Research Article

A Master-Slave Calibration Algorithm with Fish-Eye Correction

J. C. Neves, J. C. Moreno, and H. Proença

Instituto de Telecomunicações (IT), Department of Computer Science, University of Beira Interior, 6201-001 Covilhã, Portugal

Correspondence should be addressed to J. C. Moreno; jcmb@ubi.pt.

Surveillance systems capable of autonomously monitoring vast areas are an emerging trend, particularly when wide-angle cameras are combined with pan-tilt-zoom (PTZ) cameras in a master-slave configuration. The use of fish-eye lenses allows the master camera to maximize the coverage area while the PTZ acts as a foveal sensor, providing high-resolution images of regions of interest. Despite the advantages of this architecture, the mapping between image coordinates and pan-tilt values is the major bottleneck in such systems, since it depends on depth information and fish-eye effect correction. In this paper, we address these problems by exploiting geometric cues to perform height estimation. This information is used both for inferring 3D information from a single static camera deployed on an arbitrary position and for determining lens parameters to remove fish-eye distortion. When compared with the previous approaches, our method has the following advantages: (1) fish-eye distortion is corrected without relying on calibration patterns; (2) 3D information is inferred from a single static camera disposed on an arbitrary location of the scene.

1. Introduction

The coexistence of humans and video surveillance cameras in outdoor environments is becoming commonplace in modern societies. This new paradigm has raised the interest in automated surveillance systems capable of inferring useful information from the scene (e.g., person identification, action recognition, and abnormal event detection). However, these systems are designed for monitoring vast areas, which highly decreases the resolution of regions of interest.

To address this issue, several approaches have exploited pan-tilt-zoom (PTZ) cameras, since the mechanical properties of these devices allow zooming in on arbitrary scene locations. Most PTZ-based methods adopt a master-slave configuration, where a static camera monitors a large surveillance area to instruct the PTZ camera to zoom in on regions of interest. While several advantages can be outlined, intercamera calibration is the major bottleneck of this configuration, since an accurate mapping from image coordinates to pan-tilt space requires depth information and distortion correction, as illustrated in Figure 1. The existing approaches [1–4] rely on rough approximations or on the use of multiple static devices to perform triangulation. Also, they assume the pinhole model for the static camera. Such assumption is highly

restrictive, since in surveillance scenarios fish-eye lenses are commonly used to increase the coverage area and the distortion introduced by these lenses is nonnegligible.

In this paper, we propose a master-slave calibration algorithm capable of both removing the fish-eye distortion and inferring an accurate mapping between the static and the active camera without requiring calibration patterns. Our approach exploits geometric cues—which are typically available in urban environments—to measure objects in the scene. As in [5], the vanishing line of a reference plane in the scene and one vertical vanishing point are used to infer the height of static objects or subjects walking throughout a surveillance scenario. This information has a twofold goal: (1) to determine the properties of a fish-eye lens and (2) to determine the 3D position of a subject. In the former, the height of an object is exploited to determine the angle of view and the projection type of the lens to rectify the image coordinates according to the pin-hole projective transform. In the later, subjects height is imposed to the projective transform to determine its 3D location, enabling the correct estimation of pan-tilt values.

When compared with the previous approaches, our method has the following advantages: (1) fish-eye distortion is corrected without relying on calibration patterns; (2) 3D
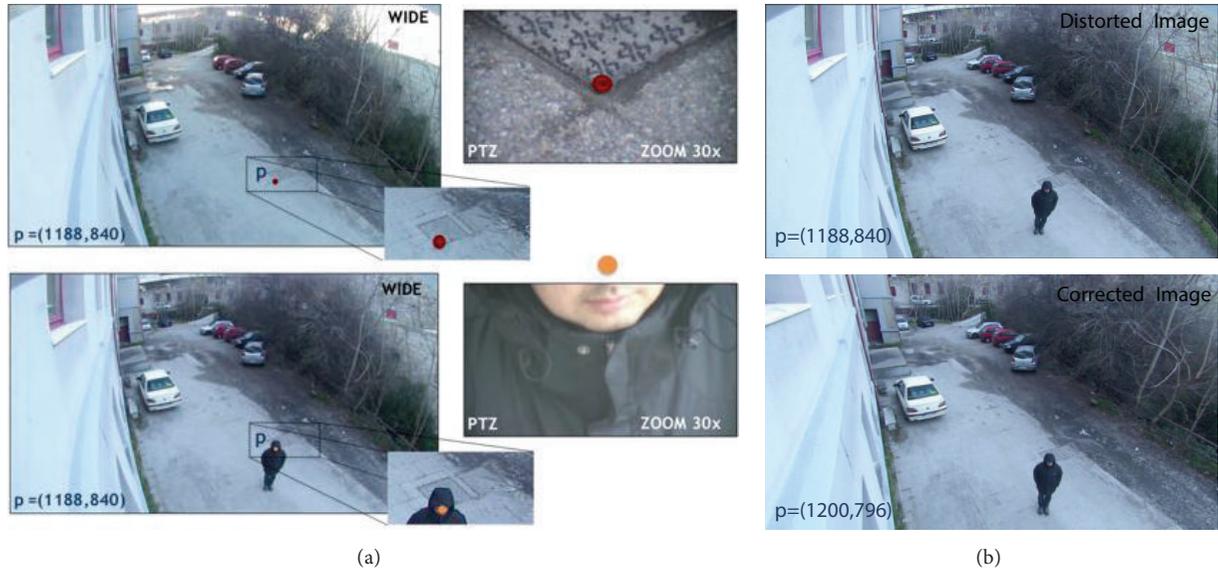
FIGURE 1: The major problems in the calibration of master-slave systems. (a) Inaccurate estimation of the pan and tilt angles when depth information is not considered. (b) Nonnegligible distortion introduced by the use of fish-eye lenses.

information is inferred from a single static camera; (3) cameras can be disposed on an arbitrary location of the scene.

The remainder of this paper is organized as follows. Section 2 summarizes the most relevant master-slave approaches as well as the existing fish-eye correction strategies. Section 3 describes the proposed method. The experimental evaluation of the proposed algorithm is presented and discussed in Section 4. Finally, Sections 5 and 6 outline the major conclusions of this work and its future direction.

2. Related Work

Most fish-eye correction approaches focus on defining a mapping from the viewing sphere to the view plane using polynomial functions or fish-eye projection models [6–8]. Straight line preservation is a strategy commonly used to infer the correction models, for which two distinct approaches have been proposed: (1) the use of planar calibration patterns [9–14] and (2) automatic extraction of geometric constraints from the scene. In the former, a set of calibration points, arranged in straight lines, are used to minimize the lines curvature or the reprojection error when full calibration is considered. The later uses a set of automatically detected key points to impose epipolar geometry constraints in multiple views of the scene [15–18]. Another strategy is to use a semiassisted straight line detection [19].

Regarding the integration of fish-eye correction in master-slave systems, [20] is the only work that proposed a full integrated system. However, this approach does not take into account depth information which turns the mapping between both devices as an ill-posed problem. To alleviate the mapping inaccuracies, the cameras are assumed to be side-by-side.

To address the lack of depth information in master-slave systems, a large number of approximations have been proposed. The use of manually constructed look-up tables [21] or linear interpolations [22, 23] is one alternative to perform the static to pan-tilt mapping. To alleviate the burden of manual mapping, automatic calibration approaches infer an approximate relation between camera images using feature point matching [20].

Some alternative approaches have also been presented in [2, 24, 25]. In [24] multiple consecutive frames were used to approximate target depth. However, this strategy is time-consuming and, consequently, increases the delay between issuing the order and directing the PTZ. You et al. [25] estimated the relationship between the static and the active camera using a homography for each image of the mosaic derived from the slave camera. Del Bimbo et al. [2] relied on feature point matching to automatically estimate a homography (H), relating the master and slave views with respect to the reference plane. H is used to perform an online mapping between the feet locations in the master to the slave camera and also determine the reference plane vanishing line from the one manually marked on the static view. Despite being capable of determining head location, this strategy has to set the active camera in an intermediate zoom level to cope with the uncertainties of vanishing line location. In contrast to the previous approaches, the use of multiple static cameras has also been introduced to solve the lack of depth information in master-slave systems. However, these systems either rely on stereographic reconstruction [26], which is computationally expensive, or dispose the cameras in a specific configuration to ease object triangulation [3, 4], which is not practical for real-world scenarios.

3. Proposed Method

In this section the proposed method is divided into two distinct phases: the fish-eye correction method and
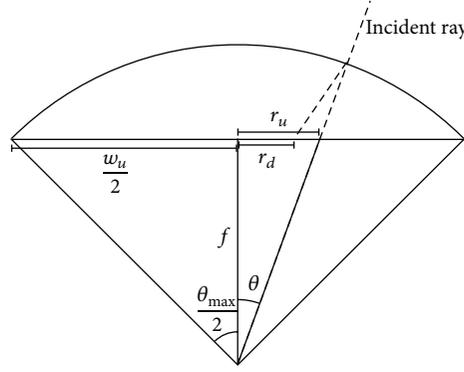
Figure 2: Illustration of the fish-eye projection model. In the pin-hole camera model, a ray of light defining an angle θ with the optical axis is projected at a distance r_u from the principal point. The use of a fish-eye lens forces the ray to be projected at r_u according to the projection type of the lens.

Table 1: Projection models of the fish-eye lenses.

Projection	Model
Equidistance	$r = f\theta$
Orthogonal	$r = 2f \sin(\theta)$
Equisolid angle	$r = 2f \sin\left(\dfrac{\theta}{2}\right)$
Stereographic	$r = 2f \tan\left(\dfrac{\theta}{2}\right)$

the master-slave calibration algorithm. The former is used to rectify the image coordinates to the projective projection, on which our master-slave calibration depends. The later shows how to determine the 3D position of a subject's head in the scene and the correspondent pan and tilt values.

3.1. Fish-Eye Correction. While the pin-hole camera projection can be modelled by the perspective projection $r = f \tan(\theta)$, fish-eye lenses introduce one of the following projections described in Table 1, r being the distance to the principal point, f the focal distance, and θ the angle between the incident ray and the optical axis. Figure 2 illustrates the effect of fish-eye lenses on the projection of an incident ray when compared with the projective projection of the pin-hole model. r_d and r_u represent the radial positions where a ray is projected when a fish-eye lens is used and when it is not, respectively. This model provides evidence that the radial position yielded by a projective projection model can be recovered by establishing a relation between r_u and r_d. Although a more general model exists—the polynomial fish-eye transform (PFET) [27]—they require a larger amount of ground truth data, and for the majority of the lenses, these models are a good approximation of the fish-eye projective models described in Table 1 [28].

Given the pin-hole camera projection model $r_u = f \tan(\theta)$ and a fish-eye projection model $r_d = f\kappa(\theta)$, a relation between r_u and r_d is given by

$$r_u = f \tan\left(\kappa^{-1}\left(\frac{r_d}{f}\right)\right), \tag{1}$$

where κ is one fish-eye projection function.

Considering that f is necessary to define (1), it can be determined by

$$f = \frac{w_u}{\kappa\left(\theta_{\max}/2\right)}, \tag{2}$$

and thus

$$r_u = \frac{w_u}{\kappa\left(\theta_{\max}/2\right)} \tan\left(\kappa^{-1}\left(\frac{r_d \cdot \kappa\left(\theta_{\max}/2\right)}{w_u}\right)\right), \tag{3}$$

where $\theta_{\max}$ is the horizontal angle of view and w_u represents the image width in pixels. While determining w_u being trivial, $\theta_{\max}$ and κ require knowledge about the lens properties, which are often unavailable.

As such, we argue that the height of scene objects can be used to estimate $\theta_{\max}$ and κ. The insight behind this idea is that image-based height estimation methods rely on the pin-hole camera model and thus yields incorrect height measurements in distorted images. Therefore, fish-eye correction is regarded as a minimization problem, where the correct lens parameters are the ones which minimize the height estimation error in the corrected image.

In order to perform height estimation from a single camera, we build on the work of Criminisi et al. [5]. We use three vanishing points $(\mathbf{v}_x, \mathbf{v}_y, \mathbf{v}_z)$ for the x-, y- and z-axis, determined by the intersection of parallel lines (points at infinite) drawn manually in the image scene. $\mathbf{v}_x$ and $\mathbf{v}_y$ are determined from parallel lines contained in the reference plane, so that the line $\mathbf{l}$ defined by these points represents the plane vanishing line. The point $\mathbf{v}_z$ does not belong to reference plane since it is the intersection of two parallel lines perpendicular to the reference plane.

Given $\mathbf{l}$, $\mathbf{v}_z$, the top $(\mathbf{p}_t)$, and bottom $(\mathbf{p}_b)$ points in an image, the height of an object can be obtained by

$$Z = -\frac{\|\mathbf{p}_b \times \mathbf{p}_t\|}{\alpha\left(\mathbf{l} \cdot \mathbf{p}_b\right)\|\mathbf{v} \times \mathbf{p}_t\|}, \tag{4}$$

where $\alpha = -\|\mathbf{p}_{rb} \times \mathbf{p}_{rt}\|/(Z_r(\mathbf{l} \cdot \mathbf{p}_{rb})\|\mathbf{v} \times \mathbf{p}_{rt}\|)$, whereas $\mathbf{p}_{rt}$ and $\mathbf{p}_{rb}$ are the top and base points of a reference object in the image with height equal to Z_r.

Considering that the vanishing points are marked on the original image, (3) is used to correct their locations and estimate the height of an object with respect to the lens parameters, hereinafter denoted by $Z(\theta, \kappa)$. Given the height Z_o of an object in the scene, the angle of view ($\theta_{\max}$) and the projection type (κ) can be estimated by

$$\{\theta_{\max}^*, \kappa^*\} = \arg \min_{\theta, \kappa} \left| Z_o - Z(\theta, \kappa) \right|. \tag{5}$$

3.2. Master-Slave Calibration. First, we introduce the notation used to describe the proposed master-slave calibration algorithm:

  (i) (X, Y, Z): the 3D world coordinates.

 (ii) (X_s, Y_s, Z_s): the 3D coordinates in the static camera referentiality.

(iii) (X_p, Y_p, Z_p): the 3D coordinates in the PTZ camera referentiality.

 (iv) (x_s, y_s): the 2D coordinates in the static camera referentiality.

  (v) (x_p, y_p): the 2D coordinates in the PTZ camera referentiality.

 (vi) (x_h, y_h): the head position of a subject in the static camera image plane.

(vii) (θ_p, θ_t): the pan, tilt parameters of the PTZ camera.

In the pin-hole camera model, the projective transformation of 3D scene points onto the 2D image plane is governed by

$$\lambda \begin{pmatrix} x_h \\ y_h \\ 1 \end{pmatrix} = \underbrace{\mathbf{K}\,[\mathbf{R} \mid \mathbf{T}]}_{:=\mathbf{P}} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}, \tag{6}$$

where λ is a scalar factor and $\mathbf{K}$ and $[\mathbf{R} \mid \mathbf{T}]$ represent the intrinsic and extrinsic camera matrices, which define the projection matrix $\mathbf{P}$.

Let $\mathbf{p}_h = (x_h, y_h)$ denote the head position of a subject in the static camera image plane. Solving (6) for (X, Y, Z) yields an underdetermined system, that is, infinite possible 3D locations for this point. As such, we propose to solve (6) by determining one of the 3D components previously.

By assuming a world coordinate system (WCS) where the XY plane corresponds to the reference ground plane of the scene, the Z component of a subject's head corresponds to its height (Z_x). The use of height information reduces (6) to

$$\lambda \begin{pmatrix} \mathbf{p}_h \\ 1 \end{pmatrix} = \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & Z_x \mathbf{p}_3 + \mathbf{p}_4 \end{bmatrix} \begin{pmatrix} X \\ Y \\ W \end{pmatrix}, \tag{7}$$

where λ is a scalar factor and $\mathbf{p}_n$, $n \in \{1, 2, 3, 4\}$, is the set of column vectors of the projection matrix P (refer to Appendix for the demonstration of (7)). In consequence, our algorithm

works on the static camera to determine $\mathbf{p}_h$ and infer the subject position in the WCS using its height.

Assuming that there is no displacement between the PTZ center of rotation and the optical center, the coordinates of 3D world point (X, Y, Z) in the PTZ referentiality are given by

$$\begin{pmatrix} X_p \\ Y_p \\ Z_p \end{pmatrix} = [R \mid T] \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}. \tag{8}$$

The correspondent pan and tilt angles can be therefore obtained by

$$\theta_p = \arctan\left(\frac{Y_p}{X_p}\right),$$

$$\theta_t = \arccos\left(\frac{Y_p}{\sqrt{\left(X_p\right)^2 + \left(Y_p\right)^2 + \left(Z_p\right)^2}}\right). \tag{9}$$

Considering that both fish-eye correction and master-slave calibration algorithms depend on an accurate height estimation, it is important to note that the ground is assumed to be approximately plane. The validity of our method in approximately plane scenarios has been assessed in Section 4.

4. Experimental Results

In this section, the evaluation of the proposed method was divided into two distinct phases: (1) fish-eye correction and (2) estimation of the image coordinate to pan-tilt mapping.

4.1. Performance Evaluation: Fish-Eye Correction. The proposed fish-eye correction method was tested using a surveillance camera equipped with a fish-eye lens installed in an outdoor parking lot. Three pairs of parallel lines were manually annotated on the distorted image to estimate the location of one vertical and two horizontal vanishing points. Additionally, two reference objects were annotated and measured as depicted in Figure 3(a). These data were used to estimate the height deviation with respect to $\theta_{\max}$ using the different fish-eye projection functions, and the attained results are presented in Figure 3(b). The comparative analysis between the different fish-eye projection types supports the idea that lens parameters can be inferred by minimizing the error of automatic height estimation. According to (3), the pair $(\kappa, \theta_{\max}) = \{\text{orthogonal}, 112.5°\}$ would be chosen as the lens parameters, which constitutes a good approximation to the real angle of view of the lens, $112.6°$.

In order to validate the effectiveness of our approach, a comparison with pattern-based approaches (CB) was conducted by determining the average reprojection error when calibrating the camera using images corrected with the different strategies. For this purpose, a checkerboard was used and 60 marks were disposed in the scene and their
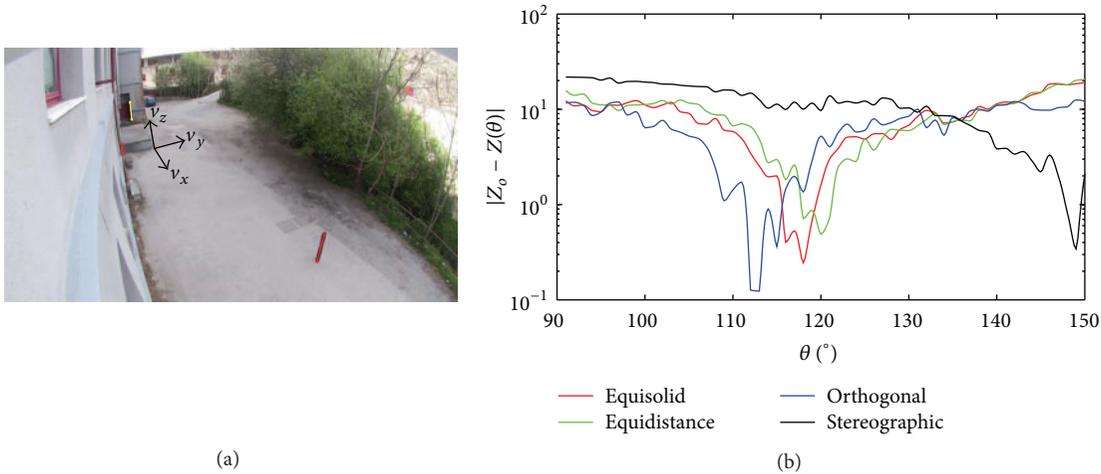
FIGURE 3: Evaluation of the proposed fish-eye correction method. (a) The surveillance scenario used to perform the experimental validation of the proposed approach. The reference object is annotated in yellow while the objects used to perform distortion correction are presented in red. Lines representing the direction of the vanishing points are shown in black color. (b) The height estimation error, in centimeters, when using different fish-eye projections and angles of view to correct the image distortion.

image and world coordinates were manually determined. In both strategies, the intrinsic and extrinsic parameters of the camera were determined with the method described in [29]. The distribution of the reprojection error using both strategies is presented in Figure 4(a), whereas Figures 4(b) and 4(c) illustrate the displacement between the correct positions (in green) and the projected positions (in red) for CB and our method, respectively. The comparative analysis of the reprojection error of both approaches provides evidence that the proposed method provides a good approximation to typical fish-eye removal approaches without requiring the use of a planar calibration pattern.

Additionally, a comparative analysis of the height estimation performance was conducted. This performance was measured with respect to the deviation Δh to the true height of the target. The height of a human being was used to assess Δh in 50 different scene locations, as illustrated in Figure 4(e). As shown in Figure 4(d), the distribution of Δh is highly similar for both approaches and in average an accurate height estimation is attained.

4.2. Performance Evaluation: Intercamera Calibration. To assess the accuracy of the proposed approach, we used the following procedure: given (x_s, y_s) and its corresponding (x_p, y_p) point, the algorithm error $(\Delta\theta)$ was determined by the angular difference between the estimated (X_p, Y_p, Z_p) and the 3D ray associated with (x_p, y_p). When compared with the typical reprojection error, this strategy is advantageous since it allows a direct comparison with the camera angle of view.

To assess the overall performance of our approach, three different persons were recorded—comprising more than 300 frames—byboth the static and the active camera while walking throughout a surveillance scenario. Both PTZ and wide-view images were annotated to mark the pixel location of the head and feet. Using these data, the system was evaluated with

respect to $\Delta\theta$, which was useful to determine if an object of interest will be successfully imaged when using the PTZ at the maximum zoom.

Figure 4(f) illustrates the attained results for the proposed method with respect to the pan and tilt error. The obtained results provide evidence that in the majority of the cases the displacement between the estimated pan-tilt values and the center of the region of interest is less than the field of view of the PTZ camera when using a 30-time zoom magnification, which corresponds to the maximum capability of state-of-the-art PTZ cameras.

5. Conclusions

In this paper, we introduced a master-slave calibration algorithm capable of removing fish-eye distortion and accurately estimating the mapping from the image coordinates to pan-tilt space without depending on calibration patterns. The geometrical cues typically available in urban scenes were exploited to perform height estimation, which can be used to infer the parameters of fish-eye lenses and also the 3D position of subjects in the scene.

An experimental evaluation in a real surveillance scenario provided evidence that fish-eye correction based on height estimation attains highly similar results to typical pattern-based approaches. Regarding the master-slave calibration algorithm, the pan and tilt errors of the method are confined to a tight range of values which in the majority of the cases do not exceed the PTZ field of view.

6. Further Work

In the future, we aim at determining how this approach can be extended to more general fish-eye correction models while maintaining the amount of ground truth data as low as possible. For that purpose, we will investigate how multiple

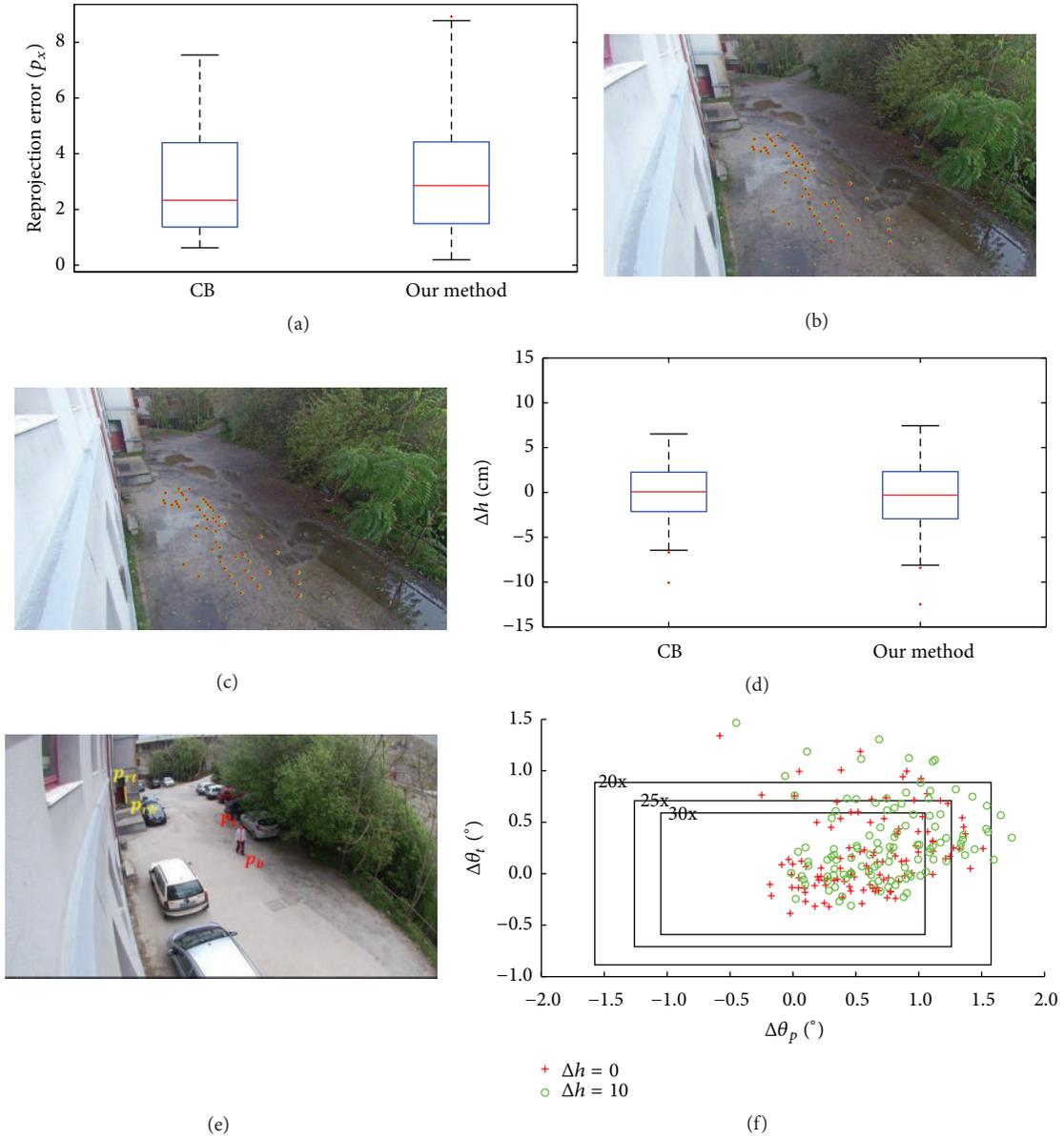(a)

(b)

(c)

(d)

(e)

(f)

FIGURE 4: Comparative analysis between the proposed fish-eye correction approach and fish-eye correction with calibration patterns. (a) The reprojection errors attained using a calibration pattern (CB) to correct fish-eye distortion and using the proposed method. Notice the residual difference with the traditional calibration approach. The reprojected pixel locations are illustrated in (b) and (c) for CB and for our approach, respectively. (d) Comparative analysis of the height estimation error in different locations of the scene. (f) The pan-tilt angle error when observing a human being in different scene locations.

height measurements extracted from a walking human can be informative enough to infer the correct parameters of a PFET.

Appendix

Determining 3D Position from the Inverse Projective Transform

An explanation of the relation between (6) and (7) is given below.

A complete representation of (6) is given by

$$
\lambda \begin{pmatrix} x_h \\ y_h \\ 1 \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}, \quad \text{(A.1)}
$$

from where we get the following equations:

$$x_h = \frac{p_{11}X + p_{12}Y + p_{13}Z + p_{14}}{p_{31}X + p_{32}Y + p_{33}Z + p_{34}},$$

$$y_h = \frac{p_{21}X + p_{22}Y + p_{23}Z + p_{24}}{p_{31}X + p_{32}Y + p_{33}Z + p_{34}}. \tag{A.2}$$

Equation (A.2) can be equivalently written using homogeneous coordinates as

$$\lambda \begin{pmatrix} x_h \\ 1 \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{13}Z + p_{14} \\ p_{31} & p_{32} & p_{33}Z + p_{34} \end{pmatrix} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix},$$

$$\lambda \begin{pmatrix} y_h \\ 1 \end{pmatrix} = \begin{pmatrix} p_{21} & p_{22} & p_{23}Z + p_{24} \\ p_{31} & p_{32} & p_{33}Z + p_{34} \end{pmatrix} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}, \tag{A.3}$$

which can be combined in

$$\lambda \begin{pmatrix} \mathbf{p}_h \\ 1 \end{pmatrix} = \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & Z_x\mathbf{p}_3 + \mathbf{p}_4 \end{bmatrix} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}. \tag{A.4}$$
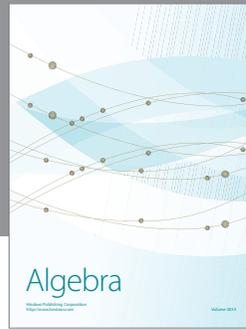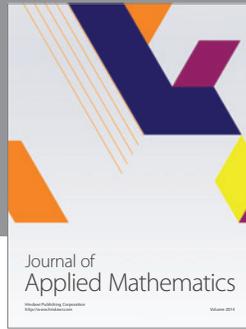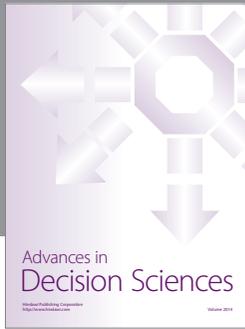
Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] F. W. Wheeler, R. L. Weiss, and P. H. Tu, "Face recognition at a distance system for surveillance applications," in *Proceedings of the 4th IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS '10)*, pp. 1–8, IEEE, Washington, DC, USA, September 2010.

[2] A. Del Bimbo, F. Dini, G. Lisanti, and F. Pernici, "Exploiting distinctive visual landmark maps in pan-tilt-zoom camera networks," *Computer Vision and Image Understanding*, vol. 114, no. 6, pp. 611–623, 2010.

[3] H.-C. Choi, U. Park, and A. K. Jain, "PTZ camera assisted face acquisition, tracking & recognition," in *Proceedings of the 4th IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS '10)*, pp. 1–6, Washington, DC, USA, September 2010.

[4] U. Park, H.-C. Choi, A. K. Jain, and S.-W. Lee, "Face tracking and recognition at a distance: a coaxial and concentric PTZ camera system," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 10, pp. 1665–1677, 2013.

[5] A. Criminisi, I. Reid, and A. Zisserman, "Single view metrology," *International Journal of Computer Vision*, vol. 40, no. 2, pp. 123–148, 2000.

[6] D. E. Stevenson and M. M. Fleck, "Robot aerobics: four easy steps to a more flexible calibration," in *Proceedings of the 5th International Conference on Computer Vision*, pp. 34–39, Cambridge, Mass, USA, June 1995.

[7] H. Bakstein and T. Pajdla, "Panoramic mosaicing with a 180° field of view lens," in *Proceedings of the IEEE 3rd Workshop on Omnidirectional Vision*, pp. 60–67, IEEE, Copenhagen, Denmark, June 2002.

[8] C. Hughes, E. Jones, M. Glavin, and P. Denny, "Validation of polynomial-based equidistance fish-eye models," in *Proceedings of the IET Irish Signals and Systems Conference (ISSC 2009 '09)*, pp. 1–6, IET, Dublin, Ireland, June 2009.

[9] S. Shah and J. K. Aggarwal, "Intrinsic parameter calibration procedure for a (high-distortion) fish-eye lens camera with distortion model and accuracy estimation," *Pattern Recognition*, vol. 29, no. 11, pp. 1775–1788, 1996.

[10] F. Devernay and O. Faugeras, "Straight lines have to be straight," *Machine Vision and Applications*, vol. 13, no. 1, pp. 14–24, 2001.

[11] C. Bräuer-Burchardt and K. Voss, "A new algorithm to correct fish-eye- and strong wide-angle-lens-distortion from single images," in *IEEE International Conference on Image Processing (ICIP '01)*, pp. 225–228, Thessaloniki, Greece, October 2001.

[12] S. Ramalingam, P. Sturm, and S. K. Lodha, "Towards complete generic camera calibration," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, vol. 1, pp. 1093–1098, June 2005.

[13] J. Kannala and S. S. Brandt, "A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 8, pp. 1335–1340, 2006.

[14] R. Hartley and S. B. Kang, "Parameter-free radial distortion correction with center of distortion estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 8, pp. 1309–1321, 2007.

[15] A. W. Fitzgibbon, "Simultaneous linear estimation of multiple view geometry and lens distortion," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. I125–I132, December 2001.

[16] B. Mičušik and T. Pajdla, "Estimation of omnidirectional camera model from epipolar geometry," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 485–490, June 2003.

[17] D. Claus and A. W. Fitzgibbon, "A rational function lens distortion model for general cameras," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '05)*, vol. 1, pp. 213–219, IEEE, San Diego, Calif, USA, June 2005.

[18] J. P. Barreto and K. Daniilidis, "Fundamental matrix for cameras with radial distortion," in *Proceedings of the 10th IEEE International Conference on Computer Vision (ICCV '05)*, vol. 1, pp. 625–632, Beijing, China, October 2005.

[19] J. Wei, C.-F. Li, S.-M. Hu, R. R. Martin, and C.-L. Tai, "Fisheye video correction," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 10, pp. 1771–1783, 2012.

[20] Y. Liu, S. Lai, C. Zuo, H. Shi, and M. Zhang, "A master-slave surveillance system to acquire panoramic and multiscale videos," *The Scientific World Journal*, vol. 2014, Article ID 491549, 11 pages, 2014.

[21] X. Zhou, R. T. Collins, T. Kanade, and P. Metes, "A master-slave system to acquire biometric imagery of humans at distance," in *Proceedings of the 1st ACM International workshop on Video Surveillance*, pp. 113–120, Berkeley, Calif, USA, November 2003.

[22] H. C. Liao and Y. C. Cho, "A new calibration method and its application for the cooperation of wide-angle and Pan-Tilt-Zoom cameras," *Information Technology Journal*, vol. 7, no. 8, pp. 1096–1105, 2008.

[23] L. Marchesotti, S. Piva, A. Turolla, D. Minetti, and C. S. Regazzoni, "Cooperative multisensor system for real-time face detection and tracking in uncontrolled conditions," in *Proceedings of the Image and Video Communications and Processing*, vol. 5689 of *Proceedings of SPIE*, pp. 100–114, Miami, Fla, USA, January 2005.

[24] Y. Xu and D. Song, "Systems and algorithms for autonomous and scalable crowd surveillance using robotic PTZ cameras assisted by a wide-angle camera," *Autonomous Robots*, vol. 29, no. 1, pp. 53–66, 2010.

[25] L. You, S. Li, and W. Jia, "Automatic weak calibration of master-slave surveillance system based on mosaic image," in *Proceedings of the 20th International Conference on Pattern Recognition (ICPR '10)*, pp. 1824–1827, IEEE, Istanbul, Turkey, August 2010.

[26] A. Hampapur, S. Pankanti, A. Senior, Y.-L. Tian, L. Brown, and R. Bolle, "Face cataloger: multi-scale imaging for relating identity to location," in *Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS '03)*, pp. 13–20, Miami, Fla, USA, July 2003.

[27] A. Basu and S. Licardie, "Alternative models for fish-eye lenses," *Pattern Recognition Letters*, vol. 16, no. 4, pp. 433–441, 1995.

[28] C. Hughes, E. Jones, M. Glavin, and P. Denny, "Validation of polynomial-based equidistance fish-eye models," in *Proceedings of the IET Irish Signals and Systems Conference (ISSC '09)*, pp. 1–6, Dublin, Ireland, June 2009.

[29] R. Cipolla, T. Drummond, and D. Robertson, "Camera calibration from vanishing points in image of architectural scenes," in *Poceedings of the British Machine Vision Conference (BMVC '99)*, pp. 382–391, Nottingham, UK, September 1999.