Determining the Geographical Location of Image Scenes based on Object Shadow Lengths

Frode Eika Sandnes

Abstract Many studies have addressed various applications of geo-spatial image tagging such as image retrieval, image organisation and browsing. Geo-spatial image tagging can be done manually or automatically with GPS enabled cameras that allow the current position of the photographer to be incorporated into the meta-data of an image. However, current GPS-equipment needs certain time to lock onto navigation satellites and these are therefore not suitable for spontaneous photography. Moreover, GPS units are still costly, energy hungry and not common in most digital cameras on sale. This study explores the potential of, and limitations associated with, extracting geo-spatial information from the image contents. The elevation of the sun is estimated indirectly from the contents of image collections by measuring the relative length of objects and their shadows in image scenes. The observed sun elevation and the creation time of the image is input into a celestial model to estimate the approximate geographical location of the photographer. The strategy is demonstrated on a set of manually measured photographs.

Keywords Geo-spatial tagging · Image content analysis · Image classification · Webcam analysis

F. E. Sandnes (✉)
Faculty of Engineering, Oslo University College,
P.O. Box 4, St. Olavs plass,
0130 Oslo, Norway
e-mail: frodes@hio.no

1 Introduction

Automatic image classification, labelling and retrieval are active research topics [29, 30]. Most photographers do not have the time and patience to manually catalogue single photographs and label these with textual descriptions. Instead, most users are often able to memorize approximately when a photo was taken, say "during the summer of 2008", or "in the winter holiday after the September 11 event". Moreover, users will have few problems associating a particular image with a location, such as "our holiday in Puerto Rico", "the business trip to Cape Town" or "the PCM 2009 conference in Bangkok". These are all possible because cameras not only store the images recorded by the camera chips but also store the time and date when the photos were taken using a digital clock built into the camera. Some cameras also store camera settings such as exposure time, aperture, focal distance, focal length, etc, using EXIF (Exchangeable Image File Format) [1] initiated by the Japan Electronics and Information Technology Industries Association (JEITA). This meta-information can also be used to organize images [2].

Geo-spatial information is an emerging image attribute that is used in addition to the time and date of an image. Combined time and geo-spatial attributes make it easier to organise, retrieve and browse large image collections [3, 4]. Moreover, image collections are growing rapidly and often viewed on mobile devices. Falling costs have resulted in most people owning digital cameras, and the quality of the camera equipment is constantly improving. Currently, even mobile phone cameras have megapixel resolution. Low cost digital storage has eliminated cost and time barriers previously associated with the development of film.

Still, GPS technology is not commonplace in most digital cameras as they add to the cost in a very competitive market.

Moreover, although the idea of using GPS technology is attractive in theory, it may not always be practical. A photographer may have to react spontaneously to a given situation and quickly take a shot. However, GPS enabled devices often need certain amount of time to lock onto the available overhead GPS satellites. In fact, the process of obtaining a reasonable GPS reading can sometimes take several minutes. Next, imagine that very response GPS enabled cameras became commonplace, then there will still be huge collections of digital photographs in existence taken with older digital cameras without geo-spatial capabilities. Finally, the current GPS-infrastructure is reaching the end of its lifetime and one does not have any guarantees for publically available satellite navigation systems in the future [5].

1.1 Direct Sun Elevation Measurements

GPS technology is a relatively new phenomenon. Prior to GPS technology navigation and positioning was achieved using the position of celestial bodies such as the sun, moon and the stars. During days with clear skies the sun provides a good reference point for estimating ones position. Based on the time of year, the sun follows a sinusoidal path across the skies relative to an observer on earth. On the northern hemisphere the sun goes up in the east and sets in the west and is located at a southern direction at midday. On the southern hemisphere the sun goes from east to west via a northern route. Generally, the elevation of the sun is higher at midday for small latitudes compared to high latitudes where the maximum elevation of the sun is lower. Moreover, during winter the elevation is lower than the summer, and while it is winter on the northern hemisphere it is summer on the northern hemisphere and vice versa.

Seafarers have exploited this phenomenon for hundreds of years. For instance, the sextant was used to measure the elevation of the sun above sea level by aligning two adjustable views. One view was centred on the horizon and another view was centred on the sun, such that the two views were aligned. Then, an accurate angular reading of the suns elevation was taken. Next, the height of the observer above sea level was compensated for. By the means of an accurate watch, a compass and an astronomical almanac the position of the observer was estimated with a very high accuracy of close to 0.1 nautical miles which is approximately 200 meters.

These traditional celestial navigation techniques have inspired researchers working on autonomous robot navigation where a digital camera was used to measure the approximate elevation the sun as a kind of digital sextant [6]. Related research includes the development a sun sensor [22].

A lens is usually characterised in terms of its focal length f. A simplified explanation of focal length is how much magnification a lens provides. A lens with a large focal length magnifies an image more than a lens with a smaller

focal length. However, with more magnification the lens field of view is smaller. The field of view covered by a lens with focal length f is given by

$$a = 2\tan^{-1}\left(\frac{d}{2f}\right) \tag{1}$$

where d represents the width of the image sensor inside the camera. Classic 35 mm film has a dimension of 36×24 mm, while digital camera sensors often are smaller. For instance, cameras in the Nikon's DX series have dimensions of about 23.6×15.5 mm, Cannon APS-C has dimensions of 22.2× 14.8 mm, and pocket camera sensors can be as small as 2.4× 1.8 mm (1/6″ sensors). Usually the lenses are rectilinear, that is, all straight edges in the scene appear straight in the captured image. The field of view can be measured along the horizontal (width), vertical (height) or along the diagonal. It is the dimensions of the sensor (or digital) film that determines the field of view along the vertical and horizontal dimensions. A 35 mm camera with a 50 mm lens will therefore have a horizontal view of 46.8° and a vertical view of 27°. It has been shown that the lens focal length for a camera can be determined using a sequence of outdoor images where the position of the sun is hand labelled [7].

Given a camera configuration with a resolution of P_x x P_y pixels and a field of view of V_x x V_y degrees along the horizontal and vertical positions, respectively. Then the degrees per pixel are given by:

$$\partial a = \frac{V_x}{P_x} \approx \frac{V_y}{P_y} \tag{2}$$

The vertical degrees per pixel should be approximately the same along the horizontal and vertical axis. Given an optimal image scene comprising clear skies, a sun and a distinct horizon, the distance in pixels between the sun and the horizon are easily measured, and hence the elevation e of the sun can be calculated as

$$e = \partial a|y_{sun} - y_{horizon}| \tag{3}$$

Where y_{sun} is the vertical pixel value for the centre of the sun and $y_{horizon}$ a representative vertical pixel value of the horizon assuming the camera is level. Several methods for horizon extraction have been proposed, including the use of orientation projection [8, 9]. These are robust methods aimed at micro aircraft control with unfocused rapidly moving images. Given the elevation of the sun and the current solar time an astronomer's almanac can be used to determine the geographical location [13].

The direct sun elevation measurement technique is not well suited for the analysis of digital image collections. First, the calculations are dependent on the characteristics of the physical camera design. Second, most camera lenses have a limited field of view and will only work when the

sun is at low elevations. For example, with a 50 mm lens and 35 mm digital film the maximum theoretical elevation is 26°. With a 100 mm lens and 35 mm digital film the maximum theoretical elevation is 14°, and for a 200 mm lens and 35 mm digital film the maximum theoretical elevation is 6°. Next, with the exception of beautiful sunrises and sunsets, it is uncommon to take direct photographs of the sun. Finally, although accurate horizon detection algorithms exist for small aircrafts flying at certain altitudes, it is much harder to determine the altitude from a photographer's perspective as he or she may be located in a city, in a valley or next to other tall objects that obstructs the view of the horizon [20].

1.2 Indirect Sun Elevation Measurements

Direct sun observations can be avoided by measuring the sun elevation indirectly. In particular, the position of the sun has also been measured indirectly by investigating the lighting condition of a scene [25], represented using the exposure level. The lighting conditions are related to the elevation of the sun, where in general solar noon is the brightest time of day. The exposure level can be computed using the aperture, shutter speed and film speed settings that many digital cameras store in the image EXIF headers [1, 2]. Experiments have shown that a brightness representation of the suns trajectory can be sufficiently mapped for image collections. Based on these trajectories rough estimates of solar noon and day-lengths can be made. Solar noon and day-length measurements can again be used to estimate the longitude and altitude of the observer. This approach has been demonstrated to yield a longitudinal accuracy of 15° and a latitudinal accuracy of 30° with arbitrary holiday photo collections [25]. A problem with this strategy is that it requires a sufficiently large set of outdoor images with a sufficiently large temporal spread. For images without exposure metadata, it has been demonstrated that a very rough indication of longitude can be determined by simply taking the mean time for a sequence of images within a 24 h window as the solar noon. The achieved accuracy for arbitrary collections of holiday photos was about 30° [26]. An advantage of both these indirect methods is that they also work under cloudy conditions, and the latter strategy even works indoors.

1.3 Webcam Measurements

Another branch of related research attempts to determine the geographical location of webcams [23, 26, 28]. Webcams are often used to acquire sequences of regularly spaced images for monitoring purposes. The cameras are usually located in a fixed location and often pointing in a constant direction. On the downside, few webcams store meta-information in EXIF headers and analysis can therefore only be performed using actual image contents. Webcam image sequences have been used to determine the relative position of webcams and their orientation [23, 24]. Moreover, an accuracy of about 2° was achieved using a contents-based intensity measure of webcam images sampled every 5 to 11 min [28]. This approach allowed the sunrise and sunsets to be determined, and hence the solar noon and length of day could be calculated. However, webcam images represent a special case and webcam techniques are not applicable to general image collections.

1.4 Landmark Recognition

Another novel approach to geo-tagging involves automatically recognizing known landmarks in image scenes. Given knowledge about the location of the landmarks the location of the image scene can therefore be inferred [21]. Such strategies clearly depend on both an extensive landmark database and a powerful landmark matching algorithm.

1.5 Object-Shadow Lengths and Sun Elevation

This study proposes a new strategy for deriving the geographical origin of image scenes based on both the image contents and image meta-information. The proposed strategy relies on the fact that the lengths of shadows cast by vertical objects on horizontal surfaces indirectly reveal the elevation of the sun. If such sun elevation measurements are obtained together with the time at which photographs were taken it is possible to derive the geographical location where the images were captured. There are several locations at which one can observe the sun at a given elevation at a given time. Therefore, up to three images taken at different times at the same location are used to identify a single and unique geographical location. This study investigates the practicality, reliability and accuracy of such object-shadow length sun elevation measurements for determining geographical location of image scenes. Although this strategy will not work on cloudy days it has potential for much greater accuracy than previous indirect methods based on scene brightness.

2 Shadows and Sun Elevation

Shadows provide an indirect clue to the elevation of the sun as the sun at a high elevation will cast a short shadow while the sun at a low elevation will cast a long shadow. Given an

object with a height H and a shadow with length L, the elevation e of the sun is simply

$$e = \tan^{-1}\left(\frac{H}{L}\right) \qquad (4)$$

This is illustrated in Fig. 1. A convenient property of this equation is that it is based on a ratio and any units associated with the object and shadow length measurements are cancelled. Hence, the shadow based sun elevation measurements are close to independent of the technical properties of the camera and the relative dimensions of the scene with the exception of distortions caused by low quality lenses.

Next, it can be shown that the relationship between the elevation of the sun e and the geographical location of the observer (see Fig. 2) is given by:

$$\sin(e) = \sin\delta\sin\varphi + \cos\delta\cos\varphi\cos w \qquad (5)$$

where φ is the latitude of the observer, w is the sun angle of the observer and δ is the declination of the sun at the given date which can be approximated by:

$$\delta = -0.4092797\cos\left(\frac{2\pi}{365}(M+10)\right) \qquad (6)$$

Here, the declination of the sun is represented in radians and M denotes the day of the year. The constant 0.4092797 represents the maximum declination angle of the sun, or earth tilt, in radians (23.45°) that occurs during the two solstices (see Fig. 3). Note that this is a rough approximation of the sun declination angle, i.e., a simple sinusoidal with a period of 365 days, and that more accurate approximations exist. However, the author's experimentation has shown that this expression provides sufficient accuracy for the purpose of this study.
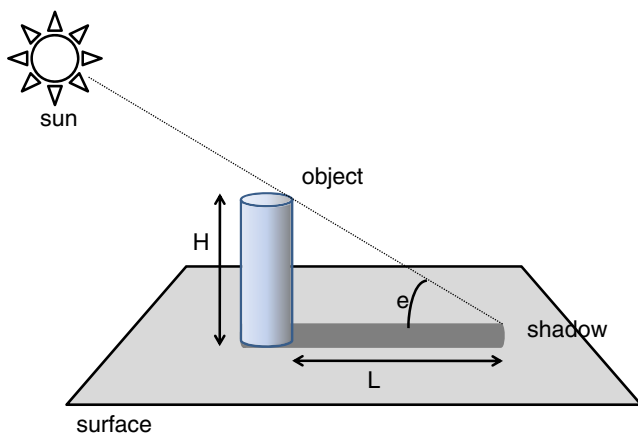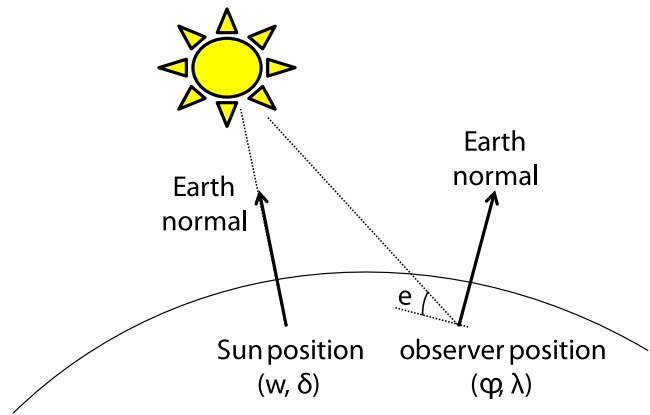

Figure 2 The sun position is the point on earth where the sun elevation is 90°, namely the position with latitude δ (sun declination angle) and longitude w (solar angle). For an observer the elevation angle e depends on the observers position, i.e., latitude ϕ and longitude λ.

Next, the longitude λ of the observer is related to the solar time t_{sun} as follows

$$t_{sun} = t_{utc} - \frac{12}{\pi}\lambda \qquad (7)$$

and solar time t_{sun} is related to the sun angle w as follows:

$$w = \frac{180}{12}(t_{sun} - 12) \qquad (8)$$

Given an elevation measurement e_1 at UTC time t_1 one can find all observation points with the given sun elevation for the given time. In this study we traversed the Earth's surface with a resolution of 1°, giving, 360×180 points and stored all locations in L_1 which satisfied the sun elevation criteria for the given time. For high elevations the possible locations form a circle-like shape on the Earth's surface as shown in Fig. 4.
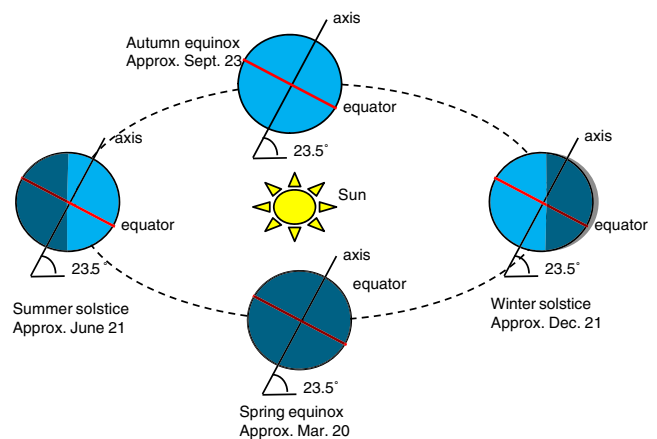

Figure 3 The northern hemisphere is more exposed to the sun during the summer and the southern hemisphere is exposed to the sun during the winter and the maxima occur during the two solstices as the earth's tilt is then parallel to the direction of the sun. Both hemispheres are equally exposed to the sun during the two equinoxes as the earth's tilt is then perpendicular to the direction of the sun.

Figure 1 The relationship between the sun elevation e, object height H and shadow length L.

Figure 4 The location traces for three elevation observations from Cape Town, South Africa at 7.59, 10.48 and 13.18 UTC during February 27, 2009. The three traces cross approximately in one location.
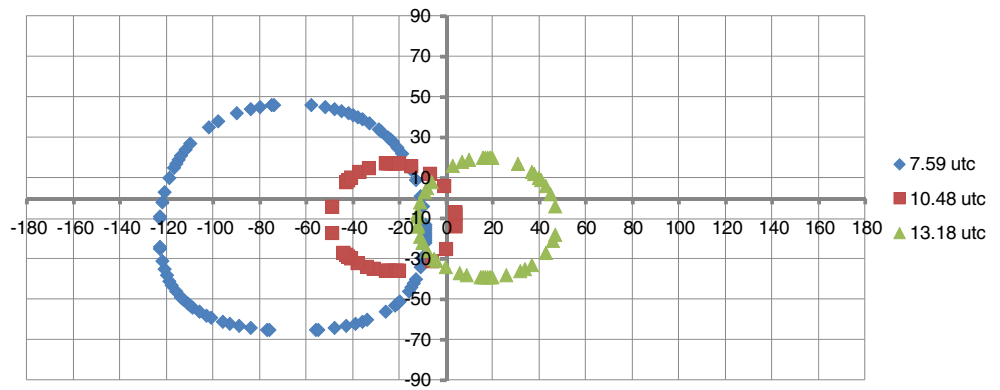
In order to get a more accurate fix on the actual location a second sun elevation e_2 at a different image taken at time t_2 is obtained, giving rise to a second trace of locations L_2 (see Fig. 4). These, two traces cross in two locations (φ_1, λ_1) and $(\varphi_2, \lambda_{12})$—one on the southern and one on the northern hemisphere.

In order to determine which of the two estimated locations that represents the true location a third sun elevation e_3 from a third image taken at time t_3 is needed. This gives rise to a third trace of location points L_3. Then, in most situations there will be only be one point where all the three traces L_1, L_2 and L_3 cross simultaneously, namely the true location (φ, λ) of the observer. Note that also the correct hemisphere is determined in these cases.

The feasibility of this approach is dependent on the season. It will work especially well during the winter and during the summer where the declination of the sun is large, while it will work less well during the spring and autumn when the declination of the sun is small. With a large declination the length of day is very different on the two hemispheres and the sun elevation paths are very distinct (see Fig. 5). On the contrary, with a small sun declination the differences between the sun elevation paths on the hemisphere are small and it is harder to distinguish between the two (see Fig. 6). In other words, the approach works best closest to the two solstices (generally 21st of June and 21st of December) and the strategy will not be able to

distinguish between the two hemispheres during the two equinoxes (approximately 20th March and 23rd September). This hemisphere ambiguity is illustrated in Fig. 6. With small sun declinations it is necessary with additional clues in order to determine which hemisphere the observer was located at.

The ability to successfully identify the correct hemisphere is also dependent on the angle between the latitude and the declination of the sun. With a large solar angle and a latitude close to the declination of the sun angle, it is more difficult to determine on which hemisphere the observer is located, while this is much easier when the angle between the latitude and the sun declination is large. Yet, if the observer's latitude is close to the declination of the sun and an observation is made close to the solar noon, that is, with a small solar angle, then the location can also be determined quite accurately as the sun can only be observed at elevations of close to 90° at a limited area on the Earth's surface. Moreover, traces for sun elevations taken at different times will also only cross in one point. This is illustrated in Fig. 7. The plot shows that all the traces only cross through one point. Therefore, images taken at latitudes close to the sun declination line can be determined with one image if the sun angle is small and with two images otherwise. The plot shows that the diameter of the trace 12:30 is only 15°, while at 12:00 the trace is simply one point. One hour before and after noon the diameter of

Figure 5 The location traces for sun elevation observations at Oslo, Norway during January 1 at 9:00, 11:00 and 13:00 UTC, respectively. The traces only cross in one point—the location of the observations.
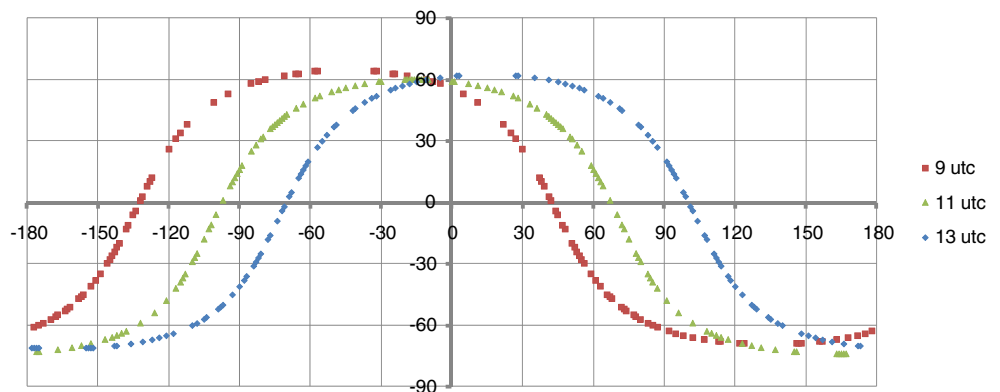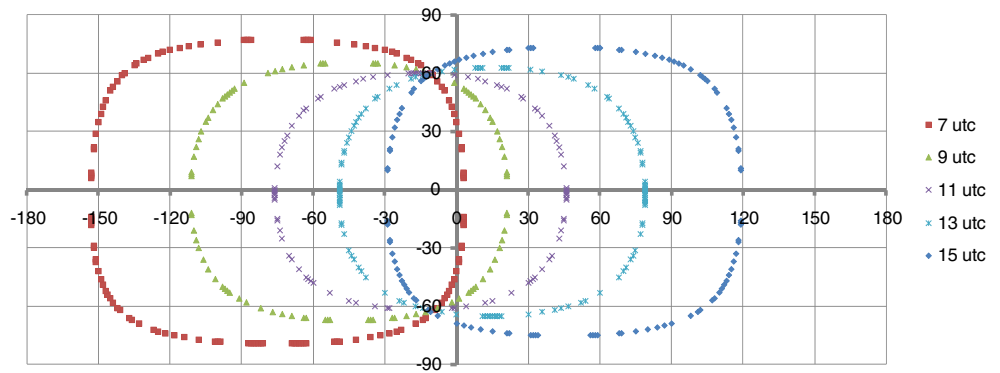
Figure 6 The location traces for sun elevation observations at Oslo, Norway during the March 20 equinox at 7:00, 9:00, 11:00, 13:00 and 15:00 UTC, respectively. Note that the different traces cross in two points—one on the southern and northern hemisphere.

the traces are 30° and grows with 30° for each hour in either direction away from the solar noon.

2.1 Land Test

Previous sections have demonstrated that it may be difficult to determine the correct hemisphere when images are taken close to the equinoxes or if shadows from only two images are used. For this purpose a simple land test is proposed. It comprises mapping the two points onto a simple world map to determine if the points hit land or water. The one that hits land is chosen.

Imagine for example that two images are taken in Oslo, Norway (59.9° north, 10.7° east) during the spring equinox of March 20th. These will yield the coordinates (59.9°, 10.7°) and (−59.9°, 10.7°). Figure 8 shows these coordinates plotted onto a world map. Clearly, the former is located at Oslo, while the latter is located in the ocean south of the African continent. Unless the photograph was taken onboard a ship it is natural to reject the latter coordinate and conclude that the coordinate on the northern hemisphere is correct. By inspecting the world map in Fig. 8 it is obvious that the simple map test works for most locations in Northern Europe, North America and Asia. This is because approximately 70% of the Earth's surface is covered in water.
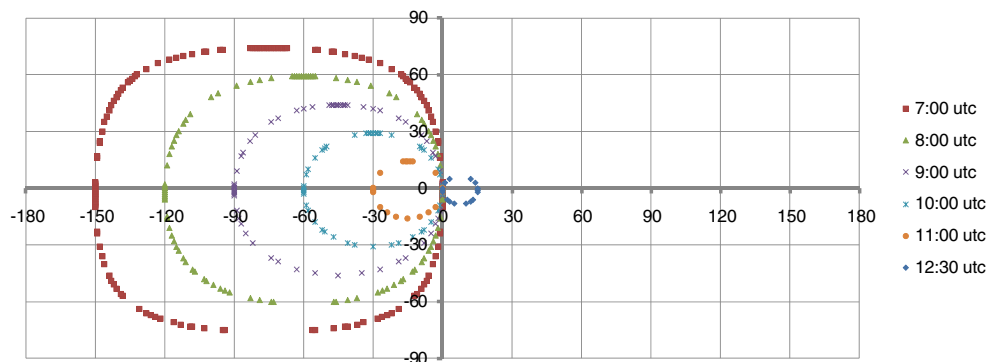
Figure 9 summarizes the proposed strategy for determining the geographical location of a set of image scenes. Input to the algorithm are three sun elevation measurements obtained from the object-shadow length ratios, the times the three images were captured and the date of the event. The output of the algorithm is the approximate geographical location of the place the images where captured.

2.2 Automatic Object-Shadow Length Measurements

This study focuses on how to determine the approximate geographical location given a set of object-shadow length measurements. Obtaining accurate object-shadow length measurements is indeed a non-trivial problem as one has to identify objects, identify shadows and determine which objects relate to which shadows. Therefore, only a rough speculation on how this may be achieved is attempted here. Inspiration is drawn from the literature which contains several accounts of work related to shadow detection [14, 15]. For instance shadow detection has been successfully applied to video based on colour models [16]. Segmentation of objects and background in outdoor images has also been studied [17] as well as shadows in aerial photographs [18, 19].

An image collection may be large and advanced processing of all the images is unrealistically time-

Figure 7 Sun elevation traces for observations at (0, 0) during an equinox.

Figure 8 Using a map to resolve a hemisphere ambiguity. The coordinate for the southern hemisphere is rejected as it does not refer to a land area. The map is taken from Wikipedia (Creative Commons).
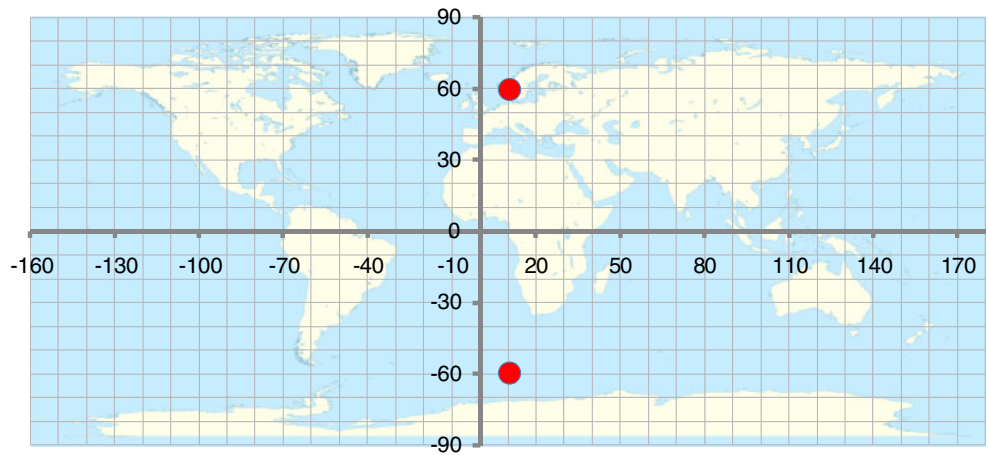
consuming. A natural first step is therefore to identify suitable image candidates, that is, images that are likely to have shadows. This is simply achieved by using the exposure attributes stored in EXIF-headers, including the aperture f (f-number), shutter speed s and film speed iso. Based on these the exposure level EV can be determined [31, 32]:

$$EV = \log_2 \frac{f^2}{s} + \log_2 \frac{iso}{100} \qquad (9)$$

Then outdoor images taken on a sunny day with sufficient shadows should have an exposure value of approximately 12 or more. If EXIF information is not

available a content based strategy can be used to identify suitable candidate images although that will be computationally more demanding than simply inspecting the EXIF-information. Several content-based strategies for classifying outdoor and indoor images have been proposed in the literature, for instance using colour space histograms [10] and support vector machines [11]. Moreover, attempts at extracting information from daytime images of the skies [12] have been proposed.

Next, candidate images can be separated into their hue and brightness components. Objects may be identified and segmented in the hue plane [27], and shadows identified and segmented in the brightness plane. Having obtained

Figure 9 Algorithm for determining the location of image scenes based on object-shadow lengths.

```
Coordinate findLocation(Angle elevationₐ, Time timeₐ,
                        Angle elevation_c, Time time_b,
                        Angle elevation_c, Time time_c,
                        Date date)
begin
    traceₐ ← locationsWithSunElevation(elevationₐ, timeₐ, date)
    trace_b ← locationsWithSunElevation(elevation_b, time_b, date)
    trace_c ← locationsWithSunElevation(elevation_c, time_c, date)
    location ← traceₐ ∩ trace_b ∩ trace_c
    If date=equinox then
            return landTest(location)
    else
            return location
end


Set locationsWithSunElevation(Angle observedElevation, Time time, Date date)
begin
    declination ← sunDeclination(date)  // Eq. 6
    for longitude ← -180 to 180 step resolution
        for latitude ← -90 to 90 step resolution
            begin
                elevation ← calcElevation(latitude, longitude, declination, time) // Eq. 5 and 8
                if observedElevation ≈ elevation then
                        trace.add(Coordinate(latitude, longitude))
            end
    return trace
end
```

Table 1 Image test suite used in the experiments.

UTC time	UTC decimal time (hours)	measured elevation (degrees)	shadow vector length (pixels)	focal length (mm)
07:33:10	7.6	34.4	2056.4	7.1
07:35:40	7.6	34.1	332.6	7.1
09:02:00	9.0	56.6	1225.4	7.1
10:21:10	10.4	62.8	478.8	7.1
10:29:13	10.5	63.0	995.4	7.1
10:36:31	10.6	73.6	526.1	28.1
13:10:58	13.2	52.4	411.8	7.1
13:11:13	13.2	60.0	300.8	36.5

these segments the object lengths and shadow lengths can be measured.

This procedure can be repeated for several images and statistical approaches can be used to assess what shadow measurements that should be accepted and which ones that should be rejected.

Clearly, the outlined strategy is challenging as one may easily detect false objects and false shadows and thus end up with erroneous sun elevation measurement. Therefore, further research is needed to identify robust extraction strategies.

2.3 Time and Date Assumptions

The strategy presented herein assumes that all images are consistently time-stamped with date and time. Further, it is

assumed that the time-zone is known such that the times can be converted to UTC (Coordinated Universal Time). All the calculations presented herein are represented in UTC. Most owners set their camera to the time zone of their home country. Few users bother to change the time of their cameras when travelling to a different country in a different time zone. Since the camera clocks usually have their own battery one may assume that for most users the time will be set to the same time-zone for the entire lifetime of the camera and that potential time drifts will affect all images equally.

2.4 Image Scene Assumptions

The shadow model is also based on two further assumptions. First, the viewing plane is approximately level. If

**Figure 10** The manual 3-point sun-elevation measurement procedure.
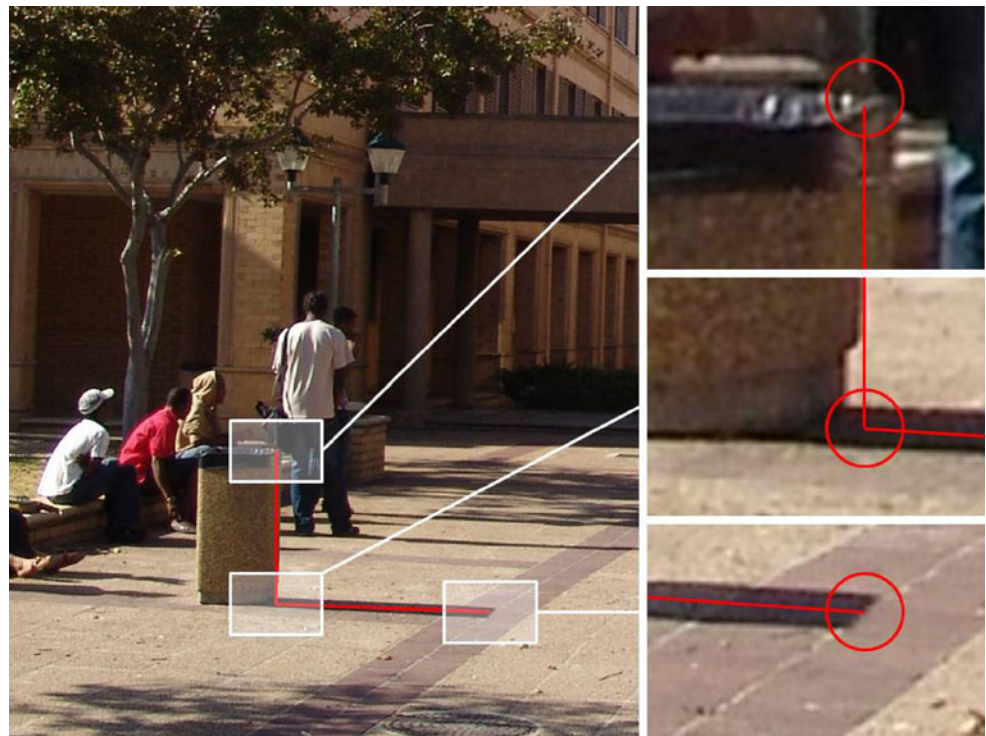
standing in a slope such as on the side of a hill the shadow angle calculations would require the model to take the slope into consideration. Given a slope of s degrees and a shadow of length L cast up the slope, then the error in the shadow due to the slope is $E=L - L \cos(s)$.

Second, the model assumes that all the objects are completely vertical with straight lines. Curved or tilting objects will cast more complex shadows and an angle extraction algorithm will have to take information about the scene into consideration. When a curved and tilted object is combined with a sloping surface the extraction of shadow information is even more complex. One strategy would be to classify images according to how tilted the ground is and the tilted or curved the objects are. Images with such

Figure 11 The images used in the experimental evaluation. Detailed attributes of these images are given in Table 2. The image resolution is 3,264× 2,448 pixels.
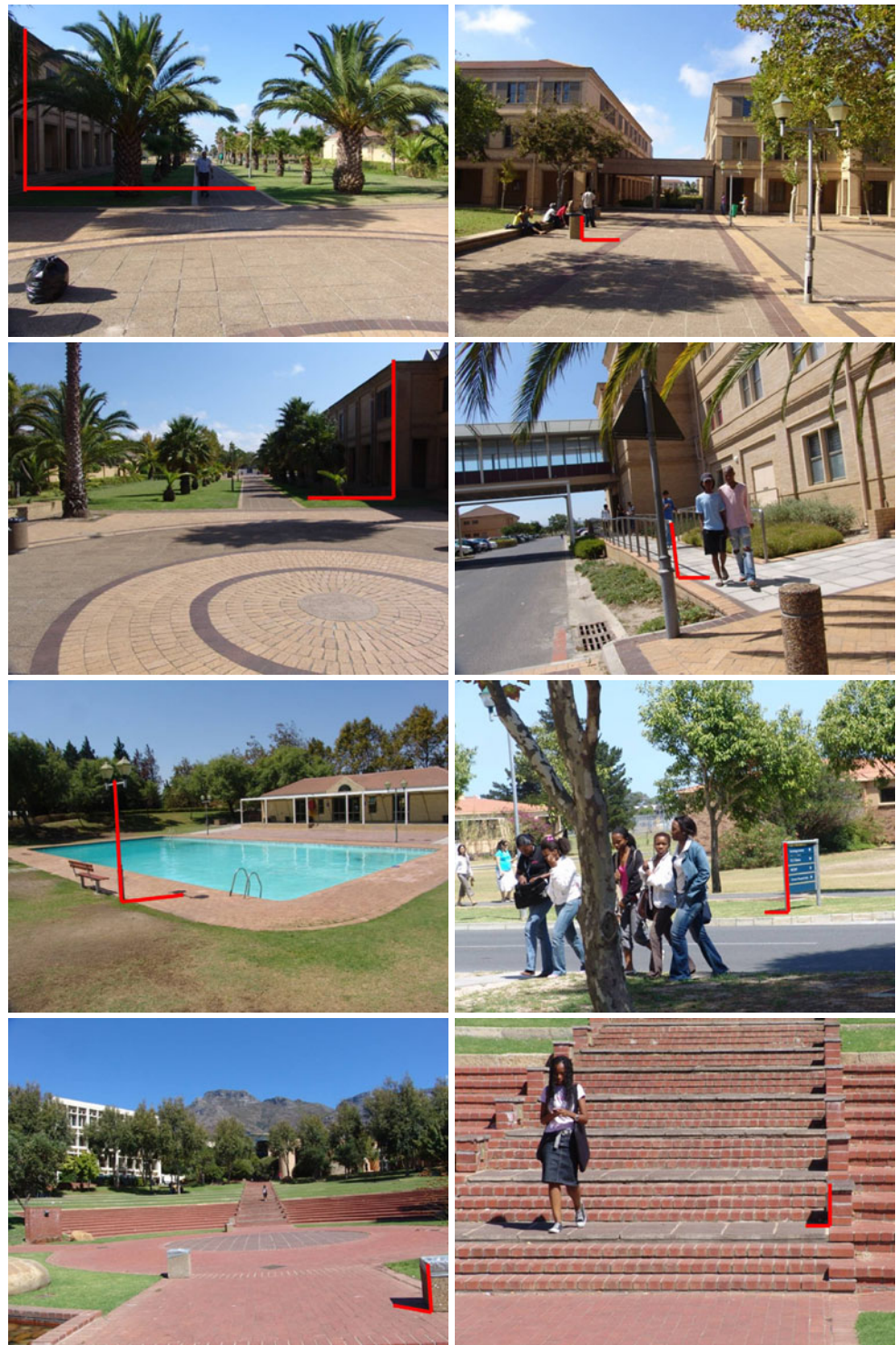
characteristics can then be eliminated from the shadow extraction procedure as their geometry is too complex for simple analysis procedures.

3 Experimental Evaluation

3.1 Test Suite

To assess the technique proposed herein a series of photographs taken at two campuses of Cape Peninsula University of Technology in Cape Town, South Africa during February 27, 2009 were used. This was a sunny day with clear skies and hence distinct shadows. The collection was photographed by the author, but without this experiment in mind. The sample therefore represents an arbitrary and natural image collection. A Sony DSC-F828 digital camera with 8 megapixel resolution and a zoom lens was used. First the image collection was manually inspected and a set of 8 photographs were selected. The following criteria had to be satisfied: The image scene had to contain a visible object and this object had to cast a visible shadow. The objects had to be vertical and straight. Only images where the shadows perceivably fell approximately perpendicular to the camera direction were selected to minimize image projection distortions. That is, images with shadows going straight left or right were selected. For each of the selected images Microsoft Paint was used to measure the exact pixel locations of three object-shadow feature points, namely the top of the object, the point connecting the object and the shadow and the shadow end point. These three points make up an L-shape, or inverted L shape as illustrated in Fig. 10. In this example the rubbish bin makes up the object and the shadow is cast on the right side of the bin. Next, EXIF-information, including the time and date of the photograph and the focal length used, were extracted using Microsoft Office Picture Manager. The images used and the associated feature points are illustrated in Fig. 11. Table 1 lists test suite details including the UTC time, measured elevation, the length of the measured shadow vector and the focal length of the lens used (degree of wide angle or zoom).

The coordinate 33.9° south, 18.8° east was used to represent Cape Town in this experiment. The date of the image collection is the 58th day of the year when the declination of the sun is approximately -9.1°. Hence, there is a significant difference between the hemispheres. This date is 21 days away from the spring equinox with no hemisphere difference and 68 days away from the winter solstice when there are maximum seasonal differences between the hemispheres.

3.2 Geographical Accuracy

Table 2 summarizes the result obtained with the proposed strategy. These results both demonstrate the accuracy of the strategy and the effects of varying the accuracy of the elevation measurements that are the input to the algorithm. First, the images were ranked according to the accuracy of their measured elevation accuracy. Then, a sliding window of 3 images was run through the ranking list to generate 5 sets of images with varying accuracy. The table therefore lists a linguistic description of accuracy, the rank of the images used, the actual index of the images used and the latitude and longitudes obtained with both the two and three image techniques.

The results show that the overall best estimate had a latitudinal error of 2.1° and longitudinal error of 1.2°. Then, as the accuracy of the sun elevation measurements decreased the largest error for this dataset was 16.9° latitude and 12.8° longitude. These results are superior to those obtained using image intensity [25] and matches the accuracy obtained using webcam image sequences [28].

Note that both the 2-image and 3-image strategies yield the same accuracy (see Table 2). The only effective difference between the two techniques is that the 3-image method was capable of automatically resolving the correct hemisphere and the 2-image solutions had to be resolved manually.

These results are much less accurate than the accuracy offered by GPS receivers. However, the purpose of this strategy is not to navigate, or survey landmass. The purpose is to geo-tag images and an accuracy of approximately 2° suffices for uniquely distinguishing continent and even country. It would, however, be interesting to investigate if

Table 2 Accuracy of latitude and longitude estimates.

accuracy	rank	images	2 images		3 images	
			latitude	longitude	latitude	longitude
high-high	6, 7, 8	3, 4, 5	−2.1°	−1.2°	−2.1°	−1.2°
high-medium	5, 6, 7	7, 5, 4	−0.1°	5.8°	−0.1°	5.8°
medium-medium	4, 5, 6	1, 7, 5	0.9°	6.8°	0.9°	6.8°
medium-low	3, 4, 5	2, 1, 7	0.9°	6.8°	0.9°	6.8°
low-low	1, 2, 3	8, 6, 2	16.9°	12.8°	16.9°	12.8°

the accuracy could be further improved by using images taken with this strategy in mind, that is, images where the photographer ensures that a clear shadow and its object is captured such that they occupy a majority of the image view and that the shadow is perpendicular to the camera.

3.3 Shadow Measurement Accuracy

Figure 12 shows that the observed sun elevations follow the theoretical sun elevations with a few exceptions. The first two elevation measurements are too low and the 6th and last elevation measurements are too large.

There are several sources of error in the above experiment. First, the camera clock may not be completely accurate. However, an inspection of the camera revealed that the clock was accurate to 2 min from the actual time. Still, the time will only affect the longitude. If the time is off by 1 h the longitudinal error will be 15°, for every minute of clock error the longitude error is 0.25° and every second of time inaccuracy affects the longitude by 0.004°. Therefore, an error of up to 2 min could have affected the longitude by up to half a degree. Note that an unsynchronized clock will not affect the latitude estimates since all the images are correctly spaced in relative time.

Distortions caused by camera projections may be a source of error (see Fig. 13). Although, all the shadows are perceived to be perpendicular to the camera direction it may not be the case in practice. In particular, for images taken with the zoon, that is, shadows that are further away will visually appear more perpendicular than shadows that are taken with wider lens configuration and that are closer to the camera. This is particularly noticeable if the plane of the shadow is close in height to the observer. Figure 13 illustrates how the shadows on a plane below the observer appear less perpendicular than shadows on a plane on similar height to the observer. The effect is that these shadows are erroneously observed as too short. This effect is further amplified by camera object distance. This
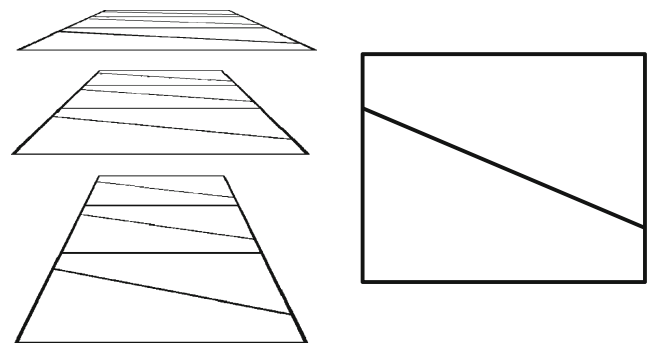

Figure 13 Perspective distortion affects the perceived shadow angle.

hypothesis is backed up by the results where sun elevation errors appear to correlate with the level of zoom (focal length). The two measurements with the largest error, that is, the sixth image and the eight image are both taken with zoom, namely focal lengths of 28.1 mm and 36.5 mm, respectively, where the latter yields the largest sun elevation error. The other images are taken using a wide angle lens with a focal length of 7.1 mm. By inspecting the last image, showing a student walking down a set of stairs, one sees that the measured shadow falls on a plateau. The projection makes the shadow appear perpendicular to the camera direction and the width of the plateau appears narrow. But, an inspection of the image as a whole will reveal that this plateau in fact is quite wide and that the shadow is at a slight angle. If one was standing closer one may have observed that the direction of this shadow is far from perpendicular to the camera angle. Consequently, the shadow measurement is too short compared to the object height resulting in a sun elevation measurement that is too high. This error is confirmed by the results in Fig. 12 where the measured sun elevation is 11.4° higher than the theoretical sun elevation. The measured shadow length was 154 pixels while the actual length should have been 235 pixels. The measurement was therefore short by about
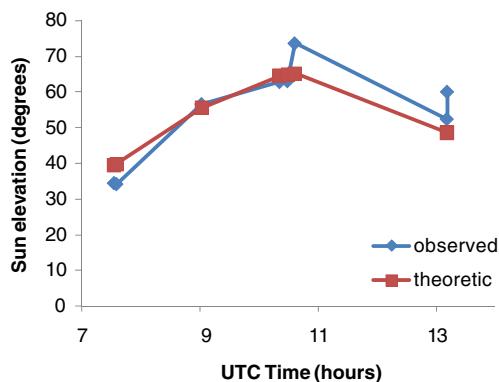

Figure 12 The measured and theoretical sun elevations for the eight images used in this experiment.
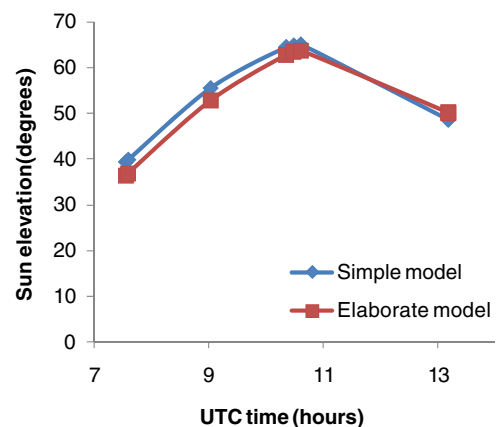

Figure 14 A comparison of the simple and elaborate sun elevation models.

81 pixels, or 34%. Future work should therefore introduce some measure to compensate for projection distortions. This involves identifying potential inaccurate shadow measurements by taking the distance into consideration where the distance is related to the focal length of the lens, the actual length of the shadow in number of pixels and the position of the shadow within a scene. A small shadow may indicate a shadow further away. A shadow closer to the middle of a scene (low-medium y-value), that is, closer to the horizon, is likely to be further away from the camera compared to a shadow towards the bottom of a scene (high y-value) that is likely to be closer to the camera.

3.4 Celestial Model Accuracy

The celestial model used in this study is simplistic as it is purely based on the geometric properties of the sun and earth orbits. Advantages of this model include that it is simple to implement, easy to describe and involves little computational effort. However, other more elaborate and complex models exist that take other factors into consideration such as atmospheric refraction [33]. Figure 14 illustrates differences between the simple and a more elaborate model. The data for the elaborate model was acquired using an online sun-elevation calculator (http://www.satellite-calculations.com/Satellite/suncalc.htm) that is implemented according to a procedure described in [33]. The plot seems to suggest a minor time discrepancy, that is, the simple model is slightly ahead in time of the more elaborate model.

When comparing the simple and elaborate model with the actual measurements it was found that the simple model yielded a mean sun elevation error of 4.9° (SD=3.6) and the elaborate model resulted in a mean sun elevation error of 3.9° (SD=3.9). Hence, the elaborate model had an overall better fit to the measurements compared to the simple model, although the spread in error was also larger. Therefore, for any real applications of this approach the simple model should be replaced with a more elaborate celestial model such as the one described in [33]. Note that the strategy presented herein is general and works with any celestial model.

4 Conclusions

A framework for determining the location a series of photographs based on the contents of the images was presented. The elevation of the sun is determined indirectly using the shadows cast by vertical objects. The advantage of shadow based sun elevation extraction is that it can be performed without knowledge about the optical properties of the camera or the absolute scale of objects in the scene.

Experimental results revealed that the location of images could be found with an accuracy of down to 2° in latitude and longitude given shadow measurements with an error below 2° of sun elevation. The meter-level accuracy provided by GPS technology is usually not needed for image browsing and cataloguing applications as an overall positioning accuracy of a few degrees is sufficient to identify approximately where in the world the photographs are taken. The strategy therefore has potential for content based geo-spatial information retrieval. However, its success is reliant on the progress of future research into automatic accurate object-shadow length measurement algorithms.

References

1. Alvarez, P. (2004). Using Extended File Information (EXIF) file headers in digital evidence analysis, International Journal of Digital Evidence 2(3).
2. Jang, C.-J., Lee, J.-Y, Lee, J.-W., & Cho, H.-G. (2007). Smart management system for digital photographs using temporal and spatial features with EXIF metadata, presented at 2nd International conference on digital information management, pp. 110-115.
3. Ahern, S., Naaman, M., Nair, R., & Hui-I Yang, J. (2007). World explorer: visualizing aggregate data from unstructured text in geo-referenced collections, presented at 7th ACM/IEEE-CS joint conference on Digital libraries, pp. 1-10.
4. D. Carboni, S. Sanna, and P. Zanarini, GeoPix: image retrieval on the geo web, from camera click to mouse click, presented at Proceedings of the 8th conference on Human-computer interaction with mobile devices and services, pp. 169-172, 2006.
5. GAO, GLOBAL POSITIONING SYSTEM. (2009). Significant Challenges in Sustaining and Upgrading Widely Used Capabilities, United States Government Accountability Office.
6. Cozman, F., & Krotkov, E. (1995). Robot localization using a computer vision sextant, presented at IEEE International Conference on Robotics and Automation.
7. Lalonde, J. F., Narasimhan, S. G., & Efros, A. A. (2008). Camera parameters estimation from hand-labelled sun positions in image sequences., Robotics Institute, Carnegie Mellon University. Technical Report CMU-RI-TR-08-32.
8. Bao, G.-Q., Xiong, S.-S., & Zhou, Z.-Y. (2005). Vision-based horizon extraction for micro air vehicle flight control. *IEEE Transactions on Instrumentation and Measurement, 54*(3), 1067–1072.
9. Ettinger, S. M., Nechyba, C., & lfju, P. G. (2002). Towards Flights autonomy: Vision-based horizon detection for micro air vehicles, presented at IEEE International Conference on Robotics and Automation.
10. Szummer, M., & Picard, R. W. (1998). Indoor-outdoor image classification, presented at IEEE International Workshop on Content-Based Access of Image and Video Database.
11. Serrano, N., Savakis, A., & Luo, A. (2002). A computationally efficient approach to indoor/outdoor scene classification, presented at 16th International Conference on Pattern Recognition.
12. Lalonde, J.-F., Narasimhan, S. G., & Efros, A. A. (2008). What does the sky tell us about the camera? Presented at European Conference on Computer Vision.
13. Michalsky, J. J. (1988). Astronomer's Almanac algorithm (1950–2050). *Solar Energy, 40*(3), 227–235.

14. Horprasert, T., Harwood, D., & Davis, L. S. (1999). A statistical approach for real-time robust background subtraction and shadow detection, presented at IEEE ICCV.

15. Levine, M. D., & Bhattacharyya, J. (2005). Removing shadows. *Pattern Recognition Letters, 26*(3), 251–265.

16. KaewTraKulPong, P., & Bowden, R. (2001). An improved adaptive background mixture model for realtime tracking with shadow detection, presented at 2nd European Workshop on Advanced Video Based Surveillance Systems, AVBS01.

17. Lefèvre, S., Mercier, L., Tiberghien, V., & Vincent, N. (2002). Multiresolution color image segmentation applied to background extraction in outdoor images, presented at IS&T European conference on color in graphics, image and vision, 2002.

18. Li, Y., Sasagawa, T., & Gong, P. (2004). A system of the shadow detection and shadow removal for high resolution city aerial photo, presented at XXth ISPRS Congress.

19. Wang, J. M., Chung, Y. C., Chang, C. L., & Chen, S. W. (2004). Shadow detection and removal for traffic images, presented at IEEE International Conference on Networking, Sensing and Control.

20. Sandnes, F. E. (2009). Sorting holiday photos without a GPS: what can we expect from contents-based geo-spatial image tagging? *Lecture Notes on Computer Science, 5879*, 256–267.

21. Zheng, Y.-T., Ming, Z., Yang, S., Adam, H., Buddemeier, U., Bissacco, A., et al. (2009). Tour the world: building a web-scale landmark recognition engine, in the proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2009), pp. 1085-1092.

22. Trebi-Ollennu, A., Huntsberger, T., Cheng, Y., & Baumgartner, E. T. (2001). Design and analysis of a sun sensor for planetary rover absolute heading detection. *IEEE Transactions on Robotics and Automation, 17*(6), 939–947.

23. Jacobs, N., Satkin, S., Roman, N., Speyer, R., & Pless, R. (2007). Geolocating static cameras, in the proceedings of IEEE 11th International Conference on Computer Vision (ICCV 2007), pp. 1-6.

24. Jacobs, N., Roman, N., & Pless, R. (2008). Toward fully automatic geo-location and geo-orientation of static outdoor cameras, in the proceedings of IEEE Workshop on Applications of Computer Vision, pp. 1–6, 2008.

25. Sandnes, F. E. (2010). Where was that photo taken? Deriving geographical information from image collections based on temporal exposure attributes. *Multimedia Systems, 16*(4–5), 309–318 (This publication is not in print).

26. Sandnes, F. E. (2010). Unsupervised and fast continent classification of digital image collections using time, in Proceedings of ICSSE 2010, IEEE CS Press, pp. 516–520 (This publication is not in print).

27. Y-P, Huang, T-W, Chang, Y-R, Chen, & F E, Sandnes. (2008). A back propagation based real-time license plate recognition system. *International Journal of Pattern Recognition and Artificial Intelligence, 22*(2), 233–251.

28. Sandnes, F. E. (2010). A simple content-based strategy for estimating the geographical location of a webcam, in Proceedings of PCM2010, Springer Lecture Notes on Computer Science, LNCS, Vol. 6297 (pp. 36-45) (This publication is not in print).

29. Huang, W., Gao, Y, & Chan, K. L. (2010). A review of region-based image retrieval. *Journal of Signal Processing Systems, 59* (2), 143–161. doi:10.1007/s11265-008-0294-3 (This publication is not in print).

30. Heesch, D., & Petrou, M. (2010). Markov random fields with asymmetric interactions for modelling spatial context in structured scene labelling. *Journal of Signal Processing Systems, 61*(1), 95–103. doi: 10.1007/s11265-009-0349-0 (This publication is not in print).

31. Jones, L. A., & Condit, H. R. (1941). The brightness scale of exterior scenes and the computation of correct photographic exposure. *Journal of the Optical Society of America, 31*(11), 651–678.

32. Ray, S. F. (2000). Camera exposure determination, in the manual of photography: photographic and digital imaging, R. E. Jacobson, S. F. Ray, G. G. Atteridge, & N. R. Axford, (Eds.). Focal Press.

33. Schlyter, P. (2010). Computing planetary positions - a tutorial with worked examples, Downloaded March 26, 2010 from http://www.stjarnhimlen.se/comp/tutorial.html#5 (This publication is not in print).

Frode Eika Sandnes received a B.Sc. in computer science from the University of Newcastle upon Tyne, U.K., and a Ph.D. in computer science from the University of Reading, U.K. He is currently a Professor in the department of Computer Science at Oslo University College, Norway. His research interests include multimedia processing, error-correction and human computer interaction.