Design and assessment of a 360° panoramic and high-performance capture system with two tiled catadioptric imaging channels

Song, Weitao; Liu, Xuming; Lu, Peng; Huang, Yetao; Weng, Dongdong; Zheng, Yuanjin; Liu, Yue; Wang, Yongtian

2018

Song, W., Liu, X., Lu, P., Huang, Y., Weng, D., Zheng, Y., . . . Wang, Y. (2018). Design and assessment of a 360° panoramic and high-performance capture system with two tiled catadioptric imaging channels. *Applied Optics*, 57(13), 3429-3437.
doi:10.1364/AO.57.003429

<https://hdl.handle.net/10356/84892>

<https://doi.org/10.1364/AO.57.003429>

© 2018 Optical Society of America. All rights reserved. This paper was published in *Applied Optics* and is made available with permission of Optical Society of America.

Downloaded on 26 Apr 2023 18:59:27 SGT

Design and assessment of a 360° panoramic and high-performance capture system with two tiled catadioptric imaging channels

WEITAO SONG,^{1,*} XUMING LIU,⁵ PENG LU,⁴ YETAO HUANG,⁴ DONGDONG WENG,^{2,3} YUANJIN ZHENG,¹ YUE LIU,^{2,3} AND YONGTIAN WANG^{2,3}

¹School of Electrical and Electronic Engineering, Nanyang Technological University, S2.1 B6-02 50 Nanyang Dr, Singapore 639798, Singapore

²AICFVE of Beijing Film Academy, 4, Xitucheng Rd., Beijing 100088, China

³Beijing Engineering Research Center of Mixed Reality and Advanced Display, School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China

⁴Beijing MADV Technology Co., Ltd., JIA-1 Qinghe Yongtai Garden, Beijing 100092, China

⁵Jiangxi Lianchuang Electronic Co., Ltd., 1699 Jingdong Road Nanchang Hi-tech Development Zone, Nanchang, Jiangxi 330000, China

*Corresponding author: bitsongweitao@gmail.com

It has always been a challenge to obtain a 360° panoramic imaging system with a high-resolution and non-visible seam. In this study, a 360° panoramic and high-performance system has been developed based on two tiled imaging channels with a field of view of 190° and $F/\#$ of 2.0. In this study, some issues regarding 360° panoramic tiled methods, such as relative illuminance and ghost images, were analyzed in detail and taken into consideration during the design procedure. In the final design, the stop has been fulfilled for each field angle, and the ghost images have also been well optimized. A proof-of-concept prototype producing a high-performance 360° panoramic video has been developed.

1. INTRODUCTION

Capture systems with an ultrawide angle can provide a view of the entire surrounding environment, and a high-immersion image source, which can give the visual information of the world. Due to recent rapid developments in the industries of virtual reality, surveillance systems, and driverless vehicles, the 360° panoramic imaging system has become essential to further progress [1,2]. Due to the many necessary applications of capture systems, there has been much research on the design of these and similar systems.

The fish-eye lens was the first typical type of imaging system with an ultrawide angle to be developed. To realize the large inclination changes in incident rays of a large field of view (FOV), the front group of the fish-eye always consists of several pieces of the meniscus lens, and thus, the rays can be limited into and pass the stop [3,4]. With the development of the processing and design technique of an aspherical lens, the field of view of the fish-eye lens can achieve more than 180°. The commercial Nikkor 6 mm $F/\#$ 2.8 fish-eye lens has a FOV and diameter ratio of 220° and 10.6, respectively [5]. Martin has also provided a hyperfield fish-eye lens with a $F/\#$ of 5.6

and a FOV of 310° [6]. Some complex structures based on fish-eye lenses, which can achieve a large field of view with high performance, have been summarized in [7]. A panoramic annular lens (PAL), which is always achieved by a catadioptric structure, is another typical type of ultrawide angle lens [8,9]. The first type of PAL system consists of a fish-eye lens and a convex reflective mirror, and the panoramic FOV can be captured despite a blind zone in the central region. Powell described the basic design principle of the PAL in detail, and many effective designs have been developed [10]. By optimizing the surface of the fish-eye lens and the reflective mirror, two or more omnidirectional views from a different viewpoint can be imaging [11]. Thus, many calibration models and computer vision methods have been provided to make the structure commercialized [12]. Semireflective surfaces and stitched aspherical surfaces have been employed to achieve a PAL system without the blind zone in the middle [13,14]. Pernechele presented a hyper-hemispheric lens using semireflective surfaces offering a FOV of 260° [15]; Cheng developed a lens with a FOV 360°(azimuth) × 270°(elevation) by stitching three channels of the optical path. However, by using either the fish-eye lens

or PAL system, it cannot be possible to achieve a view of 360° panoramic imaging at one time [16].

To expand the FOV to the whole space, spatial and tiled methods have been employed. The tiled method is also valid for applications in other imaging systems, such as multiple-projector displays [17], head-mounted displays [18], light field display systems [19], and holographic displays [20]. One of the most popular approaches to produce a 360° panoramic capture system employs six small-size cameras with a relative large FOV on the cubic frame. Smooth transition between adjacent videos is required for the spherical surface based on blending method images [21]. Samsung separated the cameras into two groups and developed 3D 360° panoramic imaging for virtual reality (VR) applications [22]. Lytro also presented a panoramic light field capturing system through the use of multiple cameras [23]. As the FOV of a fish-eye lens can reach above 180°, two capturing channels can be enough to obtain the 360 deg imaging. Some commercial products, such as Theta [24], Insta360 [25], and Samsung Gear360 [26], have been found in markets. The ghost images, the illuminance, and the pattern of the stitched zone have been reasons for poor performance. Moreover, the design strategies and the tiled method of the 360° panoramic capture with two channels have not been analyzed in detail.

In this paper, a 360° panoramic imaging method with two capturing channels has been presented in detail, which includes optical design strategies, relative illumination correction, and tiled methods. Relative illumination, ghost images, and other performance issues were analyzed in detail and taken into consideration during the optical design. In the final design, the imaging performance has almost achieved the diffraction limit, the stop has been fulfilled for each FOV, and the ghost images have also been well optimized. Optical distortion, relative illuminance, and the structure of the system have also been considered in the tiled method. Therefore, the final proof-of-concept prototype produced a high-performance 360° panoramic picture of 7 K resolution.

The rest of this paper is organized as follows. The structure of a 360° panoramic capture system with two capturing channels and the specifications of the overall optical system are shown in Section 2. Then, in Section 3, we present the design, optimization, and analysis results of the optical lens. In Section 4, tiled methods are presented in detail. Section 5 gives the experimental setup and the results of the 360° panoramic system. Our conclusions are provided in Section 6.

2. SYSTEM SPECIFICATION

A. 360° Panoramic System with Two Tiled Imaging Channels

Cameras are tools to simulate the functioning of the human eye. In the image produced by common cameras, 3D scenes within the FOV of the camera are projected on a 2D plane by the perspective projection method. Cameras with a 360° panoramic view are tools to extend the FOV of common eyes and to simulate the process observing the surrounding world. Different from that of standard cameras, a scene of 360° cannot be projected on a single plane. Figure 1 shows the process of the eyes [Fig. 1(a)] and the tiled imaging channels [Fig. 1(b)] obtaining the world. From the top view as shown in Fig. 1(a),

Fig. 1. Schematic diagram of the process of (a) the human eyes and (b) the tiled imaging channels observing the world.

the panoramic image can be shown for different depths from one point (the human eye). A 360° panoramic picture is formed by projecting the 3D world on the sphere from the central point. In the developed system with two tiled imaging channels, the optical centers of the two imaging systems are separated by a distance due to the structure. To tile the images from the two cameras on the sphere surface on a definite distance R , there will be a tiled error on the panoramic image in the tiling zone for the objects apart from the tiled depth. As shown in Fig. 1(b), the object within the depth will be lost, and the objects far from the depth will be repeated in the panoramic image. That is to say, the panoramic view cannot be captured correctly for all depth by multiple tiled imaging channels. The tiled error δ for the particular object distance L can be defined as the incorrect length (imaging missing or imaging repeating) on the circle of this object distance. Based on the basic geometric principle, the tiled error can be expressed as follows, which depends on the distance of object L and the distance between the optical centers of the two cameras ϵ :

$$\delta = \frac{\epsilon}{2\pi R} \frac{|R - L|}{R}. \quad (1)$$

That is to say, the tiled error may be reduced by decreasing the distance between the optical centers of the two cameras.

Many approaches have been proposed to reduce the gap between these optical centers, which can improve the final results. For example, multiple planar mirrors were employed to ensure the optical centers of the reflective cameras overlapped with each other [27]. However, part of the field angle was blocked by the body of cameras, and the number of employed cameras also increases the complexity of the system. It is better to find a structure to eliminate this error. In order to reduce the tiled

Fig. 2. Schematic diagram of different types of structures to arrange the two imaging channels.

error in the tiled imaging channels, we have proposed a catadioptric structure to reduce the gap between the two optical centers. Figure 2 shows the different types of structure to arrange the two imaging channels. We can see a catadioptric structure with a reflective element inside the optical path can make the distance between the two optical centers much smaller than the commercial 360° panoramic cameras. Thus, the tile error will also be small according to Eq. (1). The catadioptric structure has been employed in telescopes and periscopes [28] to reduce the distance of the optical system from one direction, and in our developed system, a right-angle prism with reflective coating was used in either image channel.

B. System Specifications of the Developed 360° Panoramic System

The overall specifications of the developed 360° panoramic system are summarized in Table 1. Each imaging channel is made up of seven glass lenses and one piece of right-angle prism, and the lenses consist of five spherical lenses and two aspherical lenses. Two prisms from the different channels are coated with a reflective film and cemented with each other. Thus, all optical elements are assembled into a crossing type of structure with the prisms in the center. To obtain a 360° panoramic view, the FOV of either channel should be above 180°, which can capture the whole world with two channels. A lower F -number

should be chosen to get more illuminance on the image sensor, and it is better to obtain the equal relative illuminance from all FOVs because all the FOVs should be treated equally in the panoramic view. Therefore, a FOV of 190° and an F -number of 2.0 were set for the single display channel; no vignetting within all FOVs was set, and that is to say, the rays were set to fulfill the stop for all field angles. This significantly improves the relative illuminance compared to other designs but also makes the design more challenging.

3. OPTICAL DESIGN STRATEGIES

The imaging performance and the structure of one channel play decisive roles in a tiled capture system. To obtain a tiled panoramic system with high performance and compact structure, the optical system for one single display channel needs to be well optimized. As mentioned above, designing a lens with a large FOV has been discussed for a long time. Therefore, many related works [29–33] can be analyzed and treated as a reference during the design process. However, due to the specification of the proposed method, some issues are still required to be taken into account.

A. Design Methods and Considerations

To achieve a large FOV, large inclination should be made by the front group in incident rays of large FOV, which can let the rays pass the stop. In our design, a reversed telephoto structure was chosen with a combination of negative and positive lenses.

As shown above, a right-angle prism is employed. Thus, an off-axis system should be designed and optimized, whose design process is more difficult than that of the on-axis system. As the reflective surface of the right-angle prism will not introduce optical power into the system, so the function can be replaced by a cube (a planar glass in the design software), which can simplify the problem to an on-axis lens design (shown in Fig. 3). During the optimization process, four representative wavelengths, 486.1, 546.1, 587.6, and 656.3 nm, were used with equal weights. Due to the on-axis lens and the coaxial rotationally symmetric optics, the optical performance of the optimized system was assessed at the representative field angles for the four design wavelengths. During the optimization process, some strategies should also be controlled to obtain a better structure with high performance.

To reduce high-order aberration, the structure for the positive and negative lenses is relatively complex. Also, two aspherical lenses were employed in the negative and positive groups separately. One is used to refract the rays, and the other to reduce the distortion. As aspherical surfaces were employed, the field angles should be chosen to make sure the beams cover the whole aspherical surface. In actual design process, seven representative field angles were chosen as follows: (0°, 0°), (0°, 15°), (0°, 30°), (0°, 45°), (0°, 60°), (0°, 75°), and (0°, 95°).

To make the function of planar glass available in the actual catadioptric system, the rays on the surface of planar glass should be controlled. From the view of the structure, the thickness of the planar glass should be larger than the active aperture of it. The thickness of the planar glass L_p should be controlled to be longer than the height of the intersection point OA between the planar glass and the meridional ray of the

Table 1. Specifications of the System

	Single Display Channel
Sensor	
Diagonal size	1/2.3 in. (7.77 mm)
Active display area	6.206 mm × 4.679 mm
Resolution	4768 (H) × 3516 (V) pixels
Imaging system	
Effective focal length	1.43 mm
$f/\#$	2.0
Number of optical elements	(7lens + 1prism) × 2
Wavelength	423–660 nm
Field of view (FOV)	190° × 2
MAX image height	4.4 mm
Total optical length	27.0 mm
Length of two vertex	24.0 mm
Performances	
Vignetting	No vignetting in the whole field
$f - \theta$ Distortion	<7.1%

maximum field angle. That is to say, OA should be controlled to be shorter than half of L_p , which has been shown in Fig. 3.

To make the image performance consistent in a large field angle when employing computational correction, the f -theta distortion should also be controlled as little as possible for the whole field. Thus, the image height for the maximum field angle should be set within the short side of the image sensor, and the focal length was also set according to the specifications.

To improve the relative illuminance over the imaging zone, the rays were set to fill the stop for all field angles. Despite the relative illuminance correction which may be introduced to get a uniform seam, to obtain an image with a high illuminance is also required. That is because it can increase the noise-signal ratio when elevating the gray scale during the illuminance correction. During the design process, the structure will be changed, and the rays may not fulfill the stop. Therefore, the entrance pupil was calculated after each cycle of optimization, and rays for each selected field angle were selected to make sure that the marginal area of the entrance pupil was analyzed during the design process.

The two cameras are assembled individually, and these two sets of optical paths are separated by a mechanical part. Thus, the stray light between the image channels can be avoided. Stray light existing within the developed system is formed within one camera, and it can be divided into ghost images and the one eliminating the stray/useful light ratio similar to the white noise in electronic systems. In the developed system, ghost images may introduce unexpected point in the panoramic image, so that is the main stray light required to be eliminated. Ghost paths are introduced mainly by double reflections on optical surfaces. The common method used to reduce the ghost images on the imaging zone is to improve the coating on the optical surface to increase the transmission, but this method is expensive, and the results are not consistent. In our design, the stray light paths have been analyzed and avoided by controlling the lens shape during the design process. For example, a structure with the front group of two meniscus lenses is likely to introduce the ghost images between the surfaces between these two pieces of lens. Therefore, the front group was set combining a meniscus lens and a biconcave lens. In the final design, no obvious ghost image was generated in our design during the analysis process.

Fig. 3. Layout of the optics for one imaging channel.

B. Optical Performance

Based on the design methods and consideration above, one imaging channel with a 190 deg field of view and a high performance can be obtained. The detail of the developed 360° panoramic capture system can be seen in [34]. The final design for one imaging channel is shown in Fig. 4. As discussed above, the cube in the layout plays the same function as a right-angle prism. It consists of one prism and seven pieces of lenses, of which two are aspherical, and no vignetting exists for all selected field angles. The layout of the developed compact panoramic system with two imaging channels is shown in Fig. 4. Because the pixel size of the image sensor used in the prototype was 1.34 μm , the Nyquist frequency was close to 370 lp/mm. Figure 5(a) shows that the polychromatic modulation transfer function (MTF) is above 0.1 at a spatial frequency of 370 lp/mm, which satisfies the requirement of the system design.

The f -theta distortion is also low, 7.1% for the whole field and the image performance will be slightly distorted in a large field angle when computational correction is employed to generate the panoramic image by two imaging channels. The relationship between the field angle and the image height is shown in Fig. 5(b) in the blue line and this design parameter can act as initial parameters during the calibration process, which will be discussed in Section 4.B. In contrast, $f - \tan \theta$ and $f - \theta$ relationships are also shown in Fig. 5(b) along with the actual design result.

C. Tolerance Analysis

Tolerance analysis is directly related to the optical fabrication, alignment, and cost. Therefore, it plays an important role in optical design. The tolerance shown in Table 2 shows the commercial tolerance requirements, which have been discussed with a processing plant. It can act as a reference in the analysis, the yield rate, and the final cost. The merit function is constructed by the MTF of the imaging system; the analyzed frequency is 185 lp/mm, which is 50% of the Nyquist frequency of the image sensor. Figure 6 shows the results of tolerance

Fig. 4. Layout of the optics for a capture system with two tiled catadioptric imaging channels.

Fig. 5. Performance of the designed system. (a) Polychromatic MTF plot of the imaging system. (b) Relationship between the field angle and the image height.

Table 2. Tolerance Values of the Panoramic Optical System

Tolerance Item	Value
Surface radius(fringes)	3
Surface irreg. (fringes)	1
Surface decenter (mm)	± 0.01
Element decenter(mm)	± 0.01
Thickness (mm)	± 0.02
Tilt (')	± 3
Glass material (nd, vd)	0.0005, 0.3%

analysis for the selected FOV. The tolerance results show that, after processing and alignment, most of the final systems will maintain MTF values above 0.1. That is to say, the system has relatively low tolerance sensitivity according to the analysis results, and proper fabrication and alignment are feasible.

D. Prototype and Experimental Results

Figure 7 shows the module of the employed sensor, optical systems, and final imaging part. To test the display performance, an experimental setup was developed with an image sensor and the developed optical system. Image data, which was received by the image sensor, were transferred to the circuit board, and a field-programmable gate array (FPGA) acted as the main chip

Fig. 6. Tolerance analysis results of the panoramic optical system.

Fig. 7. Modules of the imaging system.

Fig. 8. Imaging performance of the developed imaging system. (a) Full image captured by one imaging channel. (b), (c) Enlarged images of the full image.

to make the digital process on the board. In terms of the buildup, five charts with black-and-white stripes were placed at the same distance of 1100 mm, and the charts were set at the center and a field angle of 75 deg, separately. The high-performance image transferred from the sensor is shown in Fig. 8, and the enlarged images are shown in Figs. 8(b) and 8(c).

A ghost image test was also set up with a spotlight at different field angles including the center and relatively large FOV, and no obvious ghost image can be seen within the whole image (shown in Fig. 9).

Fig. 9. Ghost image tests of the developed system.

4. DIGITAL TILED METHOD

A. Relative Illuminance Correction

As mentioned above, the relative illuminance for the large FOV is less than that for the center. That may not be a problem in common cameras as it is hard for users to notice due to the gradual variations, but in the tiled image, the seam will be clear due to the inconsistency from different channels. Increasing the relative illuminance by fulfilling the stop during the design process is an effective approach, but relative illuminance correction before tiling is also essential. Figure 10(a) shows the experimental setup for relative illuminance correction. To correct relative illuminance, a uniform light source should be provided. In terms of the common cameras, objects can be projected on a 2D plane. Thus, a light box with a uniform backlight can be employed. But in a panoramic imaging system, a uniform sphere should be used, so an integrating sphere was employed in our developed system. The developed camera was placed close to the center of the integrating sphere, and raw images could be obtained. The Lambertian scattered material

Fig. 10. (a) Experimental setup for relative illuminance correction. (b) Gradation curve between the gray scale for three channels and the field angle. (c) Raw image obtained inside the integrating sphere.

was coated on the surface of the sphere, and rays from the light source were reflected multiple times inside it. The more times the rays reflect, the more uniform illumination can be obtained. Therefore, the half surface of the sphere with the entrance of the light source cannot be acted as a standard surface with uniform illuminance, and only one channel can be corrected at one time.

To obtain a panoramic image, two imaging channels will capture the outside simultaneously, and two images with a FOV of 190 deg for different directions can be obtained at one time. These two images were stitched into one image to simplify the calculation during the correcting and calibration processes, which were completed by the FPGA inside the camera. Two spherical images are shown in Fig. 10(c) as a raw image in the correction process. Figure 10(b) shows a gradation curve between the gray scales for three color channels and the field angle (radius in the right part of the raw image), which is calculated according to the raw image. As it is shown, the relative illumination at large field angles is higher than for the common fish-eye lens due to filling the stop. Moreover, the gray scales of red, green, and blue channels are also different from each other due to the coating film on the lens and the response curve of the image sensor for different wavelengths. In the final panoramic image, the gray scale of three color channels for each image within the imaging zone is elevated by calculated times according to the coordinated value in the obtained raw image.

B. Computational Model of the System and Its Calibration

The lens and prisms are manufactured and assembled to form two separated catadioptric imaging systems with a FOV of above 180 deg from different directions. However, in the actual system, difficulties arise during lens and sensor alignment, due to the precision limit of the mechanical system. In order to obtain a 360° panoramic capture system, the position relationship and the intrinsic parameters of the cameras should be calibrated.

Multiple basic calibration models for cameras of large FOV have already been discussed in [35]. As for the developed system, a computational model is needed to establish a mapping between the points in the world space and the pixels in the image, which has been discussed as follows. In terms of the intrinsic parameters, the image distortion coefficients for the fish-eye must be taken into account, and in this computational model, only radial distortions (the relationship between the image height and the FOV) are considerable to simplify the calibration process. As two fish-eye imaging systems are employed to generate a 360° panoramic image, two sets of intrinsic and extrinsic parameters are required in the model, which can be marked as i . The relationship between the point coordinates $[x_c, y_c, z_c]^T$ in the camera coordinate system and the pixel coordinates $[x_p, y_p]^T$ in the captured image can be expressed as follows:

$$x_{pi} = \frac{d_i(z_{ci}/r_{ci})}{p_{xi}} \frac{x_{ci}}{\sqrt{x_{ci}^2 + y_{ci}^2}} + x_{0i} \quad (i = 1, 2), \quad (2)$$

$$y_{pi} = \frac{d_i(z_{ci}/r_{ci})}{p_{yi}} \frac{y_{ci}}{\sqrt{x_{ci}^2 + y_{ci}^2}} + y_{0i} \quad (i = 1, 2), \quad (3)$$

$$r_{ci} = \sqrt{x_{ci}^2 + y_{ci}^2 + z_{ci}^2} \quad (i = 1, 2). \quad (4)$$

In the computational model, $d(\cdot)$ is the function of calculating the image height given the field angle z_{ci}/r_{ci} , and (x_o, y_o) are the pixel coordinates of the optical axis on the image, which corresponds to the image height of zero:

$$\frac{d_i(\theta)}{p_i} = f_1\theta + f_2\theta^2 + \dots = \gamma_i D(\theta) \quad (i = 1, 2). \quad (5)$$

In terms of the extrinsic parameters, a rotation matrix and a transfer vector can express the relationship between the pixel coordinates $[X_W, Y_W, Z_W]^T$ in the world coordinate system and those $[x_c, y_c, z_c]^T$ in the camera coordinate system:

$$\begin{pmatrix} x_{ci} \\ y_{ci} \\ z_{ci} \end{pmatrix} = \begin{pmatrix} R_i & T_i \\ 0 & 0 \end{pmatrix} \begin{pmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{pmatrix} \quad (i = 1, 2). \quad (6)$$

The 3×3 matrices R_i can be determined by three individual coefficients, θ_{i1} , θ_{i2} , and θ_{i3} (for example, Rodrigues' rotation formula). The translation vector T_i is also described by three coefficients t_{i1} , t_{i2} , and t_{i3} . Therefore, in the computational model of the developed system, there are in total nine coefficients (including three intrinsic parameters and six extrinsic parameters) for one camera of large FOV.

Many approaches have been developed to obtain a mapping between the pixel coordinates on the image and the points in the world coordinate system (camera-world map). Figure 11 shows the experimental setup for the calibration process,

(a)

(b)

Fig. 11. (a) Experimental setup for the calibration process. (b) Obtained raw image during the calibration process.

and in the raw image [shown in Fig. 11(b)], the corner points can be seen on both sides of the captured image. Based on the physical position of the corner points and corresponding pixels on the captured images, the camera-world map can be obtained. The key to the calibration procedure is to solve a set of non-linear equations of this computational model [Eqs. (6) and (7)] based on the camera-world map. To solve the equations, given a set of initial values, a non-linear iterative method should be employed to obtain a final solution.

The resolution for either image channel as a whole is $U_{xi} \times U_{yi}$, and the initial intrinsic parameters (x_o, y_o) are given as $x_{oi} = U_{xi}/2$ and $y_{oi} = U_{yi}/2$, which is the center of the sensor. In terms of the distortion parameters, the initial parameters can be obtained using the relationship between the field angle and the image height in the design result, which is shown in Fig. 5(b). Thus, $D(\theta)$ can be fitted to a polynomial with coefficients $f_1, f_2, f_3 \dots$ with this relationship, which is similar to the fish-eye computational model in [36]. The initial value of γ_i is set as one, which will be optimized during the non-linear iterative process. Therefore, the initial values of the intrinsic parameters are given as follows:

$$\begin{cases} x_{o1} = U_{x1}/2, & y_{o1} = U_{y1}/2 \\ x_{o2} = U_{x2}/2, & y_{o2} = U_{y2}/2. \\ \gamma_1 = \gamma_2 = 1 \end{cases} \quad (7)$$

Based on the intrinsic parameters given above, the initial extrinsic parameters can be obtained by solving the linear equations if finding a sufficient number of corresponding points. Iterative methods, such as the Levenberg-Marquardt algorithm, can be employed to refine all parameters based on the obtained initial values, and thus, the two sets of nine initial coefficients can be found, which includes the intrinsic parameters of the main camera, the rotation matrix and the translation vector for the two cameras.

In summary, the presented calibration procedure above can be described as follows:

- Obtain a camera-world map. Many approaches can achieve this function, such as finding the corner points, finding the center of a circular point, or decoding the structured light images. In the actual process, corner points are used to set up the camera-world map (shown in Fig. 11).
- Given the initial parameters of the two cameras, calculate the extrinsic parameters of two cameras based on Eq. (7), using the least-squares method.
- Refine all of the parameters using Eqs. (2)–(6), according to the Levenberg-Marquardt algorithm [37].
- Obtain a correspondence map between the pixel coordinate on the two imaging channels and the points in the world coordinate. Thus, the position relationship and the intrinsic parameters of the cameras are obtained.

C. Tiled Methods

Let a Cartesian coordinate system be defined as the middle point between the origins of the two camera coordinates as its origin. The XYZ axis is set to be parallel with the XYZ axis of the first camera coordinate. This newly defined coordinate system is treated as a panoramic one, which can be expressed by polar coordinates. Supposing the tiled surface is set at the distance of R , the relationship between the point

coordinates $[X_p, Y_p, Z_p]^T$ in the panoramic system and the polar coordinates $[R, \eta, \varphi]^T$ on the tiled sphere can be given as Eq. (8):

$$\begin{pmatrix} X_p \\ Y_p \\ Z_p \end{pmatrix} = \begin{pmatrix} R \sin \eta \cos \varphi \\ R \cos \eta \\ R \sin \eta \sin \varphi \end{pmatrix}. \quad (8)$$

Based on the calibration procedure, the coefficients of the two cameras can be solved. Thus, the position relationship between these two camera coordinate systems can be expressed as follows:

$$\begin{pmatrix} x_{c2} \\ y_{c2} \\ z_{c2} \end{pmatrix} = R_2 R_1^{-1} \begin{pmatrix} x_{c1} \\ y_{c1} \\ z_{c1} \end{pmatrix} - R_2 R_1^{-1} T_1 + T_2. \quad (9)$$

Based on the definition, the panoramic coordinate system can be treated as the one moved a half of the transfer vector in equation from the first camera coordinate system. Thus, the relationship between the polar coordinates $[\eta, \varphi]^T$ and the point coordinates $[x_{c1}, y_{c1}, z_{c1}]^T$ in the first camera coordinate system can be expressed as Eq. (8), and $[x_{c2}, y_{c2}, z_{c2}]^T$ in the second camera coordinate system can be given based on Eq. (10):

$$\begin{pmatrix} x_{c2} \\ y_{c2} \\ z_{c2} \end{pmatrix} = \begin{pmatrix} R \sin \eta \cos \varphi \\ R \cos \eta \\ R \sin \eta \sin \varphi \end{pmatrix} - \frac{1}{2} T_2 + \frac{1}{2} R_2 R_1^{-1} T_1. \quad (10)$$

The intrinsic parameters show the relationship between the pixel coordinates $[x_p, y_p]^T$ and the point coordinates $[x_c, y_c, z_c]^T$ in the two camera systems, and the methods to obtain these parameters have also been presented above along with the extrinsic parameters:

$$x_{pi} = \gamma_i D(z_{ci}/r_{ci}) \frac{x_{ci}}{\sqrt{x_{ci}^2 + y_{ci}^2}} + x_{0i} \quad (i = 1, 2), \quad (11)$$

$$y_{pi} = \gamma_i D(z_{ci}/r_{ci}) \frac{y_{ci}}{\sqrt{x_{ci}^2 + y_{ci}^2}} + y_{0i} \quad (i = 1, 2). \quad (12)$$

Based on these equations, we can get a correspondence map between the pixel coordinates on the image and the points in the panoramic coordinate system. As for the polar coordinates $[\eta, \varphi]^T$ in the panoramic system, the corresponding pixel coordinates $[x_p, y_p]^T$ in the two camera systems can be expressed as $[x_{p1}(R, \eta, \varphi), y_{p1}(R, \eta, \varphi)]^T$ and $[x_{p2}(R, \eta, \varphi), y_{p2}(R, \eta, \varphi)]^T$, respectively. In the developed system, the gray scale on the panoramic view of (η, φ) can be the interpolation value of the adjacent four points on the capture images. After the interpolation, the gray scale $[x_p, y_p]^T$ on the image of the first and the second cameras are $S_1(x_p, y_p)$ and $S_2(x_p, y_p)$, respectively. The calculation of the gray scale can be applied to the RGB channels. Also, the area around the seam can be a merged value of the two values from the two images, which can decrease the tiled error in the panoramic picture. Thus, the gray scale for (R, η, φ) in the panoramic system can be expressed as follows [the unit of angle in Eq. (13) is the radian]:

$$I(R, \eta, \varphi) = \begin{cases} S_2(x_{p2}, y_{p2}), & \varphi \in [-\pi/2, -\Delta] \\ \alpha S_2(x_{p2}, y_{p2}) + (1 - \alpha) S_1(x_{p1}, y_{p1}), & \varphi \in [-\Delta, \Delta] \\ S_1(x_{p1}, y_{p1}), & \varphi \in [\Delta, \pi - \Delta] \\ \alpha S_2(x_{p2}, y_{p2}) + (1 - \alpha) S_1(x_{p1}, y_{p1}), & \varphi \in [\pi - \Delta, \pi + \Delta] \\ S_2(x_{p2}, y_{p2}), & \varphi \in [\pi + \Delta, 3\pi/2] \end{cases}. \quad (13)$$

Here, $x_{pi}(i = 1, 2)$ stands for $x_{pi}(R, \eta, \varphi)$, and $y_{pi}(i = 1, 2)$ stands for $y_{pi}(R, \eta, \varphi)$. 2Δ stands for the merged zone in the panoramic picture, which cannot exceed the imaging zone of the cameras. α is the average factor in the merging algorithm, which can be 1/2 or a variable function subject to Δ .

Based on the panoramic picture on the sphere, an image of a definite solid angle can be picked up, and thus, an image of a normal view can be shown. Similar to the map projection, locations from the surface of a sphere can be projected to form the panoramic view on a 2D plane. Thus, the panoramic image or video of all the information can be shown.

D. 360° Panoramic Image Based on the Developed Prototype

As described above, the panoramic image and video can be shown by projecting the spherical image on the 2D plane. Here, 360° panoramic video is shown by setting the η, φ as the horizontal and vertical axes, respectively. Figure 12 shows a stitched photograph taken by the prototype after the calibration of the two imaging channels. The horizontal axis for the panoramic image covers from $-\pi/2$ to $3\pi/2$ [φ in Eq. (13)], which is divided equally, and the two seams will be at the position of 0 and π . On the other hand, the vertical axis of the final image covers 0 to π [η in Eq. (13)], which is similar to the longitude of the Earth. Figures 12(b) and 12(c) show the enlarged images of the panoramic image [Fig. 12(a)], which can show the high imaging performance of the developed system. Shown in the picture, the seam is hard to see in the 360° panoramic image, and the brightness is also corrected. The 360° panoramic video is also shown in Visualization 1.

Fig. 12. Stitched image produced by two imaging channels through the developed prototype after calibration (Visualization 1). (a) Full panoramic image captured by the developed system. (b), (c) Enlarged images of the panoramic image.

5. CONCLUSIONS

In this paper, a tiled panoramic capture system that uses two compact catadioptric imaging channels was developed. Each imaging channel achieved a FOV of 190°, an F -number of 2.0, with low distortion over the entire field. The rays from all field angles with 190° fulfilled the stop in the system, and the image performance was designed to close to the diffraction limit. Design methods for the imaging channel and the tiled method were discussed in detail. The performance of the capture system was shown with high imaging quality and no visible ghost images, along with a demonstration of a proof-of-concept experiment.

Funding. National Key Research and Development Program of China (2016YFB1001502); National Research Foundation Singapore (NRF) (NRF-CRP11-2012-01); National Natural Science Foundation of China (NSFC) (61727808).

REFERENCES

1. S. Thibault and J. C. Artonne, "Panomorph lenses: a low cost solution for panoramic surveillance," *Proc. SPIE* **6203**, 62030S (2006).
2. K. H. Heng, W. D. Zhong, T. H. Cheng, N. Liu, and Y. He, "Beam divergence changing mechanism for shortrange inter-unmanned aerial vehicle optical communications," *Appl. Opt.* **48**, 1565–1572 (2009).
3. W. J. Smith, *Modern Lens Design* (McGraw-Hill, 2005), Vol. 2.
4. J. M. Geary, *Introduction to Lens Design: With Practical ZEMAX Examples* (Willmann-Bell, 2002).
5. Y. Shimizu, "Wide angle fisheye lens," U.S. patent 3,737,214 (5 June, 1973).
6. C. B. Martin, "Design issues of a hyper-field fisheye lens," *Proc. SPIE* **5524**, 84–92 (2004).
7. S. Thibault, "Panoramic lens applications revisited," *Proc. SPIE* **7000**, 70000L (2008).
8. Z. Huang, J. Bai, and X. Y. Hou, "Design of panoramic stereo imaging with single optical system," *Opt. Express* **20**, 6085–6096 (2012).
9. D. Hui, M. Zhang, Z. Geng, Y. Zhang, J. Duan, A. Shi, L. Hui, Q. Fang, and Y. Liu, "Designs for high performance PAL-based imaging systems," *Appl. Opt.* **51**, 5310–5317 (2012).
10. I. Powell, "Design study of an infrared panoramic optical system," *Appl. Opt.* **35**, 6190–6194 (1996).
11. S. Yi and N. Ahuja, "An omnidirectional stereo vision system using a single camera," in *International Conference on Pattern Recognition*, Y. Y. Tang, S. P. Wang, G. Lorette, D. S. Yeung, and H. Yan, eds. (IEEE Computer Society, 2006), pp. 861–865.
12. W. Li and Y. F. Li, "Single-camera panoramic stereo imaging system with a fisheye lens and a convex mirror," *Opt. Express* **19**, 5855–5867 (2011).
13. W. Stürzl, D. Soccol, J. Zeil, N. Boeddeker, and M. V. Srinivasan, "Rugged, obstruction-free, mirror-lens combination for panoramic imaging," *Appl. Opt.* **47**, 6070–6078 (2008).
14. G. I. Kweon, K. T. Kim, G. H. Kim, and H. S. Kim, "Folded catadioptric panoramic lens with an equidistance projection scheme," *Appl. Opt.* **44**, 2759–2767 (2005).
15. C. Pernechele, "Hyper-hemispheric and bifocal panoramic lenses," *Proc. SPIE* **8896**, 889603 (2013).
16. D. Cheng, C. Gong, C. Xu, and Y. Wang, "Design of an ultrawide angle catadioptric lens with an annularly stitched aspherical surface," *Opt. Express* **24**, 2664–2677 (2016).
17. B. Sajadi and A. Majumder, "Autocalibration of multi-projector CAVE-like immersive environments," *IEEE Trans. Vis. Comput. Graphics* **18**, 381–393 (2012).
18. W. Song, D. Cheng, Z. Deng, Y. Liu, and Y. Wang, "Design and assessment of a wide FOV and high-resolution optical tiled head-mounted display," *Appl. Opt.* **54**, E15–E22 (2015).
19. W. Li, H. Wang, M. Zhou, S. Wang, S. Jiao, X. Mei, and J. Kim, "Principal observation ray calibration for tiled-lens-array integral imaging display," in *Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2013), pp. 1019–1026.
20. G. Xue, J. Liu, X. Li, J. Jia, Z. Zhang, B. Hu, and Y. Wang, "Multiplexing encoding method for full-color dynamic 3D holographic display," *Opt. Express* **22**, 18473–18482 (2014).
21. J. Lee, B. Kim, K. Kim, Y. Kim, and J. Noh, "Rich360: optimized spherical representation from structured panoramic camera arrays," *ACM Trans. Graphics* **35**, 63 (2016).
22. <http://thinktankteam.info/beyond/>.
23. <https://www.lytro.com/immerge>.
24. <https://theta360.com/>.
25. <https://www.insta360.com/>.
26. <http://www.samsung.com/global/galaxy/gear-360/>.
27. K. Tan, H. Hua, and N. Ahuja, "Multiview panoramic cameras using mirror pyramids," *IEEE Trans. Pattern Anal. Mach. Intell.* **26**, 941–946 (2004).
28. P. R. Yoder, Jr., "Mounting multiple lenses," in *Opto-Mechanical Systems Design* (CRC Press, 2005), Chap. 5, pp. 229–235.
29. M. Laikin, "Wide angle lens systems," *Proc. SPIE* **237**, 530–533 (1980).
30. G. I. Kweon, Y. H. Choi, and M. Laikin, "Fisheye lens for image processing applications," *J. Opt. Soc. Korea* **12**, 79–87 (2008).
31. W. S. Sun, C. L. Tien, Y. H. Chen, and P. Y. Chu, "Ultra-wide angle lens design with relative illumination analysis," *J. Eur. Opt. Soc.* **11**, 16001 (2016).
32. A. Ning, "Compact fisheye objective lens," U.S. patent 7,023,628 B1 (25 September, 2008).
33. M. Kawada, "Fisheye lens unit," U.S. patent 7,283,312 B2 (25 April, 2006).
34. W. Song, Y. Huang, X. Liu, and J. Zeng, "Camera and imaging apparatus," CN patent 201610290697.X (4 May 2016).
35. A. M. G. Tommaselli, J. Marcato, Jr., M. V. A. Moraes, S. L. A. Silva, and A. O. Artero, "Calibration of panoramic cameras with coded targets and a 3D calibration field," in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences EuroCOW* (ISPRS, 2014), Vol. **XL-3/W1**, pp. 137–142.
36. J. Bouguet, "Camera calibration toolbox for Matlab," <http://www.vision.caltech.edu/bouguetj>.
37. R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision* (Cambridge University, 2003), pp. 597–628.