UNIVERSITY *of York*

This is a repository copy of *Height from Photometric Ratio with Model-based Light Source Selection*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/105332/

Version: Submitted Version

White Rose
university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Height from Photometric Ratio with Model-based Light Source Selection

William Smith, Fufu Fang

Department of Computer Science, University of York, York, UK

Abstract

In this paper, we present a photometric stereo algorithm for estimating surface height. We follow recent work that uses photometric ratios to obtain a linear formulation relating surface gradients and image intensity. Using smoothed finite difference approximations for the surface gradient, we are able to express surface height recovery as a linear least squares problem that is large but sparse. In order to make the method practically useful, we combine it with a model-based approach that excludes observations which deviate from the assumptions made by the image formation model. Despite its simplicity, we show that our algorithm provides surface height estimates of a high quality even for objects with highly non-Lambertian appearance. We evaluate the method on both synthetic images with ground truth and challenging real images that contain strong specular reflections and cast shadows.

Keywords: photometric stereo, surface integration, non-Lambertian reflectance, albedo estimation

1. Introduction

Photometric stereo has a long history in computer vision [1]. In recent years it has begun to find practical application in areas such as face recognition [2, 3], object capture [4, 5], medical imaging [6] and surface texture classification [7]. Photometric stereo uses the intensity of reflected light under varying illumination direction to infer the orientation and reflectance properties of a surface. Usually, the only reflectance property that is estimated is the diffuse albedo, although more exotic experimental setups allow estimation of additional reflectance properties such as specular albedo [8], surface roughness [9] and index of refraction [10].

The advantages of photometric stereo are well known: observations are dense (measurements are made at every pixel and resolution is limited only by the resolution of the camera); it can be applied to smooth surfaces devoid of matchable features; estimated surfaces can be relit since both shape and reflectance properties are estimated; and it is able to recover fine scale surface detail. However, surface orientation is only a 2.5D shape representation and the estimated normal field must be integrated in order to recover surface depth or used to refine a 3D mesh captured using other cues [11]. Often, modelling errors mean that the estimated surface normals are subject to low frequency bias leading to distortion in the 3D surface and low global accuracy. Also, photometric methods are usually much more demanding in data capture terms and also in their requirement for controlled conditions.

Despite sustained research effort on the topic, many methods are surprisingly still heavily based on the original approach proposed by Woodham [1]. One of the reasons for this is that Woodham's least squares solution "averages out" errors due to inaccurate assumptions or model inaccuracies. This leads to robust performance even in the presence of significant deviations from the assumptions of Lambertian reflectance with no shadows or specularities, or when there are errors in the light source directions and intensities.

A popular class of approaches that has arisen in the past decade are based on selecting a subset of observed intensities for each pixel [12, 13, 14, 15]. We refer to these as *selection-based approaches*. The idea is to exclude observations that are believed to deviate from the modelling assumptions made by the algorithm, for example by excluding shadowed pixels. This transforms the photometric stereo problem into one of labelling pixels according to photometric phenomena. These approaches are motivated by the fact that greater than three observations leads to redundancy which can be exploited by exclusion of noisy data. Almost all of these selection-based approaches rely on ad hoc heuris-

tics that require careful tuning of parameters.

1.1. Related work

The photometric stereo literature is large and here we review only the most relevant work, including selection-based techniques, alternative lighting models and methods that make use of photometric ratios.

The classical approach of Woodham [1] uses each observation to construct a linear equation based on the Lambertian equation (neglecting self shadows). The resulting system of equations is solved in a least squares sense. Hence, a minimum of three observations at each pixel are required to recover the surface normal scaled by the albedo. Woodham's algorithm does not take shadows and highlights into account, they are treated as noise. In practice, this is a strong assumption and it is difficult to image many objects without encountering shadows or highlights. Hence, three image photometric stereo is very sensitive to noise.

Using more than three images results in an overdetermined system and greater robustness to noise. However, systematic noise caused by specularities, shadows, light source attenuation or even a non-linear camera lead to systematic bias in the surface estimates. Since the system of equations at each pixel is independent, Woodham's algorithm also imposes no constraints on surface smoothness or integrability.

The classical approach to this problem is to incorporate smoothness or integrability priors into an energy term expressed as a function of the field of surface normals [16, 17, 18, 19, 20]. The drawback of such approaches is that these constraints are only softly satisfied and the result is dependent upon the weight assigned to each prior. There have been some attempts to incorporate photometric constraints into the surface integration process. For example, Chandraker et al. [21] use cast shadows to impose inequality constraints on the surface height estimates.

Selection-based photometric stereo. Barsky and Petrou [12] presented the first selection-based photometric stereo algorithm. Their idea was to use the minimal set of images for which there is redundancy (i.e. four images). A threshold on the reconstruction error using the four image photometric stereo result is used to label pixels where it is believed one observation in the input quadruple is corrupted by the presence of a shadow or specularity. The presence of a highlight is detected using one of two methods. One method is to check if the chromatic distance between the estimated diffuse albedo and the colour of the light sources is above a threshold.

Another method is to check if the recovered surface normal could cause specular reflection based on the viewing direction of the camera and the light source direction, i.e. it is close to the perfect specular configuration. If the quadruple does not contain a highlight, it must contain a shadow.

Sun et al. [13] suggested a hierarchical selection strategy for eliminating highlights and shadows. Their method requires images under six illumination conditions and can recover diffuse albedo and surface normals for non-Lambertian surfaces. They first order the observations for a pixel in descending order of brightness. They then check if the first observation contains a highlight, and if the fifth and the sixth observations contain shadows. Highlight and shadow detection is done in the same way as Barsky and Petrou [12].

Hernández et al. [15] proposed a three image photometric stereo algorithm that is robust to shadowing. Their method makes the assumption that for a surface patch, only one pixel from the input triplet is affected by shadow. They use a segmentation algorithm to detect shadowed regions. For the segmented shadowed regions, albedo and surface normal estimates are obtained using only the remaining two observations with integrability being used to resolve the resulting ambiguity.

A number of approaches can be viewed as outlier removal. Miyazaki et al. [22] compute surface normals from all possible observation triplets and then compute the median direction as a robust estimate of the surface normal. Yu et al. [23] propose an algorithm based on the maximum feasible subset framework. The idea is to select the maximum subset of observations that satisfy the Lambertian constraint. Mukaigawa et al. [24] used a random sampling approach to eliminate non-Lambertian observations. These approaches are related to our selection strategy, in the sense that we also detect and remove outliers. The difference is that our notion of outliers is based on deviation from predicted appearance using an initial (possibly naive) model whereas theirs are based on statistical analysis of the input data.

Ikehata et al. [14] posed the problem of selection as imposing sparsity on a Lambertian error matrix. However, strictly imposing sparsity leads to a non-convex optimisation problem. While not selection per se, Higo et al. [25] use the idea of "consensus" in photometric stereo. Namely, observations merely reduce the size of the solution space for a given surface normal. This allows them to use very general models, making relatively weak assumptions about reflectance and camera properties. However, their approach is data-heavy, requiring very large numbers of images to build a sufficient consensus on the correct normal direction. Moreover, cast

shadows cannot be modelled and act as noise in the consensus process.

Rather than make per pixel observation selections, recently, Argyriou et al. [26] tackled the problem of choosing the *globally* optimal configuration of light source directions. The idea is that for a class of objects, e.g. faces, there is likely to be a configuration of light sources that minimises the number of shadowed observations over the whole image set on average. They pose the problem of searching for the lighting configuration as one of sparse optimisation.

Alternative illumination models. When a surface is lit by a point source (and assuming no inter-reflections), shadowing is a binary function, i.e. a point on the surface is either lit or it is shadowed. Under a continuous field of illumination, this is not the case as some portion of the illumination may be visible and observations need not be selected and excluded in a binary fashion. The analogue of shadowing under continuous illumination is occlusion, where part of the local upper hemisphere is occluded by other parts of the surface.

The special case of a uniform (i.e. ambient) continuous field of illumination was studied by Prados et al. [27] who demonstrated that shading under such illumination can be ambiguous. Allowing arbitrary illumination but neglecting occlusions, the most general approach is due to Basri et al. [28]. They use a spherical harmonic model of Lambertian appearance to develop a photometric stereo algorithm for complex illumination environments. More recently, such a setting has been considered in a single image, shape-from-shading setting. For example, Huang and Smith [29] provide a linear approach for shape recovery under a first order spherical harmonic lighting model.

Rather than considering natural illumination environments, Ma et al. [8] developed a variant of photometric stereo which operates under continuous linear gradient illumination fields. They pose surface normal recovery as estimation of the centre of mass of the reflectance lobe. To achieve this, they compute the first moment of the reflectance function by integration with a linear gradient in each direction of the coordinate system. Although they do not model occlusions, their approach degrades gracefully with increasing occlusion. In fact, their method estimates the "bent normal" direction often used in graphics (the mean unoccluded direction).

Another relaxation of assumptions is to allow near-field (i.e. local) lights. Here, the attenuation with distance must be modelled. Recently, Papadhimitri and Favaro[30] showed that uncalibrated near-field photometric stereo suffers from reduced ambiguities in comparison to the distant case.

Photometric Ratios. The idea of using ratios between photometric measurements to aid surface analysis was first proposed by Davis and Soderblom [31]. In contrast to most later work, they used intensity measurements from a single image. By finding profiles with assumed equal topographic and albedo variations, they take ratios to cancel out surface reflectivities.

The use of ratios between images under two different illumination conditions was first proposed by McEwen [32] in the context of remote sensing. Lee and Brady [33] were the first to exploit the resulting constraint in a computer vision context. Specifically, they took ratios between observations from different viewpoints leading to an equation in one of the components of the surface gradient. They use this to estimate depth which they subsequently combine with depth estimates from binocular stereo. This was followed by another hybrid photometric/geometric shape estimation method from Wolff and Angelopoulou [34]. Instead of computing depth from photometric measurements, they use photometric ratios to derive invariants that improve the robustness of stereo matching. In a similar vein, both Lee and Kuo [35] and Wu et al. [36] use photometric ratios to eliminate albedo from their image formation model. However, they do not exploit the linearisation in depth and instead estimate surface normals.

Like Davis and Soderblom [31], Zhao and Chellappa [37] also use ratios extracted from a single image. In this case it is between symmetric points on bilaterally symmetric surfaces (specifically, faces). As for the two image case, this leads to an expression that is independent of albedo and linear in terms of surface gradients. This allows them to obtain a ratio from a single image and hence to use the constraint for single image shape-from-shading. Wöhler and Hafezi [38] introduce a ratio-based error term into a classical variational approach to photometric surface estimation. This leads to an iterative scheme in which surface gradients are recovered independently of the unknown albedo.

The advantage of photometric ratios is that they lead to linear equations in depth. Other linearisations of shading-based error terms have also been used in single image shape-from-shading. This was first proposed by Pentland [39] who used a local Taylor expansion of the reflectance function to provide linear equations in the surface gradients. This was extended by Tsai and Shah [40] who solve directly for surface height by substituting in numerical approximations to the surface gradients. This leads to a highly efficient and practical algorithm. An alternative manipulation of the Lambertian

reflectance equation was proposed in Ecker and Jepson's [41] polynomial shape-from-shading algorithm. Here, the non-linear normalisation factor in the surface normal is removed by working with the square of the intensity. This leads to an optimisation problem containing the squares of the surface gradients which they solve using homotopy solvers or via SDP relaxations.

Recently, there has been renewed interest in exploiting ratios for photometric stereo, led by Mecca and coworkers [42, 43, 44, 45, 46]. Their line of work uses photometric ratios to construct PDEs for which they can prove well-posedness and solve using upwind or semi-Lagrangian schemes. Initially they considered the classical photometric stereo set up of orthographic viewer and distant point light sources [43], proving the uniqueness of the two image case in the presence of boundary conditions. This was extended to a practically useful situation [42] by considering cast shadows in three image photometric stereo. In shadow regions, a shadowed observation is given zero weight (it is assumed that at least two observations are unshadowed), allowing the surface to be recovered in spite of the missing data. However, the shadow weights assume that a cast shadow has zero intensity so this does not represent a robust selection strategy. The same weighting strategy was extended to perspective viewing [44]. More recently, the case of near-field illumination [45] has been considered.

While these methods target mathematical well-posedness and minimal solutions, they have a number of practical drawbacks. First, they assume a Lambertian reflectance model but do not incorporate a robust observation selection procedure. For this reason, the methods are evaluated on either synthetic data or real images that do not contain shadows or significant specularities. Second, since all of these methods are variational, they require boundary conditions to be known (or estimated using photometric stereo with stricter assumptions). Third, they consider either minimal (e.g. two or three image) setups or place restrictions on the positioning of lights (e.g. requiring the illuminant to be at the optical centre [45]).

There have been two attempts to use image ratios for photometric stereo in a non-Lambertian setting. Chandraker et al. [47] propose a differential photometric stereo framework in which ratios with a frontally or ambiently lit image are used to cancel dependence on albedo. Their approach is quite general and works with any isotropic reflectance model. The state of the art in this direction is recent work in which Mecca et al. [46] work with very general assumptions (arbitrary attenuation models, general light positions, non-Lambertian reflectance) to achieve very high quality reconstructions

from specular reflections.

1.2. Contributions

In this paper we make a number of novel contributions that lead to a practically useful and robust photometric stereo algorithm. Our approach combines the robustness of selection-based approaches with the attractive properties of solving in the surface height domain via photometric ratios. Our first contribution is to propose a model-based approach to observation selection. The idea is to obtain an initial estimate of the surface normals and albedo using any existing photometric stereo algorithm. This estimate is used to predict the appearance of the object under the illumination conditions originally observed. Noisy observations that deviate from the model assumptions can then be automatically detected and removed via a robust estimate of the noise variance.

Our second contribution is to incorporate this selection strategy into a photometric stereo algorithm based on photometric ratios. This approach allows us to solve directly for surface height, whilst still only requiring the solution of a large (sparse) linear system of equations. We are not the first to observe that photometric ratios lead to linear equations in depth, nor are we the first to exploit them for photometric stereo. We use the same orthographic, Lambertian ratio model as in [43]. However, we believe that this is the first work to use photometric ratios in the context of energy minimisation.

Taken together, these contributions lead to an approach that we believe is the first practically useful height-from-photometric ratio algorithm, i.e. which is applicable to images of real objects exhibiting non-Lambertian appearance. This enables us to recover high quality depth maps from datasets with an arbitrary number of images in a manner that is highly tolerant to corruption by shadows and specularities. Compared to previous attempts to use photometric ratios for height recovery, our approach is simpler and more efficient, being based on discrete approximations to the surface gradients and posing the problem as a large system of linear equations. In addition, we propose a robust method for computing per-pixel albedo from the valid observations using simple linear regression. This combination of robust performance on real world, noisy images with computational efficiency provides a method of real practical utility. We demonstrate this in our experimental evaluation by including results on objects with highly non-Lambertian appearance and severe shadowing. Moreover, we compare our approach to state of the art algorithms and show a quantitative improvement in perfor-

mance, particularly with respect to the accuracy of the recovered height map.

1.3. Paper organisation

In Section 2 we begin by reviewing relevant theory from the photometric stereo literature. In particular, the least squares approach and the notion of photometric ratios, leading to linear equations in the surface gradients. In Section 3 we present our model-based selection strategy for excluding unreliable observations. In Section 4 we show how to use the photometric ratio equations to build a large, sparse linear system of equations for photometric stereo. In Section 5 we present experimental results including quantitative comparison to the state-of-the-art and qualitative reconstructions on challenging images containing both cast shadows and specular reflections.

2. Preliminaries

We assume that a surface with smooth (i.e. C^2 continuous) height function $z(x, y)$ is viewed orthographically. We observe the surface K times under varying illumination and in the kth image the observed image intensity is denoted $i_k(x, y)$. In each image, we assume that the surface is lit by a distant point light source with direction and intensity given by the vector $\mathbf{s}_k \in \mathbb{R}^3$ for image $k = 1 \ldots K$.

A diffuse Lambertian surface reflects incident light equally in all directions. Hence, the observed intensity depends only on the foreshortening of the light source with respect to the surface normal $\mathbf{n}$:

$$i = \max(0, \rho \mathbf{n} \cdot \mathbf{s}), \tag{1}$$

where $\rho \in [0, 1]$ is the diffuse albedo (intrinsic reflectivity of the surface). Clamping to zero simulates the effect of self shadow (i.e. where the angle between surface normal and light source exceeds $90°$). Hence, an observation of a Lambertian surface provides an equation that is linear in the scaled surface normal ($\rho \mathbf{n}$) for incident angles between $0°$ and $90°$.

2.1. Least squares photometric stereo

This linear relationship was exploited in the original photometric stereo algorithm of Woodham [1]. His idea was to use observations with calibrated light sources (i.e. where $\mathbf{s}_k$ are known) to form a system of linear equations that is solved in a least squares sense. Since the scaled surface normal has three degrees of freedom, a minimum of three observations are required. Moreover, the $\mathbf{s}_k$ must be linearly independent. If more than three observations are available, the system is overdetermined and the least squares solution becomes increasingly robust to Gaussian noise.

The observed intensities for a pixel (x, y) observed under K different illumination directions can be stacked to form the observation vector $\mathbf{i}(x, y) \in \mathbb{R}^K$. The light source vectors $\mathbf{s}_k$ can also be stacked to form the illumination matrix $\mathbf{S} \in \mathbb{R}^{K \times 3}$. The scaled surface normal vector can be found by solving the linear least squares problem:

$$\mathbf{m}(x, y) = \operatorname*{argmin}_{\mathbf{x}} \|\mathbf{S}\mathbf{x} - \mathbf{i}(x, y)\|^2. \tag{2}$$

The estimated albedo is given by the length of the vector: $\rho(x, y) = \|\mathbf{m}(x, y)\|$ and the estimated surface normal by the normalised vector: $\mathbf{n}(x, y) = \frac{\mathbf{m}(x,y)}{\rho(x,y)}$. Such a linear system is built and solved independently for each pixel.

Woodham's approach is optimal in the sense that it minimises the average reillumination error under the Lambertian model. However, the accuracy of the estimated surface normals and diffuse albedo values breaks down in the presence of systematic deviations from the model. There are two particularly problematic such deviations. The first are cast shadows. These are a non-local effect where remote parts of the surface block the light source from reaching a point on the surface. The result is an intensity close to zero (inter-reflections, ambient light and noise mean that the intensity is rarely exactly zero). The second is additive specular reflection. This is a local effect for shiny surfaces occurring when the viewer, surface normal and light source are close to the perfect specular configuration. This leads to an intensity much greater than predicted by the Lambertian model. Since the approach is based on linear least squares and is therefore not robust, the result is strongly affected by these outlying observations. Finally, using all observations also neglects the effect of self-shadowing, i.e. it ignores the fact that Equation 1 is only piecewise linear.

2.2. Photometric Ratio

Classical photometric stereo solves for the surface normal at each pixel independently. There are two disadvantages to this approach. First, the recovered normals may not correspond to a physically-realisable surface, that is to say the surface normals may not satisfy the integrability condition. Second, if the ultimate goal is to recover surface height then the normals must be integrated into a surface. Performing this as a two step process means that errors accumulate over the two steps.

An alternative is to solve directly for surface height. To do this, the problem must be posed in terms of surface height. This idea is not new and has recently gained popularity with a number of methods [42, 43, 44, 45, 46, 47] exploiting the simplified relationship between surface height and image intensity that arises when photometric ratios are taken.

The surface normals and surface height are related via the surface gradients, $p = \partial_x z$ and $q = \partial_y z$:

$$\mathbf{n} = \frac{[-p \ -q \ 1]^T}{\sqrt{p^2 + q^2 + 1}}. \qquad (3)$$

Substituting surface gradients into the Lambertian photometric stereo equations above leads to a non-linear system of equations. However, consider two non-shadowed observations under illumination directions $\mathbf{s}$ and $\mathbf{t}$:

$$i_1 = \rho \frac{-s_1 p - s_2 q + s_3}{\sqrt{p^2 + q^2 + 1}}, \quad i_2 = \rho \frac{-t_1 p - t_2 q + t_3}{\sqrt{p^2 + q^2 + 1}}. \quad (4)$$

By taking ratios between the two images, the non-linear term in the surface gradients cancels out, as does the albedo:

$$\frac{i_1}{i_2} = \frac{-s_1 p - s_2 q + s_3}{-t_1 p - t_2 q + t_3}. \qquad (5)$$

Rearranging leads to a linear equation in the surface gradients that is independent of albedo:

$$(i_2 s_1 - i_1 t_1)p + (i_2 s_2 - i_1 t_2)q = i_2 s_3 - i_1 t_3. \qquad (6)$$

This formulation was derived by Mecca and Falcone [43]. In a photometric stereo setting, each pair of images provides one such equation per pixel. Given the advantages over Woodham's original expression in terms of surface normals, it is surprising that this formulation has not been adopted as standard in the photometric stereo literature. Later, by taking finite difference approximations to the surface gradient, we use this expression to pose photometric stereo as a large sparse system of linear equations in terms of the unknown surface heights.

3. Model-based observation selection

The goal of selection-based photometric stereo is to detect and exclude observations that are believed to deviate from the modelling assumptions. We take a model-based approach to this problem. By this, we mean that we use a model of image formation (namely Lambertian with self shadowing and varying albedo) to predict appearance. We use errors between predicted

and observed appearance to perform a statistical test to separate inliers (observations fitting our model) from outliers (those deviating from the model, e.g. shadows or specularities).

We begin by assuming that we have to hand an initial estimate of the surface normals and albedo. We denote these $\tilde{\mathbf{n}}(x, y)$ and $\tilde{\rho}(x, y)$ respectively. In our experimental results, these are either provided by Woodham's least squares approach applied to all observations or, in the case of light stage data, estimated using spherical gradient photometric stereo [8] on a set of images under spherical gradient illumination.

Observations that do adhere to our modelling assumptions are expected to differ from the predicted intensity by only an additive noise term:

$$\max(0, \tilde{\rho}(x, y)\tilde{\mathbf{n}}(x, y) \cdot \mathbf{s}_k) = i_k(x, y) + \epsilon, \qquad (7)$$

where $\epsilon \sim \text{Norm}[0, \sigma^2]$ is the noise term which we assume is normally distributed with zero mean and variance σ^2 and accounts for error in $\tilde{\mathbf{n}}$ and $\tilde{\rho}$. Under this assumption, we expect the difference between observed and predicted intensities to follow a normal distribution:

$$\max(0, \tilde{\rho}(x, y)\tilde{\mathbf{n}}(x, y) \cdot \mathbf{s}_k) - i_k(x, y) \sim \text{Norm}[0, \sigma_k^2]. \quad (8)$$

Observations that deviate from our modelling assumption, i.e. those that are shadowed or are specular highlights, will not differ from predicted intensity by additive, zero mean noise. Our idea is to estimate the variance of the noise and use this to perform a Z-test to exclude unreliable observations. We begin by robustly estimating the standard deviation of the noise for the kth image via the median absolute deviation:

$$\text{MAD}_k = \underset{x,y}{\text{median}}(|\max(0, \tilde{\rho}(x, y)\tilde{\mathbf{n}}(x, y) \cdot \mathbf{s}_k) - i_k(x, y)|).$$
$$(9)$$

Note that since we expect the differences to come from a zero mean distribution, we do not need to subtract an estimated mean value. Our estimate of the standard deviation is now given by:

$$\sigma_k \approx 1.4826 \, \text{MAD}_k. \qquad (10)$$

For each observation, we now compute a Z-score:

$$Z_{x,y,k} = \frac{\max(0, \tilde{\rho}(x, y)\tilde{\mathbf{n}}(x, y) \cdot \mathbf{s}_k) - i_k(x, y)}{\sigma_k}. \qquad (11)$$

Finally, observations whose absolute Z-score is greater than a threshold are excluded as unreliable. In addition to the Z-score criterion, we also make sure that selected observations are not self-shadowed. Since

these values lie outside the linear region of the Lambertian reflectance function, they do not adhere to the assumptions made in Equation 6 and cannot be used in the height-from-photometric ratio method in the following section. Self-shadowed pixels are those where $\tilde{\mathbf{n}}(x, y) \cdot \mathbf{s}_k \leq 0$.

The selection strategy above does not guarantee that there will be any valid observations for a pixel. This may lead to the linear system being under constrained during height estimation. Hence, for such pixels we force the selection of a minimum of three observations. To do this, we sort the observations by their absolute Z-score and pick values in ascending order, excluding self-shadowed observations and those already chosen. We continue this process until all pixels have at least three observations.

4. Height from photometric ratio

Suppose that the set of valid observations (i.e. those chosen by our selection procedure) at pixel (x, y) is denoted $\mathcal{V}(x, y)$. We wish to solve the following minimisation problem:

$$\min_z \sum_{x,y} \sum_{j,k \in \mathcal{V}(x,y)} \begin{aligned} &\| \left(i_k(x, y) s_{j,1} - i_j(x, y) s_{k,1} \right) \partial_x z(x, y) + \\ &\left(i_k(x, y) s_{j,2} - i_j(x, y) s_{k,2} \right) \partial_y z(x, y) - \\ &i_k(x, y) s_{j,3} + i_j(x, y) s_{k,3} \|^2. \end{aligned}$$
$$(12)$$

The surface is observed K times in total. This means that there may be a maximum of $K(K + 1)/2$ unique and valid pairs of observations for a pixel. In practice, we use only K pairs obtained by a single cycle through the valid observations for a pixel.

The surface gradients can be approximated numerically from the discretised surface height function using finite differences. The central difference approximation to the first derivative in the horizontal direction can be obtained by convolution with the kernel $[-1\ 0\ 1]$ (assuming unit spacing of the pixel grid). Note that higher order approximations to the surface derivatives could in principle be used [48] with a potential improvement in accuracy for surfaces in which an assumption of local planarity is poor.

To reduce sensitivity to noise and improve robustness, the depth values are first smoothed using a centre-weighted kernel approximating a Gaussian with standard deviation 0.6. By associativity of the convolution operator we can pre-convolve the finite difference and smoothing kernels leading to the following kernels for computing the surface gradients in the horizontal and vertical directions:

$$\partial_x z \approx z * \frac{1}{12} \begin{bmatrix} -1 & 0 & 1 \\ -4 & 0 & 4 \\ -1 & 0 & 1 \end{bmatrix}, \ \partial_y z \approx z * \frac{1}{12} \begin{bmatrix} 1 & 4 & 1 \\ 0 & 0 & 0 \\ -1 & 4 & 1 \end{bmatrix}.$$
$$(13)$$

At boundary pixels, or when using a binary mask to exclude background pixels, not all neighbours may be available for a given pixel. In this case, we use unsmoothed central differences (where both horizontal or vertical neighbours are available) or, where only a single neighbour is available, single forward/backward differences. Substituting these expressions into Equation 6 therefore leads to a linear equation with between 3 and 8 unknown values of z (depending on which combination of numerical gradient approximations are used). Hence, we can solve Equation 12 using linear least squares.

Suppose our image contains N pixels of interest (i.e. pixels that are part of the foreground mask) then we write the unknown height values as the vector $\mathbf{z} \in \mathbb{R}^N$. Substituting the discrete gradient approximations into Equation 12 and rewriting in matrix form, we can express our objective as:

$$\min_{\mathbf{z}} \|\mathbf{Az} - \mathbf{b}\|^2.$$
$$(14)$$

Differentiating and setting equal to zero leads to the linear system

$$\mathbf{Az} = \mathbf{b},$$
$$(15)$$

where $\mathbf{A}$ has N columns and a maximum of NK rows. Each row of $\mathbf{A}$ evaluates the left hand side of Equation 6 by selecting the appropriate elements from the vector of height values to approximate the appropriate surface gradients. The vector $\mathbf{b}$ contains the corresponding right hand side constant values. Although $\mathbf{A}$ is very large, it is also sparse. Each row has at most 8 non-zero values. In order to resolve the unknown constant of integration (i.e. applying an arbitrary offset to z does not affect its orthographic images), we fix the height of one pixel to zero.

Efficient methods for solving large sparse systems such as Equation 15 are available and we use the method of Davis [49]. This starts by computing the QR decomposition of $\mathbf{AP} = \mathbf{QR}$, where $\mathbf{P}$ is a fill-reducing permutation for $\mathbf{A}$. For efficiency, we avoid calculating $\mathbf{Q}$ explicitly, and compute $\mathbf{c} = \mathbf{Q}^T \mathbf{b}$ directly during factorisation. Also, we compute the "economy-size" $\mathbf{R}$ with only N rows. To find $\mathbf{z}$ we then solve the smaller linear least squares problem $\mathbf{Rz} = \mathbf{c}$.

4.1. Albedo Estimation

Equation 6 is independent of albedo. Hence, the approach described above recovers surface height without
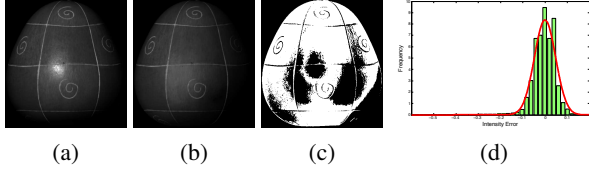
Figure 1: Mask generation example. An input image (a) is compared to its predicted Lambertian appearance (b). A normal distribution is robustly fitted to the distribution of errors (d) and a Z-test used to generate a binary observation mask (c).

providing any information about albedo. Computing albedo amounts to a simple inverse rendering problem. However, as above, we must only use selected observations for which we believe our modelling assumptions are valid.

We begin by computing the surface normals of the recovered surface height function. To do so, we apply Equation 13 to the estimated surface height function (using appropriate alternatives for pixels with fewer than 8 neighbours) to approximate the surface gradients. We substitute these estimates into Equation 3 to provide our estimated field of surface normals $\mathbf{n}(x, y)$.

We seek the albedo estimate that minimises the squared difference between predicted and observed intensity for those observations that were selected. Hence, we wish to solve the following minimisation problem at each pixel:

$$\min_{\rho(x,y)} \sum_{k \in \mathcal{V}(x,y)} \|\rho(x, y)\mathbf{n}(x, y) \cdot \mathbf{s}_k - i_k(x, y)\|^2. \quad (16)$$

The solution to this problem is obtained in closed form by simple linear regression:

$$\rho(x, y) = \frac{\sum_{k \in \mathcal{V}(x,y)} \mathbf{n}(x, y) \cdot \mathbf{s}_k i_k(x, y)}{\sum_{k \in \mathcal{V}(x,y)} (\mathbf{n}(x, y) \cdot \mathbf{s}_k)^2}. \quad (17)$$

5. Experiments

In this section we present results of applying our algorithm to three datasets, two previously available and one collected for this study. We experiment with two different algorithms for providing the initial surface normal and albedo estimates: Woodham's least squares approach and the recent spherical gradient photometric stereo algorithm of Ma et al. [8].

5.1. Selection strategy

We begin in Figure 1 by demonstrating our process of observation selection. We compute the difference between an input image (a) and a Lambertian rendering
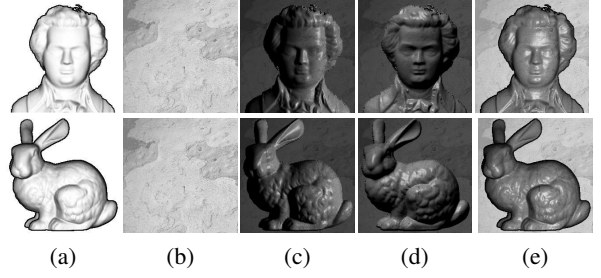

Figure 2: Synthetic data used in quantitative evaluation. (a) Lambertian rendering of ground truth shape. (b) Albedo map. (c)-(e) 3 of the 40 input images.

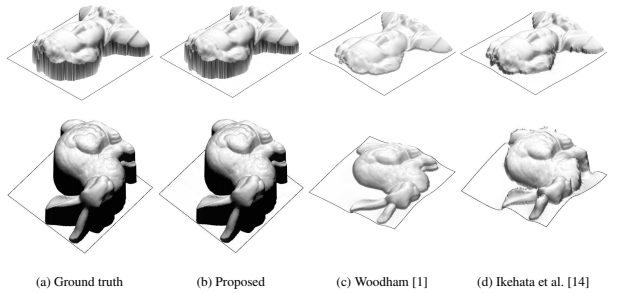(a) Ground truth   (b) Proposed   (c) Woodham [1]   (d) Ikehata et al. [14]

Figure 3: Comparison of estimated surface height maps (b-d) against ground truth (a). In the case of (c) and (d), surface height is obtained by integrating estimated normals using [48].

(b) under the same illumination conditions, computed using surface normal and albedo estimates provided by a naive algorithm such as Woodham's. The distribution of errors is shown as a histogram in (d). It is clear that most of the errors can be approximated by a normal distribution. We show our robust estimate of this distribution in red. Outlying errors can be identified as not adhering to our model assumptions. In this example, the large negative tail in the histogram corresponds to the specular highlight (in which the predicted intensity is much smaller than the observed). We perform a Z-test using this distribution and also label pixels computed as being self shadows in the predicted image. This allows us to generate an observation mask (c) which is used subsequently to select pairs of intensities for photometric shape estimation.

5.2. Quantitative comparison

In order to provide quantitative evaluation, we use two synthetic datasets. From a ground truth depth map, we compute surface normals (Figure 2(a)). We also generate a textured albedo map (Figure 2(b)). We then render 40 images using the Blinn-Phong reflectance model with a shininess parameter of 75 and the same lighting directions as in [14]. This provides images with strong

Method	Bunny							Mozart						
No. images	10	15	20	25	30	35	40	10	15	20	25	30	35	40
[1]	53.72 ± 1.67	52.27 ± 1.68	52.70 ± 0.96	52.79 ± 0.85	52.69 ± 0.53	52.72 ± 0.24	52.65	20.60 ± 0.88	20.14 ± 1.11	20.29 ± 0.49	20.10 ± 0.48	20.02 ± 0.34	19.76 ± 0.18	19.83
	5.19 ± 3.33	4.44 ± 2.89	3.01 ± 1.08	2.78 ± 0.96	2.99 ± 1.04	3.40 ± 0.72	3.25	4.50 ± 2.32	3.62 ± 1.49	3.95 ± 1.17	3.27 ± 0.96	3.33 ± 0.81	3.29 ± 0.62	3.18
[14]	47.46 ± 3.44	46.25 ± 3.32	45.45 ± 3.86	46.10 ± 2.76	46.41 ± 1.37	47.24 ± 1.08	47.41	16.48 ± 1.60	16.02 ± 1.79	16.80 ± 1.30	16.09 ± 1.33	16.24 ± 1.00	16.67 ± 0.73	16.57
	2.21 ± 3.96	0.82 ± 1.85	0.80 ± 0.76	**0.05 ± 0.46**	**0.40 ± 0.51**	**0.20 ± 0.33**	**0.18**	4.97 ± 4.77	**0.96 ± 2.01**	**1.01 ± 1.89**	1.48 ± 1.34	**0.79 ± 1.44**	**1.09 ± 0.90**	**1.13**
HfPR-[1]	45.31 ± 6.40	46.07 ± 2.76	48.77 ± 2.57	48.81 ± 2.05	50.01 ± 2.03	49.21 ± 1.18	49.67	16.59 ± 2.16	16.04 ± 1.98	16.28 ± 1.98	16.08 ± 1.99	14.98 ± 1.63	15.61 ± 2.20	15.62
	6.08 ± 3.87	4.16 ± 3.37	3.93 ± 1.50	2.95 ± 0.91	3.74 ± 0.98	3.66 ± 0.86	3.07	7.11 ± 3.61	6.27 ± 2.72	5.94 ± 2.55	6.33 ± 2.62	6.18 ± 1.94	5.57 ± 1.43	4.49
HfPR-[14]	49.93 ± 2.12e7	43.57 ± 40.21	42.88 ± 7.07	45.16 ± 5.53	42.68 ± 3.16	40.90 ± 2.84	39.69	15.67 ± 2.00	13.01 ± 2.18	14.91 ± 1.68	12.58 ± 2.40	12.18 ± 2.10	14.06 ± 3.29	11.48
	5.84 ± 2.75	1.75 ± 1.72	1.95 ± 0.70	1.43 ± 0.65	1.72 ± 0.42	1.69 ± 0.49	1.58	4.88 ± 3.46	4.14 ± 1.59	5.03 ± 2.02	4.90 ± 1.71	4.87 ± 1.76	4.33 ± 1.15	4.08
HfPR-[50]	**22.49 ± 11.13**	**0.85 ± 7.16**	**0.82 ± 0.66**	**0.54 ± 0.33**	**0.76 ± 0.29**	**0.76 ± 0.20**	**0.56**	**4.31 ± 4.35**	**4.26 ± 2.87**	**4.60 ± 1.90**	**5.33 ± 1.90**	**4.92 ± 1.83**	**4.55 ± 1.66**	**4.31**
	2.84 ± 1.12	**0.70 ± 0.89**	**0.61 ± 0.21**	0.49 ± 0.12	0.50 ± 0.09	0.50 ± 0.06	0.45	**3.89 ± 2.03**	3.37 ± 1.33	2.58 ± 0.85	3.19 ± 0.85	2.99 ± 0.63	3.03 ± 0.58	2.81

Table 1: Quantitative results for varying numbers of input images. For each method and number of input images we show the RMSE height map error in pixels (first row) followed by median angular error of surface normals in degrees (second row). For fewer than 40 images, we repeat the experiment multiple times, each time using a randomly selected subset of the appropriate size. We show the median error and median absolute deviation over all trials.

Method	Bunny							Mozart						
Noise	5%	10%	15%	20%	25%	30%	50%	5%	10%	15%	20%	25%	30%	50%
[1]	52.64 ± 0.00	52.63 ± 0.01	52.62 ± 0.01	52.62 ± 0.01	52.60 ± 0.01	52.59 ± 0.01	52.54 ± 0.02	19.83 ± 0.00	19.82 ± 0.00	19.81 ± 0.00	19.81 ± 0.01	19.80 ± 0.01	19.80 ± 0.01	19.77 ± 0.01
	3.50 ± 0.00	3.70 ± 0.00	3.88 ± 0.01	4.04 ± 0.01	4.17 ± 0.01	4.30 ± 0.01	4.76 ± 0.01	3.32 ± 0.00	3.46 ± 0.00	3.59 ± 0.01	3.72 ± 0.01	3.84 ± 0.01	3.95 ± 0.01	4.38 ± 0.01
[14]	46.71 ± 0.20	46.20 ± 0.15	46.07 ± 0.11	46.01 ± 0.25	46.13 ± 0.14	46.34 ± 0.09	**47.33 ± 0.13**	16.40 ± 0.04	16.32 ± 0.04	16.24 ± 0.04	16.16 ± 0.06	16.18 ± 0.06	16.20 ± 0.12	16.34 ± 0.05
	0.14 ± 0.01	**0.09 ± 0.01**	**0.11 ± 0.01**	**0.19 ± 0.01**	**0.34 ± 0.01**	**0.50 ± 0.01**	**1.49 ± 0.01**	**0.63 ± 0.00**	**0.61 ± 0.00**	**0.60 ± 0.00**	**0.62 ± 0.00**	**0.66 ± 0.00**	**0.74 ± 0.00**	**1.78 ± 0.02**
HfPR-[1]	50.25 ± 1.02	50.88 ± 1.18	50.89 ± 0.73	51.12 ± 1.33	50.84 ± 0.91	50.58 ± 1.08	52.11 ± 0.93	16.44 ± 1.14	16.01 ± 1.35	16.58 ± 1.35	17.31 ± 0.88	17.34 ± 0.94	17.60 ± 1.34	18.06 ± 0.82
	3.32 ± 1.02	2.98 ± 0.74	3.35 ± 0.56	3.94 ± 0.61	4.70 ± 0.72	5.13 ± 0.75	6.16 ± 1.05	6.15 ± 1.39	5.99 ± 1.44	5.09 ± 1.52	6.12 ± 1.10	5.32 ± 1.51	5.82 ± 1.44	5.75 ± 0.97
HfPR-[14]	41.16 ± 1.96	42.56 ± 1.47	44.31 ± 1.29	45.31 ± 0.91	46.05 ± 0.80	46.48 ± 0.52	49.02 ± 0.43	12.67 ± 1.04	12.59 ± 1.38	13.44 ± 1.15	13.58 ± 0.72	13.89 ± 0.68	13.82 ± 0.90	15.83 ± 0.65
	1.36 ± 0.15	1.88 ± 0.22	2.08 ± 0.17	2.50 ± 0.16	2.87 ± 0.18	3.41 ± 0.19	4.81 ± 0.39	3.90 ± 0.86	3.94 ± 0.52	4.38 ± 0.56	4.64 ± 0.55	4.84 ± 0.64	5.06 ± 0.64	5.13 ± 0.50
HfPR-[50]	**12.54 ± 1.48**	**22.61 ± 1.90**	**31.78 ± 1.28**	**38.00 ± 0.99**	**41.24 ± 0.92**	**44.25 ± 0.91**	48.81 ± 0.37	**5.86 ± 1.29**	**7.44 ± 0.97**	**8.55 ± 1.03**	**9.65 ± 0.70**	**10.57 ± 0.81**	**11.63 ± 1.02**	**15.01 ± 0.58**
	0.86 ± 0.05	1.23 ± 0.09	1.58 ± 0.12	2.16 ± 0.22	2.61 ± 0.23	3.11 ± 0.23	4.87 ± 0.36	3.66 ± 0.55	3.96 ± 0.61	4.36 ± 0.70	4.85 ± 0.59	4.63 ± 0.77	5.08 ± 0.76	5.23 ± 0.47

Table 2: Quantitative results for varying levels of additive noise. Errors are presented in the same way as in Table 1.

specular reflections. We then apply our algorithm, providing surface height and albedo estimates. We also apply the algorithms of Woodham [1] and Ikehata et al. [14], providing surface normal estimates which we integrate into a surface height map using a state-of-the-art surface integration method [48].

In Figure 3 we show estimated surfaces in a novel view. Here, we use all 40 images. The surface recovered by the proposed method is almost indistinguishable from ground truth, whereas the integrated surfaces from [1] and [14] are significantly distorted (in particular, the large gradients at the edge of the objects have been overly smoothed).

Next, in Table 1 we provide quantitative results when the number of input images is varied between 10 and 40. We compute two errors. First, the median angular difference between ground truth and estimated surface

normals. For the proposed method, surface normals are computed from the estimated height map. Second, the root mean square error between ground truth and estimated surface height. To account for the constant of integration, we subtract the mean value from all height maps before computing this error. We experiment with three different guide normals for our selection method: those provided by the least squares method of Woodham (referred to as HfPR-[1]), Ikehata et al. (referred to as HfPR-[14]) and a robust version of the least squares method using RANSAC (referred to as HfPR-[50]).

The state of the art method [14] is optimal in terms of the accuracy of the estimated surface normals in all but three cases. However, when the normals are integrated into a height map, the accuracy of the recovered surface is far inferior to estimating surface height directly using our method HfPR-[50]. In many cases, even when we

use Woodham guide normals (HfPR-[1]) we still obtain a more accurate surface than [14].

In Table 2 we repeat the 40 image experiment but introduce additive noise in the same manner as [14]. We observe the same effect that, even in the presence of noise, our method yields the most accurate height map in all but one case.

5.3. Qualitative evaluation

In this section we provide qualitative results on three different datasets exhibiting different sources of systematic noise. For each dataset (in Figures 4, 6 and 7), we show an example input image in the first column, a rotated view of the recovered surface in the second column, a colourised representation ($R = \frac{n_x+1}{2}$, $G = \frac{n_y+1}{2}$ and $B = \frac{n_z+1}{2}$) of the surface normals of the recovered surface in the third column and the estimated albedo map in the fourth column.

Shadowed images. We begin in Figure 4 by showing results on the Harvard dataset [51]. This dataset was collected for the purpose of evaluating shape-from-shading algorithms. For this reason, the objects have approximately uniform albedo and reflectance is primarily diffuse. However, the images include strong cast shadows that lead to artefacts in the output of naive algorithms. Each object is captured under 20 different illumination conditions.

The recovered surfaces are both globally accurate whilst also containing finescale detail. We emphasise that our method recovers these height maps directly, i.e. the step of integrating normals into a height map is not required. Estimated albedo is uniform for much of the surface, as would be expected. Bright patches correspond to concave regions where inter-reflections raise the observed intensity above that predicted by a model with only direct lighting.

Our selection strategy ensures that shadowed observations have not corrupted the surface in shadowed regions. In Figure 5 we show a region in which many of the input images contain cast shadows. Woodham's algorithm (re-rendered surface normals shown top right) leads to corruption of the surface normals. Shadowed regions are being explained by adjusting the surface normal directions. On the other hand, our result (re-rendered surface shown bottom right) contains no shadowing artefacts.

Specular images. To evaluate performance on highly specular objects, in Figure 6 we show results for the UCSD reflectance database [52]. This database is challenging because the objects were selected for their
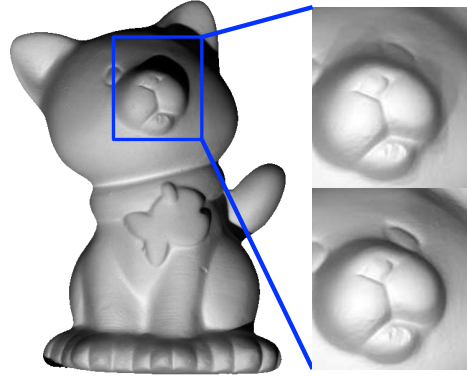

Figure 5: Close up of a reconstruction in a heavily shadowed region. On the right we show reilluminations of the recovered surface with frontal lighting. Top right: naive result using Woodham's approach. Note that the surface normals are corrupted by cast shadows. Bottom right: our result. Our estimated surface contains no artefacts due to shadowing.

highly non-Lambertian reflectance properties. Indeed, they were originally used for the purposes of reflectance estimation. The dataset contains high dynamic range images which are necessary to capture the order of magnitude difference in brightness between diffuse and specular reflections for these objects.

Note that despite dramatic deviations from our modelling assumptions, we are able to recover smooth height fields and normal maps that are both globally stable yet contain finescale detail. The object in the third row demonstrates an interesting failure case. The albedo in the lower, right of centre segment of the gourd has a diffuse albedo close to zero. The useful observations in this region are the specularities and a non-Lambertian photometric stereo algorithm would be required to recover meaningful shape. Nevertheless, since our algorithm solves globally for the surface height, the height estimates surrounding the region impose constraints that lead to a reasonable interpolation over this area.

Shadowed and specular images. For our final experiments, in Figure 7 we show results on a dataset that we collected in our own lightstage [53]. The objects were chosen as they exhibit both specular reflection and contain cast shadows. The objects are also highly complex containing finescale shape and albedo variations as well as both smooth and abrupt changes in surface gradient. We captured 18 images with varying illumination for each object. In this case, the initial shape and albedo estimates are provided by the spherical gradient photometric stereo algorithm of Ma et al. [8]. The bias in this algorithm is different to the results of Wood-

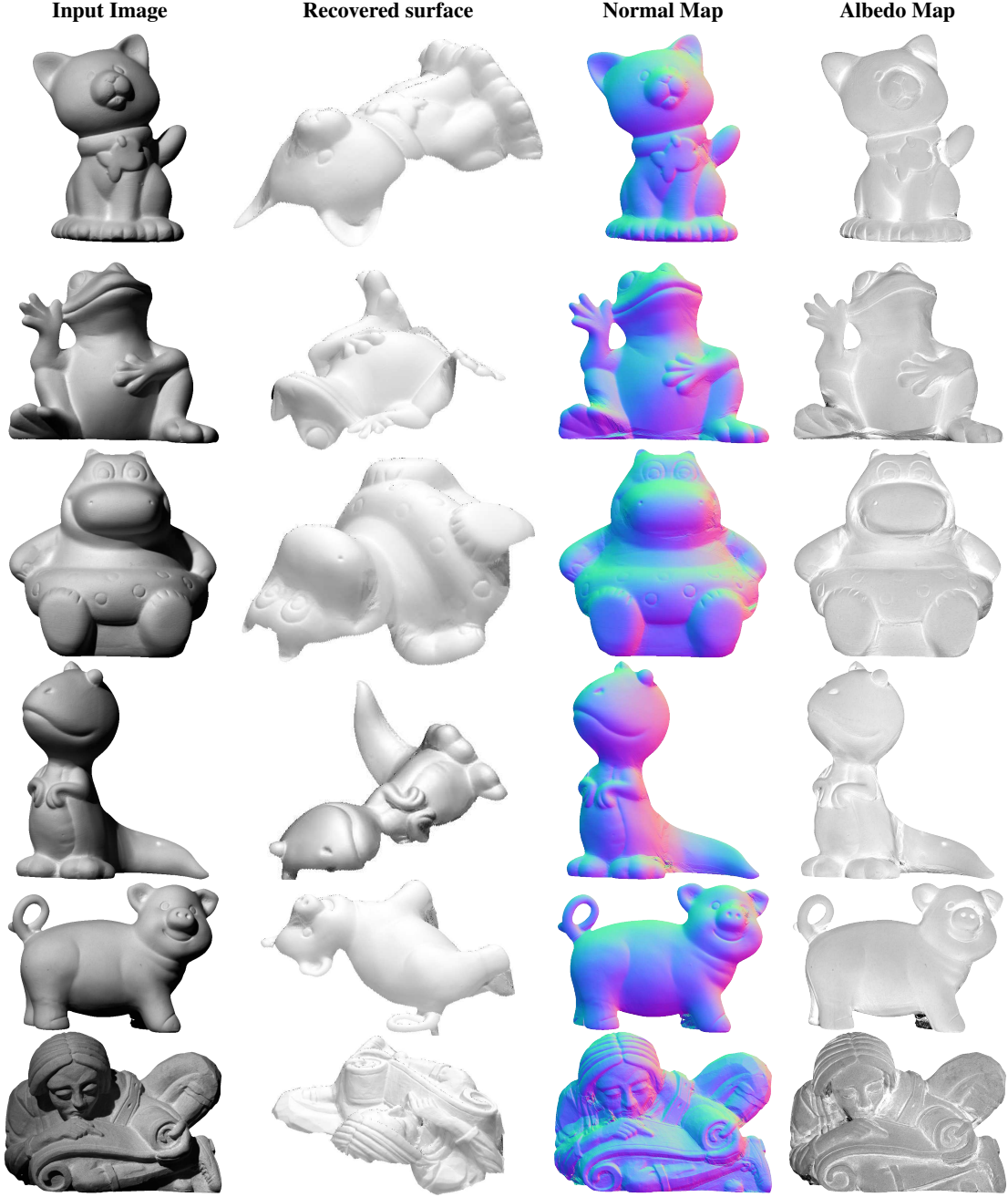| Input Image	Recovered surface	Normal Map	Albedo Map

Figure 4: Qualitative results for the Harvard dataset [51]. Input images (col. 1) contain numerous cast shadows. Col. 2 shows reconstructed surfaces in a novel view. Col. 3 shows the surface normals of the reconstructed surface and col. 4 the albedo map.

ham's algorithm. Albedo estimates are corrupted by ambient occlusion and surface normal estimates in occluded regions are biased towards the "bent normal" direction. Note that despite the complexity of the object shape and appearance, we are still able to recover stable height estimates with no visible corruption caused by non-Lambertian effects.

6. Conclusions

We have presented a photometric stereo algorithm that combines model-based selection with solving for surface height via photometric ratios. Our approach has a number of attractive properties. The selection algorithm is robust to any source of noise causing non-
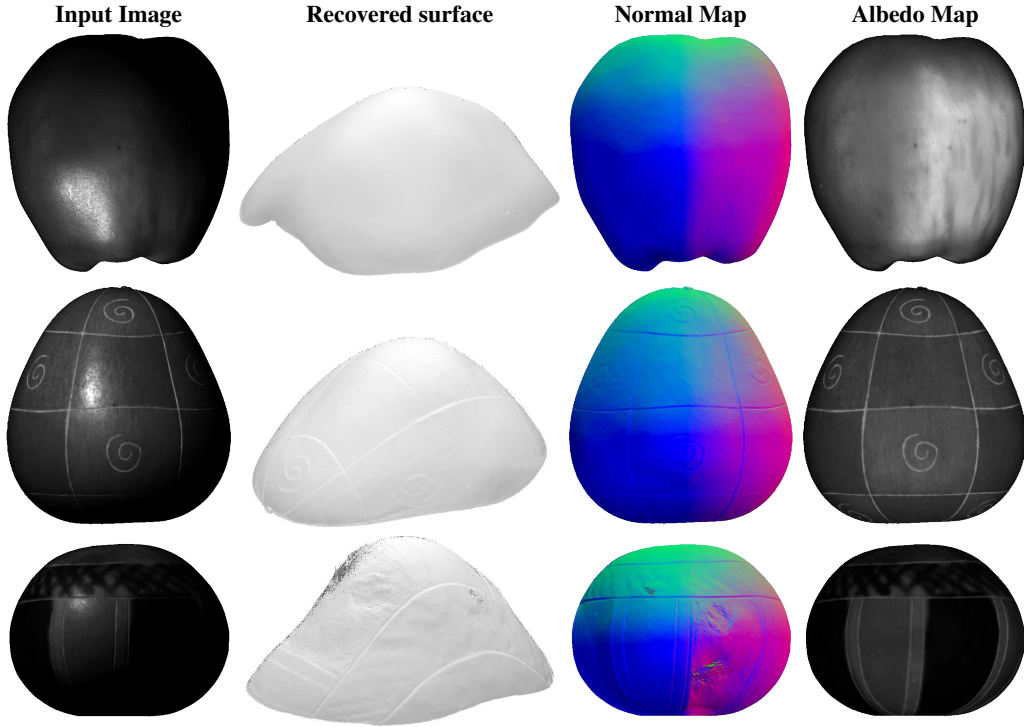
11

Figure 6: Qualitative results for the UCSD dataset [52]. Input images (col. 1) contain strong specular reflections. Col. 2 shows reconstructed surfaces in a novel view. Col. 3 shows the surface normals of the reconstructed surface and col. 4 the albedo map.

Gaussian deviations from the expected intensity. Our experimental results show that this is useful for detecting shadowed and specular observations. By solving in the surface height domain, we guarantee an integrable solution. Moreover, surface normals at each pixel are no longer independent since the surface gradients depend on surface height at neighbouring pixels. This introduces non-local constraints and improves robustness. Finally, our method requires only a single parameter. Even this, the threshold parameter, is a dimensionless quantity (indicating the number of standard deviations of noise that we will consider as valid observations) and is therefore easy to set (we used the same value in all of our experimental results). This means that our method does not require tuning to a particular dataset.

There are a number of avenues for future work. Concerning the selection strategy, we currently make selections independently across images and then use a single cycle through the selections to form pairs. It may be better to define a criterion for a good pair and make selections pairwise. We would also like to experiment with weighting pairs to improve robustness. Our selection strategy is currently based on a very simple statistical test. In future, we intend to consider more appropriate directional distributions for uncertainty in the surface normal direction and use this to rigorously define the probability distribution for an observation. Moreover, our selection relies on a reliable initial estimate. If this is systematically biased, then the selections themselves are biased. Hence, in future we will explore whether selection and reconstruction can be combined without relying on an initial reconstruction.

Concerning height recovery from photometric ratios, Equation 6 can be made invariant to ambient illumination by introducing a third image. Specifically, the idea is to take ratios of *differences* between images so that the ambient term is cancelled. Although this makes the method more general, it would also be more sensitive to noise. We also believe that robustness could be improved by introducing priors such as smoothness into Equation 12. In low albedo regions where all observations are noisy, it may be better to rely on smooth interpolation from more reliably reconstructed regions.

Reproducible Research

A Matlab implementation of our algorithm and the synthetic datasets used are available from http://www-users.cs.york.ac.uk/~wsmith/HfPR.

Input Image **Recovered surface** **Normal Map** **Albedo Map**

Figure 7: Qualitative results for our own data collected in a lightstage. Input images (col. 1) contain both cast shadows and strong specular reflections. Col. 2 shows reconstructed surfaces in a novel view. Col. 3 shows the surface normals of the reconstructed surface and col. 4 the albedo map.

References

[1] R. J. Woodham, Photometric method for determining surface orientation from multiple images, Optical engineering 19 (1) (1980) 191139–191139.

[2] S. K. Zhou, G. Aggarwal, R. Chellappa, D. W. Jacobs, Appearance characterization of linear lambertian objects, generalized photometric stereo, and illumination-invariant face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 29 (2) (2007) 230–245.

[3] A. S. Georghiades, P. N. Belhumeur, D. Kriegman, From few to many: Illumination cone models for face recognition under variable lighting and pose, IEEE Trans. Pattern Anal. Mach. Intell. 23 (6) (2001) 643–660.

[4] A. Hertzmann, S. M. Seitz, Shape and materials by example: A photometric stereo approach, in: Proc. CVPR, Vol. 1, 2003, pp. I–533.

[5] D. Vlasic, P. Peers, I. Baran, P. Debevec, J. Popović, S. Rusinkiewicz, W. Matusik, Dynamic shape capture using multi-view photometric stereo, ACM Trans. Graph. 28 (5) (2009) 174.

[6] N. Tsumura, Y. Miyake, F. H. Imai, Medical vision: measurement of skin absolute spectral-reflectance image and the application to component analysis, in: Proc. International Conference on Multispectral Color Science, 2001, pp. 25–28.

[7] M. L. Smith, The analysis of surface texture using photometric stereo acquisition and gradient space domain mapping, Image Vis. Comput. 17 (14) (1999) 1009–1019.

[8] W.-C. Ma, T. Hawkins, P. Peers, C.-F. Chabert, M. Weiss, P. Debevec, Rapid acquisition of specular and diffuse normal maps from polarized spherical gradient illumination, in: Proc. Eurographics Symposium on Rendering, 2007, pp. 183–194.

[9] A. Ghosh, T. Chen, P. Peers, C. A. Wilson, P. Debevec, Estimating specular roughness and anisotropy from second order spherical gradient illumination, Comput. Graph. Forum 28 (4) (2009) 1161–1170.

[10] A. Ghosh, T. Chen, P. Peers, C. A. Wilson, P. Debevec, Circularly polarized spherical illumination reflectometry, ACM Trans.

Graph. (Proceedings of SIGGRAPH Asia) 29 (6) (2010) 162.

[11] D. Nehab, S. Rusinkiewicz, J. Davis, R. Ramamoorthi, Efficiently combining positions and normals for precise 3D geometry., ACM Trans. Graph. 24 (3) (2005) 536–543.

[12] S. Barsky, M. Petrou, The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows., IEEE Trans. Pattern Anal. Mach. Intell. 25 (10) (2003) 1239–1252.

[13] J. Sun, M. Smith, L. Smith, S. Midha, J. Bamber, Object surface recovery using a multi-light photometric stereo technique for non-lambertian surfaces subject to shadows and specularities, Image Vis. Comput. 25 (7) (2007) 1050–1057.

[14] S. Ikehata, D. Wipf, Y. Matsushita, K. Aizawa, Robust photometric stereo using sparse regression, in: Proc. CVPR, 2012, pp. 318–325.

[15] C. Hernández, G. Vogiatzis, R. Cipolla, Overcoming shadows in 3-source photometric stereo, IEEE Trans. Pattern Anal. Mach. Intell. 33 (2) (2011) 419–426.

[16] B. K. Horn, M. J. Brooks, The variational approach to shape from shading, Computer Vision, Graphics, and Image Processing 33 (2) (1986) 174–208.

[17] R. T. Frankot, R. Chellappa, A method for enforcing integrability in shape from shading algorithms, IEEE Trans. Pattern Anal. Mach. Intell. 10 (4) (1988) 439–451.

[18] B. K. Horn, Height and gradient from shading, Int. J. Comput. Vis. 5 (1) (1990) 37–75.

[19] A. Yuille, D. Snow, Shape and albedo from multiple images using integrability, in: Proc. CVPR, 1997, pp. 158–164.

[20] A. P. Harrison, D. Joseph, Maximum likelihood estimation of depth maps using photometric stereo, IEEE Trans. Pattern Anal. Mach. Intell. 34 (7) (2012) 1368–1380.

[21] M. Chandraker, S. Agarwal, D. Kriegman, Shadowcuts: Photometric stereo with shadows, in: Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on, 2007, pp. 1–8.

[22] D. Miyazaki, K. Hara, K. Ikeuchi, Median photometric stereo as applied to the segonko tumulus and museum objects, Int. J. Comput. Vis. 86 (2–3) (2010) 229–242.

[23] C. Yu, Y. Seo, S. W. Lee, Photometric stereo from maximum feasible lambertian reflections, in: Proc. ECCV, 2010, pp. 115–126.

[24] Y. Mukaigawa, Y. Ishii, T. Shakunaga, Analysis of photometric factors based on photometric linearization, JOSA A 24 (10) (2007) 3326–3334.

[25] T. Higo, Y. Matsushita, K. Ikeuchi, Consensus photometric stereo, in: Proc. CVPR, 2010, pp. 1157–1164.

[26] V. Argyriou, S. Zafeiriou, B. Villarini, M. Petrou, A sparse representation method for determining the optimal illumination directions in photometric stereo, Signal Processing 93 (11) (2013) 3027–3038.

[27] E. Prados, N. Jindal, S. Soatto, A non-local approach to shape from ambient shading, in: Proc. SSVM, 2009, pp. 696–708.

[28] R. Basri, D. Jacobs, I. Kemelmacher, Photometric stereo with general, unknown lighting., International Journal of Computer Vision 72 (3) (2007) 239–257.

[29] R. Huang, W. A. P. Smith, Shape-from-shading under complex natural illumination, in: Proc. ICIP, 2011, pp. 13–16.

[30] T. Papadhimitri, P. Favaro, Uncalibrated near-light photometric stereo, in: Proc. BMVC, 2014.

[31] P. A. Davis, L. A. Soderblom, Modeling crater topography and albedo from monoscopic viking orbiter images: 1. methodology, Journal of Geophysical Research: Solid Earth 89 (B11) (1984) 9449–9457.

[32] A. S. McEwen, Topography and albedo of ius chasma, mars, Lunar and Planetary Science Conference 16 (1985) 528–529.

[33] S. Lee, M. Brady, Integrating stereo and photometric stereo to monitor the development of glaucoma, in: Proc. BMVC, 1990.

[34] L. B. Wolff, E. Angelopoulou, 3-d stereo using photometric ratios, in: Proc. ECCV, 1994, pp. 247–258.

[35] K. M. Lee, C.-C. J. Kuo, Shape from photometric ratio and stereo, Journal of Visual Communication and Image Representation 7 (2) (1996) 155–162.

[36] T.-P. Wu, K.-L. Tang, C.-K. Tang, T.-T. Wong, Dense photometric stereo: A markov random field approach, IEEE Trans. Pattern Anal. Mach. Intell. 28 (11) (2006) 1830–1846.

[37] W. Y. Zhao, R. Chellappa, Symmetric shape-from-shading using self-ratio image, Int. J. Comput. Vis. 45 (1) (2001) 55–75.

[38] C. Wöhler, K. Hafezi, A general framework for three-dimensional surface reconstruction by self-consistent fusion of shading and shadow features, Pattern Recognition 38 (7) (2005) 965–983.

[39] A. P. Pentland, Linear shape from shading, Int. J. Comput. Vis. 4 (2) (1990) 153–162.

[40] P. S. Tsai, M. Shah, Shape from shading using linear approximation, Image Vis. Comput. 12 (8) (1994) 487–498.

[41] A. Ecker, A. D. Jepson, Polynomial shape from shading, in: Proc. CVPR, 2010, pp. 145–152.

[42] R. Mecca, A. Wetzler, R. Kimmel, A. M. Bruckstein, Direct shape recovery from photometric stereo with shadows, in: Proc. 3DV, 2013, pp. 382–389.

[43] R. Mecca, M. Falcone, Uniqueness and approximation of a photometric shape-from-shading model, SIAM Journal on Imaging Sciences 6 (1) (2013) 616–659.

[44] R. Mecca, A. Tankus, A. Wetzler, A. M. Bruckstein, A direct differential approach to photometric stereo with perspective viewing, SIAM Journal on Imaging Sciences 7 (2) (2014) 579–612.

[45] R. Mecca, A. Wetzler, A. M. Bruckstein, R. Kimmel, Near field photometric stereo with point light sources, SIAM Journal on Imaging Sciences 7 (4) (2014) 2732–2770.

[46] R. Mecca, E. Rodolà, D. Cremers, Realistic photometric stereo using partial differential irradiance equation ratios, Computers & Graphics 51 (2015) 8–16.

[47] M. Chandraker, J. Bai, R. Ramamoorthi, On differential photometric reconstruction for unknown, isotropic brdfs, IEEE Trans. Pattern Anal. Mach. Intell. 35 (12) (2013) 2941–2955.

[48] M. Harker, P. O'Leary, Least squares surface reconstruction from measured gradient fields, in: Proc. CVPR, 2008, pp. 1–7.

[49] T. A. Davis, Algorithm 915, SuiteSparseQR: Multifrontal multithreaded rank-revealing sparse QR factorization, ACM Trans. Mathematical Software 38 (1) (2011) 8:1–8:22.

[50] M. A. Fischler, R. C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, Communications of the ACM 24 (6) (1981) 381–395.

[51] Y. Xiong, A. Chakrabarti, R. Basri, S. J. Gortler, D. W. Jacobs, T. Zickler, From shading to local shape, IEEE Trans. Pattern Anal. Mach. Intell. 37 (1) (2015) 67–79.

[52] N. Alldrin, T. Zickler, D. Kriegman, Photometric stereo with non-parametric and spatially-varying reflectance, in: Proc. CVPR, 2008.

[53] A. Dutta, W. A. P. Smith, Minimal image sets for robust spherical gradient photometric stereo, Technical Sketch, SIGGRAPH Asia (2010) 22:1–22:2.