ORIGINAL PAPER

Video image mosaic implement based on planar-mirror-based catadioptric system

Lin Zeng · Weigang Zhang · Shengping Zhang · Dingwen Wang

Abstract Planar-mirror-based catadioptric method is one of the most hot topics in recent years. To overcome the disadvantages of the planar-mirror-based catadioptric panoramic camera, described by Nalwa (1996, 2001, 2000), such as the requirement for high-precision optical device designing and the stitching lines in the resulting images, we proposed a planar-mirror-based video image mosaic system with high precision for designing. Firstly, we designed a screw nut on our system, which can be adjusted to locate the viewpoints of the cameras' mirror images at a single point approximately. It provides a method for those who have difficulties in their designing and manufacturing for high precision. Then, after the image distortion correction and cylinder projection transforms, we can stitch the images to get a wide field of view image by template matching algorithm. Finally, nonlinear weighting fusion is adopted to eliminate the stitching line effectively. The experimental results show that our system has good characteristics such as video rate capture, high resolution, no stitching line and without affection by the depth of field of view.

L. Zeng
School of Electronic Information,
Wuhan University, Wuhan, 430072, China
e-mail: zenglin.whu@gmail.com

W. Zhang · S. Zhang
School of Computer Science and Technology,
Harbin Institute of Technology at Weihai,
Weihai, 264209, China
e-mail: wgzhang@jdl.ac.cn

S. Zhang
e-mail: shengping.zhang@gmail.com

D. Wang (✉)
Institute of Microelectronics and Information Technology,
Wuhan University, Wuhan, 430072, China
e-mail: wangdw@whu.edu.cn

1 Introduction

Wide field of view (FOV) and high-resolution image acquisition have become an important subject for research in vision-based applications. Due to its high practicability, it has been widely used in intelligent video surveillance systems, such as moving object detection [4–7] and tracking [8–11]. Techniques for constructing wide FOV image can be classified into two categories: dioptric methods and catadioptric methods [12].

Dioptric methods mean that only refractive elements (such as lens) are employed. Fisheye lens [13–15], rotating cameras [16,17], and camera clusters [18–22] are most widely used in this kind of methods. Although real-time wide FOV images can be acquired by fisheye lens, the resolutions of the images are usually low due to the limitation of the single image sensor in the fisheye lens. Also, the distortion of near lens scene caused by fisheye lens cannot be resolved. Rotating cameras can achieve high resolution of the FOV images. However, as the single camera can only get a limited visual angle of the scene, it is impossible to acquire real-time wide FOV images. Camera clusters that capture images simultaneously can acquire wide FOV images at real-time video rate. But the optic center of the cameras cannot locate at a single point because of the cameras' space limitations, thus the parallax caused by these limitations always leads to ghost phenomenon in the mosaic image as affected by the depth of field of view.

Catadioptric methods mean a combination of refractive components and reflective components. There are two specific methods: the curved mirror method and the planar-

mirror-based catadioptric method. The curved mirror methods [23–27] use the combination of curved mirror and single camera, enlarge the visual angle of the single camera by the reflective effects of the curved mirror. These methods are able to acquire wide FOV images in real time without stitching line. But as the same as the fisheye lens, it has the disadvantages of low resolution. The planar-mirror-based catadioptric methods are widely used to acquire wide FOV images in recent years. These methods are firstly proposed by Nalwa [1–3]. He obtained 360° cylindrical panoramic video images by planar-mirror-based catadioptric system consisting of 4 cameras and 4 trapezoid mirrors and achieved ideal performance. According to Nalwas' idea, Hua [28,29] obtained cylindrical panoramic video images with the wide vertical view angle by bilayer planar mirrors. Gao [30] also obtained real-time hemisphere panoramic video images using hexahedral planar mirrors. The innovation of planar-mirror-based catadioptric methods is that after focusing different cameras' optic centers into a single point by planar mirrors, parallax between images captured by different cameras can be eliminated and the final wide FOV images have the advantages of good real-time performance, high resolution and not affected by the depth of field. However, Nalwas' method uses planar-mirror-based catadioptric system to focus optical axis of different cameras into a single point strictly; the accuracy of the mirrors and the equipments is really high. Because there is no overlap region between images, it is impossible to make image fusion at the stitch areas. Thus, there will be obvious stitching lines in the resulting wide FOV images.

Image registration is the basis of video image mosaic. Commonly used registration methods are mainly divided into three classes: template matching, mutual information, and features-based methods. Template matching methods firstly get a matching template by selecting a window of gray information from the overlap areas directly [31] and then search in another image until reach the highest matching score. This kind of methods can solve the image registration with low computational complexity. Mutual information methods were firstly proposed by Kughn and Hines [32] in 1975. The early mutual information methods can only be used for image registration when there is only pure translations between two images. Then, Keller et al. [33] introduced the polar Fourier transform into the mutual information methods, which can achieve the improved image registration stability with translation, rotation, and scale transformation. Zhou et al. [34] propose a Bayesian-based mutual information technique, combined with an established affine transformation model, which can register images with affine transformation efficiently and accurately. However, mutual information methods often require the mosaic images have larger overlap area and would fail in a smaller percentage of overlap situation. Corner features algorithm was proposed by
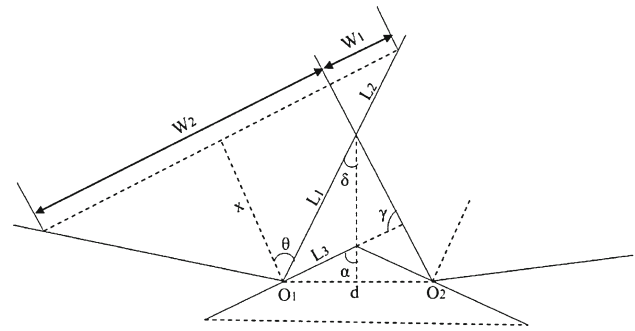

Fig. 1 The geometric relationship between the adjacent cameras

Harrs [35] in 1998, which is the most representative method in features-based registration. This method is invariant with respect to difference of rotation and image brightness and can be widely used in 2D image registration [36] and 3D reconstruction [37]. But it has higher computational complexity, which make it not suitable for real-time demanding registration occasion.

In this paper, to describe the fault of the traditional planar-mirror-based catadioptric methods, which need high optical designing precision, we first establish a mathematical model to demonstrate the influence of the parallax to the image mosaic and then proposed a mirror pyramid camera which has undemanding requirement of machining precision and can focus 3 cameras' optic centers into one point approximately by adjusting the screw nuts. Thus, the final FOV images can be achieved by image registration. Finally, to eliminate the stitch line caused by optical difference effectively, we proposed an improved algorithm based on nonlinear weighting fusion.

2 Design of planar-mirror-based catadioptric system

Although the influence of the parallax on image mosaic caused by the variance of viewpoints was illustrated in [38] by comparing the results of the image mosaic system, it cannot be proved by the mathematical theory. In order to proposed a new image mosaic method which has the characters of no stitching line and invariance to depth of field, we first analyze the causes of viewpoint variance and then prove the inevitability by establishing a mathematics model based on the geometric relationships. According to the mathematical model, a planar-mirror-based system can be designed and the ideal image mosaic results can be achieved.

Let $\langle \text{Camera}_1, \text{Camera}_2, \ldots, \text{Camera}_n \rangle$ be the camera cluster where every camera has a vertical line VL_n, which is perpendicular to the direction of its viewpoint. As shown in Fig. 1, taking one of the adjacent cameras $\langle \text{Camera}_1, \text{Camera}_2 \rangle$, for example, O_1 and O_2 are the optic centers of Camera$_1$ and Camera$_2$, respectively. The angle between $\langle VL_1, VL_2 \rangle$ is 2α. The distance between two optic centers
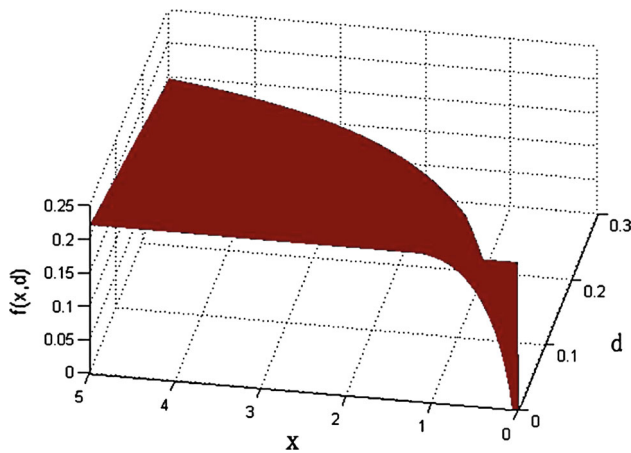
Fig. 2 The 3-dimensional plot of function $f(x, d)$

is d. The viewpoint of Camera$_n$ is 2θ, and the depth of the field of Camera$_n$ is x. L_1, L_2, L_3, w_1, w_2 and γ are shown in Fig. 1. According to the principles of geometry, we have the following equations:

$$\delta = \alpha + \theta - 90° \quad \gamma = 270° - 2\alpha - \theta \tag{1}$$

$$w_1 + w_2 = 2x \tan \theta \tag{2}$$

$$L_1 = \frac{d}{2 \sin \delta}, \quad L_2 = \frac{x}{\cos \theta} - L, \quad \frac{L_1}{\sin \gamma} = \frac{L_3}{\sin 2\delta} \tag{3}$$

$$\frac{w_1}{L_1} = \frac{L_2}{L_1} \tag{4}$$

$$f(x, d) = \frac{w_1}{w_1 + w_2} \tag{5}$$

According to Eqs. (1), (2), (3), (4) and (5), we have:

$$f(x, d) = \frac{\sin 2\delta}{2 \sin \theta \sin \gamma} - \frac{d \cos \delta}{2x \sin \gamma \tan \theta}$$

Because α and θ are known for a specific fixed panoramic camera system, the equation above can be modified to give

$$f(x, d) = t_1 - t_2 \frac{d}{x}$$

where $t_1 = \frac{\sin 2\delta}{2 \sin \theta \sin \gamma}$, $t_2 = \frac{\cos \delta}{2 \sin \gamma \tan \theta}$.

For the regular decagon camera system in this paper, $\theta = 22.5°$, $\alpha = 72°$, we can obtain

$$t_1 = 0.211, t_2 = 1.238.$$

$$f(x, d) = 0.221 - 1.238 \frac{d}{x}$$

The 3-dimensional plot of function $f(x, d)$ is shown in Fig. 2.

As shown in Fig. 2, supposing $f(x, d)$ denotes the normalized overlap region between 2 nearby cameras, x and d indicate the distance from the scene being photographed to the cameras and the distance between the nearby cameras' optic centers, respectively. Then, we will come to the following conclusions:
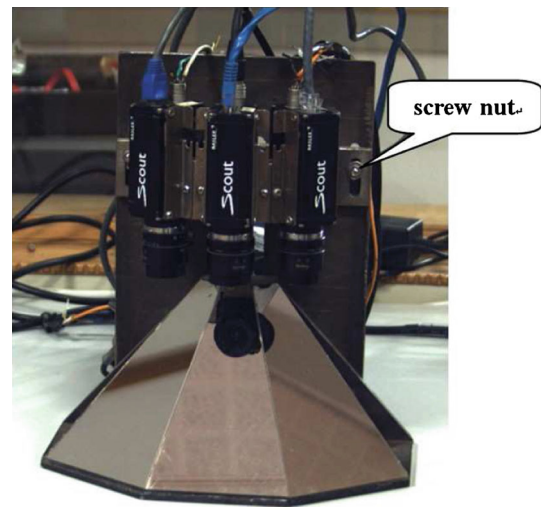

Fig. 3 The planar-mirror-based catadioptric system consisting of 3 trapezoid plane mirrors, 3 gigabit Ethernet cameras, and a screw nut to adjust the camera's height

1. If $d = 0$, then $f(x, d)$ is a constant value. It means that depth of field has no effect on $f(x, d)$;
2. If $d > 0$ (the nearby cameras' optic centers are not at the same point), then the value of $f(x, d)$ changes with depth of field x. As x increases, $f(x, d)$ tends to be a constant value (the value is 0.221 in this paper). While the smaller the value of d is, the faster the change rate of $f(x, d)$ to the constant value will be as x increases.

So, the overlap region approximates the constant value 0.221 as the distance between two cameras' optic centers approaches 0. The overlap region can be eliminated when the distance is 0. However, the distance between optic centers of multi-cameras cannot be 0 in engineering and reality. Thus, the overlap region between adjacent cameras cannot be avoided and so as its influence to image mosaic.

According to the mathematical equations above, the planar-mirror-based system we designed set 3 viewpoints of the cameras into an approximate one point by reflection principle of the plane mirrors. Then, the problem of space-limited viewpoint variance of the 3 cameras can be resolved. As shown in Fig. 3, this system is mainly composed by 3 trapezoid plane mirrors, 3 cameras, and a screw nut to adjust the camera's height. The optic axel of each cameras goes through the bisector of its corresponding plane mirror and is perpendicular to the pedestal. Figure 4 shows the system profile's optical path diagrams which passes through the bisector of trapezoid plane mirror. C is the optic center of one of the cameras, and C' is its image in the mirror. The angle between the mirror and the pedestal is 45°. Figure 5 shows the cross-section view of our catadioptric system's base, B1, B2, and B3, is the intersecting lines between the 3 mirrors and the
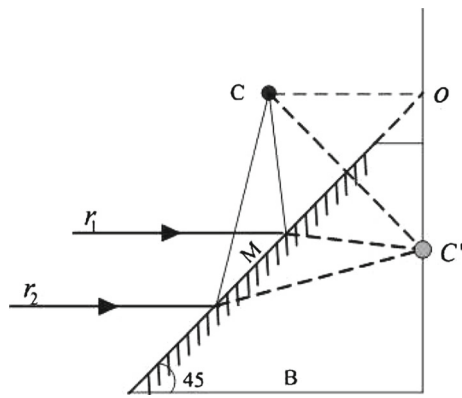
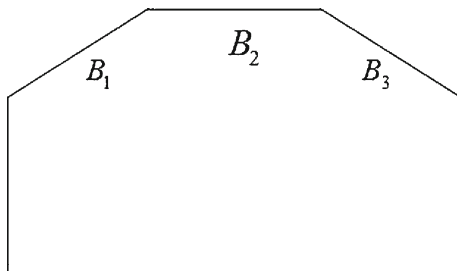**Fig. 4** The optical path diagrams of the planar-mirror-based catadioptric system's profile

Fig. 5 The cross-section view of the planar-mirror-based catadioptric system's base, *B1*, *B2*, and *B3*, represents to be 3 adjacent sides of a regular decagon

pedestal. The horizontal view angle of wide FOV image is 108° in this paper. As 360°/(108°/3) = 10, B1, B2 and B3 are designed to be 3 adjacent sides of a regular decagon.

The height between the cameras and the pedestal in the planar-mirror-based catadioptric system can be adjusted by the screw nut (as shown in Fig. 3). As shown in Fig. 4, when the height of camera's optic center *C* is adjusted to be level to point *O*, the optic centers of the 3 cameras can be focused into a common point in the mirror image. In this paper, the height of the 3 cameras is adjusted to be a little lower than the ideal *O* point to ensure there will be some overlap regions (approximately 20 pixels) between the adjacent cameras. This setting can guarantee that the optic centers can focus into one point approximately as well as fusing the images to eliminate the stitching lines by the overlap information between adjacent images.

3 The processes of wide FOV images composition

There are three steps for computing the wide FOV images: First, correct the distortion of the video images captured by the cameras and project these corrected images onto a common cylindrical surface. Second, image registration can be made according to template matching algorithm, and the

three images can be stitched into a wide FOV image. Third, fuse the stitching lines of the wide FOV image by nonlinear weighting fusion algorithm.

3.1 Image distortion correction and cylindrical projection transformation

The optical lens cannot satisfy the requirement of the ideal pinhole imaging model; then, there will be a certain degree of radial distortion and slight tangential distortion. So it is necessary to make image distortion correction before cylindrical projection transformation. This paper adopts the distortion model [39] as follows:

$$x' = \frac{x}{f_x}, \ y' = \frac{y}{f_y} \tag{6}$$

$$x'' = x' \cdot (1 + k_1 r^2 + k_2 r^4) + 2p_1 x' y' + p_2 (r^2 + 2x'^2) \tag{7}$$

$$y'' = y' \cdot (1 + k_1 r^2 + k_2 r^4) + p_1 (r^2 + 2y'^2) + 2p_2 x' y' \tag{8}$$

where

$$r^2 = x'^2 + y'^2 \tag{9}$$

Finally, the following equation can be obtained:

$$u = f_x \cdot x'' + c_x, v = f_y \cdot y'' + c_y \tag{10}$$

From Eqs. 6 to 10, (x, y) is the ideal coordinate of formatted image. (u, v) is the coordinate of distorted image. (c_x, c_y) is the reference point of the image. k_1 and k_2 are the coefficients of radial distortion. p_1 and p_1 are the coefficients of tangential distortion. (f_x, f_y) is the focus length of the camera in pixels. In this paper, we use the OpenCV library[1] to correct the distortion, and high order coefficients are not considered.

We use the cylindrical transform equation in [40] to project the three distortion-corrected images to the same cylindrical surface before performing image mosaic. Figure 6a is the original image captured by polygonal mirrors catadioptric camera. Figure 6b is the image after distortion correction and cylindrical projection. The lookup table of distortion correction and cylindrical projection can be calculated during the initialization of the system. Further transformation of images can be obtained through bilinear interpolation calculation from the lookup table. Thus, real-time computing can be guaranteed.

3.2 Image registration

The overlap region between the adjacent images of the catadioptric system is small (approximately 20 pixels in horizontal direction). Besides, rotation and zoom of the image can be ignored and only the horizontal and vertical offsets are
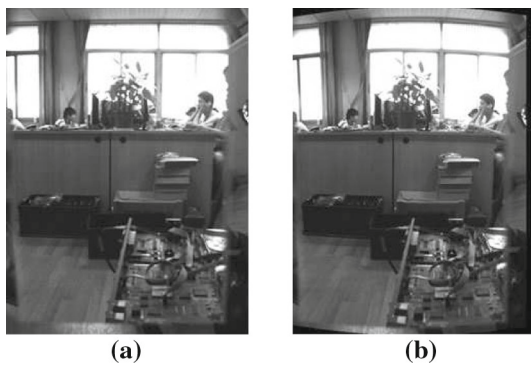
---

[1] http://sourceforge.net/projects/opencvlibrary/

Fig. 6 The comparison between the original image and the processed image. **a** The original image. **b** The image after distortion correction and cylindrical projection
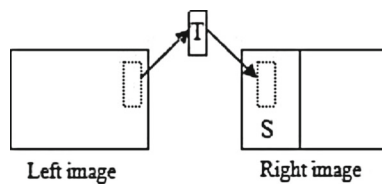

Fig. 7 The schematic diagram of template matching algorithm, T indicates a template from the left image

Fig. 8 The schematic diagram of nonlinear weighting algorithm

considered. So the position relationship of the images can be calculated by template matching algorithm.

As shown in Fig. 7, suppose the overlap region between two images exists. T is the template image taken from left image, and S is the search region in the right image for T. Suppose template T is overlapped on S and scans over the whole region. The searching image under the template image T is called sub-image $S^{(i,j)}$, where (i, j) is the left-top coordinate of sub-image $S^{(i,j)}$ in the search region. If the size of S is $X \times Y$ and the size of T is $M \times N$, then the ranges of i, j are $0 < i < X - M + 1$ and $0 < j < Y - N + 1$.

Compare each $S^{(i,j)}$ in search region S and template T, there will always be a sub-image $S^{(i,j)}$ which is identical to T under ideal condition (the variance between T and $S^{(i,j)}$ is 0). However, as they are subject to lightning condition, noise, and the differences of sensors, the variance cannot be 0. Therefore, squared error $D(i, j)$ can be used to demonstrate the resemblance of T and $S^{(i,j)}$.

$$D(i, j) = \sum_{x=1}^{M} \sum_{y=1}^{N} [T(x, y) - S^{(i,j)}(x, y)]^2 \qquad (11)$$

The equation can be expanded as

$$D(i, j) = \sum_{x=1}^{M} \sum_{y=1}^{N} [T(x, y)]^2 + \sum_{x=1}^{M} \sum_{y=1}^{N} [S^{(i,j)}(x, y)]^2$$
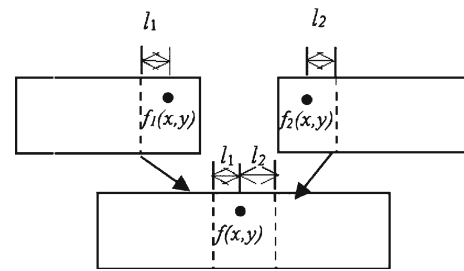$$-2 \sum_{x=1}^{M} \sum_{y=1}^{N} T(x, y) \times S^{(i,j)}(x, y) \qquad (12)$$

The first item in Eq. (12) means the energy of the template, and it is a constant value which is not affected by (i, j). The second item means the energy of $S^{(i,j)}$ and changes slowly with (i, j). The third item is the correlation between template T and $S^{(i,j)}$, and its value changes with (i, j). When the images are registered, it has the maximum value. So the normalized correlation function $R(i, j)$ can be used to measure the resemblance:

$$R(i, j) = \frac{\sum_{x=1}^{N} \sum_{y=1}^{N} T(x, y) \times S^{(i,j)}(x, y)}{\sqrt{\sum_{x=1}^{N} \sum_{y=1}^{N} [T(x, y)]^2} \sqrt{\sum_{x=1}^{N} \sum_{y=1}^{N} [S^{(i,j)}(x, y)]^2}} \qquad (13)$$

According to Schwarz's inequality, as shown is Eq. (13), the range of $R(i, j)$ is : $0 < R(i, j) < 1$. The best match point is the place which has the greatest value of $R(i, j)$ when template T scans over search region S.

It is easy to mosaic three images into a wide FOV image from the matching result between adjacent images by the template matching algorithm. The matching algorithm is calculated only once at the initialization, and the following video stream mosaicing can use its standard parameter repeatedly.

3.3 Image fusion

There is brightness variance among captured images, due to many factors such as lighting condition, exposure compensation. It lead to obvious stitching lines in the fused images. In this paper, an self-adaptive nonlinear weighting is adopted to fuse the variance gradually.

Supposing that the width of the overlap region between two image is $l_1 + l_2$ as shown in Fig. 8, and the left, right, and fused images are denoted as $f_1(x, y)$, $f_2(x, y)$, and $f(x, y)$. If the average brightness of $f_1(x, y)$ and $f_2(x, y)$ is $\overline{M_1}$ and $\overline{M_2}$. The average of them is $\overline{M} = \frac{\overline{M_1} + \overline{M_2}}{2}$. Let $\theta_1 = \frac{l_2 \cdot \pi}{l_1 + l_2}$, $\theta_2 = \frac{l_1 \cdot \pi}{l_1 + l_2}$, the nonlinear weighting function is

$$d_1 = \frac{\sin(\theta_1 - \frac{\pi}{2})}{2} + \frac{1}{2} \qquad (14)$$

$$d_2 = \frac{\sin(\theta_2 - \frac{\pi}{2})}{2} + \frac{1}{2} \qquad (15)$$
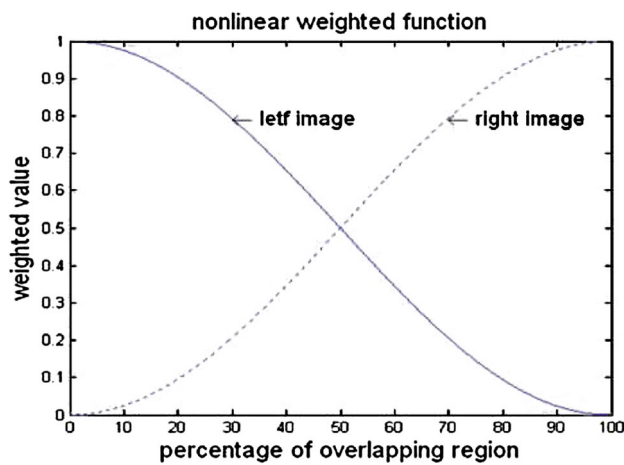
Fig. 9 The nonlinear weighting coefficient function

The weighting coefficients above are used to adjust the brightness of the overlap region. Figure 9 shows the function figure for Eqs. (14) and (15). Meanwhile, we take advantage of the following expressions to deal with absolute luminance difference

$$f'_1(x, y) = f_1(x, y) + d_1 \times (\overline{M} - \overline{M_1}) \tag{16}$$

$$f'_2(x, y) = f_2(x, y) + d_2 \times (\overline{M} - \overline{M_2}) \tag{17}$$

In Eqs. (16) and (17), $f_1(x, y)$ and $f_2(x, y)$ represent the original images, and $f'_1(x, y)$ and $f'_2(x, y)$ are images after we adjusted the luminance that needed to be fused, so the fused image $f(x, y)$ can be expressed as Eq. (18)

$$f(x, y) = d_1 \times f'_1(x, y) + d_2 \times f'_2(x, y) \tag{18}$$

4 Experiment results and analysis

The experimental results are obtained by executing on a computer with Pentium Dual Core E6500 2.93 GHz CPU integrating with three gigabit network cards. The planar-mirror-based catadioptric system includes three BASLER scA780-54gm gigabit Ethernet cameras which image capture achieve 54 fps at the 782×582 resolution. In order to expand the vertical viewpoint of the camera, the camera is rotated 90° to get 582×782 resolution. In order to validate the video image mosaic results by our designing method, we will design our experiments by the following three procedures:

– Presenting the flow diagram of image mosaic process;
– Mosaic results at short range;
– Mosaic results at distant range.

First, we take three video images of our laboratory scene by three cameras at the same time, and then we carry on
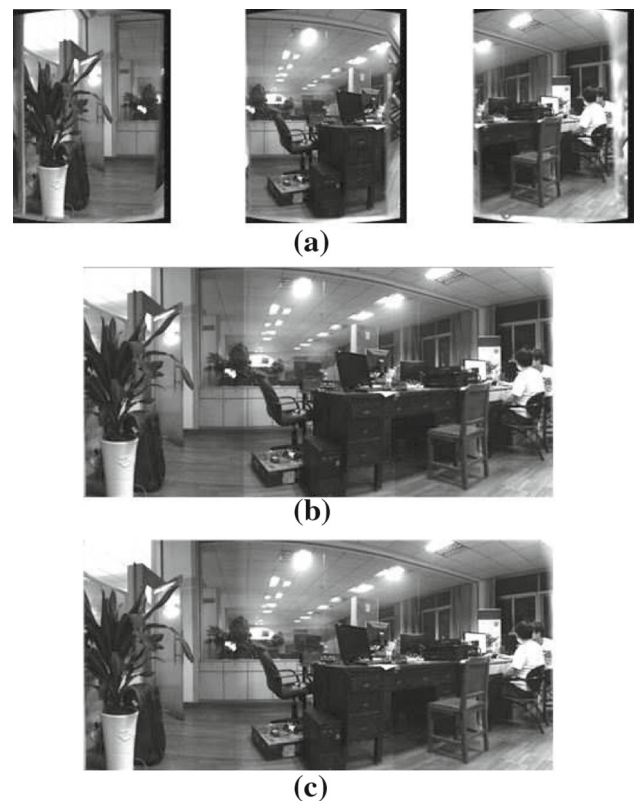

Fig. 10 The flow diagram of image mosaic process. **a** Images after distortion correction and cylindrical projection. **b** Wide FOV image by means of template matching algorithm registration. **c** Wide FOV image after nonlinear weighting image fusion

image correction and cylinder transformation on those pictures, the result of this experimental procedure is as shown in Fig. 10a, we can see there are some certain overlap regions among these three pictures, the overlap area is about 20 pixels. Second, template matching is performed on these three ones, the mosaic result is shown in Fig. 10b, for the inevitable luminance difference between those three cameras, the stitching line of these stitched pictures is also unavoidable. That is why we use nonlinear weighted fusion on these video images, which can get better mosaic result that is shown in Fig. 10c. To verify the feasibility of our method, we take some another image mosaic result as comparison, which is shown in Fig. 11, it comes from FullView company's catadioptric panoramic camera FC-100. According to the contrast the Figs. 10c and 11, the nonlinear weighted fusion method has a better removal effect on the stitch line existence and can get a more natural visual effects.

The mosaic effect of an excellent video image mosaic system should not be influenced by the scene's depth of field, for the result of image mosaics at short range and distant range are both approached to experiment and test, by that we can verify our system's mosaic effects in different depth of field. First, we also take a synthetic image of our laboratory
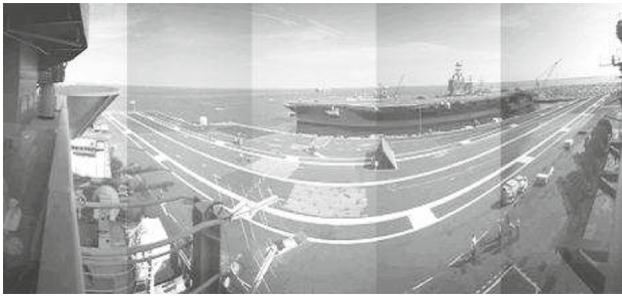
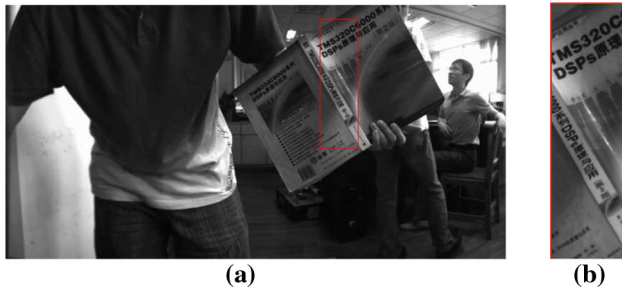**Fig. 11** Mosaic video image of FC-100 from FullView company

(a) **(b)**

Fig. 12 **a** A wide FOV video image at close range. **b** An enlarged image from the area around *stitching line*

Fig. 13 A wide FOV video image at distant range
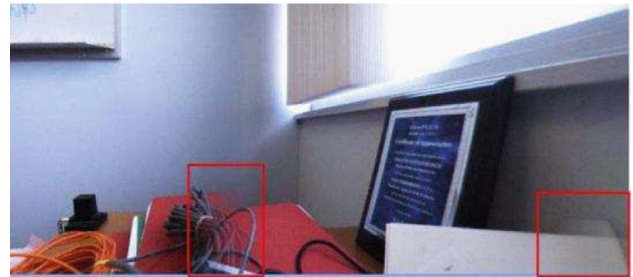

Fig. 14 Mosaic video image at close range from LadyBug2 of Point *Grey* Company, the *red rectangle* in this image indicates the faultiness of obvious residual ghosting and blurring

0.4 meters away from our video image mosaic system, as shown in Fig. 12a, we can see one of our co-authors stand in the front of it, with a book in his hand, the area enclosed by a red box is the location of images stitch line, to give a better view of it, we zoom in it as shown in Fig. 12b. Then, we take a synthetic image of our laboratory about 4 meters away from the system, as shown in the Fig. 13, it is very hard to locate where the stitch line is, that proves the feasibility of our proposed method. We also give another image mosaic result from a short distance as a comparison, it is taken by Point Grey's panorama camera named LadyBug2, and it is a panorama camera using refractive methods with multiple cameras. As shown in the Fig. 14, this product has obvious residual ghosting and blurring problems, and we also enclose it in a red box to make it is easier to perceive it. Compared it again with Fig. 12, our system with catadioptric methods can avoid this problems effectively.

We obtained 54 fps mosaicing speed at 45 % CPU utilization rate(mostly used in network data acceptance), and the generated wide FOV image has validate resolution of 1450×720 and view angle of $108° \times 52°$.

5 Conclusion

To solve the problems of mirror pyramid cameras, such as too high precision required for designing and video images with inevitable stitching line, we proposed a mirror pyramid camera with undemanding precision for designing and

processing. First, we devised a screw nut on our design system to adjust the cameras' height to make their optic centers located at a single point. Then, by the method of nonlinear weighting fusion, we can eliminate the stitching line effectively. The experimental results indicate that our system has the characteristics of good real time, high resolution, without stitching line and not affected by depth of field. According to the theory in this paper, it is easy to design a panoramic camera with ideal mosaicing performance.

However, our catadioptric image mosaicing device has a limited imaging range (in this paper, the range is from 0.4 meter to infinity); if the users want to get the better mosaicing result within 0.4 meter, we need to improve our system to reduce the optical parallax between cameras. On the other hand, when we design our catadioptric image mosaicing system, we do not consider the cameras' difference, which would be an adverse factor in image mosaicing. So in our future work, we will make more precise calibration for the camera selection to solve this problem.

References

1. Nalwa, V. S.: A true omnidirectional viewer. Technical report, Bell Laboratories (1996)
2. Nalwa, V.S.: Panoramic viewing system with offset virtual optical centers. US Patent 6,219,090, 17 Apr (2001)
3. Nalwa, V.S.: Stereo panoramic viewing system. US Patent 6,141,145, 31 Oct (2000)
4. Zhang, S., Yao, H., Liu, S.: Dynamic background subtraction based on local dependency histogram. Int. J. Pattern Recognit. Artif. Intell. **23**(7), 1397–1419 (2009)
5. Zhang, S., Yao, H., Liu, S.: Spatial-temporal nonparametric background subtraction in dynamic scenes. In: Proceedings of IEEE International Conference on Multimedia & Expo, pp. 518–521 (2009)
6. Zhang, S., Yao, H., Liu, S.: Dynamic background modeling and subtraction using spatio-temporal local binary patterns. In: Proceedings of IEEE International Conference on Image Processing, pp. 1556–1559 (2008)
7. Zhang, S., Yao, H., Liu, S., Chen, X., Gao, W.: A covariance-based method for dynamic background subtraction. In: Proceedings of IEEE International Conference on, Pattern Recognition, pp. 3141–3144 (2008)
8. Zhang, S., Yao, H., Sun, X., Lu, X.: Sparse coding based visual tracking: review and experimental comparison. Pattern Recognit. (2013). doi:10.1016/j.patcog.2012.10.006
9. Zhang, S., Yao, H., Zhou, H., Sun, X., Liu, S.: Robust visual tracking based on online learning sparse representation. Neurocomputing **100**, 31–40 (2013)
10. Zhang, S., Yao, H., Sun, X., Liu, S.: Robust visual tracking using an effective appearance model based on sparse coding. ACM Trans. Intell. Syst. Technol. **3**(3), 1–18 (2012)
11. Zhang, S., Yao, H., Liu, S.: Robust visual tracking using feature-based visual attention. In: Proceedings of IEEE International Conference on Acoustics, Speech, and, Signal Processing, pp. 1150–1153 (2010)
12. Gledhill, D., Tian, G.Y., Taylor, D., Clarke, D.: Panoramic imaging—a review. Comput. Graph. **27**, 435–445 (2003)
13. Xiong, Y.L., Turkowski, K.: Creating image-based vr using a self-calibrating fisheye lens. In: Proceedings of IEEE International Conference on Computer Vision and, Pattern Recognition, pp. 237–243 (1997)
14. Miyamoto, K.: Fish eye lens. J. Opt. Soc. Am. **54**, 1060–1061 (1964)
15. Shah, S., Aggarwal, J.K.: A simple calibration procedure for fisheye (high distortion) lens camera. In: Proceedings of IEEE International Conference on Robotics and Automation, pp. 3422–3427 (1994)
16. Aggarwal, M., Ahuja, N.: A new imaging model. In: Proceedings of IEEE International Conference on Computer Vision, vol. 1, pp. 82–89 (2001)
17. Peleg, S., Herman, J.: Panoramic mosaics by manifold projection. In: Proceedings of IEEE International Conference on Computer Vision and, Pattern Recognition, pp. 338–343 (1997)
18. Tang, W.K., Wong, T.T., Heng, P.A.: A system for real-time panorama generation and display in tele-immersive applications. IEEE Trans. Multimed. **7**, 280–292 (2005)
19. Swaminathan, R., Nayar, S.K.: Nonmetric calibration of wide-angle lenses and polycameras. IEEE Trans. Pattern Anal. Mach. Intell. **22**, 1172–1178 (2000)
20. Santos, C.T.P., Santos, C.A.S.: 5cam: a multicamera system for panoramic capture of videos. In: Proceedings of the 12th Brazilian Symposium on Multimedia and the web, pp. 99–107 (2006)
21. Foote, J., Kimber, D.: Flycam: practical panoramic video and automatic camera control. In: Proceedings of IEEE International Conference on Multimedia and Expo, vol. 3, pp. 1419–1422 (2000)
22. Agarwala, A., Agrawala, M., Cohen, M., Salesin, D., Szeliski, R.: Photographing long scenes with multi-viewpoint panoramas. ACM Trans. Graph. **25**(3), 853–861 (2006)
23. Jiyong, Z., Xianyu, S.: Hyperboloidal catadioptric omnidirectional imaging system. Acta Optica Sinica **9**, 1138–1142 (2003)
24. Hicks, R.A., Bajcsy, R.: Catadioptric sensors that approximate wide-angle perspective projections. In: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, Vol. 1, pp. 545–551 (2000)
25. Peleg, S., Ben-Ezra, M., Pritch, Y.: Omnistereo: panoramic stereo imaging. IEEE Trans. Pattern Anal. Mach. Intell. **23**, 279–290 (2001)
26. Baker, S., Nayar, S.K.: A theory of catadioptric image formation. In: Proceedings of IEEE International Conference on Computer Vision, pp. 35–42 (2004)
27. Hong, J., Tan, X., Pinette, B., Weiss, R., Riseman, E.M.: Image-based homing. Control Syst. **12**, 38–45 (1992)
28. Tan, K.H., Hua, H., Ahuja, N.: Multiview panoramic cameras using mirror pyramids. IEEE Trans. Pattern Anal. Mach. Intell. **26**, 941–946 (2004)
29. Hua, H., Ahuja, N., Gao, C.Y.: Design analysis of a high-resolution panoramic camera using conventional imagers and a mirror pyramid. IEEE Trans. Pattern Anal. Mach. Intell. **29**, 356–361 (2007)
30. Gao, C.Y., Hua, H., Ahuja, N.: A hemispherical imaging camera. Comput. Vis. Image Underst. **114**(2), 168–178 (2010)
31. Lewis, J.P.: Fast template matching. In: Proceedings of the Vision Interface, pp. 120–123 (1995)
32. Kughn, C.D., Hines, D.C.: The phase correlation image alignment method. In: Proceedings of the IEEE 1975 International Conference on Cybernettcs and Society (Sept.), pp. 163–165. New York (1975)
33. Yosi, K: Pseudo-polar based estimation of large translations rotations and scalings in images. In: Proceedings of IEEE Workshop on Motion and Video Computing, pp. 201–206 (2005)
34. Zhou, H., Liu, T., Lin, F., Pang, Y., Wu, J., Wu, J.: Towards efficient registration of medical images. Comput. Med. Imaging Graph. **31**(6), 374–382 (2007)
35. Harris, C., Stephens, M.: A combined corner and edge detection. In: Proceedings of the Fourth Alvey Vision Conference, pp. 147–151 (1998)
36. Zhou, H., Wallace, A., Green, P.: Efficient tracking and ego-motion recovery using gait analysis. Signal Process. **89**(12), 2367–2384 (2009)
37. Zhou, H., Li, X., Sadka, A.: Nonrigid structure-from-motion from 2-d images using markov chain Monte Carlo. IEEE Trans. Multimed. **14**(1), 168–177 (2012)
38. Nalwa, V.S.: Outwardly pointing cameras. Technical report, FullView, Inc. (2008)
39. Zhang, Z.Y.: A flexible new technique for camera calibration. IEEE Trans. Pattern Anal. Mach. Intell. **22**, 1330–1334 (2000)
40. An, W.H., Zhou, B.F.: A simple and efficient approach for creating cylindrical panorama. Proc. Fourth Int. Conf. Virtual Real. Appl. Ind. **5444**, 123–129 (2004)