More information about this series at http://www.springer.com/series/7412
Foreword

Welcome to the proceedings of the 2016 edition of the European Conference on Computer Vision held in Amsterdam! It is safe to say that the European Conference on Computer Vision is one of the top conferences in computer vision. It is good to reiterate the history of the conference to see the broad base the conference has built in the 13 editions. First held in 1990 in Antibes (France), it was followed by subsequent conferences in Santa Margherita Ligure (Italy) in 1992, Stockholm (Sweden) in 1994, Cambridge (UK) in 1996, Freiburg (Germany) in 1998, Dublin (Ireland) in 2000, Copenhagen (Denmark) in 2002, Prague (Czech Republic) in 2004, Graz (Austria) in 2006, Marseille (France) in 2008, Heraklion (Greece) in 2010, Florence (Italy) in 2012, and Zürich (Switzerland) in 2014.

For the 14th edition, many people worked hard to provide attendees with a most warm welcome while enjoying the best science. The Program Committee, Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling, did an excellent job. Apart from the scientific program, the workshops were selected and handled by Hervé Jégou and Gang Hua, and the tutorials by Jacob Verbeek and Rita Cucchiara. Thanks for the great job. The coordination with the subsequent ACM Multimedia offered an opportunity to expand the tutorials with an additional invited session, offered by the University of Amsterdam and organized together with the help of ACM Multimedia.

Of the many people who worked hard as local organizers, we would like to single out Martine de Wit of the UvA Conference Office, who delicately and efficiently organized the main body. Also the local organizers Hamdi Dibeklioglu, Efstratios Gavves, Jan van Gemert, Thomas Mensink, and Mihir Jain had their hands full. As a venue, we chose the Royal Theatre Carré located on the canals of the Amstel River in downtown Amsterdam. Space in Amsterdam is sparse, so it was a little tighter than usual. The university lent us their downtown campuses for the tutorials and the workshops. A relatively new thing was the industry and the sponsors for which Ronald Poppe and Peter de With did a great job, while Andy Bagdanov and John Schavemaker arranged the demos. Michael Wilkinson took care to make Yom Kippur as comfortable as possible for those for whom it is an important day. We thank Marc Pollefeys, Alberto del Bimbo, and Virginie Mes for their advice and help behind the scenes. We thank all the anonymous volunteers for their hard and precise work. We also thank our generous sponsors. Their support is an essential part of the program. It is good to see such a level of industrial interest in what our community is doing!

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

October 2016

Theo Gevers
Arnold Smeulders
Preface

It is our great pleasure to present the workshop proceedings of the 14th European Conference on Computer Vision, which was held during October 8–16, 2016, in Amsterdam, The Netherlands. We were delighted that the main conference of ECCV 2016 was accompanied by 26 workshops. The workshop proceedings are presented in multiple Springer LNCS volumes.

This year, the 2016 ACM International Conference on Multimedia was collocated with ECCV 2016. As a synergistic arrangement, four out of the 26 ECCV workshops, whose topics are of interest to both the computer vision and multimedia communities, were held together with selected 2016 ACM Multimedia workshops.

We received 44 workshop proposals on a broad set of topics related to computer vision. The high quality of the proposals made the selection process rather difficult. Owing to space limitation, 27 proposals were accepted, among which two proposals were merged to form a single workshop due to overlapping themes.

The final 26 workshops complemented the main conference program well. The workshop topics present a good orchestration of new trends and traditional issues, as well as fundamental technologies and novel applications. We would like to thank all the workshop organizers for their unreserved efforts to make the workshop sessions a great success.

October 2016

Hervé Jégou
Gang Hua
Organization

General Chairs

Theo Gevers  University of Amsterdam, The Netherlands
Arnold Smeulders  University of Amsterdam, The Netherlands

Program Committee Co-chairs

Bastian Leibe  RWTH Aachen, Germany
Jiri Matas  Czech Technical University, Czech Republic
Nicu Sebe  University of Trento, Italy
Max Welling  University of Amsterdam, The Netherlands

Honorary Chair

Jan Koenderink  Delft University of Technology, The Netherlands and KU Leuven, Belgium

Advisory Program Chair

Luc van Gool  ETH Zurich, Switzerland

Advisory Workshop Chair

Josef Kittler  University of Surrey, UK

Advisory Conference Chair

Alberto del Bimbo  University of Florence, Italy

Local Arrangements Chairs

Hamdi Dibeklioglu  Delft University of Technology, The Netherlands
Efstratios Gavves  University of Amsterdam, The Netherlands
Jan van Gemert  Delft University of Technology, The Netherlands
Thomas Mensink  University of Amsterdam, The Netherlands
Michael Wilkinson  University of Groningen, The Netherlands
Workshop Chairs

Hervé Jégou  
Facebook AI Research, USA

Gang Hua  
Microsoft Research Asia, China

Tutorial Chairs

Jacob Verbeek  
Inria Grenoble, France

Rita Cucchiara  
University of Modena and Reggio Emilia, Italy

Poster Chairs

Jasper Uijlings  
University of Edinburgh, UK

Roberto Valenti  
Sightcorp, The Netherlands

Publication Chairs

Albert Ali Salah  
Boğaziçi University, Turkey

Robby T. Tan  
Yale-NUS College and National University of Singapore, Singapore

Video Chair

Mihir Jain  
University of Amsterdam, The Netherlands

Demo Chairs

John Schavemaker  
Twnkls, The Netherlands

Andy Bagdanov  
University of Florence, Italy

Social Media Chair

Efstratios Gavves  
University of Amsterdam, The Netherlands

Industrial Liaison Chairs

Ronald Poppe  
Utrecht University, The Netherlands

Peter de With  
Eindhoven University of Technology, The Netherlands

Conference Coordinator, Accommodation, and Finance

Conference Office

Martine de Wit  
University of Amsterdam, The Netherlands

Melanie Venverloo  
University of Amsterdam, The Netherlands

Niels Klein  
University of Amsterdam, The Netherlands
Workshop Organizers

W01 — Datasets and Performance Analysis in Early Vision
Michael Goesele  
Bernd Jähne  
Katrin Honauer  
Michael Waechter  
TU Darmstadt, Germany  
Heidelberg University, Germany  
Heidelberg University, Germany  
TU Darmstadt, Germany

W02 — Visual Analysis of Sketches
Yi-Zhe Song  
John Collomosse  
Metin Sezgin  
James Z. Wang  
Queen Mary University of London, UK  
University of Surrey, UK  
Koç University, Turkey  
The Pennsylvania State University, USA

W03 — Biological and Artificial Vision
Kandan Ramakrishnan  
Radoslaw M. Cichy  
Sennay Ghebreab  
H. Steven Scholte  
Arnold W.M. Smeulders  
University of Amsterdam, The Netherlands  
Free University Berlin, Germany  
University of Amsterdam, The Netherlands  
University of Amsterdam, The Netherlands  
University of Amsterdam, The Netherlands

W04 — Brave New Ideas For Motion Representations
Efstratios Gavves  
Basura Fernando  
Jan van Gemert  
University of Amsterdam, The Netherlands  
The Australian National University, Australia  
Delft University of Technology, The Netherlands

W05 — Joint ImageNet and MS COCO Visual Recognition Challenge
Wei Liu  
Genevieve Patterson  
M. Ronchi  
Yin Cui  
Tsung-Yi Lin  
Larry Zitnick  
Piotr Dollár  
Olga Russakovsky  
Jia Deng  
Fei–Fei Li  
Alexander C. Berg  
The University of North Carolina at Chapel Hill, USA  
Brown University, USA  
California Institute of Technology, USA  
Cornell Tech, USA  
Cornell Tech, USA  
Facebook AI Research, USA  
Facebook AI Research, USA  
Carnegie Mellon University, USA  
University of Michigan, USA  
Stanford University, USA  
The University of North Carolina at Chapel Hill, USA

W06 — Geometry Meets Deep Learning
Emanuele Rodolà  
Jonathan Masci  
Pierre Vandergheynst  
Università della Svizzera Italiana, Switzerland  
Università della Svizzera Italiana, Switzerland  
Ecole Polytechnique Fédérale de Lausanne, Switzerland
Sanja Fidler University of Toronto, Canada
Xiaowei Zhou University of Pennsylvania, USA
Kostas Daniilidis University of Pennsylvania, USA

W07 — Action and Anticipation for Visual Learning
Dinesh Jayaraman University of Texas at Austin, USA
Kristen Grauman University of Texas at Austin, USA
Sergey Levine University of Washington, USA

W08 — Computer Vision for Road Scene Understanding and Autonomous Driving
Jose Alvarez NICTA, Australia
Mathieu Salzmann Ecole Polytechnique Fédérale de Lausanne, Switzerland
Lars Petersson NICTA, Australia
Fredrik Kahl Chalmers University of Technology, Sweden
Bart Nabbe Faraday Future, USA

W09 — Challenge on Automatic Personality Analysis
Sergio Escalera Computer Vision Center (UAB) and University of Barcelona, Spain
Xavier Baró Universitat Oberta de Catalunya and Computer Vision Center (UAB), Spain
Isabelle Guyon Université Paris-Saclay, France, and ChaLearn, USA
Hugo Jair Escalante INAOE, Mexico
Víctor Ponce López Computer Vision Center (UAB) and University of Barcelona, Spain

W10 — BioImage Computing
Patrick Bouthemy Inria Research Institute, Switzerland
Fred Hamprecht Heidelberg University, Germany
Erik Meijering Erasmus University Medical Center, The Netherlands
Thierry Pécot Inria, France
Pietro Perona California Institute of Technology, USA
Carsten Rother TU Dresden, Germany

W11 — Benchmarking Multi-Target Tracking: MOTChallenge
Laura Leal-Taixé TU Munich, Germany
Anton Milan University of Adelaide, Australia
Konrad Schindler ETH Zürich, Switzerland
Daniel Cremers TU Munich, Germany
Ian Reid University of Adelaide, Australia
Stefan Roth TU Darmstadt, Germany
W12 — Assistive Computer Vision and Robotics

Giovanni Maria Farinella  University of Catania, Italy
Marco Leo  CNR – Institute of Applied Sciences and Intelligent Systems, Italy
Gerard G. Medioni  University of Southern California, USA
Mohan Trivedi  University of California, San Diego, USA

W13 — Transferring and Adapting Source Knowledge in Computer Vision

Wen Li  ETH Zürich, Switzerland
Tatiana Tommasi  University of North Carolina at Chapel Hill, USA
Francesco Orabona  Yahoo Research, NY, USA
David Vázquez  CVC and Universitat Autònoma de Barcelona, Spain
Antonio M. López  CVC and Universitat Autònoma de Barcelona, Spain
Jiaolong Xu  CVC and Universitat Autònoma de Barcelona, Spain
Hugo Larochelle  Twitter Cortex, USA

W14 — Recovering 6D Object Pose

Tae-Kyun Kim  Imperial College London, UK
Jiri Matas  Czech Technical University, Czech Republic
Vincent Lepetit  Technical University Graz, Germany
Carsten Rother  Technical University Dresden, Germany
Ales Leonardis  University of Birmingham, UK
Krzysztof Wallas  Poznan University of Technology, Poland
Carsten Steger  MVTec GmbH, Germany
Rigas Kouskouridas  Imperial College London, UK

W15 — Robust Reading

Dimosthenis Karatzas  CVC and Universitat Autònoma de Barcelona, Spain
Masakazu Iwamura  Osaka Prefecture University, Japan
Jiri Matas  Czech Technical University, Czech Republic
Pramod Sankar Kompalli  Flipkart.com, India
Faisal Shafait  National University of Sciences and Technology, Pakistan

W16 — 3D Face Alignment in the Wild and Challenge

Jeffrey Cohn  Carnegie Mellon University and University of Pittsburgh, USA
Laszlo Jeni  Carnegie Mellon University, USA
Nicu Sebe  University of Trento, Italy
Sergey Tulyakov  University of Trento, Italy
Lijun Yin  Binghamton University, USA
W17 — Ego-centric Perception, Interaction, and Computing
Giuseppe Serra
Rita Cucchiara
Walterio Mayol-Cuevas
Andreas Bulling
Dima Damen
University of Modena and Reggio Emilia, Italy
University of Modena and Reggio Emilia, Italy
University of Bristol, UK
Max Planck Institute for Informatics, Germany
University of Bristol, UK

W18 — Local Features: State of the Art, Open Problems, and Performance Evaluation
Jiri Matas
Krystian Mikolajczyk
Tinne Tuytelaars
Andrea Vedaldi
Vassileios Balntas
Karel Lenc
Czech Technical University, Czech Republic
Imperial College London, UK
KU Leuven, Belgium
University of Oxford, UK
University of Oxford, UK
University of Oxford, UK

W19 — Crowd Understanding
François Brémond
Vít Libal
Andrea Cavallaro
Tomas Pajdla
Petra Palatka
Jana Trojanova
Inria Sophia Antipolis, France
Honeywell ACS Global Labs Prague, Czech Republic
Queen Mary University of London, UK
Czech Technical University, Czech Republic
Neovision, Czech Republic
Honeywell ACS Global Labs Prague, Czech Republic

W20 — Video Segmentation
Thomas Brox
Katerina Fragkiadaki
Fabio Galasso
Fuxin Li
James M. Rehg
Bernt Schiele
Michael Ying Yang
University of Freiburg, Germany
Google Research, USA
OSRAM GmbH, Germany
Oregon State University, USA
Georgia Institute of Technology, USA
Max Planck Institute Informatics and Saarland University, Germany
University of Twente, The Netherlands

W21 — The Visual Object Tracking Challenge Workshop
Matej Kristan
Aleš Leonardis
Jiri Matas
Michael Felsberg
Roman Pflugfelder
University of Ljubljana, Slovenia
University of Birmingham, UK
Czech Technical University in Prague, Czech Republic
Linköping University, Sweden
Austrian Institute of Technology, Austria
### W22 — Web-Scale Vision and Social Media

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lamberto Ballan</td>
<td>Stanford University, USA</td>
<td>USA</td>
</tr>
<tr>
<td>Marco Bertini</td>
<td>University of Florence, Italy</td>
<td>Italy</td>
</tr>
<tr>
<td>Thomas Mensink</td>
<td>University of Amsterdam, The Netherlands</td>
<td></td>
</tr>
</tbody>
</table>

### W23 — Computer Vision for Audiovisual Media

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jean-Charles Bazin</td>
<td>Disney Research, USA</td>
<td>USA</td>
</tr>
<tr>
<td>Zhengyou Zhang</td>
<td>Microsoft Research, USA</td>
<td>USA</td>
</tr>
<tr>
<td>Wilmot Li</td>
<td>Adobe Research, USA</td>
<td>USA</td>
</tr>
</tbody>
</table>

### W24 — Computer Vision for Art Analysis

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joao Paulo Costeira</td>
<td>Instituto Superior Técnico, Portugal</td>
<td>Portugal</td>
</tr>
<tr>
<td>Gustavo Carneiro</td>
<td>University of Adelaide, Australia</td>
<td>Australia</td>
</tr>
<tr>
<td>Alessio Del Bue</td>
<td>Istituto Italiano di Tecnologia (IIT), Italy</td>
<td>Italy</td>
</tr>
<tr>
<td>Ahmed Elgammal</td>
<td>Rutgers University, USA</td>
<td>USA</td>
</tr>
<tr>
<td>Peter Hall</td>
<td>University of Bath, UK</td>
<td>UK</td>
</tr>
<tr>
<td>Ann-Sophie Lehmann</td>
<td>University of Groningen, The Netherlands</td>
<td>Netherlands</td>
</tr>
<tr>
<td>Hans Brandhorst</td>
<td>Iconclass and Arkyves, The Netherlands</td>
<td></td>
</tr>
<tr>
<td>Emily L. Spratt</td>
<td>Princeton University, USA</td>
<td>USA</td>
</tr>
</tbody>
</table>

### W25 — Virtual/Augmented Reality for Visual Artificial Intelligence

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antonio M. López</td>
<td>CVC and Universitat Autònoma de Barcelona, Spain</td>
<td>Spain</td>
</tr>
<tr>
<td>Adrien Gaidon</td>
<td>Xerox Research Center Europe (XRCE), France</td>
<td>France</td>
</tr>
<tr>
<td>German Ros</td>
<td>CVC and Universitat Autònoma de Barcelona, Spain</td>
<td>Spain</td>
</tr>
<tr>
<td>Eleonora Vig</td>
<td>German Aerospace Center (DLR), Germany</td>
<td>Germany</td>
</tr>
<tr>
<td>David Vázquez</td>
<td>CVC and Universitat Autònoma de Barcelona, Spain</td>
<td>Spain</td>
</tr>
<tr>
<td>Hao Su</td>
<td>Stanford University, USA</td>
<td>USA</td>
</tr>
<tr>
<td>Florent Perronnin</td>
<td>Facebook AI Research (FAIR), France</td>
<td>France</td>
</tr>
</tbody>
</table>

### W26 — Joint Workshop on Storytelling with Images and Videos and Large-Scale Movie Description and Understanding Challenge

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gunhee Kim</td>
<td>Seoul National University, South Korea</td>
<td>Korea</td>
</tr>
<tr>
<td>Leonid Sigal</td>
<td>Disney Research Pittsburgh, USA</td>
<td>USA</td>
</tr>
<tr>
<td>Kristen Grauman</td>
<td>University of Texas at Austin, USA</td>
<td>USA</td>
</tr>
<tr>
<td>Tamara Berg</td>
<td>University of North Carolina at Chapel Hill, USA</td>
<td>USA</td>
</tr>
<tr>
<td>Anna Rohrbach</td>
<td>Max Planck Institute for Informatics, Germany</td>
<td>Germany</td>
</tr>
<tr>
<td>Atousa Torabi</td>
<td>Disney Research Pittsburgh, USA</td>
<td>USA</td>
</tr>
<tr>
<td>Tegan Maharaj</td>
<td>École Polytechnique de Montréal, Canada</td>
<td>Canada</td>
</tr>
<tr>
<td>Marcus Rohrbach</td>
<td>University of California, Berkeley, USA</td>
<td>USA</td>
</tr>
<tr>
<td>Christopher Pal</td>
<td>École Polytechnique de Montréal, Canada</td>
<td>Canada</td>
</tr>
<tr>
<td>Aaron Courville</td>
<td>Université de Montréal, Canada</td>
<td>Canada</td>
</tr>
<tr>
<td>Bernt Schiele</td>
<td>Max Planck Institute for Informatics, Germany</td>
<td>Germany</td>
</tr>
</tbody>
</table>
Contents – Part II

W08 - Computer Vision for Road Scene Understanding and Autonomous Driving (Continued)

Real-Time Semantic Segmentation with Label Propagation ......................... 3
   Rasha Sheikh, Martin Garbade, and Juergen Gall

W11 - Benchmarking Multi-target Tracking: MOTChallenge

Performance Measures and a Data Set for Multi-target,
Multi-camera Tracking .............................................................................. 17
   Ergys Ristani, Francesco Solera, Roger Zou, Rita Cucchiara,
   and Carlo Tomasi

POI: Multiple Object Tracking with High Performance Detection
   and Appearance Feature ........................................................................... 36
   Fengwei Yu, Wenbo Li, Quanquan Li, Yu Liu, Xiaohua Shi,
   and Junjie Yan

Long-Term Time-Sensitive Costs for CRF-Based Tracking by Detection ...... 43
   Nam Le, Alexander Heili, and Jean-Marc Odobez

Tracking Multiple Persons Based on a Variational Bayesian Model .......... 52
   Yutong Ban, Sileye Ba, Xavier Alameda-Pineda, and Radu Horaud

Multi-class Multi-object Tracking Using Changing Point Detection ........ 68
   Byungjae Lee, Enkhbayar Erdenee, Songguo Jin, Mi Young Nam,
   Young Giu Jung, and Phill Kyu Rhee

Online Multi-target Tracking with Strong and Weak Detections ............ 84
   Ricardo Sanchez-Matilla, Fabio Poiesi, and Andrea Cavallaro

Multi-person Tracking by Multicut and Deep Matching ......................... 100
   Siyu Tang, Bjoern Andres, Mykhaylo Andriluka, and Bernt Schiele

W12 - Assistive Computer Vision and Robotics

Visual and Human-Interpretable Feedback for Assisting Physical Activity ... 115
   Michel Antunes, Renato Baptista, Girum Demisse, Djamila Aouada,
   and Björn Ottersten
Mobile Mapping and Visualization of Indoor Structures to Simplify Scene Understanding and Location Awareness ........................................ 130
   Giovanni Pintore, Fabio Ganovelli, Enrico Gobbetti, and Roberto Scopigno

Automatic Video Captioning via Multi-channel Sequential Encoding .......... 146
   Chenyang Zhang and Yingli Tian

Validation of Automated Mobility Assessment Using a Single 3D Sensor ... 162
   Jiun-Yu Kao, Minh Nguyen, Luciano Nocera, Cyrus Shahabi, Antonio Ortega, Carolee Winstein, Ibrahim Sorkhoh, Yu-chen Chung, Yi-an Chen, and Helen Bacon

Deep Eye-CU (DECU): Summarization of Patient Motion in the ICU ......... 178
   Carlos Torres, Jeffrey C. Fried, Kenneth Rose, and B.S. Manjunath

Fall Detection Based on Depth-Data in Practice .............................. 195
   Christopher Pramerdorfer, Rainer Planinc, Mark Van Loock, David Fankhauser, Martin Kampel, and Michael Brandstötter

A Real-Time Vehicular Vision System to Seamlessly See-Through Cars ...... 209
   Francois Rameau, Hyowon Ha, Kyungdon Joo, Jinsoo Choi, and InSo Kweon

Solving Rendering Issues in Realistic 3D Immersion
for Visual Rehabilitation ............................................................ 223
   Tristan Carrier-Baudouin, Claude Chapdelaine, Marc Lalonde, Philippe Quinn, and Samuel Foucher

Human-Drone-Interaction: A Case Study to Investigate the Relation
Between Autonomy and User Experience ...................................... 238
   Patrick Ferdinand Christ, Florian Lachner, Axel Hösl, Bjoern Menze, Klaus Diepold, and Andreas Butz

Feasibility Analysis of Eye Typing with a Standard Webcam ................. 254
   Yi Liu, Bu Sung Lee, Andrzej Sluze, Deepu Rajan, and Martin Mckeown

A Technological Framework to Support Standardized Protocols
for the Diagnosis and Assessment of ASD .................................. 269
   Marco Leo, Marco Del Coco, Pierluigi Carcagni, Pier Luigi Mazzeo, Paolo Spagnolo, and Cosimo Distante

Combining Human Body Shape and Pose Estimation for Robust Upper
Body Tracking Using a Depth Sensor ........................................... 285
   Thomas Probst, Andrea Fossati, and Luc Van Gool
Multi-level Net: A Visual Saliency Prediction Model. ........................................ 302
Marcella Cornia, Lorenzo Baraldi, Giuseppe Serra, and Rita Cucchiara

Learning and Detecting Objects with a Mobile Robot to Assist Older Adults in Their Homes ................................................................. 316
Markus Vincze, Markus Bajones, Markus Suchi, Daniel Wolf, Astrid Weiss, David Fischinger, and Paloma da la Puente

An Interactive Multimedia System for Treating Autism Spectrum Disorder. . 331
Massimo Magrini, Ovidio Salvetti, Andrea Carboni, and Olivia Curzio

Vision-Based SLAM Navigation for Vibro-Tactile Human-Centered Indoor Guidance ................................................................. 343
Thomas Gulde, Silke Kärcher, and Cristóbal Curio

Perfect Accuracy with Human-in-the-Loop Object Detection. .................. 360
Rorry Brenner, Jay Priyadarshi, and Laurent Itti

Using Computer Vision to See ................................................................. 375
Bogdan Mocanu, Ruxandra Tapu, and Titus Zaharia

Brazilian Sign Language Recognition Using Kinect .................................. 391
José Elías Yauri Vidalón and José Mario De Martino

Human Interaction Prediction Using Deep Temporal Features .................. 403
Qiuhong Ke, Mohammed Bennamoun, Senjian An, Farid Boussaid, and Ferdous Sohel

Human Joint Angle Estimation and Gesture Recognition for Assistive Robotic Vision ................................................................. 415
Alp Guler, Nikolaos Kardaris, Siddhartha Chandra, Vassilis Pitsikalis, Christian Werner, Klaus Hauer, Costas Tzafestas, Petros Maragos, and Iasonas Kokkinos

A 3D Human Posture Approach for Activity Recognition Based on Depth Camera ................................................................. 432
Alessandro Manzi, Filippo Cavallo, and Paolo Dario

ISANA: Wearable Context-Aware Indoor Assistive Navigation with Obstacle Avoidance for the Blind ................................. 448
Bing Li, J. Pablo Muñoz, Xuejian Rong, Jizhong Xiao, Yingli Tian, and Aries Arditi

An Integrated Framework for 24-hours Fire Detection .......................... 463
Jongwon Choi and Jin Young Choi
Smart Toothbrushes: Inertial Measurement Sensors Fusion with Visual Tracking .................................................. 480
Marco Marcon, Augusto Sarti, and Stefano Tubaro

Evaluation of Infants with Spinal Muscular Atrophy Type-I Using Convolutional Neural Networks .......................... 495
Bilge Soran, Linda Lowes, and Katherine M. Steele

W16 - 3D Face Alignment in the Wild and Challenge

The First 3D Face Alignment in the Wild (3DFAW) Challenge ........... 511
László A. Jeni, Sergey Tulyakov, Lijun Yin, Nicu Sebe, and Jeffrey F. Cohn

3D Face Alignment Without Correspondences .......................... 521
Zsolt Sánta and Zoltan Kato

Bi-Level Multi-column Convolutional Neural Networks for Facial Landmark Point Detection ............................. 536
Yanyu Xu and Shenghua Gao

Fully Automated and Highly Accurate Dense Correspondence for Facial Surfaces .............................................. 552
Carl Martin Grewe and Stefan Zachow

Joint Face Detection and Alignment with a Deformable Hough Transform Model ................................................. 569
John McDonagh and Georgios Tzimiropoulos

3D Face Alignment in the Wild: A Landmark-Free, Nose-Based Approach ... 581
Flávio H. de Bittencourt Zavan, Antônio C.P. Nascimento, Luan P. e Silva, Olga R.P. Bellon, and Luciano Silva

Fast and Precise Face Alignment and 3D Shape Reconstruction from a Single 2D Image ............................................ 590
Ruiqi Zhao, Yan Wang, C. Fabian Benitez-Quiroz, Yaojie Liu, and Aleix M. Martinez

Shape Augmented Regression for 3D Face Alignment .................. 604
Chao Gou, Yue Wu, Fei-Yue Wang, and Qiang Ji

Two-Stage Convolutional Part Heatmap Regression for the 1st 3D Face Alignment in the Wild (3DFAW) Challenge .......... 616
Adrian Bulat and Georgios Tzimiropoulos
## W19 - Crowd Understanding

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi-person Pose Estimation with Local Joint-to-Person Associations</td>
<td>627</td>
</tr>
<tr>
<td><em>Umar Iqbal and Juergen Gall</em></td>
<td></td>
</tr>
<tr>
<td>Density-Aware Pedestrian Proposal Networks for Robust People Detection in Crowded Scenes</td>
<td>643</td>
</tr>
<tr>
<td><em>Sangdoo Yun, Kimin Yun, Jongwon Choi, and Jin Young Choi</em></td>
<td></td>
</tr>
<tr>
<td>People Counting in Videos by Fusing Temporal Cues from Spatial Context-Aware Convolutional Neural Networks</td>
<td>655</td>
</tr>
<tr>
<td><em>Panos Sourtzinos, Sergio A. Velastin, Miguel Jara, Pablo Zegers, and Dimitrios Makris</em></td>
<td></td>
</tr>
<tr>
<td>Abnormal Crowd Behavior Detection Based on Gaussian Mixture Model.</td>
<td>668</td>
</tr>
<tr>
<td><em>Oscar Ernesto Rojas and Clesio Luis Tozzi</em></td>
<td></td>
</tr>
<tr>
<td>Unsupervised Deep Domain Adaptation for Pedestrian Detection</td>
<td>676</td>
</tr>
<tr>
<td><em>Lihang Liu, Weiyao Lin, Lisheng Wu, Yong Yu, and Michael Ying Yang</em></td>
<td></td>
</tr>
<tr>
<td>Pixel Level Tracking of Multiple Targets in Crowded Environments</td>
<td>692</td>
</tr>
<tr>
<td><em>Mohammadreza Babaee, Yue You, and Gerhard Rigoll</em></td>
<td></td>
</tr>
<tr>
<td>LCrowdV: Generating Labeled Videos for Simulation-Based Crowd Behavior Learning</td>
<td>709</td>
</tr>
<tr>
<td><em>Ernest Cheung, Tsan Kwong Wong, Aniket Bera, Xiaogang Wang, and Dinesh Manocha</em></td>
<td></td>
</tr>
<tr>
<td>Anomaly Detection and Activity Perception Using Covariance Descriptor for Trajectories</td>
<td>728</td>
</tr>
<tr>
<td><em>Hamza Ergezer and Kemal Leblebicioğlu</em></td>
<td></td>
</tr>
<tr>
<td>Automatic Calibration of Stationary Surveillance Cameras in the Wild.</td>
<td>743</td>
</tr>
<tr>
<td><em>Guido M.Y.E. Brouwers, Matthijs H. Zwemer, Rob G.J. Wijnhoven, and Peter H.N. de With</em></td>
<td></td>
</tr>
<tr>
<td>Data-Driven Motion Pattern Segmentation in a Crowded Environments</td>
<td>760</td>
</tr>
<tr>
<td><em>Jana Trojanová, Karel Křehnáč, and François Brémont</em></td>
<td></td>
</tr>
</tbody>
</table>
W21 - The Visual Object Tracking Challenge Workshop

The Visual Object Tracking VOT2016 Challenge Results .......................... 777

Matej Kristan, Aleš Leonardis, Jiří Matas, Michael Felsberg,
Roman Pflugfelder, Luka Čehovin, Tomáš Vojíř, Gustav Háger,
Alan Lukežič, Gustavo Fernández, Abhinav Gupta, Alfredo Petrosino,
Alireza Memarmoghadam, Alvaro García-Martin, Andrés Solís Montero,
Andrea Vedaldi, Andreas Robinson, Andy J. Ma, Anton Varfolomieiev,
Aydin Alatan, Aykut Erdem, Bernard Ghanem, Bin Liu, Bohyung Han,
Brais Martinez, Chang-Ming Chang, Changsheng Xu, Chong Sun,
Daijin Kim, Dapeng Chen, Dawei Du, Deepak Mishra, Dit-Yan Yeung,
Erhan Gundogdu, Erkut Erdem, Fahad Khan, Fatih Porikli, Fei Zhao,
Filiz Bunyak, Francesco Battistone, Gao Zhu, Giorgio Roffo,
Gorthi R.K. Sai Subrahmanyam, Guilherme Bastos, Guna Seetharaman,
Henry Medeiros, Hongdong Li, Honggang Qi, Horst Bischof,
Horst Possegger, Huchuan Lu, Hyemin Lee, Hyeonseob Nam,
Hyung Jin Chang, Isabel Drummond, Jack Valmadre, Jae-chan Jeong,
Jae-il Cho, Jae-Yeong Lee, Jianke Zhu, Jiayi Feng, Jin Gao,
Jin Young Choi, Jingjing Xiao, Ji-Wan Kim, Jiyeoup Jeong,
João F. Henriques, Jochen Lang, Jongwon Choi, Jose M. Martinez,
Junliang Xing, Junyu Gao, Kannappan Palaniappan, Karel Lebeda,
Ke Gao, Krystian Mikolajczyk, Lei Qin, Lijun Wang, Longxin Wei,
Luca Bertinetto, Madan Kumar Rapuru, Mahdieh Poostchi,
Mario Maresca, Martin Danelljan, Matthias Mueller, Mengdan Zhang,
Michael Arens, Michel Valstar, Ming Tang, Mooyeol Baek,
Muhammad Haris Khan, Naiyan Wang, Nana Fan, Noor Al-Shakarji,
Ondrej Miksik, Osman Akin, Payman Moallem, Pedro Senna,
Philip H.S. Torr, Pong C. Yuen, Qingming Huang, Rafael Martin-Nieto,
Rengarajan Pelapur, Richard Bowden, Robert Laganière, Rustam Stolkin,
Ryan Walsh, Sebastian B. Krah, Shengkun Li, Shengping Zhang,
Shizeng Yao, Simon Hadfield, Simone Melzi, Siwei Lyu, Siyi Li,
Stefan Becker, Stuart Golodetz, Sumithra Kakanuru, Sunglok Choi,
Tao Hu, Thomas Mauthner, Tianzhu Zhang, Tony Pridmore,
Vincenzo Santopietro, Weiming Hu, Wenbo Li, Wolfgang Hübner,
Xiangyuan Lan, Xiaomeng Wang, Xin Li, Yang Li, Yiannis Demiris,
Yifan Wang, Yuankai Qi, Zejian Yuan, Zexiong Cai, Zhan Xu, Zhenyu He,
and Zhizhen Chi
The Thermal Infrared Visual Object Tracking VOT-TIR2016 Challenge Results ......................................................... 824


Fully-Convolutional Siamese Networks for Object Tracking ............... 850

Luca Bertinetto, Jack Valmadre, João F. Henriques, Andrea Vedaldi, and Philip H.S. Torr

W23 - Computer Vision for Audio–Visual Media

Speech-Driven Facial Animation Using Manifold Relevance Determination ... 869

Samia Dawood, Yulia Hicks, and David Marshall

GeThR-Net: A Generalized Temporally Hybrid Recurrent Neural Network for Multimodal Information Fusion ................................. 883

Ankit Gandhi, Arjun Sharma, Arijit Biswas, and Om Deshmukh

Suggesting Sounds for Images from Video Collections ....................... 900

Matthias Solèr, Jean-Charles Bazin, Oliver Wang, Andreas Krause, and Alexander Sorkine-Hornung

Author Index ....................................................................... 919