TRANSCODING AND STATISTICAL MULTIPLEXING OF MPEG4 (H.264) BROADCAST VIDEO

John Hartung, Ph.D.
EGT

Santhana Krishnamachari, Ph.D.
EGT

Abstract

The bandwidth demand of HD content is driving the use of more efficient video compression such as MPEG4 (H.264) encoding for satellite distribution, and statistical multiplexing for MSO access networks. Transcoding and statistical multiplexing are usually implemented independently; however, in this paper we show that this is not the best approach. We show that integrating the transcoding and statistical multiplexing operations will result in improved video quality, reduced operational complexity and lower cost. The paper is organized into four sections: Introduction, Transcoding, Statistical Multiplexing, and Conclusion.

INTRODUCTION

MSOs are planning to increase the number of HD programs they offer from around 25 today to more than 100 over the next couple of years. This increase is placing a tremendous strain on the available access bandwidth in the MSO HFC networks. Three approaches are being taken to solve this bandwidth problem: analog channels are being converted to more efficient digital transmission, switched digital broadcast is being deployed, and HD content is being statistically multiplexed so that 3 or more HD channels are carried in a QAM. The statistical multiplexing approach has the advantage of not requiring the additional network infrastructure and software needed to support switched broadcast or of turning off existing analog services. In addition, the cost of statistical multiplexing can be shared across multiple service groups and nodes locally, regionally, or nationally.

The new HD channels will be received in various encoding formats and bit rates from satellite distribution and terrestrial broadcasters. MPEG2 format is typically received at rates around 15 Mbps or higher, and satellite distributors are beginning to use MPEG4 encoding at 8 Mbps for new programming. Although MPEG4 capable set top boxes are beginning to be deployed, the large numbers of legacy MPEG2 set top boxes require the conversion of all content into MPEG2 format. This paper describes various approaches for transcoding and statistical multiplexing with quantitative comparisons. A novel approach that combines transcoding with statistical multiplexing is shown to have the best compression efficiency and quality.

TRANSCODING

Overview

In general, transcoding from MPEG4 to MPEG2 requires a full decode and re-encode because many of the tools available in the MPEG4 standard are incompatible with MPEG2. Examples of these tools include advanced prediction algorithms such as the use of multiple reference frames and intraframe prediction, and filtering in the prediction loop to reduce blocking artifacts. In some specific instances the MPEG2 parameters can be determined or estimated from the MPEG4 parameters leading to higher quality and lower complexity.

Independent Decode-Encode

One approach to transcoding from MPEG4 to MPEG2 is to fully decode the MPEG4 frames and then re-encode with an MPEG2 encoder. This can be implemented with an entirely separate decoder and encoder; however, this approach does not produce the highest possible MPEG2 encoding quality and is computationally expensive. One reason that quality is compromised is frame coding types are not preserved, and therefore high quality reference frames, such as I and P frames, are not re-encoded with the same types in MPEG2. This results in lower quality reference frames and propagation of coding distortion when they are used for prediction of P and B frames. Separation of decode and encode functions also prevents the original MPEG4 encoding parameters from being re-used as initial MPEG2 encoding parameter estimates to reduce complexity. Reuse of these parameters is especially useful in motion estimation where initial estimates can be used to reduce the search complexity by limiting the search range.

Integrated Transcoding

An alternative approach is to decode the MPEG4 input, and at the same time pass the MPEG4 encoding parameters to the MPEG2 encoding stage. This is shown in Figure 1. In addition to preserving frame types and reducing encoding complexity, the MPEG4 parameters are also used by the First Pass
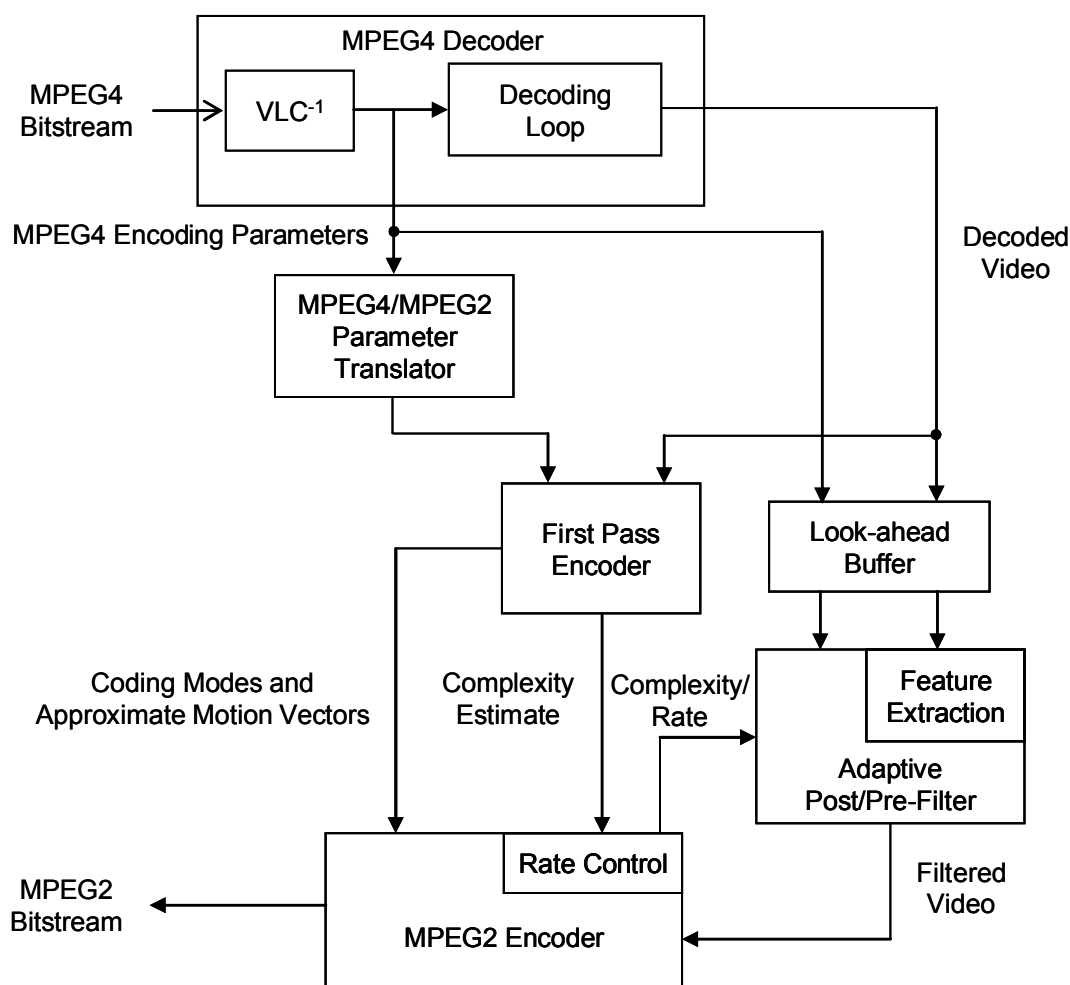

Figure 1

Encoder and Adaptive Post/Pre-Filter. The First Pass Encoder determines the MPEG2 encoding modes and approximate prediction residuals from both the decoded video and MPEG4 encoding parameters, where possible. A relative complexity is determined for each frame within a group of pictures (GOP), and this in turn is used by the second pass MPEG2 encoder Rate Control to determine an optimal target encoding rate for each frame within the Look-Ahead Buffer, thereby achieving the best overall quality.

Integration of the decoding and encoding functions also enables advanced Adaptive Post/Pre-Filtering of the decoded video. This filtering serves two purposes: removal of encoding artifacts from the decoded MPEG4 bit stream, and filtering to reduce MPEG2 encoding artifacts. For both types of filtering feature extraction is used to identify areas having characteristics that mask distortion due to the response of the human visual system (HVS). For example, distortion in textured areas is difficult to perceive so those areas can be highly filtered to reduce the required number of coding bits, while areas with edges

need to be preserved in order to retain image details. The original MPEG4 encoding parameters are used to adaptively remove encoding artifacts by estimating the encoding distortion from prediction parameters and quantization step sizes. The MPEG2 Encoder allocates a coding rate to each frame based on its' complexity and the bits available for all frames within the GOP. This ratio of complexity /rate indicates the amount of pre-filtering needed to minimize artifacts in the MPEG2 encoded frame. Optimal filtering and reduced complexity result from the availability of both MPEG4 and MPEG2 encoding parameters along with feature extraction of the decoded video.

Comparison

Figure 2 shows a plot comparing the quality of Independent Decode-Encode with Integrated Transcoding for 1920 x 1080i HD video. The original MPEG4 video is encoded at 10 Mbps. Peak signal to noise ratio (PSNR) is used as an objective measure of the difference between the original and encoded video. A higher PSNR represents better
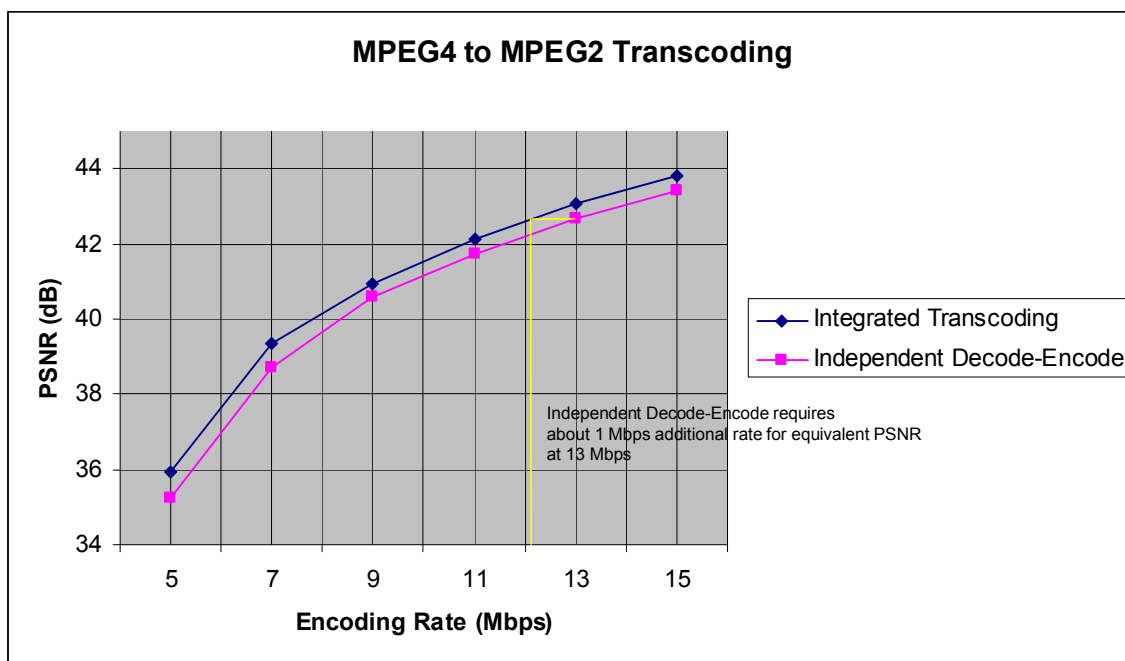

Figure 2

quality with about a .5 dB change resulting in a perceived quality difference. It can be seen from the plot that Independent Decode-Encode achieves, on average, about .5 dB lower PSNR than Integrated Transcoding. This translates into about a 1 Mbps higher rate to achieve equivalent performance. The next section also shows that the average PSNR is not the whole story when it comes to statistical multiplexing. Integrated Transcoding also results in lower frame to frame PSNR variance and therefore a more uniform and lower rate to achieve a constant quality.

STATISTICAL MULTIPLEXING

Overview

HD channels are delivered to a head end at a constant bit rate using either MPEG4 or MPEG2 encoding. The bit rate is chosen to produce good quality for the most difficult sequences, even though a lower rate would be sufficient most of the time. For MPEG2 HD content this rate is 15 Mbps, or higher, allowing only two channels to be transmitted within a 6 MHz QAM channel. Statistical multiplexing increases the number of

programs that can be carried by re-encoding each input at a lower rate that varies as a function of the channel's complexity. The individual rates are controlled in order to maintain the original video quality at an aggregate rate that allows additional channels to be carried within a QAM. For a 38.8 Mbps QAM channel this corresponds to an average encoding rate of below 13 Mbps for three or more HD channels.

The challenge to achieving multiplexing gain is to combine channels such that their instantaneous encoding rate remains close to their average rate. For SD this requirement is met because of the large number (>12) of channels transmitted in a QAM. However, with only three HD channels transmitted in a QAM, the channel characteristics must also be considered. One approach is to combine two low complexity channels, such as progressive movie content, with one high action channel, such as sports. A second factor that limits the number of channels that can be multiplexed is the efficiency of transcoding and rate shaping. These translate directly into the rate required to encode individual channels at high quality.
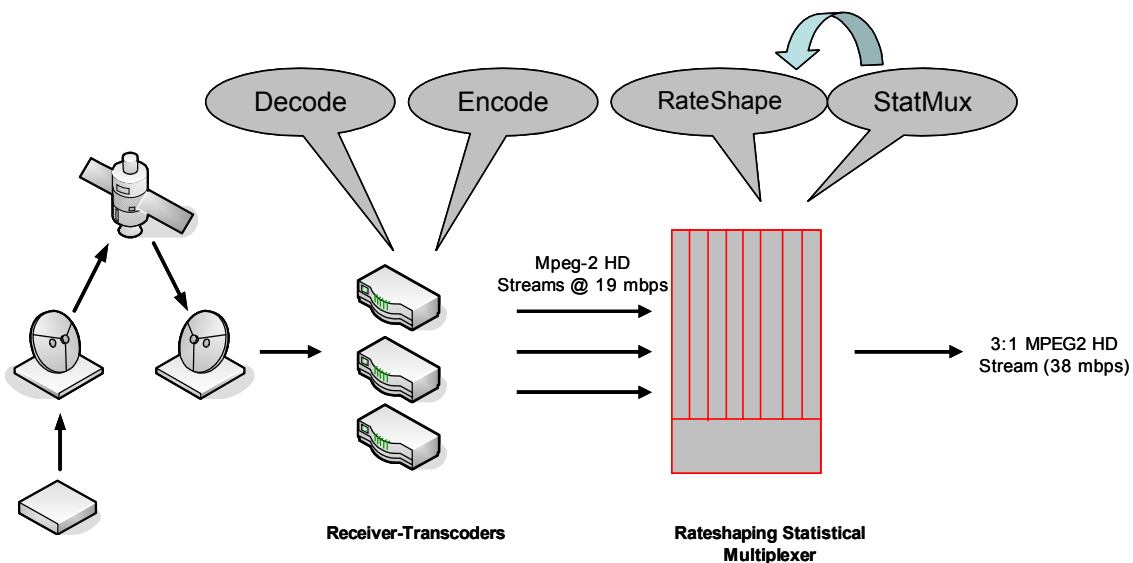
Two methods have been used to transcode and

Figure 3

rate shape content for statistical multiplexing as described below. These are transcoding from MPEG4 to MPEG2 followed by rate shaping, and MPEG4 decoding and MPEG2 re-encoding with a closed loop statistical multiplexer. A third approach, MPEG4 to MPEG2 transcoding integrated with closed loop statistical multiplexing, is shown to produce the best quality.

Transcoding and Rate Shaping

In this architecture the MPEG4 input is first transcoded to MPEG2 in the Receiver-Transcoders. For HD MPEG4 delivered at 8 Mbps this first stage of transcoding produces an MPEG2 output of about 15 Mbps. The MPEG2 programs are then statistically multiplexed in a second stage of rate shaping to form an MPTS meeting the QAM rate as shown in Figure 3. The second stage is usually implemented using a rate shaper that modifies the original MPEG2 encoding parameters without performing a full decode and re-encode. This approach runs into problems when the rate reduction for any channel exceeds about 15%; significant video quality reduction occurs under these conditions. A rate reduction of significantly greater than 15% is fairly common, and occurs whenever two of the channels need a bandwidth above 13 Mbps to achieve adequate quality. If we consider the case where two channels require 14 Mbps, then the third channel must be reduced by greater than 33% (15 mbps to 10 mbps) to meet the total rate of 38.8 Mbps. The performance of this approach is fundamentally limited by the fact that is uses two stages of MPEG processing, transcoding followed by rate shaping. In the comparison section we show that the performance falls well below the two other approaches.

Decoding and Closed Loop Encoding

A second approach begins by decoding the input MPEG4 bitstreams using Receiver-Decoders as shown in Figure 4. This output is then re-encoded using MPEG2 encoders within a closed loop statistical multiplexer. A single stage of re-encoding, decode followed by encode, introduces less distortion than the previous method, however separation of the decoder and encoder prevents reuse of the original MPEG4 encoding parameters. This in turn leads to a lower PSNR for the target rates determined by the statistical multiplexer. The Comparison section also shows that both this approach and the previous one, introduce greater variance in the frame to frame PSNR. This shows up as artifacts in the statistically multiplexed video that degrade the video more than would be reflected in the average PSNR comparisons.
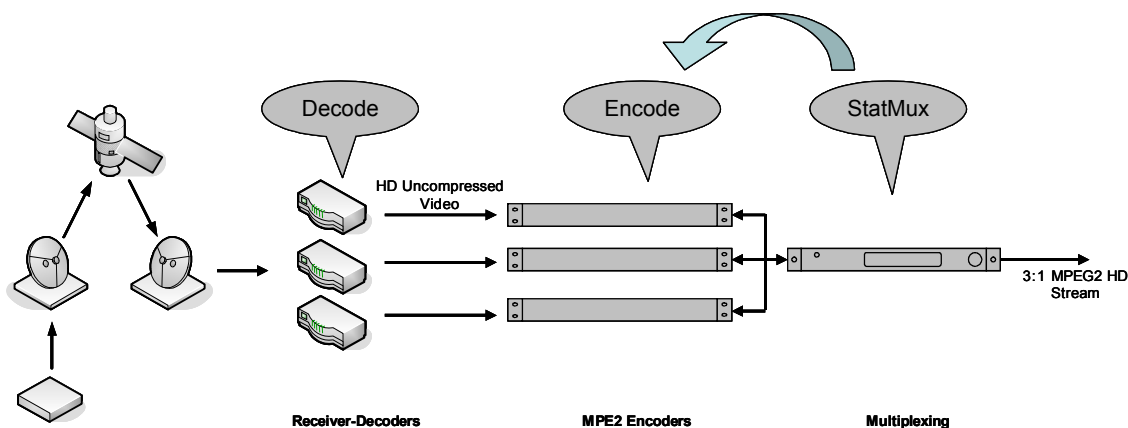

Figure 4

Closed Loop Transcoding

The third approach converts the MPEG4 output from a receiver directly to MPEG2 using an Integrated Transcoder, as shown in Figure 5. This approach achieves the best performance by transcoding directly to the statistical multiplexing rate in a single stage of MPEG encoding. The rate feedback also enables the integrated transcoder to adapt the post/pre filters for the target rate, rather than an intermediate rate. The perceptual quality is also improved by the lower encoding rate variance achieved in this implementation as shown in the next section.
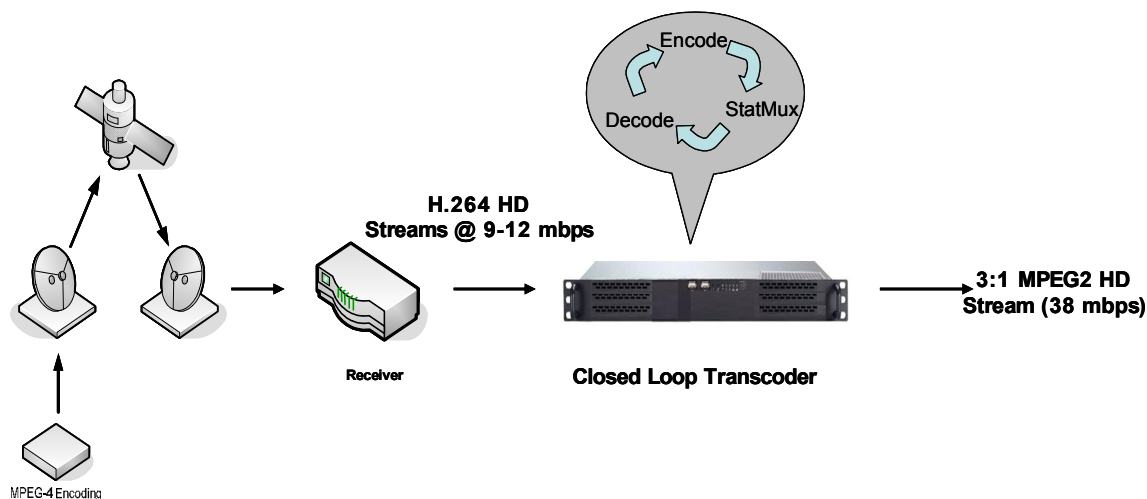
one channel peaks. Lower target rates can be chosen for the easier channels, allowing a higher rate to be allocated for the complex channel.

Figure 7 shows the MPEG2 output frame PSNRs for 1920 x 1080i video transcoded using the three approaches. The original MPEG4 video was encoded at 10 Mbps and the output is at 13 Mbps. The important consideration here is the variance of the PSNR for each frame. A lower PSNR indicates that the frame is more complex and would need to be coded at a higher bit rate to achieve equal quality. Rate shapers having a

Figure 5

Comparisons

Figures 6 and 7 show the single channel PSNR performance for the three statistical multiplexing approaches described above. The results are for 1920 x 1080i HD video originally encoded using MPEG4 at 10 Mbps. Integrated Transcoding achieves a 2.75 Mbps advantage over Transcoding and Rate Shaping at rates around 13 Mbps as shown in Figure 6. It achieves a 1 Mbps advantage over Decoding and Closed Loop Encoding as shown in a previous section. These gains produce higher overall quality, but are particularly important when the complexity of

high variance produce frequent artifacts because there is a higher probability that individual target rates exceed the aggregate available rate. The plot shows that the Closed Loop Transcoder achieves the lowest variance, followed by the Decoder with Closed Loop Encoding and Transcoding and Rate Shaping implementations.

CONCLUSION

Integrating transcoding and statistical multiplexing produces several benefits over competing approaches, the most important being optimum compression efficiency and
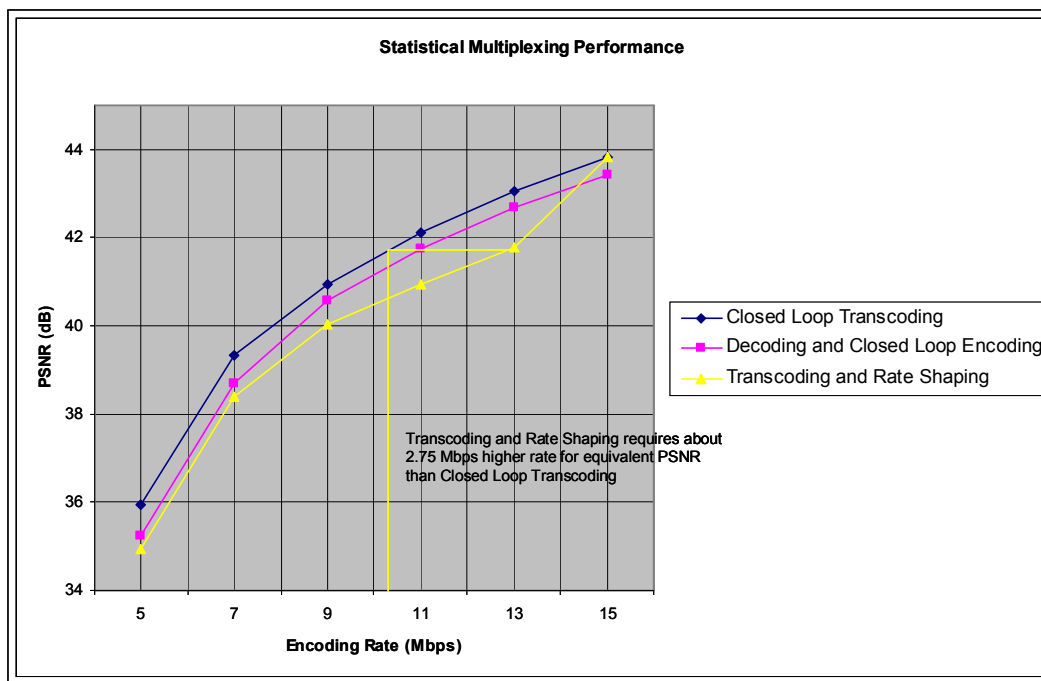
Statistical Multiplexing Performance
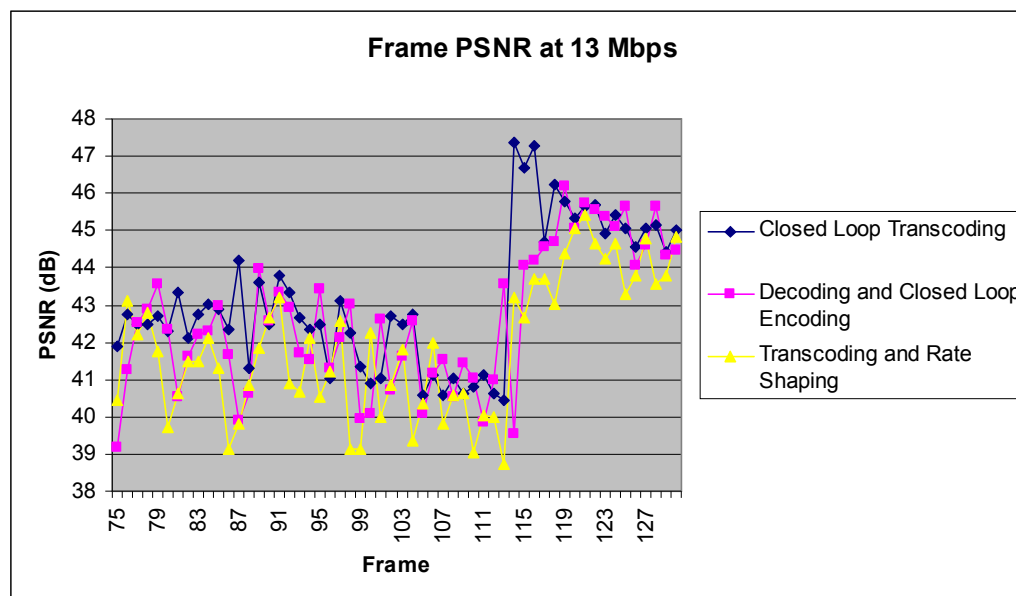

Figure 6

Frame PSNR at 13 Mbps

Figure 7

video quality. This architecture achieves these benefits through the reuse of the original encoding parameters, both for conversion between input and output formats, and for encoding at the statistical multiplexing rate. Integrating the two functions also enables a single stage of encoding thereby avoiding the distortion due to two generations of processing. Although this paper has focused on transcoding from MPEG4 to MPEG2 standards, similar gains are achieved when constant bit rate MPEG2 content is statistically multiplexed into an MPEG2 output.