Special Section: Best Papers from the 19th ACM Symposium on Virtual Reality Software and Technology (VRST 2013)

Guest Editors' Introduction

New technical innovations in telepresence allow people to attend meetings or gather together anywhere they are and to share the same virtual space using 3D avatars. More sophisticated systems have been developed with 3D capturing and rendering, leading to a true 3D telepresence experience. If the 3D avatar is remotely guided by the real participant, it is possible to replace this real participant by his or her autonomous virtual counterpart, as an i-substitute. The i-substitute is supposed to give a partial illusion that the real person is present. This implies that the i-substitute should look the same as the real human, speak with the same intonation, and be somehow aware of the real situation, the real participants, and the task currently being performed. The i-substitute should react at the right time based on the perception he/she has from the real participants. It implies an evaluation of what each real participant is doing. Perception will be obtained by visual and audio input and recognition. The i-substitute reacts according to the input and its current knowledge; its reactions encompass animation (body and facial gestures) and speech synthesis. The i-substitute could be an autonomous virtual human as well as a social robot, depending on the type of meeting and environment.

This special section contains the four best papers on this innovative topic from the 19th ACM Symposium on Virtual Reality Software and Technology (VRST 2013), held by the Institute for Media Innovation, Nanyang Technological University, October 6–8, 2013 on the Nanyang Technological University (NTU) campus. These four papers have been considerably improved and expanded from the conference papers. Each extended paper has been peer reviewed carefully. The editors would like to thank the external reviewers for their careful work.

In the first paper, "A United Framework for Individualized Avatar-Based Interactions," Arjun Nagendran, Remo Pillat, Adam Kavanaugh, Greg Welch, and Charles Hughes propose a framework to interactively control avatars in remote environments. The system, called AMITIES, serves as the central component that connects people controlling avatars (inhabiters), various manifestations of these avatars (surrogates), and people interacting with these avatars (participants). A multi-server-client architecture, based on a low-demand network protocol, connects the participant environment(s), the inhabiter station(s), and the avatars. A human-in-the-loop metaphor provides an interface for remote operation, with support for multiple inhabiters, multiple avatars, and multiple participants. Custom animation blending routines and a gesture-based interface provide inhabiters with an intuitive avatar control paradigm. This gesture control is enhanced by genres of program-controlled behaviors that can be triggered by events or inhabiter choices for individual or groups of avatars. This mixed (agency and gesture-based) control paradigm reduces the cognitive and physical loads on the inhabiter while supporting natural bi-directional conversation between participants and the virtual characters or avatar counterparts, including ones with physical manifestations, for example robotic surrogates.

In the second paper, "Human–Robot Interaction by Understanding Upper Body Gestures," a human–robot interaction system based on a novel combination of sensors is proposed by Yang Xiao, Zhijun Zhang, Aryel Beck, Junsong Yuan, and Daniel Thalmann. It allows one person to interact with a humanoid social robot using natural body language. The robot understands the meaning of human upper body gestures and expresses itself by using a combination of body movements, facial expressions, and verbal language. The gestures are characterized by the head, arm, and hand posture information. The wearable Immersion CyberGlove II is employed to capture the hand posture. This information is combined with the head and arm posture captured from the Microsoft Kinect. The gesture understanding approach based on an innovative combination of sensors is the main contribution of this paper. In addition, for robot motion generation and control, a novel online motion planning method is proposed. In order to generate appropriate dynamic motion, a quadratic programming (QP)-based dual-arms kinematic motion generation scheme is proposed, and a simplified recurrent neural network is employed to solve the QP problem. The integration of handshake within the HRI system illustrates the effectiveness of the proposed online generation method.

The third paper, “A Study on High-Level Autonomous Navigational Behaviors for Telepresence Applications,” by Wee Ching Pang, Gerald Seet, and Xiling Yao, presents a framework enabling navigational autonomy for a mobile platform with application scenarios specifically requiring a humanoid telepresence system. The proposal promises reduced operator workload and safety during robot motion. In addition, the framework enables the inhabitor (the human controlling the platform) to provide inputs for head and arm gesticulation. This allows the inhabitor to focus more on interactions in the remote environment than being engrossed in controlling robot navigation. This paper discusses the development of higher-level, human-like navigational behaviors such as following, accompanying, and guiding a person autonomously. A color histogram comparison and position-matching algorithm has been developed to track the person using the Kinect sensors. In addition to providing a safe and easy to use system, the high-level behaviors are also required to be “human-like,” in that they obey the laws of proxemics and other human interaction norms such as walking speed. An obstacle avoidance function has also been implemented using the virtual potential field method. A preliminary evaluation was conducted to validate the algorithm and to support the claim of reducing operator cognitive load due to navigation.

In the last paper in this special section, “Modelling Multi-Party Interactions among Virtual Characters, Robots, and Humans,” Zerrin Yumak, Jianfeng Ren, Nadia Magnenat Thalmann, and Junsong Yuan show that 3D virtual humans and physical human-like robots can be used to interact with people in a remote location in order to increase the feeling of presence. In a telepresence set-up, their behaviors are driven by real participants. We envision that in the absence of the real users, when they have to leave or they do not want to perform a repetitive task, the control of the robots can be handed to an artificial intelligence component to sustain the ongoing interaction. At the point when human-mediated interaction is required again, the control can be given back to the real users. One of the main challenges in telepresence research is the adaptation of 3D position and orientation of the remote participants to the actual physical environment for appropriate eye-contact and gesture awareness in a group conversation. In case the human behind the robot and/or virtual human leaves, multi-party interaction should be handed to an artificial intelligence component. This paper discusses the challenges in autonomous multi-party interaction among virtual characters, human-like robots, and real participants and describes a prototype system to study these challenges.

In keeping with past practice, the four papers included in the special section are designated in the Table of Contents with an “S”.

Nadia Magnenat Thalmann
Daniel Thalmann
Nanyang Technological University, Singapore

Guest Editors