

# Enhancing Social Communication and Belonging by Integrating TV Narrativity and Game-Play

Marian F Ursu

NIM, Department of Computing  
Goldsmiths, University of London  
London, SE14 6NW, UK  
m.ursu@gold.ac.uk

Pablo Cesar

CWI  
P.O. Box 94079, NL-1090 GB  
Amsterdam, The Netherlands  
pablo.cesar@cwi

Doug Williams

Broadband Applications Research  
BT  
Ipswich, IP5 3RE, UK  
doug.williams@bt.com

## ABSTRACT

This paper provides a brief overview of the workshop on *Enhancing Social Communication and Belonging by Integrating TV Narrativity and Game-Play*. Its main objective is to explore new forms of TV-mediated communication between groups of people (such as family and friends) separated by space and/or time that would lead to improving the maintenance and fostering of their social relationships and, ultimately, of their feeling of being together. The investigation considers three perspectives: *socio-cognitive* – defining facets of the experience(s) of togetherness, when group interaction is mediated by the TV screen; *system design* – specifying requirements for TV-centric systems that support social interaction; *system implementation* – analysing enabling technologies.

## Categories and Subject Descriptors

C.2.4 [Computer-Communication Networks]: Distributed Systems; H.1.2 [Models and Principles]: User/Machine Systems—*Human factors*; H4.3 [Information Systems Applications]: Communications Applications; H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems; H.5.2 [User interfaces]: Theory and methods; H.5.3 [Group and Organization Interfaces]: Theory and models; I.2.4 [Artificial Intelligence]: Knowledge Representation Formalisms and Methods; I.7.2. [Document Preparation]: Multi/mixed media; J.5 [Arts and Humanities]

## General Terms

Design, Documentation, Experimentation, Human Factors, Languages, Theory

## Keywords

Interactive Television, Social Television, Mediated Social Interaction, Mediated Communication, Narrativity, Game Play, User Experience, Presence, Personalisation, Enhanced Reality, Networked Media, Video Conferencing, Cinematography, Virtual Directing, Streaming, Multimedia Composition, Knowledge Representation

## 1. INTRODUCTION

Much research into *social interactive television* looked at the development of interactive services, offered whilst watching the TV content, that enable people to enjoy the social experience of TV even when they are in separate locations. Approaches include being able to share the viewed content – by seeing what friends are watching when programmes are broadcast or synchronising the watching of time-shifted content, sending recommendations and sharing favourites, or even sharing the remote control – and to comment on what is seen – by text chat, video conference or through avatars. Examples of prototype systems include AmigoTV (Alcatel-Lucent), CollaboraTV (AT&T), ConnectTV (TNO), and CoSe (Siemens); research endeavours along these lines have been presented and discussed, for example, at the Social Interactive Television Workshop at EuroITV 2007 [1].

In this approach, the TV programme, created by professionals, is the centre-point, and the other activities revolve around and are complementary to watching it. The programmes' performers are different from the audience. The TV programmes (both recordings and live transmissions) are distributed (e.g. broadcast) in a star model – from the distributor to each member of the audience – whilst the aforementioned complementary services provide for direct connections between the members of the audience.

What if the focus of the social interaction is shifted on other activities (i.e. not on watching TV programmes), such as social game-play, recounting of memories or even informal chats, between small groups of people who know each other (such as family and friends) but happen to be in separate locations? What if the professionally prepared TV content is replaced by an audio-visual communication, virtually directed, between them? Could the central role of the television screen be exploited when the performers become the same with the audience? Could the television's well established narrative forms and cinematic techniques be exploited in such a context? Could such an approach lead to a better sense of social belonging and togetherness? Could systems that support such forms of social interaction be built?

This workshop seeks to address these questions, and others that arise when exploring possibilities for devising enjoyable social experiences for people in separate locations when the screen is seen “merely” as a generic display device and not a “television” in its traditional sense.

## 2. AIM

This workshop aims to investigate how existing forms of social communication can be supported by, and new forms can be developed through, a *TV-centred communication*, with a view to enhancing the feeling of social belonging and togetherness between groups of people separated by space and time. More specifically, it aims to explore how family and friends, who are in separate locations, can share moments of fun together whilst playing social games, seeing and hearing each other whilst they laugh with and at each other, sharing instant impressions or recounting past memories, via a virtually directed TV-centred communication.

The workshop will have a particular bias on interaction and communication that is stimulated by or framed within *social game-play*. This should not be understood as merely attempting to transfer computer games onto the TV platform, by substituting the PC screen with that of the TV's, but rather as considering *the complete experience* of playing a game together: the game is just the pretext for socializing together; it is the reactions, the jokes, the laughter, the arguments, the parallel conversations, the recounted memories, and so on, that are in focus.

Social game-play could range from being very structured and rich in audio-visual content, such as electronic games, supported by a game engine, through structured but simpler in terms of interfaces, such as board games, to less structured and requiring no interface, such as guessing games.

The term *TV-centred communication* suggests that the audio-visual communication between groups of people goes beyond the standard face-to-face model of the current video-conferencing systems, aspiring to reach the aesthetic quality of good TV narratives, through employing cinematic techniques in both the capturing and editing of the content. It subsumes both *live* (real-time) and *catching-up* (off-line) communication. For live communications the content processing delays should not influence the fluency of the interaction.

In terms of the live communication, the aim is to create a *seamless connection* between the participating groups/locations. Each location, normally accommodating a *group* of people, must get the best account of what happens in (all) the other locations. This can be achieved only if there is significant automation in terms of both content capture and editing. This function can be denoted as automatic or virtual directing, or *interaction orchestration*. Cues, such as laughter, movement, talking person, ought to be extracted automatically, possibly helped by sensory information and direct instructions from the participants. They determine, on the basis of some embedded interaction or communication intelligence – denoted here as *orchestration intelligence* – the cinematic techniques that are to be applied for both content capture (cameras and microphones) and content delivery (screens and speakers).

Real-time orchestration can have two functions: to simply, but accurately and effectively, support the *recounting* of events happening in other locations; and to *moderate* the interaction.

The off-line communication is about semi-automatically constructing *TV narratives* which best capture key experiences. The content may be captured semi-automatically, during a live interaction, as described above, or manually, during events that happen outside such interactions (for example, as video recording or still pictures). Once the content was captured, TV narratives, possibly interactive, will be assembled semi-automatically. Such narrations could then be incorporated in live communications.

Orchestration, in this context too, has the meaning of a virtual director and editor: it is about deciding which content to record and select, and subsequently, about how to edit it in meaningful (interactive) TV narratives. This route is founded in the ShapeShifting Media Technology [2].

**Scope.** The setting is that of a limited number of households, each containing a *group* of people who know each other and want to stay connected. The focus is on interaction via moving picture and sound that are automatically captured and automatically edited. As output devices, the central role of the TV screen is preserved, possibly accompanied by secondary screens, but other devices, such as surround sound systems and ambient devices, may also be included. For input, each location will have a number of cameras, arrays of microphones and possibly other types of input such as from game consoles and sensors. All the devices should be such that they can be integrated in a household environment.

## 3. INVESTIGATION STRANDS

The proposed investigation considers three perspectives.

**Socio-Cognitive and Perceptual** – *defining facets of the experience of social belonging and togetherness.* Which are the facets of the experience of being together between people sharing a physical space at the same time, and possibly being engaged in social games? Which aspects of the social communication foster that feeling? Which could be transferred and supported through TV-mediated communication? Which TV formats and cinematic techniques (visual: types of shots, camera movement, edits, effects, etc.; and aural: spatial placement of sound, voice synthesis, sound effects, etc.) could enhance the experience of being together in a video conference like experience, but between groups of people in more than two locations? Are there new ways of communication, not possible without the technological support, which could foster the feeling of togetherness? Can there new communication techniques be proposed for *near* real time communication, when the delays due to audio-visual content processing are too large to be unnoticed?

**System Design** – *specifying requirements for TV-centric systems that support social interaction.* What related systems or prototypes already exist and how are they received by the end users? Which communication platforms do they employ? Which of their features could be adopted for the current aim? Are there new requirements/features refined through simulation and user evaluation? What is the economical feasibility of such propositions? What kinds of input and output devices are required, to implement this kind of social communication, and, particularly, emphasising TV based devices, what should their spatial layout be (in the household)?

**Enabling Technologies** – *analysing the capabilities of existing core technologies that could be employed in the development of TV centric systems for social interaction and, at the same time, refining new requirements for them.* Are there existing representation schemes which could (partly) express the envisaged communication intelligence for interaction orchestration? What are the capabilities of their associated reasoning techniques? Are there examples of AI mechanisms that could understand social situations related to the ones described here? Are there models which can *predict* aspects of social interaction? Are there any relevant personalisation techniques? Which features can be extracted automatically from audio-visual streams, possibly helped by sensory input and/or game states, and with what efficiency (processing time) and accuracy (precision)? What

description notations (ontologies) are there available? What multimedia notations for adaptive content exist? What are the capabilities of the low-delay decoding, transmission and encoding algorithms for both audio and video? What are the performances of the multimedia composition and rendering algorithms? How could game engines be integrated in TV based communication? How could different communication platforms be integrated?

The questions listed above are indicative: they guide but not restrict the proceedings of the workshop.

## ACKNOWLEDGEMENTS

The organization of this workshop was partly supported by the Integrated Project TA2, Together Anytime, Together Anywhere (website: <http://www.ta2-project.eu>). TA2 receives funding from the European Commission under the EU's Seventh Framework Programme, grant agreement number 214793. The authors

gratefully acknowledge the European Commission's financial support and the productive collaboration with the other TA2 consortium partners.

## REFERENCES

- [1] Lugmayr, A., and Gelobiowski, P. 2007. Interactive TV: A Shared Experience. TICSP Adjunct Proceedings of EuroITV 2007, Amsterdam (refer to the section on Social Interactive Television, edited by Geerts, D., Harboe, G. and Massey, N.)
- [2] Ursu, M.F., Kegel, I., Williams, D., Thomas, M., Mayer, H., Zsombori, V., Tuomola, M.L., Larsson, H., and Wyver, J. 2008. ShapeShifting TV: Interactive Screen Media Narratives, *Multimedia Systems* 14 (2), ACM/Springer, 115-132.