Pacific Graphics 2014
J. Keyser, Y. J. Kim, and P. Wonka
(Guest Editors)

Volume 33 (2014), Number 7

Time-Lapse Photometric Stereo and Applications

Fangyang Shen[1,2] Kalyan Sunkavalli[3] Nicolas Bonneel[1,4] Szymon Rusinkiewicz[5] Hanspeter Pfister[1] Xin Tong[6]

[1]School of Engineering and Applied Sciences, Harvard University
[2]State Key Laboratory of VR Technology and Systems, Beihang University
[3]Adobe Research
[4]CNRS
[5]Princeton University
[6]Microsoft Research

Abstract

This paper presents a technique to recover geometry from time-lapse sequences of outdoor scenes. We build upon photometric stereo techniques to recover approximate shadowing, shading and normal components allowing us to alter the material and normals of the scene. Previous work in analyzing such images has faced two fundamental difficulties: 1. the illumination in outdoor images consists of time-varying sunlight and skylight, and 2. the motion of the sun is restricted to a near-planar arc through the sky, making surface normal recovery unstable. We develop methods to estimate the reflection component due to skylight illumination. We also show that sunlight directions are usually non-planar, thus making surface normal recovery possible. This allows us to estimate approximate surface normals for outdoor scenes using a single day of data. We demonstrate the use of these surface normals for a number of image editing applications including reflectance, lighting, and normal editing.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Color, shading, shadowing, and texture I.4.8 [Image Processing and Computer Vision]: Scene Analysis—Photometry, Shading, Shape, Time-varying imagery

1. Introduction

The variation of scene appearance with illumination provides a strong indication of surface shape [Hor70], and is one of the few cues that is able to yield dense (per-pixel) estimates of geometry. Such fine-scale geometric detail, in the form of normal maps, is of great interest for applications ranging from rendering under novel illumination to editing materials or local geometry. In this paper, we focus on a practical and easily-deployed passive imaging configuration — time-lapse photography under daylight illumination — and demonstrate recovery of approximate per-pixel normal maps allowing us to provide geometry-aware image editing operations.

Outdoor time-lapse sequences present a compelling opportunity for visual analysis, because of their ease of acquisition and constrained variation. The latter is a key point: although the lighting varies in intensity and color throughout the day, we have domain-specific information about the range of possible variation. Specifically, we know that the illumination is due to a combination of low-spatial-frequency skylight and the point-like sun, which moves through the sky in a constrained arc. These constraints are sufficiently strong that Jacobs et al. have shown that even simple PCA is enough to identify local scene structure in many cases [JRP07].

Our goal, however, is normal recovery, and we build upon the techniques of *Photometric Stereo*. Photometric Stereo techniques estimate per-pixel surface normals by analyzing the variation in observed intensities under varying illumination. Extending these techniques to outdoor scenes is not trivial. The illumination in outdoor scenes is complex and uncontrolled; it consists of temporally-varying sunlight and skylight components that combine in different proportions at every scene point depending on both local shading effects and non-local effects like shadowing. In addition, Photometric Stereo techniques require that there be a wide variation in the lighting in the scene because this variation decides how well-conditioned the surface normal recovery is. Over the course of a single day, the sun moves in a planar arc across

the sky, and because of this restricted motion, previous work has deemed a single day of time-lapse data as inadequate for normal estimation.

In this work, we show that it is possible to recover surface normals for an outdoor scene from images captured over the course of a single day. This result is based on the observation that, while the sun does move in a near-planar arc over the course of a day, *this plane is offset from the center of the earth*. As a result, the directions of the rays from the sun over the day do not lie on a plane, making Photometric Stereo-based surface acquisition possible. Our method includes compensation for shadows and skylight, and we demonstrate both analysis of synthetic scenes and results on captured data. Finally, we show how the estimated normals can be used to realistically edit the material properties, lighting, and geometry of the scene.

2. Related Work

Photometric Stereo The original Photometric Stereo algorithms developed by Woodham and others relied on assumptions of Lambertian (perfect diffuse) shading and multiple light sources of known position and brightness [Woo78, Sil80]. In recent years, researchers have extended the fundamental methods to allow for uncalibrated lighting [Hay94, BJ01, BJK07], inter-reflections and shadows [CKK05, CAK07, SZP10], and non-Lambertian surfaces [NIK90, Geo03], including arbitrary spatially-varying materials [GCHS05]. The ultimate goal of this avenue of research is an explicit decomposition into scene properties, which completely explain the variation present in the image sequences by plausible geometry, (spatially-varying) materials, and (temporally-varying) illumination [Wei01, MNIS03]. Among the many approaches for stable estimation of (parts of) this decomposition are: estimation of illumination and materials, given known geometry [YM98, RZ10]; estimation of reflectance and geometry/shadow profiles, given known illumination [MLP04, ON12]; and clustering of regions of constant normal [KN06].

Outdoor scene analysis In spite of all the advances made in generalizing photometric stereo beyond the original assumptions, many of these methods are still restricted to indoor illumination conditions. One of the earliest attempts to extend these methods to outdoor scenes is the work of Sato and Ikeuchi [SI95] that decomposes outdoor time-lapse sequences into a skylight and sunlight terms and estimates normals from the sunlight term using constraints based on surface integrability. Later work has explored factoring temporal pixel profiles into scaled and offset combinations of basis curves for sun and sky [SMPR07], and analyzed the variation of sun and sky color throughout the day [SRM*08]. Shen and Tan [ST09] applied Photometric Stereo methods to estimate weather conditions from internet images. Ackerman et al. [ALFG12] presented one of the first works to explicitly recover surface normals and non-Lambertian scene reflectance from time-lapse data captured over many months. Abrams et al. [AHP12] presented an alternate technique that uses months of data to reconstruct Lambertian scenes. In subsequent work, they proposed using shadows to recover sparse depth from time-lapse sequences [AMP13, ASP14]. Yu et al. [YYT*13] use a small number of images captured in a single day along with light probes capturing the incident illumination to reconstruct surface normals.

In this work, we build upon the idea of sun/sky decomposition and recover a complete surface normal at every pixel of large-scale outdoor scenes from a single day's time-lapse date. This is in contrast with previous work, that relies on many months of data [ALFG12, AHP12], recovers only one component of the surface normal [SMPR07]), or assumes surface continuity [SI95].

3. Time-Lapse Photometric Stereo

The goal of this work is to estimate per-pixel normals from a sequence of images of a static scene captured over the course of a single day to allow for editing operations. We denote this image sequence by $I(x,t)$, where x denotes the pixel location and t the time. For an image sequence with m pixels and n frames, $I(x,t)$ represents an $m \times n$ matrix. We assume that the incident irradiance at any point in the scene is due to only two sources – direct illumination from the sun and highly diffuse illumination from the sky. Under this assumption the observed radiance at any pixel can be expressed as a sum of a sky component I^{sky} (i.e radiance due to illumination from the sky) and a sun component I^{sun} (i.e. radiance due to illumination from the sun) :

$$I(x,t) = I^{sky}(x,t) + S(x,t)I^{sun}(x,t), \qquad (1)$$

where S is a shadow function that indicates if a pixel x is illuminated by the sun at time t.

This representation indicates the difficulties in applying photometric stereo directly to outdoor time-lapse images. First, the observed radiance at every pixel results from complex illumination consisting of both directional and smooth low frequency lighting. Second, even if the sun component were to be isolated, applying photometric stereo requires that the images be captured under at least three non-planar light sources. It has been widely believed in the vision community that the sun moves in a planar arc across the sky, making surface normal estimation from images captured on a single day an under-constrained problem [Woo78, SI95, SRM*08].

In the following sections we show how we resolve each of these problems. Starting with the time-lapse sequence, we iteratively estimate the shadows and the sky component (Sec. 3.1), thereby isolating the sun component. We show that the sunlight directions over the course of a day are non-planar (Sec. 3.2), and use this result to recover per-pixel surface normals the sun component (Sec. 3.3). We finally use

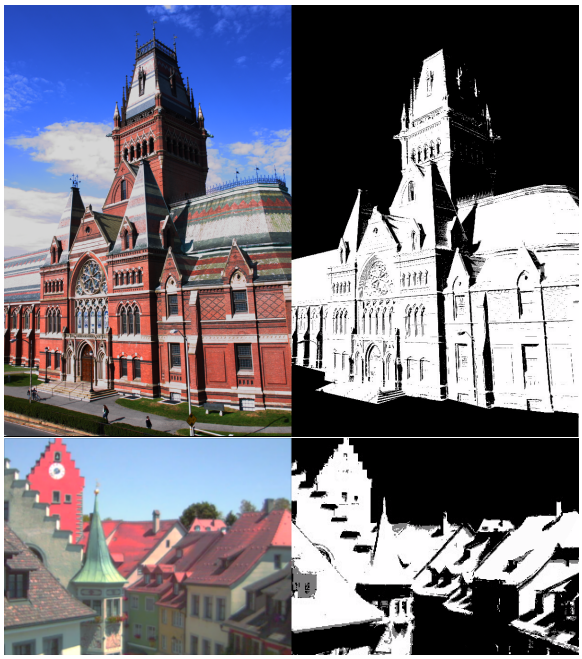


Figure 1: *Input frames from time-lapse sequences (left) and the corresponding recovered shadow masks (right).*

this information to alter material properties, illumination and geometry in an image (Sec. 4).

3.1. Estimating shadows, and sky component initialization

Illumination from the sky is highly diffuse and we assume that the sky component at any surface point depends primarily on the albedo and the amount of visible sky, while the surface normal is of secondary importance [LZ94]. We model the sky component at every pixel using a low-rank approximation

$$I^{sky}(x,t) = \sum_{i=1}^{N} N_i^{sky}(x) L_i^{sky}(t), \qquad (2)$$

i.e., we assume that the sky component is well approximated by N basis time-varying curve $L_i^{sky}(t)$ (that encodes the variation in the intensity of skylight) that is scaled by the per-pixel factor $N_i^{sky}(x)$ (that encodes the albedo, the ambient occlusion and the normal dependent effects at that surface point). We found that using $N = 2$ provides good accuracy while avoiding over-fitting.

Solving Eqn. 2 is equivalent to factorizing a matrix with missing data and we use Alternating Least Squares (ALS) to fit the model. However, to fit Eqn. 2 to only the shadowed pixels in the time-lapse data we need to know the shadow function S. Conversely, if we know N_i^{sky} and L_i^{sky}, we can recover S by reconstructing the sky component I^{sky} and comparing it to the time-lapse data. We leverage this dependence by using an interactive scheme to compute both I^{sky} and S together. We initialize S using a simple shadow detection approach proposed in previous work [SMPR07] and refine it using the following scheme:

1. Solve for N_i^{sky} and L_i^{sky} by factorizing I using pixels that are in shadow, with ALS.
2. Reconstruct the sky component I^{sky} using Eqn. 2.
3. Update the shadow function: a pixel is in shadow when $I(x,t) < k_1 I^{sky}(x,t)$ (i.e., $S(x,t) = 0$), in sunlight when $I(x,t) > k_2 I^{sky}(x,t)$ (i.e., $S(x,t) = 1$), and is set as unknown (i.e., $S(x,t) = 0.5$) otherwise. In practice, we found that $k_1 = 1.1$ and $k_2 = 1.6$ worked well.
4. Filter the shadow function, set $S(x,t_0)$ to be unknown if $0 < \sum_{t=t_0-k}^{t_0+k} S(x,t) < 2k+1$.

For each pixel, step 4 excludes the frames in which it may lie on shadow edges or in soft shadows, which break the assumptions of our shading model.

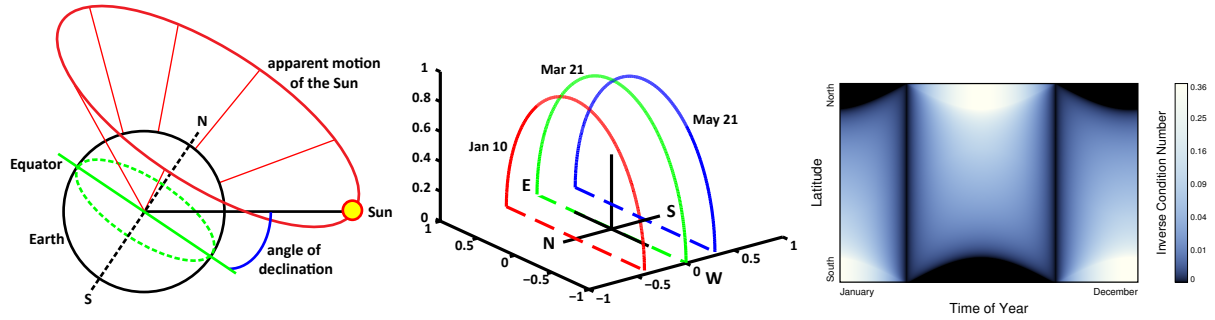
We repeat these steps for 3 iterations to get the final shadow and sky component estimates. As with any iterative solution, the quality of the final result depends on how accurate the shadow and L^{sky} initializations are. In particular, if the original shadow functions marks a large number of non-shadow pixels as shadowed, the estimates of N_i^{sky} and L_i^{sky} leading to grossly incorrect results. To avoid this, we first fit a rank-1 approximation of the skylight component, L_1^{sky} (i.e., $N = 1$ in Equation 2), to a set of K (typically $5-20$) user-selected pixels that are mostly in shadow. We then use L_1^{sky} to estimate N_1^{sky} at every pixel in the scene. This gives us a rank-1 fit to the sky component at every pixel that is a good initialize for the final rank-2 fitting. We then initialize $N_i^{sky} = N_1^{sky}$ for both components and L_i^{sky} to sine and cosine curves with the same frequency as the sun position, and use the iterative refinement technique described earlier to estimate the shadow masks and sky component.

3.2. Using the sun for Photometric Stereo

By removing the contribution of sky illumination we can estimate the appearance of the scene under directional lighting as $I^{sun} = \max(I - I^{sky}, 0)$. Under the Lambertian photometric stereo assumption, the sun component at pixel x at time t is given by

$$I^{sun}(x,t) = \rho(x)[\hat{N}^{sun}(x)]^T \hat{L}^{sun}(t) l(t), \qquad (3)$$

where $\rho(x)$ and $\hat{N}^{sun}(x)$ are the albedo and unit surface normal (a 3-vector) at pixel x and $L^{sun}(t)$ and $l(t)$ are the direction (another 3-vector) and intensity of the sun at time t. We can combine the albedo and surface normal into N^{sun} and the direction and intensity of the sun into L^{sun}. The normals and/or lighting directions can then be recovered by factorizing the $m \times n$ sun component matrix I^{sun} (formed by stacking



(a) (b) (c)

Figure 2: *(a) An earth-centric illustration of the motion of the sun. The axis of the earth's rotation is inclined at an angle of $23^\circ 27'$ to the plane of its orbit around the sun. This creates an angle of declination (shown in blue in (a)) between the N-S axis and the earth-sun line. As a result, the sun moves in a great plane (shown in red) that is offset from the equatorial plane, and the sunlight rays are non-planar. As the earth revolves around the sun, the angle of declination varies from a $\pm 23^\circ 27'$ during the solstices to 0° at the equinoxes. This leads to different solar planes as illustrated in (b), where we have plotted the position of sun at $0^\circ N$, $0^\circ E$ for three different days. The planes are different and the plane on Mar. 21 (the vernal equinox) passes through the origin, making all sunlight directions co-planar. While the sunlight directions are non-planar during all other times of the year, the stability of the photometric stereo problem depends on the conditioning of the sunlight matrix. This is illustrated in (c), that shows a visualization of the inverse of the condition number (ratio of minimum to maximum eigenvalues) of the light direction matrix, for varying latitude and date. The matrix is singular in the polar regions during winter (since the sun is not visible) and worldwide on the vernal and autumnal equinoxes (since the path of the sun is planar on those dates). Otherwise, the matrix is nonsingular, indicating that normal estimation is possible.*

per-pixel time profiles) into the $m \times 3$ normal matrix N^{sun} and the $3 \times n$ lighting matrix L^{sun}:

$$I^{sun} = N^{sun} L^{sun}. \tag{4}$$

Previous work [Woo78, Hay94] has shown that this factorization can be solved using singular value decomposition as long as the image data matrix, I^{sun} is rank-3, i.e., there are at least three non-planar normals in the scene and the scene has been imaged under at least three non-planar lighting directions.

The motion of the sun over the course of a single day across the sky at any location on the earth is well approximated by a planar arc. This has lead to the belief that surface normals cannot be unambiguously recovered from a time-lapse sequence [Woo78, SI95, SRM*08]. The key insight in this paper is that *while the sun does move in a plane in the sky, in general, this plane is offset from the center of the earth*. The earth rotates about an axis that is inclined at an angle of $23^\circ 27'$ to the plane of its orbit around the sun. This angle ensures that the declination of the earth – i.e., the angle between the plane of the earth's equator and the rays of the sun – varies and goes to zero only twice a year (during the equinoxes). As illustrated in Fig. 2(a), the angle of declination creates an offset between the solar plane and the center of earth. This ensures that *in spite of the planar motion of the sun, the directions of the rays from the sun over the day do not lie on a plane*, but rather on the lateral surface of the cone whose apex lies at the center of the earth and whose base is described by the solar plane.

The sun's position over the course of a day at any location on the earth can be described by the time-varying 3-vector:

$$L^{sun}(t) = [\sin\theta\cos(\omega t - \phi_0), \sin\theta\sin(\omega t - \phi_0), \cos\theta] \tag{5}$$

where ω is the frequency of the sun's motion, θ depend on the location of the scene and the angle of the earth's declination, and ϕ_0 is an offset for the local date and time of observations. While the x and y components of these direction vectors define a plane, the z-component is the offset of this plane from the origin. In general, these direction vectors form a rank-3 matrix. However, for the special case of $\theta = 90^\circ$ (as observed during the spring and autumn equinox when the angle of declination becomes 0°), the offset equals 0 and the rank of the light matrix reduces to 2. For this special case, the radiance for a Lambertian surface illuminated by sunlight reduces to the model used in previous work [SI95, SRM*08] and the recovery of surface normals from a single day time-lapse sequence is ambiguous.

Even if the light source matrix L^{sun} is rank-3, robust surface normal recovery also requires that it be well-conditioned. To explore how well-behaved L^{sun} is, we estimated L^{sun} for every latitude on the earth and every day in the year and computed the ratio of the third and first eigen values of its covariance matrix. Large ratios indicate that the photometric stereo problem is well-conditioned, while ratios close to zero indicate a degenerate rank-2 system. As illus-

trated in Fig. 2(c), L^{sun} is well-conditioned over many parts of the earth for many days of the year.

3.3. Estimating normals from the sun component

Since the light sources are non-planar it follows that the sun component I^{sun} is a rank-3 matrix and we can solve Eqn. 4 for the 3D normal and light source directions. If the position of the light sources is unknown this is a problem of uncalibration photometric stereo [Hay94] and the surface normals and light positions can be recovered only upto an unknown linear transform. In our case, because the sun follows a highly constrained path across the sky, it acts as a semi-calibrated light source whose the lighting positions can be easily estimated if the location and date and time of capture are known [Ast09].

Using the location (latitude/longitude) of capture and the timestamps of each frame, we set the light source directions L^{sun} to the true sunlight directions in the celestial coordinate system. Using these light source directions $\hat{L}^{sun}$, we recover the surface normals $\hat{N}^{sun}$ by using Alternating Least Squares to solve Eqn. 3. In our optimization we update the surface normals N^{sun} and the magnitude of the light directions $||L^{sun}||$. This allows us to fit the changes in the light source intensity while ensuring that the basis curves are not biased towards any set of dominant normals in the scene. In practice, we initialize the sunlight intensities by performing this optimization at a set of K user-selected ground pixels; since ground pixels face up, this gives us good initializations for the sunlight intensities and helps the optimization converge better.

3.4. Refining all the components

Estimating the shadows, sky component, and sun component, as described above, gives us good estimates of the scene normals. In practice, we have found that we can improve the estimated normals further by repeating the entire process to refine each component. Given the estimated sky and sun component, we re-estimate the shadow masks by using the heuristic that if the observed intensity is closer to the reconstructed sky component, the pixel is probably in shadow, and if it is closer to the sum of the reconstructed sky and sun components, it is probably lit. We set $S(x,t) = 0$ if $I(x,t) < I^{sky}(x,t) + \frac{1}{3}(I^{sun}(x,t) - I^{sky}(x,t))$, $S(x,t) = 1$ when $I(x,t) > I^{sky}(x,t) + \frac{2}{3}(I^{sun}(x,t) - I^{sky}(x,t))$, and set it to unknown otherwise. We also filter the shadow function by setting

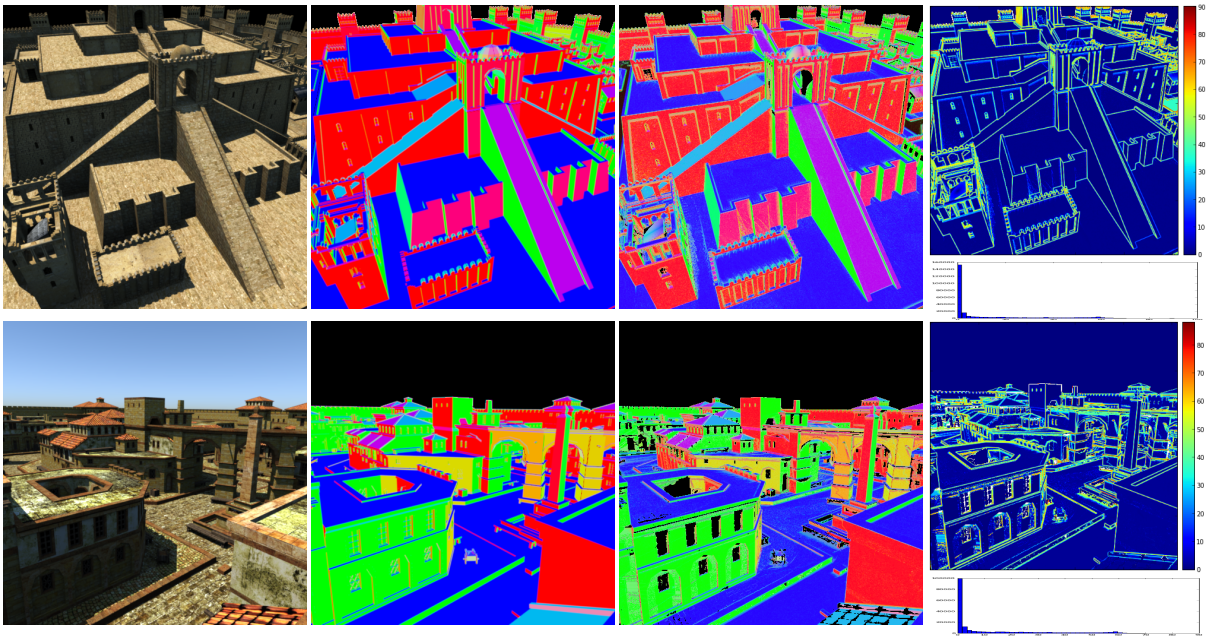
4. Results and Discussion

Synthetic data We tested the proposed algorithm on synthetic time-lapse sequences rendered with Mitsuba to obtain a full global illumination solution. Our sequences have 55 frames each and use the Hosek et al. daylight model [HW12] to simulate the time-varying color and intensity of sunlight and skylight in Tokyo, Japan (35.6895°N, 139.6917°E) on June 20, 2012. Fig. 3 shows one of these frames and compares the ground truth normal map and the estimated normal map for these scenes, and visualizes the angular error per-pixel. The median angular error in our estimates was 1.36° and 1.24° for these two scenes. While the recovered normals are very faithful to the the ground truth, there are two main causes of error in this result. First, we cannot estimate the normals reliably at pixels that were lit by the sun for very short periods of time and at pixels that were almost never in shadow. The second source of error results from our rank-2 approximation of the sky component. Analytical models for skylight distribution [PSM93] show that, even though sky illumination is largely diffuse, it tends to be brightest in the direction of the sun. This results in a slight normal dependence in the sky component that is fully not captured by our rank-2 approximation and influences the accuracy of the normal estimation.

Captured data To test our algorithm on real data, we captured two Lambertian objects (a cylinder and a mouth shaped figure) in Cambridge, USA (42.37°N, 71.11°W) on Nov 9, 2010. Fig. 4 shows four example frames from the original sequence and reconstructed images. Note that the reconstructed image preserve the shading in the original images very accurately. The fact that we are able to capture the smooth variation in the surface normals on the cylinder indicate that the light source directions are sufficiently non-planar and well-conditioned. As in all our results, our estimated normals were in the celestial coordinate framework (in which we have the light positions). For this data, we also captured the ground truth geometry using LIDAR scans, and computed a transformation that best aligned the estimated normals to the ground truth normals, in order to match the coordinate frames. As can be seen from the result, our transformed normals compare very favorably with the ground truth normals, even though they were computed from images captured under uncontrolled outdoor illumination.

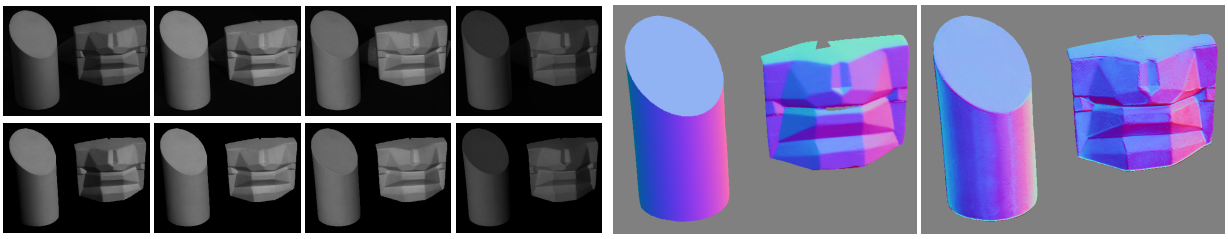
We also tested our algorithm on two real world time-lapse sequences downloaded from online webcam archives. The first one is a time-lapse sequence from the AMOS webcam archive [JRP07] (`http://amos.cse.wustl.edu/camera?id=10870`) captured at Meersburg, Germany (47.69°N, 9.27°E) on June 27, 2011. This sequence was also analyzed by Abrams et al. [AHP12], and we linearized the pixel intensities using their estimated radiometric response curve. In Fig. 5, we compare our results from a single day of data to their normals estimated from a few months of time-lapse images. While our result is noisy in regions that are in shadow for most of the day, qualitatively, it compares favorably with their results computed from 2-4 months of data.

The second sequence was captured at Trutnov, Czech



(a) Rendered frames (b) Ground truth normals (c) Estimated normals (d) Error distribution

Figure 3: *Normal recovery on synthetic datasets. We evaluate our method on two synthetic time-lapse sequences (a) of large-scale scenes with a large amount of variation in the surface normals (b). The surface normals estimated by our method (c) are very close to the ground truth normals at most pixels (d) (as is clear from the histogram distribution of the errors at the bottom). Most of the errors are concentrated in regions that are in shadows for most of the frames, and edges.*

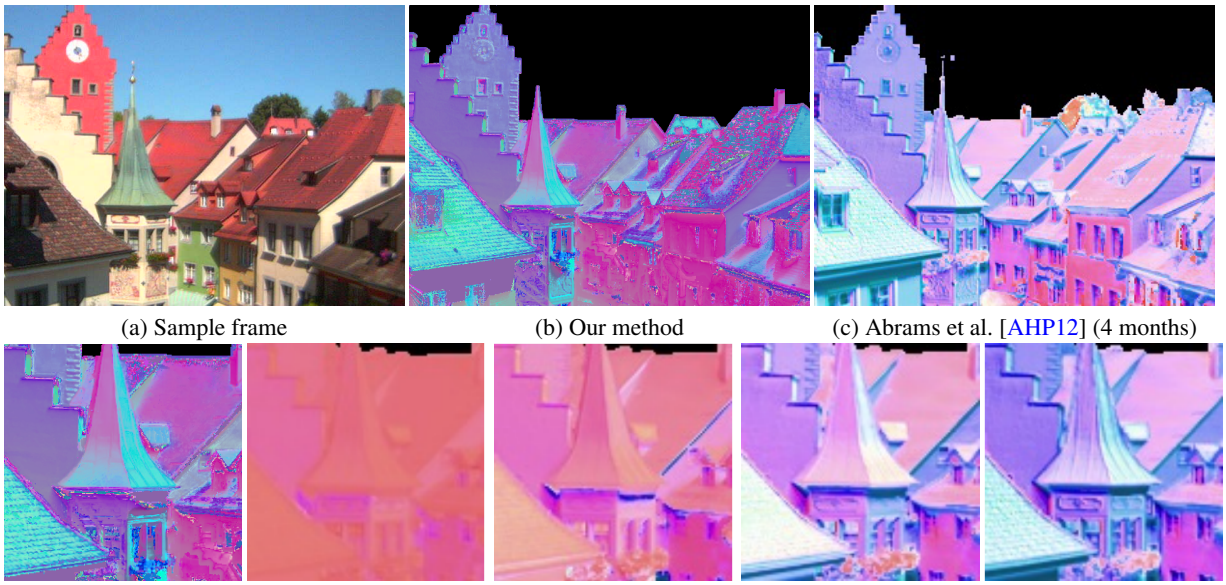


(a) Input (top) and Reconstructed (bottom) images (b) Ground truth normals (c) Estimated normals

Figure 4: *Normal recovery on a simple captured dataset. The left panel shows the four frames from the captured time-lapse sequence (top) along with the reconstructed images from our analysis (bottom). The right panel shows the ground truth surface normals (b) and the per-pixel normals estimated using our technique (c).*

Republic (50.56°N, 15.91°E) on June 24, 2011 (http://kamery.humlnet.cz/en/kamery/). Fig. 6 shows one shadow frame, and the computed surface normals. Since we did not have access to the radiometric calibration for this camera, we assumed a gamma of 2.2 to linearize this data.

Applications Separating a time-lapse sequence into normals, shading, shadowing and albedo enables a number of applications that manipulate the appearance of these time-lapse images. We demonstrate a set of these applications on two time-lapse sequences – one that we captured ourselves in Cambridge, USA on October 28, 2010 (Fig. 7), and the other from the work of Abrams et al. [AHP12] (Fig. 9). We computed the shadows, sky components, sun components, and normals for these sequences using our technique; Once estimated, we can directly edit each of these components for different applications. In particular, it becomes easy to simulate clear or overcast skies by changing the ratio of sunlight to skylight (Fig. 7), to change the albedo component while accounting for shading variations (Fig. 8), or to alter the apparent surface normals in the scene via bump-mapping techniques that take into account the existing orientation of surfaces (Figs. 8 and 9). In all these results, we masked the sky out while computing the normals and editing the images; the sky is reproduced from the original image.



(a) Sample frame (b) Our method (c) Abrams et al. [AHP12] (4 months)

(d) Our results (e) [AHP12] (1 month) (f) [AHP12] (1.5 months) (g) [AHP12] (2 months) (h) [AHP12] (4 months)

Figure 5: *We compare our results computed from a single day's data (b) with those reported by Abrams et al. [AHP12] from many months of data (c) on a dataset from the AMOS time-lapse archive. The detail images in the second row show close-ups of the results. While our results (d) are noisier in regions that are in/out of shadows in very few frames (for e.g., the red roofs), qualitatively they capture as much detail as the results from Abrams et al. [AHP12] that computed from 2-4 months of data (g/h). To make this comparison, we show our estimated normals using the same color map defined in their paper.*

(a) Sample frame (b) Estimated shadow mask (c) Estimated normals

Figure 6: *Normal estimation on a webcam sequence. From this time-lapse sequence captured in Trutnov (a), we are also to estimate shadow masks (b), and per-pixel surface normals (c).*

Limitations The quality of the normal reconstruction from our technique is closely related to the variation in the lighting in the input time-lapse sequence. In particular, our method depends on a sufficient number of observations of the scene both in and out of shadow. As a result, we are not able to reliably reconstruct surface normals at pixels that are almost always lit by the sun (because we can not disambiguate the contribution of the the sky illumination to the intensity), or are almost always in shadow (because the variation in the sun position at these pixels is not sufficient to make normal recovery robust). This is reflected in the erroneous normals estimated at these pixels in our results.

Our method assumes that the scene is static and Lambertian, and deviations from these assumptions (for e.g., specular reflections from windows, moving people and cars, etc.) introduce errors to the profiles of these pixels. Our shadow refinement algorithm often classifies these outliers as uncertain, therefore not using them in curve fitting, but they can corrupt the normal estimates we get. More robust algorithms that explicitly model and handle these sources of noise can greatly expand the range of applicable scenes.

(a) Input frame (b) Edited frames

Figure 7: *Illumination editing. Once sunlight/skylight components are obtained from a time-lapse sequence (a), we can alter their ratio and simulate a clearer sky (b).*

(a) Input frames

(b) Edited frames

Figure 8: *Albedo and geometry editing. By recombining the shading component with a new albedo, we can edit the albedo while preserving consistent lighting. Here, we added a "'PG14"' logo on the monument. In addition, by combining the estimated normals with a bump map, we can alter the perception of the geometry. We reuse the same logo to do this.*

Finally, our image formation model assumes a binary shadow mask, and the shadow refinement stage discards pixels that have soft shadowing from the fitting. While it is a challenging problem, the ability to handle soft shadows would be an interesting extension of our technique. In general, shadows contain useful information about scene geometry [ASP14], and leveraging this information could make our technique more robust.

5. Conclusions

In this paper we have shown that Photometric Stereo can be extended to the problem of estimating per-pixel surface normals from single-day time-lapse sequences captured under daylight illumination. Our work is built on the insight that the directions of sunlight over a single day are in fact non-planar, allowing us to estimate approximate per-pixel normals for large scale outdoor scenes. We validated our method on both synthetic and captured data, and showed that the estimated normals can be used for a variety of image editing tasks.

In the future we would like to explore more accurate skylight models to improve the accuracy of the estimated normals. Also, while we have not applied any additional constraints in our solution, an interesting avenue of future work is to combine this method with additional information such as sparse reconstructions from multi-view stereo to produce accurate high resolution surface geometry.

Acknowledgements

We would like to thank Austin Abrams for help with the Meersburg time-lapse sequence, and the comparisons in Fig. 5. This work was partially supported by NSF grant IIS 1110955, China National NSF grant 61073078, 61272348, 61202235, Ph.D. Program Foundation of Ministry of Education of China (No. 20111102110018), National Key Technology R&D Program of China (2014BAK18B01) and the China Scholarship Council.

References

[AHP12] ABRAMS A., HAWLEY C., PLESS R.: Heliometric stereo: Shape from sun position. In *Proc. ECCV* (2012). 2, 5, 6, 7

[ALFG12] ACKERMANN J., LANGGUTH F., FUHRMANN S., GOESELE M.: Photometric stereo for outdoor webcams. In *Proc. CVPR* (2012). 2

[AMP13] ABRAMS A., MISKELL K., PLESS R.: The episolar constraint: Monocular shape from shadow correspondence. In *Proc. CVPR* (2013). 2

[ASP14] ABRAMS A., SCHILLEBEECKX I., PLESS R.: Structure from shadow motion. In *Proc. ICCP* (2014). 2, 8

[Ast09] ASTRONOMICAL APPLICATIONS DEPARTMENT OF THE U.S. NAVAL OBSERVATORY: Sun or moon altitude/azimuth table, 2009. http://aa.usno.navy.mil/data/docs/AltAz.php. 5

[BJ01] BASRI R., JACOBS D.: Lambertian reflectance and linear subspaces. In *Proc. ICCV* (2001), vol. 2, pp. 383–390. 2

(a) Input frames (b) Edited frames

Figure 9: *Geometry editing. We can alter the perception of surface shape by using a bump map in conjunction with our reconstructed normals. Reconstructing the images with these altered normals results in images that properly account for lighting changes including cast shadows (third row).*

[BJK07] BASRI R., JACOBS D., KEMELMACHER I.: Photometric stereo with general, unknown lighting. *Int. Journal of Computer Vision 72*, 3 (2007), 239–257. 2

[CAK07] CHANDRAKER M., AGARWAL S., , KRIEGMAN D.: Shadowcuts: Photometric stereo with shadows. In *Proc. CVPR* (2007). 2

[CKK05] CHANDRAKER M., KAHL F., KRIEGMAN D.: Reflections on the generalized bas-relief ambiguity. In *Proc. CVPR* (2005). 2

[GCHS05] GOLDMAN D. B., CURLESS B., HERTZMANN A., SEITZ S.: Shape and spatially varying BRDFs from photometric stereo. In *Proc. ICCV* (2005). 2

[Geo03] GEORGHIADES A.: Incorporating the Torrance and Sparrow model of reflectance in uncalibrated photometric stereo. In *Proc. ICCV* (2003), vol. 1, pp. 816–823. 2

[Hay94] HAYAKAWA H.: Photometric stereo under a light source with arbitrary motion. *J. Opt Soc. Am. 11*, 11 (1994). 2, 4, 5

[Hor70] HORN B. K. P.: *Shape from Shading: A Method for Obtaining the Shape of a Smooth Opaque Object from One View*. PhD thesis, Massachusetts Institute of Technology, 1970. 1

[HW12] HOSEK L., WILKIE A.: An analytic model for full spectral sky-dome radiance. *ACM Trans. Graphics (Proc. ACM SIGGRAPH) 31*, 4 (July 2012). To appear. 5

[JRP07] JACOBS N., ROMAN N., PLESS R.: Consistent temporal variations in many outdoor scenes. In *Proc. CVPR* (2007), pp. 1–6. 1, 5

[KN06] KOPPAL S. J., NARASIMHAN S. G.: Clustering appearance for scene analysis. In *Proc. CVPR* (June 2006), vol. 2, pp. 1323 – 1330. 2

[LZ94] LANGER M., ZUCKER S. W.: Shape-from-shading on a cloudy day. *J. Opt Soc. Am.* (1994), 467–478. 3

[MLP04] MATUSIK W., LOPER M., PFISTER H.: Progressively-refined reflectance functions from natural illumination. In *Rendering Techniques* (2004), pp. 299–308. 2

[MNIS03] MATSUSHITA Y., NISHINO K., IKEUCHI K., SAKAUCHI M.: Illumination normalization with time-dependent intrinsic images for video surveillance. In *Proc. CVPR* (2003), pp. 3–10. 2

[NIK90] NAYAR S. K., IKEUCHI K., KANADE T.: Determining shape and reflectance of hybrid surfaces by photometric sampling. *IEEE Trans. Robotics and Automation 6*, 4 (1990), 418–431. 2

[ON12] OXHOLM G., NISHINO K.: Shape and reflectance from natural illumination. In *Proc. ECCV* (2012). 2

[PSM93] PEREZ R., SEALS R., MICHALSKY J.: All-weather model for sky luminance distribution - preliminary configuration and validation. *Solar Energy 50*, 3 (March 1993), 235–245. 5

[RZ10] ROMEIRO F., ZICKLER T.: Blind reflectometry. In *Proc. ECCV* (2010). 2

[SI95] SATO Y., IKEUCHI K.: Reflectance analysis under solar illumination. In *IEEE Workshop on Physics-Based Modeling and Computer Vision* (June 1995), pp. 180–187. 2, 4

[Sil80] SILVER W.: *Determining Shape and Reflectance Using Multiple Images*. Master's thesis, MIT, 1980. 2

[SMPR07] SUNKAVALLI K., MATUSIK W., PFISTER H., RUSINKIEWICZ S.: Factored time-lapse video. *ACM Trans. Graphics (Proc. ACM SIGGRAPH) 26*, 3 (July 2007), 101:1–101:10. 2, 3

[SRM*08] SUNKAVALLI K., ROMEIRO F., MATUSIK W., ZICKLER T., PFISTER H.: What do color changes reveal about an outdoor scene? In *Proc. CVPR* (2008). 2, 4

[ST09] SHEN L., TAN P.: Photometric stereo and weather estimation using internet images. In *Proc. CVPR* (2009). 2

[SZP10] SUNKAVALLI K., ZICKLER T., PFISTER H.: Visibility subspaces: Uncalibrated photometric stereo with shadows. In *Proc. ECCV* (2010), pp. 251–264. 2

[Wei01] WEISS Y.: Deriving intrinsic images from image sequences. In *Proc. ICCV* (2001), pp. II: 68–75. 2

[Woo78] WOODHAM R.: Photometric stereo: A reflectance map technique for determining surface orientation from image intesity. In *Proc. SPIE* (1978), vol. 155, pp. 136–143. 2, 4

[YM98] YU Y., MALIK J.: Recovering photometric properties of architectural scenes from photographs. In *Proc. ACM SIGGRAPH* (1998), pp. 215–224. 2

[YYT*13] YU L.-F., YEUNG S.-K., TAI Y.-W., TERZOPOULOS D., CHAN T. F.: Outdoor photometric stereo. In *Proc. ICCP* (2013). 2