Toward robust and physically plausible shaded stereoscopic segmentation

Dejun Wang wangdj@cs.ucla.edu
Emmanuel Prados eprados@cs.ucla.edu
Stefano Soatto soatto@cs.ucla.edu

UCLA CSD-TR 050013, Apr.04.2005.

Abstract

We address the multi-view shape from shading problem, that is the recovery of 3-D shape, lighting configuration and surface albedo from multiple calibrated views. Previous approaches to this problem relied on physically impossible illumination models (negative light), and resulted in biased estimates of shape and lighting positions. Furthermore, since the solution involves infinite-dimensional optimization, existing approaches were quite slow. We develop a new model that explicitly enforces positivity in the light sources, and show that it significantly improves the accuracy and robustness relative to existing approaches. Furthermore, we show that the most computationally expensive step in the optimization can actually be solved in closed form. This significantly improves speed of convergence over existing schemes. We illustrate the behavior of our algorithm directly on the same data used by previous authors, so direct comparison is possible.

1 Introduction

We propose a method to estimate the shape, albedo and illumination direction from a collection of multiple views of a Lambertian scene. This problem has been named multi-view shape from shading, although it also relates to stereo reconstruction. It is usually assumed that the view are calibrated, that is the mutual position and orientation of the cameras is known. This problem has been introduced by Jin et al. [5], who have shown that the problem can be formulated as a global optimization task with respect to all the unknowns, but the optimization is ill-posed. Being inspired by some classical technics of the SFS literature (see [8], for example), they have proposed an auxiliary vector field, that can be interpreted as a relaxation of the model constrained, and they have demonstrated their approach on a collection of real and synthetic images. In their work, Jin et al. [5] recovered an illumination model that is not physically plausible. In fact, the illumination consisted of point light sources, some with negative radiance. Therefore, the recovered illumination had no relation to the actual illumination in the scene.

Here we build on the work of Jin et al. and we show improvements in a number of ways. First, we formulate the multi-view shading problem in a way that guarantees that the recovered illumination is physically plausible, i.e. has positive radiance. This model is more complex than [5] because it enforces inequality constraints. Second, we show that one of the steps of the optimization, indeed the most computationally intensive one where the (infinite-dimensional) auxiliary vector field is estimated using a gradient flow, can be actually solved in closed form. This results in significant improvements in speed and robustness. We illustrate our results on the same dataset used by Jin et al. (made available on the web), and show radical improvement both in the estimation of shape, radiance, and the position of the light, all with significantly reduced computational complexity.

While one may argue that such improvements are incremental, the increment is quite significant: we improve existing approaches in accuracy, robustness and speed at the same time. While existing approaches could not be easily integrated with other reconstruction modalities, for instance multi-view stereo, because of the presence of the auxiliary vector field and the non-physical nature of the light, that impinges on the estimate of reflectance, our model has the potential of being integrated with multi-view stereo, at least for Lambertian scenes, since the auxiliary vector field can be solved for in closed-form and therefore can be factored out of the reconstruction process.

1.1 Prior Work and our contribution

The work on shape from shading is generally written in [19, 6, 3, 15]. More specifically, the readers are referred to papers including [9, 14, 20, 11, 7, 10, 16].

Belhumer et al.[17, 1] analyzed the effect of changing lighting on the object appearance for fixed viewpoint. Estimating the light direction and shape in an alternating way was done by Samaras and Metaxas [2], which differs from our work mainly in that we consider multiple views, which enable us to recover the whole shape instead of a depth map only. Using variational methods in shape from shading dates back to the eighties [6, 12]. Yu and Malik [18] showed the work on establishing illumination configuration with known scene geometry and reflectance. In this article, we only consider the Lambertian objects.

A closely related work is that of Jin et al [5],who combine the reconstruction of shape and light configuration together and solve the variational problems via level set methods. This work differs from ours in that the authors allows the light, whatever point light source or ambient term, to be negative, which is not physically possible. They introduced an auxiliary vector field to replace the normal direction for increasing stability, but fails to yield an estimate of the auxiliary vector field, which affect the reconstruction of light configuration.

In this article, our approach mainly differs from [5] at three aspects: first, we add the positive constraint for the light sources of both ambient light and point light source; second, we give a closed-form function for the auxiliary vector field, which is proved strictly in this article; third, we optimize the energy function with respect to the ambient term and point light sources simultaneously, which avoids the local minimum problem in the alternative method. We demonstrate its performance on the experiments both on synthesized data and real data and prove the stability, accurateness and robustness of the proposed algorithm for even sophisticated shapes having sharp changes.

2 Problem Formalization

Let $S \in \mathbb{R}^3$ be a smooth surface. We denote with $X = [x, y, z]^T$ the coordinates of a generic point on S with respect to a fixed reference frame. The goal is to reconstruct the surface S and the light sources from a set of n images $I_i : \Omega_i \to \mathbb{R}, i = 1, ..., n$, where $\Omega_i \subset \mathbb{R}^2$. The intrinsic and extrinsic calibration parameters for each image are assumed to be known [4]. Thus each camera can be modeled as a perspective projection $\pi_i : \mathbb{R}^3 \to \Omega_i; X \to x_i = \pi_i(X) = \pi(X_i)$, where X_i is the coordinates of X in the i-th camera reference frame. X and X_i are related by a rigid body transformation, that is $X_i = R_i\ X + T_i$. We assume that there is a background B covering the field of view of each camera. We also assume B to be a sphere with infinite radius and define the foreground projection to be the region $Q_i = \pi_i(S) \subset \Omega_i$ and denote its complement in Ω_i by Q_i^c. We also define the back-projection $\pi_i^{-1} : \Omega_i \to \mathbb{R}^3$ of x_i onto S, which could be the first intersection point with S of a ray starting from the i-th camera center and passing through x_i.

We assume that both the foreground and background are Lambertian. The radiances are modeled as scalar-valued functions:

$$\rho : S \to \mathbb{R}, \quad \text{and} \quad h : B \to \mathbb{R}, \tag{1}$$

The surface is assumed to have constant albedo, without loss of generality, to be 1. Therefore, the varying image appearance is only generated by the lighting configuration and the scene geometry. We assume that the true light configuration could be approximated by a superposition of two different component: ambient term E_0 and distant positive point light sources.

For distant point light sources, $\rho(X) = k\langle N(X)\ , L\rangle\xi(X)$, where k denotes the intensity of the light source, L the unit vector pointing in the direction of the light, N the surface unit outward normal and $\xi : S \to \{0, 1\}$ the visibility of the light. In the case of convex objects, the visibility is given by $\xi = \mathcal{H}(\langle N, L\rangle)$, where $\mathcal{H}$ denotes the Heaviside step function.

Thus we can write the image formation model as:

$$I(X) = k\langle N(X), L\rangle\xi(X) + E_0. \tag{2}$$

Our model differs from [5] mainly in that we add a positive constraint on both ambient term and intensity of point light source for the sake of guaranteeing physical constraint, that is, $E_0 \geq 0$ and $k \geq 0$. Actually in [5], its light model introduced a negative light (i.e. the intensity k of the light can be negative), which is not physically correct. The proposed light model in this article fulfills the physical constraints and works well for recovering shape and light configuration with contributions on the numerical solution.

3 Optimization Problem formulation

For the sake of recovering the shape and light configuration simultaneously, we propose to minimize the following energy function.

$$E_{total} = E_{data} \ + \ \alpha E_{prior} \ + \ \beta E_{coupling} = \sum_{i=1}^{n} \int_S \chi_i (I_i(\pi_i(X)) - max(\langle kV(X), L\rangle, 0) - E_0)^2 \sigma_i dA$$
$$+ \sum_{i=1}^{n} \int_{Q_i^c} \big(I_i(\pi(X)) - h\big)^2 d\Omega_i + \alpha \int_S dA + \beta \int_S \big(1 - \langle V(X), N(X)\rangle\big) dA.$$

We have two more concepts here: χ_i is the visibility function of a point X on surface S with respect to the i-th camera, i.e. $\chi_i(X) = 1$ for points on S that are visible from the i-th camera and $\chi_i(X) = 0$ otherwise. σ_i represents the change of coordinates from $d\Omega_i$ to dA, i.e $\sigma_i = \frac{d\Omega_i}{dA} = \langle X_i, N\rangle / Z_i^3$. α, β are the coefficients for $E_{prior}, E_{coupling}$. h is the estimation of background intensity.

V is the auxiliary vector field as a relaxed version of normal vector N [5]. V takes the place of the *unit* normal in modeling the shading effects and is supposed to resolve the instability of the approach which directly use normal direction in modeling shading. Although [5] introduced the concept V, it did not work well in optimizing E_{data} with respect to V and even result in the failure in recovering lights. In this article we propose a closed-form solution for V in section 3.2 and verify its great impact on the energy minimization.

3.1 Updating the Surface S

By fixing $V,\ k,\ L,\ E_0,\ h$, we first propose to update shape S to decrease E as the following gradient descent flow:

$$S_t = \Big(\sum_{i=1}^{n} \frac{1}{Z_i^3} \big((I_i - (k\langle V, L\xi\rangle + E_0))^2 - (I_i - h)^2\big) \cdot \big\langle \nabla\chi_i, R_i^T X_i \big\rangle$$
$$- \sum_{i=1}^{n} 2\chi_i \big(I_i - k\langle V, L\xi\rangle - E_0\big)\big(\xi k L^T \nabla_S V R_i^T X_i + k \langle V, L\rangle \langle \nabla\xi, R_i^T X_i\rangle\big) + (2H(\alpha+\beta) - \beta\nabla_S \cdot V)\Big) N.$$

The numerical implementation of the flow is carried out in the level set framework [13]. The next step would be to fix S and minimize E with respect to $V,\ L,\ E_0,\ h$. We divided this step into the following two parts while considering the positive constraints.

3.2 Updating the Auxiliary Vector Field V

Fix $k,\ L,\ E_0,\ h$, for each point X on surface S, we search for V to minimize E . In the minimization process, we need to guarantee that $\|V\| = 1$. [5] proposed to update V with an iterated process. Unfortunately there is no strict proof that the strategy would guarantee the convergence of V. Besides the experiments in this article would verify that its instability would prevent the algorithm from recovering the true shape and light configuration.

We propose to give a closed-form representation for auxiliary vector field V. In order to parameterize V, without loss of generality, we let $V = (p, q, \sqrt{1-p^2-q^2})$. Thus the minimization problem of E over vector field V is transformed into the minimization over the vector field (p, q). The gradient of E over (p, q) is $\nabla E(p, q) = (\frac{dE}{dp}, \frac{dE}{dq})$. For each point X on shape S, the minimization condition is :

$$\begin{cases} \frac{dE}{dp} = 0, \\ \frac{dE}{dq} = 0. \end{cases} \tag{3}$$

Besides, according to the derivation over compound function, we have

$$\begin{cases} 0 = \frac{dE}{dp} = \frac{dE}{dV}|_{V=(p,q,\sqrt{1-p^2-q^2})} \cdot \frac{dV}{dp} \\ \quad\ = \frac{dE}{dV}|_{V=(p,q,\sqrt{1-p^2-q^2})} \cdot (1, 0, -\frac{p}{\sqrt{1-p^2-q^2}})^T, \\ 0 = \frac{dE}{dq} = \frac{dE}{dV}|_{V=(p,q,\sqrt{1-p^2-q^2})} \cdot \frac{dV}{dq} \\ \quad\ = \frac{dE}{dV}|_{V=(p.,q,\sqrt{1-p^2-q^2})} \cdot (0, 1, -\frac{q}{\sqrt{1-p^2-q^2}})^T. \end{cases} \tag{4}$$

Accordingly, there exists a real scalar γ , such that

$$\frac{dE}{dV}|_{V=(p,q,\sqrt{1-p^2-q^2})} = \gamma.(p, q, \sqrt{1-p^2-q^2})^T. \tag{5}$$

According to the definition of E, we have that

$$\frac{dE}{dV}|_{V=(p,q,\sqrt{1-p^2-q^2})} = \sum_{i=1}^{n} \chi_i(2I_i - 2\langle(p, q, \sqrt{1-p^2-q^2})^T, kL\xi\rangle - 2E_0)kL\xi\sigma_i + \beta N.$$

By combining the above two equations, we have:

$$\sum_{i=1}^{n} \chi_i(2I_i - 2\langle(p, q, \sqrt{1-p^2-q^2})^T, kL\xi\rangle - 2E_0)kL\xi\sigma_i + \beta N = \gamma(p, q, \sqrt{1-p^2-q^2}).$$

The above nonlinear equation system for (p, q, γ) is with three unknowns and three equations actually. To solve it, we introduce an equivalent and easier form of the above equation as follows:

$$\begin{aligned}
&\sum_{i=1}^{n} \chi_i(2I_i - 2\langle V, kL\xi\rangle - 2E_0)kL\xi\sigma_i + \beta N = \gamma V, \\
&\quad \Leftrightarrow \gamma V = -\langle V, kL\xi\rangle(kL\xi)(\sum_{i=1}^{n} 2\chi_i\sigma_i) \\
&\qquad +(\beta N + kL\xi \sum_{i=1}^{n} 2\chi_i(I_i - E_0)\sigma_i), \\
&\|V\| = 1, V \in R^3.
\end{aligned} \tag{6}$$

If $\xi = 0$, obviously the solution is $V = \vec{N}/\|\vec{N}\|$; else we let

$$\begin{cases} \vec{G} = (L\xi)(\sum_{i=1}^{n} 2\chi_i\sigma_i), \\ \vec{H} = \beta N + L\xi \sum_{i=1}^{n} 2\chi_i(I_i - E_0)\sigma_i. \end{cases} \tag{7}$$

Thus we could get that

$$\begin{cases} \|(\gamma.I_{3*3} + \vec{G}.L^T)^{-1}\vec{H}\|_2 = 1, \\ V = (\gamma.I_{3*3} + \vec{G}.L^T)^{-1}\vec{H}. \end{cases} \tag{8}$$

In the above Equations, $\vec{G},\ L^T,\ \vec{H}$ are known, thus we could simply get γ via Newton-methods with the initial value γ determined by the previous V and other parameters in the above equation on γ. Actually as for the initialization of γ, since we have known the previous V^*, we could just simply get it by solving the following linear minimization problem

$$argmin_\gamma \|\sum_{i=1}^{n} \chi_i(2I_i - 2\langle V^*, kL\xi\rangle - 2E_0)kL\xi\sigma_i + \beta N - \gamma V^*\|. \tag{9}$$

After calculating γ, we could get V from the second equation of (7), which minimizes E and fulfill the requirement of $\|V\| = 1$.

3.3 Updating $k,\ L$ and E_0

In this process, we fix S, V, and want to minimize E with respect to $k,\ L,\ E_0$ and h.
Minimization with respects to h:
For the sake of minimizing E over h, h should satisfy

$$h = \frac{\sum_{i=1}^{n} \int_{Q_i^c} I_i d\Omega_i}{\sum_{i=1}^{n} \int_{Q_i^c} d\Omega_i}. \tag{10}$$

Minimization with respects to $k,\ L$ and E_0:
In [5], the model introduces a negative light which is not physically correct. In particular, the variable k representing the

intensity of the light is optimized in $\mathbb{R}$. Here we propose a light model which really fulfills the physical constraints. Also, we enforce k to be nonnegative. As a consequence, k must be optimized in $\mathbb{R}^+$. In order to simplify this step, we introduce a variable $\tilde{L} \in \mathbb{R}^3$:

$$\tilde{L} = kL.$$

Thanks to this new variable, we can simply and equivalently rewrite the positivity constraint as:

$$k \geq 0 \Leftrightarrow k\langle N(X), L\rangle \xi(X) = \langle N(X), \tilde{L}\rangle \tilde{\xi}(X),$$

where $\tilde{\xi}(X) = \mathcal{H}(\langle N, \tilde{L}\rangle)$. Thus, instead of processing an optimization on $(k, L) \in S(0,1) \times \mathbb{R}^+$ which may be relatively sophisticated, we process an optimization on $\tilde{L} \in \mathbb{R}^3$ (which is considerably easier to implement).

Let us remind that [5] proposed to retrieve E_0 and $\tilde{L}$ respectively, that is easily trapped into local minimum, i.e. the approach may always keep away from the true intensity of light source all along. To see the failure resulted in by the approach, we conduct comparison experiments.

In this article, we propose to minimize *simultaneously* the energy function with respect to $\tilde{L}$ and E_0 with the positive constraint of $E_0 \geq 0$. We let the Lagrange function be as follows:

$$\begin{aligned} &\sum_{i=1}^{n} \int_S \chi_i (I_i - \langle V, \tilde{L}\tilde{\xi}\rangle - E_0)^2 \sigma_i dA + \sum_{i=1}^{n} \int_{Q_i^c} \big(I_i - h\big)^2 d\Omega_i \\ &+\alpha \int_S dA \ + \ \beta \int_S \big(1 - \langle V, N\rangle\big) dA + \lambda E_0 = 0. \end{aligned} \tag{11}$$

The Kuhn-Tucker condition for the above Lagrange equation is :

$$\begin{cases} 1) E_0 \geq 0, \\ 2) \lambda E_0 = 0, \\ 3) \sum_{i=1}^{n} \int_{Q_i} (I_i - \langle V, \tilde{L}\tilde{\xi}\rangle - E_0) V\tilde{\xi} d\Omega_i = 0, \\ 4) \sum_{i=1}^{n} \int_{Q_i} (I_i - \langle V, \tilde{L}\tilde{\xi}\rangle - E_0) d\Omega_i + \lambda = 0. \end{cases} \tag{12}$$

$\lambda \geq 0$; We now focus on the definition field of the above problem. If the optimization is retrieved within the field of $\{E_0 \mid E_0 > 0\}$. Then $\lambda = 0$. Thus according to (12.3) and (12.4) we have

$$\begin{cases} E_0 = \sum_{i=1}^{n} \int_{Q_i} (I_i - \langle V, \tilde{L}\tilde{\xi}\rangle) d\Omega_i / \sum_{i=1}^{n} \int_{Q_i} d\Omega_i, \\ \tilde{L} = (\sum_{i=1}^{n} \int_{Q_i} VV^T \tilde{\xi} d\Omega_i)^{-1} (\sum_{i=1}^{n} \int_{Q_i} (I_i - E_0) V\tilde{\xi} d\Omega_i). \end{cases} \tag{13}$$

In order to simply the above equations, we let

$$\begin{cases} M = \sum_{i=1}^{n} \int_{Q_i} VV^T \tilde{\xi} d\Omega_i, \\ A = \sum_{i=1}^{n} \int_{Q_i} I_i d\Omega_i, \\ B = \sum_{i=1}^{n} \int_{Q_i} d\Omega_i; \\ V^* = \sum_{i=1}^{n} \int_{Q_i} V\tilde{\xi} d\Omega_i, \\ p^* = \sum_{i=1}^{n} \int_{Q_i} I_i V\tilde{\xi} d\Omega_i. \end{cases} \tag{14}$$

By using the following abbreviations, which are all known, we could simplify the equations as follows:

$$\begin{cases} E_0 = (A - \langle V^*, \tilde{L}\rangle)/B, \\ \tilde{L} = M^{-1}(p^* - E_0 V^*). \end{cases} \tag{15}$$

The above abbreviations are all known. By combining the two equations of the above, we have that :

$$\tilde{L} = M^{-1}(p^* - (A - \langle V^*, \tilde{L}\rangle)V^*/B).$$

So

$$\tilde{L} = \langle V^*, \tilde{L}\rangle (M^{-1}V^*)/B + M^{-1}(p^* - AV^*/B). \tag{16}$$

We introduce two new vector variables to simplify the above equation: one is $\vec{\alpha}$, another is $\vec{\beta}$

$$\vec{\alpha} = (\alpha_1, \alpha_2, \alpha_3) = (M^{-1}V^*)/B; \qquad \vec{\beta} = (\beta_1, \beta_2, \beta_3) = M^{-1}(p^* - AV^*/B). \tag{17}$$

Thus we have

$$(\tilde{L}_1, \tilde{L}_2, \tilde{L}_3)^T = (V_1^* \tilde{L}_1 + V_2^* \tilde{L}_2 + V_3^* \tilde{L}_3)(\alpha_1, \alpha_2, \alpha_3)^T + (\beta_1, \beta_2, \beta_3)^T. \quad (18)$$

We let

$$W_{3*3} = \begin{bmatrix} (V_1^* \alpha_1 - 1) & V_2^* \alpha_1 & V_3^* \alpha_1 \\ V_1^* \alpha_2 & (V_2^* \alpha_2 - 1) & V_3^* \alpha_2 \\ V_1^* \alpha_3 & V_2^* \alpha_3 & (V_3^* \alpha_3 - 1) \end{bmatrix}. \quad (19)$$

Thus we have

$$W_{3*3} \tilde{L} = -\vec{\beta}. \quad (20)$$

The above is a simple 3D linear algebra problem for $\tilde{L}$, we can easily derive $\tilde{L}$. Now, from the representation of E_0 with respect to $\tilde{L}$, we can derive E_0. If it is greater than zero, then it is what we are searching for and thus $\tilde{L}$ is also fixed. Otherwise, we have to look for the minima within the field $\{E_0 \mid E_0 = 0\}$. Then the problem is much easier. We can derive $\tilde{L}$ from (12.3) as follows because $E_0 = 0$.

$$\tilde{L} = \left(\sum_{i=1}^{n} \int_{Q_i} V V^T \tilde{\xi} d\Omega_i \right)^{-1} \left(\sum_{i=1}^{n} \int_{Q_i} I_i V \tilde{\xi} d\Omega_i \right). \quad (21)$$

4 Experiments

For the sake of testing the performance of recovering shape and light configuration, we compare the enhanced stereo algorithm with positive constraints and closed-form of Vector field and the non-enhanced algorithm [5] via experiments on both synthesized data and real data. We first did experiments on a synthesized data: 14 images from different viewpoints for a ball with radius 8 lighted by a directional light source $\tilde{L}$ with direction L(0,0,1) and an ambient term. The 1st one is: The intensity k = 100 and and ambient light source intensity $E_0 = 100$. We also conducted experiments on other light configuration for the synthesized data: one strong point light source of k = 100 with weak ambient term of $E_0 = 30$;

For the 1st experiment on synthesized data, we compare the shape result of the enhanced algorithm and the non-enhanced algorithm in Figure 1, which also compares the vector field, an important factor for evolution. It is shown from figure 1 that the enhanced algorithm leads to much more accurate shape and vector field. We also show the light configuration reconstruction results including the ambient term intensity and the intensity of point light source, in figure 2. With the enhanced algorithm, the light configuration quickly converges to the ground-truth value, while the existing algorithm always keep away from the true value. The experiments indicate that the positive constraints and the simultaneous updating of E_0 and $\tilde{L}$ help to recover the light configuration better. The table shown in figure 4 lists the final results comparison of the light configuration.

For 2nd experiment on different light configuration on synthesized data, we show the iteration process of ambient intensity and intensity of point light in Figure 3. It is given that the enhanced algorithm generates fairly accurate light configuration while the alternative one rarely converges to the true result or even fails to converge. Furthermore, we conduct the experiments on the real data and compare the result in a very direct and obvious way with the existing algorithm. In this experiment, there are 28 calibrated images of a dancer doll data of approximately uniform albedo. The doll is illuminated with a standard fluorescent overhead lamps and by an additional strong light. Figure 5 shows 4 representative views, which shows the strong point light source on the head of the doll.

We ran our enhanced algorithm with one positive light source and one ambient term. We choose 3 representative views to compare the shape results of the two algorithms. The reconstructed object shape is shown in the lower row of Figure6. The comparison result of the existing method is shown in the upper row of Figure 6. ¿From the experiment, the enhanced algorithm gains better performance than the alternative one, especially in the parts including the two hands, the two hair buns, the buttocks and the eyes, where the normal direction of the points vary sharply. The proposed algorithm extracts the shape of the thumbs and the rest four fingers clearly. Besides, the two hair buns and the buttocks are recovered obviously and could be easily discerned with respect to the neighborhood parts. We draw different lines to help to show the difference of performance of the two algorithms.

We carefully analyze the root of the proposed algorithm's better performance over alternative approaches. We find out that in the enhanced algorithm, the closed-form vector field helps to build up an accurate and robust description of normal

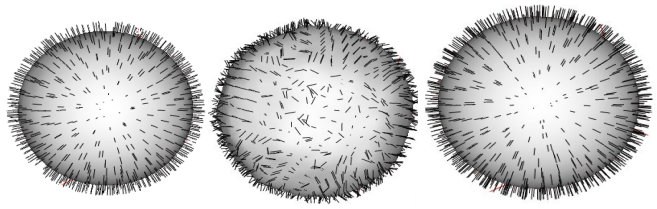

Figure 1: *The final shape and vector field. The left one is of the enhanced algorithm;The middle one is of the non-enhanced algorithm; The right one is of the ground truth object shape and its vector field.*

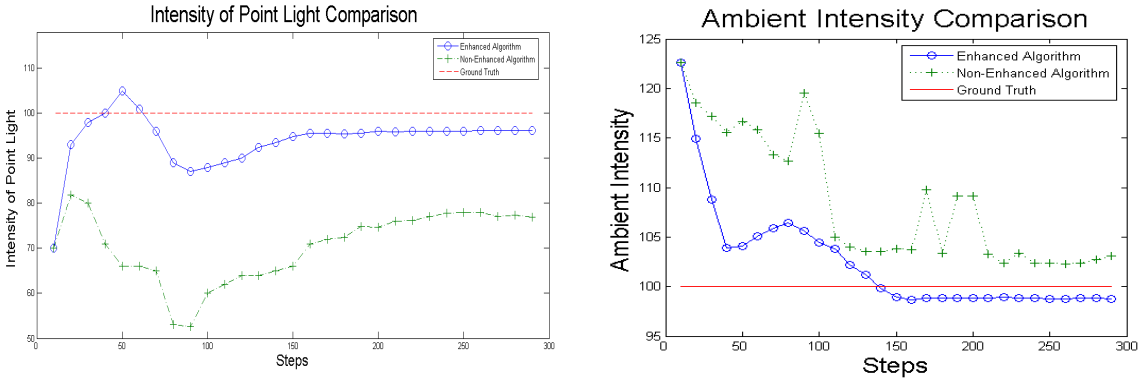

Figure 2: *Ground truth ambient intensity is 100 and the intensity of point light is 100 also. The left figure shows the iteration process of the intensity of point light source; At step 100, the enhanced algorithm converges to satisfying result, while the alternative one keeps away from the ground-truth value. The right figure shows the iteration process of the ambient intensity. At the 150th step, the enhanced algorithm converges to satisfying result, while the alternative one keeps oscillating.*

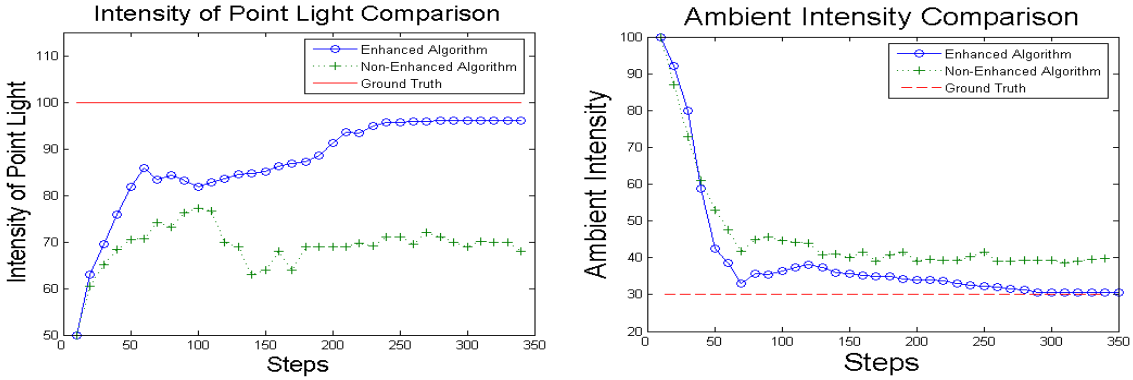

Figure 3: *Ground truth ambient intensity is 30 and the intensity of point light is 100. The left figure shows the iteration process of the intensity of point light source; The enhanced algorithm finally converges to 98.0, while the non-enhanced one keeps far away from the ground-truth all along.The right figure shows the iteration process of the ambient intensity. At the 300th step, the enhanced algorithm converges to satisfying result 29.5, while the alternative one keeps away from the true value ,30.*

direction and thus makes the shapes converge to the ground-truth result. The vector field results of hands,buttocks and hair buns are shown in figure 7 and 8. In further, the final light configuration converges to : Ambient term: 85.1; Directional Light source : 89.5 *(0.009,-0.252, 0.959).

Light parameters	Enhanced Algorithm	Non-enhanced algorithm	Ground Truth.
L	(0.000,0.001,0.999)	(0.015,-0.010,1.000)	(0.000,0.000,1.000)
E_0	98.4	105.2	100.00
Intensity of point light	98.5	77.2	100.00

Figure 4: The table compares the results of light parameters with the enhanced algorithm and the non-enhanced algorithm

Figure 5: *Example views of the input data set consisting 28 views of a dancer model on a table. The head is much brighter than the rest part because the head is facing the spot light.*

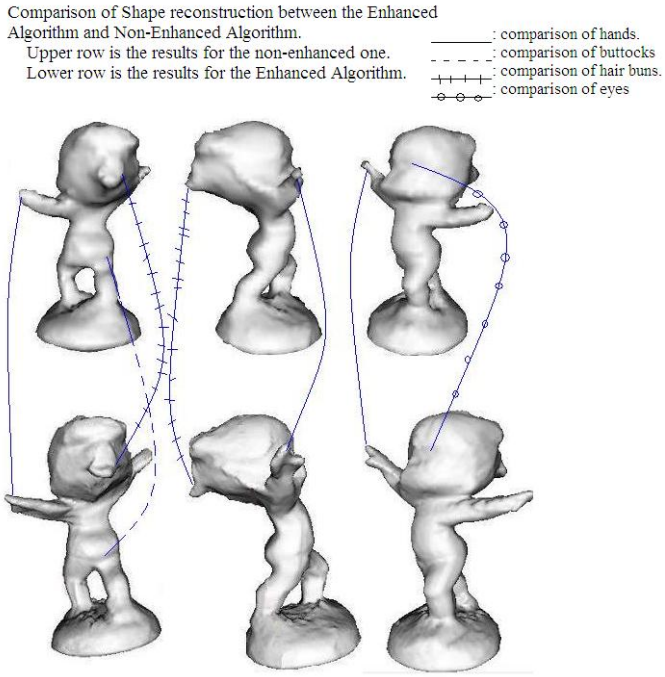

Figure 6: *The comparison of the reconstructed shape of the dancer data between the enhanced algorithm and the non-enhanced algorithm. We give images of the object from 3 representative viewpoints. The upper row is images from the non-enhanced algorithm. The lower one is images from the Enhanced algorithm.*

5 Conclusion

We have introduced an image-formation model for Lambertian scenes viewed from multiple calibrated viewpoint under a collection of point light sources. Unlike previously proposed models, we explicitly enforce positivity of the light sources, and therefore our model is physically plausible. We develop a novel algorithm for the estimation of shape, albedo and lighting position. This algorithm significantly improves the prior art on all grounds: accuracy, robustness and speed. More specifically, we develop a novel solution for the auxiliary vector field that represents the normal to the surface in closed form. This was the most computationally expensive step of existing algorithm. Furthermore, the positivity constraint that

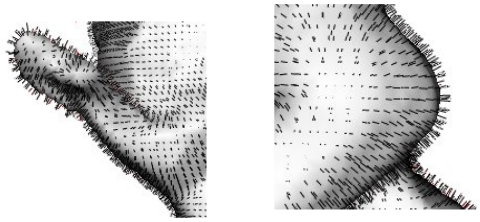

Figure 7: *The vector field result for Hand(Left image) and the buttock(right one)*

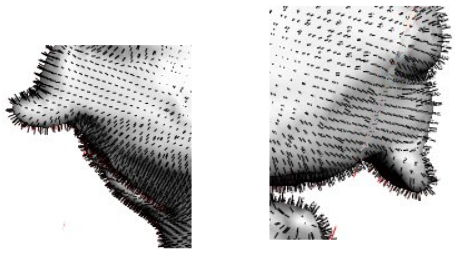

Figure 8: *The vector field result for the Hair buns*

we enforce allows us to recover an illumination model that is physically plausible, and we show that it allows for fairly accurate estimation of lighting position. Of course, in laboratory environments the light is never ideal, as there are inter-reflections, diffuse illumination etc., but our model approximates the real illumination in the best possible way, as measured by the discrepancy between the real images and those generated by our model. We have demonstrated the performance of our algorithm directly on the same data set of previous approaches to the same problem, which shows a significant improvement on all fronts.

References

[1] H. F. Chen, P. N. Belhumeur, and D. W. Jacobs. In search of illumination invariants. In *Proc. IEEE Conf. on Comp. Vision and Pattern Recogn.*, 2000.

[2] D.Samaras. Illumination constraints in deformable models for shape and light direction estimation. *IEEE Trans. Pattern Anal. Mach. Intell.*, pages 247–264, 2003.

[3] J-D. Durou, M. Falcone, and M. Sagona. A survey of numerical methods for shape from shading. Research report 2004-2-R, IRIT, January 2004.

[4] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge University Press, 2000.

[5] H.Jin, D.Cremers, A.Yezzi, and S.Soatto. Shedding light on stereoscopic segmentation. In *Proc. IEEE Conf. on Comp. Vision and Pattern Recogn.*, pages 36–42, 2004.

[6] B. Horn and M. Brooks (eds.). *Shape from Shading*. MIT Press, 1989.

[7] B. Horn, R. Szeliski, and A. Yuille. Impossible shaded images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15(2):166–170, 1993.

[8] B. K.P. Horn. Height and Gradient from Shading. *Int. J. of Computer Vision*, 5(1):37–75, August 1990.

[9] J. Koenderink and A. van Doorn. Photometric invariants related to solid shape. *Optica Acta*, 27(7):981–996, 1980.

[10] M. Langer and S. Zucker. Shape from shading on a cloudy day. *J. Opt. Soc. Am. A*, 11(2):467–478, 1994.

[11] S. Nayar, K. Ikeuchi, and T. Kanade. Surface reflection: physical and geometrical perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(7):611–634, 1991.

[12] J. Oliensis and P. Dupuis. A global algorithm for shape from shading. In *Proc. of the Intl. Conf. on Computer Vision*, pages 692–710, 1993.

[13] S. Osher and J. Sethian. Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi equations. *J. of Comp. Physics*, 79:12–49, 1988.

[14] S. Peleg and G. Ron. Nonlinear multiresolution: a shape from shading example. *IEEE Trans. Pattern Anal. Mach. Intell.*, 1(12):1206–1210, 1990.

[15] E. Prados. *Application of the theory of the viscosity solutions to the Shape From Shading problem*. PhD thesis, Univ. of Nice-Sophia Antipolis, 2004.

[16] H. Ragheb and E.R. Hancock. A probabilistic framework for specular shape-from-shading. *Pattern Recognition*, 36:407–427, 2003.

[17] A. L. Yuille, D. Snow, R. Epstein, and P. Belhumeur. Determining generative models of objects under varying illumination: shape and albedo from multiple images using svd and integrability. *Int. J. of Computer Vision*, 35:203–222, 1999.

[18] Y.Yu and J.Malik. Recovering photometric properties of architectural scenes from photographs. In *Proc. of ACM SIGGRAPH*, pages 207–217, July 1998.

[19] R. Zhang, P. Tsai, J. Cryer, and M. Shah. Analysis of shape from shading techniques. In *Proc. of the Intl. Conf. on Computer Vision and Pattern Recognition*, pages 377–384, 1994.

[20] Q. Zheng and R. Chellappa. Estimation of illuminant direction, albedo and shape from shading. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(7):680–702, 1991.