CONSTANT BITRATE CONTROL FOR A DISTRIBUTED VIDEO CODING SYSTEM

Mariusz Jakubowski[1], João Ascenso[2] and Grzegorz Pastuszak[1]
[1]Institute of Radioelectronics, Warsaw University of Technology, 15/19 Nowowiejska Str., Warsaw, Poland

[2]Instituto Superior de Engenharia de Lisboa – Instituto de Telecomunicações
R. Conselheiro Emídio Navarro, 1, Lisbon, Portugal

Keywords: Wyner-Ziv coding, distributed video coding, rate control.

Abstract: In some distributed video coding (DVC) systems, the total bitrate depends mainly on the key frames (Intra coded) quality and on the side information accuracy. In this paper, a rate control (RC) mechanism is proposed to achieve and maintain a certain target bitrate for the overall Intra and WZ bitstream, mainly by adjusting online the Intra frames quality through the quantization parameter (QP). In order to obtain a similar decoded quality of Intra and WZ frames, the relevant parameters: QP for the key frames and the quantization index (Q_{Index}) for WZ frames are controlled jointly. The major novelty of this work is a statistical model that expresses the relationship between Q_{Index} and WZ frames bitrate. The proposed rate control solution is integrated into the VISNET2 WZ codec and the experimental results demonstrate the efficiency of the proposed algorithm to reach and maintain the target bitrate.

1 INTRODUCTION

Around 2002, a new video coding paradigm known as distributed video coding (DVC) has emerged, inspired by two Information Theory results from the 70's: the Slepian-Wolf theorem (Slepian and Wolf, 1973) and the Wyner-Ziv theorem (Wyner and Ziv, 1976). The main advantage of DVC lies in emergent application scenarios such as wireless video surveillance, low-power video sensor networks and mobile camera phones. In such applications, there are strong requirements in terms of low encoding complexity or a more balanced complexity distribution between the encoder and decoder. Improved error resilience is also a desired feature since most of the considered channels are quite noisy (e.g. wireless channels). In such scenarios, DVC fits well, since it explores the video statistics, partially or totally, at the decoder not at the encoder side, as in traditional video coding solutions, e.g. in MPEG-x and H.26x standards. In DVC, one of the most interesting cases is the source coding of a source *X*, while a source *Y*, known as side information, is available at the decoder only. Wyner and Ziv showed that for lossy coding under certain conditions (Wyner and Ziv, 1976), there is no loss of coding efficiency if the dependency between *X* and *Y* is explored at the decoder with reference to the case where joint encoding is performed (i.e. *X* and *Y* are available at the encoder). This interesting result opens the possibility to design a system where two statistically dependent signals are compressed in a distributed way (separate encoding, joint decoding) while still achieving the coding efficiency of conventional predictive coding schemes (joint encoding and decoding). However, practical DVC codecs did not yet achieve this target performance, especially when low complexity encoding is a major requirement.

One of the most interesting and used DVC architectures is based on turbo codes and a feeback channel (FC) to perform rate control at the decoder. The feedback channel has a key role, since the decoder, knowing the available side information, can test for successful decoding (i.e. if most of the errors were corrected) and ask for the necessary bitrate to achieve a certain target quality (established by the encoder). Actually, in this solution there is no bitrate control. A certain quality is established by the encoder and the decoder just spends the necessary rate to achieve it.

However, when the video transmission occurs in constant bandwidth or bandwidth limited channels, it is necessary to have a fixed target encoding bitrate

Jakubowski M., Ascenso J. and Pastuszak G. (2008).
CONSTANT BITRATE CONTROL FOR A DISTRIBUTED VIDEO CODING SYSTEM.
In *Proceedings of the International Conference on Signal Processing and Multimedia Applications*, pages 131-138
DOI: 10.5220/0001937501310138

for the whole transmission. In this case, the encoder must allocate the bitrate among each coding unit (e.g. frame) and control the encoder parameters, i.e. adjust the quantization parameter, in order to spend the allocated bits efficiently.

In this context, this paper presents an encoder rate control technique which achieves a constant bitrate while minimizing changes in the quality of the decoded sequence.

2 VISNET2 WZ VIDEO CODEC

The overall Wyner-Ziv (WZ) coding architecture for the VISNET2 video codec is illustrated in Figure 1. This codec follows the architecture proposed in (Brites *et. al.*, 2006), except for the encoder rate control module which is proposed in this paper.

The coding process starts by the division of the video frames into key frames and Wyner-Ziv (WZ) frames. Then, one or two key frames are encoded using the H.264/AVC Intra mode (Wiegand *et. al.*, 2003) in order to guarantee that each GOP is delimited by key frames. The quality and thus the rate of each key frame is defined mainly by the quantization parameter (QP).

The frames in between are WZ frames, which are simply coded with a H.264/AVC 4×4 block-based discrete cosine transform (DCT) followed by the aggregation of DCT coefficients in 16 frequency bands b_k. Each band is uniformly quantized and bitplanes are created and sent to the turbo encoder. The encoder establishes the final decoded quality by defining for each band b_k the respective number of bitplanes M_k for which WZ bits are generated, i.e. the amount of bitplanes that will have a small error probability after turbo decoding. There are 8 4×4 quantization matrices (Brites *et. al.*, 2006), which define different M_k values for each DCT band b_k allowing to achieve different rate-distortion (RD) performances. The quantization matrices used by both encoder and decoder are defined by the Q_{Index} parameter.

At the decoder, for each WZ frame, the side information Y_i, an estimate of the X_i frame, is created by motion compensated interpolation (MCI) based on two references, one temporally in the past and another in the future (for GOP = 2 the references correspond to the key frames). Then, the DCT transform is applied to the side information and, with a Laplacian correlation model, soft-input information is obtained for the turbo decoder. The iterative turbo decoder uses the received parity bits and the soft-input side information and attempts to generate the decoded (with small error probability P_e) quantized symbol stream. If the decoding is not successful ($P_e > 10^{-3}$) the decoder requests via the feedback channel for more parity bits, until successful decoding ($P_e < 10^{-3}$) is achieved. Each bitplane of each band is turbo decoded starting from the most significant biplane and the DC coefficient band. A zig-zag scan order is followed for the DCT bands. After turbo decoding all bitplanes of all DCT bands for which WZ bits were sent, the quantized symbol stream is obtained. Next, in the reconstruction module, the side information is used together with the decoded quantized symbol stream, to obtain the decoded X_i frame after the IDCT transform.

Finally, the key frames and WZ coded frames are mixed again to generate the decoded video sequence with a quality defined by the QP (for key frames) and Q_{Index} (for WZ frames) encoding quantization parameters. The bitrate is spent according to the side information quality, i.e. the accuracy of the MCI estimation.

A novel encoder rate control module is proposed in this paper (see Figure 2) which needs as an input the bits spent on the WZ and key frames and allocates the available bitrate among the WZ and Intra key frames by changing the QP (for key frames) and Q_{Index} (for WZ frames) according to the rate control algorithm proposed in the next Section.

3 PROPOSED RC ALGORITHM

In Figure 2, the flowchart of the proposed rate control algorithm is presented. For each GOP of the sequence, the WZ encoder is run with some initial value of Q_{Index} and generates the parity bits. Next, key frames are encoded with a new QP value which is selected based on the bitrate of the previous GOP and the predicted bitrate in the current and next GOP. The WZ decoder, invoked in the next step, uses these key frames and the parity bits produced by the WZ encoder. In the last step, a new value of Q_{Index} parameter is selected according to the QP value in order to obtain similar WZ and Intra frames quality. The procedure is repeated up to the last GOP.

In the next subsections, a detailed description of QP and Q_{Index} selection procedures is given.

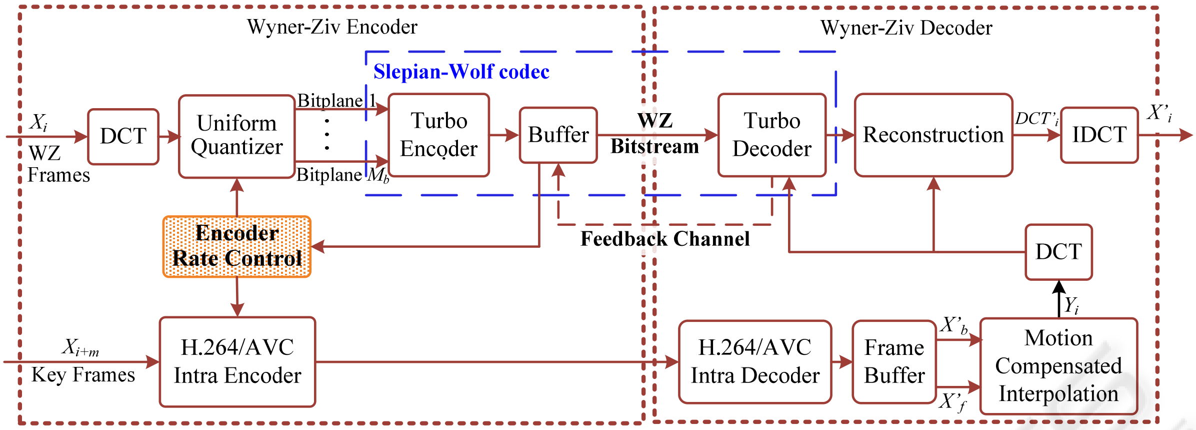


Figure 1: VISNET2 WZ Video Codec Architecture.

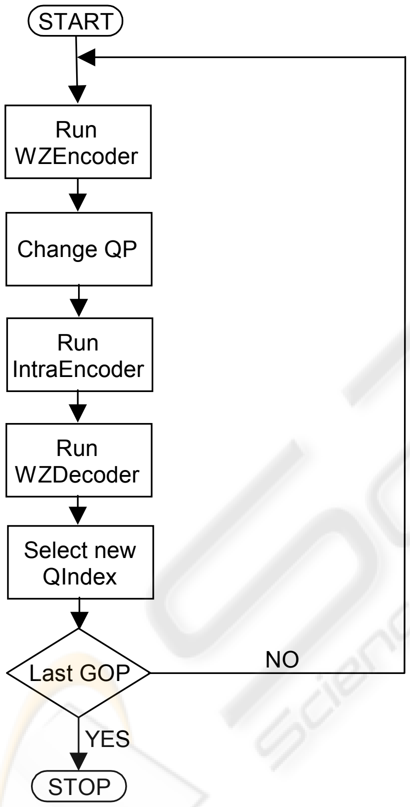


Figure 2: Rate control algorithm flowchart.

3.1 QP Selection Procedure

The QP selection procedure is a key element of the algorithm. With the QP value is possible to control the bitrate and quality of the Intra frames and, indirectly, of the WZ frames, since the side information (critical in the WZ codec RD performance) is created based on the Intra frames. The selection algorithm takes into account the bitrate of the previous GOP and a predicted increase or decrease of the WZ frames bitrate caused by a possible change in Q_{Index}. If the previous bitrate is greater or smaller than the target one, QP should be modified. It is known that in H.264/AVC a change of QP by 1 corresponds to a change in bitrate of approximately 12% (change of QP by 6 means that bitrate is halved or doubled) (Wiegand *et al.*, 2003). According to this rule, the relationship between the bitrate of Intra frames and the QP parameter can be expressed as:

$$R_{I1} = R_{I0} * 2^{-\frac{QP_1 - QP_0}{6}} \tag{1}$$

where R_{I0}, R_{I1} are the previous and the predicted Intra frames bitrate, respectively. QP_0 and QP_1 are the previous and predicted QP, respectively. First, the target bitrate of Intra frames in the current GOP, R_{IT}, is estimated by

$$R_{IT} = \frac{R_T * IP}{FR} - R_{WZ1} \tag{2}$$

where R_T is the target bitrate, IP is the period of Intra frames in an encoded sequence, FR is the frame rate, and R_{WZ1} is the predicted bitrate of WZ frames in the current GOP. To obtain a good estimation of R_{WZ}, a couple of experiments were performed to study the influence of the Q_{Index} and QP parameters on the overall WZ rate. In Figure 3, the average bitrate of WZ frames according to the Q_{Index} for the Foreman QCIF sequence with fixed QP equal to 24 is shown. In Figure 4, the average bitrate of WZ frames for the same sequence according to the QP parameter, maintaining fixed Q_{Index} equal to 4, is presented. It can be seen that from QP = 0 up to about 30, the bitrate of WZ frames is almost constant and starts to increase faster approximately from QP = 36. An important conclusion follows from this study: if Q_{Index} does not change, the bitrate of WZ frames in the current GOP will remain similar to the previous one. If Q_{Index} changes, the bitrate of WZ frames in the current GOP will change , but in a more difficult way to predict. As seen in Figure 3, when Q_{Index} is incremented from 7 to 8 the bitrate is almost doubled, but from 4 to 5 only an increase of 10% is observed.

Several experiments were performed on the Coastguard, Foreman, Hall Monitor, and Soccer QCIF sequences in order to obtain a model which describes the dependence of the bitrate of WZ frames on Q_{Index} when it changes from one value (Q_{I0}) to another (Q_{I1}). This model can be presented in the form of transition table (*TT*) which models the WZ frames bitrate for different Q_{I0} and Q_{I1} pairs. Each element of the table is described by the following equation:

$$TT[Q_{I0}][Q_{I1}] = \frac{R_{WZ}[Q_{I1}] - R_{WZ}[Q_{I0}]}{R_{WZ}[Q_{I0}]} \quad (3)$$

where $R_{WZ}[Q_{I0}]$ and $R_{WZ}[Q_{I1}]$ are the bitrates of WZ frames at two different quantization indexes. For each coded sequence it is possible to evaluate (3) and obtain a different table. Table 1 and Table 2 show the model obtained for the Soccer and Coastguard sequences. The strategy proposed here to cope with this inter sequence variation is to use the transition table with average values from the four previously defined sequences, and update it with actual values during the coding process. These initial values are shown in Table 3. Now, the R_{WZ1} term in (2) can be expressed as

$$R_{WZ1} = R_{WZ0} + R_{WZ0} * TT[Q_{I0}][Q_{I1}] \quad (4)$$

where R_{WZ0} is the previous bitrate of WZ frames, Q_{I0} and Q_{I1} are the previous and current Q_{Index} values which are used to index the transition table. When R_{IT} is estimated, QP_1 can be estimated by substituting R_{I1} by R_{IT} in (1)

$$QP_1 = QP_0 - 6 * \log_2 \frac{R_{IT}}{R_{I0}} \quad (5)$$

However, in order to ensure a fast approach to the target bitrate set at the encoder and to maintain relatively smooth changes in quality, some additional constraints are established for the QP variation. First of all, QP can change freely between 0 and 51 only in the second GOP. The first GOP is coded with some initial QP value and the outcome is unknown. It can be far above or far below the target bitrate, so it is necessary to get close to the target as fast as possible. However, a rapid change of QP can cause a temporary peak of bitrate within one GOP and in reaction, QP in the next GOP would have to change strongly again in the opposite direction. The result would be rapid changes of PSNR between GOPs which causes a flickering artefact, i.e. a negative subjective impact in video quality. To avoid such an instability, it is proposed to limit the QP variation between consecutive GOPs: QP can be increased by five and decreased by three at maximum. It was found that the ability to decrease the bitrate is more critical than increasing. Additionally, the algorithm tries to predict the consequences of decreasing the QP in the next GOP.

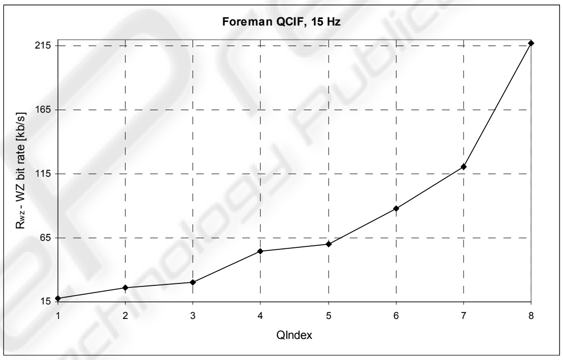


Figure 3: Bitrate characterization in WZ frames for different Q_{Index} values.

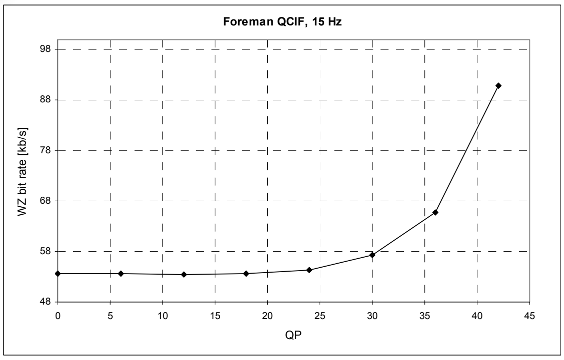


Figure 4: Bitrate characterization in WZ frames for different QP values.

For example, if QP variation causes a change of Q_{Index} in the next GOP which will lead to a large increase of the WZ frames bitrate, the algorithm will prevent such a situation and reduce the amount of QP decrease. This approach allows a relatively stable quality (PSNR) and the bitrate close to the target after a few first GOPs.

Table 1: Bitrate transition table of WZ frames for the Soccer sequence.

QI0 \ QI1	1	2	3	4	5	6	7	8
1	0	0.25	0.48	1.77	1.69	2.60	3.72	6.79
2	-0.20	0	0.18	0.99	1.15	1.87	2.77	5.22
3	-0.32	-0.15	0	0.69	0.82	1.43	2.19	4.26
4	-0.60	-0.50	-0.41	0	0.08	0.44	0.89	2.12
5	-0.63	-0.53	-0.45	-0.07	0	0.34	0.75	1.90
6	-0.72	-0.65	-0.59	-0.30	-0.25	0	0.31	1.17
7	-0.79	-0.73	-0.68	-0.47	-0.43	-0.24	0	0.65
8	-0.87	-0.84	-0.81	-0.68	-0.65	-0.54	-0.39	0

Table 2: Bitrate transition table of WZ frames for the Coastguard sequence.

QI0 \ QI1	1	2	3	4	5	6	7	8
1	0	0.47	0.75	1.87	2.16	3.83	5.86	13.68
2	-0.32	0	0.19	0.94	1.15	2.28	3.65	8.96
3	-0.43	-0.16	0	0.63	0.80	1.76	2.91	7.37
4	-0.65	-0.49	-0.39	0	0.10	0.69	1.39	4.12
5	-0.68	-0.53	-0.45	-0.09	0	0.53	1.17	3.64
6	-0.79	-0.69	-0.64	-0.41	-0.35	0	0.42	2.04
7	-0.85	-0.78	-0.74	-0.58	-0.54	-0.29	0	1.14
8	-0.93	-0.90	-0.88	-0.80	-0.78	-0.67	-0.53	0

Table 3: Averaged transition table.

QI0 \ QI1	1	2	3	4	5	6	7	8
1	0	0.44	0.72	1.82	2.11	3.54	5.19	10.75
2	-0.30	0	0.19	0.96	1.16	2.14	3.28	7.09
3	-0.41	-0.16	0	0.64	0.81	1.64	2.59	5.79
4	-0.64	-0.49	-0.39	0	0.10	0.60	1.18	3.14
5	-0.68	-0.54	-0.45	-0.09	0	0.45	0.98	2.74
6	-0.78	-0.68	-0.62	-0.37	-0.31	0	0.36	1.56
7	-0.83	-0.76	-0.72	-0.54	-0.49	-0.26	0	0.88
8	-0.91	-0.87	-0.85	-0.75	-0.72	-0.60	-0.46	0

3.2 Q_{Index} Selection Procedure

The Q_{Index} parameter is selected according to the QP value in order to maintain a similar quality of WZ and Intra frames, reducing the flickering effect which is quite important from the subjective point of view. In (Areia *et al*, 2008), Q_{Index} for four sequences (Coastguard, Foreman, Hall Monitor, and Soccer) are matched with QP parameters which give similar Intra frames quality. These values are collected in Table 4. Because an appropriate model which relates these two parameters is difficult to find, it is proposed to use the QP average values for these four sequences (Table 4) and select the Q_{Index} which is matched with the QP equal or smaller than the current QP value. For example, if current QP equals 38, the closest value in Table 4 is 39, which means that Q_{Index} 3 is selected.

Table 4: Points of equivalent quality for the key frames and WZ frames. C. – Coastguard, F. – Foreman, H. M. – Hall Monitor, S. – Soccer.

Q_{Index}	C.	F.	H. M.	S.	**Avg.**
1	39	42	37	45	**41**
2	38	40	36	44	**40**
3	38	39	35	42	**39**
4	35	36	33	38	**36**
5	34	35	32	38	**35**
6	33	33	31	35	**33**
7	31	31	29	31	**31**
8	27	26	25	26	**26**

4 EXPERIMENTAL RESULTS

The proposed algorithm was integrated into the VISNET2 DVC codec, as shown in Section 2. For the experiments, six QCIF format sequences were used. Four of them had already been used in previous experiments: Coastguard, Foreman, Hall Monitor, and Soccer. Two additional sequences, Paris and Stefan, were taken to verify if the used parameters are not sequence dependent. The test conditions are shown in Table 5.

Table 5: Test conditions.

Sequences	Coastguard, Foreman, Hall Monitor, Paris, Soccer, Stefan
Intra Period	2
Domain	Transform
Initial Q_{Index}/QP	3/39
Key Frames Codec	H.264
Frame Rate	15
Target Bitrate [kb/s]	250

In Figure 5, the resulting total bitrate for each GOP and the PSNR for each frame is shown for all test sequences. The initial Q_{Index}/QP value is set at 3/39 for all sequences. For most of them, it gives the initial bitrate far below the target one. It makes possible to demonstrate the capability to reach fast the target bitrate and maintain it fixed as desired. It can be seen that in most cases the bitrate approaches the target one after a few first GOPs. Except for Hall Monitor, after the five first GOPs the difference is not greater than 20%. A slow increase of the bitrate in case of the Hall Monitor sequence is a price for a relatively stable PSNR and a result of a lack of an accurate model to express the relationship of the QP and Q_{Index} parameters. The algorithm does not decrease the QP parameter further because it would change the Q_{Index} value, in such a way that a rapid increase of the WZ frames bitrate would lead to an overall bitrate overflow. The bitrate overflow for the Coastguard sequence is visibly correlated with a rapid decrease of PSNR at the frame number 38. This frame is very blurred and causes an unexpected increase of the WZ bitrate without the change of Q_{Index}, together with a large decrease of PSNR. The mechanism built in the algorithm efficiently compensates the excess of bitrate by increasing QP immediately. For the remaining sequences, bitrate is very close to the target one but never exceeds it.

The second subject of interest was the differences in the quality of Intra and WZ frames within one GOP. In general, they remain within a range of 1 - 2 dB after coding of a few first GOPs. However, in more dynamic regions of a sequence these differences can achieve 3-5 dB (Coastguard, Foreman, Soccer) or in extreme cases even more than 10 dB (Coastguard, Soccer).

5 CONCLUSIONS

The proposed method for DVC rate control confirms its efficiency in terms of achieving and maintaining the required bitrate. Thanks to the limitations imposed on the QP variation, differences in the quality between Intra and WZ frames fall, in general, within a range of 1-2 dB. However, there is still a lot of room for improvement. In future work, a more accurate mechanism for the Q_{Index} selection and for the bitrate prediction of WZ frames should be developed.

ACKNOWLEDGEMENTS

The work presented was developed within activities of VISNET II, the European Network of Excellence, (http://www.visnet-noe.org), founded under the European Commission IST 6FP programme.

REFERENCES

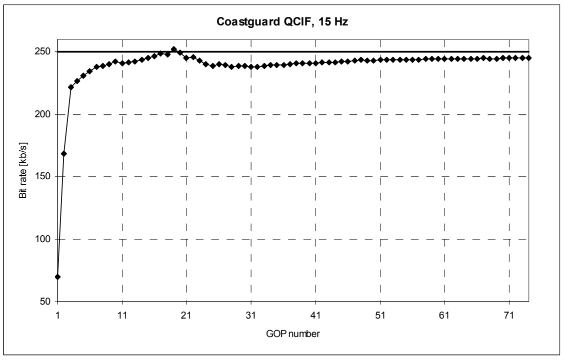
Slepian, J., Wolf., J., 1973, "Noiseless Coding of Correlated Information Sources", *IEEE Trans. on Information Theory*, vol. 19, nº 4, pp. 471 - 480.

Wyner, A., Ziv, J., 1976. "The Rate-Distortion Function for Source Coding with Side Information at the Decoder", *IEEE Trans. on Information Theory*, vol. 22, nº 1, pp. 1 - 10.

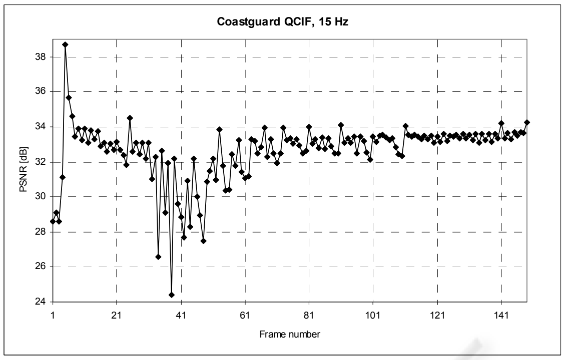
Brites, C., Ascenso, J., Pereira, F., 2006. Improving Transform Domain Wyner-Ziv Video Coding Performance. In *ICASSP'2006, IEEE Int. Conf. on Acoustics, Speech and Signal Processing,* May 14-19, Toulouse, France.

Areia, J. D., Pereira, F., Fernando, W.A.C., 2008. Impact of the Key Frames Quality on the Overall Wyner-Ziv Video Coding Performance. In *ELMAR-2008, 50th International Symposium ELMAR-2008, September 10-13. Zadar, Croatia* (submitted).

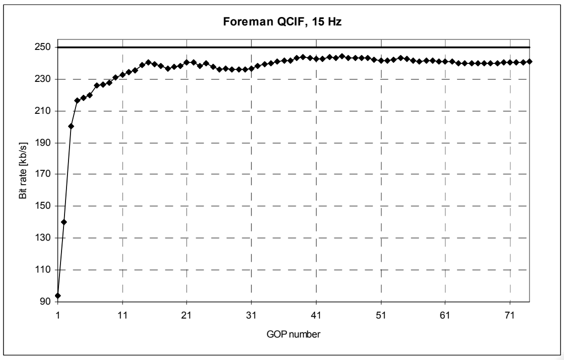
Wiegand, T, Sullivan, G. J., Bjontegaard, G., and Luthra, A., 2003. Overview of the H.264/AVC Video Coding Standard. In *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560-576.


Coastguard QCIF, 15 Hz
Bit rate [kb/s]
250
200
150
100
50
1 11 21 31 41 51 61 71
GOP number

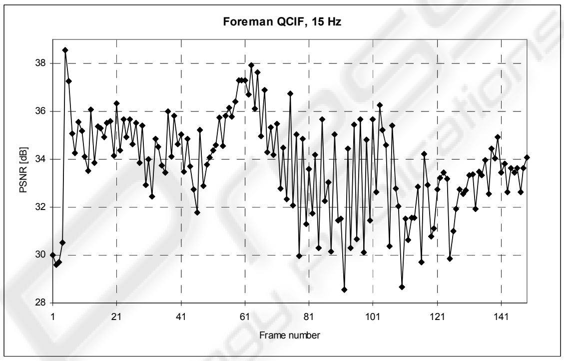
(a)


Coastguard QCIF, 15 Hz
PSNR [dB]
38
36
34
32
30
28
26
24
1 21 41 61 81 101 121 141
Frame number

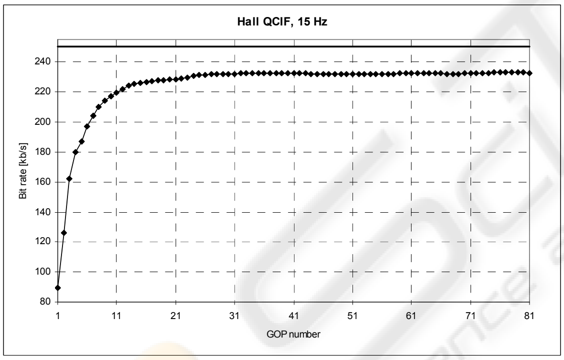
(b)


Foreman QCIF, 15 Hz
Bit rate [kb/s]
250
230
210
190
170
150
130
110
90
1 11 21 31 41 51 61 71
GOP number

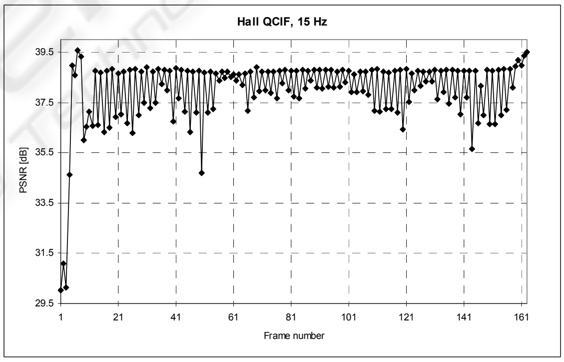
(c)


Foreman QCIF, 15 Hz
PSNR [dB]
38
36
34
32
30
28
1 21 41 61 81 101 121 141
Frame number

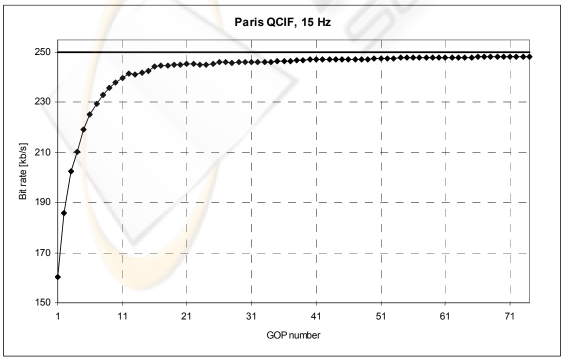
(d)


Hall QCIF, 15 Hz
Bit rate [kb/s]
240
220
200
180
160
140
120
100
80
1 11 21 31 41 51 61 71 81
GOP number

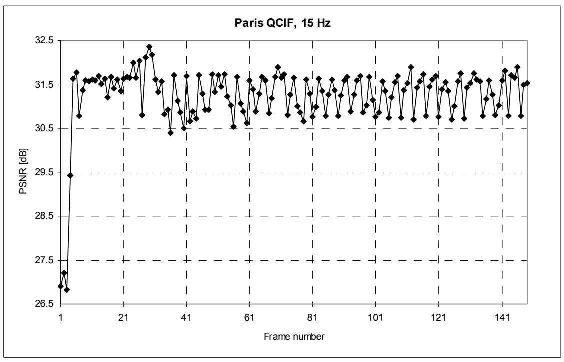
(e)


Hall QCIF, 15 Hz
PSNR [dB]
39.5
37.5
35.5
33.5
31.5
29.5
1 21 41 61 81 101 121 141 161
Frame number

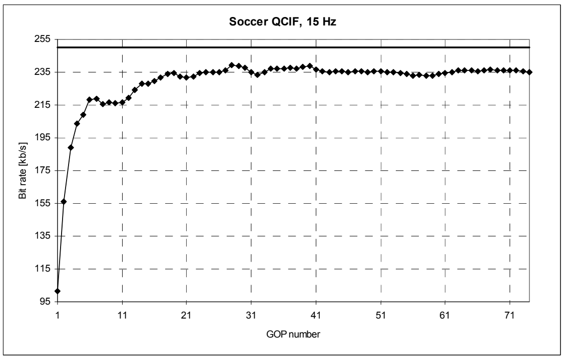
(f)


Paris QCIF, 15 Hz
Bit rate [kb/s]
250
230
210
190
170
150
1 11 21 31 41 51 61 71
GOP number

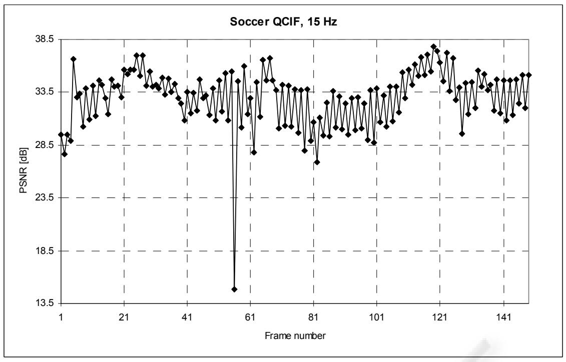
(g)


Paris QCIF, 15 Hz
PSNR [dB]
32.5
31.5
30.5
29.5
28.5
27.5
26.5
1 21 41 61 81 101 121 141
Frame number

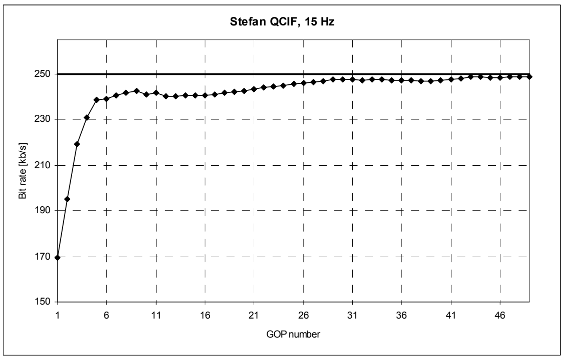
(h)



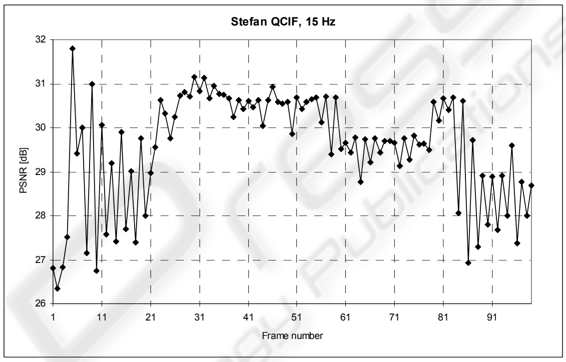
(i)



(j)



(k)



(l)

Figure 5: Bitrate and PSNR changes: (a), (b) Coastguard, (c), (d) Foreman, (e), (f) Hall Monitor, (g), (h) Paris, (i), (j) Soccer, and (k), (l) Stefan.