USC–SIPI REPORT #113

New Algorithms for Reconstruction of a 3-D Depth Map from One or More Images

by

Min Shao, Tal Simchony, and Rama Chellappa

November 1987

Signal and Image Processing Institute
UNIVERSITY OF SOUTHERN CALIFORNIA
Department of Electrical Engineering-Systems
3740 McClintock Avenue, Room 400
Los Angeles, CA 90089-2564 U.S.A.

New Algorithms For Reconstruction Of A 3-D Depth Map From One Or More Images*

M. Shao, T. Simchony and R. Chellappa

November 10, 1987

Signal and Image Processing Institute

Phe 324 , Dept of EE-Systems

Univ Park , MC-0272

University of Southern California

Los Angeles , CA90089-0272 .

Tel: (213)-743-8559

Index terms :

- Shape from shading
- Surface reconstruction

*Partially supported by the NSF Grant MIP-84-51010 and matching funds from IBM and AT&T Information Systems.

New Algorithms For Reconstruction Of A 3-D Depth Map From One Or More Images

Abstract

New algorithms are developed to recover the depth and orientation maps of a surface from its image intensities. They combine the advantages of stereo vision and shape-from-shading (SFS) methods. These algorithms generate dense surface depth and orientation maps accurately and unambiguously. Previous SFS algorithms can not be directly extended to combine stereo images because the recovery of surface depth and that of orientation are separated in these formulations. A new SFS algorithm is proposed to couple the generation of the depth and orientation maps. The new formulation also ensures that the reconstructed surface depth and its orientation are consistent. The SFS algorithm for a single image is next extended to utilize stereo images. The correspondence over stereo images is established simultaneously with the generation of surface depth and orientation. An alternative approach is also suggested for combining stereo and SFS techniques. This approach can be used to combine needle maps which are directly available from other sources such as photometric stereo. Finally we present an algorithm to combine sparse depth measurements with an orientation map to reconstruct a surface. This algorithm, based on transform methods, obtains the solution in $O(N^2 log N)$ operations for an $N \times N$ lattice. The same algorithm can be combined with the above algorithms for solving SFS problem with sparse depth measurements. Thus various information sources can be used to reconstruct a surface accurately.

1 Introduction

One of the goals of a computer-vision system is to recover the three dimensional shape of a surface from its image intensities. There exist several approaches to this problem. Conventional stereo methods [1] are characterized by matching certain feature points in stereo images. As stereo vision can determine only a sparse set of surface depths, it is often followed by a surface interpolation process [2]. The surface orientation can not be recovered directly from matching. The most difficult problem with this method is that of identifying corresponding feature points. Algorithms for recovery of shape from shading have been investigated extensively [3,4,5,6]. Shape-from-shading (SFS) techniques explore the information contained in image intensities by reconstructing a surface that is consistent with observed image intensities. The SFS techniques are formulated for a single image so correspondence is not necessary. The results of SFS algorithms for a single image may not be accurate or robust. Sometimes it is ambiguous to recover a surface from a single image [7]. Thus stereo images are often needed to recover a surface accurately and unambiguously.

In this paper new algorithms are proposed to fill the gap between conventional stereo and SFS techniques. They combine the advantages of stereo vision and SFS methods. SFS techniques are used to recover the depth and orientation maps of a surface from its stereo images which are taken from different viewing directions with fixed light source. Uniform matching is performed over these stereo images in order to obtain dense depth and orientation maps. Both the stereo correspondence and surface depth are established simultaneously under two constraints. The first is the geometrical constraint of stereo vision. The second constraint is provided by the irradiance equations so that the reconstructed surface is everywhere consistent with observed image intensities. These algorithms ensure

full use of shading information and recover both surface depth and orientation. Because stereo images are used in recovering shape from image intensities the solution is accurate, and can not be ambiguous as it can be when a single image is used.

Stereo vision and irradiance equations in SFS problem provide constraints on surface depth and orientation respectively. In order to combine these two methods we need a natural way of incorporating surface depth into orientation constraints. Existing SFS algorithms can not be used for this purpose because all these algorithms recover the orientation map in a separate step, prior to recovering the surface depth. Also some of them do not even enforce the integrability constraint so that the reconstructed surface orientation and depth may not be consistent [6]. Consistent surface reconstruction requires that the reconstructed needle map always correspond to the orientation map of the reconstructed surface depth. In this work we first suggest a new SFS algorithm which provides a novel solution to these two difficulties. It allows a natural incorporation of geometric stereo into SFS methods. The SFS problem is formulated here as that of solving a coupled set of first-order partial differential equations. The first one of these equations is the irradiance equation. The other two enforce the consistency between the reconstructed surface depth and orientation. Because surface depth and orientation are reconstructed simultaneously, stereo images can be easily incorporated.

An alternative algorithm is also suggested to generate surface depth and orientation maps from stereo images by combining stereo method and SFS techniques. In this approach, different needle maps are generated for each image, using the SFS technique proposed in [7]. The depth map can be generated from these needle maps by establishing the correspondence so that the disparity over these needle maps is minimized. The integrability constraint is

also enforced. This approach is related to Ikeuchi's work of combining needle maps [8]. However Ikeuchi's formulation of combining needle maps has some errors. First of all, the orientation of a surface is different with respect to different coordinate systems. The needle maps reconstructed from different images can not be compared directly. Instead they should be mapped into the same coordinate system. Secondly the derivatives of the needle maps used in the formulation should be obtained before the transformation of the needle maps. Although the transformation of needle maps between different coordinates is independent of the position of surface points, the transformation of the derivatives of needle maps is a non-linear function of the surface heights. We will show how to combine needle maps correctly.

Photometric stereo seems to be the only practical solution for obtaining an orientation map for surface with varying albedo. The result may not be accurate because of modeling problems associated with the imaging process. One can improve in accuracy by using sparse depth measurements from geometric stereo . In this paper we show how to combine photometric stereo and geometric stereo information to improve the reconstruction of the surface height.

The organization of the paper is as follows:

Section 2 introduces a new method to recover the depth and orientation maps of a surface simultaneously and consistently. In Section 3 we show how the needle maps of a surface viewed from different directions are related, and extend the results in Section 2 and the algorithm presented in [7] to combine stereo images. The use of sparse information to recover surface is discussed in Section 4, followed by summary in Section 5.

2 Recovery of Depth and Needle Maps From A Single Image

The SFS problem is to extract the shape information from image intensities. Formally, given a 2-D intensity distribution $E(x, y)$, and a reflectance map $R(p, q)$ with constant albedo, it may be regarded as a problem of recovering a surface, $Z(x, y)$, satisfying the *image irradiance equation* :

$$E(x, y) = R(p, q) \tag{1}$$

where

$$p = Z_x \tag{2}$$

$$q = Z_y \tag{3}$$

and $(-p(x, y), -q(x, y), 1)$ is the surface orientation at $(x, y, Z(x, y))$.

Almost all SFS algorithms recover the needle map (p, q) in a separate step, prior to recovering the depth map. The needle map (p, q) is obtained by minimizing the brightness error under the constraint that the surface is smooth. Then the depth Z is recovered from p and q [5]. As p and q are treated as independent variables, the recovered surface needle map (p, q) may not correspond to the orientation of the underlying surface. The integrability constraint [6,9,7] is needed to ensure the solution to the SFS problem is the correct one. This constraint is often expressed in terms of the gradient space as

$$p_y = q_x \tag{4}$$

Horn and Brooks [6] used the constraint as a penalty term in their formulation to enforce the integrability constraint. This formulation is not satisfying for our problem. Mathematically

(2) and (3) do not always imply (4). Most of all, the depth information contained in stereo images is not coupled into the recovery of a needle map. This makes it difficult to generalize the algorithm to combine stereo images.

Instead of using the integrability constraint $p_y = q_x$ as a penalty term, we formulate the SFS problem in a different way so that surface depth and needle maps are coupled and the recovered needle map is always consistent with the reconstructed surface depth.

The above equations (1), (2) and(3) can be considered as a coupled set of first-order differential equations of independent unknown functions $p(x, y)$, $q(x, y)$ and $Z(x, y)$. Equation (1) is the irradiance equation which enforces the recovered surface to correspond to the given image intensities. Equations (2) and (3) are called *consistency constraints.* Consistency constraints ensure that the reconstructed needle map (p, q) always corresponds to the orientation map of the reconstructed surface depth. Thus the SFS problem is now reduced to solving equations (1), (2) and (3) for orientation p and q, and depth Z.

Unfortunately we are not aware of numerically stable methods for solving these nonlinear partial differential equations. So we reformulate the problem as one of finding p, q and Z such that equations (1), (2) and (3) are satisfied under some criterion. Specifically p, q and Z should be chosen to minimize the error functional:

$$\int\int_{\Omega}[(E(x,y) - R(p,q))^2 + (Z_x - p)^2 + (Z_y - q)^2]dxdy \tag{5}$$

Solving for Z, p and q is still an ill-posed problem in the sense of Hadamard [10] as there is no unique solution. To overcome this difficulty we regularize it by assuming that the surface is smooth. According to Ikeuchi and Horn [5], the measure of "lack of smoothness"

is given by

$$\int\int(p_x^2 + p_y^2 + q_x^2 + q_y^2)dxdy \tag{6}$$

Adding this term to the error functional term, one has the following functional to be minimized with respect to p, q and Z:

$$\int\int_\Omega[(E(x,y) - R(p,q))^2 + (Z_x - p)^2 + (Z_y - q)^2 + \lambda(p_x^2 + p_y^2 + q_x^2 + q_y^2)]dxdy \tag{7}$$

Here λ is a weighting factor for the smoothness term.

Using Euler formula [11] one obtains the following equations:

$$\begin{cases} \nabla^2 p = -\frac{1}{\lambda}[(Z_x - p) + (E - R(p,q))R_p] \\ \nabla^2 q = -\frac{1}{\lambda}[(Z_y - q) + (E - R(p,q))R_q] \\ \nabla^2 Z = p_x + q_y \end{cases} \tag{8}$$

where R_p and R_q are the partial derivatives of $R(p,q)$ with respect to p and q. And

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

is the Laplacian operator.

Thus we get a coupled set of non-linear Poisson equations. In order to solve these equations, boundary conditions for p, q and Z are needed. The boundary conditions can be obtained in two steps. First the surface depth along the zero crossing boundaries can be found by Marr-Poggio-Grimson [2] stereo algorithm. Then derivatives of the depth and image intensities can be used to find the surface orientation along the boundaries [8,12].

Because of the non-linear nature of the equations one can not get a closed-form solution.

They are solved by using Jacobbi Picard iterations :

$$\begin{cases} \nabla^2 p^{n+1} = -\frac{1}{\lambda}[(Z_x^n - p^n) + (E - R(p^n, q^n))R_p(p^n, q^n)] \\ \nabla^2 q^{n+1} = -\frac{1}{\lambda}[(Z_y^n - q^n) + (E - R(p^n, q^n))R_q(p^n, q^n)] \\ \nabla^2 Z^{n+1} = p_x^n + q_y^n \end{cases} \tag{9}$$

The iterations can be continued until there is little change in p, q and Z between two consecutive iterations.

At each step of the iterations three Poisson equations have to be solved. We used direct methods [7] to solve these Poisson equations. Experiments with synthetic images show that the direct methods are fast and accurate and that they work on both rectangular and irregular regions. Due to space limitations, we have not included these results.

In this formulation of the SFS problem the integrability constraint is enforced implicitly. The reconstructed surface depth and orientation are always consistent. As the depth Z and the orientation p and q are coupled at every step, it is easy and natural to generalize this algorithm to combine stereo images as described in the following section.

3 Combining Stereo Images

Recovering a surface from a single image is sometimes ambiguous [7]. And the solution of SFS problem is not accurate and robust. Stereo methods use multiple images to overcome these difficulties. They do not make full use of the shading information. Instead they establish correspondence at certain feature points and use geometric relation over stereo images to recover the surface depth. Thus only a sparse set of surface depths can be recovered. In the following we show how SFS techniques can be combined with stereo methods

to overcome the above difficulties. The coupling between surface depth and orientation in the SFS formulation given in Section 2 enables us to combine the geometric constraint on surface depth and the irradiance constraint on orientation. Thus global correspondence can be established over stereo images. Furthermore the surface orientation and depth are recovered simultaneously with the global correspondence.

In the following we first show how the needle maps from stereo images are related. Then a method is presented to combine multiple needle maps to obtain an accurate surface depth. It is found that this technique can be applied to recover surface orientation and depth directly from stereo images by combining the SFS algorithm presented in [7]. As this formulation is highly non-linear we present a simple and elegant algorithm for combining stereo images by extending the results in Section 2. For simplicity all the formulas are derived for the case of two images. They can be easily extended to the case when more than two images are available.

3.1 Camera Set-up and Needle Map Transformation

We use the same camera set-up as in Ikeuchi [13], see Figure 1. The left image plane is perpendicular to the spatial z axis, while the right image plane is inclined with respect to z-axis so that the two optical axes intersect with each other at the origin of the global coordinate system, which is fixed on the surface. Let (u^l,v^l,w^l) and (u^r,v^r,w^r) be the left and right camera coordinate systems and (x, y, z) be the global coordinate system. Assume the object is far away from the cameras so that orthographic projection can be used. Using

the parameters in Figure 1, one has the following coordinate transformations:

$$\begin{bmatrix} u^l \\ v^l \\ w^l \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -d^l \end{bmatrix} \tag{10}$$

$$\begin{bmatrix} u^r \\ v^r \\ w^r \end{bmatrix} = \begin{bmatrix} cos\theta & 0 & -sin\theta \\ 0 & 1 & 0 \\ sin\theta & 0 & cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -d^r \end{bmatrix} \tag{11}$$

$$\begin{bmatrix} u^r \\ v^r \\ w^r \end{bmatrix} = \begin{bmatrix} cos\theta & 0 & -sin\theta \\ 0 & 1 & 0 \\ sin\theta & 0 & cos\theta \end{bmatrix} \begin{bmatrix} u^l \\ v^l \\ w^l \end{bmatrix} + \begin{bmatrix} -d^l sin\theta \\ 0 \\ -d^r + d^l cos\theta \end{bmatrix} \tag{12}$$

By defining a=$cos\theta$, b=$sin\theta$, c=$-d^l\ sin\theta$, d=$-d^r + d^l\ cos\theta$, one obtains

$$\begin{bmatrix} u^r \\ v^r \\ w^r \end{bmatrix} = \begin{bmatrix} a & 0 & -b \\ 0 & 1 & 0 \\ b & 0 & a \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -d^r \end{bmatrix} \tag{13}$$

and

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a & 0 & b \\ 0 & 1 & 0 \\ -b & 0 & a \end{bmatrix} \begin{bmatrix} u^r \\ v^r \\ w^r \end{bmatrix} + \begin{bmatrix} bd^r \\ 0 \\ ad^r \end{bmatrix} \tag{14}$$

Suppose the underlying surface can be expressed in the global coordinate system as:

$$Z = Z(x, y)$$

and the gradient map is

$$p = Z_x(x, y)$$

$$q = Z_y(x, y)$$

Then the orientation of the surface at $(x, y, Z(x, y))$ is $(-p, -q, 1)$.

As the left camera coordinate system is identical to the global coordinate system except the translation along the z-axis, it is easy to see that the gradient in the left camera coordinate system is

$$p^l = p \tag{15}$$

$$q^l = q \tag{16}$$

In the right coordinate system, the orientation at the same point corresponds to different p and q because of the relative rotation of the coordinate systems. The relation between (p^r, q^r) and (p, q) can be found using the coordinate transformation.

Vector $(-p^r, -q^r, 1)$ in the right coordinate system is transformed into a vector in the global coordinate system by

$$\begin{bmatrix} -p^* \\ -q^* \\ z^* \end{bmatrix} = \begin{bmatrix} a & 0 & b \\ 0 & 1 & 0 \\ -b & 0 & a \end{bmatrix} \begin{bmatrix} -p^r \\ -q^r \\ 1 \end{bmatrix} = \begin{bmatrix} -ap^r + b \\ -q^r \\ bp^r + a \end{bmatrix} \tag{17}$$

Comparing with the standard orientation expression $(-p, -q, 1)$, one can find that in the global coordinate system the needle map from the right image can be expressed as

$$\hat{p}^r = \frac{p^*}{z^*} = \frac{ap - b}{a + bp} \tag{18}$$

$$\hat{q}^r = \frac{q^*}{z^*} = \frac{q}{a + bp} \tag{19}$$

Note the surface point $(x, y, Z(x, y))$ is imaged in the right image plane at

$$\begin{cases} u^r = ax - bZ \\ v^r = y \end{cases} \tag{20}$$

As the left camera coordinate system is parallel to the global coordinate system, we prefer to use the global coordinate system instead of the left camera coordinate system in the following subsections.

3.2 Combining Needle Maps

Suppose multiple needle maps for stereo images are available from photometric stereo, how can they be combined into a consistent and accurate depth map? We start from the cost function suggested by Ikeuchi [8,13]:

$$e = \int\int s + \lambda(d_p + d_q)dxdy \tag{21}$$

where

$$s = (Z_x - p^l)^2 + (Z_y - q^l)^2$$

$$d_p = (\hat{p}^r(ax + bZ + c, y) - p^l(x, y))^2$$

$$d_q = (\hat{q}^r(ax + bZ + c, y) - q^l(x, y))^2$$

Values of $\hat{p}^r$ and $\hat{q}^r$ can be obtainted from the needle map for the right image by using the transformations (18) and (19). Thus

$$d_p = \{p^l(x, y) - \hat{p}^r(ax - bZ, y)\}^2 = \{p^l(x, y) - \frac{ap^r(ax - bZ, y) - b}{bp^r(ax - bZ, y) + a}\}^2 \tag{22}$$

$$d_q = (q^l(x, y) - \hat{q}^r(ax - bZ, y))^2 \ = (q^l(x, y) - \frac{q^r(ax - bZ, y)}{bp^r(ax - bZ, y) + a})^2 \tag{23}$$

Ikeuchi [8,13] used p^r and q^r instead of $\hat{p}^r$ and $\hat{q}^r$ in (22) and (23). This is not correct, because the needle maps of a surface in different coordinate systems are generally different.

Note that d_p and d_q are functions of $Z(x,y)$, s in (21) corresponds to the integrability constraint , d_p and d_q couple the information from the two needle maps so that they match each other . The Euler term corresponding to d_p and d_q is :

$$2\{p^l - \frac{ap^r - b}{bp^r + a}\}\frac{ab(bp^r + a)\frac{\partial p^r}{\partial u^r} - b^2(ap^r - b)\frac{\partial p^r}{\partial u^r}}{(bp^r + a)^2} + 2\{q^l - \frac{q^r}{bp^r + a}\}\frac{b(bp^r + a)\frac{\partial q^r}{\partial u^r} - b^2 q^r \frac{\partial p^r}{\partial u^r}}{(bp^r + a)^2}$$

The Euler equation corresponding to (21) is

$$\begin{aligned}\nabla^2 Z = (p^l_x + q^l_y) \quad &+ \quad \lambda[\{p^l - \frac{ap^r - b}{bp^r + a}\}\frac{(a^2b + b^3)\frac{\partial p^r}{\partial u^r}}{(bp^r + a)^2}] \\ &+ \quad \lambda[\{q^l - \frac{q^r}{bp^r + a}\}\frac{b(bp^r + a)\frac{\partial q^r}{\partial u^r} - b^2 q^r \frac{\partial p^r}{\partial u^r}}{(bp^r + a)^2}]\end{aligned}$$

This is an asymmetric formulation based on the left needle map with a correction term from the right needle map. The quantities p^r , q^r , $p^r_{u^r}$ and $q^r_{u^r}$ are functions of $Z(x,y)$ and make up the nonlinear part of the equation. A similar equation can be found to reconstruct the surface depth from the right needle map with a correction term from the left needle map.

If the stereo image intensities are directly available we can couple the needle maps and depth maps at every step. In this case one needle map and one depth map are generated in each coordinate system. But these needle and depth maps are coupled. We follow the ideas presented in [7] , and solve for the right and left depth maps using the algorithm given below:

1. calculate p^l_{n+1} , q^l_{n+1} and p^r_{n+1} , q^r_{n+1} using Lee's [14] algorithm independently for the left and right camera

2. reconstruct the depth in the global coordinate system by solving

$$\nabla^2 Z^l_{n+1} = (p_x + q_y) \quad + \quad \lambda[\{p^l - \frac{ap^r - b}{bp^r + a}\}\frac{(a^2b + b^3)\frac{\partial p^r}{\partial u^r}}{(bp^r + a)^2}]$$
$$+ \quad \lambda[\{q^l - \frac{q^r}{bp^r + a}\}\frac{b(bp^r + a)\frac{\partial q^r}{\partial u^r} - b^2 q^r \frac{\partial p^r}{\partial u^r}}{(bp^r + a)^2}]$$

3. reconstruct the depth in the right camera coordinate system by solving

$$\nabla^2 Z^r_{n+1} = (p^r_u + q^r_v) \quad + \quad \lambda[\{p^r - \frac{ap^l - \hat{b}}{\hat{b}p^l + a}\}\frac{(a^2\hat{b} + \hat{b}^3)\frac{\partial p^l}{\partial x}}{(\hat{b}p^l + a)^2}]$$
$$+ \quad \lambda[\{q^r - \frac{q^l}{\hat{b}p^l + a}\}\frac{\hat{b}(\hat{b}p^l + a)\frac{\partial q^l}{\partial x} - \hat{b}^2 q^l \frac{\partial p^l}{\partial x}}{(\hat{b}p^l + a)^2}]$$

where

$$\hat{b} = -sin\theta, p^l = p^l(au^r - \hat{b}Z^r + \hat{c}, y)\ ,\ \hat{c} = d^r\ sin\theta$$

4. use the central difference approximation to determine $\tilde{p}^l_{n+1}$ and $\tilde{q}^l_{n+1}$ from Z^l_{n+1}, and replace p^l_{n+1} and q^l_{n+1} by $\tilde{p}^l_{n+1}$ and $\tilde{q}^l_{n+1}$.

5. use the central difference approximation to determine $\tilde{p}^r_{n+1}$ and $\tilde{q}^r_{n+1}$ from Z^r_{n+1}, and replace p^r_{n+1} and q^r_{n+1} by $\tilde{p}^r_{n+1}$ and $\tilde{q}^r_{n+1}$.

6. return to stage 1

The suggested algorithm is symmetric, enforces integrability , and couples the information from the two cameras in an early stage . All the steps in the algorithms which involve solving a Poisson equation are implemented using the direct methods [7] .

3.3 A New Method for Coupling Stereo Images

In this section we look at the coupling problem differently. In the previous algorithm two needle maps and two depth maps were generated. And the formulation is highly non-linear.

One can use the global coordinate system as a reference coordinate system to recover the surface orientation and depth from stereo image intensities.

Suppose two images are taken from the left and right cameras respectively, resulting in two irradiance equations. The irradiance equation for the left image is:

$$E^l(u^l, v^l) = R^l(p^l, q^l) \tag{24}$$

By using (10), (15) and (16) one obtains:

$$E^l(x, y) = R^l(p, q) \tag{25}$$

And the irradiance equation for the right image is:

$$E^r(u^r, v^r) = R^r(p^r, q^r) \tag{26}$$

From (18), (19) and (20) the irradiance equation for the right image can be written as:

$$E^r(ax - bZ, y) = R^r(\frac{ap + b}{a - bp}, \frac{q}{a - bp}) \tag{27}$$

Note that R^l, R^r are the same surface reflectance map, except that the viewing direction is different. In the Lambertian case the reflectance map is independent of the viewing direction so

$$R^r(\frac{ap + b}{a - bp}, \frac{q}{a - bp}) = R^l(p, q) = R(p, q) \tag{28}$$

Thus (27) can be reduced to

$$E^r(ax - bZ, y) = R(p, q) \tag{29}$$

For simplicity, in the rest of this section we only consider the Lambertian case. But the formulation can be readily applied to non-Lambertian cases.

Following the algorithm described in Section 2, one formulates the following cost functional:

$$\int\int [(E^l(x,y) - R(p,q))^2 + \alpha(E^r(ax - bZ, y) - R(p,q))^2 + (Z_x - p)^2 + (Z_y - q)^2 + \lambda(p_x^2 + p_y^2 + q_x^2 + q_y^2)]dxdy$$

Note that another weighting factor α is introduced which can control the use of information contained in the other images. For example, when the correspondence is not accurately established we prefer to keep α small.

Using the variational principle one obtains the following Euler equations:

$$\nabla^2 p = -\frac{1}{\lambda}[(Z_x - p) + (E^l(x,y) - R(p,q))R_p + \alpha(E^r(ax - bZ, y) - R(p,q))R_p] \quad (30)$$

$$\nabla^2 q = -\frac{1}{\lambda}[(Z_y - q) + (E^l(x,y) - R(p,q))R_q + \alpha(E^r(ax - bZ, y) - R(p,q))R_q] \quad (31)$$

$$\nabla^2 Z = p_x + q_y + \alpha b E^r_{u^r}(ax - bZ, y)(E^r(ax - bZ, y) - R(p,q)) \quad (32)$$

These equations are also solved using Jacobbi Picard iterations. If there is no correspondence between two images one simply leave out the term $(E^r - R)$.

The correspondence between the two images is established during the generation of the depth map and at the same time it is used to compute the depth map. A good initial estimate is required for the iterative solution. It can be computed by applying the algorithm described in section 2 to the left image to generate the initial depth and needle maps. Thus one can increase α in (30),(31) and (32) from 0 to 1 as the correspondence is increasingly accurate, to control the coupling of stereo images. If the images are too noisy to compute the derivatives of the image intensities, the weighting factor α in (32) can be set to zero. That is, the depth map can be recovered from p and q without directly coupling the stereo images. But α in (30) and (31) need not be zero.

4 Combining Photometric and Geometric Stereo Information

Photometric stereo seems to be the only practical solution for obtaining the orientation map for a surface with varying albedo. Still error occurs in orientation map due to modeling problems of the imaging process. In order to improve the reconstruction of the surface from the needle map, additional information is needed. Sparse measurements obtained from geometrical stereo can serve this purpose [2]. The sparse measurements are available in arbitrary locations on the surface. We use the concept of direct minimization to solve this problem. First the following functional is introduced:

$$\int\int[(Z_x - p)^2 \ + \ (Z_y - q)^2 \ + \ \lambda(Z - Z^m)^2]dxdy \tag{33}$$

where Z^m represents the known depth values obtained from geometric stereo, and

$$\lambda(x, y) = c\sum_{s\in S}\delta(x - x_s, y - y_s)$$

where S is the set of points where measurements exist.

The function that minimizes (33) is chosen as the solution. The presence of the delta function in the integral creates problems in the formulation of the Euler equation. Hence we choose to use a direct method for minimizing the functional. We discretize the integral using finite differences:

$$Z_x = \frac{Z_{i+1,j} - Z_{i,j}}{\Delta x}$$
$$Z_y = \frac{Z_{i,j+1} - Z_{i,j}}{\Delta y}$$

The delta function becomes a Kroneker delta function and one obtain the following cost function:

$$\sum_{i,j}\left\{\left[\frac{(Z_{i+1,j}-Z_{i,j})}{\Delta x}-\frac{p_{i,j}+p_{i+1,j}}{2}\right]^2+\left[\frac{(Z_{i,j+1}-Z_{i,j})}{\Delta y}-\frac{q_{i,j}+q_{i,j+1}}{2}\right]^2\right\}\Delta x\Delta y+ \quad (34)$$
$$\sum_{i,j}\left\{\hat{\lambda}_{i,j}(Z_{i,j}-Z^m_{i,j})^2\right\}$$

where

$$\hat{\lambda}_{i,j}=\begin{cases} c & \text{if } Z^m_{i,j} \text{ exists} \\ 0 & \text{otherwise} \end{cases}$$

Taking the derivative of the cost function (34) with respect to $Z_{i,j}$, $i,j = 1\cdots n$ one obtains a set of linear equations of the form:

$$\alpha^2(Z_{i-1,j}-2Z_{i,j}+Z_{i+1,j})+Z_{i,j-1}-(2+\alpha\hat{\lambda}_{i,j})Z_{i,j}+Z_{i,j+1}=\tilde{f}_{ij} \quad (35)$$

where

$$\tilde{f}_{ij}=\left(\frac{p_{i+1,j}-p_{i-1,j}}{2\Delta x}+\frac{q_{i,j+1}-q_{i,j-1}}{2\Delta y}\right)(\Delta y)^2-\alpha\hat{\lambda}_{i,j}Z^m_{i,j}$$

where $\alpha^2 = (\Delta y/\Delta x)^2$. We now show that Neuman boundary conditions [7] are available for this problem. The normal derivative on the boundary is a linear combination of the Z_x and Z_y. In the case of rectangular boundary contours (like in images), $\frac{dZ}{d\eta} = Z_x$ along the contours $x = constant$ and $\frac{dZ}{d\eta} = Z_y$ along the contours $y = constant$. Brooks and Horn [6] show that using the variational principle on the boundary, we simply have to replace Z_x by its estimate p and Z_y by q. Thus, $\tilde{f}$ is modified for the Neuman boundary conditions. An efficient algorithm, based on transform methods is now derived for solving the set of equations in (35) with Neuman boundary conditions.

Expanding Z_{ij} in terms of sinusoidal basis functions one gets:

$$Z_{ij} = \sum_{n=0}^{N} a_{nj} cos \frac{in\pi}{N} \tag{36}$$

Assuming $\Delta x = \Delta y$ and substituting (36) into (35) we get

$$\sum_{n=0}^{N} [(2cos\frac{n\pi}{N} - 2)a_{nj} + a_{nj+1} - (2 + \hat{\lambda}_{ij})a_{nj} + a_{nj-1}]cos\frac{in\pi}{N} = \tilde{f}_{ij} \tag{37}$$

Expanding $\tilde{f}_{ij}$ in terms of sinusoidal basis functions

$$\tilde{f}_{ij} = \sum_{n=0}^{N} F_{nj} cos\frac{n\pi i}{N} \tag{38}$$

and substituting (38) into (37) and equating coefficients of the cosine terms

$$(2cos\frac{n\pi}{N} - 2)a_{nj} + a_{nj+1} - (2 + \hat{\lambda}_{ij})a_{nj} + a_{nj-1} = F_{nj} \tag{39}$$

The tri-diagonal system of equations in (39) can now be solved using LU decomposition. Since sparse measurements are available the system is no longer singular. The algorithm can now be stated as follows:

1. calculate $\tilde{f}_{ij}$ using p and q and the measurements Z^m modifying $\tilde{f}$ for the Neuman boundary conditions

2. get F_{nj} from $\tilde{f}_{ij}$ using discrete cosine transformation

3. solve the system of tri-diagonal equations for the $a_{nj}s$, taking the additional term in the diagonal element into account when a measurement of Z is available

4. back transform $a \rightarrow Z$

The algorithm solves the problem directly in $O(N^2 logN)$ operations.

5 Summary

Several new algorithms have been presented to recover 3-D surface depth and orientation from image shading information. The SFS algorithm proposed in the paper works in the single image domain. It has several advantages over earlier SFS algorithms. This algorithm can be easily combined with conventional stereo method to utilize stereo images. Starting with the idea of combining multiple needle maps we derived another method to combine stereo vision and SFS techniques. These algorithms have all the advantages of stereo and SFS methods to recover a surface from image intensities. The accuracy of these algorithms can be improved by using sparse depth measurements.

References

[1] D. Marr and T. Poggio, "Understanding Image Intensity", *Proc. Roy. Soc. London*, vol B 204, 1979; see also,"A theory of human stereo vision," AI Lab.,MIT Memo451,1977.

[2] W. E. L. Grimson, *"From Images to Surfaces"*, MIT Press, Cambridge, 1981.

[3] B.K.P. Horn, "Shape-from-Shading:A Method for Obtaing the Shape of a Smooth Opaque Object from a Single View", *Artificial Intelligence Lab,M.I.T*, MAC-TR-79 and AI-TR-232, November 1970.

[4] B.K.P. Horn, "Understanding Image Intensity", *Artificial Intelligence*, vol 8,No. 2, 1977.

[5] K. Ikeuchi and B. Horn, "Numerical Shape from Shading and Occluding Boundaries", *Artificial Intelligence*, vol.17, pp. 141–185, 1981.

[6] B. K. P. Horn and M. J. Brooks, "The Variational Approach to Shape from Shading", *Computer Vision , Graphics , and Image Processing*, vol.33, pp. 174–208, 1986.

[7] T. Simchony and R. Chellappa, "Direct Analytical Methods for Solving Poisson Equations in Computer Vision Problems", In *IEEE Computer Society Workshop on Computer Vision*, Miami Beach,Florida, Nov. 1987.

[8] K. Ikeuchi, "Constructing A Depth Map from Images", *Artificial Intelligence Lab , M.I.T*, A.I.M 744, August 1983.

[9] R. T. Frankot and R. Chellappa, "A Method for Enforcing Integrability in Shape from Shading Algorithms", *IEEE Transactions on Pattern Analysis and Machine Inteligence*, Accepted for Publication, (Also presented at the First Int. Conf. on Computer Vision , London June 1987).

[10] T. Poggio and V. Torre, "Ill posed Problems and Regularization Analysis in Early Vision", *Artificial Intelligence Lab , M.I.T*, A.I.M 773, 1984.

[11] R. Courant and D. Hilbert, *"Methods of Mathematical Physics"*, Volume vol. 1, Interscience, New York, 1953.

[12] A. Blake, "On the Geometric Information Obtainable from Simultaneous Observation of Stereo Contour and Shading", In *Report CSR 205-86*, Department of Computer Science , University of Edinbugh, 1986.

[13] K. Ikeuchi, "Determining a Depth Map Using a Dual Photometric Stereo", *The International Journal of Robotics Research*, vol. 6 , No. 1, pp. 15–31, 1987.

[14] D. Lee, A Provably Convergent Algorithm for Shape from Shading, In *Proc. DARPA Image Undestanding Workshop*, pages 489–496, Miami Beach,Florida, December 1985.

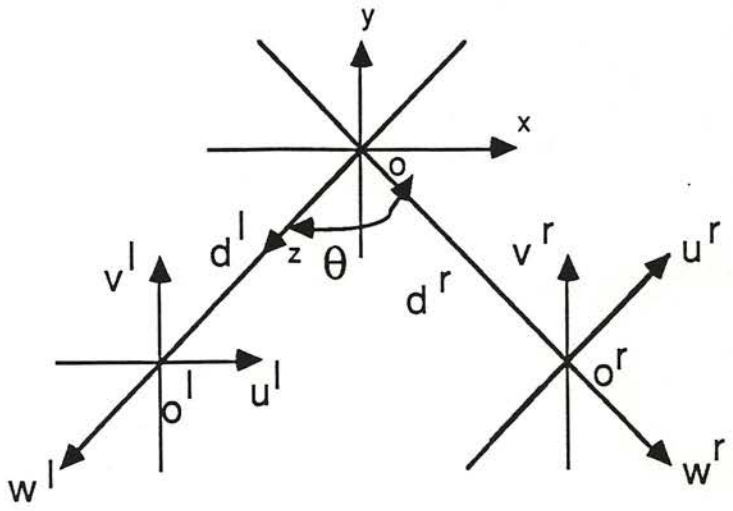

Figure 1. Camera Set-up for Stereo Images

New Algorithms For Reconstruction Of A 3-D Depth Map From One Or More Images*

M. Shao, T. Simchony and R. Chellappa

Signal and Image Processing Institute
Phe 324 , Dept of EE-Systems
University of Southern California
Los Angeles , CA90089-0272 .

Abstract

New algorithms are developed to recover the depth and orientation maps of a surface from its image intensities. They combine the advantages of stereo vision and shape-from-shading (SFS) methods. These algorithms generate dense surface depth and orientation maps accurately and unambiguously. Previous SFS algorithms can not be directly extended to combine stereo images because the recovery of surface depth and that of orientation are separated in these formulations. A new SFS algorithm is proposed to couple the generation of the depth and orientation maps. The new formulation also ensures that the reconstructed surface depth and its orientation are consistent. The SFS algorithm for a single image is next extended to utilize stereo images. The correspondence over stereo images is established simultaneously with the generation of surface depth and orientation. An alternative approach is also suggested for combining stereo and SFS techniques. This approach can be used to combine needle maps which are directly available from other sources such as photometric stereo. Finally, we discuss the use of embedding techniques to combine sparse depth measurements.

1 Introduction

One of the goals of a computer-vision system is to recover the three dimensional shape of a surface from its image intensities. There exist several approaches to this problem. Conventional stereo methods [1] are characterized by matching certain feature points in stereo images. As stereo vision can determine only a sparse set of surface depths, it is often followed by a surface interpolation process [2]. The surface orientation can not be recovered directly from matching. The most difficult problem with this method is that of identifying corresponding feature points. Algorithms for recovery of shape from shading have been investigated extensively [3, 4, 5, 6]. Shape-from-shading (SFS) techniques explore the information contained in image intensities by reconstructing a surface that is consistent with observed image intensities. The SFS techniques are formulated for a single image so correspondence is not necessary. The results of SFS algorithms for a single image may not be accurate or robust. Sometimes it is ambiguous to recover a surface from a single image [7]. Thus stereo images are often needed to recover a surface accurately and unambiguously.

*Partially supported by the NSF Grant MIP-84-51010 and matching funds from IBM and AT&T Information Systems.

In this paper new algorithms are proposed to fill the gap between conventional stereo and SFS techniques. They combine the advantages of stereo vision and SFS methods. SFS techniques are used to recover the depth and orientation maps of a surface from its stereo images which are taken from different viewing directions with fixed light source. Uniform matching is performed over these stereo images in order to obtain dense depth and orientation maps. Both the stereo correspondence and surface depth are established simultaneously under two constraints. The first is the geometrical constraint of stereo vision. The second constraint is provided by the irradiance equations so that the reconstructed surface is everywhere consistent with observed image intensities. These algorithms ensure full use of shading information and recover both surface depth and orientation. Because stereo images are used in recovering shape from image intensities the solution is accurate, and can not be ambiguous as it can be when a single image is used.

Stereo vision and irradiance equations in SFS problem provide constraints on surface depth and orientation respectively. In order to combine these two methods we need a natural way of incorporating surface depth into orientation constraints. Existing SFS algorithms can not be used for this purpose because all these algorithms recover the orientation map in a separate step, prior to recovering the surface depth. Also some of them do not even enforce the integrability constraint so that the reconstructed surface orientation and depth may not be consistent [6]. Consistent surface reconstruction requires that the reconstructed needle map always correspond to the orientation map of the reconstructed surface depth. In this work we first suggest a new SFS algorithm which provides a novel solution to these two difficulties. It allows a natural incorporation of geometric stereo into SFS methods. The SFS problem is formulated here as that of solving a coupled set of first-order partial differential equations. The first one of these equations is the irradiance equation. The other two enforce the consistency between the reconstructed surface depth and orientation. Because surface depth and orientation are reconstructed simultaneously, stereo images can be easily incorporated.

An alternative algorithm is also suggested to generate surface depth and orientation maps from stereo images by combining stereo method and SFS techniques. In this approach, different needle maps are generated for each image, using the SFS technique proposed in [7]. The depth map can be generated from these needle maps by establishing the correspondence so that the disparity over these needle maps is minimized. The

integrability constraint is also enforced. This approach is related to Ikeuchi's work of combining needle maps [8]. However Ikeuchi's formulation of combining needle maps has some errors. First of all, the orientation of a surface is different with respect to different coordinate systems. The needle maps reconstructed from different images can not be compared directly. Instead they should be mapped into the same coordinate system. Secondly the derivatives of the needle maps used in the formulation should be obtained before the transformation of the needle maps. Although the transformation of needle maps between different coordinates is independent of the position of surface points, the transformation of the derivatives of needle maps is a non-linear function of the surface heights. We will show how to combine needle maps correctly.

Photometric stereo seems to be the only practical solution for obtaining an orientation map for surface with varying albedo. The result may not be accurate because of modeling problems associated with the imaging process. One can improve in accuracy by using sparse depth measurements from geometric stereo . In this paper we show how to combine photometric stereo and geometric stereo information to improve the reconstruction of the surface height.

The organization of the paper is as follows:

Section 2 introduces a new method to recover the depth and orientation maps of a surface simultaneously and consistently. In Section 3 we show how the needle maps of a surface viewed from different directions are related, and extend the results in Section 2 and the algorithm presented in [7] to combine stereo images. The use of sparse information to recover surface is discussed in Section 4, followed by a summary in Section 5.

2 Recovery of Depth and Needle Maps From A Single Image

The SFS problem is to extract the shape information from image intensities. Formally, given a 2-D intensity distribution $E(x,y)$, and a reflectance map $R(p,q)$ with constant albedo, it may be regarded as a problem of recovering a surface, $Z(x,y)$, satisfying the *image irradiance equation* :

$$E(x,y) = R(p,q) \tag{1}$$

where

$$p = Z_x \tag{2}$$

$$q = Z_y \tag{3}$$

and $(-p(x,y), -q(x,y), 1)$ is the orientation of the surface at $(x,y,Z(x,y))$.

Almost all SFS algorithms recover the needle map (p,q) in a separate step, prior to recovering the depth map. The needle map (p,q) is obtained by minimizing the brightness error under the constraint that the surface is smooth. Then the depth Z is recovered from p and q [5]. As p and q are treated as independent variables, the recovered surface needle map (p,q) may not correspond to the orientation of the underlying surface. The integrability constraint [6, 9, 7] is needed to ensure the solution to the SFS problem is the correct one. This constraint is often expressed in terms of the gradient space as

$$p_y = q_x \tag{4}$$

Horn and Brooks [6] used the constraint as a penalty term in their formulation to enforce the integrability constraint. This formulation is not satisfying for our problem. Mathematically (2) and (3) do not always imply (4). Most of all, the depth information contained in stereo images is not coupled into the recovery of a needle map. This makes it difficult to generalize the algorithm to combine stereo images.

Instead of using the integrability constraint $p_y = q_x$ as a penalty term, we formulate the SFS problem in a different way so that surface depth and needle maps are coupled and the recovered needle map is always consistent with the reconstructed surface depth.

The above equations (1), (2) and(3) can be considered as a coupled set of first-order differential equations of independent unknown functions $p(x,y)$, $q(x,y)$ and $Z(x,y)$. Equation (1) is the irradiance equation which enforces the recovered surface to correspond to the given image intensities. Equations (2) and (3) are called *consistency constraints*. Consistency constraints ensure that the reconstructed needle map (p,q) always corresponds to the orientation map of the reconstructed surface depth. Thus the SFS problem is now reduced to solving equations (1), (2) and (3) for orientation p and q, and depth Z.

Instead of directly solving these nonlinear equations, we reformulate the problem as one of finding p, q and Z such that equations (1), (2) and (3) are satisfied under some criterion. Specifically p, q and Z should be chosen to minimize the error functional:

$$\int\int_\Omega [(E(x,y) - R(p,q))^2 + (Z_x - p)^2 + (Z_y - q)^2]dxdy \tag{5}$$

Solving for Z, p and q is still an ill-posed problem in the sense of Hadamard [10] as there is no unique solution. To overcome this difficulty we regularize it by assuming that the surface is smooth. According to Ikeuchi and Horn [5], the measure of "lack of smoothness" is given by

$$\int\int (p_x^2 + p_y^2 + q_x^2 + q_y^2)dxdy \tag{6}$$

Adding this term to the error functional term, one has the following functional to be minimized with respect to p, q and Z:

$$\int\int_\Omega [(E(x,y) - R(p,q))^2 + (Z_x - p)^2 + (Z_y - q)^2 + \lambda(p_x^2 + p_y^2 + q_x^2 + q_y^2)]dxdy \tag{7}$$

Here λ is a weighting factor for the smoothness term.

Using Euler formula [11] one obtains the following equations:

$$\begin{cases} \nabla^2 p = -\frac{1}{\lambda}[(Z_x - p) + (E - R(p,q))R_p] \\ \nabla^2 q = -\frac{1}{\lambda}[(Z_y - q) + (E - R(p,q))R_q] \\ \nabla^2 Z = p_x + q_y \end{cases} \tag{8}$$

where R_p and R_q are the partial derivatives of $R(p,q)$ with respect to p and q. And

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

is the Laplacian operator.

Thus we get a coupled set of non-linear Poisson equations. In order to solve these equations, boundary conditions for p, q and Z are needed. The boundary conditions can be obtained in two steps. First the surface depth along the zero crossing boundaries can be found by Marr-Poggio-Grimson [2] stereo algorithm. Then derivatives of the depth and image intensities can be used to find the surface orientation along the boundaries [8, 12].

Because of the non-linear nature of the equations one can not get a closed-form solution. They are solved by using Jacobi Picard iterations :

$$\begin{cases} \nabla^2 p^{n+1} = -\frac{1}{\lambda}[(Z_x^n - p^n) + (E - R(p^n, q^n))R_p(p^n, q^n)] \\ \nabla^2 q^{n+1} = -\frac{1}{\lambda}[(Z_y^n - q^n) + (E - R(p^n, q^n))R_q(p^n, q^n)] \\ \nabla^2 Z^{n+1} = p_x^n + q_y^n \end{cases} \tag{9}$$

The iterations can be continued until there is little change in p, q and Z between two consecutive iterations.

At each step of the iterations three Poisson equations have to be solved. We used direct methods [7] to solve these Poisson equations. Experiments with synthetic images show that the direct methods are fast and accurate and that they work on both rectangular and irregular regions. Due to space limitations, we have not included these results.

In this formulation of the SFS problem the integrability constraint is enforced implicitly. The reconstructed surface depth and orientation are always consistent. As the depth Z and the orientation p and q are coupled at every step, it is easy and natural to generalize this algorithm to combine stereo images as described in the following section.

3 Combining Stereo Images

Recovering a surface from a single image is sometimes ambiguous [7]. And the solution of SFS problem is not accurate and robust. Stereo methods use multiple images to overcome these difficulties. They do not make full use of the shading information. Instead they establish correspondence at certain feature points and use geometric relation over stereo images to recover the surface depth. Thus only a sparse set of surface depths can be recovered. In the following we show how SFS techniques can be combined with stereo methods to overcome the above difficulties. The coupling between surface depth and orientation in the SFS formulation given in Section 2 enables us to combine the geometric constraint on surface depth and the irradiance constraint on orientation. Thus global correspondence can be established over stereo images. Furthermore the surface orientation and depth are recovered simultaneously with the global correspondence.

In the following we first show how the needle maps from stereo images are related. Then a method is presented to combine multiple needle maps to obtain an accurate surface depth. It is found that this technique can be applied to recover surface orientation and depth directly from stereo images by combining the SFS algorithm presented in [7]. As this formulation is highly non-linear we present a simple and elegant algorithm for combining stereo images by extending the results in Section 2. For simplicity all the formulas are derived for the case of two images. They can be easily extended to the case when more than two images are available.

3.1 Camera Set-up and Needle Map Transformation

We use the same camera set-up as in Ikeuchi [13], see Figure 1. The left image plane is perpendicular to the spatial z axis, while the right image plane is inclined with respect to z-axis so that the two optical axes intersect with each other at the origin of the global coordinate system, which is fixed on the surface. Let (u^l,v^l,w^l) and (u^r,v^r,w^r) be the left and right camera coordinate systems and (x,y,z) be the global coordinate system. Assume the object is far away from the cameras so that orthographic projection can be used. Using the parameters in Figure 1, one has the following coordinate transformations:

$$\begin{bmatrix} u^l \\ v^l \\ w^l \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -d^l \end{bmatrix} \tag{10}$$

$$\begin{bmatrix} u^r \\ v^r \\ w^r \end{bmatrix} = \begin{bmatrix} cos\theta & 0 & -sin\theta \\ 0 & 1 & 0 \\ sin\theta & 0 & cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -d^r \end{bmatrix} \tag{11}$$

$$\begin{bmatrix} u^r \\ v^r \\ w^r \end{bmatrix} = \begin{bmatrix} cos\theta & 0 & -sin\theta \\ 0 & 1 & 0 \\ sin\theta & 0 & cos\theta \end{bmatrix} \begin{bmatrix} u^l \\ v^l \\ w^l \end{bmatrix} + \begin{bmatrix} -d^l sin\theta \\ 0 \\ -d^r + d^l cos\theta \end{bmatrix} \tag{12}$$

By defining a=$cos\theta$, b=$sin\theta$, c=$-d^l\, sin\theta$, d=$-d^r + d^l\, cos\theta$, one obtains

$$\begin{bmatrix} u^r \\ v^r \\ w^r \end{bmatrix} = \begin{bmatrix} a & 0 & -b \\ 0 & 1 & 0 \\ b & 0 & a \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -d^r \end{bmatrix} \tag{13}$$

and

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a & 0 & b \\ 0 & 1 & 0 \\ -b & 0 & a \end{bmatrix} \begin{bmatrix} u^r \\ v^r \\ w^r \end{bmatrix} + \begin{bmatrix} bd^r \\ 0 \\ ad^r \end{bmatrix} \tag{14}$$

Suppose the underlying surface can be expressed in the global coordinate system as:

$$Z = Z(x,y)$$

and the gradient map is

$$p = Z_x(x,y)$$
$$q = Z_y(x,y)$$

Then the orientation of the surface at (x,y,Z) is $(-p,-q,1)$.

As the left camera coordinate system is identical to the global coordinate system except the translation along the z-axis, it is easy to see that the gradient in the left camera coordinate system is

$$p^l = p \tag{15}$$

$$q^l = q \tag{16}$$

In the right coordinate system, the orientation at the same point corresponds to different p and q because of the relative rotation of the coordinate systems. The relation between (p^r,q^r)

and (p, q) can be found using the coordinate transformation.

Vector $(-p^r, -q^r, 1)$ in the right coordinate system is transformed into a vector in the global coordinate system by

$$\begin{bmatrix} -p^* \\ -q^* \\ z^* \end{bmatrix} = \begin{bmatrix} a & 0 & b \\ 0 & 1 & 0 \\ -b & 0 & a \end{bmatrix} \begin{bmatrix} -p^r \\ -q^r \\ 1 \end{bmatrix} = \begin{bmatrix} -ap^r + b \\ -q^r \\ bp^r + a \end{bmatrix} \quad (17)$$

Comparing with the standard orientation form $(-p, -q, 1)$, one can find that in the global coordinate system the needle map from the right image can be expressed as

$$\hat{p}^r = \frac{p^*}{z^*} = \frac{ap - b}{a + bp} \quad (18)$$

$$\hat{q}^r = \frac{q^*}{z^*} = \frac{q}{a + bp} \quad (19)$$

Note the surface point $(x, y, Z(x, y))$ is imaged in the right image plane at

$$\begin{cases} u^r = ax - bZ \\ v^r = y \end{cases} \quad (20)$$

As the left camera coordinate system is parallel to the global coordinate system, we prefer to use the global coordinate system instead of the left camera coordinate system in the following subsections.

3.2 Combining Needle Maps

Suppose multiple needle maps for stereo images are available from photometric stereo, how can they be combined into a consistent and accurate depth map? We start from the cost function suggested by Ikeuchi [8, 13]:

$$e = \int\int s + \lambda(d_p + d_q)dxdy \quad (21)$$

where

$$s = (Z_x - p^l)^2 + (Z_y - q^l)^2$$

$$d_p = (\hat{p}^r(ax + bZ + c, y) - p^l(x, y))^2$$

$$d_q = (\hat{q}^r(ax + bZ + c, y) - q^l(x, y))^2$$

Values of $\hat{p}^r$ and $\hat{q}^r$ can be obtainted from the needle map for the right image by using the transformations (18) and (19). Thus

$$\begin{aligned} d_p &= \{p^l(x, y) - \hat{p}^r(ax - bZ, y)\}^2 \\ &= \{p^l(x, y) - \frac{ap^r(ax - bZ, y) - b}{bp^r(ax - bZ, y) + a}\}^2 \end{aligned} \quad (22)$$

$$\begin{aligned} d_q &= (q^l(x, y) - \hat{q}^r(ax - bZ, y))^2 \\ &= (q^l(x, y) - \frac{q^r(ax - bZ, y)}{bp^r(ax - bZ, y) + a})^2 \end{aligned} \quad (23)$$

Ikeuchi [8, 13] used p^r and q^r instead of $\hat{p}^r$ and $\hat{q}^r$ in (22) and (23). This is not correct, because the needle maps of a surface in different coordinate systems are generally different.

Note that d_p and d_q are functions of $Z(x, y)$, s in (21) corresponds to the integrability constraint , d_p and d_q couple the information from the two needle maps so that they match each other . The Euler term corresponding to d_p and d_q is :

$$2\{p^l - \frac{ap^r - b}{bp^r + a}\}\frac{ab(bp^r + a)\frac{\partial p^r}{\partial u^r} - b^2(ap^r - b)\frac{\partial p^r}{\partial u^r}}{(bp^r + a)^2} + 2\{q^l - \frac{q^r}{bp^r + a}\}\frac{b(bp^r + a)\frac{\partial q^r}{\partial u^r} - b^2 q^r \frac{\partial p^r}{\partial u^r}}{(bp^r + a)^2}$$

The Euler equation corresponding to (21) is

$$\nabla^2 Z = (p_x^l + q_y^l) + \lambda[\{p^l - \frac{ap^r - b}{bp^r + a}\}\frac{(a^2b + b^3)\frac{\partial p^r}{\partial u^r}}{(bp^r + a)^2}] + \lambda[\{q^l - \frac{q^r}{bp^r + a}\}\frac{b(bp^r + a)\frac{\partial q^r}{\partial u^r} - b^2 q^r \frac{\partial p^r}{\partial u^r}}{(bp^r + a)^2}]$$

This is an asymmetric formulation based on the left needle map with a correction term from the right needle map. The quantities p^r , q^r , $p_{u^r}^r$ and $q_{u^r}^r$ are functions of $Z(x, y)$ and make up the nonlinear part of the equation. A similar equation can be found to reconstruct the surface depth from the right needle map with a correction term from the left needle map.

If the stereo image intensities are directly available we can couple the needle maps and depth maps at every step. In this case one needle map and one depth map are generated in each coordinate system. But these needle and depth maps are coupled. We follow the ideas presented in [7] , and solve for the right and left depth maps using the algorithm given below:

1. calculate p_{n+1}^l , q_{n+1}^l and p_{n+1}^r , q_{n+1}^r using Lee's [14] algorithm independently for the left and right camera
2. reconstruct the depth in the global coordinate system by solving

$$\nabla^2 Z_{n+1}^l = (p_x + q_y) + \lambda[\{p^l - \frac{ap^r - b}{bp^r + a}\}\frac{(a^2b + b^3)\frac{\partial p^r}{\partial u^r}}{(bp^r + a)^2}] + \lambda[\{q^l - \frac{q^r}{bp^r + a}\}\frac{b(bp^r + a)\frac{\partial q^r}{\partial u^r} - b^2 q^r \frac{\partial p^r}{\partial u^r}}{(bp^r + a)^2}]$$

3. reconstruct the depth in the right camera coordinate system by solving

$$\nabla^2 Z_{n+1}^r = (p_u^r + q_v^r) + \lambda[\{p^r - \frac{ap^l - \hat{b}}{bp^l + a}\}\frac{(a^2\hat{b} + \hat{b}^3)\frac{\partial p^l}{\partial x}}{(bp^l + a)^2}] + \lambda[\{q^r - \frac{q^l}{bp^l + a}\}\frac{\hat{b}(\hat{b}p^l + a)\frac{\partial q^l}{\partial x} - \hat{b}^2 q^l \frac{\partial p^l}{\partial x}}{(bp^l + a)^2}]$$

where

$$\hat{b} = -sin\theta, p^l = p^l(au^r - \hat{b}Z^r + \hat{c}, y) \ , \ \hat{c} = d^r sin\theta$$

4. apply the central difference approximation to determine $\bar{p}_{n+1}^l$ and $\bar{q}_{n+1}^l$ from Z_{n+1}^l, and replace p_{n+1}^l and q_{n+1}^l by $\bar{p}_{n+1}^l$ and $\bar{q}_{n+1}^l$.
5. apply the central difference approximation to determine $\bar{p}_{n+1}^r$ and $\bar{q}_{n+1}^r$ from Z_{n+1}^r, and replace p_{n+1}^r and q_{n+1}^r by $\bar{p}_{n+1}^r$ and $\bar{q}_{n+1}^r$.
6. return to stage 1

The suggested algorithm is symmetric, enforces integrability , and couples the information from the two cameras in an early stage . All the steps in the algorithms which involve solving a Poisson equation are implemented using the direct

methods [7] .

3.3 A New Method for Coupling Stereo Images

In this section we look at the coupling problem differently. In the previous algorithm two needle maps and two depth maps were generated. And the formulation is highly non-linear. One can use the global coordinate system as a reference coordinate system to recover the surface orientation and depth from stereo image intensities.

Suppose two images are taken from the left and right cameras respectively, resulting in two irradiance equations. The irradiance equation for the left image is:

$$E^l(u^l, v^l) = R^l(p^l, q^l) \tag{24}$$

By using (10), (15) and (16) one obtains:

$$E^l(x, y) = R^l(p, q) \tag{25}$$

And the irradiance equation for the right image is:

$$E^r(u^r, v^r) = R^r(p^r, q^r) \tag{26}$$

From (18), (19) and (20) the irradiance equation for the right image can be written as:

$$E^r(ax - bZ, y) = R^r(\frac{ap+b}{a-bp}, \frac{q}{a-bp}) \tag{27}$$

Note that R^l, R^r are the same surface reflectance map, except that the viewing direction is different. In the Lambertian case the reflectance map is independent of the viewing direction so

$$R^r(\frac{ap+b}{a-bp}, \frac{q}{a-bp}) = R^l(p, q) = R(p, q) \tag{28}$$

Thus (27) can be reduced to

$$E^r(ax - bZ, y) = R(p, q) \tag{29}$$

For simplicity, in the rest of this section we only consider the Lambertian case. But the formulation can be readily applied to non-Lambertian cases.

Following the algorithm described in Section 2, one formulates the following cost functional:

$$\int\int [(E^l(x, y) - R(p, q))^2 + \alpha(E^r(ax - bZ, y) - R(p, q))^2 + (Z_x - p)^2 + (Z_y - q)^2 + \lambda(p_x^2 + p_y^2 + q_x^2 + q_y^2)]dxdy$$

Note that another weighting factor α is introduced which can control the use of information contained in the other images. For example, when the correspondence is not accurately established we prefer to keep α small.

Using the variational principle one obtains the following Euler equations:

$$\nabla^2 p = -\frac{1}{\lambda}[(Z_x - p) + (E^l(x, y) - R(p, q))R_p + \alpha(E^r(ax - bZ, y) - R(p, q))R_p] \tag{30}$$

$$\nabla^2 q = -\frac{1}{\lambda}[(Z_y - q) + (E^l(x, y) - R(p, q))R_q + \alpha(E^r(ax - bZ, y) - R(p, q))R_q] \tag{31}$$

$$\nabla^2 Z = \alpha b E^r_{u^r}(ax - bZ, y)(E^r(ax - bZ, y) - R(p, q)) + p_x + q_x \tag{32}$$

These equations are also solved using Jacobi Picard iterations. If there is no correspondence between two images one simply leave out the term $(E^r - R)$.

The correspondence between the two images is established during the generation of the depth map and at the same time it is used to compute the depth map. A good initial estimate is required for the iterative solution. It can be computed by applying the algorithm described in section 2 to the left image to generate the initial depth and needle maps. Thus one can increase α in (30),(31) and (32) from 0 to 1 as the correspondence is increasingly accurate, to control the coupling of stereo images. If the images are too noisy to compute the derivatives of the image intensities, the weighting factor α in (32) can be set to zero. That is, the depth map can be recovered from p and q without directly coupling the stereo images. But α in (30) and (31) need not be zero.

4 Combining Photometric and Geometric Stereo Information

Photometric stereo seems to be the only practical solution for obtaining the orientation map for a surface with varying albedo. Still error occurs in orientation map due to modeling problems of the imaging process. In order to improve the reconstruction of the surface from the needle map, additional information is needed. Sparse measurements obtained from geometrical stereo can serve this purpose [2]. The sparse measurements are available in arbitrary locations on the surface. We use the concept of direct minimization to solve this problem. First the following functional is introduced:

$$\int\int [(Z_x - p)^2 + (Z_y - q)^2 + \lambda(Z - Z^m)^2]dxdy \tag{33}$$

where Z^m represents the known depth values obtained from geometric stereo, and

$$\lambda(x, y) = c\sum_{s \in S} \delta(x - x_s, y - y_s)$$

where S is the set of points where measurements exist.

The function that minimizes (33) is chosen as the solution. The presence of the delta function in the integral creates problems in the formulation of the Euler equation. Hence we choose to use a direct method for minimizing the functional. We discretize the integral using finite differences:

$$Z_x = \frac{Z_{i+1,j} - Z_{i,j}}{\Delta x}$$

$$Z_y = \frac{Z_{i,j+1} - Z_{i,j}}{\Delta y}$$

The delta function becomes a Kroneker delta function and one obtain the following cost function:

$$\sum_{i,j}\left\{\hat{\lambda}_{i,j}(Z_{i,j} - Z^m_{i,j})^2\right\} + \sum_{i,j}\left[\frac{(Z_{i+1,j} - Z_{i,j})}{\Delta x} - \frac{p_{i,j} + p_{i+1,j}}{2}\right]^2 \Delta x\Delta y + \sum_{i,j}\left[\frac{(Z_{i,j+1} - Z_{i,j})}{\Delta y} - \frac{q_{i,j} + q_{i,j+1}}{2}\right]^2 \Delta x\Delta y \tag{34}$$

where

$$\tilde{\lambda}_{i,j} = \begin{cases} c & \text{if } Z^m_{i,j} \text{ exists} \\ 0 & \text{otherwise} \end{cases}$$

Taking the derivative of the cost function (34) with respect to $Z_{i,j}$, $i,j = 1 \cdots n$ one obtains a set of linear equations of the form:

$$\alpha^2(Z_{i-1,j} - 2Z_{i,j} + Z_{i+1,j}) + Z_{i,j-1} - (2 + \alpha\tilde{\lambda}_{i,j})Z_{i,j} + Z_{i,j+1} = \tilde{f}_{ij} \tag{35}$$

where

$$\tilde{f}_{ij} = \left(\frac{p_{i+1,j} - p_{i-1,j}}{2\Delta x} + \frac{q_{i,j+1} - q_{i,j-1}}{2\Delta y}\right)(\Delta y)^2 - \alpha\tilde{\lambda}_{i,j}Z^m_{i,j}$$

where $\alpha^2 = (\Delta y/\Delta x)^2$. We now show that Neuman boundary conditions [7] are available for this problem. The normal derivative on the boundary is a linear combination of the Z_x and Z_y. In the case of rectangular boundary contours (like in images), $\frac{dZ}{d\eta} = Z_x$ along the contours $x = constant$ and $\frac{dZ}{d\eta} = Z_y$ along the contours $y = constant$. Brooks and Horn [6] show that using the variational principle on the boundary, we simply have to replace Z_x by its estimate p and Z_y by q. Thus, $\tilde{f}$ is modified for the Neuman boundary conditions. The solution to (35) may be obtained using embedding techniques discussed in [7].

5 Summary

Several new algorithms have been presented to recover 3-D surface depth and orientation from image shading information. The SFS algorithm proposed in the paper works in the single image domain. It has several advantages over earlier SFS algorithms. This algorithm can be easily combined with conventional stereo method to utilize stereo images. Starting with the idea of combining multiple needle maps we derived another method to combine stereo vision and SFS techniques. These algorithms have all the advantages of stereo and SFS methods to recover a surface from image intensities. The accuracy of these algorithms can be improved by using sparse depth measurements.

References

[1] D. Marr and T. Poggio, "Understanding Image Intensity", *Proc. Roy. Soc. London*, vol B 204, 1979; see also,"A theory of human stereo vision," AI Lab.,MIT Memo451,1977.

[2] W. E. L. Grimson, *"From Images to Surfaces"*, MIT Press, Cambridge, 1981.

[3] B.K.P. Horn, "Shape-from-Shading:A Method for Obtaing the Shape of a Smooth Opaque Object from a Single View", *Artificial Intelligence Lab,M.I.T*, MAC-TR-79 and AI-TR-232, November 1970.

[4] B.K.P. Horn, "Understanding Image Intensity", *Artificial Intelligence*, vol 8,No. 2, 1977.

[5] K. Ikeuchi and B. Horn, "Numerical Shape from Shading and Occluding Boundaries", *Artificial Intelligence*, vol.17, pp. 141–185, 1981.

[6] B. K. P. Horn and M. J. Brooks, "The Variational Approach to Shape from Shading", *Computer Vision , Graphics , and Image Processing*, vol.33, pp. 174–208, 1986.

[7] T. Simchony and R. Chellappa, "Direct Analytical Methods for Solving Poisson Equations in Computer Vision Problems", In *IEEE Computer Society Workshop on Computer Vision*, Miami Beach,Florida, Nov. 1987.

[8] K. Ikeuchi, "Constructing A Depth Map from Images", *Artificial Intelligence Lab , M.I.T*, A.I.M 744, August 1983.

[9] R. T. Frankot and R. Chellappa, "A Method for Enforcing Integrability in Shape from Shading Algorithms", *IEEE Transactions on Pattern Analysis and Machine Inteligence*, Accepted for Publication, (Also presented at the First Int. Conf. on Computer Vision , London June 1987).

[10] T. Poggio and V. Torre, "Ill posed Problems and Regularization Analysis in Early Vision", *Artificial Intelligence Lab , M.I.T*, A.I.M 773, 1984.

[11] R. Courant and D. Hilbert, *"Methods of Mathematical Physics"*, Volume vol. 1, Interscience, New York, 1953.

[12] A. Blake, "On the Geometric Information Obtainable from Simultaneous Observation of Stereo Contour and Shading", In *Report CSR 205-86*, Department of Computer Science , University of Edinbugh, 1986.

[13] K. Ikeuchi, "Determining a Depth Map Using a Dual Photometric Stereo", *The International Journal of Robotics Research*, vol. 6 , No. 1, pp. 15–31, 1987.

[14] D. Lee, A Provably Convergent Algorithm for Shape from Shading, In *Proc. DARPA Image Undestanding Workshop*, pages 489–496, Miami Beach,Florida, December 1985.

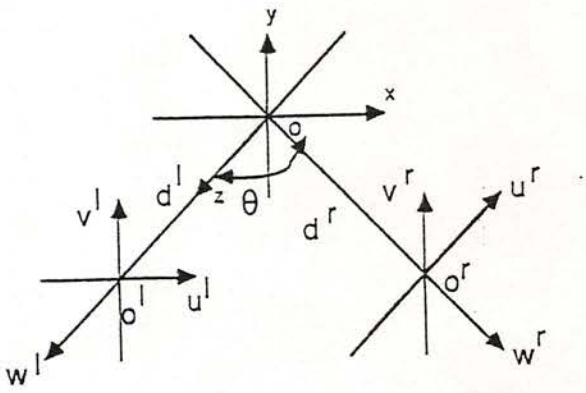

Figure 1. Camera Set-up for Stereo Images