H.264/AVC and SVC Video Trace Library:*
A Quick Reference Guide
`http://trace.eas.asu.edu`

Martin Reisslein, Lina Karam, and Patrick Seeling†

May 2009

1 What are Video Traces and How are They Useful for Networking Research?

There are three ways to characterize encoded video for the purpose of networking research: (*i*) *video traffic model*, (*ii*) *video bit stream*, and (*iii*) *video traffic trace*. Video traffic models strive to capture the essential properties of the real traffic in parsimonious mathematical models. Video traffic models are typically developed based on the statistical properties of samples of the real traffic, or, in many cases, video traces of the real traffic. Video traces are therefore typically a prerequisite for model development. Video traffic models for H.264/AVC and H.264 SVC are largely an open research area. The *bit stream* contains the complete video information and allows for networking experiments where the quality of the video—after suffering losses in the network—is evaluated (a capability that can be emulated with offset distortion video traces [6]. Key limitations of the bit streams are that they require expertise in video coding, are very large in size, and are usually proprietary and/or protected by copyright.

Video traces are an attractive alternative to traffic models and bit streams in that they represent the traffic and quality of the videos. While the bit streams contain the actual bits carrying the video information, the traces are simple files that contain the number of bits needed to encode each video frame and the quality level (e.g., in PSNR) of the encoding [7]. Thus, there are no copyright issues with traces and networking researchers do not need to acquire detailed video coding expertise nor video coding equipment to conduct video networking research.

2 Why Should I Use H.264/AVC and SVC Traces?

The H.264/AVC and H.264 SVC [5] codecs achieve substantially more efficient video encoding through a number of advanced coding techniques, such as multiple reference frames and generalized B-pictures that exploit redundancies over longer time horizons as well as Intra picture prediction of macroblocks that exploits spatial redundancies in a given video frame [4]. With these advanced coding techniques, the H.264/AVC and H.264 SVC codecs generate traffic that is very different

*This material is based upon work supported by the National Science Foundation under Grant No. 0750927. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

†M. Reisslein and L. Karam are with the Dept. of Electrical Eng. at Arizona State Univ., Tempe, {`reisslein, karam`}`@asu.edu`. P. Seeling is with the University of Wisconsin—Stevens Point, `Patrick.Seeling@uwsp.edu`.

from the older MPEG codecs: with the new compression tools, the H.264/AVC and H.264 SVC codecs typically compress video down to roughly half the average bit rate of the MPEG-4 Part 2 codec [1, 4, 5]. On the other hand, the H.264/AVC and H.264 SVC codecs produce a much higher traffic variability. The coefficient of variation (standard deviation normalized by mean) of the frame sizes reaches levels above 2.5, whereas it was typically in the range from 0.7 to 1.4 with MPEG-4 Part 2 [7]. These new traffic characteristics demand networking protocols that accommodate high traffic variations, much higher variations than have been considered for MPEG-4 video networking.

Aside from the vastly improved compression efficiency, the widespread use of H.264/AVC and H.264 SVC for video compression for networked video applications is promoted by the widespread adoption of these encoding standards in DVB, ATSC, 3GPP, 3GPP2, MediaFLO, DMB, DVD Forum (HD-DVD), and Blu-Ray Disc Association (BD-ROM).

3 How to Use H.264/AVC and SVC Video Traces?

H.264/AVC is designed for single-layer (non-scalable) video compression. Also, by default, H.264/AVC does not employ hierarchical B frames; thus, essentially preserving the frame dependency structures that were present in MPEG-4 Part 2. Hence, video traces for H.264/AVC have the same structure and can be used in the same manner as MPEG-4 Part 2 traces. The general instructions on how to employ video traces in networking research in [7] fully apply.

H.264 SVC [5] is primarily designed for scalable video encoding, but can also be used for encoding video into a single layer. The resulting single-layer encoding has typically temporal scalability built in, that is, a single-layer SVC stream can be played out with different frame rates (in frames displayed per second). An important advance that is used by default in H.264 SVC is the concept of hierarchical B frames, that is the concept of B frames being used as a prediction reference for other B frames. These B frame dependencies introduce novel timing constraints for the network transport of H.264 SVC video, which are explained in [2].

4 Which H.264/AVC and SVC Video Traces are Available?

As of May 2009, three sections of the video trace library at `http://trace.eas.asu.edu` provide H.264/AVC and H.264 SVC video traces: The section on H.264 Main Profile (AVC) section initiated in March 2007 provides traces for five CIF (352x288 pixel) format videos as well as four High Definition (HD) videos. The set of five CIF format videos includes *NBC News*, *Silence of the Lambs*, *Star Wars IV*, *Sony Demo*, and *Tokyo Olympics* and provides a good sample of videos of different genres. Each of the CIF videos is encoded using four different GoP structures. H.264/AVC achieves typically the best compression performance for the GoP pattern with three B frames between I/P frames [1], denoted by G16-B3, and this pattern is therefore recommended. For each GOP pattern, each video is encoded with a range of quantization parameter settings to cover a range of PSNR video quality levels and corresponding mean bit rates.

The library section on H.264 Scalable (SVC) encodings initiated in October 2008 provides traces of H.264 SVC encodings for the set of five representative CIF format videos. Specifically, this section provides H.264 SVC encodings into a single layer [1], as well as into temporal and spatial scalable layers [3]. The encodings for both the sections March 2007 H.264/AVC and October 2008 H.264 SVC employ overall encoding settings that achieve high rate-distortion (RD) efficiency at the expense of a relatively high computational effort for encoding as would be typical for offline encodings for video streaming in the Internet. For H.264 SVC, the G16-B15 GoP pattern achieves generally the highest compression efficiency [1] and is therefore recommended. Also, both library sections follow the conventional library structure based on static traces files, which has been used for the MPEG-4 Part 2 trace library [7].

The library section on H.264 Scalable (SVC) Single Layer encodings initiated in April 2009 provides H.264 SVC single layer encodings for eleven CIF format videos, which include the set of five representative videos. These videos are encoded with moderate RD efficiency at a moderate computational effort as would be typical for low to moderate complexity video streaming with the relatively small (by the standards of 2009) CIF resolution. This new library section follows a novel library structure that dynamically extracts the desired video trace format from an underlying database. This dynamic video trace library structure is inspired by Web 2.0 applications and we therefore refer to it as "Video Traces 2". The key motivation for this novel Video Traces 2 library structure is to flexibly accommodate a wide range of levels of detail in the video traces for the different encoding types.

We briefly explain some key features of the Video Traces 2 structure for the H.264 SVC single layer traces as follows. Each video has been encoded for a range of GoP structures and in turn there are encodings for a range of sets of quantization scales for each GoP structure. For a given video encoded with a given GoP structure and a given set of quantization scales, the library provides the single-layer video at the full frame rate, which is typically 25 or 30 frames/second. In addition, the lower frame rates that can be extracted using the temporal scalability feature of H.264 SVC are provided. The lowest frame rate encoding is denoted by `(0,0,0)` and the middle index in this triple counts up as the frame rate increases. The number $k+1$ of encodings with different frame rates is related to the number β of B frames in the GoP as $\beta = 2^k - 1$. Thus, with $\beta = 15$ B frames in the GoP, i.e., with the G16-B15 pattern, $k + 1 = 5$ different frame rates are possible ($k = 4$ different frame rates are possible due to the $k = 4$ levels in the B frame hierarchy, and one additional frame rate is achieved by dropping all B frames and displaying only I and P frames). For this pattern, the $(0, 4, 0)$ trace provides the full frame rate (30 frames/second), while the $(0, 0, 0)$ trace provides the lowest frame rate of $30/2^{k+1} = 1.875$ frames/second. For each encoding at a given frame rate, the library provides a preview of the of the available traces. The preview contains basic statistical characteristics of the encoding, which are explained in [1, 7], followed by the trace data for the first GoP. A trace extraction window is provided that allows the user to specify the desired trace format. Aside from the frame traces, the library provided the offset distortion traces [6] for the full frame rate encoding.

5 How Do I Cite the Video Trace Library?

When using traces from the video trace library in your research, please cite [1, 7], i.e.,

`P. Seeling, M. Reisslein, and B. Kulapala. Network performance evaluation with frame size and quality traces of single-layer and two-layer video: A tutorial.` *`IEEE Communications Surveys and Tutorials`*`, 6(3):58-78, Third Quarter 2004.`

and

`G. Van der Auwera, P.T. David. and M. Reisslein. Traffic and Quality Characterization of Single-Layer Video Streams Encoded with H.264/MPEG-4 Advanced Video Coding Standard and Scalable Video Coding Extension,` *`IEEE Transactions on Broadcasting`*`, 54(3):698-718, September 2008.`

6 Questions? Comments? Feedback?

Please e-mail any questions, comments, and feedback to `reisslein@asu.edu`. We look forward to hearing from you and addressing all your video trace needs.

7 Outlook

As of May 2009, the video trace library provides a basic set of five representative CIF format videos encoded with H.264/AVC and H.264 SVC as well as a larger collection of eleven videos encoded into a single layer with H.264 SVC. The next steps in the video trace library expansion are higher format videos up to High Definition encoded with H.264 SVC into a single layer as well as encodings with the new scalability features of H.264 SVC. Of particular interest is the new Medium Grain Scalability feature of H.264 SVC, which provides SNR-quality scalability with only a 10% overhead compared to corresponding single-layer encodings. Throughout these expansions, the Video Traces 2 structure will form the basis for the dissemination of the vast amounts of trace data. Comments on the usability of the new Video Traces 2 structure are therefore especially appreciated.

References

[1] G. Van der Auwera, P.T. David, and M. Reisslein. Traffic and quality characterization of single-layer video streams encoded with H.264/MPEG-4 advanced video coding standard and scalable video coding extension. *IEEE Transactions on Broadcasting*, 54(3):698–718, September 2008.

[2] G. Van der Auwera and M. Reisslein. Implications of smoothing on statistical multiplexing of H.264/AVC and SVC video streams. *IEEE Transactions on Broadcasting*, 55(3), September 2009.

[3] G. Van der Auwera, P.T. David, L. Karam, and M. Reisslein. Traffic and quality characterization of scalable video streams encoded with the H.264/AVC Scalable Video Coding extension. *Advances in Multimedia*, Article ID 164027, 2008.

[4] D. Marpe, T. Wiegand, and G. Sullivan. The H.264/MPEG-4 advanced video coding standard and its applications. *IEEE Communications Magazine*, 44(8):134–143, August 2006.

[5] 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*, 17(9):1103–1120, September 2007.

[6] P. Seeling, M. Reisslein, and F.H.P. Fitzek. Offset trace-based video quality evaluation after network transport. *Journal of Multimedia*, 1(2):1–12, May 2006.

[7] P. Seeling, M. Reisslein, and B. Kulapala. Network performance evaluation with frame size and quality traces of single-layer and two-layer video: A tutorial. *IEEE Communications Surveys and Tutorials*, 6(3):58–78, Third Quarter 2004.