

Topic Review: Multimedia Encoding Techniques

Review: MPEG: A Video Compression Standard for Multimedia Applications

“MPEG: A Video Compression Standard for Multimedia Applications” was published in April, 1991 in the journal, Communications of the ACM (CACM). CACM is a broadly focused journal covering nearly all topics in computer science; as such, articles typically present higher-level overviews of significant developments that are investigated in greater depth in other publications. This article fits this model. It provides a relatively high-level summary of the history, development, core requirements, and technical details of the MPEG video compression standard. The goals of the article itself are not explicitly stated; however, ostensibly the article aims to provide a synopsis of the MPEG working group’s recent (at the time) draft standard.

By any account, the MPEG standard has been highly successful, evidenced by its near ubiquity throughout the multimedia landscape – even almost twenty years after its publication. Didier Le Gall, the author of this article and chair of the MPEG video group, clearly had this type of impact in mind for the MPEG standard. In its opening paragraphs, the article sites the industry impact of the Group 3 facsimile compression algorithm as a model for the type of effect the MPEG standard could have. Le Gall follows with a description of a three phased methodology used by the MPEG group to develop the standard. He describes the following three phases: Requirements, Competition, and Convergence. Each of these phases reveals an awareness of the potential impact the standard could have and a focus on ensuring it would remain relevant to end-users and achieve broad adoption. I believe this was instrumental to its success.

The requirements phase focused on the practical applications and use cases for MPEG video, and derived feature requirements from these. The following broad requirements were developed during this phase:

1. **Random Access.** Playback must be able to be initiated at any point within the MPEG bit stream.
2. **Fast Forward/Reverse Searches.** It must be possible to deliver a fast forward or fast reverse “effect” in which selected frames are displayed during a sped up scan over the compressed bit stream.
3. **Reverse Playback.** Full quality playback of video in reverse
4. **Audio-Visual Synchronization.** Mechanisms must be provided for the accurate synchronization of the audio and video data streams.
5. **Robustness to Errors.** The decoding algorithm must be tolerant of errors in the bit stream and handle them gracefully.
6. **Coding/Decoding Delay.** The algorithm should perform well over a range of acceptable delays

7. **Editability.** The compressed video should be editable with reasonable accuracy (e.g. 1/5 second)
8. **Format Flexibility.** The video must adapt to various raster sizes and frame rates.
9. **Cost Tradeoffs.** The algorithms must be implementable in a relatively small number of processing chips and using available technology of the time.

The next phase of development was the competition phase, in which 14 different proposals were submitted by various companies. These proposals were analyzed, tested, and evaluated for contribution to the final standard. This competitive approach was beneficial to the success of the standard in two ways: 1) it solicited and evaluated a large set of ideas, and elements from the best of these could be used to form the standard; and 2) it involved direct participation and investment from industry leaders, whose support would certainly be needed if the standard was to be broadly adopted. Involving these critical players at this developmental stage was a strategically smart decision.

Though this phase of the standard's development appears to have been well designed and executed, the author's treatment of it, I believe, is the weakest part of the article. Le Gall could have done a much better job establishing the advantages of the ultimate MPEG video standard over other competing solutions. The article does cite CCITT's H.261 recommendation as a significant contributor to the MPEG standard; however, it does not discuss the technical differences between H.261 and MPEG. More so, it would have been helpful if the article had given greater insight into the different proposals that were evaluated during competitive phase of the standard's development. Specifically, what alternative approaches were proposed and rejected, and what were the reasons? While detailed experimental analyses are out of the scope of an article of this type, some form quantitative evaluation of MPEG video against previous or alternative methods would have been helpful. It is also notable that the list of contributors (see Table 1) is devoid of any academic institutions, despite the author's claim that the standard "must bring together the best of academic and industrial research."

The author describes the convergence phase as essentially a synthesis of the best ideas produced from the competition phase. There are few details given regarding this phase or any information about the specific issues that were encountered. As such, there is little to comment on, other than this appears to be a weaker point in the article.

The next major portion of the article provides an general overview of the MPEG compression algorithm. The algorithm is described as consisting of the following two essential techniques:

1. Block-based motion compensation
2. Discrete Cosine Transform (DCT) based compression

Briefly, motion compensation exploits the high degree of temporal redundancy typical of video streams. MPEG uses predictive pictures (or P-frames) are derived using information from a previous frame, and introduces the concept of interpolated pictures (or backward predictive B-frames), which can make predictions using both the previous and future frames. A high degree of compression can be achieved using P-frames and B-frames (both forms of inter-frame compression); however, they require data from previous or future frames, and thus cannot be independently decompressed. This works against the

core requirements of random access and editability. As a solution, intra-frame DCT based compression is utilized to compress specific frames at regular and frequent intervals.

DCT based compression exploits spatial redundancy within a single frame. These frames, called Intraframes or I-frames, do not provide as high degree of compression as P-frames or B-frames; however, they can be decoded without reference to other frames, and serve essentially as anchors for random access into the compressed bit stream. MPEG's DCT based compression techniques borrow heavily from the techniques used by JPEG compression.

Overall, the author's treatment of the MPEG algorithm is reasonably thorough for an overview article. However, particularly with regard to the spatial compression algorithms, the article does assume quite a bit of knowledge and familiarity with existing techniques for the compression of visual data. Considering its broad audience, it would have been helpful if the author had provided additional references for those who are not as familiar with these technologies.

The last portion of the article briefly mentions what might be the most significant factor contributing to the success of the MPEG video standard. The standard specifies only a syntax for the encoded bit stream and the process for decoding it. It does not specify any specific decoder implementation or the mechanism for encoding the bit stream. This approach has several key advantages:

1. It reduces the complexity of the standard to only what is essential. It does not attempt to over specify and thus decreases the chances that some portion of the standard will prove infeasible or prohibitive in some way.
2. It gives implementers the a high degree of flexibility when applying the standard, opening up competition and allowing manufacturers to differentiate their products as they see fit. This makes it an attractive platform to develop solutions against because there is a large opportunity to innovate.
3. It is end-user focused. It ensures that decoders built according to the standard will continue to perform their essential function (decoding MPEG video) even as manufactures iterate on specific implementation details or encoding techniques change.

In all, this was a well written and well organized overview of the MPEG standard. It is somewhat lacking in its discussion of alternatives approaches and justification of this standard as the best solution. However, considering the relatively quick adoption, wide use, and undeniably significant contribution of the MPEG standard to the field of multimedia technology of the past two decades, perhaps such comparisons are mute and be overlooked.