Technical Paper

"It's Not Dead Yet!" - MPEG-2 Video Coding Efficiency Improvements

Learn about significant new improvements to MPEG-2 Video coding, how this enables more services to legacy set-top boxes and receivers, and about the revolutionary new TANDBERG EN8100 Advanced MPEG-2 Broadcast Encoder.

Matthew S. Goldman
TANDBERG Television, part of the Ericsson Group
Bedford, New Hampshire
www.tandbergtv.com

Originally published in the 2009 NAB Broadcast Engineering Conference Proceedings.

ABSTRACT

Broadcasters are required by federal regulation, and locked into in practice due to the millions of consumer digital televisions (DTVs) purchased, to use the MPEG-2 Video format for free over-the-air DTV service to the home. The 19.39Mbps bandwidth is already a constraint for those broadcasters that offer multiple DTV services in a single 6MHz channel. With the advent of new Mobile/Handheld services being introduced in 2009, the need for more efficient coding is paramount. Any available bandwidth in a 6MHz channel is at a premium; thus, improvements in MPEG-2 Video coding efficiency are extremely valuable.

Contrary to widespread belief, MPEG-2 Video has not yet reached its maximum performance. Improvements in integrated circuit speed and intelligence mean that the amount of information that can be processed in real-time is enabling new approaches to coding MPEG-2 Video that before were not possible. Examples include fully exhaustive motion searches and predictive processing, rate-distortion optimization (RDO) algorithms, and lessons learned from implementing MPEG-4 AVC coding retrospectively applied to MPEG-2 Video to improve the use of the compression algorithms. This paper explores new approaches to coding MPEG-2 Video efficiently. Wherever possible, simulation results have be used to demonstrate the improvements.

EFFICIENCY IMPROVEMENTS POSSIBLE

The MPEG-2 Video standard was ratified in 1994 and the selected set of algorithmic tools and the decisions made on the theoretical processing models were based on the state of integrated circuit technology of that time. Over the past decade, there has been an estimated 40 fold increase in integrated circuit processing power, yet the past several years have not produced improvements in MPEG-2 Video coding efficiency of more than approximately five percent. As such, the general industry belief is that further improvements to MPEG-2 Video coding efficiency is limited by the standard itself.

However, taking a fresh look at MPEG-2 Video, by combining the new integrated circuit processing power with knowledge obtained through the design of MPEG-4 AVC encoders, has shown that it is possible to increase the coding efficiency of MPEG-2 Video by more than 15%. Such a *Revolutionary Optimized* MPEG-2 Video Encoder is practical to implement today.

Three areas are investigated:

- Look-ahead encoding
- Two-stage motion estimation
- Pre-processing

Look-Ahead Encoding

Look-ahead encoding is the process of using a pre-encoder (first stage encoder) to analyze the incoming video to capture encoding metrics and then passing those metrics to a second/final stage encoder that uses those metrics to optimize coding efficiency better. It has been used in premium MPEG-2 Video encoders for years. Due to practical processing constraints, previous look-ahead implementations only were capable of providing a bit-rate demand calculation.

With recent advances in integrated circuit processing power, the look-ahead encoding concept now can be expanded to have multiple encodes in the look-ahead stage. As shown in Figure 1, the Revolutionary Optimized MPEG-2 Video Encoder uses multiple encoders in the look-ahead stage. Each look-ahead encoder is set to a different operating point and passes frame size information to a buffer size predictor. This, in turn, is fed to a rate control function.

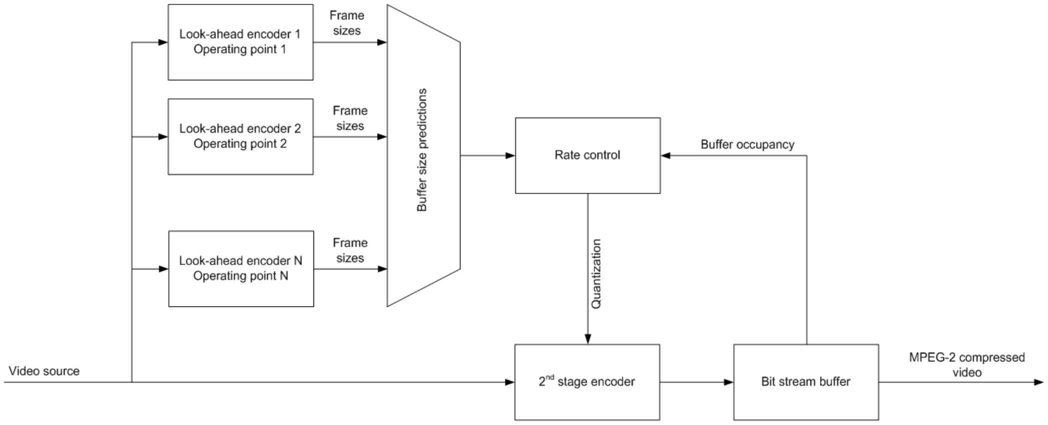

Figure 1. Multiple look-ahead encoders enable better predictions, which results in better rate control.

The look-ahead encoders are set to a common GOP (Group of Pictures) structure, which is determined during a pre-processing step. Since these encoders are analyzing the bit-stream to produce metrics for the second stage encoder their operation is simpler and need not be bound by bit-stream conformance restrictions imposed on the second stage encoder. For example, rate control is not necessary and a fixed quantization level (Qp) may be used. The goal is to determine what the bit-rate would be at a particular quantization step to achieve a particular picture "quality".

By using multiple operating points to produce more encoding metrics, better stability of the rate control is achieved, which results in better overall picture quality because bit allocation is planned better. The encoder is not forced into large swings in quantization levels in order to maintain proper buffer occupancy.

The Revolutionary Optimized MPEG-2 Video Encoder is able to make better predictions, which results in improved rate control and helps determine exactly what quantization is needed to achieve optimum picture quality for each frame.

Two-Stage Motion Estimation

Motion estimation is the process of determining motion vectors that describe the transformation of motion from one picture to another, typically adjacent pictures in a video sequence. MPEG-2 Video does not define how to perform motion estimation. This is a value-add feature that differentiates encoder implementations. Motion estimation is computationally intensive. Traditional MPEG-2 Video encoders perform a block-matching motion estimation over a pre-defined search range.

There are a variety of mechanisms employed such as hierarchical and local exhaustive [1, 2].

- Hierarchical: In order to achieve a large enough search range, a down-sampled, lower resolution picture is used to pinpoint the general area of the likely best match, and then an exhaustive search on a small region is performed.
- Local exhaustive: Based on research, a smaller range that is a subset of the full picture is chosen and then exhaustively search of each block within the range is performed to find the best match.

With advances in technology, a much more optimized motion estimation can be performed. The Revolutionary Optimized MPEG-2 Video Encoder uses the set of look-ahead encoders described in the previous section to give indication of motion. The set of motion and operating point data from the first stage are fed to the second stage encoder, but they are refined in time for a second stage motion estimation to be performed during the final encode stage. This is shown in Figure 2.

Pre-processor

With advances in technology, it is possible now to perform more comprehensive analysis of the source video - also known as pre-processing - in advance of the compression stage. Traditional MPEG-2 Video encoders typically do not do all of these because of processing limitations. Examples of useful picture analysis pre-processing functions:

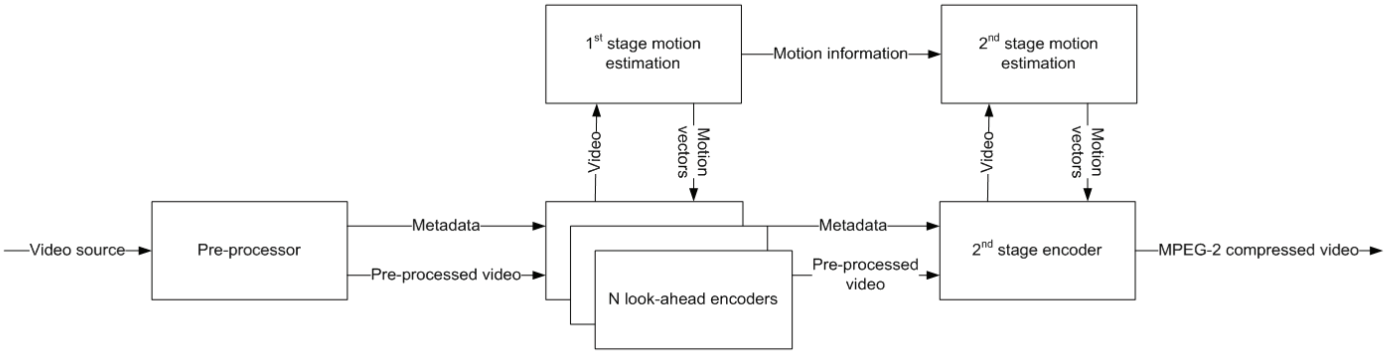

Figure 2. Two-Stage motion estimation

- Field-frame decision: Measure field dominance as well as field and frame picture activity to select between field and frame picture coding modes
- Scene-cut detection: Detect single-picture scene changes to prepare buffer allocation to handle the large picture data to be required for the start of the next scene
- Fade detection: Detect fade to/from black to use coding parameters more appropriate for this special effect
- Flash detection: Detect rapid chrominance changes and luminance saturation to use coding parameters more appropriate for this special effect and prevent picture break-up
- Adaptive GOP structures and GOP length: Vary the combination of P-pictures and B-pictures within a GOP to match the picture type to the content better

The Revolutionary Optimized MPEG-2 Video Encoder will be able to perform all of the above comprehensive pre-processing functions.

RATE DISTORTION OPTIMIZATION

Rate-distortion optimization (RDO) is the process by which the loss of video quality (the distortion) is compared against the data required to encode the video (the rate) and an optimal choice is made against a measure of constraint. While the concept has been understood since the early 1990s, the technique was never applied to MPEG-2 Video encoders because of its high processing demands. Interest in RDO has increased in recent years as an implementation method for dealing with the increased complexity of the MPEG-4 AVC algorithmic toolset within the cost-constrained processing power available [3].

Traditional real-time MPEG-2 Video encoders choose the result that yields the highest picture quality per macroblock; however, such a choice typically requires more bits to encode and may give comparatively little quality benefit. The "sum" of all these decisions may exceed the bit allocation possible for the overall picture and actually decrease the overall picture quality that would have occurred if some of the "best" choices for a macroblock were not selected. Up to recently, there has been insufficient processing resources available to traditional MPEG-2 Video encoders to choose tools and settings that optimize the overall visual quality. Instead, intelligent mode decisions have been part of the value-add that separates one vendor's encoder implementation from another.

Only recently, due to increases in processing power coupled with lower cost points for the same, has RDO been practical to implement for MPEG-2 Video. With RDO, the Revolutionary Optimized MPEG-2 Video Encoder is able to encode each macroblock using every encoding tool combination in parallel, then determine which method worked best in terms of bit-rate cost or by measuring the bit-rate combined with total visual distortion from the source video signal. In essence, the Revolutionary Optimized MPEG-2 Video Encoder produces the best visual quality by picking the best result among the set of possible encoding results that have been produced in parallel. This process is repeated for every macroblock of a picture.

There still is a practical trade-off between visual quality and bit-rate for each macroblock so the encoder implementation still needs to find the optimum balance. Processing power alone will not deliver the maximum potential gains.

RDO, in conjunction with look-ahead encoding and two-stage motion estimation enables the Revolutionary Optimized MPEG-2 Video Encoder to optimize bit-rate allocation far beyond that exhibited by today's best MPEG-2 Video encoders, with the end result being significant bit-rate reduction for the same visual quality.

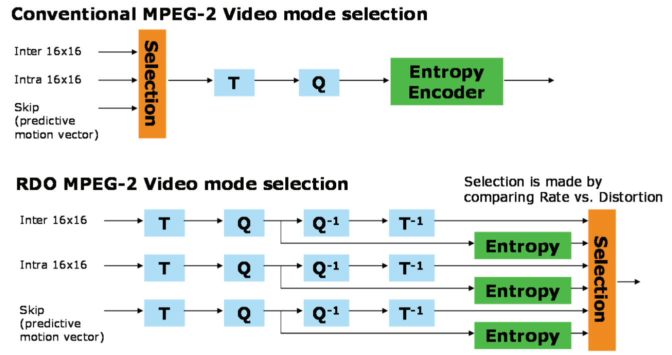

Figure 3. Rate distortion optimization example.

Figure 3 illustrates an example of RDO applied to an MPEG-2 Video encoder and how it compares to a conventional mode selection performed by a traditional MPEG-2 Video encoder.

SIMULATION RESULTS

In order to prove out the theory stated in this paper, experiments were conducted in the TANDBERG Television encoder laboratory. An accurate research simulation model was used to represent the Revolutionary Optimized MPEG-2 Video Encoder and this was compared against the best traditional MPEG-2 Video encoder product available on the market. Operational settings were chosen to be representative of what is expected to be used in the field.

Due to the processing power required, experiments were performed for standard definition only. However, the general consensus is that significant gains can be achieved for high definition as well.

Summary

Testing demonstrated that a 15% improvement in MPEG-2 Video compression efficiency is realistic. All of these techniques are compatible with the MPEG-2 Video specification, guaranteeing interoperability with legacy set-top boxes and TV receivers.

Figure 4 contains the PSNR results of a comparison between the Revolutionary Optimized MPEG-2 Video Encoder and a traditional MPEG-2 Video encoder for a typical broadcast sports sequence (soccer). At a 6Mbps operating point, a scenario typical of constant bit-rate encoding, a bit-rate reduction of approximately 15% has been realized for an equivalent picture quality. At a 3.5Mbps operating point, a scenario typical of statistically multiplexed channels, a saving of approximately 18% has been realized.

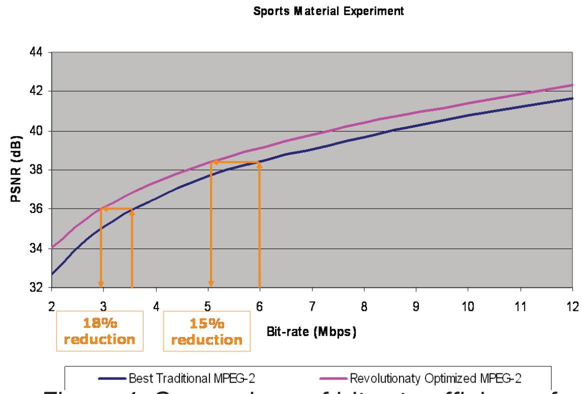

Figure 4. Comparison of bit-rate efficiency for real-time sports.

Figure 6. Soccer test sequence.

Figure 7. Kiel test sequence.

Constant Bit-Rate

Three common test sequences (Figures 5-7) were considered at various bit-rates to show how the Revolutionary Optimized MPEG-2 Video Encoder compares to a traditional MPEG-2 Video encoder. The CBR tests were carried out using the Tektronics PQA200. This provided accurate and repeatable results on short test sequences because it uses a double-ended measurement algorithm with a full reference source.

Figure 5. Mobile and Calendar test sequence

Encoder settings were as follows:

- GOP size: 12 frames
- GOP structure: IBBP adaptive
- Buffer size: standard (1.6Mb)
- Intra_DC_precision: 8b
- Alternate_scan: Auto
- Bandwidth: Sharp
- Horizontal resolution: 720

Results are shown in Figures 8 and 9.

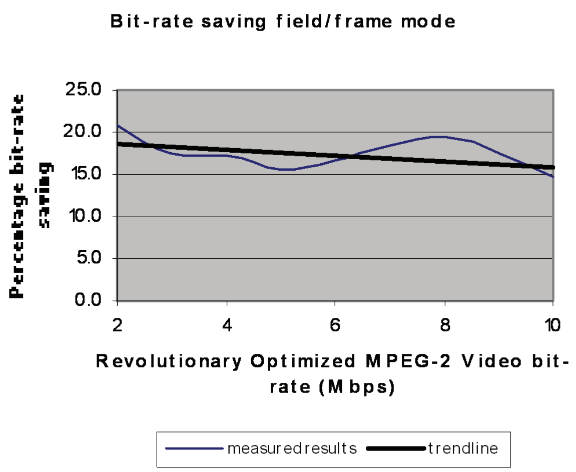

Figure 8. Bit-rate savings in field/frame mode.

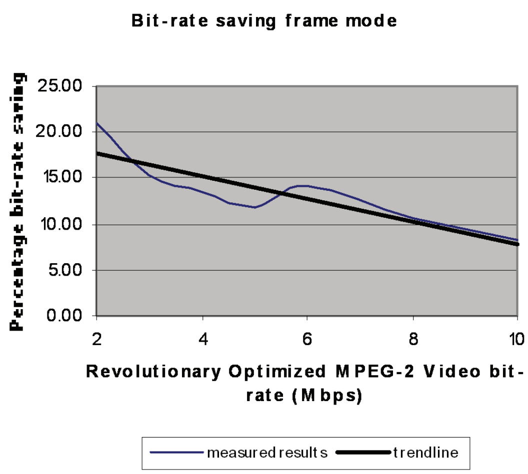

Figure 9. Bit-rate savings in frame only mode.

As can be seen, the bit-rate saving in picture field/frame mode is more than 15% over the entire bit-rate range. Although the bit-rate saving in frame-only mode is slightly lower, at the critical bit-rate range of 2-3Mbps, the bit-rate saving is still above 15%.

Statistical Multiplexing

Because many direct-to-home applications use statistical multiplexing (statmux), these experiments were performed as well, comparing the Revolutionary Optimized MPEG-2 Video Encoder to a traditional MPEG-2 Video encoder.

A five channel statmux encoding system was used for comparison. All five input video channels were time synchronized sequences, representing typical genres. The same sequences were compressed using the research simulation model simulating statistical multiplexing.

The investigation compares a like-for-like and a 15% reduction in group rate for the Revolutionary Optimized MPEG-2 Video Encoding system.

Five critical sequences were chosen for the statmux experiment.

1. Sport, Football (Soccer)
2. Sport, Rugby
3. Film
4. Studio, Sitcom
5. Studio, Music video

These are shown in Figure 10 below:

Figure 10. Five statmux test sequences of varying genres.

Encoder settings were as follows:

- GOP size: 12 and 36 frames
- Group bit-rate: 12.5Mbps (traditional MPEG-2)
- Group bit-rate: 10.625Mbps (Revolutionary Optimized MPEG-2)
- Adaptive GOP - Revolutionary Optimized MPEG-2 only
- GOP structure: IBBP adaptive
- Picture field/frame: Frame mode
- Bandwidth: Medium and Sharp
- Intra_DC_precision: 8b
- Alternate_scan: Auto
- Bandwidth: Sharp
- Horizontal resolution: 720

The picture quality was measured by DVQ, using Rohde & Schwarz DVQ Digital Video Quality Analyzer. The DVQ method was adopted since it is most suitable for statmux picture quality measurements. This measurement technique has the advantage that it does not need a reference signal, but it requires longer test sequences to produce more accurate statistical results.

Tables 1 and 2 show the DVQ difference between the traditional MPEG-2 Video statmux system and the Revolutionary Optimized MPEG-2 Video statmux system (at a 15% bit-rate reduction) at two different GOP sizes. Figures 11 and 12 show the average DVQ results for 12 and 36 frame GOPs respectively.

Table 1. *Revolutionary Optimized Statmux vs. Traditional Statmux (GOP=12).*

	sport1	sport2	film3	studio4	studio5
Revolutionary Optimized (15% less bit-rate)	68.17	60.16	74.9	62.69	74.66
Traditional MPEG-2	64.78	57.96	75.07	63.03	74.3
DVQ diff	3.39	2.2	-0.17	-0.34	0.36

Table 2. Revolutionary Optimized Statmux vs. Traditional Statmux (GOP=36).

	sport1	sport2	film3	studio4	studio5
Revolutionary Optimized (15% less bit-rate)	75.4	66.02	80.92	71.96	76.61
Traditional MPEG-2	70.77	63	78.18	68.01	77.56
DVQ diff	4.63	3.02	2.74	3.95	-0.95

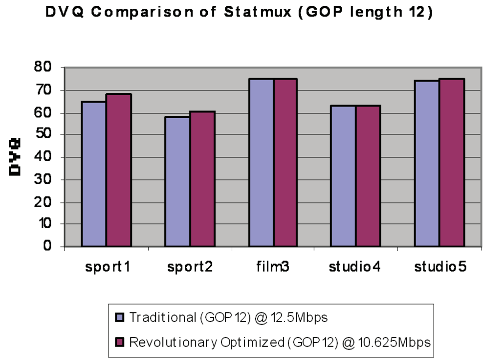

Figure 11. DVQ comparison of statmux, GOP length 12.

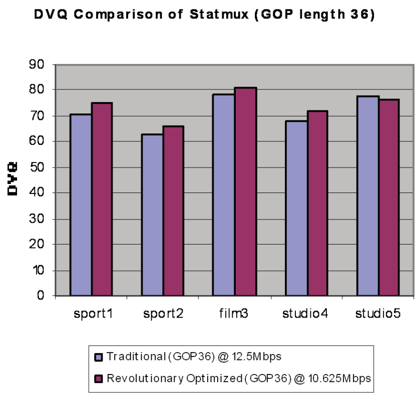

Figure 12. DVQ comparison of statmux, GOP length 36.

As can be seen from the above charts, the picture quality with a longer GOP is better than a 12 frame GOP, both in the Revolutionary Optimized MPEG-2 Video Encoder statmux and in the traditional MPEG-2 Video encoder statmux system. This is due to the bit-rate savings in P and B pictures compared to I pictures. In both cases, the Revolutionary Optimized MPEG-2 Video Encoder is better than the traditional MPEG-2 Video encoder on average as well as in the majority of sequences. Although there is small penalty in studio5 in the Revolutionary Optimized MPEG-2 Video Encoder (for GOP36), the difference is insignificant since the picture quality of all other sequences have been improved, despite the fact that the group bit-rate has been reduced by 15%.

CONCLUSIONS

Double-ended picture quality measurements at constant bit rates have shown that the Revolutionary Optimized MPEG-2 Video Encoder provides a substantial bit-rate saving over a wide a range of bit-rates compared with the best available traditional MPEG-2 Video encoder. Although the bit-rate saving in frame mode drops toward higher bit-rates compared to picture field/frame mode, at critical bit-rates in the range 2-3Mbps, the Revolutionary Optimized MPEG-2 Video Encoder provides a bit-rate saving of at least 15%.

The DVQ results have shown that the average picture quality of the Revolutionary Optimized MPEG-2 Video Encoder with a 15% group bit-rate reduction is higher both in picture adaptive field/frame mode and in frame mode. Visually, the Revolutionary Optimized MPEG-2 Video Encoder is clearly better, especially in plain areas and less blocky in the background.

Over the past several years, there has been little coding efficiency gains with MPEG-2 Video and many in the industry believed that MPEG-2 Video had reached its theoretical coding efficiency asymptote. However, the Revolutionary Optimized MPEG-2 Video Encoder clearly demonstrates that this is not so. With 15% or more in bit-rate reduction possible, it provides a significant step-change in commercial applications. Service operators will realize fewer transponders/channels for the same amount of services or more services in the same amount of bandwidth. For direct-to-home satellite or cable operators, this frees up spectrum for additional high definition services. For digital terrestrial television operators, this will free up spectrum for new services such as mobile/handheld television. In all of these cases, these improvements are compatible fully with the hundreds of millions of legacy digital televisions and set-top boxes that decode the MPEG-2 Video standard only.

For more information on TANDBERG Television's EN8100 Advanced MPEG-2 Broadcast Encoder, please see the attached datasheets or visit www.tandbergtv.com.

REFERENCES

[1] M. Goldman et al, "ATSC Encoding, Transport, and PSIP Systems", Chapter 5.22, NAB Engineering Handbook, 10th Edition, 2007.

[2] A. Bock, "What Factors Affect the Coding Performance of MPEG-2 Video Encoders?", DVB '99 Conference Proceedings, March 1999.

[3] G. Sullivan, T. Wiegand, "Rate Distortion Optimization for Video Compression", IEEE Signal Processing Magazine, November 1998.

ACKNOWLEDGEMENTS

Many thanks to my TANDBERG Television colleagues Alois Bock and Matthew Compton for their assistance with this paper.