Journal of Network and Computer Applications 37 (2014) 25–32

Contents lists available at ScienceDirect

Journal of Network and Computer Applications

journal homepage: www.elsevier.com/locate/jnca

Multimedia contents adaptation by modality conversion with user preference in wireless network

Soocheol Lee [a], Seungmin Rho [b,*], Jong Hyuk Park [c]

[a] Electronic Documentation Division, Korean Intellectual Property Office, Seoul, Republic of Korea
[b] Department of Multimedia, Sungkyul University, Korea
[c] Department of Computer Science and Engineering, Seoul National University of Science and Technology, Seoul, Republic of Korea

ARTICLE INFO

Article history:
Received 8 June 2010
Received in revised form
22 December 2010
Accepted 30 March 2011
Available online 19 May 2011

Keywords:
Multimedia content adaptation
Quality of service
Transcoding
Transmoding
User preference
Wireless network

ABSTRACT

The rapid evolution in the hardware sector has brought up various media end user devices like PDAs or smart phones on which online multimedia content can be consumed. Due to different capabilities of these devices as well as individual user preferences, the original multimedia resources have to be adapted in order to fit the specific devices' constraints and to meet the user's requirements. Adapting multimedia content through user preferences is an important aspect of Quality of Service (QoS). There has been a vast amount of activities in research and standard development in this area. In this paper, we present methods for adapting multimedia contents to accommodate user preference. In addition, we propose the user preference content value model (u-CVM) that relates the content value of different modalities with resources. Experiments demonstrate that modality conversion brings a wider range of adaptation for QoS support. Moreover, the proposed approach has to be effective to be applied in Digital Multimedia Broadcasting (DMB) and Telematics.

1. Introduction

In ubiquitous computing environments, users may access and interact with multimedia content on different types of terminals and networks (Febiansyah et al., 2010; Guo et al., 2011). Such an environment includes a rich variety of multimedia terminals such as PC, PDA, or smart phones. One critical need in such a ubiquitous environment is the ability to handle the huge variation of resource constraints such as bandwidth, display, capability, CPU speed, power, and so on. This diverse set of terminal devices with different capabilities and heterogeneous networks with dynamically changing conditions in many cases results in interoperability problems. Therefore, one major goal in today's multimedia research is the development of Universal Multimedia Access (UMA) capabilities (MPEG, 2003a, 2003b; Vetro and Timmer, 2005), where a user is enabled to consume any resource, anywhere, and anytime. To achieve this goal, the multimedia content has to be adapted to meet the limitations of a user's terminal and network. Such multimedia adaptation could be, for example, transcoding from one video format to another or scaling a video in the spatial domain such that is fits on the terminal's screen. Furthermore, the content must also be adapted such that a user has an equivalent, informative experience anytime, that is, the end point of universal multimedia consumption is the end user and not the terminal.

Multimedia content adaptation is a key means for addressing the challenges of UMA (Mohan et al., 1999; Lei and Georganas, 2001). The general idea is to transform characteristics such as network bandwidth as well as capabilities of the end user's device being dynamically taken into account. In practice, QoS management can be done at both the network level and the application level. This paper is concerned with the application level, where content adaptation is an important solution to provide the QoS support.

Content adaptation (Nam et al., 2005; Thang et al., 2005; Fox et al., 1996, 1998; Vetro et al., 2003) has two major aspects: one is modality conversion as transmoding, which converts the content from one modality to a different modality, and the other is content scaling, which changes the amount of resource of the content without converting its modality.

Modalities can be derived from different modes of content coding. Even different coding formats for images (GIF, JPEG) are sometimes referred to as modalities or submodalities.

There are various conditions that may affect the decision on modality conversion (Thang et al., 2005; Ma et al., 2000). They can be grouped into four main factors: (i) the modality-presenting capability, which is the support to display certain modalities. This factor can be determined from the characteristics of a terminal or

* Corresponding author.
E-mail address: smrho@sungkyul.edu (S. Rho).

1084-8045/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jnca.2011.03.034

the surrounding environment, (ii) user preference, which shows the user's levels of interest in different modalities, (iii) the resource constraints of terminals or networks, such as the connection bit-rate, or the memory size available for the requested contents, and (iv) the semantics of the content itself. For example, between a news video and a ballet video, the provider would be more willing to convert the former to a stream of text.

The MPEG-7 standard (ISO, 2000; Ramanujan et al., 1997; Magalhaes and Pereira, 2004) provides an important metadata framework for describing multimedia summaries and user preference. As a result, there has been tremendous interest in building technology for UMA based on the MPEG-7 standard.

While MPEG-7 can be used to describe user's content preferences for personalized filtering, searching and browsing, MPEG-21 (MPEG, 2003a) expands the scope to address seamless access of multimedia across devices, networks, and users. In particular, MPEG-21 Part 7 "Digital Item Adaptation (DIA)" (Vetro and Timmer, 2005) supports user-centric adaptation of multimedia content to the usage environment. For example, MPEG-21 DIA addresses the customization of the presentation of multimedia content based on user's presentation preferences for content display/rendering, QoS, and configuration and conversion with regard to multimedia modalities. In addition, MPEG-21 DIA supports user accessibility and personalization. For example, MPEG-21 DIA facilitates adaptation of multimedia resources according to the description of a user's audio-visual accessibility characteristics.

In this paper, we propose a systematic approach that can convert the content objects of a multimedia document subjected to the resource constraint and the user preference. We formulate the content adaptation process of multimedia contents as a general constrained optimization problem and then extend it to support modality conversion. The content value model is employed to describe the relationship between the resource and the content values of different modalities.

The process of content scaling can be represented by rate–quality curve, which shows the quality of a scaled content according to the bit-rate. A recent trend is to use the rate–quality curve as metadata to automate content scaling. Usually, the curve is obtained for a particular modality because each modality has its own scaling characteristics. Based on the concept of the rate–quality curve, we proposed user preference content value model (u-CVM), which shows that the content value depends on the user preference. Through experiments, we verify that when modality conversion is applied, the user will receive high quality multimedia services. The proposed model is shown to be effective in supporting the user preference on modality conversion.

The paper is organized as follows. In Section 2, we describe the related work. In Section 3, we present an adaptation framework and profiles for adaptation. In Section 4, modeling of modality conversion is shown by the u-CVM, which relates the content qualities of different modalities. Experimental results, illustrating the performance of the proposed approach, are presented in Section 5. Section 6 concludes this paper.

2. Related work

There have been many research activities and advances in this field. Earlier works (such as Ma et al., 2000; Snoek and Worring, 2004; Halal et al., 2003) have explored some interesting aspects of adaptation like bandwidth reduction, format conversion, and modality replacement for Web browsing applications. Recently, international standards such as MPEG-7, MPEG-21, W3C, and TV Anytime have developed related tools and protocols to support the development and deployment of video adaptation applications. The modality capability of a terminal can be determined from Usage Environment descriptions of MPEG-21 Digital Item Adaptation (DIA). MPEG-21 DIA also provides the Conversion Preference tool to specify the user preference on modality conversion, and the Universal Constraints description tool to define the constraints of the adaptation. Meanwhile, W3C creates the Composite Capability/Preference Profile (CC/PP) protocol (W3C, 2004) to exchange the characteristics of users and terminals.

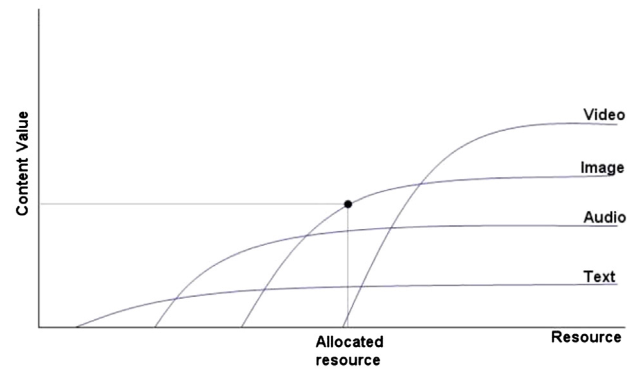

Fig. 1. OCV model.

Many previous researches have tackled specific cases of modality conversion under the constraint of modality capability. For example, text to speech technologies have been used for many applications such as receiving emails through telephones or reading textual contents on PC screens for blind users, and video to image conversion can be employed to send an image sequence to a terminal that does not support video playback.

In order to find the conversion boundaries between modalities, Thang et al. (2005) represented the relationship between the resource and content values of different modalities using the overlapped content value (OCV) model as shown in Fig. 1. In the figure, the OCV model of a video content consists of video, image, audio, and text curves. The modality curves are normally non-decreasing and saturate when the amount of a resource is large enough. Each point on the modality curves corresponds to a content version of that modality. The intersection points of the modality curves represent the conversion boundaries between the modalities. Though not really mandatory, each modality curve should cut another one, at most, at one point. Similar to rate–quality curves in image/video coding, the OCV model can be either operational or parametric.

The concept of UMA is to enable access to any multimedia content over any type of network, such as Internet, Wireless LAN, or others, from any type of terminal with varying capabilities such as mobile phones, personal computer, and television sets. The primary function of UMA services is to provide the best QoS or User experience by either selecting appropriate content formats or adapting the content formation directly to meet to playback environment or to adapt the content playback environment to accommodate the content.

Most content adaptation systems (Fox et al., 1996, 1998; Bickmore et al., 1997; Spyglass; QuickWeb; Avantgo) are HTTP proxy-based. The proxy intercepts client device's request for Web pages fetches the requested content, adapts it, and sends the adapted version to the client; this adaptation is often termed as "transcoding."

In the TranSend project (Fox et al., 1996, 1998), a proxy transcodes Web content on the fly. The adaptation, which they term "distillation," is primarily limited to image compression and reduction of image size and color space. Video is also converted