GPS coordinates estimation and camera calibration from solar shadows

Imran N. Junejo [a,*], Hassan Foroosh [b]

[a] Department of Computer Science, University of Sharjah, P.O. Box 27272, Sharjah, United Arab Emirates
[b] School of Electrical Engineering and Computer Science, University of Central Florida, Orlando, FL 32816, USA

ARTICLE INFO

ABSTRACT

In this paper, we discuss the issue of camera parameter estimation (intrinsic and extrinsic parameters), along with estimation of the geo-location of the camera by using only the shadow trajectories. By observing stationary objects over a period of time, it is shown that only six points on the trajectories formed by tracking the shadows of the objects are sufficient to estimate the horizon line of the ground plane. This line is used along with the extracted vertical vanishing point to calibrate the stationary camera. The method requires as few as two shadow casting objects in the scene and a set of six or more points on the shadow trajectories of these objects. Once camera intrinsic parameters are recovered, we present a novel application where one can accurately determine the geo-location of the camera up to a longitude ambiguity using only three points from these shadow trajectories without using any GPS or other special instruments. We consider possible cases where this ambiguity can also be removed if additional information is available. Our method does not require any knowledge of the date or the time when the images are taken, and recovers the date of acquisition 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.

© 2010 Elsevier Inc. All rights reserved.

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 [6], and navigation [35,15], 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 requires surprisingly sophisticated astronomical knowledge [20,22,36]. 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 [5,10], or to recover camera intrinsic and extrinsic parameters [2,8]. Determining the GPS coordinates and the date of the year

from shadows in images is a new concept that we introduce in this paper.

Our approach is a two step process: auto-calibration and geo-temporal localization. Camera auto-calibration is a vast area of research and it is beyond the scope of the current work to summarize the related work, we refer the readers to [18] for a review of the techniques existing in this area. Briefly put, starting from the initial work using known configuration of points (or calibration rigs) in 2D or 3D [34,38,39,31], the camera calibration techniques have evolved to a stage where calibration objects are no longer required and rely only on scene information or point-correspondences [13,16,33,9,27,1]. The proposed method is based on this later category of calibration techniques and the most related work to ours are those of Cao and Foroosh [7] and Lu et al. [26]. The authors in [7] use multiple views of objects and their cast shadows for camera calibration, requiring the object's that cast shadows to be visible in each image and typically from parallel objects perpendicular to the ground plane. Similarly, [26] 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 and (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 [19]. Therefore, our auto-calibration method does not exploit camera motion as in [17,21,28] but rather uses shadows to deduce scene structures

* Corresponding author.
E-mail addresses: imran.junejo@inria.fr, ijunejo@cs.ucf.edu (I.N. Junejo).

that constrain the geometric relations in the image plane [25,33,38].

For geo-temporal localization, recently Jacobs et al.[23] use 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. Recently [32] present a method where color changes in a scene observed for an extended period of time are used to learn the geometry of the scene. Given the date and the UTC time-stamp for each frame, they are able to compute the geo-location of the camera. In contrast to these works, the proposed 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 estimated 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 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.

The rest of the paper is organized as follows: A brief introduction to the process of shadow formation and the projective camera are given in the next section. In order to perform camera calibration, we need to recover the horizon line (or the line at infinity) of the ground plane. We propose two such methods, varying in their applicability, in Section 3. Due to noise in the images, or uneven ground plane, the estimation of the horizon line might not be very accurate. To deal with this situation, a robust solution is proposed in Section 4 and the camera calibration method is described in Section 5. The main task, i.e. the estimation of the geo-temporal location of the camera, is described in Section 6. Often, we have to deal with very few images; we present a solution in Section 7 where only two images can be used to estimate the GPS coordinates of the camera. We rigorously test the proposed method on synthetic data and on several real datasets, as shown in Section 8. Encouraging results indicate the practicality of the proposed method.

2. Preliminaries and the setup

2.1. Camera model

The projection of a 3D scene point $\mathbf{X} \sim [X \ Y \ Z \ 1]^T$ onto a point in the image plane $\mathbf{x} \sim [x \ y \ 1]^T$, for a perspective camera can be modeled by the central projection equation:
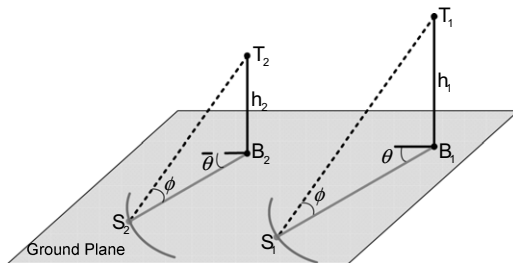
$$\mathbf{x} \sim \underbrace{\mathbf{K}[\mathbf{R} \ | -\mathbf{R}\mathbf{C}]}_{\mathbf{P}}\mathbf{X}, \quad \mathbf{K} = \begin{bmatrix} \lambda f & \gamma & u_o \\ 0 & f & v_o \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

where $\sim$ indicates equality up to a non-zero scale factor and $\mathbf{C} = [C_x \ C_y \ C_z]^T$ represents the position of the camera center. Here $\mathbf{R} = \mathbf{R}_x\mathbf{R}_y\mathbf{R}_z = [\mathbf{r}_1 \ \mathbf{r}_2 \ \mathbf{r}_3]$ is the rotation matrix and $-\mathbf{R}\mathbf{C}$ is the relative translation between the world origin and the camera center. The upper triangular 3×3 matrix $\mathbf{K}$ encodes the five intrinsic camera parameters: focal length f, aspect ratio λ, skew γ and the principal point at (u_o, v_o) [12,18,24].

2.2. Shadow formation

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 is a smooth curve that depends only on the altitude and the azimuth angles of the sun in the sky and the vertical distance h of the object from its footprint. This geometric configuration is rather interesting, since the object point $\mathbf{T}$ together with the ground plane act as an artificial pinhole camera, where the camera projection center is the object point, the image plane is the ground plane, the focal length is the vertical distance h, and the principal point is the footprint $\mathbf{B}$.

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

$$\overline{\mathbf{S}}_i = \overline{\mathbf{B}}_i + h_i \cot\phi \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \quad (2)$$
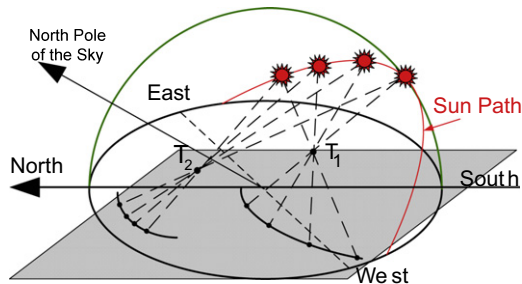
where $\overline{\mathbf{S}}_i = [S_{ix} \ S_{iy}]^T$ and $\overline{\mathbf{B}}_i = [B_{ix} \ B_{iy}]^T$ are the inhomogeneous coordinates of the shadow position $\mathbf{S}_i$, and the object's footprint $\mathbf{B}_i$ on the ground plane, ϕ is sun altitude, and θ the sun azimuth. Eq. (2) is based 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

$$\mathbf{H}_s^{12} \sim \begin{bmatrix} h_2/h_1 & 0 & B_{2x} - B_{1x}h_2/h_1 \\ 0 & h_2/h_1 & B_{2y} - B_{1y}h_2/h_1 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

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

2.3. Shadow detection and tracking

Although many techniques can be adopted to successfully track shadows and thus obtain shadow trajectories under varying light-

Fig. 1. Two objects T_1 and T_2 casting shadows 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.

ing conditions [37,4], we adopt a very simple and practical semi-automatic approach. For a set of images $\mathbf{S}_l = \{\mathbf{I}_1, \mathbf{I}_2, \ldots, \mathbf{I}_m\}$, we construct a background image $\mathbf{I}$ where each pixel (x, y) contains the brightest pixel value from our set of images $\mathbf{S}_l$. After background subtraction, the most prominent shadow points are detect manually. Mean shift tracking algorithm [11] is then applied to track the shadow points in the subsequent frames. Although fairly accurate in tracking, it fails due to certain weather conditions, such as when clouds regularly block the sunlight. These conditions then warrant a manual re-localization of the shadow points.

3. Recovering the vanishing line

The goal in the calibration step is to recover the vanishing line of the ground plane from the shadow trajectories. Once the vanishing line ($\mathbf{l}_\infty$) is recovered, it is used together with the vertical vanishing point, found by fitting lines to vertical directions, to recover the image of the absolute conic (IAC). There are two cases that need to be considered:

3.1. When shadow casting object is visible

This case requires that the bottom point of the shadow casting object be visible in the image, generally selected manually by a user. An example of this case is the light pole visible in image sequence shown in Fig. 12. Fig. 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 time instance $t = 1$, the sun located at vanishing point $\mathbf{a}_1$ 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 the vanishing point $\mathbf{v}_x^1$ on the ground plane. Similarly, for time instance $t = 2$ and $t = 3$, we obtain the vanishing points $\mathbf{v}_x^2$ and $\mathbf{v}_x^3$, respectively. These vanishing points all lie on the vanishing line of the ground plane on which the shadows are cast, i.e. $\mathbf{v}_x^{i\mathsf{T}}\mathbf{l}_\infty = 0$, 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 of at least two vertical objects 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 shadow casting object point might not always be visible in a video sequence. Fig. 13 shows a picture of downtown Washington, DC, 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
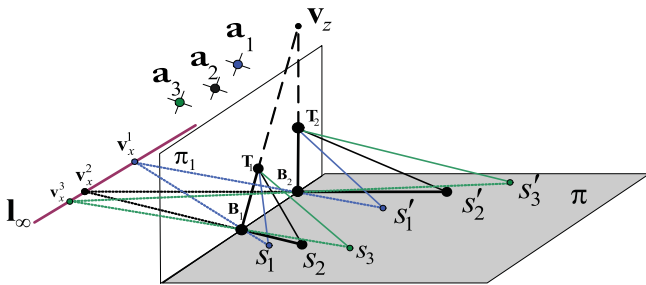
the ground plane cannot be determined. In this setup, $\mathbf{l}_\infty$ can not be recovered as described in the previous case. However, the vertical vanishing point can be obtained by other vertical structures in the scene, not necessarily the shadow-casting structures.

Note: In this case, we use only shadow trajectories to recover the horizon line $\mathbf{l}_\infty$. However, as described in Section 6, 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 1 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 [18]. The two distinct and unique image conics $\mathbf{C}_1$ and $\mathbf{C}_2$ are related by $\mathbf{C}_2 \sim \left(\mathbf{H}\mathbf{H}_s^{12}\mathbf{H}^{-1}\right)^{-T} \mathbf{C}_1 \left(\mathbf{H}\mathbf{H}_s^{12}\mathbf{H}^{-1}\right)^{-1}$, 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, and we describe it in the following subsections for the sake of completeness.

3.2.1. Computing conic intersections

We now present the method for computing conic intersections and expand on its relation to the recovery of the vanishing line $\mathbf{l}_\infty$.

All conics passing through the four points of intersection can be written as

$$\mathbf{C}_\mu \sim \mathbf{C}_1 + \mu\mathbf{C}_2 \tag{4}$$

Eq. (4) defines a pencil of conics parameterized by μ, where all the conics in the pencil intersect at the same four points $\mathbf{m}_i$, $i = 1, \ldots, 4$. 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

$$\det(\mathbf{C}_\mu) = \det(\mathbf{C}_1 + \mu\mathbf{C}_2) = 0 \tag{5}$$

It can then be readily verified that (5) is a cubic equation in terms of μ. Therefore upon solving (5), we obtain at most three distinct values μ_i, $i = 1, \ldots, 3$, which provide the three corresponding degenerate conics

$$\mathbf{C}_{\mu_i} \sim \mathbf{C}_1 + \mu_i\mathbf{C}_2, \quad i = 1, \ldots, 3 \tag{6}$$

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

Fig. 2. The setup used for estimating geo-temporal information when the bottom and the top locations of the object are visible. Three instances are shown above. The top of the object is denoted by $\mathbf{T}_1$ and $\mathbf{T}_2$, and their corresponding bottoms as $\mathbf{B}_1$ and $\mathbf{B}_2$, respectively. The shadows cast by object **1** are denoted by $\mathbf{S}_i$ while that of object **2** at the same time instance is denoted by $\mathbf{S}_i'$. The sun location at these time instances is denoted by $\mathbf{a}_i$.
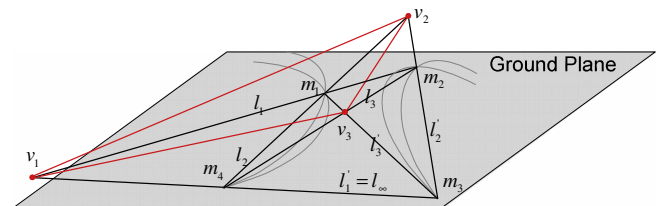

Fig. 3. The two gray conics are fitted by two sets of five distinct shadow positions on the ground plane cast by two world points. Generally, the two conics intersect at four points $\mathbf{m}_i$, $i = 1, \ldots, 4$. The four points form a quadrangle inscribed to any one of the gray conics. The diagonal triangle $\triangle\mathbf{v}_1\mathbf{v}_2\mathbf{v}_3$ is self-polar [30].

$$\mathbf{C}_{\mu_i} \sim \mathbf{l}_i \mathbf{l}_i'^{T} + \mathbf{l}_i' \mathbf{l}_i^{T}, \quad i = 1, \ldots, 3 \tag{7}$$

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 (7) that

$$\mathbf{C}_{\mu_i}^* \mathbf{l}_i = \mathbf{C}_{\mu_i}^* \mathbf{l}_i' = 0, \quad i = 1, \ldots, 3 \tag{8}$$

which yields (by using the property that the cofactor matrix is related to the way matrices distribute with respect to the cross product [18])

$$\mathbf{C}_{\mu_i}^* \mathbf{l}_i \times \mathbf{C}_{\mu_i}^* \mathbf{l}_i' = \mathbf{C}_{\mu_i}\left(\mathbf{l}_i \times \mathbf{l}_i'\right) = 0, \quad i = 1, \ldots, 3 \tag{9}$$

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 [30].

Theorem 3.1 (*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 lie on* $\mathbf{l}_\infty$.

This theorem follows directly from the projective geometry and we omit the proof here. Thus the triangle $\Delta \mathbf{v}_1 \mathbf{v}_2 \mathbf{v}_3$ is the diagonal triangle of the quadrangle composed of points $\mathbf{m}_i$, $i = 1, \ldots, 4$ inscribed in a conic. There also exists a harmonic relationship between any two sides of the quadrangle and $\mathbf{v}_i$ of $\Delta \mathbf{v}_1 \mathbf{v}_2 \mathbf{v}_3$ that meets that side. Exploring this harmonic relationship for obtaining further constraints is the topic of our future research.

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

$$\left(\mathbf{l}_i \times \mathbf{l}_i'\right)^{T} \mathbf{C}_\mu \left(\mathbf{l}_j \times \mathbf{l}_j'\right) = 0, \quad i \neq j, \ i, j = 1, \ldots, 3 \tag{10}$$

This means that any pair of right null vectors of the degenerate conics $\mathbf{C}_{\mu_i}$, $i = 1, \ldots, 3$ 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

$$\mathbf{C}_{\mu_i} \sim \mathbf{l}_i \mathbf{l}_i'^{T} + \mathbf{l}_i' \mathbf{l}_i^{T} \tag{11}$$

Eq. (11) provides four 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

$$\mathbf{l}_i \times \mathbf{l}_i' \sim \text{null}(\mathbf{C}_{\mu_i}) \tag{12}$$

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.

4. Robust estimation of $\mathbf{l}_\infty$

The shadow cast on the ground plane might not be very accurately localized. This is due to the difficulty of the problem, occurring mainly as a result of irregularities in the road, or the shadow not being very sharp due to a cloudy weather, etc. Therefore some scheme needs to be adopted to minimize the influence of outliers

and noise in the true data points so that accurate results may be obtained.

As shown in Fig. 3 and discussed above, two of the four points of intersection are at infinity (without loss of generality m_3 and m_4), and therefore one of the vertices, $\mathbf{v}_1$ of the self-polar triangle $\Delta \mathbf{v}_1 \mathbf{v}_2 \mathbf{v}_3$ must also be a vanishing point. These three points m_3, m_4 and $\mathbf{v}_1$ lie on the horizon line of the ground plane, denoted by $\mathbf{l}_\infty$ in the figure. Therefore given six or more corresponding image points on the shadow paths of any two objects, we can get six or more self-polar triangles. The computed *points of intersection* along with the vertices of the self-polar triangle are used to recover the horizon line of the ground plane. Fig. 4 illustrates the horizon line fitted to many points obtained through experiments on synthetic data, to be described shortly. Therefore, the system of overdetermined set of equations needed to solve for $\mathbf{l}_\infty$ can be given as:

$$\Phi^{T} \mathbf{l}_\infty = 0 \tag{13}$$

where Φ is a matrix containing the estimated vanishing points. Note that for $n \geqslant 6$ 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. Φ is therefore a $\frac{3n!}{(n-5)!5!} \times 3$ matrix and we have to robustly estimate $\mathbf{l}_\infty$.

The main goals of robust statistics is to recover the best structure that fits the majority of the model while rejecting the outliers. We need to recover the best $\mathbf{l}_\infty$ such that $\mathbf{K}$ is closest to the actual calibration matrix. The popular standard least squares (LS) estimation, which minimizes the Euclidean norm of the residuals, is extremely sensitive to outliers, i.e. it has a breakdown point of zero. Total Least Squares (TLS) method, on the other hand, minimizes the Frobenius norm. Given an over-determined system of equations, TLS problem is to find the smallest perturbation to the data and the observation matrix to make the system of equations compatible. A suitable function also needs to be selected that is less forgiving to outliers, one such example is the *truncated quadratic* [3], commonly used in computer vision (cf. Fig. 5). The errors are weighted up to a fixed threshold, but beyond that, errors receive constant penalty. Thus the influence of outliers goes to zero beyond the threshold.

In order to remove the outlier influence, we use the truncated Rayleigh quotient. The quotients are estimated as:
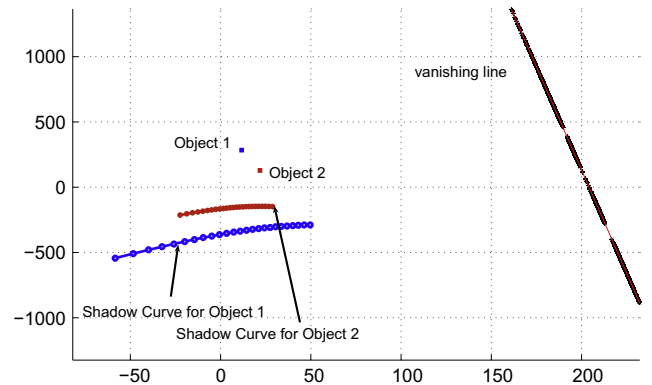

Fig. 4. Experiments with synthetic data: Shadows of two objects are projected on the ground plane, plotted in blue and red. More than 20 points are sampled on the shadow trajectory of each of the object. Using 5 points at a time to fit a conic to these shadow points, the horizon line is recovered by using the points of intersection of these conics and the vertices of the self-polar triangles, as described in the text.
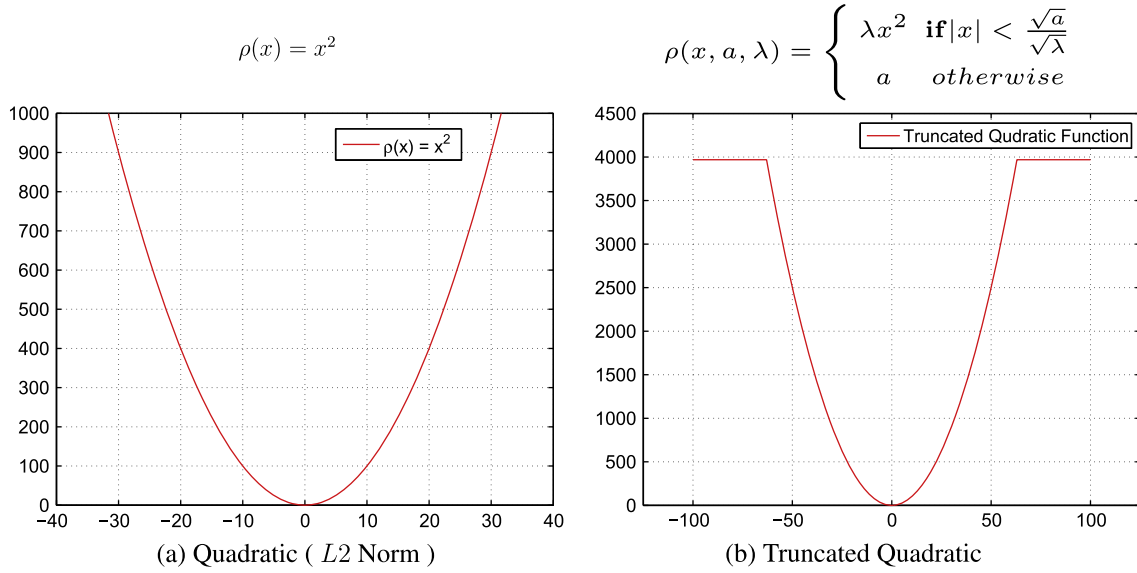
$$\rho(x) = x^2$$

$$\rho(x, a, \lambda) = \begin{cases} \lambda x^2 & \mathbf{if}\, |x| < \frac{\sqrt{a}}{\sqrt{\lambda}} \\ a & otherwise \end{cases}$$

(a) Quadratic ($L2$ Norm)

(b) Truncated Quadratic

Fig. 5. Two commonly used minimization cost functions.

$$\rho(\mathbf{l}_\infty) = \sum^N \frac{\mathbf{q}^{\mathbf{T}} A \mathbf{q}}{\mathbf{q}^{\mathbf{T}} \mathbf{q}} < \xi \tag{14}$$

where $\mathbf{q}$ represent the three parameters of $\mathbf{l}_\infty$, $A = \begin{bmatrix} v_x^i & v_y^i & 1 \end{bmatrix}^{\mathbf{T}}$ $\begin{bmatrix} v_x^i & v_y^i & 1 \end{bmatrix}$ contains the determined vanishing points, and ξ is the threshold. The Rayleigh quotients are estimated from the observation points and the residual errors are estimated. The threshold ξ is set to the median of all the residual errors. Observation points obtained from Eq. (13) having residual errors greater than ξ are removed as outliers. After outlier removal, the outlier-free remaining observation points Φ are used to construct the overdetermined system of Eq. (13). The system is then solved using the Singular Value Decomposition (SVD). The correct solution is the eigenvector corresponding to the smallest eigenvalue.

In summary, in order to minimize the influence of noise on our observation matrix Φ, we apply the Rayleigh quotient to filter out the noisy data points. Once the outliers are removed, the Total Least Squares method is applied to the remaining observation points to estimate an accurate $\mathbf{l}_\infty$.

5. Camera calibration

Using pole–polar relationship $\mathbf{l}_\infty \sim \omega \mathbf{v}_z$, the horizon line $\mathbf{l}_\infty$ and the vertical vanishing point $\mathbf{v}_z$ provide two constraints on the image of the absolute conic [18]. Assuming a camera with zero skew, and unit aspect ratio, the IAC would be of the form

$$\omega \sim [\omega_1\ \omega_2\ \omega_3] \sim \begin{bmatrix} 1 & 0 & \omega_{13} \\ 0 & 1 & \omega_{23} \\ \omega_{13} & \omega_{23} & \omega_{33} \end{bmatrix} \tag{15}$$

In the existing literature on camera calibration the role of IAC is primarily investigated in terms of its relationship with other geometric entities in the image plane, i.e. the vanishing points and the vanishing line. The relation between IAC and the internal parameters is often limited to equation $\omega \sim (\mathbf{KR})^{-T}\mathbf{I}\ (\mathbf{KR})^{-1} \sim \mathbf{K}^{-T}\mathbf{K}^{-1}$. In this section we present a relation that is more intrinsic to the IAC. Geometric interpretation for this relation allows us to gain more insight in to widely used the "closeness-to-the-center" constraint [8,18].

Theorem 5.1 (*Invariance*). *Let ω be the image of the absolute conic. The principal point $\tilde{\mathbf{p}}$ satisfies*

$$\omega\tilde{\mathbf{p}} \sim \mathbf{l}_\infty \tag{16}$$

where $\mathbf{l}_\infty \sim [0\ 0\ 1]^T$ is the line at infinity.

The proof is straightforward and follows by performing the Cholesky factorization of the Dual Image of the Absolute Conic (DIAC), ω^*, and direct substitution of $\tilde{\mathbf{p}}$.

5.1. Geometric interpretation

The result in (16) is better understood if we provide its geometric interpretation. Clearly, (16) is independent of the image points, i.e. for any two points $\tilde{\mathbf{m}}_1$ and $\tilde{\mathbf{m}}_2$, we have $\tilde{\mathbf{p}}^T\omega\tilde{\mathbf{m}}_1 = \tilde{\mathbf{p}}^T\omega\tilde{\mathbf{m}}_2$. Therefore, it reflects some intrinsic property of the IAC. This intrinsic property is better understood if we rewrite (16) as:

$$\tilde{\mathbf{p}}^T\omega_1 = 0 \tag{17}$$
$$\tilde{\mathbf{p}}^T\omega_2 = 0 \tag{18}$$

where ω_i are the rows of the IAC (or equivalently its columns due to symmetry).

This shows that

$$\tilde{\mathbf{p}} \sim \omega_1 \times \omega_2 \tag{19}$$

which is true for a general camera model, i.e. no particular assumptions made about the aspect ratio, or the skew.

A geometric interpretation (see Fig. 6) of this result is that the two rows ω_1 and ω_2 of the IAC correspond to two lines in the image plane that always intersect at the principal point regardless of the other intrinsic parameters.

Using the two constraints provided by the pole–polar relationship, we express the IAC in terms of only one of its parameters, e.g. ω_{33}, and solve for it by enforcing the constraint that the principal point is close to the center of the image by minimizing

$$\hat{\omega}_{33} = \arg\min \|\omega_1 \times \omega_2 - \mathbf{c}\| \tag{20}$$

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 (15). It must be noted that the pole–polar relationship could also be used on its own to recover a more simplified IAC without using the minimization in (20). Note
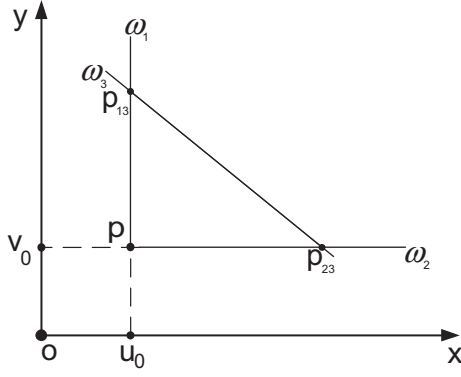
Fig. 6. The geometry associated with the IAC: ω_1, ω_2, and ω_3 represent the lines associated with the IAC when the skew is zero. The principal point is located at the intersection of the first two lines, providing two linear constraints on the IAC.

Fig. 7. The setup used for estimating geo-temporal information.

also that the proposed auto-calibration method is independent of any scene structure [25,33,38], or (special) camera motions [17,21,28]. We only require the vertical vanishing point and that the shadow be cast on a plane without requiring any further information.

6. 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 h. 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, an imaginary line that extends through the north and south geographic poles, tilts approximately to an 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 a camera by observing shadow trajectories only. We next describe the methods for determining these quantities.

6.1. Calculating latitude

An overview of the proposed method is shown in Fig. 7. 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 $\mathbf{a}_i$ and $\mathbf{v}'_i$, $i = 1, 2, 3$ be the sun and the shadow vanishing points, respectively [*note*: we are using $\mathbf{v}'_i$ for the shadow vanishing points, while we used $\mathbf{v}^1_x$ in the previous sections. This change of notation is necessary, as shall be clear in what follows]. For a calibrated camera, the following relations hold for the altitude angle ϕ_i and
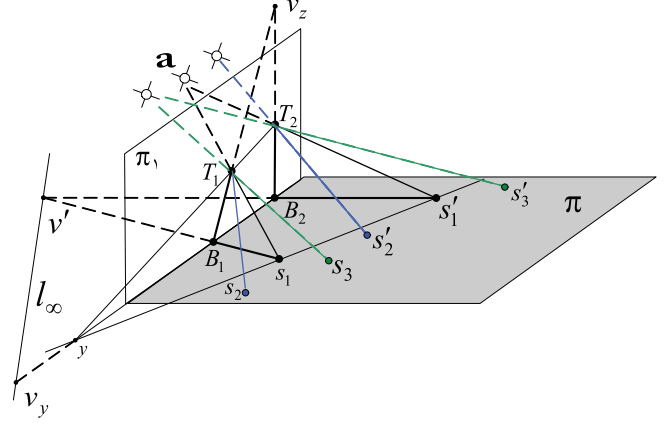
the azimuth angle θ_i of the sun orientations in the sky, all of which are measured directly in the image domain:

$$\cos \phi_i = \frac{\mathbf{v}'^T_i \boldsymbol{\omega} \mathbf{v}_i}{\sqrt{\mathbf{v}'^T_i \boldsymbol{\omega} \mathbf{v}'_i}\sqrt{\mathbf{v}^T_i \boldsymbol{\omega} \mathbf{v}_i}} \tag{21}$$

$$\sin \phi_i = \frac{\mathbf{v}^T_z \boldsymbol{\omega} \mathbf{v}_i}{\sqrt{\mathbf{v}^T_z \boldsymbol{\omega} \mathbf{v}_z}\sqrt{\mathbf{v}^T_i \boldsymbol{\omega} \mathbf{v}_i}} \tag{22}$$

$$\cos \theta_i = \frac{\mathbf{v}^T_y \boldsymbol{\omega} \mathbf{v}'}{\sqrt{\mathbf{v}^T_y \boldsymbol{\omega} \mathbf{v}_y}\sqrt{\mathbf{v}'^T \boldsymbol{\omega} \mathbf{v}'}} \tag{23}$$

$$\sin \theta_i = \frac{\mathbf{v}^T_x \boldsymbol{\omega} \mathbf{v}'}{\sqrt{\mathbf{v}^T_x \boldsymbol{\omega} \mathbf{v}_x}\sqrt{\mathbf{v}'^T \boldsymbol{\omega} \mathbf{v}'}} \tag{24}$$

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 as the image of the world origin. The vanishing point $\mathbf{v}_y$ along the *y*-axis is then given by $\mathbf{v}_y \sim \boldsymbol{\omega}\mathbf{v}_x \times \boldsymbol{\omega}\mathbf{v}_z$. Now, let ψ_i be the angles measured clockwise that the shadow points make with the positive *x*-axis as shown in Fig. 7. We have

$$\cos \psi_i = \frac{\mathbf{v}'^T_i \boldsymbol{\omega} \mathbf{v}_x}{\sqrt{\mathbf{v}'^T_i \boldsymbol{\omega} \mathbf{v}'_i}\sqrt{\mathbf{v}^T_x \boldsymbol{\omega} \mathbf{v}_x}} \tag{25}$$

$$\sin \psi_i = \frac{\mathbf{v}'^T_i \boldsymbol{\omega} \mathbf{v}_y}{\sqrt{\mathbf{v}'^T_i \boldsymbol{\omega} \mathbf{v}'_i}\sqrt{\mathbf{v}^T_y \boldsymbol{\omega} \mathbf{v}_y}} \quad i = 1, 2, 3 \tag{26}$$

Next, we define the following ratios, which are readily derived from spherical coordinates, and also used in sundial construction [20,22,36]:

$$\rho_1 = \frac{\cos \phi_2 \cos \psi_2 - \cos \phi_1 \cos \psi_1}{\sin \phi_2 - \sin \phi_1} \tag{27}$$

$$\rho_2 = \frac{\cos \phi_2 \sin \psi_2 - \cos \phi_1 \sin \psi_1}{\sin \phi_2 - \sin \phi_1} \tag{28}$$

$$\rho_3 = \frac{\cos \phi_2 \cos \psi_2 - \cos \phi_3 \cos \psi_3}{\sin \phi_2 - \sin \phi_3} \tag{29}$$

$$\rho_4 = \frac{\cos \phi_2 \sin \psi_2 - \cos \phi_3 \sin \psi_3}{\sin \phi_2 - \sin \phi_3} \tag{30}$$

For our problem, it is clear from (21)–(26) that these ratios are all determined directly in terms of image quantities. 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

$$\alpha = \tan^{-1} \left(\frac{\rho_1 - \rho_3}{\rho_4 - \rho_2} \right) \tag{31}$$

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

$$\lambda = \tan^{-1}(\rho_1 \cos \alpha + \rho_2 \sin \alpha) \tag{32}$$

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

6.2. Calculating day number

Once the latitude is determined from (32), we can also determine the exact day when the images are taken. For this purpose, let δ denote the declination angle, i.e. the angle of the sun's rays to the equatorial plane (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

$$\sin \hbar \cos \delta - \cos \phi \sin \theta = 0 \tag{33}$$
$$\cos \delta \cos \lambda \cos \hbar + \sin \delta \sin \lambda - \sin \phi = 0 \tag{34}$$

Again, note that the above system of equations depend only on image quantities defined in (21)–(26). Upon finding the declination and the hour angles by solving the above equations, the exact day of the year when the images are taken can be found by

$$N = \frac{365}{2\pi} \sin^{-1}\left(\frac{\delta}{\delta_m}\right) - N_o \tag{35}$$

where N is the day number of the date, with January 1st 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 1st.

6.3. Calculating 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 images 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 (32) 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 images. Then the GPS longitude of the location where the images are taken is given by

$$\gamma = \gamma_l + (\hbar - \hbar_l) \tag{36}$$

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 sensor, 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.

7. Using only two shadow points

At any location on the globe, the relationship between the location of the sun and the shadow is unique. This relationship can be graphically represented through sun-path diagrams. Also, the exact position of the sun can be determined for any given time of the day using only the azimuth and altitude angle of that site. Fig. 8 shows an example of vertical projection of sun-path as observed from earth. The vertical axis denotes the altitude and the horizontal axis denotes the azimuth angle. This plot is an earth base view of the sun's movement across the celestial sphere. The exact form of the curve depends on the location (latitude and longitude) and the time of the year. The question now is, can we estimate the GPS coordinates from just two points, whereas in previous sections we used three points?

Let (ϕ_1, θ_1) and (ϕ_2, θ_2) be the estimated azimuth and altitude angles of these two points. The method presented in Section 6 requires azimuth and altitude angles, θ and ϕ respectively, of at least three shadow points, so we need to determine (ϕ_3, θ_3). We also need to estimate the four ratios, i.e. (27)–(30), which depends on the angle, ψ. This angle ψ is measured between the shadow point $\mathbf{v}'$ and the +ve x-axis, as shown in Fig. 7. Therefore, we need to first, estimate the azimuth and altitude angle of the sun for any time of the day, and second, estimate the vanishing point $\mathbf{v}'$ of the shadows cast at that particular time.

It becomes clear upon observing Fig. 8 that the sun-path curve is symmetric. The axis of symmetry is exactly at 180° azimuth angle. This corresponds to the solar noon, that is, when the sun is at its highest point. Now consider the case when we have only two images i.e. we have only two shadow points. This is shown in Fig. 9. The axis of symmetry is plotted by a vertical line at $\theta = 180°$. The two shadow points obtained from the images are plotted on the left of this axis. These two points are then reflected across the axis, as shown in the figure. The problem now reduces to fitting a polynomial curve to these four points. A polynomial of kth degree is given as:

$$y = a_0 + a_1 x + \cdots a_k x^k \tag{37}$$

where the goal is to minimize the residual

$$R = \sum_{i=1}^{n} \left[y_i - \left(a_0 + a_1 x + \cdots a_k x^k \right) \right]^2$$

to fit the model as close to the data as possible. In matrix notation, the solution to the polynomial fit is given by:

$$\mathbf{y} = \Gamma \mathbf{g} \tag{38}$$

where $\mathbf{y}$ contains the LHS of (37) evaluated for all data points, the matrix Γ contains the x values of the data points from the RHS of
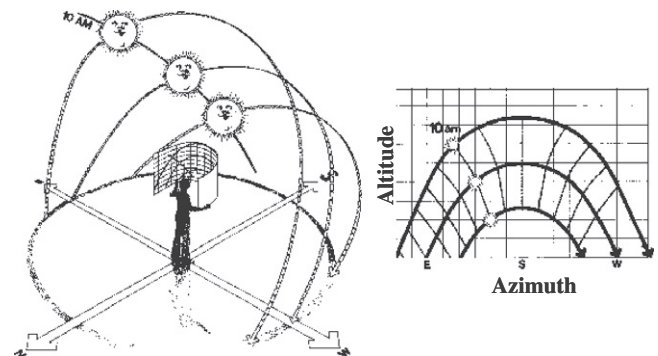

Fig. 8. The Cylindrical of Sun Path Diagram (Mazria, Edward, The Passive Solar Energy Book). The shadow of an object throughout the course of a day follows a curve on the ground plane.
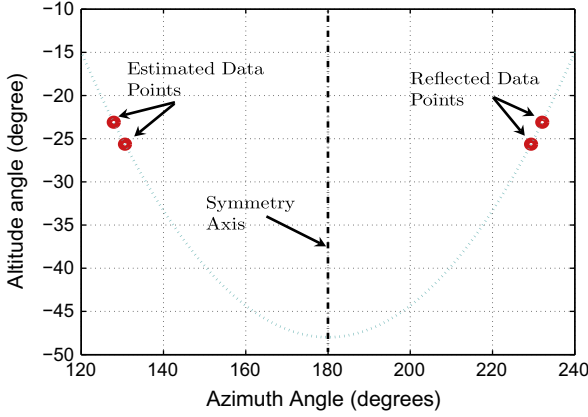
Fig. 9. A 2nd – $degree$ polynomial fitted to the estimated altitude and azimuth angles.

(37), and **a** contains the unknown parameters g_i [29]. (38) can be solved as:

$$\mathbf{g} = (\Gamma^T \Gamma)^{-1} \Gamma^T \mathbf{y} \qquad (39)$$

In our experiments, the polynomial that best fits the shadow data is that of degree 2. This is plotted in Fig. 9 as a dotted green curve. Once this curve is obtained, altitude ϕ_3 of any azimuth θ_3 of our choice can be estimated and vice versa.

Once (ϕ_3, θ_3) are obtained from the fitted shadow curve, the shadow point $\mathbf{v}'$ is obtained by solving the two equations:

$$\mathbf{v}'^T \mathbf{l}_\infty = 0 \qquad (40)$$

$$\cos \theta_3 = \frac{\mathbf{v}_y^T \boldsymbol{\omega} \mathbf{v}'}{\sqrt{\mathbf{v}_y^T \boldsymbol{\omega} \mathbf{v}_y} \sqrt{\mathbf{v}'^T \boldsymbol{\omega} \mathbf{v}'}} \qquad (41)$$

Once $\mathbf{v}'$ is obtained, (25) is used to estimate ψ to determine the four ratios, i.e. (27)–(30). This enables us to use the method described in Section 6 to estimate the GPS coordinates.

In the next section, we validate our method and evaluate the accuracy of both the self-calibration and the 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).
  - Obtain a robust estimation of the $\mathbf{l}_\infty$ by using the Rayleigh quotient, as described in Section 4.
- Perform camera calibration, as described in Section 5.
- Estimate the altitude and the azimuth angles, Eqs. (21)–(24). Estimate the ratios (27)–(30) to estimate the latitude λ and N.
- If only two images are available, use the method described in Section 7 to calculate the required parameters.
- If time of image acquisition is know, estimate the longitude γ, discussed in Section 6.3.

---

8. 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.

8.1. Synthetic data

Two vertical objects of different heights were randomly placed on the ground plane. Using the online version of SunAngle Software [14], we generated altitude and azimuth angles for the sun corresponding to our own geo-location with latitude 28.51° and longitude 81.3°. The data was generated for the 315th day of the year, i.e. the 11th of November 2007 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 the 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 Fig. 10. Note that, as explained in [33], 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 Fig. 10d. 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 −17.932°, and the day of the year is found to be 314.52.

8.2. Two point case

Fig. 11 shows the error curves of the estimated latitude, solar declination angle, and the day of the year when only two shadow points are used, as described in Section 7. For a maximum noise level of 1.54 pixels, the error for $\delta < 1.1\%$, $N < 1.7$, and $\lambda < 2$. This demonstrates that even in the presence of significant noise and using just two points, the proposed method gives very satisfactory results.

8.3. Real data

Several experiments on real datasets are reported below for demonstrating the effectiveness of the proposed method. Depending on the quality and resolution of the obtained images, we divide the datasets into `lowRes` and `highRes`.

8.3.1. `lowRes` dataset

This dataset is characterized by poor image resolution, generally 320×240. All the sequences in this dataset are acquired by traffic monitoring cameras. An example is shown in Fig. 12: 11 images were captured live from downtown Washington, DC, area, using one of the webcams available online at http://trafficland.com/. As shown in Fig. 12, a lamp post and a traffic 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

$$\mathbf{K} = \begin{bmatrix} 700.357 & 0 & 172 \\ 0 & 700.357 & 124 \\ 0 & 0 & 1 \end{bmatrix}$$

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
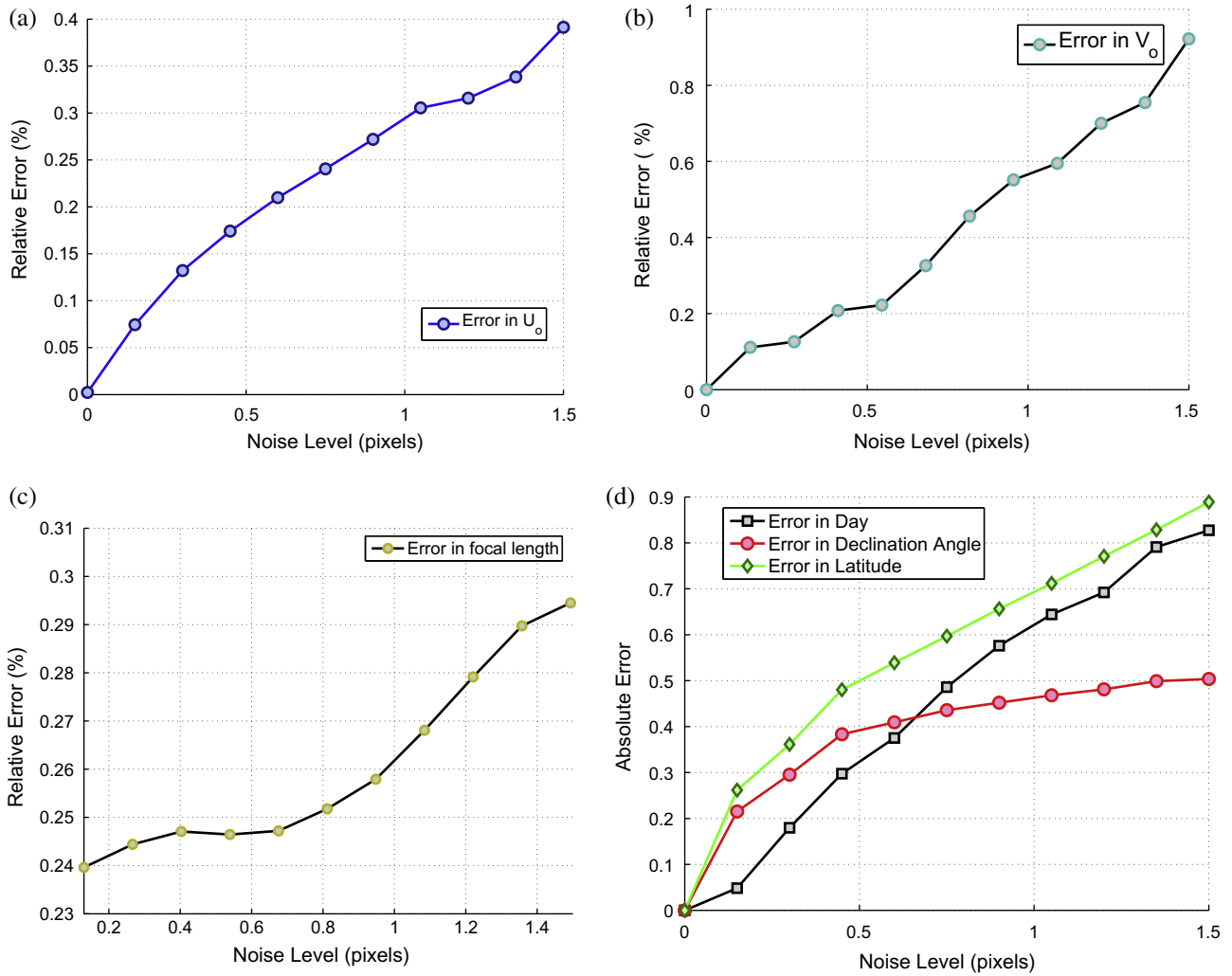
Fig. 10. Performance averaged over 1000 independent trials: (a and b) relative error in the coordinates of the principal point (u_o, v_o), (c) the relative error in the focal length f, and (d) result for average error in latitude, solar declination angle, and day of the year.
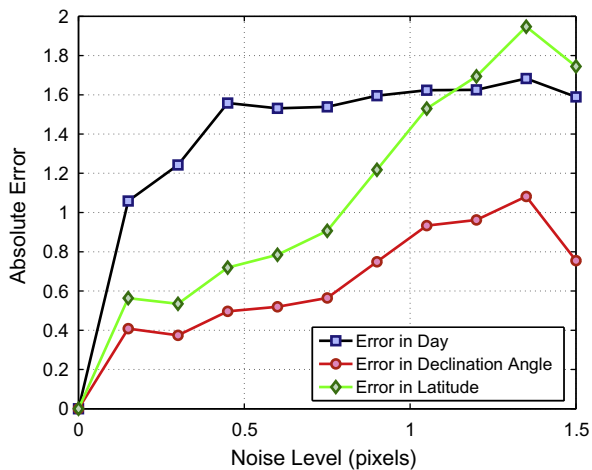

Fig. 11. Result for average error in latitude, solar declination angle, and day of the year when only two shadow points are used to estimate these quantities. The error is slightly higher than the three point case, which is understandable.

points and averaged the results. This dataset was captured on the 15th 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 1. The low standard deviations can be interpreted as small uncertainty, indicating that our method is consistently providing reliable results.

A second example set is shown in Fig. 13. The ground truth for this data set is as follows: longitude 77.02°, latitude 38.53°, day number of 331, and the declination of −21.8°. For this dataset 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 37.791°, the day number as 323.0653, and the declination angle as −17.29°.

Sample images from the remaining eleven sequences are shown in Fig. 14, captured from different places around the globe. Table 2 shows the results obtained from applying the proposed method to all these sequences. The first row of Table 2, i.e. Napoleonville, Lou-

Table 1
Results for 11 sets of 10-image combination, with mean value and standard deviation.

	$\mathscr{C}_1$	$\mathscr{C}_2$	$\mathscr{C}_3$	$\mathscr{C}_4$	$\mathscr{C}_5$	$\mathscr{C}_6$	$\mathscr{C}_7$	$\mathscr{C}_8$	$\mathscr{C}_9$	$\mathscr{C}_{10}$	$\mathscr{C}_{11}$	Mean	STD
λ	33.73	35.70	37.03	36.1	35.72	38.21	39.23	45.78	41.84	40.88	41.96	38.743	3.57
δ	−14.47	−15.78	−15.93	−16.54	−17.25	−16	−16.70	−18.94	−15.87	−16.99	−16.24	−16.43	1.11
N	328.64	332.26	331.09	326.87	330.15	331.37	331.32	332.56	326.81	331.72	326.72	329.95	2.28

Fig. 12. Few of the images taken from one of the live webcams in downtown Washington, DC. 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.

Fig. 13. 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, DC. The objects that cast shadows on the ground are highlighted. Shadows move to the left of the images as time progresses.

isiana – R$_1$, shows the results of applying our method on the sequence in row 1 of Fig. 14 and so on. The 2nd column of the table contains the ground truth latitude (λ), the 3rd column contains the computed latitude, the 4th contains the ground truth δ while the 5th column contains the estimated δ. The ground truth day is displayed in 6th column while the computed value is shown in 7th column. From these results, the average error in λ is 1.53° with a standard deviation of 0.75°. The average error in δ is 1.72° with a standard deviation of 0.6°.The average error in N is 20.36 with a standard deviation of 10.16.

8.3.2. `highRes` dataset

The three sequences in this dataset is characterized by a good resolution image, generally a resolution higher than 640×480. The first sequence is shown in Fig. 15, captured in Orlando, FL, USA. The top row shows some sample images, the second image from left in row 2 is the accumulated background image. The third image from left in row 2 shows one of the shadow points being tracked. Latitude of the location is 28.6°, whereas we obtained 28.74° from these images. The ground truth for the declination angle (δ) and the day number (N) is −22.31° and 338. From the proposed method, these quantities are calculated to be −21.79° and 329, respectively.

Second and third sequences are shown in Fig. 16, in row 1 (Sharjah, UAE) and row 2 (Phoenix, AZ), respectively. Ground truth latitude for the top row is 25.29°, whereas we obtained 24.81° from these images. The ground truth for the declination angle (δ) and the day number (N) is −21.80° and 11. From the proposed method, these quantities are calculated to be −22.17° and 24, respectively. Ground truth latitude for the 2nd row is 33.43°, whereas we obtained 34.17° from these images. The ground truth for the declination angle (δ) and the day number (N) is −22.91° and 2. From the proposed method, these quantities are calculated to be −23.35° and 19, respectively. The results obtained from these images are fairly accurate compared to low-resolution images, owing to a more precise shadow point localization.

In general, one of the reason for noise in results is the low quality of the images, in addition to the weather effects, such as mud on the road or the cloudy weather. Nonetheless, the obtained param-eters are very close to the ground truth values. Whereas [23] observe this scene for weeks and months, we perform the same task of estimating the GPS coordinates of this location by using only a few images obtained during the course of a day. Some of the cases, where the proposed method fails are shown in Fig. 17. The figure shows the limitations of the approach, i.e. cloudy conditions, reflections on the road surface, shadow falling on non-planar surfaces, or shadow casting objects not visible in the captures images.

9. Discussion and 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 images are taken without using any prior information. Our approach consists of two steps: auto-calibration from shadows, and geo-temporal localization. The auto-calibration step requires only the shadow trajectories of two objects on the ground plane to be visible in the images, along with the vertical vanishing point. Unlike shadow-based calibration methods such as [2,8], this step does not require the objects themselves to be seen in the images.

It is, however, important that the shadow trajectories can be used to fit conics. An exception, which leads to a degenerate case, happens twice a year during equinox, when the lengths of the day and the night are equal. As a result, it can be shown that, the shadow trajectories degenerate to straight lines. Two cases may occur: if the two objects casting shadows are not aligned along the east–west direction, then their shadow trajectories will be two distinct straight lines that are parallel in the world. Therefore, their intersection would provide only a single point at infinity, which is insufficient to determine the horizon line. Furthermore, if the two objects are aligned along the east–west direction, then the shadow lines will coincide and no vanishing point can be found. In both cases auto-calibration cannot be performed using our method. However, this degenerate case is rather rare and happens only twice a year.
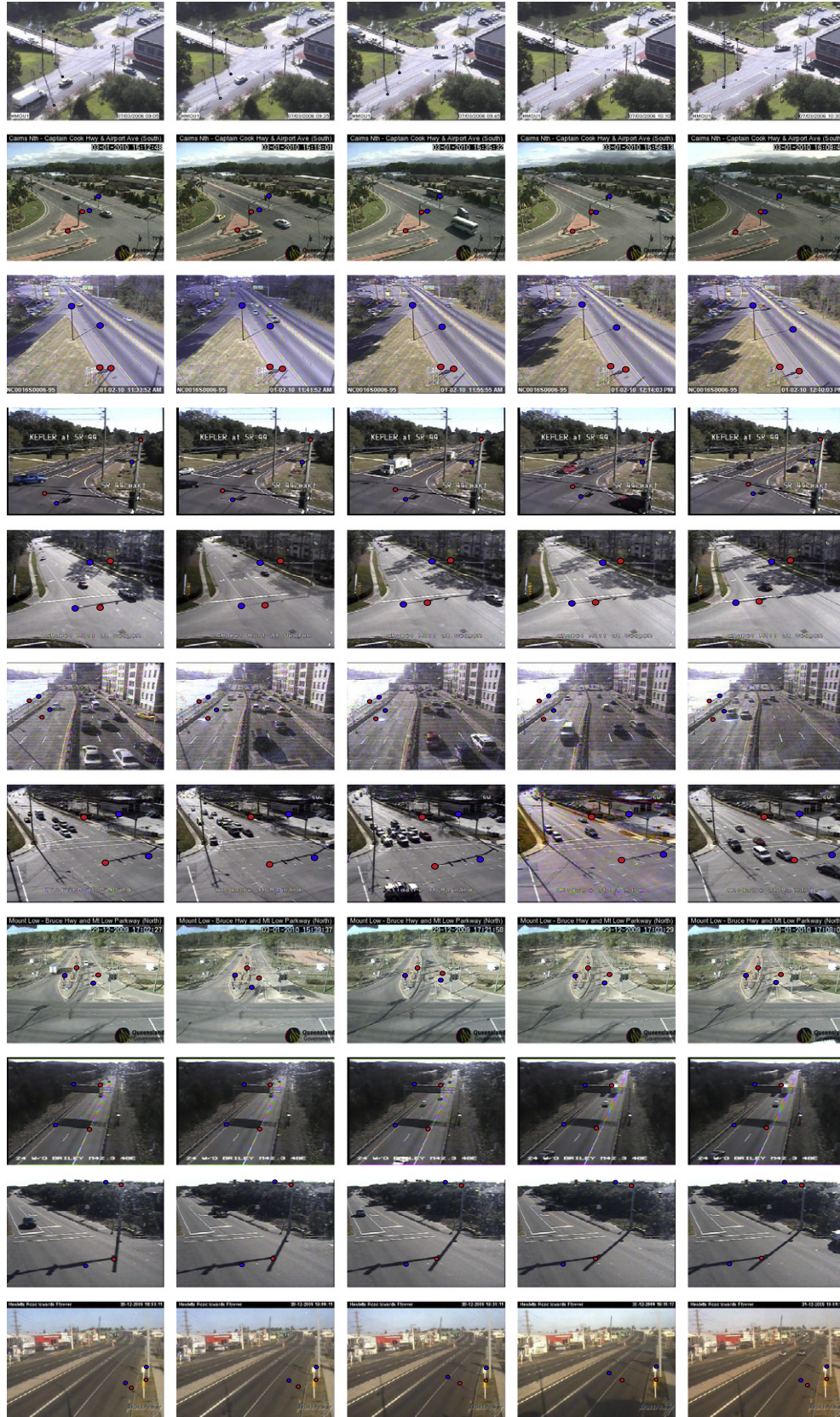
Fig. 14. Examples from the `lowRes` data set.

Table 2
Obtained results vs. ground truth for Fig. 14.

Sequence – Row$_\#$	λ	λ'	δ	δ'	N	N'
Napoleonville, Louisiana – R$_1$	29.91°	27.56°	22.08°	20.04°	228	274
Cairns North, Australia – R$_2$	16.89°	17.75°	−22.78°	−23.39°	3	11
Charlotte , NC – R$_3$	35.26°	36.89°	−22.83°	−20.95°	2	358
Daytona Beach , FL – R$_4$	29.06°	30.94°	−22.83°	−20.9°	2	22
Morrisville, NC – R$_5$	35.81°	36.12°	−22.83°	−20.62°	2	348
New York City, NY – R$_6$	40.77°	37.65°	−23.17°	−21.47°	363	19
Cary, NC – R$_7$	35.77°	34.21°	−22.83°	−21.34°	2	23
Deeragun, Australia – R$_8$	19.24°	20.91°	−22.73°	−19.88°	3	345
Nashville, TN – R$_9$	36.25°	35.02°	−22.86°	−21.15°	2	16
Kitty Hawk, NC – R$_{10}$	36.13°	37.22°	−23.20°	−22.19°	363	337
Bay Of Plenty, NZ – R$_{11}$	38.79°	36.11°	−23.05°	−21.56°	365	17

Fig. 15. `highRes`: Top row shows a few images in the data set overlooking a bus stand. The 2nd image in row 2 shows the accumulated background of the scene and the 3rd image in row 2 depicts a shadow being tracked. The last image in row 2 has two vertical lines shown which are used for computing the vertical vanishing point.

Fig. 16. `highRes`: Two sequences obtained from higher resolution cameras. See text for more details.

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

Fig. 17. Some of the examples where the proposed method fails due to weather and other varying conditions. (a) This is the case when the shadow casting object does not appear in the scene but the shadow does, (b) when the shadow no longer lies on a plane. (c) Shows a common case where the shadows disappear due to overcast conditions, finally (d) depicts a situation where other circumstances, such as object occlusion or reflections from the road prevents accurate shadow localization.

References

[1] L.D. Agapito, E. Hayman, I. Reid, Self-calibration of rotating and zooming cameras, Int. J. Comput. Vision 45 (2) (2001) 107–127.

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

[3] M.J. Black, P. Anandan, The robust estimation of multiple motions: parametric and piecewise-smooth flow fields, J. Comput.Vision Image Underst. 63 (1) (1996) 75–104.

[4] M.J. Black, D.J. Fleet, Y. Yacoob, Robustly estimating changes in image appearance, Comput. Vision Image Underst. (CVIU) 78 (2000).

[5] J. Bouguet, P. Perona, 3D photography on your desk, in: Proc. ICCV, 1998, pp. 43–50.

[6] T.E. Boult, X. Gao, R. Micheals, M. Eckmann, Omni-directional visual surveillance, IEEE Trans. Pattern Anal. Mach. Intell. 22 (7) (2004) 515–534.

[7] X. Cao, H. Foroosh, Camera calibration and light source orientation from solar shadows, J. Comput. Vision Image Underst. (CVIU) 105 (2006) 60–72.

[8] X. Cao, M. Shah, Camera calibration and light source estimation from images with shadows, in: Proc. IEEE CVPR, 2005, pp. 918–923.

[9] B. Caprile, V. Torre, Using vanishing points for camera calibration, Int. J. Comput. Vision 4 (2) (1990) 127–140.

[10] Y. Caspi, M. Werman, Vertical parallax from moving shadows, in: Proc. CVPR, 2006, pp. 2309–2315.

[11] D. Comaniciu, V. Ramesh, P. Meer, Kernel-based object tracking, IEEE Trans. Pattern Anal. Mach. Intell. (PAMI) 25 (5) (2003) 564–575.

[12] F. Devernay, O.D. Faugeras, Straight lines have to be straight, Mach. Vision Appl. 13 (1) (2001) 14–24.

[13] O. Faugeras, T. Luong, S. Maybank, Camera self-calibration: theory and experiments, in: Proc. of ECCV, 1992, pp. 321–334.

[14] C. Gronbeck, Sunangle software. <www.susdesign.com/sunangle/>.

[15] A. Guiducci, Camera calibration for road applications, Comput. Vision Image Underst. (CVIU) 79 (2000) 250–266.

[16] R.I. Hartley, Self-calibration from multiple views with a rotating camera, in: Proc. ECCV, 1994, pp. 471–478.

[17] R.I. Hartley, Self-calibration of stationary cameras, Int. J. Comput. Vision 22 (1) (1997) 5–23.

[18] R.I. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision, second ed., Cambridge University Press, 2004. ISBN: 0521540518.

[19] J. Heikkila, Geometric camera calibration using circular control points, IEEE Trans. Pattern Anal. Mach. Intell. 22 (10) (2000) 1066–1077.

[20] A. Herbert, Sundials Old and New, Methuen & Co. Ltd., 1967.

[21] A. Heyden, K. Astrom, Euclidean reconstruction from image sequences with varying and unknown focal length and principal point, in: Proc. IEEE CVPR, 1997, pp. 438–443.

[22] Frederick W. Sawyer III, A Three-Point Sundial Construction, Bulletin of the British Sundial Society, 94, 1, 22–29, 1994, Feb.

[23] N. Jacobs, S. Satkin, N. Roman, R. Speyer, R. Pless, Geolocating static cameras. in: Proc. of ICCV, 2007, pp. 469–476.

[24] O. Lanz, Automatic lens distortion estimation for an active camera, in: International Conference on Computer Vision and Graphics (ICCVG), vol. 13, no. 1, September 22–24, 2004, Warsaw, Poland, pp. 14–24.

[25] D. Liebowitz, A. Zisserman, Combining scene and auto-calibration constraints, in: Proc. IEEE ICCV, 1999, pp. 293–300.

[26] F. Lu, Y. Shen, X. Cao, H. Foroosh, Camera calibration from two shadow trajectories. in: Proc. ICPR, 2005, pp. 1–4.

[27] M. Pollefeys, R. Koch, L.V. Gool, Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters, Int. J. Comput. Vision 32 (1) (1999) 7–25.

[28] M. Pollefeys, R. Koch, L.V. Gool, Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters, Int. J. Comput. Vision 32 (1) (1999) 7–25.

[29] W. Press, B. Flannery, S. Teukolsky, W. Vetterling, Numerical Recipes in C, Cambridge University Press, 1988.

[30] J.G. Semple, G.T. Kneebone, Algebraic Projective Geometry, Oxford Classic Texts in the Physical Sciences (1979).

[31] P. Sturm, Critical motion sequences for the self-calibration of cameras and stereo systems with variable focal length, in: British Machine Vision Conference, Nottingham, England, Sep 1999, pp. 63–72.

[32] K. Sunkavalli, F. Romeiro, W. Matusik, T. Zickler, H. Pfister, What do color changes reveal about an outdoor scene? in: IEEE Conference on Computer Vision and Pattern Recognition, 2008, pp. 1–8.

[33] B. Triggs, Autocalibration from planar scenes, in: Proc. ECCV, 1998, pp. 89–105.

[34] R. Tsai, A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses, IEEE J. Robot. Autom. 3 (4) (1987) 323–344.

[35] A. Veeraraghavan, R. Chellappa, M. Srinivasan, Shape-and-behavior encoded tracking of bee dances, IEEE Trans. Pattern Anal. Mach. Intell. 30 (3) (2008) 463–476.

[36] A. Waugh, Sundials: Their Theory and Construction, Dover Publications, Inc., 1973. ISBN 0-486-22947-5.

[37] J. Yao, Z. Zhang, Hierarchical shadow next term detection for color aerial images, Comput. Vision Image Underst. (CVIU) (2005) 102.

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

[39] Z. Zhang, Camera calibration with one-dimensional objects, IEEE Trans. Pattern Anal. Mach. Intell. 26 (7) (2004) 892–899.