**Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG**
**(ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6)**
23rd Meeting: San Jose, California, USA, 21–27 April, 2007

Document: JVT-W083
Filename: JVT-W083.doc

*Title:* **CE6: View Interpolation Prediction for Multi-view Video Coding**

*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

This contribution describes a ‘VIP P-picture’ coding which uses the synthesized image as the additional reference frame. The proposed view interpolation method can make an intermediate image by using initial disparity estimation, variable block-based disparity estimation, and pixel-level disparity estimation based on the adjusted search range. In addition, motion vector prediction scheme is modified and vertical displacement is compensated to maximize the efficiency of ‘VIP P-picture’ coding.

# 1. Introduction

Multi-view video simultaneously capture the several scenes from two or more adjacent cameras. Most prediction structures are trying to exploit inter-view correlation of multi-view video. Since S1, S3, and S5 views in the prediction structure in Fig. 1 have two inter-view frames [1], these three B views 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 briefly describe the proposed view interpolation scheme for multi-view video coding, and we propose ‘VIP P-picture’ coding scheme to use the interpolated image.

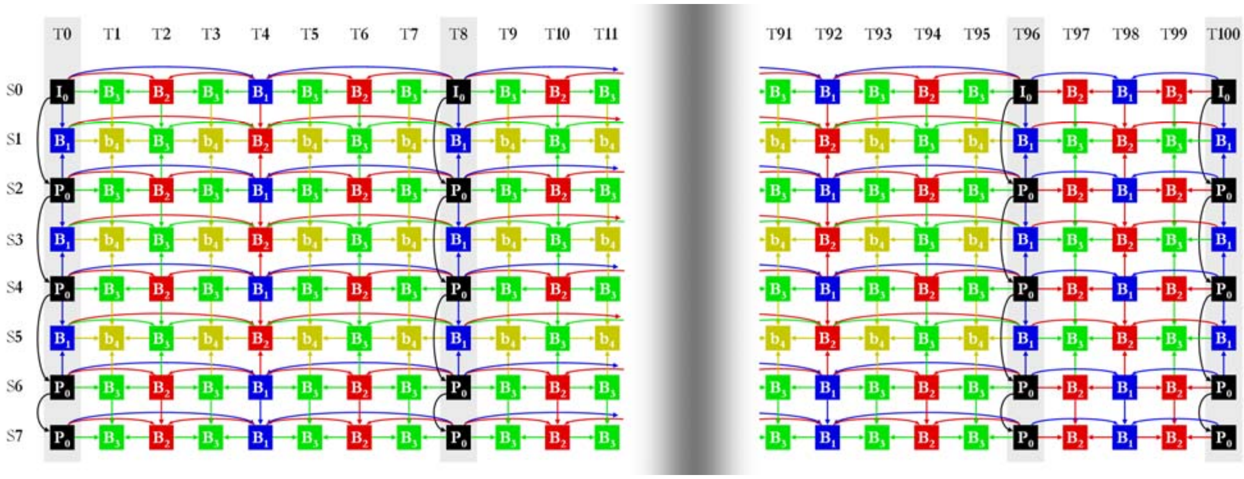



Fig. 1: Inter-view-temporal Prediction Structure

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

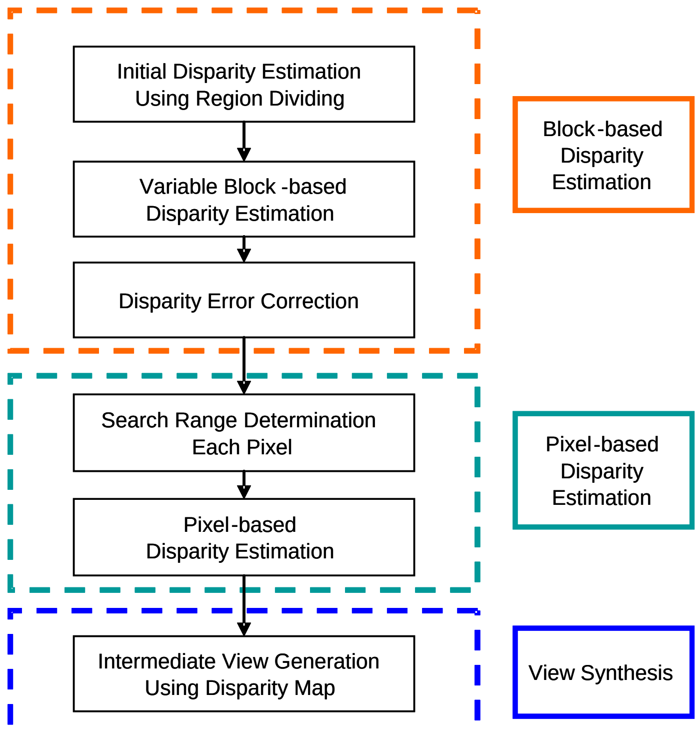



Fig. 2. Proposed View Interpolation Method

Previous interpolation methods have several weaknesses: (a) the quality of the synthesized image depends on the maximum disparity search range, and (b) the accuracy of disparity estimation is unstable 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 based on the region dividing scheme. We also use variable block sizes for disparity estimation, which is more effective in complex region such as 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 disparities. 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 B views such as S1, S3, and S5 in Fig. 1 because these views have already coded adjacent views. After generating the intermediate image, it can be used as an additional reference frame. The intermediate image mostly overlaps with the image to be coded. We propose the view interpolated prediction P-picture (VIP P-picture) coding scheme and it performs like 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 to the co-located block in the intermediate image and it does not need motion vector prediction. Figure 3 shows the overall coding structure using the VIP scheme.

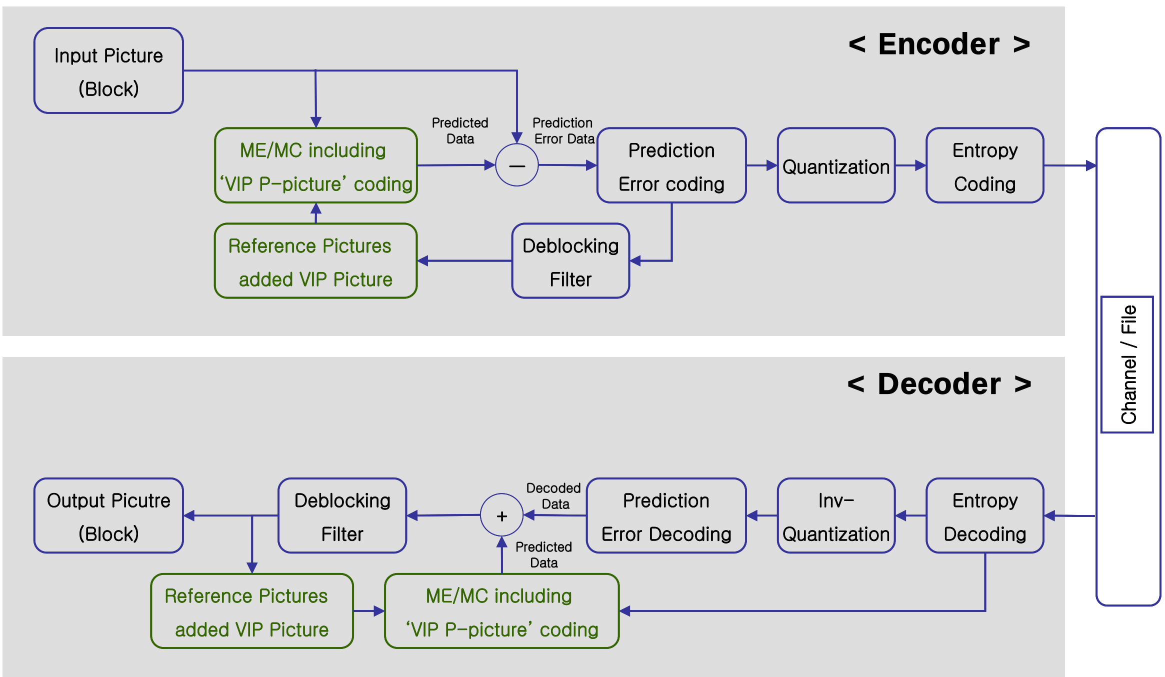
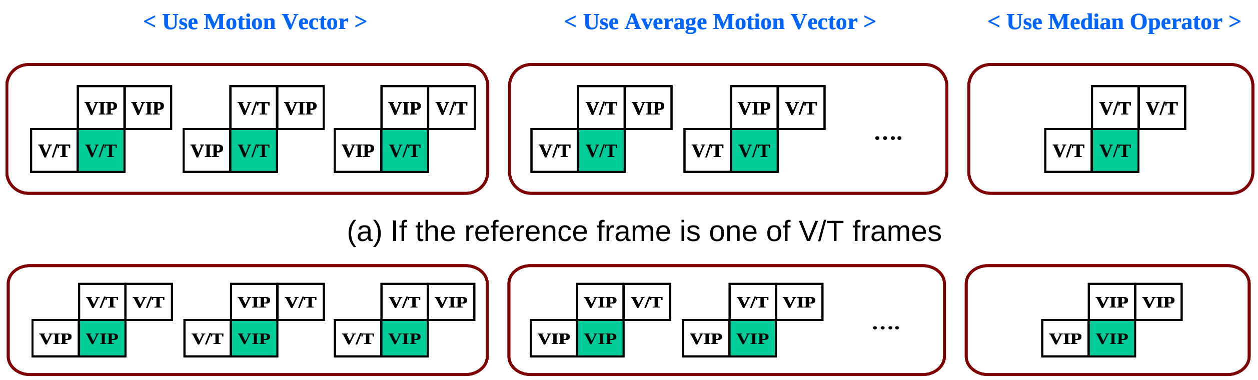



Fig. 3 Coding Structure added 'VIP P-picture' coding

## 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 have 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 a VIP frame, motion vector predictor only considers neighboring blocks referring VIP frame as 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

### 3.3 Compensation of Vertical and Unequal Horizontal Displacements

Some test sequences show vertical displacement and unequal horizontal displacements which decrease the accuracy of stereo matching. We compensate these displacements by handling them manually. Figure 5 shows the example of the compensation of vertical displacement. We attached the excel file related to vertical and unequal horizontal displacement. For the vertical displacement, we calculated pixel based disparity. We represent the unequal horizontal displacement by using the disparity rates between the center view and its left and right views.

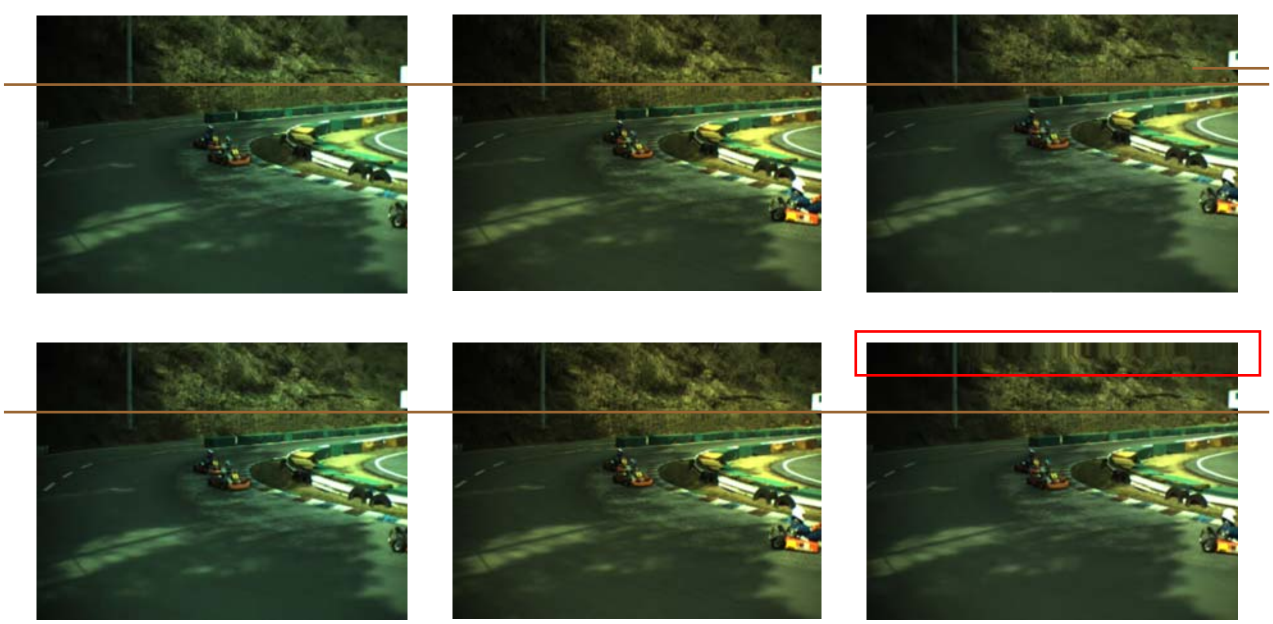


Fig. 5 Compensation of the Vertical Displacement

## 4. Experimental Results

According to the description in [3], four QPs (22, 27, 32, 37) have been used for seven sequences. We have implemented the proposed view interpolation scheme and 'VIP P-picture' coding method using JMVM 1.0 software following the CE description [4]. VIP scheme shows interesting results for dense sequences such as 'Akko&Kayo', 'Rena', and 'Breakdancers'. However it does not show gains for the other sequences. Whole experimental results are contained in the attached excel file.

Table 1. Performance Evaluation for '*Akko&Kayo'* Sequence

| Basic QP | Avg. PSNR (dB) | | Avg. Bitrate (kbps) | | Gain | Bit saving |
|---|---|---|---|---|---|---|
| | JMVM | Proposed | JMVM | Proposed | | |
| 37 | 32.85 | 32.84 | 167.56 | 153.89 | 0.16dB | 3.12% |
| 32 | 35.84 | 35.82 | 276.78 | 264.94 | | |
| 27 | 38.79 | 38.77 | 494.75 | 485.64 | | |
| 22 | 41.46 | 41.44 | 942.11 | 937.35 | | |

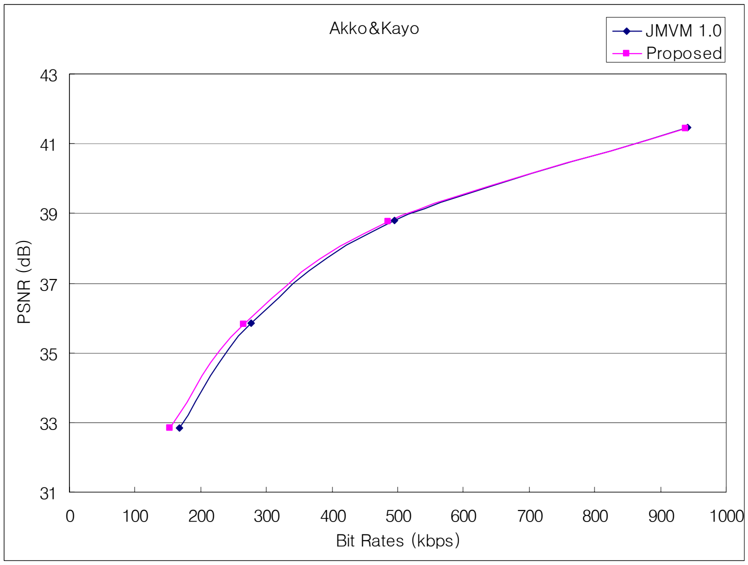



Fig. 5 Rate-Distortion Curves for '*Akko&Kayo'* Sequence

Table 2. Performance Evaluation for '*Akko&Kayo'* Sequence (VIP views only)

| Basic QP | Avg. PSNR (dB) | | Avg. Bitrate (kbps) | | Gain | Bit saving |
|---|---|---|---|---|---|---|
| | JMVM | Proposed | JMVM | Proposed | | |
| 37 | 32.61 | 32.59 | 120.35 | 86.18 | 0.74dB | 11.93% |
| 32 | 35.50 | 35.45 | 189.58 | 159.97 | | |
| 27 | 38.36 | 38.31 | 335.68 | 312.92 | | |
| 22 | 41.02 | 40.99 | 649.01 | 637.09 | | |

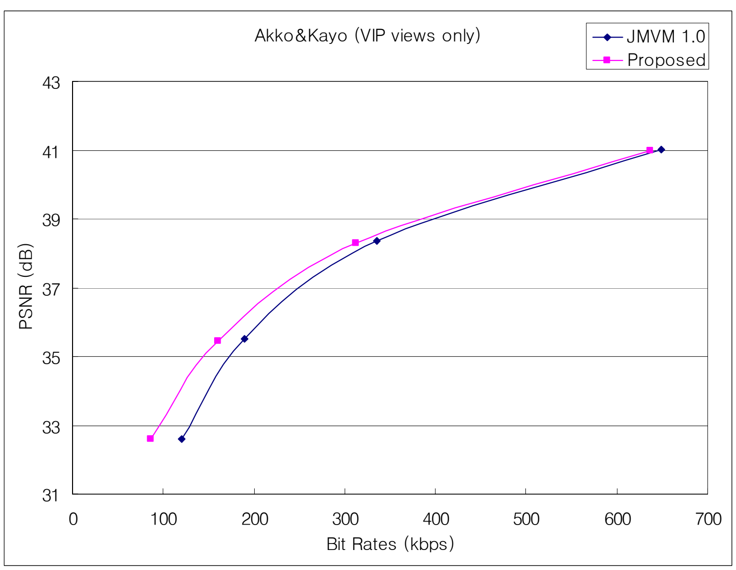



Fig. 6 Rate-Distortion Curves for '*Akko&Kayo'* Sequence (VIP views only)

Table 3. Performance Evaluation for '*Rena'* Sequence

| Basic QP | Avg. PSNR (dB) | | Avg. Bitrate (kbps) | | Gain | Bit saving |
|---|---|---|---|---|---|---|
| | JMVM | Proposed | JMVM | Proposed | | |
| 37 | 35.47 | 35.49 | 77.51 | 72.65 | 0.24dB | 5.16% |
| 32 | 38.48 | 38.47 | 137.12 | 129.29 | | |
| 27 | 41.57 | 41.55 | 273.93 | 256.91 | | |
| 22 | 44.40 | 44.39 | 571.48 | 566.72 | | |

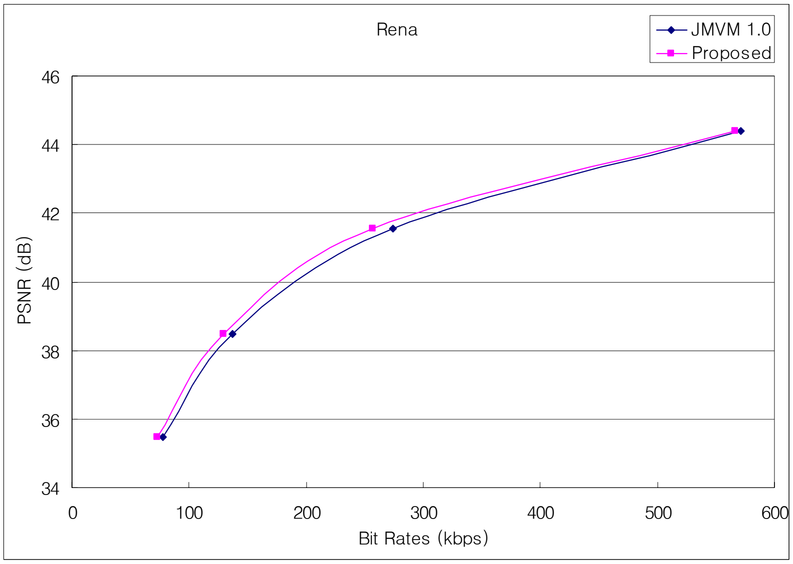



Fig. 7 Rate-Distortion Curves for '*Rena* Sequence

Table 4. Performance Evaluation for '*Rena'* Sequence (VIP views only)

| Basic QP | Avg. PSNR (dB) | | Avg. Bitrate (kbps) | | Gain | Bit saving |
|---|---|---|---|---|---|---|
| | JMVM | Proposed | JMVM | Proposed | | |
| 37 | 35.21 | 35.26 | 61.81 | 50.70 | 0.57dB | 11.99% |
| 32 | 38.14 | 38.13 | 104.75 | 86.85 | | |
| 27 | 41.13 | 41.09 | 204.58 | 187.31 | | |
| 22 | 43.96 | 43.94 | 425.95 | 415.08 | | |

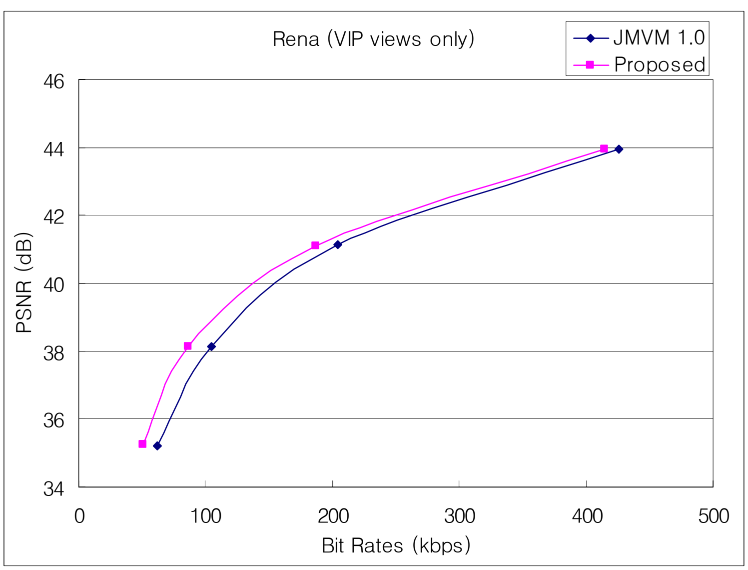



Fig. 8 Rate-Distortion Curves for '*Rena'* Sequence (VIP views only)

Table 5. Performance Evaluation for '*Breakdancers'* Sequence

| Basic QP | Avg. PSNR (dB) | | Avg. Bitrate (kbps) | | Gain | Bit saving |
|---|---|---|---|---|---|---|
| | JMVM | Proposed | JMVM | Proposed | | |
| 37 | 35.01 | 35.03 | 117.42 | 118.80 | 0.03dB | 1.25% |
| 32 | 37.08 | 37.11 | 203.20 | 204.39 | | |
| 27 | 38.67 | 38.71 | 399.72 | 398.99 | | |
| 22 | 39.88 | 39.92 | 989.92 | 974.54 | | |

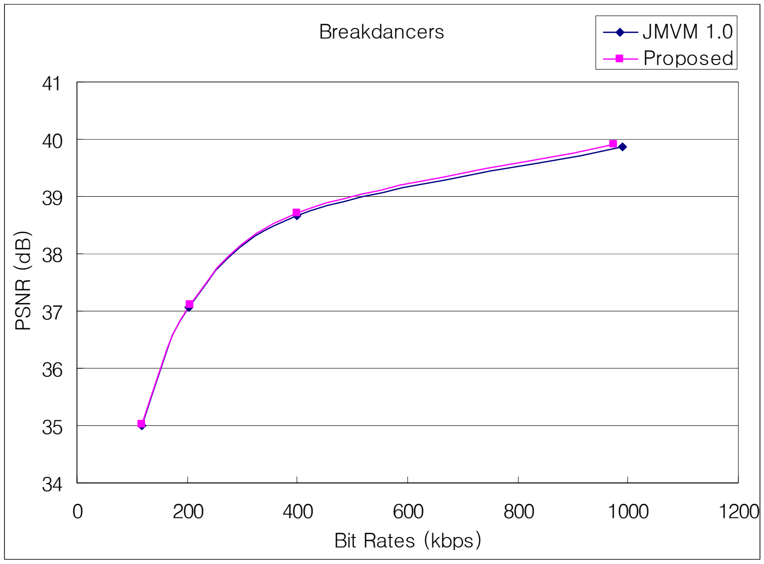



Fig. 9 Rate-Distortion Curves for '*Breakdancers'* Sequence

Table 6. Performance Evaluation for '*Breakdancers'* Sequence (VIP views only)

| Basic QP | Avg. PSNR (dB) | | Avg. Bitrate (kbps) | | Gain | Bit saving |
|---|---|---|---|---|---|---|
| | JMVM | Proposed | JMVM | Proposed | | |
| 37 | 34.48 | 34.54 | 92.86 | 96.55 | 0.08dB | 3.09% |
| 32 | 36.62 | 36.70 | 157.96 | 161.11 | | |
| 27 | 38.37 | 38.49 | 298.89 | 296.96 | | |
| 22 | 39.71 | 39.83 | 689.40 | 648.40 | | |

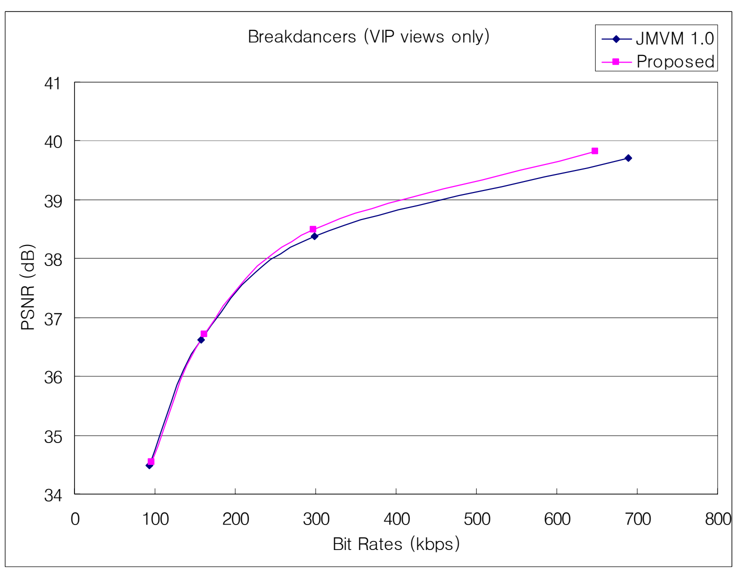



Fig. 10 Rate-Distortion Curves for '*Breakdancers'* Sequence (VIP views only)

# 5. Conclusion

'VIP P-picture' coding method has proposed to use interpolated images for MVC. It is composed of additional macroblock modes for VIP coding scheme and modified motion vector prediction scheme. Unequal horizontal displacements are compensated to maximize the quality of the synthesized image. This scheme shows average 0.66dB gain for B-views and average 0.2dB gain for whole views for dense sequences: 'Akko&Kayo' and 'Rena'. It also shows 0.03dB gain for B-views and 0.08dB gain for whole views for 'Breakdancers'. However, 'VIP P-picture' coding dose not show any gains for other sequences. We think these results are caused by sparse camera arrangement and some problems [5] of test sequences.

# 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"
[4] ISO/IEC JTC1/SC29/WG11 JVT-V306, "CE6: View Interpolation Prediction for MVC"
[5] ISO/IEC JTC1/SC29/WG11 JVT-W084, "Observations of Multi-view Test Sequences"

(Append for Proposal Documents)

# JVT Patent Disclosure Form

| International Telecommunication Union<br>Telecommunication Standardization Sector | International Organization for Standardization | International Electrotechnical Commission |
|---|---|---|

## Joint Video Team - *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 Team (JVT) with information about the patent status of techniques used in or proposed for incorporation in a Recommendation | Standard. The 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 the 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 the 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 formal *Patent Statement and Licensing Declaration form (see http://www.itu.int/ITU-T/ipr/index.html)*, which should be submitted by Patent Holders to the ITU TSB Director, ISO Secretary General, and IEC General Secretary at the time the patent holder believes that the patent is essential to the implementation of a draft or approved Recommendation | International Standard (in addition to the less formal reporting in the earlier proposal/contribution stages of work within the JVT).**

**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 | San Jose, California, 21-27 April, 2007 |

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

| | |
|---|---|
| Name (ex: "JVT") | **JVT** |
| Title | **CE6: 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) – The 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: