Scalable Video Coding Guidelines and Performance Evaluations for Adaptive Media Delivery of High Definition Content

Michael Grafl, Christian Timmerer, Hermann Hellwagner
Alpen-Adria-Universität (AAU)
Klagenfurt, Austria
e-mail: {firstname.lastname}@itec.aau.at

Daniel Negru
CNRS LaBRI Lab.
University of Bordeaux
Bordeaux, France
e-mail: daniel.negru@labri.fr

Wael Cherif
Viotech Communications
France
e-mail: wael.cherif@viotech.net

Stefano Battista
bSoft Ltd.
Macerata, Italy
e-mail: bautz@bsoft.net

***Abstract*—Scalability within media coding allows for content adaptation towards heterogeneous user contexts and enables in-network adaptation. However, there is no straightforward solution how to encode the content in a scalable way while maximizing rate-distortion performance. In this paper we provide encoding guidelines for scalable video coding based on a survey of media streaming industry solutions and a comprehensive performance evaluation using four state of the art scalable video codecs with a focus on high-definition content (1080p).**

Keywords-scalable video coding; adaptation; high-definition video; encoding; adaptive media streaming; content-aware networking

I. INTRODUCTION

The need for scalability (e.g., spatial, temporal, signal-to-noise ratio) in video coding is often motivated to address heterogeneous environments in terms of terminal characteristics (e.g., different resolutions) and network conditions (e.g., varying available bandwidth). Recently, the call for proposals on scalable video coding extensions of High Efficiency Video Coding (HEVC) was issued [1]. Todays' state of the art solution is Scalable Video Coding (SVC), an extension to the Advanced Video Coding (AVC) standard which employs a cumulative layered coding approach [2]. In addition to temporal scalability of AVC, SVC supports spatial and quality scalability. Quality scalability can be achieved through coarse-grain scalability (CGS), which uses the same mechanisms as spatial scalability but at a single resolution, or through medium-grain scalability (MGS), which enables a finer granularity for adaptation per video frame. For the MGS mode, most encoders, such as the reference software Joint Scalable Video Model (JSVM) [3], perform requantization, the quantization parameter (QP) for which is configured manually.

The deployment of SVC has an important role in adaptive media streaming. In particular, it allows the adaptation to the users' contexts and enables in-network adaptation in emerging content-aware networks [4]. Media-Aware Network Elements (MANEs) can adapt SVC streams on-the-fly during the delivery to accommodate fluctuating network conditions (e.g., congestion) [5]. For this technique to work, the content has to be encoded appropriately, taking expected terminal capabilities (such as resolution) and characteristics of the codec into account.

This paper devises encoding recommendations for SVC for adaptive media streaming applications based on a survey of media streaming industry solutions. The rate-distortion (RD) performance of these recommendations is validated for various encoders and several further encoding configurations for adaptive media streaming are evaluated focusing on high-definition (HD) content.

The remainder of this paper is structured as follows. Background and related work are highlighted in Section II. In Section III, we develop recommendations for SVC streaming. These recommendations and a set of encoding configurations are evaluated in Section IV. Section V concludes the paper and gives an outlook on future work.

II. BACKGROUND AND RELATED WORK

Various studies of SVC performance have been performed, incorporating either objective evaluations [2][6] or subjective evaluations for different application areas [7]. However, most studies are restricted to settings with only two SVC layers and are only concerned with the performance of the highest layer. A broader range of SVC settings is assessed in [8], including an evaluation of the best extraction path (i.e., whether to adapt in spatial, temporal, or quality direction). A survey of subjective SVC evaluations is given in [9]. SVC-based adaptation techniques are investigated in [10] and [11]. To the best of our knowledge, no research has been conducted to evaluate different SVC encoding configurations for adaptive media streaming of HD (1080p) content. Also, the available performance evaluations have used arbitrary spatial resolutions and bitrates rather than

considering configurations that are actually applied by industry solutions.

In addition to the reference software, JSVM, several proprietary SVC encoders exist. To the best of our knowledge, the most prominent ones are MainConcept[1], VSS[2], and bSoft[3]. Note that the encoders exhibit different encoding configuration options and yield individual bitstream characteristics. Performance tests of all these encoders will be presented throughout the paper.

Peak Signal-to-Noise Ratio (PSNR) is one of the most widely used full reference metrics for objective video quality assessment due to its simplicity and its low computational requirements.

The NTIA Video Quality Metric (VQM) [12] is a standardized full-reference objective method. VQM compares an original video sequence to a distorted sequence in order to estimate the video quality by combining perceptual effects of several video impairments such as blurring, jerky/unnatural motion, global noise, block distortion, and color distortion. VQM was specifically designed to correlate better with the human visual system than PSNR [13]. Therefore, we also use VQM results in addition to PSNR in our performance tests.

III. ENCODING RECOMMENDATIONS

A. Multi-Bitrate Streaming of Single-Layer Formats

Despite academic activity and performance studies of SVC, scalable media coding has only recently gained attention by the industry. In order to establish recommendations for SVC-based video streaming, we take a look at existing industry recommendations for multi-bitrate streaming of single-layer video formats. Among the most prominent streaming solutions and platforms are: Apple HTTP Live Streaming (HLS), Adobe HTTP Dynamic Streaming (HDS), Microsoft Smooth Streaming, YouTube, Netflix, Hulu, and MTV. Several of these technologies (namely Apple HLS[4,5], Adobe HDS [14][15], Microsoft Smooth Streaming [16][17], YouTube[6,7], and MTV [16]) provide recommendations for content encoding. We briefly analyze those recommendations and deduce suggestions for SVC streaming.

The spatial resolutions listed in those recommendations range from QCIF (176x144) at bitrates around 50 kbps (even 112x64 for thumbnail display, to be precise) up to 1920x1080 at maximum bitrates around 8 Mbps. All investigated recommendations suggest progressive scan. In general, around one to four streams per resolution are suggested. Resolutions 1280x720 and 1920x1080 are common to most platforms, but at lower resolutions, both the exact resolution and aspect ratio are different across platforms. A list of resolutions and bitrates of the discussed recommendations is provided on a dedicated Web page[8].

TABLE I. DEVISED BITRATE RECOMMENDATIONS FOR SVC STREAMING.

Resolution	Bitrate suggestions		Dyadic spatial scalability
	4 bitrates [kbps]	*2 bitrates [kbps]*	
1920x1080	10400, 7200, 5500, 4000	8800, 6050	down
1280x720	7800, 4800, 2750, 1500	5000, 2750	down
704x576		2200, 1350	down
960x540		2475, 1980	up
640x360		1760, 660	up
352x288	1950, 1080, 500, 270	1320, 330	up & down
176x144		110, 55	up

Since AVC and other common video codecs use macroblock sizes of 16x16 block luminance samples [2], resolutions divisible by 16 are better suited for optimizing coding performance (known as *mod-16* rule). Less than half of the resolutions adhere to this rule. Note that some encoders, e.g., the bSoft encoder, try to optimize coding performance by removing those incomplete macroblocks and, thus, cropping a small part of the video.

Less than a quarter of the investigated resolutions support dyadic downscaling, the same holds for dyadic upscaling, but none meets both criteria. This means that the used resolutions would not support SVC encoding with three dyadic spatial resolutions. Furthermore, the CIF resolution (352x288), which is commonly used in research literature, is only used in one encoding recommendation; most other streaming solutions prefer 512x288, which has a wider aspect ratio. None of the recommendations lists the 4CIF resolution (704x576).

Since all recommendations target single-layer formats, the support of dyadic spatial scalability is irrelevant in their scenarios, which is reflected by the choice of recommended resolutions.

Although many of the investigated industry solutions deploy HTTP streaming, the coding guidelines we devise in this paper are applicable to SVC media streaming in general.

B. Deduced Recommendations for SVC

Based on the encoding recommendations of industry solutions for multi-bitrate streaming, Table I comprises a list of typical resolutions and bitrates for SVC streaming. These guidelines take the popularity of resolutions among streaming solutions, top and bottom bitrates, as well as bitrate steps into account. We placed special emphasis on assembling meaningful resolutions, reducing the number of different resolutions from Section III.A. With the exception of 352x288 and 704x576, each listed resolution is mentioned in at least two streaming solution recommendations.

Table I focuses on two and four bitrates per resolution. For those resolutions, for which the examined recommendations do not list sufficient different bitrates, the column for *4 bitrates* in Table I was left blank.

[1] http://mainconcept.com/

[2] http://www.vsofts.com/technology/scalable-video-coding.html

[3] http://bsoft.net/

[4] http://developer.apple.com/library/ios/#technotes/tn2224/_index.html

[5] http://developer.apple.com/library/mac/#technotes/tn2218/_index.html

[6] http://support.google.com/youtube/bin/answer.py?hl=en&answer=1722171&topic=1728573

[7] http://support.google.com/youtube/bin/answer.py?hl=en&answer=1723080&topic=1731151

[8] http://alicante.itec.aau.at/am1.html

To account for SVC overhead, bitrates are increased by 10% compared to the examined recommendations (with some rounding where appropriate). Research literature typically assumes a coding overhead of 10% per enhancement layer compared to single-layer AVC [2][6]. We argue that this bitrate increase can be safely applied as network traffic forecasts [18] show a continuous increase of video network traffic and connection speeds that would easily accommodate the proposed bitrate increase. Furthermore, the use of SVC enables dynamic bitrate adaptation, alleviating the risk of stalling. For streams with 2 layers, we propose to add 10% overhead for both bitrates. However, with 4 layers, we keep the original bitrate for the base layer in order to support low bandwidths, increase the bitrate for the first enhancement layer by 10%, for the second by 20% and for the third by a total of 30%.

We added the 4CIF resolution in order to better support dyadic spatial scalability. That is, three spatial layers with resolutions QCIF, CIF, and 4CIF are possible. The table also indicates whether dyadic spatial scalability (up- or downscaling) is supported by the listed resolutions.

Depending on the scenario and targeted client devices, we suggest streams with a total of six to twelve extraction points out of the possible combinations in Table I. We also suggest allocating proportionally more bitrates per resolution for higher resolutions. For example, a configuration may contain the four bitrates indicated for 1920x1080 and two bitrates for 960x540.

IV. HIGH-DEFINITION SVC ENCODING PERFORMANCE

In this section, extensive performance evaluations of SVC with a focus on 1080p resolutions are presented, including various SVC configurations and different encoders (JSVM, MainConcept, VSS, and bSoft). The goals of these evaluations are (1) to provide RD performance results in terms of PSNR and VQM, (2) to investigate various encoding configurations, (3) to highlight the characteristics of different SVC encoders, and (4) to validate the encoding recommendations devised in Section III.B.

Four different video sequences were selected for performance evaluations, based on their Spatial Information (SI) and Temporal Information (TI) [19]: *PedestrianArea* (low SI, low TI), *Dinner* (low SI, high TI), *DucksTakeOff* (high SI, low TI), and *CrowdRun* (high SI, high TI). Due to space constraints, not all sequences are shown for each test.

We first evaluate rate control modes (i.e., constant bitrate vs. fixed QP) for different encoders in order to compare their RD performance and to validate whether the devised bitrate recommendations yield consistent qualities at all resolutions. Then, we test the combination of spatial and quality scalability to decide whether to encode one stream per resolution or all resolutions in one stream for media streaming scenarios. Another factor to adaptive streaming configurations is the number of quality layers for a given resolution, which affects the flexibility of adaptations at the cost of coding overhead. This aspect is evaluated with 1 to 4 quality layers for various encoders. Finally, we investigate the impact of requantization on the RD performance, which controls the bitrate distances between quality layers.

Unless noted otherwise, the resolution was set to 1920x1080 and the deltaQP (dQP) for requantization between MGS layers was set to 2 (the bSoft encoder does not need any requantization configuration, as explained later on).

Note that the following performance results are implementation-dependent and provide a snapshot of current SVC encoder performances. Nevertheless, we strive to highlight performance characteristics of SVC as a coding scheme that we expect to remain valid beyond the mere comparison of encoder implementations.

A. Rate Control Modes

We validate the bitrate recommendations of Section III.B in the following test, comparing constant bitrate (CBR) vs. fixed QP rate control modes of several encoders.

For each resolution (from 1920x1080 down to 176x144), bitstreams were encoded with 2 MGS layers. In CBR mode, target bitrates were set to the values stated in Table I (for 2 bitrates). For encoding with fixed QP, we selected for each sequence the two QPs that resulted in bitrates just above and just below the target bitrate of the enhancement layer for the respective resolution in Table I (for 2 bitrates).

The bSoft encoder requires an initial QP value even for CBR encoding. We noticed that fixed QP settings always yielded better RD performance than any CBR setting with that initial QP. Thus, only fixed QP rate control results are shown for the bSoft encoder.

The tested version of the MainConcept encoder has some limitations concerning supported bitrates for CBR mode. Therefore, we were only able to obtain results for 1920x1080 with target bitrates of 4,400 kbps for the base layer and 8,800 kbps for the enhancement layer.

The JSVM encoder was only evaluated for a resolution of 1920x1080 at fixed QP mode. Although the tested version of the JSVM provides basic CBR support, it only supports CBR mode at the base layer, making it unsuitable for our tests.

The PSNR results at a resolution of 1920x1080 are shown in Fig. 1. The bitrate ranges from the suggestions in Table I (for 2 bitrates) are indicated as green background. Corresponding VQM results for *PedestrianArea* and *CrowdRun* sequences are given in Fig. 2. Note that the y-axis of VQM results is an impairment scale from 1 (high distortion) to 0 (no distortion), indicating the expected quality of a sequence.

In terms of encoder comparison, JSVM outperforms the other encoders with respect to RD performance, followed by MainConcept and VSS. The bSoft encoder, while having somewhat lower PSNR results, has good VQM results (especially for sequences with high SI), which are on par with the other encoders in terms of rate-distortion. Since VQM correlates better with the human visual system, these results suggest that the actual visual quality of the bSoft encoder is significantly higher than indicated by the PSNR values. Similar to the behavior of the bSoft encoder, CBR modes of the MainConcept and VSS encoders tend to have better VQM than corresponding PSNR results.

When comparing rate control modes, we see that the MainConcept encoder achieves higher quality in fixed QP mode than in CBR mode. In contrast, the VSS encoder yields

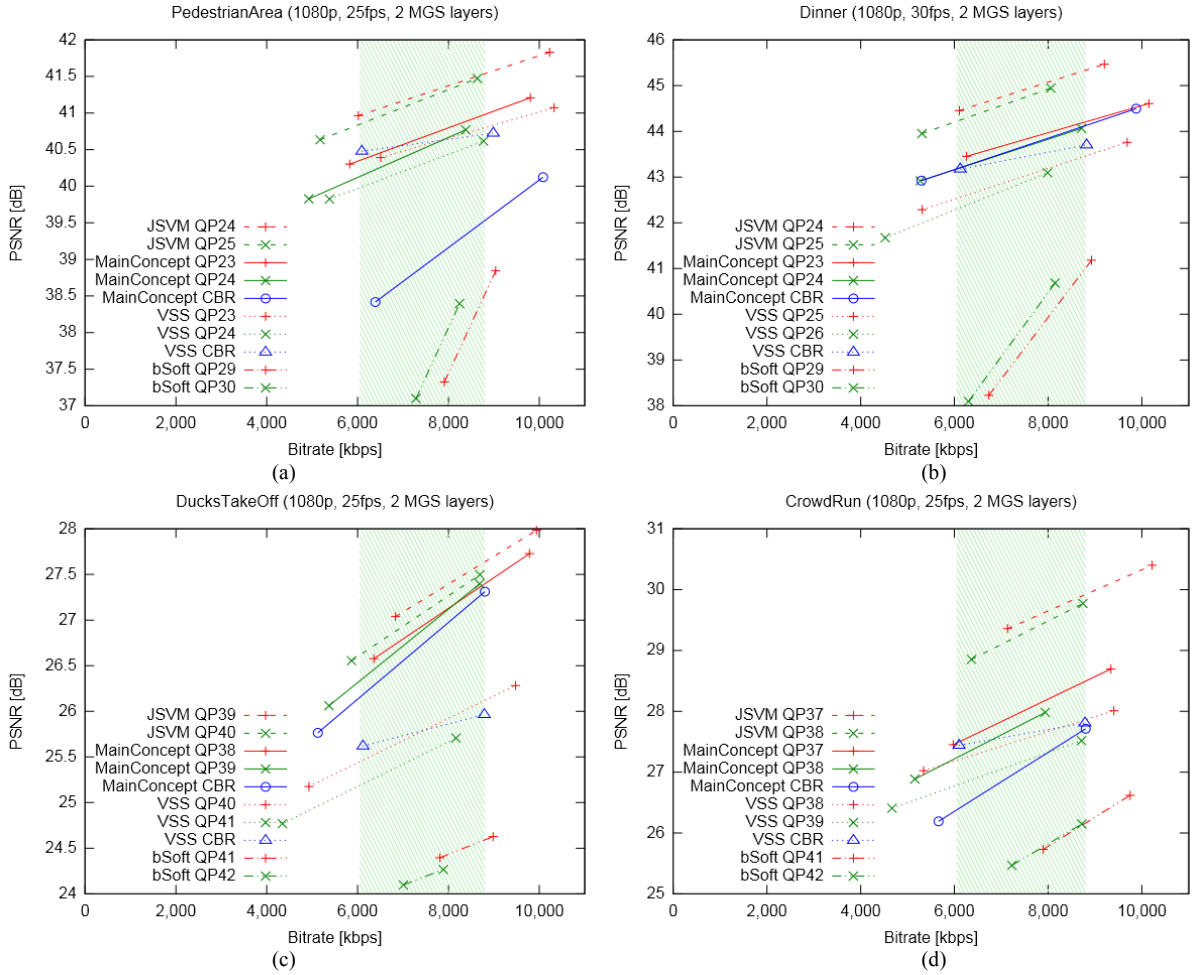


Figure 1. PSNR results of rate control modes for (a) PedestrianArea, (b) Dinner, (c) DucksTakeOff, and (d) CrowdRun sequences.

equal or slightly lower quality in fixed QP mode compared to CBR mode. Among the tested encoders and rate control modes, the VSS encoder in CBR mode shows the lowest decrease of RD performance towards the base layer.

VQM results for lower resolutions are presented for the *PedestrianArea* sequence in Fig. 3. Again, the suggested bitrate ranges from Table I (for 2 bitrates) are indicated as green background. The suggested bitrates yield quite constant qualities across all resolutions, except for the lowest resolution QCIF. The rationale behind low target bitrates for QCIF is to enable (a fallback) video transmission even for very low bandwidths.

As already observed for 1920x1080, the VSS encoder in CBR mode and the bSoft encoder (in fixed QP mode) perform consistently better in terms of VQM results compared to PSNR results. The VSS encoder tends to have better VQM results in CBR mode for all resolutions and is almost on par with JSVM.

B. *Combining Spatial Scalability and MGS*

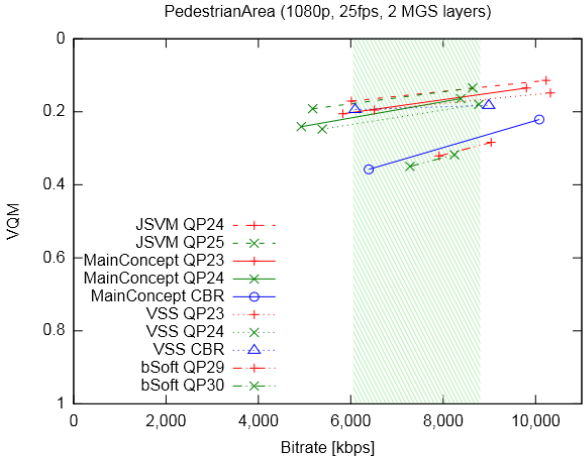
In the following test, we investigate the RD performance of spatial scalability at two resolutions combined with two MGS layers. We compare the RD performance to bitstreams with two MGS layers at each resolution.

This configuration is also relevant for determining whether to use one SVC bitstream for multiple resolutions or to use separate SVC bitstreams featuring quality scalability for each resolution in SVC streaming scenarios. Due to space constraints, only the results for the bSoft encoder with the following configuration are presented: resolution 1 is 960x528, resolution 2 is 1920x1056.

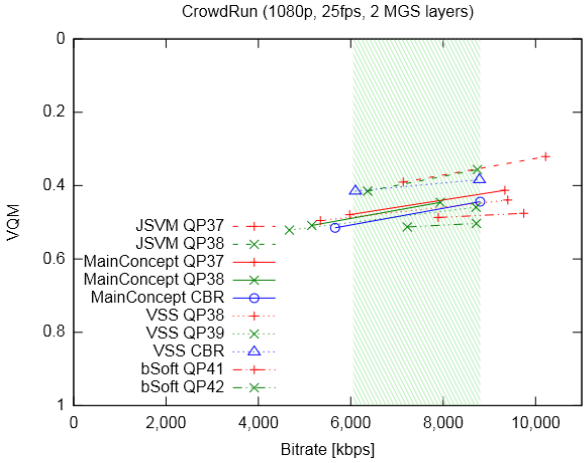
The bSoft encoder requires resolutions divisible by 16, which is the reason for the slightly cropped vertical resolutions in this test. For this test, we aimed for bitrates conforming to the recommendations of Table I. The PSNR results for *PedestrianArea* are shown in Fig. 4. Note that Fig. 4 (a) shows extraction points for resolution 960x528, while Fig. 4 (b) shows extraction points for resolution 1920x1056. Note that the line labeled *spatial scalability* ranges over both resolutions.

There is a small overhead at the lower resolution for enabling spatial scalability. Since the layers of the higher resolution depend on the lower resolution ones, the RD performance at 1920x1056 is worse. To achieve the same quality, the single-resolution bitstreams need around 18% to 26% less bitrate. Conversely, a single-resolution bitstream of the same bitrate as the spatial scalability bitstream achieves roughly 1-1.5 dB higher PSNR at 1920x1056. It requires almost the same disk space to store the spatial scalability bitstream or two separate bitstreams for the two respective resolutions with the same quality.

The bitstream with spatial scalability loses slightly less



(a)



(b)

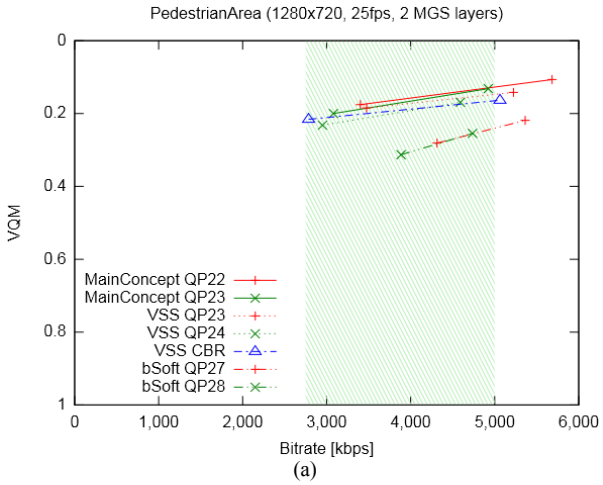
Figure 2. VQM results of rate control modes for (a) *PedestrianArea* and (b) *CrowdRun* sequences.

quality towards lower layers (indicated by the lower slopes in Fig. 4), resulting in smoother in-network adaptation. However, we consider the bitrate overhead to be a more relevant factor in favor of using separate SVC bitstreams for each resolution.

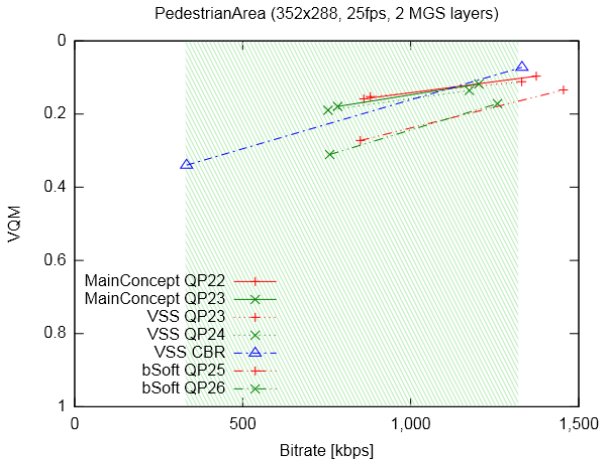
C. Number of MGS Layers

The following test investigates the impact of the number of SVC layers in MGS mode on the RD performance. Intuitively, higher numbers of layers come with some bitrate penalties. We tested the JSVM, MainConcept, and bSoft encoders with the following configuration: the QP of the highest layer was set to 28. PSNR results are shown in Fig. 5. Results for the VSS encoder are similar to the results for the MainConcept encoder but are not included in order not to overload the figure.

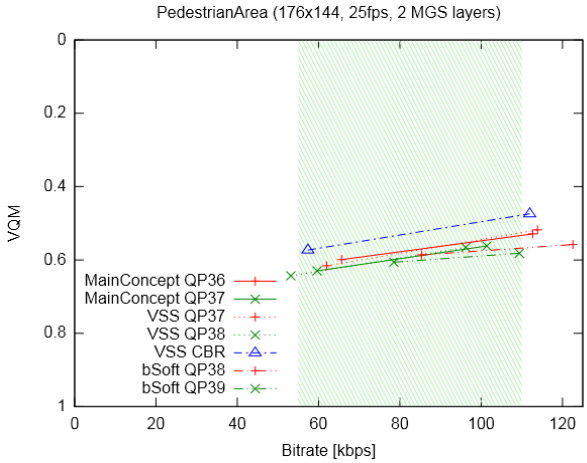
The JSVM and MainConcept encoders exhibit rather constant decrease in RD performance for higher number of layers. For the bSoft encoder, bitstreams with 2 and 3 MGS layers (labeled *bSoft 2MGS* and *bSoft 3MGS* respectively) have almost the same RD performance, the base layer of *bSoft 2MGS* has even lower bitrate and PSNR than the base layer of *bSoft 3MGS*.



(a)



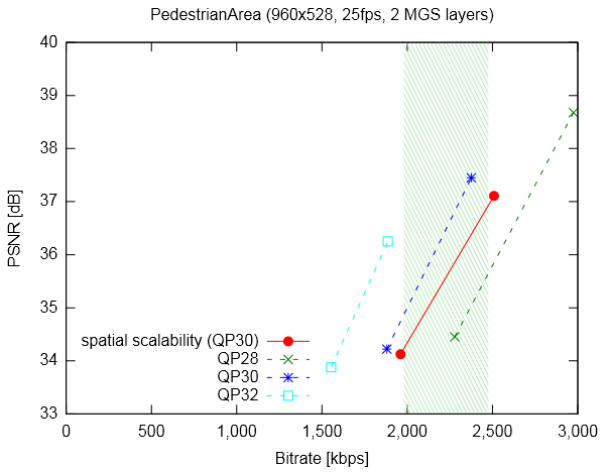
(b)



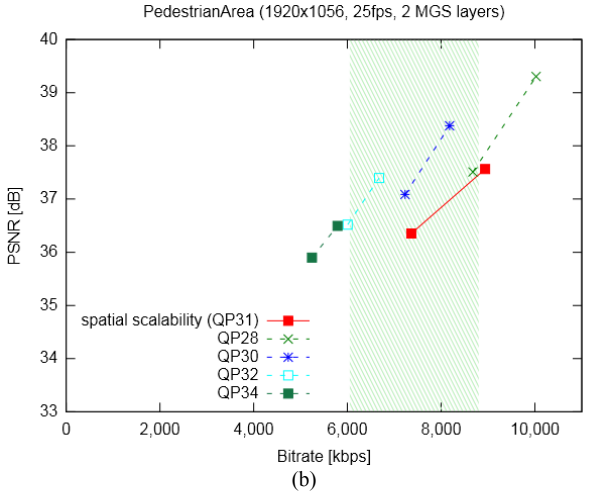
(c)

Figure 3. VQM results of rate control modes at (a) 1280x720, (b) 352x288, and (c) 176x144 resolutions.

The PSNR results of the highest layers remain relatively static across the number of MGS layers for all encoders (although they slightly decrease for MainConcept). Instead, encoders allocate less quality to the base layers for each additional MGS layer due to the applied configuration.



(a)



(b)

Figure 4. PSNR results of spatial scalability for the bSoft encoder. The line labeled spatial scalability represents a single bitstream ranging over both resolutions (a) 960x528 and (b) 1920x1056.

On average across all sequences, the JSVM encoder requires around 11.7% more bitrate for adding one MGS layer, the MainConcept encoder around 15.2% more bitrate, the VSS encoder around 19.7%, and the bSoft encoders only around 8.2% more bitrate. The bitrate penalty for additional layers generally decreases with the number of MGS layers used as starting point. The overhead for the JSVM roughly confirms the findings of previous studies on lower resolutions [2][6], overheads for the MainConcept and VSS encoder are a bit higher than expected.

D. *Requantization of MGS Layers*

In this test, the encoding performance of SVC encoders with four MGS layers and varying dQP between those layers was evaluated. For example, a dQP of 2 denotes QPs of MGS layers (from highest to lowest layer) of 28, 30, 32, and 34. The VQM results for *PedestrianArea* are shown in Fig. 6. Note that in contrast to other encoders, which use requantization for MGS layers, the bSoft encoder distributes transform coefficients automatically across layers, eliminating the need for different dQP encodings in this test.

It can be observed that a dQP of 2 is sufficient for

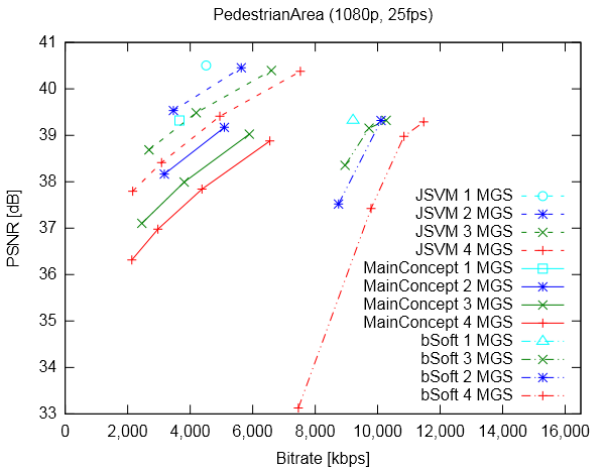


Figure 5. PSNR results for varying number of MGS layers.

serving a decent range of bitrates at 4 layers, while having the best RD performance. Higher dQP values cause such a strong quantization of the base layer that VQM results drop even below poor quality.

We compared PSNR vs. VQM for *PedestrianArea* in Fig. 7. The plot clearly shows that especially the lower layers of the bSoft encoder yields better VQM results than other encoders at the same PSNR. For other encoders, the plots show a strong (but not linear) correlation between PSNR and VQM for the respective sequence. Note however, that this correlation is content-dependent.

Mean encoding durations per frame are as follows: 46,990 ms/frame for the JSVM, 226 ms/frame for MainConcept, 251 ms/frame for VSS, and 1,916 ms/frame for bSoft. MainConcept and VSS are two orders of magnitude faster than the JSVM. The bSoft encoder is still one order of magnitude faster than the JSVM. For all encoders, encoding speeds are slightly slower at lower dQP values. Also, the *Dinner* sequence yields shortest encoding durations across all encoders (probably because it is a synthetic scene), followed by *PedestrianArea*.

As industrial SVC encoders are optimized for encoding speed, they typically sacrifice some RD performance, e.g., by using fast block search algorithms for motion estimation. On the other hand, the JSVM accepts high computational complexity throughout the video coding tool chain to ensure high RD performance.

Based on our findings, we conclude that, out of the industrial encoders, the MainConcept encoder is better suited for good RD performance at the highest layer, while the VSS encoder yields a more stable RD performance across layers. For the bSoft encoder, the bitrate should be considered; on the other hand, it performs better for more complex sequences and its VQM results indicate higher RD performance than the PSNR results do.

V. CONCLUSIONS

In this paper, we have investigated encoding guidelines of dominant industry solutions for AVC-based media streaming and devised SVC encoding guidelines therefrom. We have validated these guidelines together with further

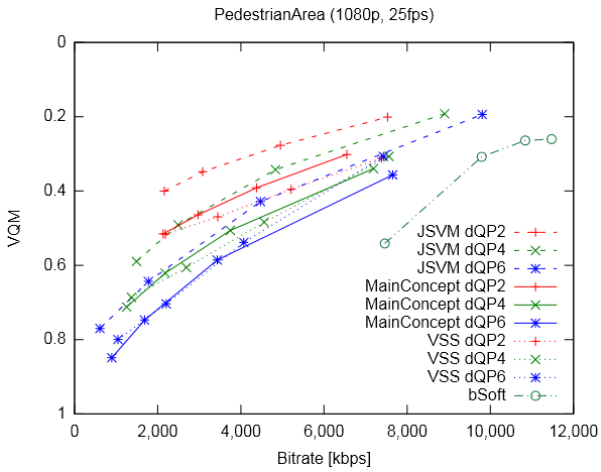


Figure 6: VQM results for varying dQP between MGS layers.

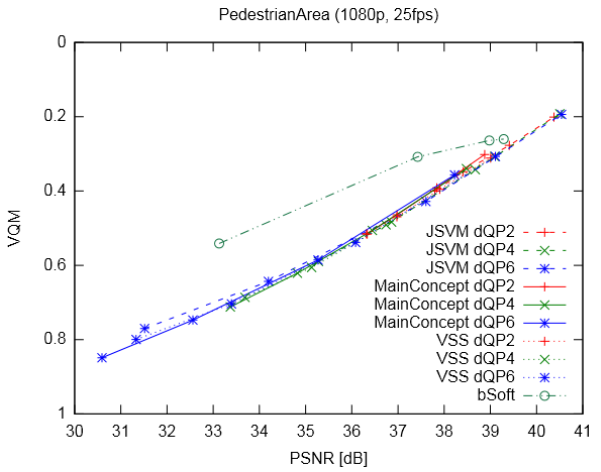


Figure 7. Correlation between PSNR and VQM for varying dQP of MGS layers.

evaluations of encoding configurations of high-definition video content relevant for adaptive media streaming in content-aware networks. Furthermore, we have tested and highlighted characteristics of various encoders.

Our evaluations show that CBR as well as fixed QP rate control modes yield solid quality for the devised bitrate suggestions for all resolutions. Our findings also indicate that it is more suited for media streaming to encode one SVC stream per resolution rather than a single stream comprising all resolutions. For several encoders, the number of SVC layers at 1080p resolution induces higher bitrate overheads than anticipated. We also found that some encoders yield better RD performance in VQM than PSNR results.

Our future work will include evaluations of further SVC encoding configurations directed at realistic streaming scenarios, the integration with adaptive media streaming systems, as well as the elaboration of the presented coding guidelines towards HTTP-based streaming.

ACKNOWLEDGMENT

This work was supported in part by the EC in the context of the ALICANTE project (FP7-ICT-248652). The authors thank MainConcept GmbH, Vanguard Software Solutions Inc., and bSoft Ltd. for providing their encoders.

REFERENCES

[1] MPEG output document N12957, "Joint Call for Proposals on Scalable Video Coding Extensions of High Efficiency Video Coding (HEVC)," Stockholm, Sweden, July 2012. URL: http://mpeg.chiariglione.org/sites/default/files/files/standards/parts/docs/w12957-v2-w12957.zip. Accessed February 1, 2013.

[2] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the Scalable Video Coding Extension of the H. 264/AVC Standard." *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 17, no. 9, pp. 1103–1120, Sep. 2007.

[3] Joint Video Team (JVT), "Joint Scalable Video Model (JSVM)", Version 9.19.15, 2011.

[4] M. Grafl, C. Timmerer, H. Hellwagner, D. Negru, E. Borcoci, D. Renzi, A.-L. Mevel, and A. Chernilov, "Scalable Video Coding in Content-Aware Networks: Research Challenges and Open Issues," in *Trustworthy Internet*, L. Salgarelli, G. Bianchi, and N. Blefari-Melazzi, Eds. Milano: Springer Milan, pp. 349–358, 2011.

[5] R. Kuschnig, I. Kofler, M. Ransburg, and H. Hellwagner, "Design options and comparison of in-network H. 264/SVC adaptation," *J. Vis. Commun. Image R.*, vol. 19, no. 8, pp. 529–542, Dec. 2008.

[6] M. Wien, H. Schwarz, and T. Oelbaum, "Performance Analysis of SVC," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 17, no. 9, pp. 1194–1203, Sep. 2007.

[7] T. Oelbaum, H. Schwarz, M. Wien, and T. Wiegand, "Subjective performance evaluation of the SVC extension of H.264/AVC," in *ICIP 2008*, pp. 2772 –2775, Oct. 2008.

[8] F. Niedermeier, M. Niedermeier, and H. Kosch, "Quality Assessment of the MPEG-4 Scalable Video CODEC," in *Image Analysis and Processing – ICIAP 2009*, vol. 5716, P. Foggia, C. Sansone, and M. Vento, Eds. Springer, pp. 297–306, 2009.

[9] J.-S. Lee, F. De Simone, and T. Ebrahimi, "Subjective quality assessment of scalable video coding: A survey," in *QoMEX 2011*, pp. 25 –30, 2011.

[10] G. Nur, H. K. Arachchi, S. Dogan, and A. M. Kondoz, "Advanced Adaptation Techniques for Improved Video Perception," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 22, no. 2, pp. 225 –240, Feb. 2012.

[11] G. Nur, H. K. Arachchi, S. Dogan, and A. M. Kondoz, "Seamless video access for mobile devices by content-aware utility-based adaptation," *Multimedia Tools and Applications*, pp. 1–31, May 2012.

[12] ITU-R Rec. BT.1683, "Objective perceptual video quality measurement techniques for standard definition digital broadcast television in the presence of a full reference," 2004.

[13] M. H. Pinson and S. Wolf, "A new standardized method for objectively measuring video quality," *IEEE Transactions on Broadcasting,* vol. 50, no. 3, pp. 312 – 322, Sep. 2004.

[14] A. Kapoor, "Dynamic streaming on demand with Flash Media Server 3.5 | Adobe Developer Connection", blog entry, URL: "http://www.adobe.com/devnet/flashmediaserver/articles/dynstream_on_demand.html", January 12, 2009. Accessed February 1, 2013.

[15] M. Levkov, "Video encoding and transcoding recommendations for HTTP Dynamic Streaming on the Adobe® Flash® Platform", White Paper, Adobe Systems Inc., Oct. 2010.

[16] J. Ozer, "Adaptive Streaming in the Field," *Streaming Media Magazine*, vol. Dec. 2010/Jan. 2011.

[17] J. Ozer, "Encoding for Adaptive Streaming," presented at *Streaming Media West 2011*, Los Angeles, CA, USA, Nov. 2011.

[18] Cisco, "Cisco Visual Networking Index: Forecast and Methodology, 2011-2016," White Paper, Cisco, May 30, 2012.

[19] ITU-T Rec. P.910, "Subjective video quality assessment methods for multimedia applications," 2008.