MULTIRESOLUTION SHAPE FROM SHADING*

Gad Ron
Shmuel Peleg

Dept. of Computer Science
The Hebrew University of Jerusalem
91904 Jerusalem, Israel

Abstract

Multiresolution approaches are often used in computer vision to speed up computationally extensive tasks. A solution is computed for a smaller image, and that solution helps to guide towards the complete solution on a larger image. Trying to adopt this multiresolution approach for the computationally extensive shape from shading problem proved to be harder than expected.

Reduction of the grey level resolution of an image, as often done in image "pyramids", does not correspond to images obtained by reducing the *shape resolution:* the 3-D resolution of an object.

The correspondence between grey level and shape resolutions is discussed, and a method is proposed to use multiresolution the case of shape from shading.

1 Introduction

The problem of shape from shading is to reconstruct a 3-D surface from a given intensity image, when the surface reflectance properties are known. This paper treats surfaces with Lambertian reflectance properties.

In this section we will briefly review the Lambertian reflectance model and a proposed solution to the shape from shading problem, and describe the multiresolution pyramid.

1.1 The Lambertian Reflectance Model

Given a surface S and a point light source at direction $\vec{L}$, the surface is called *Lambertian* if the amount of light R reflected from the surface S is a function of the angle between the light direction $\vec{L}$ and the surface normal $\vec{N_s}$. In this paper we will always assume that the reflectance coefficient of the surface is 1, and will omit this coefficient from all equations. The reflectance function of the surface S at every point can thus be written as follows:

$$R(\vec{N_s}, \vec{L}) = \cos(\alpha_{\vec{N_s},\vec{L}}) = \frac{< \vec{N_s} \cdot \vec{L} >}{||\vec{N_s}|| \cdot ||\vec{L}||} \qquad (1)$$

By denoting $p = \frac{\partial S}{\partial x}$ and $q = \frac{\partial S}{\partial y}$, the surface normal can be written as $\vec{N_s} = (-1, p, q)$. Given a light source in the direction $\vec{L} = (-1, p_L, q_L)$, the Lambertian reflectance function at a point with surface derivatives (p, q) can be written as

$$R_{\vec{L}}(\vec{N_s}) = R_{\vec{L}}(p, q) = \frac{1 + p \cdot p_L + q \cdot q_L}{\sqrt{1 + p^2 + q^2} \cdot \sqrt{1 + p_L^2 + q_L^2}} \qquad (2)$$

For a light source in the direction $\vec{L} = (-1, 0, 0)$, i.e. illumination from above, the reflectance function is

$$R(\vec{N_s}) = R(p, q) = \frac{1}{\sqrt{1 + p^2 + q^2}} \qquad (3)$$

1.2 A Shape From Shading Algorithm

In shape from shading the inverse problem of the reflectance model is solved. Given a grey level image I of an object with a Lambertian surface, the surface S is to be reconstructed. This problem is underconstrained, and at any point the observed reflectance R can correspond to many (p, q) pairs. Brooks and Horn [1] added a smoothness constraint on the surface derivatives. They developed an iterative algorithm to compute the surface derivatives (p, q) that minimize the error between the observed intensity and the computed reflectance, giving weight to the smoothness term. The error term at every image point is

$$E(p, q) = (I - R(p, q))^2 + \lambda \cdot V(p, q) \qquad (4)$$

*This paper has been supported by a grant from the Israeli NCRD

CH2752-4/89/0000/0350$01.00 © 1989 IEEE

where I is the given intensity at the point, R is the intensity computed from the suggested (p,q) using the reflectance model, $V(p,q)$ is the smoothness term for (p,q) at the point, and λ is the strength given to the smoothness term. The smoothness term was computed using a discrete approximation of the surface Laplacian:

$$V(p,q) = (p-\hat{p})^2 + (q-\hat{q})^2 \qquad (5)$$

where $\hat{p}$ and $\hat{q}$ are the local averages of p and q. The average we used in this paper to compute $\hat{p}$ and $\hat{q}$ is a convolution with

$$\frac{1}{20}\begin{pmatrix} 1 & 4 & 1 \\ 4 & 0 & 4 \\ 1 & 4 & 1 \end{pmatrix}$$

The iterative step to compute p_{i+1} from p_i [1] is as follows:

$$p_{i+1} = \hat{p}_i + \frac{1}{\lambda}\cdot(I - R(p_i,q_i))\cdot\frac{\partial R}{\partial p}(p_i,q_i) \qquad (6)$$

where $\hat{p}_i$ is the local average of p at iteration i, I is the observed grey level, and $R(p_i,q_i)$ is the intensity predicted by the reflectance model. A similar equation holds for q_{i+1}. This algorithm needs boundary conditions on (p,q) in order to converge. This iterative algorithm can give a non-integrable solution: integrating (p,q) on a closed path does not give zero as should be for derivatives of a true surface. Ikeuchi and Horn [4] proposed to use Gaussian sphere coordinates (f,g), instead of the (p,q) coordinates. This coordinate system enables to enforce boundary conditions for any slope, unlike the (p,q) system that has a singularity point when p or q are close to vertical. This approach was not used in this paper, as the (p,q) space gives much simpler equations for our case.

Frankot and Chellappa [3] proposed a closed form solution for enforcing integrability on given (p,q) estimates, by projecting them on a space of integrable functions. This algorithm uses FFT for the projection. The two algorithms are computationally expensive. On a picture with N^2 pixels the iterative algorithm needs $O(N)$ iterations, and each iteration takes $O(N^2)$ operations. When an iteration uses the integrability constraint, an additional $O(N^2 log(N))$ operations are needed for each iteration.

This computational complexity suggests the introduction of a pyramidal scheme for speedup. This turned out to be a non trivial question to solve, as is shown in Section 2.

1.3 The Multiresolution Pyramid

The image pyramid has been studied extensively in the past decade [6,2,5]. A brief review of basic pyramid techniques will be given.

An image pyramid is a sequence of copies of an original image in which resolution and sample density are decreased, usually in regular steps. Let $\mathbf{G}_0$ be the original image, which form the 'basis' of the pyramid. $\mathbf{G}_0$ is convolved with a low pass filter $\mathbf{w}$, then subsampled by discarding every other row and column to form $\mathbf{G}_1$, the 'first' level of the pyramid. $\mathbf{G}_1$ is then filtered and subsampled to form $\mathbf{G}_2$, and so on. In general, for $\ell > 0$, this can be written concisely in terms of a convolution operator:

$$\mathbf{G}_\ell = [\mathbf{w} * \mathbf{G}_{\ell-1}]_{\downarrow 2}. \qquad (7)$$

Here the notation $[.]_{\downarrow 2}$ indicates the image contained within brackets is subsampling by a factor of 2 in each spatial dimension. In typical pyramids $\mathbf{w}$ is similar to a Gaussian function, and the pyramid is called a *Gaussian Pyramid.*

The filter $\mathbf{w}$ is called the *generating kernel.* In practice this is chosen to be small and separable, so that the computation cost of the filter convolution is kept to a minimum. Pyramid construction is a fast algorithm that generates a full set of filtered images at a cost typically less than 10 operations per pixel of the original image.

Many iterative algorithms take advantage of the pyramidal structure. The first iteration is computed on a reduced image $\mathbf{G}_\ell$ with little computational cost. The result is than expanded in size, and an additional iteration is performed on $\mathbf{G}_{\ell-1}$. This process continues until the final iteration is performed on the original image $\mathbf{G}_0$. The advantage of this process is that most iterations are performed on images of reduced size, significantly decreasing the computational cost.

2 Surface Resolution and Intensity Resolution

For the intensity pyramid to be of any help in the computation of shape from shading, the following assumption should be true: Let I_0 be the grey level image of the surface S_0. Reducing the resolution of I_0 into I_1 should yield the same image as the image of surface S_1, which is obtained by reducing the resolution of surface S_0. In earlier work, Terzopoulos [7] used a simple greylevel pyramid for multiresolution shape from shading, with reasonable results on very simple shapes. Unfortunately, the operators of reduction of resolution and imaging (using the reflectance model) do not commute, and in general this simple approach can fail badly. Following is a one dimensional example. In this example we will view the reduction of resolution as filtering, and will omit the subsampling stage.

EXAMPLE:

Let the 1-D surface be $S_0 = \cos(2\pi x)$. Blurring S_0 into S_1 using a Gaussian filter $\frac{1}{2\pi\sigma} \cdot e^{-x^2 2\sigma}$ gives the smoother surface $S_1 = \frac{\cos(2\pi x)}{e^{2\sigma}}$. In the 1-D case Equation (3) becomes

$$R(p) = \frac{1}{\sqrt{1+p^2}} \tag{8}$$

where $p = \frac{\partial S}{\partial x}$.

As S_1 has smaller derivatives than S_0 at all points, its image R_1 should be brighter than the image of S_0 everywhere (Equation (8)). And indeed S_1 is closer to a smooth plane which gives highest intensities. On the other hand a smooth (blurred) version of I_0, I_1, can never have that $I_0 < I_1$ for every point. I_1 will just be closer to the average of I_0 at every point, and so $R_1 \neq I_1$.

As using the intensity pyramid does not solve the problem of speeding up the shape from shading algorithm, other schemes were tested.

3 Blurring Surface Reflectance

In this section we will describe an estimate to the reflectance R_1 of a reduced resolution surface S_1 from the grey level image $I_0 = R_0$.

We will assume that for our surfaces the blur operator and the derivative operator commute (both are convolutions). Therefore, if the surface derivatives (p, q) could be estimated from I_0 they could be blurred, and R_1 be generated from these smoothed derivatives. But as our main problem is to compute these surface derivatives, we will rather compute $T^2 = \tan(\alpha_{\vec{N}_s})^2$, where $\alpha_{\vec{N}_s}$ is the angle between the surface normal and the (vertical) lighting direction. From Equation (3) we can get for every image point:

$$T^2 = \tan(\alpha_{\vec{N}_s})^2 = \frac{1-R^2}{R^2} = p^2 + q^2 \tag{9}$$

and

$$T = |\tan(\alpha_{\vec{N}_s})| = \sqrt{\frac{1-R^2}{R^2}} = \sqrt{p^2+q^2}. \tag{10}$$

T is computed from the image intensities R. Given T, R can be computed by using the following equation:

$$R = \sqrt{\frac{1}{1+T^2}} \tag{11}$$

As $T = ||(p,q)||$ is the norm of (p,q), then $\alpha T = ||(\alpha p, \alpha q)||$. i.e., multiplying T by a constant corresponds to multiplying the surface derivatives by a constant. We use this to estimate the function T of the reduced resolution surface given the original surface. As resolution reduction involves blur followed by subsampling, the critical stage is the blur. We would like to estimate the function T of the blurred surface given the function T of the original surface.

When at two adjacent points the derivatives (p_1, q_1) and (p_2, q_2) are similar, then $||(p_1,q_1) + (p_2,q_2)||$, the norm of the smoothed surface, can be approximated by $||(p_1,q_1)|| + ||(p_2,q_2)||$: smoothing the two surface norms. This holds, of course, for any local convolution operator. So blurring the surface derivatives (p,q) by convolving with a Gaussian, and then computing T of the blurred surface, can be approximated by blurring the original T. This approximation is more accurate when the derivatives (p,q) do not change quickly, i.e. the surface is smooth. If on the other hand the derivatives (p,q) do change quickly, then the estimate will no longer be accurate. Blurring T, the norm of (p,q), will give higher values than the norm of $(blurred(p), blurred(q))$. But even in this case the image obtained after blurring T will be closer to the correct image than the grey level pyramid.

4 Pyramidal Scheme for Shape from Shading

Based on the value T as computed in equation (10), the suggested algorithm for building the grey level pyramid for shape from shading purposes is as follows:

1. Calculate T_0 from the given input image I_0 using Equation (10).

2. Build a Gaussian pyramid $T_0 \cdots T_{n-1}$ whose basis is T_0, as described in Section 1.3. The pyramid has n levels, chosen such that T_{n-1} will still have significant shape information. We have experimented with pyramids whose smallest levels were of size 32x32.

3. Using Equation (11) calculate from every T_i the grey level image R_i. The images R_i constitute a grey level pyramid, which is the estimate of the reduced resolution surface reflectance.

4. Using the R_i pyramid, preform the multiresolution shape from shading algorithm. First compute (p_{n-1}, q_{n-1}) from R_{n-1} using any shape from shading algorithm, like that described in Section 1.2. Than repeatedly expand the obtained low-resolution derivatives (p_i, q_i) into (p_{i-1}, q_{i-1}), to be used as an initial guess for the computation of the (p, q) derivatives at the level $i-1$ of the pyramid. The process is repeated until the (p, q) derivatives are computed for the full resolution image.

Using this pyramidal scheme will result in most iterations being performed on low resolution images, saving significant computations.

5 Experimental Results

The proposed multi resolution scheme has been tested on two simulated surfaces of size 128x128, displayed in Figure 1. From the surfaces I_0 was calculated numerically using equation (3). A 3-level reflectance pyramid was than built using the scheme of the previous section: $R_0 = I_0$ of size 128x128, R_1 of size 64x64, and R_2 of size 32x32. The blurring of the T-pyramid was done using the following mask

$$\begin{pmatrix} 0 & 0.125 & 0 \\ 0.125 & 0.5 & 0.125 \\ 0 & 0.125 & 0 \end{pmatrix}$$

The surfaces, as well as the intensity images of the reduced surfaces, are displayed in Figure 1. It is important to note that the overall brightness of the reduced resolution images is significantly higher that the brightness of the higher resolution images. This occurs as the reduced resolution surfaces are more flat, their brighness should thus increase, and this effect is captured in our scheme for resolution reduction.

Boundary conditions for (p, q) where calculated from S_0 on the surface boundaries, and reduced in resolution for the other levels. The coefficient λ of Equation (6) was taken to be 7000 in all the runs.

Figure 2 shows the surfaces reached by shape from shading algorithms without using pyramids. It shows both the regular shape from shading without using the integrability constraint (*sfs*), and using the integrability constraint (*sfsi*). In the case without using the intergability constraint during the iterations, it was applied once after convergence to enable the creation of a displayable surface.

Figures 3 and 4 display the surfaces obtained using the pyramidal algorithm with and without the integrability constraint. Let *msfs* stand for 'multi resolution shape from shading'. Figure 3 displays the surfaces in the multiresolution pyramid in all three levels.

Let *msfsi* stand for 'multi resolution shape from shading with the integrability constraint'. In this approach the integrability constraint was enforced after every iteration. Results are displayed in Figure 4.

All algorithms gave results of similar quality, differing mainly in the computational cost. Table 1 summarizes the number of iterations needed to reach convergence, and the error between the calculated and the known (p, q). The computational speedup of the multiresolution approach is evident, and this speedup improves with the complexity of the images.

results for the simpler image:

algorithm	size	no. of iterations	$E^2(p)$	$E^2(q)$
sfs	128x128	70	0.08	0.08
msfs	32x32	13		
	64x64	4		
	128x128	6	0.136	0.136
sfsi	128x128	13	1.006	1.006
msfsi	32x32	11		
	64x64	6		
	128x128	5	0.495	0.495

results for the complicated image:

algorithm	size	no. of iterations	$E^2(p)$	$E^2(q)$
sfs	128x128	81	6.3	1.7
msfs	32x32	18		
	64x64	6		
	128x128	15	8.9	3.11
sfsi	128x128	54	3.5	10.3
msfsi	32x32	8		
	64x64	7		
	128x128	8	3.11	5.17

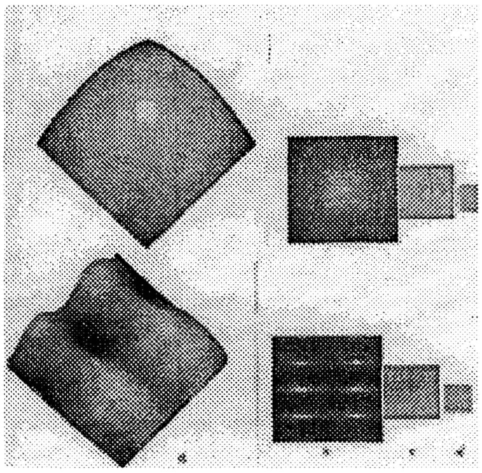

Figure 1: The reflectance pyramid.
a) A perspective view of the original surfaces, size 128x128.
b) The intensity images of the original surfaces.
c,d) Reduced resolution intensity images to sized 64x64 and 32x32. The resolution is reduced according to the algorithm described in Section 4.

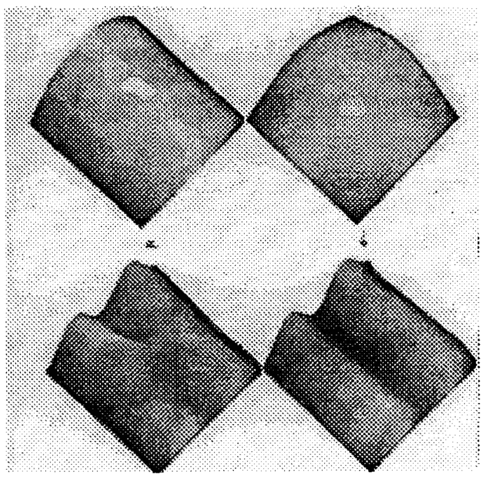

Figure 2: Perspective view of the surfaces obtained using non-pyramid algorithms.

a) Without using the integrability constraint (sfs).
b) Using the integrability constraint at each iteration (sfsi).

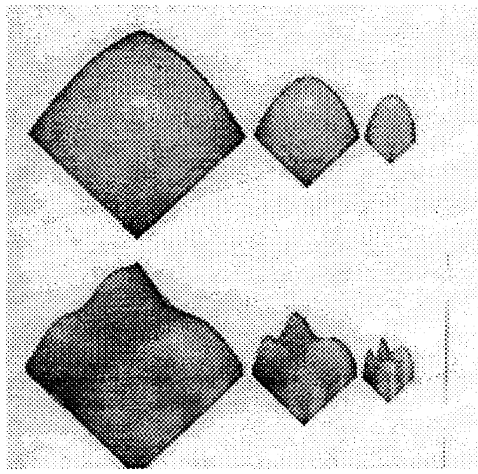

Figure 3: Perspective view of the surfaces obtained in the multiresolution approach of Section 4, without using the integrability constraint. The lowest resolution was computed first, and then the higher resolution surfaces.

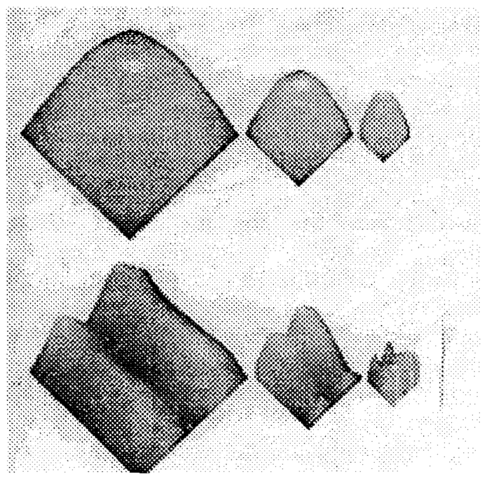

Figure 4: Same as Figure 3, but with applying the integrability constraint at every iteration.

6 Concluding Remarks

In this paper we have presented a method to reduce image resolution that simulates the reduction of shape resolution. Such resolution reduction is critical for shape from shading algorithms. We have used this method of resolution reduction in accelerating the shape from shading by using multiresolution pyramids.

We have used a Labmertian reflectance model in this paper, but a similar approach can be developed for other reflectance models. Our goal is to show that the resolution reduction based only on image intensities is inferior to resolution reduction using knowledge on the surface reflectance.

Since optical reduction of images, as done when moving a camera away from the object, is simple gerylevel operation, a question is cast on the general scheme of shape from shading. As we have shown that in general reduction of greylevel resolution does not correspond to scale reduction of the object, it follows that for accurate shape from shading we need to know the distance from the object, and actually the shape should also be know in advance. This problem does not exist for fractal object, as their properties do not change with resolution, but there is currently little interest in shape from shading for fractal surfaces.

References

[1] M.J. Brooks and B.K.P. Horn. Shape and source from shading. In *International Joint Conference on Artificial Intelligence*, pages 932–936, Los Angeles, California, August 1985.

[2] P.J. Burt. Fast filter transforms for image processing. *Computer Vision, Graphics, and Image Processing*, 16:20–51, 1981.

[3] R.T. Frankot and R. Chelappa. A method for enforcing integrability in shape from shading algorithms. *IEEE Trans. on Pattren Analysis and Machine Intelligence*, 10:439–451, July 1988.

[4] K. Ikeuchi and B.K.P. Horn. Numerical shape from shading and occluding boundaries. *Artificial Intelligence*, 17:141–184, 1981.

[5] A. Rosenfeld, editor. *Multiresolution Image Processing and Analysis*. Springer-Verlag, 1984.

[6] S.L. Tanimoto and T. Pavlidis. A hierarchical data structure for picture processing. *Computer Graphics and Image Processing*, 4:104–109, 1975.

[7] D. Terzopoulos. Efficient multiresolution for computing lightness, shape from shading, and optical flow. In *Proceedings of the fourth national conference on artificial intelligence (AAAI-84)*, pages 314–317, 1984.