Univ Access Inf Soc (2011) 10:151–163
DOI 10.1007/s10209-010-0195-7

LONG PAPER

Accessible interactive television using the MPEG-21 standard

Evangelos Vlachogiannis · Damianos Gavalas · George E. Tsekouras · Christos N. Anagnostopoulos

Published online: 6 July 2010

Abstract In this paper, the accessibility of interactive television (iTV) is discussed as a primary factor for its satisfactory adoption and commercial success. The work presented here is undertaken in the context of a research project that focuses on delivering iTV services to disabled children. This objective is accomplished through the utilization of the arising MPEG-21 standard. Based on that standard, iTV accessibility is investigated in terms of metadata and content adaptation. The novelty of the contribution lies on a systematic methodology that deals with a wide range of accessibility problems, as opposed to previous studies that focus mostly on users with only one specific disability.

Keywords Accessible interactive TV · MPEG-21 · Content adaptation approach · Metadata · Pervasive environments · Collaborative filtering

E. Vlachogiannis (✉)
Department of Product and Systems Design Engineering, University of the Aegean, Hermoupolis, Syros 84100, Greece
e-mail: evlach@aegean.gr
URL: http://www.syros.aegean.gr

D. Gavalas · G. E. Tsekouras · C. N. Anagnostopoulos
Department of Cultural Informatics, University of the Aegean, Mytilene, Lesvos Island 81100, Greece
e-mail: dgavalas@aegean.gr
URL: http://www.ct.aegean.gr/

G. E. Tsekouras
e-mail: gtsek@ct.aegean.gr
URL: http://www.ct.aegean.gr/

C. N. Anagnostopoulos
e-mail: canag@ct.aegean.gr
URL: http://www.ct.aegean.gr/

1 Introduction

"Winky Dink And You" is considered as "the first interactive TV show". It was a television children's show that aired from 1953 to 1957, and allowed interaction through the use of the "Official Winky Dink Kit". In this show, for the first time the role of television program consumers had been extended from passive viewing to active participation. This can be compared with the very recent move from the passive Web site consumption to the social—participative Web, known as Web 2.0. Wellens [30, p. 119] stated that "Interactive television represents means of linking individuals together by providing each with an electronically mediated representation of the other's voice and visual presence".

The RNIB Scientific Research Unit's Web site (Tiresias)[1] put the threshold between interactive TV and Enhanced TV as follows: "Enhanced TV is probably a better term to refer to one-way applications such as teletext, EPG access, etc., and it could be advantageous to restrict the term 'interactive TV' to two-way services reliant on some form of return path.". Whatever form iTV has taken (including WebTV, internet TV, Video on Demand, cable, satellite, digital terrestrial) during its long trial period, its adoption has been far away from the expectations [5]. According to Suzanne Stefanac of RespondTV, "the single greatest stumbling block iTV faces is the lack of a clear standard" [36]. Choi et al. [5, p. 163] developed a technology adoption model for iTV and discussed that "iTV may have different critical factors compared to conventional information systems because it is mainly used in home environment and it has never been used before".

Currently, iTV comes again to the front with more advanced technologies and a more mature audience. The iTV field has adopted techniques and technologies

[1] http://www.tiresias.org/research/guidelines/television/idtv.htm.

initially developed for the World Wide Web [8, 9]. This is more apparent in the case of IPTV, but generally applies to all kinds of iTV. Considering also that the number of TV sets is considerably larger to that of PCs worldwide [33], it becomes evident that the interaction requirements and specifically the need for accessibility are crucial. For instance, an iTV user now is in front of a large number of services (term used for TV channels) with amazing possibilities. A similar "explosion" occurred in the past in the World Wide Web and search engines; later on it was the portals (equipped with search engine facilities) and the adaptation mechanisms that made this huge amount of information manageable.

The MPEG-21 standard [11] recently released by ISO, aiming at defining an open framework for multimedia applications, seems to find a natural fit in the world of iTV [14].

This paper presents the work undertaken in the context a Greek national project aiming toward developing a MPEG-21-based framework for adapting iTV's content with respect to disabled children requirements.

The authors propose an approach for iTV accessibility focusing on the interaction of the stakeholders through adaptation. Contrary to the majority of the approaches found in literature, this approach investigates iTV accessibility in a wider manner, without focusing on a specific user group such as users with low vision.

The paper starts by presenting the related research and sets up its contribution roadmap, followed by a discussion of the identified requirements. Subsequently, a higher level approach is proposed with an accompanying architecture. Finally, the last section concludes the paper and draws directions for future work.

2 iTV accessibility research and contribution roadmap

Even since 1997, RNIB has provided recommendations for the accessibility of iTV [7]. Carmichael et al. [3] discovered similarities between the directions of iTV and that of the Web, and further noted that the gained experience from the later has to be transferred to the domain of iTV in order to avoid similar mistakes. Piccolo et al. [18] claim that the convergence between the two media (i.e., Web and iTV) is able to lead to the appropriation, with the necessary adjustments, of Web accessibility knowledge that has already been acquired, and of proposed recommendations for the design of accessible interfaces.

Toward approaching iTV accessibility, the following main elements can be identified (see Fig. 1):

- *Hunan actors*: The consumers (end users), the authors and the providers that either consumes/produces/provides the service.
- *User/Controller Interface (or Direct UI)*: The interface with which the user interacts directly.
- *Controller/Content Interface (or Indirect UI)*: The interface that often is being provided through the set-top box and displayed on the TV monitor; i.e., indirect user interaction.
- *Content:* The actual digital content (e.g., movie) with its accompanying metadata.

One way of reaching iTV for all is the satisfaction of the design requirements of the identified elements from an accessibility point of view.

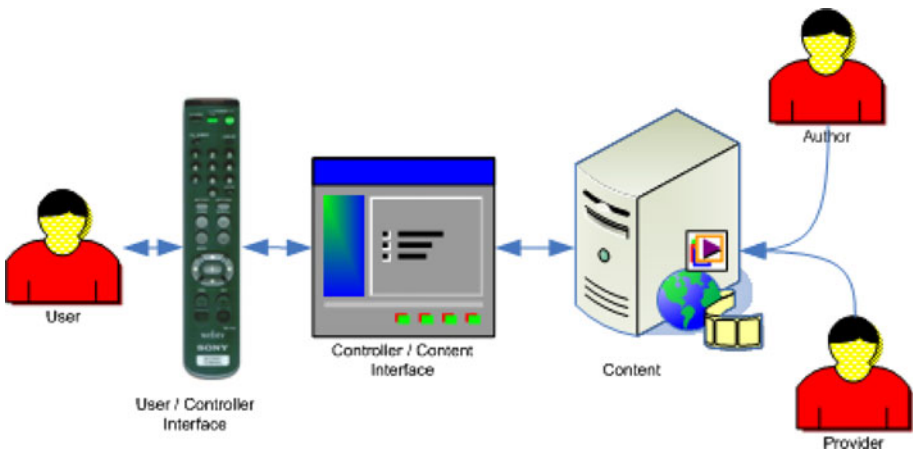
Such design requirements elements for accessible iTV are discussed in the subsequent subsections, outlining the contribution of this paper.

2.1 Human actors' accessibility role

The main stakeholders that act during the life cycle of an iTV broadcasting relating to accessibility are as follows:

- *the content providers*: The content/service providers should define an accessibility policy and also provide inspection procedures that would guaranty fidelity.

Fig. 1 iTV main elements



- *the authors*: the content authors are the ones that need to create content having in mind the aforementioned accessibility policy. Thus, through an appropriate authoring tool, they will be able to produce accessible content (i.e., multimodal metadata-enabled content).
- *the consumers (end users)*: In order to consume effectively an iTV program, the end users with some impairment need (1) to use *well* their assistive technologies and (2) to provide appropriate feedback to the system through the EPG interface (update profile, rate content, etc.)

2.2 Accessibility of the user/controller interface

Cesar et al. [4] distinguished two essential pillars in interactive digital television systems: user interaction and social communication. The first pillar, the most interesting from the perspective of accessibility, concerns the design and development of user interfaces, as the old-fashioned passive remote controls do not seem adequate and usable enough. Three related issues are distinguished in [4]: (1) extension of traditional remote controls, including voice and gestures; (2) augmentation of everyday objects, including natural ways to interact with media content and nonintrusive methods; and (3) repurposing of other devices, including handheld devices, as universal remote controls.

An alternative approach is through the use of abstract user interfaces. "People with different types of disabilities find it difficult or impossible to directly use electronic devices and services because the device's/service's user interface cannot accommodate the special needs of certain user groups (such as users with visual, hearing, or mobility impairments)" [34]. The former research suggests that users have to rely on service and device implementations that are specifically designed for them. In other words, a single individualized–universal user interface should be able to deal with as much interactive devices as possible. A standardization effort has been taking place in the context of the URC consortium in order to come up with a versatile user interface description for products, a "User Interface Socket" to which any URC can connect to discover, access and control the remote product [35]. The URC approach has been recently adopted by ISO (ISO/IEC 24752:2008) [12].

Abstract user interface seems even more challenging from an adaptation point of view. The user interface design process is very important, since it is the subsystem that directly interacts with the users and their contexts. For several years, Human–Computer Interaction (HCI) has been struggling to develop an abstract representation of user interface [21, 25]. This would offer the possibility of UI adaptation according to the system's environment (including user and context of use, e.g., [28].

Such an adaptation could be realized through user interface abstraction. Based on that, context sensitive pipelines would be introduced so that the interface can be adapted according to the system's environment [29].

2.3 Accessibility of the controller/content interface (EPG/IPG)

It turns out that the most significant accessibility difficulties concerning the iTV are related to the use of Electronic Programming Guide (EPG) by the users with visual, motor or cognitive disabilities [18]. Thus, an important step toward accessible iTV is providing a well-designed EPG. The EPG is a vital component of interactive television, allowing viewers to navigate through available programs and services. This is often a complex interface influenced by the design of WIMP (Windows, Icons, Menus, Pointers) on-screen application [2] where users explore a huge number of programs and services. Such interfaces are very different from the traditional analog TV menus, which need to handle no more that 5–10 passive channels. The Vista project [2] aimed at developing a virtual assistant, embodying a speech-based interface, between digital television viewers and the content and functions of the EPG. In order to address the needs of preschool children, Joly et al. [13] developed a set of specific requirements, which were based on a range of existing guidelines on interactive television applications, personalized recommendation systems and interaction design for children, in the context of theories of child development.

Rice and Alm [19] attempted to support older people who have difficulties in using current interface models for Digital TV. Their research indicated that "navigational techniques that mimic aspects of real-world artefacts in a manner that individual's can quickly relate to present possible new directions in DTV design. However, the success of such systems depends on research strategies that take the impact of both an appropriate input control and on-screen interaction into account" [19, p. 134].

2.4 Accessibility of content

In the related literature, there have been several attempts to incorporate accessibility issues into the MPEG-21. The majority of them are focused on visual disabilities (e.g., [20, 22, 24, 31]). Rice [20] presented the difficulties that visually disabled users face while consuming iTV services. This work gave emphasis into parameters like screen size, font size and color, icons' identification and screen layout. The derived conclusion was that the best facing approach

of the problem situation is personalization, due to diverging requirements. Choi et al. [5] mentioned the fact that the TV compared to the PC is a home appliance, and therefore is not personal but shared, which directly implies that the opinions of family members are very influential. Thang et al. [24] proposed a systematic contrast-enhancement method to improve content visibility for low-vision users, through MPEG-21 content adaptation. Yang et al. [31] proposed a technique for the accessibility of iTV for people with visual impairments, and especially color blindness. This technique involves both the incorporation of descriptive metadata in MPEG-21 and the design of an adaptive system. Berglund and Johansson [1] studied the benefits of the usage of speech—dialog in the domain of iTV, and derived several design considerations. Carmichael et al. [1, 3] concluded that the accessibility characteristics that have not yet been given the necessary attention in the context of iTV are subtitles, captions and audio description [1, 3], whereas they have been addressed in the context of the Web (WCAG2.0, SMIL,[2] SVG[3]).

3 Specific issues of the application domain

Having set a general approach for the accessibility of interactive television, this section aims at raising some more specific issues on current system design in relation to specific users target group. As already mentioned in the introduction, this work focuses on the delivery of interactive television content to disabled children. Disabled children requirements are considered as the set of requirements that comes out from the blending of disabled people requirements and children requirements relating to interactive television.

Clarkson et al. [6] identify four basic types of disabilities and related issues (in parenthesis):

- Visual Impairment (Recognizing and locating buttons on the remote control; Reading the on-screen display)
- Hearing Impairment (Subtitles, Volume, Literacy)
- Dexterity Impairment (Button sensitivity; Compact layout, Remote Control Complexity)
- Cognitive Impairment (Time delays between cause and effect; understanding the way in which elements of the on-screen display are intended to correspond to the buttons on the remote control, literacy).

From the side of children requirements, Hynd [10], while studying the responses of young children to interactive television programs, has focused on television's immediate effects on attention, comprehension, engagement and enjoyment. Hynd examined the characteristics of television that have been found to influence these outcomes for young children, as well as individual factors like gender and age.

Combining the two aforementioned research results, one can come up with some questions that can lead to (1) the potential parameters that a disabled child's profile should incorporate and (2) the technical characteristics that interactive television broadcasting should provide:

- What program the child wants to consume?
- What interaction capabilities/possibilities are being provided through a specific program so that the last one would be able to gain the attention of the child?
- How such programs need to be communicated to the child—using which modalities?
- How simple dialogs and texts should be and how much time is optimum to persist (the dialogs and texts) in order to be comprehensive?
- How simple and attractive both the remote controller and the EPG should be?
- How the context could help or distract the child?

This list is by no means exhaustive, and of course not all of these questions relate to proposed software architecture. Nevertheless, they illustrate the approach developed in order to extract the IN PARAMS and OUT PARAMS discussed in next section (see Fig. 2). It should be noticed that several technical requirements coming from the "disabled people" perspective often intersect with some coming from the "children" perspective. For instance, both an individual with a cognitive impairment and a child with literacy limitations require simple text and dialogs.

4 An approach toward accessible interactive television focusing on the content

This section aims at presenting the proposed approach for enhancing the accessibility of interactive television. This approach is focused on the following factors:

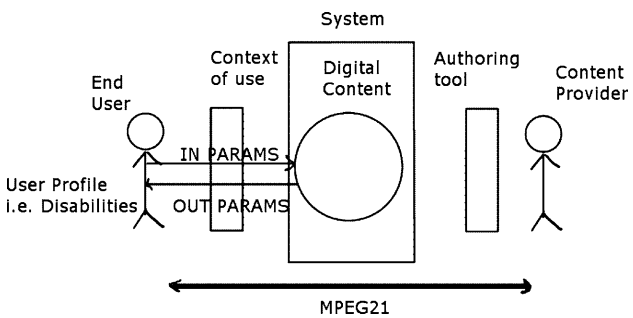


Fig. 2 MPEG-21 involvement in iTV: a possible scenario

[2] Accessibility Features of SMIL: http://www.w3.org/TR/SMIL-access/.

[3] Accessibility Features of SVG: http://www.w3.org/TR/SVG-access.

- the requirements of the content for allowing accessibility,
- the appropriate communication (e.g., subtitles, audio description, sign language, etc.) to the end user through adaptation mechanisms and
- the delivery to the end user of the appropriate programs (program recommendations), depending on user's characteristics (e.g., impairment and age) and program preferences.

MPEG-21 is able to provide to the iTV designer a framework that can offer a big, integrative picture of an iTV system. Based on that, an indicative scenario has been devised, including production, delivery and consumption of the digital content, aiming at identifying the primary entities and the way these are involved in the overall design outcome (see Fig. 2). According to that:

- The content designer (CD) identifies the target groups.
- The CD, supported by MPEG-21 metadata, describes the target groups using their characteristics (e.g., blindness) and associates interaction modes (e.g., auditory description) using an appropriate authoring tool.
- The CD develops the required content components (digital items) based on the above-decided interaction modes. These are integrated into the metadata by using the authoring tool.
- End user A, say blind, wants to consume developed content. She/he has already stored her/his profile. The context of use is described using attributes such as access device capabilities, audio configuration, time and location of the end user.
- The context of use is delivered to the serving system accompanied by the user request.
- The system inferences and maps the user's context of use with an appropriate composition of the components of the content. If, while consuming, the context of use is being modified, the system needs to be aware so that it can adapt to new requirements.

Even though MPEG-21 addresses adaptation and specifically accessibility by including several relating XML elements into its schema, it seems that it cannot ensure the accessibility of delivered content. Instead, this is a fundamental condition for providing accessibility output of the systems involved. In other words, it should be able to provide the required infrastructure so that digital content would be able to obtain the requisite variety for both the content designer (e.g., content versions), to be able to design accessible content, and the involved systems (e.g., adaptation mechanism), to have the required information to deliver an accessible result. Figure 3 presents the stakeholders related to accessibility. From such a point of view, the content provider, the author (also referred to as content designer), the authoring tools, the systems of the content provider and of course the consumer with her/his interaction profile [29] (preferences, device capabilities etc.) are identified and all play a major and cascading role in iTV accessibility.

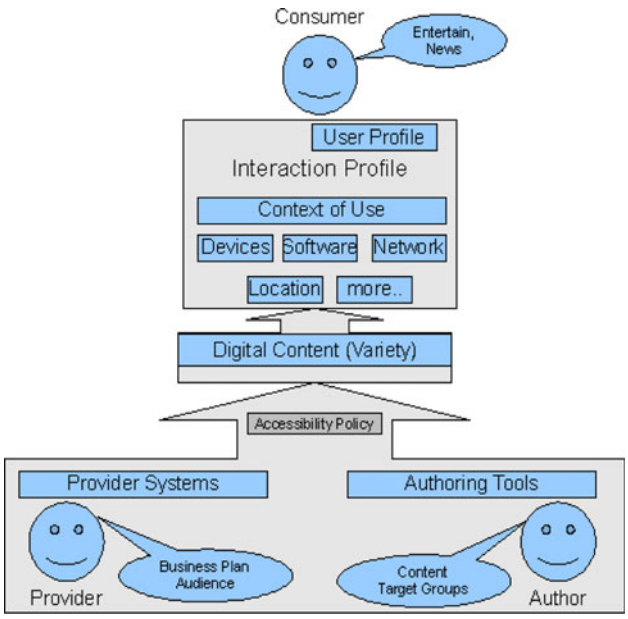


Fig. 3 Multimedia delivery stakeholders related to accessibility

Briefly, the role of the MPEG-21 toward the accessibility of iTV is revealed through the following dimensions:

Alternative content: MPEG-21 offers metadata that allows content providers to provide the content in one or more alternative ways. These often refer to different modalities and can include captions, audio descriptions, etc.

Digital Content Navigation: In iTV environments, navigation facilities within available content are provided through an Electronic Program Guide (EPG). This is actually the interactive portion of the system that offers the required functionality to the user, including service (channel) selection/retrieval, programs information and scheduling, profiling/personalizing, rating and/or even acting upon the content.

Description of context of use (IN PARAMS): The usage context actually refers to all the information that needs to be taken into account to adapt digital content according to the user's requirements.

Description of presentation parameters of digital content (OUT PARAMS): This determines what technical characteristics need to be adapted. An important implementation consideration is the transformation of MPEG-21 to SMIL as an intermediate solution to ensure media players' compatibility. This involves the mapping between those two infrastructures realized using XSLT.

Device accessibility: This refers to the accessibility of the involved hardware, including remote controls and set-top boxes.[4]

Content provider accessibility policy: Probably, an important contribution to the field of accessibility of MPEG-21 is the capability of applying and claiming for an accessibility policy. In other words, content providers need to be capable of applying a kind of accessibility policy based on the target consumer group and the former's requirements for quality assurance. For instance, such a policy could provide for digital content to be accompanied by subtitles of two languages (e.g., English, Greek) and every image with an alternative text between two and ten words. Applying such policies requires a mechanism for validating digital content to a policy description and could be for instance implemented based on Schematron[5] (an XML structure validation language for making assertions about the presence or absence of patterns in trees).

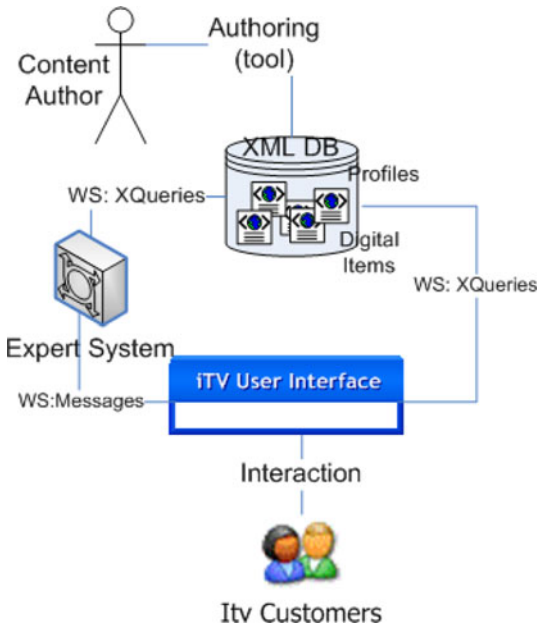


Fig. 4 iTV adaptation architecture

5 The system architecture

Figure 4 illustrates the system's overall architecture that came out following the aforementioned approach to accessibility. The overall system consists of the following:

- the accessibility-enabled authoring tool (developer21), which allows content providers to easily author a diversity of multimedia resources supporting a MPEG-21 compliant metadata model;
- the user interface (iTVSimu), which is the component through which the end user will experience the services;
- the expert (content recommendation) system, which uses an algorithm originally devised for clustering Web documents [26] in order to classify digital items and user profiles based on their attributes, and enable intelligent TV program recommendations;
- the backend infrastructure, which consists of (1) a persistence subsystem based on a native XML database where the digital content descriptor is located and (2) Web services infrastructure for the communication between distributed subsystems.

5.1 Overview of the authoring tool (developer21)

Developer21 for MPEG-21 serves as a multimedia authoring tool adding or extracting MPEG-21 descriptors and metadata in various multimedia assets, as shown in Fig. 4. Once created, these descriptors (in XML schema files) are locally stored in an XML metadata database. Users have the possibility to create a new MPEG-21 Digital Item, and to edit, delete, convert or send this metadata document to the database.

The tool is designed to support 6 different XML schemas, each one dedicated to the respective MPEG-21 part. The MPEG-21 descriptors that are provided by Developer21 are:

- Digital Item Declaration (DID)
- Digital Item Identification (DII)
- Intellectual Property Management and Protection (IPMP)
- Rights Expression Language (REL)
- Rights Data Dictionary (RDD)
- Digital Item Adaptation (DIA)

Basic information about descriptors is provided through a graphical representation: the type of descriptor (DID, DII, IPMP, RDD, REL, DIA), the type of program information and general information or only audio and video attributes. In parallel with the editing and browsing capabilities of the tool, metadata management is also supported. Binding of metadata and XML descriptors with the actual multimedia content is performed in order to create the integrated Digital Item that contains the actual content and the descriptive information. A Digital Item processed with Developer21 (see Fig. 4) is represented in a form suitable to interact with an Expert System that is used for increasing interactivity in IPTV or iTV. The expert system assigns a TV viewer to a specific category based on social parameters and then matches the appropriate audiovisual content

[4] http://www.tiresias.org/equipment/settop_boxes.htm.

[5] http://xml.ascc.net/schematron/schematron1-5.sch.

according to its respective MPEG-21 descriptors. Hence, Developer21 supports content personalization by MPEG-21 metadata, profiling the target groups with their characteristics (e.g., blindness) and associating the appropriate interaction modes (e.g., auditory description) in the multimedia content.

In general, personalization allows users to browse programs much more efficiently according to their preferences. Specifically, using Digital Item Adaptation (DIA) the authoring tool captures user characteristics such as:

- Usage Preferences and history
- Content characteristics preferences such as Audio, Display Color and Graphics presentation, Presentation Priority and Stereoscopic Video Conversion.
- Accessibility issues like focus of attention, auditory impairment, visual impairment, color vision deficiency
- Terminal technology such as codec capabilities, display capabilities, audio output capabilities, user interaction inputs and device class
- Network capabilities and condition
- Location, time and environment.

5.2 Overview of the user interface prototype (iTVSimu)

In the context of the conducted work, the need for designing and developing a simulation platform, acting as an interaction interface between our iTV architecture and the prospective viewer, was evident. Therefore, a user interface prototype has been implemented to enable users to effectively browse, search, download and consume the provided audiovisual content. In the case of disabled people, 'effectively' means that both the content and the added-value services need to be accessible to the user. However, the accessibility of the EPG was not addressed, as this was out of the scope of the project. Nevertheless, the GUI has been developed using Java Accessibility API/Java Access Bridge.[6] As a consequence, the prototype is accessible at a satisfactory level.

The developed User Interface comprises an EPG simulator. It should be noted that the choice of the implementation technologies has not been straightforward, considering the plethora of available standards and technologies like MHP,[7] GEM-IPTV, TV-Anytime, DVB-IP, Java TV and more. Given the requirement for incorporating networking functionality into the EPG subsystem, a Web-based approach instead of a standalone application has been adopted. This approach ensures the execution of the EPG through a standard browser interface.

The design approach was conducted as follows. During the early phases of the design of the prototype system, an identification of the stakeholders took place:

- *The end user*: he/she interacts with the ITV interface browsing and consuming digital content. The end user is associated with an XML-based user profile which includes personal data, preferences upon the audiovisual content (e.g., sports, news, movies) and potential disabilities (hearing problems, visual impairments, etc.)
- *The Service Provider*: The analogous of traditional TV channels.
- *The TV Guide Provider*: A service that informs end users about the offered services and their availability time schedule.

Occasionally, the Service Provider and the TV Guide Provider coincide. For simplicity reasons, this assumption was made while designing the prototype. The focus has been on the interaction of the end user with the iTV interface, since that will affect the overall functionality of a personalized system, with particular emphasis on disabled users.

Figure 5 illustrates the three elementary subsystems of the iTV user interface: the player (left panel), the EPG (right panel) and the logger (bottom panel). The EPG panel consists of three panels: (1) the "My ITV" panel, where the content recommendations appear and the user can also trigger a reminder, (2) the "Program" panel, where the user can select between services and view the program of the selected service, and (3) the "My Profile", panel where the user can modify her profile.

The three elementary subsystems presented above are supported by auxiliary services for enhancing the functionality of the iTV simulator. The functional and interactivity requirements of the above-mentioned subsystems are analyzed in the following subsections, discussing the solutions adopted in the developed prototype.

5.2.1 *iTVSimu subsystems*

5.2.1.1 Logger subsystem This is the simplest, although crucial, software module, as it provides feedback to the user for the "hidden" operations. It records and displays all (implicit or explicit) user actions (e.g., profile modification, starting/pausing/resuming a TV program, etc.). It has been implemented through the Java Observer pattern whose actions activate the logger.

5.2.1.2 Player subsystem This subsystem reproduces iTV programs (digital items) and records the user's interaction history. Its elementary module is the digital content player. Such player should support, in addition to basic

[6] http://java.sun.com/javase/technologies/accessibility/accessbridge/index.jsp.

[7] http://www.mhp.org/.

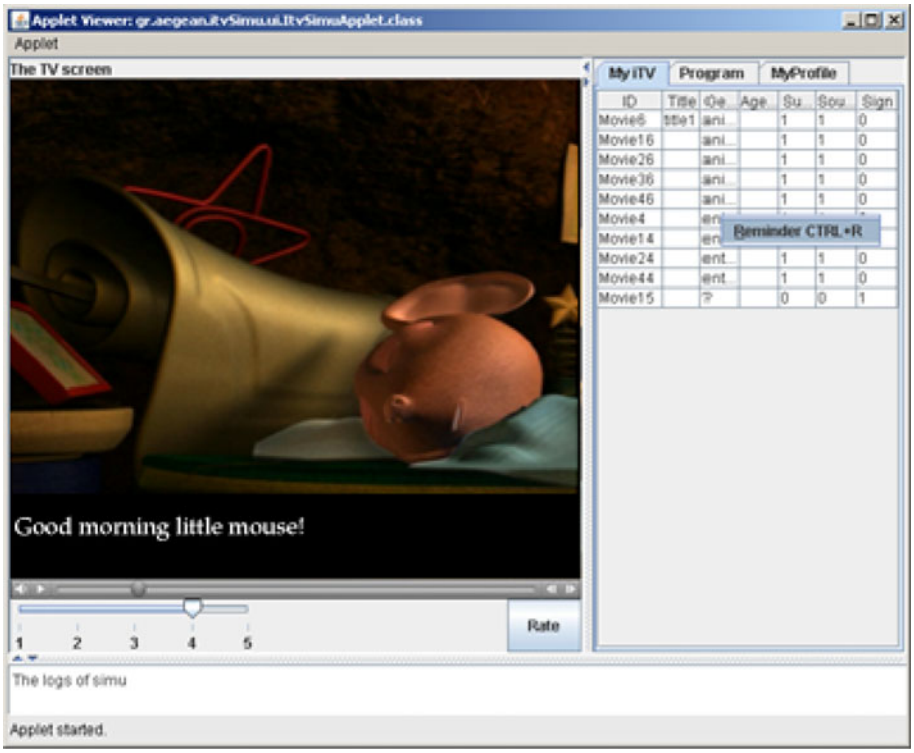


Fig. 5 A screenshot of the iTV user interface: recommendations panel

functionality (play, pause, rewind, etc.), accessibility-related functionality such as subtitles, audio descriptions, etc. Given that no MPEG-21 player is currently available, SMIL was used as an intermediate technology, mainly due to the numerous available SMIL players (e.g., X-Smiles,[8] QuickTime player). In particular, the MPEG-21 digital item declarations are transformed into SMIL format through an appropriate XSLT transformation, and subsequently the SMIL markup code is parsed by the SMIL player. This approach ensures the iTV interface's interoperability, since SMIL is now considered a mature Web technology. In the developed prototype, the SMIL player has been implemented using the QuickTime for Java API.[9] As illustrated in Fig. 6, the XSLT transformation of MPEG-21 digital items into SMIL documents depends on the user profile, taking into account potential user disabilities. An example of such digital item declaration and its SMIL representation is illustrated in Fig. 7.

The second function of the Player subsystem is the provision of user interaction information to the expert (recommendation) system. An XML-based description of the user interaction is first stored into an XML native database located on the iTV's server, and subsequently

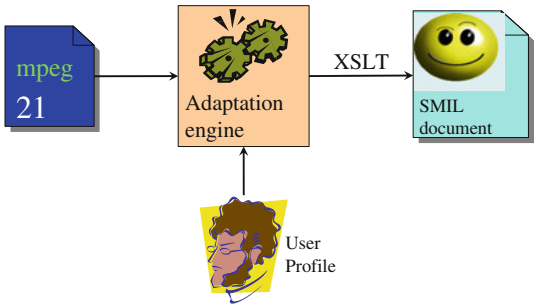


Fig. 6 XSLT transformation of MPEG-21 digital items into SMIL documents

retrieved by the expert system to enable more effective and reliable reasoning. The user interaction history comprises a function $f(x, y, \ldots, z)$, wherein $x, y, \ldots, z$ are the values of interaction parameters. Such parameters are either explicitly provided by the user or implicitly inferred by the player. Examples of implicit parameters are the playing time of a video over the video duration ratio, while the rating of a TV program (in a 0–10 scale) could be explicitly provided by the viewer. The interaction history function could be expressed as $f(x) = aX + bY$, where a, b represent weights based on the designer's priorities, which could either be static or dynamically specified (through training). As depicted in Fig. 8, the user's interaction history and the TV programs ratings posted by users that belong to the same users' cluster (the concept of user cluster will be discussed later on in this paper) comprise the input of the

[8] X-Smiles SMIL player, http://www.xsmiles.org/xsmiles_smil.html.

[9] QuickTime for Java (QTJ) is a software library that allows software written in Java to provide multimedia functionality, by making calls into the native QuickTime library. QTJ offers SMIL support and also can handle a larger variety of multimedia formats than the 'traditional' Java Media Framework (JMF) API.

```
<?xml version="1.0" encoding="UTF-8"?>
<DIDL xmlns:xsi=http://www.w3.org/…="....\MPEG-21-DI\DIDL.xsd">
<ITEM>
    <DESCRIPTOR>
        <STATEMENT TYPE="text/plain">
            Movie for normal, blind or deaf individuals
        </STATEMENT>
    </DESCRIPTOR>
    <ITEM>
        <DESCRIPTOR>
            <STATEMENT TYPE="text/plain">
                Movie for normal individuals
            </STATEMENT>
        </DESCRIPTOR>
        <COMPONENT>
            <RESOURCE REF="video.mov" TYPE="video/mov"/>
        </COMPONENT>
    </ITEM>
    <ITEM>
        <DESCRIPTOR>
            <STATEMENT TYPE="text/plain">
                Movie for blind individuals
            </STATEMENT>
        </DESCRIPTOR>
        <COMPONENT>
            <RESOURCE REF="video.mov" TYPE="video/mov"/>
            <RESOURCE REF="audiodescription.mov" TYPE="audio/mp3"/>
        </COMPONENT>
    </ITEM>
    <ITEM>
        <DESCRIPTOR>
            <STATEMENT TYPE="text/plain">
                Movie for deaf individuals
            </STATEMENT>
        </DESCRIPTOR>
        <COMPONENT>
            <RESOURCE REF="video.mov" TYPE="video/mov"/>
            <RESOURCE REF="captions.txt" TYPE="text/plain"/>
        </COMPONENT>
    </ITEM>
</ITEM>
```

```
<?xml version="1.0" encoding="UTF-8" ?>
<smil xmlns:qt="http://www.apple.com/...." time-slider="true">
<head>
    <layout>
        <root-layout width="320" height="350" background-color="black" />
        <region id="captions" backgroundColor="yellow"
            top="250" height="100" left="1" width="310" />
        <region id="movie" left="0" top="0" width="620" height="740" />
    </layout>
</head>
<body>
    <par>
        <textstream src="captions.txt" region="captions"
            systemCaptions="on" />
        <video src="video.mov" alt="Movie title" region="movie"
            begin="00:00.0" dur="00:14:02.000" />
    </par>
</body>
</smil>
```

Fig. 7 A digital item declaration document (*left*) transformed to SMIL format (*right*) which synchronizes a video with captions (appropriate for hearing impaired individuals)

expert system. The latter recommends—among the available digital content—those programs that suit the user's profile and the user cluster XML descriptions.

5.2.1.3 EPG subsystem This is the most 'interactive' subsystem, since it is used by the user to browse, navigate and download audiovisual content. In the context of the conducted research, several use cases were identified, according to which the iTV end user may use EPG in order to:

- navigate within iTV available services (zapping);
- personalize the audiovisual content based on potential individual disabilities and content preferences;
- schedule a reminder for a TV program.

An important consideration during the EPG's development has been the representation and retrieval of the TV schedule. To address this issue, the TV-Anytime Program metadata [27] has been used, along with TV-Anytime Java API developed by BBC.[10] The overall functionality of the EPG has been based upon the specifications of the JAVA TV API (JSR-000927) in a nonstrict manner. The result of the BBC TV schedule retrieval on the iTV interface is shown in Fig. 5.

[10] http://www.bbc.co.uk/opensource/projects/tv_anytime_api/.

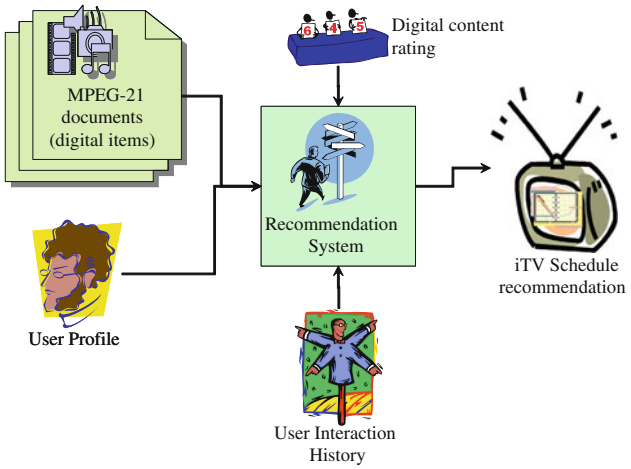


Fig. 8 TV schedule recommendation

The most important part of content personalization has been the modeling of user characteristics (e.g., disabilities) and preferences. To address this issue, the Interaction Profile of the DAWIS framework for the design of adaptive Web information systems [29] has been adopted. The most abstract layer of the DAWIS Interaction Profile consists of the Service Interaction Profile, the Delivery Context Interaction Profile, the User Interaction Profile and the Platform Interaction Profile. Based on that, an iTVProfile schema has

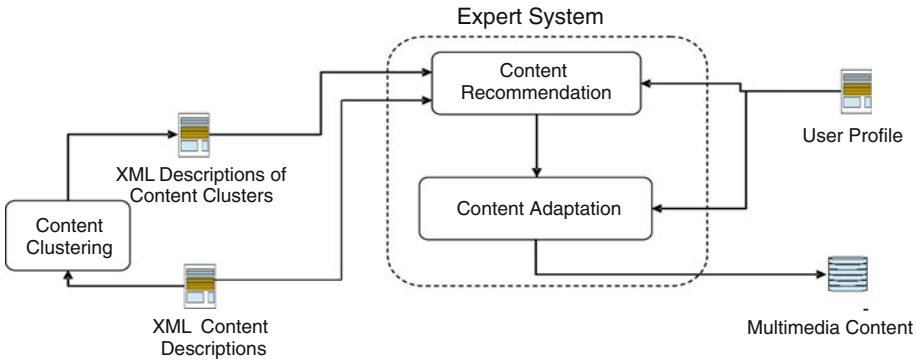


Fig. 9 Basic structure of the expert system

been developed and serialized in XML syntax including elements like LanguageNative, Languages, ContentPreferences, Disabilities, Subtitles, Captions, AudioDescription and SignLanguage. The iTVProfile instances are stored in a separate collection in the XML database storage through XQuery.[11]

5.3 Overview of the expert system

The system aims to increase the accessibility of the iTV platform through content recommendation and content adaptation, both based on user profile and content metadata (with emphasis on age and accessibility).

Thus, the Expert system consists of two subsystems, namely the content recommendation unit and the content adaptation unit. Figure 9 presents the architecture of the system.

5.3.1 The content recommendation unit

The basic functionality of the content recommendation unit is to provide program recommendations to the users based on their user profile and the multimedia content metadata, following a collaborative filtering process [15–17, 32]. This unit operates in a hierarchical structured algorithm, which is described below.

5.3.1.1 Collaborative filtering algorithm

Step 1. Cluster analysis of the content XML documents is performed to create document clusters, called content clusters. Specifically, based on the XML documents associated with the content Digital Items, the significant attributes are selected, each of which is assigned to a specific document. Since the preselected attributes are categorical in nature (like subtitle, etc.), the set of XML documents define a categorical data set. Then, an algorithm is applied to partition this data set into a number of clusters, where documents that belong to the same cluster are as similar as possible, while documents belonging to different clusters are as dissimilar as possible.

Step 2. Similarly, a number of significant attributes, including the usage history, are selected from the user profile XML documents. These attributes define a feature space where each user is represented by one point. Thus, a number of categorical data are generated, each of which corresponds to a specific user. Then, the clustering algorithm is applied to partition the set of users into a number of clusters, called user clusters.

Step 3. The system prompts the users to rate the programs they have consumed. Each user cluster is assigned to a specific content cluster. This assignment is carried out by taking into account the sum of the ratings of the users that belong to same user cluster. Then, the content cluster that corresponds to the higher rating is assigned to that user cluster. It should be emphasized that each user cluster may be mapped to multiple content clusters. Herein, one-to-one mapping is performed in order to reduce computational cost.

Step 4. The system classifies the current user to a user cluster and recommends the programs that belong to the content cluster assigned to that user cluster (see previous step).

Step 5. If needed, the recommendation list may be shortened by including the most interesting programs. This may be accomplished by applying a threshold upon the ratings of the individual programs that belong to the recommended content cluster.

5.3.2 The content adaptation unit

The content adaptation is of major importance for the efficient presentation of multimedia content [23]. Content

[11] XQuery 1.0: An XML Query Language: http://www.w3.org/TR/xquery/.

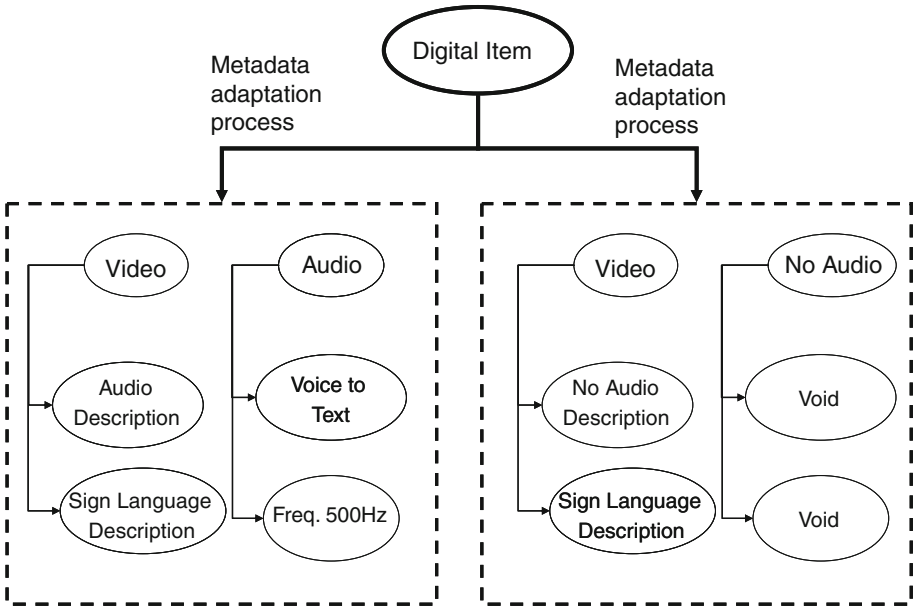


Fig. 10 Example of two different versions of a DI as a result of the metadata adaptation process

adaptation is performed by using a number of inference rules. To design the set of rules, the Digital Items are stored according to three abstraction levels. The first level stores the original multimedia object and its respective locators, which include all information required to download the object (e.g., the path, etc.). The second level includes the descriptions of the object, which mainly concern the type of the content (e.g., audio file, video file, image file, etc.). Finally, the third level includes all the subelements of the digital item. An example of the three abstraction levels is depicted in Fig. 10.

Based on the above figure, the inference rules used to adapt the content are derived as follows.

First, the symbols O1, O2 and O3 describe the objects at the first, second and third abstraction level, respectively. Thus, the object at the first level can be described in terms of second level objects:

$$O1 = \{O2(\text{Video})/O2(\text{Audio})\}$$

Likewise, the second level objects are described in terms of third level objects:

$$O2(\text{Video}) = \{O3(\text{Video,Audio Description})/ O3(\text{Video, SLD})\}$$

$$O2(\text{Audio}) = \{O3(\text{Audio, VtT})/O3(\text{Audio, Freq500 Hz})\}$$

where SLD and VtT stand for Sign Language Description and Voice-to-Text, respectively.

In the next step, the "Accessibility" attribute of the user profile is considered. An example of the domain of values for this attribute's values is:

$$\begin{aligned} PR(\text{Access}) = \{&PR(\text{Access, Total Blindness})/ \\ &PR(\text{Access, Partial Blindness})/ \\ &PR(\text{Access, Total Deafness})/ \\ &PR(\text{Access, Partial Deafness})\} \end{aligned}$$

where PR and Access stand for the Profile and the Accessibility Attribute, respectively. Based on the above analysis, the adaptation inference rules for the above example are as follows:

$$O2(\text{Video}) \cap PR(\text{Access, Total Blindness}) \rightarrow O3(\text{Video, Audio Description})$$

$$O2(\text{Video}) \cap PR(\text{Access, Partial Blindeness}) \rightarrow O3(\text{Video, Audio Description})$$

$$O2(\text{Video}) \cap PR(\text{Access, Total Deafness}) \rightarrow O3(\text{Video, SLD})$$

$$O2(\text{Video}) \cap PR(\text{Access, Partial Deafness}) \rightarrow O3(\text{Video, SLD})$$

$$O2(\text{Audio}) \cap PR(\text{Access, Total Deafness}) \rightarrow O3(\text{Audio, VtT})$$

$$O2(\text{Audio}) \cap PR(\text{Access, Partial Deafness}) \rightarrow O3(\text{Audio, Freq500Hz})$$

where $\cap$ is the conjunction operator. To this end, it should be noticed that the adaptation inference rules apply either

to the recommended program or the program that the user takes the initiative to view.

6 Conclusions and future work

So far, the developed system is at a prototype stage and its components (i.e., expert system, authoring tool, iTV simulator) have not been evaluated as a whole by end users, due to project's time limitation. Nevertheless, iTVSimu seems to offer an interesting and simplified architecture that can realize a primitive IPTV platform and serve as benchmarking software for further research in the field of content adaptation and accessibility. Currently, the prototype has implemented only a portion of the foreseen user groups. The reason is that the evaluation of the adaptation behavior requires a considerable number of users with diverse profiles, and an analogous number of digital items. Such an evaluation is planned as future work. In addition, a future work could consider more runtime parameters (implicit profile) and more effective models for multiplexing them, for example through AI techniques and simulation. Finally, a separate version of iTVSimu interface, optimized for users with hearing difficulties (e.g., incorporating auditory menus functionality), will be implemented.

From the point of view of standardization efforts, it turns out that the selection of standards was a difficult task as there are many of them, often overlapping and/or contradicting each other. Consequently, even if some designer uses open standards, the final overall design results in a proprietary solution composed of several open standards.

Finally, it should be mentioned that the proposed approach and architecture contribute to the compensation of digital divide offering accessible services to different groups of people. At the same time, having in mind that the number of disabled and elderly people is increasing and alternative access devices (e.g., mobile phones) are proliferated in everyday life, the benefits of incorporating accessibility in iTV is an opportunity for businesses to grow their market share.

Acknowledgments This work is supported by the General Secretariat of Research and Technology (Project "Software Applications for Interactive Kids TV-MPEG-21", project framework "Image, Sound, and Language Processing", project number: ΕΗΓ-16). The participants are the University of the Aegean, the Hellenic Public Radio and Television (ERT) and the Time Lapse Picture Hellas.

References

1. Berglund, A., Johansson, P.: Using speech and dialogue for interactive TV navigation. Univers Access Inf Soc **3**(3–4), 224–238 (2004)
2. Carmichael, A., Petrie, H., Hamilton, F., Freeman, J.: The vista project: broadening access to digital TV electronic programme guides. Psychnol J **1**(3), 229–241 (2003)
3. Carmichael, A., Rice, M., Sloan, D.: Inclusive design and interactive digital television: has an opportunity been missed? In 3rd Cambridge Workshop on Universal Access and Assistive Technology. Fitzwilliam College, Cambridge, 10–12 April (2006)
4. Cesar, P., Chorianopoulos, K., Jensen, J.F.: Social television and user interaction. Comput. Entertain. 6(1), 1–10. doi:10.1145/1350843.1350847 (2008)
5. Choi, H., Choi, M., Kim, J., Yu, H.: An empirical study on the adoption of information appliances with a focus on interactive TV. Telemat. Inf. 20(2),161–183. doi:10.1016/S0736-5853(02)00024-2 (2003)
6. Clarkson, J., Karger, S., Sinclair, K.: Digital television for all: report on usability and accessible design. DTI, London (2003)
7. Darby, S.: Introduction to enhancing the accessibility of digital television. In RNIB (1997)
8. Ferguson, D.A., Perse, E.M.: The world wide web as a functional alternative to television. J. Broadcast. Electron. Media **44**, 155–174 (2000)
9. Gil, A., Pazos, J., Lopez, C., Lopez, J., Rubio, R., Ramos, M., Diaz, R.: Surfing the web on TV: the MHP approach, Multimedia and Expo, 2002. ICME '02. Proceedings. 2002 IEEE International Conference, vol. 2, pp. 285–288 (2002)
10. Hynd, A.: Evaluating four and five-year old children's responses to interactive television programs. PhD thesis, Murdoch University, 2006. https://wwwlib.murdoch.edu.au/adt/pubfiles/adt-MU20070824.134316/02Whole.pdf (2006)
11. ISO: MPEG-21, Part I: Information technology—multimedia framework (MPEG-21)—vision, technologies and strategy, ISO/IEC TR 21000-1 (2004)
12. ISO/IEC: 21000-7, Information technology—multimedia framework (MPEG-21)—Part 7: Digital item adaptation, 1st edn (2004)
13. Joly, A.V., Pemberton, L., Griffiths, R.: Electronic programme guide design for preschool children. In Changing Television Environments, pp. 263–267. doi:10.1007/978-3-540-69478-6_35 (2008)
14. Lugmayr, A., Samuli, N., Seppo, K.: Digital Interactive TV and Metadata: Future Broadcast Multimedia. Springer, Berlin (2004)
15. Melville, P., Mooney, R.J., Nagarajan, R.: Content-boosted collaborative filtering for improved recommendations. In Proceedings of the 18th National Conference on Artificial Intelligence, pp. 187–192 (2002)
16. Niiranen, S.: Broadcast multimedia personalization. Licentiate thesis, Tampere University of Technology (2003)
17. Petrie, H., Weber, G.: Personalization of interactive systems. In: Proceedings of the 9th International Conference on Computers Helping People with Special Needs (ICCHP'2004), pp. 117–120 (2004)
18. Piccolo, L., Melo, A., Baranauskas, M.: Accessibility and interactive TV: design recommendations for the brazilian scenario. In: Baranauskas, M.C.C. et al. (eds.) INTERACT 2007: 11th IFIP TC 13 International Conference—Proceedings, Part I. Rio de Janeiro, Brazil, pp. 361–374 (2007)
19. Rice, M., Alm, N.: Designing new interfaces for digital interactive television usable by older adults. Comput. Entertain. 6(1):1–20. doi:10.1145/1350843.1350849 (2008)
20. Rice, M.: Personalisation of interactive television for visually impaired viewers. In 2nd Cambridge Workshop on Universal Access and Assistive Technology, Fitzwilliam College, Cambridge, 22–24 March (2004)
21. Souchon, N., Vanderdonckt, J.: A review of xml compliant user interface description languages. In Conference on Design, Specification, and Verification of Interactive Systems, vol. 2844, pp. 377–391. Lecture Notes in Computer Science, Springer, Berlin (2003)
22. Springett, M., Griffiths, R.: Accessibility of interactive television for users with low vision: learning from the web. In Proceedings

of the 5th European Conference on Interactive TV (EuroITV'2007), LNCS 4471, pp. 76–85, May (2007)
23. Sun, H., Vetro, A., Asai, K.: Resource adaptation based on MPEG-21 usage environment descriptions. In Proceedings of the IEEE International Conference on Circuits and Systems, pp. 536–539 (2003)
24. Thang, T.C., Yang, S., Ro, Y.M., Wong, E.K.: Media accessibility for low-vision users in the MPEG-21 multimedia framework. IEICE Trans. Inf. Syst. **E90-D**(8), 1271–1278 (2007)
25. Trewin, S., Zimmermann, G., Vanderheiden, G.: Abstract representations as a basis for usable user interfaces. Interact. Comput. **16**, 477–506 (2004)
26. Tsekouras, G., Anagnostopoulos, C., Gavalas, D., Economou, D.: Classification of web documents using fuzzy logic categorical data clustering. Proceedings of the 4th IFIP Conference on Artificial Intelligence Applications & Innovations (AIAI'2007), pp. 93–100 (2007)
27. TV-Anytime, ETSI TS 102 822: Broadcast and On-line Services: Search, Select and Rightful Use of Content on Personal Storage Systems
28. Velasco, A., Mohamad, Y., Gilman, S., Viorres, N., Vlachogiannis, E., Arnellos, A., Darzentas, S.: Universal access to information services—the need for user information and its relationship to device profiles. Univers. Access Inf. Soc. 3(1):88–95. doi: 10.1007/s10209-003-0075-5 (2004)
29. Vlachogiannis, E.: Methodological framework, analysis, development and creation of design support environments of adaptive systems. PhD Thesis, University of the Aegean, Department of product and Systems design Engineering (2008)
30. Wellens, A.R.: An interactive television laboratory for the study of social interaction. J. Nonverbal Behav. **4**(2), 119–122 (1979)
31. Yang, A., Ro, Y., Nam, J., Hong, J., Choi, S., Lee, J.: Improving visual accessibility for color vision deficiency based on MPEG-21. ETRI J. **26**(3), 195–202 (2004)
32. Zeng, C., Xing, C., Zhou, L.: Similarity measure and instance selection for collaborative filtering. In Proceedings of the 12th International World Wide Web Conference, pp. 652–658 (2003)
33. Zillmann, D.: The coming of media entertainment. In: Zillmann, D., Vorderer, P. (eds.) Media entertainment: the psychology of its appeal, pp. 1–20. Lawrence Erlbaum Associates, Mahwah (2000)
34. Zimmermann, G., Vanderheiden, G., Gilman, A.: Universal remote console—prototyping for the alternate interface access standard. Universal Access. Theoretical Perspectives, Practice, and Experience. 7th ERCIM International Workshop on User Interfaces for all. Revised Papers, 24–25 Oct. 2002, pp. 524–531 (2003)
35. Zimmermann, G., Vanderheiden, G., Trewin, S.: Interface Sockets, Remote Consoles, and Natural Language Agents: A V2 URC Standards Whitepaper, http://myurc.org/whitepaper.php. (2005)
36. Pignetti, L., Capria, F.: Interactive television today. DV Magazine, 1–9 (2001)