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

ISO/IEC JTC1/SC29/WG11
MPEG2008/M15852
October 2008, Busan, Korea

Source: GIST (Gwangju Institute of Science and Technology)

Status: Report

Title: Results of Experiment on Temporal Enhancement for Depth Estimation

Author: Sang-Beom Lee and Yo-Sung Ho

1. Introduction

This document reports the results of experiment on temporal enhancement for depth estimation. We obtained two views of depth sequences by using Nagoya University's depth estimation software in which the temporal enhancement is added [1]. Then, we have synthesized the intermediate view by using Nagoya University's view synthesis software [1]. In order to evaluate the updated software, we compared the original view with the synthesized view.

2. Temporal Enhancement for Depth Estimation

2.1. Weighting Function

Since the conventional depth estimation methods separately estimate the depth value for frame by frame, the resultant depth sequence has a low temporal consistency. Therefore, we have described a new matching function that refers to the depth value of the previous frame when estimating the depth of the current frame as described in the previous document [2]. The modified matching function adds the weighting function considering the depth value of the previous frame. The temporally weighted matching function is defined by

$$C_{updated}(x, y, d) = C_{original}(x, y, d) + \lambda \mid d - D_{prev}(x, y) \mid \quad (3)$$

where λ represents the slope of the weighting function and $D_{prev}(x,y)$ represents the previous depth value.

Figure 1 illustrates an example of the matching functions. As shown in Fig. 1, in case of the previous depth value is around 70, the probability that the current depth value is around 70 is very high. Therefore, we apply the weighting function which increases the matching score as the difference of the current and the previous depth value become larger. Although the weighting function determines the current depth value compulsorily, we can improve the temporal consistency of the depth sequence.

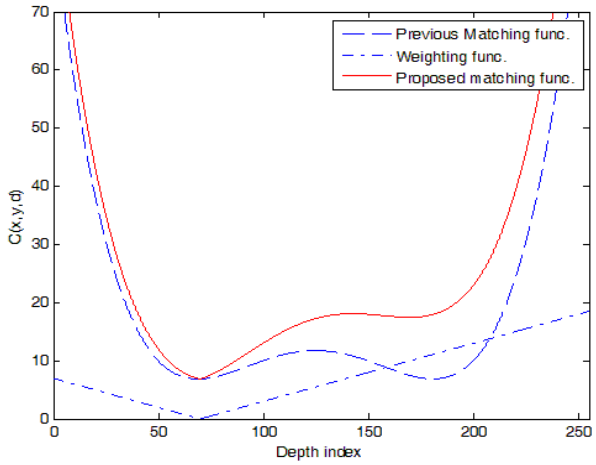
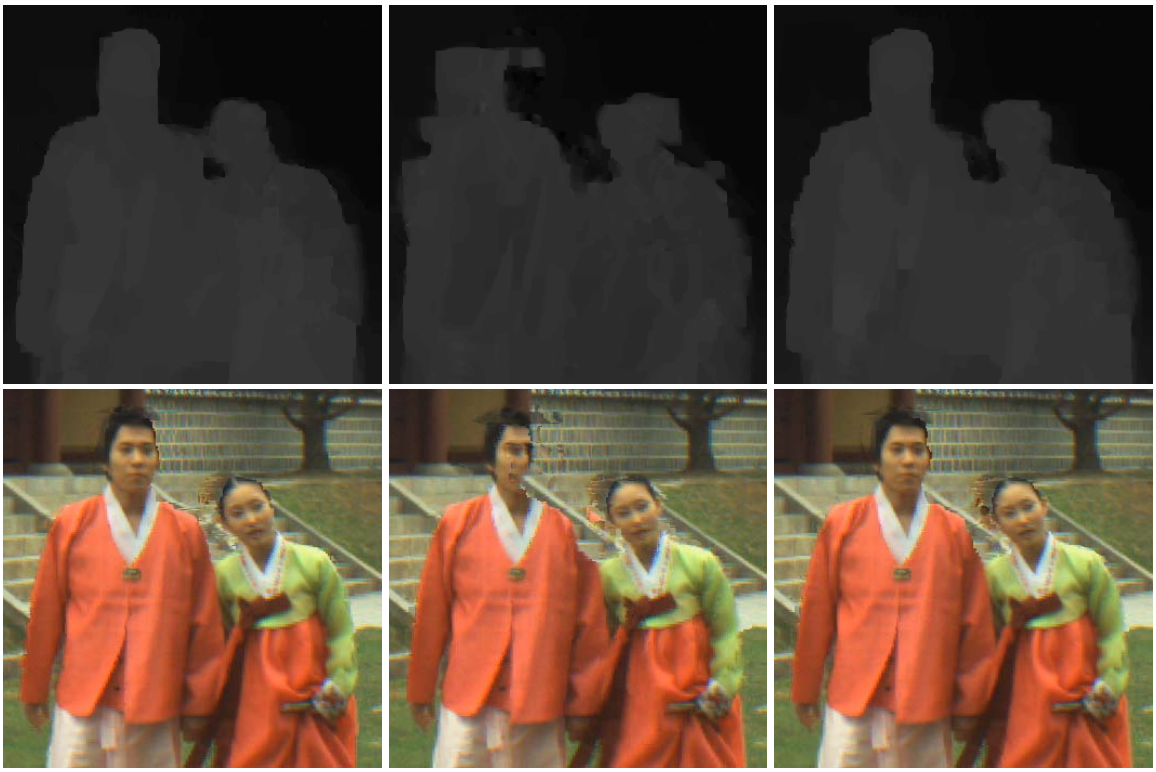


Fig. 1. Graph of the weighted matching function

2.2. Key Frame Insertion

When we added the temporal enhancement in the depth estimation software, we have noticed that the results had ghosting artifacts. This problem is caused by the propagation of the wrong depth values. In order to prevent the problem, it is necessary to insert the key frame. The temporal enhancement is restarted for each key frame.

Figure 2 shows the key frame insertion results for "Lovebird1" (view 5, 83rd frame). As shown in Fig. 2(b), the depth map has ghosting artifact. However, as shown in Fig. 2(c), these artifacts are removed by the insertion of the key frame.



(a) temporal enhancement off (b) no key frame insertion (c) after key frame insertion

Fig. 2. Key frame insertion results (Slope = 1.00)

3. Experimental Results

In order to evaluate the temporal enhancement, we have tested four test sequences, which are “Laptop”, “Lovebird”, “Newspaper”, and “Pantomime”. The experiment is performed for 100 frames for each sequence. In this experiment, all parameters were set as default except for parameters which are related to the temporal enhancement. After obtaining depth maps for the left and right views, we synthesized the intermediate view. Then, we compared the virtual view with the original view.

3.1. Experimental Results for “Slope”

Figure 3 shows the intermediate views according to the parameter “Slope”. The upper and lower figures are results for no temporal enhancement and for temporal enhancement, respectively. When applying the temporal enhancement, we noticed that the difference of two consecutive frames is lowered as shown in Fig. 3(c). It means that the flickering artifacts are efficiently removed in the intermediate views, especially at background.

(a) 12nd frame (b) 13rd frame (c) difference image

Fig. 3. Intermediate views and the difference value.

Table 1 and Fig. 4 show PSNR results according to “Slope”. As shown in Fig. 4, average PSNRs are almost the same but decreased for the case of sequence “Laptop” and “Pantomime” as “Slope” is increased. The drop in PSNR is because of the ghosting artifacts since the key frame was only the first frame in this experiment.

Table 1. Average PSNR according to “Slope”

Slope	Off	On 0.25	On 0.50	On 0.75	On 1.00
Laptop	35.0194	34.8546	34.1548	33.5133	31.5795
Lovebird1	30.1775	30.1589	30.1151	30.0499	29.9620
Newspaper	25.7963	25.9787	25.9656	25.9133	25.8467
Pantomime	34.0903	34.3704	34.3148	34.2090	28.3658

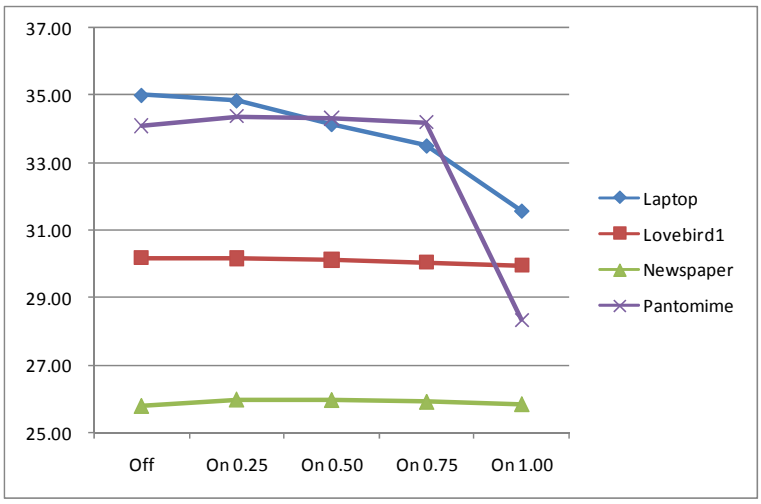


Fig. 4. Average PSNR according to "Slope"

3.2. Experimental Results for "Keyframe"

Table 2 and Fig. 5 show PSNR results according to "Keyframe". As shown in Fig. 5, average PSNRs are increased or almost the same as the interval of the key frame is decreased. From the results, we can infer that the ghosting problem is efficiently reduced since the propagation of the wrong depth values is prevented.

Table 2. Average PSNR according to "Keyframe"

Keyframe	Only the 1st	10	5
Laptop	31.5795	33.0978	33.7327
Lovebird1	29.9620	30.0511	30.1194
Newspaper	25.8467	25.6540	25.7531
Pantomime	28.3658	33.9488	34.0758

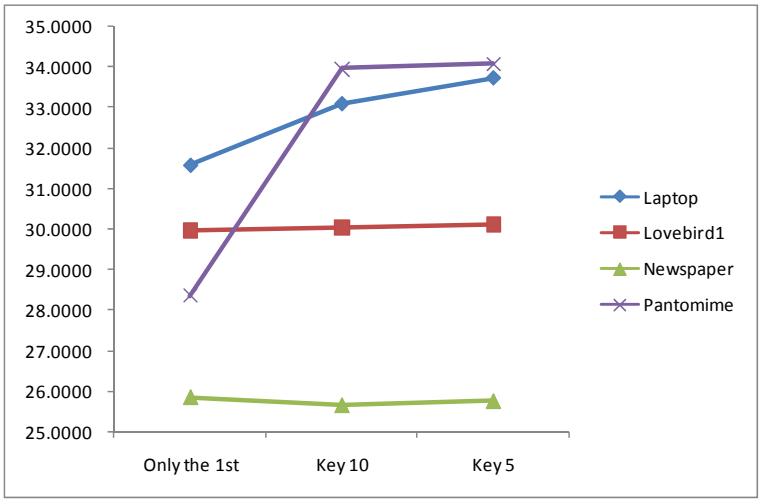


Fig. 5. Average PSNR according to "Keyframe"

4. Summary

In this document, we have reported the experiment in the temporal enhancement for depth estimation. Since we added a temporally weighted matching function, experimental results had the higher temporal consistency compared to the previous works, especially at background. The optimum determination of parameters is remained as a future work.

Acknowledgements

This work was supported in part by ITRC through RBRC at GIST (IITA-2008-C1090-0801-0017).

References

[1] ISO/IEC JTC1/SC29/WG11 "Reference Softwares for Depth Estimation and View Synthesis," M15377, April 2008.

[2] ISO/IEC JTC1/SC29/WG11, "Enhancement of Temporal Consistency for Multi-view Depth Map Estimation," M15594, July, 2008.