THE FUTURE OF TRANSCODING – THE NEED FOR MPEG-2 AND MPEG-4 TO COEXIST

John Hartung, Santhana Krishnamachari
EGT

Abstract

Virtually all commercial digital video content is stored and distributed using the MPEG-2 encoding standard introduced about 12 years ago. Although this standard enabled a large increase in the number of programs that could be carried in access networks, the capacity of those networks has not kept up with the explosion of content and the increased bandwidth requirement for high definition. The introduction of switched access networks, both HFC and IP, will help alleviate that bottleneck and allow the introduction of new clients that make use of the more efficient MPEG-4 part 10 (a.k.a. H.264) encoding standard. The challenge in realizing these gains, however, will depend on transcoding technology that is both cost effective and maintains the quality of the originally distributed program.

THE LIMITS OF MPEG-2 ENCODING

The MPEG-2 standard, like all media encoding standards, is defined by the encoded bit stream syntax, and the semantics, or operations, signified at the decoder by these bit stream elements. The main semantic elements include block based motion compensated prediction, quantized transform based encoding of prediction residuals and reference blocks, and entropy coding of encoding parameters. MPEG-2 encoder operation is not specified by the standard and encoders are not required to make use of all the operations available at the decoder. The standard was designed with complexity in mind so that real time encoders could be economically deployed. Early MPEG-2 encoders were limited in performance due to the required computational complexity needed to generate optimal bit streams. In general, this optimality requires a global search over a large number of encoding parameters, and processors capable of this computation would have made their cost prohibitive. Since the introduction of this standard the performance of integrated circuits has dramatically increased, and algorithms have been developed that achieve near optimal performance at greatly reduced complexity.

Today’s MPEG-2 encoders achieve near optimal performance in terms of objective performance measures such as the peak signal to noise ratio (PSNR) as illustrated in Figure 1. This plot shows the distortion, relative to the original content, averaged over a representative set of 18 video sequences including 24 fps film content, 30 fps interlaced content, and clean and noisy sequences. PSNR of about 34 dB results in high quality encoding that is nearly indistinguishable from the original content. As seen in the plot, the MPEG-2 encoding algorithm can achieve this result, over a broad range of sequences, at about 3.5 Mbps for standard definition video.

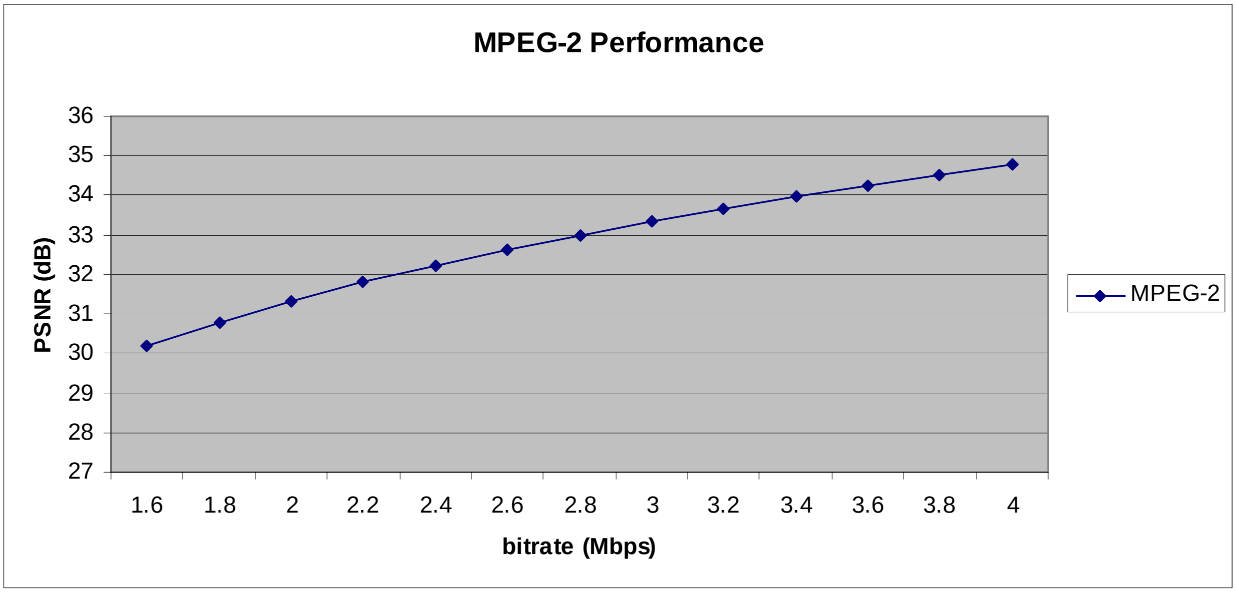

Figure 1

Objective measures, however, do not tell the entire story of encoder performance. In addition to the purely objective distortion measures, today's encoders take into account subjective evaluation through the incorporation of human visual system (HVS) models. These models estimate the masking of distortion by content so that encoding bits are preferentially spent to reduce the distortion that is most visible. Encoders also make use of preprocessing to improve their performance on noisy sequences. These adaptive filters remove noise from the original sequence before encoding, both to restore the original content and to avoid allocating excessive bits to encode the typically high frequency noise. Although the use of HVS models and preprocessing are not covered in the encoding standards, they are important to the practical deployment of video encoding.

MPEG-4 PERFORMANCE

The MPEG-4 and MPEG-2 encoding algorithms are similar in that they both use block based motion compensated prediction, quantized transform coding of residuals, and entropy coding. However, MPEG-4 introduces basic differences in these tools along with additional modes of operation.

The MPEG-4 motion estimation tools have been expanded to include additional block shapes and sizes, and multiple reference frames can be used to predict a macroblock. Field and frame prediction can also be varied on a macroblock basis. A new intraframe spatial prediction mode has also been introduced that has no correspondence to the MPEG-2 coding modes. The DCT transform used in MPEG-2 has been replaced by a smaller integer transform, and VLC entropy coding has been augmented with an optional adaptive binary arithmetic entropy coder (CABAC). The MPEG-4 standard also allows a filter in the encoding loop that helps mitigate encoding artifacts, such as blocking at low encoding rates. Although this does not improve the PSNR performance, it results in more acceptable subjective artifacts. Making full use of these new features enables a 30% to 50% reduction of coding rate for equivalent video quality at the expense of a

5-6 times increase in complexity. Figure 2 shows the performance of MPEG-4 averaged over a broad range of sequences and encoding rates. The input sequences are full D1 resolution and the encoder is operating at main profile@level 3.

It can be seen from the plot that a PSNR of about 34 dB is achieved at an encoding rate of about 2.1 Mbps as compared to 3.5 Mbps for MPEG-2, or a rate reduction of about 40%.

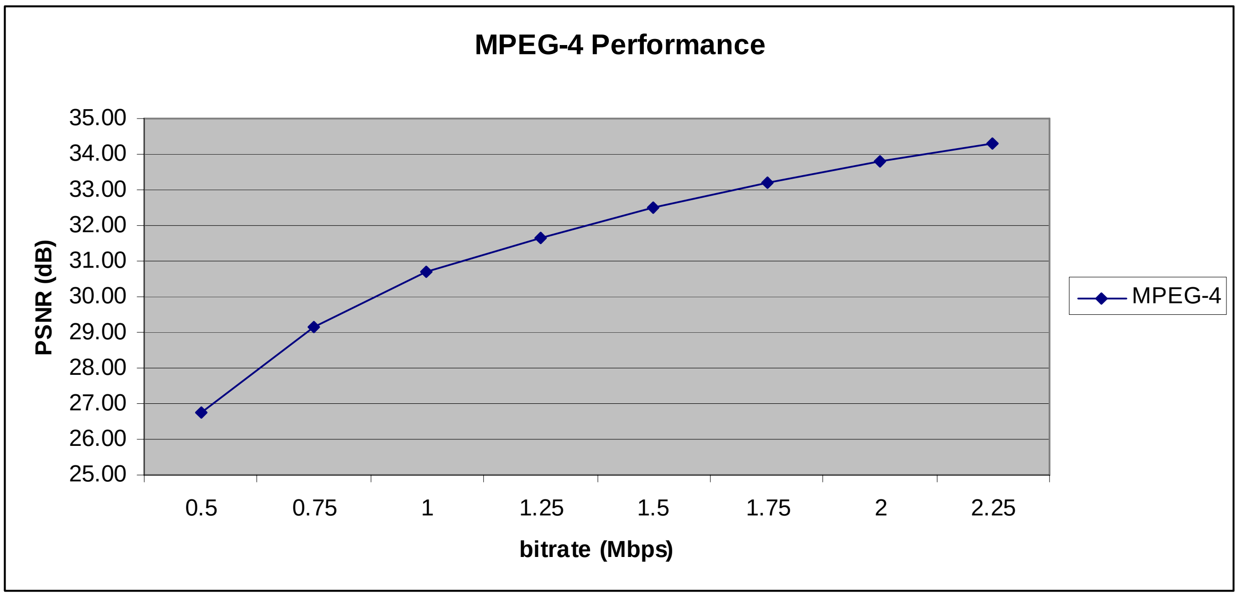

Figure 2

As with MPEG-2, MPEG-4 encoding benefits from preprocessing and HVS modeling and identical techniques can be applied to improve the subjective quality.

SWITCHED NETWORKS AND MULTIPLE CODECS

Switched networks have been deployed in cable networks for video on demand (VOD) applications, and are being deployed to increase the effective capacity for broadcast applications. Switching for VOD enables content to be switched to an individual set top box based on a subscriber request. Switched broadcast enables broadcast content to occupy a portion of the access network only when it is requested by one or more set top box clients. In this case, capacity is increased beyond broadcasting all channels because only a small number of the programs offered are actually requested concurrently.

Switching both for VOD and broadcast can also be used to enable further bandwidth savings by tailoring the requested content based on the capabilities of the requesting set top box. Because of the large number of deployed MPEG-2 set top boxes, both MPEG-2 and MPEG-4 set top boxes will co-exist in cable networks for some time. In order to take advantage of the additional bandwidth savings of MPEG-4 set top boxes, content needs to be available for either on the network when requested. Since switched applications are aware of the requesting client, the delivery can be tailored to the capability of that client, e.g. for VOD only the single requestor need be considered. In the case of switched broadcast all requestors must be considered in order to avoid transmitting the same broadcast content in multiple formats. This can be avoided by transmitting the content in MPEG-2 format only since most MPEG-4 set top boxes can also decode MPEG-2 content.

In the case of VOD, content can be stored in multiple formats on the server, however, broadcast applications require transcoding from the predominant MPEG-2 format to MPEG-4 in real time.

APPROACHES TO TRANSCODING

Several approaches can be taken to transcode from MPEG-2 to MPEG-4. The lowest complexity approach involves mapping the MPEG-2 encoding parameters into MPEG-4 equivalent representations. This is similar to the encoded domain rate shaping that is used for statistically multiplexing MPEG-2 streams. Encoding gains can be obtained through the intraframe prediction modes and improved entropy encoding available in MPEG-4, however, a large set of tools would be restricted from use limiting the ultimate performance.

An alternative approach is to decode the MPEG-2 content and apply the decoded baseband video directly to an MPEG-4 encoder. This approach does not produce high quality results, as illustrated in Figure 3. This plot compares the average PSNR of original sources that have been encoded with an MPEG-4 encoder, vs. the PSNR of decoded MPEG-2 sequences that have been encoded with the same MPEG-4 parameters. In this case the MPEG-2 sequences were coded at 4 Mbps and 3 Mbps, and their PSNRs were about equal, or greater, than that of the MPEG-4 encoding. This result uses the same set of sequences used in Figures 1 and 2. As the plot shows, the PSNR degrades up to 1 dB in the decode/encode case when the MPEG-2 was coded at 4 Mbps. One of the main reasons for the degradation is due to the fact that the frame type is not maintained between the MPEG-2 and MPEG-4 encodings. The B frames are typically encoded at lower rate,

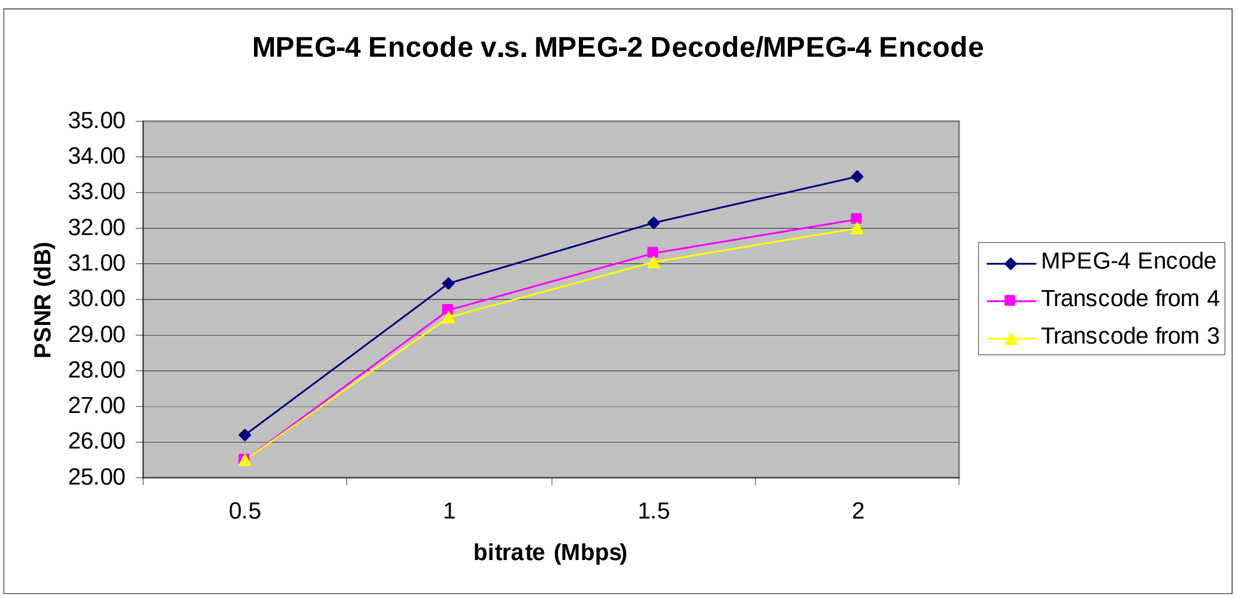

Figure 3

and PSNR, than I and P frames in the original MPEG-2 encoding. Without knowledge of the frame type, the MPEG-4 encoder can re-encode the B frames as I or P frames and subsequently use these as reference frames for prediction. This results in lower PSNR in the predicted frames and propagation of this distortion. The results shown at 4 Mbps are for high quality MPEG-2 encoding that maintains good quality for I, B, and P frames. Lower MPEG2 rates, and/or lower quality encoders produce a larger variation in the quality of the different frame types resulting in poorer transcoding results as shown in the 3 Mbps result in Figure 3. As illustrated in the plot, the resulting PSNR of the transcoded sequence degrades further at the lower MPEG-2 rates.

A final method for transcoding also performs decoding to baseband video and re-encoding using an MPEG-4 encoder, however, the MPEG-4 encoder makes use of the MPEG-2 encoding parameters. One example of this is to maintain the frame coding type to avoid the degradation described in the previous method. Referring to the results in Figure 3 again, approximately .2 dB improvement can be

Figure 4

gained when transcoding from 4 Mbps MPEG-2 to 2 Mbps MPEG-4. For lower rate and/or lower quality MPEG-2 encoding even larger gains are possible. Passing additional parameters allows for further improvements in transcoding performance, and a reduction in complexity. An example of this is the use of bit allocation in the MPEG-2 encoding as a complexity estimator that can be used for rate allocation in the transcoded sequence. This is similar to two-pass encoding algorithms, however, the necessary information already exists in the MPEG-2 bitstream. This technique improves the transcoding quality without the complexity of a two-pass MPEG-4 implementation.

Overall, high quality transcoding and reduced complexity is achieved through full decode/encode with reuse of encoding parameters. This argues for a tightly coupled system that receives MPEG-2 programs, either in SPTS or MPTS, and converts directly into MPEG-4 transport streams.

A final consideration in transcoding is the mitigation of source noise and coding artifacts in the original MPEG-2 encoding. This can be accomplished by filtering the decoded MPEG-2 sequence, either in the transform domain, or in the reconstructed baseband for encode/decode transcoders.

Figure 5

High quality MPEG-2 encoders typically apply sophisticated prefilters that remove noise, however, noise filtering improves transcoding when this is not the case. A second source of noise can be introduced by the MPEG-2 encoder. This structured noise can be effectively estimated and adaptively removed to improve the subjective quality of the transcoded sequence. Figures 4 and 5 show the corresponding frame in a transcoded MPEG-2 sequence. Both frames use the same original and transcoded parameters, however, Figure 5 demonstrates the improvements gained through post-filtering of the MPEG-2 sequence.

CONCLUSION

The need for bandwidth efficiency in cable plants continues to be driven by the increase in service and content offerings. Increasing amounts of high definition content further add to the need for improved efficiency. Switched services will help provide this additional bandwidth and enable a transition to the more efficient MPEG-4 encoding standard, however, this transition will require the coexistence of both MPEG-2 and MPEG-4 services. This coexistence will be enabled by integrated transcoders that provide a high quality, low complexity solution.

CONTACTS

John Hartung
john@egtinc.com
Santhana Krishnamachari
santhana@egtinc.com