The potential and pitfall of Interactive TV Technology: An Empirical Study

Soha Maad

Fraunhofer
Institute For Media Communication (IMK)
Germany
Soha.Maad@imk.fhg.de

Abstract.

Research and development in the area of pervasive technologies and in particular interactive television (ITV) did not deliver so far the aimed objectives. Providing ITV programs that capture the attention and edutain (educate and entertain) the ITV viewer while attending his/her preferences, needs, requirements, and limitations (e.g. physical handicap) is still far from our reach. This is attributed to several factors including the limitation of the prevalent paradigms of interaction with digital TV, the limitation of the ITV technology, and the lack of the development of truly interactive and universally accessible multimodal ITV content. In this paper, the potential and pitfall of interactive TV are exposed by over-viewing state of the art research in the development of paradigms of interaction with digital TV and by discussing the results of an empirical study on the use 3D, virtual reality (VR), and artificial intelligence (AI) technologies in developing novel paradigms of interaction with digital TV.

1. Introduction

This paper presents the potential and pitfall of ITV by considering an empirical study that supports a holistic design of ITV. The second section introduces interactive TV by giving a device description, a feature description, and by depicting a conceptual model of the ITV production delivery chain. The factors determining the growth of ITV, the outlook of the future ITV content, and the promise of ITV are also presented in section 2. The third section overviews the challenges of prevalent paradigms of interaction with ITV and the potential added value of using 3D, virtual reality (VR), and artificial intelligence technologies to develop novel paradigms of interaction involving virtual human characters in the presentation of ITV programs. Section 4 assesses the potential and pitfall of using 3D, VR, and AI technologies in developing the new paradigm of interaction mentioned in the previous section. This assessment is

supported by an empirical study conducted as a part of a holistic design for ITV. The empirical study is motivated by MARLIYN (Multimodal Avatar Responsive Live News Caster) developed at Fraunhofer Institute of Media Communication in Germany as a prototype for intelligent interactive TV. Section 4 is divided into three subsections introducing the aim of the empirical study, presenting MARILYN and the developmental issues related to it, and reporting the results of the empirical study. Section 5 discusses the major findings of the empirical study and draws general guidelines for the direction of future research to develop ITV and promote its universal accessibility. The conclusion summarizes the paper and motivates a discussion concerning the potential and pitfall of the use of emerging technologies for authoring ITV content.

2. About Interactive TV

2.1 Definition

Several attempts are made to define the term interactive TV. This section presents three definitions that aim at describing the device, the functionality, and the associated conceptual model.

Using a *device description*, a digital TV is defined in [1] as a device which features Digital Video Broadcasting (DVB) reception, persistent local storage (Hard Disk Drive — HDD), and data processing abilities.

Using a *feature description* [2], interactive TV is considered as video programming which incorporates some style of interactivity — be it with data on video , graphics on video, video within video, or retrieving video programming and possibly recording it on a digital hard disk drive for further use. To the viewer, enhancements appear as graphical and sometimes purely informational elements on the screen overlaying and interactive or accessible information data. The ITV makes it possible to use the data bandwidth in providing services unrelated to video data, such as stocks and news tickers, e-mail, chat, games, etc.. These value-added services give broadcasters and content providers the potential to differentiate their products and to attract a wider range of viewers. As such, interactive TV aims at combining the traditional style of TV watching with the interactivity of the personal computer and the internet.

A *conceptual model* of ITV [3] may describe the associated production-delivery chain. The latter can be conceived as consisting of three parts (depicted in Figure 1): i. content authoring[1] and encoding at the production end; ii. transmission[2] which refers

---

[1] An ITV content is composed of text, audio, photos, videos, and the corresponding paradigm of interactivity with this content. ITV content authoring varies from the use of packages that helps in end user development of this content to the use of high level programming languages such as Java and C for developing ITV content that can access the full functionality of the hardware at the ITV viewer end.

[2] Special hardware, software and middleware are used by content providers for performing appropriate digital compression of the ITV content (in the form of 0's and 1's) according to some digital compression standards (such as MPEG standard) prior to the transmission of this

to the broadcast of the digitally encoded content; and iii. delivery, decoding, and play-out at the ITV viewer end (the hardware at the ITV viewer end may be a set-top box[3] connected to a TV or a TV enabled PC.). The entire ITV production-delivery chain involves systems operators, set-top box technology providers, broadcasters, and content providers.
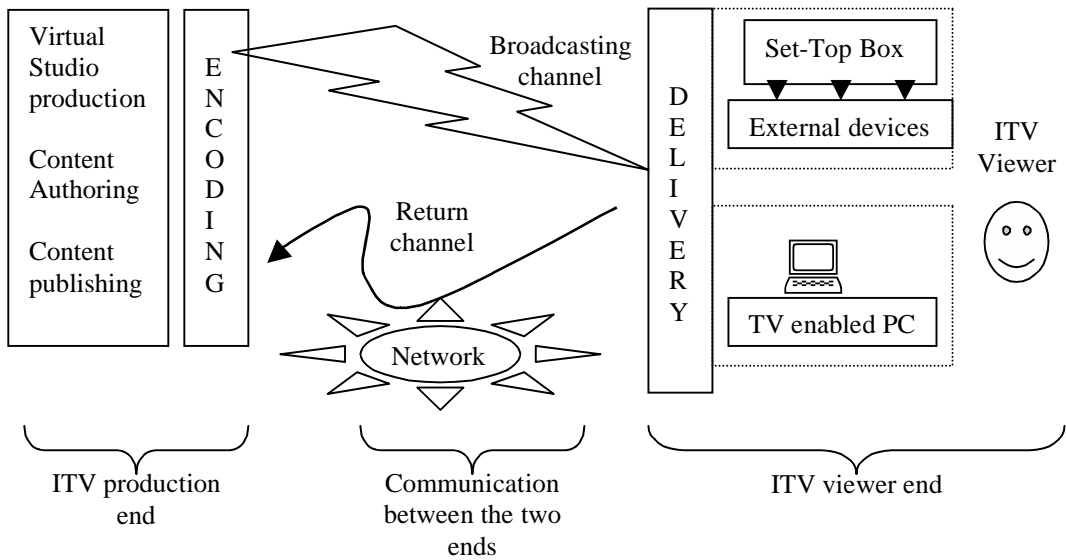
Fig. 1. The general model of the ITV production — delivery chain

2.2 The growth of ITV

A rising concern in the digital TV market is the role of interactive television (ITV) in boosting the growth of this market and constituting a new industry. The growth of ITV depends on several factors including [4, 5]:
First, the improvement in the ITV production-delivery chain This chain involves producing the media content of a TV program, broadcasting it on a particular TV channel and receiving it by the ITV viewer system.
Second, the convergence of different media Several technology and management issues are associated with the integration of traditional analog information services,

content. This digital compression is also known as encoding. The encoded ITV content is then transmitted (broadcasted) as digital signals to the ITV viewer hardware.
[3] Set-top boxes are devices that originated from the idea of the Nintendo game box. A set-top box involves four components: a network interface to receive and send data, a decoder to decode the encoded broadcasted digital signals, a buffer to ensure continuous play back to the viewer, and a synchronization hardware to synchronize audio and video streams.

particularly voice and video, with digital data services. This integration involves both converged applications (streaming media, integrated messaging, digital Telephony, and videoconferencing) as well as the underlying technology infrastructure required for successful deployment.

Third, the rising interest of TV viewers in having access to interactive TV services and programs. This involves the interest of ITV viewers in gaining more knowledge about the TV programs and shows presented (e.g. knowing the biography of the players) and in participating in special TV programs that demand feedback and commentaries from ITV viewers on a real time basis (life broadcasting). Examples of the latter type of TV programs include TV shows, games, discussions, and surveys.

Fourth, the advance in the medium and paradigms of interactivity between the ITV viewer and the TV.

2.3 The outlook of the future ITV content

Media digitization has changed the whole TV production-delivery chain. Digital broadcast applications and interactive internet services are converging to interactive multimedia services for a broad audience [6].

The ITV content evolved from video-on-demand[4], enhanced TV[5], Teletext[6], Electronic Program Guides (EPG)[7] to TV with more advanced paradigms of interactivity.

Emerging ITV content includes: personalized TV news, personalized advertising, interactive TV program guides, and interactive TV games and shows, as well as internet style of application such as e-mail, home banking etc.

2.4 The promise of ITV

Swedlow [2] sees the promise of ITV as a shift from a broadcast, passive, linear, entertainment viewing experience for millions around the world to an on demand, participatory, non-linear, infotainment, advertising targeted, broadband, two-way communications platform. The ITV viewer will have the opportunity to control this delivery medium and use it for information and communications. At home the ITV viewers may have their own mini-TV production studio in their living rooms.

---

[4] Video-on Demand (VOD) services allow people to digitally record programs and/or watch programs whenever they wish [7].

[5] Enhanced digital television involves the fusion of video, audio, images and text. Viewers can simultaneously watch an enhanced TV production and, in the background, receive hundreds of megabytes of additional video, audio, text and images related to the program [7].

[6] Teletext consists of information, such as news and sports, which are viewed on a television set capable of viewing it, or via a Set-top Box attached to a standard TV set [7].

[7] EPGs are onscreen guide that helps make it easier for viewers to choose and/or go to the large number of channels, Pay Per View events, and shows available particularly in a digital environment [7].

3. Challenges to prevalent paradigms of interaction with ITV

Several limitations are acknowledged in the prevalent paradigms of interaction with digital TV regarding: i. the degree of interaction; ii. the medium of interaction; iii. the availability of the TV programs; iv. the viewer satisfaction; and v. the coverage of a wide audience. The following items justify the enumerated limitations.

i.   A preconceived interaction is only possible
ii.   The possible medium of interaction is limited to the use of remote control by the ITV viewer. Clickable interactivity with pre-recorded video frames is so far the sole paradigm of interaction.
iii.   A real time video recording of the TV programs, if needed, imposes limitation on the 24 hours availability of the ITV program and the continuous on demand engagement and interactivity with the ITV viewers.
iv.   Watching TV programs differs in spirit from the surf of internet and the interaction with computer programs. This difference stems from the ITV viewer's needs and preferences. An ITV viewer looks for more entertaining means to gain knowledge and enjoy his/her time. Hence, a minimum level of action is aspired from the ITV viewer to get what he/she wants. Preconceived clickable interaction with pre-recorded video tends to be less attractive to the ITV viewer.
v.   Physically handicapped ITV viewers (visually impaired, deaf, etc.) might have difficulty in adapting to the preconceived clickable interactivity with the ITV program, hence the wider audience including physically handicapped ITV viewers is not reached.
vi.   TV programs are presented in a specific language and there is no option for the ITV viewer to choose a preferred or understandable language for him/her.

There is a promise of enhanced interactivity and universal accessibility[8] when resorting to emerging 3D, virtual reality (VR), and artificial intelligence (AI) technologies in authoring ITV content. Maad [3, 9] identified a prominent role for 3D virtual human avatars in providing a potential support for the development of an intelligent interactive TV (IITV) platform. The use of 3D virtual human avatars in authoring ITV content can potentially:

- help in bypassing the rigidity of traditional styles of TV production that imposes limitations on the 24 hour availability of life broadcasted TV content;
- facilitate on-demand continuous availability and engaged interactivity with the ITV viewer;

---

[8] The concept of universal access is supported, among others, by research in the area of User Interface for All (UI4ALL). This research is described in [8] as "*being rooted in the concept of Design for All in HCI, and aiming at efficiently and effectively addressing the numerous and diverse accessibility problems in human interaction with software applications and telematic services*".

- provide greater accessibility to the handicapped ITV viewer by providing speech interaction for the blind and hand gestures for the deaf;
- engage the ITV viewer in the ITV production by giving him/her an on-demand choice of virtual human characters, of the mode of interaction, and of the speaking language used in the presented TV program. The multilingual feature of ITV programs can potentially widen the scope of ITV audience.

4. Empirical Study

An Empirical study on the use of 3D, VR, and AI technologies to improve the paradigms of interaction with digital TV was conducted during CEBIT, the large computer and telecommunication trade fair held in Hanover.

A population of ITV viewers, with various background and ages, were interviewed on their preferences for various paradigms of interaction with digital TV and their opinion on the advantages and disadvantages of the use of virtual reality techniques and of the latest breakthrough in the development of media technology to support this interaction. The population of ITV viewers was exposed to MARLIYN, Multimodal Avatar Responsive Live News Caster, developed at Fraunhofer Institute of Media Communication [3]. MARILYN gives insight on the use of 3D, VR, and AI technologies for developing ITV programs presented by virtual human characters.

MARILYN involves a virtual human avatar presenting interactively multilingual, live financial news content. The medium of interaction with the avatar is a special remote control with red, green, yellow, gray, up, down, left, and right pushbuttons pressed on demands for particular content requests.

4.1 The aim of the empirical study

The aim of the above described empirical study is to support a holistic design for ITV applications. A holistic design for ITV merges the already existing three important disciplines of design for digital television identified in [1] as:

1) Broadcasting and consumer electronics engineering design that involves the design of the broadcasting computing model. This may also involve designing on-screen displays and remote-controls.
2) Ethnographic study of media consumption at home that explores different designs in a situated context (e.g. The ethnographic study of a set-top-box use [10] identified the need for different design for family viewing in the living room vs. single users in their bedrooms)
3) Interactive multimedia content creation that supports both interactive and passive users. Examples include prototyping personalized TV news [11], and usability testing of electronic program guides (EPG).

The empirical study helped in developing new concepts for ITV programs involving 3D virtual human avatars, and proposed new uses for 3D virtual human avatars for devices other than home TV's.

4.2 MARILYN

The empirical study was motivated by MARILYN (an Interactive TV content authored at Fraunhofer Institute of Media Communication — ITV competence in Germany). MARILYN "Multimodal Avatar Responsive Live News Caster" is an innovative prototype for Interactive Business Television. It involves a 3D virtual human avatar presenting a particular financial news content selected by the ITV viewer in a chosen language. MARILYN (depicted in Figure 2) features:

- A 3D virtual human avatars responding to the ITV viewers with voice and facial expressions. The 3D virtual human avatar character can be chosen. Marilyn Monroe was the default virtual character used. With a suitable almost natural expression the Marilyn avatar reads out the news from the world's stock exchanges;
- A financial content consisting of live financial indicators from various international stock exchanges, and of answers to pre-conceived frequently asked questions and live news;
- An interactive interface template operated through a remote control. The specially designed remote control includes red, green, blue, up, down, left, and right keys for navigation;
- Bilingual option (English — Deutsch);
- Content generation at the content production end (that consists of a server providing audio and facial & body animation streams on demand as well as the live financial content), digital transmission of the generated and encoded content, and play-out at the viewer end.

Five stages are adopted in the development of MARILYN [3, 9]:

Stage 1. This stage involves the development of a template for an animated virtual human avatar (3D humanoid) enriched with facial expressions, voice replies, and hand gestures. The template would also give the option to choose a desired virtual human character.

Stage2. This stage involves developing several multilingual templates for live financial news content (tickers, and news) taken from various international stock exchanges.

Stage 3. This stage involves establishing a bi-directional communication medium between the virtual human avatar (TV presenter) and the financial content. The virtual human avatar is conceived as a mediator between the ITV viewer and the financial content, thus providing the mean for interactivity with this content. The virtual human avatar has two responsibilities: (1) to serve the ITV viewer intelligently by channeling the ITV viewer's demand to the appropriate ITV content; and (2) to edutain (educate and entertain) the ITV viewer through voice replies, hand gestures, and facial expressions.

Stage 4. This stage involves establishing an interactive communication medium between the animated virtual human avatar and the ITV viewer. In this respect various levels of interaction can be identified: the constrained, the semi-constrained, and the open-ended levels of interaction.

- The constrained interaction reduces the interaction between the ITV viewer and the virtual human TV presenter to preconceived modes of interaction that follows a menu driven style of interaction with audio and visual multimedia enhancement.
- The semi-constrained interaction relies on conventional artificial intelligence techniques for retrieving and serving the ITV financial content on demand. The boundary of the financial content and the boundary of the intelligent interaction are pre-defined in this case.
- The open-ended medium of interaction suggests access to on-the-fly and on-demand authored ITV financial content. As such the virtual human avatar gives the ITV viewer an indirect mean for the access to the ITV content and the potential to intervene in authoring the ITV content to suit his/her edutainment needs.

Stage 5. This stage involves checking the compliance of the used technology with prevalent standards for ITV technology**.**

Figure 2. MARILYN: " Multimodal Avatar Responsive Live News Caster"

4.3 The results of the empirical study

The results of the empirical study are framed into various profiles for ITV viewers and producers, and a qualitative assessment of the viability of virtual human avatars for ITV and their future uses.

Nine categories of ITV viewers were identified these are: the conservative, the novice, the busy, the lazy, the humanist, the techno-phobic, the old generation, the young generation, and the ambitious. A sample reply from an ITV viewer in each category is as follows:

For the *conservative*, 3D technology is not yet fully mature for authoring ITV content.

For the *novice*, 3D Virtual avatars presenting ITV programs is an interesting and a fancy idea.

For the *busy*, 3D virtual human avatars are acceptable for presenting TV programs as long as he/she can see these programs at his/her convenient time.

For the *lazy*, all medium of interaction with the ITV are needed including voice interaction as well as interaction with virtual human avatars.

For the *humanist*, the human element should never be taken out from a TV program. A mixture of both (human and virtual) may be ok.

For the *techno-phobic*, virtual human avatars are not viable for TV programs no matter what is the reason to use them.

For the *old generation*, there is no need for virtual human avatars or for new modes of interaction. Grandmothers and grandfathers are happy with the TV as it is.

For the *young generation*, 3D virtual human avatars are useful for TV on demand and for personalized TV programs.

For the *ambitious*, there is no point of developing something especially for ITV since all devices will converge to a single edutainment box.

Three categories of ITV producers were identified these are: the trendy, the humanist, and the technology selective. A sample reply from an ITV producer in each category is as follows:

The *trendy* ITV producer sees that the use of 3D avatars is a way to differentiate his/her TV content and pretend delivering something new.

The *humanist* ITV producer says that the human TV presenter will not be happy to be replaced by a virtual avatar. This will put him/her off.

The *technology selective* ITV producer acknowledges the fact that despite the huge 3D development efforts, avatars are failing to take the place of the human.

Overall, all replies were really personal and it was difficult to generalize. However, we can conclude the preference of the human TV presenter over the virtual TV presenter. Regarding the potential uses of 3D virtual avatars and the preferred modes of interaction with digital TV, we have a novelty in various replies*.*

It is foreseen that 3D virtual avatars will be useful for TV on demand for the young generation, and for developing personalized TV programs. 3D Virtual avatars will also find prospect for devices other than home TV such as: Mobile TV, PDAs, and home devices.

5. Future Research

This empirical study asserted previous findings in similar conducted studies [10, 12, 13] and suggested new ones:

1) The need to study the use of the technology in its situated context

2) The need for a new metaphor for the viewer interaction with ITV. Chorianopoulos et al (2002) [1] argued that the currently dominant metaphor for the personal computer i.e. the desktop metaphor is not appropriate for television, because it is

adapted to fundamentally different users and task sets. He added that there is a need to design a user interface for digital television that considers the user as a television viewer. There is a need for rethinking the traditional set of human interface paradigms in light of an evolving complex set of features, such as digital broadcasts coupled with rich metadata, digital recording of programs and random access playback, local processing logic, dynamic presentation of content and services, and most crucially a different context of use.

3) The viewer sees the ITV as a medium of entertainment and passive discovery [13]. Hence, minimum efforts should be taken by the viewer to access the desired ITV content.

4) The need to study the preferences of the ITV viewer and involving him/her in the ITV content design phase:

> "*If companies want to restore growth to their brands, they have to drop the one size fits all, mass marketing approach and find out how they can make their brands more meaningful to different kinds of people. They tend to want to understand their brands, when what they really need to do is understand consumer lives and fit their brands into them ? They need to be trend watchers, ethnographers, they need to be more culturally savvy ?"*
>   *Susan Fournier, Harvard Business School*

In light of the results of the empirical study, a future research agenda is drawn for the development of ITV involving:

- The development of novel software system development (SSD) approaches for ITV Interface design that differs from the desktop interface design, and for the support of the user (ITV viewer)/developer (content provider) collaboration
- The development of novel foundations for artificial intelligence (AI) aiming at supporting open-ended/non preconceived modes of interaction with various categories of ITV viewers. This would involve the use of diverse medium of interaction that deploys voice recognition, natural language understanding, and pattern recognitionTechniques, and the support of the decision making activity in meeting the ITV viewers edutainment needs
- The development of novel ITV production-delivery chain models that support the ITV viewer involvement in the ITV content authoring.

6. Conclusion

This paper discussed the potential and pitfall of ITV with reference to an empirical study conducted as a part of a holistic design for ITV involving the use of the latest developments in 3D, VR, and AI technologies for authoring ITV content. The results of the empirical study were framed in a classification of ITV viewers and ITV producers and in the proposal of a research agenda for future development of ITV. A motivating discussion emerges out of the results of the empirical study. This involves the exploration of new means for assessing and exploring the prospect of the use of emerging technologies such as VR, AI, and 3D technologies in authoring ITV content.

6. Acknowledgement

References

1. Chorianopoulos, K., Spinellis, D., A Metaphor for Personalised Television Programming, ***SPRINGER LNCS proceedings of the***7th ERCIM Workshop "User Interfaces for All", Paris, France 2002.
2. Swedlow, T., 2000: Interactive Enhanced Television: A Historical and Critical Perspective, Whitepaper, commissioned by the American Film Institute-Intel Enhanced Television Workshop for distribution to participants of its 2000 activities, and on its Web site, **http://www.itvt.com**.
3. S. Maad**,** MARILYN: A Novel Platform For Intelligent Interactive TV (IITV), ***In proceedings*** of Human Computer Interaction International HCII2003 Conference, Crete, Greece, July 2003.
4. Meyer L., Fontaine, G., Development of Digital Television in the European Union, Reference report/1999, IDATE Institut de l'audiovisuel et des telecommunications en Europe, June 2000
5. Peters, J.J., A history of television, European Broadcasting Union, 2000.
6. http://www.imk.fraunhofer.de
7. http://www.itvdictionary.com/
8. Stephandis, C., From User interfaces for all to an information society for all: Recent achievements and future challenges, Proceedings of the 6th ERCIM Workshop "User Interfaces for All", October 2000, Italy
9. Maad, S., Universal Access For Multimodal ITV Content: Challenges and Prospects, ***SPRINGER LNCS proceedings of the*** 7th ERCIM Workshop "User Interfaces for All", Paris, France 2002.
10. O'Brien, J., Rodden, T., Rouncefield, M. and Hughes, J., At Home with the Technology: An Ethnographic Study of a Set-Top-Box Trial, ACM Transactions on Computer-Human Interaction 6, 3 (September 1999), 282-308.
11. Merialdo, B., Lee, K.T., Luparello, D., Roudaire, J., Automatic Construction of Personlaised TV News Programs, Proceedings of the 7th ACM international conference on Multimedia, ACM Press, New York, NY USA (1999) 323-331.
12. Eronen, L., Digital Television for All: User Preferences and Designer' Views on What Would Suit the User, ***SPRINGER LNCS proceedings of the*** 7th ERCIM Workshop "User Interfaces for All", Paris, France 2002.
13. Ibrahim, A., Johansson, P., Multimodal Dialogue Systems: A case study for interactive TV; ***SPRINGER LNCS proceedings of the***7th ERCIM Workshop "User Interfaces for All", Paris, France 2002.