Reasoning about Photo Collections using Models of Outdoor Illumination

Daniel Hauagge
hauagge@cs.cornell.edu
Scott Wehrwein
swehrwein@cs.cornell.edu
Paul Upchurch
paulu@cs.cornell.edu
Kavita Bala
kb@cs.cornell.edu
Noah Snavely
snavely@cs.cornell.edu

Cornell University
Ithaca, NY, USA

Abstract

Natural illumination from the sun and sky plays a significant role in the appearance of outdoor scenes. We propose the use of sophisticated outdoor illumination models, developed in the computer graphics community, for estimating appearance and timestamps from a large set of uncalibrated images of an outdoor scene. We first present an analysis of the relationship between these illumination models and the geolocation, time, surface orientation, and local visibility at a scene point. We then use this relationship to devise a data-driven method for estimating per-point albedo and local visibility information from a set of Internet photos taken under varying, unknown illuminations. Our approach significantly extends prior work on appearance estimation to work with sun-sky models, and enables new applications, such as computing timestamps for individual photos using shading information.

1 Introduction

Natural illumination plays a critical role in the appearance of outdoor scenes, and in the variation of scene appearance over time; for example, Figure 1 shows images from a photo collection of the Statue of Liberty illustrating appearance changes under different illumination conditions. Many vision tasks, such as photometric stereo and intrinsic image decomposition, require reasoning about this illumination and how it interacts with the scene.

Figure 1: The Statue of Liberty under a variety of natural illumination conditions

Although outdoor illumination is highly variable, it is far from arbitrary; in fact, it is dominated by a few

elements—sun, sky, and weather—which in turn depend fundamentally on scene location, time, and atmospheric conditions. Indeed, the computer graphics community has developed increasingly sophisticated models of outdoor illumination that, given parameters such as geolocation and time, compute a predicted outdoor environment map. Surprisingly, such illumination models are not yet widely used in computer vision despite illumination's importance in the appearance of outdoor scenes.

Our work explores the connection between community photo collections of an outdoor scene at a given location on Earth, and the distribution of lighting conditions for that scene predicted by these illumination models. The usage of these predictive models to reason about scenes from unstructured photos is still a major challenge, in part because timestamps are often missing or erroneous—community photos represent a "soup" of different observations of the scene under varying but unknown illumination. Our insight is to match *statistics* of outdoor illumination with pixel statistics derived from photo collections. We build on the photometric ambient occlusion work of Hauagge et al. [7], which explored the connection between pixel statistics and simple illumination distributions in relation to the local visibility (or ambient occlusion) of each scene point. Our work generalizes this model to handle the more realistic scenario of varying illumination in outdoor scenes.

This paper includes an analysis of how the geographic position, surface normal, and local geometry of a point interact with illumination models, and how multiple measurements of a point's appearance over time can be used to estimate albedo and local visibility for points in a scene. Since our photometric approach relies on varying illumination, we analyze the conditions under which a lack of sufficient variation can arise, and use this analysis to detect areas in a given scene (e.g., surface points that are almost always pointed away from the sun) where estimates of albedo and other appearance properties will be unreliable.

Our albedo estimation has further practical value in estimating sun positions in uncalibrated Internet photos with missing or erroneous timestamps (when geolocation is known, sun position follows from timestamps and vice versa). Such estimated sun positions are useful in analysis of outdoor illumination in photos, such as in photometric stereo, shadow detection, or grouping photos by light similarity. Timestamps are useful in correcting clocks on consumer cameras, and discovering patterns of photography (e.g., which time of day is most popular for taking photos of a given landmark).

In summary, our work has three main contributions:

- A model for relating distributions of outdoor illumination to properties of points in a scene, including albedo, surface normal, and local visibility.
- An algorithm for estimating surface properties from photo collections under varying, unknown illumination.
- A method for determining sun position in Internet photos based on shading.

2 Related work

Our paper fits in the theme of recent work that reasons about outdoor illumination to improve scene understanding from one or more images.

Estimating illumination from outdoor imagery. Several vision methods have been proposed to recover illumination from single images [14]. Lalonde et al. [13] use an analytic sun-sky model as a cue for determining sun direction, by using the model to predict the appearance of the visible sky in a single outdoor photograph. Our method uses a sun-sky

model to estimate and predict the appearance of *objects* using statistics across many images. Both methods can be used to timestamp images, and we compare to [13] in Section 5.

Haber et al. [6] pose the problem of estimating reflectance and illumination of a scene from Internet photos as an explicit inverse rendering problem, which results in a complex optimization procedure. Their work assumes arbitrary (smooth) illumination; in contrast, we leverage strong models of outdoor illumination to derive a much simpler statistical approach.

Outdoor photometric stereo. Several techniques estimate scene geometry and appearance from outdoor illumination over time, particularly from webcam data. Sunkavalli et al. decompose webcam video into components modeling albedo, geometry, and shading using a factoring approach [17, 18]. Ackermann et al. [2] and Abrams et al. [1] estimate scene albedos and normals using photometric stereo, using the sun as the varying light source. These methods rely on images captured from a single, static, georegistered camera, with known timestamps (and hence sun position). In contrast, we work with images taken from many viewpoints and cameras with largely incorrect timestamps. Finally, Yu et al. solve photometric stereo problems for images using environment light measured using light probes [20]. Again, their data is more structured in that they directly measure illumination.

Intrinsic image decomposition. Intrinsic image techniques have also been used to estimate albedo and illumination maps from single or multiple images [16, 19]. Laffont et al. also work with multiple images from varying viewpoints from one [11] or many [12] points in time. However, their approach either requires extra input (e.g., a light probe [11]), or additional smoothness priors [12]. Our statistical approach yields a per-pixel estimate that avoids smoothness priors and requires only measurements from the images themselves. Our work builds on the method of Hauagge et al. [7], generalizing it to work under significantly more practical illumination.

3 Modeling illumination for outdoor points

As discussed above, outdoor illumination exhibits great variability, but is nonetheless highly structured. The illumination reaching an outdoor scene point is influenced by a few key factors. The primary source of illumination during the day is the sun, whose position in the sky is a function of **geographic location** and the **time and date** of an observation. The location and date constrain the sun position to a well-defined path, while the time of day determines where the sun lies on that path. We denote location using latitude (ϕ) and longitude (λ), and time and date as t. The intensity and color of the sun, as well as the light from the sky caused by atmospheric scattering, vary as a function of both sun position and weather.

Weather adds immense complexity to outdoor illumination; the degree of variation increases greatly with the variety of clouds, fog, haze, and other atmospheric effects. State-of-the-art outdoor illumination models largely ignore weather and assume clear skies; since we are taking advantage of these models, we leave the incorporation of more varied weather conditions as future work. Furthermore, clear, sunny skies provide the most informative illumination for photometric methods such as ours that rely on varying and strongly directional illumination. We discuss later how to use weather records to discard cloudy images.

Having discussed the factors affecting the illumination coming from the sun and sky, we now consider scene-related properties that affect how much of this light hits a given point in a scene. The **surface orientation** at a point affects how much and which portion of the sky's illumination reaches it. For instance, if the normal is facing away from the sun's path it will

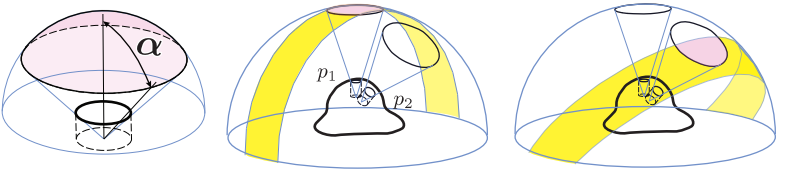

Figure 2: Influence of geolocation, date and time, orientation, and local visibility on the illumination at a point. Left: the cylindrical crevice model illustrating the local visibility angle α. In the middle, an object located near the equator sees a band of sun paths (shaded in yellow) that is centered directly overhead. In this location, the point p_1 at the bottom of a crevice can sometimes see the sun, whereas p_2 cannot. On the right, we see a location farther from the equator, with a different sun path, where the reverse is true.

never receive direct sunlight (see Figure 2). Further, a point in an open field pointing upwards towards the sky will see the entire sky dome, while a point facing downward will see less of the sky dome and more of the ground.

Finally, the illumination arriving at a point can be affected by its **local visibility**—the extent to which surrounding geometry occludes its view of the sky dome. As a simple way to describe the potentially complex local geometry around a point, we adopt the crevice model proposed by Hauagge et al. [7]. By modeling local geometry as a single cylindrical hole, we are able to describe the extent of occlusion using a single parameter, α, representing the angle from the point's surface normal to the opening of the crevice (see Figure 2).

In summary, our model considers the illumination of an outdoor scene point on a clear day as a function $L(\phi,\lambda,t,\alpha,\vec{n})$ where (ϕ,λ) are the geographic latitude and longitude, t is the time and date, $\vec{n}$ is the normal vector, and α is the local visibility angle given by our crevice model. To make predictions based on our model, we use the physically-based sun/sky model proposed by Hosek and Wilkie [8], which produces a sunny environment map given geographic location and time of day (ϕ,λ,t). We can then choose any surface normal and visibility angle $(\vec{n},\alpha)$, and integrate the irradiance over the visible portion of the environment map to acquire a value for $L(\phi,\lambda,t,\alpha,\vec{n})$, as illustrated in the top part of Figure 3.

4 Albedo and sun position in photo collections

In this section, we describe how to use the model described above to estimate local visibility (ambient occlusion) and albedo of scene points; we then describe a method for using the albedo to estimate the illumination and timestamp of individual photos.

Our method takes as input a set of photos of an outdoor scene from different viewpoints and varying, unknown times.[1] We first create a sparse 3D reconstruction using SfM and multi-view stereo, and then georegister the reconstructed scene. We project each reconstructed point into the images in which it appears to retrieve a set of observed color values for that point. The geo-registered model gives us the location (ϕ,λ) and surface normal $(\vec{n})$ for each point in the scene.

4.1 Estimating albedo for sunlit outdoor scenes

We assume that all surfaces are Lambertian and that a point's albedo does not change over time. Our method works on color images by treating each channel independently, so for

[1] Later, we assume we know the date for each photo, but not the time of day. This is consistent with our experience of errors in image timestamps.

simplicity we refer only to intensity. We use a simplified image formation model where a single observation of a point x is given by

$$I_x = \rho_x L(\phi_x, \lambda_x, t, \alpha_x, \vec{n}_x) \tag{1}$$

where I_x is the observed intensity for scene point x in image I, and ρ_x denotes its albedo. If we could find accurate values for ϕ_x, λ_x, t, α_x, and $\vec{n}_x$, we could recover the albedo ρ_x by dividing the observed intensity by the predicted illumination. The 3D reconstruction provides values for ϕ_x, λ_x, and $\vec{n}_x$, but images from Internet photo collections generally have unknown or uncertain time (t) and local visibility (α_x). For this reason, we cannot directly predict illumination for a single image in practice.

However, we have many observations of x across different images, which can provide insight about the *distribution* of intensity values observed at that point. Likewise, our lighting model can be used to predict the expected distribution of illumination conditions over the course of a year. Building upon the method proposed by Hauagge et al. [7], we match predicted statistics to observed statistics in order to estimate the local visibility and albedo of each point in the scene. Given many images that view x distributed over the year, we can estimate the expected intensity of x, $\mathcal{E}[I_x]$, by averaging the observed samples. If ρ is constant over time, then Eq. 1 implies:

$$\mathcal{E}[I_x] = \mathcal{E}[\rho_x L(\phi_x, \lambda_x, t, \alpha_x, \vec{n}_x)] = \rho_x \mathcal{E}[L(\phi_x, \lambda_x, t, \alpha_x, \vec{n}_x)] \tag{2}$$

Suppose that we know x's local visibility angle $\alpha_x = 90°$. In this case, we have all the information we need to compute $\mathcal{E}[L]$ using the sun/sky model, and we can compute ρ_x as:

$$\rho_x = \frac{\mathcal{E}[I_x]}{\mathcal{E}[L(\phi_x, \lambda_x, t, 90°, \vec{n}_x)]} \tag{3}$$

where the expectation of L is computed over a set of times t sampled throughout a full year.

4.2 Estimating local visibility angle

We can now compute albedo for a point if its local visibility angle is known, but in practice α_x is unknown and must be estimated as well. Hauagge et al. propose a technique to estimate α directly from image observations by computing a statistic over image observations I_x that is independent of ρ [7]:

$$\kappa_x = \frac{\mathcal{E}[I_x]^2}{\mathcal{E}[I_x^2]} = \frac{\mathcal{E}[L_x]^2}{\mathcal{E}[L_x^2]} \tag{4}$$

By assuming point-source illumination that moves uniformly over the hemisphere, they derive an analytical relationship between κ and α, which allows them to compute α based on the observed value of κ.

Under our more sophisticated illumination model, a closed-form relationship is harder to find; κ now depends on location, α, and $\vec{n}$. However, for a given scene location, our model allows us to predict the value of κ for a given normal and visibility angle. In particular, for a fixed location (ϕ, λ), we can compute the relationship between κ, α and $\vec{n}$ as

$$\kappa(\alpha, \vec{n}) = \frac{\mathcal{E}[L(\phi, \lambda, t, \alpha, \vec{n})]^2}{\mathcal{E}[L(\phi, \lambda, t, \alpha, \vec{n})^2]} \tag{5}$$

For a given point x with normal $\vec{n}_x$, we compute its observed κ value κ_x using Equation 4. The visibility angle α_x is chosen to be the value of α such that the predicted $\kappa(\alpha_x, \vec{n}_x)$ most closely matches the observed κ_x. Figure 3 (c) shows images of κ and expected illumination $\mathcal{E}[L]$ for several values of α; (d) and (e) show examples of κ and predicted illumination curves for three different normals. For a monotonically increasing κ curve such as the blue curve in Figure 3(d), we can simply take the observed value κ_x and look up the corresponding value of α to assign an estimated local visibility.

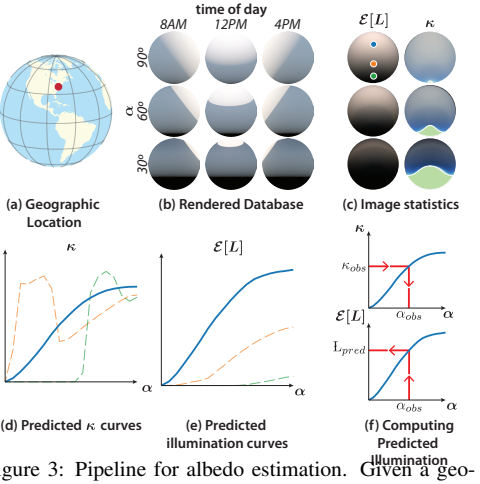

Figure 3: Pipeline for albedo estimation. Given a geographic location (a), we tabulate $L(\phi, \lambda, t, \alpha, \vec{n})$ over all parameter values (b, visualized as spheres). (c) We compute $\mathcal{E}(L)$ and κ for each α, producing curves showing (d) $\kappa(\alpha)$, and, (e) $\mathcal{E}(L)$ for each normal. (f) For an observed value of κ, we look up α and then predicted average illumination $\mathcal{E}(L)$, allowing us to estimate albedo. Green regions in (c) correspond to combinations of normal direction and crevice depth for which we cannot reliably recover albedo.

4.2.1 κ and sun visibility

Note, however, that the κ curves for normals that do not directly face the sun path, such as the orange curve in Figure 3(d), do not monotonically increase over all values of α. For each normal, there is some angle α_{min} below which the point is in a deep enough crevice that the sun never appears within its visibility cone. Just above α_{min}, κ is small because the high relative intensity of the sun's rare appearances causes the denominator ($\mathcal{E}[I^2]$) to increase more than the numerator ($\mathcal{E}[I]^2$). As α increases, more frequent sun visibility causes κ to increase monotonically (for details on the behavior of κ, see [7]). If $\alpha \geq \alpha_{min}$, we can invert this well-behaved portion of the κ function and look up the α value corresponding to an observed κ, as shown in Figure 3(f).

If $\alpha < \alpha_{min}$, a fundamental ambiguity exists—the value of κ is not explained by variation in sun visibility and is thus not informative in the context of our model. As a result, we discard the portion of the curve that lies below α_{min} to achieve an invertible monotonic curve explaining crevices with local visibility greater than α_{min} (see Figure 4). In practice, very deep crevices are not commonly observed in natural scenes and we do not see these as a common failure mode in our results. Further, we can choose a value α_0 and make a strong guarantee that the model will not break down for crevices with $\alpha \geq \alpha_0$ (that is, shallower than α_0), by discarding points with normals for which $\alpha_{min} < \alpha_0$. It is important to note that this limitation, imposed by lack of variability in illumination, will affect any photometric method, because the observations lack sufficient information to disambiguate between albedo and illumination.

Fortunately, α_{min} is a well-defined function of normal direction and latitude, so it can be determined *a priori* for each normal. For robustness, we let α_{min} be the smallest α such that the point sees the sun at least 10% of the time. We then threshold the κ curves at this value and use only the monotonic portion (above the threshold) to look up α. Figure 4 shows α_{min} for all normals in two different geographic locations, and some sample thresholded κ curves. Once we have estimated α_x, we have values for all parameters of L and we can compute ρ_x using Equation 3.

4.3 Estimating time of day

We now consider the task of determining the time a given image in the input set was captured, making use of the estimated albedos ρ_x. Under our simple image formation model, we can compute lighting by simply dividing the observed intensity by the albedo. Our approach to timestamping is to compare such a lighting estimate for a single image to a set of predicted illumination conditions over a range of times and choose the timestamp where the lighting matches most closely. In particular, we first estimate the illumination at a given point by dividing the observed intensity by our estimated albedo:

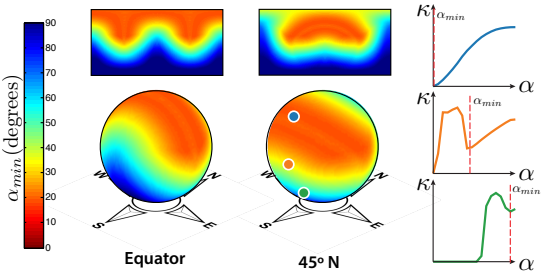

Figure 4: An illustration of how α_{min} varies with geographic latitude and normal. Each point on the sphere represents a surface normal direction, and its color encodes the α angle of a crevice that sees the sun 10% of the time during daylight. Normals pointing near the sun path (as determined by latitude) have lower α_{min} values and are more informative because κ is meaningful over a greater range of α. The rightmost column shows κ curves and α_{min} for the three different surface normals.

$$L_x^{obs} = \frac{I_x}{\rho_x + \varepsilon} \tag{6}$$

where ε is a small constant to achieve robustness to noise. We collect the estimated illumination for all visible points in an image I into a vector L_I^{obs}, and generate a corresponding predicted illumination vector $L_I^{pred}(t)$ for each hypothesized timestamp t using our model. To constrain the search space, we assume that the date in the image metadata is known and accurate, but not the time of day.

We found it important to perform a normalization before comparing illumination vectors to increase contrast and overcome noise in our lighting estimates. The most effective strategy was to normalize each vector so that the bottom and top 10 percentiles span the range $[0,1]$. We compute the cost $c(I,t)$ for time t as a robustified L^2 distance: we sort the element-wise differences $\vec{L}_I^{obs} - \vec{L}_I^{pred}(t)$ and discard the top and bottom 10%. Our final cost function is

$$c(I,t) = \|R(\vec{L}_I^{obs} - \vec{L}_I^{pred}(t))\|_2 \tag{7}$$

where $R(\cdot)$ is the robustification operator above. The need for robustness in this distance measure is mainly due to phenomena not captured by our model, such as cast shadows (where L^{obs} is darker than L^{pred}) and specular highlights (where L^{obs} is brighter than L^{pred}). Finally, the timestamp for an image is chosen by finding the time t that minimizes the cost $c(I,t)$.

5 Experiments and applications

Evaluating our method is a challenging task, as ground truth albedo and timestamps are difficult to acquire for photo collections. To evaluate our method in a more controlled setting we created the TENTACLE dataset with 100 images of a 3D-printed object taken outdoors over the course of a sunny day. We also created an analogous synthetic dataset, TENTACLER, by rendering the same object using a physically-based renderer under the sun-sky model [8] at times sampled throughout a day.

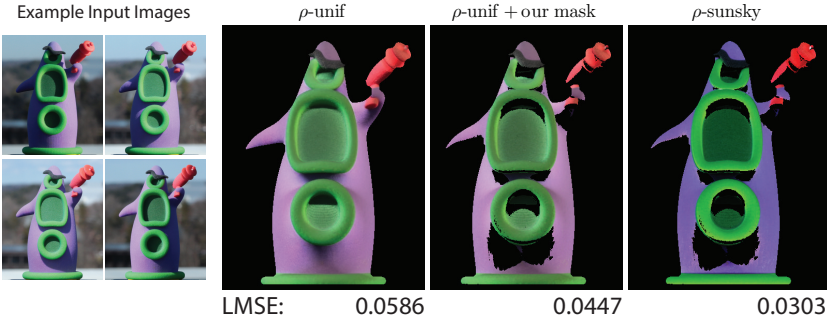

Figure 5: Results of our algorithm compared to Hauagge et al., with and without the mask generated using our approach. We also list the grayscale local mean squared error for each result.

We also collected two photo collection datasets from Flickr: STATUE and CASTLE. STATUE contains 78K images of the Statue of Liberty in New York, USA. CASTLE contains 33K images of a theme park attraction in Florida, USA. Ground truth timestamps for 347 CASTLE images were manually entered by reading the time from a clock in the scene. For STATUE we evaluate our timestamping method on an additional set of 265 images from the AMOS webcam dataset [9] with known ground truth timestamps (STATUEA). STATUEA is distinct from STATUE in that we did not use it for albedo estimation.

5.1 Albedo

A comparison of the albedo obtained with the technique proposed by Hauagge et al. [7] (ρ-unif) and our method (ρ-sunsky) for the TENTACLE dataset is shown in Fig. 5. Our technique recovers a significantly flatter albedo and successfully identifies and discards points which cannot be recovered accurately. We quantitatively compare our performance to ρ-unif using the Local Mean Squared Error (LMSE) metric [5]. Note that the TENTACLE "ground truth" is the albedo sent to the 3D printer, but colors are not reproduced perfectly in 3D printing. Therefore, we use the LMSE on grayscale images to evaluate the piecewise constant albedo without penalizing the color mismatch. We also computed LMSE on the result from [7] with the mask generated by our method, showing that a significant improvement can be made by identifying for which points the albedo cannot be estimated accurately using photometric methods, as discussed in our analysis in Section 4. Our method, which combines this analysis with our more realistic lighting model, performs best.

5.2 Timestamps

Quantitative results on the timestamping task are shown in Table 1. We compare the performance of our timestamping method with the albedo from [7] (t-unif) versus our albedo method (t-sun). As baselines we also compare to raw Exif timestamps (CASTLE only) and "chance" error (Rand), the average expected error given by guessing a random time between sunrise and sunset. Restricting Rand to daylight hours allows its average error to outperform the Exif timestamps. Comprehensive results can be found on the project webpage [2].

As the table shows, our method performs best on all datasets, although errors get larger as the datasets become less structured. We significantly outperform the single-image method of

[2] http://www.cs.cornell.edu/projects/photo_outdoor_illum

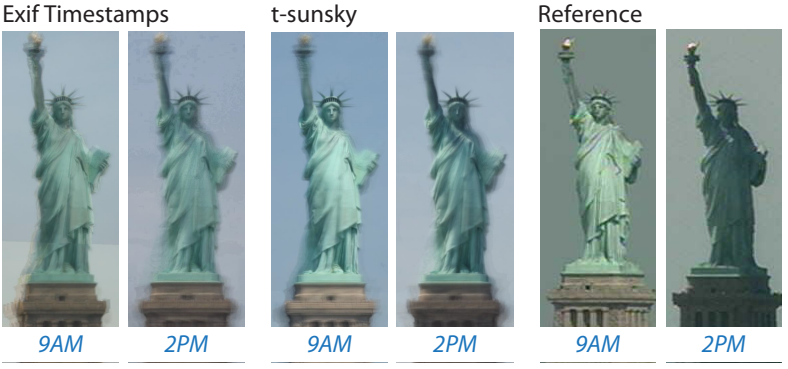

Figure 7: Alternative visualization of timestamping results. We take a set of images taken in July with a given timestamp, reproject them to a single viewpoint using a homography, and average them. We show that Exif timestamps (left) do not produce coherent lighting when averaged due to the timestamp errors. Using our timestamps (middle), the average images match the lighting in the corresponding reference images (right), taken from the AMOS dataset where timestamps are known. Note distinct cast shadow at 2PM and the clear change from 9AM to 2PM.

Lalonde et al. [13], demonstrating the value of using many images to reason about a scene. The difference between t-sun and t-unif is smaller on Internet datasets, where noise in the input data affect t-sun more. For example, our method relies on the surface normal estimates from multiview stereo, which we have observed to contain significant noise. Another source of error on Internet datasets is tone mapping, which we have modeled using a simple $\gamma = 2.2$ approach. Future improvements in radiometric calibration and better normals will improve results for our method. Finally, a common failure case occurs when our algorithm mistakenly assigns a cloudy image to either sunset or sunrise, when shadows are also very diffuse.

Figure 6 shows a sample result, including an input photo along with the scoring curve used to determine the timestamp. For an alternative visualization of our timestamping results see Fig. 7, where we show the average of all images for a given time (after registration using a homography) as estimated using either raw Exif tags or our method.

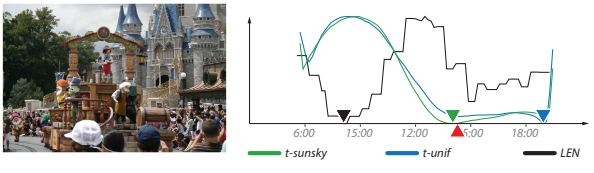

Figure 6: An example timestamping result (full results in supplemental). The plot on the right shows timestamp error metrics over the full day for the image on the left. Colored markers show the global minimum (estimated timestamp) for each method, and the red marker indicates the ground truth timestamp.

Dataset (# images)	TentacleR (200)		Tentacle (100)		StatueA (265)		Castle (347)	
	Avg	Med	Avg	Med	Avg	Med	Avg	Med
Rand	248.6	237.0	218.2	208.6	230.8	216.3	231.6	219.4
EXIF	–	–	–	–	–	–	287.5	211.0
LEN	–	–	–	–	249.3	223.0	195.0	150.0
t-unif	27.6	29.3	58.0	61.8	133.3	80.5	114.7	74.4
t-sun	9.9	9.8	53.1	46.9	136.9	49.3	87.0	57.3

Table 1: Average and median timestamp error (in minutes) for various methods on our datasets. Rand represents chance over daylight hours, while LEN is the method of Lalonde et al. [13]. Sun position error in degrees can be calculated approximately by dividing minutes by 4.

5.3 Implementation details

We start with a large collection of Internet photos, and use SfM to obtain camera extrinsics and intrinsics along with a sparse set of 3D points [3], which we manually georegister. We then use PMVS to compute a larger point set with surface normals [4] and recompute the visibility list for each camera using a z-splatting algorithm (to increase the number of observations per point). We approximate the response of all cameras as a gamma curve with $\gamma = 2.2$, as is common in prior work [6]. We discard pixels with intensities very near 0 and 1. We use $\varepsilon = 10^{-6}$ when dividing by albedo to find illumination in Eq. (6).

Because current sun-sky illumination models are limited to clear skies, we restrict the input to our albedo estimation phase to images taken on days with limited cloud cover. For STATUE, we use NOAA weather records [15] to select days when cloud cover in the area did not exceed 25%. CASTLE had few cloudy images, so weather-based pruning was unnecessary.

To render our database of predicted illumination spheres we use a physically based renderer [10], illuminating spheres with the sun-sky model in [8], and adding a ground plane with an albedo of 0.15. To simulate different α angles in the renderings, we use a modified diffuse shader that integrates only over the appropriate solid angle of the hemisphere. We generate predictions sampled every 10 minutes over the daylight hours of a full year, for α angles in 5° increments from 5° to 90°. When computing statistics over these samples, we adjust the distribution in the predicted statistics to more closely match the observed data. We do this by weighting each predicted illumination value by the portion of images whose Exif timestamps indicate that they were taken on that day. For TENTACLE, we used only the L values from the times when images were captured.

6 Conclusion

We have presented a new way to utilize models of sun-sky illumination in computer vision, by leveraging them to predict statistics of illumination that can be matched to pixel statistics in large, unstructured image collections. This matching process results in a per-point algorithm for estimating albedo and local visibility. These illumination models can then be further used to timestamp images based on shading. The algorithm's limitations stem from its photometric approach: we require multiple images of a scene under different illumination, and cannot recover albedo or estimate timestamps in the absence of such variation.

We believe there are many interesting avenues for future work in connecting natural illumination to image statistics. For instance, it would be useful to explicitly model weather (especially in cloudy parts of the world), so that our distributions over illumination account for this important phenomenon. It would also be interesting to relate image statistics to other surface properties, such as gloss, so as to handle more general materials. Eventually, we hope to use our techniques to model temporally varying appearance, such as the weathering effects on the Statue of Liberty. There are also many interesting future applications of automatic time-stamping photos, such as modeling temporal behavior of people in the aggregate.

7 Acknowledgements

This work was supported in part by NSF grants IIS-0964027, IIS-1111534, IIS-1149393, by the Intel Science and Technology Center for Visual Computing, and by the Sloan Foundation. We are grateful to Daniel Carpenter for the ground truth CASTLE timestamps.

References

[1] Austin Abrams, Christopher Hawley, and Robert Pless. Heliometric stereo: Shape from sun position. In *ECCV*, 2012.

[2] J. Ackermann, F. Langguth, S. Fuhrmann, and M. Goesele. Photometric stereo for outdoor webcams. In *CVPR*, 2012.

[3] Sameer Agarwal, Noah Snavely, Ian Simon, Steven M. Seitz, and Richard Szeliski. Building Rome in a day. In *ICCV*, 2009.

[4] Yasutaka Furukawa and Jean Ponce. Accurate, dense, and robust multi-view stereopsis. In *CVPR*, 2007.

[5] Roger Grosse, Micah K. Johnson, Edward H. Adelson, and William T. Freeman. MIT Intrinsic Images, 2009. http://people.csail.mit.edu/rgrosse/intrinsic/.

[6] Tom Haber, Christian Fuchs, Philippe Bekaer, H-P Seidel, Michael Goesele, and Hendrik PA Lensch. Relighting objects from image collections. In *CVPR*, 2009.

[7] Daniel Hauagge, Scott Wehrwein, Kavita Bala, and Noah Snavely. Photometric ambient occlusion. *CVPR*, 2013.

[8] Lukas Hosek and Alexander Wilkie. An analytic model for full spectral sky-dome radiance. *ACM Transactions on Graphics*, 2012.

[9] Nathan Jacobs, Nathaniel Roman, and Robert Pless. Consistent temporal variations in many outdoor scenes. In *CVPR*, 2007.

[10] Wenzel Jakob. Mitsuba renderer, 2010. http://www.mitsuba-renderer.org.

[11] Pierre-Yves Laffont, Adrien Bousseau, Sylvain Paris, Frédo Durand, and George Drettakis. Coherent intrinsic images from photo collections. *SIGGRAPH Asia*, 2012. URL http://www-sop.inria.fr/reves/Basilic/2012/LBPDD12.

[12] Pierre-Yves Laffont, Adrien Bousseau, and George Drettakis. Rich intrinsic image decomposition of outdoor scenes from multiple views. *IEEE Trans. Visualization and Computer Graphics*, 2013. URL http://www-sop.inria.fr/reves/Basilic/2013/LBD13.

[13] Jean-François Lalonde, Alexei A. Efros, and Srinivasa G. Narasimhan. Estimating the natural illumination conditions from a single outdoor image. *International Journal of Computer Vision*, 2011.

[14] Peter Nillius and Jan-Olof Eklundh. Automatic estimation of the projected light source direction. In *CVPR*, 2001.

[15] NOAA. Quality controlled local climatological data. http://cdo.ncdc.noaa.gov/qclcd/QCLCD. 01/2013.

[16] Li Shen and Chuohao Yeo. Intrinsic images decomposition using a local and global sparse representation of reflectance. In *CVPR*, 2011.

[17] Kalyan Sunkavalli, Wojciech Matusik, Hanspeter Pfister, and Szymon Rusinkiewicz. Factored Time-Lapse Video. In *SIGGRAPH*, 2007.

[18] Kalyan Sunkavalli, Todd Zickler, and Hanspeter Pfister. Visibility Subspaces: Uncalibrated Photometric Stereo with Shadows. In *ECCV*, 2010.

[19] Y. Weiss. Deriving intrinsic images from image sequences. In *ICCV*, 2001.

[20] Lap-Fai Yu, Sai-Kit Yeung, Yu-Wing Tai, Demetri Terzopoulos, and Tony Chan. Towards Outdoor Photometric Stereo. In *ICCP*, 2013.