

Where was that photo taken? Deriving geographical information from image collections based on temporal exposure attributes

Frode Eika Sandnes

Faculty of Engineering, Oslo University College, P.O. Box 4, St. Olavs plass, 0130 Oslo, Norway

e-mail: frodes@hio.no, tel: +47 22 45 32 49, fax:+47 22 45 32 05

Abstract

This paper demonstrates a novel strategy for inferring approximate geographical information from the exposure information and temporal patterns of outdoor images in image collections. Image exposure is reliant on light and most photographs are therefore taken during daylight which again depends on the position of the sun. Clearly, the sun results in different lighting conditions at different geographical location at different times of the day and hence the observed intensity patterns can be used to deduce the approximate location of the photographer at the time the photographs were taken. Images taken inside or at night are temporally connected to the daylight images and the geographical information can therefore be transferred to related "dark" photographs. The strategy is efficient as it only considers meta information and not image contents. Large databases can therefore be indexed efficiently. Experimental results demonstrate that the current approach yields a longitudinal error of 15.7 degree and a latitudinal error of 30.5 degrees for authentic image collections comprising a mixture of outdoor and indoor images. The strategy determined the correct hemisphere in all the tests. Although not as accurate as GPS receiver, the geographical information is sufficiently detailed to be useful. Applications include improved image retrieval, image browsing and automatic image tagging. The strategy does not require a GPS receiver and the strategy can be applied to existing digital image collections.

1 Introduction

Current digital cameras provide high quality images at a low cost compared to just a decade ago. Advances in storage technology allow amateur photographers to take thousands of photographs without having to consider the cost of developing film and printing and physical storage space. Consequently, personal image collections are growing at an exploding rate. Most people neither have the time to carefully sort images into suitable categories, nor manually annotate images with textual information to help future retrieval. GPS technology allows images to be tagged with the geographical coordinates where an image is taken. This is very useful when taking many photographs in different parts of the world. The geo-spatial tagged images are therefore much easier to manage, store, retrieve and browse as images can be classified according to location and time [1, 2].

However, the GPS approach to image tagging is problematic for several reasons. First, very few low-cost digital cameras are equipped with GPS technology. Second, GPS devices typically need several minutes to lock onto overhead navigation satellites. Third, the GPS navigation infrastructure is reaching the end of its lifetime and one has no guarantee that there will be a publically available navigation infrastructure in the future [3]. Fourth, most existing digital image collections are not tagged with geo-spatial information.

This paper presents a different approach to geo-spatial image tagging which is not reliant on GPS technology or similar geographical information systems. Moreover, the strategy can be applied to existing digital image collections without geo-spatial information. The strategy is based on analysing the temporal camera usage dynamics and exposure values embedded in image files.

2. Background

Prior to GPS technology seafarers navigated according the celestial bodies such as the sun, the moon and the stars. For instance, the compass can be used to obtain the orientation, or azimuth, of the sun, the sextant is typically used to obtain the elevation of the sun above the horizon and a chronograph is used to get an accurate reading of the current time. The idea of celestial navigation has also been attempted in modern times for robot navigation using a digital camera as a digital sextant [4] and absolute sun orientation measurements [5]. Related research has also used image contents to determine relative camera locations [6] and camera orientation [7] in camera networks. A camera network is a collection of webcams located at various geographical locations. Unlike a photographer, webcams can take continuous sequences of photographs at regular intervals at fixed locations.

Life is organised around the celestial motion of the sun. Humans are physiologically linked to daylight in a cyclic pattern. We get up in the morning when the sun rises and we typically go to bed at midnight to get approximately eight hours of sleep until the sun is re-emerging in the horizon. When on holiday we typically go on sightseeing during the day so that we are able to see sights in bright daylight. Moreover, photographs are also taken during daylight as good images require sufficient lighting. Based on this one can assume that photography correlates with the presence of the sun, i.e., that one takes more pictures during the day than during the night. Obviously, images taken during the day are taken at brighter conditions than images taken indoors or at night.

Most digital cameras are equipped with an internal clock that is usually set according to the local time-zone once the camera is first purchased. When travelling to different time-zones, most users do

not bother to set the cameras clock to the local time-zone¹. All images taken with the camera are tagged with the time and date of the camera clock, often using the EXIF-format [1, 8, 9].

More advanced cameras capture additional information, especially the optical camera settings for the photograph such as the focal length of the lens (if adjustable), the exposure time, aperture and whether flash is used or not. Combined, the camera exposure time, aperture, film speed and flash information can be used as features to deduce information about the lighting level of the scene in the image without actually having to assess the actual image contents.

Time and geo-spatial attributes make it easier to organise, retrieve and browse large image collections [2, 10] and this is especially important as digital image collections are growing at an exploding rate.

3 Method

Given an image collection C with N images denoted I_i where $i \in \{1, 2, \dots, N\}$ represents the temporally ordered images and a function $t(I_i)$ that gives the time image I_i was taken, then the exposure value $EV(I_i)$ of image I_i is defined as [11-14]:

$$EV(I_i) = \log_2 \frac{\text{aperture}^2(I_i)}{\text{shutter}(I_i)} + \log_2 \frac{\text{iso}(I_i)}{100} \quad (1)$$

where $\text{aperture}(I_i)$ is the aperture of image I_i , represented as an f-number, i.e., f/2.8, f/4, f/5.6, f/8, f/11, f/16, etc, $\text{shutter}(I_i)$ is the shutter speed of image I_i measured in seconds, i.e., 1/1000 s, 1/500 s, 1/250 s, 1/125 s, 1/60 s, 1/30 s, 1/15 s, etc and $\text{iso}(I_i)$ is the iso value (film speed) for image I_i , typically 100 or 200. These are all obtained from the EXIF meta information recorded in the image file by the digital camera.

The exposure value can be used to make intelligent guesses about the scene contents and a summary of how the exposure value can be interpreted is provided in Table 1. For example, a sunny day is characterised by an exposure value in the range of 14-16, while cloudy days are represented by exposure values of 12-14. Sunsets are often represented by exposure values of 12. Night scenes have an exposure value of less than 11. Similarly, indoor images often have an exposure value of less than

¹ One reviewer insisted that he always set the clock of all his devices once he arrives in a new country. The Reviewer probably represents a small minority of very tech-savvy users.

12. Well lit places such as galleries may have an exposure value of 11, while a typical home may have an exposure value in the range of 5 to 7. One may subtract 2 from the exposure value if the scene is in the shadows.

Direct camera settings provide a more efficient and objective means of determining lighting conditions compared to content based strategies proposed in the literature. Firstly, images are the result of applying optical camera settings, and valuable information about the original scene may be lost. Moreover, computational effort is required to process each image. Such contents based strategies have typically been used for classifying outdoor and indoor images, using colour space histograms [15] and support vector machines [16], or for extracting information from the skies [17].

Table 1. Interpretation of scenes according to the exposure value (EV) of images. The data originates from [18, 19].

EV	Interpretation
16	Bright Sunlight distinct shadows
15	Sunlight distinct shadows
14	Before sunset, Hazy sunlight soft shadows
13	Before sunset, Cloudy day soft shadows
12	At sunset, Shady scene in sunlight
11	Just after sunset
10	Just after sunset, night neon signs, indoor bright room
9	Just after sunset, night arena sports, indoor sport event
8	Bright street, Indoor offices
7	Indoor home
6	Indoor home
5	Indoor home
4	Outdoors at night

The strategy presented herein assumes that all the date and time is set once correctly according to the owners' locale for the image collection. This is a realistic assumption as most users will only set the time and date once they purchase the camera and use it for the first time. Most people do not bother to subsequently alter it, or may not even know how to set the time. Moreover, most digital cameras contain a separate secondary internal battery which sole purpose is to power the internal clock. Such batteries may power the clock for many years irrespective of the state of the main camera battery, which may remain discharged for long periods when the camera is not in use.

The first step is to translate the timestamps of the images into universal time (UTC) using:

$$t_{UTC} = t_{local} - Z \quad (2)$$

where Z is the camera time zone measured in hours before or after universal time or it is zero if the clock is set according to UTC. Then images are temporally clustered using the log-difference between consecutive images:

$$d(i) = \log_{10}[t(I_i) - t / I_{i-1}] \quad (3)$$

The log-difference $d(i)$ gives an indication of the temporal separation between images and an interpretation is given in Table 2 based on t measured in milliseconds.

Table 2. Temporal clustering of images (in milliseconds)

$d(i)$	Temporal image separation	Interpretation
1-2	milliseconds	Multishot of single scene
3-4	seconds	Same scene
5-6	minutes	Same event
7	hours	Same day
8	days	Same journey
9	weeks	Unrelated
10	months	Unrelated
11	years	Unrelated
12	decades	Unlikely

Figure 1 demonstrates how the log time differences are used to temporally interpret a collection of 37,625 amateur images. The number of occurrences for each difference category is also plotted using a log scale as the various categories differ greatly. The graph shows that there are just over 100 differences that are 9 or larger suggesting that the image collection comprises about 111 clusters of related images. About 10,000 images are taken with only millisecond pauses and represent multi shots. Moreover, about 20,000 images are taken with a few seconds apart signaling that they represent different images taken of the same scene. Next, about 6,000 images are taken a few minutes apart suggesting that these belong to similar events. Finally about 300 images are separated by a few hours suggesting that these belong to the same day.

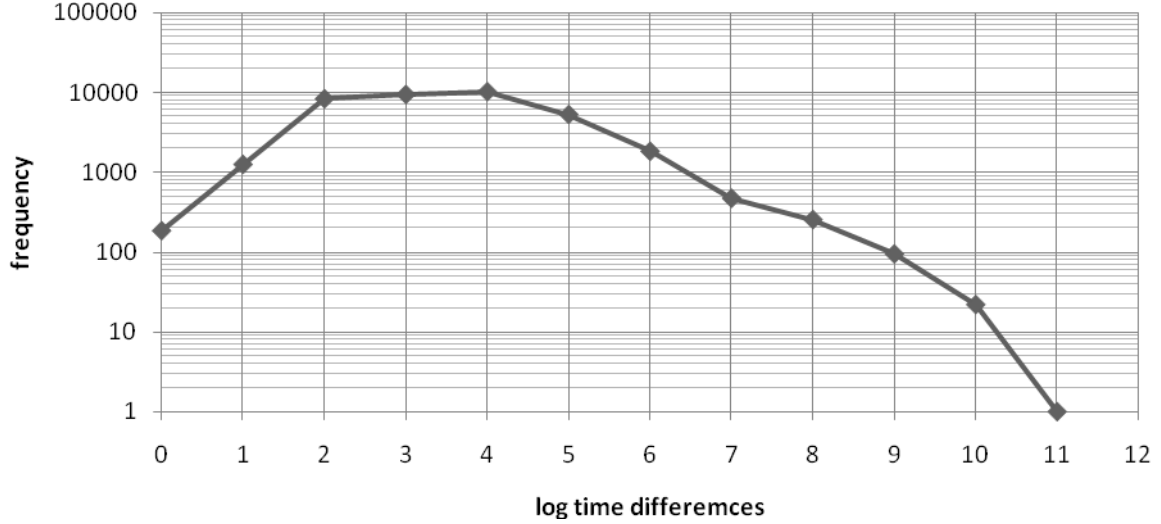


Figure 1. A log-log plot of time differences between consecutive images for a collection of 37,625 amateur images. The horizontal axis shows the log time differences and the vertical axis shows the log-frequency of occurrence.

Next, a cluster is split into days representing a window of 24 hours centered on midday universal time, i.e., centered around 12:00 UTC. The images in the 24 hour chunk is divided into hourly bins, i.e., the images taken the first hour are placed into one bin, the images taken the second hour are placed in the second bin, and so forth.

For each bin j the image with the maximum exposure value is denoted $EV_{max,j}$ taken at time t_j is determined. If the image has an exposure value greater than 15 (a direct shot of the sun) then the next largest exposure value is used instead. Next, all exposure values below 10 are discarded as images with such exposure values usually are indoor or night images. Moreover, all maximum exposure values that are more than one exposure value units smaller their adjacent values are discarded, i.e., $EV_{max,j}$ is discarded if $EV_{max,j+1} - EV_{max,j} > 1$ or $EV_{max,j-1} - EV_{max,j} > 1$.

The remaining time and exposure value pairs are used to fit a sinusoidal to represent the suns motion across the skies of the form

$$f(t) = A\sin(Bx + C) + D \tag{4}$$

Clearly, the period of the sinusoidal is 24 hours representing one rotation of the Earth and B is therefore given by

$$B = \frac{\pi}{12} \quad (5)$$

The remaining coefficients A , C and D are found by least squares fitting using:

$$\min \sum_i (EV_{\max,i} - f(t_i))^2 \quad (6)$$

where the following constraints are set: $2 < A < 8$, $0 < C < \pi/12$ and $0 < D < 14$.

The C coefficient signals the time of the maximum sun elevation, i.e., midday occurs when:

$$Bt_{\text{midday}} + C = \frac{\pi}{2} \quad (7)$$

That is, midday occurs at

$$t_{\text{midday}} = \frac{1}{B} \left(\frac{\pi}{2} - C \right) \quad (8)$$

Then, the longitude is:

$$\lambda = 2\pi \frac{12 - t_{\text{midday}}}{24} \quad (9)$$

measured in radians where positive values represent degrees west and negative values represent degrees east.

Once the longitude is estimated the approximate latitude of the observer can be estimated. The number of hour degrees of the sunset is given by:

$$t_{\text{sunset}} = \frac{\sin^{-1} \left(\frac{10 - D}{A} \right) - C}{B} \quad (10)$$

Note that the sunset exposure value is set to 10 according to the ANSI exposure value interpretations [18, 19]. Knowing the local time of sunset in radians then the classic sun equation can be used to estimate the latitude as follows

$$\varphi = \tan^{-1} \left(\frac{-\cos(t'_{\text{sunset}})}{\tan(\delta)} \right) \quad (11)$$

where t'_{sunset} is in degree angles relative to midday, namely

$$t'_{sunset} = \frac{\pi}{12} |t_{midday} - t_{sunset}| \quad (12)$$

and the declination of the sun (in degrees) is approximated by

$$\delta = -0.4092797 \cos\left(\frac{2\pi}{365}(M + 10)\right) \quad (13)$$

and M is the day in the year the pictures were taken. The estimation process is illustrated in Figures 2 and 3. Note that usually one refers to latitude and longitude in that particular order. However, in this study we refer to longitude first as longitude is a primary measure and latitude is a secondary measure.

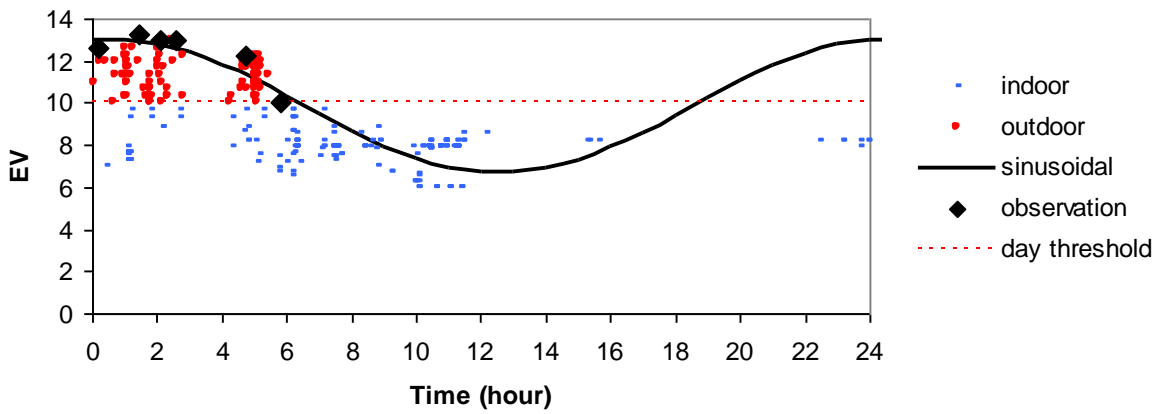


Figure 2. Determining the approximate sun elevation path from temporal exposure values for Tokyo, Japan – 24 hour view. Midday is estimated at 0.56 UTC and sunset at 6.73 UTC.

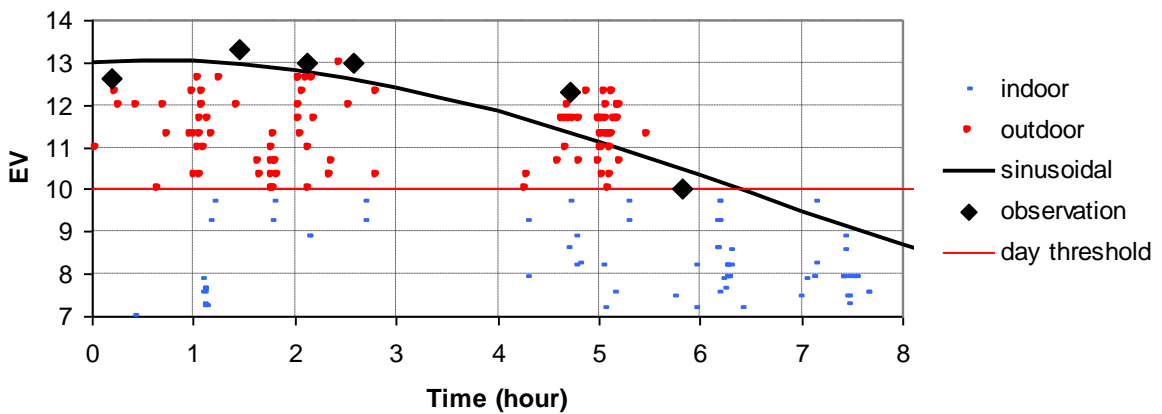


Figure 3. Determining the approximate sun elevation path from temporal exposure values for Tokyo, Japan – 8 hour detail.

4 Experimental evaluation

The strategy was applied to the author's personal image collection which at the time of writing comprises 37,625 8-megapixel images which are all taken with the same digital camera - a Sony FSC Sony DSC-F828. The collection has been manually clustered into events with a brief explanation of where and under what circumstances the pictures were taken. Next, all events involving indoor images such as conferences and laboratory photographs were omitted. Moreover, very small collections comprising outdoor images taken during very narrow time intervals were also discarded. Finally, only one set from each city was used to ensure a geographical spread as a majority of the photographs in the author's collection were taken in Taiwan. The result comprised 3,046 unique images taken over a period of four years at various locations around the world. Table 3 lists attributes of the image collection including the city, country and continent of the where the images were taken, the official longitude and latitude for these cities, the number of images in each collection, the number of days spanned by each event and the date the images were taken (start-date).

The geographical classifier was implemented in java and was run on a Dell personal computer with an AMD Athlon Dual core processor and 4 Gb RAM running Windows Vista Personal edition. Drew Noakes' freely available (EXIF) metadata-extractor library (available at <http://drewnoakes.com/>) was used for extracting EXIF information from the images. It took 2 minutes and 25 seconds to traverse the image collection and extract the EXIF information. The geographical clustering step took 20 seconds. This yields a processing delay of 0.05 seconds per image. Note that the Java code was not optimized. Significant time savings could be achieved by simply removing console output and employing a more efficient least-squares optimization engine.

Table 3. The image collection used for the experiments, comprising 3,046 images taken during a four year period.

City	Country	Continent	Longitude	Latitude	No. images	days	Start date
Brisbane	Australia	Oceania	153° East	27° South	198	2	8 Jul 2009
Cape Town	South Africa	Africa	18° East	35° South	641	6	21 Feb 2009
Indiana	USA	Americas	79° West	40° North	99	1	8 Oct 2007
Kaohsiung	Taiwan	Asia	121° East	25° North	136	1	5 Feb 2008
Oregon	USA	Americas	122° West	45° North	98	1	10 Oct 2007
Oslo	Norway	Europe	10° East	59° North	387	3	23 Jun 2008
Paris	France	Europe	2° East	48° North	290	4	10 Aug 2008
San Juan	Puerto Rico	Americas	66° West	18° North	780	8	21 Jul 2006
Seoul	South Korea	Asia	127° East	37° North	197	2	27 Apr 2007
Tokyo	Japan	Asia	139° East	35° North	220	2	16 Apr 2006
Wuhan	China	Asia	108° East	30° North	358	5	1 Sep 2007

Table 4. Experimental results

	Estimated location		Error		
	Longitude	Latitude	Longitude	Latitude	Overall
Brisbane	134° East	35° South	19°	8°	21°
Cape Town	16° East	77° South	2°	42°	42°
Indiana	66° West	71° North	13°	31°	34°
Kaohsiung	100° East	52° North	21°	27°	34°
Oregon	108° West	79° North	14°	34°	37°
Oslo	34° West	51° North	44°	8°	45°
Paris	8° West	74° North	10°	26°	28°
San Juan	76° West	67° North	10°	49°	50°
Seoul	106° East	63° North	21°	26°	33°
Tokyo	124° East	68° North	15°	33°	36°
Wuhan	104° East	82° North	4°	52°	52°
		Mean	15,7°	30,5°	37,4°
		SD	11,3°	14,2°	9,3°

The classifier successfully identified the events according to the temporal patterns. Table 4 summarizes the results of the experiments, including the estimated longitudes and latitudes for the collections, the longitudinal and latitudinal errors as well as an overall error. Note that for simplicity Euclidian distance was used to compute the overall error. The results show that the mean longitudinal error was 15.7 degrees. Cape Town and Wuhan were determined with the highest longitudinal accuracy while Oslo had the lowest longitudinal accuracy. Clearly, an error of roughly 15 degrees means that the continent of an image can be determined with high confidence, and, in many instances, the country can also be determined. In context of the longitudinal range of 360 degrees a longitudinal error of 15.7 degrees which equates to a relative longitudinal error of approximately 4.6%.

The mean latitudinal error was 30.5 degrees, or twice that of the longitudinal error. In context of the latitudinal range of 180 degrees the overall latitudinal error is 16.9%. Clearly, it was harder to determine the latitude compared to the longitude. However, for all the tests run the approach resulted the correct classification of hemisphere, i.e., all locations were determined to be on the Northern hemisphere apart from Cape Town and Brisbane which were successfully classified as belonging on the Southern hemisphere. A latitudinal error of 30.5 degrees provides only a very coarse grained geographic accuracy. The longitude estimates are predominantly linked to the temporal patterns of the images. If the temporal patterns have certain traits, such as being limited to a narrow time window during the day, or focused on an unusual time of day, say night photographs, then this can significantly affect the longitudinal accuracy.

Next, the calculation of latitude is closely tied to the observed length of day. If the length of day is inaccurate the latitude will be greatly affected. The results show that it is harder to estimate the length of day compared to estimating the local midday. With too few high-exposure-value measurements, or given measurements that are temporally too close, the days may be incorrectly observed as too short. Another problem occurs for events spanning several days where the photographer has travelled long distances. For example imagine a photographer travelling across USA during a week. The photographer may start up in Florida and finish in Washington State. Firstly, these locations are separated by multiple time-zones and Florida is in the south while Washington State is in the north. Locations closer to the equator has less variations in day length while the day length of locations further away from the equator vary significantly according to the season. If these days are treated as one event with one overall longitude measurement, then one may end up with day length observations that are too large since the difference between the overall longitude measurement and the day length measurements of the west-most location may be larger than that of a single day. Consequently, in addition to resulting in large errors, one may erroneously estimate the incorrect hemisphere. This is because days are shorter on the hemisphere with winter and longer on the hemisphere with summer, and while there is winter on one hemisphere it is summer on the other, and vice versa. The confidence intervals of the estimations are illustrated in Figures 4 and 5.

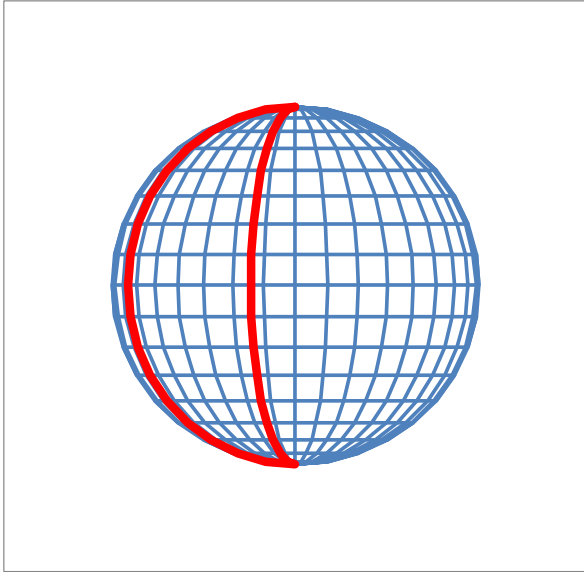


Figure 4. An orthographic projection of the confidence interval for geographic estimates based on just temporal information. It is only possible to estimate the longitude.

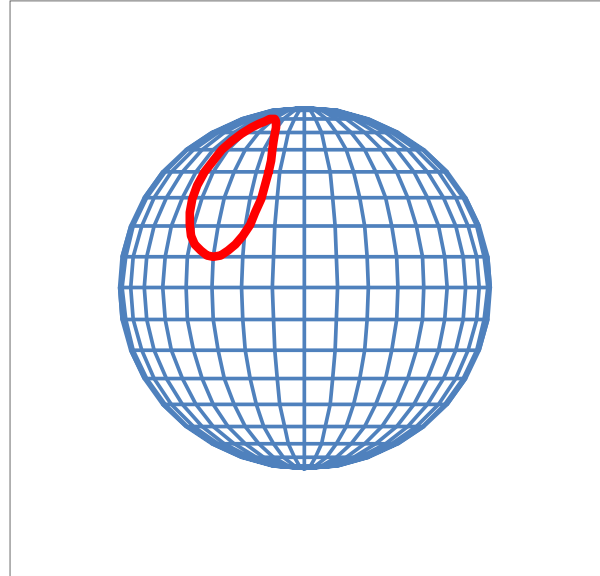


Figure 5. An orthographic projection of the confidence interval for geographic estimates based on both temporal and exposure information. Both longitude and latitude estimates can be made. The longitudinal estimates are more accurate.

A consequence of these problems is that estimates should be based on observations taken over a daily window to avoid the problems of long travels within a multiple day journey. Moreover, for each day one has to assess the validity of the measurements. If the measurements comprise a low ratio of images with high exposure values the observations for that particular day need to be discarded and the measurements of neighboring days should be used instead.

This is illustrated in Table 5 which shows the details for a 9-day car journey around the Midwest, USA, during July, 2005, involving sightseeing and indoor meetings with representatives from universities. The table lists the estimated longitude, latitude, the number of images for each day and the ratio of outside images, that is, images with an exposure value greater than 10. The table reveals that the hemisphere predictions for days 2, 4, 5, 6 and 9 are correct, while the hemisphere predictions for day 1, 3, 7 and 8 are incorrect. Moreover, when comparing these with the ratio of outside images it is clear that days 3, 7 and 8 have the lowest ratio of outside images of 1 %, 5 % and 13 %, respectively. Similarly, the days with correct hemisphere predictions all have high ratios of outside images ranging from 21-43 %. The only exception is the first day, which has a high ratio of outside images. The

reason why this day is different may be due to other factors such as jetlag after arriving into the USA on a flight. Clearly, using the ratio of outside images one can assess the validity of the hemisphere predictions. Finally, any disagreement, such as the first day, can be resolved through a majority vote, where the predictions for each day count as a vote. In this example day 3, 7 and 8 would lose their vote because of low outside ratio, and North would receive a majority vote for North of 5:1. In this example the majority vote of 5:4 would also give the correct hemisphere prediction without considering the ratio of outside photographs.

Table 5. Detailed for a 9-day car journey around the Midwest, USA, during July, 2005.

Day	city	Longitude	Latitude	No. images	Outside ratio
1	Ann Arbour	155° West	51° South	37	27 %
2	Ann Arbour	129° West	11° North	67	28 %
3	South Bend	153° West	35° South	91	1 %
4	Chicago	96° West	50° North	96	45 %
5	Chicago	164° West	65° North	60	33 %
6	Southern Illinois	75° West	57° North	91	21 %
7	Southern Illinois	71° West	47° South	85	5 %
8	Iowa	147° West	5° South	123	13 %
9	Chicago	107° West	52° North	110	43 %

Table 6. Test suite collected from photograph collections shared via Picasa web album.

Location	Source (owner)	Camera	Time-zone	No. images	Days	Start date
Brishbane	manoharpala	Olympus FE3010 X895	5	91	2	9 Jan 2010
Cape Town	aga.moodley	Nikon D60	-1	774	6	13 Dec 2008
Indiana	anne.raker	Cannon SD790	5	102	1	14 Sep 2009
Kaohsiung	michael.alling	Sony W300	-8	307	1	18 Jul 2009
Oregon	nevda	Cannon SD750	5	154	7	4 Oct 2009
Oslo	kOKSak	Cannon EOS 400D	-2	236	2	17 Sep 2007
Paris	haaann	Panasonic DMC-L27	-7	204	2	17 Apr 2008
San Juan	ravisharma	Canon EOS REBEL Xti	3	113	3	30 Aug 2008
Seoul	Theos766	Nikon S520	6	115	1	23 Dec 2009
Wuhan	mahmoodkhan77	Sony DSC P73	3	68	1	15 Jun 2004

Finally, to verify the generality of the approach a collection of photographs taken by people unknown to the author were collected from the Picasa web album (<http://picasaweb.google.com/>). This is a service that allows users to share photographs. Moreover, images can be easily downloaded in bulk from the searchable database. The 2,164 image test suite were acquired using the place names listed in Table 3 as search keywords and visually inspecting whether these collections were representative of the respective locations. Table 6 lists collection details including the Picasa web-album account

name of the photographer, the size of each collection and as the estimated relative timezone of the camera. The selected collections were used without any manual intervention. The relative timezones were estimated in three ways. For most of the collections sunset photographs were selected and the time of these photographs were compared to the sunset times obtained using a sunset calculator (<http://www.timeanddate.com/>) with the location and day of year as parameters. In the Seoul collection an obvious daytime photograph and nighttime photograph closely spaced in time were used to measure the sunset time and the result aligned with the actual sunset time. The Kaohsiung and Oslo Collections were easily aligned with the camera time as they contained images of clocks. The Kaohsiung collection contained a nighttime image of a train station with a huge LED-style neon digital clock while the Oslo collection contained several photographs of the city hall in Oslo with the analogue clock tower face clearly visible on several of the images. It is natural to assume that these public clocks represent accurate local times. Next, these collections are obtained with several different models of cameras from different vendors including an Olympus FE3010, Nikon D60, Panasonic DMC-L27, Cannon EOS 400D, Sony Cybershot DSC-W300, Cannon Powershot SD750, Canon EOS REBEL Xti, Panasonic DMC-L27 and Nikon COOLPIX S520. This variation gives support to the claim that the proposed strategy is camera independent. It is assumed that the date settings in the collections obtained are correct.

Table 7. Results obtained with the test suite from the Picasa web-album.

location	estimated location			error		
	longitude	latitude		longitudinal	latitudinal	overall
Brisbane	142 East	60 South		11	33	35
Cape Town	3 West	65 South		21	30	37
Indiana	86 West	85 North		7	45	46
Kaohsiung	161 East	29 North		40	4	40
Oregon	125 West	80 South		3	125	125
Oslo	16 West	88 North		26	29	39
Paris	38 West	72 South		40	120	126
San Juan	89 West	67 North		23	49	54
Seoul	89 East	31 North		38	6	38
Wuhan	110 East	58 North		2	28	28

Table 7 lists the obtained results. The results are less accurate than those obtained using the authors own collection as the mean longitudinal error is 21 degrees and the mean latitudinal error is 46 degrees. However, no longitude estimate deviates by more than 40 degrees and the most accurate is

within 2 degrees of the true longitude. In terms of latitude then the correct hemisphere is correctly determined in 8 of 10 cases, indicating that the estimates are much better than random guesses. Only Oregon and Paris are assigned the incorrect hemisphere and hence greatly affect the mean error. When omitting these latitude estimates the mean latitudinal error is just 28 degrees, which is close to what was obtained with the author's own images. One explanation for the false hemisphere classifications could be that the date settings of the camera used were incorrect as the incorrect day of year will result in an erroneous length of day estimate. Little is known about the process with which the images were taken and there could also be other sources of errors. On the positive side, the latitudes of Kaohsiung and Seoul are within 4 and 6 degrees of the true latitude, respectively.

Future work will focus on reducing the latitudinal error by improving the sunrise/sunset estimation. In order to achieve this it may be necessary to also exploit image contents. Rough estimates have shown that analysis of shadows in images can result in geographic estimates with a high accuracy [20].

5 Limitations of this study and future work

The approach presented in this study depends on the availability of EXIF meta-information and the approach will consequently only work with photographs taken with digital cameras that provide such information. Very low-cost cameras and older digital cameras may not have these facilities such as web-cams. However, most digital cameras on sale today provide detailed EXIF information, even camera enabled mobile phones. A minimum requirement for obtaining geo-spatial information is that the images are time stamped. Rough longitudinal estimates can be obtained based on solely the time-stamps. However, to obtain latitudinal estimates images without EXIF meta-information would have to be analyzed based on image contents. Future research will therefore focus on combining the proposed approach with content based strategies to both obtain latitudinal estimates from images without EXIF information and to improve the geo-spatial estimation precision of images with EXIF meta-information. One obvious approach is to automatically analyze the lighting conditions in the contents of the images by particularly focusing on the sky.

Moreover, the proposed method will not work if the meta-information is compromised through editing, automatic distortion or camera clock disruptions. However, most people have too many photographs and have no time to modify or alter the meta-information. This is in fact one of the fundamental problems of large image collections, i.e., there is no time to manually label or tag the data. Second, digital cameras are fashion accessories and most people have relatively updated

cameras. Most cameras, and even quite old digital cameras, have a separate internal battery for the camera clock that draw little power and can last for many years. Third, most photo editing software will leave the meta-information intact apart from parameters that are directly affected by the editing operation such as reduced pixel resolution after a down-sampling operation.

Next, the proposed method relies on collections of photographs taken with temporal spread, i.e., photographs taken throughout a significant portion of the day. The method will not work well if there are too few images or the images are taken during a very narrow time-interval. However, in many realistic holiday situations the photographer is active throughout the day and the temporal spread thus emerges quite naturally.

Finally, the method requires a portion of the photographs to be taken outdoors, but not all. In the absence of photographs without sun related exposure values it is not possible to make day-length observations and consequently no latitude estimates. However, it is possible to make rough longitude estimates based on the time-stamps alone for collections of indoor-only photographs.

Future work includes improving the accuracy and precision of the strategy by utilizing the contents of the images. For instance, by analyzing the colors and the intensity of the sky in the images a more accurate determination of the sunset times can be obtained.

Although the precision of the current approach is low it may be combined with other approach such as geo-tagging based on landmark recognition [21] where the current approach can be used to determine the approximate location and limit the search conducted using landmark recognition.

6 Conclusions

A strategy for estimating the geographic origin of photograph collections based on temporal patterns and image exposure values was presented. Temporal exposure information is recorded by most digital cameras. The strategy assumes that the images are taken with one camera and that the settings of the internal clock remain unaltered for the duration the photographs were taken. The temporal information is used to group images into days and the exposure values are used to estimate the time midday (sun peak) and sunrise/sunset. Having obtained these, the longitude and latitude is determined using classic celestial equations. Dark images, i.e., images without sufficiently large exposure values are also successfully classified due to their temporal similarity to images with sufficiently high exposure values. Experiments involving real-world amateur image collections revealed that the strategy resulted in longitudinal and latitudinal errors of 15.7 and 30.5 degrees,

respectively. The current strategy is an alternative to GPS image tagging and can be used to classify existing images, retrospectively. The proposed strategy is computationally inexpensive as only image attributes are used and no time-consuming image content analysis is performed. On average it took approximately 50 milliseconds to successfully classify each image with modest hardware and an inefficient implementation. Moreover, the statistical nature of the approach means that the results are robust to variations in photographing behaviour. Future work involves reducing the errors and increasing the accuracy by combining the meta information with information extracted from the image contents. One drawback of the proposed approach is that it is unable to classify collections exclusively comprising indoor images.

Acknowledgements

The author is grateful to the insightful comments of the anonymous reviewers, invaluable help from the special issue guest editors, in particular Professor Zhiwen Yu, and the comments made by Professor Bendik Bygstad that put me on the path of Viking navigation and sunstones.

References

- [1] C.-J. Jang, J.-Y. Lee, J.-W. Lee, and H.-G. Cho, "Smart Management System for Digital Photographs using Temporal and Spatial Features with EXIF metadata," in the proceedings of 2nd International Conference on Digital Information Management, pp. 110-115, 2007.
- [2] D. Carboni, S. Sanna, and P. Zanarini, "GeoPix: image retrieval on the geo web, from camera click to mouse click," in the proceedings of Proceedings of the 8th conference on Human-computer interaction with mobile devices and services, pp. 169-172, 2006.
- [3] GAO, "GLOBAL POSITIONING SYSTEM: Significant Challenges in Sustaining and Upgrading Widely Used Capabilities," United States Government Accountability Office 2009.
- [4] F. Cozman and E. Krotkov, "Robot localization using a computer vision sextant," in the proceedings of IEEE International Conference on Robotics and Automation, pp. 106-111, 1995.
- [5] A. Trebi-Ollennu, T. Huntsberger, Y. Cheng, and E. T. Baumgartner, "Design and analysis of a sun sensor for planetary rover absolute heading detection.," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 6, pp. 939 - 947, 2001.
- [6] N. Jacobs, S. Satkin, N. Roman, R. Speyer, and R. Pless, "Geolocating Static Cameras," in the proceedings of IEEE 11th International Conference on Computer Vision (ICCV 2007), pp. 1-6, 2007.
- [7] N. Jacobs, N. Roman, and R. Pless, "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.
- [8] P. Alvarez, "Using Extended File Information (EXIF) File headers in Digital Evidence Analysis," *International Journal of Digital Evidence*, vol. 2, no. 32004.
- [9] N. L. Romero, V. V. G. C. G. Chornet, J. S. Cobos, A. A. S. C. Carot, F. C. Centellas, and M. C. Mendez, "Recovery of descriptive information in images from digital libraries by means of EXIF metadata," *Library Hi Tech*, vol. 26, no. 2, pp. 302-315, 2008.

- [10] S. Ahern, M. Naaman, R. Nair, and J. Hui-I Yang, "World explorer: visualizing aggregate data from unstructured text in geo-referenced collections," in the proceedings of 7th ACM/IEEE-CS joint conference on Digital libraries, pp. 1-10, 2007.
- [11] L. A. Jones, "Sunlight and skylight as determinants of Photographic exposure. I. Luminous density as determined by solar altitude and atmospheric conditions," *Journal of the Optical Society of America*, vol. 38, no. 2, pp. 123-178, 1948.
- [12] L. A. Jones, "Sunlight and skylight as determinants of Photographic exposure. II. Scene structure, directional index, photographic efficiency of daylight, safety factors, and evaluation of camera exposure," *Journal of the Optical Society of America*, vol. 39, no. 2, pp. 94-135, 1949.
- [13] L. A. Jones and H. R. Condit, "The Brightness Scale of Exterior Scenes and the Computation of Correct Photographic Exposure," *Journal of the Optical Society of America*, vol. 31, no. 11, pp. 651-678, 1941.
- [14] S. F. Ray, "Camera Exposure Determination," in *The Manual of Photography: Photographic and Digital Imaging*, R. E. Jacobson, S. F. Ray, G. G. Atteridge, and N. R. Axford, Eds.: Focal Press, 2000.
- [15] M. Szummer and R. W. Picard, "Indoor-outdoor image classification," in the proceedings of IEEE International Workshop on Content-Based Access of Image and Video Database, pp. 42-51, 1998.
- [16] N. Serrano, A. Savakis, and A. Luo, "A computationally efficient approach to indoor/outdoor scene classification," in the proceedings of 16th International Conference on Pattern Recognition, pp. 146-149, 2002.
- [17] J.-F. N. Lalonde, S. G. Efros, A. A., "What does the sky tell us about the camera?," in the proceedings of European Conference on Computer Vision, 2008.
- [18] ANSI, "ANSI PH2.7-1973 American National Standard Photographic Exposure Guide," American National Standards Institute, New York 1973.
- [19] ANSI, "[ANSI PH2.7-1986. American National Standard for Photography - Photographic Exposure Guide," American National Standards Institute, New York 1986.
- [20] F. E. Sandnes, "Sorting holiday photos without a GPS: What can we expect from contents-based geo-spatial image tagging?," *Lecture Notes on Computer Science*, vol. 5879, no., pp. 256-267, 2009.
- [21] Y.-T. Zheng, Z. Ming, S. Yang, H. Adam, U. Buddemeier, A. Bissacco, F. Brucher, T.-S. Chua, and H. Neven, "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, 2009.