Networking for Multimedia
Tech Topic #02: Multimedia Encoding Techniques

D. Le Gall, “MPEG: A Video Compression Standard for Multimedia Applications,” Communications of the ACM, vol. 34, num. 4, April 1991.

Half tutorial, half semi-technical introduction to a contemporary new compression and encoding standard, this journal article foregos most traditional paper organization and format guidelines in favor of attempting to provide a readable overview. As a result a review will necessarily consist mostly of commentary on the presentation of the material and comparisons to the present day.

The first two sections provide some Gauguin-like informational structure: where MPEG came from, what it is, where it is going. The description of related standards bodies is dated when considering younger readers whose main exposure to CCITT is their CRC polynomial. But it was probably helpful at the time. What isn't dated, in fact is remarkably relevant, is the relationship between MPEG and JPEG, for two reasons: that JPEG is still one of the most widely used image compression formats, and that the spatial redundancy techniques leveraged in all the subsequent generations of MPEG are just as closely tied to JPEG as the original MPEG standard was.

The section describing the standardization logistics (and table 1) seemed included more than a bit to appease other participating parties, considering the author was, after all, from a particular company who had submitted a proposal. But the reviewer supposes that everyone deserves a pat on the back once in a while.

The laundry list of requirements which MPEG aims to satisfy stands up to the test of time amazingly well—one would expect the comparable journal article on H.264/AVC to replicate this section point for point. The “VCR-like” capabilities ring especially strongly given the prevalence of “Tivo-like” functionality today. It is somewhat surprising that although ISDN and LAN communication channels are mentioned by name, DS1s are not, despite their obvious relevance (at 1.544 Mbps) for the choice of 1.5 Mbps as a high-end bit-rate goal. Without this one might question why 1.5 Mbps was chosen over the going 10BASE-T Ethernet rate of 10 Mbps.

The meat of the article is in its description of the compression algorithm, which in general does a very nice job toeing the line between abstract summaries and dense mathematical formulas. The discussion of MPEG's use of temporal redundancy, in particular, is succinct and effective in its explanation of what is a very technical topic. The author avoids completely any formulas at all until finally near the end of the section when discussing motion estimation!

The spatial redundancy description could have examined more the relationship between MPEG and JPEG in this area. Also the consistency of content could potentially have been improved. As an example, the reasons for choosing the DCT are discussed and even bulleted, but the rationale behind the use of zig-zag scan is not touched upon at all.

The genericity (which the reviewer does not believe is a word) goal repeated again in the layered structure section forecasts just how extensible later generations of the standard will become.

It was noted that the “standard and quality” section is relatively short, and quality is only mentioned in the final 3 or 4 paragraphs along with bit rate and resolution. But this leads nicely to another paper...

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, num. 9, September 2007.

Again primarily a description of standardization efforts, this article incorporates more aspects of a “standard” conference or journal paper. When compared to the above MPEG overview, the tone and content of this work make significantly more assumptions about audience knowledge and familiarity with the subject. This is in part due to the fact that 20 years of research have occurred since the original MPEG standard was first presented, and partly because the publication venue is more specific. It says something in itself that in 2007 there is an entire IEEE journal devoted to video technology systems.

The applications section serves as a nice motivator for why the reader ought to care about scalable video encodings. Unfortunately, with the possible exceptions of this section and the one that follows on SVC history, the paper as a whole reads like a set of notes from MPEG members' musings on scalable video put into sentences. In general sections are organized seemingly effectively, however there is little in the way of flow from section to section and the article overall lacks the clarity of presentation present in the previously discussed overview of MPEG-1.

After a brief background on H.264/AVC, most of the pages are spent exploring 3 basic concepts for SVC: temporal, spatial and quality scalability. Each concept is presented in sufficient detail to make it clear that the authors' know well what they are talking about, but common linkage between individual sub-sections (and sub-sub-sections) is often hard to come by. In other words, they might have spent more effort connecting the dots.

Experimentation appears suddenly and without warning. The parameters which are used for comparison in all evaluations are never justified, although this may be an artifact of publishing in such a specific arena—it may just happen to be the case that peak SNR vs. bit rate are the de facto standards by which objective video quality is judged.

The section on encoder control for spatial scalability was especially interesting given that encoding techniques are not specified by H.264/AVC. This suggestion and corresponding evaluation was one of the few places the authors directly referenced their other work outside of the standard.

In general the work highlights a glaring truth about the MPEG standards: they have become so flexible so as to hardly be considered standards. The sheer number of profiles defined practically guarantees that no decoder will be able to support even most of them. This fact takes a certain amount of significance away from the present work, despite the fact that the SVC extensions throughout are shown to be improvements over everything that has come before.

Given that a major goal of the paper is to describe how these extensions solve the existing SVC problem of loss of coding efficiency, the work to this extent is a major success. The use of inter-layer prediction techniques are quite clever and are shown to relate directly to a gain in coding efficiency. Of course, as with any standard, the ultimate measure of success will be take-rate, but given the head-start that basing off of MPEG-4 part 10 has provided, anything is possible.