Interactivity and User Participation in the Television Lifecycle: Creating, Sharing, and Controlling Content

Pablo Cesar
CWI: Centrum Voor Wiskunde en Informatica
Amsterdam, the Netherlands
P.S.Cesar@cwi.nl

Konstantinos Chorianopoulos
Ionian University
Corfu, Greece
choko@ionio.gr

ABSTRACT

Interactive TV research encompasses a rather diverse body of work (e.g. multimedia, HCI, CSCW, UIST, user modeling, media studies) that has accumulated over the past 20 years. In this article, we highlight the state-of-the-art and consider two basic issues: What is interactive TV research? Can it help us reinvent the practices of creating, sharing and watching TV? We survey the literature and identify three concepts that have been inherent in interactive TV research: 1) interactive TV as content creation, 2) interactive TV as a content and experience sharing process, and 3) interactive TV as control of audiovisual content. We propose this simple taxonomy (create-share-control) as an evolutionary step over the traditional hierarchical produce-distribute-consume paradigm. Moreover, we highlight the importance of sociability in all phases of the create-share-control model.

Categories and Subject Descriptors

H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems, J.7 [**Computers in Other Systems**]: Consumer Products, H.4.3 [**Information Systems Applications**]: Communications Applications

General Terms

Human Factors.

Keywords

Interactive television, taxonomy, research survey, interactivity, social TV, user participation.

1. INTRODUCTION

Despite widespread use in academia and industry the term "interactive television" (also referred to as interactive TV, iTV, ITV) is quite ambiguous.

In this work, we define interactive TV (ITV) as an audiovisual user experience that involves at least one user and one, or more audiovisual and networked devices. In contrast to video games, ITV research is more concerned about linear narrative, mash-ups of fixed video-clips, and low to mild levels of user input, while dynamic graphics are employed mostly for overlays rather than

content itself. Nevertheless, the borderline between the two is sometimes vague. Previous definitions were focused on the technological aspects and ignored the fact that even traditional TV is potentially interactive. For example, viewers compete mentally with quiz show participants, or co-operate between collocated groups. Moreover, viewers react emotionally to TV content, they record and share TV content with friends and discuss about shows either in real-time, or afterwards.

Although several ITV developments have followed parallel or even competing paths (e.g. PC versus TV), the convergence of network and rendering platforms has made such distinctions somewhat superficial (e.g. YouTube is available on a TV through AppleTV). There are still significant differences, but those differences regard mostly the context and preferences of the user, rather than the capabilities of the technology. In the following sections, we organize ITV research into three concepts, which stand as an evolutionary step over the traditional model of production-distribution-consumption: 1) creating, 2) sharing and 3) controlling content. In particular we consider the end-user having an active role in each one of these activities, instead of being just a "viewer."

2. CONTENT CREATION

2.1 Authoring Tools

Similarly to any new medium [4], researchers tried to shape the future of interactive digital television using traditional development techniques from the PC and the Web. Apart from a number of innovative solutions (e.g., [25]), the respective authoring tools provided limited support for television-specific issues that will permit the production of innovative television content and services.

TV audiences have become familiar with a visual grammar that requires all programs, as well as presentation styles to be dynamic and surprising, which is in sharp contrast with the traditional usability principle of consistency. An ITV UI might not look like a button or a dialog box. Instead, it could be an animated character, which features multimodal behaviors (e.g., text, motion, and speech). Furthermore, user selections that activate scene changes should be performed in accordance with the established TV visual grammar (e.g., dissolves, transitions, fade-outs). Unfortunately, the majority of the development tools and techniques for a software user interface assume that the final product will be some variation of the WIMP paradigm.

2.2 Content and Metadata Modeling

Content and services authoring involves semantic modeling and presentation integration. Semantic modeling is needed for content

description, thus it is essential for content selectivity and searching. The result of the semantic modeling stage is the metadata modeling of the content. Presentation integration, on the other hand, corresponds to the composition of enriched content from different media assets. For example, presentation integration is the inclusion of subtitles or audio track accompanying a video.

There are a number of content description standards including TV-Anytime, MPEG-2, and MPEG-7. While MPEG-7 is mostly a researchers-oriented solution, TV-Anytime is widely accepted and promoted by the industry. MPEG-2 content stream description is already used in digital television deployment to, for example, populate the electronic program guide. In terms of presentation description, there are a number of standards such as MHP/OCAP.

2.3 User Generated Content

For a long time, the target audience for interactive digital television authoring tools has been professionals working in broadcast companies. The functionality of these tools includes the aggregation of different media objects such as subtitles overlaying the video content and some event handling mechanism for user interaction. In addition to original user generated content, ITV hold a great opportunity for user edits and mash-ups on public content (e.g. subtitles, alternative edits).

The success of social web sites that distribute users' content (e.g., YouTube) has demonstrated that end-users want to change the way they find and consume audiovisual media. Researchers have been developing novel ITV systems that allow users to create their own media. A promising direction includes the development of lightweight authoring tools [5] targeted to end-users for content enriching, such as annotations or virtual edits. Similarly, the ITV community should take into account initiatives from the web, such as interfaces for reusing existing material of in web [24].

3. SHARING CONTENT

There has been a significant body of computer supported co-operative work (CSCW) research on supporting interaction among geographically distributed co-workers, but there is limited investigation in the context of leisure activities, and in particular distributed use of audiovisual content, such as TV.

3.1 Beyond hierarchical content distribution

In the past, TV content in the living room has been provided either by broadcast, or optical discs. There is wide agreement over the specifications for the digital video broadcasting (e.g., DVB-S/C/T/H specifications satellite, cable, terrestrial, and mobile). Furthermore, TV content can be efficiently distributed over peer-to-peer (P2P) networks and broadband connected TV boxes. Alternative and complementary devices and distribution methods have been also considered, such as mobile phones (mobile DTV).

Mobile DTV systems have been designed to complement mobile networks with broadcast and multicast capabilities for spectrum-efficient delivery of multimedia services on mobile devices in both outdoor and indoor environments. The presence of multiple situated base-stations allows personalization to fit the terminal's and physical location's preferences. Several guidelines for the quality of TV reception on the move have been provided [17] and comprehensive user studies on video mobile consumption have been published [20]. Research on P2P technology for television transmission has provided interesting results booth from the

academia perspective [12] and from the industry (e.g., Joost and BBC's iPlayer). The major benefits of this approach are the efficiency of transmission, the diminution of traffic for a specific node (the server), and the social nature of P2P that can be utilized for enhancing the process of sharing and watching television.

3.2 Content enriched communication

Content enriched communication systems consist of technological solutions for integrated interpersonal communication and content distribution. The study of social interactive television is not new. Wellens [28] wrote: "interactive television represents means of linking individuals together by providing each with an electronically mediated representation of the other's voice and visual presence". Social TV systems have been popular in the industry with system such as Amigo TV [9] and Social TV [15]. Most of those systems focus on synchronous communication mechanism between television viewers. User studies include Geerts [13] that compares voice and text chat and highlights the difficulty of text entry for a television environment, Weisz et al. [27] that indicates the connectedness factor, when providing chat capabilities for television viewers, and Shamma et al. [23] that study user behavior when sharing videos on an IM session. Apart from communication methods, sharing solutions such as direct recommendations have been developed [5].

3.3 Presence, awareness and seamless social bonding

If TV watching takes place over a distance and in particular during different times, then the main requirement is to facilitate the communication of basic information that discloses status, preference and activity of distant viewers. Indeed, an important functionality of a Social TV system would be to create the impression of watching TV alongside a group of friends. Social TV provides a shared social context for conversations about the media that they have enjoyed, although not at the same time or place. These types of communication could be considered as the non-verbal part of the content enriched communication, described in the previous section.

Harboe et al. [14] performed an ethnographic study of a traditional TV set-up, which is enhanced with novel communication devices that support lightweight remote awareness. They reported that the ambient communication devices have enhanced the shared experience of TV watching over a distance. TV watching in groups is governed by a set of cultural practices and interaction rules, which have evolved in a way so that co-located viewers can enjoy each other's company. These rules should be reflected in the case of mediated sociability [10].

4. CONTENT CONTROL

Content control corresponds to the activity, human-driven or automatic, of controlling the television content and services.

4.1 Interactivity in content

Jensen [16] has proposed a simple taxonomy of television services from the perspective of the end-user. He differentiates three different interactive television forms: enhanced information that is sent via the broadcast channel (e.g., banners), personalized, and complete interactive (i.e., return channel provision). He points out that, currently, only "low-technology discount solutions" are

provided. The most popular contemporary solution has been the mobile phone text message, which can evolve in the future to Multimedia Messaging Service (MMS) solutions.

In addition to television services and learning applications, the research community has studied interactive narrative. Interactive narrative research addresses the production and creation of story-driven TV content. From that respect, it is probably the most innovative proposition toward ITV content, since story-driven content is the most popular type of TV content. For example, Agamanolis [2] and Ursu [25] have proposed authoring tools and development frameworks for the production of interactive programs, in which the end-users are capable of modifying the outcome of a television program. Nevertheless, there is evidence that increased consumer interaction is not always appropriate for story-driven media programming [26].

4.2 Automation and Personalization

Research on personalized television has brought together researchers from the communities of user modelling and adaptive hypermedia to discuss the applicability of previous theories in the field of ITV. Within that paradigm, there are a few discreet sub-streams of research: 1) personalized EPG, 2) personalized TV content, and 3) personalized advertising.

During the 90's there had been a lot of speculation about the five hundred channels future of ITV. At that time, communication scientists were reporting that viewers recall fewer than a dozen of TV channels [11]. In above cases, the research approach follows the IT perspective and assumes that the viewers, each time they open the TV, need to select a channel or a program out of the available set of broadcasts. Instead, Chorianopoulos and Spinellis [7] proposed the virtual channel metaphor for television viewers. Much of the personalization research has been exploiting meta-data provided during the authoring of content [1]. A complementary strategy is to exploit the activity of the users in order to better describe and understand multimedia content [23] and to consider that television watching is normally a shared experience [19].

4.3 Usability and accessibility of content and services

Television and film have been established as important equalization mechanisms for the dissemination and distribution of cultural materials. In the past, closed captioning has allowed people who are deaf and hard of hearing to be included as audience members. More recently, researchers have enhanced closed captioning with an affective visual grammar [18]. To include some of non-visual information (e.g. music, sound effects, and speech prosody) in closed captions, researchers have proposed graphical representations of the emotive information that is normally represented with non-dialog sound.

In addition to the requirement for inclusive design with regard to senior citizens and people with disabilities, it is necessary to consider the uses and gratifications of the TV experience in the evaluation process [8]. Indeed, the traditional user interface evaluation paradigm involving efficiency and task completion may not be adequate for the assessment of ITV applications. Unless ITV applications are evaluated with consideration for the ordinary TV viewer, they are going to be appropriate only for the computer literate user, thus excluding the TV audience from easy access to information society services.

4.4 Input Devices

Traditional interactive television settings consider the remote control as the single entry point, which imposes a number of restrictions on how people can interact with content at home. We can divide the research on input devices into three major directions: extension of traditional remote controls, re-utilization of everyday objects such as pillows or paper, and adaptation of other personal devices such as mobile phones [6].

Besides extending traditional remote controls, interesting findings have resulted by rethinking the possibilities of everyday objects such as paper and pillows. For example, Berglund et al. [3] have presented an extensive user study about using digital paper and ink for selecting television programs. Finally, apart from current interest of using mobile phones as secondary screens, digital devices such as pen-enabled tablet PCs have been used for annotating a manipulating audiovisual content. For example, [5] presented the "authoring from the couch" paradigm, which allows viewers to enrich television content by using a tablet PC or a mobile phone from the sofa.

5. CONCLUSION

Although we have put much effort into summarizing previous research, we find that the following quote by Dan O'Sullivan[1] has been the most comprehensive definition of interactive television:

> *'Interactive Television is an oxymoron. On the other hand, television provides the most common ground in our culture for ordinary conversation, which is arguably the most enjoyable interaction a person has. We should try to leverage the power of television while creating some channel back from the audience to provide content, control or just a little conversation.'*

In summary, after a comprehensive survey of existing literature, three distinct concepts, which are inherent in interactive television research, can be differentiated: 1) content creation, 2) content sharing, and 3) content control. The taxonomy is proposed as an evolutionary step over the traditional hierarchical produce-distribute-consume value chain. In addition, the taxonomy is not meant for categorizing future research issues or commercial products into one of the three concepts. In contrast, we find that many current efforts might be described by a combination of the three basic concepts. For example, online video sharing systems (e.g. YouTube) are intended for content sharing, but also provide several features for sociability, control of content, as well as content creation.

Finally, we propose that the popular social TV research stream is a higher-level effort with a broader impact on all three concepts. For example: we can have sociability during control (e.g. presence, chat), during distribution (e.g. social recommendation), during creation (e.g. user tags, folksnomies). Further research should emphasize that the television/audiovisual viewer is not just a passive entity, but an active node in a technological and social network that creates, edits, shares and controls content. Interactive social

---

[1] Dan O'Sullivan, http://itp.nyu.edu/~dbo3/proj/#tele retrieved July 2008.

television, ambient intelligence, and user-generated content are three major directions that should be combined in order to provide next-generation interactive television systems. Nevertheless, we do not foresee the end of professionally produced content. Although user-generated content has gained importance, viewers will keep on enjoying strong storytelling in high-quality professional productions.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] Ardissono, L., Kobsa, A., and Maybury, M. 2004 Personalized Digital Television: Targeting Programs to Individual Viewers (Human-Computer Interaction Series, 6). Kluwer Academic Publishers.

[2] Agamanolis, S., & Bove V.M. 2003. Viper: A Framework for Responsive Television. IEE Multimedia. 10 (3): 88-98.

[3] Berglund, A., Berglund, E., Larsson, A., and Bang, M. 2006. Paper Remote: an augmented television guide and remote control. Universal Access Information Society 4: 300-327.

[4] Bolter, J. D. and Grusin, R. 1999. Remediation: Understanding New Media. MIT Press.

[5] Cesar, P., Bulterman, D.C.A., Geerts, D., Jansen, A.J., Knoche, H., and Seager, W. 2008. Enhancing Social Sharing of Videos: Fragment, Annotate, Enrich, and Share. In Proceedings of ACM Multimedia.

[6] Cesar, P., Chorianopoulos, K., and Jensen, J.F. 2008. Social Television and User Interaction. ACM Computers in Entertainment 6(1).

[7] Chorianopoulos, K. and Spinellis, D. 2004. User Interface Development for Interactive Television: Extending a Commercial DTV Platform to the Virtual Channel API. Computers and Graphics, 28(2):157–166.

[8] Chorianopoulos, K. and Spinellis, D. 2006. User interface Evaluation of Interactive TV: a Media Studies Perspective. Universal Access in the Information Society 5(2): 209-218.

[9] Coppens, T., Trappeniers, L., and Godon, M. 2004. AmigoTV: Towards a Social TV Experience. In Proceedings of the European Conference on Interactive Television.

[10] Ducheneaut, N., Moore, R.J., Oehlberg, L., Thornton, J.D., and Nickell, E. 2008. SocialTV: Designing for Distributed, Social Television Viewing. International Journal of Human-Computer Interaction, 24(2).

[11] Ferguson, D. A. and Perse, E. M. 1993. Media and Audience Influences on Channel Repertoire. Journal of Broadcasting and Electronic Media, 37(1):31–47.

[12] Fokker, J., Ridder, H., Westendorp, P., and Pouwelse, J. 2007. Psychological Backgrounds for Inducing Cooperation in Peer-to-Peer Television. In Proceedings of the European Conference on Interactive Television, pp. 136-145.

[13] Geerts, D. (2006). Comparing Voice Chat and Text Chat in a Communication Tool for Interactive Television. In Proceedings of the Nordic Conference on Human-computer Interaction, pp. 461-464.

[14] Harboe, G., Metcalf, C. J., Bentley, F., Tullio, J., Massey, N., and Romano, G. 2008. Ambient Social TV: Drawing People into a Shared Experience. In Proceeding of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1-10.

[15] Harboe, G., Massey, N., Metcalf, C.G., Wheatley, D., and Romano, G. 2008. The Uses of Social Television. ACM Computers in Entertainment 6(1).

[16] Jensen, J.F. 2005. Interactive Television: New Genres, New Format, New Content. In Proceedings of the Australasian Conference on Interactive Entertainment, pp. 89–96.

[17] Knoche, H, McCarthy, J., and Sasse, M. 2005. Can Small Be Beautiful? Assessing Image Size Requirements for Mobile TV. In Proceedings of ACM Multimedia, pp. 829 – 838.

[18] Lee, D. G., Fels, D. I., and Udo, J.P. 2007. Emotive Captioning. ACM Computers in Entertainment 5(2).

[19] Masthoff, J. 2004. Group Modeling: Selecting a Sequence of Television Items to Suit a Group of Viewers. User Modeling and User-Adapted Interaction 14(1): 37-85.

[20] O'Hara, K., Mitchell, A. S., and Vorbau, A. 2007. Consuming video on mobile devices, In Proceedings of CHI, 857-866.

[21] Putnam, R. 2001. Bowling Alone: The Collapse and Revival of American Community. Simon & Schuster.

[22] Silverstone, R. 1994. Television and Everyday Life. Routledge.

[23] Shamma, D. A., Bastea-Forte, M., Joubert, N., and Liu, Y. 2008. Enhancing Online Personal Connections Through the Synchronized Sharing of Online Video. In Extended Abstracts on Human Factors in Computing Systems, pp, 2931-2936.

[24] Shaw, R., and Schmitz, P. 2006. Community Annotation and Remix: A Research Platform and Pilot Deployment. In Proceedings of the ACM International Workshop on Human Centered Multimedia, pp. 89–98.

[25] Ursu, M.F., Kegel, I.A., 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 Journal 14(2).

[26] Vorderer, P., Knobloch, S., and Schramm, H. 2001. Does Entertainment Suffer from Interactivity? The Impact of Watching an Interactive TV Movie on Viewers' Experience of Entertainment. Media Psychology, 3(4):343–363.

[27] Weisz, J.D., Kiesler, S., Zhang, H., Ren, Y., Kraut, R.E., and Konstan, J.A. 2007. Watching Together: Integrating Text Chat with Video. In Proceedings of the SIGCHI Conference on Human factors in computing systems, pp, 877-886.

[28] Wellens, A.R. 1979. An Interactive Television Laboratory for the Study of Social Interaction. Journal of Nonverbal Behavior 4(2): 119-122.