SVC Paper Review

“Overview of the Sacalable Video Coding Extension of the H.264/AVC Standard” was published by Heiko Schwarz, Detlev Marpe,m and Thomas Wiegand in the IEEE Transactions on Circuts and Systems for Video Technology, Volume 17, Number 9, September 2007.

The paper opens explaining that older digital video transmission and storage systems are based on H.222 / MPEG-2 for broadcast over channels like satellite or cable and stored on DVDs. Modern video transmission uses IP and real-time services, like streaming, and storage in computer file formats like mp4 and 3gp.

The paper reviews the standardized extension of H.264/AVC, Scalable Video Coding (SVC). It explains SVC as the ability to decode a high quality data stream into a lower quality by selecting a subset of bits.

Examples of the different applications SVC are looked into. Clients with restrictions (display resolution, processing power, battery power), multi-cast servers, surveillance applications. The paper speaks of clients with less battery power and less processing power, but does not using SVC increase the decoding complexity compared to single-layer decoding? It would only be beneficial to clients where the the increase in complexity outweighs the benefits of using SVC elsewhere in the system.

The paper next explains the history of SVC and then about the H.264 standard. Hybrid video coding, as found in H.264 and all common standards, is based on motion-compensated temporal differential pulse code modulation (DPCM) together with spatial de-correlating transformations. Explains history of the standardization of the SVC extension, where MPEG issued a call for proposals for efficient SVC technology in October 2003.

Section IV is a basic overview of the H.264 standard, and was included to “keep the paper self-contained” and only went over “key features that are relevant for understanding the concepts of extending H.264/AVC towards SVC”. However, it would have been beneficial to include notes on where SVC will come in later when describing the parts of H.264.

The next section of the paper, Section V, is the bulk of the paper, describing the extending of H.264 towards an SVC standard. The topics covered under temporal scalability include: hierarchical prediction structures and coding efficiency of hierarchal prediction structures. The topics covered under spacial scalability include inter-layer prediction, (inter-layer motion prediction, inter-layer residual prediction, inter-layer intra-prediction), generalized spacial scalability, complexity considerations, coding efficiency, and encoder control. The topics covered under quality scalability include controlling drift in quality scalable coding, encoder control, bit stream extraction, coding efficiency, and SVC-to-H.264/AVC Rewriting.

Supporting temporal scalability with a reasonable number of temporal layers, no changes to the design of H.264 were required. This is very important, having the changes of SVC not interfere with the H.264 standard. In the first experiment of the paper the coding efficiency for hierarchal B-pictures was analyzed. Figure 2 shows the results, showing that by enlarging the GOP size up to about 1s, with gains over more than 1dB.

Figure 4 shows a multilayer structure with inter-layer predictions that SVC H.264 uses. The inter-layer predictions is a common video coding method, and using it with the multilayer data improves efficiency. A more complex figure of the multilayer structure with inter-layer predictions is shown in Figure 8, showing different methods.

Figure 6 shows the efficiency of inter-layer prediction in SVC for different sequences and prediction structures. The figure shows that multiple loop decoding can further improve the coding efficiency, but the gain is often minor and not worth the significant increase in decoder complexity. Two of the three test videos shows that a GOP size of 16 pictures is very close in results to using a single-layer coding scheme.

Finally, before the conclusion, the paper goes over the high-level design of the SVC standard, primarily focusing on the SVC bit stream. Figure 12 shows a high-level overview of the SVC encoding structure. This could have been included at the beginning of the previous section as an outline for the detailed steps.

The higher level topics covered include the dependent laying structure shown if Figure 12, bit stream manipulation, switching between different dependency layers.

The conclusion sums up the paper with the differences that the new SVC extension provides than older scalable profiles: hierarchical prediction structures, interlayer prediction of motion, key pictures fro controlling drift, and single motion compensation. The paper then lists the discovered differences shown by the experiments.

Overall, the paper is detail packed, almost to the point where it became difficult to read. The largest section, Section V – Basic Concepts for Extending H.264/AVC Towards an SVC Standard, is almost too packed. The section's three sub-headers, Temporal Scalability, Spatial Scalability and Quality Scalability, take up nine pages, with eight of them coming from the last two sub-headers. While the length is not the problem, the readability is.