INTERACTIVITY IN MULTIMEDIA DOCUMENTS AND SYSTEMS

Synonyms: *User interaction; user input*

Definition: *Multimedia interactivity describes the set of possible actions a user can do to change the state of a multimedia system, e. g., the course of a multimedia document's playback.*

The rapid growth of network capabilities in the last decade has fostered the diffusion of a wide set of interactive multimedia documents and systems, ranging from multimedia portals and distributed e-learning applications to interactive games and systems for multimedia data retrieval. These applications usually allow users to actively interact with the system rather than being passive recipients of information.

The problem of describing, supporting and helping user interaction has been largely investigated in literature. Multimedia systems usually allow two types of navigation facilities to control a presentation [3]: to adjust the current time reference in a presentation playback or to follow a hyperlink. In the first case, the user interaction is similar to the VCR functionalities for videocassettes: the user can start, stop, pause, fast-forward or rewind a presentation playback. In the second case, the user may jump to a completely different section of the document.

This second situation is more difficult to achieve and support, since the interaction may affect and change the overall structure of the presentation. Moreover, multimedia applications contain both static objects, like text pages and images, and continuous media items, like video and audio files. A hyperlink can be easily inserted into a static object, as an example, the user easily recognizes the presence of an anchor in an underlined word in a text page, whereas in multimedia system, the association of links with e within components is considerably complex for continuous media, for which a common way of defining anchor is not already available. Consider, for example, the problem of inserting a hyperlink inside an audio file. A proposed solution is the use of *hotspots*, i. e. an icon which remains on the user screen for a determined time interval, on which the user may click to follow the link. Despite many efforts, this problem has not already found a suitable solution, and the discovery of hyperlinks inside continuous media is often still hard for the user.

From the synchronization point of view, a multimedia application may have different behavior when a user follows a link, since multimedia systems do not clearly define a notion of how much information the reader leaves: the link can bring the user to a new document, which completely replaces the source, or both destination and source (or a part of) may coexist in the final presentation.

The Amsterdam Hypermedia Model [3] solves this problem by defining the notion of *context of a hyperlink* which clearly states which components survive after the user interaction, and which others are replaced and added.

The standard SMIL, *Synchronized Multimedia Integration Language* [2, 5], whose third version is currently under definition, is a markup language which allows authors to design highly interactive multimedia presentations. In fact, multimedia documents

designed with SMIL can contain hyperlinks to other documents, or to some of their components, but they can also modify their behavior according to user interactions. A SMIL document contains both the description of the spatial layout of media items and their temporal synchronization.

The user can start the playback of a SMIL presentation and passively follows its natural evolution, or she can play a more active role. The interactions allowed are not limited to the choice of a link, but the user can freely move the mouse around the user interface and click on the media items displayed. The SMIL standard in fact allows to synchronize the behavior of an item to external events: as an example, the playback of a media item can be started, or stopped, when the user clicks on a particular image, or moves the mouse over, or out, of a video file.

Moreover, unlike other encodings like MPEG-4, SMIL can be considered an integration format, since it does not store the entire presentation in a composite "sealed container", but each stream is a single file which can be distributed across the network and easily reused: the SMIL specification simply describes how the components are displayed on the user screen and synchronized. In MPEG-4 all the files are defined and controlled by the content creator, while SMIL allows the specification of a set of alternatives which can be chosen by the user according to her preferences and settings: as an example, an user can personalize a multimedia document playback by selecting the language or the presence of subtitles like in a DVD environment. In this sense, according to Bulterman [1], the standard SMIL enhances the possibilities of interactions offered to the user who can now partially control and select the multimedia contents. Moreover, since the insertion of new media items in a SMIL presentation is very easy, it helps the implementation of tools which allow the user to annotate multimedia items.

We must note here that the new possibilities offered by interactivity in multimedia systems bring also some drawbacks. The presence of too much information in the user screen may lead to their cognitive overload and disorientation [4]: multimedia systems offer users freedom to navigate into a very large information space by selecting an own path, but the pluralism of available choices may overawe the user that is no longer able to manage it without proper navigation aids. Therefore with freedom comes complexity, and so, disorientation.

For this reason, interactivity enhances user experience only when implemented properly, and must be carefully designed for each multimedia application: the user should always be able to see what are the consequences of the available choices and receive a suitable feedback on the taken action. Moreover the author must provide guided tours or navigational aids that guide the user through the fulfillment of her goal.

Finally, since interactivity in multimedia systems deals with creating experiences which allow the user to do or make something, a new form of interaction between users and multimedia systems is the content adaptation process. By content adaptation process we mean the set of actions performed to adapt a multimedia presentation to the user context, i. e., the device, the screen resolution, the network connection, the user preferences as well as the situation in which she is immersed: as an example, audio files cannot be played in a silent ambient like a library.

This is a particular form of interaction, since this process is usually transparent to the user, who simply receives, in response to her request, a suitable document to be rendered in the surrounding environment, and does not take note of the selection process that has built the final result from a set of rather content equivalent alternatives. Nevertheless, it

is not less important since a multimedia application which is not able to adapt its content to the user context (or at least, to her device) cannot be played in any given situation.

See: Multimedia Synchronization – Area Overview; Multimodal Interfaces; Multimedia Content Adaptation.

References

1. **D.** Bulterman, "User-centered control *within* multimedia presentations", Multimedia Systems Journal, Vol. 12, No. 4-5, March 2007, pp. 423-438.
2. Bulterman et al, "Synchronized Multimedia Integration Language (SMIL) 3.0 Candidate Recommendation", http://www.w3.org/TR/SMIL3/, 2008.
3. L. Hardman, D. Bulterman and G. van Rossum, "The Amsterdam hypermedia model: adding time and context to the Dexter model", Communications of ACM, Vol. 37, No. 2, Feb. 1994, pp. 50-62.
4. D. Kirsh, "Interactivity and MultiMedia Interfaces", Instructional Sciences, Springer Netherlands, Vol. 25, No. 2, 1997, pp. 79-96.
5. D. F. Zucker and D. Bulterman, "Open standard and open sourced SMIL for interactivity", ACM interactions, Vol. 14, No. 6, Nov. 2007, pp. 41-46.