Geometric error correction of convergent multi-view images

Yun-Suk Kang[✉] and Yo-Sung Ho

A geometric error correction of convergent multi-view images is proposed. Shape factors and camera parameters of the captured multi-view image are used to estimate an ideal multi-camera arrangement and its camera parameters. A correction transform of each viewpoint to reduce the geometric error is calculated based on the original and estimated camera parameters. The error-corrected multi-view image is obtained by applying the transform to the captured images. The experimental results show that the proposed method efficiently reduces the geometric error of the convergent multi-view images.

Introduction: A multi-view image is one of the most appropriate data formats to realise the free viewpoint and the auto-stereoscopic three-dimensional (3D) videos [1]. To capture the multi-view image, multiple cameras are densely arranged in a certain shape. However, it is hard to obtain the ideal multi-view image due to a geometric error, which has appeared in the captured images. As shown in Fig. 1*a*, the geometric error is caused by a misalignment in the camera arrangement. It has appeared as vertical mismatches between corresponding pixels of adjacent viewpoints. Moreover, it is expressed as non-regular disparity values for the same depth plane. This geometric error shown in the multi-view image decreases not only the performance of 3D image processing such as depth estimation and view synthesis, but also the visual quality of the images as 3D contents.

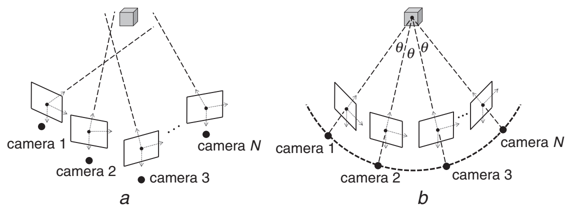

Fig. 1 *Convergent-type multi-camera arrangements*
a Practical arrangement
b Ideal arrangement

In the case of parallel-type multi-camera arrangements, multi-view image rectification becomes a solution to reduce the geometric error [2]. However, it cannot be applied to convergent-type multi-camera arrangements since the image rectification makes all image planes coplanar and all camera centres collinear. For the convergent-type multi-camera arrangements, it is important to regulate an angle between the adjacent viewpoints to be the same as well as to reduce the geometric error, as shown in Fig 1*b*. Although the convergent-type multi-camera arrangements are frequently used in the field, there have been few algorithms to reduce the geometric error of the convergent multi-view images. In this Letter, a correction method of the geometric error of the convergent multi-view images is proposed.

Geometric error correction of convergent multi-view images: In this Letter, the proposed method is used to estimate the ideal multi-camera arrangement and its corresponding camera parameters based on the shape factors and camera parameters of the captured multi-view image. Then, a correction transform $\boldsymbol{T}_n$ for the *n*th view is calculated by combining the original and estimated camera parameter sets. By applying $\boldsymbol{T}_n$ to the original multi-view image, the error-corrected multi-view image is obtained. Fig. 2 shows the procedure of the proposed method.

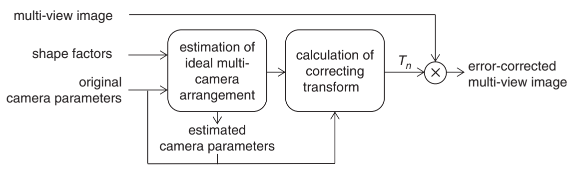

Fig. 2 *Procedure of proposed geometric error correction*

The estimation of the ideal multi-camera arrangement and its corresponding camera parameters is composed of three steps. The ideal conditions of the convergent-type multi-camera arrangement are an equal distance d and an equal angle θ between two neighbouring cameras, and identical camera internal characteristics. After calibrating the captured multi-view image, the original camera parameters and the shape factors d and θ are used for the estimation.

In the first step, reference axes $\boldsymbol{x}_\mathrm{r}$, $\boldsymbol{y}_\mathrm{r}$ and z_r are decided. In the proposed method, the horizontal, vertical and principal axes of the *n*th camera are denoted by $\boldsymbol{x}_n$, $\boldsymbol{y}_n$ and z_n, where $1 \le n \le N$. Since the vertical axes in the convergent-type multi-camera arrangement have similar directions, r is calculated by (1), where $\boldsymbol{y}'_\mathrm{avg}$ is the unit vector of the average direction of the filtered $\boldsymbol{y}_n$. The filtering of $\boldsymbol{y}_n$ calculates an average direction $\boldsymbol{y}_\mathrm{avg}$ and standard deviation σ_y of all $\boldsymbol{y}_n$, and then exclude $\boldsymbol{y}_n$ that is located outside the cubic region bounded by $2\sigma_y$ based on $\boldsymbol{y}_\mathrm{avg}$. Fig. 3 shows this outlier filtering process of $\boldsymbol{y}_n$ [2]

$$r = \arg\min_n \left(\left| \boldsymbol{y}_n - \boldsymbol{y}'_\mathrm{avg} \right| \right) \tag{1}$$

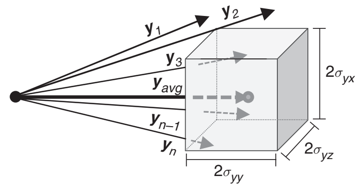

Fig. 3 *Outlier filtering process*

In the second step, camera extrinsic parameters are estimated based on the reference axes and the shape factors. In the proposed method, the estimated parameters have a subscript *i* which means the ideal multi-camera arrangement. To estimate the camera centres $\boldsymbol{C}_{i,n}$, one of the original camera centres is selected as a basis point $\boldsymbol{C}_{i,b}$, i.e. $\boldsymbol{C}_{i,b} = \boldsymbol{C}_b$, where $1 \le b \le N$. Moreover, this basis point has the reference axes as its coordinate system, i.e. $\boldsymbol{x}_{\iota,b} = \boldsymbol{x}_\mathrm{r}$, $\boldsymbol{y}_{\iota,b} = \boldsymbol{y}_\mathrm{r}$ and $z_{\iota,b} = z_\mathrm{r}$. From $\boldsymbol{C}_{i,b}$, $\boldsymbol{C}_{i,n}$ is recursively obtained by (2) for all n except for the case of $n = b$, where $\alpha = \sqrt{(1/\cos^2\theta) - 1}$. b of $\boldsymbol{C}_{i,b}$ is decided beforehand as (3) using (2) to minimise the sum of the absolute distances between the original and estimated camera centres

$$\boldsymbol{C}_{i,n} = \boldsymbol{C}_{i,n\pm1} - d\left(\frac{\alpha z_{\iota,n} \pm \boldsymbol{x}_{\iota,n}}{\left\| \alpha z_{\iota,n} \pm \boldsymbol{x}_{\iota,n} \right\|} \right) \tag{2}$$

$$b = \arg\min_m \sum_{\substack{n=1 \\ n \ne m}}^{N} \left| \boldsymbol{C}_{i,n} - \boldsymbol{C}_m \right| \tag{3}$$

After estimating the camera centres, a rotation matrix $\boldsymbol{R}_{i,n}$ of the *n*th camera is also defined by (4). However, the case of $n = b$ is exempted since $\boldsymbol{R}_{i,b}$ is already defined as $\left[\boldsymbol{x}^\mathrm{T}_{\iota,b}\, \boldsymbol{y}^\mathrm{T}_{\iota,b}\, z^\mathrm{T}_{\iota,b}\right]^\mathrm{T}$. $\boldsymbol{x}_{\iota,n}$ in (4) is defined as (5). Fig. 4 describes the whole process in the second step

$$\boldsymbol{R}_{i,n} = \left[\boldsymbol{x}^\mathrm{T}_{\iota,n} \quad \boldsymbol{y}^\mathrm{T}_{\iota,b} \quad \left(\boldsymbol{x}^\mathrm{T}_{\iota,n} \times \boldsymbol{y}^\mathrm{T}_{\iota,b} \right) \right]^\mathrm{T} \tag{4}$$

$$\boldsymbol{x}_{\iota,n} = \frac{\alpha \boldsymbol{z}_{\iota,n\mp1} \pm \boldsymbol{x}_{\iota,n\mp1}}{\left\| \alpha \boldsymbol{z}_{\iota,n\mp1} \pm \boldsymbol{x}_{\iota,n\mp1} \right\|} \tag{5}$$

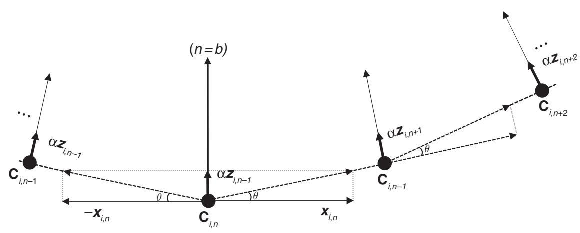

Fig. 4 *Estimation of camera extrinsic parameters*

In the third step, the camera intrinsic parameters are estimated as components of a common intrinsic matrix $\boldsymbol{A}_i$. These parameters are defined as certain values between the minimum and maximum of their original values to minimise the sum of the absolute differences between the original and estimated values [2].

Through these three steps, the camera projection matrix $\boldsymbol{P}_{i,n}$ of the nth camera is finally defined as (6). By using $\boldsymbol{P}_{i,n}$ and the pseudo-inverse of the original projection matrix $\boldsymbol{P}_n$, $\boldsymbol{T}_n$ is calculated by (7). By applying $\boldsymbol{T}_n$ to the nth viewpoint image, the error-corrected multi-view image is obtained

$$\boldsymbol{P}_{i,n} = \boldsymbol{A}_i[\boldsymbol{R}_{i,n}| - \boldsymbol{R}_{i,n}\boldsymbol{C}_{i,n}] \tag{6}$$

$$\boldsymbol{T}_n = \boldsymbol{P}_{i,n}\boldsymbol{P}_n^+ \tag{7}$$

Experimental results: To test the proposed method, two sets of convergent multi-view images are captured and calibrated [3]. The first one is composed of the five-view and has an angle of ~8°. The other one is the nine-view with an approximate angle of 4°. These two image sets are shown in Fig. 5*a*.The overlapped images shown in Fig. 6*a* show the geometric error in the captured images.

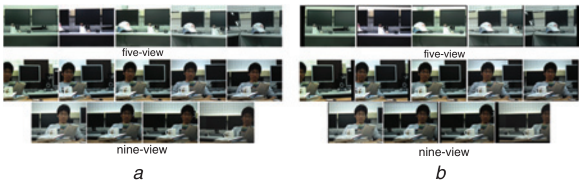

Fig. 5 *Convergent multi-view images for experiments*

a Captured multi-view images
b Error-corrected multi-view images

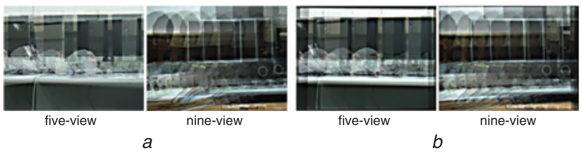

Fig. 6 *Overlapped multi-view images*

a Captured multi-view images
b Error-corrected multi-view images

The results of the proposed method are shown in Fig. 5*b*. The geometric errors in the vertical and horizontal directions are reduced as shown in Fig. 6*b*. For an objective evaluation, three feature points from the five-view image and four feature points from the nine-view image are manually selected in all the viewpoints. Fig. 7*a* shows the average vertical pixel mismatches with respect to the feature points were reduced by the proposed method. Moreover, the disparity deviation of each feature point for all the viewpoints was decreased, as shown in Fig. 7*b*.

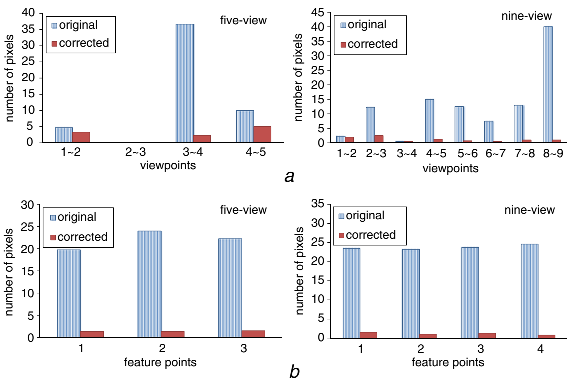

Fig. 7 *Objective evaluation of proposed method*

a Average vertical pixel mismatches for feature points
b Average disparity deviation for feature points

Conclusion: In this Letter, a geometric error correction of convergent multi-view images is proposed to reduce the geometric error and obtain the error-corrected multi-view image. From the experimental results, it is confirmed that the proposed method can efficiently reduce the geometric error such as the vertical mismatches and the non-regular horizontal displacements in pixels.

Acknowledgment: This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (no. 2011-0030079).

2 October 2014
doi: 10.1049/el.2014.3541
One or more of the Figures in this Letter are available in colour online.

Yun-Suk Kang and Yo-Sung Ho (*Gwangju Institute of Science and Technology, 123 Cheomdangwagi-ro, Buk-gu, Gwangju, Republic of Korea*)

✉ E-mail: yunsuk@gist.ac.kr

References

1 Merkle, P., Smolic, A., Muller, K., and Wiegand, T.: ‘Multi-view video plus depth representation and coding’. Proc. Int. Conf. on Image Processing, San Antonio, TX, USA, September 2007, pp. 201–204

2 Kang, Y.S., and Ho, Y.S.: ‘An efficient image rectification method for parallel multi-camera arrangement’, *IEEE Trans. Consum. Electron.*, 2011, **57**, (3), pp. 1041–1048

3 Zhang, Z.: ‘A flexible new technique for camera calibration’, *IEEE Trans. Pattern Anal. Mach. Intell.*, 2000, **22**, (11), pp. 1330–1334