

Ubiquitous Realities through Situated Social Agents

Mauro Dragone (Mauro.Dragone@ucd.ie)

Thomas Holtz (Thomas.Holtz@ucd.ie)

Brian R. Duffy (Brian.Duffy@ucd.ie)

Gregory.M.P. O'Hare (Gregory.OHare@ucd.ie)

Department of Computer Science, University College Dublin
Belfield, Dublin 4, Ireland

Abstract

In order to take the developing field of Social Robotics a stage further, this work investigates mixed reality in embodied social agents. The Virtual Robotic Workbench is employed which provides a versatile framework for experimentation in interoperability and cooperation between heterogeneous robots (real, simulated and virtual) and humans in multi-reality domains. Explicit social interaction, both between virtual and real robots and between robots and people, is supported.

Keywords: Virtual reality and augmented reality, Social and Conversational agents

1. Introduction

In recent years, the field of socially capable artificial entities - whether embodied as physical robots, as avatars in virtual worlds, or in the form of digital assistants - has grown significantly. Not only has this facilitated differing forms of ubiquitous artificial intelligence, it has also allowed us to develop what can be termed *ubiquitous realities*.

Recent studies suggest that behavioural realism should match visual realism in order to satisfy user expectations and generate believable experiences [1],[2] for social interaction. This statement holds true for intelligent characters as well as autonomous robots [3],[4]. As a result, a lot of research has been invested into achieving more lifelike avatar behaviour [5],[6] and robots that are situated in a social environment [7]. However, the research presented in this paper primarily focuses on the functional mechanisms required to develop socially situated agents in ubiquitous realities. Issues, such as the preservation of identity in mixed reality social interactions, have been addressed but will not be discussed in this paper.

In order to identify the differences and similarities between situated embodiment in the real and the virtual world, one has to observe what happens when the boundaries begin to blur. This work has realised a mixed reality application that enables observers to visualise and interact with virtual and mixed reality characters through a Head Mounted Display (HMD).

Section 2 reviews work on socially situated embodied agents, and specifically outlines the Socially Situated Agent Architecture (SoSAA) and the Virtual Robotic Workbench (VRW) - a Collaborative Virtual Environment (CVE) for social robotics. Section 3 describes the use of these tools for the implementation of mixed reality agents and presents some practical demonstrations. Finally, future work is discussed in section 4.

2. Situated Social Agents

Recent research has focused beyond individual robots and has begun to consider robot collectives. One of these endeavours is the Social Situated Agent Architecture (SoSAA), a framework that extends our previous work on social robotics [8] allowing mixed reality experiences.

Situated social agents constitute autonomous entities equipped with both explicit social capability and a strong provision for both their physical and social context, i.e. their embodiment. Core to achieving an integrated approach to the agent and its physical and social environment is a synthesis of reactive controllers and traditional planning architectures. In this work, we embrace the behaviour-based control paradigm [9] to achieve reactive-deliberative synthesis.

The SoSAA architecture is based upon the combination of BDI [10] (Belief, Desire and Intention) agents - a reactive behavioural system, and explicit social infrastructure. Key components include a hardware abstraction

layer for heterogeneous robot platform applications, a coherent reactive-deliberative control synthesis, a BDI deliberative level developed through Agent Factory [11], and a FIPA (<http://www.fipa.org>) compliant social level to support explicit social interaction.

The *physical level* provides for an ease of portability of the SoSAA architecture between differing physical robot platforms and between simulated and real robots. The Reactive Controller component at the *reactive level* supervises the physical layer, managing a library of behavioural modules. Behaviours implement both reflex robot responses to unexpected or dangerous events and more complex actions. The *deliberative level* of SoSAA follows a Multi-Agent System (MAS) organization with several agents supervising the different functional levels of the robot. At any given time, a number of agents share the control of the robotic platform. These agents vary in complexity from simple procedural knowledge modules that deal with lower level capabilities of the platform (i.e. sensorial organization, configuration and behavioural arbitration) to means-ends reasoning. The reasoning capability is delivered through Agent Factory [11], an integrated and tooled environment for the rapid prototyping of social agents.

To enable collaboration among social robots, Agent Factory agents, developed at the *social level*, make use of the Speech Act Theory [12] for accurate and expressive communication mechanisms in Multi-Agent Systems. This is undertaken by performing a speech act (such as *requesting*, *ordering*, *informing* or *promising*) that sends a message to one or more of their socially capable acquaintances in order to affect their mental states. In addition, SoSAA implements a number of more sophisticated coordination protocols responding to the semantics described in FIPA specifications, among them, the Contract-Net-Protocol for group formation.

Specific tools are needed in order to work toward true interoperability and cooperation between social heterogeneous robotic agents (real, simulated & virtual) and effective human-robot interaction. The Virtual Robotic Workbench (VRW), which works synergistically with the SoSAA architecture, is an instance of one such tool and is detailed in the following section.

2.1 The Virtual Robotic Workbench

The Virtual Robotic Workbench [13][14] is an agent-based Collaborative Virtual Environment (CVE) for embodied social agents interacting with other agents and humans. It facilitates interoperability between physical robots, users, and virtual avatars.

The VRW was developed to ease configuration and networking for large scale heterogeneous real and virtual robotic teams, thus enabling dynamic composition of sensors and actuators while supporting dynamic discovery of resources and peers. In addition, the VRW offers a framework to support social integration, implemented in a re-usable and standardised form. The core features of the VRW are: the immersion of robots in a shared collaborative environment, and the adoption of mature Multi-Agent Systems technology in order to enable robot-robot and human-robot interaction. The Workbench employs indirect communication through augmented sensing, cooperation and migration or mutation of mobile agents among different robotic platforms.

The Virtual Robotic Workbench offers:

- A communication medium - based upon a XML multicast protocol - which is exploited for dynamic resource discovery and to exchange information and control among humans and robots.
- A visualisation medium (Virtual Robotic Workbench Visualisation Suite), which offers real-time, multimedia visualisation facilitating behavioural scrutiny and situational awareness.
- A FIPA (<http://www.fipa.org>) compliant Agent technology, which supervises the social interface between each user and the shared environment.

As in DIS [15] the Workbench uses UDP multicast as the preferred transport protocol, but we also developed adapters that support multiple protocols to communicate with remote nodes through UDP or TCP/IP over Ethernet links and Bluetooth.

The Workbench Messaging System includes a mechanism that allows the run-time definition of secondary channels, each with a separate transport layer. VRW XML messages can contain special URL tags, which are used to publish details about these secondary

channels. Secondary channels can be used, for example, to publish static attributes (like a three-dimensional model or avatar used to represent the agent in the virtual or augmented space). When such a tag is encountered for the first time, the receiving node launches a background thread to access its content. In this manner, a node connecting to the workbench is gradually informed about the entities participating in the action without the need for more complex session-oriented protocols.

The VRW Visualization Suite uses different tools for scene rendering. Only the actual graphic presentation is delegated to such tools, while functionalities like smoothing and collision detection are retained within the VRW. The Virtual Reality Modelling Language (VRML) was selected as the medium for model description, however, alternate versions were developed, among them, an OpenGL-GLUT and a Java3D component. Visualisation on the Internet is achieved through the Java3D renderer or via the interaction with a VRML plug-in through scripting or via the External Authoring Interface (EAI) [16].

3. Mixed Reality Agents

During a SoSAA mixed reality experience, users, each wearing a computer and a see-through Head Mounted Display (HMD) with associated digital camera, are immersed within an augmented space obtained by superimposing synthetic imagery on the HMD video. Ideally, users should be free to move in this space. They may interact with the entities (robots) through animated virtual characters visualized on top of them (see Figure 2) or monitor simulated robots that appear to be working alongside real ones. SoSAA realizes the overlay between real and virtual images by tracking the position and orientation of the head of each observer.

In tracking an observer's viewpoint, a traditional marker-based method was chosen. Marker-based techniques compute the camera's position relative to a physical marker by analysing its appearance in the camera's image. These methods offer an easy solution to the registration problem between virtual and real imagery; at least as long as the necessary markers are visible and recognizable in the processed video. Markers offer a cheap alternative to other systems based on, for

example, magnetic tracking devices but require a degree of engineering of the experimental environment.

Marker-less techniques for motion estimation [17] could be used to complement traditional approaches through, for instance, updating the observer position during gaps in the observation of markers. With this in view, the application was designed as a service-oriented framework defining abstract, customisable layers like *tracking*, *filtering*, *data distribution* and *rendering*.

As in [18] the tracking layer of SoSAA is a shared tracking system in which multiple observers communicate with each other in order to collaborate in the tracking effort. Tracking and rendering in the current implementation are based upon ARToolkit [19] while the VRW Messaging Service is exploited for the dynamic discovery of peer observers and the dissemination of data through the network. This data includes tracking, robot telemetry (positional updates and sensing), and other information regarding the behaviour and the cognitive state of the robotic agents and the virtual characters.

3.1 Tracking with ARToolkit

ARToolkit provides a software library, which enables the recognition and the 3D pose estimation of physical markers. These patterns can be arbitrary, but to improve the identification we used the system proposed by Owen *et al.* [20], generating the pattern from a direct cosine transformation (DCT) in order to minimize the physical similarity of different markers.

To improve ARToolkit's accuracy over distance, we use a Kalman Filter [21] to improve the tracking results (i.e. fusing multiple sources of data) and thus allow recognition over a larger distance.

As the user can move freely around the robot, distance is not the only issue. A single marker on top of the robot, for example, would be rarely seen from a favourable angle. One can overcome this obstacle by using multiple markers to track a single object [22]. The robot is therefore equipped with five markers aligned in a cube, so the application observes at least one marker from all angles. A transformation matrix for each of the four sides relates the respective coordinate system to that of the marker on top of the cube. If two or three sides

of the cube are visible, the Kalman filter is used to fuse the measurements into a single, improved result.

3.2 Shared Tracking

The limited viewing range of the cameras that we employ (typically 50 degrees) results in limitations to the freedom of movement of the human operators wearing the HMDs. Whenever other objects occlude the cubic marker or when it exits from the field of view of the observer, the associated robot ceases to be tracked and the virtual character disappears even if the majority of the robot is still visible.

The tracking capabilities of the system are improved by placing additional static markers (see figure 1) in the environment to increase the probability of observing at least one marker from any position of the observer.

SoSAA then relies on the distribution of observation among multiple observers. Each observation, for instance, produces camera coordinates, which locate the observer in relation to the observed marker. Combining observations from multiple observers, together with any known relations between static markers, produces a graph where the vertexes are frames of reference, and the edges are transformations from one frame to the other. This graph can be used to deduce the coordinates of tracked objects (by finding a path in the graph) even in the absence of direct observation thus augmenting the visualisation capabilities of each observer.

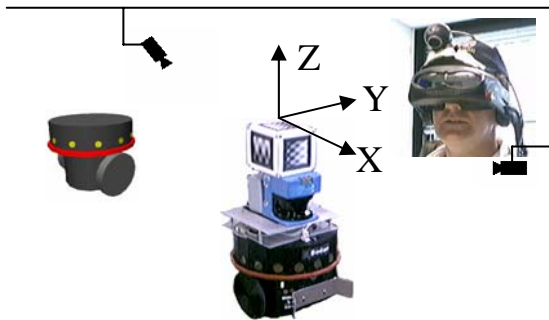


Figure 1 : Main components in a typical experimental setting.

SoSAA adopts Dijkstra's algorithm [23] to find the optimum path from the observer node to every other node in the observation graph. By weighing the edges with error covariance, reliable paths are chosen over unreliable ones.

3.3 Behavioural Rendering

The SoSAA also defines a framework for distributed behavioural rendering, managing the appearance of both simulated robots and virtual characters associated with real robots (i.e. virtual heads). Virtual and real elements form a distributed hybrid system ought to exhibit cohesion and behavioural consistency to the observer. To do this, the SoSAA supports animated characters through the activation and management of their animation cycle in line with the behaviour of their robotic counterparts.

In each observer a copy of a *Virtual Translator Agent* is in charge of the behavioural rendering for each robot in the scene. This agent is responsible for reading the BDI state transmitted by a particular robot through the VRW and for keeping the behaviour of the associated virtual character in line with the cognitive state of the real robotic agent. The Virtual Translator Agent maps subsets of robot beliefs to gesture-type animations available to the virtual counterpart.

Each avatar offers a predefined set of gestures (i.e. *greet, smile, laugh, frown, nod*) implemented in the VRML model through time-dependent position interpolators. Whenever the Virtual Translator Agent selects an appropriate gesture, the ARToolkit VRML access layer is used in order to activate the script that initiates the time-cycle for the corresponding animation.

3.4 Demonstrations

In order to illustrate the flexibility of the Socially Situated Agent Architecture, a number of demonstrations have been undertaken. The following figures show mixed-reality autonomous characters that demonstrate the functionality of the tracking system and the intentional deliberative reasoning mechanisms functioning in a social context.

Figure 2 shows an animated character situated on top of the physical robot. Experiments are currently underway to employ "virtual clothes" to real devices, which, as illustrated in the picture, hide such functional components as the 3D laser scanner and the tracking cube (visible in figure 3).

Figure 3 shows a robot delivering a soccer ball to the observer [24], whose relative position is notified to the robot through the shared tracking framework.

IBM's ViaVoice (is used as speech recognition system for simple vocal control



Figure 2: (a) An Augmented Reality socially situated agent, (b) a head mounted display used to view the AR scene.



Figure 3: Example of child playing football with SoSAA using augmented reality.

4. Conclusions and Future Work

The Socially Situated Agent Architecture provides a unique instrument to explore mixed-reality situated agent research. The system supports the embedding of virtual faces with expressive capabilities upon real robots, without the serious engineering complexities and constraints that would be associated with doing this in the real world. We have also begun to use these mixed reality features to label the current behavioural mode in order to reduce mode misunderstanding for shared control [25]. Our system facilitates the exploration of artificial identity in virtual agents where different observers (child or adult) may be presented with very different avatars. This provides for significantly more flexibility than systems that use a fixed video screen on top of the robot (e.g. [26]) as they can present just one character for all social participants.

Continuing work has already started to integrate occasional observations of the

markers with robot telemetry information and other devices (optical and magnetic) to track the viewpoint of the observer. Kalman filtering has improved the reliability of the tracking and it appears to be the best candidate for merging tracking with these additional sources of information. A model of the dynamic of the observer could also be used for the prediction step in the Kalman Filter.

The system needs to have a better strategy in calculating the coordinates of a simulated robot, through, for example, anchoring them to a more stable marker in the graph to reduce the shaking of their image in the augmented scene.

In the longer term, we intend to use the capabilities we are developing for the construction of a *Mixed Reality Robotic Laboratory*, in which simulated entities can be added (from remote network sites) to the real scenes and thereafter be evaluated in their performance of prescribed tasks.

References

- [1] M. Slater and A. Steed. Meeting People Virtually: Experiments in Shared Virtual Environments. R. Schroeder (ed), *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*, pages 146-171. Springer, London. 2001.
- [2] M. Garau, M. Slater, V. Vinayagamoorthy, A. Brogni, A. Steed and M.A. Sasse. The Impact of Avatar Realism and Eye Gaze Control on Perceived Quality of Communication in a Shared Immersive Virtual Environment. SIGCHI conference on human factors in computing systems, pages 529-536. 2003.
- [3] C. Bartneck and J. Forlizzi. A Design-Centred Framework for Social Human-Robot Interaction. 13th IEEE International Workshop on Robot and Human Interactive Communication. 2004.
- [4] B.R. Duffy. Anthropomorphism and the social robot. *Robotics and Autonomous Systems*, 42:3-4. 2003.
- [5] H. Prendinger and M. Ishizuka. Introducing the cast for social computing: Life-like characters. H. Prendinger and M. Ishizuka, editors, *Life-like Characters. Tools, Affective Functions and Applications*. Cognitive Technologies Series, pages 3-16. Springer, 2004.

- [6] J. Rickel, S. Marcella, J. Gratch, R. Hill, D. Traum and W. Swartout. Toward a New Generation of Virtual Humans for Interactive Experiences. *IEEE Intelligent Systems*, 17(4):32-38. 2002.
- [7] C. Breazeal. Toward sociable robots. *Robotics and Autonomous Systems*, 42(3-4):167-175. 2003.
- [8] B.R. Duffy. *The Social Robot*. Ph.D. Thesis, University College Dublin, Department of Computer Science. 2000.
- [9] R.C. Arkin. Motor Schema Based Navigation for a Mobile Robot: An Approach to Programming for Behavior. *IEEE Int. Conference on Robotics & Automation*, pages 264-271. 1987.
- [10] A.S. Rao and M.P. Georgeff. Modelling Rational Agents within a BDI Architecture. 2nd Int. Conference on Principles of Knowledge Representation & Reasoning, pages 473-484. 1991
- [11] R.W. Collier, *Agent Factory: A Framework for the Engineering of Agent-Oriented Applications*. Ph.D. Thesis, University College Dublin, Department of Computer Science. 2001..
- [12] J. Searle. *Speech Acts: An essay in the philosophy of language*, London, Cambridge University Press. 1969.
- [13] B.R. Duffy, G.M.P. O'Hare, R.P.S. O'Donoghue, C.F.B. Rooney and R.W. Collier. Reality and virtual reality in mobile robotics. 1st International Workshop on Managing Interactions in Smart Environments. 1999.
- [14] M. Dragone. *The Virtual Robotic Workbench*. Technical Report No. UCD-PRISM-05-11. University College Dublin, Department of Computer Science. 2005.
- [15] IEEE Std 1278.1-1995. *IEEE Standard for Distributed Interactive Simulation - Application Protocols*. IEEE, Inc. 1995.
- [16] C. Marrin. *External Authoring Interface (EAI) Proposal*. Silicon Graphics Inc. Mountain View California. 1997. Available via <http://www.web3d.org>.
- [17] A.J. Davison. Real-Time Simultaneous Localisation and Mapping with a Single Camera. 9th International Conference on Computer Vision. 2003.
- [18] OpenTracker Website, visited June 11, 2005. www.studierstube.org/opentracker.
- [19] H. Kato, M. Billinghurst and I. Poupyrev. *ARToolKit*. Technical report, Hiroshima City University. 2000.
- [20] C.B. Owen, F. Xiao and P Middlin. What is the best fiducial? 1st IEEE International Augmented Reality Toolkit Workshop, pages 98-105. 2002.
- [21] R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, pages 35-45, 1960.
- [22] W. Piekarski. and B.H. Thomas. Using ARToolKit for 3D Hand Position Tracking in Mobile Outdoor Environments. 1st Int. Augmented Reality Toolkit Workshop. 2002.
- [23] E.W. Dijkstra. A note on two problems in connection with graphs. *Numerische Mathematik* 1:269-271. 1959.
- [24] Dragone, M., O'Donoghue, R., Leonard, J.J., O'Hare, G.M.P., Duffy, B.R., Patrikalakis, A., Leederkerken, J. "Robot Soccer Anywhere: Achieving Persistent Autonomous Navigation, Mapping and Object Vision Tracking in Dynamic Environments". *In Proceedings of SPIE Opto Ireland*, Dublin, Ireland, April, 2005.
- [25] R. Ross, H. Shi, T. Vierhuff, B. Krieg-Bruckner and J. Bateman. Towards Dialogue Based Shared Control of Navigating Robots. *International Conference on Spatial Cognition*. 2004.
- [26] E. Paulos and J. Canny. Social Tele-Embodiment: Understanding Presence. *Autonomous Robots*, 11(1): 87-95 . 200

