen
75. Weakly Supervised Summarization of Web Videos, Rameswar Panda, Abir Das, Ziyan Wu, Jan Ernst, Amit K. Roy-Chowdhury
76. FCN-rLSTM: Deep Spatio-Temporal Neural Networks for Vehicle Counting in City Cameras, Shanghang Zhang, Guanhang Wu, João P. Costeira, José M. F. Moura

Vision for X
77. Fast Face-Swap Using Convolutional Neural Networks, Iryna Korshunova, Wenzhe Shi, Joni Dambre, Lucas Theis
78. Towards a Visual Privacy Advisor: Understanding and Predicting Privacy Risks in Images, Tribhuvanesh Orekondy, Bernt Schiele, Mario Fritz

1230–1400 Doctoral Consortium (Palazzo del Casinò) (by invitation only)
- Federica Arrigoni (Univ. of Udine)
- Monami Banerjee (Univ. of Florida)
- Jawadul Hasan Bappy (Univ. of California)
- Lorenzo Baraldi (Univ. degli Studi di Modena e Reggio Emilia)
- Rudrasis Chakraborty (Univ. of Florida)
- Gyeongmin Choe (Korea Advanced Inst. of Science and Tech.)
- Ionut Cosmin Duta (Univ. of Trento)
- Dong Gong (Northwestern Polytechnical Univ.)
- Kai Han (The Univ. of Hong Kong)
- Ben Harwood (Monash Univ.)
- Wei Ke (Univ. of Chinese Academy of Sciences)
- Seungryong Kim (Yonsei Univ.)
- Ziwei Liu (The Chinese Univ. of Hong Kong)
- Julieta Martinez (Univ. of British Columbia)
- Pascal Mettes (Univ. of Amsterdam)
- Xi Peng (Rutgers Univ.)
- Yifan (Evan) Peng (The Univ. of British Columbia)
- Debaditya Roy (Indian Inst. of Technology Hyderbad)
- Christian Rupprecht (Technische Universität München)
- Swami Sankaranarayanan (Univ. of Maryland)
- Yibing Song (City Univ. of Hong Kong)
- David Joseph New Tan (Technische Universität München)
- Wei Wang (Univ. of Trento)
- Xiangyu Xu (Tsinghua Univ.)
- Xi Yin (Michigan State Univ.)
- Zijing Zhao (The Hong Kong Polytechnic Univ.)

1030–1230 Demos (Salone Adriatico)

1030–1230 Exhibitors (Various locations)
- See exhibitor list on page 12.

1230–1330 Lunch
1330–1430 Session O5-S5: Face and Human Behaviour Analysis (Sala Grande, Sala Darsena, & Sala Perla)

Papers in this session are also in Poster Session P6.

**Chairs:** Yusuke Sugano (Osaka Univ.)
Yiannis Patras (Queen Mary Univ. of London)

1330 Orals (O5)

Format (9 min. for presentation + 3 min. for questions)

1. [1330] First-Person Activity Forecasting With Online Inverse Reinforcement Learning, Nicholas Rhinehart, Kris M. Kitani
2. [1342] Binarized Convolutional Landmark Localizers for Human Pose Estimation and Face Alignment With Limited Resources, Adrian Bulat, Georgios Tzimiropoulos
5. [1418] Temporal Non-Volume Preserving Approach to Facial Age-Progression and Age-Invariant Face Recognition, Chi Nhan Duong, Kha Gia Quach, Khoa Luu, Ngan Le, Marios Savvides

1430 Spotlights (S5)

Format (3 min. & 30 sec. for presentation; no questions)

2. [1433] Unlabeled Samples Generated by GAN Improve the Person Re-Identification Baseline in Vitro, Zhedong Zheng, Liang Zheng, Yi Yang
3. [1437] Egocentric Gesture Recognition Using Recurrent 3D Convolutional Neural Networks With Spatiotemporal Transformer Modules, Congqi Cao, Yifan Zhang, Yi Wu, Hanqing Lu, Jian Cheng
4. [1441] Recursive Spatial Transformer (ReST) for Alignment-Free Face Recognition, Wanglong Wu, Meina Kan, Xin Liu, Yi Yang, Shiguang Shan, Xilin Chen
5. [1445] Learning Discriminative Aggregation Network for Video-Based Face Recognition, Yongming Rao, Ji Lin, Jiwen Lu, Jie Zhou
6. [1448] Synergy Between Face Alignment and Tracking via Discriminative Global Consensus Optimization, Muhammad Haris Khan, John McDonagh, Georgios Tzimiropoulos
7. [1452] SVDNet for Pedestrian Retrieval, Yifan Sun, Liang Zheng, Weijian Deng, Shengjin Wang
8. [1456] Towards More Accurate Iris Recognition Using Deeply Learned Spatially Corresponding Features, Zijing Zhao, Ajay Kumar

1500–1700 Poster Session P6 (Sala Laguna, Salone Adriatico, & Sale Mosaici 1–2)

See poster locations map on pg. 1.

**Oral O5 Posters**

1. First-Person Activity Forecasting With Online Inverse Reinforcement Learning, Nicholas Rhinehart, Kris M. Kitani
2. Binarized Convolutional Landmark Localizers for Human Pose Estimation and Face Alignment With Limited Resources, Adrian Bulat, Georgios Tzimiropoulos
4. RPAN: An End-To-End Recurrent Pose-Attention Network for Action Recognition in Videos, Wenbin Du, Yali Wang, Yu Qiao
5. Temporal Non-Volume Preserving Approach to Facial Age-Progression and Age-Invariant Face Recognition, Chi Nhan Duong, Kha Gia Quach, Khoa Luu, Ngan Le, Marios Savvides

**Spotlight S5 Posters**

7. Unlabeled Samples Generated by GAN Improve the Person Re-Identification Baseline in Vitro, Zhedong Zheng, Liang Zheng, Yi Yang
8. Egocentric Gesture Recognition Using Recurrent 3D Convolutional Neural Networks With Spatiotemporal Transformer Modules, Congqi Cao, Yifan Zhang, Yi Wu, Hanqing Lu, Jian Cheng
9. Recursive Spatial Transformer (ReST) for Alignment-Free Face Recognition, Wanglong Wu, Meina Kan, Xin Liu, Yi Yang, Shiguang Shan, Xilin Chen
10. Learning Discriminative Aggregation Network for Video-Based Face Recognition, Yongming Rao, Ji Lin, Jiwen Lu, Jie Zhou
11. Synergy Between Face Alignment and Tracking via Discriminative Global Consensus Optimization, Muhammad Haris Khan, John McDonagh, Georgios Tzimiropoulos
12. SVDNet for Pedestrian Retrieval, Yifan Sun, Liang Zheng, Weijian Deng, Shengjin Wang
13. Towards More Accurate Iris Recognition Using Deeply Learned Spatially Corresponding Features, Zijing Zhao, Ajay Kumar

**3D Computer Vision**
15. BB8: A Scalable, Accurate, Robust to Partial Occlusion Method for Predicting the 3D Poses of Challenging Objects Without Using Depth, Mahdi Rad, Vincent Lepetit
17. Parameter-Free Lens Distortion Calibration of Central Cameras, Filippo Bergamasco, Luca Cosmo, Andrea Gasparetto, Andrea Albarelli, Andrea Torsello
18. Pose Guided RGBD Feature Learning for 3D Object Pose Estimation, Vassileios Baltas, Andreas Doumanoglou, Caner Sahin, Joël Sock, Rigas Kouskouridas, Tae-Kyun Kim
19. Efficient Global Illumination for Morphable Models, Andreas Schneider, Sandro Schönborn, Lavrenti Frobelen, Bernhard Egger, Thomas Vetter
20. Low Compute and Fully Parallel Computer Vision With HashMatch, Sean Ryan Fanello, Julien Valentin, Adarsh Kowdle, Christoph Rhemann, Vladimir Tankovich, Carlo Ciliberto, Philip Davidson, Shahram Izadi
21. Dense Non-Rigid Structure-From-Motion and Shading With Unknown Albedos, Mathias Gallardo, Toby Collins, Adrien Bartoli
22. From Point Clouds to Mesh Using Regression, Lubor Ladický, Olivier Saurer, SoHyeon Jeong, Fabio Maninchedda, Marc Pollefeys
23. Stereo DSO: Large-Scale Direct Sparse Visual Odometry With Stereo Cameras, Rui Wang, Martin Schwörer, Daniel Cremers
24. Space-Time Localization and Mapping, Minhaeng Lee, Charless C. Fowlkes

**Computational Photography**
26. Attention-Aware Deep Reinforcement Learning for Video Face Recognition, Yongming Rao, Jiwen Lu, Jie Zhou
27. Learning to Fuse 2D and 3D Image Cues for Monocular Body Pose Estimation, Bugra Tekin, Pablo Márquez-Neila, Mathieu Salzmann, Pascal Fua
28. Deep Facial Action Unit Recognition From Partially Labeled Data, Shan Wu, Shangfei Wang, Bowen Pan, Qiang Ji
29. Pose-Driven Deep Convolutional Model for Person Re-Identification, Chi Su, Jianing Li, Shiliang Zhang, Junliang Xing, Wen Gao, Qi Tian
30. Recognition of Action Units in the Wild With Deep Nets and a New Global-Local Loss, C. Fabian Benitez-Quiroz, Yan Wang, Aleix M. Martinez
32. Towards Large-Pose Face Frontalization in the Wild, Xi Yin, Xiang Yu, Kihyuk Sohn, Xiaoming Liu, Manmohan Chandraker

**Face & Gesture**
33. A Joint Intrinsic-Extrinsic Prior Model for Retinex, Bolun Cai, Xianming Xu, Kailing Guo, Kui Jia, Bin Hu, Dacheng Tao
34. Going Unconstrained With Rolling Shutter Deblurring, Mahesh Mohan M. R., A. N. Rajagopalan, Gunasekaran Seetharaman
36. From Square Pieces to Brick Walls: The Next Challenge in Solving Jigsaw Puzzles, Shir Gur, Ohad Ben-Shahar
37. Online Video Deblurring via Dynamic Temporal Blending Network, Tae Hyun Kim, Kyoung Mu Lee, Bernhard Schölkopf, Michael Hirsch
38. Supervision by Fusion: Towards Unsupervised Learning of Deep Salient Object Detector, Dingwen Zhang, Junwei Han, Yu Zhang
39. Fast Multi-Image Matching via Density-Based Clustering, Roberto Tron, Xiaowei Zhou, Carlos Esteves, Kostas Daniilidis
40. Characterizing and Improving Stability in Neural Style Transfer, Agrim Gupta, Justin Johnson, Alexandre Alahi, Li Fei-Fei

Recognition
41. Cross-Modal Deep Variational Hashing, Venice Erin Liong, Jiwen Lu, Yap-Peng Tan, Jie Zhou
42. Spatial Memory for Context Reasoning in Object Detection, Xinlei Chen, Abhinav Gupta
43. Deep Binaries: Encoding Semantic-Rich Cues for Efficient Textual-Visual Cross Retrieval, Yuming Shen, Li Liu, Ling Shao, Jingkuan Song
44. Learning a Recurrent Residual Fusion Network for Multimodal Matching, Yu Liu, Yanming Guo, Erwin M. Bakker, Michael S. Lew
45. Rotational Subgroup Voting and Pose Clustering for Robust 3D Object Recognition, Anders Glent Buch, Lilita Kiforenko, Dirk Kraft
46. CoupleNet: Coupling Global Structure With Local Parts for Object Detection, Yousong Zhu, Chaoyang Zhao, Jinqiao Wang, Xu Zhao, Yi Wu, Hanqing Lu
47. Speaking the Same Language: Matching Machine to Human Captions by Adversarial Training, Rakshith Shetty, Marcus Rohrbach, Lisa Anne Hendricks, Mario Fritz, Bernt Schiele
50. Situation Recognition With Graph Neural Networks, Ruiyu Li, Makandar Tapaswi, Renjie Liao, Jiaya Jia, Raquel Urtasun, Sanja Fidler
51. Learning Visual N-Grams From Web Data, Ang Li, Allan Jabri, Armand Joulin, Laurens van der Maaten
52. Attention-Based Multimodal Fusion for Video Description, Chiori Hori, Takaaki Hori, Teng-Yok Lee, Ziming Zhang, Bret Harsham, John R. Hershey, Tim K. Marks, Kazuhiko Sumi
53. Learning the Latent “Look”: Unsupervised Discovery of a Style-Coherent Embedding From Fashion Images, Wei-Lin Hsiao, Kristen Grauman
54. Aligned Image-Word Representations Improve Inductive Transfer Across Vision-Language Tasks, Tanmoy Gupta, Kevin Shih, Saurabh Singh, Derek Hoiem
55. Learning Discriminative Latent Attributes for Zero-Shot Classification, Huajie Jiang, Ruiping Wang, Shiguang Shan, Yi Yang, Xilin Chen
56. PPR-FCN: Weakly Supervised Visual Relation Detection via Parallel Pairwise R-FCN, Hanwang Zhang, Zawlin Kyaw, Jinyang Yu, Shih-Fu Chang

Segmentation, Grouping & Shape
57. Higher-Order Minimum Cost Lifted Multicuts for Motion Segmentation, Margret Keuper
58. Deep Free-Form Deformation Network for Object-Mask Registration, Haoyang Zhang, Xuming He
59. Region-Based Correspondence Between 3D Shapes via Spatially Smooth Biclustering, Matteo Denitto, Simone Melzi, Manuele Bicego, Umberto Castellani, Alessandro Farinelli, Mário A. T. Figueiredo, Yanir Kleiman, Maks Ovsjanikov

Statistical Methods & Learning
60. Learning Discriminative ab-Divergences for Positive Definite Matrices, Anoop Cherian, Panagiotis Stanitsas, Mehrtash Harandi, Vassilios Morellas, Nikolaos Papanikolopoulos
61. Consensus Convolutional Sparse Coding, Biswarup Choudhury, Robin Swanson, Felix Heide, Gordon Wetzstein, Wolfgang Heidrich
63. Self-Supervised Learning of Pose Embeddings From Spatiotemporal Relations in Videos, Ömer Sümer, Tobias Dencker, Björn Ommer
64. Approximate Grassmannian Intersections: Subspace-Valued Subspace Learning, Calvin Murdock, Fernando De la Torre
65. Side Information in Robust Principal Component Analysis: Algorithms and Applications, Niannan Xue, Yannis Panagakis, Stefanos Zafeiriou
1545–1630 Afternoon Break

1700–1800 Session O6: Video Analysis (Sala Grande, Sala Darsena, & Sala Perla)

Papers in this session are also in Poster Session P7.

Chairs: Timothy Hospedales (Univ. of Edinburgh)
Shin’ichi Satoh (National Inst. of Informatics)

Format (9 min. for presentation + 3 min. for questions)

1. [1700] Spatial-Aware Object Embeddings for Zero-Shot Localization and Classification of Actions, Pascal Mettes, Cees G. M. Snoek
2. [1712] Semantic Video CNNs Through Representation Warping, Raghudeep Gadde, Varun Jampani, Peter V. Gehler
4. [1736] Detail-Revealing Deep Video Super-Resolution, Xin Tao, Hongyun Gao, Renjie Liao, Jue Wang, Jiaya Jia
5. [1748] Learning Video Object Segmentation With Visual Memory, Pavel Tokmakov, Karteek Alahari, Cordelia Schmid

1900–2200 Dinner (Aresenale)

Directions: The banquet will be held at the Aresenale in Venice, an historical venue which once was the Venetian Arsenal and shipyard. The venue is now also hosting the Biennale exhibition and it will be possible for ICCV guests to visit the adjacent pavilions (South Korea, China and Lebanon) for free.

Transportation will be organized for all participants from the conference venue to the banquet (you’ll be asked to show your badge or banquet ticket). At the end of the banquet, a transportation service will be organized back from Aresenale to San Marco Square and Lido Santa Maria Elisabetta.

Please note that we strongly discourage you to get to the Aresenale on your own, as it will take much more time to get there than with our dedicated transfers. More info will be available onsite for all banquet ticket holders.
Friday, October 27

0800–1800 Registration (Palazzo del Casinò Terrace)

0900–1000 Session O7-S6: Low-Level Vision (Sala Grande, Sala Darsena, & Sala Perla)

Papers in this session are also in Poster Session P7.

Chairs: Paolo Favaro (Univ. of Bern)
Theo Gevers (Univ. of Amsterdam)

0900 Orals (O7)

Format (9 min. for presentation + 3 min. for questions)

2. [0912] Makeup-Go: Blind Reversion of Portrait Edit, Ying-Cong Chen, Xiaoyong Shen, Jiaya Jia
3. [0924] Shadow Detection With Conditional Generative Adversarial Networks, Vu Nguyen, Tomas F. Yago Vicente, Maozheng Zhao, Minh Hoai, Dimitris Samaras
4. [0936] Learning High Dynamic Range From Outdoor Panoramas, Jinsong Zhang, Jean-François Lalonde

1000 Spotlights (S6)

Format (3 min. & 30 sec. for presentation; no questions)

1. [1000] MemNet: A Persistent Memory Network for Image Restoration, Ying Tai, Jian Yang, Xiaoming Liu, Chunyan Xu
2. [1003] Structure-Measure: A New Way to Evaluate Foreground Maps, Deng-Ping Fan, Ming-Ming Cheng, Yun Liu, Tao Li, Ali Borji
4. [1011] Practical and Efficient Multi-View Matching, Eleonora Maset, Federica Arrigoni, Andrea Fusiello
5. [1015] Unrolled Memory Inner-Products: An Abstract GPU Operator for Efficient Vision-Related Computations, Yu-Sheng Lin, Wei-Chao Chen, Shao-Yi Chien

1030–1115 Morning Break

1030–1230 Poster Session P7 (Sala Laguna, Salone Adriatico, & Sale Mosaici 1–2)

See poster locations map on pg. 1.

Oral O6 Posters

1. Spatial-Aware Object Embeddings for Zero-Shot Localization and Classification of Actions, Pascal Mettes, Cees G. M. Snoek
2. Semantic Video CNNs Through Representation Warping, Raghuudeep Gadde, Varun Jampani, Peter V. Gehler
4. Detail-Revealing Deep Video Super-Resolution, Xin Tao, Hongyun Gao, Renjie Liao, Jue Wang, Jiaya Jia
5. Learning Video Object Segmentation With Visual Memory, Pavel Tokmakov, Karteek Alahari, Cordelia Schmid

Oral O7 Posters

7. Makeup-Go: Blind Reversion of Portrait Edit, Ying-Cong Chen, Xiaoyong Shen, Jiaya Jia
8. Shadow Detection With Conditional Generative Adversarial Networks, Vu Nguyen, Tomas F. Yago Vicente, Maozheng Zhao, Minh Hoai, Dimitris Samaras
9. Learning High Dynamic Range From Outdoor Panoramas, Jinsong Zhang, Jean-François Lalonde
10. DCTM: Discrete-Continuous Transformation Matching for Semantic Flow, Seungryong Kim, Dongbo Min, Stephen Lin, Kwanghoon Sohn
**Spotlight S6 Posters**

11. MemNet: A Persistent Memory Network for Image Restoration, Ying Tai, Jian Yang, Xiaoming Liu, Chunyan Xu
12. Structure-Measure: A New Way to Evaluate Foreground Maps, Deng-Ping Fan, Ming-Ming Cheng, Yun Liu, Tao Li, Ali Borji
14. Practical and Efficient Multi-View Matching, Eleonora Maset, Federica Arrigoni, Andrea Fusiello
15. Unrolled Memory Inner-Products: An Abstract GPU Operator for Efficient Vision-Related Computations, Yu-Sheng Lin, Wei-Chao Chen, Shao-Yi Chien
16. Learning to Push the Limits of Efficient FFT-Based Image Deconvolution, Jakob Kruse, Carsten Rother, Uwe Schmidt
17. Learning Spread-Out Local Feature Descriptors, Xu Zhang, Felix X. Yu, Sanjiv Kumar, Shih-Fu Chang

**3D Computer Vision**

19. Joint Estimation of Camera Pose, Depth, Deblurring, and Super-Resolution From a Blurred Image Sequence, Haesol Park, Kyoung Mu Lee
20. 2D-Driven 3D Object Detection in RGB-D Images, Jean Lahoud, Bernard Ghanem
21. Ray Space Features for Plenoptic Structure-From-Motion, Yingliang Zhang, Peihong Yu, Wei Yang, Yuanxi Ma, Jingyi Yu
22. Depth Estimation Using Structured Light Flow — Analysis of Projected Pattern Flow on an Object's Surface, Ryo Furukawa, Ryusuke Sagawa, Hiroshi Kawasaki
23. Monocular Dense 3D Reconstruction of a Complex Dynamic Scene From Two Perspective Frames, Suryansh Kumar, Yuchao Dai, Hongdong Li
25. Dynamics Enhanced Multi-Camera Motion Segmentation From Unsynchronized Videos, Xikang Zhang, Bengisu Ozbay, Mario Sznaier, Octavia Camps
26. Taking the Scenic Route to 3D: Optimising Reconstruction From Moving Cameras, Oscar Mendez, Simon Hadfield, Nicolas Pugeault, Richard Bowden
27. FLaME: Fast Lightweight Mesh Estimation Using Variational Smoothing on Delaunay Graphs, W. Nicholas Greene, Nicholas Roy

**Biomedical Image Analysis**


**Computational Photography**

29. From RGB to Spectrum for Natural Scenes via Manifold-Based Mapping, Yan Jia, Yinqiang Zheng, Lin Gu, Art Subpa-Asa, Antony Lam, Yoichi Sato, Imari Sato

**Face & Gesture**

31. Learning Dense Facial Correspondences in Unconstrained Images, Ronald Yu, Shunsuke Saito, Haoxiang Li, Duygu Ceylan, Hao Li
32. Jointly Attentive Spatial-Temporal Pooling Networks for Video-Based Person Re-Identification, Shuangjie Xu, Yu Cheng, Kang Gu, Yang Yang, Shiyu Chang, Pan Zhou

**Low-Level Vision & Image Processing**

33. Automatic Content-Aware Projection for 360° Videos, Yeong Won Kim, Chang-Ryeol Lee, Dae-Yong Cho, Yong Hoon Kwon, Hyeok-Jae Choi, Kuk-Jin Yoon
35. Non-Linear Convolution Filters for CNN-Based Learning, Georgios Zoumpourlis, Alexandros Doumanoglou, Nicholas Vretos, Petros Daras
36. AOD-Net: All-In-One Dehazing Network, Boyi Li, Xiulian Peng, Zhangyang Wang, Jizheng Xu, Dan Feng
37. Simultaneous Detection and Removal of High Altitude Clouds From an Image, Tushar Sandhan, Jin Young Choi
38. Understanding Low- and High-Level Contributions to Fixation Prediction, Mathias Kümmerer, Thomas S. A. Wallis, Leon A. Gatys, Matthias Bethge
39. Image Super-Resolution Using Dense Skip Connections, Tong Tong, Gen Li, Xiejie Liu, Qinquan Gao
40. Convergence Analysis of MAP Based Blur Kernel Estimation, Sunghyun Cho, Seunghyung Lee
42. Deep Generative Adversarial Compression Artifact Removal, Leonardo Galteri, Lorenzo Seidenari, Marco Bertini, Alberto Del Bimbo

**Motion & Tracking**
43. Online Multi-Object Tracking Using CNN-Based Single Object Tracker With Spatial-Temporal Attention Mechanism, Qi Chu, Wanli Ouyang, Hongsheng Li, Xiaogang Wang, Bin Liu, Nenghai Yu

**Recognition**
44. Mutual Enhancement for Detection of Multiple Logos in Sports Videos, Yuan Liao, Xiaqing Lu, Chengcui Zhang, Yongtao Wang, Zhi Tang
45. Referring Expression Generation and Comprehension via Attributes, Jingyu Liu, Liang Wang, Ming-Hsuan Yang
46. RoomNet: End-To-End Room Layout Estimation, Chen-Yu Lee, Vijay Badrinarayanan, Tomasz Malisiewicz, Andrew Rabinovich
47. SSH: Single Stage Headless Face Detector, Mahyar Najibi, Pouya Samangouei, Rama Chellappa, Larry S. Davis
49. Boosting Image Captioning With Attributes, Ting Yao, Yingwei Pan, Yehao Li, Zhaofan Qiu, Tao Mei
50. Learning to Estimate 3D Hand Pose From Single RGB Images, Christian Zimmermann, Thomas Brox
51. Locally-Transferred Fisher Vectors for Texture Classification, Yang Song, Fan Zhang, Qing Li, Heng Huang, Lauren J. O’Donnell, Weidong Cai
52. Object-Level Proposals, Jianxiang Ma, Anlong Ming, Zilong Huang, Xinggang Wang, Yu Zhou
53. Extreme Clicking for Efficient Object Annotation, Dim P. Papadopoulos, Jasper R. R. Uijlings, Frank Keller, Vittorio Ferrari
54. WordSup: Exploiting Word Annotations for Character Based Text Detection, Han Hu, Chengquan Zhang, Yuxuan Luo, Yuzhuo Wang, Junyu Han, Erui Ding
55. Illuminating Pedestrians via Simultaneous Detection & Segmentation, Garrick Brazil, Xi Yin, Xiaoming Liu

56. Generalized Orderless Pooling Performs Implicit Salient Matching, Marcel Simon, Yang Gao, Trevor Darrell, Joachim Denzler, Erik Rodner

**Segmentation, Grouping & Shape**
58. RDFNet: RGB-D Multi-Level Residual Feature Fusion for Indoor Semantic Segmentation, Seong-Jin Park, Ki-Sang Hong, Seungyong Lee
59. The Mapillary Vistas Dataset for Semantic Understanding of Street Scenes, Gerhard Neuhold, Tobias Ollmann, Samuel Rota Bulò, Peter Kontschieder
60. Self-Organized Text Detection With Minimal Post-Processing via Border Learning, Yue Wu, Prem Natarajan

**Statistical Methods & Learning**
61. Sparse Exact PGA on Riemannian Manifolds, Monami Banerjee, Rudra Sundaresha Chakraborty, Baba C. Vemuri
62. Tensor RPCA by Bayesian CP Factorization With Complex Noise, Qiong Luo, Zhi Han, Xi’ai Chen, Yao Wang, Deyu Meng, Dong Liang, Yandong Tang
63. Multimodal Gaussian Process Latent Variable Models With Harmonization, Guoli Song, Shuhui Wang, Qingming Huang, Qi Tian
64. Segmentation-Aware Convolutional Networks Using Local Attention Masks, Adam W. Harley, Konstantinos G. Derpanis, Iasonas Kokkinos
65. Rotation Equivariant Vector Field Networks, Diego Marcos, Michele Volpi, Nikos Komodakis, Devis Tuia
67. AutoDIAL: Automatic Domain Alignment Layers, Fabio Maria Carlucci, Lorenzo Porzi, Barbara Caputo, Elisa Ricci, Samuel Rota Bulò
69. Unsupervised Object Segmentation in Video by Efficient Selection of Highly Probable Positive Features, Emanuela Haller, Marius Leordeanu
70. Nonparametric Variational Auto-Encoders for Hierarchical Representation Learning, Prasoon Goyal, Zhiting Hu, Xiaodan Liang, Chenyu Wang, Eric P. Xing
71. Dense and Low-Rank Gaussian CRFs Using Deep Embeddings, *Siddhartha Chandra, Nicolas Usunier, Iasonas Kokkinos*

Video
72. A Multimodal Deep Regression Bayesian Network for Affective Video Content Analyses, *Quan Gan, Shangfei Wang, Longfei Hao, Qiang Ji*
73. Moving Object Detection in Time-Lapse or Motion Trigger Image Sequences Using Low-Rank and Invariant Sparse Decomposition, *Moein Shakeri, Hong Zhang*

74. A Multilayer-Based Framework for Online Background Subtraction With Freely Moving Cameras, *Yizhe Zhu, Ahmed Elgammal*

75. Dynamic Label Graph Matching for Unsupervised Video Re-Identification, *Mang Ye, Andy J. Ma, Liang Zheng, Jiawei Li, Pong C. Yuen*

76. Spatiotemporal Modeling for Crowd Counting in Videos, *Feng Xiong, Xingjian Shi, Dit-Yan Yeung*

Vision for X
77. Personalized Cinemagraphs Using Semantic Understanding and Collaborative Learning, *Tae-Hyun Oh, Kyungdon Joo, Neel Joshi, Baoyuan Wang, In So Kweon, Sing Bing Kang*

78. What Is Around the Camera?, *Stamatios Georgoulis, Konstantinos Rematas, Tobias Ritschel, Mario Fritz, Tinne Tuytelaars, Luc Van Gool*

1030–1230 Demos (Salone Adriatico)

1030–1230 Exhibitors (Various locations)
- See exhibitor list on page 12.

1230–1330 Lunch

Notes:
1330–1500 Session O8-S7: Recognition 3 (Sala Grande, Sala Darsena, & Sala Perla)

Papers in this session are also in Poster Session P8.

Chairs: Kristen Grauman (Univ. of Texas at Austin)  
Kostas Daniilidis (Univ. of Pennsylvania)

1330 Orals (O8)

Format (9 min. for presentation + 3 min. for questions)

3. [1354] 3D Graph Neural Networks for RGBD Semantic Segmentation, Xiaojuan Qi, Renjie Liao, Jiaya Jia, Sanja Fidler, Raquel Urtasun
4. [1406] Learning Multi-Attention Convolutional Neural Network for Fine-Grained Image Recognition, Heliang Zheng, Jianlong Fu, Tao Mei, Jiebo Luo
5. [1418] Learning 3D Object Categories by Looking Around Them, David Novotny, Diane Larlus, Andrea Vedaldi

1430 Spotlights (S7)

Format (3 min. & 30 sec. for presentation; no questions)

2. [1433] Towards End-To-End Text Spotting With Convolutional Recurrent Neural Networks, Hui Li, Peng Wang, Chunhua Shen
4. [1441] Learning From Video and Text via Large-Scale Discriminative Clustering, Antoine Miech, Jean-Baptiste Alayrac, Piotr Bojanowski, Ivan Laptev, Josef Sivic
5. [1445] TALL: Temporal Activity Localization via Language Query, Jiyang Gao, Chen Sun, Zhenheng Yang, Ram Nevatia
6. [1448] End-To-End Face Detection and Cast Grouping in Movies Using Erdős-Rényi Clustering, SouYoung Jin, Hang Su, Chris Stauffer, Erik Learned-Miller
7. [1452] Active Decision Boundary Annotation With Deep Generative Models, Miriam Huijser, Jan C. van Gemert
8. [1456] Convolutional Dictionary Learning via Local Processing, Vardan Popyan, Yaniv Romano, Jeremias Sulam, Michael Elad

1500–1700 Poster Session P8 (Sala Laguna, Salone Adriatico, & Sale Mosaici 1–2)

See poster locations map on pg. 1.

Oral O8 Posters

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

Spotlight S7 Posters

7. Towards End-To-End Text Spotting With Convolutional Recurrent Neural Networks, Hui Li, Peng Wang, Chunhua Shen
9. Learning From Video and Text via Large-Scale Discriminative Clustering, Antoine Miech, Jean-Baptiste Alayrac, Piotr Bojanowski, Ivan Laptev, Josef Sivic
10. TALL: Temporal Activity Localization via Language Query, Jiyang Gao, Chen Sun, Zhenheng Yang, Ram Nevatia
11. End-To-End Face Detection and Cast Grouping in Movies Using Erdős-Rényi Clustering, SouYoung Jin, Hang Su, Chris Stauffer, Erik Learned-Miller
12. Active Decision Boundary Annotation With Deep Generative Models, Miriam Huijser, Jan C. van Gemert
13. Convolutional Dictionary Learning via Local Processing, Vardan Papyan, Yaniv Romano, Jeremias Sulam, Michael Elad

**Oral & Posters**


15. One Network to Solve Them All — Solving Linear Inverse Problems Using Deep Projection Models, J. H. Rick Chang, Chun-Liang Li, Barnabás Póczos, B. V. K. Vijaya Kumar, Aswin C. Sankaranarayanan

16. Representation Learning by Learning to Count, Mehdi Noroozi, Hamed Pirsiavash, Paolo Favaro

17. StackGAN: Text to Photo-Realistic Image Synthesis With Stacked Generative Adversarial Networks, Han Zhang, Tao Xu, Hongsheng Li, Shaoting Zhang, Xiaogang Wang, Xiaolei Huang, Dimitris N. Metaxas

18. Unsupervised Learning of Object Landmarks by Factorized Spatial Embeddings, James Thewlis, Hakan Bilen, Andrea Vedaldi

**3D Computer Vision**

19. Editable Parametric Dense Foliage From 3D Capture, Gaurav Chaurasia, Paul Beardsley

20. Refractive Structure-From-Motion Through a Flat Refractive Interface, François Chadebecq, Francisco Vasconcelos, George Dwyer, René Lacher, Sébastien Ourselin, Tom Vercauteren, Danail Stoyanov

21. Submodular Trajectory Optimization for Aerial 3D Scanning, Mike Roberts, Debadeepta Dey, Anh Truong, Sudipta Sinha, Shital Shah, Ashish Kapoor, Pat Hanrahan, Neel Joshi

22. Camera Calibration by Global Constraints on the Motion of Silhouettes, Gil Ben-Artzi

23. Deltille Grids for Geometric Camera Calibration, Hyowon Ha, Michal Perdoch, Hatem Alismail, In So Kweon, Yaser Sheikh

**Computational Photography**


25. Reflectance Capture Using Univariate Sampling of BRDFs, Zhuo Hui, Kalyan Sunkavalli, Joon-Young Lee, Sunil Hadap, Jian Wang, Aswin C. Sankaranarayanan

26. Estimating Defocus Blur via Rank of Local Patches, Guodong Xu, Yuhui Quan, Hui Ji

**Face & Gesture**

27. RGB-Infrared Cross-Modality Person Re-Identification, Ancong Wu, Wei-Shi Zheng, Hong-Xing Yu, Shaogang Gong, Jianhuang Lai

28. Intrinsic 3D Dynamic Surface Tracking Based on Dynamic Ricci Flow and Teichmüller Map, Xiaokang Yu, Na Lei, Yalin Wang, Xianfeng Gu

29. Multi-Scale Deep Learning Architectures for Person Re-Identification, Xuelin Qian, Yanwei Fu, Yu-Gang Jiang, Tao Xiang, Xiangyang Xue

30. Range Loss for Deep Face Recognition With Long-Tailed Training Data, Xiao Zhang, Zhiyuan Fang, Yandong Wen, Zhifeng Li, Yu Qiao


**Low-Level Vision & Image Processing**

32. Realistic Dynamic Facial Textures From a Single Image Using GANs, Kyle Olszewski, Zimo Li, Chao Yang, Yi Zhou, Ronald Yu, Zeng Huang, Sitao Xiang, Shunsuke Saito, Pushmeet Kohli, Hao Li

33. Pixel Recursive Super Resolution, Ryan Dahl, Mohammad Norouzi, Jonathon Shlens

34. Recurrent Color Constancy, Yanlin Qian, Ke Chen, Jarno Nikkanen, Joni-Kristian Kämäräinen, Jiří Matas


**Motion & Tracking**

36. Monocular Video-Based Trailer Coupler Detection Using Multiplexer Convolutional Neural Network, Yousef Atoum, Joseph Roth, Michael Bliss, Wende Zhang, Xiaoming Liu

37. Parallel Tracking and Verifying: A Framework for Real-Time and High Accuracy Visual Tracking, Heng Fan, Haibin Ling

38. Non-Rigid Object Tracking via Deformable Patches Using Shape-Preserved KCF and Level Sets, Xin Sun, Ngai-Man Cheung, Hongxun Yao, Yiluan Guo

**Optimization Methods**

39. A Discriminative View of MRF Pre-Processing Algorithms, Chen Wang, Charles Herrmann, Ramin Zabih
<table>
<thead>
<tr>
<th>Session</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recognition</td>
<td>Offline Handwritten Signature Modeling and Verification Based on Archetypal Analysis, Elias N. Zois, Ilias Theodorakopoulos, George Economou</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Long Short-Term Memory Kalman Filters: Recurrent Neural Estimators for Pose Regularization, Huseyin Coskun, Felix Achilles, Robert DiPietro, Nassir Navab, Federico Tombari</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Learning Spatio-Temporal Representation With Pseudo-3D Residual Networks, Zhaofan Qiu, Ting Yao, Tao Mei</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Deeper, Broader and Artier Domain Generalization, Da Li, Yongxin Yang, Yi-Zhe Song, Timothy M. Hospedales</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Deep Spatial-Semantic Attention for Fine-Grained Sketch-Based Image Retrieval, Jifei Song, Qian Yu, Yi-Zhe Song, Tao Xiang, Timothy M. Hospedales</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Soft-NMS — Improving Object Detection With One Line of Code, Navaneeth Bodla, Bharat Singh, Rama Chellappa, Larry S. Davis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Semantic Jitter: Dense Supervision for Visual Comparisons via Synthetic Images, Aron Yu, Kristen Grauman</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Video Scene Parsing With Predictive Feature Learning, Xiaojie Jin, Xin Li, Huaxin Xiao, Xiaohui Shen, Zhe Lin, Jimei Yang, Yunpeng Chen, Jian Dong, Luoqi Liu, Zequn Jie, Jiashi Feng, Shuicheng Yan</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Understanding and Mapping Natural Beauty, Scott Workman, Richard Souvenir, Nathan Jacobs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Human Pose Estimation Using Global and Local Normalization, Ke Sun, Cuilng Lan, Junliang Xing, Wenjun Zeng, Dong Liu, Jingdong Wang</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HashNet: Deep Learning to Hash by Continuation, Zhangjie Cao, Mingsheng Long, Jianmin Wang, Philip S. Yu</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Scaling the Scattering Transform: Deep Hybrid Networks, Edouard Oyallon, Eugene Belilovsky, Sergey Zagoruyko</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Flip-Invariant Motion Representation, Takumi Kobayashi</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Scene Categorization With Spectral Features, Salman H. Khan, Munawar Hayat, Fatih Porikli</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Image2song: Song Retrieval via Bridging Image Content and Lyric Words, Xuelong Li, Di Hu, Xiaoqiang Lu</td>
<td></td>
</tr>
<tr>
<td>Segmentations, Groupings &amp; Shapes</td>
<td>Deep Functional Maps: Structured Prediction for Dense Shape Correspondence, Or Litany, Tal Remez, Emanuele Rodolà, Alex Bronstein, Michael Bronstein</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Training Deep Networks to Be Spatially Sensitive, Nicholas Kolkin, Eli Shechtman, Gregory Shakhnarovich</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3DCNN-DQN-RNN: A Deep Reinforcement Learning Framework for Semantic Parsing of Large-Scale 3D Point Clouds, Fangyu Liu, Shuaipeng Li, Liqiang Zhang, Chenghu Zhou, Rongtian Ye, Yuebin Wang, Jiwen Lu</td>
<td></td>
</tr>
<tr>
<td>Statistical Methods &amp; Learning</td>
<td>Efficient Low Rank Tensor Ring Completion, Wenqi Wang, Vaneet Aggarwal, Shuchin Aeron</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Semantic Image Synthesis via Adversarial Learning, Hao Dong, Simiao Yu, Chao Wu, Yike Guo</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Unified Deep Supervised Domain Adaptation and Generalization, Saeid Motiyan, Marco Piccirilli, Donald A. Adjeroh, Gianfranco Doretto</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Interpretable Transformations With Encoder-Decoder Networks, Daniel E. Worrall, Stephan J. Garbin, Daniyar Turmukhambetov, Gabriel J. Brostow</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Deep Clustering via Joint Convolutional Autoencoder Embedding and Relative Entropy Minimization, Kamran Ghasedi Dizaji, Amirhossein Herandi, Cheng Deng, Weidong Cai, Heng Huang</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Deep Scene Image Classification With the MFAFVNet, Yunsheng Li, Mandar Dixit, Nuno Vasconcelos</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Learning Bag-Of-Features Pooling for Deep Convolutional Neural Networks, Nikolaos Passalis, Anastasios Tefas</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adversarial Examples Detection in Deep Networks With Convolutional Filter Statistics, Xin Li, Fuxin Li</td>
<td></td>
</tr>
<tr>
<td>Video</td>
<td>Joint Prediction of Activity Labels and Starting Times in Untrimmed Videos, Tahmida Mahmud, Mahmudul Hasan, Amit K. Roy-Chowdhury</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R-C3D: Region Convolutional 3D Network for Temporal Activity Detection, Huijuan Xu, Abir Das, Kate Saenko</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Temporal Context Network for Activity Localization in Videos, Xiyang Dai, Bharat Singh, Guyue Zhang, Larry S. Davis, Yan Qiu Chen</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Localizing Moments in Video With Natural Language, Lisa Anne Hendricks, Oliver Wang, Eli Shechtman, Josef Sivic, Trevor Darrell, Bryan Russell</td>
<td></td>
</tr>
</tbody>
</table>
Friday, October 27 (Afternoon)

Program

72. Tube Convolutional Neural Network (T-CNN) for Action Detection in Videos, Rui Hou, Chen Chen, Mubarak Shah

73. Learning Action Recognition Model From Depth and Skeleton Videos, Hossein Rahmani, Mohammed Bennamoun


Vision for X

75. GPLAC: Generalizing Vision-Based Robotic Skills Using Weakly Labeled Images, Avi Singh, Larry Yang, Sergey Levine

76. Semi-Global Weighted Least Squares in Image Filtering, Wei Liu, Xiaogang Chen, Chuanhua Shen, Zhi Liu, Jie Yang

77. Scale Recovery for Monocular Visual Odometry Using Depth Estimated With Deep Convolutional Neural Fields, Xiaochuan Yin, Xiangwei Wang, Xiaoguo Du, Qijun Chen

1700–1800 Session O9: Machine Learning (Sala Grande, Sala Darsena, & Sala Perla)

Papers in this session are also in Poster Session P8.

Chairs: Iasonas Kokkinos (Univ. College London)
Ying Wu (Northwestern Univ.)

Format (9 min. for presentation + 3 min. for questions)


3. [1724] Representation Learning by Learning to Count, Mehdi Noroozi, Hamed Pirsiavash, Paolo Favaro

4. [1736] StackGAN: Text to Photo-Realistic Image Synthesis With Stacked Generative Adversarial Networks, Han Zhang, Tao Xu, Hongsheng Li, Shaoting Zhang, Xiaogang Wang, Xiaolei Huang, Dimitris N. Metaxas

5. [1748] Unsupervised Learning of Object Landmarks by Factorized Spatial Embeddings, James Thewlis, Hakan Bilen, Andrea Vedaldi

1500–1700 Demos (Salone Adriatico)

1500–1700 Exhibitors (Various locations)
- See exhibitor list on page 12.

1545–1630 Afternoon Break

END OF MAIN ICCV 2017 CONFERENCE

Notes:
Platinum Donors

Google
amazon
intel
Baidu 百度
DiDi
lyft
facebook
Sensetime
SenseTime
meitu
MMT Lab
Adobe
CrowdFlower
KLA
Tencor
Microsoft
Tencent

Gold Donors

UBER
ATG
ebay
FIVE
vuforia
nuTonomy