A Simple Content-based Strategy for Estimating the Geographical Location of a Webcam

Frode Eika Sandnes

Faculty of Engineering, Oslo University College, P.O. Box 4 St. Olavs Plass, N-0130 Oslo, Norway
frodes@hio.no

Abstract. This study proposes a strategy for determining the approximate geographical location of a webcam based on a sequence of images taken at regular intervals. For a time-stamped image sequence spanning 24 hours the approximate sunrise and sunset times are determined by classifying images into day and nighttime images based on the image intensity. Based on the sunrise and sunset times both the latitude and longitude of the webcam can be determined. Experimental data demonstrates the effectiveness of the strategy.

Keywords: image analysis, geographical information system, webcam

1 Introduction

Geographical information systems are becoming increasingly important in computer science. One avenue of geographical information systems relates to images. Some photographers attach GPS devices to their digital cameras in order to geo-tag images with the location where the images were taken. Geo-tagged images can simplify, speed up and enhance photo browsing activities – especially with very large image collections [1, 6]. However, there are several problems with GPS technology. First, current GPS devices may need several minutes to lock onto overhead satellites. This may be unacceptable if a photographer needs to shoot sudden scenes. Second, current GPS receivers consume a lot of power. Third, still few digital cameras are equipped with built in GPS receivers. Fourth, the GPS infrastructure is reaching the end of its lifetime and there is no guarantee that this service will be available in the future [9].

Several non-GPS approaches have been attempted. For instance, landmark recognition has been used to identify image scene locations [23]. If one recognizes a known landmark in an image and knows the location of the landmark, then it follows where the image was photographed.

Direct sun observations have also been used to determine the geographical location of the observer [7, 22]. In particular, a digital camera has been used to implement a digital sextant for robot navigation where the sun elevation is obtained by measuring the distance between the sun and the horizon. There are several problems with this approach. First, one requires direct sun observations, and the sun is not visible on cloudy days. Second, a very wide angle lens is required to measure high sun

elevations close to 90 degrees that occur close to the equator. Third, knowledge about the optical characteristics of the lens is needed to translate pixels distances into angular distances. Four, although several good horizon extraction algorithms exist [5, 8], it may be difficult to accurately identify if the horizon is obstructed by objects such as trees, small hills and buildings.

In order to omit some of the problems with direct sun measurements it has been proposed to measure the sun elevation based on the height of objects and the length of the shadows cast by these objects [19], although no automatic systems employing this approach have been demonstrated yet.

An alternative non-content based approach that also works on cloudy days has been proposed where the camera exposure characteristics have been used to estimate the midday and sunrise or sunset in collections of related image series [18]. Most digital cameras embed camera exposure characteristics such as shutter speed, film speed and aperture [3, 4, 13-16] in images using the EXIF format [2, 12, 17]. This approach achieved a longitudinal accuracy of approximately 15 degrees and a latitudinal accuracy of approximately 30 degrees. However, this strategy relies on extra meta-information.

Image contents have also been used to determine the relative geographical position of a network of webcams [10, 11] where images are taken at regular images with statically positioned webcams.

This study proposes another image based geographical positioning system. The strategy assumes a geographically fixed webcam accessible via the Internet. Assuming that regular time stamped images can be taken using this webcam the proposed approach can determine the approximate location of the web-cam purely based on the contents of the images returned. Applications of this include the determination of webcam locations for unlabelled webcams, the corroboration and confirmation of the published location for particular webcams or self-configuring mobile webcams that can autonomously determine their own location.

2 Method

The proposed approach is based on regularly sampling a webcam for 24 hours. For each image the strategy determines if the image is a nighttime or daytime shot. Several related algorithms for this have been proposed such as indoor-outdoor classification strategies [20, 21]. A content based method is needed since most webcams do not embed EXIF exposure data in the images. For the purpose of webcams a very simple intensity based strategy was employed where the overall intensity of an image I sampled at time t measured in universal time (UTC) is calculated:

$$s = \frac{1}{X \cdot Y} \sum_{x=1}^{X} \sum_{y=1}^{Y} I_{x,y} \quad (1)$$

where $I_{x,y}$ is the combined red, green and blue component of each pixel, namely:

$$I_{x,y} = \frac{I_{x,y,r} + I_{x,y,g} + I_{x.y.b}}{3} \quad (2)$$

Next, candidate sunrise and sunset times are found as follows. A sunrise image candidate satisfies $s_i < s_{mid}$ and $s_{i+1} > s_{mid}$ and a sunset image candidate satisfies $s_i > s_{mid}$ and $s_{i+1} < s_{mid}$, where s_i is the intensity of image i and s_{mid} is the midpoint between the minimum and maximum intensity for the 24-hour image sequence, namely:

$$s_{mid} = \frac{s_{\min} + s_{\max}}{2} \quad (3)$$

where s_{min} and s_{max} are the minimum and maximum intensity values for the 24-hour image sequence.

If there are several sunrise or sunset candidates then the ones that result in the largest day and night sequences are chosen. Finally, to obtain a more accurate sunrise/sunset estimate linear interpolation is employed by finding the time $t_{sunrise/sunset}$ where the line that passes through (t_i, s_i) and (t_{i+1}, s_{i+1}) has a height of s_{mid}, namely:

$$t_{sunrise/sunset} = \frac{t_{mid} - b}{a} \quad (4)$$

where

$$b = s_i - at_i \quad (5)$$

and

$$a = \frac{s_{i+1} - s_i}{t_{i+1} - t_i} \quad (6)$$

An alternative strategy is to identify sunrise and sunset points where the intensity timeline passes a threshold relative to the maximum value. A sunrise point is detected at time t_i if $s_i < s_{max}\ T$ and $s_{i+1} > s_{max}\ T$. Similarly, a sunset is detected at time t_i if $s_i > s_{max}\ T$ and $s_{i+1} < s_{max}\ T$, where T is the threshold and s_{max} is the maximum intensity value in the 24-hour image sequence. In this study the threshold was set to 0.8 which was found through experimentation.

Having established the sunrise time $t_{sunrise}$ and sunset time t_{sunset}, then midday is simply calculated as occurring at

$$t_{midday} = \begin{cases} \frac{t_{sunset} - t_{sunrise}}{2} & if\ t_{sunrise} < t_{sunset} \\ \frac{24 + t_{sunset} - t_{sunrise}}{2} \bmod 24 & otherwise \end{cases} \quad (7)$$

The angular sunrise time is then

$$a_{xunrise} = \begin{cases} \pi \dfrac{t_{sunrise} - t_{midday}}{12} & if\ t_{sunrise} < t_{midday} \\ \pi \dfrac{t_{midday} - t_{sunrise}}{12} & otherwise \end{cases} \quad (8)$$

And the angular sunset time is

$$a_{sunset} = -a_{sunrise} \quad (9)$$

The angular sunset can be used to determine the latitude using the classic sunrise equation with solar disk correction, namely

$$\cos(a_{sunset}) = \frac{\sin(-0.83) - \sin(\varphi) \cdot \sin(\delta)}{\cos(\varphi) \cdot \cos(\delta)} \quad (10)$$

which is numerically solved for latitude φ. Here the declination of the sun can be approximated by

$$\delta = -0.4092797 \cos\left(\frac{2\pi}{365}(M + 10)\right) \quad (11)$$

where M is the current day number of the year. Next, the longitude is simply

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

2.1 Accuracy

Sample rate is the major factors affecting the accuracy of the location estimates although there also are other factors. Imagine that a webcam is observed at a rate of s images a day (24 hours) then the longitudinal accuracy is given by

$$\lambda_{accuracy} = \frac{\pi}{s} \quad (13)$$

and the latitudinal accuracy is

$$\varphi_{accuracy} = \tan^{-1}\left(\frac{-\cos(\lambda_{accuracy})}{\tan(\delta)}\right) \quad (14)$$

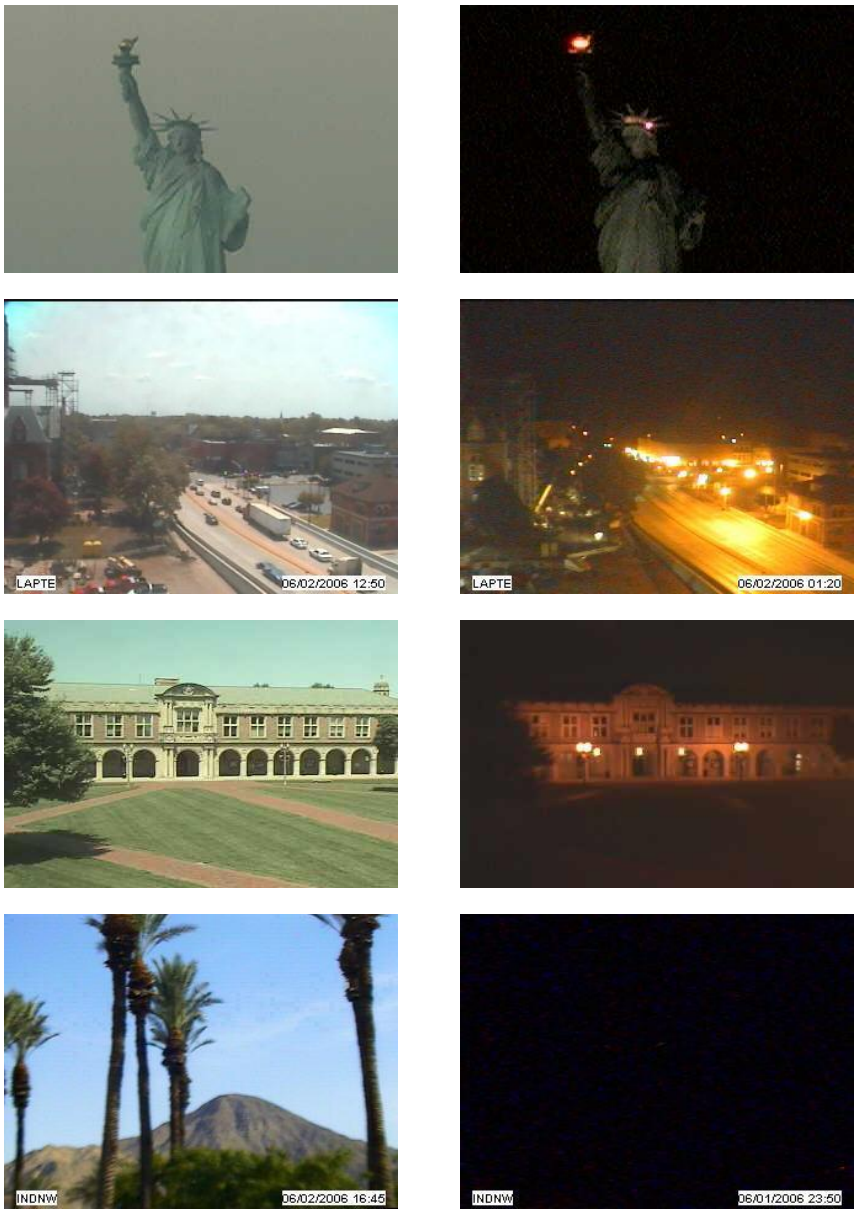

Fig. 1. The test-suite used in this study were taken from four webcams. Both daytime and nighttime images are shown.

3 Experimental evaluation

The approach presented herein was assessed using a test-suite of images taken from the AMOS Archive of Many Outdoor Scenes [10] (see Fig. 1). The subset of images used comprised data obtained from four webcams during 2-7 June 2006 recorded in Central Time (UTC-5 or UTC-6). One of the webcams is located in New York overlooking the statue of liberty.

Table 1 shows the results obtained with the proposed strategy for the four webcams. Both the results obtained for the individual days (June 2-6) and mean for all the five days are listed. Webcam 15 is the only one which is known (New York City). The true latitude and longitude for New York City are 40.7 degrees north and 74.0 degrees west, respectively. The true sunrise and sunset times for June 2, 2006 is 4.43 and 19.35, respectively, or 4:26 and 19:21 in hour:minute format. These values were corroborated using two online sunrise-sunset calculators. Note that this is local time which is at UTC+5.

The results show that the mean sunrise is 8 minutes early and the sunset is 6 minutes late yielding a day length error of 14 minutes. The standard deviation for the sunrise and sunset estimates are 9 and 11, respectively.

Next, the estimated latitude and longitudes for webcam 15 is 42.6 degrees north and 73.1 degrees west, respectively. This yields a latitudinal error of 1.9 degrees, and longitudinal error of 0.9 degrees. The coordinates found points to Cheshire, Massachusetts, US. The achieved results are more than ten times more accurate than what has been achieved with arbitrary image collections [18] which achieved latitudinal and longitudinal errors of 30 and 15 degrees, respectively. Note also that the latitudinal error is nearly twice that of the longitudinal error, which is consistent with previous research.

The results for the remaining three cameras with unknown location are equally consistent and the standard deviations for these measurements, especially webcam 190, are smaller. This suggests that these results may be even more accurate than the one for the New York City webcam.

When plotting the coordinates obtained using Google maps it is found that webcam 190 at (47.4°, -88.1°) is at Eagle Hauk, Michigan, US, webcam 4 at (40.8°, -90.2°) is at Maquon, Illinois, US and webcam 82 at (41.4°, -116.3°) is at Humboldt National Forrest, Nevada, US.

3.1 Effects of threshold

Fig. 2 shows how the threshold T affects the results as the obtained latitude and longitude are plotted against threshold. The plot shows that both the latitude and longitude is closest to the actual latitude and longitude with a threshold of 0.8. Then the threshold exceeds 0.85 the accuracy decays rapidly.

Table 1. Geographical results obtained for the four web-cams.

webcam	day	sunrise	sunset	lat.	Long.
15	June 2	4.33	19.38	42.1°	72.9°
	June 3	4.52	19.40	40.4°	74.4°
	June 4	4.30	19.22	40.5°	71.4°
	June 5	4.20	19.65	45.0°	73.9°
	June 6	4.12	19.60	45.1°	72.9°
	mean	**4.29**	**19.45**	**42.6°**	**73.1°**
	SD	**0.15**	**0.18**	**2.3**	**1.2**
190	June 2	5.10	20.83	47.6°	89.5°
	June 3	4.98	20.80	48.1°	88.4°
	June 4	4.93	20.72	47.6°	87.4°
	June 5	4.98	20.70	47.0°	87.6°
	June 6	5.00	20.70	46.7°	87.8°
	mean	**5.00**	**20.75**	**47.4°**	**88.1°**
	SD	**0.06**	**0.06**	**0.5**	**0.9**
4	June 2	5.60	20.55	41.2°	91.1°
	June 3	5.57	20.57	41.5°	91.0°
	June 4	5.67	20.45	39.2°	90.9°
	June 5	5.38	20.40	41.3°	88.4°
	June 6	5.48	20.50	41.1°	89.9°
	mean	**5.54**	**20.49**	**40.8°**	**90.2°**
	SD	**0.11**	**0.07**	**0.9**	**1.2**
82	June 2	7.20	22.30	42.6°	116.3°
	June 3	7.15	22.28	42.7°	115.8°
	June 4	7.30	22.23	40.7°	116.5°
	June 5	7.30	22.37	41.7°	117.5°
	June 6	7.27	22.10	39.4°	115.3°
	mean	**7.24**	**22.26**	**41.4°**	**116.3°**
	SD	**0.07**	**0.10**	**1.4**	**0.8**

Longitude is least affected by the threshold. This is probably because a change in threshold affects the sunrise and sunset estimation times equally and since the longitude is based on the midpoint, the error cancels out. However, the latitude is

more strongly affected since it is based on the length of day. Low thresholds result in too long day estimates, that is, too early sunrises and too late sunsets. Consequently, the latitude estimates are too large. Similarly, with a too high threshold the day length estimates will be too short, that is, a too late sunrise and too early sunset, which again leads to too small latitude estimates.

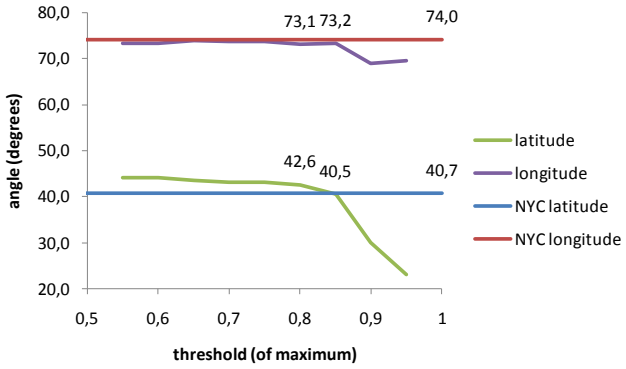

Fig. 2. Effects of threshold on mean longitude and latitude accuracy.

3.2 Intensity

Fig. 3 shows intensity traces obtained for June 2, 2006 using the four webcams. The plots confirm that the image series fall into two distinct categories of day and nighttime images. Although there are some variations for each group, the groups are significantly different from each other. The graphs also show that there are more variations during the day compared to the night. This is what we would expect as there are more activities in the scenes at night and there are varying lighting conditions according to the cloud conditions, etc.

Moreover, there were more image variations for the statue of liberty webcam compared to the others. An inspection of the webcam images reveals that this is because this webcam zooms in and out on the statue of liberty. Although the position and orientation is constant the scale is not.

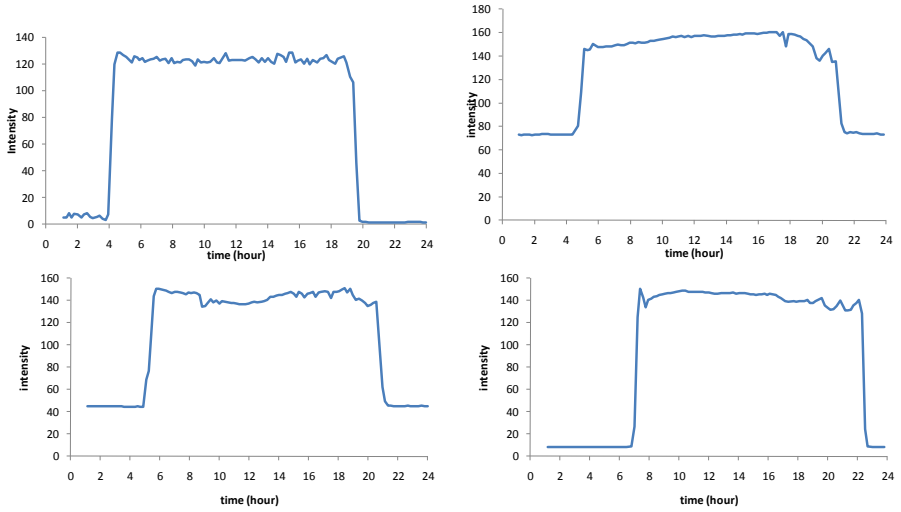

Fig. 3. Intensity traces for June 2, 2006 for the four webcams.

3.3 Effects of sample rate

If the webcams takes S regularly spaced images during a 24 hour period, then the mean interval between two neighboring image is $w = 24 \times 60 / S$ minutes. In the worst case the sunrise-sunset points may be off by $w/2$ minutes. Table 2 shows this in terms of the measurements obtained in this study. Clearly, webcam 4 and 82 has nearly twice as high sample rate as webcam 15 and 190 and consequently the errors for webcam 4 and 82 are smaller than those for webcam 15 and 190. The potential error related to sample rate for the New York City measurements are a latitudinal error of 2.1 degrees and a longitudinal error of 1.43 degrees. The true coordinates are well within these limits.

4 Conclusions

This study explored the possibility of combining content-based information with celestial mathematics to determine the geographical location of webcams. The proposed strategy is computationally effective and simple to implement and an accuracy of about 2 degrees was achieved.

Table 2. Effect of sample rate.

set	Images per day	mean interval (min.)	max midday error (min.)	max sunset error (min.)	lat. error (deg.)	long. error (deg.)
15	136,54	11,43	5,71	5,71	2,07	1,43
190	141,28	11,04	5,52	5,52	1,62	1,38
4	226,63	6,88	3,44	3,44	1,41	0,86
82	245,35	6,36	3,18	3,18	1,32	0,79

References

1. Ahern, S., Naaman, M., Nair, R., and Hui-I Yang, J.: 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)
2. Alvarez, P.: Using Extended File Information (EXIF) File headers in Digital Evidence Analysis. International Journal of Digital Evidence, 2, 3 (2004)
3. ANSI: ANSI PH2.7-1973 American National Standard Photographic Exposure Guide. American National Standards Institute, New York (1973)
4. ANSI: ANSI PH2.7-1986. American National Standard for Photography - Photographic Exposure Guide. American National Standards Institute, New York (1986)
5. Bao, G.-Q., Xiong, S.-S., and Zhou, Z.-Y.: Vision-based horizon extraction for micro air vehicle flight control. IEEE Transactions on Instrumentation and Measurement, 54(3), pp. 1067-1072 (2005)
6. Carboni, D., Sanna, S., and Zanarini, P.: 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)
7. Cozman, F. and Krotkov, E.; Robot localization using a computer vision sextant. In the proceedings of IEEE International Conference on Robotics and Automation, pp. 106-111 (1995)
8. Ettinger, S. M., Nechyba, C., and lfju, P. G.; Towards Flights autonomy: Vision-based horizon detection for micro air vehicles. In the proceedings of IEEE International Conference on Robotics and Automation, (2002)
9. GAO: GLOBAL POSITIONING SYSTEM: Significant Challenges in Sustaining and Upgrading Widely Used Capabilities. United States Government Accountability Office (2009)
10. Jacobs, N., Roman, N., and Pless, R.: 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)
11. Jacobs, N., Satkin, S., Roman, N., Speyer, R., and Pless, R.: Geolocating Static Cameras. In the proceedings of IEEE 11th International Conference on Computer Vision (ICCV 2007), pp. 1-6 (2007)
12. Jang, C.-J., Lee, J.-Y., Lee, J.-W., and Cho, H.-G.: 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)

13. Jones, L. A.: 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, 38(2), pp. 123-178 (1948)
14. Jones, L. A.: 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, 39(2), pp. 94-135 (1949)
15. Jones, L. A. and Condit, H. R.: The Brightness Scale of Exterior Scenes and the Computation of Correct Photographic Exposure. Journal of the Optical Society of America, 31(11), pp. 651-678 (1941)
16. Ray, S. F.: 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)
17. Romero, N. L., Chornet, V. V. G. C. G., Cobos, J. S., Carot, A. A. S. C., Centellas, F. C., and Mendez, M. C.: Recovery of descriptive information in images from digital libraries by means of EXIF metadata. Library Hi Tech, 26(2), pp. 302-315 (2008)
18. Sandnes, F. E.: Geo-Spatial Tagging of Image Collections using Temporal Camera Usage Dynamics. In the proceedings of I-SPAN 2009, pp. 160-165 (2009)
19. Sandnes, F. E.: Sorting holiday photos without a GPS: What can we expect from contents-based geo-spatial image tagging?. Lecture Notes on Computer Science, 5879, pp. 256-267 (2009)
20. Serrano, N., Savakis, A., and Luo, A.: A computationally efficient approach to indoor/outdoor scene classification. In the proceedings of 16th International Conference on Pattern Recognition, pp. 146-149 (2002)
21. Szummer, M. and Picard, R. W.: Indoor-outdoor image classification. In the proceedings of IEEE International Workshop on Content-Based Access of Image and Video Database, pp. 42-51 (1998)
22. Trebi-Ollennu, A., Huntsberger, T., Cheng, Y., and Baumgartner, E. T.: Design and analysis of a sun sensor for planetary rover absolute heading detection. IEEE Transactions on Robotics and Automation, 17(6), pp. 939 – 947 (2001)
23. Zheng, Y.-T., Ming, Z., Yang, S., Adam, H., Buddemeier, U., Bissacco, A., Brucher, F., Chua, T.-S., and Neven, H.: 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)