

# Vision for human-computer interaction and virtual reality systems

One Day BMVA Symposium in London, UK on Wednesday 6<sup>th</sup> May, 2015

**Chair:** Manuela Chessa (University of Genoa, Italy)

The chair, Manuela Chessa, opened the BMVA meeting at 10am on 6<sup>th</sup> May 2015 by welcoming speakers and attendees. She introduced the main motivation of the BMVA meeting titled "*Vision for human-computer interaction and virtual reality systems*". The meeting was dedicated to the integration of two important topics in the development of human-computer interaction systems based on virtual reality (VR). On one hand, the meeting addressed the investigation and development of Computer Vision techniques for the detection, tracking and analysis of human body pose and movement. In conjunction with such technological development, scientific investigators must study the perceptual issues that may arise with the use of VR systems.

Following the introduction, the first keynote speaker, Guido Maiello from University College London, opened the morning session dedicated to psychologists and to the study of visual perception in (stereoscopic) virtual reality. Guido Maiello's talk, titled "Depth Perception and Binocular Vision in Naturalistic Virtual Reality", focused on the importance of investigating and monitoring eye movements in virtual reality settings. The speaker also introduced a gaze-contingent depth-of-field display and demonstrated how such a display, as well as other virtual reality techniques, can be employed to modify binocular eye movements.

Focus cues for VR applications were further discussed in the excellent talk given by Dr Simon Watt of Bangor University, titled "Presenting natural focus cues in virtual/augmented reality". The speaker eloquently outlined the vergence-accommodation mismatch and its consequences for immersion in VR. Dr Watt went on to describe the merits and drawbacks of several solutions which have been proposed in the literature to properly simulate accommodation: from software solutions, to volumetric and fixed-viewpoint displays, ending with light-field displays. The speaker concluded that presenting natural focus cues is worthwhile and that future displays will certainly render approximately correct focus cues, yet the vision-science community must provide a more thorough understanding of how the visual system employs focus cues in order to build generalizable models to predict how novel displays will perform.

To conclude the morning session Dr Paul Hibbard from the University of Essex discussed ways to quantitatively define the quality of simulated 3D environments. He proposed that the optimal solution is task dependent. For each VR application, developers and experimenters must determine the best trade-off among appearance, performance and experience, since maximizing any of these three metrics will often lead to a decrease in the others.

The afternoon session of the meeting, dedicated to Computer Vision topics, commenced following the lunch break. The second keynote speaker Dr Tae-Kyun Kim from Imperial College London opened the afternoon session. Dr Kim gave a talk titled "Articulated Hand Pose Estimation by Decision Forests", in which he presented a detailed overview on the latest techniques for hand action/gesture recognition via random forest classification algorithms.

The following presenter, Victoria Bloom from Kingston University, described a method for real-time human action recognition. The action recognition system outlined by the speaker was developed to

perform online detection of multiple complex actions from multiple users, which is an essential requirement of natural user interfaces.

Muhammad Asad of City University London presented a talk titled "Hand pose and orientation estimation for ego-centric devices". The work was motivated by the fact that the diffusion of mobile and wearable devices requires the development of novel interaction methods.

Following these three technically challenging talks, a coffee break allowed meeting attendees to assimilate and discuss the presented content.

The meeting commenced once again with a talk by Prof. Francesca Odone from the University of Genoa. Prof. Odone presented joint work with colleagues from both the University of Genoa and from the Italian Institute of Technology on "Understanding human motion and its qualities". The presentation was well received as it discussed biological motion perception from the perspective of cognitive science with a computer vision approach. Specifically the multi-centre research team of which Prof. Odone is part is applying computational models to understand the cognitive processes underlying biological motion perception. Their findings are integrated into a robotic system, the iCub, which is designed, among other uses, to study how human perceptual development can be applied to humanoid robotic systems.

The following speaker, Dr Nicoletta Noceti from the University of Genoa, presented "Good practices of hand gestures recognition for the design of customized NUI". The talk focused on a dynamic gestures recognition method for Natural User Interfaces, i.e. in all interfaces where a user may employ arms and hands to perform dynamic gestures in front of a camera to direct a system's behaviour. To demonstrate the robustness of the method Dr Noceti presented a photo browsing system as an example application.

The final talk of the meeting, titled "You-Do, I-Learn: Unsupervised Multi-User Egocentric Approach Towards Video-Based Guidance", was given by Dr Dima Damen of the University of Bristol. She presented an approach for discovering task relevant objects and their common modes of interaction from multi-user egocentric video. The fully automated method compares appearance, position and motion features of objects within a scene and combines these features with gaze fixations as a proxy for attention. Once a task relevant object is identified, the system provides the user with a video guide on how the object is to be used. Following her presentation Dr Damen invited meeting attendees to test a working prototype of her developed system implemented on Google glasses.

The meeting officially ended at 5pm, but scientific discussion continued both in the hallways of the BCS building and at dinner engagements among meeting attendees. Overall, the meeting was well received, highly productive, and concluded with the potential of interesting collaborations to be commenced in the future.

Manuela Chessa