Canonical processes for creating personalized semantically rich multimedia presentations

Ansgar Scherp

Abstract  Authoring of multimedia content can be considered as composing media assets such as images, videos, text, and audio in time, space, and interaction into a coherent multimedia presentation. Personalization of such content means that it reflects the users’ or user groups’ profile information and context information. Enriching the multimedia content with semantically rich metadata allows for a better search and retrieval of the content. To actually create personalized semantically-rich multimedia content, a manual authoring of the many different documents for all the different users’ and user groups’ needs is not feasible. Rather a (semi-)automatic authoring of the content seems reasonable. We have analyzed in detail today’s approaches and systems for authoring, personalizing, and semantically enriching multimedia presentations. Based on this analysis, we derived a general creation chain for the (semi-)automatic generation of such content. In this paper, we introduce this creation chain. We present our software engineering support for the chain, the component framework SemanticMM4U. The canonical processes supported by the creation chain and SemanticMM4U framework are described in detail. We also provide an explicit mapping of SemanticMM4U framework components to the processes and argue for the benefits of defining canonical processes for creating personalized semantically rich multimedia presentations.

Keywords  Multimedia content · Multimedia authoring · Multimedia personalization · Multimedia semantics · Semantics derivation · Canonical processes · Media production · Component framework · Component technology

1 Introduction

Personalization of multimedia content and the dynamic creation of such content has been object of research for more than a decade. Today, we find many different scientific approaches as well as industrial solutions that provide personalized content to their users. However, the creation of personalized multimedia content is still a challenging and tedious task [9]. A practical support for the dynamic authoring of such content is neither provided by industrial solutions nor research projects. In addition, the systems we find today exploit semantically rich metadata for selecting media assets and organizing them into personalized multimedia content. However, this semantically rich and highly valuable source of information is not considered any further [32]. Once the multimedia content has been created, the semantically rich information exploited is thrown away. The created multimedia presentations carry either none or only a small part of the metadata that actually has been exploited for the creation task.

Consequently, we developed a generic and at the same time practical support for the dynamic authoring of personalized semantically rich multimedia presentations. This support not only exploits the semantically rich metadata for the media asset selection and organization into the personalized multimedia content but makes this metadata explicit and available by integrating it into the created multimedia presentations. In addition, it allows for deriving further metadata during the actual organization of the media assets into the multimedia content. The result of this research is a software engineering approach, the component framework SemanticMM4U (short for “Semantic MultiMedia For You”) for authoring personalized semantically rich multimedia presentations [32,27]. The overall goal of the SemanticMM4U component framework is to ease the creation of personalized
multimedia content and to derive semantically rich information about it. The SemanticMM4U framework is intended as an integrated approach as we believe that the different processes can be well supported and realized by different actors in the multimedia community.

The remainder of this paper is organized as follows: First, we introduce in Sect. 2 the general creation chain for authoring personalized semantically rich multimedia presentations. In Sect. 3, we present in detail the canonical processes involved in the different phases of this creation chain. In Sect. 4, we describe the goals and architecture of the SemanticMM4U component framework and provide a mapping of the framework components to the canonical processes. Finally, we present concrete applications using the SemanticMM4U framework and argue for the benefits of defining canonical processes in order to help the different actors in the multimedia community to work together.

2 Creation of personalized multimedia content

For the creation of personalized multimedia content, we find many different research approaches and industrial solutions. Well known examples from research are the Cuypers engine [17,36], the projects Opéra [3] and WAM [23,24], the SampLe framework [14,15], and the Standard Reference Model for Intelligent Multimedia Presentation Systems [7,13]. Examples from industry are, e.g., the HotStreams system [18] and the online bookstore Amazon [2] for text-centric content. These systems exploit semantically rich information to conduct the multimedia content creation task.

Based on an extensive analysis of these and further approaches and systems [27,32], we developed a general creation chain for authoring personalized semantically rich multimedia content [27,30–32]. This creation chain is illustrated in Fig. 1. As the schematic depiction shows, it consists of four phases. These phases span from media query, media organization, to the actual publishing and delivery of the content to the users. In the following, we present the phases of the creation chain from left to right.

2.1 Select

In the first phase, mono-media assets such as images, text, audio, and video are selected from the media storages. Media storages can be any of today’s off-the-shelf media databases such as Oracle10g interMedia [25] or the image database QBIC [19] from IBM. They can also be any self-developed solution for storing and managing media assets and their metadata. The query parameters to the media storages are among others the users’ or user groups’ profile information such as knowledge, preferences, interests, and needs [8,16,22] and context information like location, time, and characteristics of the end device [12,34,35] (described in detail in [27, Sect. 6.2]). For example, in the context of a mobile tourist guide the media assets of a specific sight are selected based on the user’s current location and the display capabilities of the mobile device. As a specific user of the guide might prefer an acoustic presentation rather than a visual experience, only audio clips are selected.

Besides returning already existing media assets that fulfill the query, also new media assets can be created on the fly. For example, if an image is requested in thumbnail size but is actually stored in full resolution, a downscaled version of the image can instantly be created to fulfill the request (see [27, p. 172]).

2.2 Assemble

In the next phase, the selected media assets are organized in time and space into a coherent, structured multimedia presentation. In addition, navigational interaction in form of hyperlinks can be defined. The multimedia presentations are created in an internal, tree-based multimedia content organization.