Multikonferenz Wirtschaftsinformatik (MKWI)
University of Duisburg/Essen, March 9-11, 2004
Cherif Branki, J. Felix Hampe, Reiner Helfrich, Karl Kurbel, Gerhard Schwabe, Frank Teuteberg, Stefan Uellner, Rainer Unland, and Gerhard Wanner (Editors)
Volume 3, pages 204-215
Publisher: Aka GmbH

*mobile*MM4U – framework support for dynamic personalized multimedia content on mobile systems

Ansgar Scherp
OFFIS, Oldenburg

Susanne Boll
University of Oldenburg

Abstract: Multimedia content today can be considered as the composition of media objects into an interactive multimedia presentation. Typically, multimedia content comprises continuous and discrete media objects such as videos and images which are arranged in space and time with interaction possibilities. Personalization of such multimedia content means that the content reflects a user's personal context. To achieve such personalized multimedia content, a manual creation of many different documents for all the different user contexts, however, is not feasible let alone economical, particular in the heterogeneous mobile world. Rather a dynamic process of selecting and assembling personalized multimedia content depending on the user context seems reasonable. In this paper, we are presenting mobileMM4U, a component-based software framework, which provides the generic components for typical tasks of multimedia personalized content specifically targeted at mobile devices. The functionality provided by mobileMM4U supports the software development process of mobile multimedia applications, as it alleviates application developers from the central device-specific content assembly task and lets them concentrate on the design of their actual application. With mobileMM4U, we contribute to an easier mobile multimedia application development and allow a shorter time to market in the so fast changing and moving world of mobile applications.

Keywords: component-based software framework, mobile devices, multimedia content personalization, user context driven multimedia generation, context-aware multimedia applications, mobile tourism applications

1 Introduction

Mobility is more and more becoming an essential aspect of our life, both in our private life and our pursuit of business [IFMO02]. The new mobile devices that we see recently on the market provide capabilities for multimedia applications such as color displays, faster CPUs, higher battery power, more main memory, and local disk space. This power and resources of today's mobile devices are already sufficient enough to play complex multimedia presentations including MP3 compressed audio files and MPEG videos, like

SMIL [SMIL01] presentations on a Personal Digital Assistant (PDA) using the PocketSMIL player [INRIA02]. The increasing memory capacity on mobile devices also allows to store large amounts of data on the mobile device.

Quickly the question arises, where the multimedia content that is delivered to and presented on the user's mobile device comes from. Mostly, today's content is still text-centric and provided in a one-size-fits-all fashion. To serve mobile devices usually existing Web content is transformed into "smaller" Web pages to be viewed on a Web browser on a PDA. However, this content does neither reflect the different personal interests of the user nor specifically the various characteristics of the end user's device, and, as mentioned before does not contain multimedia data. Consequently, a dynamic creation of personalized multimedia content for the requesting user is in need. Especially in the context of mobile systems this is a special challenge. For an optimal delivery of content to the mobile user the exploitation of the device's features and the user's special interest should both lead to a sound presentation at the user's side. Therefore, we look in a little more detail in the general process for the creation of personalized multimedia content for mobile devices in Section 1.1. In Section 1.2, we present related work in the field and motivate the mobileMM4U approach before we give an overview of the paper in Section 1.3.

1.1 Personalization of mobile multimedia content

The general personalization process of multimedia content is illustrated in Figure 1. The input parameters of this process can be characterized by three groups. First, we have the content from which the personalized content is selected and assembled. For a reasonable usage of multimedia data the media data always needs to be associated with so called metadata. Second, we have the descriptions of the user's personal and technical context. These parameters comprise also the specific mobile device characteristics. The third group of input parameters are those that influence the general structure of the resulting multimedia presentation, the temporal course, the spatial layout, and subsumes other preferences a user could have for the multimedia presentation. All these input values are used by an application which we call a *personalization engine*. Generally speaking, such a personalization engine now exploits the user information to select the media data of relevance. All media objects have to be chosen to meet the quality of service for the given device characteristics at the best. The parameters for the document structure, the layout and style parameters and other possible rules or constraints are used by the personalization engine to assemble the multimedia presentation. Figure 1 sketches the temporal and spatial arrangement of the selected media objects over time in a spatial layout following the document structure and other preferences selected. Though the figure draws a very general picture of the personalization process, it illustrates the system support one would need to develop systems generating personalized, context-aware multimedia content for mobile systems. From our point of view, the growing demand of creating personalized multimedia content can only result in support for a *dynamic* creation of multimedia content, especially in the heterogeneous mobile world.

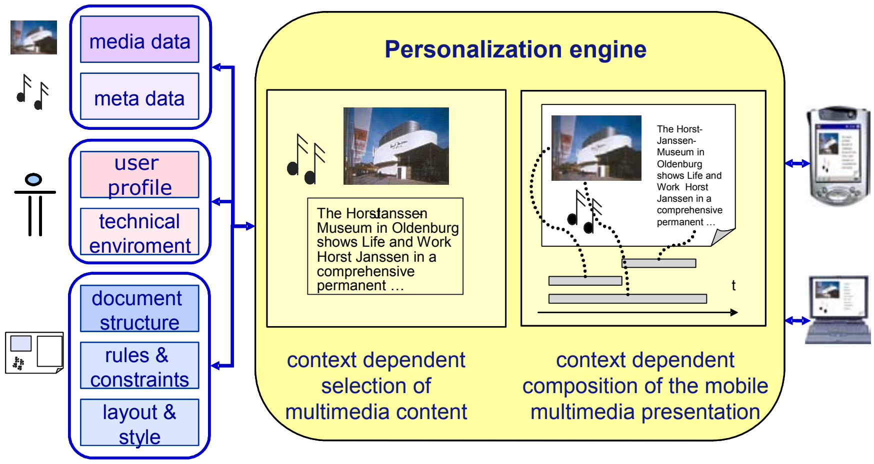

Figure 1: The general process of personalizing multimedia content

1.2 Related Work: Dynamic creation of multimedia content today

Looking on the dynamic creation of personalized content, we find different approaches which are mainly dealing with desktop PCs and Laptops as the end user device. Approaches that already dynamically create personalized *text-centric* content are typically found on the Web, e.g., the Internet book shop Amazon, and with adaptive hypermedia systems [BKV98, BBC02]. On the pathway to user-centered, on-demand generation of personalized *multimedia* content, we primarily find research approaches that address the personalization of *single* media only, like the semi-automatic home-video editor Hitchcock [GBS+01], or the personalized album MyPhotos [SZZ+02]. In the area of real dynamic generation of (personalized) *multi*media presentations, we find interesting research approaches with the Cuypers system [OCG+00, GOH01] and in the context of the OPERA project [BJK01]. Even though dealing with personalization, mobile devices are not in the focus of these approaches. For dynamic creation support of *mobile* multimedia content we find research approaches, e.g., with [MLK+01, LL03], providing first solutions for one of the most challenging tasks of current research, the development of personalized multimedia-based mobile information systems [BMBF02].

Summarizing, the approaches we see for multimedia content personalization use constraints, transformation rules and the like to achieve the generation of personalized multimedia content. However, our observation is, that these approaches are limited when it comes to more complex or very application specific personalization tasks. Therefore, we draw the conclusion that for the generation of complex personalized multimedia content, especially for mobile systems, programming is unavoidable anyway. Consequently, with mobileMM4U, we propose a software engineering approach in which we provide applications with the necessary components to dynamically generate personalized multimedia content for mobile devices. The mobileMM4U framework components relieve the developers of mobile multimedia applications from the central device-specific content assembly tasks and lets them concentrate on the analysis and design of their concrete application domain. It is embedded in our component-based software framework MM4U [Bo03]. The aim of mobileMM4U is to provide a substantial progress in the field of personalized mobile multimedia applications (more in Section 3). However, the research in the field of multimedia software engineering and especially software engineering for mobile devices is astonishingly sparsely populated.

1.3 Overview of the paper

The remainder of this paper is organized as follows: In Section 2, the MM4U framework, in which mobileMM4U is embedded, is introduced. The different components of mobileMM4U that support the dynamic creation of multimedia content on mobile devices and their special requirements are presented in Section 3 and are illustrated by our personalized city guide application. Section 4 concludes the paper.

2 The MM4U software framework

With MM4U we have developed a generic, desktop oriented software framework for the dynamic generation of personalized multimedia presentations respectively a framework for creating personalized multimedia applications. It is designed to be independent from any special domain, that is it can be used to generate personalized multimedia presentations for any application area like personalized sightseeing guide or personalized e-learning applications. The overall goal of MM4U is to simplify and cheapen the development of personalized multimedia applications. The focus of the framework so far is the creation of personalized multimedia content for stationary systems, not yet addressing the very specific requirements that arise in the context of mobile systems, we now meet with mobileMM4U. The general idea of the MM4U software framework is illustrated in Figure 2. A personalized multimedia application uses the functionality of the framework to create personalized multimedia content, and integrates this functionality in whatever application dependent functionality is needed. As illustrated in Figure 3, the framework is divided in several layers which provide modular support for the different tasks of the personalization process of multimedia content as shown in Figure 1. The framework is designed to be extensible by embedding additional functionality via well-defined interfaces as indicated by the empty boxes with dots. In the following, the features of the framework are described along the different layers of the architecture, indicated by the encircled numbers.

(1) *Connectors*: The components that bring user profile data and media data into the framework are the User Profile Connectors and the Media Data Connectors. They integrate existing systems for user profile stores, media storage, and retrieval solutions. As there are many different systems and formats available for user profile information, the User Profile Connectors abstract from the actual access to and retrieval of user profile information and provide a unified interface to the profile information. On the same level, the Media Data Connector abstracts from the access to media objects in different media storage and retrieval solutions that are available today with an unified interface.

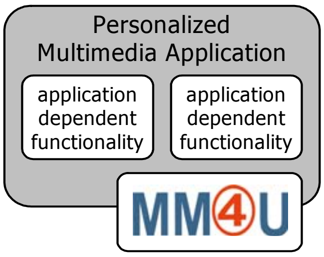

Figure 2: Framework usage

(2) *Accessors*: The Accessor components, i.e., the User Profile Accessor and the Media Pool Accessor, provide the internal data model of the user profiles and media data information within the system. Via these components the user profile information and media data needed for the desired content personalization is accessible and processable for the application.

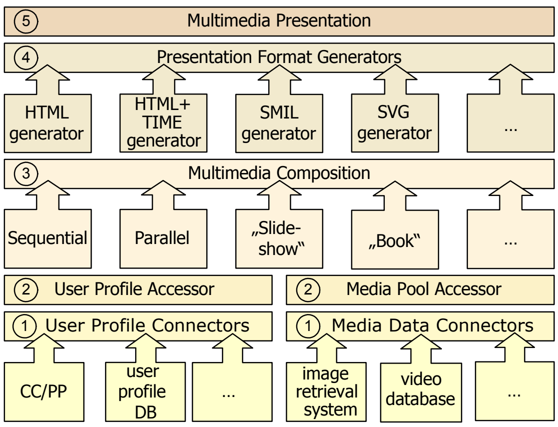

Figure 3: Overview of the multimedia personalization framework

The Connectors and Accessors are designed such that the framework's components are *not* re-inventing existing systems for user modeling or multimedia content management. They rather provide a seamless integration of the systems by distinct interfaces and comprehensive data models.

(3) *Multimedia Composition*: The Multimedia Composition component comprises abstract operators in compliance with the composition capabilities of multimedia composition models like SMIL, Madeus [JLR+98], and ZYX [Bo01] which provide complex multimedia composition functionality. It employs the information from the User Profile component and the Media Pool component for the composition task. The component is developed as such that it enables to seamlessly plug-in additional, possibly more complex or application-specific composition operators.

(4) *Presentation Format Generators*: The composition results in an internal representation of the multimedia content independent from the different end devices. The Presentation Format Generators work on this internal data model and convert it into a standardized presentation format like SMIL, SVG [SVG03] or HTML+ TIME [H+T98], that can be displayed by the corresponding multimedia player on the client device.

(5) *Multimedia Presentation*: The Multimedia Presentation component on top of the framework realizes the interface for applications to actually play the presentation of different multimedia presentation formats. The goal here is to integrate existing presentation components of the common multimedia presentation formats like SMIL, SVG, or HTML+TIME which the underlying Presentation Format Generator produces.

With this framework we have a basis for creating personalized multimedia presentations. For the specific requirements of creating such presentations for mobile systems, we have been developing mobileMM4U, the specific components for personalized mobile multimedia applications.

3 *mobile*MM4U – framework components supporting mobile multimedia applications

The aim of mobileMM4U is to provide a substantial improvement of the multimedia software development process for personalized *mobile* applications. Involved parties in the multimedia application development process is typically a heterogeneous team including media designers, computer scientists, and domain experts. The mobileMM4U framework alleviates the developer team from the device-specific content assembly task and lets them concentrate on the development of the actual application. The media designers create media objects and develop the multimedia layout which are input to the mobile framework. The domain experts provide the knowledge about the application domain and therewith the input for the computer scientists who develop the complex composition operators. For this, the computer scientists use the basic components of the mobileMM4U framework to actually develop the personalized mobile application. In the following, we concentrate on the view of the computer scientist on the media composition and generation process and present the framework components we developed to provide suitable support for the dynamic creation of multimedia content for mobile devices. We present the mobile support along the layers of the MM4U framework as presented in Section 2.

3.1 User Profile Accessor and Media Pool Accessor

User profiles contain contextual information to generate mobile multimedia presentations. This information comprises technical characteristics of the end device, the current location of the mobile user together with information about the surrounding and environment, as well as his or her personal preferences, interests and prior knowledge. The user profile information is read by the User Profile Accessor component and passed to the composition layer for further processing. The Media Pool Accessor uses contextual information from the Multimedia Composition layer as input for querying underlying media stores, such as the display size and color depth. The input parameters are mapped to the media metadata. The result is an ordered list of those media objects, which fulfill the request at the best. For example, if the display does not support colors, only greyscale media objects are selected, or if the mobile device is not able to render a particular media type, e. g., MPEG compressed videos, images will be selected instead.

3.2 Multimedia Composition

With respect to the multimedia composition for mobile systems, the corresponding composition operators of the Multimedia Composition layer have to take contextual information about the user and his or her device into account. The MM4U framework allows to develop complex application-specific composition operators. Complex operators in general exploit the basic operators the framework offers, e.g., Sequential, Parallel and Loop, and are enhanced by additional application-specific functionality. In mobileMM4U, we developed a complex operator *CityMap* addressing the requirements of mobile devices. The CityMap operator allows to adapt the mobile presentation to features of the user's end device, e. g., the screen size and the number of media used for it, and is applicable for any city where a personalized mobile tourist guide is to be developed. This complex operator is one example of extending the functionality of the MM4U framework towards mobile application support in mobileMM4U; here in the area of mobile tourism applications.

The CityMap operator, due to limited display size and computing power of mobile devices, requests only media objects from the Media Pool Accessor that meet the limitations of the mobile device as best as possible. The two screenshots of our prototypical city guide application in Figure 4 illustrate this: Figure 4a) is captured from a Tablet PC using the RealOne Player [RN03] and 4b) is taken from a Pocket PC using PocketSMIL Player [INRIA02]. In the given example, the Media Pool Accessor delivers smaller media objects of lower resolution for the presentation generation on the Pocket PC.

The limitations of mobile devices also effect the layout of the mobile presentation. Consequently, not only smaller media are selected, but also the layout of the multimedia presentation is changed by the composition operators. Since the city map image in

Figure 4 b) is automatically scaled by a factor of x < 1 to fit on the display of the mobile device, also the positions of the sightseeing spots on the map have to be recalculated. This is done by the complex operator *CityMap*. In mobileMM4U, the CityMap operator takes as input an image of arbitrary size representing the city map, a list of spots consisting of an ID and a global position in space, plus the factor x to (re-)calculate the new positions of the spots on the map. To make the most use of the limited display size of the mobile device, also the *Sightseeing4U* logo in the heading of the presentation is omitted by the composition operators, as can be seen in Figure 4 b). By this fewer media are used for the mobile presentation and, consequently, the amount of data to be transferred over the wireless network is reduced, too.

The relevant media are not only determined due to technical limitations of the mobile device, but also according to the user's environment such as location, time of day, weather, or noise. With regard to the location, the user receives only that multimedia presentation of the sight her or she actually stands in front of. Furthermore, if the environment during a city tour is too noisy, no audible media are selected for the presentation, but visual media only.

a) RealOne Player on a Tablet PC b) PocketSMIL Player on a PDA

Figure 4: Screenshots of the city guide presentation in SMIL format on different devices

Finally, the CityMap operator queries for those media objects that preferably match the users interests. For example, the complex composition operator for drawing the city map and the spots on it, will select building of renaissance style if this is one of the user's favourite epochs. If the user profile changes during the city tour, the complex operator will recognize this, and adapt the personalized route on its map by removing or adding sights on the map. The CityMap operator shows one example of how mobile application support can be integrated in mobileMM4U. As a further application example, we are currently developing a personalized multimedia sports news ticker for mobile devices.

3.3 Presentation Format Generators

In contrast to the multimedia composition operators, the Presentation Format Generators are independent of the concrete application domain and only rely on the targeted output format. In MM4U we developed generators for SMIL, SVG, and HTML [HTML98] in the context of stationary multimedia presentations. With mobileMM4U, we now provide special types of these Presentation Format Generators to map the internal representation of the multimedia presentation to the specific capabilities of the output format for mobile devices. The aforementioned multimedia document models already consider the limitation of mobile devices by means of building subsets and determining profiles that are particularly designed for usage on mobile devices.

In mobileMM4U, we developed Presentation Format Generators for the mobile versions of SMIL and SVG. The SMIL generator for mobile devices only generates a subset of the SMIL 2.0 document model, the SMIL 2.0 Basic Language Profile (see [SMIL01]), that is particularly adapted for mobile devices. The difference between both document models lies for example in the power of their layout operators. While in SMIL 2.0 nested layout regions are allowed, the Basic Language Profile of SMIL 2.0 only admits a flat layout hierarchy. This means, that in mobileMM4U the internal representation of the layout structure, which allows nested regions like in SMIL 2.0, is automatically broken down to a flat one. Another subset of SMIL we consider in MM4U is the 3GPP SMIL Language Profile [3GPPa03] targeted for Multimedia Messaging Service (MMS). This profile is a superset of the SMIL 2.0 Basic Language Profile and is used for scene description within the MMS interchange format [3GPPb03].

Figure 5: PocketSVG with SVGT on a PDA

Figure 6: Internet Explorer with HTML on a PDA

For SVG, we implemented the generators for the two profiles SVG offers for mobile devices (Mobile SVG). The profile SVG Tiny (SVGT) is intended for multimedia-ready mobile phones and the SVG Basic (SVGB) profile is aimed for pocket computers like PDAs and Handheld Computers (HHC) [MSVG03]. The mobile SVG presentations are limited with regard to the SVG modules supported on the mobile devices. Figure 5 shows the city guide application in SVGT format using PocketSVG

[PSVG02]. As we can not assume that there is a mobile multimedia player available for every "fancy" document model on the end user's device, there is always the fall back to HTML. If there is no specific multimedia player installed on the mobile device, e.g., a Pocket PC with build in Internet Explorer only, the personalized mobile application generates pure HTML as output format instead. The example in Figure 6 shows our personalized city guide in HTML for such a case.

3.4 Architectural and network issues

In this section, we draw our attention to some general architectural issues to show how the mobileMM4U framework is used by a concrete personalized mobile application to generate multimedia content for different mobile devices. The personalized application deploys an instance of the framework and runs on a server connected to the Internet. When a client sends a request to the server, the multimedia presentation is generated on demand by the framework as described in Section 2. Then the presentation is sent back to the client and displayed to the user. All application logic resides on the server-side and only the proper multimedia player needs to be installed on the client.

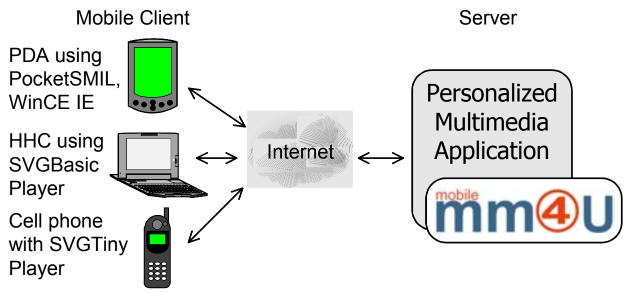

Figure 7: Usage of the framework by different clients

The architecture in Figure 7 assumes a permanent network connection between the client and the server, which is not a disadvantage in general. But for mobile applications and services this may not be appropriate because of the potential high network charges and possibly small network bandwidth. Also network failures that occur in areas with no reception, must be taken into account by the client application. The personalized city tour for example should not just stop processing and returning a general network error message to the user, when the connection is lost. To be more independent of the network connection it is inevitable that at least some parts of the application logic reside on the client device. That is the application logic is shifted from the server to the client, and that media data is buffered in client's caches. The amount of memory and processing power of today's available mobile devices are already sufficient enough for it, as we have shown in Section 1. The user profile information, media and metadata are then not read from the Internet, using the wireless network, but from locally cached databases and repositories. The mobileMM4U framework will be installed as any other application on the mobile device. The personalized multimedia application only needs to be notified where to find it. We are currently working on the specific aspects of a client-side mobileMM4U.

4 Conclusion

With the mobileMM4U framework, we presented a software engineering approach to support the development of arbitrary personalized mobile multimedia applications. mobileMM4U provides components for the automatic generation of personalized multimedia presentations in different output formats for different mobile end devices. The framework relieves the developers of mobile multimedia applications from the central device-specific content assembly tasks and lets them concentrate on their application-specific job. The developers also benefit from the fact that only players for standardized multimedia formats need to be installed on the mobile devices; assuming that mobile players for standard document models will become more popular and wide spread. The seamless integration of mobileMM4U with MM4U allows to dynamically create personalized multimedia content not only for mobile devices but also for desktop PCs and will lower the development costs of those applications that have to serve both desktop PCs and mobile clients. Due to their domain independence, MM4U and mobileMM4U can be used by arbitrary personalized multimedia applications, each application applying a different configuration of the framework of course. Consequently, for providers of mobile applications the framework approach supports a cheaper and quicker development process and by this contributes to a quicker time to market in times in which we find new types of mobile devices every six months.

Literature

[3GPPa03] 3rd Generation Partnership Project: TS 26.246; Transparent end-to-end Packet-switched Streaming Service; 3GPP SMIL language profile (Release 6), 2003.

[3GPPb03] 3rd Generation Partnership Project: TS 26.234; Transparent end-to-end Packet-switched Streaming Service; Protocols and codecs (Release 5), 2003.

[BBC02] P. D. Bra, P. Brusilovsky, and R. Conejo. Proc. of 2nd Intl. Conf. for Adaptive Hypermedia and Adaptive Web-Based Systems. Springer LNCS 2347, Malaga, Spain, 2002.

[BJK01] F. Bes, M. Jourdan, and F. Khantache. A Generic Architecture for Automated Construction of Multimedia Presentations. In Intl. Conf. on Multimedia Modeling (MMM 2001), Amsterdam, The Netherlands, Nov. 5-7 2001.

[BKV98] P. Brusilovsky, A. Kobsa, and J. Vassileva. Adaptive Hypertext and Hypermedia. Kluwer Acad. Publ., Dordrecht, 1998.

[BMBF02] Federal Ministry of Education and Research, Bonn, Germany: IT Research 2006, 2002.

[Bo01] S. Boll and W. Klas: ZYX - A Multimedia Document Model for Reuse and Adaptation. In: IEEE Transactions on Knowledge and Data Engineering, 13(3): 361-382, 2001.

[Bo03] S. Boll: MM4U – A framework for creating personalized multimedia content. In: Proc. of the Intl. Conf. on Distributed Multimedia Systems (DMS' 2003), Sept. 2003.

[GBS+01] A. Girgehnsohn, S. Bly, F. Shipman, J. Boreczky, and L. Wilcox. Home Video Editing Made Easy – Balancing Automation and User Control. In Proc. of the Human-Computer Interaction, pp. 464–471, Tokyo, Japan, 2001.

[GOH01] J. Geurts, J. v.Ossenbruggen, and L. Hardman. Applicationspecific contstraints for Multimedia Presentation Generation. In Intl. Conf. on Multimedia Modeling 2001, Amsterdam, The Netherlands, Nov. 5-7 2001.

[H+T98] P. Schmitz, J. Yu, Peter Santangeli, et al.: Timed Interactive Multimedia Extensions for HTML (HTML+TIME). W3C, version 09/18/1998

[HTML98] D. Raggett, A. Le Hors, I. Jacobs: HyperText Markup Language (HTML) version 4.0. W3C Recommendation, revised on 04/24/1998.

[IFMO02] Institute for Mobility Research (IFMO), Berlin, Germany: The Future of Mobility – Scenarios for the Year 2020, 2002.

[INRIA02] INRIA Rhône-Alpes: PocketSMIL 2.0, 2003. URL: http://opera.inrialpes.fr/pocketsmil

[JLR+98] M. Jourdan, N. Layaïda, C. Roisin, L. Sabry-Ismaïl, and L. Tardif: Madeus, and Authoring Environment for Interactive Multimedia Documents. ACM Multimedia 1998:267-272

[LL03] T. Lemlouma and N. Layaïda: Adapted Content Delivery for Different Contexts, SAINT 2003 Conf., Orlando, Florida, USA, January 27-31, 2003. IEEE, pp. 190-197.

[MLK+01] M. Metso, M. Löytynoja, J. Korva, P. Määttä, and J. Sauvola: Mobile Multimedia Services – Content Adaptation. 3rd International Conference on Information, Communications and Signal Processing, Singapore, 2001.

[MSVG03] O. Andersson, H. Axelsson, P. Armstrong, R. Berjon, et al.: Mobile SVG Profiles: SVG Tiny and SVG Basic. W3C Recommendation 01/14/2003

[OCG+00] J. van Ossenbruggen, F. Cornelissen, J. Geurts, L. Rutledge, and H. Hardman. Cuypers: a semi-automatic hypermedia presentation system. Techn. Rep. INS-R0025, CWI, The Netherlands, Dec. 2000.

[PSVG02] CSIRO Australia: PocketSVG, 2002. URL: http://www.pocketsvg.com/

[RN03] RealNetworks, USA: RealOne Player, 2003. URL: http://www.real.com/

[SMIL01] J. Ayars, D. Bulterman, A. Cohen, K. Day, E. Hodge, P. Hoschka, et al.: Synchronized Multimedia Integration Language (SMIL 2.0) Specification. W3C Rec. 08/07/2001

[SVG03] O. Andersson, H. Axelsson, P. Armstrong, C. Corporation, et al.: Scalable Vector Graphics (SVG 1.2) Specification. W3C Working Draft 11/13/2003

[SZZ+02] Y. Sun, H. Zhang, L. Zhan, and M. Li: MyPhotos - A System for Home Photo Management and Processing. In ACMMM 2002, Juan-les-Pins, France, Dec. 1-6 2002.