Camera Calibration from Two Shadow Trajectories

Fei Lu, Xiaochun Cao, Yuping Shen, Hassan Foroosh
University of Central Florida
Computational Imaging Laboratory
{feilu,xccao,ypshen,foroosh}@cs.ucf.edu

Abstract

We introduce an efficient method for recovering the camera parameters automatically from the cast shadows of two 3D points observed over time. Compared to previous related work, our method has less restrictions in the sense that object-to-shadow correspondences do not have to be available in the image. We demonstrate how the horizon line may be recovered from only shadow points, and how the camera intrinsic and extrinsic parameters are determined using the pole-polar relationship and minimizing the algebraic distance of the principal point. The approach is fully validated on both synthetic and real data, and tested against various sources of error. We finally present an application to metrology from shadows only - i.e. when the object is not visible in the image.

1. Introduction

Camera calibration is an essential task in computer vision which is aiming at extracting intrinsic and extrinsic camera parameters from 2D images. Traditional techniques ([6], [11], [12]) usually involve taking images of some special pattern with known 3D geometry, extracting features and minimizing their reprojection errors. These methods often produce very good results but require the euclidian information from the calibration patterns. Recent multiple view approaches [7] [8] avoid the use of special calibration objects, but typically require more than three views and have higher computational cost and also involve non-linear problems.

1.1 Related Work and Contributions

Objects and their cast shadows are common in the real world, especially in outdoor environments. There has been some investigation of outdoor camera calibration (e.g. [4]) while most methods attempt to eliminate shadows in order to improve the algorithms rather than exploit their properties to deduce camera parameters.

However, Shadows are interesting for camera calibration since the shadows of two parallel lines cast on a planar surface by an infinite point light source are also parallel. Calibration using cast shadows was first proposed in [2]. A major drawback of this approach is that it requires accurate geography (e.g. longitude, latitude) information as well as time stamps for each particular image, which is not always possible to obtain. Inter-image constraints on the IAC from objects and their cast shadows were introduced in [3], where the authors argue that two views of an outdoor scene with two vertical objects under sunshine are enough to calibrate a camera.Both these methods require the entire objects and their shadows to be present in the image, in order to select their top and bottom points. Furthermore, they also require that objects in the image have regular shapes such that their corresponding shadow lines are parallel. Therefore these methods are not applicable in some scenarios such as a part of a building and its shadow captured in videos, or even worse, only shadows in the scene with no object-to-shadow correspondences.

In this paper, we assume that our camera has a unit aspect ratio and zero skew as argued in [8]. In our method, we recover the horizon line in the image from the shadows only, and without the knowledge of object-to-shadow correspondences. Our method is then based on the pole-polar relationship and and the minimization of algebraic distance algorithm for the principal point. Different from [2] [3], the contribution of our method lies on eleminating the requirement for known 3D quantities, and on recovering camera parameters from information determined only by images.

This paper is organized as following: §2 introduces the fundamentals about camera calibration and the Image of Absolute Conic. §3 presents the computation of horizon line from ground shadow trajectory. §4 introduces the method for calibration from shadows. §5 shows the results of synthetic and real experiments. Conclusions are given in §6.

2. Calibration with Image of Absolute Conic

Image of the Absolute Conic ω [8] is an imaginary point conic on the plane at infinity which is invariant to camera

The 18th International Conference on Pattern Recognition (ICPR'06)
0-7695-2521-0/06 $20.00 © 2006 IEEE

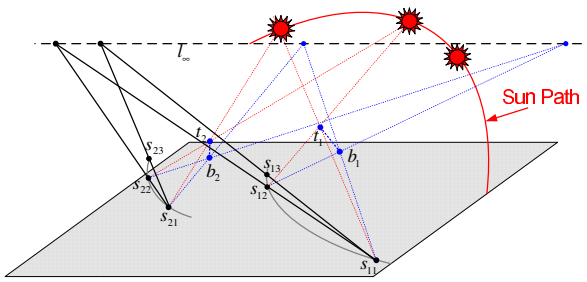


Figure 1. The 3D points $\mathbf{t}_i$ cast shadows at different time j at positions $\mathbf{s}_{ij}$

rotation and translation. It is directly related to the camera internal matrix K via

$$\omega = \mathbf{K}^{-\mathbf{T}}\mathbf{K}^{-\mathbf{1}} \tag{1}$$

where $\mathbf{K}$ is the camera calibration matrix. Assuming a unit aspect ratio and zero camera skew, ω can be expanded up to scale as

$$\omega \sim \begin{bmatrix} 1 & 0 & \omega_{13} \\ 0 & 1 & \omega_{23} \\ \omega_{13} & \omega_{23} & \omega_{33} \end{bmatrix} \tag{2}$$

The vanishing point $\mathbf{v}_z$ of the normal direction to a plane and the vanishing line $\mathbf{l}_\infty$ of that plane satisfy the pole-polar relationship with respect to the image of absolute conic ω

$$\mathbf{l}_\infty = \omega \mathbf{v_z} \tag{3}$$

After ω is determined, $\mathbf{K}$ can be easily obtained by Cholesky decomposition.

3. Horizon line from shadow trajectories

In this section , we demonstrate that the tracked shadow trajectories of two stationary objects on the ground plane can be used to recover the horizon line. The key idea is illustrated in Fig. 1 and is based on the fact that the line segments defined by corresponding shadow points, e.g. $\mathbf{s}_{11}\mathbf{s}_{12}$ and $\mathbf{s}_{21}\mathbf{s}_{22}$, are parallel in the world and therefore intersect on the horizon line $\mathbf{l}_\infty$ in the image plane.

To prove this, let $\mathbf{t}_i$ be a 3D point and $\mathbf{b}_i$ the closest point on the ground plane to $\mathbf{t}_i$. The line $\mathbf{t}_i\mathbf{b}_i$ is perpendicular to the ground plane. Now let s_{ij} be the casted shadow of point $\mathbf{t}_i$ at time $\mathbf{j}$. Because the sun is far away (1.5×10^{11} meters from the earth), it is reasonable to consider the sunlight as parallel light rays. Therefore, $\mathbf{b}_1\mathbf{s}_{1j}$ and $\mathbf{b}_2\mathbf{s}_{2j}$ are parallel and hence we have

$$\frac{b_1 s_{11}}{b_2 s_{21}} = \frac{b_1 s_{12}}{b_2 s_{12}} = \frac{t_1 b_1}{t_2 b_2} \tag{4}$$

Of course that this equation is only in 3D, and not in the image plane. However, it indicates that the two triangles $\triangle \mathbf{b}_1\mathbf{s}_{11}\mathbf{s}_{12}$ and $\triangle \mathbf{b}_2\mathbf{s}_{21}\mathbf{s}_{22}$, are similar. So line $\mathbf{s}_{11}\mathbf{s}_{12}$ and $\mathbf{s}_{21}\mathbf{s}_{22}$ are parallel to each other. Therefore, the shadows observed over time are sufficient to provide the horizon line $\mathbf{l}_\infty$ and the object top and bottom points ($\mathbf{t}_i$, $\mathbf{b}_i$) are not required to be present in the image.

Now, if we assume that the vertical vanishing point is available, e.g. using the methods [9] [10], the computed horizon line $\mathbf{l}_\infty$, together with the vertical vanishing point $\mathbf{v}_z$, provides two constraints in Eq.(3) on the IAC. The main advantage of this method, compared to [2] and [3], is the looser restriction on the scene. It only requires the input of shadows tracked over time any knowledge of the 3D points that cast the shadows. Since scenes observed over time is typical in surveillance systems and because our method requires cast shadows of unknown objects, it has potential applications in video surveillance systems as demonstrated in Section 5.

3.1 Shadow trajectory detection

Here we introduce a semi-automatic approach to detect the shadow trajectory. First, the background $\mathbf{bg}$ (Fig. 5(b)) is defined using the brightest value of each pixel (r, c) among the set of input images $\{\mathbf{I}_1, \mathbf{I}_2, ...\mathbf{I}_n\}$:

$$\mathbf{bg}(r, c) = \mathbf{max}(\mathbf{I_k}(r, c)) \qquad k \in [1, n] \tag{5}$$

Then background subtraction is applied to each image and the prominent shadow points in the first frame are selected (Fig. 5(c)). Mean shift tracking algorithm [5] is then applied to determine the shadow trajectory in different frames.

4. Calibration from shadows

From Section 3 we know that Eq.(3) provides two constraints on IAC ω. Let $\mathbf{l}_\infty \sim [l_x, l_y, 1]^T$ and $\mathbf{v}_z \sim [v_x, v_y, 1]^T$. Using equations of Eq.(2) and Eq.(3), we can express ω_{13} and ω_{23} in terms of ω_{33} as follows

$$\omega_{13} = -\frac{-l_x v_y^2 + l_y v_x v_y - v_x + l_x w_{33}}{l_x v_x + l_y v_y - 1} \tag{6}$$

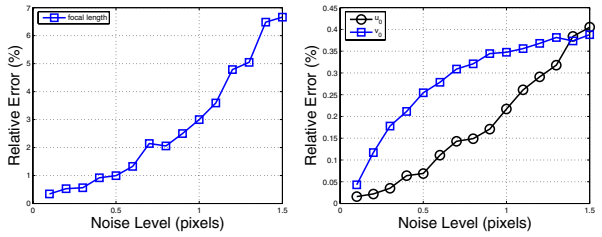
$$\omega_{23} = -\frac{-l_y v_x^2 + l_x v_x v_y - v_y + l_y w_{33}}{l_x v_x + l_y v_y - 1} \tag{7}$$

Here we introduce a cost function on the algebraic distance of the principal point from the center of the image (c_x, c_y) which gives an extra weak constraint on ω.

$$[u_0' \quad v_0'] = \mathbf{arg} \min_{u_0, v_0 \in \Omega} \sum_i (\omega_{13} + u_0)^2 + (\omega_{23} + v_0)^2 \tag{8}$$

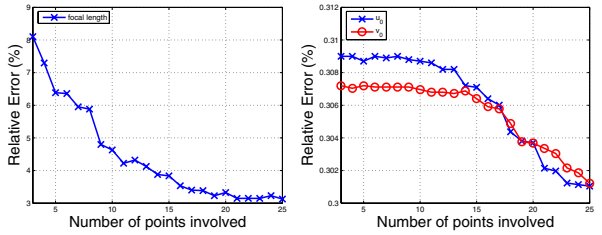
where Ω is the 2D searching space of $\mathbf{u}_0$ and $\mathbf{v}_0$. By substituting Eq.(6) and Eq.(7) into Eq.(8) and minimizing it, ω_{33} can be estimated, which in turn determines ω_{13} and ω_{23}.

Because the principal points of recent CCD cameras are very close to the center of the image, the searching space



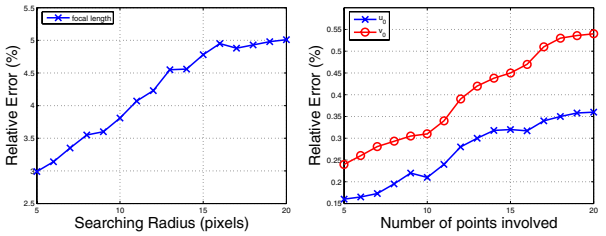
(a)focal length f (b)principal point u_0, v_0

Figure 2. Relative errors over different noise levels



(a)focal length f (b)principal point u_0, v_0

Figure 3. Relative errors as functions of number of shadow points



(a)focal length f (b)principal point u_0, v_0

Figure 4. Relative errors over different searching radius

Ω can be narrowed down to a window centered on the image center. Our experimental results in Section 5 prove the validity of this assumption.

With the knowledge of ω, vertical vanishing point and horizon line, camera extrinsic parameters, i.e. three rotation angles and translation vectors (up to scale) can also be computed similar to [3].

5. Experiments and Application

5.1 Synthetic data

The synthetic camera has fixed intrinsic parameters, with focal length $\mathbf{f} = 1000$ and the principal point at $\mathbf{u}_0 = 4$ and $\mathbf{v}_0 = 3$. 3D coordinates of two points are $(0, 0, 20)$ and $(20, 30, 40)$. The sun altitude angle and azimuth angle are real data from [1] between 10am and 2pm with 20 minutes interval (25 points total for each shadow trajectory). Each point in the image are corrupted with Gaussian noises of 16 different levels from $0 \sim 1.5$ pixel. For each level, 1000 independent trials are performed using the technique in Section 4 to compute the focal length and principal point. Estimated camera parameters were then compared with the ground truth. Fig. 2 (a) and (b) show the average relative errors of $\mathbf{f}$, $\mathbf{u}_0$ and $\mathbf{v}_0$ as a function of noise variance, with searching space $[-10, 10]$ centering in the image. It can be seen that the errors increase almost linearly as a function of the noise level. The maximum relative error of focal length reaches 6.66% when noise reaches 1.5 pixel. Meanwhile the maximum relative errors of principal point are 4.07% and 3.86% for $\mathbf{u}_0$ and $\mathbf{v}_0$, respectively.

(c) (d)

Figure 5. (a)A frame of real data (b)Computed background. (c)Shadow point tracking from background subtraction. (d)Shadow trajectories for a roof corner and the top of a tree

Theoretically, only three pairs of shadow points are sufficient to determine the vanishing line $\mathbf{l}_\infty$. In the second experiment, different numbers of shadow points are used for calibration. The noise level is fixed at 1 pixel. Otherwise, the setup is the same as the first one above. It can be seen from the results in Fig. 3 that the relative errors reduce as the number of shadow points increase.

In the third experiment, we compared performances under different searching spaces Ω, with all the 25 points in each shadow trajectory and a fixed Gaussian noises of 1 pixel. Searching radius is from 5 to 20 pixels and 1000 independent trails are performed for each case. Fig. 4 displays the average relative errors of focal length and the principal point. From the results we observe that the accuracy decreases when the searching space is larger, which is reasonable because smaller Ω makes the searching algorithm more stable and faster to converge.

5.2 Real Scene

A surveillance camera was set up near a bus station to capture surrounding environment and casted shadows between 10am and 2pm. The casted shadow trajectories of a

Table 1. Parameter estimation (All points involved)

Settings	$\mathbf{f}$	$\mathbf{u}_0$	$\mathbf{v}_0$
Our method	1453.6	576.2	434.5
Relative error	2.48%	0.54%	0.47%
[8]	1418.4	568.5	427.9

Table 2. Parameter estimation (3 pairs of shadow points involved for each computation)

Parameter	$\mathbf{f}$	$\mathbf{u}_0$	$\mathbf{v}_0$
Mean	1423.6	576.2	431.6
Standard Deviation	121.1	2.96	3.04
Relative error vs. [8]	0.32%	0.54%	0.26%

roof corner and the top of a tree in the scene were marked and selected for calibration. Fig. 5 (d) shows the two shadow trajectories. A vertical vanishing point is computed from the pillars. In order to evaluate our results, we implemented the classical method of using three pairs of vanishing points corresponding to orthogonal lines in an image [8]. Results are compared and are listed in Table 1, which indicates that the proposed algorithm provides good results.

As discussed above, three pairs of shadow points are sufficient for calibration. In this experiment there are **11** corresponding points for each trajectory and **3** of them are selected for calibration. Thus $\mathbf{C}_3^{11} = 165$ results can be computed. Table 2 shows the mean value and standard deviation of the 165 combinations and the relative error compared with [8].

5.3 Image-based Metrology Application

In this section, we show that the relative heights of the two world points $\mathbf{t}_i$ and $\mathbf{t}_j$ can be recovered from their shadow trajectories over time, even if the two points can not be seen in the images. For this purpose, we first recover the affine property (Fig. 6) of the ground plane by an affine rectification:

$$\mathbf{H}_p = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ l_1 & l_2 & l_3 \end{pmatrix}, \quad (9)$$

where $[l_1\ l_2\ l_3]^T$ is the vanishing line $\mathbf{l}_\infty$. Since the ratio of lengths on parallel lines is an affine invariant, from Eq.(4) and the fact that the two world triangles $\triangle \mathbf{b}_1\mathbf{s}_{11}\mathbf{s}_{12}$ and $\triangle \mathbf{b}_2\mathbf{s}_{21}\mathbf{s}_{22}$ are similar, we have

$$\frac{\mathbf{s}_{ik}\mathbf{s}_{il}}{\mathbf{s}_{jk}\mathbf{s}_{jl}} = \frac{\mathbf{t}_i\mathbf{b}_i}{\mathbf{t}_j\mathbf{b}_j}, \quad \forall i, j, k, l. \quad (10)$$

Therefore, given $\mathbf{n}$ observations of shadows of two object i and j, we get $\mathbf{C}_2^n$ solutions for $\frac{\mathbf{t}_i\mathbf{b}_i}{\mathbf{t}_j\mathbf{b}_j}$, from which the optimal solution can be computed as a weighted mean. The weight is dependent on the relative distance between k and l. In our experiment we have $\mathbf{n} = 11$, which produces 55 solutions for $\frac{\mathbf{t}_i\mathbf{b}_i}{\mathbf{t}_j\mathbf{b}_j}$. The optimal solution for the relative height of the two 3D points in Fig. 6(b) was computed and verified as 1.31.

(a) (b)

Figure 6. Image-based Metrology (a)Computed horizon line $\mathbf{l}_\infty$. (b)Two 3D points and their corresponding shadows in different time after affine transformation

6. Conclusions

This paper addresses the problem of camera calibration from two shadow trajectories. Prior knowledge of the 3D Euclidean coordinates are not required in our method. We also demonstrate that unlike existing methods object-to-shadow correspondences are not required. These advantages make our method an efficient calibration technique that could be especially suitable for outdoor video surveillance systems. Experimental results on both synthetic and real data verify the validity and usefulness of the method.

References

[1] Sun or moon altitude/azimuth table for one day http://aa.usno.navy.mil/data/docs/altaz.html.

[2] M. Antone and M. Bosse. Calibration of outdoor cameras from cast shadows. In *Proc. IEEE Int. Conf. Systems, Man and Cybernetics*, pages 3040–3045, 2004.

[3] X. Cao and M. Shah. Camera calibration and light source estimation from images with shadows. In *Proc. IEEE CVPR*, pages 918–923, 2005.

[4] R. Collins and Y. Tsin. Calibration of an outdoor active camera system. *In Proc. CVPR*, pages 528–534, 1999.

[5] D. Comaniciu, V. Ramesh, and P. Meer. Real-time tracking of non-rigid objects using mean shift. In *Proc. IEEE CVPR*, pages 142–151, 2000.

[6] O. Faugeras. *Computer Vision: a Geometric Viewpoint*. MIT Press, 1993.

[7] R. I. Hartley. Self-calibration from multiple views with a rotating camera. In *Proc. ECCV*, pages 471–478, 1994.

[8] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

[9] N. Krahnstoever and P. Mendonca. Bayesian autocalibration for surveillance. In *Proc. IEEE ICCV*, pages 1858 – 1865, 2005.

[10] F. Lv, T. Zhao, and R. Nevatia. Self-calibration of a camera from video of a walking human. In *Proc. ICPR*, pages 562–567, 2002.

[11] R. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE J. of Robotics and Automation*, 3(4):323–344, 1987.

[12] Z. Zhang. A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(11):1330–1334, 2000.

0-7695-2521-0/06 $20.00