Contribution proposal for the WWRF-Meeting March 6-7, 2001
Subject area: Applications, Services, and the User's Perspective

Jörg Heuer, André Kaup, Uwe Rauschenbach

Adaptive Multimedia Messaging: Application Scenario and Technical Challenges

Siemens Corporate Technology, Information and Communications
81730 Munich, Germany
{Joerg.Heuer, Andre.Kaup, Uwe.Rauschenbach}@mchp.siemens.de

Abstract

This paper proposes multimedia messaging as an application for next generation mobile systems. The basic idea is to enable mobile users to send multimedia messages to each other, which can be seen as the convergence of today's voice mail boxes and multimedia email. We envision ubiquitous access to this messaging facility – it should be possible to generate and access multimedia messages to a large degree independently from the capabilities of the device used. In this document, we will highlight some of the technical challenges posed by such a scenario and sketch ways to meet them.

1 The Multimedia Messaging Scenario

With 3G mobile networks and beyond, a tremendous growth of connected mobile multimedia communication and computing devices will take place. This new device class will include mobile multimedia phones, personal digital assistants (PDAs) with built-in mobile network access and multimedia capabilities as well as automotive information and entertainment devices. All these devices will offer the user access to electronic information while on the move, but they will differ greatly in their capabilities and the capacity of the currently-used access network.

On the other hand, more and more electronic information takes the form of multimedia (images, graphics, video and audio), and it is expected that mobile users will make extensive use of it in the future. Assuming that, the future messaging scenario will be the convergence of today's multimedia emails and voice mailbox systems.

Figure 1 illustrates the envisioned scenario. Using a device such as a (mobile) telephone, a mobile multimedia-enabled phone, a PDA or a computer, a user composes a message consisting of different media items depending on the input capabilities of the device. Additional structuring information (annotations, key frame selection, summary) can be provided to enhance the information content of the message. The message is then sent to a message center which is an evolution of today's GSM mailboxes capable of intelligently handling multimedia messages. From there, the receiver can retrieve the message. The goal is that each multimedia message is accessible independently from what device is used for generating or fetching it. To realize this goal, the message center must use adaptive multimedia representations to support the adaptation of the multimedia content to the resources of the transmission channel, the capabilities of the client device as well as the preferences and the environment of the user.

A first exploration of the messaging scenario [3] has led to promising results, but has also shown the technical problems which remain to be solved. In the following, we will highlight some of these challenges.

2 Technical Challenges

Multimedia messages are more complex than voice messages, as Figure 2 illustrates. They may contain audio, video, still images, text and links to external information. That's why one challenge of this scenario is to provide *intuitive user support* for message composition, which usually involves a sort of timeline editing. Here, it is especially important to develop techniques which can intuitively be operated on small-format mobile devices with a limited number of keys and a small display. To disburden the user, as many functions as possible (e.g., extraction of shots and key frames from the video, generation of a panorama picture from a video pan etc.) have to be carried out automatically.

Messaging is always embedded into a *context*, which includes the capabilities of the sending and the receiving device, the network resources, the user preferences and the user environment (such as location). That means that devices of different brands must be enabled to efficiently communicate the context to each other and to describe the structure of the multimedia message in a common format with small overhead. To achieve this, *international standardization* is currently underway in the areas of meta data for multimedia (MPEG-7) [1] and device capabilities respectively user preferences (CC/PP) [2]. These new standards will provide the technological basis for universal access to multimedia information.

A third challenge is to *adapt the media items* to the context. This can be achieved by using transcoding technologies or scalable media formats. Transcoding may consist of intra-format (e.g., changing the bit rate of a video without changing the file format), intra-modality (e.g., converting a color image into a black-and-white bitmap and storing it in an image file format suitable for transmitting) and inter-modality (e.g., extracting key frames from a video or converting between text and voice) media transformations. It can be performed offline after the message has been recorded or online when the message is retrieved.

An important open issue is to find a set of context parameters which is as small as possible but sufficient to efficiently control the transcoding process.

References

[1] Overview of the MPEG-7 Standard, ISO/IEC JTC1/SC29/WG11 N3752, La Baule, October 2000.

[2] Composite Capabilities/Preference Profiles Working Group Homepage, The World Wide Consortium, http://www.w3.org/Mobile/CCPP/, 2000.

[3] Heuer J., Casas J. L., Kaup A. Adaptive Multimedia Messaging Based on MPEG-7 - The M3-Box. Proc. Second International Symposium on Mobile Multimedia Systems & Applications, Delft, The Netherlands, 9-10 November 2000.

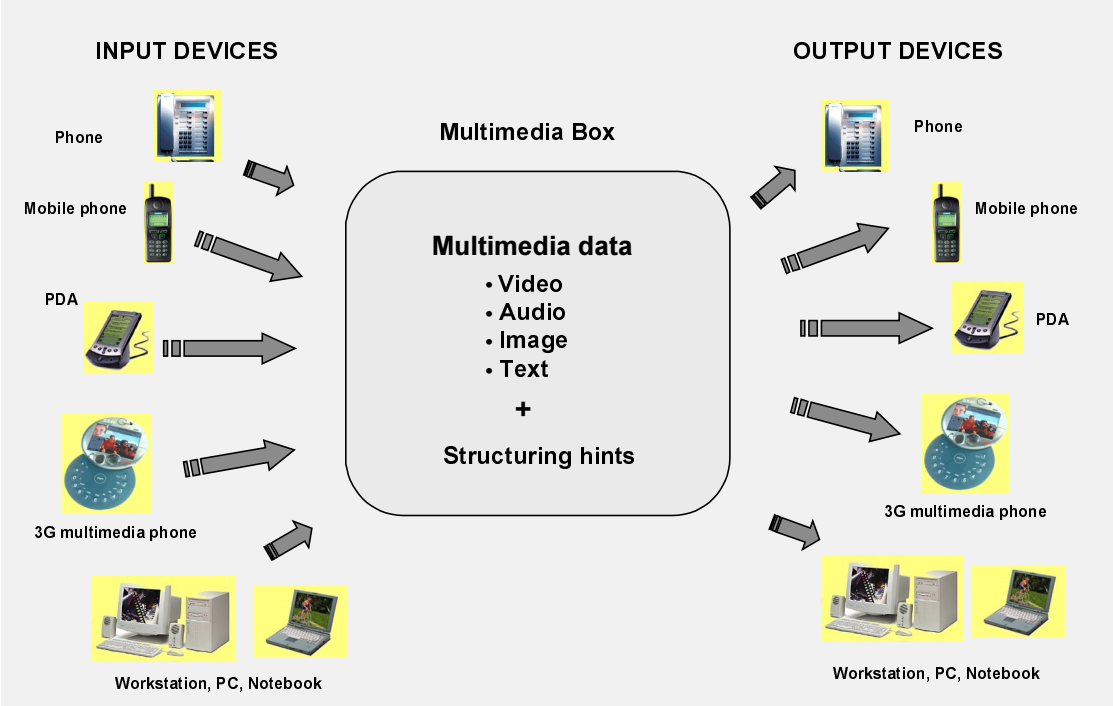

Figure 1: The messaging scenario.

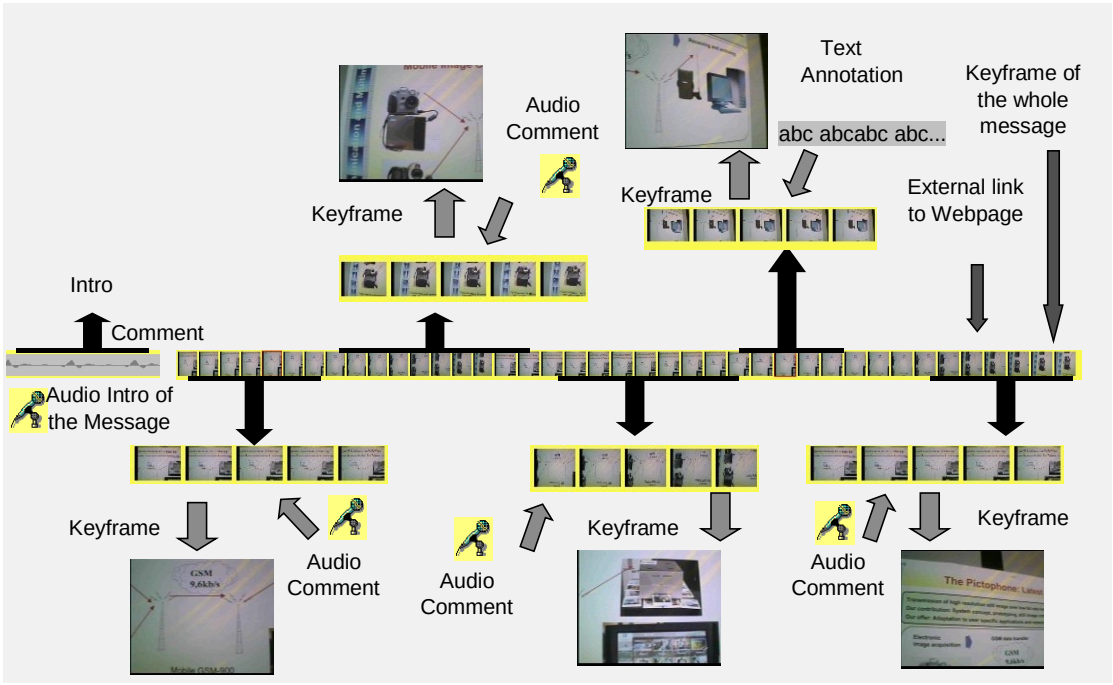

Figure 2: The possible complexity of a multimedia message, which may contain audio, video, still images, text and links to external information.