

Skew Correction of Digital Images for Stereo Matching

Ji-Hong Kim*
Dept. of Electronic Engineering
Dong-eui University
Busan, Korea

Hun Choi
Dept. of Electronic Engineering
Dong-eui University
Busan, Korea

Abstract—In the stereo vision, three-dimensional distance information of the object is extracted by using the left and right images acquired through stereo camera. The process of extracting the distance information requires the stereo matching that grasps the position of the corresponding pixel from each other in two images. The stereo matching needs a lot of calculation and the epipolar constraint is commonly used for reducing the amount of calculation. But the epipolar constraint cannot be applied accurately if two images for stereo matching have angle and/or height discrepancy. In this paper, an efficient scheme for correcting the discrepancy in stereo images is presented. The proposed method can be used as a preprocessing process for applying the epipolar constraint in stereo matching.

Keywords—Stereo Matching; Stereo Image; Epipolar Constraint; Skew Correction

I. INTRODUCTION

In the field of stereo vision, one of main issues is to extract three-dimensional distance information of the object in left and right images acquired through the stereo camera[1]. The distance information extracted from stereo images can be utilized in various applications, including image recognition, pattern recognition, and 2D/3D image conversion, *etc.* In order to extract the distance information from stereo images, it is essential to find pixels corresponding to each other in the left and right images. This process is called as matching and requires a lot of computational cost. Through a matching process, the difference in the positions of the pixels corresponding to each other, that is, disparity, can be calculated and by utilizing the disparity, it becomes possible to extract the distance information between the camera and the object. In the matching process, the epipolar constraint is generally used for reducing the amount of computation[2]. However, if the fine mismatch in the angle and/or height of the left and right cameras occurs, it may be impossible to find an accurate distance information through the epipolar constraint when the distance between the object and the camera is far away. Therefore the discrepancy in the angle and/or height in stereo images should be corrected before applying the epipolar constraint for the matching.

In the skew-correcting process, the feature points can be used for finding the corresponding pixels in the left and right images. There are some methods for extracting the feature points, including SIFT(Scale Invariant Feature Transform), Harris detector, and Moravec detector, *etc*[3, 4]. The SIFT extracts candidate feature points composed of the maximum

and minimum point, after constructing the DoG(Difference of Gaussian) images in scale space for the input image[3]. The extracted candidate feature point is revised in consideration of the stability and noise immunity, based on the direction and size of the peripheral region, to generate a descriptor. Such a descriptor has robust characteristics for the image rotation and changes in the light condition and image size.

In this paper, as a preprocessing process for the application of the epipolar constraint, an efficient method based on SIFT is presented for correcting the discrepancy in horizontal and/or vertical direction in stereo images. Feature points found through SIFT are utilized for calculating the discrepancy angle and/or height between the right and left images.

The organization of this paper is as follows. In section II, a method for extracting distance information in stereo images is briefly reviewed and the proposed method for correcting the discrepancy angle is described in section III. The simulation results using the proposed method are presented in section IV. Finally conclusions are drawn in section V.

II. DISTANCE INFORMATION EXTRACTION USING EPIPOLAR CONSTRAINT

In order to extract the distance information between the camera and the object in stereo images, it is essentially necessary to calculate the disparity of the points corresponding to each other in the left and right images. As shown in Fig. 1, assume that the points corresponding to an object X in stereo images are located at positions away x_L and x_R , respectively, from the center point of images[2].

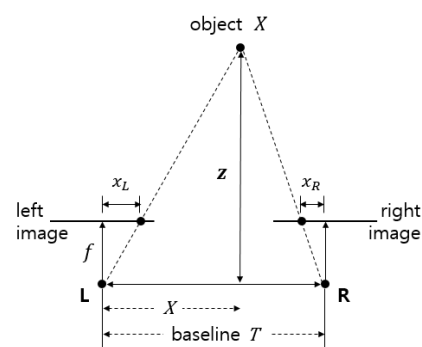


Fig. 1. Distance information and disparity

This work was supported by Dong-eui University Grant(2015AA126)

* : Corresponding Author

In Fig. 1, L and R stands for the left and right cameras, respectively, and z means the distance between the camera and the object. It can be seen from Fig. 1 that the relationship between the focal length of the camera f and the distance z is established as in (1) and (2).

$$f : x_L = z : X \quad (1)$$

$$f : (-x_R) = z : (T - X) \quad (2)$$

In (2), T stands for the baseline that means the distance between two cameras. From (1) and (2), the distance z is described as (3), where $x_L - x_R$ is defined as the disparity d .

$$z = \frac{f \cdot T}{x_L - x_R} = \frac{f \cdot T}{d} \quad (3)$$

Since f and T are known in general cases, it is necessary to calculate the disparity d for determining the distance z , and therefore the positions of corresponding points x_L and x_R for the object X should be found by the matching process.

In the matching process it takes a lot of computational time to find a pixel on the right image corresponding to the pixel on the left image. For example, for an image with $1,920 \times 1,080$ full HD resolution, the matching process for each pixel requires $1,920 \times 1,080$ comparisons. In order to reduce this computational time, the epipolar constraint can be utilized in the matching process. Epipolar constraint means that the pixel on the right image corresponding to the pixel on the left image is located in the epipolar line as shown in Fig. 2[2]. Therefore, if taking advantage of the epipolar constraint, the matching process is limited to the search process on 1D space instead of a 2D space, and therefore the computational complexity is reduced.

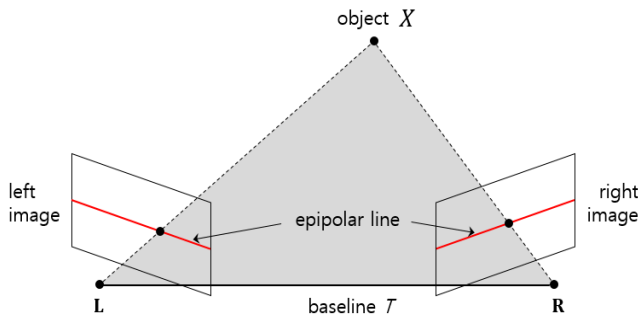


Fig. 2. Epipolar constraint

However, if there occurs a fine discrepancy in the mounting angle and/or height of the left and right cameras, an accurate distance information may not be found through the epipolar constraint when the distance between the object and the camera is far away. Therefore the angle and/or height discrepancy in stereo images should be corrected before applying the epipolar constraint in the matching process as shown in Fig. 3.

III. SKEW CORRECTION BASED ON SIFT

In this paper, an efficient method for correcting the mounting discrepancy of stereo camera using the SIFT is proposed and the overall process is shown in Fig. 4.

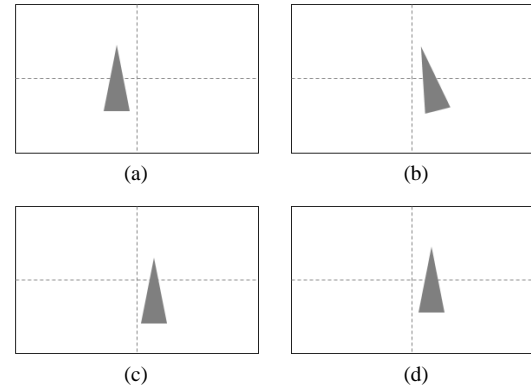


Fig. 3. (a) left image, (b) right image with angle discrepancy, (c) right image with height discrepancy, (d) right image corrected for (b) and (c)

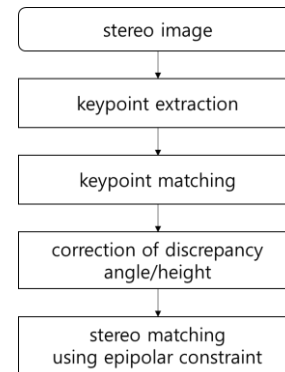


Fig. 4. Process of the skew correction for the stereo matching

A. Keypoint Extraction Using SIFT

The keypoint extraction process using the SIFT is composed of the keypoint detection and description stages[4]. In the keypoint detection stage, the keypoints are extracted from Gaussian difference between the scale spaces of stereo images and in the keypoint description stage, the orientation of the keypoints are determined. The scale image $S(x, y, \sigma)$ constructing a scale space is acquired by taking a convolution of an input image $I(x, y)$ with Gaussian operator $G(x, y, \sigma)$ as in (4), where σ is a scale parameter. The DoG image $D(x, y, \sigma)$ is determined by the difference between adjacent scale images as in (5).

$$L(x, y, \sigma) = G(x, y, \sigma) \otimes I(x, y) \quad (4)$$

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (5)$$

An example of scale space and DoG images acquired from the scale space is described in Fig. 5, where the scale space is composed of 3 octaves and 5 scales. The process of extracting keypoints is composed of two steps. In the first step, the maxima and minima which are used as candidates of the keypoint are detected. The method of detecting the maxima and minima is accomplished by comparing each pixel in the DoG image with 26 neighboring pixels, including

8 neighboring pixels in the same level scale image and 18 neighboring pixels in the higher and lower level scale images. In the second step, the process of correcting the maximum and minimum is implemented in the subpixel precision and the candidate keypoints having lower contrast are removed.

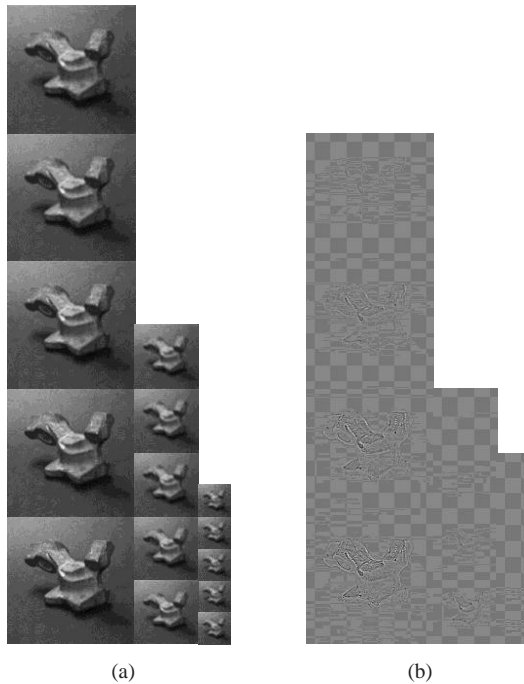


Fig. 5. Scale space and DoG images (reference image source: <http://vasc.ri.cmu.edu/idb/html/stereo/>)

B. Keypoints Matching and Discrepancy Correction

Prior to the keypoint matching, the matching area should be set up. The reason is there exists the disparity in the left and right images and by the disparity, it is not possible to implement the matching process in the boundary area of the stereo images. Therefore the matching area is set up as the image region except for the boundary region.

In the matching process, the Euclidean distance between keypoints found in the left and right images is used for determining whether or not matching. That is, if the Euclidean distance is less than a threshold value, the two keypoints are judged to be matched. An example of the keypoints matching in the left and right images is shown in Fig. 7.

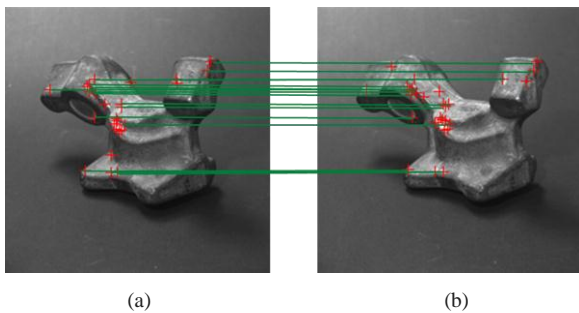


Fig. 6. Matching of keypoints (a) left image (b) right image

When the k -th keypoint matched each other in the left and right images are denoted as $I_L^k = (x_L^k, y_L^k)$ and $I_R^k = (x_R^k, y_R^k)$, the angles of the keypoints are described as (6).

$$\Theta_L^k = \tan^{-1}\left\{\frac{y_L^k - y_C}{x_L^k - x_C}\right\} \quad (6)$$

$$\Theta_R^k = \tan^{-1}\left\{\frac{y_R^k - y_C}{x_R^k - x_C}\right\}$$

In (6), (x_C, y_C) is the center point coordinate taken into account the disparity of the stereo images. That is, assuming that N keypoints are matched each other and the stereo images are of size $P \times Q$,

$$x_C = \frac{P}{2} + \frac{1}{2N} \sum_{k=0}^{N-1} (x_R^k - x_L^k) \quad (7)$$

$$y_C = \frac{Q}{2}$$

Therefore the angle discrepancy in the left and right images can be corrected by the rotation angle Θ_L and Θ_R , respectively, as in (8).

$$\Theta_L = \frac{1}{2N} \sum_{k=0}^{N-1} (\Theta_L^k - \Theta_R^k) \quad (8)$$

$$\Theta_R = \frac{1}{2N} \sum_{k=0}^{N-1} (\Theta_R^k - \Theta_L^k)$$

The height discrepancy is corrected by adjusting the vertical coordinates of the images. The adjusted vertical coordinates, \hat{y}_L and \hat{y}_R , in the left and right images are described by (10).

$$\hat{y}_L = y_L + \frac{1}{2N} \sum_{k=0}^{N-1} (y_R^k - y_L^k) \quad (9)$$

$$\hat{y}_R = y_R + \frac{1}{2N} \sum_{k=0}^{N-1} (y_L^k - y_R^k)$$

C. Stereo Matching Using Epipolar Constraint

After skew-correcting the angle and/or height discrepancy, the block matching method is used for searching for the corresponding block on the epipolar line of the left and right images as in Fig. 7. The degree of similarity between matching blocks in left and right images is calculated by using SSD (Sum of Squared Difference) of (10).

$$SSD(x, y; p) = \sum_{(i,j) \in W} [I_L(x+i, y+j) - I_R(x+i+p, y+j)]^2 \quad (10)$$

In (10), p stands for the distance between matching blocks in the left and right images, and W means the matching block.

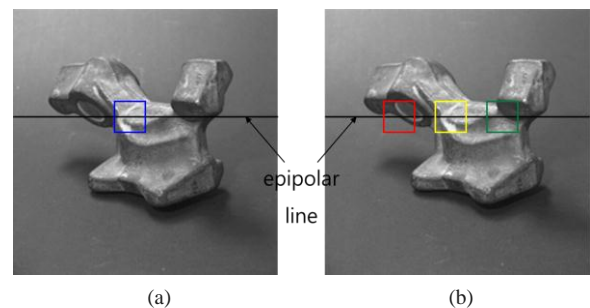


Fig. 7. Matching by block matching method (a) left image (b) right image

IV. SIMULATIONS AND ANALYSIS

In this paper, a new method for correcting the angle/height discrepancy in stereo images is proposed. In order to verify the effectiveness of the proposed method, the simulations are implemented for four stereo images as shown in Fig. 8. The stereo images are grey scale images of size 256×256 . In the simulations, the test images are made by rotating the right images by 1° in the range of -5° to $+5^\circ$ for verifying the performance of correcting the angle discrepancy. In Fig. 9, the keypoints in the left images and right images with angle discrepancy of -5° are depicted.

In the simulation, the boundary area that is not processed has the width of 8 pixels, and the window for employing the block matching method has the size of 8×8 . The simulation results for four test stereo images are depicted in Fig. 10, where the black and red lines describe the SSDs before and after the skew correction.

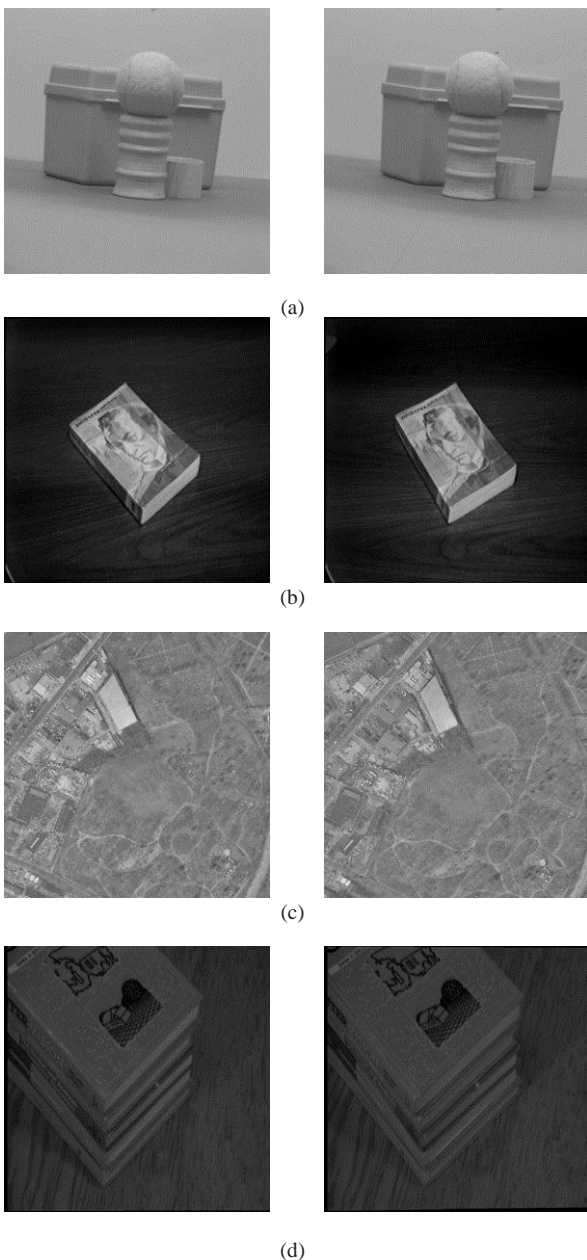


Fig. 8. Test stereo images. left side: the left image, right side: the right image. (a) ball (b) book (c) woodarea (d) books (image source:

<http://vasc.ri.cmu.edu/idb/html/stereo/>)

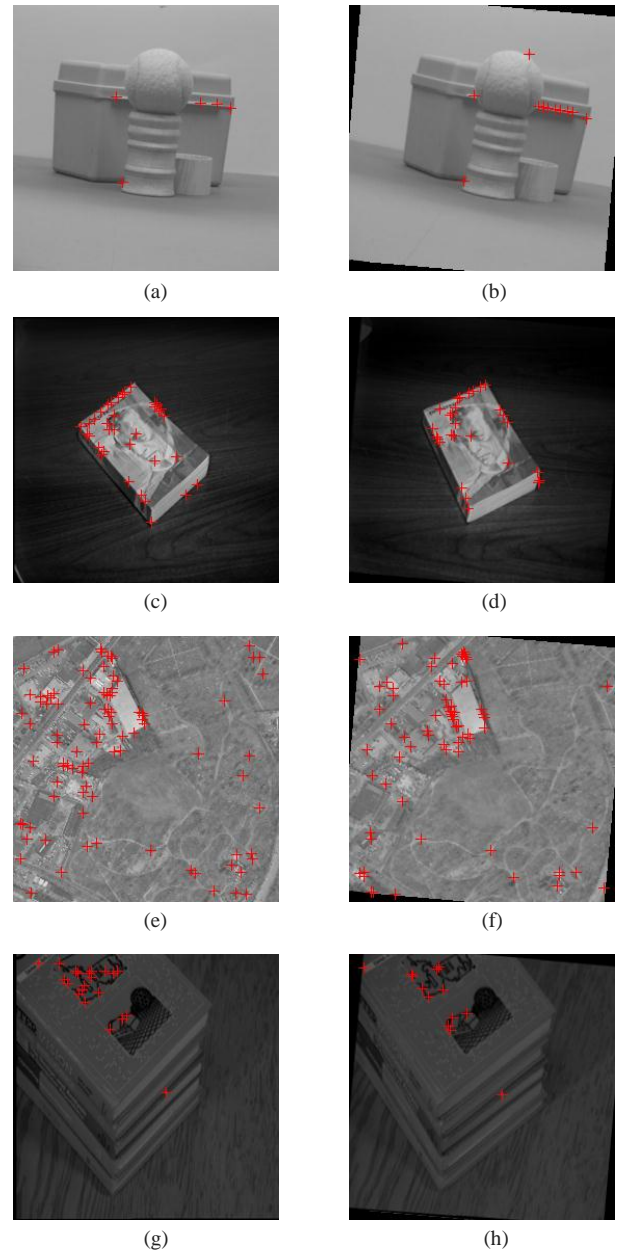


Fig. 9. Keypoints at the left images(a, c, e, g) and right images with angle discrepancy of -5° (b, d, f, h)

As shown in Fig. 10, the SSDs of the skew-corrected images have the lower value than those of the original test images. Therefore, we can see that the stereo matching using the epipolar constraint is better performed by applying the proposed skew correction method to the stereo image.

V. CONCLUSIONS

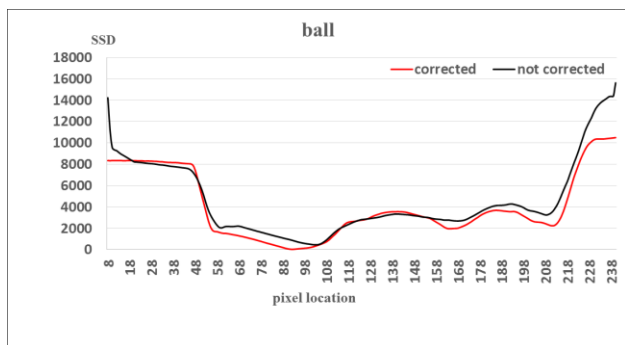
In this paper, an efficient scheme for correcting the angle/height discrepancy in the stereo image is proposed. The proposed method can be used as a preprocessing process for the application of the epipolar constraint. In the proposed method, the keypoints are detected based on the SIFT, and by comparing the location of the keypoints, the angle/height discrepancy in the stereo images are calculated and corrected. By the simulation results, the proposed method has an excellent performance.

ACKNOWLEDGMENT

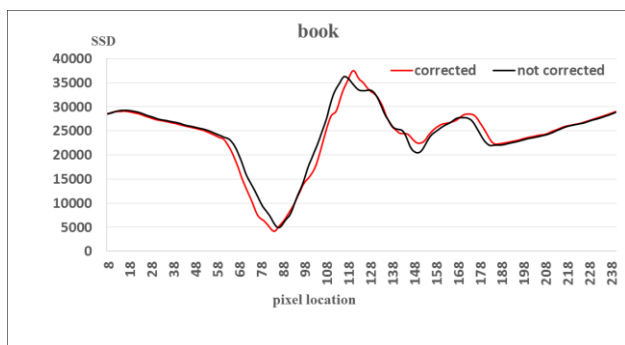
This work was supported by Dong-eui University Grant(2015AA126).

REFERENCES

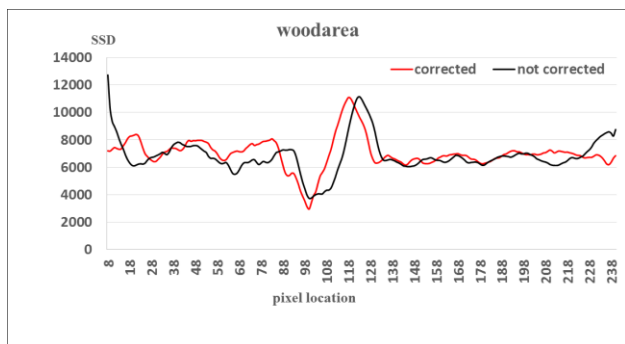
- [1] J. Kim, "Stereo Image Matching Using the SIFT," Proceeding of Korea Multimedia Society Spring Conference, Vol. 18, No. 1, pp.767-768, 2015.
- [2] H. Choi, Computer Vision, Hongrung Publishing Com., 2012.
- [3] David G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision, Vol. 60, pp. 91-110, 2004.
- [4] Y. Meng and B. Tiddeman, "Implementing the Scale Invariant Feature Transform Method," <http://huro-sift.googlecode.com/svn/tags/Final/Research/implementing.the.scale.invariant.feature.transform.sift.method.pdf>
- [5] J. Kim and S. Jang, "Correction of Rotated Region in Medical Images Using SIFT Features," Journal of Korea Multimedia Society, Vol. 18, No. 1, pp. 17-24, 2015.
- [6] Yu Meng and B. Tiddeman, "Implementing the Scale Invariant Feature Transform(SIFT) Method," <http://huro-sift.googlecode.com/svn/tags/Final/Research/implementing.the.scale.invariant.feature.transform.sift.method.pdf>
- [7] C. Harris and M. Stephens, "A Combined Corner and Edge Detector," 4-th Alvey Vision Conference, pp. 147-151, Manchester, 1988.



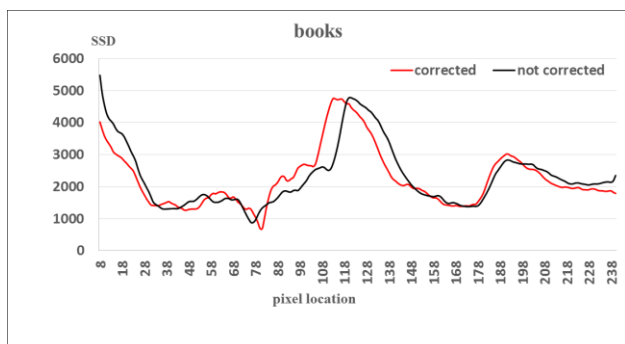
(a)



(b)



(c)



(d)

Fig. 10. SSDs before and after the skew correction (a) ball (b) book (c) woodarea (d) books