# The Performance of Self-Enhancement Method in Optical Flow-Based Motion Estimation for Sequential Fisheye Images with Sudden Movements of the Objects

Arief Suryadi Satyawan[1] Junichi Hara[2,3] Hiroshi Watanabe[1,2]

[1]Graduate School of Fundamental Science and Engineering, Waseda University
[2]Global Information and Telecommunication Institute, Waseda University
[3]RICOH Company, LTD

## 1. Introduction

The self-enhancement method in optical flow-based motion estimation (LKI) has been succeeded to be used for estimating motion from sequential fisheye images, which has massive distortion [1]. This method is a modified version of the Lucas and Kanade's optical flow concept that used to serve sequential perspective images [2]. However, the performance of LKI has not been proven to estimate sequential fisheye images that consist of sudden movements of the objects. This research focuses on evaluating the LKI scheme to handle such kind of image.

## 2. Experimental Methodology

As mentioned in [1], the improved Lucas and Kanade's (LKI) scheme applied in this experiment can be seen in table 1. To be able to implement sudden movements of objects, the input pair of fisheye images can be obtained from two (frame1 and frame2), three (frame1 and frame3), four (frame1 and frame 4), or five (frame1 and frame5) image difference of each sequential fisheye images. There are four sequential fisheye images used in this experiment. While the Hand, Man, and Truck I sequence are obtained directly from one side of the Ricoh Theta S camera [3], the Cube sequence is made synthetically from Blender software. The resolution of these sequences is 640 x 640 pixels.

The performance of motion vectors is determined from the quality of the reconstructed fisheye image (RI). The RI is generated from the first fisheye image input that transformed by the estimated motion vectors (MV).

## 3. Results

### 3.1 Quantitative evaluation

It can be seen from Fig. 1 to 4 that the performance of the LKI goes down gradually proportional to a reduction of image difference. When the image difference reaches 5, the PSNR of each sequence generally reduces until about 15 dB from the maximum value (when the image difference is 2). However, this condition is far better in comparison with the results obtained by using block-based (BB) motion estimation scheme. The gap can reach about 14 dB in the worst situation.

### 3.2 Qualitative evaluation

Due to page limitation, there are only few RI's that can be presented in this paper. They are presented in Fig. 5 (a to d). Each figure shows RI obtained from three image difference of input. It can be seen that the RI for the Man and Cube sequence is still good, while the RI for the Hand sequence is distorted slightly. In case of the Truck I, since the position of the object in the frame1 and frame3 is very different (it caused by the different speed of truck while it was captured at t=1 and t=3), the obtained motion vectors cannot be determined very well. Therefore, the RI becomes worst. This sequential fisheye images can be categorized as an unpredictable movement of the object, or it is worse than sudden movements of the object.

## 4. Conclusion

In this research, the evaluation of the LKI scheme for handling four sequential fisheye images with a sudden movement of the object has been conducted. The LKI scheme works well to produce motion vectors from at least three frames difference. This means that to some extent, the LKI can estimate motion from sequential fisheye images with sudden movements of objects.

## 5. References

[1] A.S. Satyawan, J. Hara, and H. Watanabe, "An Improvement of Motion Estimation on Sequential Fisheye Images," IEICE General Conference, BS-2-12, Mar. 2018.
[2] S. Baker and I. Matthews, "Lucas-Kanade 20 Years On: A unifying framework," International Journal of Computer Vision, vol.56, no.3, pp.221-255, Feb.-Mar. 2004.
[3] https://theta360.com/en/about/theta/s.html

Table 1. Experimental Algorithm

| *Algorithm for motion estimation calculation* |
|---|
| *Input: a pair of fish-eye images* |
| *Output: reconstructed image, vector flows (u & v), motion compensated, & PSNR* |
| *1: **for** each two successive images, **do*** |
| *2: $im1_{st}$ & $im2_{nd}$ ← convert the two images to grayscale* |
| *3: $im1_{st}$ & $im2_{nd}$ ← smooth the two grayscale images* |
| *4: u & v ← calculating vector flows* |
| *5: im1R ← $im1_{st}$ moved by u & v* |
| *6: im_prediction ← subtract $im2_{nd}$ by im1R* |
| *7: PSNR ←10 $\log_{10}$ im_prediction* |
| *8: $im1_{st}$ ← im1R* |
| *9: Do 2 to 7* |
| *10: if the new PSNR < previous PSNR Do 2 to 7* |
| *11: else end* |
| *12: **end for*** |
| *13: **return*** |

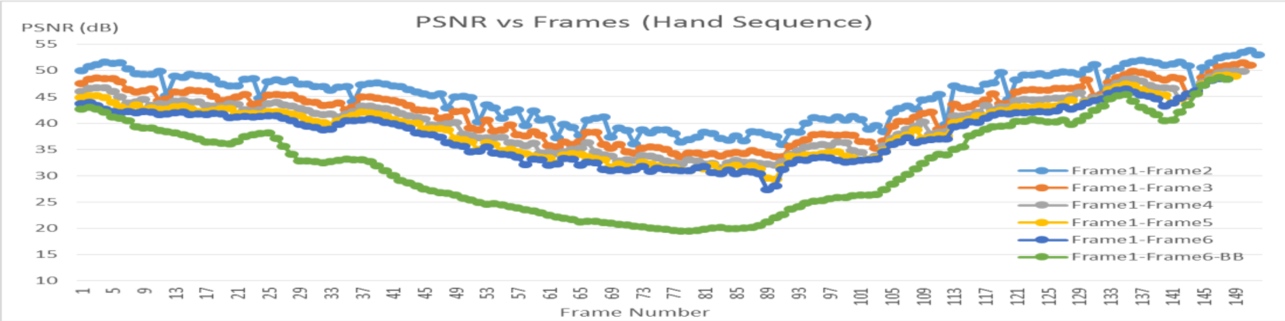



Fig. 1 PSNR versus frame number for Hand sequence

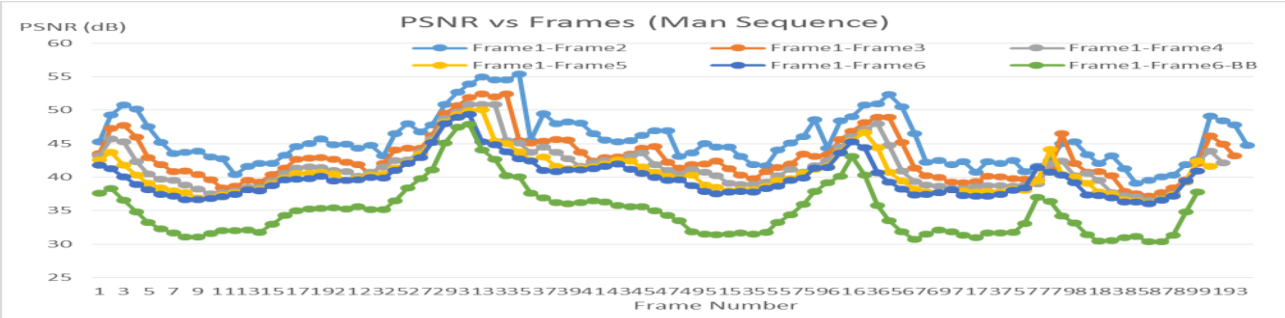



Fig. 2 PSNR versus frame number for Man sequence

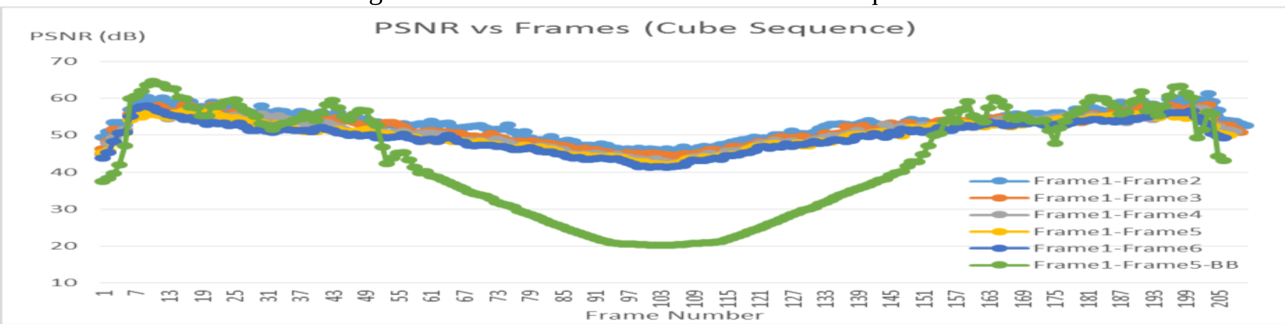



Fig. 3 PSNR versus frame number for Cube sequence

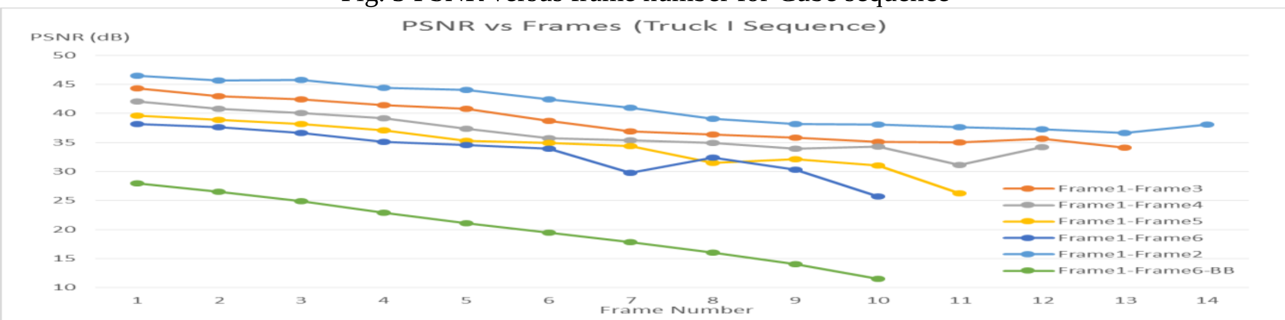
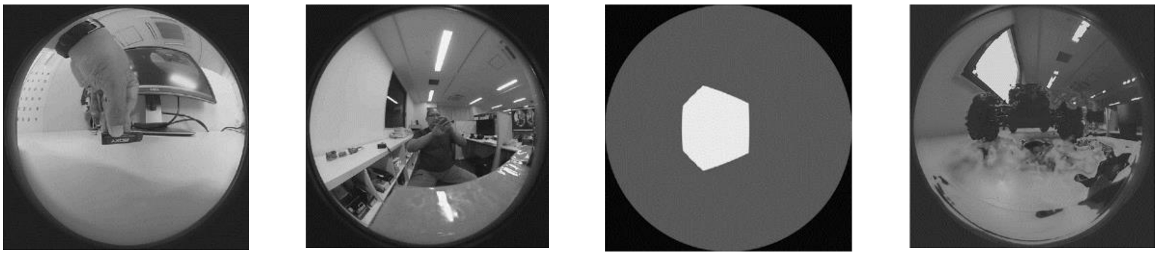



Fig. 4 PSNR versus frame number for Truck I sequence



(a) Hand sequence (b) Man sequence (c) Cube sequence (d) Truck I sequence

Fig. 5 Qualitative performance of estimated motion shown by the reconstructed images (RI)