Improvement of Spatial Resolution Using Block-Matching Based Motion Estimation and Frame Integration

Danya Sugai and Takayuki Hamamoto
Graduate School of Engineering, Tokyo University of Science, 6-3-1, Niijuku, Katsusika-ku, Tokyo, 125-8585, Japan
Email: {sugai, hamamoto}@isl.ee.kagu.tus.ac.jp

***Abstract*—High frame-rate images typically possess a high temporal correlation between their motion vectors. By use of this characteristic, high accuracy motion estimation techniques for high frame-rate images were proposed in our previous work. Compared with conventional methods, our proposed methods are relatively resistant to spatial aliasing errors. However these methods are still inadequate for resistant to spatial aliasing errors. This paper proposes a new matching based motion estimation method that further improves spatial resolution. The proposed motion estimation method is more resistant to spatial aliasing errors and is suitable for reconstructing a high spatial resolution image using multiple high frame-rate images. The method uses multiple previous frames for motion estimation. First, motion vectors are independently estimated between the current frame and multiple previous frames using the basic block matching method and parabola fitting. Next, a time-axially-averaged motion vector is approximated from multiple motion vectors using the least-squares method. Finally, the motion vector in the current frame is calculated by the approximation straight line. The results of motion estimation experiments conducted using images with spatial aliasing errors show that the proposed method improves accuracy. Further, the results of super-resolution experiments conducted using high frame-rate images show that the proposed method also improves spatial resolution.**

***Index Terms*—high frame-rate, motion estimation, super resolution, aliasing error**

I. INTRODUCTION

Image sensing devices play an important role in image processing systems, because they determine the quality of the input image. Hence, these devices are continually being improved in order to obtain high-quality images. Further, their characteristics, such as higher sensitivity, lower noise, and high spatial resolution, are constantly being upgraded. In fact, to achieve high frame-rates, low power consumption or downsizing, many researchers have considered on-sensor analog-to-digital converters. However, it is difficult to improve spatial resolution effectively such a device owing to the resolution limits of optical lens. Therefore, the improvement of spatial resolution via the digital post-processing is being actively studied [1].

Because high frame-rate images have a high temporal correlation, the amount of motion between frames is very small. This characteristic has several advantages in the motion estimation applied to imaging systems for reconstructing high spatial resolution images. For example, it improves the accuracy of gradient-based motion estimation because it can effectively estimate small motions, but cannot cope with wide-range motion [2], [3]. In addition, the amount of calculations in the block matching-based method is reduced as its search area becomes smaller.

High accuracy motion estimation techniques based on the gradient method that exploits high frame-rate images have already been proposed in our previous work [4], [5]. The techniques proposed can reduce the estimation error in the spatial gradient using the high temporal correlation characteristic of the high frame-rate images. This paper proposes a new spatial aliasing resistant motion estimation method that improves spatial resolution for reconstruction. The results of motion estimation accuracy and improvements in the spatial resolution of high frame-rates images obtained using the proposed method are also presented and evaluated.

II. MOTION ESTIMATION BASED ON THE MATCHING METHOD

A. Block Matching Method

The block matching method is a way of motion estimation. This method compares a target macroblock in the previous frame with neighborhoods of macroblocks in the current frame, and determines the best matched coordinates by using an evaluation function. One of the popular functions is the sum of squared differences (SSD). The relation between these quantities can be expressed as:

$$S_{sdd} = \sum_i \sum_j \{ f(i, j) - g(i + dx, j + dy) \}^2 \quad (1)$$

where i, j are pixel positions. $f(i, j)$ is the pixel value of a target macroblock in the previous frame, $g(i, j)$ is the pixel value of macroblock in the current frame, and dx and dy denote the motion in the i and j directions, respectively.

Manuscript received April 5, 2015; revised August 31, 2015.

doi: 10.18178/joig.3.2.117-121

Some methods have been proposed for sub-pixel motion estimation based on the block matching algorithm. One method interpolates sub-pixel values by spatial filtering processing. After expanding image size by a spatial filter, sub-pixel motion vectors are estimated using the block matching method. Another method uses parabola fitting over three points [6]. The estimated sub-pixel displacement $\hat{d}$ by parabola fitting can be expressed as:

$$\hat{d} = \frac{R(-1) - R(1)}{2R(-1) - 4R(0) + 2R(1)} \tag{2}$$

where $R(s)$ is a similarity function and s is the integral shifted value from the extremal pixel position. Further, the range of $\hat{d}$ is usually defined as $-0.5 \le \hat{d} \le 0.5$. Using the SSD, $R_x(s_x)$ and $R_y(s_y)$ are expressed as:

$$R_x(s_x) = \sum_i \sum_j \{f(i,j) - g(i + dx + s_x, j + dy)\}^2 \tag{3}$$

$$R_y(s_y) = \sum_i \sum_j \{f(i,j) - g(i + dx, j + dy + s_y)\}^2 \tag{4}$$

B. Proposed Method

Before estimating motion vectors, we first degrade the high spatial frequency by a low-pass filtering. This process is carried out to reduce random noises or aliasing effects, and improve the accuracy of motion estimation. However, when an acquired image has aliasing errors, the shapes of the acquired image are different from those of the real image, and it is difficult to correctly estimate motion vectors using the acquired image. Fig. 1 illustrates this situation. It shows how an acquired image loses the shape of the real image owing to a lower spatial resolution: a lower spatial resolution means that the pixel pitch is long. If the real image moves with sub-pixel precision, the shape of the acquired image will change as a result of aliasing effects. This results in difficulty estimating the motion vectors with sub-pixel precision. In other words, as shown in Fig. 2, if the motion vectors are known, the shape of the real image can be reconstructed using several frames.

To reduce errors due to aliasing effects, we propose the block-matching based motion estimation. This motion estimation uses multiple previous frames instead of only one previous frame. Fig. 3 and Fig. 4 indicate the concept underlying the proposed method. First, motion vectors are independently estimated between the current frame and multiple previous frames, using the basic block matching method and parabola fitting. Further, approximation straight lines in the x and y directions are calculated from the estimated motion vectors using the least-squares method. Time-axially-averaged current motion vectors are then derived from the calculated approximation straight lines. This method is based on the concept that high frame-rate images have high temporal correlation and the motion vectors are regarded as having uniform motion. The characteristic of this proposed method reduces errors due to aliasing effects and various noises.

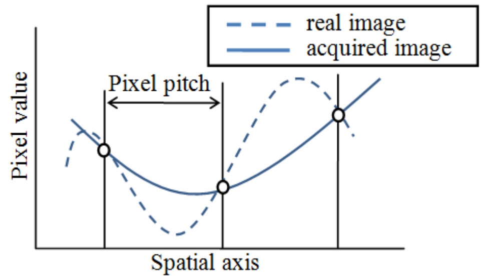


Figure 1. Acquired image with error due to aliasing effects.

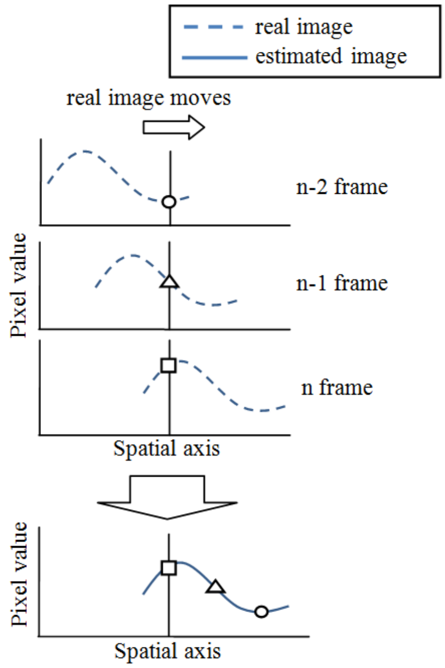


Figure 2. A real image reconstructed from several frames.

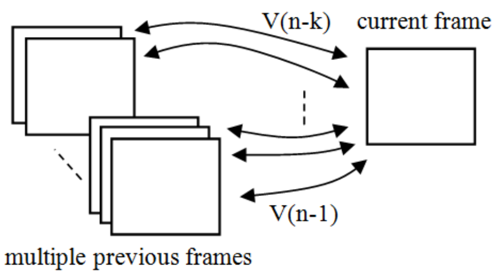


Figure 3. Concept underlying the proposed motion estimation method using multiple previous frames.

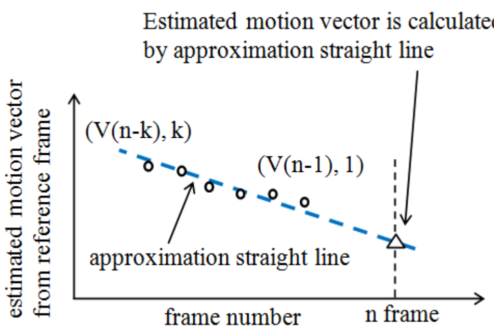
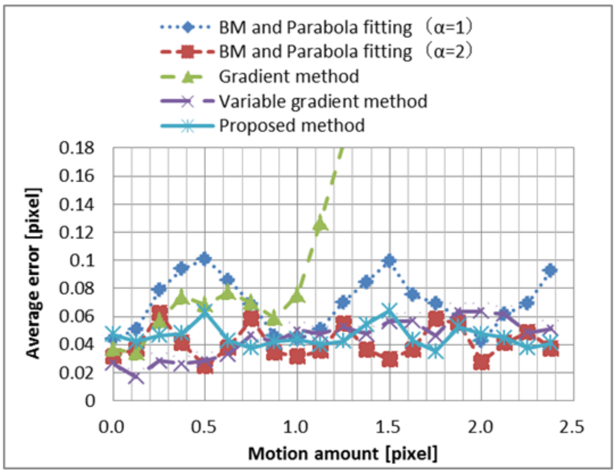


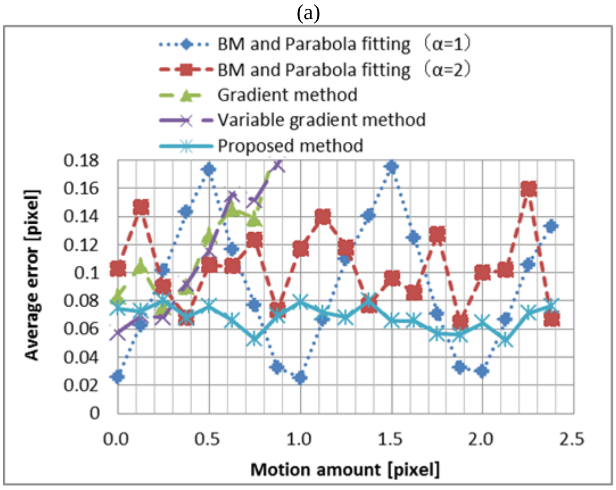
Figure 4. Proposed motion estimation method using an approximation straight line.

(a) (b)

Figure 5. 1/8 pixel shifted images: (a) Images with low aliasing noises. (b) Images with aliasing noises.



(a)



(b)

Figure 6. Evaluation of average error obtained by the various motion estimation methods: (a) Evaluation results for Fig. 5(a). (b) Evaluation results for Fig. 5(b).

III. EVALUATION

A. Evaluation of Motion Estimation

The experimental results for the average error of the various motion estimation methods using synthesized images are shown in this section. The evaluation used images that were each horizontally shifted by each 1/8 pixel, as shown in Fig. 5(a) and Fig. 5(b). Fig. 5(b) has more aliasing noises than Fig. 5(a), and Gaussian noise at approximately $PSNR = 34dB$ was added to these images. The results are shown in Fig. 6(a) and Fig. 6(b). In Fig. 6(a) and Fig. 6(b), the horizontal axis indicates the motion amount from the basic image, and the vertical axis indicates the average error of the estimated motion. The following five methods (inclusive of the proposed method) were compared:

(1) The block matching and parabola fitting (α=1)
(2) The two times expanded block matching and parabola fitting (α=2)
(3) The general gradient method
(4) The variable gradient method
(5) The proposed method

The block size M for estimating the motion was 21×21. In block matching based methods (1) and (2), the search area N was set to be 3×3, with α indicating the expansion ratio of the block matching process. In proposed method (5), the search area N was also set to 3×3. Motion estimation was calculated using 10 frames.

In Fig. 6(a), the accuracy of the motion estimation using the proposed method is the same as those obtained using the two times expanded block matching and parabola fitting, and the variable gradient method. In Fig. 6(b), the average errors of the motion estimation using block matching based methods (1) and (2) fluctuate significantly. The errors obtained using gradient based methods (3) and (4) have a lot. By contrast, the accuracy of the motion estimation using the proposed method, (5), is the highest on average. These results show that the proposed method is highly resistant to spatial aliasing noises.

B. Evaluation of Super Resolution

The super-resolution algorithm based on the integration of high frame-rate images is described in this section. As shown in Fig. 7, one high-resolution image is reconstructed using multiple images. First, registration is performed: The motion vectors are estimated from each of the reference frames to the target frame and pixel values are gathered to the target frame from reference frames based on the estimated vectors. Next, reconstruction is performed: The target frame with the pixel values of sub-pixel positions is filtered only once to reconstruct a high resolution image.

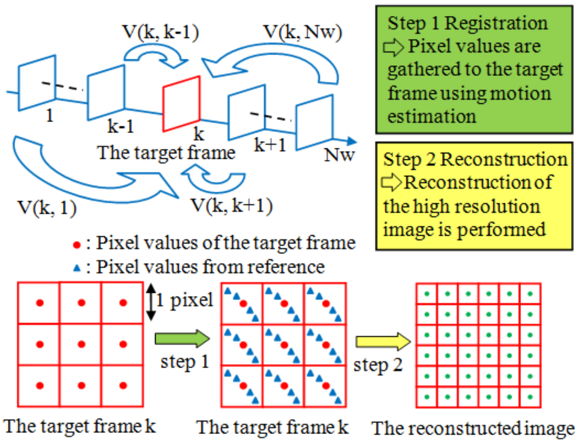


Figure 7. Flow of increasing spatial resolution.

In this experiment, a 384×512 image was reconstructed using nine 192×256 images. Fig. 8(a) shows the original image obtained at 300 fps. This image has aliasing errors, and the motion amount is approximately 0.25 pixel/frame. We used four of the motion estimation methods listed in the previous section (the block matching and parabola fitting (α=1) method was excluded). The bicubic interpolation algorithm was used to carry out the filtering for reconstruction.

(a)

(b)

(c)

(d)

(e)

Figure 8. Experimental results: (a) Original image. (b) Block matching and parabola fitting (α=2). (c) Gradient method. (d) Variable gradient method. (e) Proposed method.

TABLE I. RESULTS OF PSNR

Figure number	PSNR (dB)
8(b)	24.95
8(c)	24.90
8(d)	25.94
8(e)	26.26

Figs. 8(b)-Fig. 8(e) show the experimental results obtained using the super-resolution algorithm. The PSNR for each method is given in Table I. The results of PSNR using block matching and parabola fitting (b) and gradient method (c) are almost the same at approximately $25dB$. The results of PSNR using the variable gradient method (d) and the proposed method (e) are better than the above methods at approximately $26dB$. The PSNR using the proposed method (e) is slightly higher than that obtained using the gradient method (c). The results indicate that the accuracy of the super-resolution is improved by applying the proposed motion estimation method. The results of this experiment show that the proposed method is effective as a super-resolution algorithm for actual images that include spatial aliasing errors at a high frame-rate.

IV. CONCLUSION

This paper proposed a new motion estimation method based on block matching. The proposed method uses multiple frames and approximates motion vectors via a line to engender high resistant to spatial aliasing errors.

The results of the experiment conducted using images with low aliasing errors showed that the accuracy of the motion estimation by the proposed method is the same as that of previous methods. However, the results of the experiment conducted using images with aliasing errors show that the accuracy of the motion estimation by the proposed method is high on average. In addition, applying the proposed method to the super-resolution algorithm, the experimental results indicate that the proposed method improved the spatial resolution of high frame-rate images with spatial aliasing errors.

In future work, we will investigate not only the motion estimation method but also the super-resolution algorithm having resistant to spatial aliasing noises.

REFERENCES

[1] S. C. Park, M. K. Park, and M. G. Kang, "Super-Resolution image reconstruction: A technical overview," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 21-36, 2003.

[2] S. Lim, J. G. Apostolopoulos, and A. E. Gamal, "Optical flow estimation using temporally oversampled video," *IEEE Transactions on Image Processing*, vol. 14, no. 8, pp. 1074-1087, 2005.

[3] R. Battiti, E. Amaldi, and C. Koch, "Computing optical flow across multiple scale: An adaptive coarse-to-fine strategy," *International Journal of Computer Vision*, vol. 6, no. 2, pp. 133-145, 1995.

[4] D. Sugai, H. Katayama, and T. Hamamoto, "Image emphasis based on integration of high frame-rate images by new gradient-based motion estimation method," in *Proc. IEEE International Symposium on ISPACS*, 2012, pp. 118-123.

[5] Y. Sato, D. Sugai, and T. Hamamoto, "Gradient-Based motion estimation using characteristic of high frame-rate images," in *Proc. International Workshop on Advanced Image Technology*, 2013, pp. 23-28.

[6] M. Shimizu and M. Okutomi, "Precise sub-pixel estimation on area-based matching," in *Proc. Eighth IEEE International Conference on Computer Vision*, 2001, pp. 90-97.

Danya Sugai received the B.E. and the M.E. in electrical engineering from Tokyo University of Science in 2001, and 2003. He is currently a Ph.D. student at the Graduate School of Engineering, Tokyo University of Science. His research interests are image processing and image sensors.

Takayuki Hamamoto received the B.E. and the M.E. in Electrical Engineering from Tokyo University of Science in 1992, and 1994, respectively, and the Dr.E. in Electrical Engineering from the University of Tokyo, in 1997. Currently, he is a Professor at the department of electrical engineering of Tokyo University of Science
His current research interests are image processing and computational image sensors.