USING PHOTOMETRIC STEREO TO REFINE THE GEOMETRY OF A 3D SURFACE MODEL

Zsolt Jankó
Computer and Automation Research Institute and
Eötvös Loránd University, Budapest, Hungary

Keywords: Surface reconstruction, Bumpiness recovery, Calibration–estimation problem, RANSAC.

Abstract: In this paper we aim at refining the geometry of 3D models of real objects by adding surface bumpiness to them. 3D scanners are usually not accurate enough to measure fine details, such as surface roughness. Photometric stereo is an appropriate technique to recover bumpiness. We use a number of images taken from the same viewpoint under varying illumination and an initial sparse 3D mesh obtained by a 3D scanner. We assume the surface to be Lambertian, but the lighting properties are unknown. The novelty of our method is that the initial sparse 3D mesh is exploited to calibrate light sources and then to recover surface normals. The importance of refining the geometry of a bumpy surface is demonstrated by applying the method to synthetic and real data.

1 INTRODUCTION

Creating photorealistic 3D models of real objects is a challenging problem. For realistic appearance precise geometry is essential. 3D laser scanners are usually used to acquire the geometry of real objects; however, most of these scanners are not accurate enough to measure fine details, such as surface roughness. Realistic appearance of a model is significantly reduced when 3D roughness is not sufficiently represented, or when the lack of the roughness is concealed by textures.

Photometric stereo is a popular field of computer vision aiming at recovering surface orientation from images. The essence of photometric methods is to calculate surface normals from the changing of the intensity through altering the incoming light. The great advantage of these techniques is their accuracy, which lets us measure fine details, small bumps on the surface.

In this paper we present a novel method based on photometric stereo. As input the method obtains a number of images taken from the same viewpoint under varying illumination and a sparse 3D mesh captured by a 3D scanner. We assume that the surface is Lambertian and that the size of the object is small relative to its distance to the light source. The camera is assumed to be calibrated, but lighting properties are unknown. The novelty of our method is that the initial sparse 3D mesh is exploited to calibrate light sources. The problem is decomposed in two parts: first, lighting properties are estimated using the given initial geometry; second, geometry is refined using the already calibrated light sources.

Related Work

The main goal of photometric stereo is to estimate surface orientation from a number of images, where the direction of incident illumination varies, but the camera is fixed. In the most general situation we have information neither about the camera and the light sources nor about surface reflectance. However, in this case surface orientation cannot be perfectly extracted. One needs to set appropriate assumptions about the environment or the surface to obtain precise results.

A general assumption is that camera and light sources are calibrated. When the surface is considered to be Lambertian (Woodham, 1978; Rushmeier et al., 1997), normals of surface points can be easily extracted by solving a linear system of equa-

Jankó Z. (2007).
USING PHOTOMETRIC STEREO TO REFINE THE GEOMETRY OF A 3D SURFACE MODEL.
In *Proceedings of the Second International Conference on Computer Vision Theory and Applications - IU/MTSV*, pages 355-362

tions. A surface with more complex, spatially-varying *BRDF (Bidirectional Reflection Distribution Function)* makes recovery more difficult.

In their papers (Lensch et al., 2001; Lensch et al., 2003) Lensch et al. also discuss the problem of BRDF acquisition from photographic data. Their work relates to ours as they require scanned geometry as input. They cluster material estimates over the surface, then use the known BRDF to refine surface normals. The main differences between their approach and ours are that we assume Lambertian surface while they can handle spatially-varying BRDFs; on the other hand, they need to exactly know illumination properties to recover BRDFs, while we use uncalibrated lights.

If illumination is unknown, one needs to set suitable constraints on surface reflectance and geometry. When Lambertian reflectance is assumed (Zhang et al., 2003; Yuille et al., 2003), a few constraints are sufficient, but assuming more complex or unknown reflection models (Yuille et al., 1999; Drbohlav and Šara, 2001; Lim et al., 2005) requires further restrictions on the geometry and the environment. Typical assumptions are: the radiances of light sources are constant or known; BRDF is constant over the whole surface; surface normals are integrable, i.e., they form a consistent vector field.

Note that most of the techniques mentioned above apply the integrability constraint that forces the recovered normals to form a continuous surface. In most cases this assumption is reasonable, but in our case, when a smoothed continuous surface is given and roughness is to be recovered, this constraint is not applicable.

The structure of this paper is the following. First, we formalise the problem of surface bumpiness recovery, then the proposed method is discussed. Finally, the results of the method for synthetic and real data are presented.

2 BUMPINESS RECOVERY

2.1 Problem Formulation

Consider the input setup of m images and a surface mesh of the same object. The images are supposed to have been taken from the same view with fixed and known camera parameters but under varying and unknown lighting conditions. The mesh consists of a set $\mathcal{P}$ of 3D points and a set $\mathcal{N}$ of corresponding normals. The surface is assumed to be Lambertian.

The goal is to refine the surface and determine local bumpiness by extracting a dense and accurate normal map. According to Lambert's cosine law, intensity in surface point $\mathbf{X}$ under lighting condition labelled by $i \in [1..m]$ can be calculated as

$$I(\mathbf{X}, i) = \frac{1}{\pi} a_{\mathrm{d}}(\mathbf{X}) \underline{n}(\mathbf{X})^T \rho(i) \underline{l}(\mathbf{X}, i), \qquad (1)$$

where $a_{\mathrm{d}}(\mathbf{X})$ denotes the albedo, $\underline{n}(\mathbf{X})$ the surface normal in surface point $\mathbf{X}$, while $\rho(i)$ is the radiance, $\underline{l}(\mathbf{X}, i)$ the direction to the light source i from point $\mathbf{X}$. Vectors $\underline{n}(\mathbf{X})$ and $\underline{l}(\mathbf{X}, i)$ are unit length.

We wish to solve equation (1) when $I(\mathbf{X}, i)$ is known and we have a rough approximation for $\underline{n}(\mathbf{X})$ if $\mathbf{X} \in \mathcal{P}$. Albedos and light source attributes are completely unknown. The goal is to estimate the positions and the radiances of the light sources, as well as to extract "pixel-dense" normal and albedo maps.

Let us modify the notations of (1) and extend it over the whole image domain. Eliminating surface points that are invisible from the camera, a one-to-one correspondence exists between remaining 3D surface points and image points, thus the functions are modified to depend not on 3D points $\mathbf{X}$, but on image points $\mathbf{u}$. Furthermore, the functions of parameter i are substituted by m independent functions. For instance, I_i denotes image i taken under lighting conditions labelled by i; it is a function of image points $\mathbf{u}$. Equation (1) is then modified as

$$I_i(\mathbf{u}) = \frac{1}{\pi} a_{\mathrm{d}}(\mathbf{u}) \underline{n}(\mathbf{u})^T \rho_i \underline{l}_i(\mathbf{u}), \, i \in [1..m], \qquad (2)$$

where $\mathbf{u}$ is in the domain R_i of image i. Note that the fixed camera means that all images have the same domain, denoted by R. Practically, domain R is restricted to contain only such image points that the corresponding 3D points are visible from all light sources. This domain can easily be determined from the input images.

According to (2), the problem is posed as the minimisation of the following function:

$$\phi(a_{\mathrm{d}}, \underline{n}, \rho_1, \underline{l}_1, \ldots, \rho_m, \underline{l}_m) = \sum_{i=1}^{m} \sum_{\mathbf{u} \in R} \left(I_i(\mathbf{u}) - \frac{1}{\pi} a_{\mathrm{d}}(\mathbf{u}) \underline{n}(\mathbf{u})^T \rho_i \underline{l}_i(\mathbf{u}) \right)^2 . \qquad (3)$$

We decompose the problem in two steps: first, normals $\underline{n} \in \mathcal{N}$, which are only rough approximations of real normals, are used to calibrate light sources; second, the calibrated light sources are used to create dense albedo and normal maps.

2.2 Calibration–Estimation Problem

In (3) the direction $\underline{l}_i(\mathbf{u})$ of light source i from surface point $\mathbf{X}(\mathbf{u})$ is a function of $\mathbf{u}$, which makes it more difficult to estimate the light source's position. Fortunately, in most cases the size of the object is small compared to the distance between the object and the light source, thus direction $\underline{l}$ does not change significantly over the surface. Substituting $\mathbf{X}$ by the centroid of the surface is therefore reasonable, which implies that direction $\underline{l}_i$ is constant and depends on $\mathbf{u}$ no more. Consequently, the problem of calibrating light sources using normals $\underline{n} \in \mathcal{N}$ can be defined as the minimisation of

$$\phi'(a_{\mathrm{d}}, \rho_1, \underline{l}_1, \ldots, \rho_m, \underline{l}_m) = \sum_{i=1}^{m} \sum_{\mathbf{X}(\mathbf{u}) \in \mathcal{P}} \left(I_i(\mathbf{u}) - \frac{1}{\pi} a_{\mathrm{d}}(\mathbf{u}) \underline{n}(\mathbf{u})^T \rho_i \underline{l}_i \right)^2. \quad (4)$$

Estimates of ρ_i and $\underline{l}_i$ are possible to find up to a constant linear transformation using the Singular Value Decomposition (SVD). For this, let us collect intensities $I_i(\mathbf{u}_j)$, $i \in [1..m]$ in the $k \times m$ matrix D, where $k = |\mathcal{P}|$ is the number of points of the surface mesh and $\mathbf{u}_j$, $j \in [1..k]$, is a series of image points for which $\mathbf{X}(\mathbf{u}_j) \in \mathcal{P}$. Using the notations of (Koenderink and Van Doorn, 1997), *estimates of point properties* (such as albedo and normal vector) are collected in the $k \times 3$ matrix E, and *calibration data of light sources* are in the $m \times 3$ matrix C. Thus, the calibration–estimation problem defined in (4) is equivalent to the problem of factorising matrix D into EC^T, which requires also rank reduction, since the rank of D is usually greater than 3.

As discussed in (Koenderink and Van Doorn, 1997; Yuille et al., 1999), SVD decomposes matrix D in the form $D = UWV^T$, where W contains the singular values of D. If our model is correct, the rank of matrix D must be 3, i.e., W must have only 3 nonzero elements. Even if this usually does not hold, SVD is still guaranteed to provide the best least squares solution for factorisation and rank reduction. For mostly Lambertian objects the three largest eigenvalues and the corresponding columns of U and V represent the Lambertian component of the reflectance function (Yuille et al., 1999). If W_3 is the 3×3 diagonal matrix containing only the 3 largest singular values from W, and U_3 (V_3) is a $k \times 3$ ($m \times 3$) matrix with the columns of U (V) corresponding to the 3 largest singular values, then $E = U_3\sqrt{W_3}$ and $C = V_3\sqrt{W_3}$. Now $D \approx EC^T$.

Although the above solution seems convenient, it may fail to be a desirable one: for an invertible 3×3 matrix A, EA and CA^{-T} is also a solution. One has to set appropriate constraints to restrict the number of possible solutions. One of the most frequently used constraints is the so called *integrability* constraint (Yuille et al., 1999), where integrability conditions are used to ensure that surface normals form a consistent surface. It assumes smooth surface without roughness. However, we already have a smooth surface approximating the real one, and bumpiness is what we wish to extract, thus integrability constraint cannot be applied. Instead of it, other constraints are used as discussed below.

2.3 Support Function Based on Feasible Constraints

Let us denote the i-th row of C by $\underline{c}_i^T$, the j-th row of E by $\underline{e}_j^T$, and the q-th column of A by $\underline{a}_q$. We are looking for an invertible matrix A, for which $\underline{e}_j^T A = \frac{1}{\pi} a_{\mathrm{d}}(\mathbf{u}_j)\underline{n}(\mathbf{u}_j)^T$, $j \in [1..k]$, and $A^{-1}\underline{c}_i = \rho_i \underline{l}_i$, $i \in [1..m]$. Normals $\underline{n}(\mathbf{u}_j)$, $j \in [1..k]$ are known; they are only initial approximations of real surface normals, but they can be used to set a constraint on A as $\underline{e}_j^T A$ should be collinear with $\underline{n}(\mathbf{u}_j)^T$. Formally, we wish matrix A to satisfy equalities $\frac{\underline{e}_j^T \underline{a}_1}{n_1(\mathbf{u}_j)} = \frac{\underline{e}_j^T \underline{a}_2}{n_2(\mathbf{u}_j)} = \frac{\underline{e}_j^T \underline{a}_3}{n_3(\mathbf{u}_j)} = \frac{1}{\pi} a_{\mathrm{d}}(\mathbf{u}_j)$. This constraint can be expressed in form of a homogeneous linear system of equations $H\underline{x} = \underline{0}$, where $\underline{x} = [\underline{a}_1^T, \underline{a}_2^T, \underline{a}_3^T]^T$ and

$$H = \begin{bmatrix} \frac{\underline{e}_1^T}{n_1(\mathbf{u}_1)} & -\frac{\underline{e}_1^T}{n_2(\mathbf{u}_1)} & 0 \\ 0 & \frac{\underline{e}_1^T}{n_2(\mathbf{u}_1)} & -\frac{\underline{e}_1^T}{n_3(\mathbf{u}_1)} \\ -\frac{\underline{e}_1^T}{n_1(\mathbf{u}_1)} & 0 & \frac{\underline{e}_1^T}{n_3(\mathbf{u}_1)} \\ & \cdots & \\ \frac{\underline{e}_k^T}{n_1(\mathbf{u}_k)} & -\frac{\underline{e}_k^T}{n_2(\mathbf{u}_k)} & 0 \\ 0 & \frac{\underline{e}_k^T}{n_2(\mathbf{u}_k)} & -\frac{\underline{e}_k^T}{n_3(\mathbf{u}_k)} \\ -\frac{\underline{e}_k^T}{n_1(\mathbf{u}_k)} & 0 & \frac{\underline{e}_k^T}{n_3(\mathbf{u}_k)} \end{bmatrix}.$$

Unfortunately, uniqueness of the solution is not guaranteed, since the dimension of the null space of the linear transformation H can be greater than 1. In this case the best solution of the null space needs to be found by applying further constraints.

Two further constraints are considered: first, albedos should be greater than zero; second, the angle between the light source's direction and the camera's given direction should not exceed a given threshold. The latter is also reasonable since if the angle were great, the surface would be mostly invisible. Thus, one needs to search through the null space of linear transformation H for the element which best satisfies these two constraints.

Matrix A—and its vector form $\underline{x}$—can be considered as a *model* fitted to the given data. One needs to define an appropriate support function that gives the goodness of a model.

The first constraint, namely that the albedo is greater than zero, is true, if the angle between the two vectors $\underline{e}_j^T A$ and $\underline{n}(\mathbf{u}_j)^T$ is sufficiently small. Formally, we say that the j-th element of the estimates *supports* the model if $\angle(\underline{e}_j^T A, \underline{n}(\mathbf{u}_j)^T) \leq 1.0°$. We note that at the calculation of the angle one should consider that $\underline{e}_j^T A$ is usually not unit length.

The number of the supporting normals gives a true description to the goodness of the model, but it is useless without the second constraint. The light sources should be close in angle to the camera, otherwise most of the surface is invisible. We say that the i-th light source *supports* the model if the angle between $A^{-1}\underline{c}_i$ and the given vector towards the camera is below a threshold. Our experience shows that this threshold should be set between 30° and 45°.

The support of a model is given by the number of supporting normals and supporting light sources. However, light sources should get higher priority, since loss of a light source incurs an invalid model with larger probability than loss of a normal. Hence the support of a model $\underline{x}$ is calculated as follows:

$$S(\underline{x}) = q \cdot \text{Nof(supporting lights)} + \text{Nof(supporting normals)}, \quad (5)$$

where Nof is *Number of*, and $q = k/2$ is a reasonable choice.

2.4 Applying Ransac

The method described above is not robust. Experience shows that inaccuracy of approximated normals and noise of images lead to unreasonable results. Our method uses least squares, it is therefore very sensitive to outliers appearing among normals. We decided to apply the RANSAC (RANdom SAmple Consensus) algorithm of Fischler and Bolles (Fischler and Bolles, 1981), which is a general and efficient robust estimator.

The idea is the following: instead of using all normals from $\mathcal{N}$, a subset $\mathcal{N}_1 \subset \mathcal{N}$ is randomly selected and the method is executed only on $\mathcal{N}_1$. This is repeated a number of times for various $\mathcal{N}_1$. For each $\mathcal{N}_1$ a consensus set $\mathcal{N}_c$ is formed by the supporting normals. The consensus set defined by the best sample gives the inliers of $\mathcal{N}$. Finally, the model is re-estimated using this consensus set. Informally, the technique uses the fact that if $|\mathcal{N}_1|$ is small, and the selection is repeated sufficient times, then the probability of that a subset without any outliers has been selected is great. We have chosen to set $|\mathcal{N}_1|$ to 5, which is a little greater than the necessary 3, in order to avoid frequent under-determinedness. The number of iterations and the minimal size of an acceptable consensus set can be calculated from the estimated incidence rate of outliers, as discussed in (Hartley and Zisserman, 2000).

Notice that the above problem can be considered as clustering. Normals of $\mathcal{N}$ approximate real normals of the surface. We suppose to have a smooth surface and want to recover the real bumpy surface. Most of the normals in $\mathcal{N}$ approximate wrongly (denote this subset by $\mathcal{N}_w$) and only a few of them approximate well (denote by $\mathcal{N}_g$). However, correct normals are consistent, which helps us find their cluster. Consistency means that if $\mathcal{N}_1$ is chosen to be a subset of $\mathcal{N}_g$, then the result is correct and satisfies the constraints. On the other hand, if $\mathcal{N}_1$ contains one or more normals from $\mathcal{N}_w$, outliers deteriorate the result.

This is also the reason why further ambiguities (including the Generalised Bas-Relief Ambiguity (Belhumeur et al., 1999)) are avoided, in contrast to other methods, e.g., (Yuille et al., 2003). It is known in photometric stereo that view sources can be precisely calibrated, without any ambiguities if surface normals are known. We do not have accurate normals, but we have a set $\mathcal{N}$ of normals that has a subset of correct normals. And this subset is determined by the method described above.

2.5 Extracting Normals

After calibrating light sources, normals can be determined in each point $\mathbf{u}$ as follows. Consider again the formula of the Lambertian reflection model: $I_i(\mathbf{u}) = \frac{1}{\pi} a_\mathrm{d}(\mathbf{u})\underline{n}(\mathbf{u})^T \rho_i \underline{l}_i$, $i \in [1..m]$. Due to the light source calibration, ρ_i and $\underline{l}_i$ are known, thus $a_\mathrm{d}(\mathbf{u})$ and $\underline{n}(\mathbf{u})$ can easily be calculated, separately for each image point $\mathbf{u}$, by solving the over-determined linear system of equations $B\underline{y}_\mathbf{u} = \underline{b}_\mathbf{u}$, where

$$B = \begin{bmatrix} \rho_1 \underline{l}_1^T \\ \cdots \\ \rho_m \underline{l}_m^T \end{bmatrix} \text{ and } \underline{b}_\mathbf{u} = \begin{bmatrix} I_1(\mathbf{u}) \\ \cdots \\ I_m(\mathbf{u}) \end{bmatrix}.$$

Since $\underline{n}(\mathbf{u})$ is assumed to be unit length, $a_\mathrm{d}(\mathbf{u}) = \pi\|\underline{y}_\mathbf{u}\|$ and $\underline{n}(\mathbf{u}) = \frac{\underline{y}_\mathbf{u}}{\|\underline{y}_\mathbf{u}\|}$.

3 RESULTS

To demonstrate the efficiency, we have tested the method both on synthetic and real data. The synthetic dataset contains a 3D mesh of a bumpy sphere, a smoothed mesh without bumpiness, and five images of the bumpy sphere taken using five different light sources. For each image one of the light sources was turned on, while the others were turned off.

The method was run with the input of the smoothed mesh and the five images, yielding a dense normal map. The estimated normal map was then compared to the ground truth normal map: the angles between the corresponding normals were measured, in degrees, and the mean of the angles gave the magnitude of the error. Since the method is non-deterministic because of RANSAC, it was run 10 times resulting the error of $1.85^{\circ} \pm 0.1^{\circ}$ that is significantly smaller than the error of the normals of the input smoothed mesh, which is 7.55°. Note that the latter is also not too big. This is because the mean of the angles was considered: the surface had only a relatively small bumpy part, while larger parts of the surface were smooth. Obviously, normals of smooth parts were approximated precisely either before and after photometric estimation, and the small errors ($\approx 1^{\circ}$) of them prevented the mean from growing too large.

To demonstrate the efficiency of the method more clearly, the pictures of the normal maps are also presented. Fig. 1 contains the ground truth and the estimated normal maps, as well as their difference before and after applying the method. Pixel intensities of the normal maps represent the deviation of the normal vectors from a fixed unit vector. These intensities are calculated as $255 - 3n_d$, where n_d is the angular deviation in degrees. Pixel intensities of the difference maps show the errors in angle: when the error is n_e degrees, pixel intensity is $128 + 2n_e$. The bumps are clearly visible in the difference map before using the method, but almost perfectly disappear after that.

We have tested the method for real datasets, as well. The first dataset consists of the 3D mesh of a Plaquette and seven images of the object taken from the same viewpoint under varying illumination. To provide lighting, a simple table-lamp was used and moved in space. Fig. 2 shows three of the input images, the input 3D model and the resulting normal map.

The second dataset consists of a Frog model and five input images. (See Fig. 3.) Although it is hard to evaluate the result based on the map presented, comparing it to the input images shows that the locations of the bumps are precise.

The third dataset (Fig. 4) contain a wooden object in the shape of a Bottle. The dataset consists of the 3D mesh and seven images about the object. The results of the method applied for the three datasets demonstrate that our technique is suitable for detecting even fine roughness.

4 CONCLUSION

In this paper we presented a novel method to refine the geometry of 3D models of real objects. The photometric stereo based method uses a number of input images taken from the same viewpoint under varying lighting, supplemented with an initial sparse 3D mesh. The surface reflectance is assumed to be Lambertian, but normals, albedos and lighting properties are unknown.

The method solves the problem in two steps. First, the initial surface mesh is used to calibrate light sources. This problem is traced back to the well known calibration–estimation problem, and the solution is robustified by applying the RANSAC algorithm. Second, dense normal and albedo maps are extracted using the calibrated setup. The effectiveness of the method was demonstrated both on synthetic and real data.

ACKNOWLEDGEMENTS

The author thank Dmitry Chetverikov and Levente Hajder for their comments on this paper. This work is supported by EU Network of Excellence MUSCLE (FP6-507752).

REFERENCES

Belhumeur, P. N., Kriegman, D. J., and Yuille, A. L. (1999). The bas-relief ambiguity. *International Journal of Computer Vision*, 35(1):33–44.

Drbohlav, O. and Šara, R. (2001). Unambigous determination of shape from photometric stereo with unknown light sources. In *Proc. 8^{th} IEEE International Conference on Computer Vision*, pages 581–586.

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

Hartley, R. and Zisserman, A. (2000). *Multiple View Geometry in Computer Vision*. Cambridge University Press.

Koenderink, J. J. and Van Doorn, A. J. (1997). The generic bilinear calibration-estimation problem. *International Journal of Computer Vision*, 23(3):217–234.

Lensch, H. P., Kautz, J., Goesele, M., Heidrich, W., and Seidel, H.-P. (2001). Image-based reconstruction of spatially varying materials. In *Proc. 12^{th} Eurographics Workshop on Rendering Techniques*, pages 103–114.

Lensch, H. P., Kautz, J., Goesele, M., Heidrich, W., and Seidel, H.-P. (2003). Image-based reconstruction of spatial appearance and geometric detail. *ACM Transactions on Graphics*, 22(2):234–257.

Lim, J., Ho, J., Yang, M.-H., and Kriegman, D. (2005). Passive photometric stereo from motion. In *Proc.* 10^{th} *IEEE International Conference on Computer Vision*, pages 1635–1642.

Rushmeier, H., Taubin, G., and Guéziec, A. (1997). Applying shape from lighting variation to bump map capture. In *Proc.* 8^{th} *Eurographics Rendering Workshop*, pages 35–44.

Woodham, R. J. (1978). Photometric stereo: A reflectance map technique for determining surface orientation from image intensity. In *Image Understanding Systems and Industrial Applications, Proc. SPIE*, volume 155, pages 136–143.

Yuille, A., Coughlan, J. M., and Konishi, S. (2003). KGBR viewpoint–lighting ambiguity. *Journal of the Optical Society of America A*, 20(1):24–31.

Yuille, A. L., Snow, D., Epstein, R., and Belhumeur, P. N. (1999). Determining generative models of objects under varying illumination: Shape and albedo from multiple images using SVD and integrability. *International Journal of Computer Vision*, 35(3):203–222.

Zhang, L., Curless, B., Hertzmann, A., and Seitz, S. M. (2003). Shape and motion under varying illumination: Unifying structure from motion, photometric stereo, and multi-view stereo. In *Proc.* 9^{th} *IEEE International Conference on Computer Vision*, pages 618–625.