

Investigation of the GoP Structure for H.26L Video Streams

F. Fitzek P. Seeling M. Reisslein* M. Rossi M. Zorzi[†]

acticom GmbH – mobile networks
R & D Group
Germany
[fitzek|seeling]@acticom.de

Arizona State University
Department of Electrical Engineering
USA
reisslein@asu.edu

Dipartimento di Ingegneria
Università di Ferrara
Italy
[mrossi|zorzi]@ing.unife.it

31 December 2002

Technical Report acticom-02-004

The goal of this work is to investigate the video quality and bandwidth requirement for different Group of Picture structures for the H.26L video codec. We present bandwidth requirements and video quality in terms of the picture signal to noise values related to several different Group of Picture–structures at constant quantization settings considering a simple error model for the transmission over wireless links. These results can be used to compare different network conditions with suitable Group of Picture–structures. Especially for multicast transmission these measurements are very interesting. We employed the Group of Picture–schemes on one QCIF video sequence at a constant quantization setting, which produces a medium picture signal noise ratio at moderate frame sizes.

*The work of M. Reisslein is supported in part by the National Science Foundation through Grant No. Career ANI-0133252 and Grant No. ANI-0136774. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

[†]The work of M. Rossi and M. Zorzi has been supported by Ericsson.

Contents

1	Introduction	4
2	GoP Structures	4
3	Measurements	5
4	Results	6
5	Further Work	9
6	Acknowledgment	9

List of Figures

1	Single layer GOP (frames 1–12).	5
2	The Methodology for the GoP Measurement.	6
3	PSNR versus BEP for gop00.	7
4	Smoothed Bit Rate for gop00.	7
5	Bit Rate for gop00.	7
6	PSNR versus BEP for gop01.	7
7	Smoothed Bit Rate for gop01.	7
8	Bit Rate for gop01.	7
9	PSNR versus BEP for gop02.	7
10	Smoothed Bit Rate for gop02.	7
11	Bit Rate for gop02.	7
12	PSNR versus BEP for gop03.	7
13	Smoothed Bit Rate for gop03.	7
14	Bit Rate for gop03.	7
15	PSNR versus BEP for gop04.	7
16	Smoothed Bit Rate for gop04.	7
17	Bit Rate for gop04.	7
18	PSNR versus BEP for gop05.	8
19	Smoothed Bit Rate for gop05.	8
20	Bit Rate for gop05.	8
21	PSNR versus BEP for gop06.	8
22	Smoothed Bit Rate for gop06.	8
23	Bit Rate for gop06.	8
24	PSNR versus BEP for gop07.	8
25	Smoothed Bit Rate for gop07.	8
26	Bit Rate for gop07.	8
27	PSNR versus BEP for gop08.	8
28	Smoothed Bit Rate for gop08.	8
29	Bit Rate for gop08.	8
30	PSNR versus BEP for gop09.	8
31	Smoothed Bit Rate for gop09.	8
32	Bit Rate for gop09.	8
33	PSNR versus BEP for gop10.	9
34	Smoothed Bit Rate for gop10.	9
35	Bit Rate for gop10.	9
36	Frame sizes (averaged over GoPs) and content for the <i>Highway</i> sequence ($Q = 16$).	10
37	Comparison in terms of the video quality for GoP structure gop01, gop05, gop07, and gop10.	11

List of Tables

1	Different GoP Structures	4
---	--------------------------	---

1 Introduction

For the transmission of video in wireless environments the video quality and the bandwidth requirement are of special interest. While the network provider is interested in small bandwidth requirements, the customers wants to have a good video quality. Of course both is not achievable at the same time and there will be a trade-off between these two parameters. Using less bandwidth makes the video less robust to transmission errors, which might be a problem in the error-prone wireless environment. The goal of this work is to investigate the video quality and bandwidth requirement for different GoP structures. In [7] we present statistics related to several different GoP-structures for the H.26L codec at constant quantization settings. These results can be used to compare different network conditions with suitable GoP-structures. We employed the GoP-schemes stated below on the QCIF video sequence *Highway* [2] at a constant quantization setting of 30, which produces a medium PSNR and moderate bit rates.

For our following investigation we have chosen the H.26L codec of the Joint Video Team (JVT), which is a collaboration of the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG) for the development of a new video coding standard. The goal is to build a video coding schemes that supports both, *conversational* and *non-conversational* applications. An examples for conversational service is video telephony, while streaming is an example for the non-conversational service. The JVT/H.26L codec include a Video Coding Layer (VCL), which is responsible for the video compression, and a Network Adaptation Layer (NAL), which covers the transmission of the content. The NAL is able to transport the coded video data over existing and future networks in the wired and wireless format. A more detailed introduction to the H.26L standard and its video trace measurements is given in [3].

2 GoP Structures

As given in [7], we have chosen the GoP structures given in Table 1 for our investigations and comparisons.

Table 1: Different GoP Structures

GoP number	GoP name	Frame sequence	GoP-length
00	gop00	IBBPBBPBBPBB	12
01	gop01	IIIIIIII...	1
02	gop02	IPIPIP...	2
03	gop03	IPPIPPI...	3
04	gop04	IPPPIPPPI...	4
05	gop05	IBIBI...	2
06	gop06	IBBIBBI...	3
07	gop07	IBBBIBBBI...	4
08	gop08	IBBPBBPBBPBBPBBPBB...	whole sequence
09	gop09	IBBPBBSPBBPBBSPBBPBB...	whole sequence
10	gop10	IBBPBBSPBBPBBI...	12

There are three basic methods for encoding the original pictures in the temporal domain: I (Intra), P (Inter), and B (Bi-directional), as introduced in the MPEG-1 standard [6]. These

encoding methods are applied on the frame or block level, depending on the codec. The inter-coded frames use motion estimation relying on the previous inter- or intra-coded frame, whereas the bi-directional encoded frames rely on a previous as well as a following intra- oder inter-coded frame. Intra-coded frames or blocks are not relying on other video frames and thus are important to stop error propagation. The sequence of frames between two intra-coded frames is referred to as a *Group of Pictures* (GOP). The relationship between these different encoding types and how frames rely on each other in a typical MPEG frame sequence is illustrated in Figure 2. As can be taken from Table 1, if only the first frame is intra-coded, the whole stream is a single GoP.

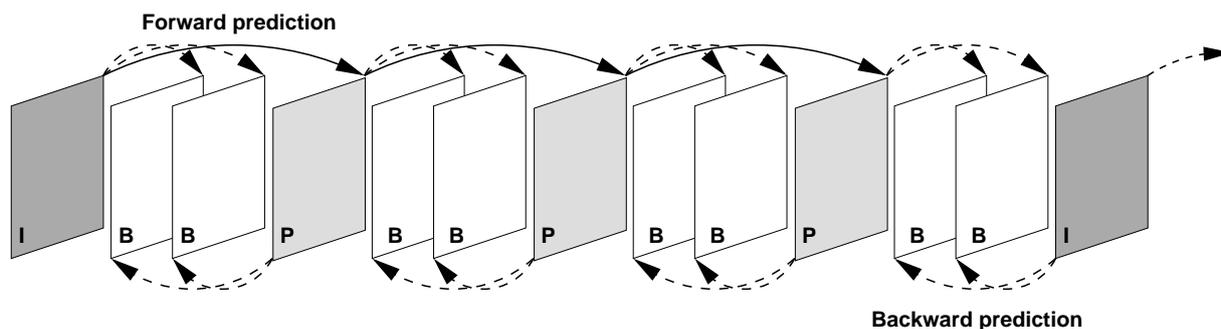


Figure 1: Single layer GOP (frames 1–12).

The new type of frames introduced with H.26L is the SP (or SI) frame. The main purpose is to allow for error resilience and stream switching possibility. The basic idea is to have SP frames to reference not only frames with the same encoding setting, but also differently encoded frames [8, 5]

3 Measurements

acticom in cooperation with Axel Meyer Fernsehproduktionen generated some new video sequences for utilization as new references. The sequence we examined shows a drive on a Czech freeway. The measurements were done in the following manner as given in Figure 2. We used the H.26L encoder version 3.6, which can be obtained from [1] to encode the original video sequence of the *Highway* sequence (available at [2]) into the H.26L bit stream with the according GoP pattern. The bandwidth requirements for the encoded video stream were already available at this point as given in Figure 2. After having generated the H.26L bit streams for different GoP structures, we did the following:

1. take each bit stream and add independent errors
2. decode the corrupted H.26L bit stream
3. compare the original video sequences with the decoded video sequence with the videometer tool (available at [4])

For each error probability we repeated that procedure 50 time. After this we compute the mean video quality and the related confidence interval of 95%. The error model is a simple function that invert bits with a certain probability. Note, error detection schemes such as in the

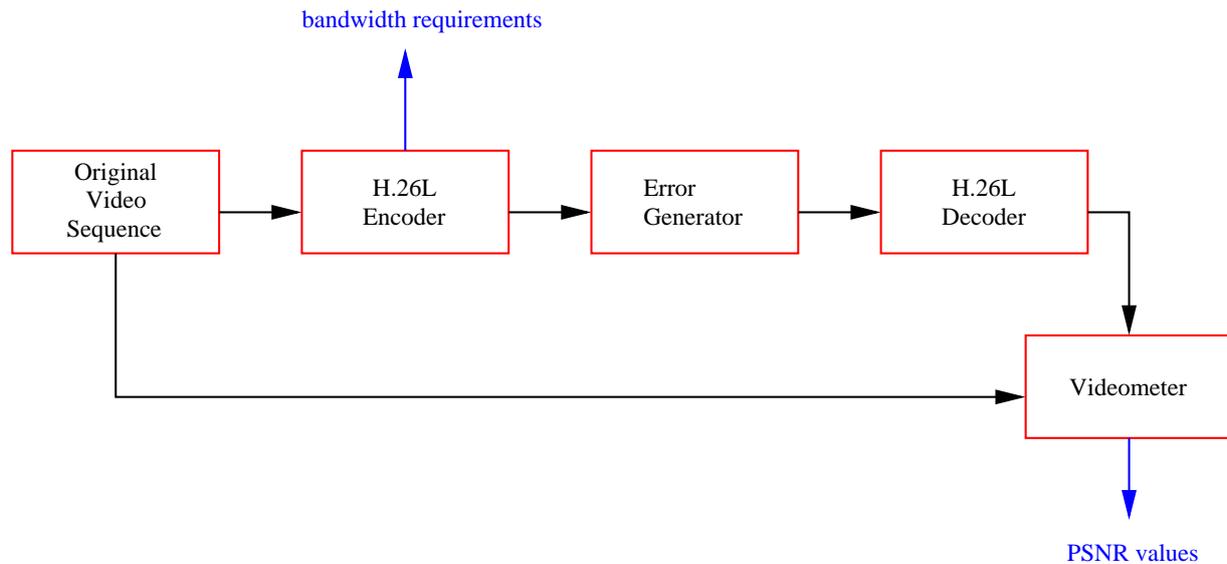


Figure 2: The Methodology for the GoP Measurement.

UDP header fields are not taken under consideration. It is assumed that each frame is passed to the decoder nevertheless how error-prone it is.

4 Results

In the following we present the results in terms of video quality and bandwidth requirements for the different GoP structures. For each GoP structure we present the PSNR¹ value versus the bit error probability (at the application layer), the smoothed bit rate, and the non-smoothed bit rate. The bit error probability varies between $10e - 09$ and $10e - 04$. For higher bit rates the PSNR values decrease dramatically and the H.26L encoder (version 3.6 [1]) shows instable behavior. Furthermore higher bit error probabilities are not well suited for video transmission. The smoothed bit rate is calculated over 12 frames, while the non-smoothed bit rate is calculated on a per frame base. The smoothed bit rate curves are introduced to increase visibility. Furthermore the mean bit rate (per video sequence) is given in the curves as a dashed line within the smoothed bit rate plot.

In the following we give a discussion of the results achieved in terms of bandwidth requirements and video quality.

bandwidth requirement GoP structure *gop01* has obviously the highest bandwidth requirements as it is composed only out of I frames containing the full frame information. Introducing more Inter frames helps to decrease the requirement in bandwidth. All bandwidth plots for the smoothed bit rate in common is the characteristic of the curves, which contain three main peaks. The peaks are caused by passing cars (remember we have the *Highway* sequence) and a bridge that we passed through as given in Figure 36. The reason why the

¹Only the Y component was taken under consideration as the human eye is more sensitive to this part of color sub-sampling.

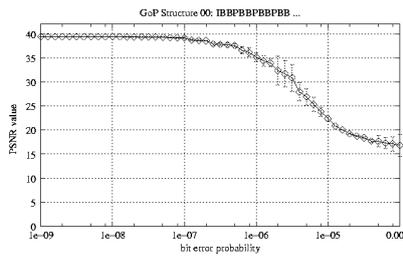


Figure 3: PSNR versus BEP for gop00.

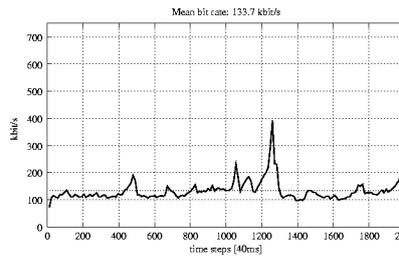


Figure 4: Smoothed Bit Rate for gop00.

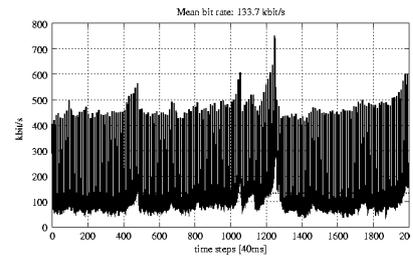


Figure 5: Bit Rate for gop00.

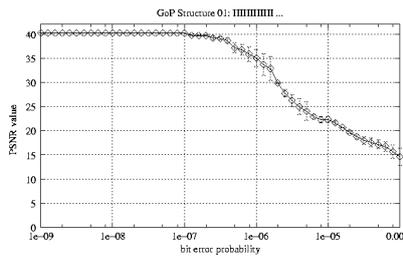


Figure 6: PSNR versus BEP for gop01.

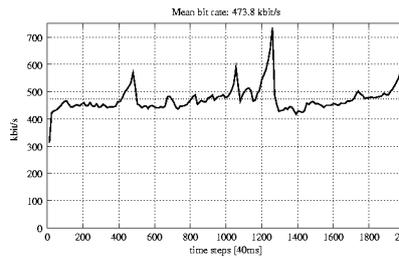


Figure 7: Smoothed Bit Rate for gop01.

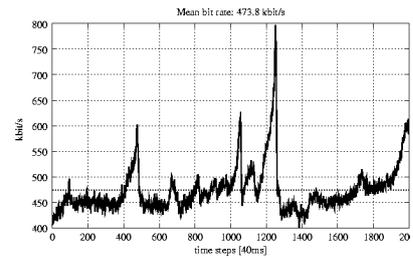


Figure 8: Bit Rate for gop01.

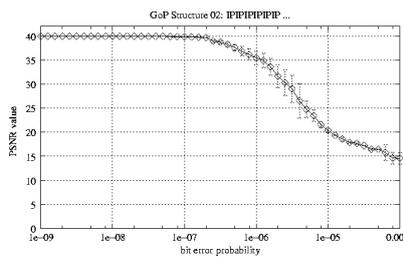


Figure 9: PSNR versus BEP for gop02.

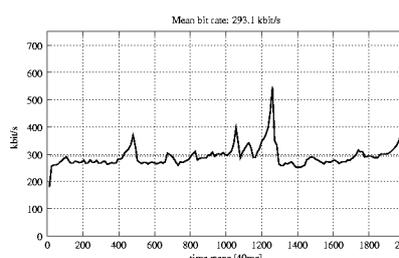


Figure 10: Smoothed Bit Rate for gop02.

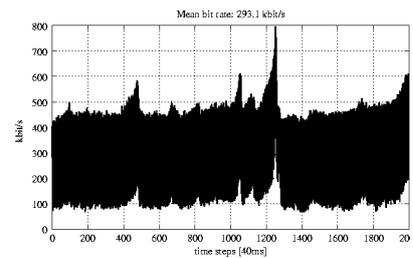


Figure 11: Bit Rate for gop02.

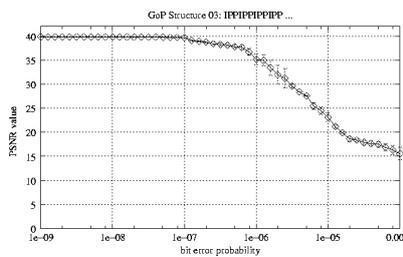


Figure 12: PSNR versus BEP for gop03.

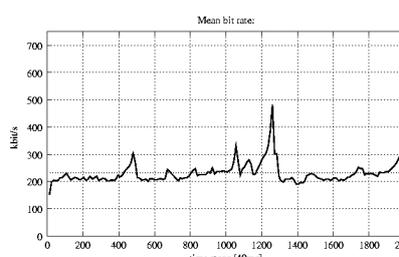


Figure 13: Smoothed Bit Rate for gop03.

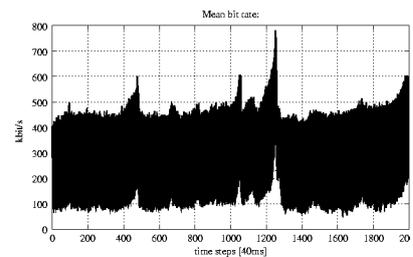


Figure 14: Bit Rate for gop03.

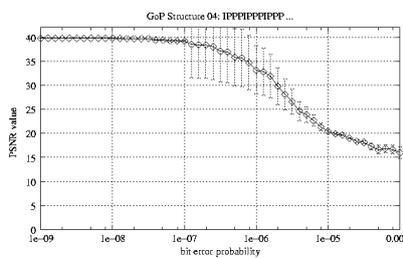


Figure 15: PSNR versus BEP for gop04.

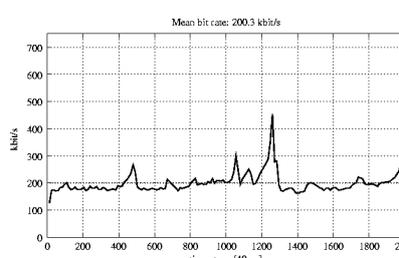


Figure 16: Smoothed Bit Rate for gop04.

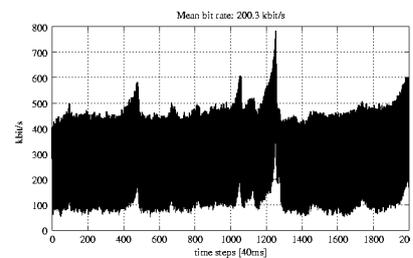


Figure 17: Bit Rate for gop04.

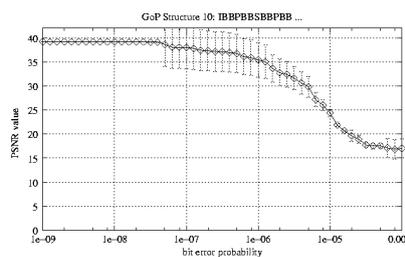


Figure 33: PSNR versus BEP for gop10.

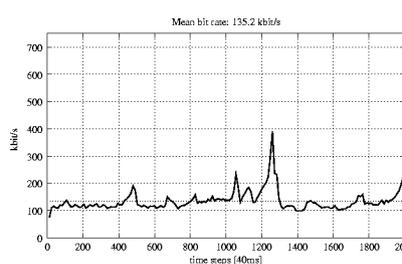


Figure 34: Smoothed Bit Rate for gop10.

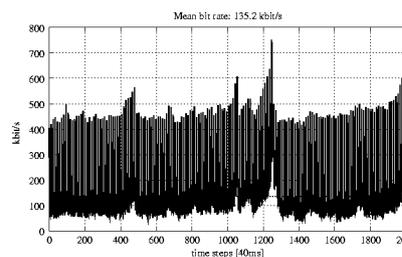


Figure 35: Bit Rate for gop10.

curves are only shifted by the mean bit rate values is that we used the variable bit rate encoding approach with fixed quality settings.

video quality The video quality of all GoP structures depends on the bit error probability. An increasing bit error probability leads to decreased PSNR values. A very interesting observation is that the GoP structure *gop01* (only I frames) does not result in the best PSNR values for all bit error probabilities. This is due to the larger frame sizes. Even if the bit error probability is the same for all video frames, the absolute number of errors is larger for the I frames. Thus, in case of independent bit error probabilities at the application layer, GoP structures with high Intra frame rates lead to high bandwidth, but the video quality is less improved than expected. This is due to the bit error characteristic. In case of bursty bit error characteristics the use of Intra frames seem to be more reasonable. In highly error prone environment the GoP structure *gop10* with the S frames leads to the highest PSNR values. In Figure 37 a comparison in terms of the video quality for GoP structure *gop01*, *gop05*, *gop07*, and *gop10* is given. For certain regions of bit error probabilities, a GoP structure with high PSNR values can be found. But the trade-off with the bandwidth requirements has to be made.

In case the GoP structure 10 is chosen throughout the whole range of bit errors a low bandwidth is guaranteed, but a maximum in quality degradation in the range of 4 dB has to be taken into account.

5 Further Work

In the future we will repeat the video quality measurements for different contents. Here we present the methodology the rest is only hard computational work. Note, the values for the PSNR as well as for the bandwidth requirements depends on the content. Thus, different values can be achieved by applying a different content. More interesting would be the investigation of the video quality in presence of a correlated error model. Therefore we use such video traces within an UMTS simulator. Furthermore, we will apply these video measurements to multicast streaming services. The trade of between bandwidth and video quality differs to that of unicast transmission.

6 Acknowledgment

We would like to thank Andrea Filippini for his help in the beginning of the project and Axel Meyer for his advice in video post-processing. Furthermore we would to thank Sterica Rein for

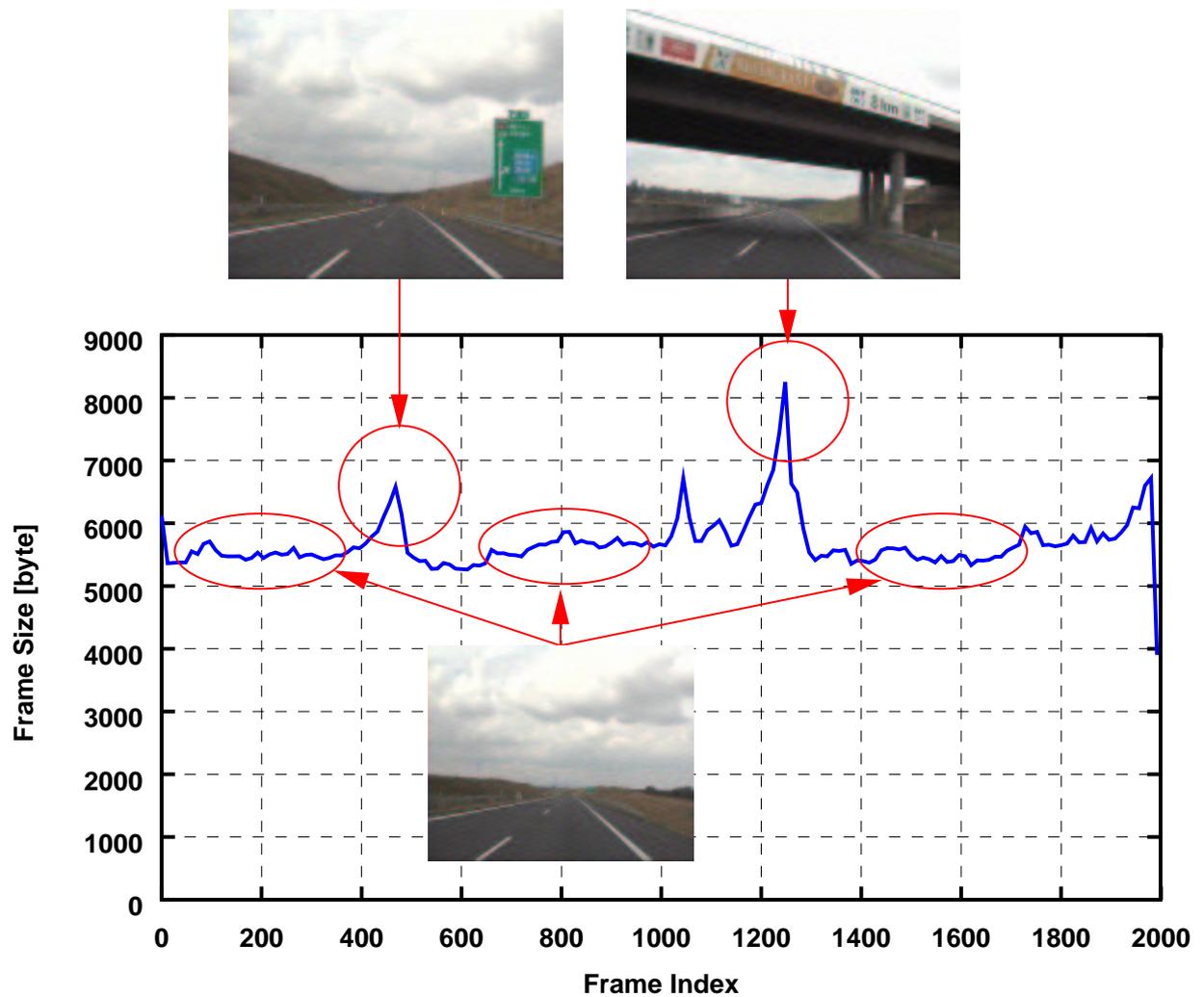


Figure 36: Frame sizes (averaged over GoPs) and content for the *Highway* sequence ($Q = 16$).

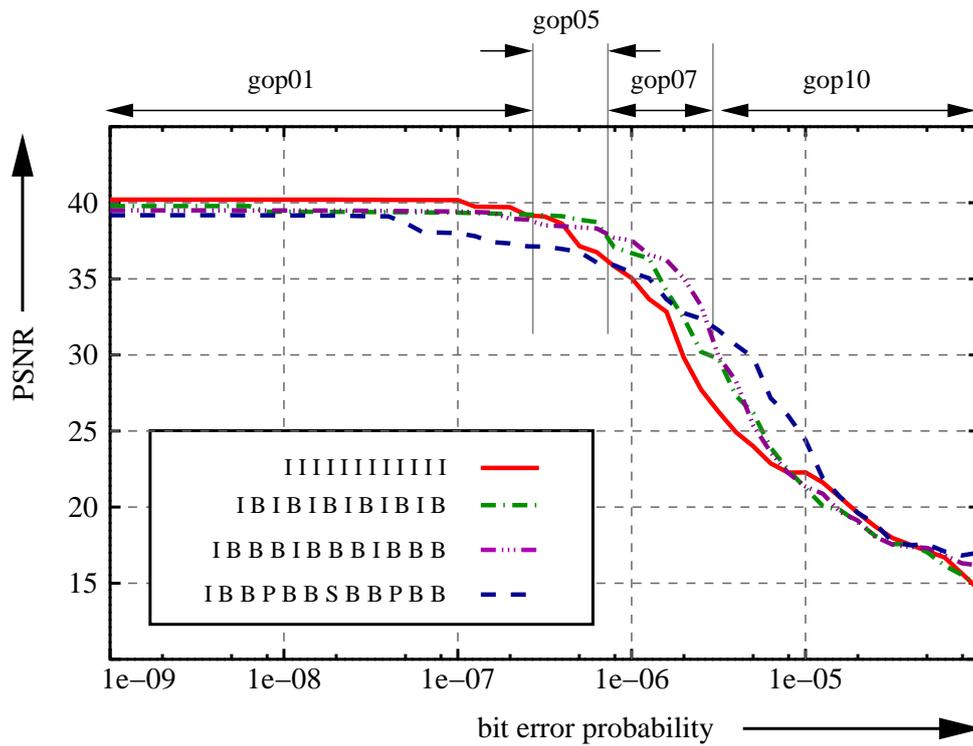


Figure 37: Comparison in terms of the video quality for GoP structure gop01, gop05, gop07, and gop10.

the technical support generating the video sequence *Highway* in Eastern Europe. Also we would like to thank all anonymous people which helped us to realize this work.

References

- [1] C. Suehring. H.26l software coordination. <http://bs.hhi.de/~suehring/tml/>. 5, 6
- [2] F.H.P. Fitzek. YUV video sequences. <http://www.acticom.info/yuv.html>. 4, 5
- [3] F.H.P. Fitzek, P. Seeling, and M. Reisslein. H.26L Pre-Standard Evaluation. Technical Report acticom-02-002, acticom – mobile networks, Germany, November 2002. 4
- [4] F.H.P. Fitzek, P. Seeling, and M. Reisslein. VideoMeter tool for YUV bitstreams. Technical Report acticom-02-001, acticom – mobile networks, Germany, October 2002. 5
- [5] Th. Wiegand G.J. Sullivan and Th. Stockhammer. Using the Draft H.26L Video Coding Standard for Mobile applications. 5
- [6] MPEG–1. Coding of moving pictures and associated audio for digital storage media at up to 1.5 Mbps. ISO/IEC 11172, 1993. 4
- [7] P. Seeling. GoP Dependency Evaluation. <http://www.acticom.info/gop.html>. 4
- [8] Th. Wiegand. H.26L Test Model Long–Term Number 9 (TML-9) draft0. ITU-T Study Group 16, Dec. 2001. 5