



One-Day Outdoor Photometric Stereo Using Skylight Estimation

Jiyoung Jung¹ · Joon-Young Lee² · In So Kweon³

Received: 14 October 2017 / Accepted: 21 December 2018

© Springer Science+Business Media, LLC, part of Springer Nature 2018, corrected publication 2019

Abstract

We present an outdoor photometric stereo method using images captured in a single day. We simulate a sky hemisphere for each image according to its GPS and timestamp and parameterize the obtained sky hemisphere into quadratic skylight and Gaussian sunlight distributions. Our previous work recovered an outdoor scene on a clear day, whereas the current paper shows that cloudy days can provide better illumination conditions for surface orientation recovery, and hence we propose a modified sky model to represent a well-conditioned skylight distribution for outdoor photometric stereo. The proposed method models natural illumination according to a sky model, providing sufficient constraints for shape reconstruction from 1-day images. We tested the proposed method to recover various sized objects and scenes from real-world outdoor daylight images and verified the method using synthetic and real data experiments.

Keywords Photometric stereo · Light modeling · Outdoor illumination · Scene reconstruction · Skylight estimation

1 Introduction

Three-dimensional (3D) scene reconstruction from images has been an ongoing computer vision research topic. Several 3D reconstruction methods have been proposed using more than two images, such as multi-view stereo using images from different viewpoints (Furukawa and Ponce 2010), photometric stereo using images under different light directions (Woodham 1980), depth from focus (Grossmann 1987) and defocus (Subbarao and Surya 1994) using images with different focal settings, etc. Many other successful approaches have used a single image, such as shape from shading (Horn 1989), with appropriate assumptions and constraints.

Related experiments were initially conducted under laboratory conditions, where equipment and environment could be controlled, but the methods were quickly improved to overcome various difficulties due to uncontrolled elements. For example, multi-view stereo evolved from using images captured by uniformly placed cameras in the laboratory (Okutomi and Kanade 1993; Yoon and Kweon 2006) to searching and downloading scene images from the web (Agarwal et al. 2011). The results are promising even when input images were captured with different cameras at different times, viewpoints, and many other conditions.

Compared to multi-view stereo, which reconstructs the scene based on feature matching, photometric stereo has the advantage of recovering surface details up to the image resolution level, since the method is based on reflected radiance determined by the incoming light direction and the surface normal of the object of interest. Photometric stereo methods originally required several specific experimental conditions, but these have been gradually alleviated.

Traditional photometric stereo methods require three or more input images under different distant point light sources with non-planar directions on the unit sphere (Woodham 1980). Approaches such as uncalibrated photometric stereo (Lu et al. 2013) and self-calibrating photometric stereo (Shi et al. 2010) deal with unknown light directions; and other methods handle less than three light directions or use ambient light (Zhang et al. 2012). In particular, natural illumination has an important role for photometric stereo

✉ Jiyoung Jung
jiyoung.jung@khu.ac.kr

Joon-Young Lee
jolee@adobe.com

In So Kweon
iskweon77@kaist.ac.kr

¹ Department of Software Convergence, Kyung Hee University, Yongin, Republic of Korea

² Adobe Research, San Jose, CA, USA

³ School of Electrical Engineering, KAIST, Daejeon, Republic of Korea

outside controlled environments (Huang and Smith 2011; Johnson and Adelson 2011). Reflectance and object shape can be recovered from the illumination map with complex light information around the object (Oxholm and Nishino 2012) and illumination maps and reflectance can be similarly recovered using objects with known shape (Lombardi and Nishino 2012). Methods combining shape-from-shading with depth sensors have recently been shown to mitigate many assumptions with promising results (Yu et al. 2013; Han et al. 2013; Jung et al. 2015a).

Outdoor environments are generally regarded as including significant unknowns and complexities. An open field appears significantly different depending on weather conditions and time of day. However, it retains a generally similar appearance, and although room appearances can be greatly influenced by lighting conditions, outdoor fields on clear days present relatively predictable scenes.

This paper extends our previous study (Jung et al. 2015b; Jung 2016) that presented an outdoor photometric stereo method using skylight estimation (Fig. 1). Since outdoor illumination is mainly contributed by the sun and clear sky and can be generally modeled, we estimate the skylight illumination map using a sky simulation model (Preetham et al. 1999) and process geo-tagged time-stamped images captured from a static camera in a single day to estimate the scene surface normal without requiring depth priors. We overcome the weak rank-3 qualification of sunlight directions for a single day by exploiting natural illumination. Since we use a handful of images, various pixels will be sunlit in less than two images, and the incomplete surface normal estimate for these pixels are refined using information from neighboring pixels with similar profiles through MRF optimization.

Our previous study was limited to recover an outdoor scene on a clear day, whereas the current study analyzes general lighting conditions for outdoor photometric stereo. We show that cloudy days can provide better illumination conditions for surface orientation recovery and propose a modified sky model to represent a well-conditioned skylight distribution for outdoor photometric stereo. Our modified sky model requires two additional scaling factors: maximum sunlight and skylight intensities, which can be estimated directly from the real sky images if available, or by direct user assistance otherwise. The modified sky model was validated by the synthetic and real data experiments.

2 Related Works

There have been several recent approaches for outdoor photometric stereo. Ackermann et al. (2012) and Abrams et al. (2012) used time lapse sequences captured by static outdoor webcams. Since both approaches modeled illumination as a distant point light (i.e., sun) with constant ambient light,

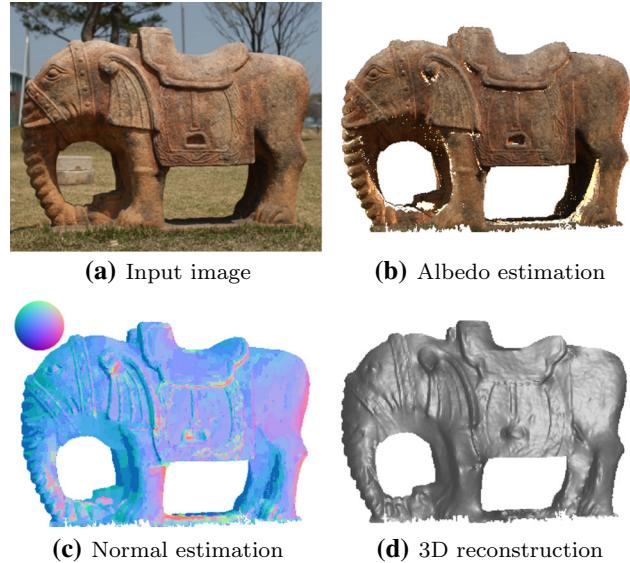


Fig. 1 An example of 1-day outdoor photometric stereo result. Fifteen input images were captured at 30-min intervals. The surface is reconstructed using Poisson solver (Simchony et al. 1990)

they required images over several months to avoid coplanar light directions in photometric stereo. They also excluded bad weather and night images and processed photometric stereo using hundreds of selected images. Abrams et al. (2012) proposed an iterative nonlinear optimization of ambient light, shadows, light color, surface normals, radiometric calibration, and exposures. However, their approach required laborious optimization using 500 carefully selected images. Ackermann et al. (2012) also proposed an image filtering and selection process to find 50 proper images from 20,000 candidate images.

Shan et al. (2013) proposed large scale image reconstruction and relighting using hundreds of thousands of images through structure-from-motion, multi-view stereo, and surface reconstruction. They first identified cloudy images to estimate albedo and ambient light, and then sunny images to estimate light including a distant point light per image and refined the scene surface normal.

A point light source with constant ambient light is an effective way to model outdoor environments using a large number of observations or depth priors. However, we considered natural illumination complexity as an aid rather than hindrance to general shape-from-shading problems (Johnson and Adelson 2011). Since outdoor scenes are covered with sky by definition, various studies have utilized the sky as an information source directly.

Sunkavalli et al. (2008) assumed that the subspace containing daylight spectra was two-dimensional and proposed a color model for outdoor image sequence temporal color changes. They explored the use of the model to analyze outdoor scenes from time lapse images. Lalonde et al. (2010)

recovered camera parameters, including its focal length and pose, from the sun position or portion of clear sky in the image with capture time and location. They also estimated natural illumination conditions from a single image using a sky model and inserted virtual objects into the image such that they appeared natural under the estimated illumination (Lalonde et al. 2012). Kawakami et al. (2013) compared the sky image with the sky model to estimate camera spectral sensitivity and white balance. Inose et al. (2013) refined multi-view stereo result for outdoor scenes using 1-day images for photometric stereo. They incorporated the sky model to separate sun and sky illumination effects on the surface using the relatively precise surface normal from multi-view stereo to estimate lighting conditions.

Hold-Geoffroy et al. (2017) proposed a CNN based technique to estimate high dynamic range (HDR) outdoor illumination from a single low dynamic range image. They leveraged a large outdoor panorama dataset to train the CNN, and then fitted a Hošek–Wilkie sky model (Hošek and Wilkie 2013) to obtain parameter set. They estimated the illumination parameters and used them to naturally render a synthetic object in the image rather than recover the scene.

Shen et al. (2014) analyzed photometric stereo stability using 1-day images. The angle of declination varies from 0° to 23° as the earth revolves around the sun, producing non-planar sunlight directions for certain days of the year. They calculated the inverse of the condition number (i.e., the ratio of minimum to maximum eigenvalues) for the light direction matrix per day for varying latitude and date. For limited cases, the light direction matrix satisfies rank-3 constraint, but with a small inverse condition number. Their analysis identified the limitation of point light source modeling for 1-day outdoor photometric stereo and hence we seek a natural illumination solution using skylight estimation.

3 Image Formation

Imaged appearance, $I(x)$, for a given pixel, x , depends on its surface orientation, \mathbf{n}_x , which determines the hemisphere of incoming light. Incident irradiance, \mathbf{E} , is calculated as

$$\mathbf{E}(\mathbf{n}_x, \mathbf{L}) = \int L(\omega_i) \max(0, \mathbf{n}_x \cdot \omega_i) d\omega_i, \quad (1)$$

where ω_i is the light incident vector to the surface and $L(\omega_i)$ is the illumination intensity at ω_i (i.e., $\mathbf{L} = L(\omega_i)\omega_i$) (Ramamoorthi and Hanrahan 2001; Oxholm and Nishino 2012). Assuming Lambertian reflectance, the albedo, ρ , is invariant to the viewing direction. Therefore, we model the pixel intensity for each color channel, c , to be proportional to \mathbf{E} on the surface,

$$I_c(x) = \rho_c(x) \mathbf{E}(\mathbf{n}_x, \mathbf{L}). \quad (2)$$

3.1 Skylight Estimation

We estimate the outdoor illumination of each incident angle of the sky hemisphere, $L(\omega_i)$, using the sky model proposed by Preetham et al. (1999), which describes the correlation between the sky brightness distribution and turbidity values. They suggested an approximation of the complex distribution coefficients to be a linear function of the single turbidity parameter, T . Intuitively, turbidity encodes the amount of scattering in the atmosphere, so lower T implies clearer sky (Lalonde et al. 2010). For example, the CIE standard clear sky light distribution can be approximated by applying $T = 2.2$ to Preetham's model.

Using the simulated sky hemisphere as an illumination map, we calculated \mathbf{E} from Eq. (1) for uniformly sampled orientations.

3.2 Shadow Handling

Shadows are crucial components for image formation, causing significant appearance changes in outdoor scenes. Several studies have investigated outdoor image formation models (Abrams et al. 2012; Ackermann et al. 2012; Sunkavalli et al. 2007), all explicitly modeling shadows. Each study estimated the shadow-or-not binary classification problem: given many images of a static scene under varying illumination, which pixels are directly illuminated at which times?

For some pixel x in an image captured at time t , the shadow mask can be expressed as

$$S(x, t) = \begin{cases} 0 & \text{if } x \text{ is in shadow in image } t, \\ 1 & \text{otherwise.} \end{cases} \quad (3)$$

We detect the shadowed pixels explicitly using Abrams et al. (2013a) shadow estimation method, an expectation maximization approach that simultaneously estimates shadows, albedo, surface normals, and ambient light. In contrast to previous shadow detection procedures, this method uses sunlight directions and solves for shadows that are most consistent with the Lambertian assumption.

A point in the scene is illuminated by all rays coming from its upper hemisphere determined by its surface orientation. The point is in shadow when its hemisphere does not contain the sun (attached shadow) or if the sun is occluded (cast shadow). The resulting shadowed pixel is illuminated by the portion of sky without the sun.

To model shadowed and non-shadowed pixels, we separate the simulated sky hemisphere into dominant sunlight, \mathbf{L}_{sun} , and diffuse skylight, \mathbf{L}_{sky} , a popular strategy for shadow modeling in several previous outdoor image formation studies (Abrams et al. 2012; Ackermann et al. 2012;

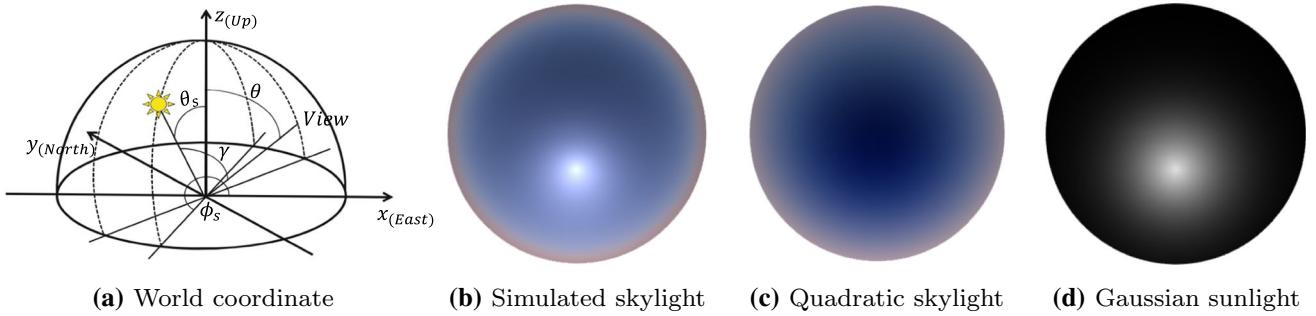


Fig. 2 Sky hemisphere simulation and parameterization. **a** World coordinate specifying the sun position and the viewing direction in the sky hemisphere. **b** A sky hemisphere is simulated using a sky

model (Preetham et al. 1999) and it is parameterized into **c** quadratic skylight distribution and **d** Gaussian sunlight distribution

Sunkavalli et al. 2007). This is a reasonable strategy because a shadowed pixel is the result of the sun occluded from directly lighting the position with some intensity maintained due to ambient illumination. Since it is difficult to distinguish if the shadow is attached or cast by its appearance, we model every detected shadowed pixel to be illuminated by the hemisphere determined by its surface normal excluding sunlight,

$$I_c(x) = \rho_c(x)(\mathbf{E}(\mathbf{n}_x, \mathbf{L}_{sky}) + S(x)\mathbf{E}(\mathbf{n}_x, \mathbf{L}_{sun})), \quad (4)$$

where $S(x)$ is a binary value that indicates if x is in shadow. The equation holds independently for each input image t , but we omit the subscript for simplicity.

3.3 Sun and Skylight Separation

We make the simulated skylight for every angle of the sky hemisphere grayscale and fit a Gaussian function of the angle with respect to the sun, γ , as shown in Fig. 2a. Then we subtract the Gaussian sunlight from the original sky hemisphere and model the remainder of the sky using a quadratic function with respect to the sky orientation as a shading equation in Johnson and Adelson (2011),

$$\mathbf{L}_{sky}(\mathbf{s}_k) = \mathbf{s}_k^T \mathbf{A}_q \mathbf{s}_k + \mathbf{b}_q^T \mathbf{s}_k + c_q \quad (5)$$

and

$$L_{sun}(\gamma) = a_g \exp(-((\gamma - b_g)/c_g)^2), \quad (6)$$

where \mathbf{s}_k is the direction of sky element k , which can be defined by γ with respect to the position of the sun; and parameters $\mathbf{A}_q \in \mathbb{R}^{3 \times 3}$ is a symmetric matrix, $\mathbf{b}_q \in \mathbb{R}^{3 \times 1}$ is a vector and $c_q \in \mathbb{R}$ is a constant. Thus, there are three (a_g, b_g, c_g) and ten ($\mathbf{A}_q, \mathbf{b}_q, c_q$) parameters for the Gaussian and quadratic function, respectively. Figure 2c and d show the quadratic skylight (\mathbf{L}_{sky}) and Gaussian sunlight (\mathbf{L}_{sun}) distributions, respectively. The fitting process of the Gaus-

sian and quadratic function is repeated until the estimated parameters converge.

Direct sunlight and ambient skylight may be simply separated if we model the sun as a point light and the ambient skylight as constant (Abrams et al. 2012; Ackermann et al. 2012; Shan et al. 2013). However, in this case, skylight distribution is complex and sunlight is not a point but rather a circle with a 0.5° angular diameter. Non-parametric separation to separate the sun and its surrounding without requiring a parametric model can be a possible solution. Although the process is intuitive and the separation may be lossless, the resulting separated illumination map of high resolution can be somewhat cumbersome in terms of computational time and memory usage. Since the sun occupies very few pixels with very high intensity, it is difficult to properly downsample the illumination map for computational necessity.

The proposed parametric model represents the complex skylight distribution effectively with a few parameters. The sun and skylight illumination maps can then be easily reproduced in any resolution with very small information loss.

When simulating a clear sky ($T = 2$), the maximum value of the absolute intensity difference before and after parameterization is less than 0.04% if the sun elevation angle is greater than 25° . The maximum of the absolute intensity difference decreases as simulated sky turbidity increases. Parameterization tends to fail when sun elevation angle is too low because the sky color distribution becomes more complex (i.e., at dawn or sunset).

3.4 Sky Model Similarity

Does the sky simulation model fully represent real skylight? If not, what is the limitation and possible solution(s)? To address these questions, we compare real sky HDR images with simulated skylight images.

The Laval HDR Sky Database (Lalonde et al. 2016) provides wide angle, high dynamic range images of the sky hemisphere. We used HDR images of the sky captured on three different days with distinct weather conditions for anal-

ysis. Images were captured using the approach described in Stumpfel et al. (2004). The sky images were captured on a clear day, a day with clear sky and mixed clouds, and a cloudy day for the same location in Quebec, Canada. We analyzed images taken every 10 min for 5 h, i.e., 31 images per day. The exact capture time and GPS coordinates for the capture location were specified, so we could simulate corresponding skylight images using Preetham's sky model. We chose $T = 2, 3$, and 5 respectively.

Rather comparing the two skylight image types by appearance, we compared their impact on surface orientation recovery. Consider a small Lambertian surface patch x with surface orientation \mathbf{n}_x and albedo ρ_x . Following Hung et al. (2015), the illumination of an environment map at time t for x can be summarized as a single point light source $\bar{\mathbf{I}}_t \in \mathbb{R}^3$. This virtual point light source is defined as the *mean light vector* and is a function of \mathbf{n}_x because it is computed over the hemisphere defined by \mathbf{n}_x . Thus, patches with different orientations define different sets of mean light vectors.

Given multiple images taken at time $t = \{1, 2, \dots, T\}$, we collect all observations for x to obtain

$$\mathbf{b} = \begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_T \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{I}}_1^T \\ \bar{\mathbf{I}}_2^T \\ \vdots \\ \bar{\mathbf{I}}_T^T \end{bmatrix} \mathbf{y} = \mathbf{Ly}, \quad (7)$$

where \mathbf{y} is the albedo scaled normal vector. Thus, the natural illumination model is similar to a model with a distant point light source, i.e., the traditional photometric stereo. If we assume that the image measurement, \mathbf{b} , is corrupted by zero-mean Gaussian noise, then the surface orientation estimate and observed pixel intensity follow normal distributions. Noise gain, λ_{max} , the maximum value of the square root of the elements on the diagonal of $(\mathbf{L}^T \mathbf{L})^{-1}$, is the only light dependent stability factor, independent of albedo and sensor noise in the confidence interval of the estimated surface orientation (Hold-Geoffroy et al. 2015a). Investigating how λ_{max} is influenced by skylight distributions will allow us to estimate photometric stereo solution stability under the given illumination. Thus, we can focus on whether the sky model generates similar illumination conditions as real skylight in terms of surface orientation recovery rather than similar appearance.

To calculate noise gain, we sampled 642 orientations by subdividing an icosahedron three times and calculated the mean light vector for each surface orientation, since different surface normals yield different integration domains. Thus, we generated one mean light vector per orientation per image. If we consider 31 environment maps, the 31 mean light vectors form a light direction matrix, \mathbf{L} , for each surface orientation. Figure 3 shows λ_{max} for the 642 orientations in spherical and landscape plots under the real sky environment maps

(the left column) and estimated skylight using Preetham's sky model (the middle column). Large λ_{max} indicates that the photometric stereo solution for the corresponding surface orientation carries large uncertainty.

Although λ_{max} under estimated skylight using Preetham's sky model has similar distribution compared with that for the real sky environment map on a clear day (Fig. 3a), the former shows a very different distribution to the latter for clear sky and mixed clouds (Fig. 3b) and cloudy (Fig. 3c) conditions.

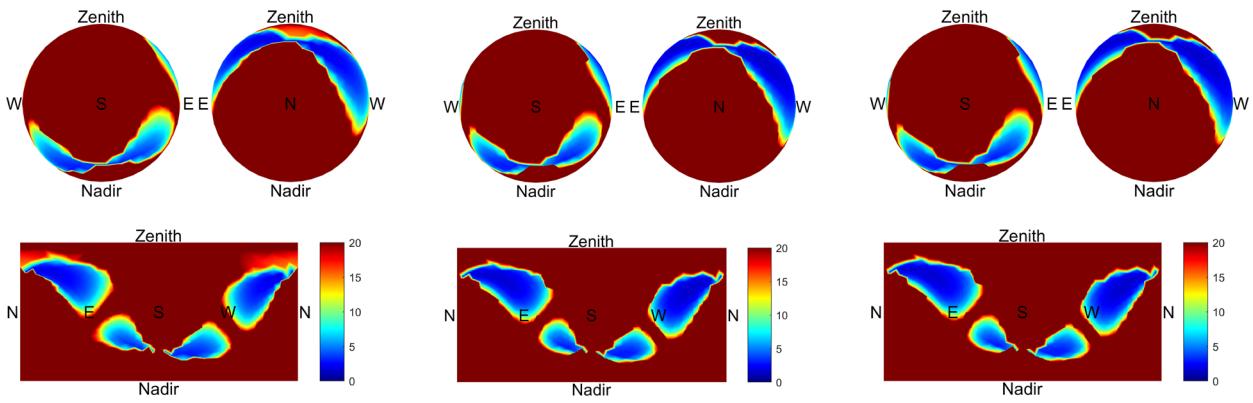
3.5 Sky Model Modification

Given the exact sun location in the sky obtained from time and GPS information, clear sky light distribution can be accurately estimated using Preetham's sky model. Figure 4a shows a set of clear sky images with no cloud and graph of the stable sun intensity. Thus, the sky model is appropriate to estimate a standard clear sky with constant sun intensity throughout the whole observation.

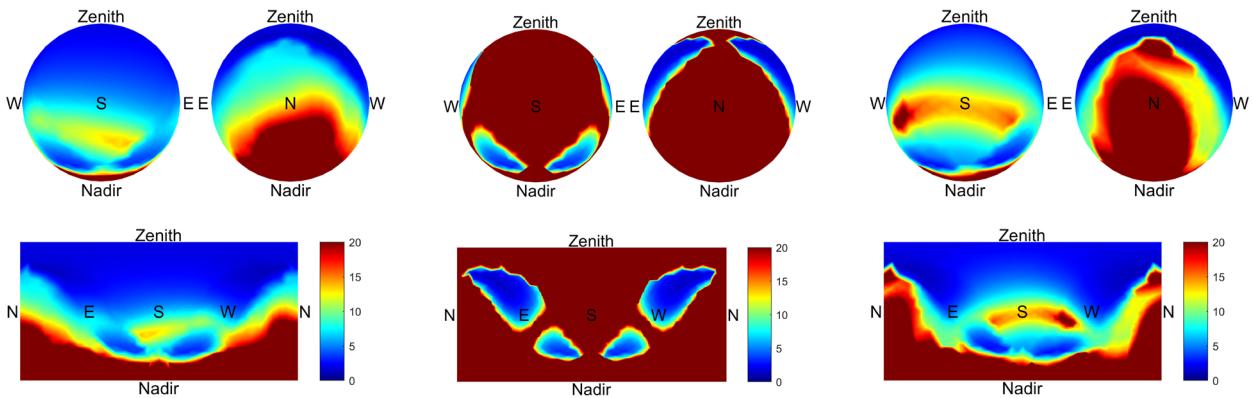
However, the clear sky yields high noise gain when recovering most surface orientations because the mean light vectors remain close to the solar trajectory. Noise gains for the other 2 days have much lower values, indicating smaller uncertainty in the photometric stereo solution. Noise gains for surface orientations to the north and ground are still high due to lack of sunlight chances throughout the day. Thus, it is important to estimate skylight distribution under good lighting conditions to achieve better quality outdoor photometric stereo and surface orientation recovery.

Although Preetham's sky model includes a turbidity parameter to consider weather conditions, the turbidity mainly affects sunlight circular size, i.e., the amount of diffuse sunlight near the sun. Consequently, noise gain under estimated skylight on a clear day (OCT-03-2014, Fig. 3a, middle) and clear sky and mixed clouds (OCT-11-2014, Fig. 3b, middle) have similar distributions, whereas noise gain under real sky environment maps for these 2 days differ significantly (Fig. 3a and b, left). Cloudy day real sky images (Fig. 4b and c) have markedly fluctuating maximum sun intensities due to cloud occlusion. This cloud behavior causes the mean light vectors to be shifted away from the solar trajectory. Turbidity cannot reproduce this fluctuation in sunlight intensity, which is an important factor in photometric stereo performance. Another ambiguity in skylight estimation is skylight intensity, which is a scaling factor per sky image reflecting sky brightness difference between consecutive images. The sky model without the scaling factor estimates skylight relative distribution in one environment map, while different cloud cover affects overall skylight intensity.

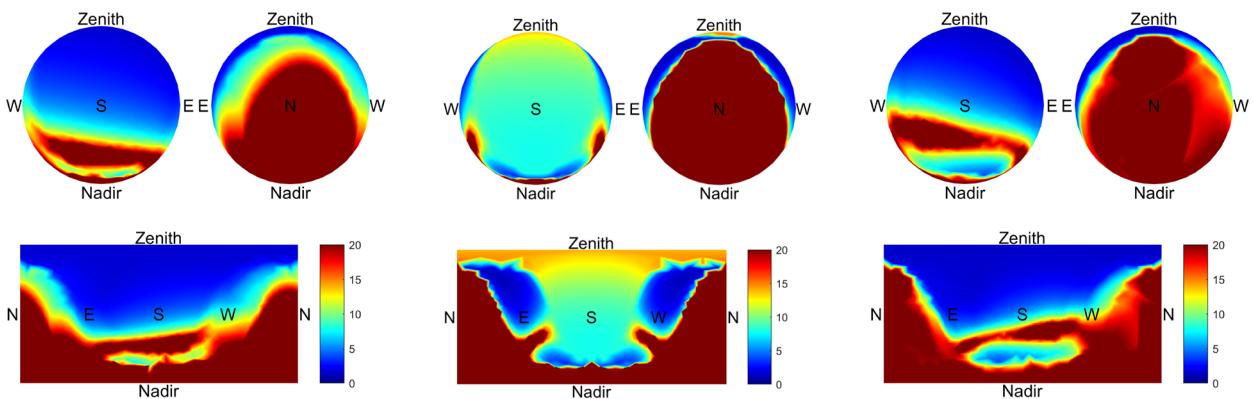
Outdoor photometric stereo using false skylight estimation fails because the estimated illumination differs significantly from the real illumination. Therefore, we propose



(a) Noise gain under the real sky environment map (left), under the estimated skylight using Preetham's sky model (middle), and under the estimated skylight using the modified sky model (right) on a clear day (OCT-03-2014)



(b) Noise gain under the real sky environment map (left), under the estimated skylight using Preetham's sky model (middle), and under the estimated skylight using the modified sky model (right) on a day with clear sky and mixed clouds (OCT-11-2014)

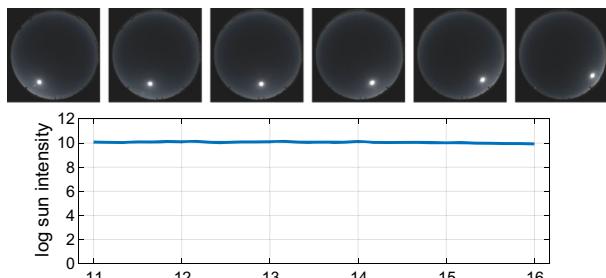


(c) Noise gain under the real sky environment map (left), under the estimated skylight using Preetham's sky model (middle), and under the estimated skylight using the modified sky model (right) on a cloudy day (NOV-08-2014)

Fig. 3 Noise gain for different surface orientations on 3 days with different weather conditions (top to bottom). The noise gain under the real sky environment map (left), the estimated skylight using Preetham's sky model (middle), and the estimated skylight using the modified sky model (right) are compared

a modified sky model including two additional factors to Preetham's sky model to generate similar illumination conditions to the real sky environment map regardless of weather

conditions. The two factors for each image are maximum sun and skylight intensities. Outdoor photometric stereo under clear sky works with or without the additional scaling fac-



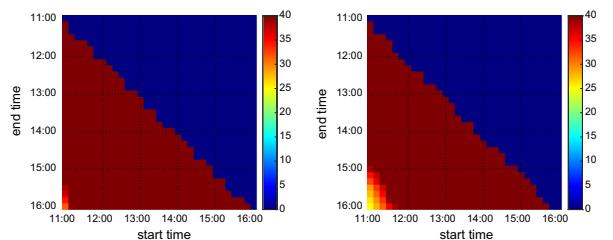
(a) Sun intensity on a clear day (OCT-03-2014)
(b) Sun intensity on a day with clear sky and mixed clouds (OCT-11-2014)
(c) Sun intensity on a cloudy day (NOV-08-2014)

Fig. 4 Real sky images and their log sun intensities on 3 days with different weather conditions. Photographs of the sky are tone-mapped for display

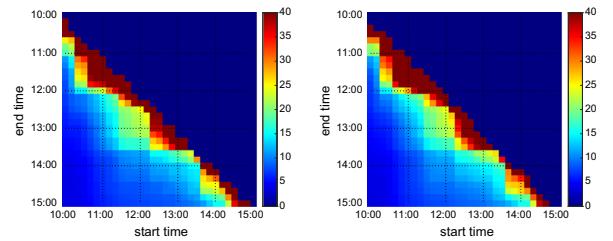
tors because they are constant throughout the whole day. These two factors are required to estimate skylight with better lighting conditions for photometric stereo. We show that adding these two factors allows the modified sky model to successfully reproduce real sky images with clouds in terms of surface normal reconstruction.

The maximum sun intensity is a_g in Eq. (6). If we parameterize estimated skylight distribution generated by the original Preetham's sky model into L_{sun} and L_{sky} as discussed in Sect. 3.3, a_g is almost constant. Hence we simply change a_g value after parameterization to be the maximum sun intensity in the modified sky model. Skylight intensity, d_{si} , is the relative scaling factor between two sky environment maps. Thus, the modified sky model is

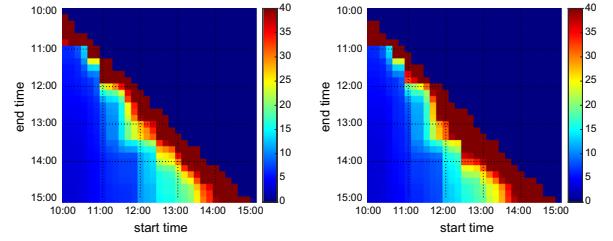
$$L_{sky}^{mod}(s_k) = L_{sky}(s_k) \cdot d_{si} = (s_k^T \mathbf{A}_q s_k + \mathbf{b}_q^T s_k + c_q) \cdot d_{si}. \quad (8)$$



(a) The median noise gain under the real sky map and the estimated skylight using the modified sky model on a clear day (OCT-03-2014)



(b) The median noise gain under the real sky map and the estimated skylight using the modified sky model on a day with clear sky and mixed clouds (OCT-11-2014)



(c) The median noise gain under the real sky map and the estimated skylight using the modified sky model on a cloudy day (NOV-08-2014)

Fig. 5 The expected uncertainty of outdoor photometric stereo as a function of time on 3 days with different weather conditions

To validate the effectiveness of the modified sky model, we found the two scaling factors directly from real sky images and applied them to the estimated skylight. We parameterized the real sky HDR image and estimated skylight using the original Preetham's sky model into L_{sky} and L_{sun} using Eqs. (5) and (6), and then changed a_g for estimated skylight to be that for the real sky image. We calculated skylight intensity d_{si} of the modified sky model as the average scalar between estimated skylight and real sky image.

Figure 3 compares noise gain behavior for real sky images and modified skylight estimation for different surface orientations and time intervals. The real sky environment maps (the leftmost column) and skylight estimation using the modified sky model (the rightmost column) show similar noise gain distribution over the surface orientation. Figure 5 compares noise gain over different time intervals (Hold-Geoffroy et al. 2015b). Noise gain drops at the same moments in both matrices, indicating that the estimated skylight distribution

has the similar ability at recovering the surface normal as its corresponding real sky environment map.

Figure 5 also shows how outdoor photometric stereo conditioning evolves over time, visualizing λ_{max} for all possible time intervals in the three different days. Since λ_{max} varies with \mathbf{n} , we plotted median gain over 321 normal vectors visible to the camera for each time interval. Figure 5a shows the clear day case yields poor outdoor photometric stereo reconstruction. Low noise gain is never achieved, regardless of start time and capture interval. The other 2 days with clouds (Fig. 5b and c) exhibit better behavior than the clear day, because clouds occluding and revealing the sun significantly affect sun intensity (Fig. 4b and c) and cause mean light vectors to shift away from the solar trajectory. The resulting light direction matrices are more likely to satisfy rank-3 constraint, producing lower noise gain in photometric stereo solution.

4 Surface Normal and Albedo Estimation

This section presents the proposed framework for outdoor photometric stereo using estimated natural illumination consisting of skylight and sunlight distributions. We detect shadowed pixels explicitly and estimate initial relative albedo using the color ratio. The surface normal is estimated using the correlation between pixel and simulated profiles generated according to the image formation model. Absolute albedo is then updated and the surface normal is refined by MRF optimization.

4.1 Albedo Estimation

A color profile is a set of RGB color values at the same pixel location in a set of photometric stereo images. It is represented as a blue line in RGB space, as shown in Fig. 6a. Shi et al. (2010) assessed color profile nonlinearity for radiometric calibration. Although this is a simple and clear method to estimate the inverse response function for a high quality camera in a dark room, it was difficult to define the exact curvature of the color profile for the current case because illumination intensity varies significantly and profiles are noisy when images are captured by low quality webcam.

Most cameras we employed were radiometrically uncalibrated. However, color profiles in the middle of RGB space exhibited sufficient linearity to estimate the color ratio ($\rho_r : \rho_g : \rho_b$). We separated relative albedo, $\hat{\rho}_c$, for each color channel, c , in Eq. (4), so the remainder of the equation refers to the common term for all three color channels,

$$I_c(x) = \hat{\rho}_c(x)\rho_s(x)(\mathbf{E}(\mathbf{n}_x, \mathbf{L}_{sky}) + S(x)\mathbf{E}(\mathbf{n}_x, \mathbf{L}_{sun})) \quad (9)$$

and

$$\begin{aligned} \boldsymbol{\rho} &= [\rho_r, \rho_g, \rho_b]^T = \rho_s[\hat{\rho}_r, \hat{\rho}_g, \hat{\rho}_b]^T = \rho_s\hat{\boldsymbol{\rho}}, \\ s.t. \quad &\sqrt{\hat{\rho}_r^2 + \hat{\rho}_g^2 + \hat{\rho}_b^2} = 1. \end{aligned} \quad (10)$$

We constructed a matrix with the color profile of a pixel, excluding any RGB set with at least one saturated channel, and then estimated relative albedo using singular value decomposition as the first eigenvector of the color profile matrix, as shown in Fig. 6a, red.

We projected the color profile onto its albedo line in RGB space obtain pixel intensity profiles, rather than converting the input image to grayscale. The projected values form a pixel profile, i.e., the term $\rho_s(x)(\mathbf{E}(\mathbf{n}_x, \mathbf{L}_{sky}) + S(x)\mathbf{E}(\mathbf{n}_x, \mathbf{L}_{sun}))$ in Eq. (9), which is the common value for all three color channels.

Saturated values had to be corrected before projecting the color profile onto its albedo line. When one channel was saturated, we estimated its proper value above saturation using the unsaturated channels and relative albedo, assuming the color set would have the same color ratio as the relative albedo. The pixel profile was then generated using the modified values. The same method was applied where two channels were saturated. Figure 6b shows an example color profile and corresponding pixel profile.

4.2 Surface Normal Estimation

The pixel profile is proportional to the term inside the parentheses in Eq. (9) by the scalar $\rho_s(x)$ per pixel. Relative albedo, specifying the color ratio of the pixel, becomes absolute albedo by multiplying $\rho_s(x)$.

We uniformly sampled 1000 orientations on a unit sphere using Saff and Kuijlaars (1997), generating the corresponding incident irradiance $\mathbf{E}(\mathbf{n}_s, \mathbf{L}_{sky})$ and $\mathbf{E}(\mathbf{n}_s, \mathbf{L}_{sun})$ for skylight and sunlight, respectively, for each unit vector \mathbf{n}_s and m input images. Then we simulated 1000 profiles for each x using skylight irradiance, sunlight irradiance, and pixel shadow, $S(x)$, using Eq. (9), for each image t without multiplying its albedo. Simulated profiles were generated as $(\mathbf{E}(\mathbf{n}_s, \mathbf{L}_{sky}) + S(x)\mathbf{E}(\mathbf{n}_s, \mathbf{L}_{sun}))$ in Eq. (9) using 1000 different orientations for x . Therefore, every simulated profile has corresponding orientation, \mathbf{n}_s , and all profiles are m -dimensional vectors.

We found the best simulated profile, i.e., the one with highest correlation with the pixel profile. To address scalar ambiguity between the simulated and pixel profiles, we employed Pearson correlation coefficient, since it is invariant to scaling. After selecting the best match, the unit vector \mathbf{n}_s which is used to generate the best simulated profile becomes the surface normal estimate for the pixel.

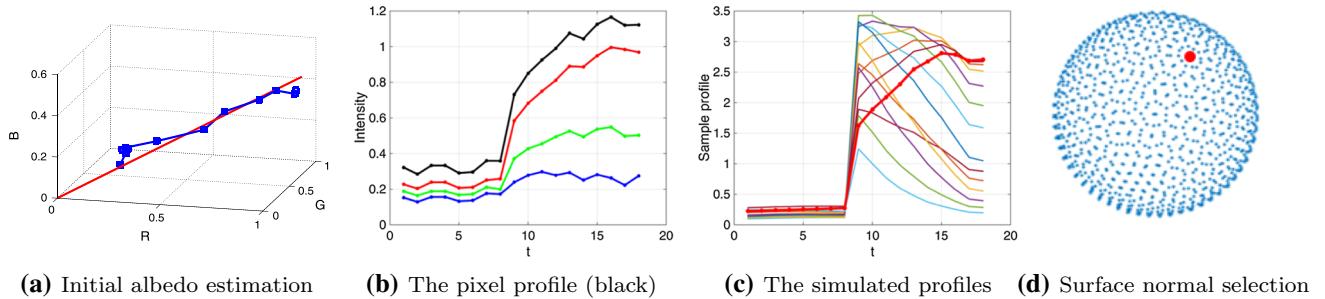


Fig. 6 An illustration of albedo and surface normal estimation process. **a** The initial relative albedo for a pixel is estimated as the first eigen vector in RGB space (red line). **b** The color values (red, green, and blue lines) are projected onto the albedo line in RGB space to be the

pixel profile (black line). **c** 1000 simulated profiles are generated and the best match with the pixel profile is selected. **d** The corresponding surface orientation is selected as the surface normal estimate (Color figure online)

After surface normal estimation, we calculated the scalar factor, $\rho_s(x)$, using least squares and hence the absolute albedo. Selecting the best surface orientation based on a measure as simple as correlation is possible because the only ambiguity in the proposed image formation model is a scalar per pixel.

The proposed framework is simple and fast. Processing 18 640 × 480 pixel images with 166,000 valid pixels took 16 s to simulate and parameterize skylight distributions, 50 s for shadow detection, and 19 s for albedo and surface normal estimation using an unoptimized Matlab implementation on a 3.40 GHz computer.

4.3 MRF Optimization

Since we process a handful of input images, some pixels are in shadow for most images, i.e., there are insufficient observations to determine their surface normals, hence their estimates are usually incorrect. We used estimates from their neighboring pixels to solve this problem. The basic concept is that if neighboring pixel color profiles are similar, their surface normal estimates should also be similar.

Given a set of pixels, \mathcal{P} , and a set of labels, \mathcal{L} , corresponding to uniformly sampled unit vectors, our goal is to find a labeling l , i.e., a mapping from \mathcal{P} to \mathcal{L} , that minimizes

$$E(l) = \sum_{p \in \mathcal{P}} D_p(l_p) + \sum_{p,q \in \mathcal{N}} V_{p,q}(l_p, l_q), \quad (11)$$

where $\mathcal{N} \subset \mathcal{P} \times \mathcal{P}$ is a neighborhood pixel system. $D_p(l_p)$ is a data term that measures the cost of assigning \mathbf{u}_{l_p} as the surface normal for pixel p ,

$$D_p(l_p) = 1 - \text{corr}(SP(l_p, p), PP(p)); \quad (12)$$

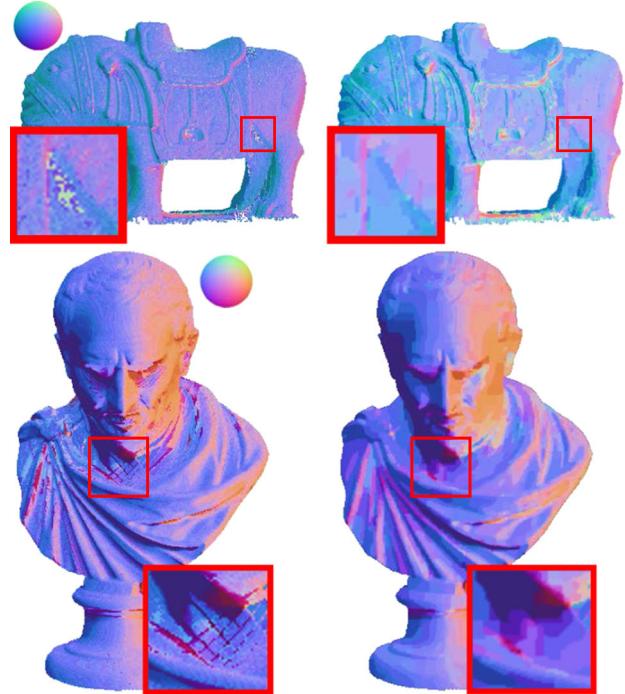


Fig. 7 Surface normal estimation before and after MRF optimization. Noisy estimates becomes smoother

$V_{p,q}(l_p, l_q)$ is a smoothness term that measures the cost of assigning \mathbf{u}_{l_p} and \mathbf{u}_{l_q} to adjacent pixels p and q ,

$$V_{p,q}(l_p, l_q) = \sum_{c=R,G,B} (1 + \text{corr}(I_c(p), I_c(q))) (1 - \mathbf{u}_{l_p}^T \mathbf{u}_{l_q}), \quad (13)$$

where $\text{corr}(a, b)$ denotes Pearson correlation coefficient for profiles a and b , $SP(l, p)$ is the simulated profile for \mathbf{u}_l using the shadow mask of pixel p , and $PP(p)$ is the pixel profile of p . $I_c(p)$ denotes the color profile of p for all

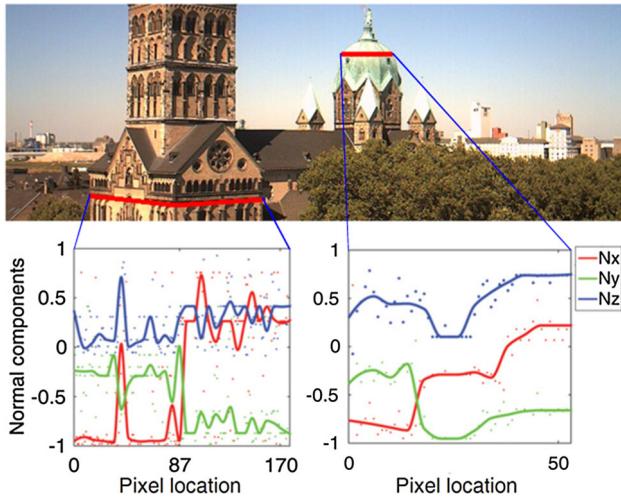


Fig. 8 Surface normal components on a scanline. x (East, red), y (North, green), and z (Up, blue) components of the normal vectors on the scanlines through the tower and the dome in *Düsseldorf* dataset are shown. The lines are fitted after 1D median filtering (Color figure online)

images. We performed multi-label optimization using graph cut (Kolmogorov and Zabih 2004) to identify the set of surface orientations that minimized $E(l)$.

5 Experiments

We demonstrated the proposed approach using five real-world datasets. *Elephant* and *Cicero* datasets were captured for two different days using the same Canon 5D Mark III camera, with each set containing 15 images taken at 30-min intervals. *Düsseldorf*, *Meersburg*, and *Arizona* datasets were sourced from the AMOS webcam archive (Jacobs et al. 2007), and sunny days were manually selected providing 18 images or less for each set.

For all experiments, we assumed Lambertian reflectance and image intensities were linear to scene radiance except

where saturated. We did not include explicit radiometric calibrations due to their unstable outcomes. However, appropriately modified radiometric calibration methods [e.g. Lee et al. (2013)] may improve the performance of our method.

5.1 Qualitative Evaluation

Figure 1 shows the proposed algorithm application followed by 3D reconstruction using Poisson solver (Simchony et al. 1990). The smooth variation of the estimated normals successfully reconstructed the surface. MRF refinement removes sparkling noise, which can cause significant surface reconstruction errors. Figure 7 shows the proposed surface normal estimation before and after MRF refinement. Although the first estimation embeds fine shape details, MRF refinement removes noisy estimates, particularly in shaded regions, yielding smooth surfaces.

Figure 8 shows surface normal estimates for two *Düsseldorf* dataset scanlines. The normals are in *xyz*-coordinate, corresponding to absolute East-North-Up. The building in the scene is facing approximately south according to its satellite image. Pixel 87 on the scanline (Fig. 8, left) corresponds to the tower corner. The red line (x , East) changes from facing west to east, whereas the blue line (z , Up) remains close to zero since both faces are vertical. The peak at pixel 40 corresponds to the arch on the wall. The red line in Fig. 8, right, shows the slightly discontinuous dome structure in the scene.

Figures 9 and 10 compare the proposed framework performance with two previous methods. Figure 10 compares albedo and surface normal estimation from the proposed method using 1-day images of *Düsseldorf* and *Meersburg* datasets. Abrams et al. (2013a) shadow estimation method which estimates albedo, surface normals, and ambient light is compared using the same 1-day images. The result of Abrams et al. (2012) using 500 images of several months are obtained from their original paper. We used the same color map defined in Abrams et al. (2012) for comparison. Since the color coding was based on absolute East-North-Up coordinates, we

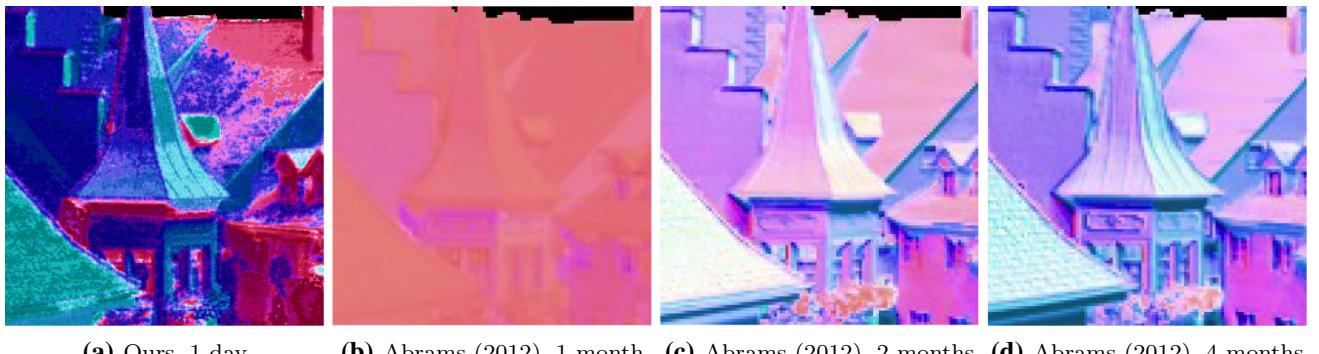


Fig. 9 Comparison with the previous method using months of images. We compare (a) our result computed from 1-day images with (b–d) those reported by Abrams et al. (2012) using several months of images

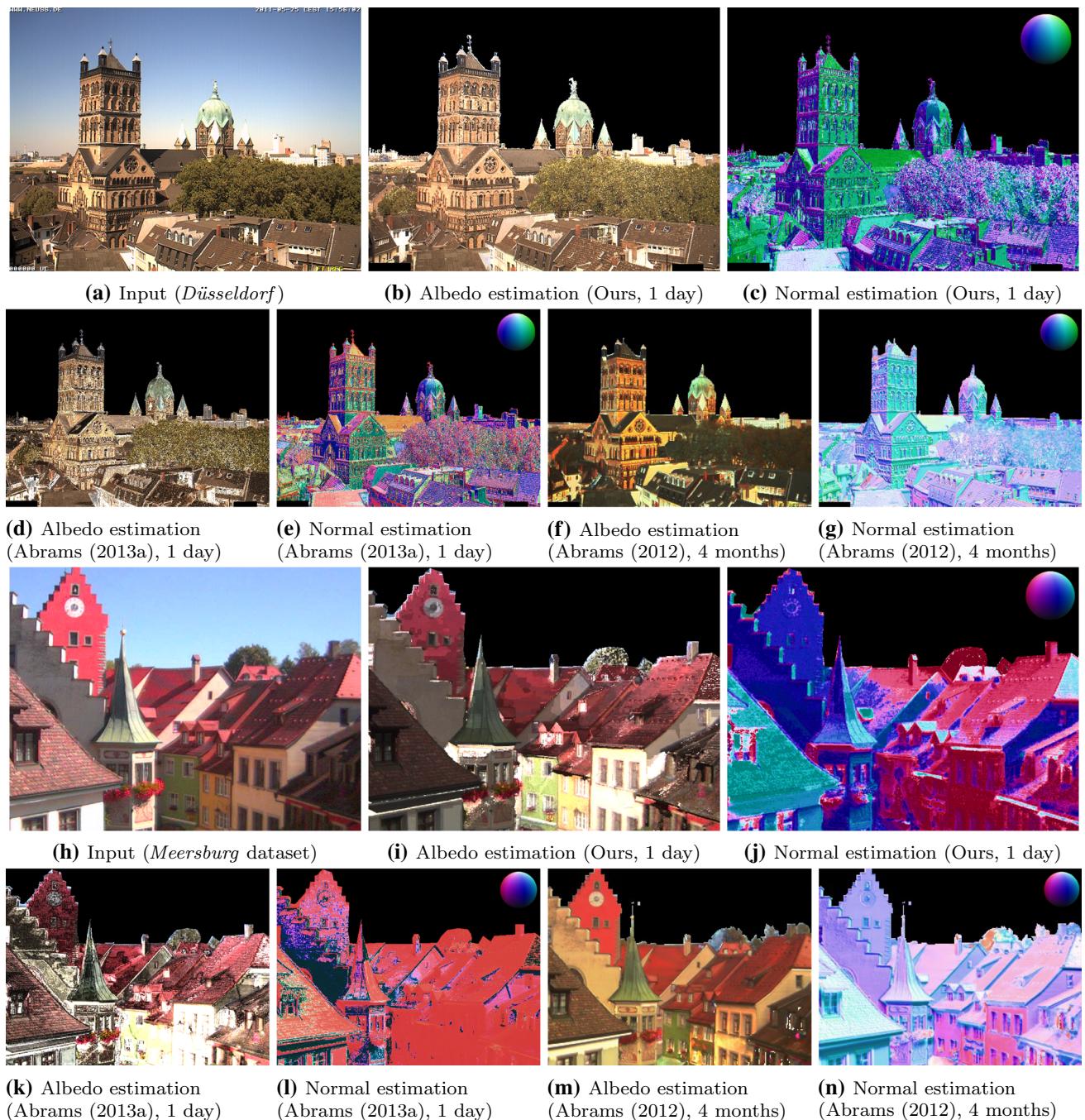


Fig. 10 Qualitative comparison on *Düsseldorf* and *Meersburg* datasets. Our results of 1-day outdoor photometric stereo are compared with two previous methods. We compare the albedo and surface normal estimation of Abrams et al. (2013a) shadow detection method explained in

Sect. 3.2 using the same images of a day as ours. The results of Abrams et al. (2012) using 4 months of images are obtained from their original paper

indicate the color sphere of the corresponding view point with the surface normal results.

The proposed method significantly outperforms other methods using 1 day or 1 month of images. The illumination model of a point light source and constant ambient light exhibits severe performance degradation when observations

do not cover several months. On the other hand, the proposed natural illumination model using skylight distribution provides sufficient constraints from small number of observations and has similar or better performance with previous methods using 4 months of images.

Table 1 Performance evaluation on fifteen different sets of 1-day images (R30) [%]

	Minimum	Average	Maximum
Abrams (2012)	0.00	2.09	4.09
Abrams (2013b)	11.8	19.1	25.9
Our method	29.9	36.1	42.2

Table 2 Performance evaluation on many days of images (R30) [%]

	1 day	1 week	2 weeks	4 months
Abrams (2012)	1.40	2.54	11.2	31.6
Abrams (2013b)	12.9	24.7	36.4	44.8
Our method	38.1	40.8	47.0	51.0

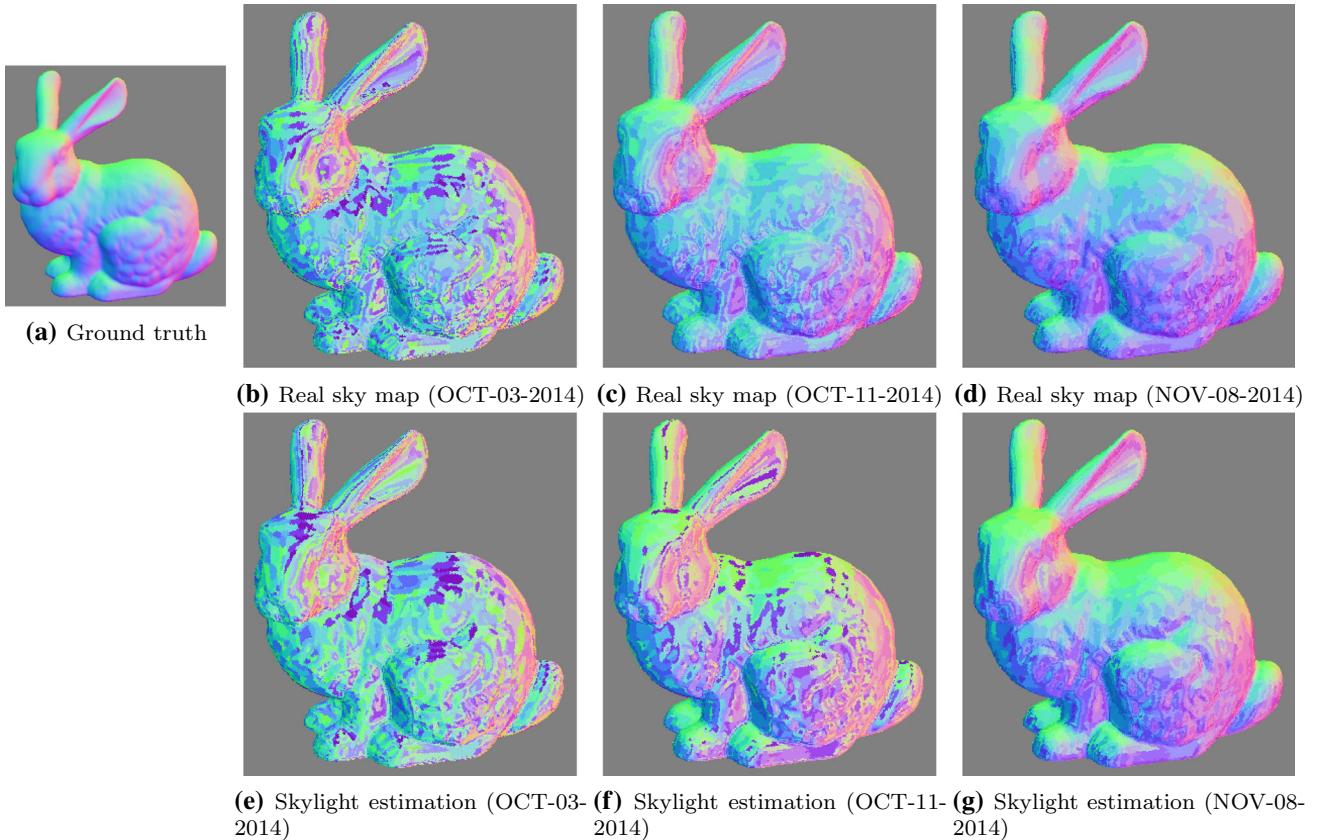


Fig. 11 Photometric stereo results of synthetically rendered images using real sky environment maps (upper row) and skylight estimation using the modified sky model with user assistance (lower row)

5.2 Quantitative Evaluation

To provide quantitative evaluation, we obtained a reference surface normal map for the *Arizona* dataset. This reference normal was from Google Earth, and many locations in the reference have substantial angular error because it does not consider trees and bushes in the scene (Abrams et al. 2012). Therefore, we computed the robustness statistics used in the Middlebury stereo dataset (Scharstein and Szeliski 2002) for evaluation. In particular, R30 denotes the percentage of pixels

that have angular error less than 30° . For the same reason, we also omitted ground plane regions in the reference from evaluation.

Table 1 compares the stability for the proposed, Abrams et al. (2012), and Abrams et al. (2013b) methods on 15 1-day datasets. Each set (day) was selected from an individual month among 15 months of data. Abrams et al. (2012) failed for some sets due to severe rank deficiency of the sun trajectory. The average performance of the proposed method was

Table 3 Performance evaluation on synthetically rendered images (Median angular error) [degree]

Date	Weather	PS using real sky environment map	PS using skylight estimation
OCT-03-2014	Clear	26.0	32.9
OCT-11-2014	Clear sky with mixed clouds	13.2	24.2
NOV-08-2014	Cloudy	10.2	14.3

Table 4 Performance evaluation on synthetically rendered images (R30) [%]

Date	Weather	PS using real sky environment map	PS using skylight estimation
OCT-03-2014	Clear	54.7	46.9
OCT-11-2014	Clear sky with mixed clouds	88.5	64.1
NOV-08-2014	Cloudy	96.6	95.7

Table 5 Performance evaluation of the modified sky model with user assistance on 30 different sets of 1-day real images (R30) [%]

	Minimum	Average	Maximum
Our method	30.8	38.7	47.2

significantly superior to the comparator methods and exhibited stable performance over different data sets.

Table 2 compares performances using many days of images. The proposed method using 1-day images is superior to Abrams et al. (2012) using 4 months of images, and improves with increasing image inclusion.

5.3 Synthetic Experiments

The proposed modified sky model was verified using outdoor photometric stereo on synthetic object images with known ground truth normals. We generated synthetic images of Stanford Bunny (Turk and Levoy 1994) under real sky environment maps. One real-world environment map was used as the sole light source for each image. Cast shadows were not simulated to isolate the analysis to photometric cues alone.

We performed calibrated outdoor photometric stereo using real sky environment maps and uncalibrated outdoor photometric stereo using skylight estimation, which is the proposed method. For calibrated outdoor photometric stereo, we employed the proposed scheme introduced in Sect. 4.2 using the real sky environment maps, but omitted shadow estimation, albedo estimation, and MRF optimization in the synthetic experiment.

In contrast to the noise gain analysis using skylight estimation in Sect. 3.5, we did not measure maximum sun and skylight intensity from real sky images, but estimated the two scaling factors from input images. This process required user assistance rather than real sky images, which are often not obtainable in practice. We assumed that sufficiently accurate ground truth normals for a couple of pixels can be provided

by the user with some intuition (Koenderink et al. (1992) presented related studies). For example, if a scene contains the ground plane, or vertical walls, the user can assign a proper surface orientation to these image portions. Then scaling factors for the modified sky model per image can be estimated by investigating the pixel profiles of those pixels, i.e., can generate the incident irradiance $\mathbf{E}(\mathbf{n}_x, \mathbf{L}_{sky})$ for pixel x , under skylight given its surface orientation. Comparing this term to the pixel profile, we can calculate the skylight intensity of the image if the pixel is in shadow for image t . Another pixel with different surface orientation that is not in shadow for image t can yield maximum sun intensity at t . Since the two scaling factors are defined for each input image, the method requires the user to assign surface orientations for one pixel in shadow and another pixel in sunlight for each input image. That is, if both of these pixels are in shadow or in sunlight for a certain image, then the surface orientation for another pixel should be identified by the user.

For the synthetic experiment, we first estimated skylight intensity per image by comparing the pixel intensities of known surface orientations which did not receive direct sunlight to the expected radiance from the quadratic skylight distribution. Maximum sun intensity was then estimated using the intensity for a sunlit pixel with different surface orientation.

Figure 11 shows photometric stereo results. Each column uses different sets of images rendered by different real sky environment maps. Column 1 shows the *Bunny* model ground truth normal, and column 2 the photometric stereo results for synthetically rendered images under clear sky. Figure 11b shows calibrated photometric stereo results using real sky environment maps. Although we used real lighting environment, many recovered surface normals show considerable angular error. This is because the clear sky was insufficient to disambiguate certain surface orientations, as discussed in Sect. 3.5. Table 3 shows that median angular error is 26° and Table 4 shows that 54.7% of pixels have angular error less than 30° (i.e., R30).



Fig. 12 Input images of *Arizona* dataset (JUL-07-2011). Three dots are the points of user assistance

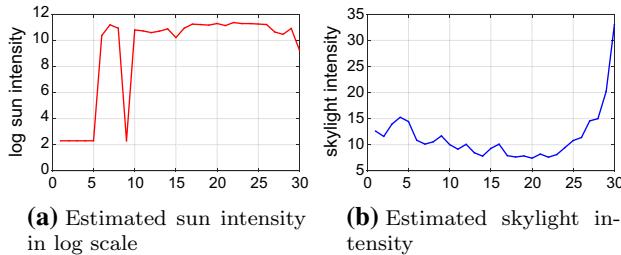


Fig. 13 Estimated sun and skylight intensity of *Arizona* dataset (JUL-07-2011)

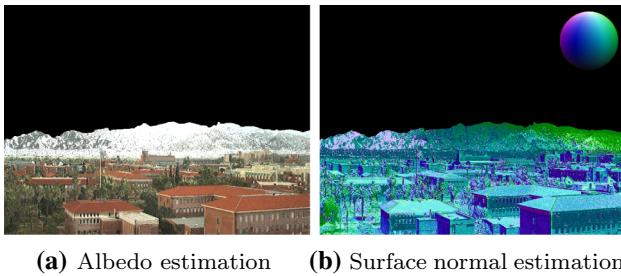


Fig. 14 Albedo and surface normal estimation of *Arizona* dataset (JUL-07-2011)

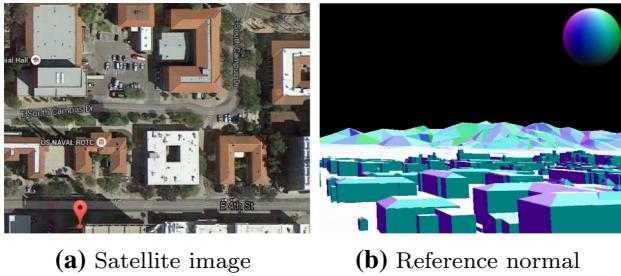


Fig. 15 Reference information of *Arizona* dataset

Figure 11c and d show the other calibrated photometric stereo results using real sky environment maps for clear sky and mixed clouds, and cloudy days, respectively. Both cases exhibit significantly improved performance compared with clear day results due to low noise gain for these lighting conditions.

Figure 11e–g show uncalibrated photometric stereo results using skylight estimation with user assistance. Figure 11g exhibits superior performance since skylight estimation using the proposed modified sky model accurately estimates the environment maps which carry low noise gain in recov-



Fig. 16 Input images of *Arizona* dataset (SEP-07-2011). Three dots are the points of user assistance

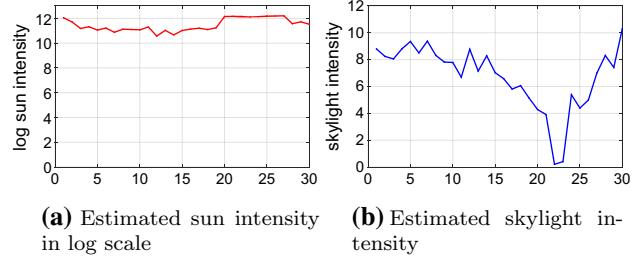


Fig. 17 Estimated sun and skylight intensity of *Arizona* dataset (SEP-07-2011)

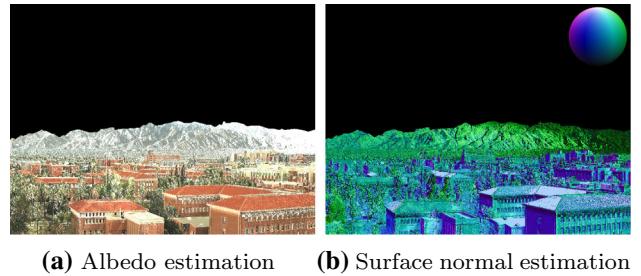


Fig. 18 Albedo and surface normal estimation of *Arizona* dataset (SEP-07-2011)

ering surface orientations. Table 4 shows that 95.7% of the pixels were recovered with angular error less than 30° for the cloudy day dataset. Clearly, skylight estimation performance cannot exceed calibrated photometric stereo performance. Hence to successfully recover outdoor scene surface normals, it is essential to process images captured under good lighting conditions and accurately estimate the skylight distribution.

5.4 Real Data Experiments

The proposed modified sky model were tested with user assistance for outdoor photometric stereo using real images. For quantitative evaluation, we used the *Arizona* dataset from Sect. 5.2 with a reference surface normal ground truth obtained from Google Earth. As above, we omitted ground plane regions in the reference.

Table 5 shows the stability of the proposed modified sky model, tested with two additional scaling factors on 30 1-day datasets selected from a 7-month period. The modified sky model shows improved average performance compared with Table 1, although the tables are based on different datasets: Table 1 dataset includes only days with clear skies, whereas



Fig. 19 Input images of *Düsseldorf* dataset (JUN-12-2011). Three dots are the points of user assistance

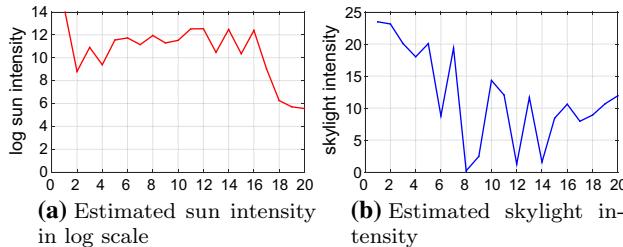


Fig. 20 Estimated sun and skylight intensity of *Düsseldorf* dataset (JUN-12-2011)



Fig. 21 Albedo and surface normal estimation of *Düsseldorf* dataset (JUN-12-2011)

Table 5 includes only cloudy days. Both sets were captured from the same camera at the same spot, but on different days.

Figure 12 shows three of the 30 input images for photometric stereo where red, green, and blue dots indicate user assistance points. As explained above, we used intensity profiles of those points to estimate the modified sky model scaling factors. Red and green pixels are vertical walls facing west and south, respectively. The blue pixel is the flat ceiling of the building whose surface is pointing upward. Figure 13 shows the estimated sun and skylight intensity for each input image. The abrupt decrease is due to cloud occlusion, causing the mean light vector to deviate from the original solar path, and hence providing improved results for outdoor photometric stereo. Figure 14 shows albedo and surface normal estimation result of *Arizona* dataset (JUL-07-2011). Refer to the satellite image in Fig. 15a).

Figures 16, 17, and 18 show results using *Arizona* dataset images captured on a different day, and Figs. 19, 20, 21, 22, 23, and 24 show results using *Düsseldorf* dataset images, which were used in qualitative evaluation in Sect. 5.1. Red, green, and blue dots in Fig. 19 indicate user assistance points, which are assumed to be vertical or horizontal planes.



Fig. 22 Input images of *Düsseldorf* dataset (SEP-02-2011). Three dots are the points of user assistance

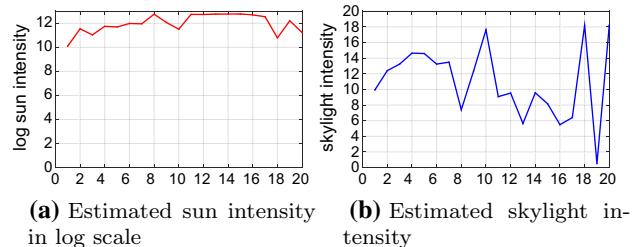
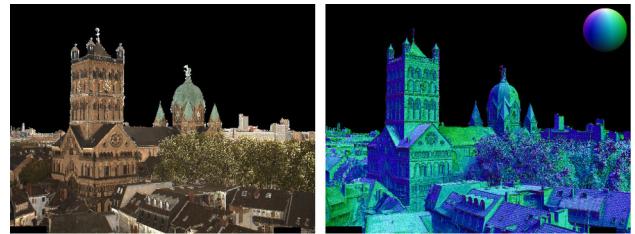


Fig. 23 Estimated sun and skylight intensity of *Düsseldorf* dataset (SEP-02-2011)



(a) Albedo estimation **(b)** Surface normal estimation

Fig. 24 Albedo and surface normal estimation of *Düsseldorf* dataset (SEP-02-2011)

Figures 21 and 24 show that the proposed modified sky model has improved performance compared to Fig. 10.

6 Conclusion

We proposed an outdoor photometric stereo method using skylight estimation, modeling the outdoor lighting environment as natural illumination. This provides more constraints to surface normal estimation than conventional point light source with constant ambient light models. Thus, significantly less images are required to recover the outdoor scene, which makes it possible to use a set of images captured in a single day. The proposed method is practical and exhibited similar or superior performance compared with current outdoor photometric stereo methods using a set of images captured over several months.

We analyzed environment lighting conditions for outdoor photometric stereo to identify essential observation characteristics to recover scene surface normals with better quality, and how to utilize such characteristics in outdoor photometric stereo.

Assuming Lambertian reflectance for the observed surface patch, we showed that a set of sky hemispheres with their mean light vectors deviated from the solar trajectory throughout its observations yield smaller surface normal estimation uncertainty. Mean light vectors are more likely to stay in the solar trajectory if the sky is clear. Therefore, intermittent cloud occlusion of the sun during observation accompanied by proper skylight estimation greatly improved outdoor photometric stereo uncertainty. We proposed a modified sky model to estimate skylight in such conditions with user assistance to disambiguate the two scaling factors in the proposed sky model.

One limitation of the proposed method is that we must choose between a fully automatic reconstruction on a clear day or user assisted semi-automatic reconstruction on a non-clear day. When real sky images are not available for real world outdoor photometric stereo problems, the proposed method requires a clear day for fully automatic reconstruction and has performance limitations due to ill-conditioned illumination. Semi-automatic reconstruction on nonclear days can achieve better results, but this requires user input to estimate the skylight using the modified sky model.

We validated the analyses and the proposed modified sky model using outdoor photometric stereo results for synthetically rendered images including an object with known ground truth normal, and real images of general outdoor scenes.

Acknowledgements This work was supported by a Grant from Kyung Hee University (KHU-20170718) and National Research Foundation of Korea (NRF-2017R1C1B5075945).

References

- Abrams, A., Hawley, C., & Pless, R. (2012). Heliometric stereo: Shape from sun position. In *Proceedings of European conference on computer vision (ECCV)*.
- Abrams, A., Hawley, C., Miskell, K., Stoica, A., Jacobs, N., & Pless, R. (2013a). Shadow estimation method for “the episolar constraint: Monocular shape from shadow correspondence”. *arXiv:1304.4112*.
- Abrams, A., Miskell, K., & Pless, R. (2013b). The episolar constraint: Monocular shape from shadow correspondence. In *Proceedings of IEEE conference on computer vision and pattern recognition (CVPR)*.
- Ackermann, J., Langguth, F., Fuhrmann, S., & Goesele, M. (2012). Photometric stereo for outdoor webcams. In *Proceedings of IEEE conference on computer vision and pattern recognition (CVPR)*.
- Agarwal, S., Furukawa, Y., Snavely, N., Simon, I., Curless, B., & Seitz, S. M. (2011). Building Rome in a day. *Communications of the ACM*, 54(10), 105–112.
- Furukawa, Y., & Ponce, J. (2010). Accurate, dense, and robust multi-view stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 32(8), 1362–1376.
- Google Earth. (2015). <https://earth.google.com>.
- Grossmann, P. (1987). Depth from focus. *Pattern Recognition Letters*, 5(1), 63–69.
- Han, Y., Lee, J. Y., & Kweon, I. S. (2013). High quality shape from a single RGB-D image under uncalibrated natural illumination. In *Proceedings of international conference on computer vision (ICCV)*.
- Hold-Geoffroy, Y., Zhang, J., Gotardo, P. F. U., & LaLonde, J. F. (2015a). What is a good day for outdoor photometric stereo? In *International conference on computational photography*.
- Hold-Geoffroy, Y., Zhang, J., Gotardo, P. F. U., & LaLonde, J. F. (2015b). x-hour outdoor photometric stereo. In *International conference on 3D vision*.
- Hold-Geoffroy, Y., Sunkavalli, K., Hadap, S., Gambaretto, E., & LaLonde, J. F. (2017). Deep outdoor illumination estimation. In *Proceedings of IEEE conference on computer vision and pattern recognition (CVPR)*.
- Horn, B. K. P. (1989). Obtaining shape from shading information. In *Shape from shading* (pp. 123–171). Cambridge, MA, USA: MIT Press.
- Hošek, L., & Wilkie, A. (2013). Adding a solar radiance function to the Hošek–Wilkie skylight model. *IEEE Computer Graphics and Applications*, 33(3), 44–52.
- Huang, R., & Smith, W. A. P. (2011). Shape-from-shading under complex natural illumination. In *Proceedings of international conference on image processing (ICIP)*.
- Hung, C. H., Wu, T. P., Matsushita, Y., Xu, L., Jia, J., & Tang, C. K. (2015). Photometric stereo in the wild. In *Proceedings of IEEE Winter Conference on Applications of Computer Vision (WACV)*. Waikoloa, HI: IEEE.
- Inose, K., Shimizu, S., Kawakami, R., Mukaigawa, Y., & Ikeuchi, K. (2013). Refining outdoor photometric stereo based on sky model. *IPSJ Transaction on Computer Vision and Applications*, 5(1), 104–108.
- Jacobs, N., Roman, N., & Pless, R. (2007). Consistent temporal variations in many outdoor scenes. In *Proceedings of IEEE conference on computer vision and pattern recognition (CVPR)*.
- Johnson, M. K., & Adelson, E. H. (2011). Shape estimation in natural illumination. In *Proceedings of IEEE conference on computer vision and pattern recognition (CVPR)*.
- Jung, J. (2016). Outdoor photometric stereo for 3D scene recovery. Ph.D. Thesis, KAIST, Daejeon, Republic of Korea.
- Jung, J., Lee, J. Y., Jeong, Y., & Kweon, I. S. (2015a). Time-of-flight sensor calibration for a color and depth camera pair. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 37(7), 1501–1513.
- Jung, J., Lee, J. Y., & Kweon, I. S. (2015b). One-day outdoor photometric stereo via skylight estimation. In *Proceedings of IEEE conference on computer vision and pattern recognition (CVPR)*.
- Kawakami, R., Zhao, H., Tan, R. T., & Ikeuchi, K. (2013). Camera spectral sensitivity and white balance estimation from sky images. *International Journal on Computer Vision (IJCV)*, 105(3), 187–204.
- Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. L. (1992). Surface perception in pictures. *Perception and Psychophysics*, 52(5), 487–496.
- Kolmogorov, V., & Zabih, R. (2004). What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 26(2), 147–159.
- Lalonde, J. F., Narasimhan, S. G., & Efros, A. A. (2010). What do the sun and the sky tell us about the camera? *International Journal on Computer Vision (IJCV)*, 88(1), 24–51.
- Lalonde, J. F., Efros, A. A., & Narasimhan, S. G. (2012). Estimating natural illumination from a single outdoor image. *International Journal on Computer Vision (IJCV)*, 98(2), 123–145.
- Lalonde, J. F., Asselin, L. P., Becirovski, J., Hold-Geoffroy, Y., Garon, M., Gardner, M. A., & Zhang, J. (2016). The Laval HDR sky database. <http://www.hdrdb.com>. Accessed 1 October 2015.

- Lee, J. Y., Matsushita, Y., Shi, B., & Kweon, I. S. (2013). Radiometric calibration by rank minimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 35(1), 114–156.
- Lombardi, S., & Nishino, K. (2012). Reflectance and illumination from a single image. In *Proceedings of European conference on computer vision (ECCV)*.
- Lu, F., Matsushita, Y., Sato, I., Okabe, T., & Sato, Y. (2013). Uncalibrated photometric stereo for unknown isotropic reflectances. In *Proceedings of IEEE conference on computer vision and pattern recognition (CVPR)*.
- Okutomi, M., & Kanade, T. (1993). A multiple-baseline stereo system. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 15(4), 353–363.
- Oxholm, G., & Nishino, K. (2012). Shape and reflectance from natural illumination. In *Proceedings of European conference on computer vision (ECCV)*.
- Preetham, A. J., Shirley, P., & Smits, B. (1999). A practical analytic model for daylight. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques (SIGGRAPH '99)* (pp. 91–100). ACM Press/Addison-Wesley Publishing Co.
- Ramamoorthi, R., & Hanrahan, P. (2001). On the relationship between radiance and irradiance: Determining the illumination from images of a convex lambertian object. *Journal of the Optical Society of America A (JOSA A)*, 18(10), 2448–2459.
- Saff, E., & Kuijlaars, A. B. J. (1997). Distributing many points on a sphere. *The Mathematical Intelligencer*, 19(1), 5–11.
- Scharstein, D., & Szeliski, R. (2002). A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal on Computer Vision (IJCV)*, 47(13), 7–42.
- Shan, Q., Adams, R., Curless, B., Furukawa, Y., & Seitz, S. M. (2013). The visual turing test for scene reconstruction. In *Proceedings of international conference on 3D vision*.
- Shen, F., Sunkavalli, K., Bonneel, N., Rusinkiewicz, S., Pfister, H., & Tong, X. (2014). Time-lapse photometric stereo and applications. *Computer Graphics Forum*, 33(7), 359–367.
- Shi, B., Matsushita, Y., Wei, Y., Xu, C., & Tan, P. (2010). Self-calibrating photometric stereo. In *Proceedings of IEEE conference on computer vision and pattern recognition (CVPR)*.
- Simchony, T., Chellappa, R., & Shao, M. (1990). Direct analytical methods for solving poisson equations in computer vision problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 12(5), 435–446.
- Stumpfel, J., Tchou, C., Jones, A., Hawkins, T., Wenger, A., & Debevec, P. (2004). Direct HDR capture of the sun and sky. In *Proceedings of the 3rd International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa (AFRIGRAPH)* (pp. 145–149). ACM.
- Subbarao, M., & Surya, G. (1994). Depth from defocus: A spatial domain approach. *International Journal on Computer Vision (IJCV)*, 13(3), 271–294.
- Sunkavalli, K., Matusik, W., Pfister, H., & Rusinkiewicz, S. (2007). Factored time-lapse video. *ACM Transaction on Graphics (TOG)*, 26(3), 101–111.
- Sunkavalli, K., Romeiro, F., Matusik, W., Zickler, T., & Pfister, H. (2008). What do color changes reveal about an outdoor scene? In *Proceedings of IEEE conference on computer vision and pattern recognition (CVPR)*.
- Turk, G., & Levoy, M. (1994). Zippered polygon meshes from range images. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques (SIGGRAPH '94)* (pp. 311–318). ACM.
- Woodham, R. (1980). Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1), 139–144.
- Yoon, K. J., & Kweon, I. S. (2006). Adaptive support-weight approach for correspondence search. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 28(4), 650–656.
- Yu, L. F., Yeung, S. K., Tai, Y. W., & Lin, S. (2013). Shading-based shape refinement of RGB-D images. In *Proceedings of IEEE conference on computer vision and pattern recognition (CVPR)*.
- Zhang, Q., Ye, M., Yang, R., Matsushita, Y., Wilburn, B., & Yu, H. (2012). Edge-preserving photometric stereo via depth fusion. In *Proceedings of IEEE conference on computer vision and pattern recognition (CVPR)*.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.