**Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG**
**(ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6)**
22nd Meeting: Marrakech, Morocco, 13 – 19 January, 2007

Document: JVT-V080
Filename: JVT-V080.doc

*Title:* **CE10: Multi-view Video Coding using View Interpolation Method**

*Status:* Input Document to JVT

*Purpose:* Proposal

*Author(s) or Contact(s):*
Yo-Sung Ho,
Cheon Lee,
Kwan-Jung Oh,
*Byeongho Choi,
and *Ji Ho Park

Gwangju Institute of Science and Technology (GIST)
1 Oryong-dong, Buk-gu, Gwangju,
500-712, Republic of Korea

*Korea Electronics and Technology Institute (KETI)
#68 Yatap-dong, Bundang-gu,
Seongnam-si, Gyeonggi-do,
463-816, Republic of Korea

Tel: 82-62-970-2211
Email: hoyo@gist.ac.kr
leecheon@gist.ac.kr
kjoh81@gist.ac.kr
bhchoi@keti.re.kr
scottie@keti.re.kr

*Source:* GIST and KETI

---


# Abstract

In this document, an additional picture coding method named 'VIP P-picture' coding is described based on the proposed view interpolation by GIST and KETI for multi-view video coding. The proposed view interpolation method can improve quality of interpolated images using initial disparity estimation, variable block-based disparity estimation, and pixel-level disparity estimation using adjusted search range. The interpolated intermediate images are used as reference frames for multi-view video coding. 'VIP P-picture' coding has been included and motion vector prediction has been modified in order to exploit the interpolated image.


# 1. Introduction

Multi-view video sequences are captured by two or more adjacent cameras simultaneously. Therefore, there is high spatial correlation among adjacent view images. Most prediction structures are trying to exploit this inter-view correlation. In the prediction structure proposed by Fraunhofer-HHI, shown in Fig. 1, each picture in S1, S3 and S5 have four reference frames: two temporal frames and two adjacent view frames [1]. Since these three views can be considered as inter-view, we can generate intermediate view images using adjacent view images. The generated intermediate view images can be used as reference frames in the disparity and motion compensation process. It is obvious that a high-quality intermediate image guarantees improvement of coding efficiency. In this document, we describe briefly the proposed view interpolation scheme for multi-view video coding, and we propose 'VIP P-picture' coding scheme to use the interpolated image.

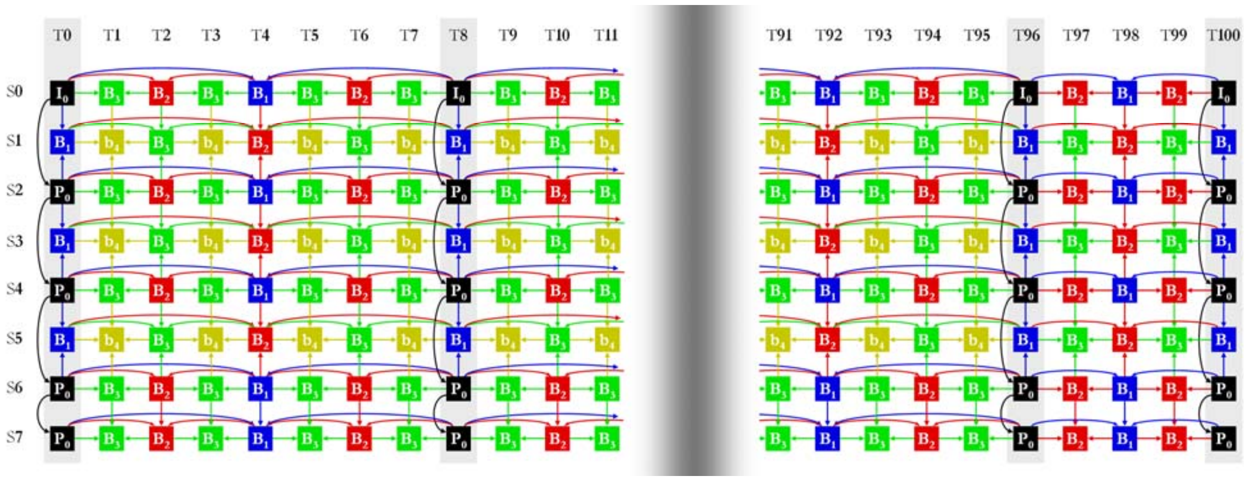



Fig. 1: Inter-view-temporal Prediction Structure

## 2. Efficient View Interpolation Method for Multi-view Coding

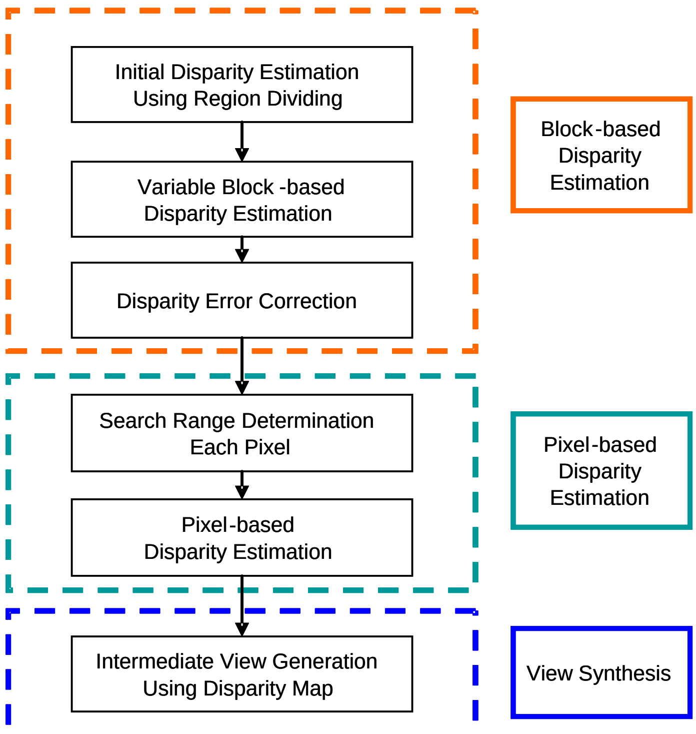



Fig. 2. Proposed View Interpolation Method

Previous interpolation methods have several weaknesses: (a) it is difficult to set the maximum search range before the disparity estimation operation, and (b) the accuracy of disparity estimation is poor in boundary regions. We have proposed an improved view interpolation scheme, as shown in Fig. 2. The first step is block-based disparity estimation which does not consider the maximum disparity search range because it is a region dividing method. We also use variable block sizes for disparity estimation, which is more effective in object boundary regions. The last step for disparity estimation is pixel-level estimation using adjusted search ranges. After these three steps, we interpolate inter-view images using the estimated disparity information. During the whole processes, the disparity errors are corrected using the median filtering. For more details, see the reference document [2].

# 3. Multi-view Video Coding Method using Interpolated Image

## 3.1 View Interpolated Prediction P-picture Coding

View interpolation prediction is applied to inter views such as S1, S3, and S5 in Fig. 1 because it is possible when adjacent views are already coded. After generating the intermediate image, it can be used as an additional reference frame. One of advantages of the intermediate image is that it is mostly overlapped with the image to be coded. We propose the view interpolated prediction P-picture (VIP P-picture) coding scheme and It performs like as P-picture coding. VIP P-picture coding is composed of additional macroblock modes: 'VIP_SKIP', 'VIP_16x16', 'VIP_8x16', 'VIP_16x8', and 'VIP_P8x8'. VIP_SKIP mode refers the co-located block in the intermediate image and it does not need a motion vector prediction. Figure 3 shows the overall coding structure using the VIP image.

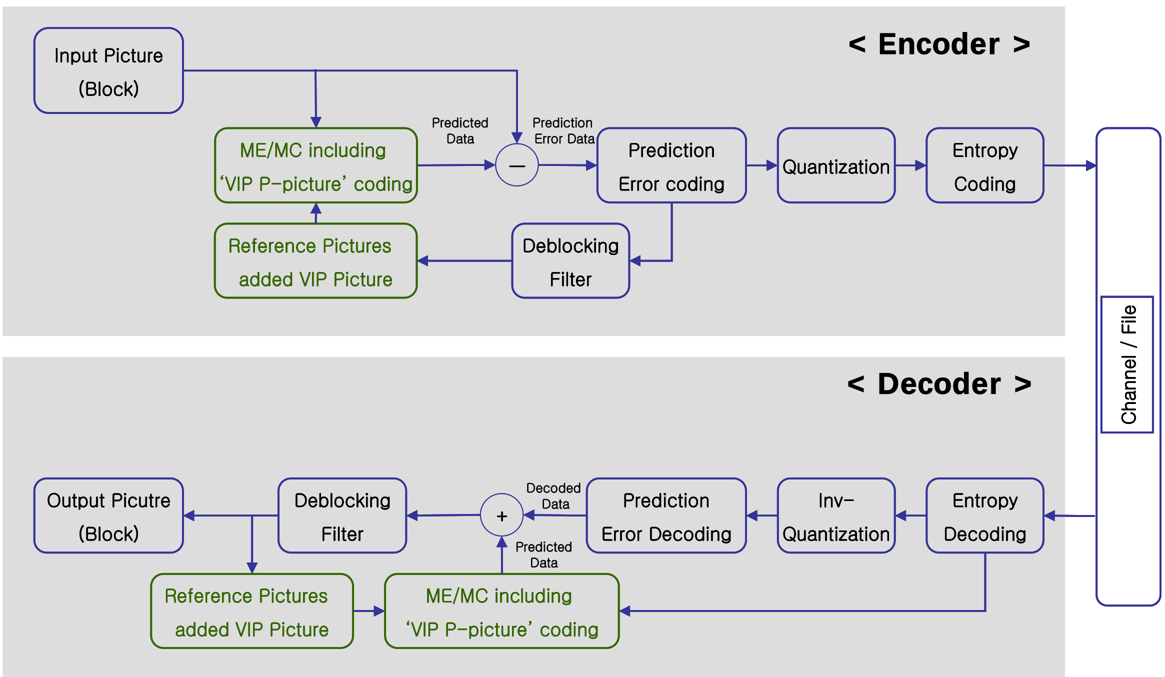
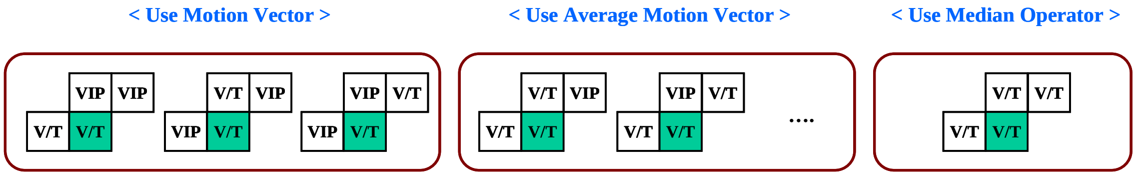


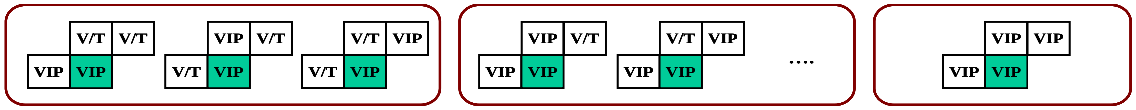

Fig. 3 Coding Structure added 'VIP P-picture' coding

## 3.2 Modified Motion Vector Prediction

The H.264/AVC encodes the difference value between the motion vector value of the current macroblock and predicted motion vector value using the motion vectors of neighboring macroblocks. The reason why this approach is efficient is that all reference frames has the similar correlation with the frame to be coded in H.264/AVC. On the other hand, the proposed coding scheme employs three types of reference frames: temporal frame (T frame), spatial frame (V frame), and VIP frame. Since three types of reference frames have different disparity values, the current motion vector prediction scheme can not improve the coding gain efficiently. We propose a modified motion vector prediction method which predicts a motion vector more using division of motion vectors between VIP frame and V/T frames. Fig. 4 shows the modified motion vector prediction methods. If a reference frame is one of V/T frames, the motion vector predictor considers only blocks referring those of V/T frames as shown in Fig. 4 (a). Similarly, if a reference frame is VIP frame, motion vector predictor only considers neighboring blocks referring VIP frame, shown in Fig. 4. (b). If the number of neighboring blocks having the same kinds of reference frame is one, motion vector predictor uses its motion vectors directly. In case of two, the average value of motion vectors is used. In case of three, the median value of motion vectors is used.




(a) If the reference frame is one of V/T frames




(b) If the reference frame is VIP frame

Fig. 4 Modified Motion Vector Prediction Method

## 4. Experimental Results

According to the description in [3], four QPs (22, 27, 32, 37) have been used for four sequences which are good enough to generate the intermediate image. We have implemented the proposed view interpolation scheme and 'VIP P-picture' coding method using JMVM 1.0 software. We have experimented four sequences only 1 second because of the lack of time. 'Rena' and 'Akko&Kayo' sequences are encoded for 31 frames, and 'Exit' and 'Ballroom' sequences are encoded for 25 frames. The search range for VIP is 48. Following 6 figures show the rate-distortion curves.

Fig. 5 and Fig. 6 show the coding results for the views adopted 'VIP P-picture' coding. According to Fig. 5 and Fig. 6, the coding efficiency has improved about 1dB in low bits rate, but it has little improved in high bits rate. The average PSNR for 'Rena' and 'Akko&Kayo' in low bits rate has improved about 0.3 dB as shown in Fig. 7 and Fig. 8. The average PSNR of interpolated images of 'Rena' and 'Akko&Kayo' are around 32 dB. 'VIP_SKIP' mode is chosen in these sequences over 30%. The results of 'Ballroom' and 'Exit' are similar to the results of the reference software because the PSNR of the interpolated image is not so good. More detail experimental results are provided in the attached Excel file.

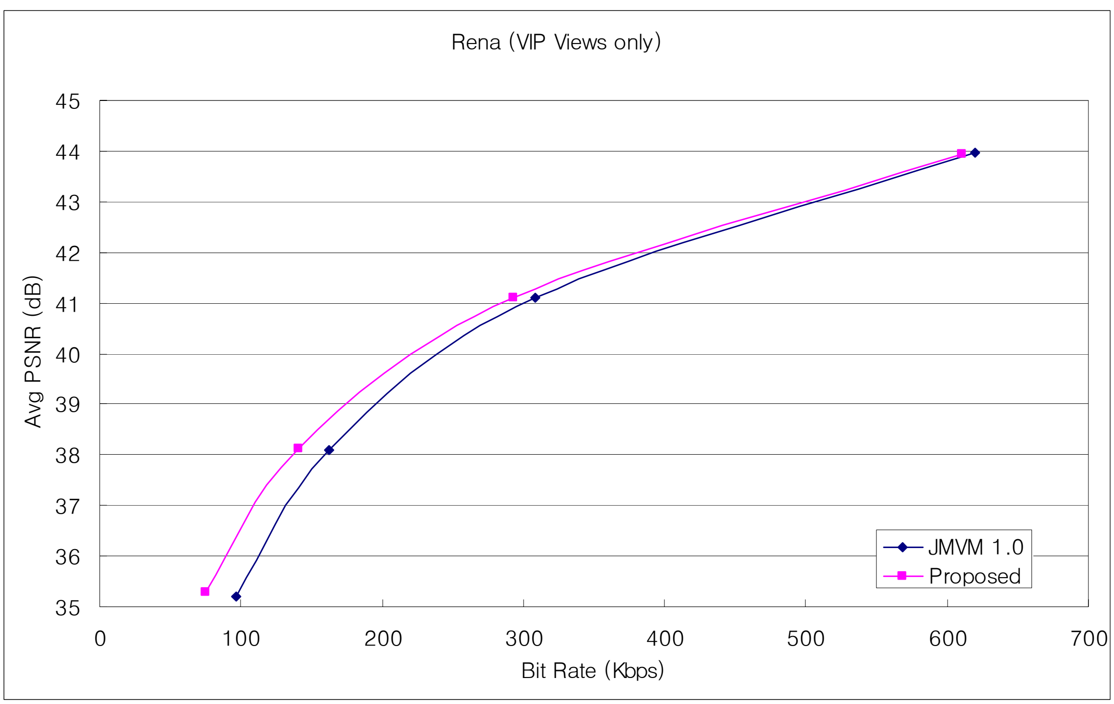



Fig. 5. Rate-Distortion Curve for Rena (only VIP views)

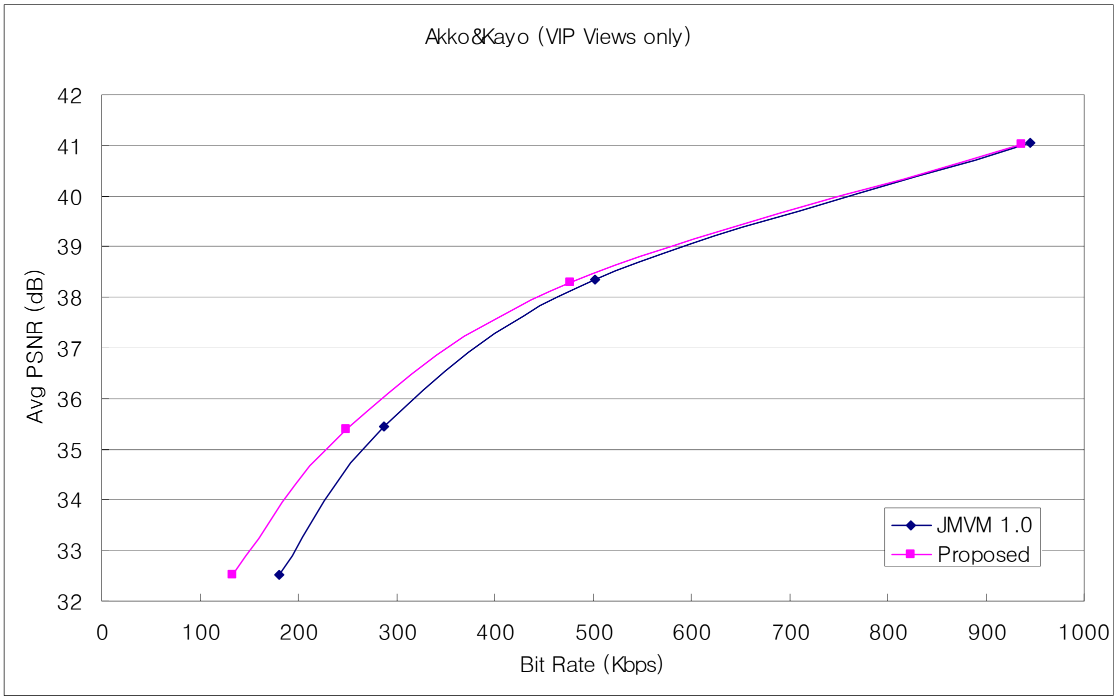



Fig. 6. Rate-Distortion Curve for Akko&Kayo (only VIP views)

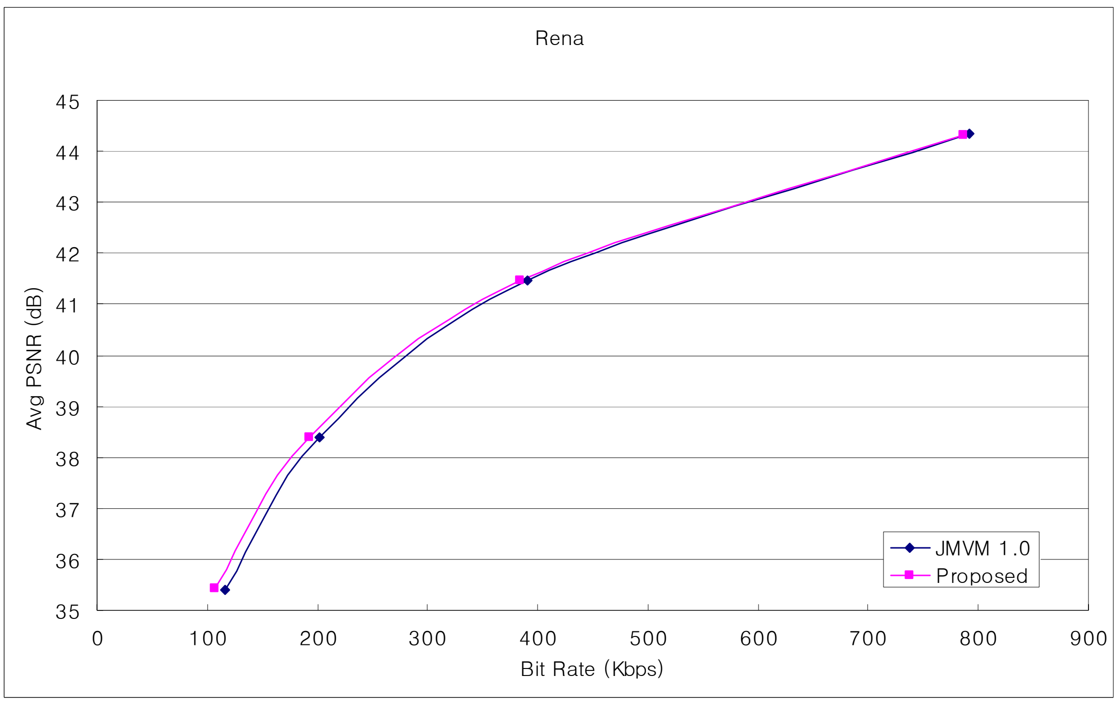



Fig. 7. Rate-Distortion Curve for Rena (for 16 views)

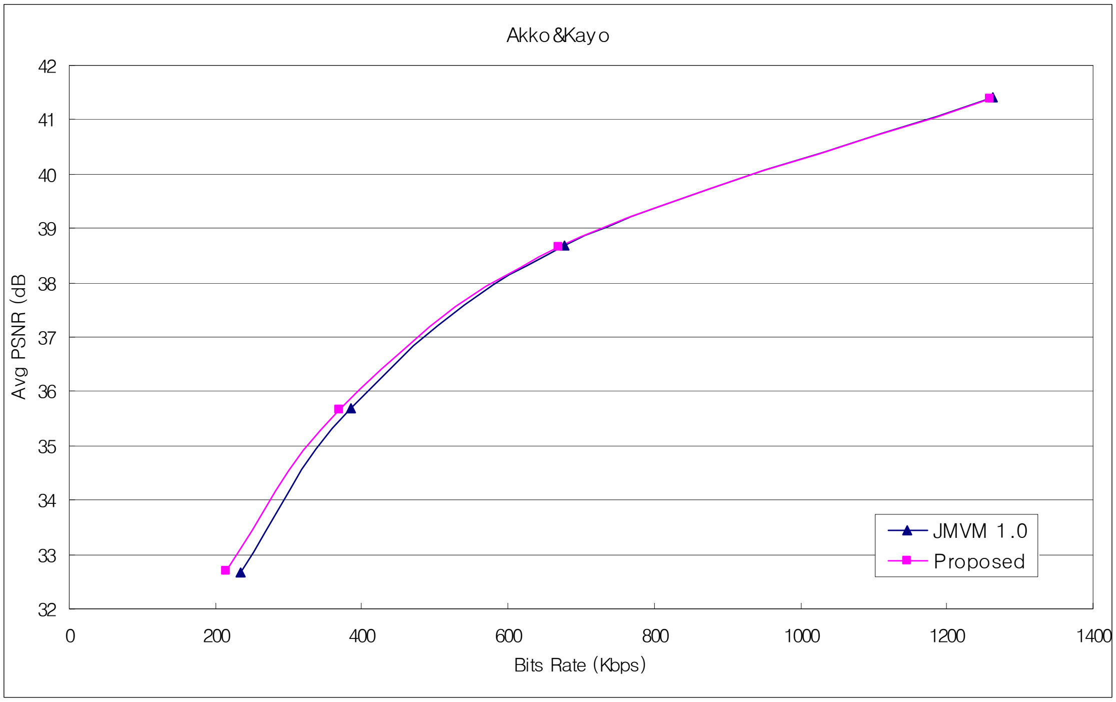



Fig. 8. Rate-Distortion Curve for Akko&Kayo (for 15 views)

# 5. Conclusion

'VIP P-picture' coding method is proposed to use an interpolated image for MVC. It is composed of additional macroblock modes for VIP coding scheme and modified motion vector prediction scheme. The experimental results show the PSNR of VIP adopted views has improved over 0.5 dB for 'Rena' and 'Akko&Kayo', and the average PSNR has improved over 0.15 dB for those sequences. Rest two 'Ballroom' and 'Exit' sequences were almost the same efficiency comparing with reference software.

# 6. Acknowledgements


This work was supported in part by the Information Technology Research Center (ITRC) through the Realistic Broadcasting Research Center (RBRC) at Gwangju Institute of Science and Technology (GIST), and in part by the Ministry of Education (MOE) through the Brain Korea 21 (BK21) project.


# 7. References


[1] ISO/IEC JTC1/SC29/WG11 W8019, "Description of Core Experiments in MVC"

[2] ISO/IEC JTC1/SC29/WG11 JVT-U102, "View Interpolation for Multi-view Video Coding"

[3] ISO/IEC JTC1/SC29/WG11 JVT-U211, "Common Test Conditions for Multiview Video coding"

(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    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.**

**Submitting Organization or Person:**

| | |
|---|---|
| Organization name | Gwangju Institute of Science and Technology (GIST)<br>Korea Electronics and Technology Institute (KETI) |
| Mailing address | C-404, Department of Information and Communications<br>1 Oryong-dong, Buk-gu, Gwangju<br>500-712 |
| Country | Republic of Korea |
| Contact person | Yo-Sung Ho |
| Telephone | +82-62-970-2211 |
| Fax | +82-62-970-2247 |
| Email | hoyo@gist.ac.kr |
| Place and date of submission | Klagenfurt, July 17-21, 2006 |

**Relevant Recommendation | Standard and, if applicable, Contribution:**

| | |
|---|---|
| Name (ex: "JVT") | **JVT** |
| **Title** | **CE10: Multi-view Video Coding using View Interpolation Method** |
| 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.

X 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)

X 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:*

3rd party name(s) ____________________

Mailing address ____________________

Country ____________________

Contact person ____________________

Telephone ____________________

Fax ____________________

Email ____________________

Patent number/status ____________________

Inventor/Assignee ____________________

Relevance to JVT ____________________

Any other comments or remarks: