

*Title:* **Extended inter-layer motion vectors prediction in scalable video coding – case study and improvement proposal**

*Status:* Input Document to JVT

*Purpose:* Proposal

*Author(s) or* Rafal Lange

*Contact(s):* Poznan University of Technology  
Piotrowo 3a, 60-965 Poznań, Poland

Tel: +48 61 665 2870

Email: [rlange@et.put.poznan.pl](mailto:rlange@et.put.poznan.pl)

*Source:* Poznan University of Technology

---

## **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. Motion vectors from the base layer are rescaled and directly taken as vectors for motion compensation in enhancement layer. The  $\frac{1}{4}$ -pel refinement can be additionally sent in the bitstream for more precise motion compensation. Those two prediction modes are explicitly signaled in the scalable bitstream by the encoder. However, there are still some cases when even better performance of motion vectors prediction can be obtained by using the motion vectors from the base layer.

The proposal is to use during standard prediction, the co-located motion vector from the base layer whenever the neighboring blocks are not available for motion prediction. Proposed techniques of extended inter-layer prediction have reduced average motion vector prediction residuals up to 25%. Overall bitrate has also decreased in all tested cases.

## 1. Introduction

Predictive coding of the motion vectors is very important for the effective compression of the motion vectors field in a hybrid video codec. In the AVC/H.264 standard [3], sophisticated median motion vector prediction is used. On the other hand, the scalable video encoder that produces a bitstream with multiple layers of spatially scalable video sequence, provides even more motion data that has to be encoded. Because of the multiple motion vectors fields, the problem of effectively representing the motion data is becoming more crucial. Beside the intra-layer AVC-like motion vector prediction, additional inter-layer motion correlation can be exploited.

In the existing draft of Scalable Video Coding standard [2] and Joint Scalable Video Model JSVM-4 [1] there are two ways of exploiting inter-layer correlation between motion vectors from the base and the spatial enhancement layers. Motion vectors from the base layer are rescaled and directly taken as vectors for motion compensation in enhancement layer. The  $\frac{1}{4}$ -pel refinement can be additionally sent in the bitstream for more precise motion compensation. Those two prediction modes are explicitly signaled in the scalable bitstream by the encoder.

## 2. Intra-layer prediction of motion vectors in AVC/H.264

There are two basic motion vector prediction schemes that are used in AVC/H.264 [3]: directional prediction of motion vector and median prediction of motion vector. Directional prediction is used for macroblock modes with partitions of  $8 \times 16$  and  $16 \times 8$  luma samples (see Fig. 1). For all other partitioning modes of the macroblocks, the scalar median prediction of motion vector is used with motion vectors from blocks A,B,D or A,B,C as an input (see Fig. 2).

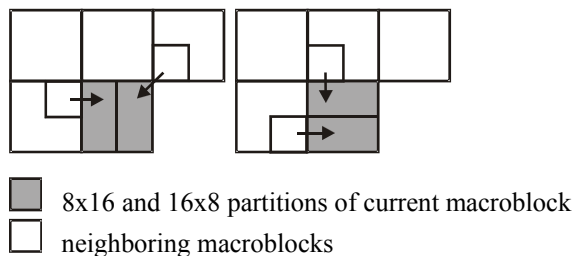


Fig. 1. Directional prediction of motion vectors

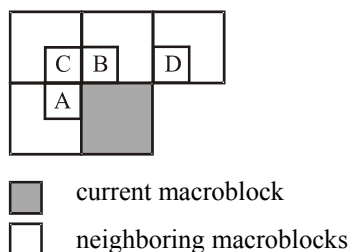


Fig. 2. Median prediction of motion vectors

Not always all neighboring blocks, of which motion vectors are used during prediction are available in the same slice as the current macroblock. Moreover, in some cases, neighboring macroblocks are intra coded or they use different reference picture for motion compensated prediction than the current block. In these cases, the standard scheme of motion vectors prediction is slightly modified. As a result for  $8 \times 16$  and  $16 \times 8$  partitioning modes, the median prediction is performed, while for other cases median prediction can be replaced by prediction from the single neighboring block.

Experiments prove that standard intra-layer prediction of motion vectors defined in AVC/H.264 works very well. On Fig. 3 and Fig. 4 histograms with absolute values of an average

motion vector prediction residual in the enhancement layer have been presented. The codec was configured to produce the scalable bitstream with 2 spatial resolution layers (QCIF and CIF). The averages of absolute values of components of motion vector residuals are given in Tab.2.

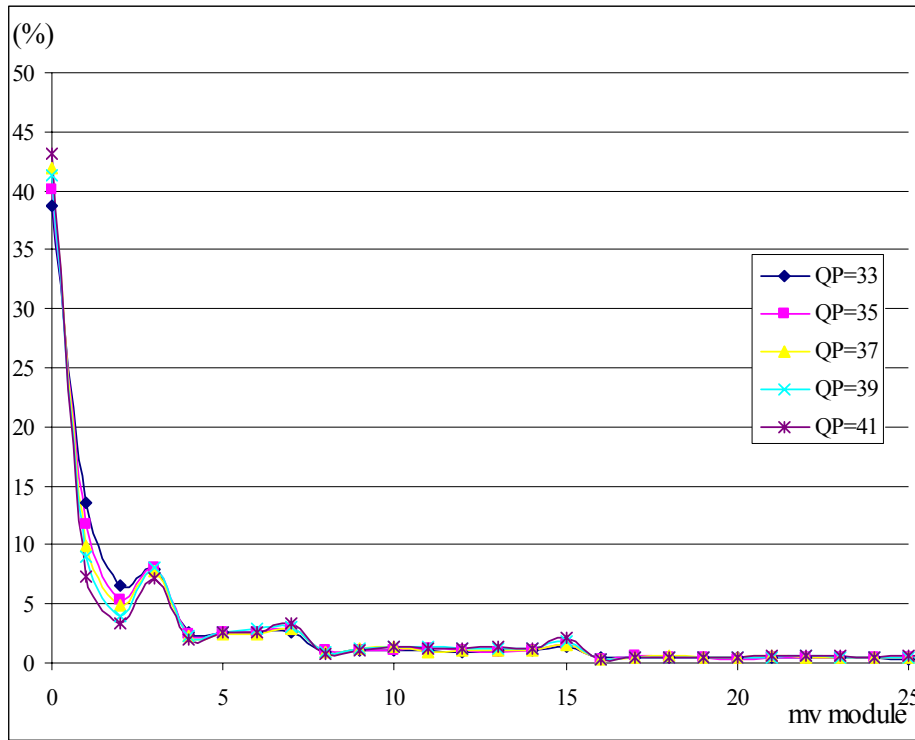


Fig. 3. Histogram of average motion vector prediction residual in P-slices, football sequence CIF 30 fps, various quality.

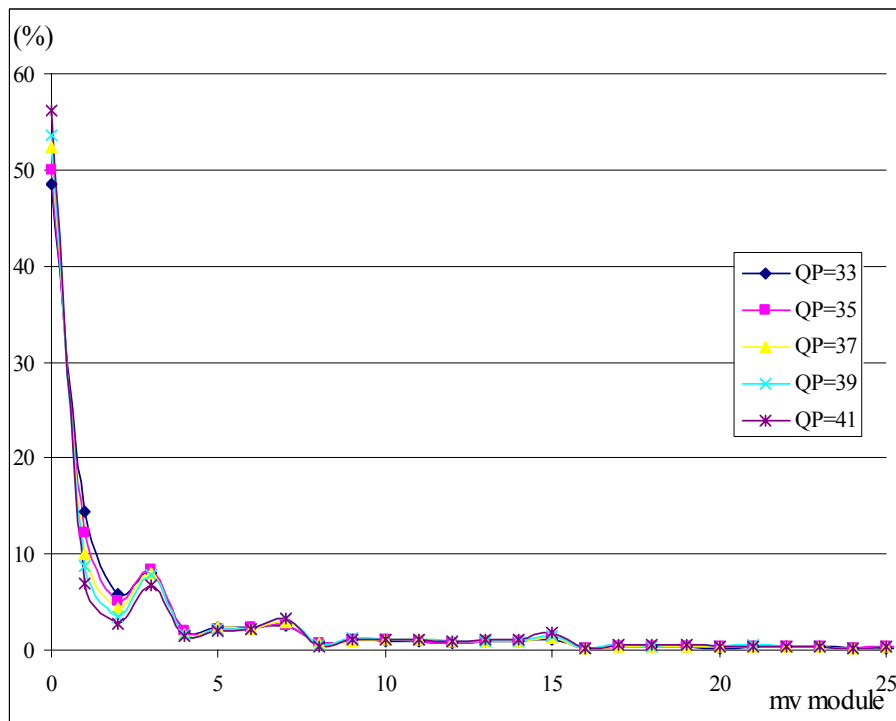


Fig. 4. Histogram of average motion vector prediction residual in B-slices, football sequence CIF 30 fps, various quality.

Depending on quantization parameter  $Q_p$  and the slice type (P or B), 40%-70% of the components of the motion vectors prediction residuals were zeros, which are very effectively represented in the bitstream. As the result, the room for improvements is very limited as it can be obtained only for remaining 60%-30% motion vectors residuals.

### 3. Inter-layer prediction of motion vectors in scalable extension of AVC/H.264

In current draft of scalable extension of AVC/H.264 (JSVM-4) [1,2] 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\_IX[mbPartIdx]** allows for switching on inter-layer prediction for selected partitions of the macroblock, with differentially encoded motion vector prediction error.

As one can see, all inter-layer motion vectors prediction modes are explicitly signaled in the bitstream.

### 4. Combined prediction of motion vectors

Here, improvements to the existing inter-layer motion vectors prediction scheme are proposed. As mentioned above, in most cases, existing intra-layer motion vectors prediction works very well. However, there are still some cases, when there is a possibility of poor prediction of motion vectors from the spatially neighboring blocks. For example, when the current macroblock is the first macroblock in the slice, there are no neighboring motion vectors for prediction. Another problem appears when neighbors are intra-coded or use different reference picture for motion compensated prediction. In such cases we can replace “missing” motion vector with the rescaled co-located motion vector from the base layer as depicted on Fig. 5.

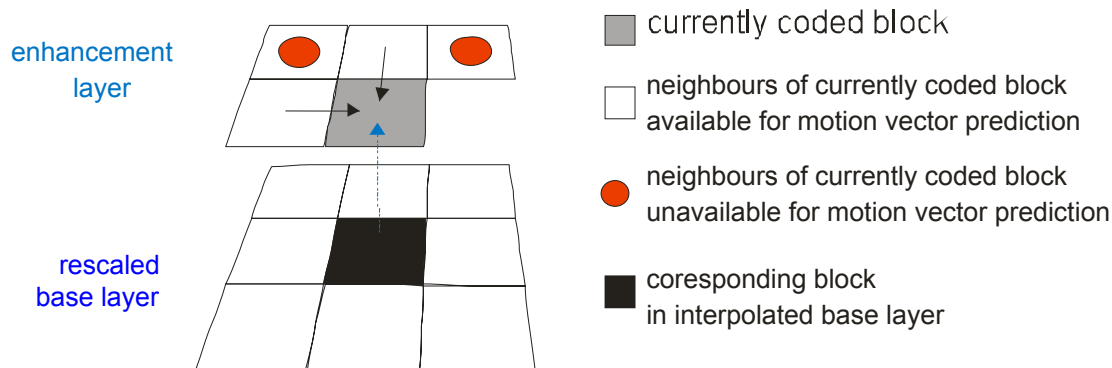


Fig. 5. Inter-layer motion prediction when neighboring blocks are unavailable for prediction.

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 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.

Tab. 1. Techniques of the proposed inter-layer motion vectors prediction.

		Vector from the base layer that is used for inter-layer prediction	
		Co-located	The nearest with the same reference frame
How many times vector from the base layer is used	Once	technique #3 (Inter-layer 3)	technique #4 (Inter-layer 4)
	Many times	technique #1 (Inter-layer 1)	technique #2 (Inter-layer 2)

We have tested the following variants of a replacement of the “missing” motion vector (see Tab. 1):

- using the motion vectors from the base layer only once during single prediction,
- using the motion vectors from the base layer as many times as many “missing” motion vectors appear during single prediction,
- using the directly co-located motion vector from the base layer,
- looking for the nearest motion vector in the base layer that uses the same reference frame index as the current motion vector (as depicted on Fig. 6).

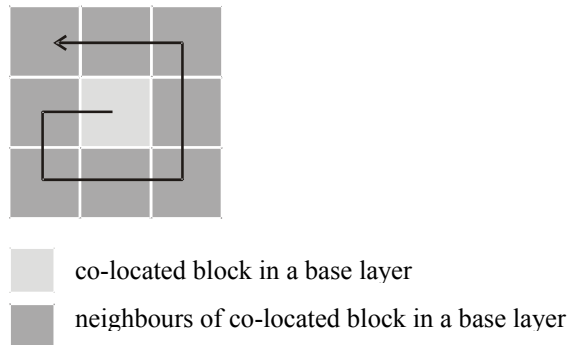


Fig. 6. Order of searching for the motion vector with the same reference frame index in the base layer

## 5. Results

The tests have been performed for the CIF sequences BUS, FOREMAN, FOOTBALL and MOBILE using AVC-based scalable video codec [4]. The coder was set up for producing two spatial layers (QCIF and CIF), both temporally scalable using B-frames. For each sequence the average motion vector prediction error have been measured (for P- and B- slices) as well as overall bitrate for various techniques of inter-layer motion vector prediction. Example test results for sequence FOOTBALL are shown in Tab. 2 and Tab.3. Average absolute values of motion vectors prediction residuals for various techniques of inter-layer motion prediction are depicted on Fig. 7 – Fig. 14. R-D curves for all test sequences are depicted on Fig. 15 – Fig. 18.

Tab. 2. Average absolute values of the components of the motion vector prediction residual in FOOTBALL sequence using various techniques of inter-layer motion prediction.

		FOOTBALL sequence, 176x144-352x288, IPBP				
		Quantization parameter				
		QP <sub>B</sub> =33, QP <sub>E</sub> =33	QP <sub>B</sub> =35, QP <sub>E</sub> =35	QP <sub>B</sub> =37, QP <sub>E</sub> =37	QP <sub>B</sub> =39, QP <sub>E</sub> =39	QP <sub>B</sub> =41, QP <sub>E</sub> =41
base layer		5.601	5.725	5.807	6.112	6.434
enhancement layer, prediction technique:	No inter-layer	6.421	6.864	7.239	7.677	8.332
	inter-layer 1	4.515	4.874	5.165	5.474	6.035
	inter-layer 2	4.912	5.160	5.481	5.692	6.160
	inter-layer 3	4.816	5.212	5.361	5.613	6.090
	inter-layer 4	4.949	5.145	5.553	5.914	6.162

Tab.3. Bitrate (kbps) and PSNR (dB) in FOOTBALL sequence using various techniques of inter-layer motion vectors prediction.

		FOOTBALL sequence, 176x144-352x288, IPBP					
		Quantization parameter					
		QP <sub>B</sub> =33, QP <sub>E</sub> =33	QP <sub>B</sub> =35, QP <sub>E</sub> =35	QP <sub>B</sub> =37, QP <sub>E</sub> =37	QP <sub>B</sub> =39, QP <sub>E</sub> =39	QP <sub>B</sub> =41, QP <sub>E</sub> =41	
base layer		320.5/30.54	240.8/30.29	182.5/27.86	133.6/26.57	94.8/25.28	
bitrate (kbps)/Y PSNR (dB)	enhancement layer, prediction technique:	no inter-layer	735.5/31.73	561.3/30.29	439.7/29.11	328.8/27.83	249.5/26.64
		inter-layer 1	724.8/31.74	553.1/30.31	431.3/29.12	322.0/27.85	242.9/26.66
		inter-layer 2	726.7/31.75	552.4/30.3	432.3/29.13	322.5/27.85	244.3/26.65
		inter-layer 3	726.1/31.74	553.1/30.31	431.9/29.13	322.8/27.86	243.3/26.66
		inter-layer 4	726.6/31.74	553.2/30.29	431.6/29.12	322.1/27.85	243.0/26.64

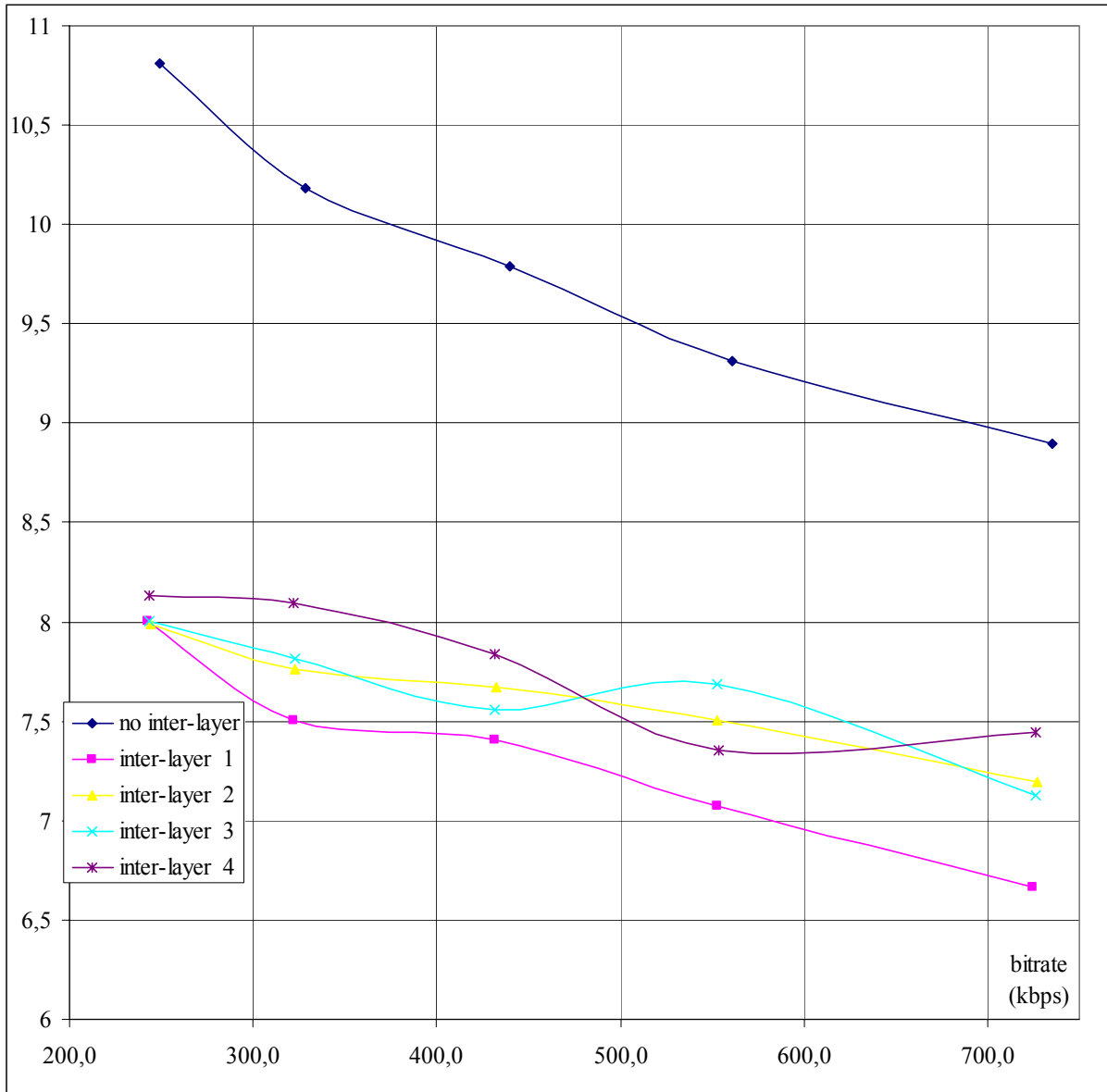


Fig. 7. Average absolute values of components of motion vector residual for various techniques of inter-layer prediction, FOOTBALL sequence, P-slices.

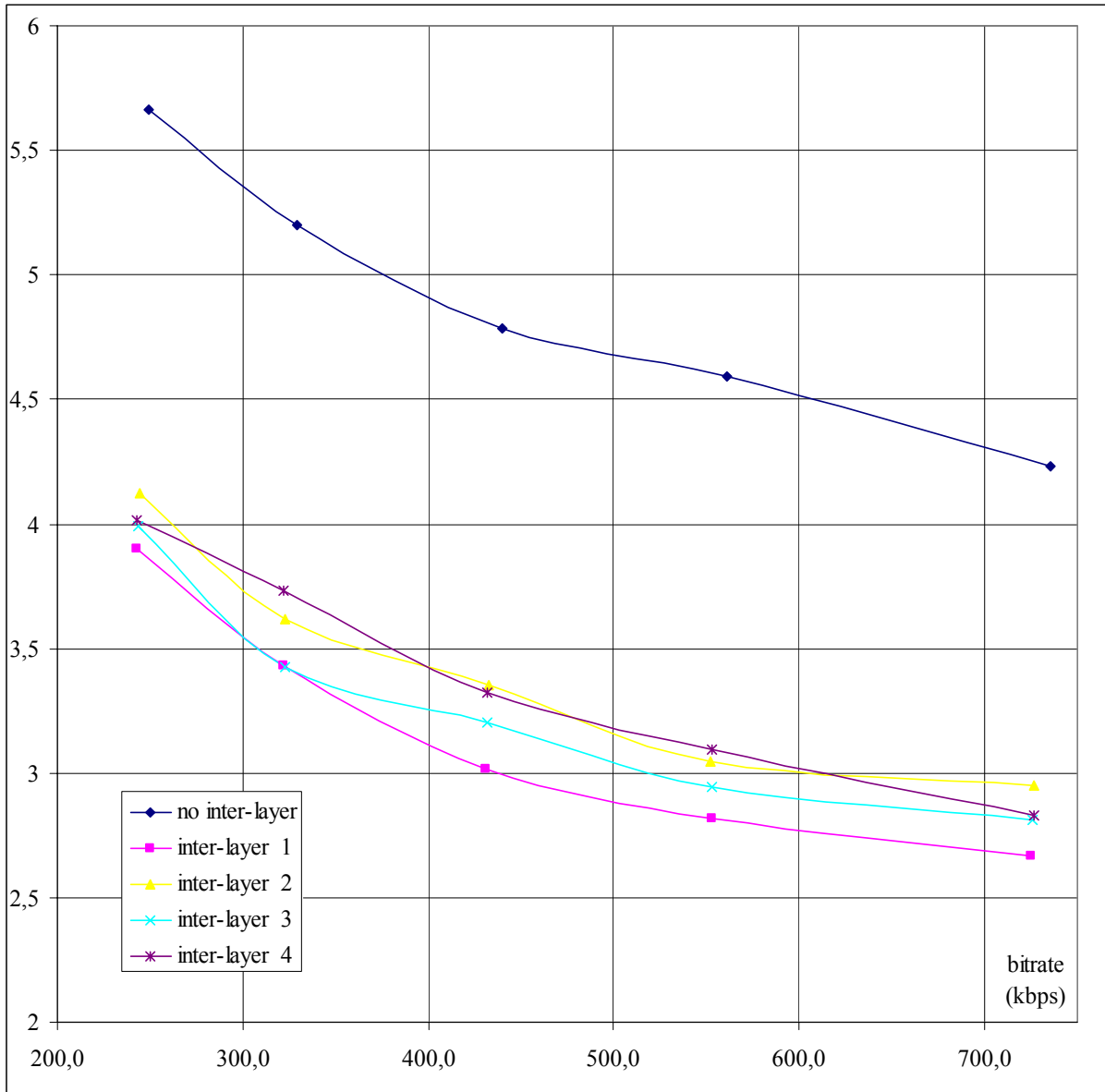


Fig. 8. Average Average absolute values of components of motion vector residual for various techniques of inter-layer prediction, FOOTBALL sequence, B-slices.

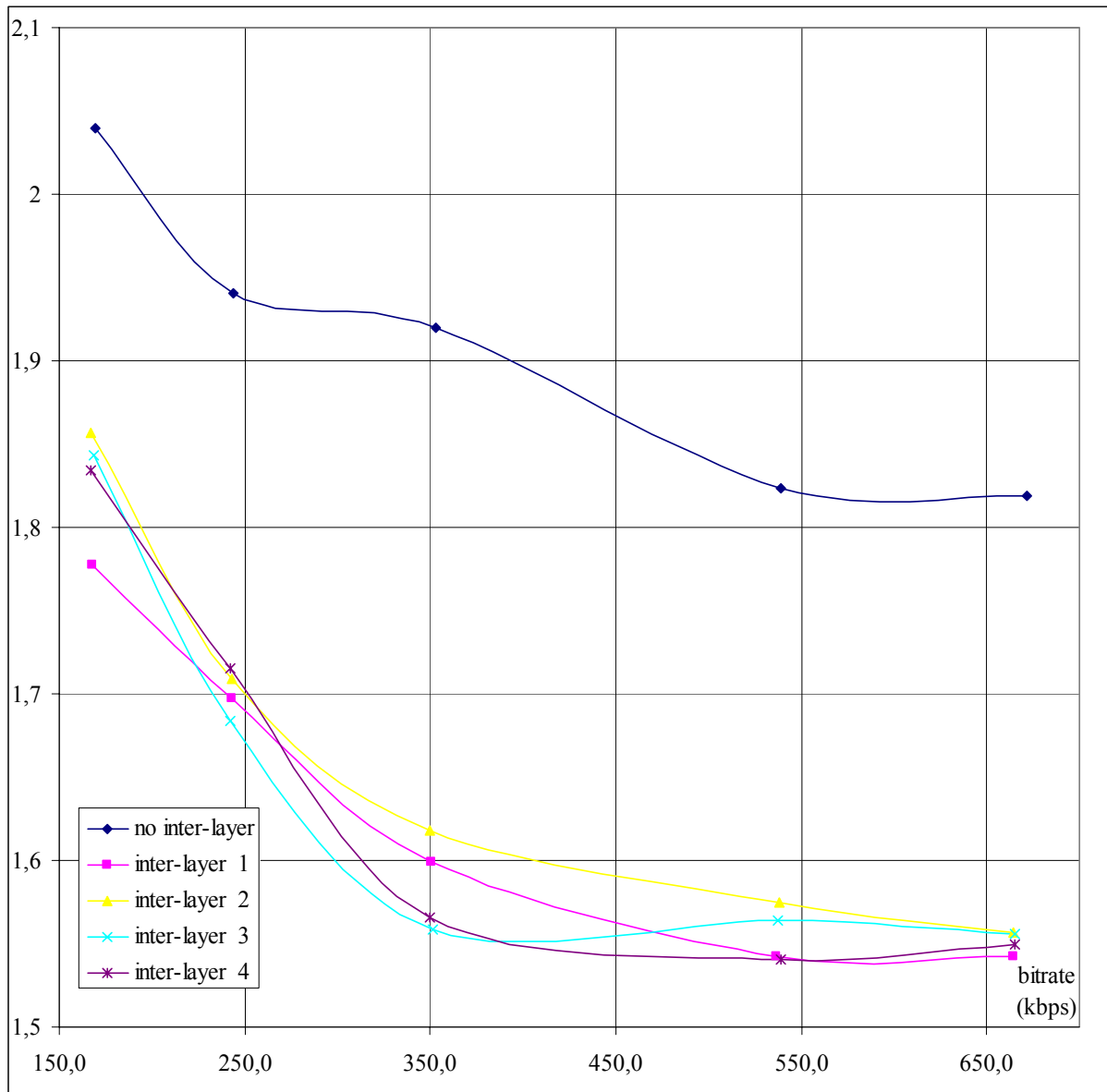


Fig. 9. Average absolute values of components of motion vector residual for various techniques of inter-layer prediction, MOBILE sequence, P-slices.

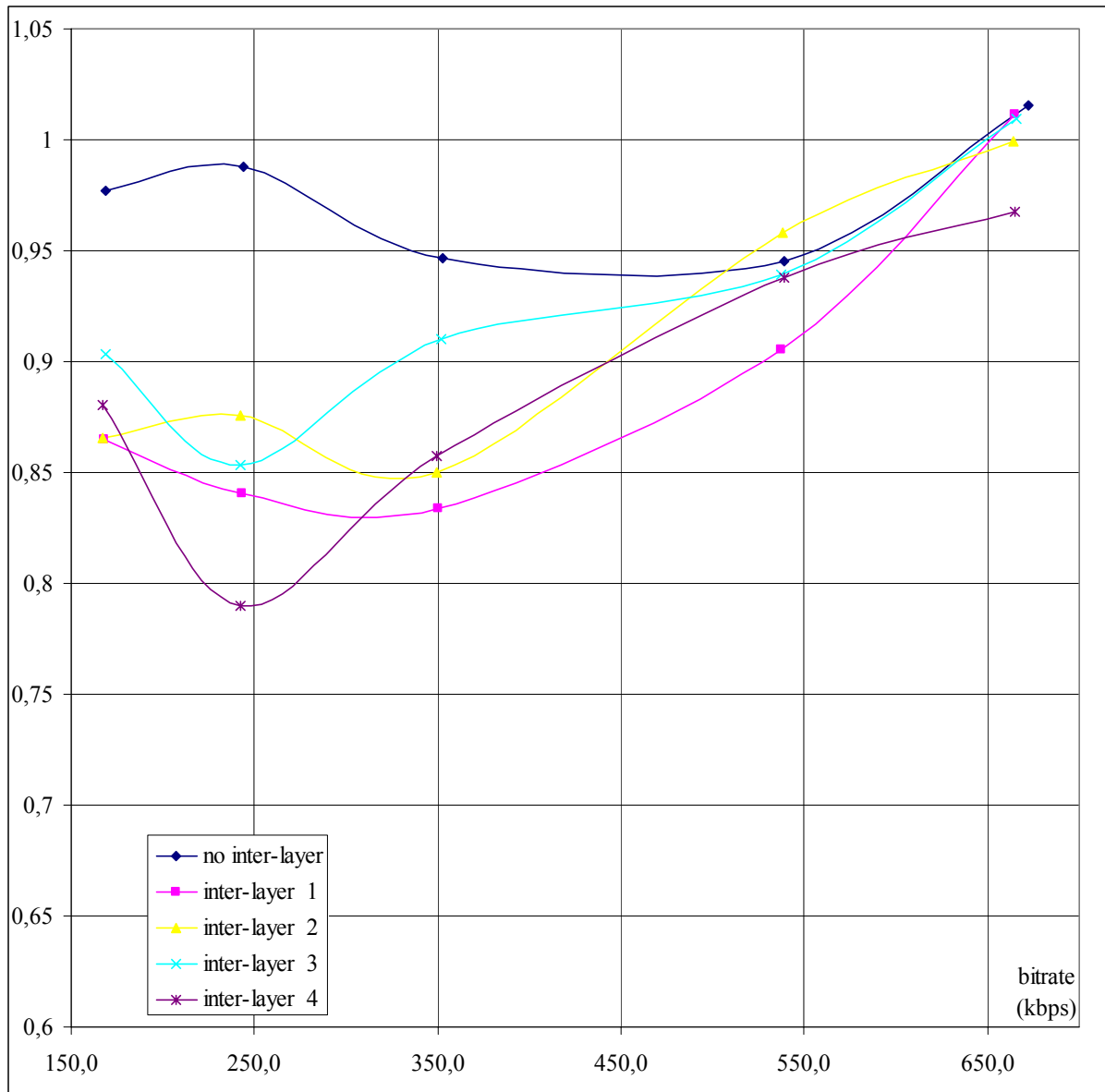


Fig. 10. Average Average absolute values of components of motion vector residual for various techniques of inter-layer prediction, MOBILE sequence, B-slices.

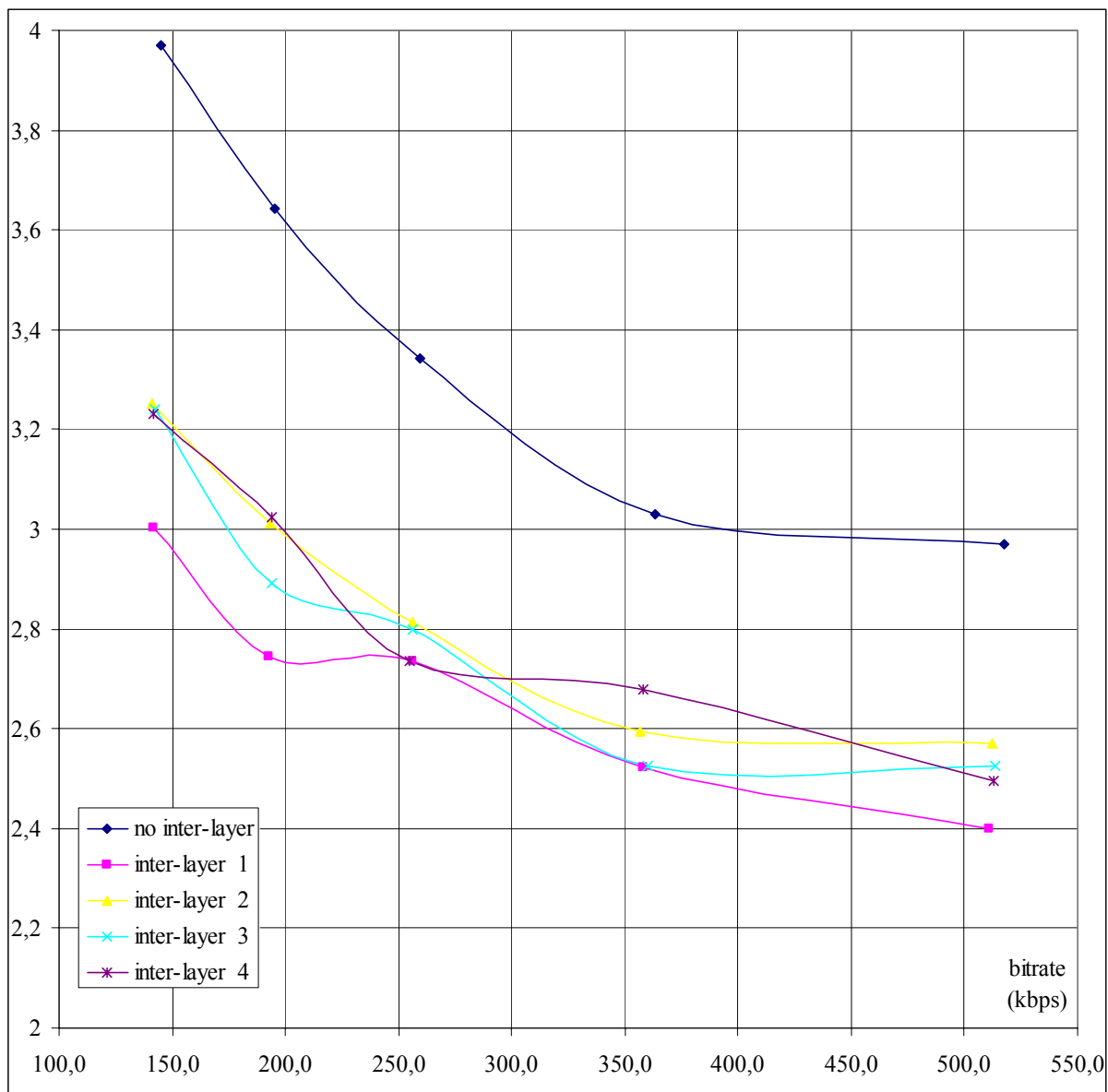


Fig. 11. Average absolute values of components of motion vector residual for various techniques of inter-layer prediction, FOREMAN sequence, P-slices.

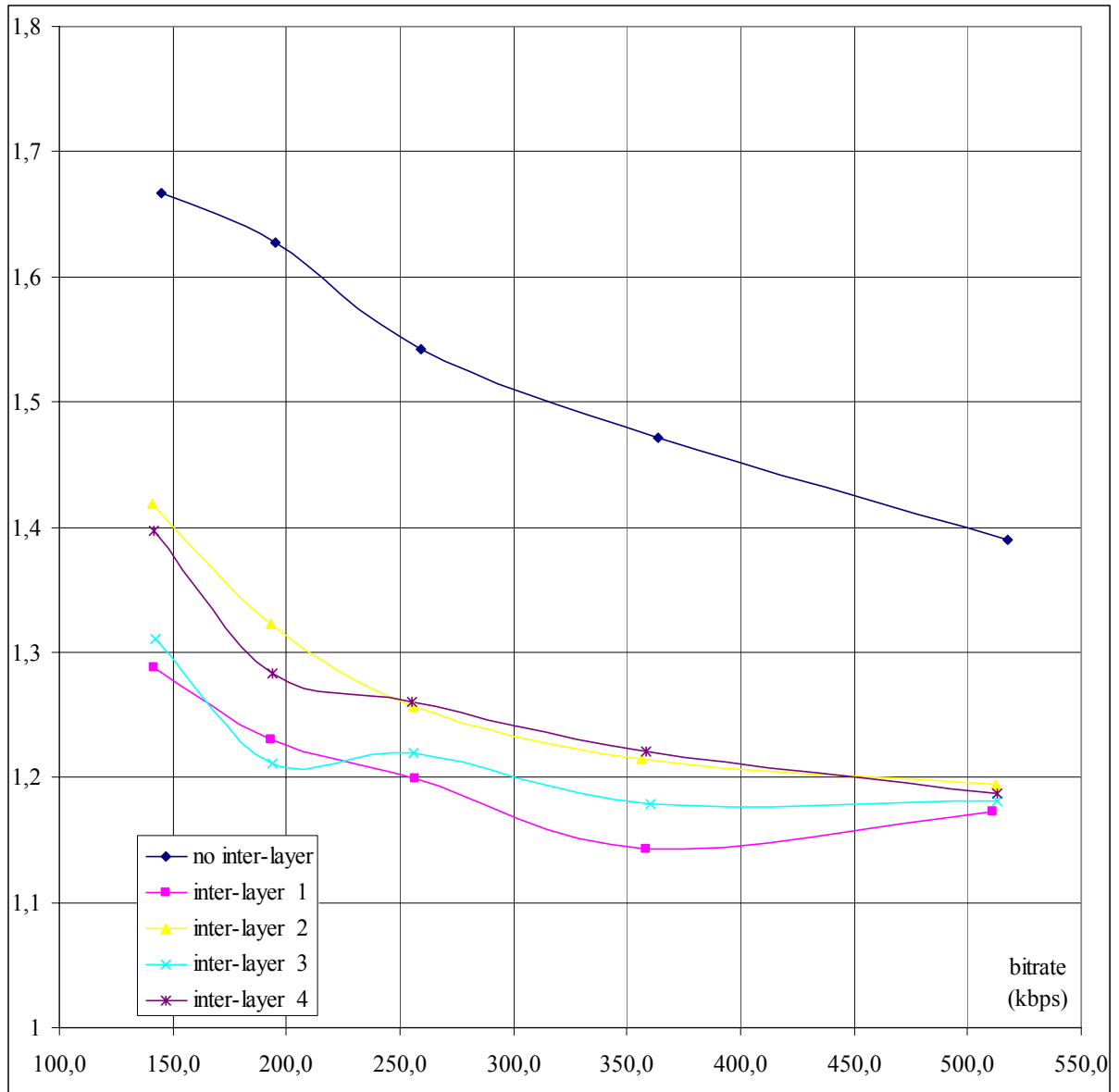


Fig. 12. Average Average absolute values of components of motion vector residual for various techniques of inter-layer prediction, FOREMAN sequence, B-slices.

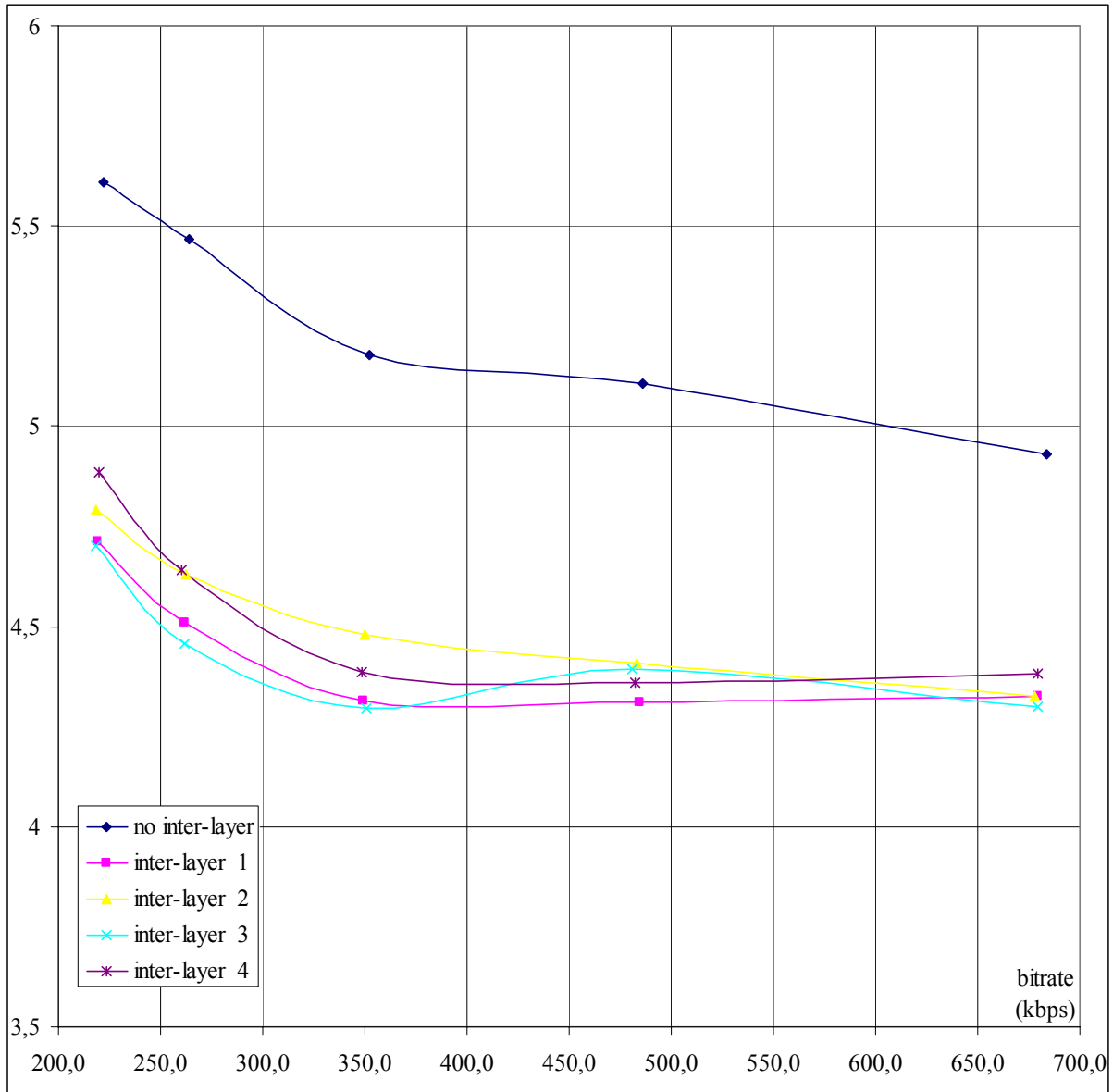


Fig. 13. Average absolute values of components of motion vector residual for various techniques of inter-layer prediction, BUS sequence, P-slices.

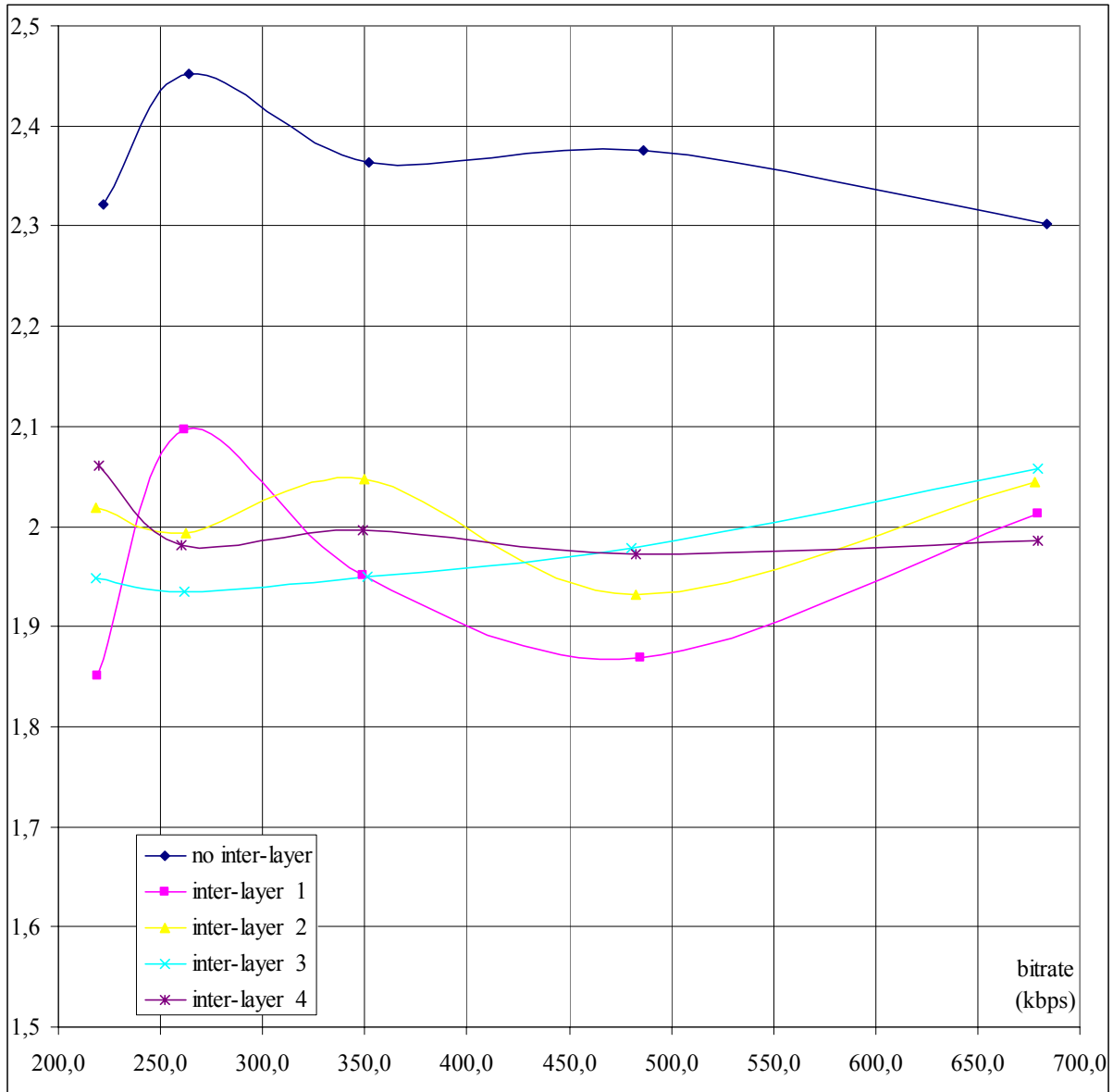


Fig. 14. Average Average absolute values of components of motion vector residual for various techniques of inter-layer prediction, BUS sequence, B-slices.

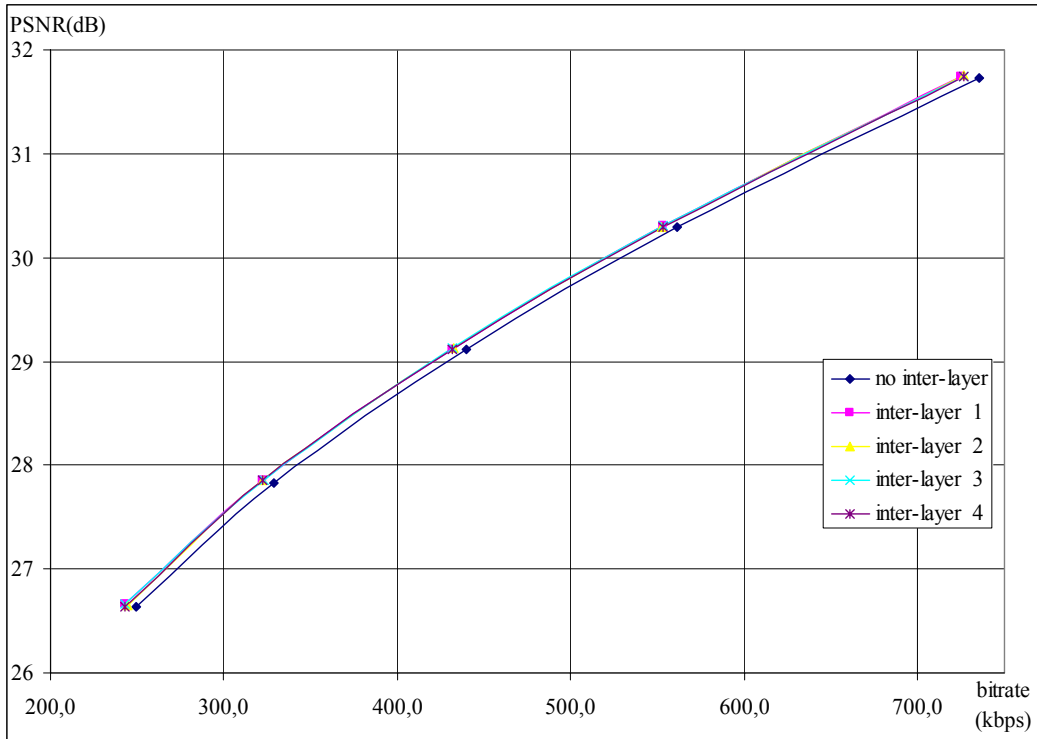


Fig. 15. R-D curves for various techniques of inter-layer predictions in FOOTBALL sequence.

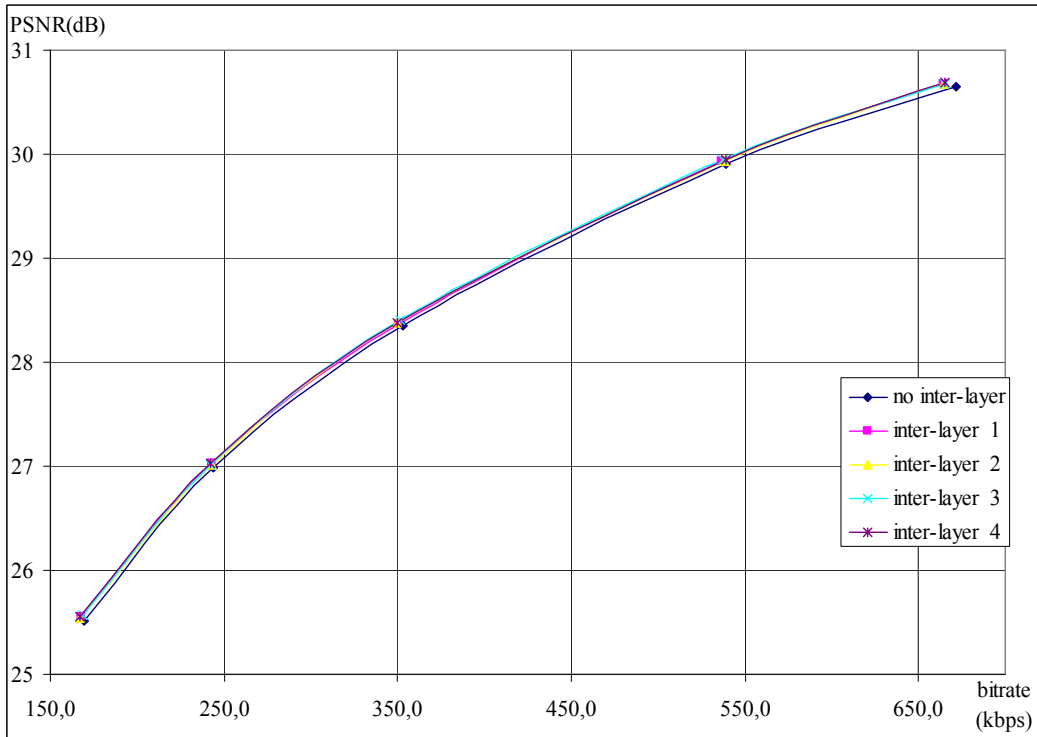


Fig. 16. R-D curves for various techniques of inter-layer predictions in MOBILE sequence.

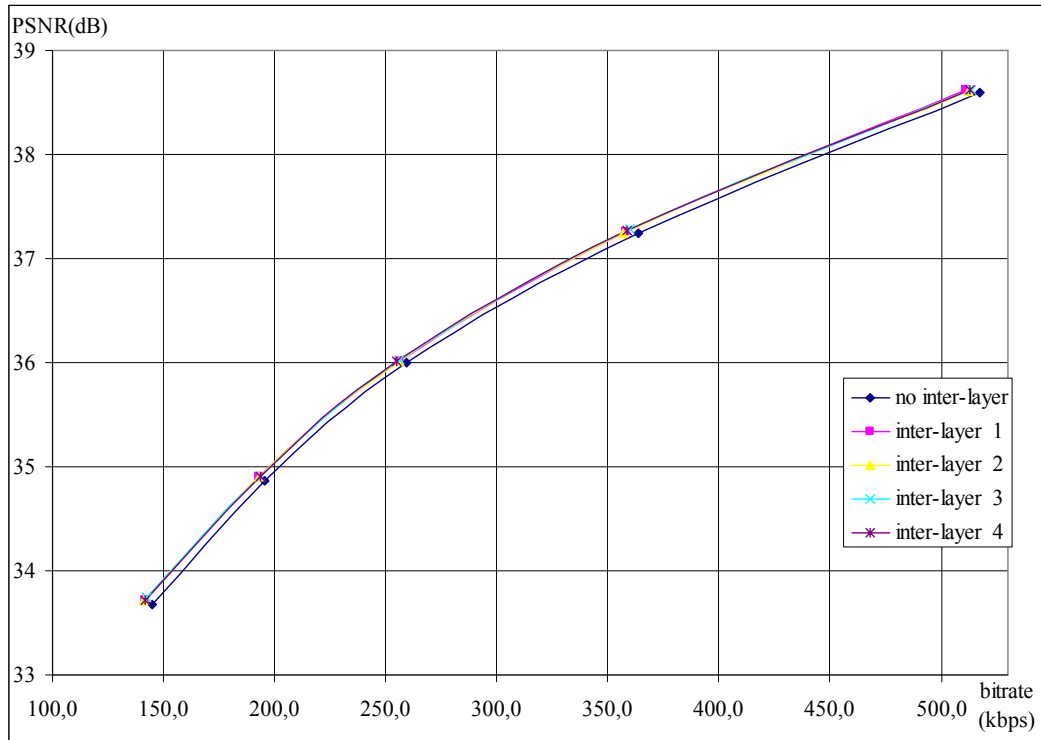


Fig. 17. R-D curves for various techniques of inter-layer predictions in FOREMAN sequence.

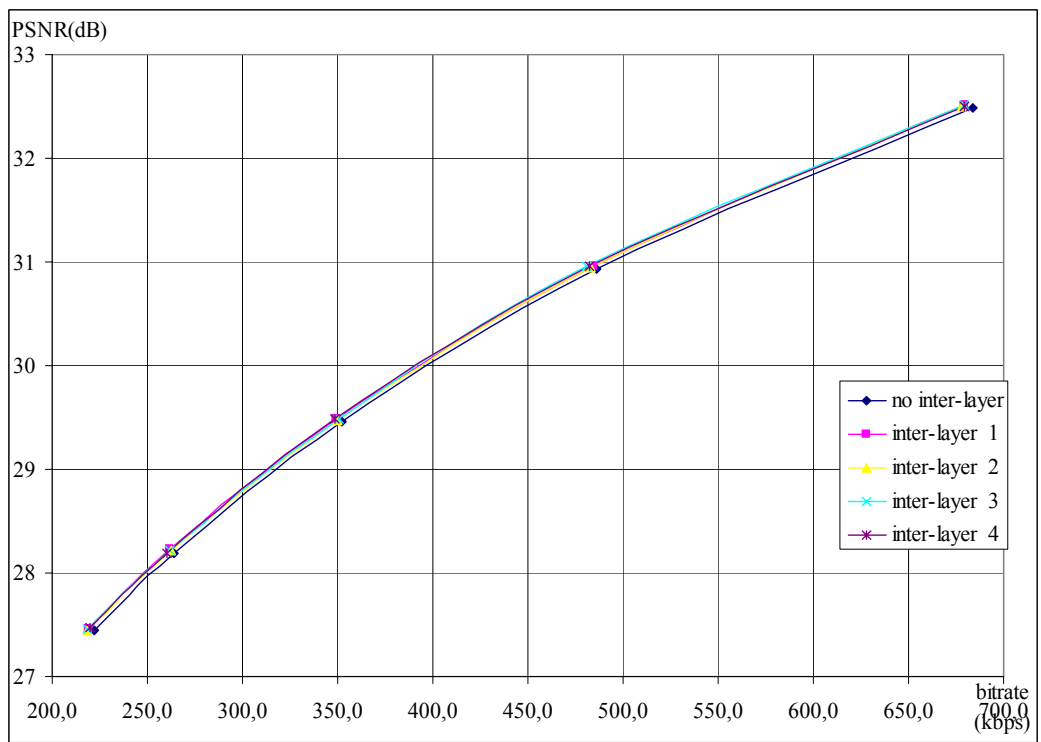


Fig. 18. R-D curves for various techniques of inter-layer predictions in BUS sequence.

Proposed techniques of inter-layer motion prediction significantly reduce motion vector residual (up to 25%, Fig. 7 – Fig. 14) and provide better performance of the scalable encoder (Fig. 15 – Fig. 18). In most cases technique #1 (inter-layer 1) outperforms other proposed techniques with the smallest motion vector prediction residual.

## 6. Conclusions

Standard prediction of motion vectors in AVC/H.264 works very efficiently. Inter-layer prediction which has been introduced into scalable extension of AVC/H.264, allows for further reducing the bitrate. Base layer motion vector prediction mode is signaled explicitly in the scalable bitstream, additionally the  $\frac{1}{4}$ -pel refinement can be sent by the encoder.

However, there are still some cases when even better performance of motion vectors prediction can be obtained by using the motion vectors from the base layer.

Our proposal is to use during standard prediction, the co-located motion vector from the base layer whenever the neighboring blocks are:

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

Proposed techniques of extended inter-layer prediction have significantly reduced motion vectors prediction residuals. Overall bitrate has also decreased in all tested cases.

The best among the proposed techniques is technique #1, which uses a motion vector of the directly co-located block in the base layer as many times as many “missing” motion vectors appear in the neighborhood of the current block.

The proposed technique of inter-layer motion prediction can be used as the extension of existing technique, which explicitly signals using base-layer motion vectors.

- [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] ISO/IEC JTC1, “Joint Scalable Video Model (JSVM) 4.0 – Annex G, Reference Encoding Algorithm Description”, ISO/IEC JTC1/WG11 Doc. N7556 – Annex G, Oct. 2005, Nice, France
- [3] ITU-T and ISO/IEC JTC1, “Advanced Video Coding for Generic Audiovisual Services”, ITU-T Recommendation H.264 - ISO/IEC 14496-10 AVC, 2003
- [4] Ł. Błaszak, M. Domański, R. Lange, A. Łuczak, „Scalable AVC Codec“, ISO/IEC JTC1/SC29/WG11 MPEG04/M10626, 15-19 March, Munich 2004.

(Append for Proposal Documents)

## JVT Patent Disclosure Form

International Telecommunication Union  
Telecommunication Standardization Sector



International Organization for Standardization



International Electrotechnical Commission



### Joint Video Coding Experts Group - *Patent Disclosure Form*

(Typically one per contribution and one per Standard | Recommendation)

Please send to:

JVT Rapporteur Gary Sullivan, Microsoft Corp., One Microsoft Way, Bldg. 9, Redmond WA 98052-6399, USA  
Email (preferred): [Gary.Sullivan@itu.int](mailto:Gary.Sullivan@itu.int) Fax: +1 425 706 7329 (+1 425 70MSFAX)

This form provides the ITU-T | ISO/IEC Joint Video Coding Experts Group (JVT) with information about the patent status of techniques used in or proposed for incorporation in a Recommendation | Standard. JVT requires that all technical contributions be accompanied with this form. *Anyone* with knowledge of any patent affecting the use of JVT work, of their own or of any other entity (“third parties”), is strongly encouraged to submit this form as well.

This information will be maintained in a “living list” by JVT during the progress of their work, on a best effort basis. If a given technical proposal is not incorporated in a Recommendation | Standard, the relevant patent information will be removed from the “living list”. The intent is that the JVT experts should know in advance of any patent issues with particular proposals or techniques, so that these may be addressed well before final approval.

This is not a binding legal document; it is provided to JVT for information only, on a best effort, good faith basis. Please submit corrected or updated forms if your knowledge or situation changes.

**This form is *not* a substitute for the *ITU ISO IEC Patent Statement and Licensing Declaration*, which should be submitted by Patent Holders to the ITU TSB Director and ISO Secretary General before final approval.**

<b><u>Submitting Organization or Person:</u></b>	
Organization name	<b>Poznan University of Technology</b>
	<b><u>Piotrowo 3a</u></b>
	<b><u>60-965 Poznań</u></b>
Mailing address	
Country	<b>Poland</b>
Contact person	<b>Rafal Lange</b>
Telephone	+48 61 665 2870
Fax	
Email	<a href="mailto:rlange@et.put.poznan.pl">rlange@et.put.poznan.pl</a>
Place and date of submission	Bangkok, Thailand, 14-20 January, 2006
<b><u>Relevant Recommendation   Standard and, if applicable, Contribution:</u></b>	
Name (ex: “JVT”)	
Title	
Contribution number	

(Form continues on next page)

**Disclosure information – Submitting Organization/Person (choose one box)**

2.0 The submitter is not aware of having any granted, pending, or planned patents associated with the technical content of the Recommendation | Standard or Contribution.

or,

The submitter (Patent Holder) has granted, pending, or planned patents associated with the technical content of the Recommendation | Standard or Contribution. In which case,

2.1 The Patent Holder is prepared to grant – on the basis of reciprocity for the above Recommendation | Standard – a free license to an unrestricted number of applicants on a worldwide, non-discriminatory basis to manufacture, use and/or sell implementations of the above Recommendation | Standard.

2.2 The Patent Holder is prepared to grant – on the basis of reciprocity for the above Recommendation | Standard – a license to an unrestricted number of applicants on a worldwide, non-discriminatory basis and on reasonable terms and conditions to manufacture, use and/ or sell implementations of the above Recommendation | Standard.

Such negotiations are left to the parties concerned and are performed outside the ITU | ISO/IEC.

2.2.1 The same as box 2.2 above, but in addition the Patent Holder is prepared to grant a “royalty-free” license to anyone on condition that all other patent holders do the same.

2.3 The Patent Holder is unwilling to grant licenses according to the provisions of either 2.1, 2.2, or 2.2.1 above. In this case, the following information must be provided as part of this declaration:

- patent registration/application number;
- an indication of which portions of the Recommendation | Standard are affected.
- a description of the patent claims covering the Recommendation | Standard;

*In the case of any box **other than 2.0** above, please provide the following:*

Patent number(s)/status \_\_\_\_\_

Inventor(s)/Assignee(s) \_\_\_\_\_

Relevance to JVT \_\_\_\_\_

Any other remarks: \_\_\_\_\_

*(please provide attachments if more space is needed)*

(form continues on next page)

Third party patent information – fill in based on your best knowledge of relevant patents granted, pending, or planned by other people or by organizations other than your own.

**Disclosure information – Third Party Patents** (choose one box)

3.1 The submitter is not aware of any granted, pending, or planned patents *held by third parties* associated with the technical content of the Recommendation | Standard or Contribution.

3.2 The submitter believes third parties may have granted, pending, or planned patents associated with the technical content of the Recommendation | Standard or Contribution.

*For box 3.2, please provide as much information as is known (provide attachments if more space needed) - JVT will attempt to contact third parties to obtain more information:*

3<sup>rd</sup> party name(s) \_\_\_\_\_

Mailing address \_\_\_\_\_

Country \_\_\_\_\_

Contact person \_\_\_\_\_

Telephone \_\_\_\_\_

Fax \_\_\_\_\_

Email \_\_\_\_\_

Patent number/status \_\_\_\_\_

Inventor/Assignee \_\_\_\_\_

Relevance to JVT \_\_\_\_\_

Any other comments or remarks: