Summer 7-15-2013

Advancements and Challenges Towards a Collaborative Framework for 3D Tele-Immersive Social Networking

Davide Andrea Mauro
Marshall University, maurod@marshall.edu

N. E. O'Connor

D. Monaghan

Marc Gowing

P. Fechteler

See next page for additional authors

Recommended Citation

Mauro, D.A., O'Connor, N.E., Monaghan, D., Gowing, M., Fechteler, P., Eisert, P., Wall, Julie, Izquierdo, E., Alexiadis, D.S., Daras, P., Mekuria, R. and Cesar, P. (2013) 'Advancements and Challenges towards a Collaborative Framework for 3D Tele-Immersive Social Networking', 4th IEEE International Workshop on Hot Topics in 3D (Hot3D). San Jose, CA, USA , July 15, 2013.

ADVANCEMENTS AND CHALLENGES TOWARDS A COLLABORATIVE FRAMEWORK FOR 3D TELE-IMMERSIVE SOCIAL NETWORKING

D.A. Mauro [1]*, N.E. O'Connor, D. Monaghan, Marc Gowing* [2]*, P. Fechteler, P. Eisert* [3]*, J. Wall, E. Izquierdo* [4]*, D.S. Alexiadis, P. Daras*[5]*, R. Mekuria, P. Cesar* [6]

[1] Institut Mines-Tèlècom, TÈLÈCOM ParisTech, CNRS-LTCI, Paris, France
[2]CLARITY: Center for Sensor Web Technology, Dublin City University, Ireland
[3]Computer Vision and Graphics Group, Fraunhofer Heinrich-Hertz-Institute (HHI), Berlin, Germany
[4]Multimedia and Vision Research Group, Queen Mary, University of London, London, UK
[5]Centre for Research and Technology - Hellas, Information Technologies Institute, Greece
[6]Centrum Wiskunde & Informatica, Amsterdam, Netherlands

ABSTRACT

Social experiences realized through teleconferencing systems are still quite different from face to face meetings. The awareness that we are online and in a, to some extent, lesser real world are preventing us from really engaging and enjoying the event. Several reasons account for these differences and have been identified. We think it is now time to bridge these gaps and propose inspiring and innovative solutions in order to provide realistic, believable and engaging online experiences. We present a distributed and scalable framework named REVERIE that faces these challenges and provides a mix of these solutions. Applications built on top of the framework will be able to provide interactive, truly immersive, photo-realistic experiences to a multitude of users that for them will feel much more similar to having face to face meetings than the experience offered by conventional teleconferencing systems.

***Index Terms*—** REVERIE, tele-immersion, social networking, 3D audio & video, avatars, autonomous behaviours

1. PRÉCIS

It is not clear who first posited that man is an inherently social animal, with the quote being ascribed to both Aristotle and Spinoza, but whatever its origin it is a truism that is reflected in virtually all aspects of our society. Nowhere is this more evident than in the online world. Email revolutionised the way in which we communicate, and the social networking phenomenon, now the most popular form of online activity over email, games and search, continues to push back communication boundaries in a manner not seen since the introduction of the telephone. Social networking sites afford everyone with an Internet connection the opportunity to have a online presence, whether this is fueled by a desire to stay in touch with our family and friends, to meet new people beyond our immediate social circle with shared interests, or even as an effective medium for professional dissemination and outreach.

And yet, despite their success, the communication supported by existing social networking platforms is a poor substitute for real human interaction. They support text, 2D images and video, with interaction being confined to a keyboard/mouse and computer screen. Inevitably, the nuances and subtleties of expression and experience that we take for granted in face-to-face human encounters are lost. As such, current online social networking is inherently limited as a communication medium in terms of the kinds of interaction it can support. You and your like-minded network of fitness enthusiasts can share information and experiences about your most recent aerobics class, but you cannot take the class communally and share the experience together in real time. You can inundate your relatives with postings of photos and videos of recent family excursions, but you cannot invite them to experience them with you as you guide them through a realistically recreated version of the trip. You can share your opinions about who performed best in last nights X-factor or American Idol, using mash-ups and audio/video clips to make your point, but you cant invite your friends to recreate the performance by joining you on stage as backing singers/dancers in a virtual rendition of the winning song.

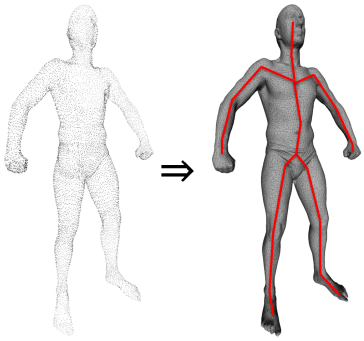
Within the REVERIE project, we believe that the time is ripe to push social networking towards the next logical step in its evolution: to immersive collaborative environments that support realistic inter-personal communication. To achieve this aim, we must break the existing tether to the mouse/keyboard as interaction mechanisms. In fact, this is already happening in the gaming world, as demonstrated by the success of Microsofts Kinect. In a similar manner, we wish bring online social networking into the real-world with interaction and communication based on what people do and say and how they act. In other words, we will bridge the phys-

ical/digital divide, making communication in social networks more reflective of how people interact in the real world.

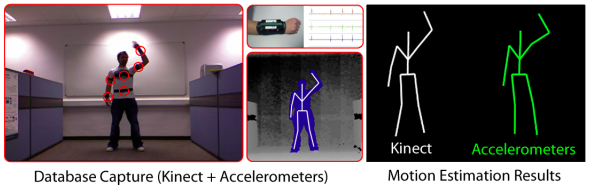
To this end, research is required is on developing/integrating multi-modal and multi-sensor signal acquisition tools for capturing, and subsequently representing and recreating both 3D environments and human activities therein. In terms of representing humans themselves, research is required to allow both full 3D reconstructions and artificial avatars in such environments, with possibly even autonomous versions of the latter. Of course, we must also focus on how users will experience this interaction and the digital online environment in which it takes place. As such, research also targets 3D and immersive audio-visual technologies. Unlike the telephone, social networking gives us the option of determining the temporal nature of our communications – they can be synchronous real-time interactions, or asynchronous off-line experiences based on previously posted content. Clearly, this presents issues not just in terms of capture and display but also with regards to data transmission and the underlying network. For this reason, networking issues also require investigation. Also the audio, and particularly 3d audio, can contribute to an enhancement of the sense of presence and immersion in the scene where as the key requirements here link both to the capturing and rendering stage as well as the streaming of the content in an integrate way with the video streams to achieve lip-synchronization. For this reason binaural spatialization techniques – namely techniques that exploit the nature of our hearing system relying only on a stereophonic signal – will be employed.

Another aspect is the artificial animation of captured humans. In cases where the capturing facilities limit the range of movements, one might like to animate its virtual representation, like walking outside the captured volume. For this reason, a photo-realistic model of the captured user should be generated and used for rendering. The control parameters can be extracted directly from a captured user in a Motion Capture manner, or fully or partly controlled by a motion control unit. Additionally, this parametric approach allows also for changing the appearance of the virtual representation. E.g. the user can wear a nice dress and makeup during the scanning for the model generation while being in casual cloth during a live capture session.

The emergence of these ever more complex 3D environments has led to a range of avatar embodiments, such as full 3D reconstructions, artificial and autonomous avatars, which must interact with the world and other virtual characters in a believable and realistic manner. To achieve this immersive interactive experience, research is being carried out into intelligent techniques which will enable both autonomous characters and users to seamlessly engage in an emotional, dynamic and expressive manner. One such tried and tested technique being investigated to achieve this natural behaviour within autonomous avatars is fuzzy logic. There are numerous rationales for using fuzzy logic: it can be used to obtain a reasonable model when real world data is too complex to be utilised in a dynamic way in real time; it will provide a degree of uncertainty which will result in more believable and realistic actions for autonomous avatar; and fuzzy logic can control behaviours at run-time speeds and meet the requirements of real-time 3D simulations.

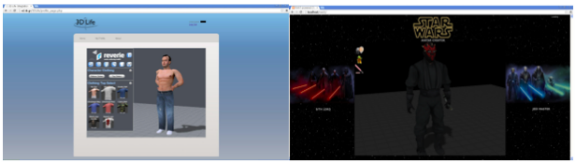


(a) Pose recovery of 3D point cloud via Kinematic ICP



(b) Anywhere motion synthesis from wearable sensors

(c) 3D reconstructions of moving humans in various poses



(d) The REVERIE avatar authoring tool

Fig. 1. Illustrations of some of the key components of the framework.

Acknowledgments

This work has been partially funded by the European Commission under contract “FP7-287723 REVERIE”.