

Calibration of a Hybrid Camera Network

Xilin Chen Jie Yang Alex Waibel
School of Computer Science
Carnegie Mellon University,
5000 Forbes Ave, Pittsburgh, PA 15213, USA
{xlchen, yang+, ahw}@cs.cmu.edu

ABSTRACT

Visual surveillance using a camera network has imposed new challenges to camera calibration. An essential problem is that a large number of cameras may not have a common field of view or even be synchronized well. We propose to use a hybrid camera network that consists of catadioptric and perspective cameras for a visual surveillance task. The relations between multiple views of a scene captured from different cameras can be then calibrated under the catadioptric camera's coordinate system. This paper addresses the important issue of how to calibrate the hybrid camera network. We calibrate the hybrid camera network in three steps. First, we calibrate the catadioptric camera using only the vanishing points. In order to reduce computational complexity, we calibrate the camera without the mirror first and then calibrate the catadioptric camera system. Second, we determine 3D positions of some points using as few as two spatial parallel lines and some equidistance points. Finally, we calibrate other perspective cameras based on these known spatial points.

1. Introduction

A camera network enables a visual surveillance system to monitor activities of targets over a larger area than a single camera. However, there are many challenges in making a camera network work properly. For example, it is very difficult to calibrate a camera network because a large number of cameras may not have a common field of view or even be synchronized well. In this research, we propose to use a hybrid camera network that consists of catadioptric and perspective cameras for a visual surveillance task. This combination will minimize possibilities of losing track of a target when it is occluded by other people and/or objects. A perspective video camera has a limited viewing angle and is directional. The human target could be occluded from multiple normal video cameras if he/she is among a crowd. An omnidirectional camera mounted on the ceiling, on the other hand, has a 360 degree viewing angle and, by virtue

of being above the “action,” is less prone to problems of occlusion. However, an omnidirectional camera has lower resolution because of its wider view angle. A hybrid camera network has several advantages. First, it provides a framework to combine multiple views of a scene captured from different cameras. In a camera network, a large number of cameras may not have a common field of view, and even be synchronized well. If we know the relationship between each camera and the catadioptric camera, we then know the relationship among all the cameras. Second, it provides a framework to fuse information at different levels. While a catadioptric camera provides coarse information about locations of the targets, perspective cameras can provide high resolution images for more precise analysis. Third, the camera network can provide not only appearance information but also 3D information by configuring multiple cameras.

In order to make a hybrid camera network functional, we have to calibrate it first. Camera calibration is to determine the intrinsic and extrinsic parameters of a camera, such as aspect ratio, focal length, skew, radial distortion, principal point, translation, and rotation. Many methods and algorithms have been developed for camera calibration over the last three decades. Yet it is still a problem for many applications. Classical calibration methods can be classified as either nonlinear or linear techniques. Nonlinear methods in photogrammetry engage full-scale nonlinear optimizations with large numbers of parameters [2, 4, 11]. These methods can deal with complex camera models and generate accurate results with a good initialization, but expensive computation has limited the use of nonlinear methods. Linear methods, such as direct linear transformation (DLT) technique [1], are the most commonly used calibration techniques in the field of computer vision. Among a large number of publications on camera calibration, few have targeted the calibration of catadioptric cameras. The catadioptric camera with a single effective viewpoint [10] is the most popular omnidirectional camera. Geyer and Daniilidis proposed a method for calibration of a parabolic mirror [7]. The method can calibrate a catadioptric camera with a parabolic mirror using three lines. Cauchois et. al presented a technique for calibration of a catadioptric

camera with a conic mirror using a calibration pattern mounted all around the sensor [3]. Kang proposed a direct self-calibration method by minimizing the epipolar constraint [9]. Strelow et. al developed a robust algorithm for calibrating various amounts of misalignment between the mirror and camera [12]. Sturm recently analyzed relations between catadioptric and perspective cameras within a scene [13]. However, to our knowledge, no previous work has directly involved in calibrating a hybrid camera network.

In this paper, we propose a method to calibrate a hybrid camera network that consists of catadioptric and perspective cameras. The calibration procedure includes three steps. First, we calibrate the catadioptric camera using only the vanishing points. In order to reduce computational complexity, we calibrate the camera without the mirror first and then calibrate the catadioptric camera system. Second, we determine 3D positions of some points using as few as two spatial parallel lines and some equidistance points. Finally, we calibrate other perspective cameras based on these known spatial points.

2. Calibration of a Hybrid Camera Network

In this research, we consider calibration of a hybrid camera network that consists of catadioptric and perspective cameras. Figure 1 depicts a visual surveillance system using the hybrid camera network. We use the catadioptric camera as a center of the scene. The catadioptric camera, at least, overlaps partial view of the scene with other perspective cameras, but we make no assumption on view overlaps of other perspective cameras and synchronization of cameras. The basic idea is to use the coordinate system of the catadioptric camera as the global coordinate system. We then obtain relationships among cameras in the network from the relationships between each camera and the global coordinate system. In this section, we present a calibration procedure that includes three steps: calibration of the catadioptric camera, localization of some spatial points, and calibration of other cameras.

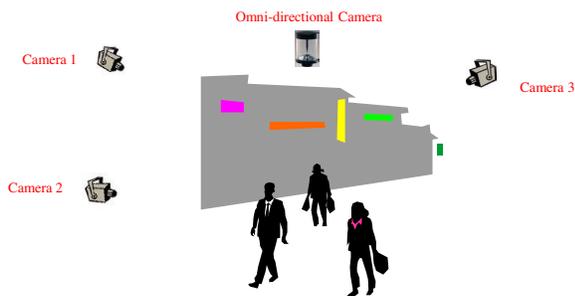


Figure 1. Illustration of a hybrid camera network

2.1. Calibration of a catadioptric camera

A catadioptric system is a combination of mirror and lenses. In this research, we employ a commercially available catadioptric system [10]. We represent the project model of the catadioptric camera as Figure 2. We assume that the camera is focused on the virtual focal plane \mathcal{F} . The image of the object from the mirror orthographically projects to the virtual focal plane \mathcal{F} , then maps perspectively onto the image plane.

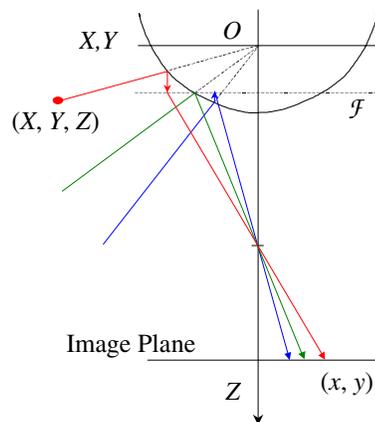


Figure 2. The catadioptric model

As shown in Figure 2, the focus of the paraboloid mirror is selected as the origin of the global coordinate, and the paraboloid can be described as:

$$Z = f_p - \frac{1}{4f_p}(X^2 + Y^2). \quad (1)$$

In a catadioptric camera system, the focal length f is a combination of paraboloid focal length and the camera focal length (Note that in the rest of this paper, the word “catadioptric camera” means a camera and associated reflector). The camera in the catadioptric system has a translation $T = (T_x \ T_y \ T_z)^T$ and rotation $\mathbf{R}_{3 \times 3}$ related to the reflector.

Given a spatial point (X, Y, Z) , its projection on the paraboloid mirror is (X_p, Y_p, Z_p) . The projection is (X_c, Y_c, Z_c) under the camera's coordinate. The final imaging of the point (X, Y, Z) is (x, y) on the image plane. If we ignore the radial distortion of the camera, the imaging model is as following:

$$Z_c \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} f_x & \beta & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix}, \quad (2)$$

where f_x and f_y are the focal length of the camera in x and y direction, and (x_0, y_0) is the center of the image. In the calibration of the catadioptric camera, we ignore the skew, i.e., $\beta = 0$.

$$\begin{pmatrix} X_C \\ Y_C \\ Z_C \end{pmatrix} = \mathbf{R} \begin{pmatrix} X_P \\ Y_P \\ Z_P \end{pmatrix} - \begin{pmatrix} T_X \\ T_Y \\ T_Z \end{pmatrix}, \quad (3)$$

$$\begin{pmatrix} X_P \\ Y_P \\ Z_P \end{pmatrix} = \frac{2f_p(\sqrt{X^2 + Y^2 + Z^2} - Z)}{X^2 + Y^2} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}, \quad (4)$$

where f_p is the focal length of the paraboloid.

In the catadioptric system, the rotation is

$$\mathbf{R} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \quad (5)$$

Considering that the focal plane of the camera in a catadioptric system is fixed, i.e., $Z_C = T'_Z = \text{constant}$, we can obtain the following equation from (2), (3), and (5).

$$\begin{aligned} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} &= \begin{pmatrix} f_x & 0 & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X_C / T'_Z \\ Y_C / T'_Z \\ 1 \end{pmatrix} \\ &= \frac{f_x}{T'_Z} \begin{pmatrix} 1 & 0 \\ 0 & \alpha \\ 0 & 0 \end{pmatrix} \begin{pmatrix} X_P - T_X \\ -Y_P - T_Y \end{pmatrix} + \begin{pmatrix} x_0 \\ y_0 \\ 1 \end{pmatrix}, \quad (6) \\ &= \frac{f_x}{T'_Z} \begin{pmatrix} 1 & 0 \\ 0 & -\alpha \\ 0 & 0 \end{pmatrix} \begin{pmatrix} X_P \\ Y_P \end{pmatrix} + \begin{pmatrix} x_0 - f_x T_X / T'_Z \\ y_0 - f_y T_Y / T'_Z \\ 1 \end{pmatrix} \end{aligned}$$

where α is the aspect ratio of the camera. The second term is the principle point of the catadioptric system.

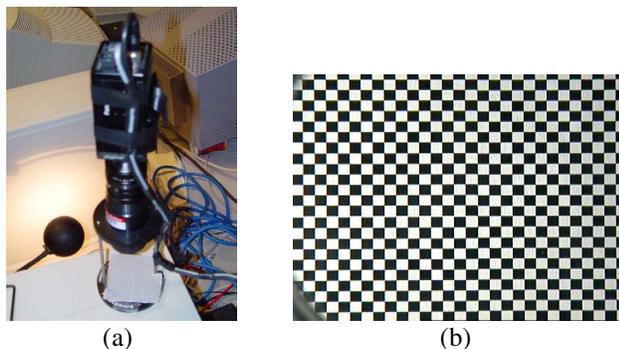


Figure 3. Calibration of catadioptric camera without the mirror: (a) setup, and (b) captured pattern image

From equation (4) and (6), we find that after aspect rectification, the combined focal length is $f = f_x f_p$. These two equations also indicate that the calibration can be divided into two parts: (1) camera calibration without the mirror, and (2) calibration of the catadioptric system. In fact, in the first part, we only need to obtain the aspect ratio α . The combined focal length and the principle point will be estimated in the second part of the calibration. A simple way to obtain the aspect ratio α is shown in Figure 3: replacing the mirror with a pattern board and estimating the aspect ratio α from the pattern image.

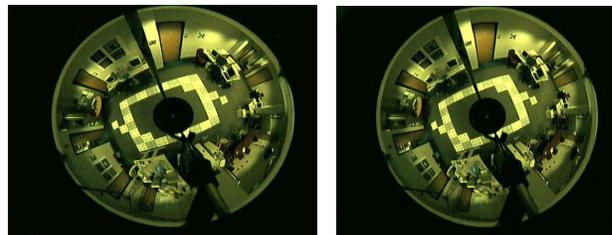


Figure 4. Using the aspect ratio to rectify an image for catadioptric camera calibration

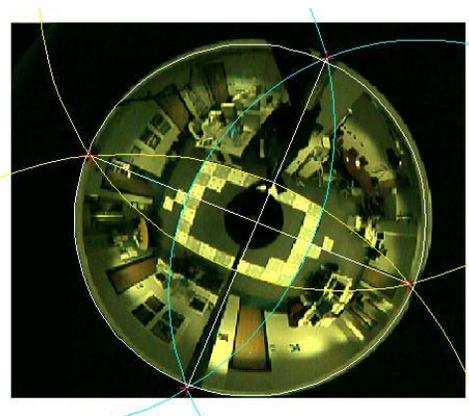


Figure 5. Calibration of a catadioptric camera using vanishing points

Once we have known the aspect ratio, we are able to rectify the image. Figure 4 shows an example of an image before and after aspect ratio rectification. We then use the algorithm proposed by Geyer and Daniilidis [6] to estimate the combined focal length. The algorithm is based on their theory on the projection of parallel lines on such a catadioptric system: if the pixel aspect ratio is equal to one, parallel lines project onto circular arcs whose centers are collinear. The projection mapping of a parabolic-orthographic system preserves incidence relationships. It can be further proved that the vanishing point of a set of parallel lines maps onto the intersection of the circular arcs. For each set of parallel lines, we can obtain two vanishing points, one for each direction. Therefore, the image center is located at the intersection of the lines connecting the vanishing points, and the vanishing points

form a circle. The radius of this circle is double of the combined focal lengths. Figure 5 illustrates an example of calibrating a catadioptric camera using vanishing points

2.2. Localization of spatial points under catadioptric camera coordinate system

In order to establish relationships between the catadioptric camera and perspective cameras, we need to know the coordinates of some spatial points under the coordinate system of the catadioptric camera. The following proposition can guide us to obtain coordinates of spatial points.

Proposition *The relationship between the spatial point $P = (X_p, Y_p, Z_p)$ and $Q = (X_q, Y_q, Z_q)$ and their imaging in the calibrated catadioptric image $\mathbf{p} = (x_p, y_p)$ and $\mathbf{q} = (x_q, y_q)$ are given by the following equations:*

$$X_p = x_p D \frac{\begin{vmatrix} n_x & n_y \\ x_q & y_q \end{vmatrix}}{\begin{vmatrix} x_q & y_q \\ x_p & y_p \end{vmatrix}}, \quad Y_p = y_p D \frac{\begin{vmatrix} n_x & n_y \\ x_q & y_q \end{vmatrix}}{\begin{vmatrix} x_q & y_q \\ x_p & y_p \end{vmatrix}},$$

$$Z_p = D \frac{\begin{vmatrix} n_x & n_y \\ x_q & y_q \\ x_p & y_p \end{vmatrix}}{\begin{vmatrix} x_q & y_q \\ x_p & y_p \end{vmatrix}} \left(f - \frac{x_p^2 + y_p^2}{4f} \right). \quad (7)$$

$$\begin{aligned} X_q &= X_p + n_x D \\ Y_q &= Y_p + n_y D \\ Z_q &= Z_p + n_z D \end{aligned} \quad (8)$$

where $\vec{n} = (n_x \ n_y \ n_z)'$ is the spatial direction from P to Q , D is the distance between P and Q , and f is the combination focal length of the catadioptric camera.

Proof: Equation (8) is straightforward. From equation (8) and the basic geometry relationship between a spatial object and its imaging, we have:

$$\begin{cases} \frac{X_p + n_x D}{x_p} = \frac{Y_p + n_y D}{y_p} = \frac{Z_p + n_z D}{z_p} \\ \frac{x_q}{x_p} = \frac{y_q}{y_p} = \frac{z_q}{z_p} \end{cases}, \quad (9)$$

We can have equation (7) by solving equation (9).

To apply this proposition, we need to know the spatial direction of the points. The spatial direction of two points can be obtained by measuring two or more spatial lines that are parallel to the line that connects these two points.

Figure 6 shows an example of 14 spatial points that have been located in the rectified image. In figure 6, “+” indicates the location of a point and numbers in parentheses are the estimated 3-D coordinates of the points.

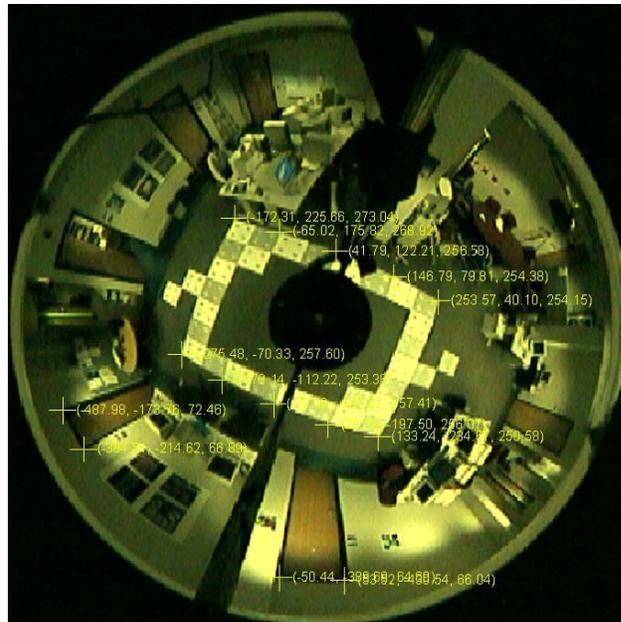


Figure 6. Locations of 14 spatial points

2.3. Calibration of perspective cameras

While an omnidirectional camera covers a large area of a scene, a perspective camera can only cover a part of it. Figure 7 shows images captured from an omnidirectional camera and a perspective camera mounted on the ceiling of a laboratory. The image boundary of the scene captured by the perspective camera has been marked in the omniview (Figure 7(b)).

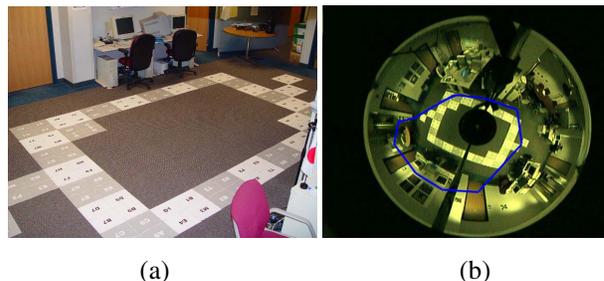


Figure 7. The perspective view (a) covers only part of the omniview (b)

Numerous calibration methods have been proposed for perspective camera [5, 8, 14, 15, 16]. Most of them use the same model or equivalent as equation (2) and (3).

Given a spatial point Q observable from both the omnidirectional camera and the perspective camera,

suppose that its position is $\mathbf{X}_O = (X_O \ Y_O \ Z_O)^T$ under the coordinate system of the omnidirectional camera, and $\mathbf{X}_P = (X_P \ Y_P \ Z_P)^T$ under the coordinate system of the perspective camera respectively, and its image coordinates in the perspective camera is $(x_p \ y_p)^T$. Then we have

$$\begin{aligned} \mathbf{X}_P &= \mathbf{R}_{3 \times 3} \mathbf{X}_O - \mathbf{T}_{3 \times 1} = \begin{pmatrix} \mathbf{R}_{3 \times 3} & -\mathbf{T}_{3 \times 1} \end{pmatrix} \begin{pmatrix} \mathbf{X}_O \\ 1 \end{pmatrix}, \\ Z_P \begin{pmatrix} x_p \\ y_p \\ 1 \end{pmatrix} &= \begin{pmatrix} f_x & \gamma & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{R}_{3 \times 3} & -\mathbf{T}_{3 \times 1} \end{pmatrix} \begin{pmatrix} \mathbf{X}_O \\ 1 \end{pmatrix}, \\ &= \mathbf{M}_{3 \times 4} \begin{pmatrix} \mathbf{X}_O \\ 1 \end{pmatrix} \end{aligned}$$

$\mathbf{M}_{3 \times 4}$ is the parameter matrix in this step. It can be determined from as few as 6 spatial points. Therefore, we can calibrate a perspective camera under the coordinate system of the omnidirectional camera as long as we have 6 or more points observable from both the omnidirectional and perspective cameras and know their coordinates under the coordinate system of the omnidirectional camera.

The procedure for calibration of a hybrid camera network is summarized as following:

1. **Calibration of a catadioptric camera:** We calibrate a catadioptric camera in two steps: first we calibrate the camera without the mirror. We obtain the aspect ratio from a pattern image captured by the camera. We use the aspect ratio to rectify images for further calibration. We use vanishing points of the spatial parallel lines to fit a circle. If fitted circle's center is C_1 , and the radius is r , the intersection C of the lines which connect the vanishing point pair is the principal point of the catadioptric system, and the focal length is

$$f = \sqrt{(r - |C - C_1|)(r + |C - C_1|)}.$$

2. **Selection of spatial points:** We use the coordinate system of the catadioptric camera as the global coordinate system. We calculate distances and spatial directions of several pair points for calibrating perspective cameras. For calibration of each perspective camera, we need at least six points observable from both the catadioptric camera and the perspective camera.
3. **Calibration of perspective cameras:** We use the known spatial points obtained from step 2 to calibrate a perspective camera at each time. Repeat this process until all perspective cameras have been calibrated.

3. Experimental Results

We have tested the proposed method by calibrating a camera network that consists of Cyclovision's omnidirectional camera and four Sony video cameras. The omnidirectional camera uses a Panasonic camera as imaging sensor.

First, we calibrated the omnidirectional camera. We captured a calibration image without mirror as shown in Figure 3(b) to obtain the aspect ratio of the camera attached to the omnidirectional system. The aspect ratio was 0.896346. With this aspect ratio, all the images were rectified before further processing. The combined focal length was 113.1364 and the principal point was at (307.6756, 227.7280). Using the calibrated parameter, we located some spatial points. Table 1 shows the difference between the estimated result and the ground truth. In Table 1, we define the Error as $|\mathbf{X} - \mathbf{X}'|/|\mathbf{X}|$, \mathbf{X} is the ground truth, and \mathbf{X}' is the estimated position. The error is less than 3.5%.

Table 1. A comparison between the estimation and ground truth (all the coordinates are in *cm*)

Estimated position	Ground Truth	Error
(-275.5, -70.3, 257.6)	(-279.24, -62.56, 259.0)	0.0226
(-173.1, -112.2, 253.4)	(-176.46, -106.19, 259.0)	0.0268
(-72.2, -157.2, 257.4)	(-73.69, -149.81, 259.0)	0.0250
(29.8, -197.5, 256.1)	(29.09, -193.44, 259.0)	0.0155
(133.2, -234.7, 250.6)	(131.86, -237.06, 259.0)	0.0235
(146.8, 79.8, 254.4)	(156.61, 80.03, 259.0)	0.0346
(41.8, 122.2, 256.6)	(42.83, 122.63, 259.0)	0.0091

To test the accuracy of the calibrated hybrid camera network, we used several points estimated from the omnidirectional to estimate $\mathbf{M}_{3 \times 4}$ of the perspective cameras. After we obtained the parameter matrix, we re-projected the other measured points onto the perspective view, and measured the difference between the virtual imaging and the real imaging. Figure 8 depicts the result. In Figure 8, X-axis and Y-axis represent pixel errors in the X direction and Y direction in the image plane. Therefore, the closer to origin a point is, the more accurate. Unfortunately, although the majority of testing points are closed to the origin, some are not. We have carefully analyzed experimental results and error sources. Some comments are in order:

1. We have discovered that 6 carefully selected points can achieve a comparable accuracy as that from more points. This suggests that we should carefully select good points rather than select more points.

2. The points near the center of the mirror in the omniview provide good accuracy. On the other hand, the points with bigger errors are those near the edge of the mirror. Possible reasons include imperfect surface near the edge, nonlinear mapping or/and low-resolution sampling. In fact, the same area in the edge corresponds to more spatial points than that of the center area. In order to achieve high accuracy, we should avoid selecting points near the edge of the mirror.
3. Since the points near the edge of the mirror have less accuracy, we should have a larger search window when we use the hybrid camera network for a visual surveillance task and the target is close to the edge area of the mirror.

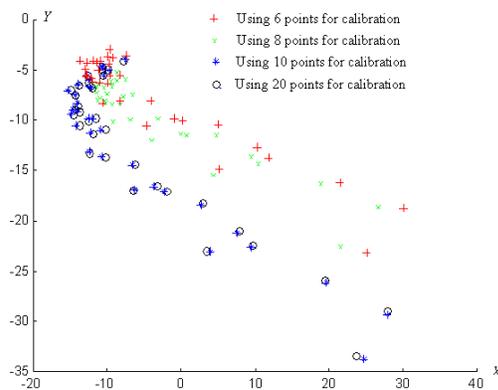


Figure 8. Errors of real image and virtual image for selected points

4. Conclusions

We have introduced a method for calibration of a hybrid camera network that consists of catadioptric and perspective cameras. We have shown a practical procedure that calibrates a hybrid camera network in three steps. The first step of the procedure is to estimate parameters for the catadioptric camera system. Then we estimate some spatial points under the catadioptric camera in the second step. The third step establishes relationship between catadioptric and perspective cameras. We have demonstrated feasibility of the proposed method by experiments. We are currently developing algorithms for visual surveillance using the calibrated hybrid camera network.

5. Acknowledgements

This research is support, in part, by the Department of Defense (United States of America) through award number N41756-03-C4024.

6. References

- [1] Y.I. Abdel-Aziz and H.M. Karara, "Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry", Proceedings of the Symposium on Close-Range Photogrammetry, 1971, pp. 1-18.
- [2] D.C. Brown, "Close-range camera calibration, Photogrammetric Engineering", August 1971, Vol. 37, No. 8, pp. 855-866.
- [3] C. Cauchois, E. Brassart, L. Delahoche, and T. Delhomme. "Reconstruction with the calibrated syclop sensor", Proceedings of IEEE International Conference on Intelligent Robots and Systems, 2000, pp.1493-1498.
- [4] W. Faig, "Calibration of close-range photogrammetry system: mathematical formulation, Photogrammetric Engineering and Remote Sensing", 1975, Vol. 41, No. 12, pp. 1479-1486.
- [5] O. D. Faugeras and G. Toscani, "The calibration problem for stereo", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1986, pp. 15-20.
- [6] C. Geyer and K. Daniilidis, "Catadioptric camera calibration", Proceedings of 7th International Conference on Computer Vision, 1999, Vol. 1, pp. 398-404.
- [7] C. Geyer and K. Daniilidis, "Para-catadioptric calibration", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, Vol. 24, No. 5, pp. 687-695.
- [8] R. Hartley, and A. Zisserman, Multiple view geometry in computer vision, Cambridge University Press, 2002.
- [9] S. Kang, "Catadioptric self-calibration", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2000, pp. 201-207.
- [10] S. K. Nayar, "Catadioptric omnidirectional camera", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 1997, pp. 482-488.
- [11] I. Sobel, "On calibrating computer controlled cameras for perceiving 3-D scenes", Artificial Intelligence, 1974, Vol. 5, No. 2, pp. 185-198.
- [12] D. Strelow, J. Mishler, D. Koes, and S. Singh, "Precise omnidirectional camera calibration", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2000, Vol. 1, pp. 689-694.
- [13] P. Sturm, "Mixing catadioptric and perspective cameras", Proceedings of the IEEE Workshop on Omnidirectional Vision, 2002, pp. 37- 44.
- [14] B. Triggs, "Camera pose and calibration from 4 or 5 known 3D points", The Proceedings of the 7th International Conference on Computer Vision, 1999, Vol. 1, pp. 278-284.
- [15] R. Y. Tsai, "An efficient and accurate camera calibration technique for 3D machine vision", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 1986, pp. 364-374.
- [16] M. Wilczkowiak, E. Boyer, P. Sturm, "Camera calibration and 3D reconstruction from single images using parallelepipeds", Proceedings of the 8th International Conference on Computer Vision, 2001, pp. 142-148.