Shape from Shading: Provably Convergent Algorithms and Uniqueness Results

Paul Dupuis[1] and John Oliensis[2]*

[1] Division of Applied Mathematics,Brown University, Providence, Rhode Island 02912, USA
[2] Department of Computer Science,University of Massachusetts at Amherst, Amherst, Massachusetts 01003, USA

Abstract. An explicit representation for the surface corresponding to a shaded image is presented and proven to be correct (under standard conditions). Uniqueness of the surface is an immediate consequence. Using this representation, various iterative algorithms for shape reconstruction are derived. It has been proven that all these algorithms converge monotonically to the correct surface reconstruction, and they have been shown experimentally to be fast and robust. Some of the results of this paper extend previous ones to the case of illumination from a general direction.

1 Introduction

Shape from shading has proven to be a difficult problem, even under the standard idealizing assumptions of a Lambertian surface, known light source, and no shadowing or occlusion. Recently, a new approach has been developed, based on relating shape from shading to an "equivalent" optimal control problem [7,8,4,9,13,14,1]. The advantages of this approach are both theoretical and practical. First, it makes possible an easy uniqueness proof for the surface corresponding to a shaded image (under standard conditions); thus shape from shading, contrary to previous belief, is often a well-posed problem. Second, the approach leads naturally to an algorithm for surface reconstruction that is simple, fast, provably convergent, and (under standard conditions) provably convergent to the correct surface. In contrast, traditional algorithms typically require thousands of iterations for reconstruction with no guarantee of convergence [5].

In this paper we prove the correctness of an optimal control representation for shape from shading, extending previous work to the case of illumination from a general direction (Section 2). Uniqueness of the surface is an immediate consequence.

From this representation, we derive in Section 3 two provably correct surface reconstruction algorithms (see also [8]). An advantage of our approach is that it gives a great deal of freedom in constructing algorithms, all of which can be proven to produce the same surface approximation. This is useful since one

* This research was supported in part by National Science Foundation grants DMS-9115762, IRI-9113690, and CDA-8922572.

of the presented algorithms has a simple theoretical interpretation and can be used to prove the correctness of the surface reconstruction [3], while the other is more efficient computationally. Note that although the algorithms described here require some information about the imaged surface in order to reconstruct it, they have recently been embedded in a global algorithm capable of surface reconstruction with no a priori information about the surface [7].

2 The Representation Theorem

We consider the idealized problem of shape from shading under the usual assumptions. Note that although we assume Lambertian surface reflectance and illumination from a single direction, our results can be extended easily to any "convex" reflectance function. Under these assumptions, the intensity at an image point $r \equiv (x, y)$ is given by $I(x, y) = \hat{L} \cdot \hat{n}$, where $\hat{L}$ is a unit vector in the light source direction, the optical axis is along the $-z$ direction, and $\hat{n}$ is the surface normal at the corresponding surface point. $I(\cdot)$ is defined on a bounded open subset $\mathcal{D}$ of $\Re^2$. In terms of the surface height function $z(x, y)$, which is assumed continuously differentiable (though this is not essential),

$$\hat{n} \equiv \frac{(-\nabla z, 1)}{(1 + ||\nabla z||^2)^{1/2}}.$$

For illumination from a general direction, we represent the surface not by $z(\cdot)$ but by its height $f(x, y) \equiv \hat{L} \cdot (x, y, z(x, y))$ measured along the light direction $\hat{L}$. Without loss of generality, assume that $L_x = 0$, $L_z > 0$. In terms of $f(\cdot)$, the image irradiance equation can be rewritten after some algebra as $H(r, \nabla f(x)) = 0$, where the Hamiltonian

$$H(r, \alpha) = I(r) \left(1 + ||\alpha||^2 - 2\alpha_y L_y\right)^{1/2} + \alpha_y L_y - 1.$$

Note that $H(r, \alpha)$ is a strictly convex function of α. The fact that the image irradiance equation can be rewritten in terms of a strictly convex H is the essential property used below. Our results can be extended to essentially any image irradiance equation that can be so written.

Singular points—those image points where the intensity achieves its maximal brightness $I(\cdot) = 1$—play a critical role in constraining the surface corresponding to a shaded image [8,11,16,2]. Only at these points is the local surface orientation determined from the intensity alone. Let $\mathcal{S}$ denote the set of singular points in the image. It is easy to see that $\nabla f = 0$ on $\mathcal{S}$, so that $\mathcal{S}$ includes all local maxima and minima of $f(x, y)$. We will focus on those singular points corresponding essentially to the local minima, and use these in determining the surface from its image. (Alternatively, our results could be derived using the local maxima.)

To specify precisely the conditions under which our results hold, we introduce some nonstandard terminology. We say that a set $A \subset \Re^2$ is smoothly connected if given any two points r and r' in A there is an absolutely continuous (" smooth") path connecting the two. We will assume that the set of singular points is a finite collection of smoothly connected sets. Then since $\nabla f = 0$ on $\mathcal{S}$, $f(\cdot)$ is constant over each connected component $\mathcal{S}_C \subset \mathcal{S}$. We will refer to a connected subset $\mathcal{S}_C$ as a set of local minima if there exists an $\varepsilon > 0$ such that $d(r, \mathcal{S}_C) < \varepsilon$ implies $f(r) \geq f(r')$ for $r' \subset \mathcal{S}_C$, i.e., if the "heights" f of nearby points are larger than the value of f on $\mathcal{S}_C$. We will refer to a point as a local minimum of f only if it is contained in such a connected subset $\mathcal{S}_C$. An analogous definition is used for local maxima. Finally, a connected subset that is neither a set of local maxima or local minima will be called a set of saddle points. Let $\mathcal{M}$ be the set of all the local minima in the above sense.

The Lagrangian corresponding to $H(\cdot)$ is:

$$\begin{aligned} L(r, \beta) &= \sup_\alpha \left[-\alpha \cdot \beta - H(r, \alpha)\right] \\ &= L_z^2 - L_y \beta_y - L_z \left(I^2(r) - |\beta_x|^2 - |\beta_y + L_y|^2\right)^{1/2} \end{aligned} \tag{2.1}$$

if $|\beta_x|^2 + |\beta_y + L_y|^2 \leq I^2(r)$ and ∞ if $|\beta_x|^2 + |\beta_y + L_y|^2 > I^2(r)$. Define

$$\mathcal{U}(r) = \left\{\beta : |\beta_x|^2 + |\beta_y + L_y|^2 \leq I^2(r)\right\}.$$

Thus $\mathcal{U}(r)$ is the domain on which $L(r, \cdot)$ is finite. The Lagrangian L serves as the running cost in the "equivalent" optimal control problem, which we now define. Consider an arbitrary path ϕ in the image plane starting at some r, and continuing for a time ρ. More precisely, the path is defined by $\phi(0) = r$, $\dot{\phi} = u(t)$, where the control $u : [0, \infty) \to \Re^2$ is any integrable function. For each such path, we define a cost which is the sum of two terms: 1) the total running cost, given by the integral of the running cost $L(\phi, u(\phi))$ over the path, and 2) a terminal cost, which depends only on the end point of the path. The control problem is to find the path giving the minimal total cost. The representation theorem states that under appropriate conditions the infimal value of the cost for starting point r is just $f(r)$.

Assume we are given an upper bound B for $\{f(r) : r \in \mathcal{D}\}$. Then define the terminal cost

$$g(r) = \begin{cases} f(r) & \text{for } r \in \mathcal{M} \\ B & \text{for } r \notin \mathcal{M} \end{cases} \tag{2.2}$$

The terminal cost imposes the large penalty B on any path terminating at a point $r \notin \mathcal{M}$. Finally, the total cost is the sum of the running and terminal costs

$$V(r) = \inf \left[\int_0^{\rho \wedge \tau} L(\phi(s), u(s)) ds + g(\phi(\rho \wedge \tau))\right], \tag{2.3}$$

where $\rho \wedge \tau$ denotes $\min(\rho, \tau)$, $\tau = \inf\{t : \phi(t) \in \partial\mathcal{D} \cup \mathcal{M}\}$ and the infimum is over all paths ϕ and stopping times $\rho \in [0, \infty)$. Thus, V is the "minimal" cost over all finite time paths, where the path terminates either at time ρ determined by the controller, or else at the first time that the path exits $\mathcal{D}$ or enters $\mathcal{M}$. We want to show that $f(r) = V(r)$.

Preliminaries. For any H, it is easy to show that the definition (2.1) implies that the running cost $L(r, \cdot)$ is convex on $\mathcal{U}(r)$: here it is strictly convex. Moreover, a direct calculation shows that $L(\cdot, \cdot) \geq 0$, and $L(r, \beta) = 0$ only for $r \in \mathcal{S}$

and $\beta = 0$. Also, since $H(r, \cdot)$ is strictly convex, it follows by standard arguments that

$$H(r, \alpha) = \sup_{\beta \in \mathcal{U}(r)} \left[-\alpha \cdot \beta - L(r, \beta) \right], \tag{2.4}$$

and for each $\alpha \in \Re^2$ there exists a unique vector $u(r, \alpha)$ such that

$$H(r, \alpha) = -\alpha \cdot u(r, \alpha) - L(r, u(r, \alpha)).$$

Define $\bar{u}(r)$ for $r \in \mathcal{D}$ by

$$0 = H(r, \nabla f(r)) = -\nabla f(r) \cdot \bar{u}(r) - L(r, \bar{u}(r)). \tag{2.5}$$

From (2.1), $\bar{u}(r)$ is given by

$$\nabla_\alpha H(r, \alpha)|_{\nabla f} = -\bar{u}(r).$$

If (as we assume) $\nabla f(r)$ is continuous, then the fact that $H(r, \cdot)$ is C^1 implies $\bar{u}(r)$ is continuous on $\mathcal{D}$. An explicit calculation shows that $\bar{u}(r)$ is proportional to the projection in the (x, y) plane of the steepest descent direction on the surface [10], where "steepest descent" is defined with respect to the light direction $\hat{L}$, rather than the vertical direction $(0, 0, 1)$.

We consider subsets $\mathcal{G}$ of $\mathcal{D}$ satisfying the following assumption.

A2.1 *Assume that S consists of a finite collection of disjoint, compact, smoothly connected sets, and that $\nabla f(\cdot)$ is continuous on the closure of $\mathcal{D}$. Let $\mathcal{G} \subset \mathcal{D}$ be a compact set, and assume $\mathcal{G}$ is of the form $\mathcal{G} = \cap_{j=1}^J \mathcal{G}_j$, $J < \infty$, where each $\mathcal{G}_j$ has a continuously differentiable boundary. Let $\mathcal{M}$ be the set of local minima of $f(\cdot)$ inside $\mathcal{G}$. Then we assume that the value of $f(\cdot)$ is known at all points in $\mathcal{M}$. Let $\bar{u}$ denote the "steepest descent" direction given by (2.5) above. Define $n_j(r)$ to be the inward (with respect to $\mathcal{G}$) normal to $\partial \mathcal{G}_j$ at r. Then we also assume that $\bar{u}(r) \cdot n_j(r) > 0$ for all $r \in \partial \mathcal{G} \cap \partial \mathcal{G}_j$, $j = 1, ..., J$.*

It will turn out that the minimizing trajectories correspond to paths of steepest descent on the surface. The assumption on $\mathcal{G}$ above thus guarantees that any minimizing trajectory that starts in $\mathcal{G}$ stays in $\mathcal{G}$. When this assumption is violated for some point r, then $f(r)$ cannot be represented as the minimal cost $V(r)$ but may be computable in terms of a maximum cost for an analogous optimal control problem. If neither of these possibilities holds, then the surface at r is not well determined. In general, this will occur only for small image sections near the image boundary [10].

Theorem 2.1 *Assume A2.1, and that B is an upper bound for $f(\cdot)$ on $\mathcal{G}$. Define $L(\cdot, \cdot)$ by (2.1), $g(\cdot)$ by (2.2), and $V(r)$ by (2.3). Then $V(r) = f(r)$ for all $r \in \mathcal{G}$.*

Proof. We first show that $V(r) \geq f(r)$. Let $u(\cdot)$ be any admissible control and define

$$\phi(t) = r + \int_0^t u(s)ds, \ \tau = \inf\{t : \phi(t) \in \partial\mathcal{D} \cap \mathcal{M}\}. \tag{2.6}$$

Since L is the Legendre transform of H and since $H(r, \nabla f(r)) = 0$ for $r \in \mathcal{G}$, $0 \geq -\nabla f(r) \cdot \beta - L(r, \beta)$ for all $\beta \in \Re^2$, and in particular $-\nabla f(\phi(t)) \cdot u(t) \leq L(\phi(t), u(t))$ for $t \in [0, \rho \wedge \tau]$. This implies that

$$-f(\phi(\rho \wedge \tau)) + f(r) = -\int_0^{\rho\wedge\tau} \nabla f(\phi(t)) \cdot u(t)dt \leq \int_0^{\rho\wedge\tau} L(\phi(t), u(t))dt,$$

and thus

$$\int_0^{\rho\wedge\tau} L(\phi(t), u(t))dt + f(\phi(\rho \wedge \tau)) \geq f(r).$$

Since $g(\phi(\rho \wedge \tau)) \geq f(\phi(\rho \wedge \tau))$, we obtain $V(r) \geq f(r)$.

Next we show $V(r) \leq f(r)$. In order to do so we will verify that for each $\varepsilon > 0$ there exists a control $u(\cdot)$ such that for ϕ and τ defined by (2.6) we have

$$\int_0^{\tau} L(\phi(t), u(t))dt + g(\phi(\tau)) \leq f(r) + \varepsilon. \tag{2.7}$$

Let $\mathcal{S}_C$ be a maximal smoothly connected component of $\mathcal{S}$. For any two points r, r' in $\mathcal{S}_C$, there exists a path $\phi(t)$ and a time $t^* < \infty$ such that $\phi(t) \in \mathcal{S}_C$ for $t \in [0, t^*]$, $r = \phi(0)$ and $r' = \phi(t^*)$. Write $\phi(t) = r + \int_0^t u(s)ds$ in terms of the control $u(t)$. Define a new control $u_\lambda(t) \equiv \lambda u(t\lambda)$, where $\lambda > 0$ is a constant, and let $\phi_\lambda(t) = \phi(t\lambda)$ be the corresponding path. Since $L(r, u)/||u|| \to 0$ as $||u|| \to 0$, for r such that $I(r) = 1$, we can choose λ such that

$$\int_0^{\lambda t^*} L(\phi_\lambda(t), u_\lambda(t))dt = \int_0^{t^*} \frac{L(\phi(t), \lambda u(t))}{\lambda} dt \leq \frac{\varepsilon}{3}.$$

Further, since $|L_y| < 1$, there exists $a > 0$ such that for any component $\mathcal{S}_C$ as above, and r such that $d(r, \mathcal{S}_C) \leq a$, we have the following. Let r' be the point in $\mathcal{S}_C$ closest to r. Then there exists a time $t_a \in [0, \infty)$, constant control $u(\cdot) = (r' - r)/t_a$ and corresponding path $\phi(t) = r + \int_0^t u(s)ds$, such that $\phi(t_a) = y$ and $\int_0^{t_a} L(\phi(t), u(t))dt \leq \varepsilon/3$. Finally, this shows that for any $\mathcal{S}_C$, and r, r' such that $d(r, \mathcal{S}_C) \leq a$ and $d(r', \mathcal{S}_C) \leq a$, there exists a control $\tilde{u}_{rr'}(t)$ and time $\sigma_{rr'} \in [0, \infty)$ such that for the corresponding path $\phi_{rr'}(t)$ we have $\phi_{rr'}(0) = r$, $\phi_{rr'}(\sigma_{rr'}) = r'$,

and

$$\int_0^{\sigma_{rr'}} L(\phi_{rr'}(t), \tilde{u}_{rr'}(t))dt \leq \varepsilon.$$

Since f is constant on $\mathcal{S}_C$, then by choosing $a > 0$ smaller if need be we can also assume that $|f(r) - f(r')| \leq \varepsilon$.

We now construct the control that satisfies (2.7). If r is a local minimum then we simply take $\tau = 0$ and are done. There are then two remaining cases: (1) r is contained in some $\mathcal{S}_C$ with $\mathcal{S}_C \cap \mathcal{M} = \emptyset$, or (2) $r \notin \mathcal{S}$. If case (1) holds then $\mathcal{S}_C \cap \mathcal{M} = \emptyset$ implies the existence of a point r' such that $f(r') < f(r)$ and $d(r', \mathcal{S}_C) \leq a$. Since A2.1 implies $\mathcal{S} \subset \mathcal{G}^0$, we can assume that $r' \in \mathcal{G}$. In this case we will set $u(t) = \tilde{u}_{rr'}(t)$ for $t \in [0, \sigma_{rr'})$.

Next consider the definition of the control for $t \geq \sigma_{rr'}$. For $c > 0$ let $b = \inf\{L(r,u) : r \in \mathcal{D}, d(r, \mathcal{S}) > c, u \in \Re^2\}$. The continuity of $I(\cdot)$ and the fact that $I(r) < 1$ for $r \notin \mathcal{S}$ imply $b > 0$. Consider any solution (there may be more than one) to

$$\dot{\phi}(t) = \bar{u}(\phi(t)), \; \phi(0) = r'. \tag{2.8}$$

According to (2.5), for any t such that $\phi(t) \in \mathcal{G} \backslash \mathcal{S}$ and $d(\phi(t), \mathcal{S}) > c$

$$\begin{aligned} \frac{d}{dt} f(\phi(t)) &= \nabla f(\phi(t)) \cdot \bar{u}(\phi(t)) \\ &= -L(\phi(t), \bar{u}(\phi(t))) \\ &\leq -b. \end{aligned} \tag{2.9}$$

A2.1 implies $\phi(t)$ cannot exit $\mathcal{G}$. Thus, since $f(r)$ is bounded on $\mathcal{G}$, (2.9) implies that $\phi(t)$ must enter the set $\{r : d(r, \mathcal{S}) \leq c\}$ in finite time, for any c. If $\phi(t) \in \mathcal{S}$ for some $t < \infty$ we define $\eta_{r'} = \inf\{t : \phi(t) \in \mathcal{S}\}$ and $w = \phi(\eta_{r'})$. Otherwise, let t_i be any sequence tending to ∞ as $i \to \infty$. Since $\mathcal{G}$ is compact we can extract a subsequence (again labeled by i) such that $\phi(t_i) \to v$ for some $v \in \mathcal{S}$. Let $\bar{i}$ be large enough that $||\phi(t_{\bar{i}}) - v|| \leq a$. Since $f(\phi(t_i)) \downarrow f(v)$, we have $f(\phi(t_{\bar{i}})) > f(v)$. For this case we define $\eta_{r'} = t_{\bar{i}}$ and $w = \phi(\eta_{r'})$.

Integrating (2.9) gives

$$f(r') - f(w) = \int_0^{\eta_{r'}} L(\phi(t), \bar{u}(\phi(t)))dt.$$

The control $u(t)$ for $t \in [\sigma_{rr'}, \sigma_{rr'} + \eta_{r'})$ is defined to be $\bar{u}(\phi(t + \sigma_{rr'}))$.

We now consider the point w. We first examine the case in which the solution to (2.8) does not enter $\mathcal{S}$ in finite time. Since $||w - v|| \leq a$, $\tilde{u}_{wv}(t)$ gives a control such that the application of this control moves $\phi(\cdot)$ from w to v with accumulated running cost less than or equal to ε. We define $u(t) = \tilde{u}_{wv}(t - (\eta_{r'} + \sigma_{rr'}))$ for $t \in [\sigma_{rr'} + \eta_{r'}, \sigma_{rr'} + \eta_{r'} + \sigma_{wv})$. If the solution to (2.8) reached $\mathcal{S}$ in finite time we define $w = v$ and $\sigma_{wv} = 0$. Let $\sigma = \sigma_{rr'} + \eta_{r'} + \sigma_{wv}$.

Let us summarize the results of this construction. Given any point $r \in \mathcal{S}$ that is not a local minimum we have constructed a piecewise continuous control $u(\cdot)$ and $\sigma < \infty$ such that if $\phi(t) = r + \int_0^\sigma u(s)ds$, then

$$\begin{aligned} f(r) - f(\phi(\sigma)) &= f(r) - f(r') + f(r') - f(w) + f(w) - f(v) \\ &\geq \int_{\sigma_{rr'}}^{\sigma_{rr'} + \eta_{r'}} L(\phi(t), u(t))dt \\ &\geq -2\varepsilon + \int_0^\sigma L(\phi(t), u(t))dt. \end{aligned}$$

We have also shown that $f(r) > f(v) = f(\phi(\sigma))$, $\phi(\sigma) \in \mathcal{S}$. Thus, either the component $\mathcal{S}_C$ containing $\phi(\sigma)$ satisfies $\mathcal{S}_C \cap \mathcal{M} \neq \emptyset$, and we are essentially done, or we are back into case (1) above, and can repeat the procedure. Let K be the number of disjoint compact connected sets that comprise $\mathcal{S}$. Then the strict decrease $f(r) > f(\phi(\sigma))$ and the fact that $f(\cdot)$ is constant on each $\mathcal{S}_C$ imply the procedure can be repeated no more than K times before reaching some $\mathcal{S}_C$ containing a point from $\mathcal{M}$. If case (2) holds we can use the same procedure, save that the very first step is omitted. Thus, in general, we have exhibited a control $u(\cdot)$ such that

$$\int_0^\tau L(\phi(t), u(t))dt + g(\phi(\tau)) \leq f(r) + (2K+1)\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, the theorem is proved. ■

3 Shape Reconstruction Algorithms

In this section, we describe how algorithms for shape reconstruction can be derived from control representations such as (2.3). It is important to note that many different algorithms can be derived, depending on how the image irradiance equation is rewritten as a Hamiltonian, and that all compute the same surface approximation from the image. Thus, for example, an algorithm can be generated from the Hamiltonian of the previous section, which we henceforth denote by $H^{(1)}$. Another possibility [8] is to write the image irradiance equation in the form $H^{(2)}(r, \nabla f(r)) = 0$, with

$$H^{(2)}(r, \alpha) \equiv \frac{1}{2}\left[I^2\alpha_x^2 + v\alpha_y^2 + 2(1-I^2)L_y\alpha_y - (1-I^2)\right], \tag{3.10}$$

where $v(r) = I^2(r) - L_y^2$. Note that when $v(r) < 0$ $H^{(2)}(r, \alpha)$ is **not** a convex function of α. Nevertheless, an algorithm can be derived from this form of the Hamiltonian, which, although it differs from the algorithm generated from $H^{(1)}$, reconstructs the same surface approximation.

The algorithms are derived using a discrete approximation of the continuous control representation. In this discrete control problem, the object is to minimize the cost over all discrete paths on a grid of pixels. A difficulty in doing this is that a discrete trajectory, in which at each time step the path jumps between neighboring pixels, is generally a poor approximation to a continuous trajectory. In order to better approximate a continuous trajectory on a discrete grid, an element of randomness is introduced. Thus the continuous optimal control problem is approximated by a discrete stochastic optimal control problem, and the cost of the continuous problem is approximated by the expectation of the cost for the discrete problem. Note that the algorithms themselves are deterministic, even though the discrete control problem is stochastic.

Thus, given a control u, we define the probabilities for the path to jump to neighboring pixels so that **on average** the discrete motion approximates the continuous motion $\dot{\phi} = u$. Let $p(r, r'|u)$ denote the transition probability for the path to move from r to a 4-nearest neighbor site r' in the current time step, given that control u is applied. We define

$$p(r, r + \text{ sign}(u_x)(1,0)|u) = \frac{|u_x|}{|u_x| + |u_y|} \tag{3.11}$$

$$p(r, r + \text{ sign}(u_y)(0,1)|u) = \frac{|u_y|}{|u_x| + |u_y|}, \tag{3.12}$$

with all other probabilities zero. We also define the size of the time step to be $\Delta t(u) \equiv 1/(|u_x| + |u_y|)$. With this definition, and assuming for example $u_x, u_y > 0$, the average motion is

$$\frac{(1,0)u_x + (0,1)u_y}{|u_x| + |u_y|} = u\Delta t(u),$$

which approximates the continuous motion. This definition actually makes sense only when $u \neq 0$. For $u = 0$, we define $p(r, r|0) = 1$, and $\Delta t(0) = 1$.

For a given sequence of controls $\{u_i\}$, let $\{\xi_i : \xi_0 = r\}$ denote the path starting at r which evolves at each time step i as determined by the control sequence $\{u_i\}$ and the transition probabilities. Then for the representation and Lagrangian (now denoted $L^{(1)}$) of the previous section, the approximating stochastic control problem is

$$V^{(1)}(r) = \inf E_x \left[\sum_{i=0}^{(N \wedge M)-1} L^{(1)}(\xi_i, u_i)\Delta t(u_i) + g(\xi_{(N \wedge M)}) \right],$$

where $N = \inf\{i : \xi_i \notin \mathcal{D} \text{ or } \xi_i \in \mathcal{M}\}$, and the minimization is over all control sequences $\{u_i\}$ and stopping times M. E_x denotes the expectation. Thus, $V^{(1)}$ is the minimum of the expectation of the cost over all finite length control sequences, where the path terminates either at discrete time M chosen by the controller, or else at the first time that the path exits $\mathcal{D}$ or enters $\mathcal{M}$.

Suppose that instead of considering paths of arbitrary length, we consider paths continuing for at most n time steps:

$$V_n^{(1)}(r) \equiv \inf E_x \left[\sum_{i=0}^{(N \wedge M \wedge n)-1} L(\xi_i, u_i)\Delta t(u_i) + g(\xi_{(N \wedge M \wedge n)}) \right], \tag{3.13}$$

Then $V_n^{(1)}(r)$ is clearly nonincreasing in n and $V_n^{(1)}(r) \downarrow V^{(1)}(r)$ as $n \to \infty$. As discussed in [8], it follows from the principle of dynamic programming that $V_n^{(1)}(r)$ and $V_{n+1}^{(1)}(r)$ are related by

$$V_{n+1}^{(1)}(r) = \min \left[\inf_{u \in \Re^2} \left(L^{(1)}(r,u)\Delta t(u) + \sum_{r'} p(r, r'|u) V_n^{(1)}(r') \right), g(r) \right] \tag{3.14}$$

Clearly, we also have the initial condition $V_0^{(1)}(r) = g(r)$. This, together with the recursive equation (3.14), gives an algorithm which converges monotonically down to V.

For the second control problem, we get a similar algorithm. $V_0^{(2)}(r) = g(r)$, and

$$V_{n+1}^{(2)}(r) = \min\left[\inf_{u\in\Re^2}\left(L^{(2)}(r,u)\Delta t(u) + \sum_{r'} p(r,r'|u)V_n^{(2)}(r')\right), g(r)\right] \text{ or}$$
$$V_{n+1}^{(2)}(r) = \min\left[\sup_{u_y \mathrm{L}_y<0}\ \inf_{u_x\in\Re}\left(L^{(2)}(r,u)\Delta t(u) + \sum_{r'} p(r,r'|u)V_n^{(2)}(r')\right), g(r)\right] \tag{3.15}$$

where the first expression applies when $v(r) \geq 0$ and the second when $v(r) < 0$. The Lagrangian $L^{(2)}$ is derived from an equation analogous to (2.1):

$$\begin{aligned} L^{(2)}(r,\beta) &= \sup_\alpha\left[-\alpha\cdot\beta - H^{(2)}(r,\alpha)\right] && \text{if } v(r) > 0 \\ &= \inf_{\alpha_y}\sup_{\alpha_x}\left[-\alpha\cdot\beta - H^{(2)}(r,\alpha)\right] && \text{if } v(r) < 0. \end{aligned}$$

(The case $v(r) = 0$ is given by the appropriate limit as $v(r) \to 0$ from either direction.) The difference from the previous algorithm is due to the nonconvexity in the Hamiltonian for image regions where $v(r) < 0$. For more detail, and experimental results obtained with the second algorithm above, consult [8, 3].

The algorithms described above are of the Jacobi type, with the surface updated everwhere in parallel at each iteration. The algorithms can also be shown to converge if implemented via Gauss–Seidel, with updated surface estimates used as soon as they are available. In fact, the Gauss–Seidel algorithms converge for any sequence of pixel updates [3]; for example, it is possible to change the direction of the sweep across the image after each pass [1]. Our experiments show that this produces a significant speedup, changing the computation time from order N to order 1 with a small constant, where N is the linear dimension of the image. Specifically, we have proven the following [3]:

Proposition 3.1 *Consider either of the recursive algorithms derived in (3.14) or (3.15). Let an initial condition $V_0^{(a)}$, where $a \in \{1,2\}$, be given and define the sequence $\{V_i^{(a)}, i \in N\}$ according to either the Jacobi iteration [e.g. (3.14)] or the Gauss-Seidel iteration, where the pixel sites are updated in an arbitrary sequence. Assume that $V_0^{(a)}(r) \geq g(r)$ for all $r \in \mathcal{D}$. Then the following conclusions hold.*

1. *For each $r \in \mathcal{D}$, $V_i^{(a)}(r)$ is nonincreasing in i and bounded from below. Define $V^{(a)}(r) = \lim_{i\to\infty} V^{(a)}(r)$. Then the function $V^{(a)}(\cdot)$ is a fixed point of (3.14) (or (3.15) if appropriate).*
2. *The function $V^{(a)}(\cdot)$ can be uniquely characterized as the largest fixed point of (3.14) (or (3.15) if appropriate) that satisfies $V^{(a)}(r) \leq V_0^{(a)}(r)$ for all $r \in \mathcal{D}$.*
3. *A function $w(r)$ is a fixed point of (3.14) if and only if it is a fixed point of (3.15).*

Remark. Thus, taking $V_0^{(a)}(r) = g(r)$, where $g(r)$ is defined by (2.2), both algorithms reconstruct the same surface approximation: $V^{(1)}(r) = V^{(2)}(r)$. The correct surface approximation is obtained by taking the largest of all the fixed points of the iterations in (3.14) or (3.15).

References

1. M. Bichsel, A. P. Pentland, , "A Simple Algorithm for Shape from Shading," *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Champaign, Illinois, pp. 459-465, June 1992.
2. A. R. Bruss, "The Eikonal Equation: Some Results Applicable to Computer Vision," *Journal of Mathematical Physics*, Vol. 23, No. 5, pp. 890-896, May 1982.
3. P. Dupuis and J. Oliensis, "An Optimal Control Formulation and Related Numerical Methods for a Problem in Shape Reconstruction," to appear in *Annals of Applied Probability.*
4. P. Dupuis and J. Oliensis, "Direct Method for Reconstructing Shape from Shading," in *IEEE Computer Vision and Pattern Recognition*, Champaign, Illinois, June 1992, pp. 453-458.
5. B. K. P. Horn and M.J. Brooks (eds.) *Shape from Shading.* MIT Press: Cambridge, MA, 1989.
6. H. J. Kushner and P. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer–Verlag: New York, 1992.
7. J. Oliensis and P. Dupuis, "A Global Algorithm for Shape from Shading," long paper, *Proc. of the Fourth International Conference on Computer Vision*, Berlin, Germany 1993, pp. 692–701.
8. J. Oliensis and P. Dupuis, "Direct Method for Reconstructing Shape from Shading," in *Physics–Based Vision: Principles and Practice, Shape Inference Volume*, L. Wolff, S. Shafer, G. Healey, editors, Jones and Bartlett, Boston, June 1992, pp. 17–28.
9. J. Oliensis and Paul Dupuis, "Direct method for reconstructing shape from shading," *Proc. SPIE Conf. 1570 on Geometric Methods in Computer Vision*, San Diego, California, July 1991, pp. 116-128.
10. J. Oliensis, "Shape from Shading as a Partially Well–Constrained Problem," *Computer Vision, Graphics, and Image Processing: Image Understanding*, Vol. 54, No. 2, September 1991, pp. 163-183.
11. J. Oliensis, "Uniqueness in Shape From Shading," *The International Journal of Computer Vision*, Vol. 6 no. 2, pp. 75-104, 1991.
12. R. T. Rockafellar, *Convex Analysis*, Princeton University Press: Princeton, 1970.
13. E. Rouy, A. Tourin, "A Viscosity Solutions Approach To Shape–From–Shading," *SIAM J. on Numerical Analysis* 29:867-884, 1992.
14. E. Rouy, A. Tourin, "A Viscosity Solutions Approach To Shape–From–Shading," unpublished report.
15. B. V. H. Saxberg, "An Application of Dynamical Systems Theory to Shape From Shading," in *Proc. DARPA Image Understanding Workshop*, Palo Alto, CA, pp. 1089–1104, May 1989.
16. B. V. H. Saxberg, "A Modern Differential Geometric Approach to Shape from Shading," MIT Artificial Intelligence Laboratory, TR 1117, 1989.

This article was processed using the LaTeX macro package with LLNCS style