No-reference H.264/AVC Statistical Multiplexing for DVB-RCS

Hamed Ahmadi Aliabad[ϕ], Sandro Moiron[ϕɣ],
Martin Fleury[ϕ]*, and Mohammed Ghanbari[ϕ]

[ϕ]University of Essex,
Colchester, United Kingdom
{hahmada, fleum,ghan}@essex.ac.uk

[ɣ]Instituto de Telecomunicações
Leiria, Portugal
smoiron@co.it.pt

Abstract. Replacement of MPEG-2 by the H.264/AVC codec for satellite video services, including aggregated video for the DVB-RCS uplink, presents an opportunity to develop efficient statistical multiplexing. In this paper, a scheme is developed that effectively models the relationship between number of non-zero coefficients and video quality. The result is the ability to equalize video quality, reduce quality fades, and smooth the overall bitrate presented. The number of channels in the multiplex can be increased through the scheme if there is a need to achieve an average of 40 dB PSNR across all inputs.

Keywords: DVB-RCS, DTH, H.264/AVC, look-ahead statistical multiplexing.

1 Introduction

In the current economic environment, there is a strong incentive to reduce costs for service providers of TV satellite broadcasting, even though revenues remain buoyant for the satellite companies themselves. To a certain extent this applies to content distribution and more especially it applies to Direct-to-Home (DTH) broadcast and its successor Digital Video Broadcasting (DVB)-Return Channel via Satellite (DVB-RCS) [1]. DVB-RCS replaces the terrestrial uplink of DTH with a satellite uplink with several Mbps available, thus further increasing the attraction of the DTH offering for certain markets in terms of massive coverage and low installation costs. One way DVB-RCS costs may be reduced is through efficient statistical multiplexing. In statistical multiplexing, a constant bitrate, the transponder's bandwidth, is allocated according to the coding complexity of the constituent video streams. Efficient statistical multiplexing can improve received video quality at the receiver, and may even increase the number of TV channels carried by a transponder [2]. In business terms it is acknowledged that the revenue that can be potentially generated from combining video streams within a multimedia channel [3] is related to the quality of the video delivered to end users. Statistical multiplexing can reduce deep quality fades [4], thus increasing the quality of experience. Though there may be times in which the content of a majority of multiplexed TV channels demands a large bandwidth allocation, the essence of well-managed statistical multiplexing is that the duration of these intervals is short.

The H.264/Advanced Video Codec (AVC) [5] has provided an opportunity for countries that are contemplating digital video broadcasting, as it can be adopted in one step, rather than in the two-step process of earlier adopters that opted for the Moving Pictures Experts Group (MPEG)-2 codec. Other countries such as Portugal and Brazil have introduced H.264/AVC-based services. H.264/AVC can use the same MPEG-2 Transport Stream (TS) [6] as previously employed by the MPEG-2 codec itself. The MPEG-2 TS [7] can assemble up to 6 or 10 or even 20 independent television programs. Programs bitrates can be constant or variable. In the case of variable data rate, these rates can be controlled based on the requirements of the system prior to multiplexing (statistical multiplexing). The H.264/AVC codec significantly improves compression ratios [8] by as much as 50%, especially for SDTV. It is reported [9] that little prior research has been conducted on statistical multiplexing of H.264/AVC streams, even though this codec is now preferred for emerging national applications of HDTV and within wireless systems such as 3GPP's MBMS.

Assuming live or pre-encoded Constant Bitrate (CBR) video as used in the paper, the basic system is based on finding the number of non-Zero coefficients (NNZC) of the input sequences to the multiplex. (The scheme can be modified to work for Variable Bitrate (VBR) video.) These coefficients are those transform coefficients in the encoded bitstream that have not been reduced to zero by the quantization process. The NNZC allows the spatial complexity of the video sequence to be judged and by implication the coding complexity required to achieve a given quality. Notice that coding complexity is a measure of the coding bits required for compression and not the computational complexity. Specifically the NNZC of an individual macroblock was found to be logarithmically proportional to the coding complexity of that macroblock. Notice that because the system is intended for broadcast quality TV, the spatial coefficients dominate the bitstream. At low bit rates, the data given over to motion vectors and headers in the compressed bitstream [3] must be taken into account. As this is a no-reference system, a method is required to estimate the video quality (Peak Signal-to-Noise Ratio (PSNR)) based on knowledge of the NNZC. To do so required an estimate of the video quality from the average quantization parameter (QP), which, in a production system, can be extracted from the encoded bitstream without full decode.

Once the relative coding complexity is determined across the sequences, the video quality is equalized across the input sequences. This operation is performed at each Group-of-Pictures (GOP) boundary, though a refined version could also include scene change detection. (A GOP in broadcast TV normally consists of 12 or 15 pictures or frames (if progressively transmitted) corresponding respectively to about half a second at 12 or 15 frames/s.) The CBR rates are subsequently adjusted on a GOP-by-GOP basis to produce what has been called 'semi-CBR-VBR' streams [10]. Research in [11] also presents a CBR multiplex of streams previously stored at a high quality. In implemented systems, such as that from Scopus [12] for the MPEG-2 codec, VBR video can be smoothed [13] prior to complexity analysis. However, it is important to note that H.264/AVC video bitstreams have been found to be significantly more variable [14] than even MPEG-4 part 2 streams, due to the variety of coding modes available in H.264/AVC. It is also reported [14] that, after H.264/AVC frame size smoothing, the output remained significantly more variable than *unsmoothed* MPEG-4 part 2 output for the same films. CBR encoding allows planning of storage capacity and in video-on-demand schemes, it allows the bandwidth from a server to be tightly controlled. If the CBR video is not pre-encoded at a high rate (prior to transcoding) then image 'dissolves', fast 'action' and scenes with camera motion (pans, zooms, tilts, ...) all suffer. However, scenes with limited motion such as head-and-shoulder news sequences are not much affected by CBR encoding.

The envisaged system is shown in Fig. 1 in which a bitrate transcoder bank modifies the input after NNZC statistics have been extracted in the compressed domain. For an example of a commercial transcoder bank for a different purpose refer to [15]. In Fig. 1, the statistical multiplexor receives n compressed bitstreams which pass through a bank of bit-rate transcoders to adjust the combined bitrate according to the output channel constraint. The bandwidth share is defined by the statistical bandwidth manager which receives content complexity measures (*parameters*) from each transcoder and returns the appropriate bandwidth share (α). Modification of the bitrate in the compressed domain is known as dynamic rate shaping [16]. Frequency domain transcoding [17] has the advantage that latency is reduced by only requiring an entropic decode. In H.264/AVC, the Context Adaptive Variable Length (CAVLC) decode and bit-stream parsing on average take only 13% of the computational complexity of a full decode [18]. The complexity of the Context Adaptive Binary Arithmetic Coding (CABAC) option is estimated [19] to add approximately 12% of the CAVLC timing to the overall time (at a potential reduction in bitrate of up to 16%).

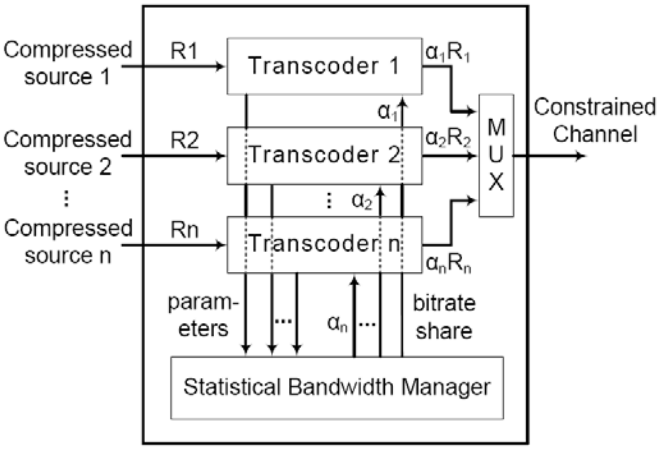

Fig. 1. Statistical multiplexor architecture.

The advantages of the proposed GOP-based dynamic rate control were found to be: a modest but worthwhile overall gain across the channels in video quality (PSNR); quality equalization across the video stream multiplex components; for the test example a stream quantity gain of about 1.5, i.e. more channels at the target PSNR of 40 dB; and a considerable reduction in bitrate burstiness of the video streams. It is known [20] that quality fluctuations have a significant impact on the subjective quality. The proposed scheme is based on empirically derived equations. Because the system is GOP-based it is probably easily integrated into GOP-based call-admission-control or bandwidth allocation systems for satellite channels [21], replacing H.264/AVC for MPEG-2 streams. In the DVB-RCS system, statistical multiplexing of aggregated user video traffic could additionally take place at the terrestrial hub, if the return channel were to be used for user video applications such as remote learning and telemedicine.

The remainder of this paper is organized as follows. Section 2 discusses related work on statistical multiplexing, covering related issues such as buffer management. Section 3 is an extended analysis of the methodology. This will be of interest to those wishing to construct their own statistical multiplexing system, based on their own likely video payload. Section 4 is an evaluation using a carousel of video sequences in each multiplexed stream. Finally, Section 5 draws some conclusions and makes suggestions for future work.

2. Related work

Statistical multiplexing techniques vary according to their complexity. In [22], a relatively simple form of statistical multiplexing was applied in which the same QP was applied to all video frames within a multiplexed group to achieve a target bit rate. A binary chop search across the range of available QPs was conducted. This procedure in the tests appeared to achieve its objective even though *no direct* account was taken of content complexity. In [11], a more computationally intensive scheme was applied in which statistical multiplexing based on spatial complexity statistics was applied to a set of rate-distortion controlled MPEG-2 video encoders. Their method is further discussed below but first we make a preliminary distinction.

There are two different ways for a statistical multiplexing algorithm to determine the bit rates of the encoders. Firstly, the feedback approach uses information from channel utilization and video coding complexity, which can be a by-product of the encoding procedure. The result of this process is a signal that is applied to the video encoders to determine their bit rates based on previous behavior. Secondly, the look-ahead approach uses the video statistics before encoding to find out each video's complexity and its assigned share of the bandwidth. It gains these statistics by pre-processing the future video frame. Then, with the information gathered from all videos, it calculates the amount of bandwidth for each of the video channels and applies the result to the encoders in order to set their bit rates. These techniques are applicable for the case of raw video as an input source.

Böröczky et al. [11] used the feedback approach in their joint rate control algorithm. They defined GOP boundaries or scene changes in a program as the point at which the MPEG-2 encoders bit rate was changed. This method provides two advantages. Firstly, during a GOP the encoder outputs at a constant bit rate. Therefore, the resulting bit streams are piecewise CBR. Secondly, this method does not need any preprocessing of the input video. The second advantage is important because processing raw video is computationally demanding and causes delay in the system. A limit was set to the bit rate changes during the same scene at GOP boundaries to prevent a noticeable change of quality in that scene. A control method was also developed to avoid buffer overflow and underflow. The method predefines two guard bounds at the top and bottom of the channel buffer. The information on channel buffer occupancy due to these guard-bands and the complexity of each input channels GOPs are jointly fed into the control algorithm to compute the bit rate of the individual encoders.

The process of obtaining statistics from video for determining the content coding complexity in the feedback approach is based on predicting the complexity from the previous frames' statistics. This can degrade the output video quality at scene changes, though some allowance for this problem can be made by using a sliding window GOP prediction method. As an alternative, He and Wu [23] used a lookahead approach in which the frame differences for each input video after encoding are found in order to estimate the variance of the number of DCT transform coefficients. These statistics in turn allow the rates for the input videos to be predicted for a desired constant distortion and a given encoder configuration. Processing is on a frame-by-frame basis but the average rate for any output video is then found for a given lookahead number of frames. For each output video, the predicted rate is normalized by the sum of all the other rates to determine the change to the encoding rate. Once a set of relative rates are found, these are then reduced step-by-step until encoder output buffer constraints are met. To avoid buffer overflow a similar two threshold system to that of [11] is used.

3 Methodology

3.1 Codec software modifications

Though a bitrate transcoder is under active development, for research purposes a decoder and encoder were used in back-to-back fashion, Fig. 2. In fact this is an example of the look-ahead method of statistical multiplexing. Again for research purposes, the H.264/AVC JM reference software was employed. The JM software (version 15.1) is written in the C programming language to ensure the fastest possible software-only processing time. In order to acquire the NNZC for each macroblock (MB) from the JM decoder, the decoder was modified as follows. The decoder obtains the coefficients from its input, arranges them in MB format, and then starts to decode each MB. The easiest method is to dump the NNZC is to obtain them prior to the decoding process of each MB. In the source codes of the decoder program in "image.c" inside "`void decode_one_slice`", the function "`read_one_macroblock(img, currSlice, currMB);`" constructs an MB coefficient matrix and stores it in "`img->cof`" for a 4×4 transform and in "`img->mb_rres`" for an 8×8 transform. After this line "`decode_one_macroblock(img, currMB, dec_picture);`" is executed which decodes the stored MB coefficient. Between these two lines is the place in which the additional code is added to dump the number coefficients and also to calculate and dump the total number of NNZC of a GOP.

In the experiments, the bitrate is changed during the encoding process. There is a feature in the encoder setting by the name of "ChannelType" which can be set to "time varying channel". By testing this feature it became clear that the CBR encoder changes the bitrate once by multiplying it by a constant factor during encoding. By exploiting this feature code was developed to oblige the encoder to change its bitrate at the start of each GOP.

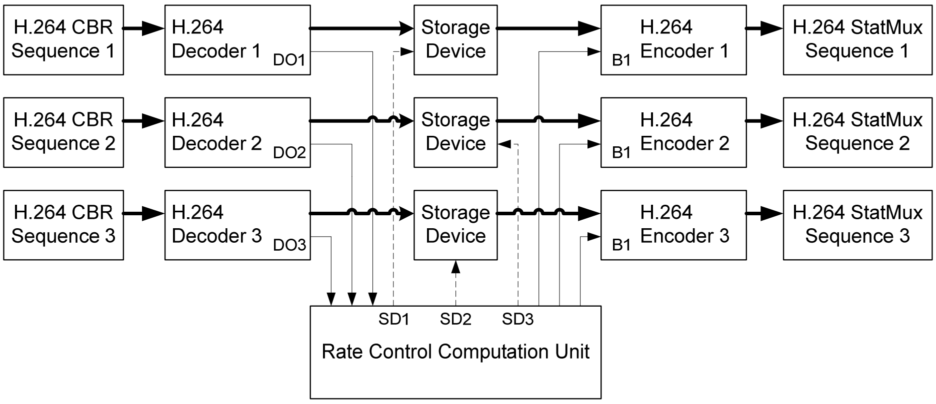

Fig. 2. Look-ahead implementation of statistical multiplexor.

For the sake of simplicity, Common Intermediate Format (CIF) resolution sequences (352 × 288 pixels/frame) test sequences were employed rather than SDTV. By 2006 there were already a variety of hardware H.264/AVC codecs available [24] though with a reduced selection of features. Blu-ray specifies H.264/AVC High Profile as one of its three formats and H.264/AVC High Profile is selected for the Memory Stick Video format.

3.2 Relationship between NNZC and PSNR

The first step toward implementation of a statistical multiplexor was to establish the relationship between PSNR and the NNZC. To estimate this relationship on a GOP basis, several sequences were encoded and decoded at different bitrates. The chosen sequences for this step were the well-known Mobile, Foreman, News, Stefan, Bus and Akiyo. The sequences were chosen based on their temporal and spatial complexity to represent different video characteristics. ‘Complexity’ in this context refers to the coding complexity involved in compressing the sequence. A more complex sequence requires more bits to compress to achieve a given video quality. Spatial complexity refers to the level of detail within each frame, whereas temporal complexity refers to the level of disparity between successive frames.

These sequences were encoded in CBR mode at 256, 512, 768 kbit/s, and 1, 1.5, 2, 2.5 and 3 Mbit/s. The PSNR of each frame were obtained from output of the encoder, and then the average PSNR of each GOP was calculated. (GOP structure IBBPBBP…, intra-refresh rate 12 frames.) The total NNZCs of a GOP, which was obtained during the decoding process, was paired with the average PSNR value of the corresponding GOP. Figure 3 plots the relationship between PSNR and NNZC per GOP for a selection of these sequences. The rates plotted are across the given bitrate ranges for all of the selected video sequences, except for Akiyo. For Akiyo, the results shown correspond to the range 512 kbit/s to 1 Mbit/s. At bitrates higher than this the QP of the encoder is reduced to its minimum level due to the simplicity of Akiyo sequence contents. Consequently, in later experimental results (though obviously not in Fig. 3) Akiyo was excluded. Other plots in Fig. 3 illustrate the differing coding complexity of the sequences, with plots lower on the vertical axis representing more complex sequences. For equal NNZC/GOP, it is reasonable and expected that quality decreases for 'complex' sequences. Equally, when NNZC increases, it is reasonable and expected that also the quality increases.

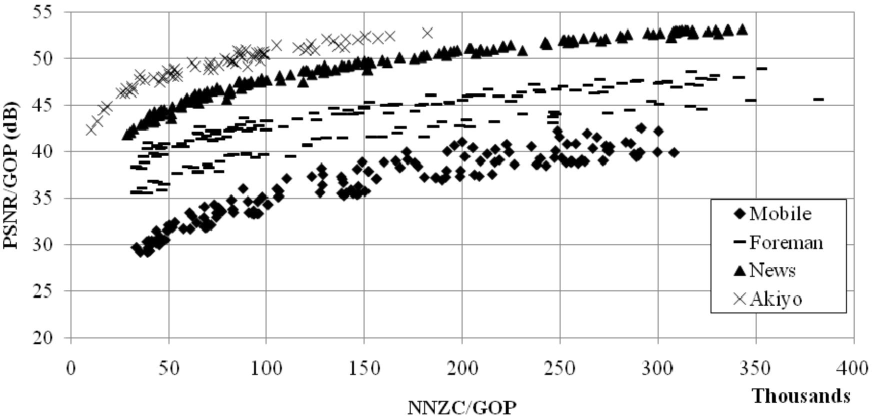

Fig. 3. NNZC per GOP versus average PSNR per GOP for selected video sequences.

An estimation equation based on these results was extracted as follows:

$$\mathrm{PSNR} = 4.8 \ln(\mathrm{NNZC}) + \mathrm{psnr}' \ (\mathrm{dB}) \quad (1)$$

Where psnr' represents the zero-crossing value of the fitted curve, which depends on the coding complexity of the sequence as a whole.

3.3 No-reference PSNR estimation

Obviously, psnr' cannot be found without reference to the input raw (YUV) video which is often not feasible. However, the QP is available from the encoded bitstream. The encoder in CBR mode allocates the bit budget for each frame in the sequence based on frame type, GOP size, GOP structure and the bitrate. The encoder assigns an initial QP and then starts the encoding process of the first frame. After the encoding of the first frame the number of bits used for this frame is observed. This number is compared with the bit budget which was allocated for this frame and as a result the QP is adjusted to maintain the allocated bitrate for the GOP. For a certain bitrate, a complex frame needs to be compressed more than a simple frame and consequently it receives a higher QP.

To find this relationship between PSNR and QP, the average PSNR of each GOP is paired with the corresponding QP. Figure 4 shows that relationship for a selection of the input sequences and bitrates previously described. This relationship can be estimated by a linear equation:

$$\text{PSNR} = (-0.89) \times \text{QP} + 60.93 \text{ (dB)}. \quad (2)$$

Equation (2) has a smaller margin of error for lower QP or equivalently higher bitrates. In the scenario envisaged in this paper, the input video streams are encoded at high bitrates to make the process of statistical multiplexing possible. Therefore, this estimate can be used with acceptable accuracy.

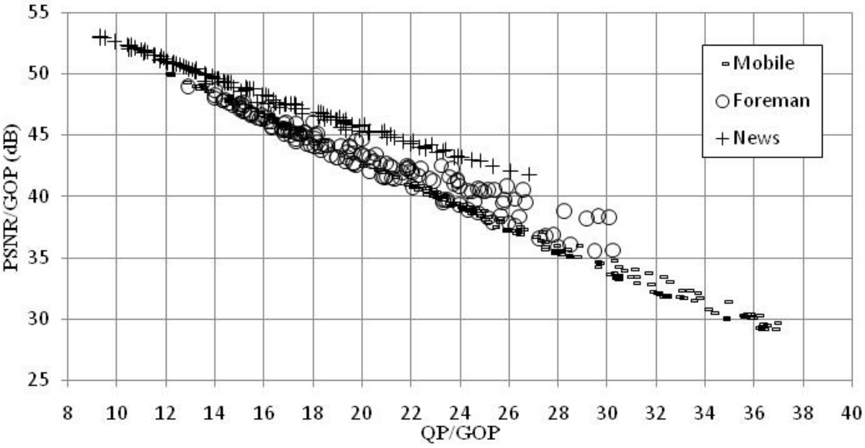

Fig. 4. QP per GOP versus average PSNR per GOP for selected video sequences.

At this point, with the estimated average PSNR (ePSNR), the total NNZCs for all GOPs so far (**t**NNZC) and the average QP (avgQP) of the decoded GOP, using (1) and (2), psnr' can be calculated as:

$$\text{ePSNR} = (-0.89) \times \text{avgQP} + 60.93 \text{ (dB)} \quad (3)$$

$$\text{psnr}' = \text{ePSNR} - 4.8 \times \ln(\mathbf{t}\text{NNZC}) \text{ (dB)} \quad (4)$$

and thus

$$\text{PSNR} = 4.8 \times \ln(\text{NNZC}) - 0.89 \times \text{avgQP} + 60.93 - 4.8 \ln(\mathbf{t}\text{NNZC}) \text{ (dB)}. \quad (5)$$

Equation (5) is the main tool of this scheme as it permits per GOP PSNR to be found.

3.4 Relating PSNR to CBR bitrate

In the encoding procedure of a codec operating in CBR mode, the best achievable way to assign the new bitrate for next encoded GOP is to oblige the encoder to change its bitrate. Therefore, (5) must be related to bitrate. After that, it becomes possible to allocate the bitrate based on the quality (PSNR) of a GOP. This requires that the relationship between NNZC and the number bits used per GOP is known. For this purpose, the total number of bit used per GOP (bits/GOP) is paired with the total NNZC for the corresponding GOP. The relationship is linear with a good accuracy at higher bitrates. Figures 5 and 6 show this relationship for three different GOPs of four different sequences (Mobile, Foreman, News and Stefan), encoded at 256 kbit/s to 3 Mbit/s. The GOP size or width does not disturb the linear relationship between bits/GOP and number of NNZC per GOP. No clear trend is discernible though for Mobile the results for a GOP size of 12 of separated from the other two GOP sizes.

By analyzing these data, the following equation (6) was found to be a reasonable estimation basis for conversion between NNZC and bits per GOP.

$$\text{NNZC} = 0.2 \times (\text{bits/GOP}) - 65536 \tag{6}$$

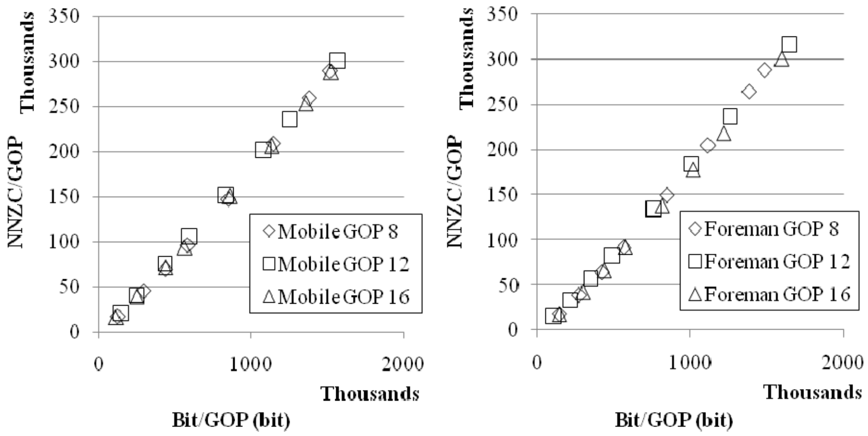

Fig. 5. NNZC per GOP versus bits/GOP for Mobile and Foreman at different GOP sizes.

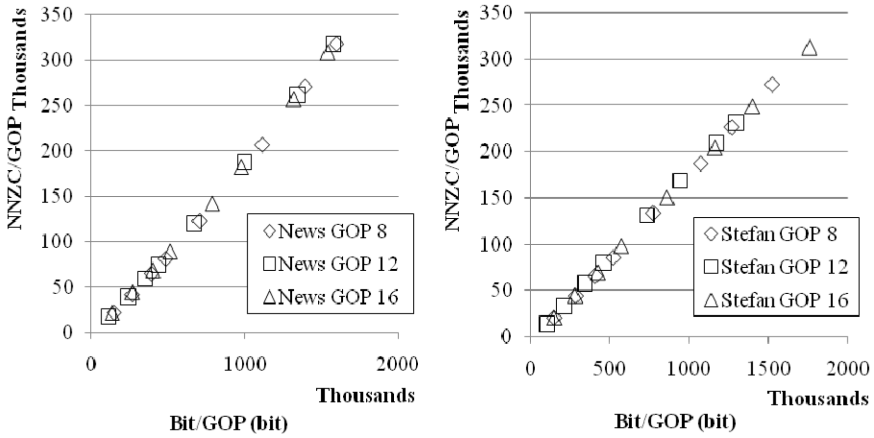

Fig. 6. NNZC per GOP versus bits/GOP for News and Stefan at different bitrates.

3.5 Calibration tests

In order to test the scheme, combined sequences of a set of the reference video sequences were CBR encoded at 3 Mbit/s. Each of three combined sequences contained 900 frames consisting of FNS (Foreman + News + Stefan), NMF (News + Mobile + Foreman) and WHB (Flower + Highway + Bus). In tests these H.264/AVC encoded combined sequences were the inputs of the statistical multiplexing system.

To apply the first step, all of the encoded sequences are decoded and the avgQP and tNNZC of each GOP was obtained. Processing continues on a per GOP basis. Therefore, the latency of the multiplexing system is one GOP or about 0.5 s. As is well known, live video is frequently delayed for a short time for the purposes of editing. Figures 7, 8 and 9 show the avgQP of each GOP for each of the FNS, NMF and WHB sequences. The ePSNR of each GOP is calculated through (3) by using avgQP.

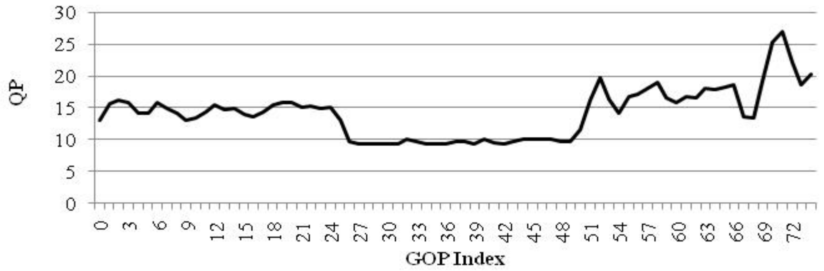

Fig. 7. Average QP per GOP for the FNS sequence set.

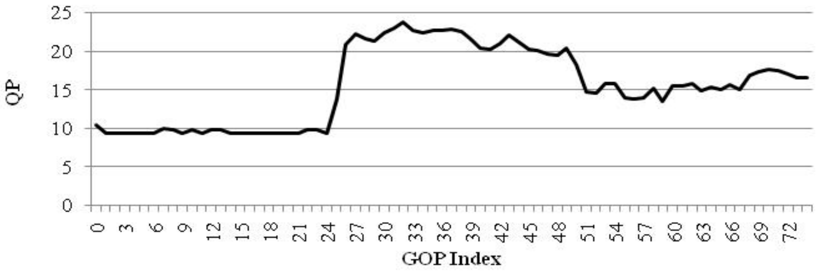

Fig. 8. Average QP per GOP for the NMF sequence set.

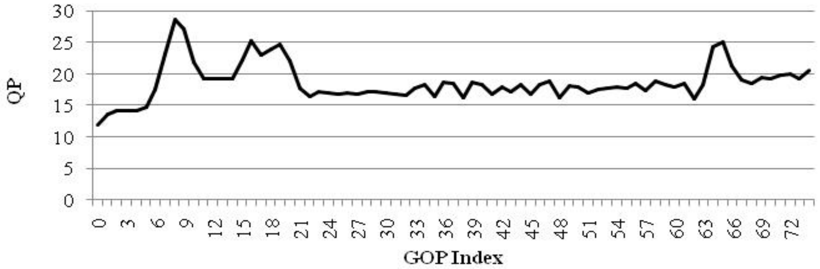

Fig. 9. Average QP per GOP for the WHB sequence set.

The result of the above calculation for all of the sequences and the correct PSNR value obtained from the encoder output are presented in Figures 10, 11 and 12. It will be apparent that there is a good match between the estimated PSNR and the calculated PSNR.

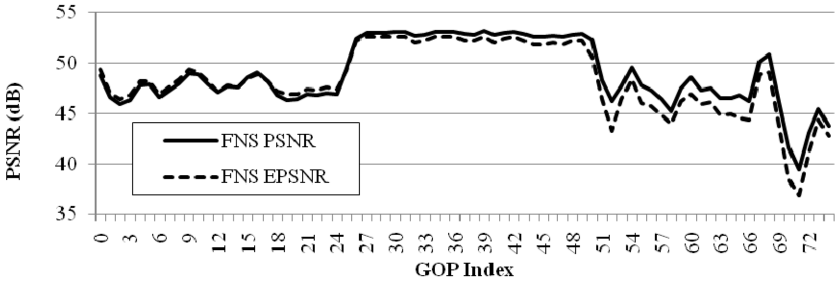

Fig. 10. Estimated PSNR (ePSNR) and calculated PSNR for the FNS combined sequence.

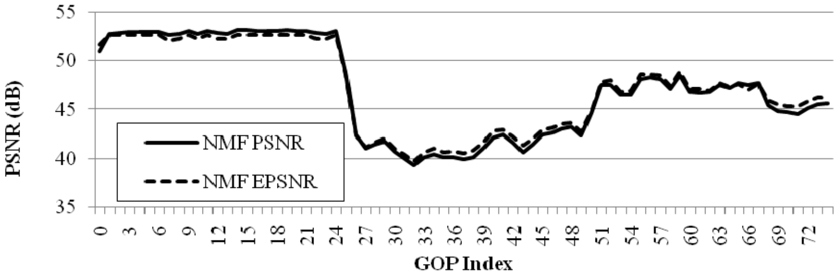

Fig. 11. Estimated PSNR (ePSNR) and calculated PSNR for the NMF combined sequence.

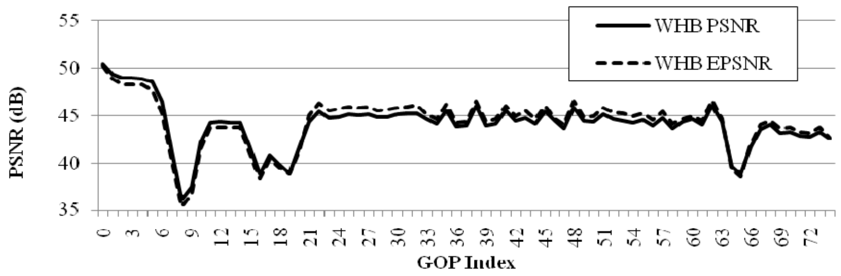

Fig. 12. Estimated PSNR (ePSNR) and calculated PSNR for the WHB combined sequence.

The estimated PSNR (ePSNR) is then used to calculate the psnr' or the constant part of eq. (1) for each GOP, by means of eq. (4). The total bitrate of the statistical multiplexing system is taken to be 3 Mbit/s for test purposes. Hence, the first condition for calculation of the bitrate allocation is defined as follows:

$$\text{Bitrate}_1(t) + \text{Bitrate}_2\,(t) + \text{Bitrate}_3(t) = 3 \text{ Mbit/s} \qquad (7)$$

where indices 1, 2 and 3, represent the FNS, NMF and WHB sequences over time t. Making use of (6), (7) can be converted to the following form, NNZC and per GOP (with GOP size = 12 frames and frame rate = 25 fps).

$$\text{NNZC}_1 + \text{NNZC}_2 + \text{NNZC}_3 = 0.2\times\ (3 \text{ Mbit/s} \times \frac{12}{25}) - 3 \times 65536 \qquad (8)$$

The second condition to satisfy a statistical multiplexing system is to normalize the overall quality, which happens if at each sampling time the quality of all of the streams becomes equal. This fact defines the second condition of this process, which is:

$$\text{PSNR}_1(\text{GOP}) = \text{PSNR}_2(\text{GOP}) = \text{PSNR}_3(\text{GOP}). \qquad (9)$$

By substituting (1) into (9) and equating, equations (10) and (11) are easily found.

$$NNZC_2 = NNZC_1 \times e^{\frac{psnr_1' - psnr_2'}{4.8}} \qquad (10)$$

$$NNZC_3 = NNZC_1 \times e^{\frac{psnr_1' - psnr_3'}{4.8}} \qquad (11)$$

Subsequently, substituting (10) and (11) into (8) results in:

$$NNZC_1 = \frac{105382}{1 + e^{\frac{psnr_1' - psnr_2'}{4.8}} + e^{\frac{psnr_1' - psnr_3'}{4.8}}} \qquad (12)$$

In the final calculation step, by using (6) again $\text{Bitrate}_{1\text{-}3}$ are calculated at the start of each GOP. These bitrates are the input of the second stage of this system which consists of the final set of encoders. The final encoders apply these bitrates through the method described in Section 3.1

4 Evaluation

The proposed scheme was evaluated through comparison of the system output quality with conventional equal allocation of bitrate multiplexing. Figures 13, 14 and 15 illustrate the average PSNR of each GOP for the three experimental sequences, according to whether the proposed scheme is applied (StatMux) or whether the available 3 Mbit/s bandwidth is allocated at 1 Mbit/s per video sequence. It is immediately clear that in Fig. 14, the constant bitrate allocation results in a deep quality fade. All of the three encoded sequences at 1 Mbit/s, contain at least one GOP with average PSNR below 30 dB. Therefore, the CBR process results are not suitable for broadcasting because of the noticeable quality degradation. On the other hand, the minimum average GOP PSNR for the statistical multiplexing scheme was about 34 dB, which is an acceptable quality.

The average PSNR over all frames was calculated for both statistical multiplexing and constant CBR mode, and demonstrated in Table 1. The overall quality difference is 0.42 (dB) gained by the proposed statistical multiplexing scheme. The standard deviation of the PSNR is also calculated as a measure of variation of the PSNR around its average. The PSNR standard deviation of the statistical multiplexing scheme is lower by a considerable extent. This result illustrates the success of the scheme in equalizing the quality across the three sequences. Because of the quality fade for constant CBR (Fig. 14 and drops below 30 dB) and because of the low overall quality of the WHB sequence (Table 1), it would be necessary to send just two sequences at a rate of 1.5 Mbit/s each for fixed rate transport. Therefore, for this example, the number of programs that can be sent at 40 dB has been increased by a factor of 1.5.

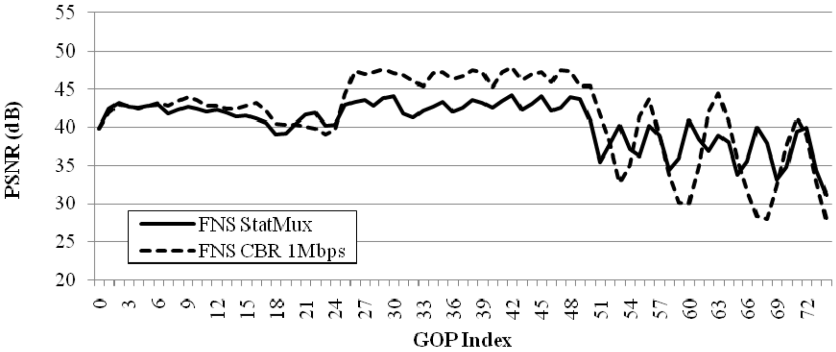

Fig. 13. Average PSNR/GOP for constant allocation of bandwidth and the proposed statistical multiplexing scheme on the FNS combined sequence.

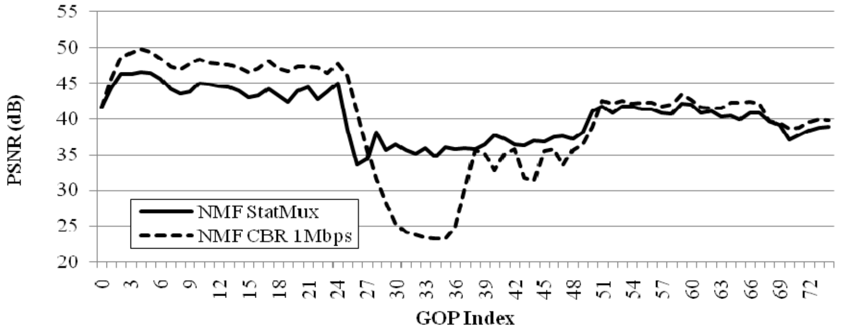

Fig. 14. Average PSNR/GOP for constant allocation of bandwidth and the proposed statistical multiplexing scheme on the NMF combined sequence.

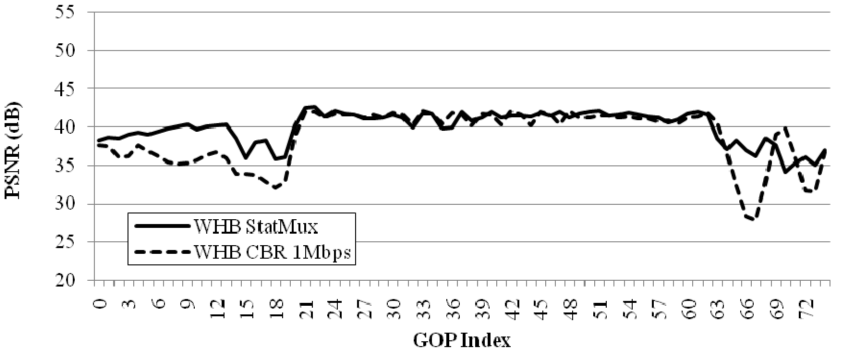

Fig. 15. Average PSNR/GOP for constant allocation of bandwidth and the proposed statistical multiplexing scheme on the WHB combined sequence.

Table 1. Per sequence average PSNR and standard deviation of the sequences.

	Statistical multiplexing		*CBR*	
Sequence	Average PSNR (dB)	PSNR standard deviation (dB)	Average PSNR (dB)	PSNR standard deviation (dB)
FNS	40.57	3.00	41.70	5.38
NMF	40.39	3.47	40.21	7.39
WHB	40.04	2.10	38.68	3.68

A further gain from using the proposed scheme was found to be a reduction in bitrate fluctuations arising from the JM codec software behavior in CBR mode. The codec selects I- and B-frame QPs intelligently to avoid too rapid transitions in quality. Unfortunately, if at scene changes the new scene is more complex, then a large fluctuation in bitrate occurs. It was observed that the proposed scheme results in smoother resulting bitrate fluctuations than when requests are made for a uniformly constant bitrate. The total overall bitrate allocation (rather than requested bitrate) is shown in Fig. 16, showing that the fixed rate allocation results in larger bitrate excursions. It is clear that rate adjustment in combination with buffering is required to ensure that the instantaneous allocation does not exceed the capacity of the satellite channel.

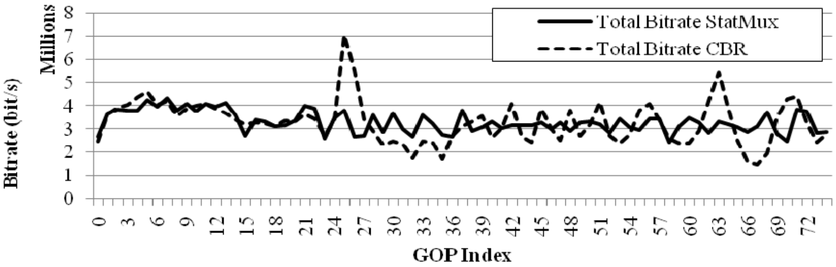

Fig. 16. Allocated bitrate/GOP for constant allocation of bandwidth and the proposed statistical multiplexing scheme for the overall multiplex.

5. Conclusion

The H.264/AVC results in approximately a doubling in compression efficiency relative to MPEG-2. The resulting bitstream is easily accommodated in an MPEG-2 TS. In this paper, we have demonstrated a statistical multiplexing scheme based on spatial complexity modeling equations parameterized by the per-GOP NNZC. For 40 dB PSNR video streams, the scheme significantly reduces PSNR fluctuations. True VBR video may result in too frequent bitrate oscillations and in this paper, a GOP-by-GOP semi-VBR scheme has been investigated. This implies that the latency of the scheme is the time taken to process a GOP, normally about 0.5 s. Aggregated user video inputs may occur in the DVB-RCS system, with the video bitrates close to those described in this paper. Therefore, this is one application of the proposed scheme. Future work should account for the impact of scene changes. The JM software implementation of H.264/AVC operating in CBR mode avoids rapid changes in quality by modifying the output bitrate. The proposed scheme additionally counters the bitrate fluctuations that occur due to the behavior of the codec at scene changes.

References

1. Neal, J., Green, R. and Landovskis, J.: Interactive channel for multimedia satellite networks. *IEEE Commun.Mag.*, vol. 39, pp. 192–198, (Mar. 2001).
2. Tandberg Television Ltd: Reflex and data Reflex statistical multiplexing system. White Paper, Slough, UK (2008)
3. Seeling, P., Reisslein, M.: The rate variability-distortion (VD) curve of encoded video and its impact on statistical multiplexing. IEEE Trans. Broadcast., vol. 51, no. 4, pp. 473–492 (2005)
4. Kuhn, M., Antkowiak, J.: Statistical multiplex what does it mean for DVB-T? FKT Fachzeitschrift für Fernsehen, Film und elektronische Medien (2000)

5. Wiegand, T., Sullivan, G. J., Bjøntegaard, G., Luthra, A.: Overview of the H.264/AVC video coding standard. IEEE Trans. Circuits Syst. Video Technol., vol. 13, no. 7, pp. 560–576, (2003)
6. Wenger, S.: H.264 over IP. IEEE Trans. Circuits Syst. Video Technol., vol. 13, no. 7, pp. 645–656, (2003)
7. Fischer, W., Digital video and audio broadcasting technology, (2nd ed.) Springer-Verlag, Berlin, 2008.
8. Ghanbari, M., Fleury, M., Khan, E. et al. Future Performance of Video Codecs, research report for Office of Communications (Ofcom), London, November, 2006
9. Vukadinovic, V., Huschke, J.: Statistical multiplexing gains of H.264/AVC video in E-MBMS. In: 3rd Int. Symp. on Wireless Pervasive Computing, pp. 468–474 (2008)
10. Ghanbari, M., Azari, J.: "Effect of bit rate variation of the base layer on the performance of two-layer video codecs. IEEE Trans. Circuits Syst. Video Technol., vol. 4, no. 1, pp. 8–17, (1994)
11. Böröczy, L, Ngai, A. Y., Westermann, E. F.: Statistical multiplexing using MPEG-2 video encoders. IBM J. of Research and Development, vol. 43, no. 4, pp. 511–520 (1999)
12. Scopus Video Networking: Advanced encoding mechanism and statistical multiplexing. White Paper, Sunnyvale, CA (2006)
13. Zhang, Z.-L., Kurose, J., Salehi, J. D, Townsley, D.: Smoothing, statistical multiplexing and call admission control for stored video. IEEE J. Sel. Areas Communs,, vol. 15, no. 6, pp. 1148 – 1166, (Aug. 1997)
14. van der Auwera, G., David, P. T., Reisslein, M.: Traffic and quality characterization of single-layer video streams encoded with the H.264/MPEG-4 Advanced Video Coding standard and Scalable Video Coding extension. IEEE Trans. Broadcast., vol. 54, no. 3, pp. 698-718, (2008)
15. Kasai, K., Nilsson, M., Jebb, T., Whybray, M., Tominaga, H.: The development of a multimedia transcoding system for mobile access to video conferencing. IEICE Trans. on Communs., vol. 10, no. 2, pp. 2171–2181, (Oct. 2002)
16. Eleftheriadis, A., Batra, P.: Dynamic rate shaping of compressed digital video. IEEE Trans. Multimedia, vol. 8, no. 9, pp. 297–314, 2006.
17. Assunção, P. A., Ghanbari, M.: A frequency-domain video transcoder for dynamic bit-rate reduction of MPEG-2 bit streams. IEEE Trans. Circuits Syst. Video Technol., vol. 8, no. 8, pp. 953–967, (1998)
18. Malvar, H., Hallapuro, A., Karczewicz, M.,Kerofsky, L.: Low complexity transform and quantization in H.264/AVC. IEEE Trans. Circuits Syst. Video Technol., vol. 13, no. 7, pp. 598–603, (July 2003)
19. Saponara, S., Blanch, C., Denolf, K., Bormans, J.: The JVT advanced video coding standard: Complexity and performance analysis on a tool-by-tool basis. Int'l Packet Video Workshop, Nantes, France, (April 2003)
20. Ghinea, G. and Thomas, J. P.: QoS impact on user perception and understanding of multimedia video clips. ACM Int'l Conf. on Multimedia, pp. 49-54, (1998)
21. De Rango, F., Tropea, M., Fazio, P., and Marano, S.: Call admission control for aggregate MPEG-2 traffic over multimedia geo-satellite networks. IEEE Trans. Broadcast., vol. 54, no. 3, pp. 612-622, (Sept. 2008)
22. Wang, L., Vincent, A.: Joint rate control for multi-program video coding, IEEE Trans. Consumer Electron., vol. 42, no. 3, pp. 300–305, (1996)
23. He, Z., Wu, D.O., Linear rate control and optimum statistical multiplexing for H.264 video broadcast. IEEE Trans. Multimedia, vol. 10, no. 7, pp. 1237–1249, (2008)
24. Marpe, D., Wiegand, T., Sullivan, G. J.: The H.264 / MPEG4 Advanced Video Coding standard and its applications. IEEE Commun. Mag., vol. 44, no. 8, pp. 134-144, (Aug. 2006)