A New Perspective [on] Shape-from-Shading

Ariel Tankus
School of Computer Science
Tel-Aviv University
Tel-Aviv, 69978
arielt@post.tau.ac.il

Nir Sochen
School of Mathematics
Tel-Aviv University
Tel-Aviv, 69978
sochen@post.tau.ac.il

Yehezkel Yeshurun
School of Computer Science
Tel-Aviv University
Tel-Aviv, 69978
hezy@post.tau.ac.il

Abstract

Shape-from-Shading (SfS) is a fundamental problem in Computer Vision. The vast majority of research in this field have assumed orthography as its projection model. This paper re-examines the basis of SfS, the image irradiance equation, under an assumption of perspective projection. The paper also shows that the perspective image irradiance equation depends merely on the natural logarithm of the depth function (and not on the depth function itself), and as such it is invariant to scale changes of the depth function. We then suggest a simple reconstruction algorithm based on the perspective formula, and compare it to existing orthographic SfS algorithms. This simple algorithm obtained lower error rates than legacy SfS algorithms, and equated with and sometimes surpassed state-of-the-art algorithms. These findings lend support to the assumption that transition to a more realistic set of assumptions improves reconstruction significantly.

1. Introduction

Shape-from-Shading (SfS) is one of the fundamental problems in Computer Vision. First introduced by Horn in the 1970s [3], its goal is to solve the image irradiance equation, which relates the reflectance map to image intensity. Due to difficulties in solving the equation, most of the works in the field add simplifying assumptions to the equation. Of particular importance is the common assumption that scene points are projected orthographically during the photographic process.

Many works in the field of Shape-from-Shading have followed the seminal work of Horn [4], and assumed orthographic projection ([2], [5], [6], [14] and many more; see [13] for a survey).

The few works that did employ the perspective projection have been too restrictive and have not addressed the

general problem. [12] and [11] assumed that distance variations between camera and surface could be ignored. [10] employed a deformable model for the SfS problem, so reconstruction took place in 3D space. Thus, during the deformation process, the image point onto which a 3D point was projected changed, and its new location should have been interpolated, resulting in a nonuniform sampling of the image.

Another approach to perspective SfS is to model the problem using piecewise planar depth functions ([8], [9]). But orthographic and perspective reflectance maps of a plane are identical, as Sect. 5.1 would show. Therefore, the difference between the two types of projection of a piecewise planar surface is confined to the edges, while at the interior of the faces the orthographic and perspective models agree with each other.

While the vast majority of the SfS literature assumes orthographic projection, and the few perspective-projection studies are limited in their scopes, no information is available on the image irradiance equation under the perspective projection model. The purpose of this research was to improve the image irradiance equation by using the more realistic, perspective projection, and then to solve the Shape-from-Shading problem under these new assumptions. In order to quantify the possible contribution of modeling the projection as perspective, we compare the locations of orthographic and perspective projections of the same surface point. The comparison would show that the distance between the two projections can become significant. Consequently, assuming perspective projection may yield a considerable improvement of the reconstructed surface, as it diminishes a major source of error in current SfS techniques.

We suggest an approximate solution of perspective SfS, which locally recovers the first and second order derivatives. From the derivatives, a global approximation of the original 3D surface is obtained. We then compare reconstruction by this algorithm with three existing algorithms.

This paper is organized as follows. Following the presentation of notation and basic assumptions (Sect. 2), we motivate the advantages of assuming a perspective projec-

tion model (Sect. 3). Section 4 then develops the image irradiance equation under the perspective projection model, and explains its dependence on the natural logarithm of the depth function. Section 5 calculates the reflectance map of simple surfaces (planes and paraboloids). Section 6 suggests a two stage SfS algorithm, which uses local approximations of the 3D surface. Section 7 describes the a comparison of the algorithm with three other well-known SfS techniques. Finally, Sect. 8 draws the conclusions.

2. Notation and Assumptions

The following notation and assumptions hold throughout this paper. Photographed surfaces are assumed representable by functions of real-world coordinates as well as of image coordinates. $\hat{z}(x, y)$ denotes the depth function in a real-world Cartesian coordinate system whose origin is at camera plane. If the real-world coordinate $(x, y, \hat{z}(x, y))$ is projected onto image point (u, v), then its depth is denoted $z(u, v)$. By definition, $z(u, v) = \hat{z}(x, y)$. f denotes the focal length, and is assumed known. The scene object is Lambertian, and is illuminated from direction $\vec{L} = (p_s, q_s, -1)$ by a point light source at infinity. $\vec{N}(x, y)$ is surface normal.

3. The Difference Between Perspective and Orthographic Projections — Motivation

We begin by introducing motivation for employing perspective rather than orthographic projection in SfS.

Figure 1(a) shows two quadrilaterals. It can be shown analytically, that perspective projection of the red quadrilateral onto the image plane is identical to orthographic projection of the blue parallelogram. Their images (under identical lighting) would also be the same, as they reside on the same plane, and therefore have identical normals. This stems from the image irradiance equation (see [4]) for a Lambertian surface illuminated by a point light source at infinity: $I(u, v) = \vec{N}(x, y) \cdot \vec{L}$. Consequently, the perspective image of the red quadrilateral is identical to the orthographic image of the blue parallelogram under the same light source (Fig. 1(b)). This implies that if the red quadrilateral was photographed by a perspective camera, but reconstructed by an ideal, orthographic algorithm, the reconstruction would be the blue parallelogram. Thus, the shape difference between the two quadrilaterals is a reconstruction error which is inherent in the orthographic model, and cannot be overcome by any specific orthographic algorithm. Furthermore, it can be proved that orthographic reconstruction of a rectangular image showing a 3D plane must yield a 3D parallelogram; this need not be the case if the projection is perspective, as Fig. 1 shows. (The proof is omitted for brevity.)
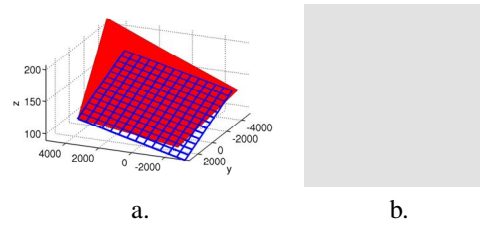

a.   b.

Figure 1: Difference in reconstruction between perspective and orthographic SfS. **a.** Perspective projection of the red quadrilateral is identical to orthographic projection of the blue parallelogram. **b.** The image produced by both surfaces ($\vec{L} = (0, 0.5, -1)$). Orthographic reconstruction of this image would produce a 3D parallelogram.

4. The Perspective Image Irradiance Equation

As a first step in solving the image irradiance equation under the perspective projection model, we convert the equation into more convenient forms.

4.1. Equation in Image Coordinates

The perspective image irradiance equation is given by:

$$I(u, v) = \vec{L} \cdot \vec{N}(x, y) \tag{1}$$

where:

$$x = -\frac{u \cdot \hat{z}(x, y)}{f} \tag{2}$$

$$y = -\frac{v \cdot \hat{z}(x, y)}{f} \tag{3}$$

Substituting Eqs. 2, 3 and $\vec{L} = (p_s, q_s, -1)$ (see Sect. 2) into Eq. 1 yields:

$$I(u, v) = \frac{1 + p_s \hat{z}_x + q_s \hat{z}_y}{\sqrt{1 + p_s^2 + q_s^2}\sqrt{1 + \hat{z}_x^2 + \hat{z}_y^2}} \tag{4}$$

We then express $\hat{z}_x$ and $\hat{z}_y$ in terms of u, v, z, z_u, and z_v, and substitute the resultant expressions along with Eqs. 2, 3 into Eq. 4. Deriving these expressions from the projection equations is nontrivial, but is outside the scope of a conference paper. We obtain:

$$I(u, v) = \frac{(u - f p_s) z_u + (v - f q_s) z_v + z}{\sqrt{1 + p_s^2 + q_s^2}\sqrt{(u z_u + v z_v + z)^2 + f^2(z_u^2 + z_v^2)}} \tag{5}$$

where $z(u, v) \stackrel{def}{=} \hat{z}(x, y)$ for (u, v) which is the perspective projection of $(x, y, \hat{z}(x, y))$. Equation 5 is the *perspective image irradiance equation*.

4.2. Dependence on $\ln(\mathbf{z}(\mathbf{u}, \mathbf{v}))$

Equation 5 shows direct dependence on both $z(u, v)$ and its first order derivatives. If one employs $\ln(z(u, v))$ instead of $z(u, v)$ itself (by definition $z(u, v) > 0$), one obtains the following equation:

$$I(u, v) = \tag{6}$$

$$= \frac{(u - fp_s)p + (v - fq_s)q + 1}{\sqrt{1 + p_s^2 + q_s^2}\sqrt{(up + vq + 1)^2 + f^2(p^2 + q^2)}} \tag{7}$$

where $p \overset{def}{=} \frac{z_u}{z} = \frac{\partial \ln z}{\partial u}$ and $q \overset{def}{=} \frac{z_v}{z} = \frac{\partial \ln z}{\partial v}$. Eq. 7 depends on the derivatives of $\ln(z(u, v))$, but not on $\ln(z(u, v))$ itself. Consequently, the problem of recovering $z(u, v)$ from the image irradiance equation reduces to the problem of recovering the surface $\ln(z(u, v))$ from Eq. 7. Because the natural logarithm is a bijective mapping and $z(u, v) > 0$, recovering $\ln(z(u, v))$ is equivalent to recovering $z(u, v) = e^{\ln(z(u,v))}$.

The image irradiance equation under orthographic projection is invariant to translation of $\widehat{z}(x, y)$, which means $\widehat{z}(x, y) + c$ (for constant c) produces the same intensity function as $\widehat{z}(x, y)$. In contrast, the perspective image irradiance equation (Eq. 5) is invariant to scale changes of $z(u, v)$. That is, the intensity functions of $c \cdot z(u, v)$ and $z(u, v)$ are identical. This follows from the properties of the natural logarithm, and can also be verified by Eqs. 5, 7. Invariance to scaling seems to be a more plausible assumption than invariance to translation when employing real cameras.

5. The Perspective Irradiance Equation of Simple Surfaces

This section examines two types of simple real-world surfaces ($\widehat{z}(x, y)$): planes and paraboloids. We calculate their representation in image coordinate system ($z(u, v)$), and their reflectance map ($R(u, v)$) under the perspective model. These calculations would sharpen the difference between the orthographic and perspective models and would also serve us later for solving the perspective SfS problem in an approximated manner.

5.1. Planes

Let us consider a general plane:

$$\widehat{z}(x, y) = z_0 + a(x - x_0) + b(y - y_0)$$

where a, b, x_0, y_0, z_0 are constants. Substituting image coordinates (u, v) according to the perspective projection equations and solving for $z(u, v)$ yields:

$$z(u, v) = z_0 \frac{f + au_0 + bv_0}{f + au + bv} \tag{8}$$

where $u_0 = -\frac{f \cdot x_0}{z_0}$, $v_0 = -\frac{f \cdot y_0}{z_0}$. The last equation states that the depth of the planar surface at point (u, v) is proportional to the *reciprocal* of $au + bv$. The opposite takes place in orthographic projection: $x \propto u$, $y \propto v$, and hence depth is proportional to $au + bv = ax + by$, by definition of $\widehat{z}(x, y)$.

Under both perspective and orthographic projections, the image irradiance equation becomes (the proof is omitted due to lack of room):

$$R(u, v) = \frac{p_s a + q_s b + 1}{\|\vec{L}\|\sqrt{a^2 + b^2 + 1}} \tag{9}$$

The image irradiance is constant (i.e., independent of u and v) for a planar object.

5.2. Paraboloids

5.2.1. Canonical Paraboloids

We first consider a canonical paraboloid of the form:

$$\widehat{z}(x, y) = ax^2 + by^2$$

Its representation in image coordinates under perspective projection is:

$$z(u, v) = \begin{cases} \frac{f}{au^2 + bv^2}, & \text{if } au^2 + bv^2 \neq 0 \\ 0, & \text{if } au^2 + bv^2 = 0 \end{cases}$$

Again, the perspective and orthographic equations are reciprocal (up to a scale factor).

The reflectance map in this case is:

$$R(u, v) = \frac{2f(p_s au + q_s bv) - (au^2 + bv^2)}{\|\vec{L}\|\sqrt{au^2 + bv^2}\sqrt{au^2 + bv^2 + 4f^2}}$$

5.2.2. General Paraboloids

For a general paraboloid of the form:

$$\widehat{z}(x, y) = z_0 + a(x - x_0) + b(y - y_0) + c(x - x_0)^2 + d(y - y_0)^2 + e(x - x_0)(y - y_0)$$

the image coordinate representation is:

$$z(u, v) = \frac{S(u, v) - \sqrt{S^2(u, v) - 4T(u, v)P}}{2T(u, v)} \tag{10}$$

where:

$$\begin{aligned} T(u, v) &\overset{def}{=} cu^2 + dv^2 + euv \\ S(u, v) &\overset{def}{=} f^2 + u(fa + 2cu_0 z_0 + ev_0 z_0) \\ &\quad + v(fb + 2dv_0 z_0 + eu_0 z_0) \\ P &\overset{def}{=} z_0(f^2 + f(au_0 + bv_0)) \\ &\quad + z_0^2(cu_0^2 + dv_0^2 + eu_0 v_0) \end{aligned}$$

(assuming $T(u, v) \neq 0$). The reflectance formula in this case is omitted due to its complexity. Even though there exists another solution to the quadratic equation, in the general case that solution is not physical. This is because substitution of z_0 into the other solution results in $z(u_0, v_0) \neq z_0$ (unless $f + au_0 + bv_0 = 0$), which contradicts the definition of z_0.

6. Approximation over Patches

This section suggests an approximate solution to the perspective SfS problem. The suggested algorithm divides up into two stages, which we hereby present.

6.1. Stage I — Independent Approximation over Patches

The first stage approximates the original surface locally by paraboloids. Since paraboloids are arbitrarily curved surfaces, they can approximate a small enough patch of $\widehat{z}(x, y)$.

Let us assume that locally $\widehat{z}(x, y)$ can be represented by a paraboloid. As delineated in Sect. 5.2.2, there exists an analytic expression for the reflectance map of a general paraboloid of the form:

$$
\begin{aligned}
\widehat{z}(x, y) = \; & z_0 + a(x - x_0) + b(y - y_0) + c(x - x_0)^2 \\
& + d(y - y_0)^2 + e(x - x_0)(y - y_0)
\end{aligned}
$$

where (x_0, y_0, z_0) is the center of the paraboloid, and a, b, c, d, e are parameters. Using the analytic expression, we minimize the double integral:

$$
\begin{aligned}
F(a, b, c, d, e, z_0, u_0, v_0) \; & \overset{def}{=} \\
= \; & \iint [R(u, v, a, b, c, d, e, z_0, u_0, v_0) - I(u, v)]^2 dudv
\end{aligned}
$$

with respect to parameters a, b, c, d, e. Minimization is currently implemented by gradient descent.

An important advantage of this stage over methods based on Calculus of Variations is a reduced dimension of the minimization space: $\mathbb{R}^5$ (five parameters: a, b, c, d, e) instead of $\mathbb{R}^n$ (where n is the image size). In addition, the dimension of the minimization space becomes independent of image size. This simplifies the functional and yields better convergence of the minimization process.

Another advantage of the paraboloidal approximation is avoiding numerical differentiation, which introduces significant error into calculations at image boundaries. When employing Calculus of Variations, numerical differentiation is inevitable, due to lack of analytic model.

The paraboloidal approximation described above can be parallelized, because the calculation over each patch is independent of other patches. In contrast, Calculus of Variations makes patches depend on their neighbors due to numerical differentiations.

6.2. Stage II — Integration of Approximations over the Patches

Due to the independence of patches in Stage I, the transition between reconstructed surface patches may not be smooth. The goal of Stage II is to ensure the consistency of the outputs of Stage I. For this, it imposes global constraints.

Stage I introduces irregularity in two ways. First, reconstruction over patches may independently converge to dual solutions. This duality is inherent in the solution of the SfS problem (see [1]). The algorithm copes with the duality by a heuristic which flips a solution to its dual based on the algorithm's earlier decisions on neighboring patches. The second phenomenon results from the invariance of the reflectance map to scaling of the depth function. Due to the independence of reconstruction, reconstruction at each patch may converge to a solution with a different scale factor. The goal of Stage II is to obtain a global solution that would have a single scale factor.

The inputs for Stage II are the coefficients of the paraboloids at each patch. At each patch, these coefficients are identical to coefficients of a second order Taylor series expansion of $\widehat{z}(x, y)$, up to a scale factor. Formally, for each patch $P_{i,j}$, there exists a constant $k_{i,j}$, the scale factor, for which:

$$
\begin{aligned}
a_{i,j} &= k_{i,j}\widehat{z}_x && (11) \\
b_{i,j} &= k_{i,j}\widehat{z}_y && (12) \\
c_{i,j} &= \frac{1}{2}k_{i,j}\widehat{z}_{xx} && (13) \\
d_{i,j} &= \frac{1}{2}k_{i,j}\widehat{z}_{yy} && (14) \\
e_{i,j} &= k_{i,j}\widehat{z}_{xy} && (15)
\end{aligned}
$$

where $a_{i,j}$, $b_{i,j}$, $c_{i,j}$, $d_{i,j}$, $e_{i,j}$ are the coefficients a, b, c, d, e at patch (i, j), respectively. The patch $P_{i,j}$ overlaps each of its neighbors by 1 pixel.

In order to obtain a global solution with a single scale factor, Stage II re-scales the reconstruction on each patch. It minimizes the distance between reconstructions of neighboring patches on the boundary between them.

Let the u and v coordinates at patch P_{ij} be: $u_1^j, \ldots, u_n^j$ and $v_1^i, \ldots, v_m^i$, respectively. The functional constraining the boundary of a patch[1] would be:

$$
\begin{aligned}
F = \; & \sum_{l=1}^{n} \left[k_{i,j} z(u_l^j, v_1^i) - k_{i+1,j} z(u_l^j, v_m^{i+1}) \right]^2 \\
& + \sum_{l=1}^{m} \left[k_{i,j} z(u_n^j, v_l^i) - k_{i,j+1} z(u_1^{j+1}, v_l^i) \right]^2
\end{aligned}
$$

---

[1]Note, that these expressions are for internal patches (i.e., patches having 4 neighbors). For the rest of the patches, the constraints at sides of image boundaries should be omitted.

$$+ \sum_{l=1}^{n} \left[k_{i,j} z(u_l^j, v_m^i) - k_{i-1,j} z(u_l^j, v_1^{i-1}) \right]^2$$

$$+ \sum_{l=1}^{m} \left[k_{i,j} z(u_1^j, v_l^i) - k_{i,j-1} z(u_n^{j-1}, v_l^i) \right]^2$$

At the minimum $\frac{\partial F}{\partial k_{i,j}} = 0$, which yields the iterative scheme:

$$
\begin{aligned}
k_{i,j}^s &= \frac{1}{A} \left[k_{i+1,j}^{s-1} \sum_{l=1}^{n} z(u_l^j, v_m^{i+1}) z(u_l^j, v_1^i) \right. \\
&\quad + k_{i,j+1}^{s-1} \sum_{l=1}^{m} z(u_1^{j+1}, v_l^i) z(u_n^j, v_l^i) \\
&\quad + k_{i-1,j}^{s-1} \sum_{l=1}^{n} z(u_l^j, v_1^{i-1}) z(u_l^j, v_m^i) \\
&\quad \left. + k_{i,j-1}^{s-1} \sum_{l=1}^{m} z(u_n^{j-1}, v_l^i) z(u_1^j, v_l^i) \right]
\end{aligned}
$$

where:

$$
\begin{aligned}
A &\overset{def}{=} \sum_{l=1}^{n} \left(z^2(u_l^j, v_1^i) + z^2(u_l^j, v_m^i) \right) \\
&\quad + \sum_{l=1}^{m} \left(z^2(u_n^j, v_l^i) + z^2(u_1^j, v_l^i) \right)
\end{aligned}
$$

and $k_{i,j}^s$ is the value of $k_{i,j}$ in the $s^{\underline{th}}$ step. For all i and j, $k_{i,j}$ is initialized to 1 (i.e., the scale factor at the end of Stage II). After each step we divide all scale factors by the scale factor of a specific patch (e.g., the central), thus fixing the scale factor of this patch at 1. This prevents the trivial solution ($\forall i, j, k_{i,j} = 0$).

7. Experimental Results

7.1. The Experiments

This section describes experiments conducted with the suggested two-stage algorithm.

To evaluate the contribution of perspective SfS, we juxtaposed it with three other SfS algorithms: Lee & Kuo [7], Zheng & Chellappa [14], and the "Fast Marching Method" of Kimmel & Sethian [6]. We selected the first two algorithms for the comparison ensuing the comparative study by Shah et al. [13], which found them best among six well-known SfS algorithms. The algorithm of Kimmel & Sethian is newer than the paper of Shah et al. [13], and is a state-of-the-art method.

We evaluated the performance of the algorithms according to three criteria adopted from Shah et al. [13]: mean depth error, standard deviation of depth error, and mean gradient error. For completeness, we also supply the standard deviation of gradient error, although it is considered not physical.

All input images used in the comparison were synthetic, and were produced from an original surface $\widehat{z}(x, y)$ in the real world. The surface was projected onto plane $[uv]$ according to the perspective projection equations (Eqs. 2, 3). A rectangular area bounded by this projection and symmetric about the optical axis was uniformly sampled. The original surface $\widehat{z}(x, y)$ was then interpolated to the sampling points. The orthographic image irradiance equation then served to create the intensity at each point. This procedure was applied to avoid direct usage of the perspective formula, which the proposed algorithm attempts to recover.

For the algorithms of Lee & Kuo and Zheng & Chellappa, we employed the implementations by Shah et al., and kept their parameters unchanged, except for the number of iterations, where we tested a wider range of values. The number of iterations for Zheng & Chellappa was: 10^k where $k = 1, \ldots, 6$, and for Lee & Kuo[2]: 10, 100, 1000, 10000, 50000. Due to lack of space, we provide the evaluation measures only for the number of iterations which obtained best results according to the mean depth error criterion. For the algorithm of Kimmel & Sethian, we extended the implementation of the "Fast Marching Method" by the Technical University of Munich[3] to accommodate the oblique light source case as well.

7.2. Comparative Evaluation

The comparison was performed on various images. We provide all evaluation measures, but for brevity we can provide merely a single reconstruction example.

Two simple surfaces we examined were: $\widehat{z}(x, y) = 3x^3 + 100$ and $\hat{z}(x, y) = 5y^4 + 100$ (light source direction: $\vec{L} = (0, 0, 1)$). Table 1A,B presents the error measures in the reconstruction of these surfaces by each of the algorithms. According to all measures, perspective SfS reconstructed the surfaces significantly better than the algorithms of Lee & Kuo and Zheng & Chellappa, and its error rates were more than an order of magnitude lower. On these simple surfaces, perspective SfS equates with the algorithm of Kimmel & Sethian, both resulting in an accurate reconstruction of the surfaces.

Fig. 2 presents a synthetic 3D model of a vase, which was employed by Shah et al. in their comparison. There, the projection was orthographic; in our case it is perspective. Similarly to them, we produced two images of this model with light source directions $\vec{L} = (0, 0, 1)$ and $\vec{L} = (1, 0, 1)$

---

[2] The narrower range of iterations for [7] is due to recursive implementation, which exhausted computer memory.

[3] Folkmar Bornemann, Technical University of Munich, WiSe 00/01, 11.12.2000, http://www-m8.mathematik.tu-muenchen.de/m3/teaching/PDE/begleit.html

| $z(x,y) =$	Algorithm:	No. of Iterations:	Mean Depth Err:	Std. Dev. Depth Err:	Mean Grad. Err:	Std. Dev. Grad. Err:
A. $3x^3 + 100$	Zheng & Chellappa:	1000	0.42352	0.32406	2.61627	2.15849
	Lee & Kuo:	50000	0.44826	0.33895	2.38544	2.02043
	Kimmel & Sethian:		0.04653	0.02846	0.21047	0.14556
	Perspective:	12000	0.03107	0.02095	0.16477	0.16158
B. $5y^4 + 100$	Zheng & Chellappa:	10	0.55947	0.46398	2.86092	3.67477
	Lee & Kuo:	100	0.65923	0.69168	3.02086	3.60862
	Kimmel & Sethian:		0.02079	0.01265	0.20461	0.17987
	Perspective:	83000	0.02180	0.01665	0.19086	0.18918
C. Vase $(\vec{L} = (0,0,1))$	Zheng & Chellappa:	100	9.68647	6.33322	12.54433	19.23475
	Lee & Kuo:	1000	9.38542	5.82842	12.92186	17.77593
	Kimmel & Sethian:		7.59557	4.29684	5.47410	11.44314
	Perspective:	275000	5.03358	3.41015	10.07100	17.00462
D. Vase $(\vec{L} = (1,0,1))$	Zheng & Chellappa:	100	7.47118	6.08315	12.09654	15.38705
	Lee & Kuo:	10	5.95292	4.23241	11.39574	13.98842
	Kimmel & Sethian:		6.52139	3.67772	17.03632	15.24355
	Perspective:	226000	5.36811	3.59976	9.80671	15.68329
E. $2\cos(\sqrt{(x^2 + (y-2)^2)})$ $+100$	Zheng & Chellappa:	100	0.84550	0.59294	3.37294	1.98929
	Lee & Kuo:	1000	0.98423	0.57303	3.97802	1.92809
	Kimmel & Sethian:		0.88806	0.61702	3.40021	1.99505
	Perspective:	180000	0.17116	0.13830	1.09588	1.05075

Table 1: Errors in reconstruction.
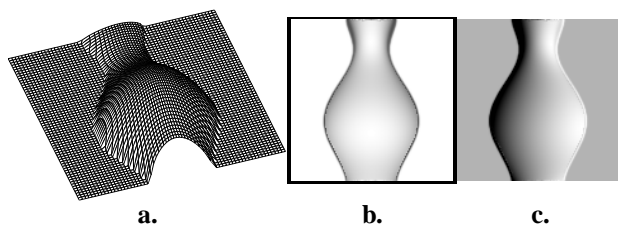
Figure 2: Synthetic vase. **a.** Depth map. **b.** The synthetic image of (a) assuming a Lambertian surface, perspective projection, and light source direction $\vec{L} = (0,0,1)$. **c.** $\vec{L} = (1,0,1)$.
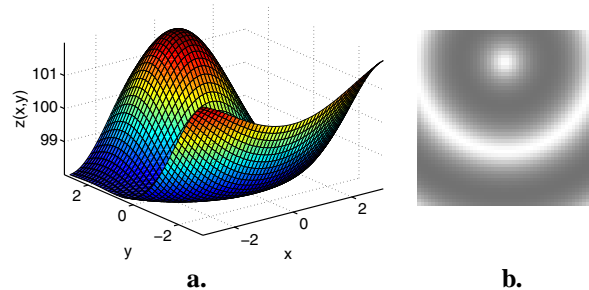

Figure 3: Original surface. **a.** $\hat{z}(x,y) = 2\cos(\sqrt{x^2 + (y-2)^2}) + 100$. **b.** The synthetic image of (a) assuming a Lambertian surface under perspective projection, $\vec{L} = (0,0,1)$.

(Fig. 2 (b),(c)). Table 1C,D presents the evaluation measures for these images. On both images, perspective SfS gained lower error rates than Lee & Kuo and Zheng & Chellappa (except for the standard deviation of gradient for $\vec{L} = (1,0,1)$, but this measure is not physical). For $\vec{L} = (0,0,1)$, perspective SfS achieved higher error rates than Kimmel & Sethian according to gradient measures, but lower error rates according to depth measures. For $\vec{L} = (1,0,1)$, perspective SfS equates with Kimmel & Sethian in the standard deviations, but achieves lower error rates for mean depth error and mean gradient error.

The last example is the trigonometric surface: $\hat{z}(x,y) = 2\cos(\sqrt{x^2 + (y-2)^2}) + 100$. Figure 3 presents the surface and its image. Figure 4 (Top Row) shows the recon-

structed surface from three points of view. The reconstruction shows similarity to the original. It seems upside-down, as the solution converged to the dual of the original surface (an inevitable problem unless more data is incorporated into the basic equation). To verify the similarity, Fig. 4 (Bottom Row) introduces the reconstruction on top of the real surface. A change of sign was applied to the reconstructed surface to avoid the duality of solutions. The surfaces were close to one another, and similar in shape.

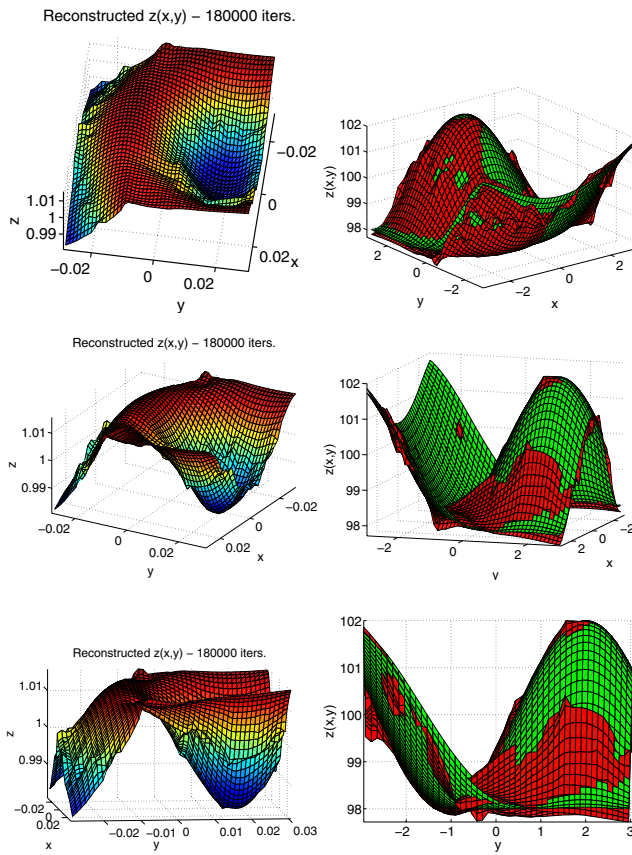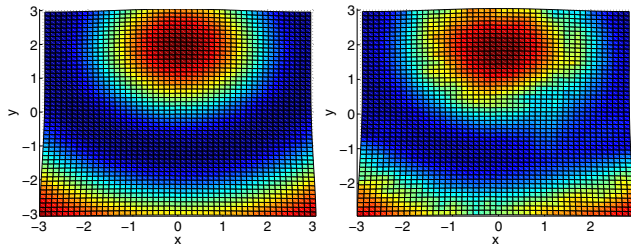Figure 5 compares the boundaries of the original and re-

Figure 4: Perspective reconstruction of surface $\widehat{z}(x,y) = 2\cos(\sqrt{x^2+(y-2)^2})+100$ (Fig. 3). **Left Column:** Reconstruction from three points of view. **Right Column:** Real surface (green) vs. its reconstruction (red), again, from three points of view.

a. Original surface.　　b. Reconstructed surface (by perspective SfS).

Figure 5: Boundaries of original and reconstructed surfaces (orthographic projections of the surfaces on plane $[xy]$). The boundaries are curved, so no orthographic algorithm can recover them correctly. The perspectively reconstructed boundary is similar in shape to the original.

constructed surfaces. In all sides of the surface, the boundaries of the original $[xy]$ domain were curved, not straight lines. This kind of boundaries cannot be recovered by orthographic SfS, because domain boundaries reconstructed orthographically must be straight lines parallel to the axes (for rectangular images). Therefore, orthographic SfS is unable to recover the boundaries of the surface of Fig. 5 correctly. Nevertheless, the perspectively recovered boundaries closely resembled the original.

Figure 6 compares the surfaces reconstructed from $\widehat{z}(x,y) = 2\cos(\sqrt{x^2+(y-2)^2})+100$ by perspective SfS and the other three algorithms. Table 1E provides the error measures. As can be seen from the figure, the distance from the original surface to the reconstructed surfaces of Lee & Kuo and Zheng & Chellappa was notably larger than to the surface recovered perspectively. The algorithm of Kimmel & Sethian reconstructed part of the surface correctly, but the other part was almost flat rather than hill-like (see top row of Fig. 6). Perspective SfS achieved lower error rates than any other method according to all error measures examined. These low error rates obtained by a very simple algorithm (gradient descent) suggest that solving the perspective instead of the orthographic image irradiance equation yields an important improvement in reconstruction.

While the three orthographic algorithms rival the best numerical way to solve the classic equation, the suggested one does not compete in the numerical domain. Instead, it shows the perspective equation is better suited for the task.

In its current implementation, our algorithm converges rather slow and is thus more suitable to offline applications where the accuracy it provides in reconstruction is crucial (while some orthographic ones could be less accurate, but real time). However, further optimization of the implementation might remove this limitation.

8. Conclusions

This research re-examined the roots of the field of Shape-from-Shading, the image irradiance equation. We developed a new formulation of the equation for the case of perspective projection and showed its relation to the logarithm of the depth function. We then suggested an algorithm for reconstruction of images based on the new equation. The algorithm was based on approximation of the original surface using independent paraboloidal patches (Stage I). Following that, global constraints were used (Stage II) in order to make all reconstructed patches refer to a single solution (of the possible dual solutions) and have a common scale factor. Finally, we supplied experimental reconstruction results on synthetic images, and a comparison of the perspective SfS algorithm with three other orthographic SfS methods (Lee & Kuo [7], Zheng & Chellappa [14], and Kimmel & Sethian [6]). It appears that despite the simplicity of the algorithm
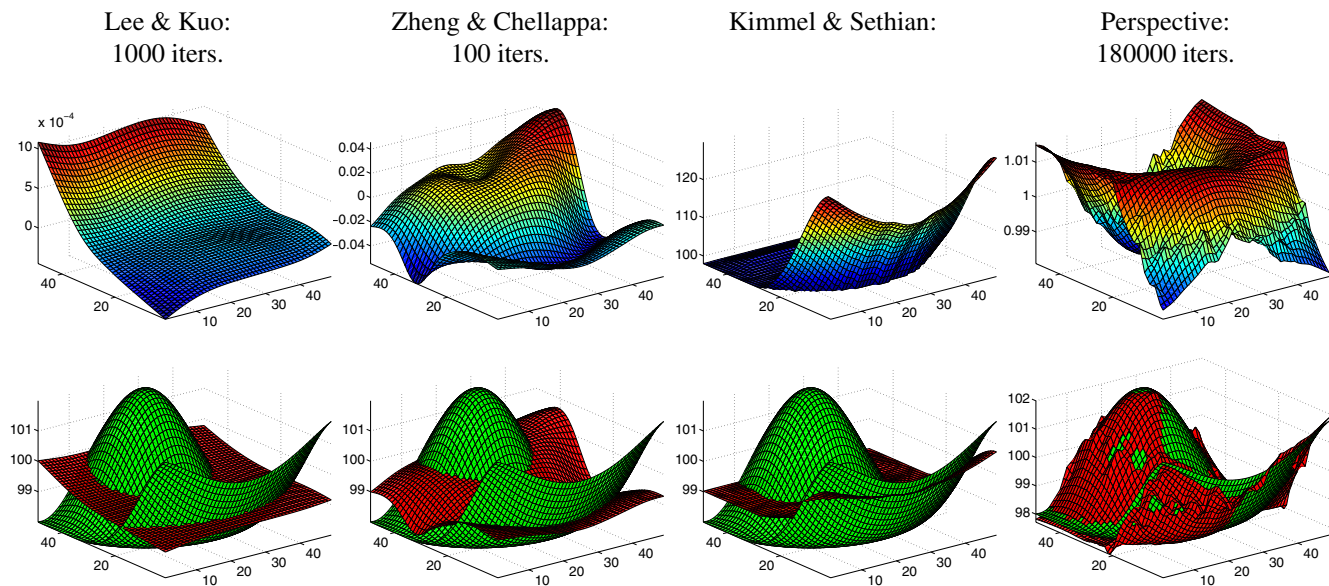
Figure 6: Comparison of perspective SfS with three other algorithms in reconstruction of surface: $\widehat{z}(x,y) = 2\cos(\sqrt{x^2 + (y-2)^2}) + 100$. **Top Row:** Reconstructed surfaces. **Bottom Row:** Comparison with original surface (green — original; red — reconstructed).

(gradient descent), perspective SfS improves performance of legacy SfS techniques, and equates with (and sometimes surpasses) state-of-the-art methods, due to the change of underlying assumptions. Thus, the transition to a more realistic projection model, the perspective model, is the cause of the improvement, rather than the specific numerical algorithm employed.

References

[1] M. J. Brooks and B. K. P. Horn. Shape and source from shading. In B. K. P. Horn and M. J. Brooks, editors, *Shape from Shading*, chapter 3, pages 53–68. The MIT Press, 1989.

[2] A. M. Bruckstein. On shape from shading. *Computer Vision, Graphics and Image Processing*, 44:139–154, 1988.

[3] B. K. P. Horn. Image intensity understanding. *Artificial Intelligence*, 8(2):201–231, Apr. 1977.

[4] B. K. P. Horn. *Robot Vision*. The MIT Press/McGraw-Hill Book Company, 1986.

[5] R. Kimmel and A. M. Bruckstein. Global shape from shading. *Computer Vision and Image Understanding*, 62(3):360–369, 1995.

[6] R. Kimmel and J. A. Sethian. Optimal algorithm for shape from shading and path planning. *Journal of Mathematical Imaging and Vision*, 14(3):237–244, 2001.

[7] K. M. Lee and C.-C. J. Kuo. Shape from shading with a linear triangular element surface model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(8):815–822, Aug. 1993.

[8] K. M. Lee and C.-C. J. Kuo. Shape from Shading with a Generalized Reflectance Map Model. *Computer Vision and Image Understanding*, 67(2):143–160, Aug. 1997.

[9] M. A. Penna. A Shape from Shading Analysis for a Single Perspective Image of a Polyhedron. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):545–54, June 1989.

[10] D. Samaras and D. Metaxas. Coupled Lighting Direction and Shape Estimation from Single Images. *Proceedings of the Seventh IEEE International Conference on Computer Vision*, 2:868–74, 1999.

[11] I. Seong, S. Hideo, and O. Shinji. A Divide-and-conquer Strategy in Shape from Shading Problem. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 413–19, 1997.

[12] S. M. Yamany, A. A. Farag, E. Rickard, D. Tasman, and A. G. Farman. A Robust 3-d Reconstruction System for Human Jaw Modeling. In *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 778–787, 1999.

[13] R. Zhang, P.-S. Tsai, J. E. Cryer, and M. Shah. Shape from shading: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(8):690–705, August 1999.

[14] Q. Zheng and R. Chellappa. Estimation of illuminant direction, albedo, and shape from shading. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):680–702, July 1991.