MPEG-21 SESSION MOBILITY FOR HETEROGENEOUS DEVICES

Frederik De Keukelaere Davy De Schrijver Saar De Zutter Rik Van de Walle Multimedia Lab Department of Electronics and Information Systems Ghent University Sint-Pietersnieuwstraat 41, B-9000 Ghent, Belgium E-mail: Frederik.DeKeukelaere@ugent.be

KEYWORDS Session Mobility, MPEG-21, Digital Item Adaptation

ABSTRACT

Nowadays, multimedia is becoming an important part of our live. Every day, the number of people that are having multimedia experiences is augmenting. To allow this growth in multimedia experiences, a new set of devices has been developed. Each of those devices has different terminal and network capabilities, or even different functionalities. As a result of this growing variety of devices, more and more people tend to use several devices for multimedia consumption. Having a wide set of devices for multimedia experiences, results in a demand for seamless switching between devices, better known as session mobility. In this paper, we discuss how session mobility can be realized between heterogeneous devices using MPEG-21 technology. First, we give an overview of the difficulties that occur when doing session mobility between heterogeneous devices. After this problem statement, we give a detailed discussion on how to overcome those difficulties by using MPEG-21. Throughout this paper, we will demonstrate how different parts of MPEG-21 can be integrated into a complete MPEG-21 compliant multimedia framework that facilitates session mobility between heterogeneous devices.

INTRODUCTION

Multimedia is becoming an important part of our lives. Every day, a lot people are having multimedia experiences. For example, while being at home, the possibility to consume multimedia tends to be ubiquitous. We can watch television, listen to the radio, record video using highly advanced camcorders, etc. Even when we are leaving our houses, multimedia stays with us during our journey. For example, multimedia appears as digital advertisements, in flight movies or simply some background music at work.



Figure 1: Universal Multimedia Access

Together with this upcoming availability of multimedia, we have a large set of new devices capable of bringing this multimedia experience to end users in a broad environment. Portable audio players, portable video players, powerful hi-fi devices and multimedia PCs are all creating an environment in which it is possible to have access to multimedia content, anywhere at any time.

This concept, accessing multimedia anywhere, at any time and on any device, is better known as Universal Multimedia Access (UMA) (Perkis et al. 2001, Vetro et al. 2003). The goal of Universal Multimedia Access, as presented in figure 1, is to create content once and afterwards allowing access to that multimedia content, anywhere, on any device, using any type of network and at any time.

People nowadays have a wide variety of devices at their disposal for multimedia consumption. Therefore, they tend to switch more often between those devices. To optimize the user experience when switching between devices, it is feasible that such transfers occur transparent without requiring complex user interactions. In this paper we describe how a low complexity interoperable framework allowing transparent transfer of multimedia sessions between devices with different capabilities can be created using MPEG-21. But before describing the MPEG-21 Session Mobility solution, let us have a look at the general problem definition and the difficulties that occur when transferring multimedia sessions between heterogeneous devices.

DIFFICULTIES WITH SESSION MOBILITY BETWEEN HETEROGENEOUS DEVICES

Transferring a multimedia session from one device to another, generally known as session mobility, has been studied in various domains. While some have focused on the mobility of sessions between applications e.g., between browsers (Song et. al 2002), others have focused on protocols that allow session mobility (Handley et. al. 1999, Schulzrinne et al. 2000).

To realize successful transfer of a multimedia session between two devices, it is possible to use the following three step protocol:

- 1. Collect session information
- 2. Transfer session information
- 3. Process information and continue session

It is straightforward to use this simple protocol in an environment where a session is transferred between devices with the same characteristics. However, currently we are moving towards a Universal Multimedia Access environment in which we have to deal with a variety of devices and session mobility between those devices tends to become more difficult. In this paper, we will address three difficulties that can occur when doing session mobility between heterogeneous devices.

The first difficulty that needs to be solved is the difference in terminal characteristics between devices. For example, watching a video on a terminal with a large screen and then transferring this video session to a terminal that does not support such a large screen will most likely require the adaptation of the video to a lower resolution. Other possible differences in terminal capabilities can cause similar problems for session mobility, such as differences in processing power, different availability of codecs and different battery capacity.

A second difficulty when using a wide set of devices, is the difference in the networks that connect different devices. Similar to differences in terminal capabilities, differences in network capabilities can result in extra complexity. For example, switching from a device with a broadband connection to a device with limited bandwidth will likely require switching to content encoded at a different bit rate in the new session. On the other hand, switching from a device with low bandwidth to a device with high bandwidth will most likely not result in the impossibility to use the same content on the new device. However using the same content would probably be a suboptimal solution, because

```
<?xml version="1.0" encoding="UTF-8"?>
<DIDL xmlns="urn:mpeg:mpeg21:2002:01-DIDL-NS">
    <Item>
        <Descriptor>
            <Statement mimeType="text/plain">
                Live: Throwing Copper
            </Statement>
        </Descriptor>
        <<u>Choice choice_id</u>="select track">
            <selection select id="track 1"/>
            <selection select id="track 2"/>
        </Choice>
        <Component>
            <Condition require="track 1">
                <Resource ref="Top.mp3" type="audio/mp3"/>
            </Condition>
        </Component>
        <Component>
            < Condition require="track 2">
                <Resource ref="Alone.mp3" type="audio/mp3"/>
            </Condition>
        </Component>
    </Item>
</DIDL>
```

Figure 2: A music album Digital Item

the new device has the possibility to receive higher quality content. Other differences in network characteristics can also have similar impact on session mobility; examples are differences in error rate of the network and different packet loss ratios.

As a final difficulty for session mobility between heterogeneous devices, we would like to address the interoperability of the messages that are sent between the different devices. To make it possible to do session mobility between devices with different characteristics, it is required to have a platform independent and lightweight language with a common, proprietary or standardized, representation for the messages carrying session information. Lack of a common representation, results in the impossibility to reconstruct the session on the new device based upon the session data of the originating device.

MPEG-21

ISO/IEC 21000, better known as MPEG-21 (Burnett et al. 2003), is the latest standard of the Moving Picture Experts Group (MPEG). In contrast to other MPEG standards, which are mostly targeted at audio and video compression, MPEG-21 is designed to become a generic framework for multimedia production and consumption.

Today, a typical multimedia production and consumption chain involves a large set of tools and standards for generating and consuming multimedia. Because MPEG-21 not only encapsulates data encoding and presentation, but also includes dynamical data adaptation, processing, and rights management, it will most likely succeed in reducing the



Figure 3: Digital Item Adaptation

multimedia production and consumption chain. MPEG-21 describes how different parts of multimedia technologies interoperate with each other and provides a set of description schemes that can be used in a process in which content is adapted to a specific set of terminal and network characteristics.

Because of the broad scope of MPEG-21, it is divided into several parts, of which Digital Item Declaration and Digital Item Adaptation are the most important parts in the context of session mobility. Most parts of MPEG-21 make intensive use of the Extensible Markup Language (XML) (Sperberg-McQueen et al. 1998) to describe their information.

MPEG-21's fundamental unit of distribution and transaction is the Digital Item (DI). A Digital Item is defined in MPEG-21 as a structured digital object with a standard XML representation, identification and associated metadata. A Digital Item can be seen as a composition of different resources like audio, video, text and metadata such as MPEG-7 descriptions, terminal capabilities and rights expressions.

In figure 2, a music album DI is described in the Digital This Digital Item Item Declaration Language. Declaration (DID) (ISO/IEC 2003) is а digital representation of a music album of the group "Live". The Item contains different components which represent album tracks. The descriptor describes the content of the Item (i.e., the music album) and the components contain resources with a reference to the place where the actual music data is stored (i.e., an mp3-file of the music track). In this example, the user has the ability to choose which track of the album he prefers to play. The choices are located in the choice-tags and based on the selections, the desired resources become available.

MPEG-21 Session Mobility is a part of ISO/IEC 21000-7 Digital Item Adaptation (DIA) (MPEG 2003). The main idea behind DIA is that Digital Items are subject to dynamic adaptation throughout their lifecycle. A high level adaptation process, which is typically used in DIA, is sketched in figure 3. In this adaptation process, the central part is the Adaptation Engine. This part performs



Figure 4: Choices in Digital Item Declarations

the adaptation of the Digital Item based on different inputs.

The first input is the original Digital Item. This Digital Item is declared in the Digital Item Declaration Language, and typically contains the multimedia data, together with some additional metadata. The second input to the Adaptation Engine is the adaptation information, typically providing information about the context in which the original Digital Item will be used. This information is standardized within the Digital Item Adaptation specification. The result of the adaptation process is an adapted Digital Item.

This adaptation process can be mapped to session mobility in the following manner. Transferring a multimedia session from one device to another device can be seen as the start of a new multimedia session on the target device (i.e., loading the original Digital Item) plus adapting that multimedia session, based on the information from the session on the originating device (i.e., the adaptation information). The result of the adaptation is the continued session on the target device, (i.e., the adapted Digital Item).

With regards to the technical aspects of session mobility, the Digital Item Declaration and the Digital Item Adaptation are the most important parts of MPEG-21. The other MPEG-21 parts that will become important when session mobility is implemented in a real life scenario are the parts of MPEG-21 that are related to Digital Rights Management (DRM). Within MPEG-21 the following parts will realize a DRM framework for multimedia: 21000-4 Intellectual Property Management and Protection, 21000-5 Rights Expression Language and 21000-6 Rights Data Dictionary.

OVERCOMING DIFFERENCES IN TERMINAL AND NETWORK CHARACTERISTICS

The discussion about solving the difficulties for session mobility between heterogeneous devices, will be split up into two sections. In this section we are going to discuss the difficulties of the differences in terminal and network characteristics between two devices. In the next section we are going to address the problem of having a common format for session information.

The main problem for differences in terminal and network characteristics is the fact that moving a session from one terminal to another one will often result in the impossibility to continue exactly the same session as the session on the first device. This is often the case due to the limitations of the hardware of the new device. Displaying a high resolution video on a low resolution display is simply impossible. When transferring in the opposite direction, it is probably possible to display the low resolution video on the high resolution video screen but this will not make optimal use of the hardware of the new device. In many cases this will not be accepted by the end user or at least it will be experienced as sub-optimal.

To solve this problem we use the Choices mechanism from the Digital Item Declaration Language. As mentioned earlier, it is possible to include choices in MPEG-21 Digital Item Declarations. Until now we have used Choices to choose between different parts of a Digital Item. For example, in figure 2 we could choose between different tracks of a music album. In the context of session mobility, Choices can be used to choose between different formats of multimedia content. For instance, in figure 4 the included choice allows the user to choose between two different movies, one in QCIF (176x144) format, and one in CIF (352x288) format. Similar to this type of choice, it is possible to include choices in the DID that are related to any type of terminal or network characteristic.

Figure 4 gives an example of a choice that is related to the terminal characteristics of a device. To be more specific, the choice in question is related to the size of the display. By including a choice in a DID, it is possible to create a Digital Item in such a way that it is possible to use the same Digital Item on devices with different capabilities. The only difference between the consumption of the multimedia content on one device and the consumption on another device will be the state of the choices that are included in the DID, i.e., the QCIF resolution will be selected on the terminal with the QCIF display, and the CIF or the QCIF resolution will be selected on the terminal with a CIF display.

For session mobility, the possibility to include choices in DIDs allows us to overcome the difficulties caused by differences in terminal and network characteristics. To allow transfer of a multimedia session with content at CIF



Figure 5: Automated Choice reconfiguration

resolution to a device with a QCIF display, it is only required to add a choice in the DID (QCIF or CIF). After transferring a session from the CIF device to the QCIF device, that choice needs to be reconfigured and after reconfiguration successful transfer has occurred.

The information included in figure 4 does not allow a machine to automatically choose between the QCIF and CIF format. The information about QCIF and CIF is contained within the select_id of the Selection element. From a machine's point of view, a select_id with the value "qcif" is semantically the same as a select_id with an arbitrary value, e.g., the value "a". The only functionality of the select_id with value "qcif" in the figure is that it makes the Component containing the QCIF video condition element with a required attribute that has the value "qcif". A full discussion on the different elements of the Digital Item Declaration (ISO/IEC 2003).

To make it possible to automatically configure Choices, it is necessary to include additional information in the Choices, more specific in the Selection elements that allow the run-time configuration of a DID. Additional information can be included in Selection elements using the Descriptor and Statement elements of the DIDL specification. Those DIDL elements allow the inclusion of metadata in Digital Item Declarations. Figure 2 contains an example of a Descriptor with a Statement element containing descriptive information about the content of the album, it contains the title of the music album.

Similar to the inclusion of descriptive data about the content it is possible to include descriptive data about the terminal and network characteristics. Those descriptions can give information about the requirements that need to be satisfied before a certain Selection can be chosen. Figure 5 gives an example of how to include additional information within a Selection element that allows machines to interpret the requirements for choosing a

Selection element correctly. The additional information included in this figure is Digital Item Adaptation A part of the Digital Item Adaptation information. specification standardizes a set of descriptors, called Usage Environment Descriptors, which allow authors to create descriptions of a usage environment. Usage Environment Descriptors make it possible to describe the terminal and network characteristics for a certain device. For session mobility, it is possible to use those Usage Environment Descriptors to allow correct interpretation, especially by machines, of the requirements for choosing a Selection element. Figure 5 contains an example of a Usage Environment Descriptor, describing the display of a terminal. In this example the Selection element can be interpreted as follows: "to choose this Selection, your terminal has to have a display with a horizontal resolution of 176 lines and a vertical resolution of 144 lines". This requirement is equal to the following requirement: "your terminal has to be able to display QCIF format". Using this additional information a machine can automatically make the Selections in a DID and configure that DID accordingly.

For session mobility, the configuration of Choices can be useful when transferring sessions between devices with different capabilities. For example, suppose we have a PC platform on which we are currently running a MPEG-21 multimedia session. The DID that is used in this session contains a set of Choices that allow the DID to be configured for both a PC and a PDA platform. Suppose now that the session from the PC needs to be transferred to the PDA. To realize such a transfer, the PC collects the information about the session, including the information about the Choices and sends that information to the PDA The PDA then processes the session platform. information, loads the Digital Item, and restores the state of the configuration of the Choices from the originating device. At this point the PDA tries to continue the session that was transferred from the PC. However, because of the differences in terminal and network characteristics of the two devices the continuing of the session fails. At this point the PDA can use the additional information in the Selection elements to reconfigure the Choices in such a way that it can actually continue the session. When the PDA tries to continue the adapted session, this time it succeeds, albeit this time with video at a lower bitrate and at a smaller resolution.

A COMMON FORMAT FOR SESSION INFORMATION

The next difficulty that needs to be addressed before session mobility can be realized successfully between heterogeneous devices is the format for the session information. Such a format needs to be lightweight and transparent. This will enable that format to be used on a variety of devices, going from powerful pc platforms, to more restricted, in terms of memory and processing power, platforms such as mobile phones. Because of the fact that we have been using the XML based DID format for the multimedia content, it would be feasible to use an XML based format for the session information as well. The ideal solution would be to use the same format, being the DID format, to store both the multimedia content and the session information. Using the same format for both purposes, allows the creator of multimedia terminals to reuse its software for both the session mobility part as for the content consumption part of the terminal and therefore reduces the required footprint for such terminals. This is especially important for devices with restrictions on processing power, memory, battery consumption, etc.

Within MPEG-21 Digital Item Adaptation a format for session information has been developed. This format, which is based on the generic DID format, allows the creation of session mobility Digital Items. Those DIs contain the necessary information about multimedia sessions allowing the reconstruction of such sessions based on the session mobility Digital Item.

For the example in figure 4, a session mobility Digital Item, conform the Digital Item Adaptation specification, can be created using the following steps. Initially an empty Digital Item is created using the DIDL elements DIDL and Item.

<?xml version="1.0" encoding="UTF-8"?> <DIDL xmlns="urn:mpeg:mpeg21:2002:01-DIDL-NS"> <Item> </Item> </DIDL>

In the second step, a Component/Resource child combination is added as a child of the Item from the first step. This Component/Resource contains a SessionMobilityTarget element which references the URI for the content DI. For the example we suppose the content has been placed on a local web server.

For each Choice element in the content DI there is an Annotation within the Item containing an Assertion (ISO/IEC 2003) that captures the current configuration-state of the Selections in the Choice of the content DI.

Suppose that the configuration of the Choice was a CIF resolution.

```
<?xml version="1.0" encoding="UTF-8"?>
<DIDL xmlns="urn:mpeg:mpeg21:2002:01-DIDL-NS"
       xmlns:sm="urn:mpeg:mpeg21:2003:01-DIA-SM-NS"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
       xmlns:dia="urn:mpeg:mpeg21:2003:01-DIA-NS">
<Item>
  <Component>
   <Resource mimeType="text/xml">
    <dia:DIADescriptionUnit
      xsi:type="sm:SessionMobilityTargetType"
      ref="http://192.168.0.1/cdi.xml" />
   </Resource>
  </Component>
  <Annotation target="http://192.168.0.1/cdi.xml#item 01">
   <Assertion target="http://192.168.0.1/cdi.xml#resolution"
             true="cif"/>
  </Annotation>
 </Item>
</DIDL>
```

As a final step in creating this session mobility Digital Item we create a Descriptor within the Item containing the information about the multimedia resource that is currently being consumed. For example, the location of the current media stream (e.g., the URI), the position in the current media stream (e.g., 50 sec) and the status of the current session (e.g., pause). This information is stored in a SessionMobilityAppInfo element of type SessionMobilityAppInfoType. The resulting session mobility Digital Item can be found in figure 6.

Based on the information in this session mobility Digital Item, it is possible to reconstruct the multimedia session that was previously running on the originating terminal. It is possible to determine what content was being consumed using the Component/Resource elements, i.e., the Digital Item located at "http://192.168.0.1/cdi.xml". It is also possible to configure the Choices in that Digital Item, i.e., CIF resolution was chosen. If necessary these choices can be reconfigured to adapt the Digital Item to the new terminal characteristics. Finally, it is also possible to restore the state of the resource consumption based on the information that is contained in the Descriptor/Statement children, i.e., the playbackstatus indicates the resource is being played, at a media time of 1.234898 seconds.

CONCLUSIONS AND FUTURE RESEARCH

In this paper we have discussed the difficulties that can rise when doing session mobility between heterogeneous devices. We have located three different causes for such difficulties. The first cause is the difference in terminal characteristics of heterogeneous devices. The second cause is the difference in network characteristics between heterogeneous devices. The final difficulty we discussed is the definition of a common format for expressing session mobility information.



Figure 6: A session mobility Digital Item

To overcome those difficulties when doing session mobility between heterogeneous devices, we investigated how MPEG-21 technology could be used to solve the different problems. MPEG-21, which is in fact a new and upcoming multimedia standard, aims to be used on a wide set of devices. To address the difficulties that occur when multimedia consumption is done on such a variety of devices, MPEG has developed the Digital Item Declaration Language. This language allows the declaration of Digital Items that are suited for consumption on devices with different terminal and network characteristics. To be more specific, the Choice elements in the DIDL make it possible for the content creator to include different Choices that allow the configuration of a Digital Item in such a way that it can be consumed on a variety of devices.

For session mobility we have investigated how the Choice mechanism can be used to allow the (re)configuration of Digital Items after a transfer of a multimedia session. We have demonstrated how this can be done by human interaction and without human interaction. To allow automated configuration of Choices, we have demonstrated how Digital Item Adaptation descriptions, more specifically, Usage Environment Descriptions, can be used for dynamic reconfiguration of the Choices without human interaction.

As a final part of the paper we addressed the problem of a common format for session information. Within MPEG-21 Digital Item Adaptation, such a common format has been developed based on the Digital Item Declaration Language. With this language, it is possible to create standardized session mobility Digital Items that will be correctly interpreted by MPEG-21 compliant terminals.

In this paper we have demonstrated how Digital Item Declaration and Digital Item Adaptation can be integrated to a complete MPEG-21 compliant multimedia framework that facilitates session mobility between heterogeneous devices.

Currently there are two other parts of MPEG-21 under development that can become more important for session mobility between heterogeneous devices. MPEG-21 Digital Item Processing and MPEG-21 Scalable Video Coding are those parts of MPEG-21 that most likely will allow MPEG-21 Session Mobility to reach its full potential. Further study on the integration of those parts of MPEG-21 in our current MPEG-21 Session Mobility framework can therefore be considered as future research.

ACKNOWLEDGMENTS

The research activities that have been described in this paper were funded by Ghent University, the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT), the Fund for Scientific Research-Flanders (FWO-Flanders), the Belgian Federal Office for Scientific, Technical and Cultural Affairs (OSTC), and the European Union.

REFERENCES

- Burnett, I., R. Van de Walle, K. Hill., J. Bormans, and Fernando Pereira, "MPEG-21 Goals and Achievements", *IEEE Transactions on Multimedia*, IEEE Computer Society, October-December 2003, p.60-70.
- ISO/IEC, "ISO/IEC 21000-2:2003 Information technology --Multimedia framework (MPEG-21) -- Part 2: Digital Item Declaration," March 2003.
- Handley, M., H. Schulzrinne, E. Schooler, and J. Rosenberg, "SIP: session initiation protocol," *Request for Comments 2543*, Internet Engineering Task Force, March 1999
- Perkis, A., Y. Abdeljaoued, C. Christopoulos, T. Ebrahimi, and J. Chicharo, "Universal Multimedia Access from Wired and Wireless Systems," *Birkhauser Boston Transactions on Circuits*, Systems and Signal Processing, Special Issue on Multimedia Communications, Volume 20, May-Aug 2001, p. 387-402.
- Moving Picture Experts Group, MPEG Home Page, http://mpeg.telecomitalialab.com/
- Moving Picture Experts Group "ISO/IEC 21000-7 FDIS Part 7: Digital Item Adaptation", ISO/IEC JTC1/SC29/WG11 N6168, Waikoloa, December 2003.
- Schulzrinne, H., and E. Wedlund, "Application-Layer Mobility using SIP", *Mobile Computing and Communications Review*, Volume 4, Number 3, pp. 47--57, July 2000.
- Sperberg-McQueen, C. M., T. Bray, and J. Paoli, The Extensible Markup Language (XML), http://www.w3.org/XML, 1998
- Song, H., H. Chu, S. Kurakake, "Browser Session Preservation and Migration," *Proceedings of WWW2002*, May 2002.
- Vetro, A., C. Christopoulos, and T. Ebrahimi, "Universal Multimedia Access," *IEEE Signal Processing Magazine*, 20 (2) 16-16, March 2003