Guest Editorial: The Computational Face

Sergio Escalera, Xavier Baró, Isabelle Guyon, Hugo Escalante, Georgios Tzimiropoulos, Michel Valstar, Maja Pantic, Jeffrey Cohn, Takeo Kanade

Guest Editorial: The Computational Face

Sergio Escalera, Xavier Baró, Isabelle Guyon, Hugo Jair Escalante, Georgios Tzimiropoulos, Michel Valstar, Maja Pantic, Jeffrey Cohn, and Takeo Kanade

1 Introduction

Automated face analysis has received special attention from the Computer Vision and Pattern Recognition communities. Research progress often gives the impression that problems such as face recognition and face detection are solved, at least for some scenarios. Several aspects of face analysis remain open problems, including the implementation of large scale face recognition/detection methods for *in the wild* images, emotion recognition, micro-expression analysis, and others. The community keeps making rapid progress on these topics, with continual improvement of current methods and creation of new ones that push the state-of-the-art. Applications are countless, including security and video surveillance, human computer/robot interaction, communication, entertainment, and commerce, while having an important social impact in assistive technologies for education and health.

The importance of face analysis, together with the vast amount of work on the subject and the latest developments in the field, motivated us to organize a special section on this theme. The scope of the compilation comprises all aspects of face analysis from a computer vision perspective. Including, but not limited to: recognition, detection, alignment, reconstruction of faces, pose estimation of faces, gaze analysis, age, emotion, gender, and facial attributes estimation, and applications among others. The special section grew out of a technical meeting of the BMVA in October 2015, from which this SI derived its name, and a workshop on face analysis collocated with CVPR2016.[1] In addition, the section was related to two academic challenges [1]. The first one on apparent age estimation, where two rounds were organized in events collocated with ICCV2015 [2] and CVPR2016 [3]. A second associated challenge on accessories, smile, and gender classification was also collocated with CVPR2016 [3]. Participants of such events were encouraged to submit articles to the special section. However, the call for papers was open to the public. In the next section we briefly summarize the contributions of the articles included in the section, after which we present a discussion on the identified trends and research opportunities on visual facial analysis.

1. http://chalearnlap.cvc.uab.es/workshop/10/description/

- S. Escalera is with the University of Barcelona and Computer Vision Center, Barcelona 08007, Spain. E-mail: sergio@maia.ub.es.
- X. Baró is with the Universitat Oberta de Catalunya and Computer Vision Center, Barcelona 08018, Spain. E-mail: xbaro@uoc.edu.
- I. Guyon is with UPSud/INRIA Université Paris-Saclay, Saint-Aubin 91190, France, and also with the Président of ChaLearn, Berkeley, CA 94708. E-mail: guyon@chalearn.org.
- H. J. Escalante is with Instituto Nacional de Astrofísica, Óptica y Eectrónica, Puebla 72840, Mexico. E-mail: hugojair@inaoep.mx.
- G. Tzimiropoulos and M. Valstar are with University of Nottingham, Nottingham NG7 2RD, United Kingdom. E-mail: {yorgos.tzimiropoulos, michel.valstar}@nottingham.ac.uk.
- M. Pantic is with Imperial College, London SW7 2AZ, United Kingdom. E-mail: m.pantic@imperial.ac.uk.
- J. Cohn is with the University of Pittsburgh, Pittsburgh, PA 15260. E-mail: jeffcohn@cs.cmu.edu.
- T. Kanade is with Carnegie Mellon University, Pittsburgh, PA 15213. E-mail: Takeo.Kanade@cs.cmu.edu.

For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TPAMI.2018.2869610

2 The Computational Face Section

This special section is formed by 13 articles of outstanding quality that together comprise a snapshot of the state of the art on the analysis of faces from visual information. Table 1 summarizes the main characteristics of the accepted papers.

The contributions focused on a diverse set of topics, techniques, and applications. Estimation of facial markers (e.g., facial landmarks, action units) was the most approached task in papers within the special issue [4], [5], [6], [7]. In fact, most papers in the issue used facial landmarks in one way or another as part of their methodology. This evidences the relevance of this task for the community. Liu et al. proposed an heterogeneous two stream deep neural net to process appearance/spatial and temporal information independently at the beginning and combining the model at a late stage [4]. A similar methodology was adopted by Wang et al. who combined a super resolution CNN and a recurrent neural net to iteratively refine landmarks in a cascade regression approach [6]. A novel deep architecture was proposed by Li et al. in which attention maps and cropping layers were included in a CNN for better detection of facial action units [7]. In contrast to previous work, Wang et al. described an iterative methodology based on deformable parts models for person-specific landmark detection. Their approach uses no learned features yet obtains outstanding performance [5].

Another problem well represented in the papers in this special section is that of facial attribute estimation [8], [9], [10]. Facial attributes include age, gender, and ethnicity. They are helpful for numerous applications (e.g., security, human-computer-interaction, and advertising). For this reason, estimation of facial attributes is another active research topic in the computer vision and pattern recognition communities. Han et al. introduces a multi-task, deep learning architecture for the simultaneous estimation of multiple facial attributes (including gender, ethnicity, age) [8]. The deep

TABLE 1
Overview of Articles in the Special Section on the Computational Face

Ref.	Task / Model	Features	Model	Dataset
[4]	Facial landmark detection in videos	Spatial/appearance and temporal features are learned form face videos via a two stream deep learning model.	Two stream deep neural network. A convolutional neural network is used to process spatial/appearance information, whereas an autoencoder process the temporal information.	300-VW and Talking Face
[5]	Facial landmark detection in videos	Facial landmarks, SIFT/HOG features	A part deformable model is obtained via multivariate Gaussians and SVMs from each person. Then a generic landmark detection method is applied (Gauss-Newton DPM). Both steps are iterated trying to improve the initial landmark detection.	300 VW, FTOVM (Own), Cats (Own)
[6]	Facial landmark detection	Features are learned with a convolutional neural network (super resolution CNN, SRCNN) from patches associated to landmarks..	SRCNN, for feature extraction coupled with a recurrent neural network with a gate recurrent unit, the latter used for iteratively refining landmarks in cascade regression	300 VW and Talking Face, HELEN, LFPW
[7]	Facial action unit detection	Facial landmarks, features learned with convolutional neural networks (EAC-Net)	An enhancing and cropping network (EAC-Net): a fine-tuned CNN with enhancing and cropping layers. Enhancing layers are formed by attention maps obtained my facial landmark information. Cropping layers crop potentially relevant image regions for AUC detection.	DISFA, BP4D
[8]	Facial attribute estimation (age, gender, ethnicity, etc.)	Features are learned with a multi task convolutional neural network from face images	Multitask deep neural network. The initial layers are shared across tasks, there are specialized layers for each facial attribute.	MORP-II, CelebA, LFWA, LAPAage2015, and LFW+ (Own)
[9]	Apparent age estimation	Features are learned with a convolutional neural network. In addition, a group coding, which can be considered as a meta-feature is obtained	Convolutional neural networks, group coding for age labels, stacking model.	MORPH-II, CACD, ChaLearnLaP Apparent age estimation data set
[10]	kinship verification and family classification	Features are learned with a convolutional neural networks	Semi-supervised clustering algorithm for generating the dataset, and pretrained CNN fine tuned with the FIW dataset for the recognition tasks.	FIW (own)
[11]	3D facial shape reconstruction	Facial landmarks	A model of facial shape combining identity and expression with an in-the-wild texture model for a 3D morphable model.	KF-ITW (own)
[12]	Head pose estimation & tracking	RGB-D video, facial landmarks	An hybrid model comprising a 3D morphable facial model automatically adapted online from a collection of samples and an online reconstruction 3D head model based on a variant of KinectFusion	BIWI, UBImpressed
[13]	Facial expression synthesis and Facial age progression	Facial landmarks, pixels	Robust joint and individual variance explained (RJIVE). RJIVE is used to discover the joint. and individual components of variation. These are used to reconstruct images.	Synthetic images, MPIE, CK+, ITW (Own)
[14]	Detection of modes of variation	Spherical light, shape, identity coefficients, Facial (ear) landmarks	A general multilinear method for tensor decomposition that does not require of labels. Extensions of the method for unconstrained data (sparsity and low-rank) are described as well.	Synthetic 3D faces (own), MultiPIE, HELEN, Ears in the wild (own)
[15]	Gaze redirection	Location and appearance features derived from pixel intensities/ positions in eye regions. Eyes are detected with standard face alignment methods.	Randomized decision trees (*eye flow forests*) and deep neural networks are used to predict gaze redirection angles, that are used to synthesize images with redirected gaze.	Gaze redirection data set (Own)
[16]	Estimation and tracking of visual focus of attention (VFOA) and gaze direction	Head position / orientation (VICON), target location	A dynamic Bayesian model using as hidden variables to gaze direction and VFOA	Vernissage dataset

model shares initial layers common to all tasks and incorporates task-specific layers at the top of the model. Tan et al. focuses on a single attribute: apparent age, and proposes another novel model based on deep learning [9]. A distinctive feature of their model is that it performs a grouping of ages that is later used in a stacking procedure prior to making age predictions. Robinson et al. address a related problem and release the largest dataset to date for kinship relationships analysis [10]. Through a semi-supervised methodology the authors collected information from 1,001 families that include 10,000 different persons. Baseline performance for kinship relation and family classification are reported as well.

Another task vastly represented in the computational face special issue is that of head pose estimation [11], [12], [13], [14]. Booth et al. introduce a new model for facial shape reconstruction in the wild [11]. They combine identity and expression with an in-the-wild texture model that is coupled with a 3D morphable model (3DMM). In addition, they release a new data set for facial shape reconstruction in the wild. A similar approach that combines a 3DMM adapted online from samples and a 3D head model similar to the one used for KinectFusion is proposed by Yu et al. [12]. The combination of models proves to be helpful for head pose estimation and tracking when people is not directly in front the camera. Related to pose estimation, two methodologies for modeling components of faces in image reconstruction were included in this issue [13], [14]. Sagonas et al. aim at accurately estimating the components of faces via an improvement to the joint individual variance explained (JIVE) approach [13]. The proposed method can be applied to more than 2 data sets and is robust to certain types of noise. The authors evaluate their method in facial expression synthesis and facial age progression. Similarly, Wang et al. introduce an unsupervised multi-linear method for tensor decomposition with the goal of detecting the modes of variation in facial imagery [14]. It is quite interesting that none of the previously described articles made direct use of deep learning models or learned features.

Finally, two contributions on gaze analysis are included in the special issue [15], [16]. Kononenko et al. propose a novel methodology that combines deep neural networks with randomized decision forests for redirecting gaze in video conferencing systems [15]. Their approach infers the degree of deviation in gaze. Synthetic corrected gaze images then are stitched to the original ones from these predictions. Similarly, Massé et al. introduce a dynamic Bayesian model that estimates and tracks visual focus of attention (VFOA) in video conferences. VFOA and gaze direction are treated as hidden variables of the probabilistic graphical model that takes as input position and orientations of head.

3 Discussion

The special issue is a compilation of cutting-edge research on the visual analysis of facial information. In the following, we draw some conclusions based on the papers in this special issue.

Deep learning is consolidated as the representation learning methodology of choice in recognition tasks, in particular. Most articles on facial markers and facial attribute estimation adopted a deep learning methodology for feature learning. A deep learning model also was used most often for recognition problems. However, it is notable that there are still tasks that do not succumb to deep learning. Methods for head pose estimation and gaze analysis bypass them. Lack of sufficiently large databases may be a contributing factor. There may be a research opportunity in the development of large databases for head pose and gaze estimation and the design of deep models to detect them.

We consider that the efficient generation of large scale resources for alternative tasks is another promising research direction. In this context, several articles in this special issue contributed with new evaluation resources. Most notably, a large scale dataset on kinship analysis and a couple of datasets for pose estimation and tracking.

All papers in this special section focused on *in the wild* imagery. This suggests the community is focusing on realistic scenarios, which will increase the impact of forthcoming methodologies. Still, we think that there are several levels of *"wildness"* when talking of data and scenarios, we think the difficulty level is somewhere in the middle of what can be expected, and in the forthcoming years extreme application scenarios will emerge. An interesting question is whether we will continue to see general purpose solutions or whether the increased difficulty levels will drive scenario-specific solutions.

Finally, in terms of approached tasks, most papers targeted popular problems that now are quite relevant. Other papers address novel tasks like synthesis of facial expressions, age progression and gaze redirection. We think novel tasks will emerge motivated by the overall progress in computer vision, machine learning and pattern recognition. This special issue presents the latest progress on the computational analysis of facial imagery.

Acknowledgments

This project has been partially supported by the Spanish projects TIN2015-66951-C2-2-R and TIN2016-74946-P (MINECO/FEDER, UE) and CERCA Programme/Generalitat de Catalunya and by INAOE, CONACyT-Mexico under grant 241306. We thank ChaLearn Looking at People sponsors for their support, including Microsoft Research, Google, NVIDIA Coorporation, Amazon, Facebook, and Disney Research.

References

[1] S. Escalera, X. Baró, H. J. Escalante, and I. Guyon, "Chalearn looking at people: A review of events and resources," in *Proc. Int. Joint Conf. Neural Netw.*, 2017, pp. 1594–1601. [Online]. Available: https://doi.org/10.1109/IJCNN.2017.7966041

[2] S. Escalera, J. Fabian, P. Pardo, X. Baró, J. Gonzàlez, H. J. Escalante, D. Misevic, U. Steiner, and I. Guyon, "Chalearn looking at people 2015: Apparent age and cultural event recognition datasets and results," in *Proc. IEEE Int. Conf. Comput. Vis. Workshop*, Dec. 2015, pp. 243–251.

[3] S. Escalera, M. Torres, B. Martinez, X. Baró, H. J. Escalante, I. Guyon, G. Tzimiropoulos, C. Corneou, M. Oliu, M. A. Bagheri, and M. Valstar, "ChaLearn looking at people and faces of the world: Face analysis workshop and challenge 2016," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, 2016, pp. 706–713. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01381152

[4] H. Liu, J. Lu, J. Feng, and J. Zhou, "Two-stream transformer networks for video-based face alignment," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018, doi: 10.1109/TPAMI.2017.2734779.

[5] G. G. Chrysos and S. Zafeiriou, "PD^{2}T: Person-specific detection, deformable tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018, doi: 10.1109/TPAMI.2017.2769654.
[6] W. Wang, S. Tulyakov, and N. Sebe, "Recurrent convolutional shape regression," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018, doi: 10.1106/TPAMI.2018.2810881.
[7] W. Li, F. Abtahi, Z. Zhu, and L. Yin, "EAC-NET: Deep nets with enhancing and cropping for facial action unit detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018, doi: 10.11069/TPAMI.2018.2791608.
[8] H. Han, A. K. Jain, S. Shan, and X. Chen, "Heterogeneous face attribute estimation: A deep multi- task learning approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018, doi: 10.1109/TPAMI.2017.2738004.
[9] Z. Tan, J. Wan, Z. Lei, R. Zhi, G. Guo, and S. Z. Li, "Efficient group-n encoding and decoding for facial age estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018, doi: 10.1109/TPAMI.2017.2738004.
[10] J. P. Robinson, M. Shao, Y. Wu, H. Liu, T. Gillis, and Y. Fu, "Visual kinship recognition of families in the wild," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018, doi: 10.1109/TPAMI.2018.2826849.
[11] J. Booth, A. Roussos, G. Trigeorgis, E. Antonakos, S. Ploumpis, Y. Panagakis, and S. Zafeiriou, "3D reconstruction of "in-the-wild" faces in images and videos," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018, doi: 10.1109/TPAMI.2018.2832138.
[12] Y. Yu, K. A. F. Mora, and J.-M. Odobez, "HeadFusion: 360 head pose tracking combining 3D morphable model and 3D reconstruction," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018, doi: 10.1109/TPAMI.2018.2841403.
[13] C. Sagonas, Y. Panagakis, and S. Zafeiriou, "Recovering joint and individual components in visual data," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018, doi: 10.11069/TPAMI.2017.2784421.
[14] M. Wang, Y. Panagakis, P. Snape, and S. Zafeiriou, "Disentangling the modes of variation in unlabelled data," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018, doi: 10.1109/TPAMI.2017.2783940.
[15] D. Kononenko, Y. Ganin, D. Sungatullina, and V. Lempitsky, "Photorealistic monocular gaze redirection using machine learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018, doi: 10.1109/TPAMI.2017.2737423.
[16] B. Massé, S. Ba, and R. Horaud, "Tracking gaze and visual focus of attention of people involved in social interaction," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018, doi: 10.1109/TPAMI.2017.2785819.

Sergio Escalera received the PhD degree in multi-class visual categorization systems at Computer Vision Center, UAB. He received the 2008 best Thesis award. He leads the Human Pose Recovery and Behavior Analysis Group at UB and CVC. He is an associate professor with the Department of Mathematics and Informatics, Universitat de Barcelona. He is also a member of the Computer Vision Center at Campus UAB. He is vice-president of ChaLearn Challenges in Machine Learning and chair of IAPR TC-12: Multimedia and visual information systems. His research interests include, between others, statistical pattern recognition, affective computing, visual object recognition, and HCI systems, with special interest in human pose recovery and behavior analysis from multimodal data.

Xavier Baró received the BS degree in computer science from the UAB, in 2003, the MS degree in computer science from the UAB, in 2005, and the PhD degree in Computer Engineering, in 2009. At the present he is associate professor and researcher with the Computer Science, Multimedia and Telecommunications Department, Universitat Oberta de Catalunya (UOC). From 2015, he is the director of the Computer Vision Master degree at UOC. He is co-founder of the Scene Understanding and Artificial Intelligence (SUNAI) group, and collaborates with the Computer Vision Center of the UAB, as member of the Human Pose Recovery and Behavior Analysis (HUPBA) group. His research interests are related to machine learning, evolutionary computation, and statistical pattern recognition, specially their applications to the field of Looking at People.

Isabelle Guyon is chaired professor in "big data" at the Université Paris-Saclay (UPSud/INRIA), specialized in statistical data analysis, pattern recognition and machine learning. Her areas of expertise include computer vision and bioinformatics. Her recent interest is in applications of machine learning to the discovery of causal relationships. Prior to joining Paris-Saclay she worked as an independent consultant and was a researcher at AT&T Bell Laboratories, where she pioneered applications of neural networks to pen computer interfaces (with collaborators including Yann LeCun and Yoshua Bengio) and co-invented with Bernhard Boser and Vladimir Vapnik Support Vector Machines (SVM), which became a textbook machine learning method. She organizes challenges in Machine Learning since 2003 supported by the EU network Pascal2, NSF, and DARPA, with prizes sponsored by Microsoft, Google, Facebook, Amazon, Disney Research, and Texas Instrument. She is action editor of the *Journal of Machine Learning Research*, program chair of the NIPS 2016 conference, and general chair of the NIPS 2017 conference.

Hugo Jair Escalante received the PhD degree in computer science, for which he received the best PhD thesis on Artificial Intelligence 2010 award (Mexican Society in Artificial Intelligence). He is researcher scientist at Instituto Nacional de Astrofisica, Optica y Electronica, INAOE, Mexico. In 2017 he received the UANL research award in exact sciences. He is secretary and member of the board of directors of ChaLearn, a non-profit organism dedicated to organizing challenges, since 2011. He is information officer of the IAPR Technical Committee 12. Since 2017, he is editor of the Springer Series on Challenges in Machine Learning. He has been involved in the organization of several challenges in machine learning and computer vision, collocated with top venues, see http://chalearnlap.cvc.uab.es/. He has served as co-editor of special issues in the *International Journal of Computer Vision*, the *IEEE Transactions on Pattern Analysis and Machine Intelligence*, and the *IEEE Transactions on Affective Computing*. He has served as area chair for NIPS 2016 - 2018, he is data competition chair of PAKDD 2018 and has been member of the program committee of venues like CVPR, ICPR, ICCV, ECCV, ICML, NIPS, IJCNN. His research interests are on machine learning, challenge organization, and its applications on language and vision.

Georgios (Yorgos) Tzimiropoulos received the diploma degree in Electrical and Computer Engineering from the Aristotle University of Thessaloniki, Greece, and the MSc and PhD degrees in Signal Processing and Computer Vision both from Imperial College London, United Kingdom. As of 2015, he is assistant professor with the School of Computer Science, University of Nottingham and member of the Computer Vision Laboratory. From 2011 to 2014, he was senior lecturer with the School of Computer Science, University of Lincoln and part-time senior research fellow with the Intelligent Behaviour Understanding Group (iBUG), Department of Computing, Imperial College London. Prior to this, he was research associate with the same group. He is currently an associate editor of the *Image and Vision Computing Journal*. His main research interests are in the areas of face and object recognition, alignment and tracking and facial expression analysis.

Michel Valstar is an associate professor with the University of Nottingham, School of Computer Science, and a researcher in Automatic Visual Understanding of Human Behaviour. He is a member of both the Computer Vision Lab and the Mixed Reality Lab. Automatic Human Behaviour Understanding encompasses Machine Learning, Computer Vision, and a good idea of how people behave in this world. He is currently coordinator of the H2020 LEIT project ARIA-VALUSPA, which will create the next generation virtual humans. He has also recently been awarded a prestigious Melinda and Bill Gates foundation award to automatically estimate babies' gestational age after birth using the mobile phone camera, for use in countries where there is no access to ultrasound scans. He was a visiting researcher at the Affective Computing group at the Media Lab, MIT, and a research associate with the iBUG group, which is part of the Department of Computing, Imperial College London. His expertise is facial expression recognition, in particular the analysis of FACS Action Units. He recently proposed a new field of research called 'Behaviomedics', which applies affective computing and Social Signal Processing to the field of medicine to help diagnose, monitor, and treat medical conditions that alter expressive behaviour such as depression.

Maja Pantic is a professor of Affective and Behavioral Computing in Imperial College and leader of the i·BUG group, working on machine analysis of human non-verbal behaviour and its applications to human-computer, human-robot, and computer-mediated human-human interaction. She published more than 200 technical papers in the areas of machine analysis of facial expressions, machine analysis of human body gestures, audiovisual analysis of emotions and social signals, and human-centered machine interfaces. She has more than 12,000 citations to her work, and has served as the Key Note Speaker, chair and co-chair, and an organization/ program committee member at numerous conferences in her areas of expertise.

Jeffrey Cohn is professor of Psychology and Psychiatry at the University of Pittsburgh and adjunct professor of Computer Science at the Robotics Institute at Carnegie Mellon University. He leads interdisciplinary and inter-institutional efforts to develop advanced methods of automatic analysis and synthesis of facial expression and prosody; and applies those tools to research in human emotion, social development, nonverbal communication, psychopathology, and biomedicine. He has served as co-chair of the 2008 and 2015 IEEE International Conference on Automatic Face and Gesture Recognition (FG2008) (FG2015), the 2009 International Conference on Affective Computing and Intelligent Interaction (ACII2009), the Steering Committee for IEEE International Conference on Automatic Face and Gesture Recognition, and the 2014 International Conference on Multimodal Interaces (ACM 2014). He has co-edited special issues of the *Journal of Image and Vision Computing* and is a co-editor of the *IEEE Transactions in Affective Computing* (TAC). His research has been supported by grants from the National Institutes of Health, National Science Foundation, Autism Foundation, Office of Naval Research, Defense Advanced Research Projects Agency, and other sponsors.

Takeo Kanade works in multiple areas of robotics: computer vision, multi-media, manipulators, autonomous mobile robots, medical robotics and sensors. He has written more than 400 technical papers and reports in these areas, and holds more than 20 patents. He has been the principal investigator of more than a dozen major vision and robotics projects at Carnegie Mellon. He has been elected to the National Academy of Engineering and the American Academy of Arts and Sciences. He is a fellow of the IEEE, a fellow of the ACM, a founding fellow of American Association of Artificial Intelligence (AAAI), and the former and founding editor of the *International Journal of Computer Vision*. Awards he received include the Kyoto Prize, the Franklin Institute Bower Prize, ACM/AAAI Newell Award, Okawa Award, C&C Award, Tateishi Grand Prize, Joseph Engelberger Award, IEEE Robotics and Automation Society Pioneer Award, FIT Accomplishment Award, and IEEE PAMI-TC Azriel Rosenfeld Lifetime Accomplishment Award.