

Experience, narrative and interaction in TV-centred communication

Frank Nack

HCS – University of Amsterdam
Science park 107,
1098 XG Amsterdam, The Netherlands
Tel: +31 20 525 6377

E-mail: lastname@uva.nl

ABSTRACT

Based on the view that changes in user expectations and social demographics alter the traditional use of TV this paper investigates the notion of experience, narrative and interaction in TV-centred communication. We have a look at synchronous and asynchronous consumption of expert as well as user-generated content. We then investigate participatory elements that enhance the feeling of social belonging and togetherness in particular when TV users are separated by time and space.

Categories and Subject Descriptors

H5.4 [Hypertext,Hypermedia]: Architectures, Navigation, User issues. I.7.2. [Document Preparation]: Hypertext/hypermedia, Multi/mixed media.

General Terms

Human Factors.

Keywords

TV-centered communication, experience, narratiity, TV genres, interaction patterns,

1. INTRODUCTION

The TV screen is nothing more than a window to access information, which, not necessarily but often, is presented in narrative form. Due to changes in the public's viewing behaviour, mainly fostering choice and interaction, as well as changing social demographics, i.e. temporal or spatial separation of people in close social relationships, there is currently much talk about expanding TV's typical one direction information stream. The aim is to facilitate a more participatory way of TV consumption, including and in fact

focusing on well-known forms of social interaction, to enhance the feeling of social belonging and togetherness in particular with those who do not share either time or space with the gatherers around the TV screen. Thus, the aim is to turn a common communication receiver, particularly used as a source of entertainment and news either in form of recorded or broadcasted material, into an interactive medium that re-establishes the lost social world. In this paper we investigate the various options and their technological requirements to support richer social interaction in a TV-centred environment. We emphasise in particular on participatory elements inherent in time, space and their influence on the representation and presentation of the content.

2. COMMUNICATIVE TV - DEFINITIONS

The technological environment to be considered in this paper consists of large flat screens as they can be found nowadays in a great number of private homes and a connection to the outside world, either through cable, setup-up box or satellite. The aim of this environment is to provide content that is generated by experts for the distribution through broadcasting, or by non-professionals where the distribution channel will be most likely web oriented. In both cases the goal is to entertain or inform.

Professionally generated content, often mirrored with respect to structure or aesthetics by non-professional content, follows in most cases distinct narrative notions [1]:

- A narrative is restricted to a more or less stringent chronological model of temporal experiences (e.g. suspense and expectation), where beginning and end have been determined and such unfolding a destiny rather than a passage into uncertain territory. Typical genres here are drama or comedy covering crime, horror, or thriller, western, science fiction, or situation comedy (sitcom). Due to their complex structures these genres usually require some authoring expertise, which makes it complicated for amateurs to establish them. However, users are willing to engage with provided

narratives, i.e. by navigating through the story space if they are assisted [2, 3].

- Almost any mode of human behaviour can be understood as narrative, representing essentially narrative forms of memory, myth, history, news, politics, or science. Those are reflected in TV genres such as news, documentaries, daily soap operas, talk shows and reality TV, and so forth. As these formats require less structured continuity there are several mechanisms how computer can automatically generate stories in such genres based on user generated content [4, 5].
- A narrative as a poststructuralist inspired attack on the dominant culture (see the two bullet points above) in form of oppositional, resistant or subversive material. Typical genres here are avant-garde documentaries or comedy shows. These types of program can be also be established in an automatic fashion, based on user generated content [6, 7].

Whatever the content, the way in which it is consumed is passive, semi-active or interactive. Viewers can sit alone or gather in groups around the screen and consume the material at the same time – or consume the same content in either a spatial and/or temporal asynchronous way. In fact, narrative universes can become reversible and no longer need to reflect the psychological cause and effect as freed social participatory environments facilitate people to consume different programs a- or synchronically and communicate exactly about that.

The essential aspect of the narrative, seen as an event, is that it establishes an experience in the audience. In the context of TV-centred media distribution we understand an experience as the nonidempotent alteration of the cognitive map and/or related cognitive processes of the one who has the experience, derived from direct observation or participation in a contextualised event or activity over a certain period (short or long-term) [8].

It is in particular the different ways of social interactions, in different temporal and spatial situations we will focus on in the ongoing discussion.

3. INTERACTION MODES IN COMMUNICATIVE TV

The essential elements that have a large influence on the users expectations with respect to expressiveness of the content as well as their willingness to actively engage or passively consume depend mainly on the idiosyncratic personality of the individual viewer. The temporal and spatial context will, however, influence his or her behaviour substantially. In this section we investigate the synchronic and asynchronous relationships between the concepts time, space and content and their inherent

possibilities and challenges for communicative and experiential TV. Their role for user modelling is mainly neglected in the discussion, even though this point is relevant. We also do not address the classical notion of watching TV at the same time, same space with the same content, as this is the state of the art.

3.1 Same time, same space, different content

This is the classical situation of people gathering in the living room for watching TV, only that here we address the current development of multiple screens in one location on which different programs can be enjoyed at the same time. As people share the same experiential space they will make use of natural observation techniques to share their experiences: they look at or talk to each other. In cases where they are immersed in the program or modality-wise shut off, e.g. by the use of headphones, they could make use of additional technological aids, such as ambient light that either describes the level of activity in the other program, as portrayed in Figure 1, or signals the change of channels on the other screen, e.g. by swiftly showing the first minute of the new program in a small semitransparent window, e.g. in the bottom right corner of the own screen.



Figure 1: Ambient light used to indicate the excitement level of a program shown on a different screen in the same space

3.2 Same time, different space, same content

The crucial aspect here is that people share the same physical time period in front of a screen to explicitly watch the same content. This does not mean that they have to be in the same location to share the same program. An example might be that parts of a family live in Amsterdam and the other temporarily in Moscow. If people in such a situation wish to share a program, e.g. a soccer game, a telenovella, or a talk show, they want to be aware of the emotional as well as reflective state of the others. This might be achieved by providing a semi-transparent window on the screen captured by a camera in the other location that portrays visual reactions, as displayed in Figure 2 on the next page. An audio commentary can additionally be overlaid on the actual sound of the program, which is especially effective in rooms with a surround sound system, as here a special location can be selected.

In cases that they wish to interact in real time with each other on interactive content, e.g. traversing through the story space of an i-narrative, the described real-time communication methods can be used for synchronisation. For example they can use the additional audio layer to communicate preferences for choices in which direction a story should develop, or make use of indirect communication mechanisms, such as sms-messages, that directly influence the story engine. The important aspect for all type of interactions, but most importantly for the indirect ones, is that enough feedback is provided for providing the user with the feeling that he or she is in control.



Figure 2: Transparent image overlay of co-watcher in a different location.

3.3 Same time, different space, different content

This case is similar to the case described in section 3.1. The difference here is that a spatial asynchrony is introduced. Imagine again the same constellation of a group of people who are temporarily situated in two cities but still wish to see what the others are doing without necessarily sharing the same TV viewing experience. For example, while one group watches the national news the other person or group consumes an automatically produced documentary about the happenings at the other location covering the previous week. The same audio-visual interaction methods as described in section 3.2 can be used here for quick validation of content or presentation of experiences. As one of the members perceives an interactive program (the automatically generated documentary) it might be necessary to install an additional ‘interaction’ layer, which allows the augmentation through annotation of the automatically generated content to facilitate highlighting, commenting, or altering of the content for later replay by the other group. This layer can also be used for direct manipulation in form of pointing so that, in cases of questions, the particular parts under consideration can be immediately localised. Here additional input mechanisms are required that could make use of double-touch technology for the screen, pointing devices or completely separated input devices, e.g. electronic drawing boards or

speech input, that are temporarily synchronised with the incoming content stream, allowing real time annotation.

3.4 Different time, same space, same content

The interesting change in parameters in this case is the asynchrony in time. The relevant situation is that people are interested in perceiving a program on the basis of experiences of others who already saw it but who are not around at the moment to be asked about it. The asynchrony of time allows a presentational flexibility not possible in the previous cases, which all required real-time presentation and hence generation. However, the challenging aspect of this scenario is that we need real-time annotation mechanisms that allow rich augmentation of the content so that later the experience can be partially re-established. At the same time we have to invent a new iconography to describe the augmentations on the screen so that the current user is in the position to perceive the experience of the other(s) but can also skip it and follow his or her own information or entertainment needs.

The representation of ‘experience’ is a crucial aspect. As mentioned earlier, there are several layers that need to be addressed. There is the context that needs to be represented. It would be interesting to see if a person watched the program fully or in a time constraint mode. It is also helpful to know what type of actions the person had performed during the consumption and why, resulting in a capture of actions, such as fast forward or –backward, slow motion, zoom in or pause with additional explanatory audio, visual or textual material. Most importantly, it is important to get an idea about the involvement of the other person with respect to interest and emotion. This material can be collected in audio, visual or textual form but can also be based on biometric data measurements. All this type of information can facilitate the recreation of experience. Capture and alignment of this data with the original content needs to be performed in temporal synchronisation (making use of technology already described in the previous sections) whereas the final instantiation and exploitation of the annotations can be done off-line. For example, we could calculate emotional expressions from captured face data and make use of that during the presentation of the experience enhanced content, e.g. in form of ambient light or additional iconic descriptions as overlays. At the same time the system could also make use of information gathered from actions, e.g. skipping scenes or collecting additional information from the web, for enhancing the augmentation layer of the original content. In that way the system could turn a documentary about World War II into an experienced enhanced documentary where the grandparents’ experiences (those during the war as well as those made during the perception of the program) provide an additional layer of information.

The representation problem is intrinsically connected with the presentation problem [9], as ways needed to be found

so that the additional information does not disturb the own experience but still allows alteration and comparison. As a result the presentation might make use of black bars at the top or bottom of the screen to present temporally available interactive icons that support the access as performed by others. Other options might be interactive visual or audio overlays on the content itself, or divergence methods in form of story maps to facilitate easy navigation, or split screen mechanisms. With respect to the emotional aspects necessary to re-establish an experience, one can use particular features in the collected material, e.g. smiles in detected faces, or can exploit biometric data measured during the perception of the content. The challenging aspect is that the annotations can serve as raw data and thus can be also annotated. The established layers of annotation can then be exploited by the system, for examples as summaries of programs, where particular expression aspects are highlighted, depending on the user's needs.

3.5 Different time, different space, same content

The situation covered here is that people in a different location would like to experience a program that has been watched by others before or who watch an automatically generated program, based on experience-enhanced material already available, that the other side has not yet seen either. For example, people in Moscow watch the summary of last week's TV-consumption by people in Amsterdam. Another example might be the more epic documentary of the children's spring holidays. In this situation the system relies even more on the representation and presentation methods already described, as a potential human-to-human information exchange, say for example a quick chat in the kitchen or hall, is not possible. Additional methods to those described in section 3.4 are not necessary but for the situations covered by the given context they have to be fully worked out. It is also important to mention here that all sort of automatically generated material needs to be clearly identifiable as such.

3.6 Different time, different space, different content

Besides the already mentioned problems in the context of experience representation, generation and presentation it is important to realize that there will be a growing divergence in interest between people if the separation covers a long period of time. The aspect that needs to be considered here is tracing the changes in people's behaviour and preferences. This is more a question of user modelling, which can be incorporated into the information network sketched out in the previous sections 3.1 – 3.5. If systems should be able to support richer social interaction in a TV-centred environment they have to have a notion about the particular individual, i.e. interest, preference, repugnance, and contextualised watching history. The analysis of challenges of user modelling in the context of

communicative TV is, however, beyond the scope of this paper.

4. CONCLUSION

In this article we outlined in particular participatory elements of experiences that enhance the feeling of social belonging and togetherness when TV users are separated by time and space. The discussion showed that methods addressing spatial separation in real-time communication ask for direct interaction support, which to some extent reduces the reproducibility of experiences, as they will not be stored. In cases of a temporal separation between TV users, beyond the time-zone differences, the capture of an experience is an essential element. The field of experience capture is still in its infancy but further advances are required to establish communication mechanisms in TV environments as described in this article. We showed, however, that experienced-enhanced content could establish forms of intimate communications freed from time and space. The visualisation of this type of material, as well as its production requires additional research, as the cognitive as well as procedural constraints are not yet clear.

5. REFERENCES

- [1] Cubitt, S. (2002). Spreadsheets, Sitemaps and Search Engines – Why Narrative is marginal to Multimedia and Networked Communication and Why Marginality is More Vital than Universality. In *New Screen Media – Cinema/Art/Narrative*, edited by Martin Rieser/Andrea Zapp, pp. 3 – 13, BFI Publishing, London.
- [2] Brooks KM (1999). *Metalinear Cinematic Narrative: Theory, Process, and Tool*. MIT Ph.D. Thesis.
- [3] Ursu, M.F., Kegel, I., Williams, D., Thomas, M., Mayer, H., Zsombori, V., Tuomola M.L., Larsson, H., Wyver, J., and Zimmer, R., (2008), *ShapeShifting TV - Interactive Programmes*, ACM/Springer Multimedia Systems Journal, special issue on Interactive Television, Vol. 14 (2), pp. 115–132, 2008
- [4] Stefano B. (2006): *Vox Populi: Semantic annotations and rhetorical editing for automatic video documentary generation*. University of Amsterdam Ph.D. Thesis
- [5] Davenport, G. and Murtaugh, M. (1995). *ConText towards the evolving documentary*. In *Proceedings of the Third ACM international Conference on Multimedia*, pp. 381 – 389, San Francisco, California, United States, November 05 - 09, 1995. ACM, New York, NY.
- [6] Harrison, C., Amento, B. (2007). *CollaboraTV: Using Asynchronous Communication to Make TV Social Again*. In *Interactive TV: A Shared Experience*, Adjunct Proceedings of EuroITV. Amsterdam. 218–222.
- [7] Terminal Time: <http://www.terminaltime.com/>
- [8] Nack, F. (2003) *Capturing experience - a matter of contextualising events*. Proceedings of the 1st ACM MM WS on Experiential Telepresence (ETP 03), Berkeley, CA, USA, November 7th, 2003, pp. 53 - 64
- [9] Douglas J.Y. and Hargadon, A (2004). *The Pleasures of Immersion and Interaction: Schemas, Scripts and the Fifth Business*. In *First Person – new Media as Story, Performance, and game*, ed. Noha Wardrip-Furin and Pat Harrigan, pp. 192 – 206, The MIT Press, London, England.