Geo-location estimation from two shadow trajectories

Lin Wu

School of Computer Science and Technology
Tianjin University, China
wulin@tju.edu.cn

Xiaochun Cao

School of Computer Science and Technology
Tianjin University, China
xcao@tju.edu.cn

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 can be estimated from two observed shadow trajectories. Our contribution is that we use the design of the 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 and number of observations.

1. Introduction

Estimating geographic information is a high-level problem in computer vision [19, 8, 12, 17, 14]. A lot of cameras are mounted over the earth to undertake video surveillance, observe weather patterns, monitor coastline or navigation. Geo-location can provide plenty of background information including local climate, average temperature, rainfall and population density [8].

Currently, there are mainly three types of geolocating methods: feature matching based methods [19, 8, 12], photometry based method [17] and geometry based methods [14]. 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 are observed objects, e.g. the solar shadows, which are able

to provide necessary geometric relationships among geographic location, scene geometry, and camera orientation. Our approach belongs to this line of research which needs to remove the distortions as well as [14].

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. Shadows are important because they provide important visual cues, e.g. depth and shape, in our perception of the world [3, 4]. They are also interesting for camera calibration [2] and useful for computer vision applications such as planar object detection [18], visual surveillance [9, 13] and inverse lighting [15]. There has been some investigation of outdoor camera calibration (e.g. [5]) while most methods attempt to eliminate shadows for better performance rather than exploiting their properties to estimate camera parameters.

In this paper, we take advantage of the fact that most of the webcams and surveillance cameras remain static, and propose a new method to recover the camera calibration and geo-location directly from shadow points. 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 the analemmatic sundial to obtain geographical information. Compared to existing methods [14, 8], our approach does not require shadow casting objects's visibility and any GPS information.

2. Related work

Vanishing points and the horizon line play central roles in perspective geometry [7] for applications such as camera calibration and image-based metrology [6]. 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 [2], 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, 2] require the entire objects and their

shadows to be present in the image.

Junejo and Foroosh [14] 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 ([14] requires the entire object to be visible since they rely on the estimation of the Azimuth and Altitude of the sun).

Sunkavalli *et al.* [17] 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 computation of camera calibration parameters and proposes a novel method to calculate a stationary camera's longitude and latitude. In Section 5, we verify the proposed method on both simulated and real data. Section 6 concludes this work.

3. Preliminaries

3.1. Introduction of the analemmatic sundial

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. 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 4.2 that from the time stamped observations of solar shadows we can recover the longitude, latitude and camera calibration.

Figure 1. Automatic shadow trajectory detection. (a) A frame of real data. (b) Background. (c) Shadow point detection. (d) Shadow trajectory.

3.2. About time: local and standard time

Sundials indicate the local solar 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, and its rotational axis is not perfectly perpendicular to its orbit, there are 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. 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}), which is either 0 or 1, shifts the official time away from solar time by an hour advanced in some countries during the summer.

3.3. Shadow trajectory detection

We employ a semi-automatic approach 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. 1 (b)). 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. 1 (c)). Shown in Fig. 1 (d) is the shadow trajectory of the left object that is made up of shadow tips at a quarter's interval.

in Fig. 3. The conic's symmetric axis is exactly the North-South direction. i.e. the analemmatic sundial's short axis.

There exists an offset between the solar time indicated by the sundial and the standard clock time because of equation of time, difference between central meridian and local meridian (i.e. t_{off}), and daylight saving time. The local longitude can be achieved by correcting the reference longitude with a longitude correction t_{lc} which is computed as

$$t_{lc} = t_{off} + t_{EOT} - t_{DST}, \quad (6)$$

Since every 1° movement of the sun is equivalent to 4 minutes, the longitude correction is applied to the sundial as

$$L = L_{meridian} - 0.25 \times t_{lc} \quad (7)$$

where $L_{meridian}$ is the central meridian of the time zone which indicates the standard time, L is the local meridian where the shadow casting object locates. In addition, the equation of time and daylight saving time are available from the date when the video is recorded.

In terms of the design of the sundial, the horizontal dimension is defined as:

$$X_{\bar{h}} = \sin \bar{h}, \quad (8)$$

where $\bar{h} = (T_{24} - 12) \times 15^\circ$ is the hour angle, T_{24} is the solar time in hours that can be calculated from the longitude correction and clock time. The vertical dimension is:

$$Y_{\bar{h}} = \sin \phi \cos \bar{h} \quad (9)$$

The footprint's location $\mathbf{E} = [0, Y_E]^T$ relative to the central point of the analemmatic sundial is:

$$Y_E = \tan \delta \cos \phi \quad (10)$$

where δ is the sun's declination which varies from -23° to 23° during a year. Therefore, the latitude ϕ is estimated as

$$\frac{|Y_{\bar{h}} - Y_E|}{|X_{\bar{h}}|} = \frac{|\sin \phi \cos \bar{h} - \tan \delta \cos \phi|}{|\sin \bar{h}|} = \frac{|\mathbf{DC}|}{|\mathbf{EC}|}. \quad (11)$$

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

5. Experimental results

5.1. 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 $t_{EOT} = 5.2$ minutes (14^{th} Dec), the latitude of $\phi = 39^\circ$. The two stationary 3D objects locating at the northern hemisphere (39° N, 117.2° 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.

Figure 4. Performance with respect to added random noise ranging from 0 to 2 pixels.

5.1.1 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. 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 geolocation estimation was measured with absolute error.

Due to the important role played by the footprint in camera calibration, we evaluate our method in situations with and without known footprint location. Fig. 4 (a) and (b) show 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, $\sigma = 2$, the relative error in focal length increases steadily to reach 3.66%. Meanwhile the maximum absolute error of latitude's mean value (Fig. 4 (c)) is 5.7 with the noise level reaching 2 pixels. In Fig. 4 (d), the absolute error of longitude compared with ground truth value is within $\pm 0.38^\circ$ when the noise level reaches 2 pixels.

5.1.2 Performance versus number of observations

Herein, the assumption is given that longitude has been determined using the proposed method in Section 4. The selected shadow points should have the same offset with re-

Figure 5. Performance with respect to the number of shadow observations.

(a) Tianjin, China (b) Zatec (c) Tianjin, China
Figure 6. 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.

spect 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. 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. 5. 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.

5.2. Real data

We also applied our method to three sets of real data as shown in Fig. 7 and Fig. 8. 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 dataset 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. 7 (a), and the footprints (magenta) were selected for the computation of geolocation. The shadow points and footprints were picked manually. Fig. 7 (b) shows the two objects' shadow points picked over time and the recovery of horizon. The time correction was determined by searching solar points having the same offsets with respect to the solar noon. The lines

Figure 7. Camera calibration and geolocation estimation on real images. (a) Two shadow trajectories are labeled with red and blue in a real frame captured by an indoor camera. (b) 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. (c) The latitude estimation illustration on the calibrated image. The green and blue are measurements along the North-South and West-East, respectively.

Figure 8. Geolocation estimation on images captured by outdoor cameras. (a) and (c) Images captured by a static webcam and 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 image.

connecting the solar points with the correct time correction converge at a vanishing point, which lies on the horizon (the red square in Fig. 7 (b)). Otherwise, the node is just a point with a certain distance from the horizon and all the searching nodes (cyan squares in Fig. 7 (b)) 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. The latitude estimating process are illustrated in Fig. 7 (c).

Another performance was evaluated on the images

captured by a webcam installed in Zätec, Czech Republic. In this set, 25 images were captured live from a square of Zätec, 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. 8 (a), 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, the latitude as 48.37° N. Therefore, the absolute error of longitude estimation is equivalents 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.8 (c), shadow trajectories are estimated automatically. Calibration and geolocation results are given in the third row of Table 1.

Note: The ground truth values of focal length for the three experiments were obtained by using calibrated objects, seeing Fig. 6. The calibration and geolocation results for the three sets of real images compared with the ground truth are shown in Table 1.¹

Table 1. Calibration and geolocation on real images

City	Date	f_{GT}	f_{est}	$long_{GT}$	$long_{est}$	lat_{GT}	lat_{est}
Tianjin	14 th Dec	663	687	117.2°	117.45°	39.1°	37.7°
Zätec	2 nd May	568	604	13.5°	13.77°	50.3°	48.37°
Tianjin	22 th Oct	804	824	117.2°	117.675°	39.1°	36.8°

6. Conclusion

This paper addresses the problem of computing the geolocation from only two shadow trajectories from uncalibrated images. Camera calibration is achieved by deriving the horizon line and the metric rectification with only four shadow observations, which demands no measurements of any distance or angle in 3D world. More importantly, it relaxes the constraints found in other geo-location estimating methods using the altitude and azimuth angles of the sun orientations, 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.

7. *Acknowledgement

This work was supported by NSFC (No. 60905019), One Hundred Talents Programme of Tianjin, Scientific Research

¹ f_{GT} and f_{est} denotes the ground truth and estimated values for focal length. Such references (lat_{GT} , lat_{est}) ($long_{GT}$, $long_{est}$) have the same meaning for latitude and longitude.

Foundation for the Returned Overseas Chinese Scholars, SEM., Open Projects Program of NLPR.

References

- [1] M. Antone and M. Bosse. Calibration of outdoor cameras from cast shadows. In *Proc. IEEE Int. Conf. Systems, Man and Cybernetics*, pages 3040–3045, 2004.
- [2] X. Cao and H. Foroosh. Camera calibration and light source estimation from images with shadows. *Computer Vision and Image Understanding*, 4(2):127–140, 2007.
- [3] Y. Caspi and M. Werman. Vertical parallax from moving shadows. In *Proc. IEEE CVPR*, pages 2309–2315, 2006.
- [4] Y. Chuang, D. Goldman, B. Curless, D. Salesin, and R. Szeliski. Shadow matting and compositing. *ACM Trans. Graph.*, 22(3):494–500, 2003.
- [5] R. Collins and Y. Tsing. Calibration of an outdoor active camera system. In *Proc. CVPR*, pages 528–534, 1999.
- [6] A. Criminisi, I. Reid, and A. Zisserman. Single view metrology. *Int. J. Comput. Vision*, 40(2):123–148, 2000.
- [7] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2004.
- [8] J. Hays and A. Efros. Im2gps: estimating geographic information from a single image. In *Proc. CVPR*, pages 1 – 8, 2008.
- [9] O. Javed and M. Shah. Tracking and object classification for automated surveillance. In *Proc. ECCV*, pages 343–357, 2002.
- [10] M. K. Johnson and H. Farid. Detecting photographic composites of people. *Proc. IWDW*, 2007.
- [11] F. Lu, X. Cao, Y. Shen, and H. Foroosh. Camera calibration from two shadow trajectories. In *Proc. ICPR*, pages II: 1–4, 2006.
- [12] N. R. R. S. N. Jacobs, S. Satkin and R. Pless. Geolocating static cameras. In *Proc. ICCV*, pages 1 – 6, 2007.
- [13] S. Nadimi and B. Bhanu. Physical models for moving shadow and object detection in video. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(8):1079–1087, 2004.
- [14] I. N. Junejo and H. Foroosh. Estimating geo-temporal location of stationary cameras using shadow trajectories. In *Proc. ECCV*, pages 318 – 331, 2008.
- [15] T. Okabe, I. Sato, and Y. Sato. Spherical harmonics vs. haar-wavelets: Basis for recovering illumination from cast shadows. In *Proc. IEEE CVPR*, pages 50–57, 2004.
- [16] R. R. Rohr. *Sundials: History, Theory, and Practice*. Dover Publications, 1996.
- [17] K. Sunkavalli, F. Romeiro, W. Matusik, T. Zickler, and H. Pfister. What do color changes reveal about an outdoor scene? In *Proc. IEEE CVPR*, pages 594–601, 2008.
- [18] Y. Yacoob and L. Davis. Segmentation of planar objects and their shadows in motion sequences. *Int. J. Comput. Vision*, 67(1):53–69, 2006.
- [19] W. Zhang and J. Kosecka. Image based localization in urban environments. In *Proc. 3DPVT*, pages 33 – 40, 2006.