INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC1/SC29/WG11
MPEG2005/M13581
July 2006, Klagenfurt, Austria

Title: Global Disparity Compensation for Multi-view Video Coding
Source: GIST
Authors: Yo-Sung Ho, Kwan-Jung Oh, Pil-Kyu Park, and Cheon Lee
(Gwangju Institute of Science and Technology)
Byeongho Choi (Korea Electronics Technology Institute)
Status: Proposal

1 Introduction

This document describes the global disparity compensation for multi-view video coding (MVC). After we explain the global disparity and its compensation in Section 2, we present our experimental results and show the effectiveness of the proposed method in Section 3.

2 Global Disparity Compensation

Due to the global disparity, current MVC schemes employ a large search range for view prediction and this makes it difficult to expand the GOP structure for view prediction [1]. The proposed algorithm compensates for the global disparity and employs the more flexible GOP structure for view prediction [2].

2.1 Global Disparity

Multi-view video coding uses the multi-view video sequences taken by several cameras. So, there exists a disparity called global disparity between adjacent views.

Fig. 1. Global Disparity between Exit_0 and Exit_1

Figure 1 shows the global disparity between Exit_0 and Exit_1. Exit_1 looks like the shifted version of Exit_0 by the shaded area.

Following simulation results show the needs of global disparity consideration. We simulate the coding efficiency for various search range. We use the IPPP GOP structure, QP=31, and view-interlaced structure as shown in Fig. 2.

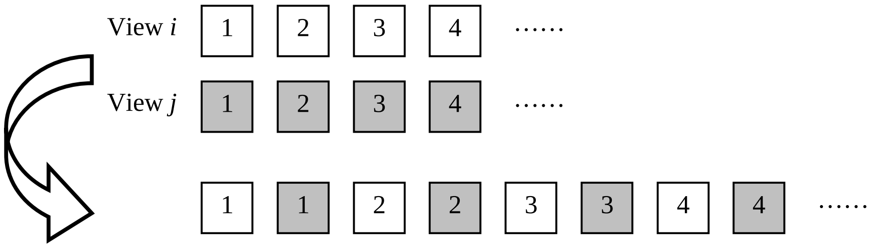

Fig. 2. View-interlaced Sequence

Table 1 shows the simulation results. As you can see, coding efficiency increases up to a certain size of search range, however it is saturated over a certain point. It means that the boundary search range is proper for that structure.

Table 1. Simulation Results of “Exit_3 and Exit_5”

Search range	16	32	64	96	128	160
PSNR	38.054	37.928	37.889	37.885	37.875	37.881
Bitrate (kbps)	1340.98	1171.45	1114.11	1111.56	1105.74	1108.33

2.2 Global Disparity Calculation

To calculate the global disparity, we define a new disparity measure like Eq. (1). *img0* and *img1* in Fig. 3 are two pictures for global disparity calculation and *R* is the number of pixels in the overlapped area.

$$(g_x, g_y) = \min_{x,y} \left[\frac{\sum_{i,j\in R} \left| img0(i,j) - img1(i-x, j-y) \right|}{R} \right] \tag{1}$$

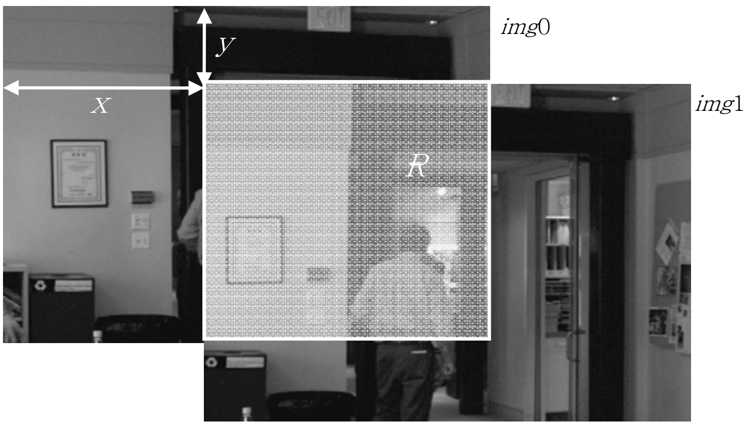

Fig. 3. Global Disparity Calculation

(g_x, g_y) is the displacement vector where the MAD (mean absolute difference) is minimum and it is chosen as the global disparity for the two views.

2.3 Global Disparity Compensation

By using the calculated global disparity, we can compensate the global disparity. Before motion estimation, we shift the reference frame as much as the global disparity. And then, we pad outside of the boundary by copying the boundary pixel value. From the Fig. 4, 5, 6, and 7, you can easily understand the procedure of global disparity compensation. The picture in Fig. 4 is the frame to be encoded as B frame, and two pictures in Fig. 5 are reference frames for that. As you can see, there exist the global disparities between pictures in Fig. 4 and Fig. 5. So, we need large search range in the motion estimation process to search the proper region.

Fig. 4. Frame to be Encoded as B frame

Fig. 5. Two Reference Frames for Fig. 4 (Top: List0, Bottom: List 1)

However, if we shift the reference frame as much as the global disparity, we do not need to the large such range anymore. Figure 6 shows the global disparity compensated pictures. As you see, all objects in the picture are located at similar positions compared with the Fig.

4. Some outer pixels do not have pixel values, because we move the reference frames. So, we copy the values of the boundary as shown in Fig. 7.

Fig. 6. Global Disparity Compensated Pictures (Top: List0, Bottom: List 1)

Fig. 7. After Padding of the Outer Pixels (Top: List0, Bottom: List 1)

To demonstrate the efficiency of the global disparity compensation, we introduce one simple experiment and show its results. We use the full search mode in motion estimation instead of the fast search mode and encode just three frames using I-B-I structure for “Exit” sequence. A quantization parameter (QP) is 31. Table 2 shows the coding results of B picture when search range varies from 32 to 128. In this case, global disparities are (-75, 0) and (126, 0). As you can see, the proposed algorithm shows better results and it is less sensitive to change of the search range. Therefore, we can know that global disparity compensation improves the coding efficiency of MVC and it also leads to shorter encoding time.

Table 2. Coding Results for Global Disparity Compensation

Search Range	Bit Rate (bits)		PSNR (dB)	
	JSVM	Proposed	JSVM	Proposed
32	56272	45096	38.1103	38.0668
64	49096	44224	37.9909	38.0532
96	45480	44360	37.9919	38.0622
128	44136	44352	38.0649	38.0589

2.4 View-temporal Prediction Structure using Global Disparity

We propose a new view-temporal prediction structure using global disparity. In general, view prediction structures cannot use the far distance prediction because their coding efficiencies are not good. Figure 8 shows the proposed view-temporal prediction structure using global disparity compensation. We apply the hierarchical-B picture structure for view prediction. Despite the hierarchical-B picture already shows the good results in temporal prediction and it has many advantages, it cannot be used for view prediction because cannot show good results in view prediction. However, we can apply the hierarchical-B picture structure for view prediction by using global disparity compensation. In addition, we arrange the high level picture like as grid type. It can effectively propagate good quality to whole pictures in view-temporal prediction structure. Unfortunately, we not yet implemented this view-temporal prediction structure perfectly. Figure 9 shows the view prediction structure. At this time, we just show the result for view prediction.

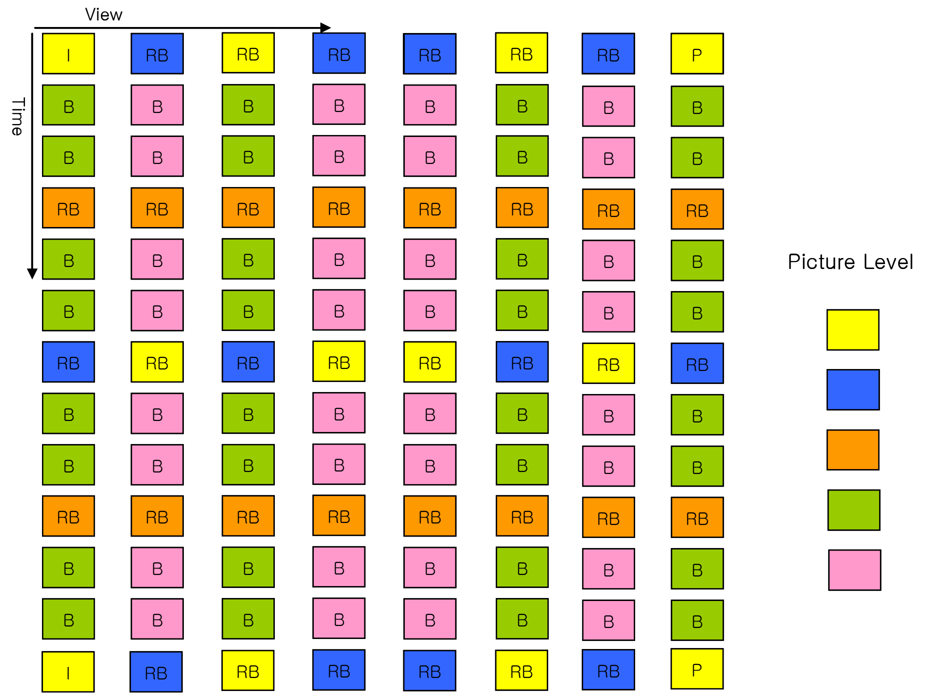

Fig. 8. View-temporal Prediction Structure using Global Disparity

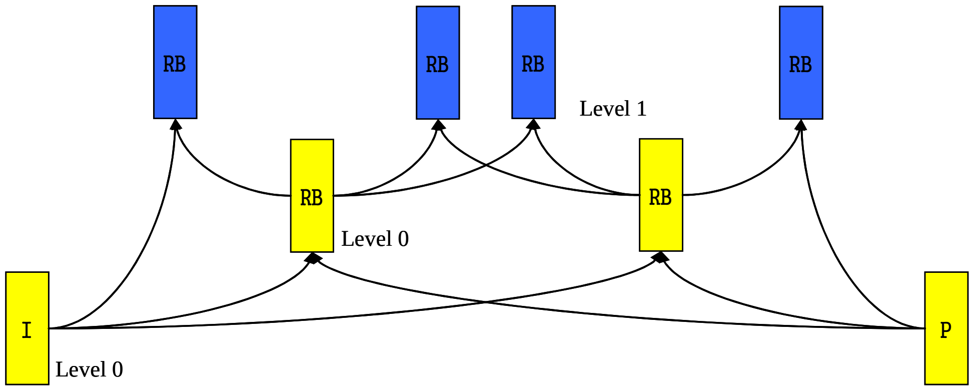

Fig. 9. View Prediction Structure using Global Disparity

3 Experimental Results and Analysis

In order to evaluate our proposed method, we have experimented with “Exit” and “Ballroom” sequences (640× 480, 8 views) provided by MERL and compared to the result of the JSVM coding. Table 3 and Fig. 10 show the coding results and the rate-distortion curve for “Exit” sequence, respectively, Table 4 shows the coding results and Fig. 11 shows the rate-distortion curve for the “Ballroom” sequence.

Table 3. Coding Results for “Exit” Sequence

Basis QP	Bit Rate (kbps)		PSNR (dB)	
	JSVM	Proposed	JSVM	Proposed
31	1001.075	908.975	37.2452	37.0375
29	1290.175	1160.000	38.0992	37.8589
26	2031.975	1804.925	39.4265	39.1562

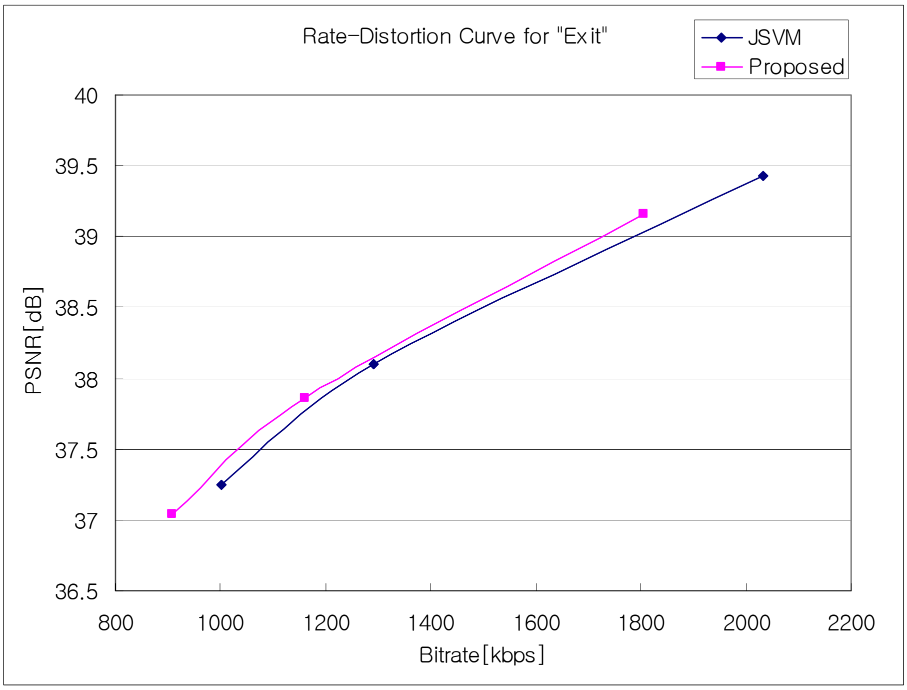

Fig. 10. Rate-Distortion Curve for “Exit”

Table 4. Coding Results for “Ballroom” Sequence

Basis QP	Bit Rate (kbps)		PSNR (dB)	
	JSVM	Proposed	JSVM	Proposed
34	1014.475	983.625	33.7435	33.8250
31	1566.675	1465.975	35.5091	35.4826
29	2022.625	1850.175	36.5929	36.4743

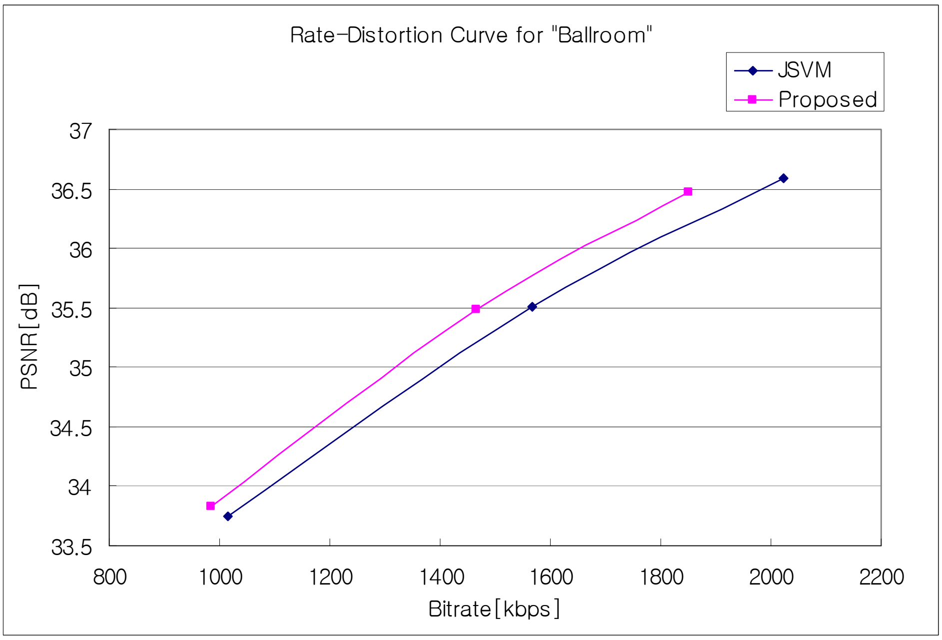

Fig. 11. Rate-Distortion Curve for “Ballroom”

4 Conclusion

In this document, we have proposed the multi-view video coding using global disparity compensation. With some test sequences, we have verified quality improvement compared to the coding scheme of the reference software. However, as mentioned before, due to the time limitation we did not yet experiment with all sequences. Next time, we can show the results.

5 References

[1] ISO/IEC JTC1/SC29/WG11 W8019, “Description of Core Experiments in MVC,” April 2006.

[2] ISO/IEC JTC1/SC29/WG11 M12542, “Multi-view Video Coding based on Lattice-like Pyramid GOP Structure,” October 2005.