

Guest Editorial: Special Issue on Vision-based Human Activity Recognition

Vanessa Testoni, Otávio Augusto Bizetto Penatti, Fernanda Alcântara Andaló, Miguel Lizarraga, Letícia Rittner, Eduardo Valle, and Sandra Avila

HUMAN activity recognition (HAR) is a challenging area of computer vision and artificial intelligence research with several applications. Most of the applications require an automated recognition of high-level activities, composed of multiple simple (or atomic) actions of persons. The implementation of *smart environments* in the IoT (*Internet of Things*) paradigm, for instance, is largely dependent on the recognition of human actions and activities.

The recognition of human activities has been approached in two different ways, namely using *external* and *wearable* sensors. In the former, the devices are fixed in predetermined points of interest and in the latter, the devices are attached to the user. External sensors can be placed in target objects which people are supposed to interact with (e.g. stove, faucet, washing machine, etc.). Cameras are also external sensors that have been employed in HAR problems, especially depth cameras, which have become popular after *Microsoft Kinect*. In the category of wearable sensors, most of the measured attributes in HAR problems are related to the user's movement (e.g. using accelerometers or GPS), environmental variables (e.g. temperature and humidity), or physiological signals (e.g. heart rate or electrocardiogram).

In this special issue, we asked for contributions to the HAR problem focusing on *cameras* as external sensors. All submissions were invited and went through the normal JCIS peer-reviewing process. As a result, we accepted only 5 papers that compose this special issue on the problem of Vision-based Human Activity Recognition.

Vision-based HAR should provide an automated analysis (or interpretation) of ongoing events and their context from video data. Performing vision-based human activity recognition is a problem that presents several challenges, such as: human detection and tracking, design of visual features for describing people, learning and recognition of actions, pose estimation, behavior recognition, location estimation and location-based services. Examples of application scenarios are: monitoring patients at hospitals, monitoring students in the classroom, surveillance systems, violence detection, detection of suspicious behaviors in events, sports visual analytics, elderly monitoring, and so on.

As one can note, this is a research field with both great

challenges and applications and we expect this to be an active area of research and development for many years to come. We are pleased that this special issue captures a broad range of these applications and approaches.

In the paper "High Level Event Detection based on Spatial Occupancy and Interpersonal Relationships" by J. Soldera, J. Bins, M. Cohen, J. Junior, S. Musse, and C. Jung, the authors describe a new approach specially suitable for security applications in several scenarios, from surveillance to the understanding of group formation and classification. The approach focus on event detection based on spatial occupancy and interpersonal relationships. The authors also propose a query language to specify complex events. The queries can be used to detect events on-the-fly as the video is processed, or applied to stored video databases.

In the paper "A proposal of a non-intrusive, global movement analysis of hemiparesis treatment" by J. Vieira, D. Carmo, Y. Jovita, and L. Oliveira, the authors focus on health applications, specially on the treatment of hemiparesis, a loss of muscle strength on one side of the body which is a common disabling condition after a stroke. The paper presents a method to perform human functional movement analysis by specific energy of movement volume using a system composed by a set of RGB-D cameras. The final goal is to assess the quality of the Physiotherapy treatment.

In the paper "Continuous Sign Recognition of Brazilian Sign Language in a Health Care Setting" by J. Vidalón and J. Martino, the authors also focus on one health application - the automatic recognition of continuous signing of Brazilian Sign Language (Libras) in healthcare settings. The paper describes the approach to recognize isolated signs and the attempt to recognize continuous signs in sentences, based on the Hidden Markov Model (HMM). The scenario of health emergency situations was chosen to illustrate the paper conversations due to its high social value. The authors worked with 58 signs and 15 sentences modeled with HMMs.

In the paper "A Variable Size Block Matching Based Descriptor for Human Action Recognition" by F. Oliveira, H. Maia, V. Mota, M. Vieira, and A. Araújo, instead of targeting a specific application, the authors focus on motion description and intend to integrate the well-known block matching technique into the field of HAR. They present a tensor self-descriptor obtained from a VSBMA (Variable Size Block Matching Algorithm). The displacement vectors computed by VSBMA are represented with histograms which are then coded into orientation tensors. Several experiments are presented and show that the method is simple yet achieves

V. Testoni, O. A. B. Penatti, F. A. Andaló, and M. Lizarraga are with Samsung R&D Institute Brazil (SRBR), Campinas, SP, 13097-160, Brazil (e-mails: {vanessa.t, o.penatti, f.andalo, m.lizarraga}@samsung.com).

L. Rittner, E. Valle, and S. Avila are with Unicamp, Campinas, SP, 13083-852, Brazil (e-mails: {lrittner, dovalle, sandra}@dca.fee.unicamp.br).

Digital Object Identifier 10.14209/jcis.2015.7

comparable results with state-of-the-art techniques.

Finally, in the paper “Kinematic-based Markerless Human Tracking in 3D Probabilistic Occupancy Grids” by R. Bem, M. Goulart, G. Simas, and S. Botelho, the authors focus on the fundamentals of the human motion tracking problem. They propose a 3D tracking algorithm without the need of markers for environments monitored by multiple cameras. The method is based on a skeletal kinematic model of the human body and is applied over a 3D probabilistic occupancy grid of the environment. Experiments were performed to evaluate the proposed method.

We now present you the papers in this special issue and hope that you find them useful and inspiring in your research.



Vanessa Testoni received the B.S. degree in computer science from the Pontifical Catholic University of Paraná (PUCPR), Brazil, in 2002, the B.S. degree in electrical engineer from the Federal University of Paraná (UFPR), Brazil, in 2004, the M.Sc. and the Ph.D. degrees in electrical engineering from the University of Campinas (UNICAMP), Brazil, respectively, in 2006 and 2011. Right after her graduation, she joined the University of California, San Diego (UCSD), USA, as a postdoctoral employee.

Since 2013 she is a researcher at SRBR (Samsung Research Institute Brazil) where her research interests include digital signal processing, information theory, image processing, computer vision, machine learning, mobile communications, and specially image and video coding. She is an IEEE and ACM member, and a .NET Microsoft Certified Professional. In 2009 she was the Brazilian PhD Student awarded the *Microsoft Research Ph.D. Fellowship Award* and in 2014 she received the *MIT TR35* (Young Innovators under 35) in the first Brazilian edition of the award.



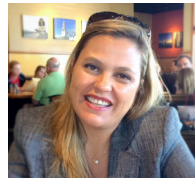
Otávio A. B. Penatti received his M.Sc. and Ph.D. degrees in Computer Science from University of Campinas – Brazil in 2009 and 2012, respectively. He is currently a research scientist at SAMSUNG Research Institute (SRBR) – Brazil. His research interests include computer vision, content-based image retrieval, pattern recognition, machine learning, and multimedia geocoding.



Fernanda A. Andaló is a Research Scientist at SAMSUNG Research Institute Brazil. She received a B.Sc. in Computer Science from University of Brasília (UnB) in 2004, a M.Sc. in Computer Science from University of Campinas (UNICAMP) in 2007, and a Ph.D. in Computer Science from the same university in 2012. In 2011 she was a visiting researcher at the Division of Engineering, Brown University, and until 2014 she was a postdoctoral researcher at VISGRAF, National Institute of Pure and Applied Mathematics (IMPA). Her research interests include Computer Vision, Image Processing, and Computer Graphics.



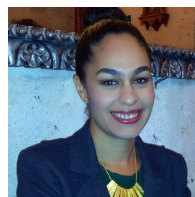
Miguel Lizarraga Miguel Lizarraga received the B.S. degree in Electrical Engineering from University of Campinas (UNICAMP), where he also received the M.Sc. and Ph.D. in 1994 and 2000, respectively. His research interests are in the area of image processing and pattern recognition. From 2001 to 2003 he worked as researcher at UNICAMP and in 2004 as professor at State University of Ceará (UECE). In 2005 started to work at SAMSUNG as the interface between academic research performed by universities and the market needs, as well as innovation management and intellectual property. Since 2012 he is R&D Manager at SAMSUNG Research Institute Brazil (SRBR) leading research projects that result in innovative solutions in the areas of Health, Education and Security.



Letícia Rittner is a professor at the School of Electrical and Computer Engineering, University of Campinas (UNICAMP) and member of the Medical Image Computing Lab (MICLab). Her main research interests are image processing and analysis, pattern recognition and medical image computing, with emphasis on diffusion MRI. She obtained MSc (2004) and PhD (2009) degrees from the University of Campinas, with a 6-month internship at the Montreal Neurological Institute, McGill University, Canada. She was a postdoc at the Medical Imaging Processing Group (MIPG) at the School of Medicine, University of Pennsylvania. She is an IEEE and SPIE member.



Eduardo Valle is a professor at the Department of Computer Engineering and Industrial Automation – DCA of the School of Electrical and Computer Engineering – FEEC at the State University of Campinas – UNICAMP. He is also a faculty member of the RECOD Lab (REasoning for COMplex Data). He’s got a Ph.D in Computer Sciences at the University of Cergy-Pontoise, in 2008. He’s got a M.Sc. and a B.Sc., also in C.S., at the Federal University of Minas Gerais, in 2003 and 2001 respectively. He works with a talented team of researchers and students on themes related to on multidimensional databases, multimedia databases, content-based information retrieval, and large-scale machine learning.



Sandra Avila is postdoctoral researcher at the School of Electrical and Computer Engineering (FEEC), at the University of Campinas (UNICAMP). In 2013 she received the Ph.D. degree in Computer Science from Federal University of Minas Gerais (UFMG), Brazil, and University Pierre and Marie Curie (UPMC Sorbonne – Paris VI), France. In 2008 she obtained the M.Sc. degree in Computer Science from UFMG. In 2006 she received the B.Sc. degree, also in Computer Science, from Federal University of Sergipe. Sandra Avila’s research focus is at Machine Learning, Computer Vision, Content-based Information Retrieval, and Medical Image Analysis.