Estimating Geo-temporal Location of Stationary Cameras Using Shadow Trajectories

Imran N. Junejo[1,⋆] and Hassan Foroosh[2]

[1] INRIA Rennes, France
[2] University of Central Florida, Orlando, U.S.A.

Abstract. Using only shadow trajectories of stationary objects in a scene, we demonstrate that using a set of six or more photographs are sufficient to accurately calibrate the camera. Moreover, we present a novel application where, using only three points from the shadow trajectory of the objects, one can accurately determine the geo-location of the camera, up to a longitude ambiguity, and also the date of image acquisition without using any GPS or other special instruments. We refer to this as "geo-temporal localization". We consider possible cases where ambiguities can be removed if additional information is available. Our method does not require any knowledge of the date or the time when the pictures are taken, and geo-temporal information is recovered directly from the images. We demonstrate the accuracy of our technique for both steps of calibration and geo-temporal localization using synthetic and real data.

1 Introduction

Cameras are everywhere. Groups, individuals or governments mount cameras for various purposes like performing video surveillance, observing natural scenery, or for observing weather patterns. As a result, a global network of thousands of outdoor or indoor cameras currently exists on the internet, which provides a flexible and economical method for information sharing. For such a network, the ability to determine geo-temporal information directly from visual cues has a tremendous potential, in terms of applications, for the field of forensics, intelligence, security, and navigation, to name a few.

The cue that we use for *geo-temporal* localization of the camera, (defined henceforth as *the physical location of the camera (GPS coordinates) and the date of image acquisition*) is the shadow trajectories of two stationary objects during the course of a day. The use of shadow trajectory of a gnomon to measure time in a sundial is reported as early as 1500 BC by Egyptians, which surprisingly requires sophisticated astronomical knowledge [1,2,3]. Shadows have been used in multiple-view geometry in the past to provide information about the shape and the 3-D structure of the scene [4,5], or to recover camera intrinsic and extrinsic parameters [6,7]. Determining the GPS coordinates and the date of the year from shadows in images is a new concept that we introduce in this paper.

⋆ This is part of the author's work at the University of Central Florida, Orlando, U.S.A.

D. Forsyth, P. Torr, and A. Zisserman (Eds.): ECCV 2008, Part I, LNCS 5302, pp. 318–331, 2008.

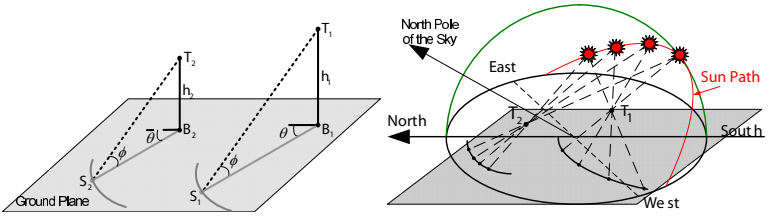


Fig. 1. Two objects T_1 and T_2 casting shadow on the ground plane. The locus of shadow positions over the course of a day is a function of the sun altitude ϕ, the sun azimuth θ and the height h_i of the object.

Our approach is a two step process: auto-calibration and geo-temporal localization. In terms of calibration, the most related work to ours are those of Cao and Foroosh [8] and Lu et al. [9]. The authors in [8] use multiple views of objects and their cast shadows for camera calibration, requiring the objects that cast shadows to be visible in each image and typically from parallel objects perpendicular to the ground plane. Similarly, [9] use line segments formed by corresponding shadow points to estimate the horizon line for camera calibration. Here our contribution is twofold: 1- develop a more flexible solution by relaxing the requirement that shadow-casting objects have to be visible or of particular geometry; 2- provide a more robust solution to estimating the vanishing line of the ground plane by formulating it as a largely overdetermined problem in a manner somewhat similar to [10]. Therefore, our auto-calibration method does not exploit camera motion as in [11,12,13] but rather uses shadows to deduce scene structures that constrain the geometric relations in the image plane [14,15,16].

For geo-temporal localization, recently Jacobs et al.[17] used a database of images collected over a course of a year to learn weather patterns. Using these natural variations, the camera is then geo-located by the correlation of camera images to geo-registered satellite images and also by correlating acquired images with known landmarks/locations. In contrast, the present work is based solely on astronomical geometry and is more flexible, requiring only three shadow points for GPS coordinates estimation. To demonstrate the power of the proposed method we downloaded some images from online traffic surveillance webcams, and determined accurately the geo-locations and the date of acquisition.

Overall two main contributions are made in this paper: *First*, we present a camera calibration method where the horizon line is extracted solely from shadow trajectories without requiring the objects to be visible: we discuss two possible cases (see below). *Second*, we present an innovative application to estimate the GPS coordinates (up to longitude ambiguity) of the location where the images were taken, along with the day of year when the images were taken (up to year ambiguity). In this step, only three points on the shadow trajectories are required, leading to a robust geo-temporal localization. Accordingly, this paper is divided into corresponding sections addressing each issue.

2 Preliminaries and the Setup

Let $\mathbf{T}$ be a 3D stationary point and $\mathbf{B}$ its footprint (i.e. its orthogonal projection) on the ground plane. As depicted in Fig. 1, the locus of shadow positions $\mathbf{S}$ cast by $\mathbf{T}$ on

the ground plane, i.e. the shadow trajectory, is a smooth curve that depends only on the altitude (ϕ) and the azimuth angles (θ) of the sun in the sky and the vertical distance h of the object from its footprint.

Without loss of generality, we take the ground plane as the world plane $z = 0$, and define the x-axis of the world coordinate frame toward the true north point, where the azimuth angle is zero. Therefore, algebraically, the 3D coordinates of the shadow position can be unambiguously specified by their 2D coordinates in the ground plane as

$$\bar{\mathbf{S}}_i = \bar{\mathbf{B}}_i + h_i \cot\phi \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \tag{1}$$

where $\bar{\mathbf{S}}_i = [S_{ix} \;\; S_{iy}]^T$ and $\bar{\mathbf{B}}_i = [B_{ix} \;\; B_{iy}]^T$ are the inhomogeneous coordinates of the shadow position $\mathbf{S}_i$, and the object's footprint $\mathbf{B}_i$ on the ground plane. (1) is based on the assumption that the sun is distant and therefore its rays, e.g. $\mathbf{T}_i\mathbf{S}_i$, are parallel to each other. It follows that the shadows $\mathbf{S}_1$ and $\mathbf{S}_2$ of any two stationary points $\mathbf{T}_1$ and $\mathbf{T}_2$ are related by a rotation-free 2D similarity transformation as $\mathbf{S}_2 \sim \mathbf{H}_s^{12}\mathbf{S}_1$, where

$$\mathbf{H}_s^{12} \sim \begin{bmatrix} h_2/h_1 & 0 & B_{2x} - B_{1x}h_2/h_1 \\ 0 & h_2/h_1 & B_{2y} - B_{1y}h_2/h_1 \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

Note that the above relationship is for world shadow positions and valid for any day time.

Shadow Detection and Tracking: In order to estimate the shadow trajectories, we adopt a semi-automatic approach. For a set of images $\mathbf{S}_I = \{\mathbf{I}_1, \mathbf{I}_2, ...\mathbf{I}_m\}$, we construct a background image $\mathbf{I}$ where each pixel (x, y) contains the brightest pixel value from our set of images $\mathbf{S}_I$. After background subtraction, the most prominent shadow points are detect manually. Mean Shift [18] tracking algorithm is then applied to track the shadow points in the subsequent frames.

3 Recovering the Vanishing Line

Once the vanishing line ($\mathbf{l}_\infty$) is recovered, it can be used together with the vertical vanishing point ($\mathbf{v}_z$), found by fitting lines to vertical directions, to recover the image of the absolute conic (IAC), which is decomposed into the camera calibration matrix $\mathbf{K}$ by using the Cholesky decomposition [19]. For recovering $\mathbf{l}_\infty$, there are two cases that need to be considered:

3.1 When Shadow Casting Object Is Visible

A situation may occur when the footprint, and optionally the shadow casting object itself are visible in the image. An example of this case is the light pole visible in image sequence shown in Figure 6. From here on, the quantities $\mathbf{T}_i, \mathbf{B}_i, \mathbf{S}_i$ and $\mathbf{S}_i'$ refer to the points projected on to the image plane.

Figure 2 illustrates the general setup for this case. The vertical vanishing point is obtained by $\mathbf{v}_z = (\mathbf{T_1} \times \mathbf{B_1}) \times (\mathbf{T_2} \times \mathbf{B_2})$. The estimation of $\mathbf{l}_\infty$ is as follows: at

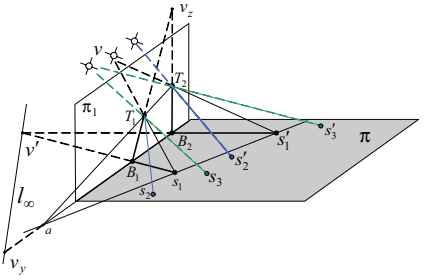


Fig. 2. The setup used for camera calibration and for estimating geo-temporal information

time instance $t = 1$, the sun located at vanishing point $\mathbf{v}_1$ casts shadow of $\mathbf{T}_1$ and $\mathbf{T}_2$ at points $\mathbf{S}_1$ and $\mathbf{S'}_1$, respectively. The sun is a distant object and therefore its rays, $\mathbf{T}_1\mathbf{S}_1$ and $\mathbf{T}_2\mathbf{S'}_1$, are parallel to each other. It then follows that the shadow rays, i.e. $\mathbf{S}_1\mathbf{B}_1$ and $\mathbf{S'}_1\mathbf{B}_2$, are also parallel to each other. These rays intersect at a vanishing point $\mathbf{v}'_1$ on the ground plane. Similarly, for time instance $t = 2$ and $t = 3$, we obtain the vanishing points $\mathbf{v}'_2$ and $\mathbf{v}'_3$, respectively. These vanishing points all lie on the vanishing line of the ground plane on which the shadows are cast, i.e. ${\mathbf{v}'_i}^T\mathbf{l}_\infty = 0$, where $i = 1, 2, \ldots n$ and n is number of instances for which shadow is being observed. Thus a minimum of two observations are required to obtain the $\mathbf{l}_\infty$.

3.2 When Shadow Casting Object Is NOT Visible

This is a more *general* case. The footprint and/or the the shadow casting object point might not always be visible in a video sequence. Figure 7 shows a picture of downtown Washington D.C, where one of the shadow casting objects is the traffic light hanging by a horizontal pole (or a cable). The footprint of this traffic light on the ground plane cannot be determined. In this setup, $\mathbf{l}_\infty$ can not be recovered as described in the previous case. Although, the vertical vanishing point can be obtained by other vertical structures in the scene, not necessarily the shadow-casting structures.

Note: In this case, we use only shadow trajectories to recover the horizon line $\mathbf{l}_\infty$. However, as described in Section 4, we do require to see the shadow casting object (although, not its footprint), in order to perform geo-temporal localization.

Assume now that we have two world points $\mathbf{T}_1$ and $\mathbf{T}_2$ that cast shadows on the ground plane. Given any five imaged shadow positions of the same 3D points ($\mathbf{T}_1$ or $\mathbf{T}_2$), cast at distinct times during one day, one can fit a conic through them, which would meet the line at infinity of the ground plane at two points. These points may be real or imaginary depending on whether the resulting conic is an ellipse, a parabola, or a hyperbola [19]. The two distinct and unique image conics $\mathbf{C}_1$ and $\mathbf{C}_2$ are related by $\mathbf{C}_2 \sim (\mathbf{H}\mathbf{H}_s^{12}\mathbf{H}^{-1})^{-T}\mathbf{C}_1(\mathbf{H}\mathbf{H}_s^{12}\mathbf{H}^{-1})^{-1}$, where $\mathbf{H}$ is the world to image planar homography with respect to the ground plane.

Since the two world conics are similar, owing to the distance of the sun from the observed objects, these two conics generally intersect at *four* points, two of which must lie on the image of the horizon line of the ground plane. The basic idea of conic intersection is illustrated in Fig. 3, for details see Appendix A. It then follows that for any conic $\mathbf{C}_\mu$ through these points of intersection we have (we omit proof here):

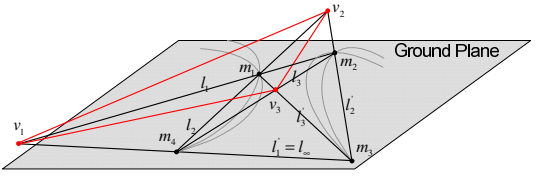


Fig. 3. The two gray conics are fitted by two sets of five distinct shadow positions on the ground plane cast by two world points. Generally, the two conics intersect at four points $\mathbf{m}_i, i = 1, ..., 4$. The diagonal triangle $\Delta\mathbf{v}_1\mathbf{v}_2\mathbf{v}_3$ is self-polar.

Theorem 1. *(Self-Polar Triangle)*
Let $\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3$ *and* $\mathbf{m}_4$ *be four points on the conic locus* $\mathbf{C}_\mu$*, the diagonal triangle of the quadrangle* $\mathbf{m}_1\mathbf{m}_2\mathbf{m}_3\mathbf{m}_4$ *is self-polar for* $\mathbf{C}_\mu$*. Since two of the points lie on* $\mathbf{l}_\infty$*, one of the vertices of* $\Delta\mathbf{v}_1\mathbf{v}_2\mathbf{v}_3$ *also lies on* $\mathbf{l}_\infty$*.*

In simple terms, for the two conics $\mathbf{C}_1$ and $\mathbf{C}_2$, two of the four points of intersection $(\mathbf{m}_1, ..., \mathbf{m}_4)$ lie on the horizon line $\mathbf{l}_\infty$. These points of intersection also define a self-polar triangle $\Delta\mathbf{v}_1\mathbf{v}_2\mathbf{v}_3$, one vertex of which also lies on $\mathbf{l}_\infty$ (cf. Fig.3).

Robust estimation of $\mathbf{l}_\infty$: Five points are required to uniquely define a conic. From two such conics ($\mathbf{C}_1$ and $\mathbf{C}_2$), we get three vanishing points (2 from intersection points and one vertex of the self-polar triangle) that lie on $\mathbf{l}_\infty$. Therefore given six or more corresponding image points on the shadow paths of the two objects, we can get six or more self-polar triangles, from which the horizon line of the ground plane can be recovered. Since, two of intersection points (2 points of the quadrangle) are also on the horizon line of the ground plane, they can be used together with one vertex of each self-polar triangle to recover the horizon line. As an example, Figure 4 illustrates the horizon line fitted to many points obtained during our experimentation. If n such vanishing points are available, we define a matrix M such that:

$$\begin{bmatrix} \mathbf{v}_1'^T \\ \mathbf{v}_2'^T \\ \vdots \\ \mathbf{v}_n'^T \end{bmatrix} \mathbf{l}_\infty = \mathbf{M}\mathbf{l}_\infty = 0$$

where the least square solution to this system of equations is the line at infinity, passing though all the vanishing points $\mathbf{v}_i'^T$.

Note that for $n \geq 6$ corresponding points on shadow paths of two objects, we obtain a total of $\frac{3n!}{(n-5)!5!}$ vanishing points. For instance, with only 10 corresponding shadow points, we would get 756 points on the horizon line. This would allow us to very accurately estimate the horizon line in the presence of noise.

3.3 Camera Calibration

In the previous section, we described a novel technique to recover the vanishing line (or the line at infinity) from using only shadow trajectories. $\mathbf{l}_\infty$, together with the vertical

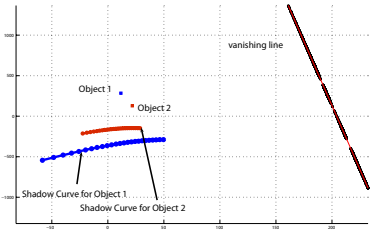


Fig. 4. The horizon line detected from a sequence of self-polar triangles and the intersection of the conics fit on shadow trajectories of two objects

vanishing point $\mathbf{v}_z$, fitted from vertical objects, provide two constraints on the image of the absolute conic (IAC) from pole-polar relationship [19]. Assuming a camera with zero skew, and unit aspect ratio [16,13], the IAC is of the form

$$\boldsymbol{\omega} \sim [\omega_1\ \omega_2\ \omega_3] \sim \begin{bmatrix} 1 & 0 & \omega_{13} \\ 0 & 1 & \omega_{23} \\ \omega_{13} & \omega_{23} & \omega_{33} \end{bmatrix} \tag{3}$$

Let $\mathbf{l}_\infty = \begin{bmatrix} l_x\ l_y\ 1 \end{bmatrix}^T$ and $\mathbf{v}_z = \begin{bmatrix} \mathbf{v}_z^x\ \mathbf{v}_z^y\ 1 \end{bmatrix}^T$, we can solve the linear constraints obtained from the pole-polar relationship to solve for w_{13} and w_{23} in terms of w_{33}:

$$w_{13} = \frac{l_x\, \mathbf{v}_z^{y2} - l_x\, w_{33} - l_y\, \mathbf{v}_z^y\, \mathbf{v}_z^x + \mathbf{v}_z^x}{l_y\, \mathbf{v}_z^y + l_x\, \mathbf{v}_z^x - 1} \tag{4}$$

$$w_{23} = -\frac{l_x\, \mathbf{v}_z^x\, \mathbf{v}_z^y - \mathbf{v}_z^{x\,2} l_y - \mathbf{v}_z^y + l_y\, w_{33}}{l_y\, \mathbf{v}_z^y + l_x\, \mathbf{v}_z^x - 1} \tag{5}$$

The remaining parameter w_{33} is estimated by minimizing the "closeness to the center" constraint [20]:

$$\hat{\omega}_{33} = \arg\min \| \begin{bmatrix} \omega_{13}\ \omega_{23} \end{bmatrix}^T - \mathbf{c} \| \tag{6}$$

where $\mathbf{c}$ is the center of the image, and $\hat{\omega}_{33}$ is the optimal solution for ω_{33}, from which the other two parameters are computed to completely recover the IAC in (3). We use Levenberg-Marquardt to minimize this equation. As most modern cameras have principal points close to the center of the image [16,8], the image center is used as the initial starting point of the minimization process.

4 The Geo-temporal Localization Step

Once we have calibrated the camera, then in order to perform geo-temporal localization, we need to estimate the azimuth and the altitude angle of the sun. At any time of the year, the exact location of the sun can be determined by these two angles. For this it is necessary that the world point casting the shadow on the ground plane be visible in the image.

The earth orbits the sun approximately every 365 days while it also rotates on its axis that extends from the north pole to the south pole every 24 hours. The orbit around the sun is elliptical in shape, which causes it to speed up and slow down as it moves around the sun. The polar axis also tilts up to a maximum angle of about 23.47° with the orbital plane over the course of a year. This tilt causes a change in the angle that the sun makes with the equatorial plane, the so called *declination angle*. Similarly, the globe may be partitioned in several ways. A circle passing through both poles is called a *Meridian*. Another circle that is equidistance from the north and the south pole is called the *Equator*. *Longitude* is the angular distance measured from the prime meridian through Greenwich, England. Similarly, *Latitude* is the angular distance measured from the equator, North $(+ve)$ or South $(-ve)$. Latitude values are important as they define the relationship of a location with the sun. Also, the path of the sun, as seen from the earth, is unique for each latitude, which is the main cue which allows us to geo-locate the camera from only shadow trajectories. Next, we describe the methods for determining these quantities.

Latitude: An overview of the proposed method is shown in Fig. 2. Let $\mathbf{s}_i$, $i = 1, 2, 3$ be the images of the shadow points of a stationary object recorded at different times during the course of a single day. Let $\mathbf{v}_i$ and $\mathbf{v'}_i$, $i = 1, 2, 3$ be the sun and the shadow vanishing points, respectively. For a calibrated camera, the following relations hold for the altitude angle ϕ_i and the azimuth angle θ_i of the sun orientations in the sky, all of which are measured directly in the image domain:

$$\cos\phi_i = \frac{\mathbf{v'}_i^T \boldsymbol{\omega} \mathbf{v}_i}{\sqrt{\mathbf{v'}_i^T \boldsymbol{\omega} \mathbf{v'}_i}\sqrt{\mathbf{v}_i^T \boldsymbol{\omega} \mathbf{v}_i}} \tag{7}$$

$$\sin\phi_i = \frac{\mathbf{v}_z^T \boldsymbol{\omega} \mathbf{v}_i}{\sqrt{\mathbf{v}_z^T \boldsymbol{\omega} \mathbf{v}_z}\sqrt{\mathbf{v}_i^T \boldsymbol{\omega} \mathbf{v}_i}} \tag{8}$$

$$\cos\theta_i = \frac{\mathbf{v}_y^T \boldsymbol{\omega} \mathbf{v'}_i}{\sqrt{\mathbf{v}_y^T \boldsymbol{\omega} \mathbf{v}_y}\sqrt{\mathbf{v'}_i^T \boldsymbol{\omega} \mathbf{v'}_i}} \tag{9}$$

$$\sin\theta_i = \frac{\mathbf{v}_x^T \boldsymbol{\omega} \mathbf{v'}_i}{\sqrt{\mathbf{v}_x^T \boldsymbol{\omega} \mathbf{v}_x}\sqrt{\mathbf{v'}_i^T \boldsymbol{\omega} \mathbf{v'}_i}} \tag{10}$$

Without loss of generality, we choose an arbitrary point on the horizon line as the vanishing point $\mathbf{v}_x$ along the x-axis, and the image point $\mathbf{b}$ of the footprint/bottom as the image of the world origin. The vanishing point $\mathbf{v}_y$ along the y-axis is then given by $\mathbf{v}_y \sim \boldsymbol{\omega}\mathbf{v}_x \times \boldsymbol{\omega}\mathbf{v}_z$.

Let ψ_i be the angles measured clockwise that the shadow points make with the positive x-axis as shown in Fig. 2. We have

$$\cos\psi_i = \frac{\mathbf{v'}_i^T \boldsymbol{\omega} \mathbf{v}_x}{\sqrt{\mathbf{v'}_i^T \boldsymbol{\omega} \mathbf{v'}_i}\sqrt{\mathbf{v}_x^T \boldsymbol{\omega} \mathbf{v}_x}} \tag{11}$$

$$\sin\psi_i = \frac{\mathbf{v'}_i^T \boldsymbol{\omega} \mathbf{v}_y}{\sqrt{\mathbf{v'}_i^T \boldsymbol{\omega} \mathbf{v'}_i}\sqrt{\mathbf{v}_y^T \boldsymbol{\omega} \mathbf{v}_y}} \quad i = 1, 2, 3 \tag{12}$$

Next, we define the following ratios, which are readily derived from spherical coordinates, and also used in sundial construction [1,2,3]:

$$\rho_1 = \frac{\cos\phi_2 \cos\psi_2 - \cos\phi_1 \cos\psi_1}{\sin\phi_2 - \sin\phi_1} \tag{13}$$

$$\rho_2 = \frac{\cos\phi_2 \sin\psi_2 - \cos\phi_1 \sin\psi_1}{\sin\phi_2 - \sin\phi_1} \tag{14}$$

$$\rho_3 = \frac{\cos\phi_2 \cos\psi_2 - \cos\phi_3 \cos\psi_3}{\sin\phi_2 - \sin\phi_3} \tag{15}$$

$$\rho_4 = \frac{\cos\phi_2 \sin\psi_2 - \cos\phi_3 \sin\psi_3}{\sin\phi_2 - \sin\phi_3} \tag{16}$$

For our problem, it is clear from (7)-(12) that these ratios are all determined directly in terms of image quantities. This is possible only because the camera has been calibrated. The angle measured at world origin between the positive y-axis and the ground plane's primary meridian (i.e. the north direction) is then given by

$$\alpha = \tan^{-1}\left(\frac{\rho_1 - \rho_3}{\rho_4 - \rho_2}\right) \tag{17}$$

from which we can determine the GPS latitude of the location where the pictures are taken as

$$\lambda = \tan^{-1}(\rho_1 \cos\alpha + \rho_2 \sin\alpha) \tag{18}$$

For n shadow points, we obtain a total of $\frac{n!}{(n-3)!3!}$ estimations of latitude(λ). In the presence of noise, this leads to a very robust estimation of λ.

Day Number: Once the latitude is determined from (18), we can also determine the exact day when the images are taken. For this purpose, let δ denote the declination angle (positive in the summer). Let also $\hbar$ denote the hour angle for a given image, i.e. the angle the earth needs to rotate to bring the meridian of that location to solar noon, where each hour time corresponds to $\frac{\pi}{12}$ radians, and the solar noon is when the sun is due south with maximum altitude. Then these angles are given in terms of the latitude λ, the sun's altitude ϕ and its azimuth θ by

$$\sin\hbar\cos\delta - \cos\phi\sin\theta = 0 \tag{19}$$

$$\cos\delta\cos\lambda\cos\hbar + \sin\delta\sin\lambda - \sin\phi = 0 \tag{20}$$

Again, note that the above system of equations depend only on image quantities defined in (7)-(12). Upon finding the declination and the hour angles by solving the above equations, the exact day of the year when the pictures are taken can be found by

$$N = \frac{365}{2\pi}\sin^{-1}\left(\frac{\delta}{\delta_m}\right) - N_o \tag{21}$$

where N is the day number of the date, with January 1^{st} taken as $N = 1$, and February assumed of 28 days, $\delta_m \simeq 0.408$ is the maximum absolute declination angle of earth in radians, and $N_o = 284$ corresponds to the number of days from the first equinox to January 1^{st}.

Longitude: Unfortunately, unlike latitude, the longitude cannot be determined directly from observing shadows. The longitude can only be determined either by spatial or temporal correlation. For instance, if we know that the pictures are taken in a particular state or a country or a region in the world, then we only need to perform a one-dimensional search along the latitude determined by (18) to find also the longitude and hence the GPS coordinates. Alternatively, the longitude may be determined by temporal correlation. For instance, suppose we have a few frames from a video stream of a live webcam with unknown location. Then they can be temporally correlated with our local time, in which case the difference in hour angles can be used to determine the longitude.

For this purpose, let $\hbar_l$ and γ_l be our own local hour angle and longitude at the time of receiving the live pictures. Then the GPS longitude of the location where the pictures are taken is given by

$$\gamma = \gamma_l + (\hbar - \hbar_l) \tag{22}$$

Therefore, by using only three shadow points, compared to 5 required for the camera calibration, we are able to determine the geo-location up to longitude ambiguity, and specify the day of the year when the images were taken up to, of course, year ambiguity. The key observation that allows us to achieve this is the fact that a calibrated camera performs as a direction tensor, capable of measuring direction of rays and hence angles, and that the latitude and the day of the year are determined simply by measuring angles in images.

In the next section, we validate our method and evaluate the accuracy of both self-calibration and geo-temporal localization steps using synthetic and real data.

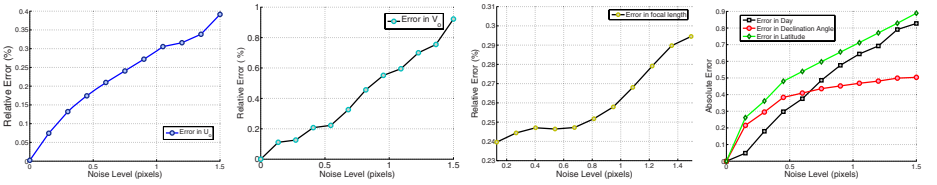
Algorithm - Geo-Temporal Localization
-Input: Shadow points of at least two objects

- Obtain the vertical vanishing point $\mathbf{v}_z$.
- Estimate the horizon line $\mathbf{l}_\infty$.
  - If the shadow casting object are visible: use the method described in Section 3.1 for $\mathbf{l}_\infty$ estimation.
  - Else, fit a conic to shadow trajectory of each object, and compute conic intersections (Section 3.2). Fit a line through the intersection points for a robust estimation of $\mathbf{l}_\infty$.
- Perform camera calibration, as described in Section 3.3.
- Estimate the altitude and the azimuth angles, eqs (7)-(10). Estimate the ratios (13)-(16) to estimate the latitude λ and N.
- If time of image acquisition is know, estimate the longitude γ.

5 Experimental Results

We rigorously tested and validated our method on synthetic as well as real data sequences for both self-calibration and geo-temporal localization steps. Results are described below.

Synthetic Data: Two vertical objects of different heights were randomly placed on the ground plane. Using the online available version of SunAngle Software [21], we generated altitude and azimuth angles for the sun corresponding to our own geo-location with



(a) (b) (c) (d)

Fig. 5. Performance averaged over 1000 independent trials: (a) & (b) relative error in the coordinates of the principal point (u_o, v_o), (c) the relative error in the focal length f. (d) Result for average error in latitude, solar declination angle, and day of the year.

Fig. 6. Few of the images taken from one of the live webcams in downtown Washington D.C. The two objects that cast shadows on the ground are shown in red and blue, respectively. Shadows move to the left of the images as time progresses.

latitude $28.51°$ *(we omit the longitude information to maintain our anonymity leaving you with one dimensional ambiguity)*. The data was generated for the 315^{th} day of the year i.e. the 11^{th} of November 2006 from `10:00am` to `2:00pm`. The solar declination angle for that time period is $-17.49°$. The vertical objects and the shadow points were projected by a synthetic camera with a focal length of $f = 1000$, the principal point at $(u_o, v_o) = (320, 240)$, unit aspect ratio, and zero skew.

In order to test resilience of the proposed self-calibration method to noise, we gradually added Gaussian noise of zero mean and standard deviation of up to 1.5 pixels to the projected points. The estimated parameters were then compared with the ground truth values mentioned above. For each noise level, we performed 1000 independent trials. The final averaged results for calibration parameters are shown in Figure 5. Note that, as explained in [15], the relative difference with respect to the focal length is a more geometrically meaningful error measure. Therefore, relative error of f, u_o and v_o were measured w.r.t f while varying the noise from 0.1 to 1.5 pixels. As shown in the figure, errors increase almost linearly with the increase of noise in the projected points. For the noise of 1.5 pixels, the error is found to be less than 0.3% for f, less than 0.5% for u_o and less than 1% for v_o.

Averaged results for latitude, solar declination angle, and the day of the year are shown in Figure 5(d). The error is found to be less than 0.9%. For a maximum noise level of 1.5 pixels, the estimated latitude is $28.21°$, the declination angle is $-17.93°$, and the day of the year is found to be 314.52.

Real Data: Several experiments on two separate data sets are reported below for demonstrating the power of the proposed method. In the first set, 11 images were captured live from downtown Washington D.C. area, using one of the webcams available online at `http://trafficland.com/`. As shown in Figure 6, a lamp post and a traffic

Table 1. Results for 11 sets of 10-image combinations. Mean value and standard deviation for latitude is found to be $(38.743°, 3.57)$, $(-16.43°, 1.11)$ for the declination angle, and $(329.95, 2.28)$ for the estimated number of the day.

	$Comb_1$	$Comb_2$	$Comb_3$	$Comb_4$	$Comb_5$	$Comb_6$	$Comb_7$	$Comb_8$	$Comb_9$	$Comb_{10}$	$Comb_{11}$
Latitude	33.73	35.70	37.03	36.1	35.72	38.21	39.23	45.78	41.84	40.88	41.96
Declination	-14.47	-15.78	-15.93	-16.54	-17.25	-16	-16.70	-18.94	-15.87	-16.99	-16.24
Day #	328.64	332.26	331.09	326.87	330.15	331.37	331.32	332.56	326.81	331.72	326.72

Fig. 7. Few of the images in the second data set that were temporally correlated with our local time, taken also from one of the live webcams in Washington D.C. The objects that cast shadows on the ground are highlighted. Shadows move to the left of the images as time progresses.

light were used as two objects casting shadows on the road. The shadow points are highlighted by colored circles in the figure. The calibration parameters were estimated as $\mathbf{K} = \begin{bmatrix} 700.36 & 0 & 172 \\ 0 & 700.36 & 124 \\ 0 & 0 & 1 \end{bmatrix}$.

Since we had more than the required minimum number of shadow locations over time, in order to make the estimation more robust to noise, we took all possible combinations of the available points and averaged the results. For this first data set the images were captured on the 15^{th} November at latitude $38.53°$ and longitude $77.02°$. We estimated the latitude as $38.74°$, the day number as 329.95 and the solar declination angle as $-16.43°$ compared to the actual day of 319, and the declination angle of $-18.62°$. The small errors can be attributed to many factors e.g. noise, non-linear distortions and errors in the extracted features in low-resolution images of 320×240. Despite all these factors, the experiment indicates that the proposed method provides good results.

In order to evaluate the uncertainty associated with our estimation, we then divided this data set into 11 sets of 10-image combinations, i.e. in each combination we left one image out. We repeated the experiment for each combination and calculated the mean and the standard deviation of the estimated unknown parameters. Results are shown in Table 5. The low standard deviations can be interpreted as small uncertainty, indicating that our method is consistently providing reliable results.

A second data set is shown in Figure 7. The ground truth for this data set was as follows: longitude $77.02°$, latitude $38.53°$, day number of 331, and the declination of $-21.8°$. For this data set we assumed that the data was downloaded in real-time and hence was temporally correlated with our local time. We estimated the longitude as $78.761°$, the latitude as $37.79°$, the day number as 323.07, and the declination angle as $-17.29°$.

6 Conclusion

We propose a method based entirely on computer vision to determine the geo-location of the camera up to longitude ambiguity, without using any GPS or other instruments, and by solely relying on imaged shadows as cues. We also describe situations where longitude ambiguity can be removed by either temporal or spatial cross-correlation. Moreover, we determine the date when the pictures are taken without using any prior information. Unlike shadow-based calibration methods such as [6,7], this step does not require the objects themselves to be seen in the images.

References

1. Herbert, A.: Sundials Old and New. Methuen & Co. Ltd. (1967)
2. III, F.W.S.: A three-point sundial construction. Bulletin of the British Sundial Society 94, 22–29 (1994)
3. Waugh, A.: Sundials: Their Theory and Construction. Dover Publications, Inc. (1973) ISBN 0-486-22947-5
4. Bouguet, J., Perona, P.: 3D photography on your desk. In: Proc. ICCV, pp. 43–50 (1998)
5. Caspi, Y., Werman, M.: Vertical parallax from moving shadows. In: Proc. CVPR, pp. 2309–2315 (2006)
6. Antone, M., Bosse, M.: Calibration of outdoor cameras from cast shadows. In: Proc. IEEE Int. Conf. Systems, Man and Cybernetics, pp. 3040–3045 (2004)
7. Cao, X., Shah, M.: Camera calibration and light source estimation from images with shadows. In: Proc. IEEE CVPR, pp. 918–923 (2005)
8. Cao, X., Foroosh, H.: Camera calibration and light source orientation from solar shadows. Journal of Computer Vision and Image Understanding (CVIU) 105, 60–72 (2006)
9. Lu, F., Shen, Y., Cao, X., Foroosh, H.: Camera calibration from two shadow trajectories. In: Proc. ICPR, pp. 1–4 (2005)
10. Heikkila, J.: Geometric camera calibration using circular control points. IEEE Trans. Pattern Anal. Mach. Intell. 22, 1066–1077 (2000)
11. Hartley, R.I.: Self-calibration of stationary cameras. Int. J. Comput. Vision 22, 5–23 (1997)
12. Heyden, A., Astrom, K.: Euclidean reconstruction from image sequences with varying and unknown focal length and principal point. In: Proc. IEEE CVPR, pp. 438–443 (1997)
13. Pollefeys, M., Koch, R., Gool, L.V.: Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters. Int. J. Comput. Vision 32, 7–25 (1999)
14. Liebowitz, D., Zisserman, A.: Combining scene and auto-calibration constraints. In: Proc. IEEE ICCV, pp. 293–300 (1999)
15. Triggs, B.: Autocalibration from planar scenes. In: Proc. ECCV, pp. 89–105 (1998)
16. Zhang, Z.: A flexible new technique for camera calibration. IEEE Trans. Pattern Anal. Mach. Intell. 22, 1330–1334 (2000)
17. Jacobs, N., Satkin, S., Roman, N., Speyer, R., Pless, R.: Geolocating static cameras. In: Proc. of ICCV, pp. 469–476 (2007)
18. Comaniciu, D., Ramesh, V., Meer, P.: Kernel-based object tracking. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) 25, 564–575 (2003)
19. Hartley, R.I., Zisserman, A.: Multiple View Geometry in Computer Vision, 2nd edn. Cambridge University Press, Cambridge (2004)
20. Junejo, I., Foroosh, H.: Dissecting the image of the absolute conic. In: 5th IEEE International Conference on Advanced Video and Signal-based Surveillance (AVSS) (2006)
21. Gronbeck, C.: Sunangle software, `www.susdesign.com/sunangle/`
22. Semple, J.G., Kneebone, G.T.: Algebraic Projective Geometry. Oxford Classic Texts in the Physical Sciences (1979)

APPENDIX

A Computing Conic Intersection

Two conics always intersect at four points, which may be real or imaginary. All conics passing through the four points of intersection can be written as

$$\mathbf{C}_\mu \sim \mathbf{C}_1 + \mu \mathbf{C}_2. \tag{A-1}$$

Equation (A-1) defines a pencil of conics parameterized by μ, where all the conics in the pencil intersect at the same four points $\mathbf{m}_i, i = 1, ..., 4$. Four such points such that no three of them are collinear also give rise to what is known as the *complete quadrangle*.

It can be shown that in this pencil at most three conics are not full rank. For this purpose note that any such degenerate conic should satisfy

$$\det(\mathbf{C}_\mu) = \det(\mathbf{C}_1 + \mu \mathbf{C}_2) = 0. \tag{A-2}$$

It can then be readily verified that (A-2) is a cubic equation in terms of μ. Therefore upon solving (A-2), we obtain at most three distinct values $\mu_i, i = 1, ..., 3$, which provide the three corresponding degenerate conics

$$\mathbf{C}_{\mu_i} \sim \mathbf{C}_1 + \mu_i \mathbf{C}_2, \quad i = 1, ..., 3. \tag{A-3}$$

In the general case (i.e. when the three parameters $\mu_i, i = 1, ..., 3$ are distinct), the three degenerate conics are of rank 2, and therefore can be written as

$$\mathbf{C}_{\mu_i} \sim \mathbf{l}_i \mathbf{l'}_i^T + \mathbf{l}'_i \mathbf{l}_i^T, \quad i = 1, ..., 3, \tag{A-4}$$

where $\mathbf{l}_i$ and $\mathbf{l}'_i$ are three pairs of lines as shown in Fig.3.

Now, let $\mathbf{C}^*_{\mu_i}$ be the adjoint matrix of $\mathbf{C}_{\mu_i}$. It then follows from (A-4) that

$$\mathbf{C}^*_{\mu_i} \mathbf{l}_i = \mathbf{C}^*_{\mu_i} \mathbf{l}'_i = 0, \quad i = 1, ..., 3, \tag{A-5}$$

which yields (by using the property that the cofactor matrix is related to the way matrices distribute with respect to the cross product [19])

$$\mathbf{C}^*_{\mu_i} \mathbf{l}_i \times \mathbf{C}^*_{\mu_i} \mathbf{l}'_i = \mathbf{C}_{\mu_i} (\mathbf{l}_i \times \mathbf{l}'_i) = 0, \quad i = 1, ..., 3. \tag{A-6}$$

In other words, the intersection point $\mathbf{v}_i$ of the pair of lines, $\mathbf{l}_i$ and $\mathbf{l}'_i$, is given by the right null space of $\mathbf{C}_{\mu_i}$. Therefore, in practice, it can be found as the eigenvector corresponding to the smallest eigenvalue of the degenerate conic $\mathbf{C}_{\mu_i}$. The triangle formed by the three vertices $\mathbf{v}_1 \mathbf{v}_2$ and $\mathbf{v}_3$ is known as the *diagonal triangle* of the quadrangle [22].

Next, we verify that for any conic $\mathbf{C}_\mu$ in the pencil

$$(\mathbf{l}_i \times \mathbf{l}'_i)^T \mathbf{C}_\mu (\mathbf{l}_j \times \mathbf{l}'_j) = 0, \quad i \neq j, \quad i, j = 1, ..., 3 \tag{A-7}$$

This means that any pair of right null vectors of the degenerate conics $\mathbf{C}_{\mu_i}, i = 1, ..., 3$ are conjugate with respect to all conics in the pencil. In other words, their intersections form the vertices of a self-polar triangle with respect to all the conics in the pencil.

To obtain the intersection points of the two shadow conics, we use the fact that all the conics in the pencil intersect at the same four points. Therefore, the intersection points can also be found as the intersection of the lines $\mathbf{l}_i$ and $\mathbf{l}'_i$ with the lines $\mathbf{l}_j$ and $\mathbf{l}'_j$ $(i \neq j)$. The lines $\mathbf{l}_i$ and $\mathbf{l}'_i$ can be simply found by solving

$$\mathbf{C}_{\mu_i} \sim \mathbf{l}_i \mathbf{l'}_i^T + \mathbf{l}'_i \mathbf{l}_i^T \tag{A-8}$$

Equation (A-8) provides 4 constraints on $\mathbf{l}_i$ and $\mathbf{l}'_i$ (5 due to symmetry minus 1 for rank deficiency). In practice it leads to two quadratic equations on the four parameters of the two lines, which can be readily solved. The solution, of course, has a twofold ambiguity due to the quadratic orders, which is readily resolved by the fact that

$$\mathbf{l}_i \times \mathbf{l}'_i \sim \text{null}(\mathbf{C}_{\mu_i}) \tag{A-9}$$

The process can be repeated for $\mathbf{l}_j$ and $\mathbf{l}'_j$, and the intersections of the lines between the two sets would then provide the four intersection points of the shadow conics.