Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG
(ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6)
19th Meeting: Geneva, CH, 31 March-7 April, 2006

Document: JVT-S063
Filename: JVT-S063.doc

Title: **Extended inter-layer motion vectors prediction in JSVM – case study and experimental results**

Status: Input Document to JVT

Purpose: Report

Author(s) or Contact(s): Rafał Lange, Marek Domański

Tel: +48 61 665 2870
Email: rlange@multimedia.edu.pl
domanski@et.put.poznan.pl

Source: Poznań University of Technology, Poznań, Poland

Abstract

In the existing draft of Joint Scalable Video Model JSVM-4 there are two ways of exploiting inter-layer correlation between motion vectors from the base and the spatial enhancement layers [1]. In the contribution JVT-R094 the extended scheme of inter-layer motion vectors prediction has been proposed for reducing motion vector prediction residual [2]. In this paper various techniques of inter-layer representation of the motion vectors in the JSVM has been investigated. Motion vectors have been encoded independently in each layer, inter-layer motion vector prediction has been performed as defined in JSVM and the combination of extended inter-layer prediction has been tested. The experimental results have shown the impact of the motion data encoding technique on the similarity of the base and the enhancement layer motion vectors. Overall bitrate, motion vector prediction residual and the average number of bits per motion vector component have been investigated.

1. Introduction

Efficient representations of the motion data in the scalable video coder are becoming very important. Beside the intra-layer AVC-like motion vector prediction, additional inter-layer motion correlation can be exploited in the enhancement layer of the scalable video encoder.

In the existing draft of Scalable Video Coding standard [1] motion vectors from the base layer are rescaled and then they can be directly taken for motion compensation in the enhancement layer. Another motion prediction mode allows for sending motion vector refinement data for more precise motion compensation. These two prediction modes are explicitly signaled in the scalable bitstream by the encoder.

Extended inter-layer motion vector prediction technique has been proposed [2] for reducing motion vector prediction residuals. The proposed scheme has been implemented in the JSVM 4.0 software and tested in order to choose the best motion vectors representation in JSVM. Extended inter-layer prediction has been compared against the independent coding of the motion vectors in each layer and against the standard method of inter-layer motion prediction as defined in JSVM.

2. Existing inter-layer prediction of motion vectors in scalable extension of AVC/H.264

In the draft of scalable extension of AVC/H.264 (JSVM-4) [1] new macroblock prediction modes are specified. They allow for inter-layer prediction of motion vectors. The syntax elements **base_mode_flag** and **base_mode_refinement_flag** defined in macroblock_layer_in_scalable_extension( ) allow for inheriting reference frame indices and motion vectors from the base layer motion vectors field. In the first mode, no further motion data is sent, however for the latter mode only 1/4 –pel motion vector refinement is sent. Additionally, **motion_prediction_flag_lX[mbPartIdx]** allows for switching on inter-layer prediction for selected partitions of the macroblock, with differentially encoded motion vector prediction error.

3. Proposal of extended inter-layer prediction of motion vectors

Despite of standard intra-layer prediction of motion vectors, there are still some cases, when there are no good predictions of the current motion vector in the enhanced layer. For example, when the current macroblock is the first macroblock in the slice, there are no neighbouring motion vectors for prediction. Another problem appears when neighbours are intra-coded or use different reference picture for motion compensated prediction. In such cases we can use “missing” or “poor” motion vector normally used for prediction with the rescaled co-located motion vector from the base layer as depicted on Fig. 1.

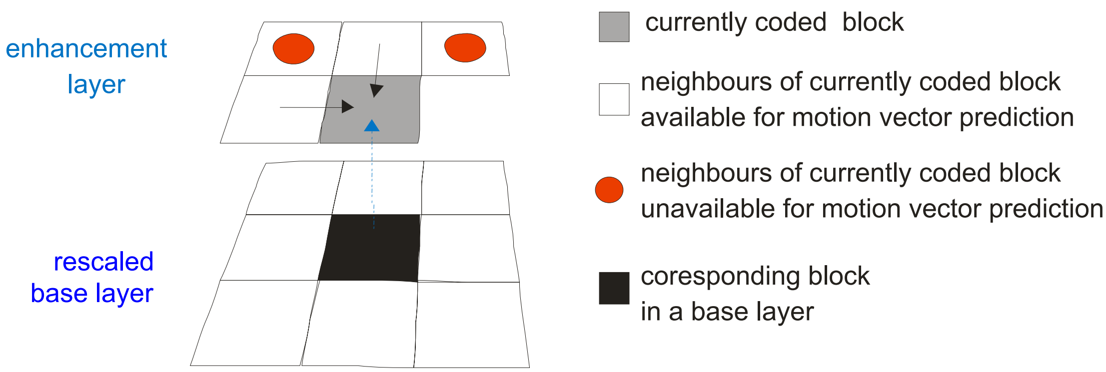

Fig. 1. Blocks used in extended inter-layer motion prediction.

The new idea is to not signal explicitly of inter-layer prediction of motion in the bitstream, but rather to use the co-located motion vector from the base layer as the possible candidate for standard prediction. This is performed only when neighboring blocks are:
- not in the same slice,
- intra coded,
- using different reference frames for prediction.

In such cases the motion vector associated with the block is marked as “missing” and can be replaced by a rescaled motion vector from the lower layer.

This extended inter-layer prediction scheme can be used jointly with the existing solution of inter-layer prediction as an optional modification. Here, the experimental results should prove what is the possible gain from using the proposed technique in JSVM.

4. Codec modifications

Some modifications have been applied to the JSVM 4.0 software that allow for testing the following cases of inter-layer motion predictions:

– no inter-layer motion dependencies are exploited and no data signalling it is sent in the bitstream (reffered as “***-_-***“ in tables and on figures),

– standard inter-layer motion prediction as defined in draft standard is used (reffered as “***-_EXP***“ on figures),

– extended inter-layer motion prediction as proposed in [2] is used (reffered as “***EXT_-***“ on figures),

– jointly used standard inter-layer prediction and proposed extended inter-layer prediction is used (reffered as “***ET_EXP***“ on figures).

5. Experimental results

The tests have been performed for the CIF sequences BUS, FOREMAN, FOOTBALL and MOBILE using modified JSVM 4.0 software. The coder was set up for producing two spatial layers (QCIF and CIF), both temporally scalable using B-frames skipping. The following parameters have been set in the configuration file:

– GOPSize = 2,
– NumberReferenceFrames = 3,
– SymbolMode = 1 (CABAC),
– UpdateStep = 0,
– InterLayerPred = 2 (adaptive).

For each sequence, the average motion vector prediction error has been measured and the average number of bits per motion vector component has been calculated. On Figures 2-5, “rate-distortion curves” are shown for various techniques of inter-layer motion prediction. On Figures 6-9 the average motion vector component residual is presented for P-slices, on Figures 10-13 the average motion vector component residual is presented for B-slices.

Additional investigation has been conducted about similarity of motion vector field in the base and in the enhancement layers. The differential motion vector field has been calculated for the upsampled motion vectors from the base layer and corresponding motion vectors from the enhancement layer. For each P- and B- slices motion vectors from the enhancement layer assigned to 4×4 luma blocks that were coded using motion compensated prediction have been subtracted from the co-located motion vectors from the base layer. The average differential motion vector length has been then calculated separately for P- and B- slices. This measures represent the similarity of motion vector fields form the base and the enhancement layer, they have been presented on Figures 14-21.

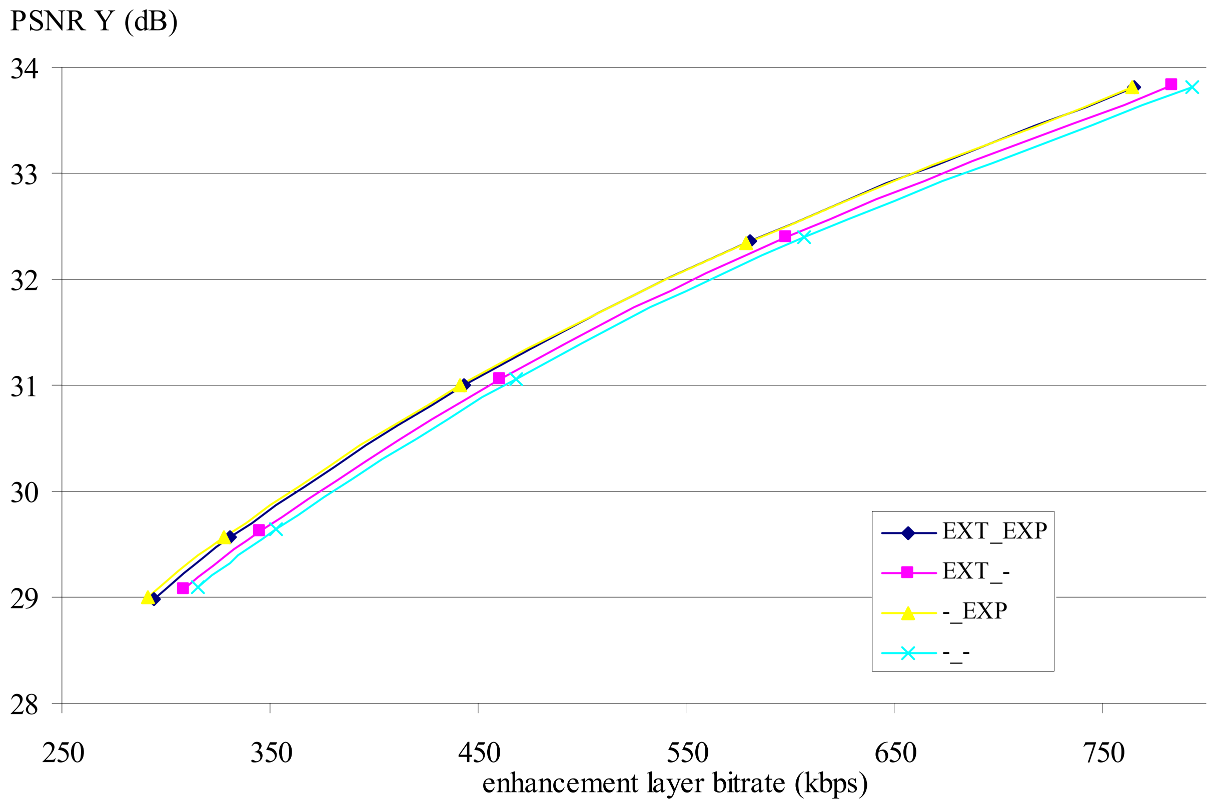

Fig. 2. R-D curves for various techniques of inter-layer predictions in BUS sequence.

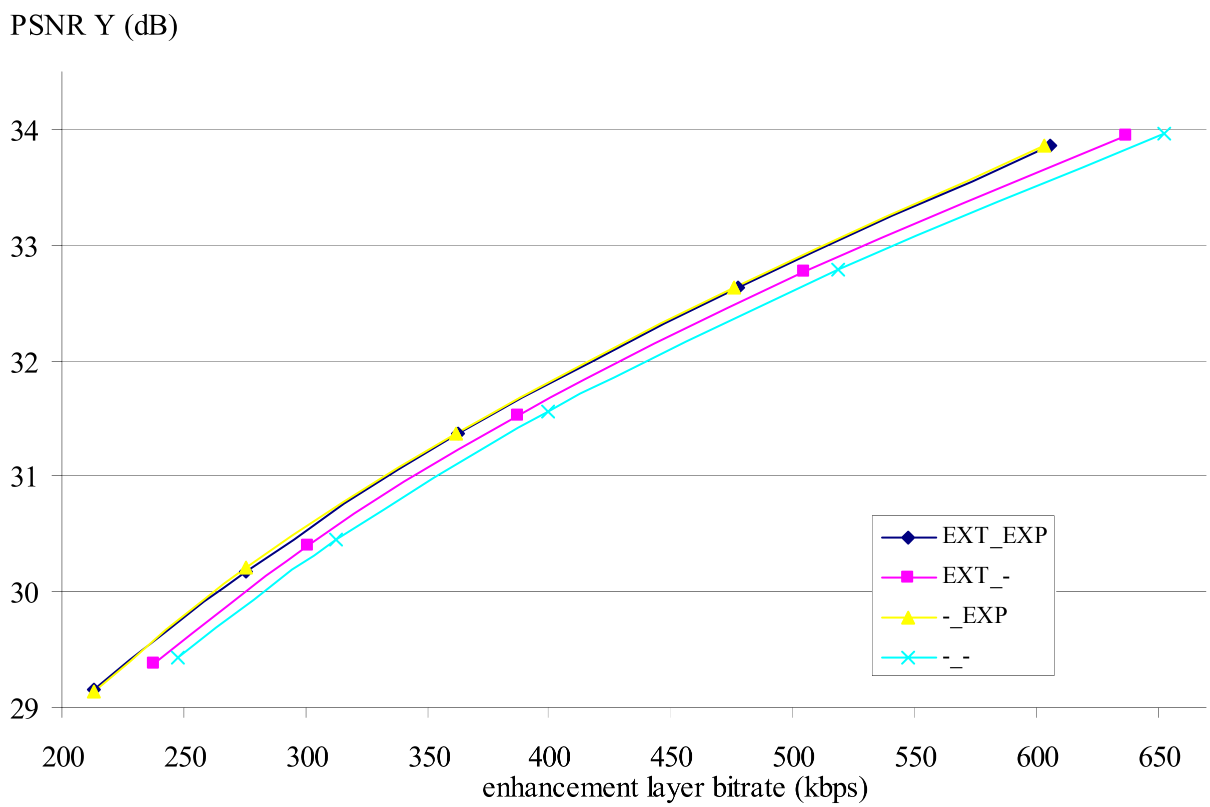

Fig. 3. R-D curves for various techniques of inter-layer predictions in FOOTBALL sequence.

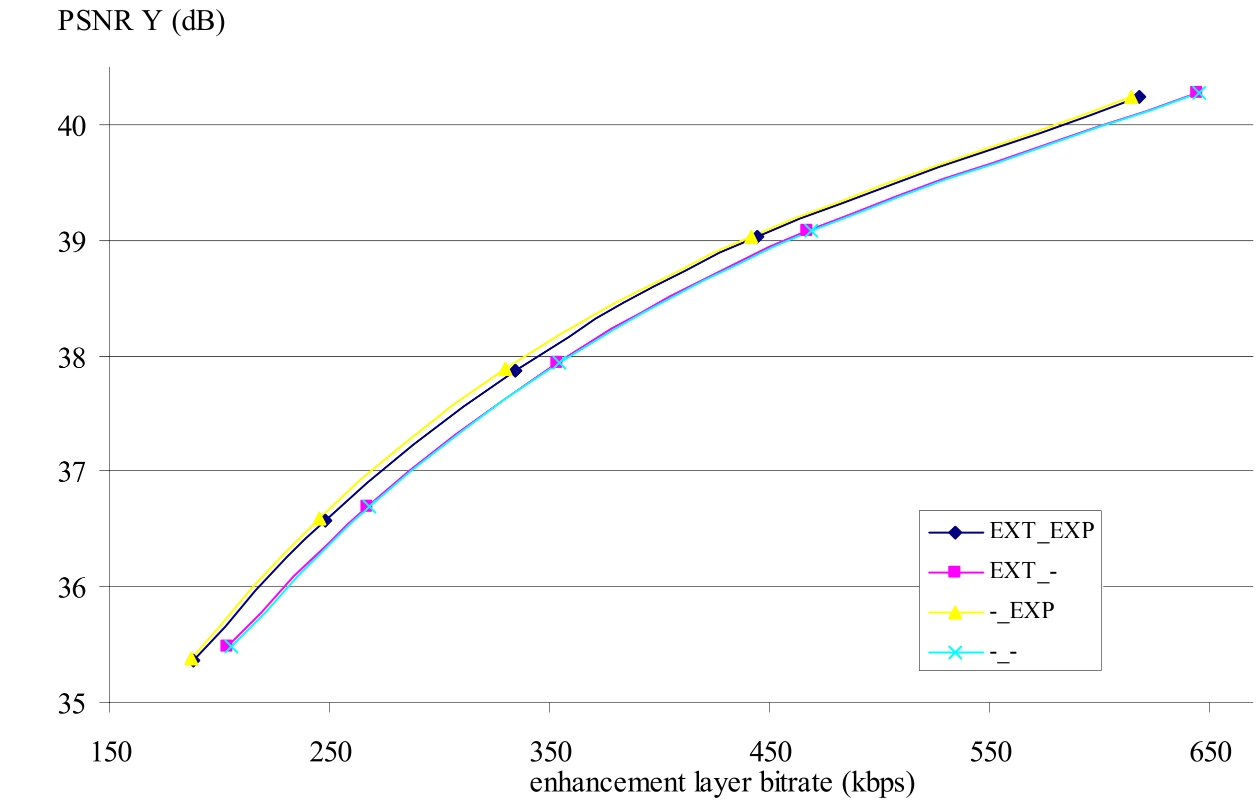

Fig. 4. R-D curves for various techniques of inter-layer predictions in FOREMAN sequence.

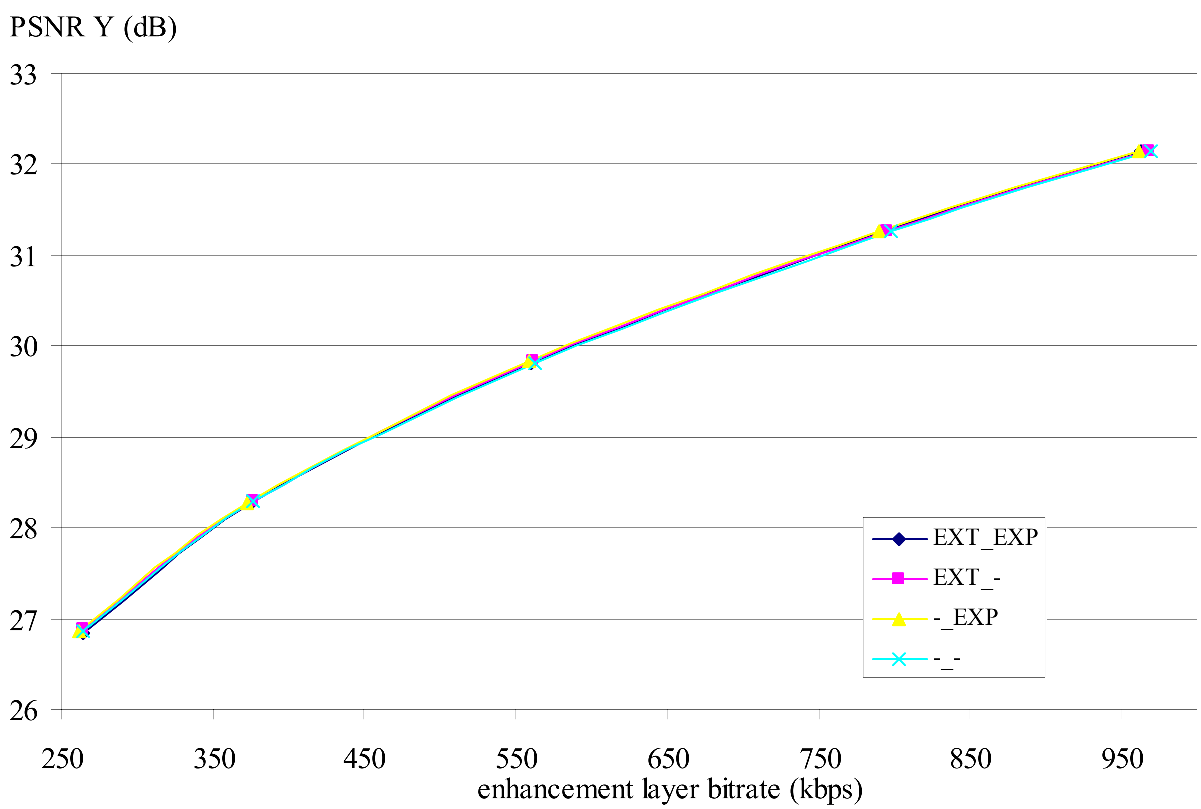

Fig. 5. R-D curves for various techniques of inter-layer predictions in MOBILE sequence.

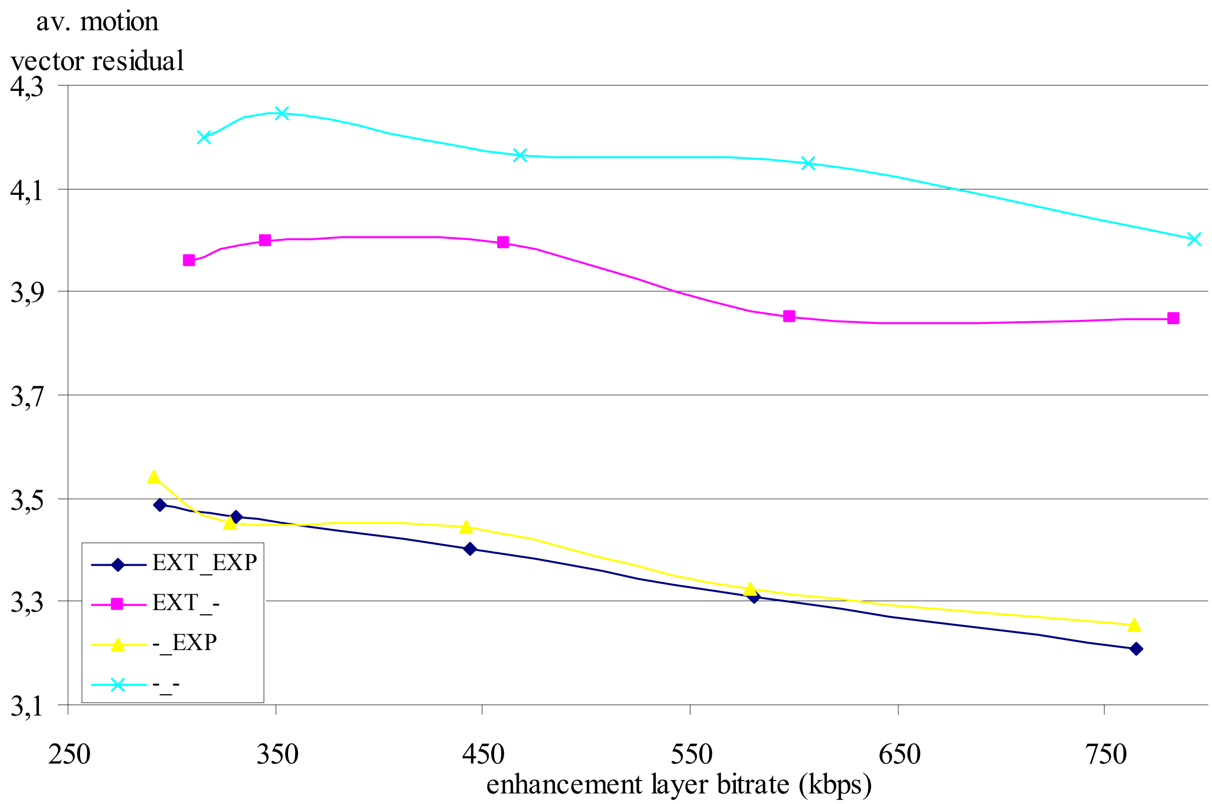

Fig. 6. Average absolute values of components of motion vector residuals for various inter-layer motion prediction, enhancement layer of BUS sequence, P-slices.

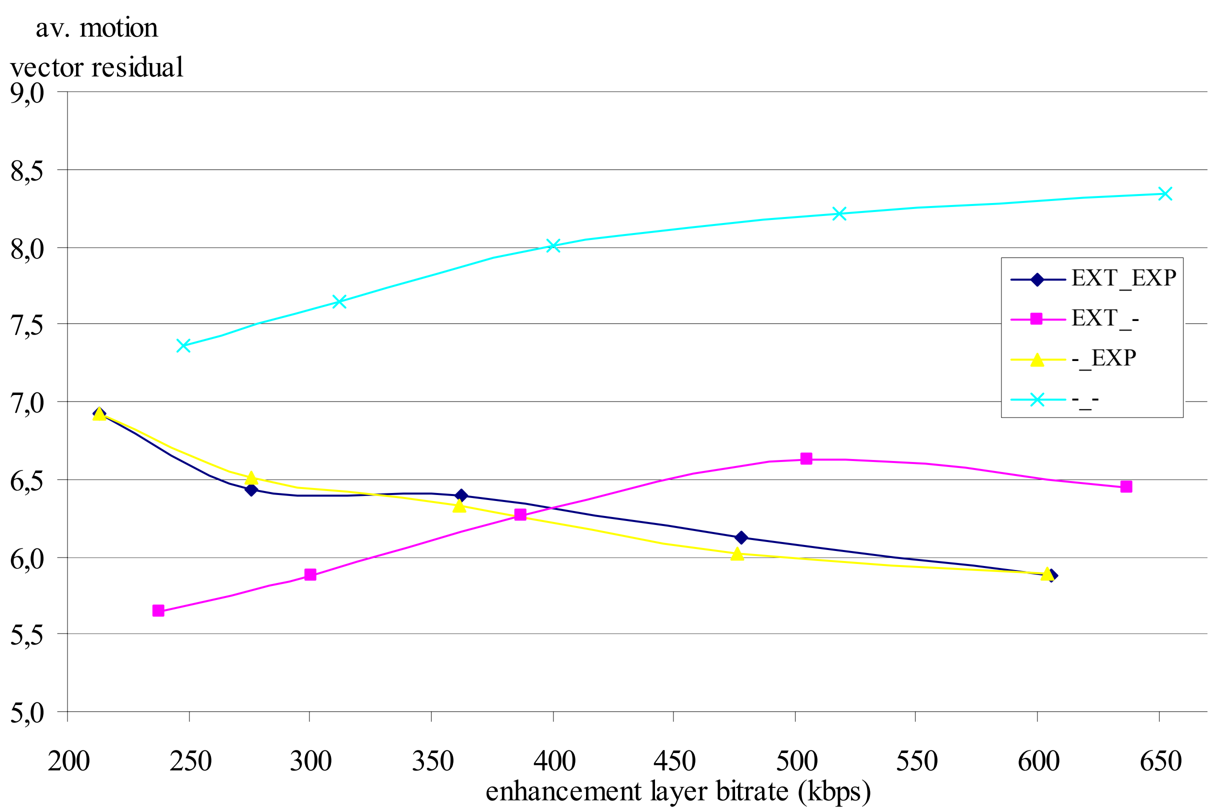

Fig. 7. Average absolute values of components of motion vector residuals for various inter-layer motion prediction, enhancement layer of FOOTBALL sequence, P-slices.

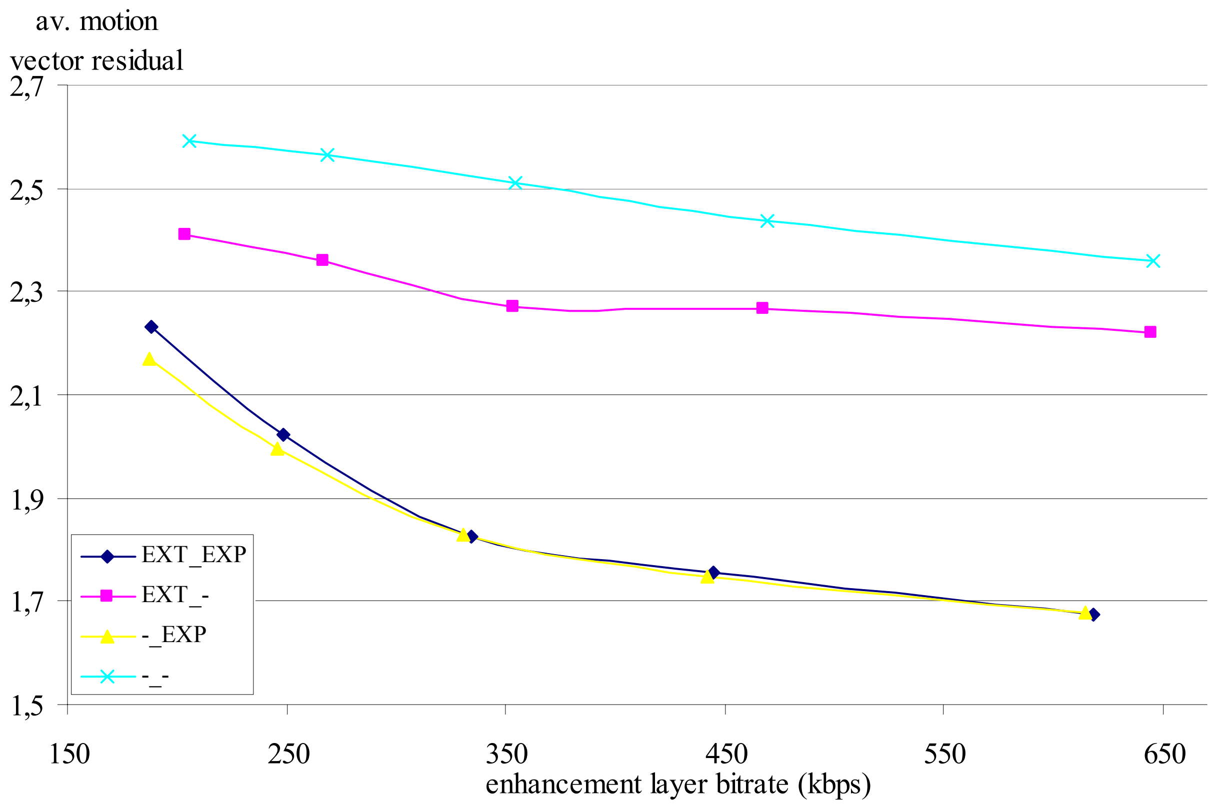

Fig. 8. Average absolute values of components of motion vector residuals for various inter-layer motion prediction, enhancement layer of FOREMAN sequence, P-slices.

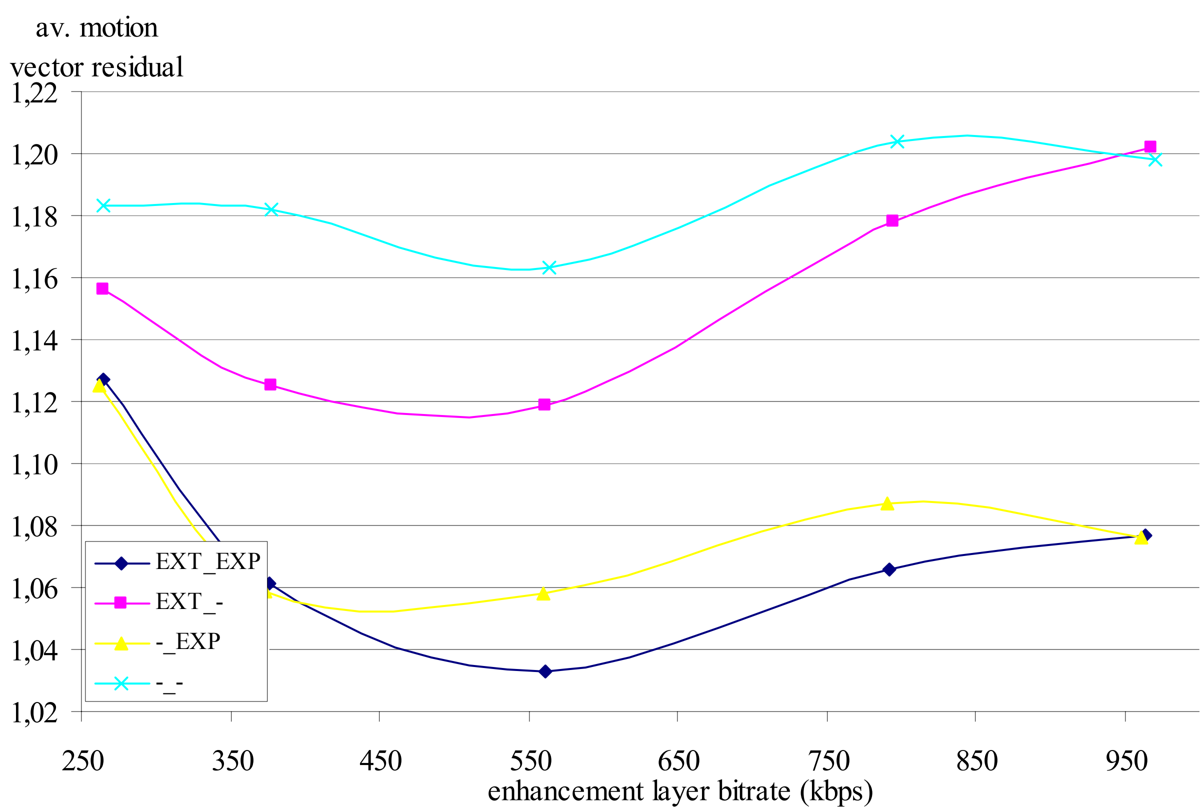

Fig. 9. Average absolute values of components of motion vector residuals for various inter-layer motion prediction, enhancement layer of MOBILE sequence, P-slices.

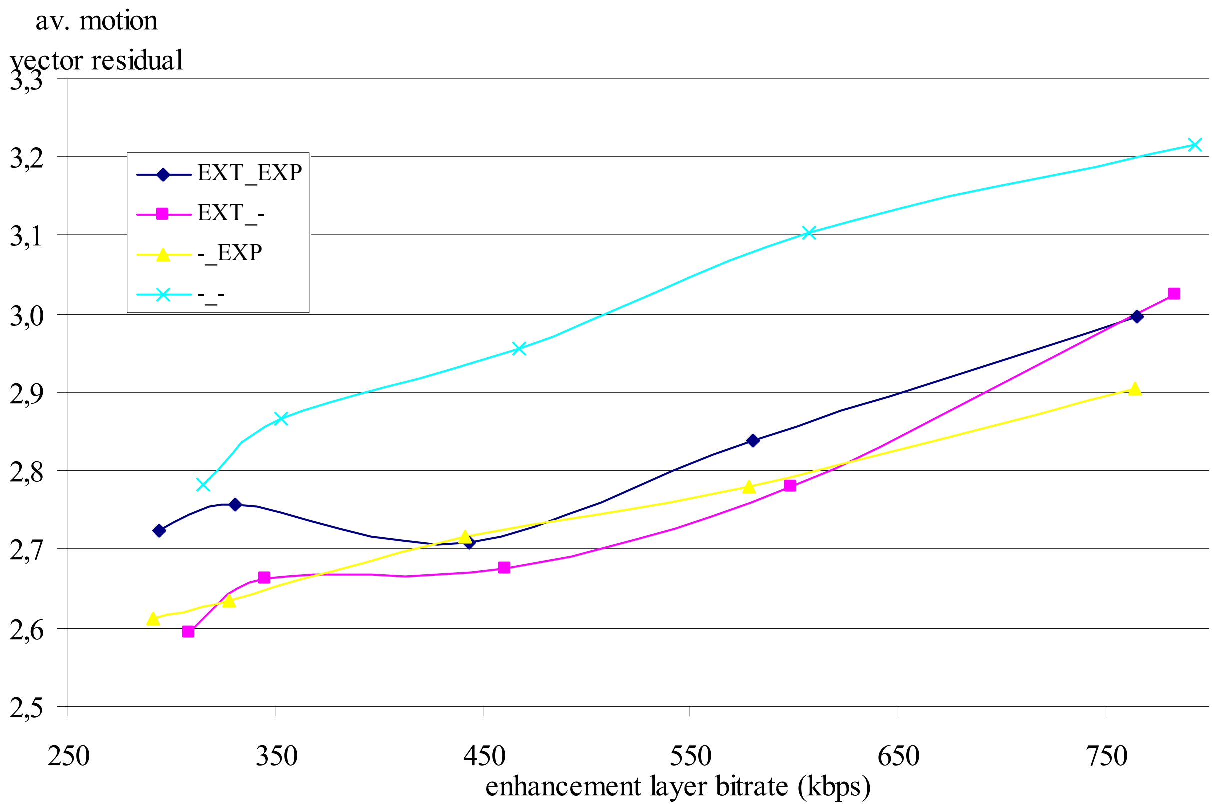

Fig. 10. Average absolute values of components of motion vector residuals for various inter-layer motion prediction, enhancement layer of BUS sequence, B-slices.

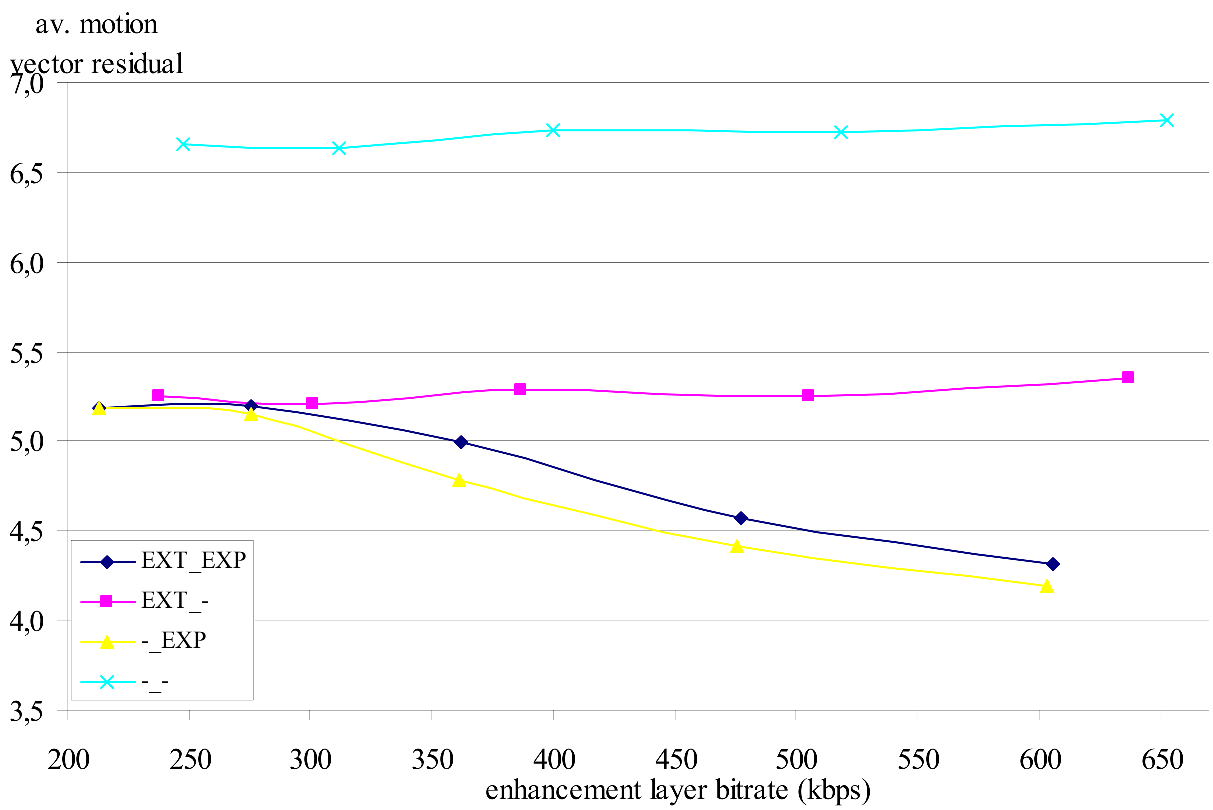

Fig. 11. Average absolute values of components of motion vector residuals for various inter-layer motion prediction, enhancement layer of FOOTBALL sequence, B-slices.

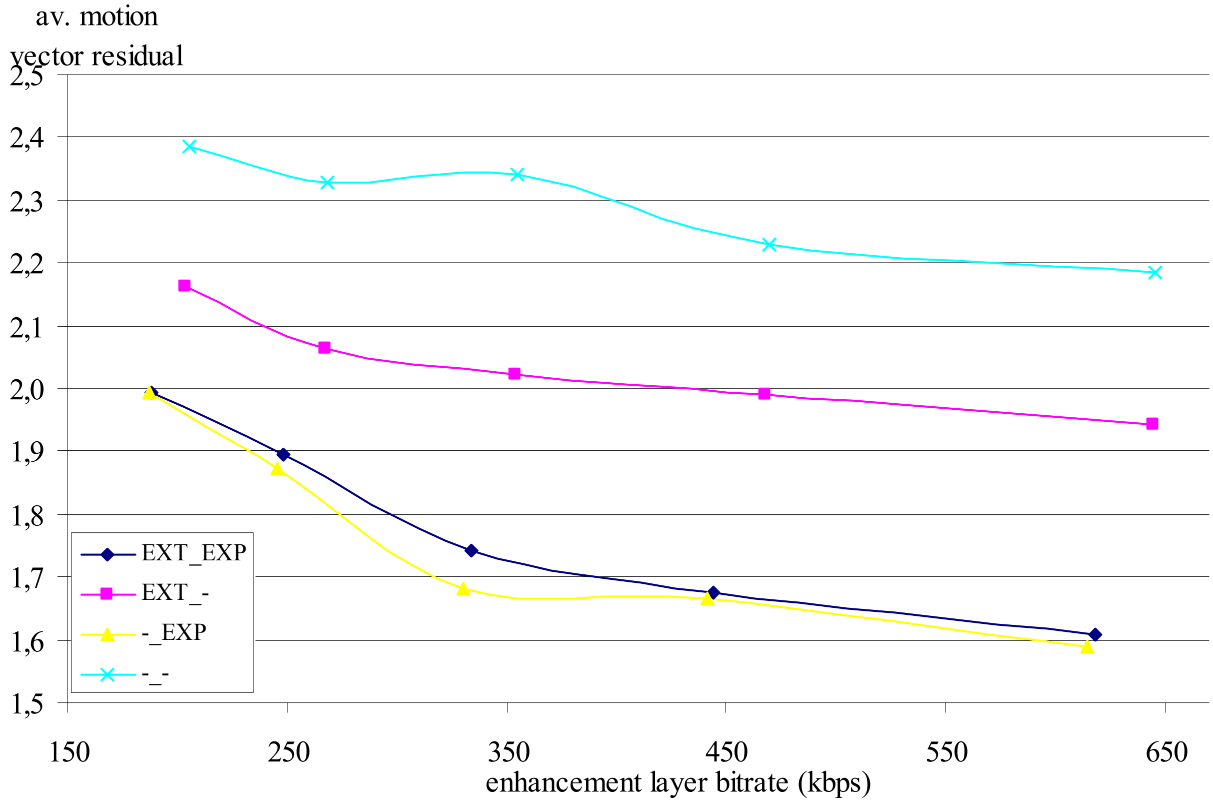

Fig. 12. Average absolute values of components of motion vector residuals for various inter-layer motion prediction, enhancement layer of FOREMAN sequence, B-slices.

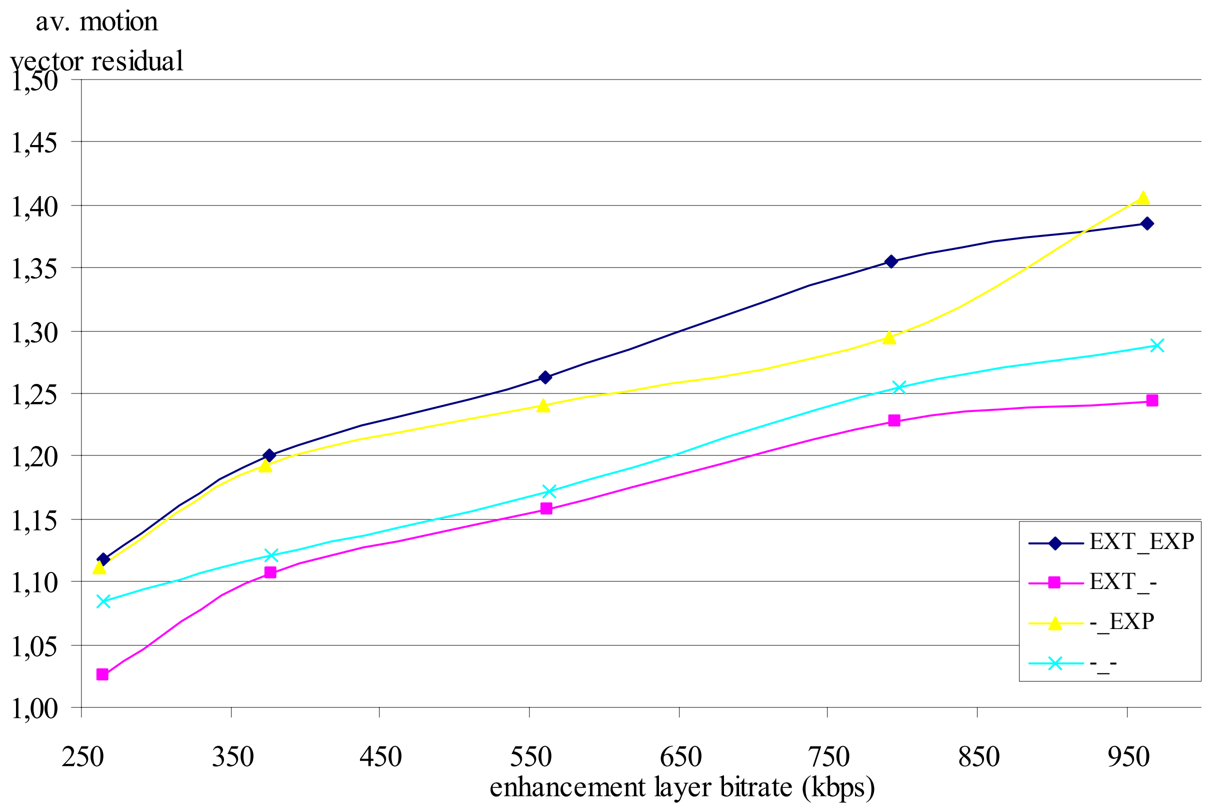

Fig. 13. Average absolute values of components of motion vector residuals for various inter-layer motion prediction, enhancement layer of MOBILE sequence, B-slices.

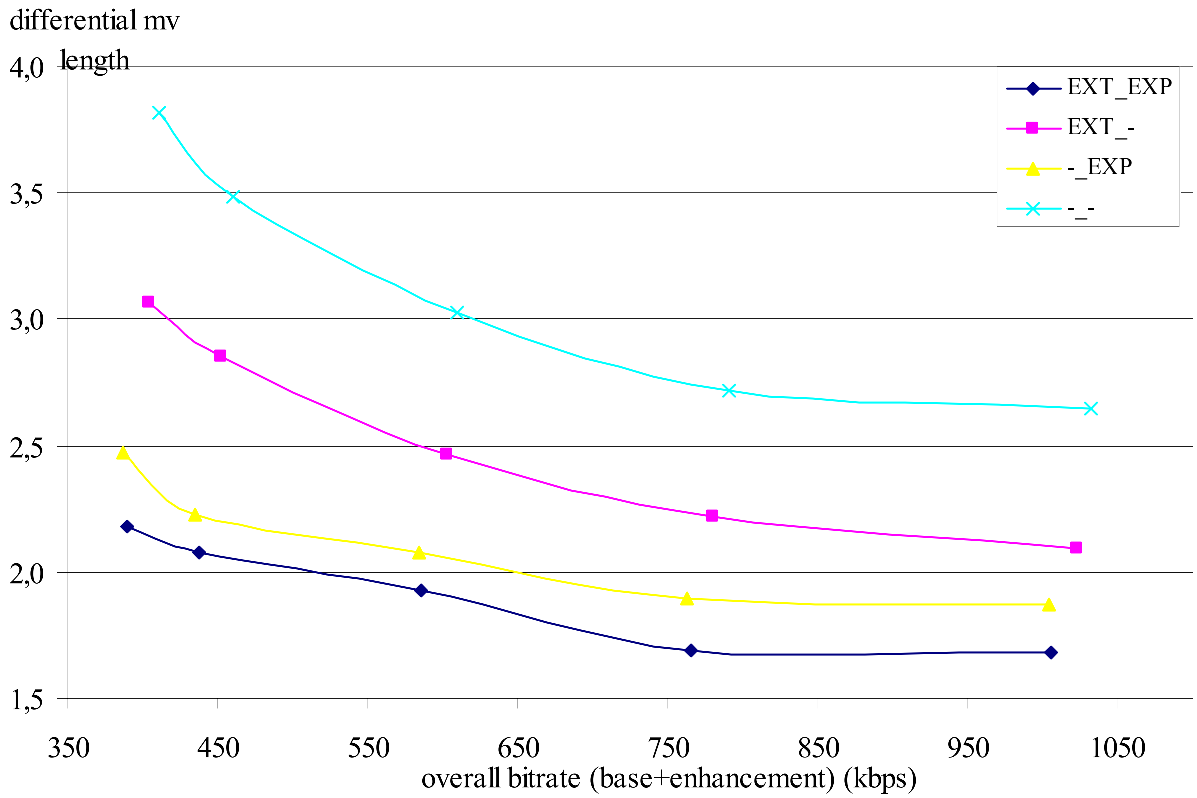

Fig. 14. Average differential motion vector length for the base and enhancement layer for various inter-layer motion prediction, BUS sequence, P-slices.

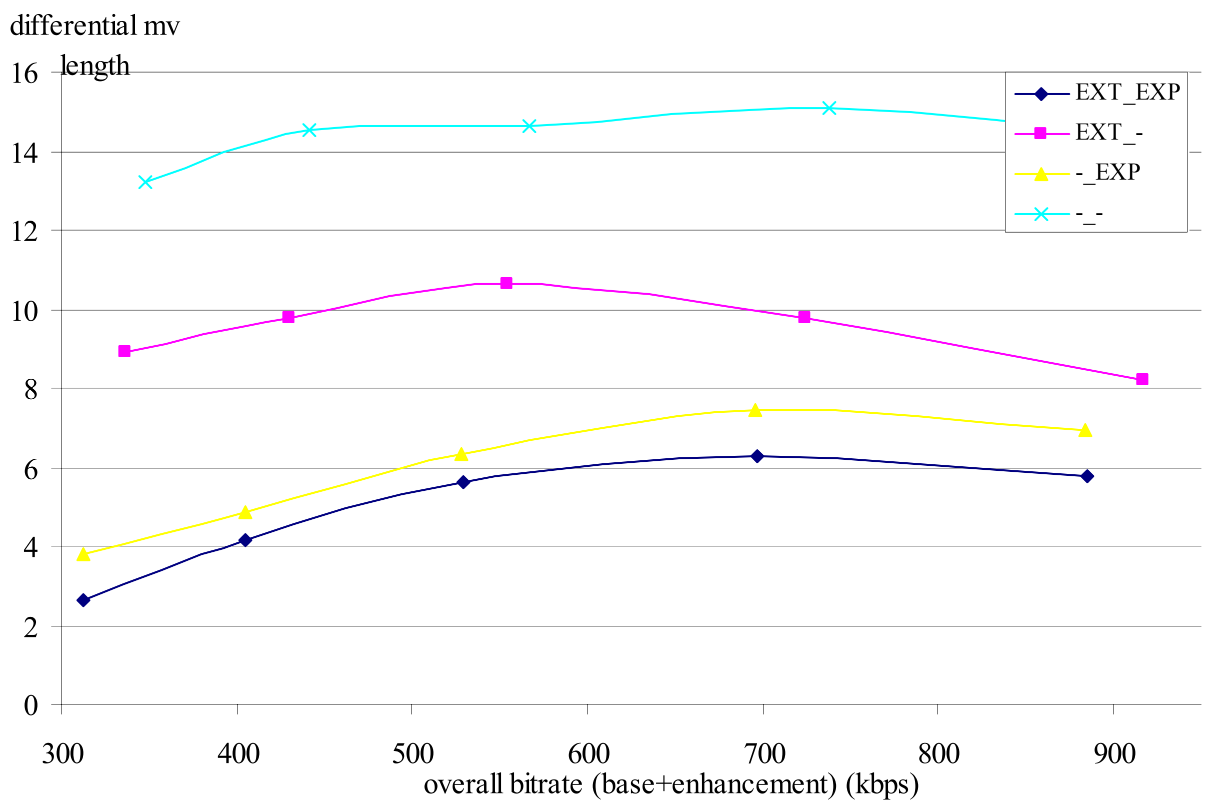

Fig. 15. Average differential motion vector length for the base and enhancement layer for various inter-layer motion prediction, FOOTBALL sequence, P-slices.

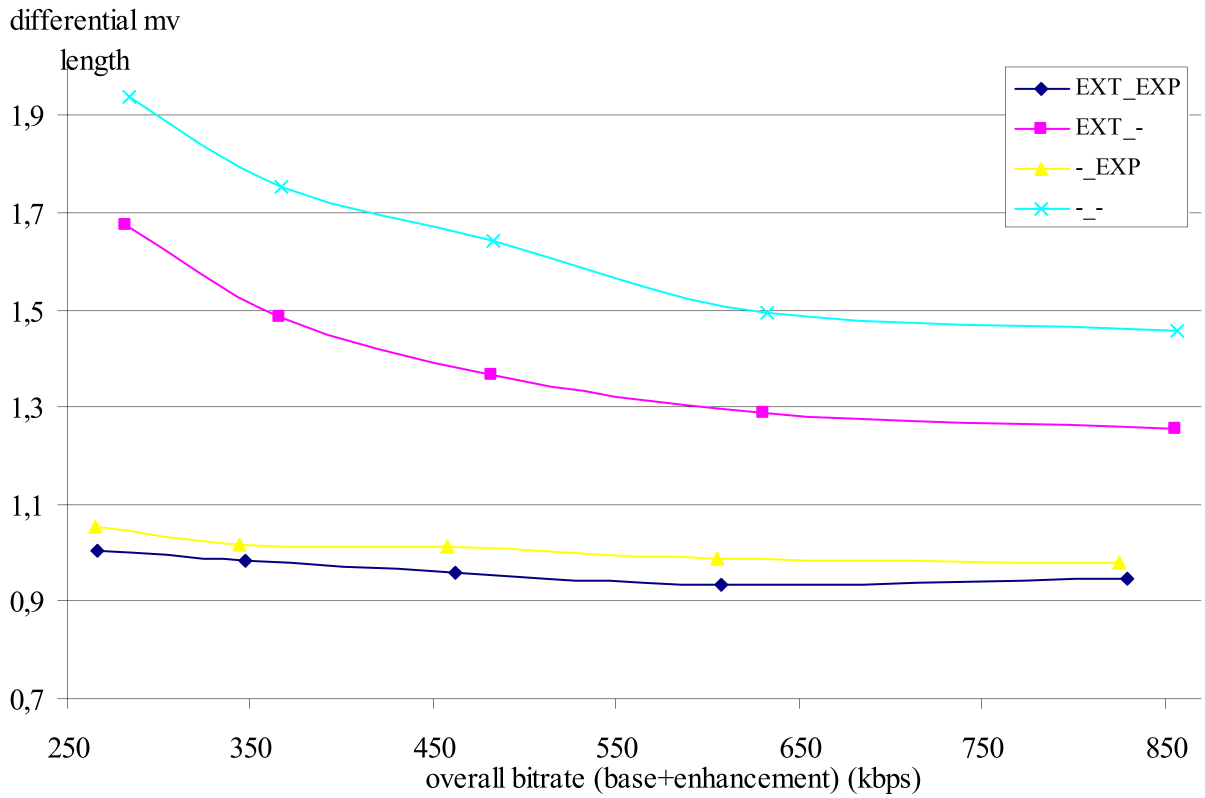

Fig. 16. Average differential motion vector length for the base and enhancement layer for various inter-layer motion prediction, FOREMAN sequence, P-slices.

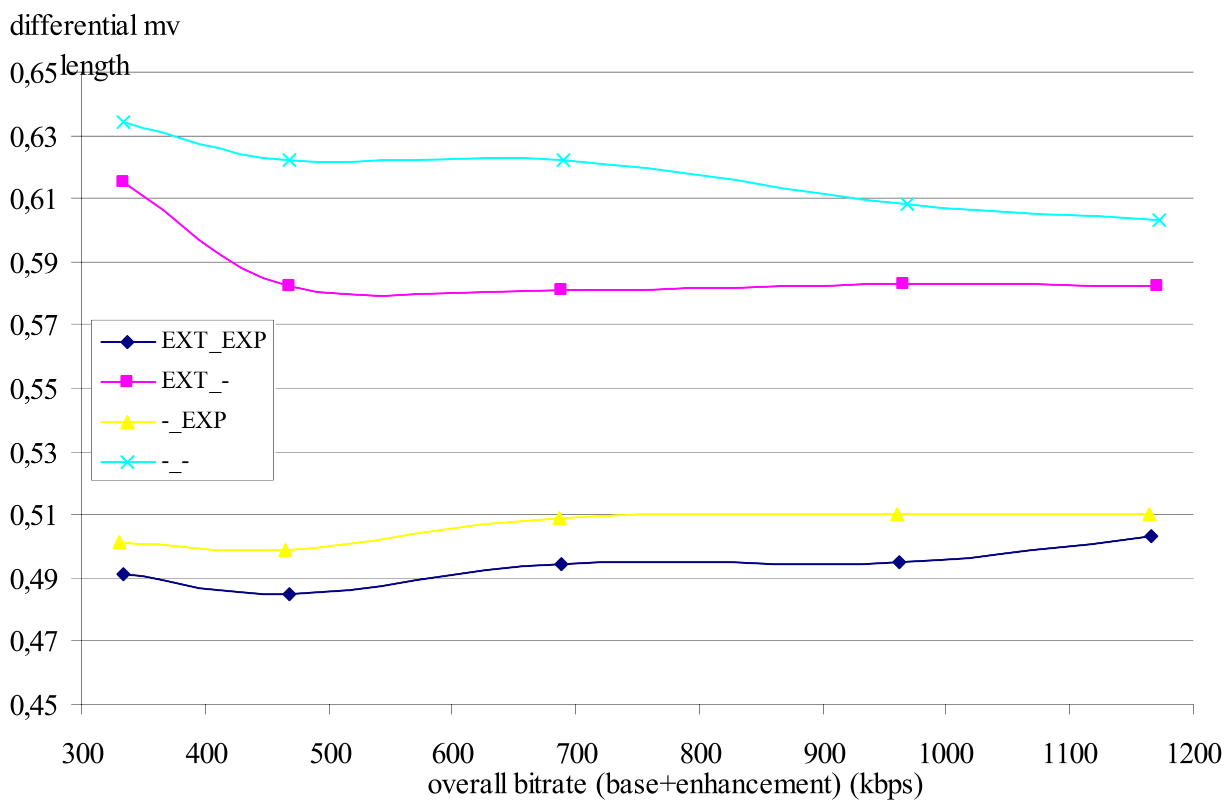

Fig. 17. Average differential motion vector length for the base and enhancement layer for various inter-layer motion prediction, MOBILE sequence, P-slices.

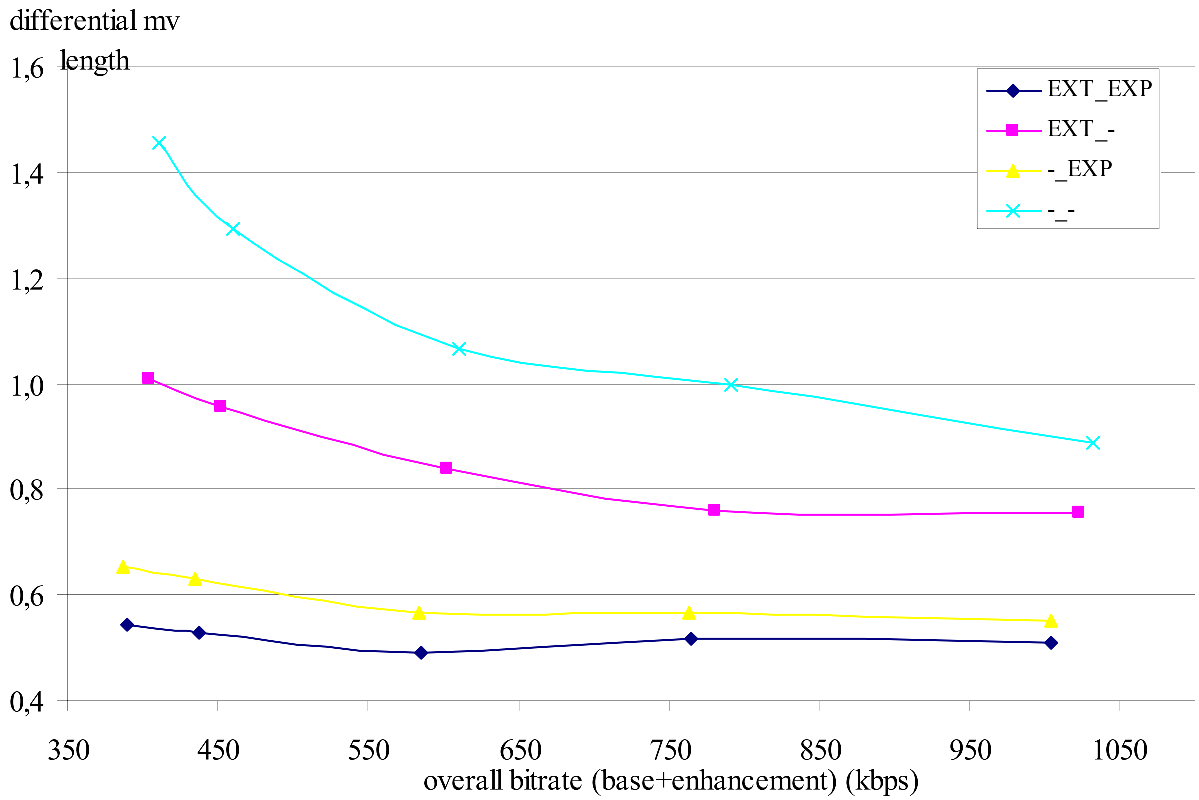

Fig. 18. Average differential motion vector length for the base and enhancement layer for various inter-layer motion prediction, BUS sequence, B-slices.

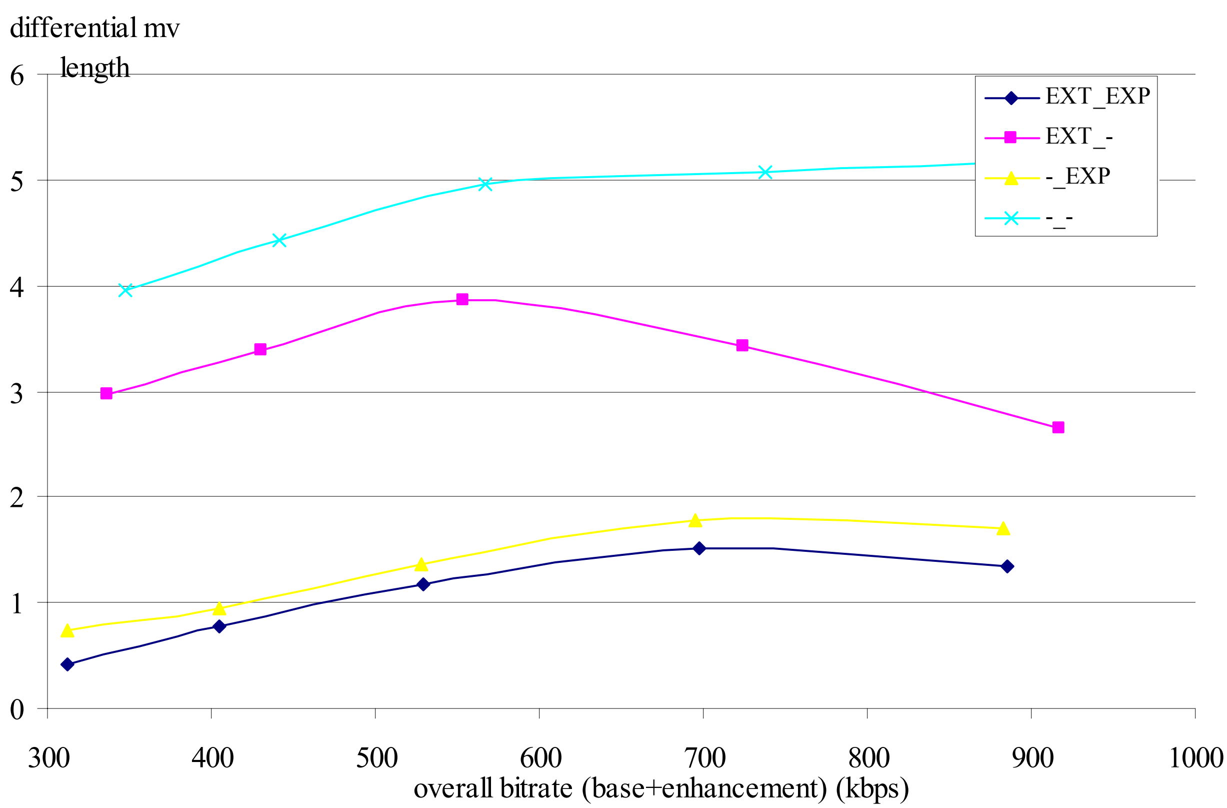

Fig. 19. Average differential motion vector length for the base and enhancement layer for various inter-layer motion prediction, FOOTBALL sequence, B-slices.

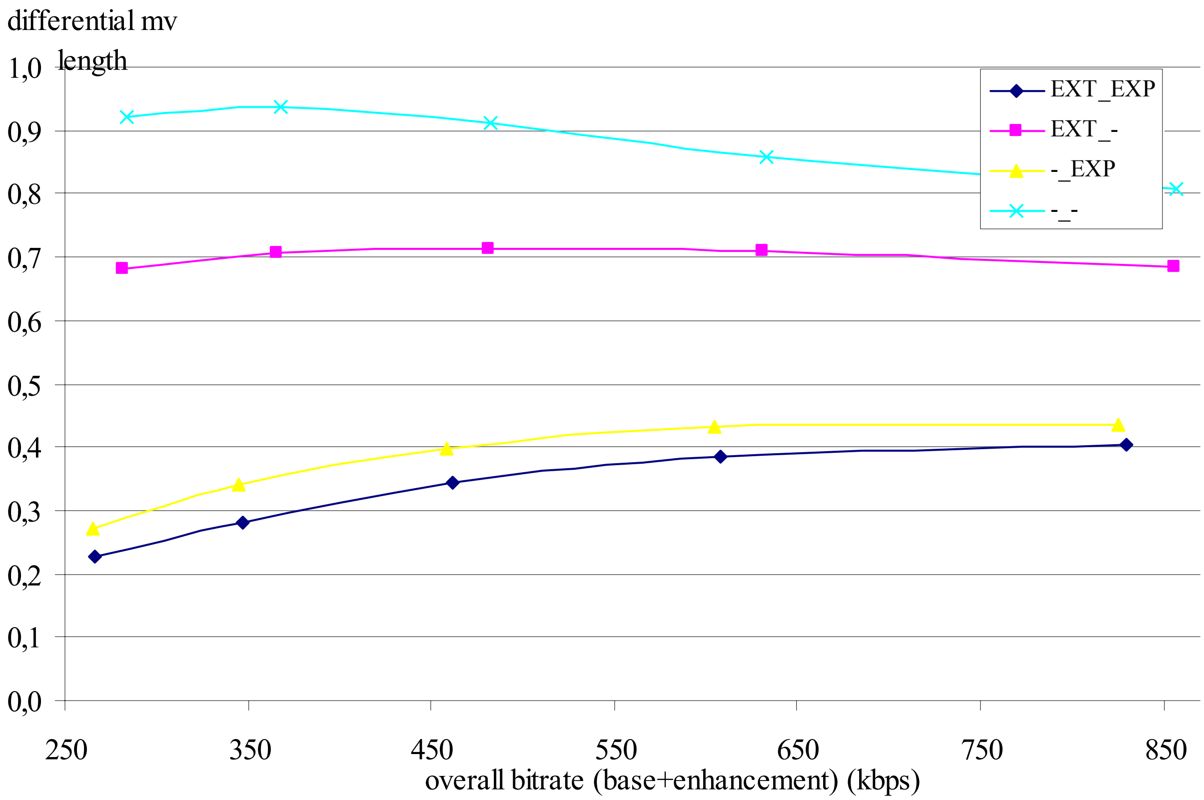

Fig. 20. Average differential motion vector length for the base and enhancement layer for various inter-layer motion prediction, FOREMAN sequence, B-slices.

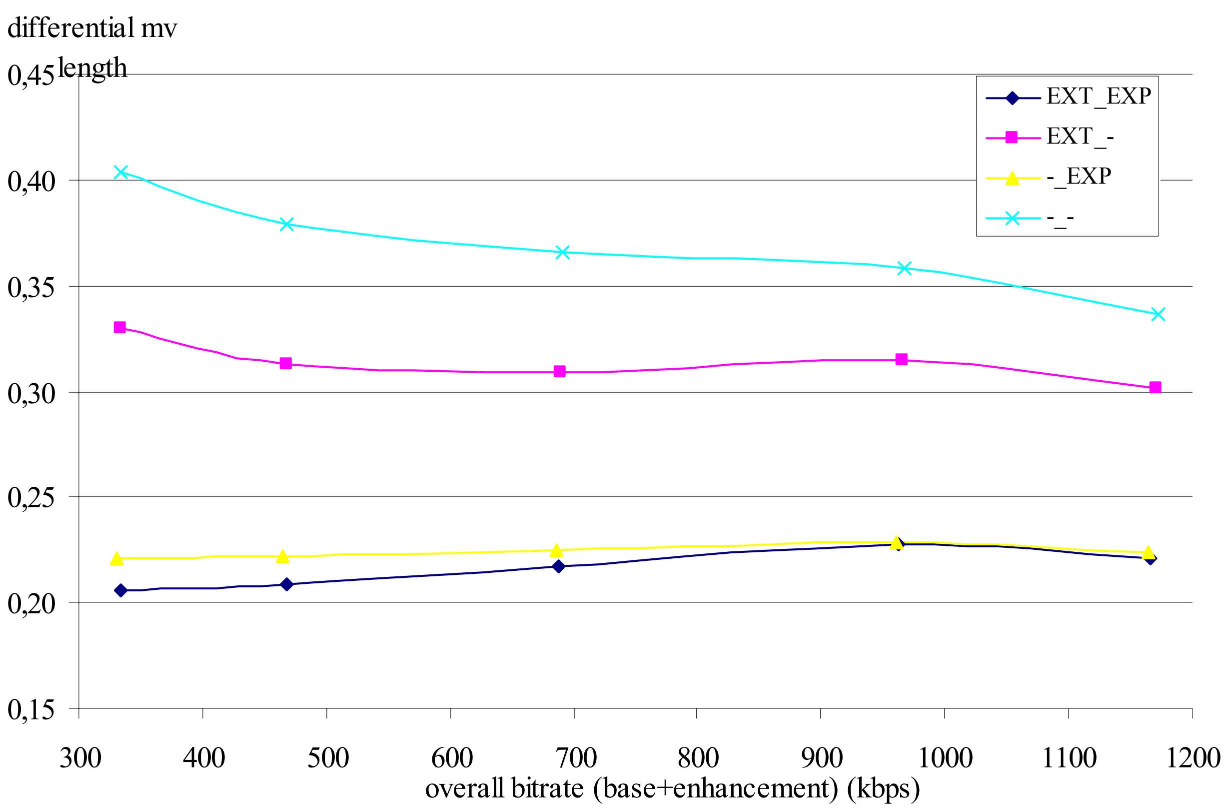

Fig. 21. Average differential motion vector length for the base and enhancement layer for various inter-layer motion prediction, MOBILE sequence, B-slices.

6. Conclusions

In the JSVM 4.0 software, extended method of inter-layer motion vector prediction has been implemented. Our proposal is to do standard intra-layer prediction with the co-located motion vector from the base layer whenever the respective blocks are:

- unavailable for motion prediction,
- intra-frame coded,
- use different reference frame for motion compensated prediction.

The proposal was tested as a standalone technique and together with the existing solution of inter-layer motion prediction. Thus four version of JSVM codec have been tested:
– codec without inter-layer motion prediction (“**-_-**“),
– codec with standard inter-layer motion prediction as defined in draft of SVC (“***-_EXP***“),
– codec with the proposed, extended technique of inter-layer motion prediction (“***EXT_-***“),
– codec with the jointly used: standard and extended technique of inter-layer motion prediction (“***EXT_EXP***“).

Surprisingly, there is no significant gain from using inter-layer motion prediction. Switching off the inter-layer motion prediction modes and removing from the scalable bitstream all syntax elements that signal inter-layer motion prediction modes (i.e. **base_mode_flag, base_mode_refinement_flag** and **motion_prediction_flag_lX**) causes the decrease of the luma PSNR up to 0.3 dB for BUS, FOOTBAL and FOREMAN sequences for the chosen configuration of the encoder. PSNR in the MOBILE sequence remains almost constant for all tested versions of inter-layer motion prediction.

In BUS and FOOTBALL sequences, our proposal produces visibly better results as compared with the encoder that does not exploit inter-layer correlation, however standard inter-layer motion prediction gives even better results (Fig. 2,3). Using jointly: proposed technique and the existing method of inter-layer motion prediction does not improve coding efficiently (Fig. 2-5).

As seen on Fig.6-9 using the proposed technique together with the existing method in most cases minimizes motion vector prediction residual in P-slices. In B- slices for BUS and FOOTBALL sequences, standard inter-layer motion prediction provides the best results, while for BUS and MOBILE sequences the best results give the proposed technique, implemented standalone (Fig. 10-13).

For all tested cases, the similarity between the motion vector field in the base layer and in the enhancement is the strongest when using the technique proposed together with the standard method (Fig. 14-21).

[1] ISO/IEC JTC1, “Joint Scalable Video Model (JSVM) 4.0, Reference Encoding Algorithm Description”, ISO/IEC JTC1/WG11 Doc. N7556, Oct. 2005, Nice, France

[2] R. Lange, “Extended inter-layer motion vectors prediction in scalable video coding – case study and improvement proposal”, JVT-R094, 18th Meeting: Bangkok, Thailand, 14-20 January, 2006