



## Camera calibration and geo-location estimation from two shadow trajectories

Lin Wu<sup>a,\*</sup>, Xiaochun Cao<sup>a</sup>, Hassan Foroosh<sup>b</sup>

<sup>a</sup>School of Computer Science and Technology, Tianjin University, Tianjin 300072, China

<sup>b</sup>Computer Vision Laboratory, University of Central Florida, Orlando, FL 32816, United States

### ARTICLE INFO

#### Article history:

Received 23 June 2009

Accepted 21 April 2010

Available online 29 April 2010

#### Keywords:

Geographical location

Metric rectification

Shadow trajectory

Analemmatic sundial

### ABSTRACT

The position of a world point's solar shadow depends on its geographical location, the geometrical relationship between the orientation of the sunshine and the ground plane where the shadow casts. This paper investigates the property of solar shadow trajectories on a planar surface and shows that camera parameters, latitude, longitude and shadow casting objects' relative height ratios can be estimated from two observed shadow trajectories. One contribution is to recover the horizon line and metric rectification matrix from four observations of two shadow tips. The other contribution is that we use the design of an analemmatic sundial to get the shadow conic and furthermore recover the camera's geographical location. The proposed method does not require the shadow casting objects or a vertical object to be visible in the recovery of camera calibration. This approach is thoroughly validated on both synthetic and real data, and tested against various sources of errors including noise, number of observations, objects locations, and camera orientations. We also present applications to image-based metrology.

© 2010 Elsevier Inc. All rights reserved.

### 1. Introduction

Estimating geographic information is a high-level problem in computer vision [40,16,24,34,26]. A lot of cameras are mounted over the Earth to undertake video surveillance, observe weather patterns, monitor coastline or navigation. Thereby, a large number of outdoor or indoor images emerge on the internet, which are valuable for people with various purposes like tourism, security, or forensics. The ability to determine geographical information from visual cues is a significant work because many images are captured at different places over the entire Earth. Geo-location can provide plenty of background information including local climate, average temperature and rainfall, population density [16].

Currently, there are mainly three types of geo-locating methods: feature matching based methods [40,16,24], photometry based method [34] and geometry based methods [26]. The limitations shared within feature matching methods are the dependence on a great quantity of GPS tagged signals and the high complexity in the matching process between input images and known signals. The photometry based method provides the access to the angular velocity of the Sun and the angle of each scene point's projection on the solar plane by analyzing the time-varying color model. The geometry based method avoids the GPS tagged information and relieves the constraints for the scene. The inputs to the geometric method are observed objects, e.g. the solar shadows, which are able to provide necessary geometric relationships among geo-

graphic location, scene geometry, and camera orientation. Belonging to this line of research, our approach also needs to remove the perspective distortions [26].

Due to the perspective distortion, the number of imaged pixels of the same target varies significantly depending on the distance between the camera and the target. In video surveillance applications, removing the distortion will not only reduce the number of pixels to be processed for speed-up, but also improve the tracking performance [8]. Metric rectification is preliminary for geo-location estimation since the achievement of latitude from the geometric relationship residing in the analemmatic sundial, holds true only in the world coordinate system. Although the traditional camera calibration methods [15] are able to remove the perspective and affine distortions, they usually involves taking images of some special pattern with known 3D geometry, extracting features and minimizing their re-projection errors. These methods often produce very good results whereas not only require the Euclidian information but also the special shapes from the calibration patterns such as sphere [39], coplanar circles [7], surface of revolution [37,11], stars in the sky [20], 2D or 1D objects [3,36,41] and symmetric objects [3]. Multiple view approaches [14,29] avoid the use of special calibration objects, but typically require more than three views, have higher computational cost, and involve non-linear problems.

Objects and their solar shadows are common in the real world [1,2,4–6,9], especially in outdoor environments. Shadows are important because they provide important visual cues, e.g. depth and shape, in our perception of the world [5,6,9,28]. They are also useful for computer vision applications such as planar object detection [38], building detection [22], visual surveillance

\* Corresponding author.

E-mail address: [jinlinwulin1986@gmail.com](mailto:jinlinwulin1986@gmail.com) (L. Wu).

[33,17,23,18,30,25] and inverse lighting [32,27]. However, shadows are not frequently addressed in 3D vision. Some few examples include the relationship between the shadows and object structure by Kriegman and Belhumeur [21,2], object recognition by Van Gool et al. [35], and shadow segmentation by Yacoob and Davis [38]. There has been some investigation of outdoor camera calibration (e.g. [10]) while most methods attempt to eliminate shadows for better performance rather than exploiting their properties to estimate camera parameters. Shadows are interesting for camera calibration [4] since the shadows of two parallel vertical lines cast on the planar ground plane by an infinite point light source are also parallel, which together with the vertical vanishing point provide the orthogonal constraint on the image of the absolute conic. Cast shadows are used in image processing algorithms for camera calibration [1,4], light source estimation [4], etc.

In this paper, a new method that takes advantage of fixed surveillance cameras is proposed to recover the horizon line and camera calibration directly from only shadow points. Horizon line is used to remove perspective distortion. Therefore, specifying the line in infinity enables metric properties to be recovered from an image of a plane by transforming the circular points to their canonical positions [15]. In this setup, only four shadow observations are required to calibrate the camera without requiring the objects to be visible. For geo-location, we employ the design of analemmatic sundial which incorporates the latitude, the position of the Sun, the time of a day to obtain geographical information. The approach explores the geometric relationship between the shadow trajectory and the analemmatic sundial to geo-locate the camera. Comparing to existing geo-location estimation method [26,16], our approach does not require shadow casting objects's visibility and any GPS information.

## 2. Related work and contributions

Vanishing points and the horizon line play central roles in perspective geometry [15] for applications such as camera calibration and image-based metrology [12]. The calibration method proposed by Antone and Bosse [1] requires accurate geography information, which is not always possible to obtain from images. Inter-image constraints on the Image of Absolute Conic from objects and their cast shadows were introduced by Cao and Foroosh [4], where the authors argue that two views of an outdoor scene with two vertical objects under sunshine are sufficient to calibrate a camera. Both these methods [1,4] require the entire objects and their shadows to be present in the image. This paper recovers the horizon line in the image from the only shadow observations without knowing the object-to-shadow correspondences. Our calibration method is then based on the geometric constraints provided by two shadow trajectories. Different from [1,4], the contribution of our method lies on relieving the requirement of visible or vertical objects that cast shadows.

Junejo and Foroosh [26] use shadow points from shadow trajectory of objects to determine the geo-location of a camera. For the camera calibration, they argue that when shadow casting objects are not visible, it requires six imaged shadow positions of the same 3D points to fit a conic which can meet the line at infinity of the ground plane at two circular points. To estimate the geo-location, they require the entire shadow casting objects to be visible. Instead, we use the method of computing latitude inspired by the design of the analemmatic sundial which just requires the footprints of shadow casting objects ([26] requires the entire object to be visible since they rely on the estimation of the azimuth and altitude of the Sun). In addition, we use the minimal geometric information, that consists of four observations of the shadow trajectory, while they [26] require six observations.

Sunkavalli et al. [34] focus on color changes and aim to separate the direct sunlight and ambient skylight components for the

recovery of partial scene information. They recover scene geometry as well as the geo-location and geo-orientation of the camera by utilizing common geometric constraints. Our method has the similar goals but different in three respects: (1) we recover the geometric constraints from only shadow trajectories while they assume the three mutually orthogonal vanishing lines are provided, i.e. a calibrated camera, (2) the inputs to our algorithm are shadow trajectories which are common in the world while they use color changes induced by variations in the spectral composition of daylight, and (3) our approach is simpler.

The paper is organized as follows: Section 3 presents the elementary knowledge of geography and analemmatic sundial. Section 4 introduces the recovery of horizon line from two shadow trajectories and the computation of camera calibration parameters. Section 5 proposes a novel method to calculate the stationary camera's longitude and latitude after rectifying the fitted conic trajectories. In Section 6, some special cases are considered in which our assumptions do not hold true. In Section 7, we verify the proposed method on both simulated and real data. We also demonstrate the application on estimating the relative height ratio between the shadow casting objects which is not carried out in [26]. Section 8 concludes this work.

## 3. Preliminaries and set up

### 3.1. Introduction of the analemmatic sundial

Any location on Earth is described by two numbers: its longitude and its latitude. Longitude is the angular position of a place on the Earth's surface measured from the prime meridian through Greenwich, England. The preferred convention is that longitudes East to the prime meridian are positive and West are negative. Similarly, latitude is the angular distance measured from the equator, North (+ve) or South (-ve). Latitude value is important since it defines the relationship of a location with the Sun. The path of the Sun that can be seen from the Earth, is unique for each latitude, which is the main cue that allows us to geo-locate a static camera by only observing shadow trajectories.

For the purpose of the design of sundials [31], it is an excellent approximation that the Sun revolves around a stationary Earth on the celestial sphere with the period of rotation as 23 h and 56 min about its celestial axis, the line connecting the celestial poles. Typically, it needs latitude information in the design for polar pointing gnomon and it is difficult to move the gnomon to a different location. An analemmatic sundial, also known as a "Human Sundial", is flexible and movable as it uses a vertical gnomon such as a person standing in the proper position for the month. Its hour lines are the vertical projection of the hour lines of a circular equatorial sundial onto a flat plane. Therefore, the analemmatic sundial is an ellipse, where the long axis is aligned East–West and the short axis points North–South. Its design involves the knowledge of declination of the Sun, latitude, direction to true North and the vertical direction. The advantage of such a sundial is that the gnomon needs not to be aligned with the celestial poles and may even be perfectly vertical with the position of the gnomon. The most important is that the hour points consisting of the sundial reflect the direction of shadows. With the availability of orthogonal projection on the plane casted by vertical objects, the varied shadow trajectories during the day can be obtained. We will show in Section 5 that from the time stamped observations of solar shadows we can recover the longitude, latitude, camera calibration and metric information.

### 3.2. About time: local and standard time

Two important concepts, related to latitude and especially longitude are local time (Solar time) and standard time (clock time).

Local time is a measure of the position of the Sun relative to a locality. It is what we use to regulate our lives locally. Standard time is mean solar time at the central meridian of a given time zone. It is the legally accepted time scale in a particular country or region.

Sundials indicate the local solar time or local apparent time, not the clock time, unless otherwise corrected. To obtain the solar time, three types of corrections need to be considered when clock time are available. First, due to the facts that the orbit of the Earth is not perfectly circular (in fact it is an ellipse) and its rotational axis is not perfectly perpendicular to its orbit, the Sun does not rotate uniformly about the Earth. These imperfections produce small variations in the sundial time throughout the year. The required correction may be as large as a quarter-hour early or late which is described by the equation of time ( $t_{EOT}$ ). Note that this correction is global and does not depend on the local latitude of sundial. The correction varies between extremes of about +14 min in February and –16 min in October, and is effectively equivalent to zero on 4 days of the year. Second, the solar time needs to be corrected for the longitude of the sundial relative to the central meridian at which the official time zone is defined. Within a given time zone, if the sundial is located at the central meridian of that time zone, it will show the clock time with the appropriate correction of the equation of time. However, if the sundial is located to the West (East), it will indicate slow (fast). This correction is often made by rotating the hour-lines by an angle equaling the difference in longitudes. (We use  $t_{off}$  to describe it.) Third, the practice of daylight saving time ( $t_{DST}$ ) shifts the official time away from solar time by an hour advanced in some countries during the summer.

### 3.3. Shadow trajectory detection

A semi-automatic approach is employed to detect the shadow point for an input video. First, we construct a background image  $\mathbf{B}$  for the set of input images  $\mathbf{V} = \{I_1, I_2, \dots, I_k\}$ :

$$\mathbf{B}(x, y) = \max_{i \in [1, k]} (I_i(x, y)). \quad (1)$$

In  $\mathbf{B}$ , each pixel  $(x, y)$  contains the brightest value among  $\mathbf{V}$ . The brightest values in  $\mathbf{B}$  are computed on gray scale (Fig. 1b). After applying the background subtraction, the shadow region stands out by thresholding the absolute differences. Then we select a point in the shadow region and enforce the flood fill algorithm to determine the area. Finally, the prominent shadow tip is highlighted by using the PCA technique (Fig. 1c). Shown in Fig. 1d is the shadow trajectory of the left object that is made up of shadow tips at a quarter's interval.

## 4. The method

### 4.1. The recovery of horizon line

The method here requires no presence of objects, but only shadows on the ground plane cast by two unknown stationary 3D points. We first utilize the tracked shadow positions to compute the horizon line. This is based on our observation that any two 3D corresponding lines on the shadow trajectories are parallel to each other. Therefore the imaged lines provide the vanishing

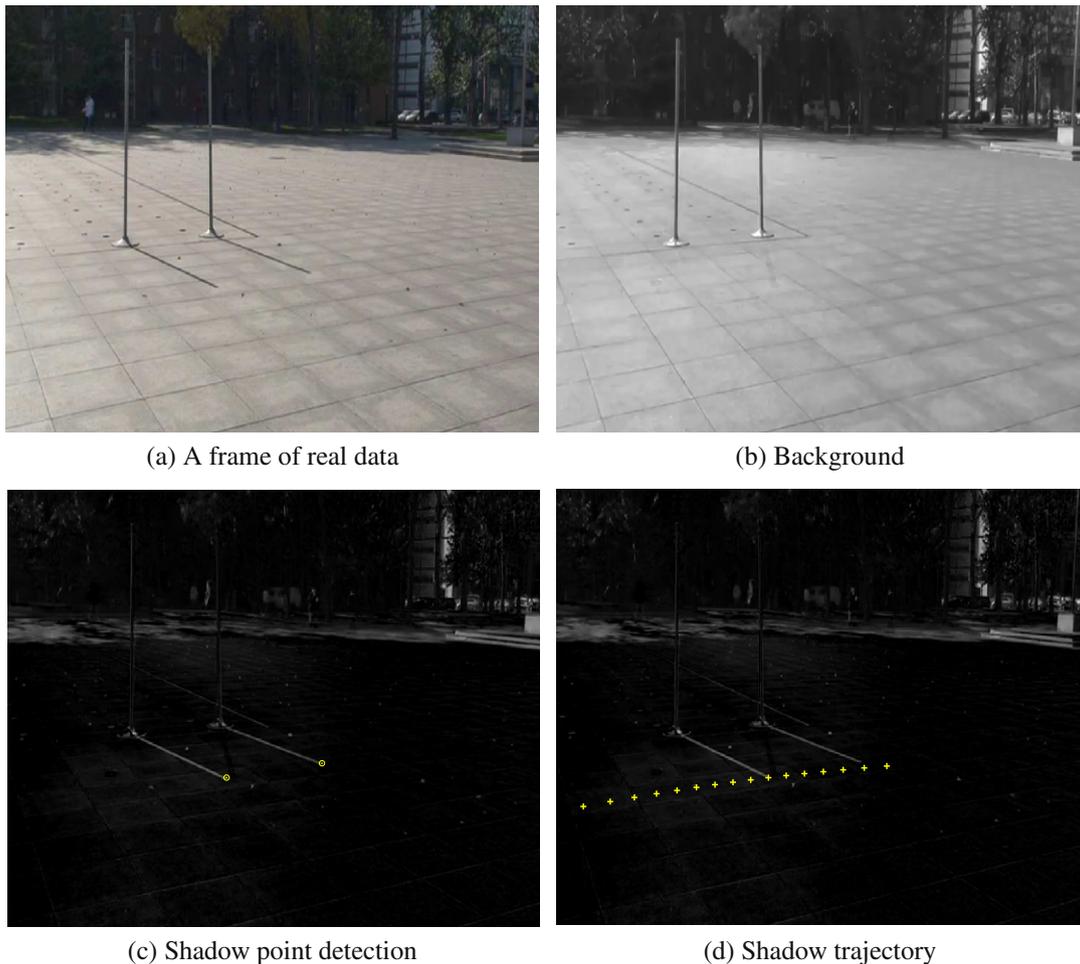
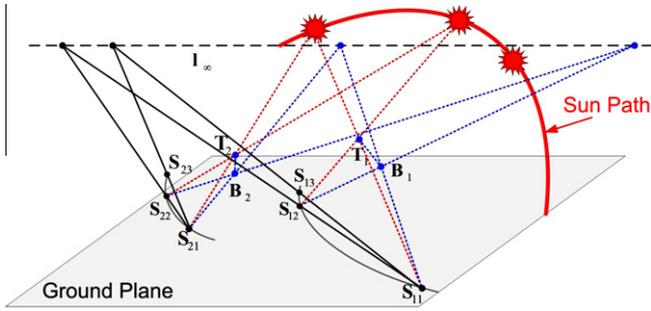


Fig. 1. Automatic shadow trajectory detection.



**Fig. 2.** The geometric relationship in two solar shadow trajectories. The 3D points  $T_i$  cast shadows at different time  $m$  at positions  $S_m$ . The blue segments  $B_1 S_{1m}$  and  $B_2 S_{2m}$  are parallel to each other and intersect at a vanishing point, as demonstrated in Eq. (2). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

points along directions perpendicular to the vertical direction. The basic idea is illustrated in Fig. 2. Suppose  $T_i$  is a 3D point and  $B_i$  is the closest point on the ground plane to  $T_i$ . Therefore, line  $T_i B_i$  is perpendicular to the ground plane.  $S_{im}$  is the cast shadow point at time  $m$ . Because the Sun is far away (approximately  $1.52 \times 10^{12}$  m) from the Earth, it is reasonable to consider the sunlight as parallel light rays. Therefore, in the 3D world coordinate, we have

$$\left. \begin{aligned} T_1 S_{1m} // T_2 S_{2m} \\ T_1 B_1 // T_2 B_2 \\ \angle T_1 B_1 S_{1m} = \pi/2 = \angle T_2 B_2 S_{2m} \end{aligned} \right\} \Rightarrow \Delta T_1 B_1 S_{1m} \sim \Delta T_2 B_2 S_{2m}, \quad (2)$$

where  $\sim$  indicates similar relationship between triangles. We can also have

$$\frac{|B_1 S_{1m}|}{|B_2 S_{2m}|} = \frac{|T_1 B_1|}{|T_2 B_2|} = \frac{|B_1 S_{1n}|}{|B_2 S_{2n}|}, \quad (3)$$

$$B_1 S_{1m} // B_2 S_{2m}, \quad (4)$$

$$B_1 S_{1n} // B_2 S_{2n}, \quad (5)$$

where  $m$  and  $n$  are two arbitrarily different time. Note that this equation is true in 3D, not in the image plane. Consequently, the two world triangles  $\Delta B_1 S_{1m} S_{1n}$  and  $\Delta B_2 S_{2m} S_{2n}$  are similar, and that the world line  $S_{1m} S_{1n}$  and  $S_{2m} S_{2n}$  are parallel to each other. Therefore, it is evident that shadows observed over time are sufficient to provide the horizon line  $l_\infty = [l_1, l_2, l_3]^T$ . In the estimation of  $l_\infty$ ,

the object top and bottom points ( $t_i, b_i$ ) are not required to be present in the image. Having the parallel constraints implying in two solar shadow trajectories, it is possible to determine the horizon line using only three pairs of observations collected from two trajectories at three different time. When more than minimally required observations are available, e.g.  $n$  observations, we can get  $C_n^2$  ( $n \geq 3$ ) vanishing points and the horizon line can be fitted by using the RANSAC algorithm [13].

#### 4.2. Camera calibration using four shadow observations

The image of absolute conic  $\omega$  is an imaginary point conic on the plane at infinity which is invariant to camera rotation and translation [15]. It depends only on the internal parameters  $K$  via

$$\omega = (KK^T)^{-1}, \quad (6)$$

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

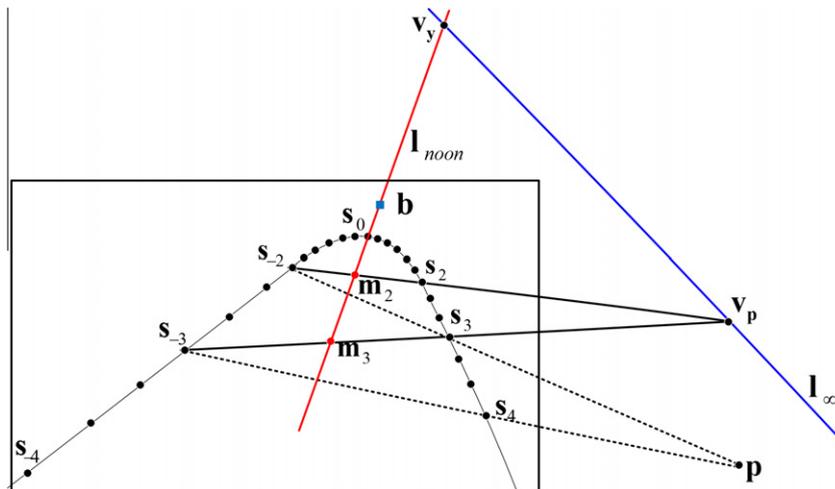
$$\omega \sim \begin{bmatrix} 1 & 0 & -u_0 \\ 0 & 1 & -v_0 \\ -u_0 & -v_0 & f^2 + u_0^2 + v_0^2 \end{bmatrix}, \quad (7)$$

where  $(u_0, v_0)$  is the principal point that is known to be approximately at the center of a image [19,15], and  $f$  is the focal length. With the principal point's availability from the captured image (it approximately locates on the center of the image), the remaining constraint, i.e.  $f$ , can be achieved from the orthogonal relationship provided by

$$v_x^T \omega v_y = 0, \quad (8)$$

where  $v_x$  and  $v_y$  are vanishing points of lines with perpendicular directions.

The orthogonal vanishing points can be obtained from the geometric constraints embodied in the shadow trajectory. By exploring the shadow trajectory's property, we find that the shadow points with the same offset to the solar noon point indicated by the solar time are symmetric with respect to the local noon line  $l_{noon}$ , see Appendix A for details. As shown in Fig. 3, shadow points  $s_{-i}$  and  $s_i$  indicated by solar time have the same offset,  $i$  hours, with respect to solar noon point  $s_0$ , and therefore are symmetric with  $l_{noon}$ .  $s_{-i}$  is  $i$  hours before  $s_0$  while  $s_i$  after it. The pair  $(s_{-i}, s_i)$  is defined as a *paired shadow points*. The lines connecting the paired shadow points are parallel to each other and intersect at a vanish-



**Fig. 3.** Symmetric relationship in the shadow trajectory.  $s_i$  is the shadow points with the offset  $i$  hours to solar time 12:00.  $b$  is the imaged footprint of  $B$  in Fig. 2.  $l_{noon}$  is the symmetric axis of the shadow trajectory and  $l_\infty$  is the horizon.



$$\frac{|Y_h - Y_E|}{|X_h|} = \frac{|\sin \phi \cos h - \tan \delta \cos \phi|}{|\sin h|} = \frac{|\mathbf{DC}|}{|\mathbf{EC}|}, \quad (16)$$

where  $|\mathbf{DC}|$  and  $|\mathbf{EC}|$  are the shadow length's measurements in North–South and West–East direction, respectively.

### 6. Degenerate cases and error analysis

#### 6.1. Latitude location

The method degenerates when the location is close to the Earth's equator, where  $\phi = 0^\circ$  and there is no shadow cast on the ground plane. In the North/South poles, where  $\phi = 90^\circ$ , the shadow trajectory takes shape of a circle which is a special type of conic. The analemmatic sundial degenerates to the equatorial sundial for which the gnomon is perpendicular to the ground plane and the equatorial is parallel with the ground plane. Meanwhile, Eq. (16) degenerates as:

$$|\cot h| = \frac{|\mathbf{DC}|}{|\mathbf{EC}|}. \quad (17)$$

Theoretically, the geo-location estimation approach will not be affected by the simplified form in Eq. (17). Practically, however, the presence of shadow (the day time) depends on the sun's declination. For example, the presence of shadow decreases with the increasing of the latitude as the Sun moves from the equator to the Tropic of Capricorn. Particularly, when the sundial locates on the high-latitude region which ranges from  $66^\circ$  to  $90^\circ$ , the Sun appears a very short time and we cannot get sufficient shadow points or even no shadow points because of polar night.

#### 6.2. Degenerate viewing direction

Without loss of generality, we specify that the North–South direction is parallel to the world  $y$ -axis, i.e. the short axis of the analemmatic sundial is the world  $y$ -axis, and the long axis is the world  $x$ -axis.

Since the two used vanishing points  $\mathbf{v}_y$  and  $\mathbf{v}_x$  in Fig. 3 are along  $x$  and  $y$  axes separately, the degenerate case occurs when the principal viewing direction is approximately perpendicular to the  $x$  or  $y$  axis. Consequently, it is necessary to evaluate the vanishing point's estimation since the uncertainty is essentially in it when the view direction for the camera produces weak perspective. As an approach based on vanishing points, our method degenerates when the vanishing points along some directions are distant. Intuitively, vanishing points go to infinity when the parallel lines in the 3D space remain parallel in the image plane and the direction of the vanishing point is parallel to the image plane for which the third component of the homogeneous coordinates becomes to zero.

As shown in Fig. 5, the lines  $\mathbf{AB}$  and  $\mathbf{CD}$  are parallel in world coordinate system, and therefore define a vanishing point  $\mathbf{v}_p$  in the image plane. To simplify the analysis, we assume that  $\mathbf{A}$ ,  $\mathbf{B}$  and  $\mathbf{C}$  are constant, and the coordinate of the shadow point  $\mathbf{D}$  is  $(x_D, y_D)$ . Without loss of generality we assume that the line  $\mathbf{AB}$  is parallel with the  $x$ -axis which can be described as  $y = c$  ( $c$  is a constant) and  $\mathbf{C}$  is  $(0, 0)$ . The position of  $\mathbf{D}$  affects the vanishing point  $\mathbf{v}_p$ 's  $x$ -coordinate  $x_p$  via a function

$$x_p = \frac{cx_D}{y_D} = f(x_D, y_D), \quad (18)$$

and the error  $\varepsilon$  along the horizontal direction for the vanishing point can be described by

$$\varepsilon = \frac{\partial f}{\partial y_D} d_y + \frac{\partial f}{\partial x_D} d_x = \frac{cx_D}{y_D^2} d_y + \frac{c}{y_D} d_x. \quad (19)$$

Eq. (19) implies that, if shadow points  $\mathbf{A}$ ,  $\mathbf{B}$  and  $\mathbf{C}$  are assumed to be approximately constant, then the error  $\varepsilon$  varies rapidly as  $y_D$  decreases when  $x_D$  is specified. We can also conclude that the error  $\varepsilon$  depends on  $y_D$  and remain invariable to  $x_D$  when  $y_D$  is specified. As Fig. 5 shows,  $d_y$  corresponds to the error  $\varepsilon$  for  $\mathbf{v}_p$  horizontally. Therefore, for calibration algorithms based on vanishing points, it is important to make use of images having large perspective distortions.

### 7. Experimental results

We show the synthetic simulations for four experiments: first, zero-mean Gaussian noise is added to the synthetic shadow points; second, we evaluate the robustness of the proposed method with respect to the number of shadow observations; third, the effect of latitude location is also considered in our experiment. The last simulation is evaluated on the viewing direction of the camera. We then show the results for three set of real data. Finally, an application to image-based metrology on height ratio estimation for shadow-casting objects is presented.

#### 7.1. Computer simulation

The simulated camera has the focal length of  $f = 2000$ , the aspect ratio of  $\lambda = 1$ , the skew of  $\gamma = 0$ , and the principal point at  $u_0 = 200$  and  $v_0 = 150$ . The synthetic light source is at infinity with the declination of  $\delta = -23^\circ$ , the equation of time of  $t_{EOT} = 5.2$  min (14th December), the latitude of  $\phi = 39^\circ$ . The two stationary 3D objects locating at the northern hemisphere ( $39^\circ\text{N}$ ,  $117.2^\circ\text{E}$ ) have heights of 10 and 20 units from the ground plane, respectively. The distance between the two footprints on the ground plane is 20 units. Therefore, the locus of shadow positions on the ground plane, i.e. the shadow trajectory, is a determined smooth curve in our experiment since it depends on the altitude and the azimuth of the Sun and the vertical distance of the object from its footprint.

##### 7.1.1. Performance versus noise level

In this experiment, the simulated shadow points are collected from 8:00 am to 4:00 pm with a frame speed at 30 frames per second. Gaussian noise with zero mean and a standard deviation of  $\sigma \leq 2.0$  pixels was added to the image shadow points. The estimated camera parameters and geo-location were then compared with the ground truth. For each noise level, 100 independent trials were performed. For the focal length  $f$ , we measured the relative error with respect to its ground truth. The geo-location estimation was measured with absolute error.

As demonstrated in Section 4, four shadow points are sufficient to determine the calibration parameter with no requirement of footprint. Due to the important role played by the footprint in

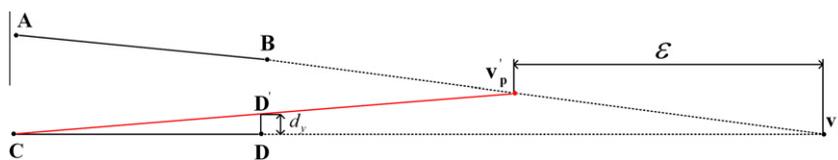


Fig. 5. Error analysis in the estimation of a vanishing point as the shadow point varies.

camera calibration, we evaluate our method in situations with and without known footprint location. Fig. 6a and b shows the average relative errors of  $f$  as functions of noise variance without and with footprint's influence. It can be seen that the two kinds of errors increase almost linearly as functions of the noise level and better results can be achieved by taking footprint into consideration. Since the footprint is required in the geo-location estimation, we employ it to get preferable results in the following experiments.

For  $\sigma = 1$ , the relative error of focal length  $f$  is 1.21%. For higher noise,  $\sigma = 2$ , which is regarded as the typical noise in practical calibration of our experiment, the relative error in focal length increases steadily to reach 3.66%. Meanwhile the maximum absolute error of latitude's mean value (Fig. 6c) is 5.7 with the noise level reaching 2 pixels. As can be seen in Fig. 6d, the absolute error of longitude compared with ground truth value is within  $\pm 0.38^\circ$  with the noise level reaching 2 pixels.

### 7.1.2. Comparison with the existing method

In [26], the authors argued that three pairs of shadow observations can determine the latitude. However, they require the world point casting the shadow on the ground plane to be visible in the image, which is not requirable in our method. As Fig. 7 shows, our method performs about as well as [26] after relieving the requirement of shadow casting object's visibility.

### 7.1.3. Performance versus number of observations

Another experiment performed was to examine the influence of numbers of shadow points selected for calibration and the

computation of latitude. Herein, the assumption is given that longitude has been determined using the proposed method in Section 4. Since we focus on the calibration and geo-location's robustness to the number of shadow observations, the time correction is assumed to be pre-determined. The selected shadow points should have the same offset with respect to the solar noon point, i.e. 12:00, and also in pairs. For example, the 4 points we used in the experiment were selected as: 8:00, 9:00, 15:00, 16:00 (solar time), (8:00, 16:00) and (9:00, 15:00) were two pairs which have the same offset with respect to 12:00. In another example, the 16 points we used were composed of 8 pairs: (8:00,16:00), (8:15,15:45), (9:00,15:00), (9:15,14:45), (10:00,14:00), (10:15,13:45), (11:00,13:00), and (11:15,12:45). We also added a noise level of  $\sigma = 2$  pixels to all projected image points. We repeated the process 100 times and computed the average errors. The number of observations varied from 4 to 32 in step of 4, and the results are shown in Fig. 8. It can be seen that the relative errors decrease as the number of shadow points increase, and the same situations hold for the latitude estimation. Both the focal length and latitude become stable when the number of observations reaches 24.

### 7.1.4. Performance versus latitude location

This experiment was to evaluate the performance of our approach with respect to the locations. This evaluation is interesting since the shape of trajectory rests on the latitude. Herein, the Sun's declination ( $\delta$ ) is  $-23^\circ$  for which the Sun irradiates at the Tropic of Capricorn and there is no sunshine or shadows at regions beyond the Arctic Circle ( $66^\circ\text{N}$ ), which is called polar night. However, when

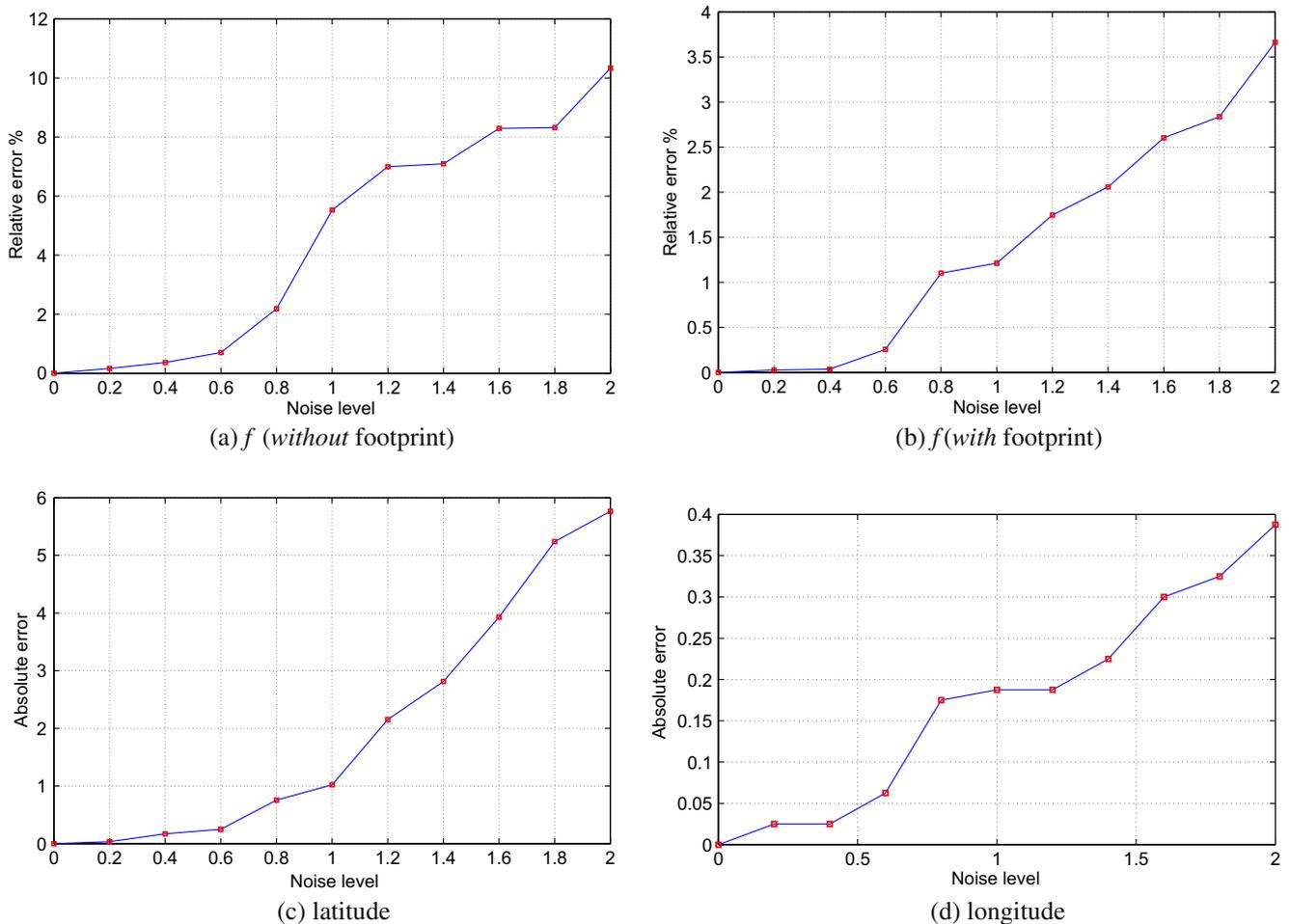


Fig. 6. Performance with respect to added random noise ranging from 0 to 2 pixels.

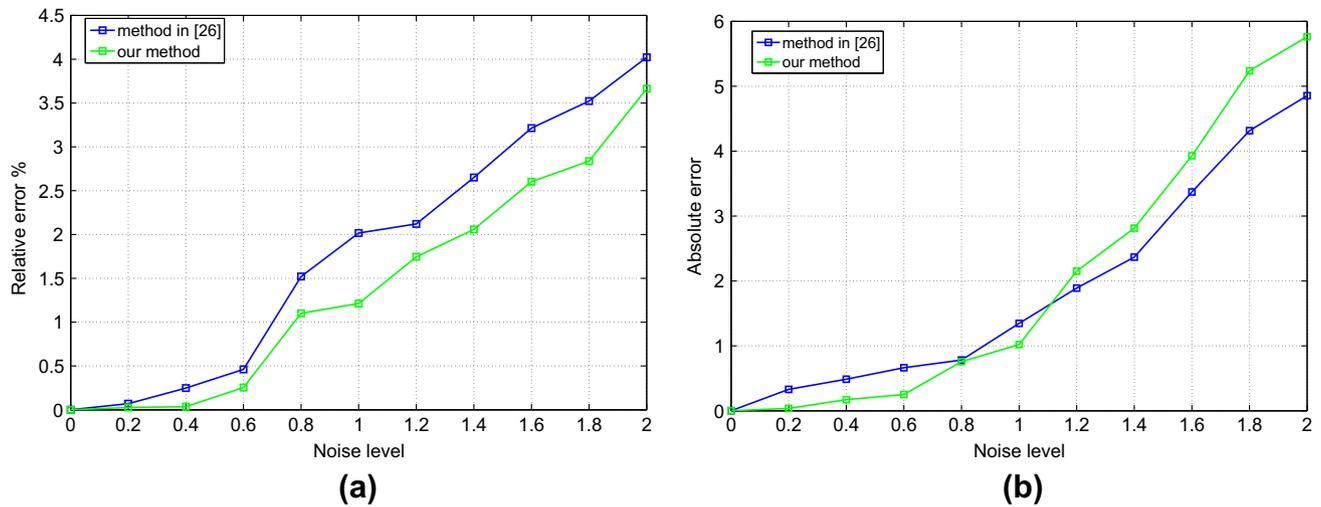


Fig. 7. Comparison with [26]. (a) and (b) are comparisons on the estimation of focal length and latitude, respectively.

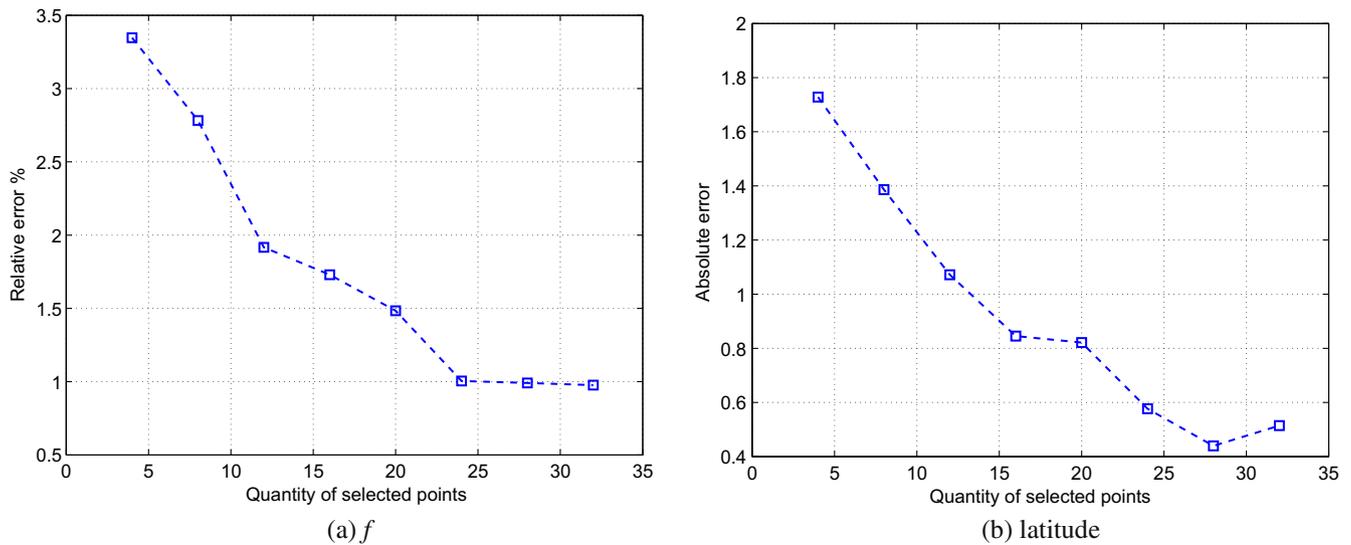


Fig. 8. Performance with respect to the number of shadow observations.

the Sun moves to the Tropic of Cancer with a  $\delta = 23^\circ$  for which the Antarctic Circle  $66^\circ\text{S}$  or even higher-latitude region has no sunshine whereas the Arctic Circle is polar day. Therefore, this experiment is evaluated on the latitude value ranging from  $20^\circ\text{N}$  to  $60^\circ\text{N}$  for  $\delta = -23^\circ$  and  $20^\circ\text{N}$  to  $80^\circ\text{N}$  for  $\delta = 23^\circ$ . The imaged points are added Gaussian noise with zero mean and a standard deviation  $\sigma = 1$ . The results in Fig. 9 validate the analysis of degenerate case in Section 6.1. From Fig. 9, we conclude that the proposed method degenerates when the objects locate on some region identified with lower latitude near to the equator such as  $20^\circ$  or even lower because of low latitude corresponds to a small-scaled trajectory which is sensitive to noise since the image have the same number of pixels. Moreover, a larger-scale trajectory can be traced out by the object locating on a high-latitude region. Consequently, high-latitude oriented calibration and geo-location estimation is less sensitive to noise and better estimating results can be achieved.

#### 7.1.5. Performance versus viewing direction

The last simulation performed was to examine the influence of the relative orientation between the principal view direction and

the ground plane. The camera internal parameters are fixed through this experiment. Shadow points are corrupted with a random Gaussian noise with zero mean and a standard deviation  $\sigma = 0.5$ . Then  $7 \times 7 = 49$  cameras (green squares in Fig. 10a) looking at the magenta star are selected to evaluate the performance. The cameras are located on a sphere centered at  $(0,0,0)$  with radius 300, which is above the ground plane. The horizontal trajectories are obtained by rotating the camera around the  $x$ - $y$  axis plane of the sphere from  $20^\circ$  to  $140^\circ$ , and rotations around the axis on  $y$ - $z$  axis plane from  $20^\circ$  to  $80^\circ$  form the vertical trajectories. As the camera rotates along the horizontal trajectory from right to left, the pitch rotation is described by a clockwise rotation which is a negative angle with a correspondingly decreasing value. As expected, the calibration and geo-location algorithms degenerate when the angle between the camera's looking-at vector and the  $x$ -axis' positive direction reaches  $80^\circ$  or above, seeing Fig. 10d and e. And here, the viewing direction is approximately parallel with the symmetric axis of the shadow conic for which the degenerate camera induces weak perspective effect. Consequently, the horizon line cannot be estimated robustly, as illustrated in Fig. 10c.

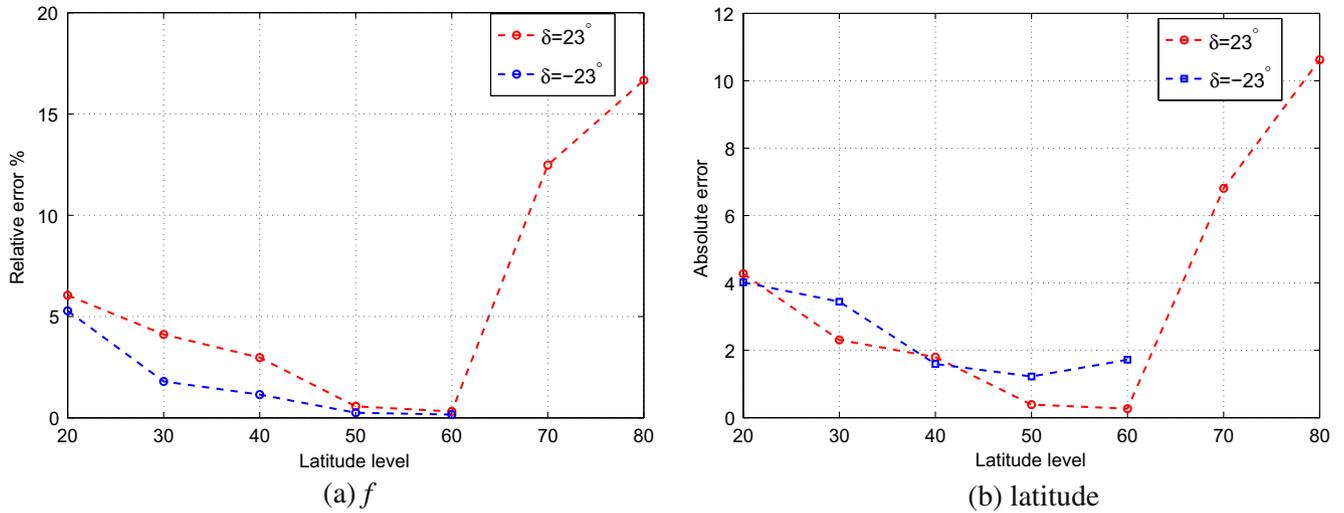


Fig. 9. Performance with respect to shadow casting objects' varying locations for  $\delta = 23^\circ$  and  $\delta = -23^\circ$ , respectively.

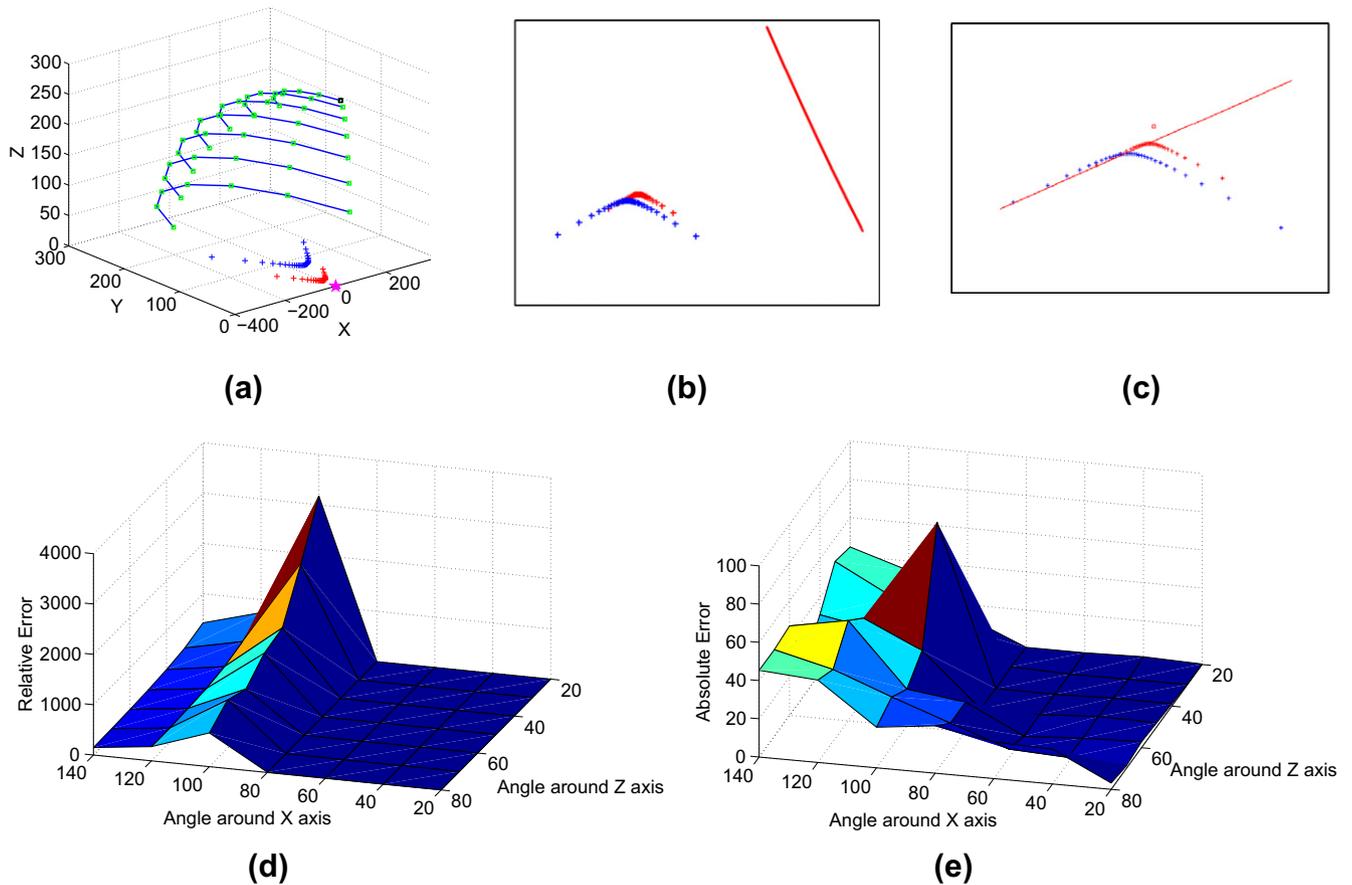


Fig. 10. Performance with respect to the relative orientation between the principal ray and the ground plane. (a) The distribution of cameras labeled with green squares used to evaluate varying camera orientation. The blue and red markers are two shadow trajectories. (b) and (c) Two image trajectories projected by the camera labeled with black and the last camera on the same horizontal trajectory in (a). The red line is horizon and the red square in (c) is the principal point. (d) and (e) are error surfaces of focal length and latitude for all of cameras in (a) and blue denotes small errors. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

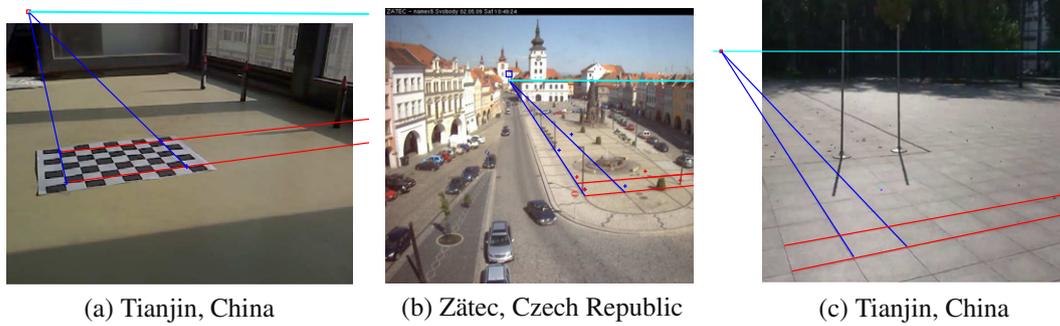
7.2. Real data

The method is also applied to three sets of real data as shown in Figs. 12 and 13. In the first test, a surveillance camera was set up indoor to capture two vertical objects and casted shadows between

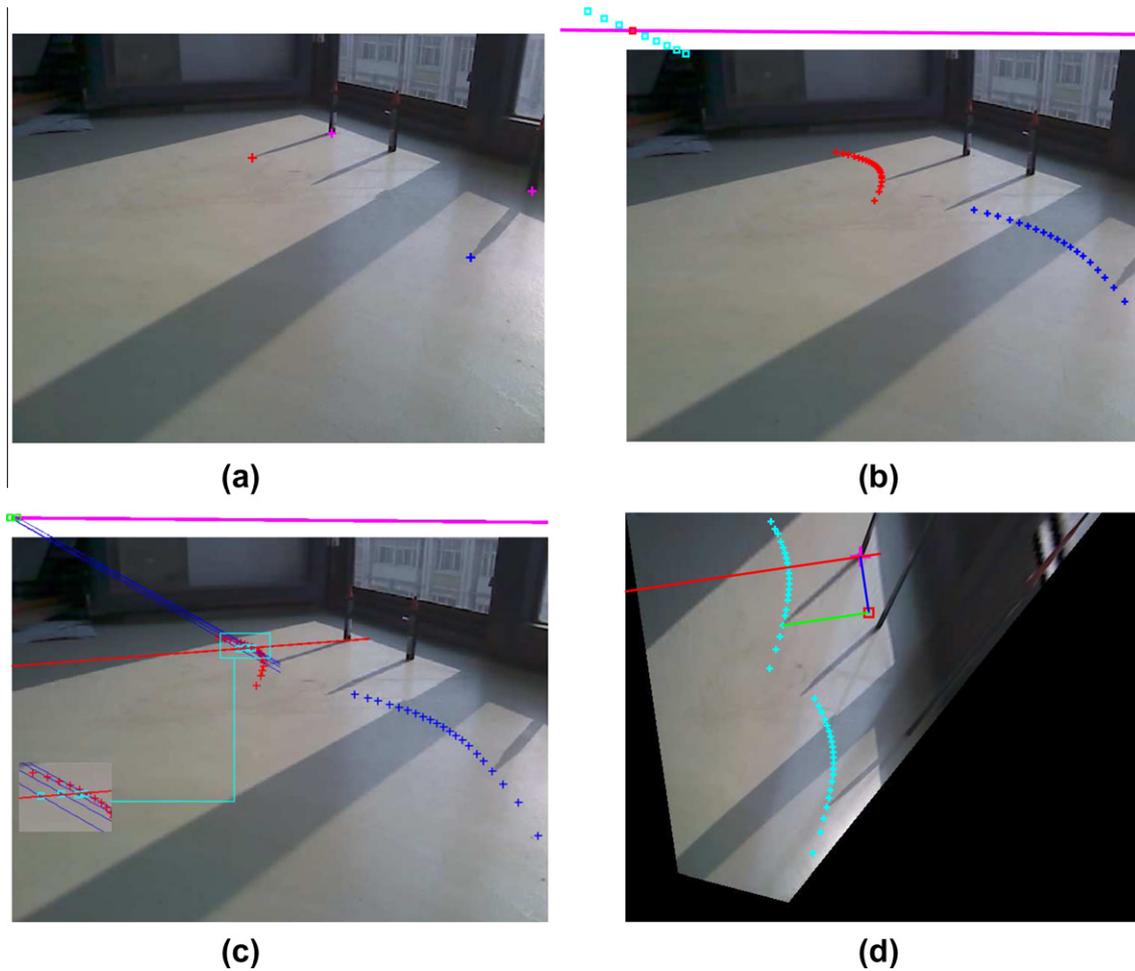
9:00 am and 3:00 pm (Beijing time). This data set was captured on the 14th December at longitude 117.2°E and latitude 39.1°N, Tianjin, China. The casted shadow trajectories of two objects in the scene were marked in red and blue in Fig. 12a, and the footprints (magenta) were selected for the computation of geo-location. The

**Table 1**  
Calibration and geo-location on real images.

City	Date	$f_{GT}$	$f_{est}$	Longitude <sub>GT</sub>	Longitude <sub>est</sub>	Latitude <sub>GT</sub>	Latitude <sub>est</sub>
Tianjin	14th December	663	687	117.2°	117.45°	39.1°	37.7°
Zátec	2nd May	568	13.5°	13.77°	50.3°	48.37°	
Tianjin	22th Oct	804	824	117.2°	117.675°	39.1°	36.8°



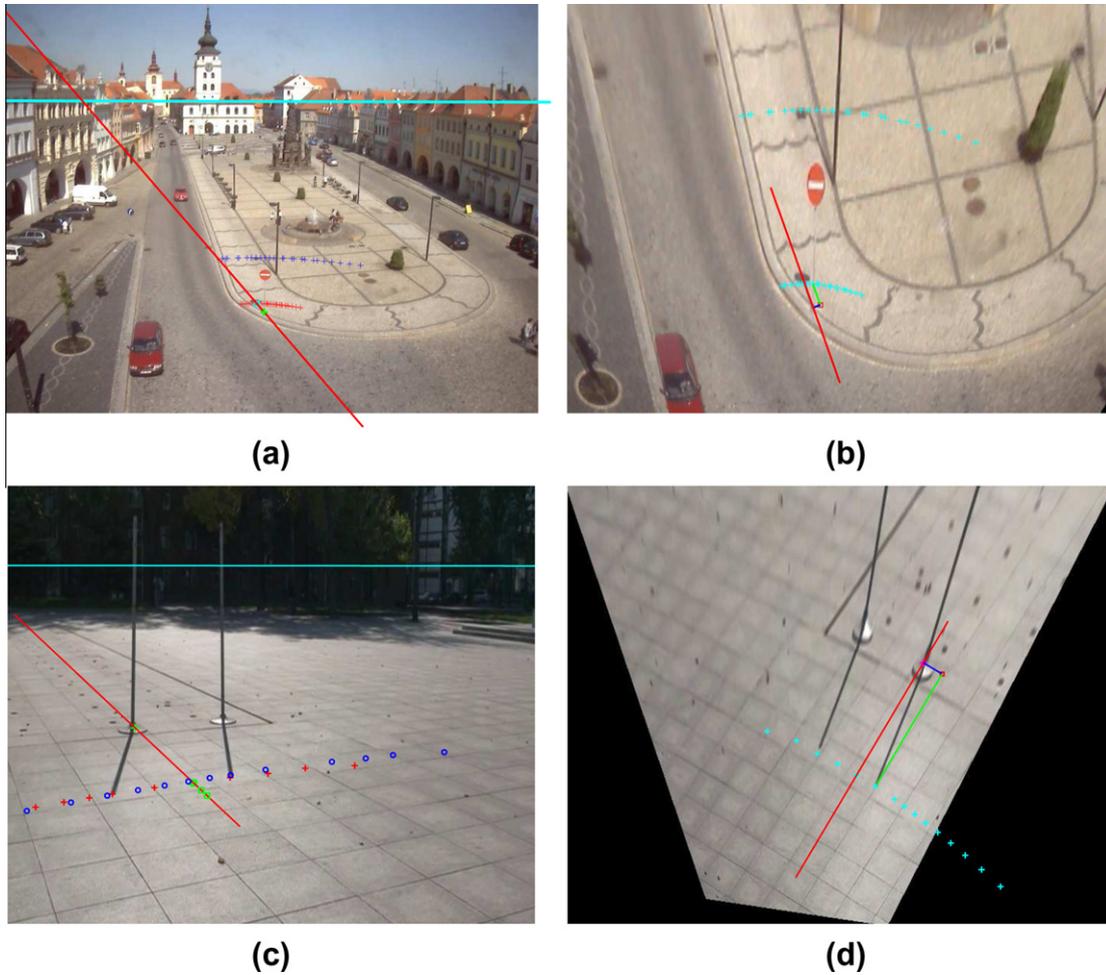
**Fig. 11.** Calibration using the orthogonal vanishing points recovered from real images. Two set of lines labeled with red and blue provide the orthogonal relationship to calibrate the camera. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 12.** Camera calibration geo-location estimation on real images. (a) A real image captured by a static indoor camera. (b) Two shadow trajectories are labeled with red and blue and the top magenta line is the estimated horizon line. (c) The red line is the estimated noon line and the cyan squares are midpoints. The blue lines connecting solar points converge at vanishing points labeled with green squares. (d) The latitude estimation illustration on the calibrated image. The green and blue are measurements along the North–South and West–East, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

shadow points and footprints were picked manually. Fig. 12b shows the two objects’ shadow points picked over time and the recovery of horizon using the proposed method. The time correc-

tion was determined by searching solar points having the same offsets with respect to the solar noon. The lines connecting the solar points with the correct time correction converge at a vanishing



**Fig. 13.** Camera calibration and geo-location estimation on images captured by outdoor cameras. (a and c) Two images captured by a static webcam and a fixed camera, respectively. Two shadow trajectories are labeled with red and blue. The cyan and red line are the estimated horizon and solar noon line. The green squares are midpoints and the green plus is the footprint. (b and d) The latitude estimation illustration on the calibrated images. Note that in (a) two trajectories are manual-picked while in (c) shadow trajectories are determined using the semi-automatic detection method in Section 3.3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

point, which lies on the horizon (the red square in Fig. 12b). Otherwise, the node is just a point with a certain distance from the horizon and all the searching nodes (cyan squares in Fig. 12b) locate on the same line. Consequently, only the correct time correction can induce the minimal distance between the point i.e. vanishing point and the horizon. Two shadow trajectories and the local noon line's estimating process are illustrated in Fig. 12c. In Fig. 12d, the rectified conic trajectory has the same shape with the real trajectory in 3D world since the transformation between the image and the rectified world plane is a similarity by fixing its circular points.

Another performance on real images was evaluated on the images captured by a webcam installed in Zâtec, Czech Republic. In this set, 25 images were captured live from a square of Z-âtec, using one of the webcams available online at <http://siemens.mesto-zatec.cz/webcam.html>. The ground truth for this data set is as follows: longitude 13.5°E, latitude 50.3°N, and the declination angle 15.15°. As illustrated in Fig. 13, two shadow trajectories were composed of manual-picked shadow points from images captured from 10:00 am to 4:00 pm (clock time). We estimate the longitude as 13.77°E with no requirement of footprint's visibility in accordance with its temporal correlation, the latitude as 48.37°N. Therefore, the absolute error of longitude estimation is equivalent to 19 km in remote geographic location. In the third real set, a static camera was installed to capture two objects in a square of Tianjin,

China on 22th October. As shown in Fig. 13c, shadow trajectories are estimated semi-automatically. Calibration and geo-location results are given in the third row of Table 1.

*Note:* The ground truth values of focal length for three experiments on real images were obtained by using calibrated objects, seeing Fig. 11. The calibration was implemented using the orthogonal relationship provided by the grid in Fig. 11a and the ground tile in Fig. 11b and c. Meanwhile, the focal length achieved from the first set of real frames using the proposed method is 687 with the relative error 3.57%. The calibration and geo-location results for the three sets of real images compared with the ground truth are shown in Table 1.<sup>2</sup>

### 7.3. Application to image-based metrology

In this section, we show that the height ratio of the two world points  $\mathbf{t}_i$  and  $\mathbf{t}_j$  can be recovered from their shadow trajectories over time see Fig. 2, even if the two points cannot be seen in the images. For this purpose, first we recover the affine property of the ground plane by a projective transformation (or affine rectification):

<sup>2</sup>  $f_{GT}$  and  $f_{est}$  denotes the ground truth and estimated values for focal length. Such references hold true for latitude and longitude.

$$\mathbf{H}_p = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ l_1 & l_2 & l_3 \end{pmatrix}, \quad (20)$$

where  $[l_1 \ l_2 \ l_3]^T$  is the vanishing line  $\mathbf{l}_\infty$ . As a result, the affine properties such as the ratio of lengths on parallel lines are invariant. Since  $\Delta \mathbf{b}_i \mathbf{s}_{im} \mathbf{s}_{in}$  and  $\Delta \mathbf{b}_j \mathbf{s}_{jm} \mathbf{s}_{jn}$  are similar as demonstrated in Section 4, we have the equivalence

$$\frac{\mathbf{s}_{ik} \mathbf{s}_{il}}{\mathbf{s}_{jk} \mathbf{s}_{jl}} = \frac{\mathbf{t}_i \mathbf{b}_i}{\mathbf{t}_j \mathbf{b}_j}, \quad \forall i, j, k, l. \quad (21)$$

Therefore, given  $n$  observations of shadows of two object  $i$  and  $j$ , we have  $C_n^2$  solutions of  $\frac{\mathbf{t}_i \mathbf{b}_i}{\mathbf{t}_j \mathbf{b}_j}$ , from which the optimal solution can be computed as a weighted mean. The weight is dependent on the relative distance between the  $k$  and  $l$ . The estimated height ratio between two shadow-casting objects in Fig. 12 is 1.06. And in Fig. 13a, the estimated height ratio between the lamp and the traffic sign (without the circle head) is 2.95. Thus, the main advantage of this method is not only the looser restriction on the scene, but also the estimation of the horizontal line and the relative height between the unknown shadow-casting objects just using two shadow trajectories. Since scenes observed over time is typical in outdoor surveillance systems and on account of the requirement in our method are cast shadows of unknown objects, it has tremendously potential applications in video surveillance systems.

## 8. Conclusion and future work

This paper addresses the problem of computing the geo-location from only two shadow trajectories and corresponding footprints on uncalibrated images. Camera calibration is achieved by deriving the horizon line and the metric rectification with only four shadow observations. Compared with the conventional methods that require images of some precisely machined calibration, our technique utilizes cast shadows by the Sun, which are frequently found in natural environments especially in outdoors, and demands no measurements of any distance or angle in 3D world. The fact that prior knowledge of the 3D coordinates of the stationary objects or light source orientation is not required, makes the proposed method a versatile utility that can be used without requiring a precisely machined calibration or other geometry structure. More importantly, it relaxes some of the constraints found in other geo-location estimating method using the altitude and azimuth angles of the Sun orientations in the sky, and hence is especially useful in cases where only limited information is available. These advantages make our method an efficient calibration technique that could be especially suitable for outdoor video surveillance systems. Experimental results on both synthetic and real data verify the validity and usefulness of the method.

## Appendix A

The position of a shadow point on the ground plane for a vertical object can be described by the equation:

$$x(h) = \cot(a) \sin(A), \quad (22)$$

$$y(h) = \cot(a) \cos(A), \quad (23)$$

where  $a$  and  $A$  are the sun's altitude and azimuth which can be defined by the following relationships:

$$a = \arcsin(\sin \phi \sin \delta + \cos \phi \cos \delta \cos h), \quad (24)$$

$$A = \arctan\left(\frac{\sin h}{\sin \phi \cos h - \cos \phi \tan \delta}\right), \quad (25)$$

and  $h = (T_{24} - 12) \times 15^\circ$  (in deg.) is the hour angle. The paired shadow points with the same offset to the solar noon point (i.e. the shadow point casted by the vertical object at 12:00 (solar time))

of the sundial can be indicated by  $(x(h), y(h))$  and  $(x(-h), y(-h))$ . Then we have

$$\begin{cases} x(-h) = \cot(a) \sin(-A) = -\cot(a) \sin(A) = -x(h), \\ y(-h) = \cot(a) \cos(-A) = \cot(a) \cos(A) = y(h). \end{cases} \quad (26)$$

Hence, the shadow points with the same offset to the solar noon point are symmetric with respect to the local noon line  $\mathbf{l}_{noon}$ . Moreover, the local noon line of the sundial is defined as the line connecting the footprint and the solar noon point.

## References

- [1] M. Antone, M. Bosse, Calibration of outdoor cameras from cast shadows, in: Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, 2004, pp. 3040–3045.
- [2] J. Bouguet, P. Perona, 3d photography on your desk, in: Proceedings of the IEEE ICCV, IEEE Computer Society, 1998, pp. 43–50.
- [3] X. Cao, H. Foroosh, Camera calibration using symmetric objects, IEEE Trans. Image Process. 15 (11) (2006) 3614–3619.
- [4] X. Cao, H. Foroosh, Camera calibration and light source estimation from images with shadows, Comput. Vis. Image Understanding 4 (2) (2007) 127–140.
- [5] X. Cao, Y. Shen, M. Shah, H. Foroosh, Single view compositing with shadows, Vis. Comput. 21 (8–10) (2005) 639–648.
- [6] Y. Caspi, M. Werman, Vertical parallax from moving shadows, in: Proceedings of the IEEE CVPR, 2006, pp. 2309–2315.
- [7] Q. Chen, H. Wu, T. Wada, Camera calibration with two arbitrary coplanar circles, in: Proceedings of the ECCV, 2004, pp. 521–532.
- [8] T.E. Choe, K. Ramnath, M.W. Lee, N. Haering, Image transformation for object tracking in high-resolution video, in: Proceedings of the ICPR, 2008, pp. 1–4.
- [9] Y. Chuang, D. Goldman, B. Curless, D. Salesin, R. Szeliski, Shadow matting and compositing, ACM Trans. Graph. 22 (3) (2003) 494–500.
- [10] R. Collins, Y. Tsin, Calibration of an outdoor active camera system, in: Proceedings of the CVPR, 1999, pp. 528–534.
- [11] C. Colombo, A.D. Bimbo, F. Pernici, Metric 3d reconstruction and texture acquisition of surfaces of revolution from a single uncalibrated view, IEEE Trans. Pattern Anal. Mach. Intell. 27 (1) (2005) 99–114.
- [12] A. Criminisi, I. Reid, A. Zisserman, Single view metrology, Int. J. Comput. Vis. 40 (2) (2000) 123–148.
- [13] Martin A. Fischler, Robert C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, Commun. ACM 24 (6) (1981) 381–395.
- [14] R.I. Hartley, Eric Hayman, L. De Agapito, I. Reid, Camera calibration and the search for infinity, in: Proceedings of the IEEE ICCV, 1999, pp. 510–517.
- [15] R.I. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge University Press, 2004.
- [16] J. Hays, A. Efros, Im2gps: estimating geographic information from a single image, in: Proceedings of the CVPR, 2008, pp. 1–8.
- [17] T. Horprasert, D. Harwood, L.S. Davis, A statistical approach for real-time robust background subtraction and shadow detection, in: Proceedings of the IEEE ICCV, Frame Rate Workshop, 1999, pp. 1–19.
- [18] O. Javed, M. Shah, Tracking and object classification for automated surveillance, in: Proceedings of the ECCV, 2002, pp. 343–357.
- [19] M.K. Johnson, H. Farid, Detecting photographic composites of people, in: Proceedings of the IWDW, 2007.
- [20] A. Klaus, J. Bauser, K. Karner, P. Elbschger, R. Perko, H. Bischof, Camera calibration from a single night sky image, in: IEEE CVPR, 2004, pp. 151–157.
- [21] D. Kriegman, P. Belhumeur, What shadows reveal about object structure, in: Proceedings of the ECCV, 1998, pp. 399–414.
- [22] C. Lin, A. Huertas, R. Nevatia, Detection of buildings using perceptual groupings and shadows, in: Proceedings of the IEEE CVPR, 1994, pp. 62–69.
- [23] I. Mikic, P. Cosman, G. Kogut, M. Trivedi, Moving shadow and object detection in traffic scenes, in: Proceedings of the ICPR, 2000, pp. 321–324.
- [24] N. Roman, R. Speyer, N. Jacobs, S. Satkin, R. Pless, Geolocating static cameras, in: Proceedings of the ICCV, 2007, pp. 1–6.
- [25] S. Nadimi, B. Bhanu, Physical models for moving shadow and object detection in video, IEEE Trans. Patt. Anal. Mach. Intell. 26 (8) (2004) 1079–1087.
- [26] Imran NJunejo, Hassan Foroosh, Estimating geo-temporal location of stationary cameras using shadow trajectories, in: Proceedings of the ECCV, 2008, pp. 318–331.
- [27] T. Okabe, I. Sato, Y. Sato, Spherical harmonics vs. haarwavelets: basis for recovering illumination from cast shadows, in: Proceedings of the IEEE CVPR, 2004, pp. 50–57.
- [28] L. Petrovic, B. Fujito, L. Williams, A. Finkelstein, Shadows for cel animation, in: Proceedings of the ACM SIGGRAPH, 2000, pp. 511–516.
- [29] M. Pollefeys, R. Koch, L.V. Gool, Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters, Int. J. Comput. Vis. 32 (1) (1999) 7–25.
- [30] A. Prati, I. Mikic, M. Trivedi, R. Cucchiara, Detecting moving shadows: algorithms and evaluation, IEEE Trans. Patt. Anal. Mach. Intell. 25 (7) (2003) 918–923.
- [31] Rene R.J. Rohr, Sundials: History, Theory, and Practice, Dover Publications, 1996.

- [32] I. Sato, Y. Sato, K. Ikeuchi, Illumination from shadows, *IEEE Trans. Patt. Anal. Mach. Intell.* 25 (3) (2003) 290–300.
- [33] J. Stauder, R. Mech, J. Ostermann, Detection of moving cast shadows for object segmentation, *IEEE Trans. Multimedia* 1 (1) (1999) 65–76.
- [34] K. Sunkavalli, F. Romeiro, W. Matusik, T. Zickler, H. Pfister, What do color changes reveal about an outdoor scene? in: *Proceedings of the IEEE CVPR*, 2008, pp. 594–601.
- [35] L. Van Gool, M. Proesmans, A. Zisserman, Planar homologies as a basis for grouping and recognition, *Image Vis. Comput.* 16 (1998) 21–26.
- [36] L. Wang, F.C. Wu, Z.Y. Hu, Multi-camera calibration with one-dimensional object under general motions, in: *Proceedings of the ICCV*, 2007, pp. 1–7.
- [37] K.Y. Wong, R.S. P Mendonca, R. Cipolla, Camera calibration from surfaces of revolution, *IEEE Trans. Patt. Anal. Mach. Intell.* 25 (2) (2003) 147–161.
- [38] Yaser Yacoob, Larry Davis, Segmentation of planar objects and their shadows in motion sequences, *Int. J. Comput. Vis.* 67 (1) (2006) 53–69.
- [39] H. Zhang, G. Zhang, K.Y.K. Wong, Camera calibration with two spheres: linear approaches, in: *IEEE ICIP*, 2005, pp. 1150–1153.
- [40] W. Zhang, J. Kosecka, Image based localization in urban environments, in: *Proceedings of the 3DPVT*, 2006, pp. 33–40.
- [41] Z. Zhang, Camera calibration with one-dimensional objects, *IEEE Trans. Patt. Anal. Mach. Intell.* 26 (7) (2004) 892–899.