Performance Analysis of Scalable Video Adaptation: Generic versus Specific Approach*

Michael Eberhard[1], Luca Celetto[2], Christian Timmerer[1], Emanuele Quacchio[2] and Hermann Hellwagner[1]

[1]*ITEC, MMC – Klagenfurt University, Austria,* [2]*STMicroelectronics, Italy*

[1]*firstname.lastname@itec.uni-klu.ac.at,* [2]*firstname.lastname@st.com*

Abstract

This paper provides a performance analysis of adaptation approaches designed for scalable media resources. In particular, we investigate the streaming of media resources compliant to the Scalable Video Coding (SVC) extensions of Advanced Video Coding (AVC) within heterogeneous environments, i.e., terminals and networks with different capabilities. Therefore, we have developed a test-bed in order to analyze two different approaches for the adaptation of scalable media resources, namely a generic approach that is applicable independently of the actual scalable coding format used and a specific approach especially built for SVC. The results show that if adaptation is required the generic approach clearly outperforms the approach specifically built for SVC.

1 Introduction

The streaming of multimedia content (audio and video) over the Internet or other heterogeneous networks raises a lot of research issues, due to problems like significant fluctuations of available bandwidth and different capabilities of terminals in terms of supported codecs, display resolution, processing power, energy consumption, etc. The provisioning of multimedia content tailored to all envisaged scenarios seems to be impractical, which calls for scalable representations of media information enabling the retrieval of degraded versions by applying simple removal (and minor editing) operations. Thus, it shall be possible to encode a multimedia stream once satisfying a plethora of application requirements ranging from High-Definition Television (HDTV) to Mobile TV. Recent advances in video coding resulted in the development of Scalable Video Coding (SVC) extensions of AVC which natively support temporal, spatial, and signal-to-noise ratio (SNR) scalability without losing coding efficiency compared to state-of-the-art single layer coding techniques and outperforming previous video coding techniques [1].

In the scope of this paper, the adaptation of scalable coding formats is defined as simple removal (and minor editing) operations of a given bitstream in best quality to a quality suited for the end user terminal. However, the growing number of scalable audio and video coding formats still requires codec-specific solutions. In order to address this issue, a generic solution adapting media resources independently of their coding format has been adopted to Part 7 of the MPEG-21 Multimedia Framework entitled Digital Item Adaptation (DIA) [2]. The adaptation is driven by XML-based metadata describing the high-level syntax of a (scalable) bitstream in terms of header, layers, and packets [3]. However, due to the description overhead introduced and due to the lack of competing implementations, such a generic solution so far does not achieve the performance of adaptation engines specifically built for a certain coding format [4].

In this paper we present a test-bed (cf. Section 1) for interoperable streaming of scalable video coding formats within heterogeneous environments. This test-bed has been used to analyze the performance of two contrary approaches enabling the adaptation of SVC: the generic approach as defined in MPEG-21 DIA [2] versus the specific one that comes along with the SVC reference software [5]. The results are presented (cf. Section 3) and discussed (cf. Section 4), showing that if adaptation is required the generic approach clearly outperforms the specific approach. Section 5 concludes the paper and points out possible future improvements.

* This work is support in part by the European Commission in the context of the P2P-Next project (FP7-ICT-216217). Further information is available under http://www.p2p-next.org/.

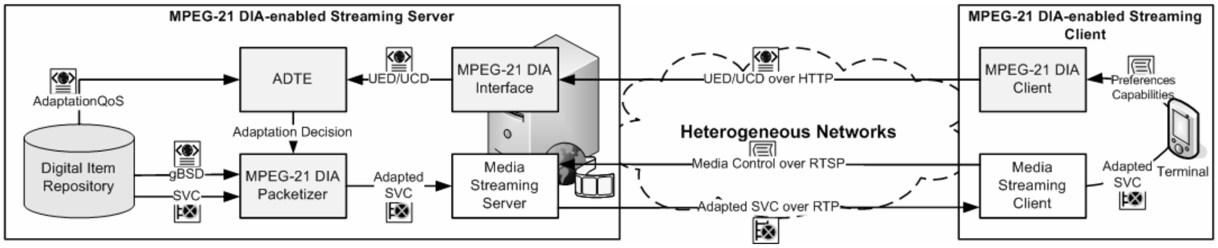

Figure 1. Test-Bed Architecture.

2 A Test-Bed for Interoperable Streaming of H.264/SVC

To demonstrate the usage of MPEG-21 DIA metadata to perform adaptations of scalable video sequences in a heterogeneous environment, the test-bed illustrated in Figure 1 was implemented.

The test-bed consists of the *MPEG-21 DIA-enabled Streaming Server* and a number of heterogeneous *Clients*, e.g., a Personal Digital Assistant (PDA), a Personal Computer (PC) and a Television-Set with a Set-Top Box (STB). The latter can decode only SVC base layer. Therefore, we have extended VideoLAN's VLC media player[2], which can be used as both, a streaming server and a streaming client, with MPEG-21 DIA capabilities for the server and the client side.

The streaming server receives the terminal capabilities and user preferences from the requesting client and provides it with SVC bitstreams in a suitable quality. To implement these functionalities, the VLC was extended by an *MPEG-21 DIA interface*, the *adaptation decision-taking engine* (ADTE) and an *MPEG-21 packetizer*. These extensions were integrated in the VLC as dynamic modules. The MPEG-21 DIA interface receives the Usage Environment Descriptions (UED) and the Universal Constraint Descriptions (UCD) [2] which contain the client's preferences and terminal capabilities and passes this information to the ADTE.

Based on the above mentioned descriptions the ADTE provides the adaptation decision by matching the information contained in the UED/UCD against the adaptation capabilities of the requested SVC sequence which are provided by means of the AdaptationQoS description [3]. The resulting adaptation decision of the ADTE is stored as *global adaptation parameters* in the VLC which can be accessed by all modules.

The adaptation of the SVC bitstream is performed by the MPEG-21 DIA packetizer. The packetizer receives the bitstream on an access unit level and performs the adaptation of the bitstream and of the corresponding parameter sets per Network Abstraction Layer Unit (NALU) utilizing the *global adaptation parameters* and the generic Bitstream Syntax Description (gBSD) [3]. Finally, the adapted SVC access units are passed to the *Media Streaming Server* which transmits the SVC bitstream to the client.

The streaming client was implemented in a portable way to make it usable on as many devices as possible. It extends the VLC with an *MPEG-21 DIA Client* that receives the user preferences and terminal capabilities as a user input. This data is wrapped into UED and UCDs and is subsequently transmitted to the streaming server using the Hypertext Transfer Protocol (HTTP) where this data is used to steer the adaptation decision as described previously.

In order to receive the adapted SVC bitstream, the *Media Streaming Client* needs to request the bitstream from the *Media Streaming Server* using the Real-Time Streaming Protocol (RTSP). Hence, the MPEG-21 DIA functionality has been added on top of an existing media streaming server/client solution without introducing legacy issues.

Finally, it should be noted that the streaming server allows each client to dynamically update its properties of the usage environment. Therefore, the terminal needs to pass new preferences or capabilities to the MPEG-21 DIA client which updates the metadata descriptions and transmits these updates to the server, again over HTTP. At the server, the adaptation decision is taken again based on the updated descriptions and the global adaptation parameters are updated. As the adaptation of the SVC media data is performed based on an access unit level and the adaptation parameters are accessed for each adaptation step, the result of the dynamic update is visible as soon as the next access unit is sent to the client.

3 Performance Evaluation

For performance evaluation, we have compared the generic adaptation approach utilizing MPEG-21 DIA metadata with an SVC specific adaptation approach utilizing the so-called *Bitstream Extractor* application of the JSVM reference software [5]. The aim of these measurements is to show that for the adaptation of

[2] http://www.videolan.org/vlc/

Table 2. Layer Properties of the Mariposa Sequence (S1)

	Layer 4	*Layer 3*	*Layer 2*	*Layer 1*	*Layer 0*
Resolution	4CIF	CIF	CIF	QCIF	QCIF
Framerate	15	30	15	30	15
Bitrate(kbps)	3264	1571	1228	376	281
Size (MB)	49	21.8	17	5.2	3.9

Table 2. Layer Properties of the Ice Sequence (S2)

	Layer 3	*Layer 2*	*Layer 1*	*Layer 0*
Resolution	4CIF	4CIF	CIF	CIF
Framerate	30	15	30	15
Bitrate(kbps)	4512	3426	1005	773
Size (MB)	4.3	3.3	0.9	0.7

SVC bitstreams a generic approach can very well compete with a coding format specific approach.

In order to perform a feasible comparison of the two adaptation approaches, the adaptation of a number of SVC bitstreams of different sizes was measured in terms of execution time and compared against each other. All performance measurements were performed on a Dell Optiplex GX620 Desktop with an Intel Pentium D 2.8 GHz processor and 1024 MB RAM. The used operating system was OpenSUSE 10.2 and the performance measurements were made utilizing the OProfile System Profiler[3] for Linux (v0.9.3). For the SVC specific adaptation, the Bitstream Extractor of the JSVM reference software 9.1 was used. For the generic adaptation approach we have used our own implementation written in C++ and, thus, comparable with the JSVM reference software.

The results of the performance measurements are presented for two representative SVC bitstreams; their properties are given in Table 1 and Table 2. Note that for other bitstreams (e.g., foreman, coastguard) similar results as presented in this paper have been observed.

For both sequences, the layer with the highest number represents the SVC sequence in best quality. The other layers provide a lower quality and can be extracted by applying an adaptation process. It should be noted that each layer contains all lower layers, i.e., layer n includes layers m < n.

A comparison of the performance of the two adaptation approaches applied to these two sequences is shown in Figure 2. The figure shows the execution time in ms for the extraction of each layer.

The results show that when an adaptation to the second highest layer is performed, the performance of the generic MPEG-21 DIA approach and the SVC-specific Bitstream Extractor is very similar. However, in case an adaptation to a lower layer is performed and a larger number of access units are truncated, the performance benefit of the generic MPEG-21 DIA approach is increasing up to a factor of 2.

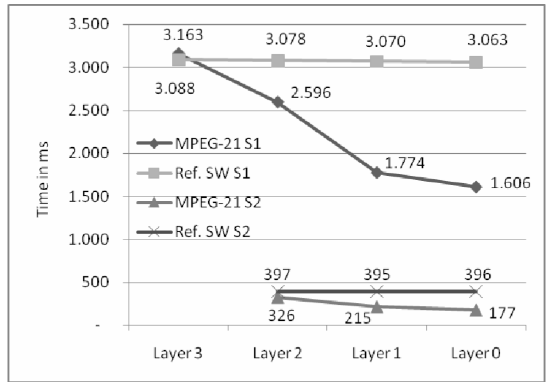

Figure 3. Performance Comparison.

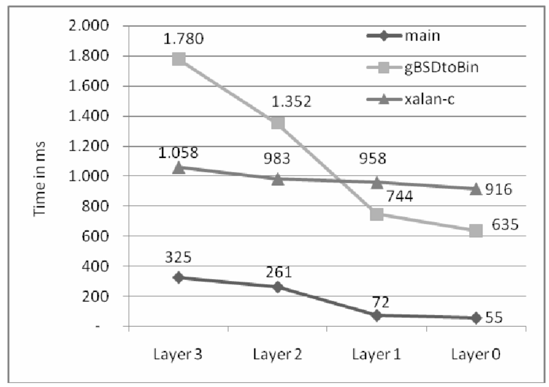

Figure 3. MPEG-21 DIA Performance.

To develop a better understanding why the generic MPEG-21 DIA approach outperforms the SVC-specific Bitstream Extractor, a more detailed illustration of the performance of the generic adaptation approach is depicted in Figure 3.

The generic MPEG-21 DIA metadata-based adaptation of the bitstream consists of three major tasks. First, the *Main* method performs the necessary preparations for the adaptation process and cleans up after the adaptation has finished. Second, the gBSD needs to be transformed based on the adaptation parameters for which the well-known Extensible Stylesheet Language for Transformations (XSLT) has been used. In particular, we have integrated the *Xalan*-C[4] XSLT engine (v1.10) due to performance reasons. Finally, the transformed gBSD is used by the *gBSDtoBin* process to extract the adapted bitstream [3].

Figure 3 shows the fraction that each major part of the generic MPEG-21 DIA approach takes to perform the adaptation for the Mariposa sequence (S1). The results show that the fraction of gBSDtoBin decreases when a lower layer is extracted whereas the fraction of the Xalan-C library remains rather constant for all lay-

[3] http://oprofile.sourceforge.net

[4] http://xml.apache.org/xalan-c/

ers. The behavior of the Main method is similar to the gBSDtoBin process.

4 Discussion

As presented in the previous section, the generic MPEG-21 DIA approach outperforms the SVC-specific Bitstream Extractor in case adaptation is performed. The main reason for the performance benefit can be recognized from Figure 3. While the execution time of the gBSDtoBin process decreases when a lower layer is extracted, the execution time of the XSLT transformation does not really change. The reason for this behavior is that the performance of the Xalan-C library remains rather constant for all layers, as the same gBSD needs to be transformed at all times. However, the gBSDtoBin process receives a smaller transformed gBSD when a lower layer is extracted and needs to extract only a smaller part of the bitstream. As the gBSDtoBin process does not need to analyze the whole bitstream but extracts only those bitstream segments which are described by the gBSD, the performance of the process improves significantly when lower layers are extracted. The Main method behaves similarly to the gBSDtoBin process, as the workload for the processing and writing of the files is reduced when lower layers are extracted.

The behavior of the gBSDtoBin process is also the main reason for the performance benefit of the generic MPEG-21 DIA approach in comparison to the SVC-specific Bitstream Extractor. On the one hand, the Bitstream Extractor always has to parse and analyze the complete bitstream, as there is no information about the size of the NALU available in its header. Thus, the Bitstream Extractor has to analyze every subsequent byte until a new NALU header is found. On the other hand, the MPEG-21 DIA approach only needs to parse and extract those segments of the bitstream which are needed for the generation of the adapted bitstream as their offsets are described by the metadata (i.e., gBSD). Thus, the performance gain of the MPEG-21 DIA approach is nearly exclusively obtained by the gBSDto-Bin process. For details about this process the interested reader is referred to [6].

Furthermore, the description of the behavior of the gBSDtoBin process also explains why the difference in terms of performance is the most significant between the layers 2 and 1 for the Mariposa sequence. As layer 2 has a size of 17 MB and layer 1 has a size of only 5.2 MB, the gBSDtoBin process needs to parse and extract a significantly smaller part compared to the Bitstream Extractor.

5 Conclusion and Future Work

A test-bed for the adaptation of SVC bitstreams based on MPEG-21 DIA metadata in a heterogeneous environment was introduced in this paper. This test-bed has been used to perform a comparison in terms of performance of the generic MPEG-21 DIA approach and the SVC-specific Bitstream Extractor of the JSVM reference software. The results show that the generic MPEG-21 DIA-based adaptation is at least as efficient as the SVC-specific reference software. If a large number of access units are dropped, the MPEG-21 DIA approach can be up to twice as fast as the reference software. In addition to this gain of performance, the MPEG-21 DIA approach is interoperable and can also be used to adapt any other scalable coding format.

Although the results of the performance comparison presented in this paper were already satisfactory, there are still some parts of the MPEG-21 DIA metadata-based adaptation process which allows for optimization, especially the transformation of the XML-based metadata. Additionally, the MPEG-21 DIA approach could be compared with a slightly modified SVC syntax where length information for NALUs is available followed by a detailed discussion of the pros and cons of this modification. Finally, different transmission strategies (e.g., adding layers instead of removing layers) and advanced buffer management techniques – specifically important in the context of dynamic adaptation – have not been discussed in this paper. Therefore, future research will investigate these issues while preserving the codec independence and general applicability of an XSLT library.

6 References

[1] M. Wien, H. Schwarz, T. Oelbaum, "Performance Analysis of SVC", *IEEE Trans. on CSVT*, vol. 17, no. 9, Sep. 2007, pp. 1194-1203.

[2] A. Vetro, "MPEG-21 Digital Item Adaptation: Enabling Universal Multimedia Access", *IEEE MultiMedia*, vol. 11, no. 1, Jan.-Mar. 2004, pp. 84-87.

[3] C. Timmerer, H. Hellwagner, "Interoperable Adaptive Multimedia Communication", *IEEE MultiMedia*, vol. 12, no. 1, Jan.-Mar. 2005, pp. 74-79.

[4] T. Zgaljic, N. Sprljan, E. Izquierdo, "Bitstream syntax description based adaptation of scalable video", *Proc. 2nd European Workshop on the Integration of Knowledge Semantic and Digital Media Technologies (EWIMT)*, London, UK, Nov.-Dec. 2005.

[5] *Joint Scalable Video Model (JSVM) 10*, Joint Video Team (JVT) of ISO/IEC and MPEG & ITU-T VCEG, N8963, San Jose, US, April 2007.

[6] I.S. Burnett et al., (eds.), *The MPEG-21 Book*, Wiley, 2006.