Project Proposal: A Multiresolution Stochastic Approach for Shape From Shading Using Intensity Gradient Constraint

Emre Uğur
Emre Akbaş

December 6, 2003

Abstract

The objective of this project is the study of *Shape from Shading* (SfS) problem in a variational framework. Among the other shape recovery methods, SfS is employed to obtain three-dimensional shape using spatial variations of brightness in an image. In this study, SfS problem will be addressed as an energy minimization problem. A constraint which establishes a connection between intensity gradient space and solution surface is imposed into the corresponding energy functional for the first time in a *direct energy minimization* scheme. A hybrid approach which combines the efficiency of deterministic methods and accuracy of stochastic methods is proposed in a multiresolution domain for the solution of this problem.

1 Introduction and Motivation

Human visual perception is based on two-dimensional images, which are used to recover the third-dimension. Different mechanisms, e.g. intensity and color information in the image and stereoscopic disparity created by the usage of two eyes are employed to accomplish this task. One of the most primitive mechanisms which is used by our visual system to extract shape information is probably *the processing of shading* information. [5] suggests an explanation for this argument by describing a principle of evolution in some animal species. *Countershading principle*[1] is employed by preys in order to neutralize the effect of sun light and conceal their shapes from predators. Experiments showed that human visual perception system has the ability to extract both light source direction and shape of the objects only from the gray-level image.

Shape recovery methods, which deal with obtaining three-dimensional shape description from one or more two-dimensional images, have long been investigated in computer vision. Being one of these methods, Shape From Shading (SfS) is concerned with recovering shape description from shading in a single two-dimensional image. Shading can be defined as the spatial variations of brightness in an image.

It is important that SfS research follow Marr's approach which regards vision as *a complex information processing system*. The fact that human visual perception can extract shape information from shading constitutes the *computational theory* of this information processing task, and deals about *what is computed?* and *why?*. From an abstract point of view, vision is a problem concerning with mapping one kind of information (2D image) to another (3D scene description). The second level, or the so-called *representational and algorithmic level* is concerned with the specific approach and the representations used. Finally, *the hardware implementation* level deals with the physical realization of the algorithm and the representation(s).

Accounting the fact that solely shape from shading, in almost all the cases, is not adequate to extract shape information and human perception of SfS interacts with other

[1] Countershading is the protective coloration of the animal, where body parts which are exposed to more light are darker colored. This principle is valid in a wide range of animals from fishes to horses. For some nice pictures, `http://members.aol.com/battyatty/count.htm`.

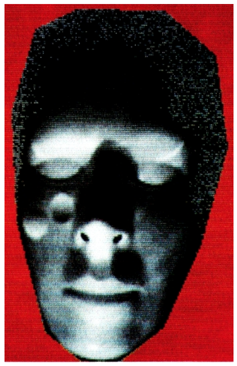


Figure 1: Human visual system has the ability to extract shape information only from the variation in brightness. Furthermore, it is possible to approximately estimate the direction of light source

visual processes such as perceptual grouping and motion perception[5], the independent study of SfS supports the modular view of the information processing system. This supports Marr's view of subdividing vision problem into smaller problems with which it is easy to tackle with.

In Section 2, different methods in literature are described briefly for the extraction of shape using variations in the brightness. In this study, a solution which is based on minimization of a functional, is proposed for shape from shading problem. Since the functional is nonlinear and there exists infinite number of solutions, some constraints are inserted into the variational formulation of the problem. After a deterministic energy minimization scheme is given, since solution tends to be stuck in local minima, a stochastic approach is proposed. The slowness of the stochastic methods leads us to consider hybrid solutions where images are processed in different resolutions. Details of the shape from shading problem and proposed method of our project can be found in Section 3.

2 Literature Work

SfS problem was first formulated by B.K.P. Horn in his doctoral thesis in 1970. Part of his thesis appeared in the *Psychology of Computer Vision* in 1975, where SfS was explained as a process for recovering the shape of an object from a single two-dimensional image using photometric cues like the amount of light reflected from the object. In his original work, image brightness was modeled as a function of surface geometry. His statement of the problem was highly underconstrainted, as no assumptions were made about the projection model and the location of either the light source, or the viewer. Expectedly, this resulted in a complex mathematical analysis which was prone to numerical instability.

In the introduction part of his book([10]), Horn mentions that the earliest work on the quantitive use of shading appears to have been used in the mid-1960s by Russian scientists for recovering the shape of parts of the lunar surface. They discovered that the brightness of the surface of the moon was a function of the ratio of the cosine of the incident angle to the cosine of the emittance angle.

Since Horn's formulation of the problem, many different techniques have been proposed for the solution of the SfS problem. As suggested in a survey ([11]) on SfS techniques, they can be grouped as the following:

- Global approaches
  - Global minimization approaches
  - Global propagation approaches
- Local approaches

Global minimization approaches obtain the solution by minimizing an energy function, whereas global propagation approaches obtain the solution by propagating the shape information from known surface points such as singular points. Local approaches derive shape information only from the intensities of the surface points in a small neighborhood.

There are two different approaches to minimize the functional created by relaxation techniques. In his original work [8], Horn produced Euler equations for each unknown in the functional. Instead, Szeliski [3] directly minimized the discrete formulation of the functional and obtained good results. The drawbacks of the first method and details of the second will be given in Section 3. Although Szeliski obtained a good convergence rate using conjugent gradient descent method, since the solution contains multiple deep local minima, in [1] an alternative stochastic approach is employed using simulated annealing. Extreme slowness of this method leads to the formulation of a hybrid approach which combines efficiency and accuracy in a multiresolutional framework.

Integrability and smoothness constraints, which will be described in detail in the next section, are the most used constraints in literature. In [6], a constraint which imposes the equality of image intensity gradient and solution surface gradients is successfully applied. They used integrability constraint additionally to enhance the validity of the surface.

3 Shape from Shading Problem and Proposed Method

As mentioned in Section 1, shape from shading is a classical problem of computer vision where various approaches include many unknowns and non-linearity for the solution. This leads to the necessity of making a number of assumptions and dealing with unrealistic/human-made situations.

Smoothly curved monochrome objects with homogeneous reflecting characteristics exhibit a variance in image radiance when shaded with one or more light sources. This variance depends on the interaction of a number of factors, namely

1. illumination of the object,
2. shape of the surface(s),
3. reflecting properties of the material,
4. and projection of the image.

SfS solutions depend on the strong argument which states that *there is a unique image intensity value for a given surface orientation* as formulated by Horn [2]. This argument requires some assumptions for interacting factors which determines the irradiance of the image. First of all, object should be homogeneous that is reflectance properties should not vary along surface. Although there are studies [7] for perspective projection case, which is a more realistic one, image is assumed to be formed in orthographic projection for simplicity. The last crucial and widely employed assumption is the position of viewer and light source, they are both thought to locate very distant with respect to the object, so light rays approach in parallel and the angle between incident ray and viewer will be constant for all points on the surface[2].

Considering the discussion above, we can conclude that surface orientation and light source direction information will be sufficient to determine the image radiance. *Reflectance map (function)* is the means of specifying the radiance of a surface patch as a function of orientation [8]. Reflectance function therefore decodes information of light intensity and distribution, and reflectance characteristics of the surface. In the context of this study, light is restricted to be emitted from a far away point source with a known direction. Thus, reflectance function serves as a mapping from surface orientation to image brightness. Formally, for a given image E, the *image irradiance equation* is,

$$E(x, y) = R(z_x(x, y), z_y(x, y))$$

where R is the reflectance function and problem is regarded as that of recovering of a smooth surface, z, and equation should be satisfied over the reconstruction domain Ω. z_x and z_y denotes the first partial derivatives of z with respect to x and y respectively. In [9], details of reflectance map calculation for various surfaces in different illumination conditions is given. From now on, short hand notation $p = \partial z/\partial x$ and $q = \partial z/\partial y$ will be used for first partial derivatives.

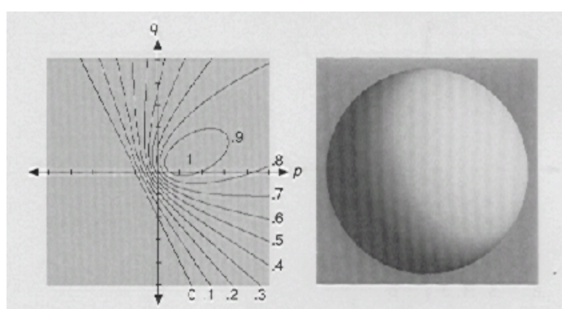


Figure 2: The reflectance map is a plot of brightness as a function of surface orientation. This plot corresponds to a Lambertian surface under point-source illumination.

[2]Due to the shape of the object(s), a small amount of the incident light may be emitted from other objects. The situations where mutual-illumination should take careful attention.

In Figure 2 a plot of the reflectance map which corresponds to a ball with *Lambertian surface* [3] properties is demonstrated. $p - q$ space is called *gradient space* and reflectance map may be defined on this space. The contours in this case turn out to be nested conic sections and each contour shows constant brightness values for (p, q) pairs.

Note that gradient space serves only as a *representation* to be processed in the shape exploration process. Marr's [12] view of vision as *a process of mapping from one representation to another* is applied in this problem. Representations of inputs or outputs or any data to be created or consumed during the *information processing task* are determined according to their performance during this process. There are different factors which affect performance such as suitability, efficiency, simplicity, time and space complexity. Gradient space is represented by an infinite plane onto which upper hemisphere of the Gaussian sphere[4] is projected from its center (Figure 3(a)). Instead of projecting only upper hemisphere, whole sphere can be projected onto the plane (*stereographic plane*) from south pole of the Gaussian sphere (Figure 3(b). For points on occluding boundary, p and q values are infinite in gradient space, so it is very difficult to deal with occluding boundaries in this representation. However, occluding boundaries as well as light sources at the back of the surfaces can be dealt with when stereographic projection is used. Orthographic projection is taken is Smith's work instead of taking perspective projection [8].

There are various shape from shading techniques to extract surface orientation, or more specifically (p, q) points over the domain Ω. As mentioned in Section 1, we will use an energy minimizing approach. For this purpose, an appropriate functional should be employed for each p and q such that constructed surface from p and q should produce the image when illuminated in same way as the actual one. Thus, the integral of the error between actual image intensity and constructed one is tried to be minimized,

$$F_1(p, q) = \iint_{(x,y)\in\Omega} (E(x, y) - R(p, q))^2 dxdy$$

The equation above is not sufficient to derive the unknowns (there would be infinite number of solutions) since for each (p, q), only one intensity value is known. Figure 2 illustrates that an intensity value corresponds to a set of (p, q) pairs (which lie on the corresponding contour). Therefore, additional constraints should be imposed in order to reach a unique solution. Different constraints are proposed in literature as mentioned in Section 2. Two of the most employed constraints are *gradient smoothness* and *integrability* constraints which provide stabilization of the iterative shape from shading and ensuring validity of the gradient field surface, respectively [3]. However we decided to use *intensity gradient constraint*[6], which forces constructed surface to be as smooth as actual surface and establish an additional connection between image intensity and gradient values. During experiments, if error reduction procedure is unsuccessful, we plan to insert additional constraints. Therefore, new functional will take the form,

$$\begin{aligned} F(p, q) = \iint_{(x,y)\in\Omega} & (E(x, y) - R(p, q))^2 \\ & + (R_p(p, q)p_x + R_q(p, q)q_x - E_x(x, y))^2 \\ & + (R_p(p, q)p_y + R_q(p, q)q_y - E_y(x, y))^2 dxdy \end{aligned} \quad (1)$$

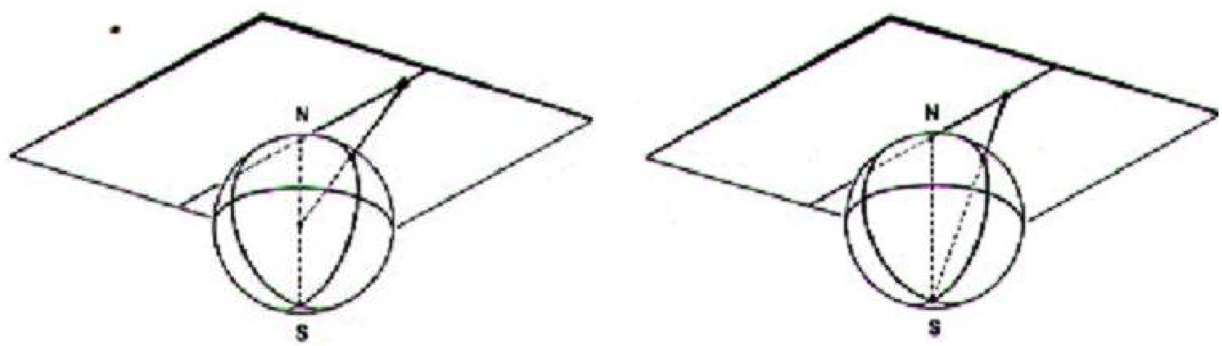
There are two main alternatives to minimize this functional, we can directly minimize the discrete form of the functional, or produce Euler equation for each unknown in discrete form [1]. Although Calculus of Variations is employed to solve this problem in general, in this study *direct energy minimization* will be used for the reasons that [1] listed:

1. All local minima and maxima conditions are solutions of Euler equations.
2. When equations are non-linear, and to be solved iteratively, there exists no proof of convergence.
3. When *direct energy minimization* is employed, no a priori knowledge on the boundary of Ω is necessary. If Euler equations are used, to make the problem well-posed, boundary conditions are necessary.

In the first part of this study, in order to minimize the energy function defined by Equation 1, we will apply the deterministic method which is described in [1]. In each iteration step k, (p^k, q^k) pairs on the domain Ω will be

[3] A Lambertian surface is an ideal surface, which reflects all of the incoming light in all directions.

[4] points on the Gaussian sphere specify directions in the space



(a) gnomic projection (b) stereographic projection

Figure 3: Different projections of the points on Gaussian sphere onto planes. (a) represents gradient space where only upper hemisphere of the sphere is projected. In order to deal with occluding boundary of an object stereographic projection can be used (b), where whole sphere is projected from South Pole.

modified according to the gradient of ϵ^k. In this project, a deterministic optimization method, classical optimal gradient descent will be used.

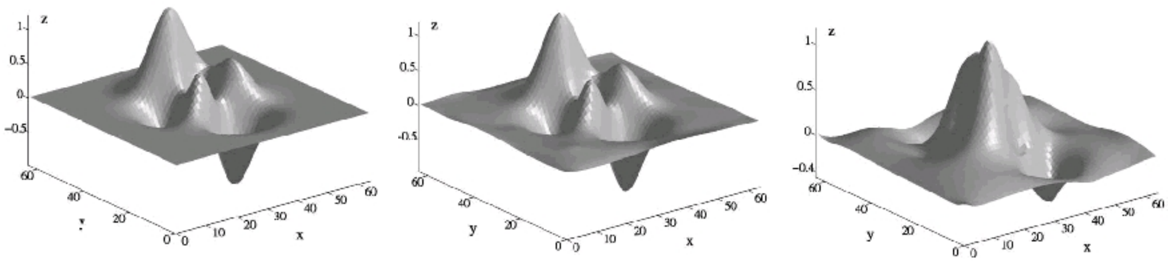
The main drawback of the relaxation methods for SfS problem is the convergence problem. Either by directly minimizing the energy function or by producing Euler equations, solution of the problem may be stuck in local extrema. Figure 4 demonstrates this fact, in which different initial configurations lead to different solutions. While solution in Figure 4(b) is very similar to original surface, Figure 4(c) seems to be stuck in local minima.

In order to escape from local minima, we will employ stochastic optimization techniques. In [1], simulated annealing is successfully applied to SfS problem. However, in this stage of the work, we haven't decided yet representation of the energy minimization problem in the stochastic framework, and which method to be used. The alternatives are simulated annealing and genetic algorithms.

As noted in [1], application of stochastic methods in SfS problem leads to a very slow convergence rate during the iteration phase. Thus, we plan to combine the efficiency of the deterministic algorithm with accuracy of the stochastic method. Referring to the *graceful degradation principle* from [12], resolution of the image will be tuned in order to speed up the slower method. Therefore, stochastic methods will be applied for an image in coarser level of the resolution axis in the first stage. After escaping from local minima, and obtaining an approximate surface of the object, this surface will be used as the starting configuration of the deterministic energy minimization method in higher resolutions.

4 Discussion

In this proposal SfS is discussed in detail and various approaches to solve the problem is given in Section 2. The problem is addressed as an energy minimization process whose functional is composed of well-known brigtness constraint and intensity gradient constraint. The reason of the usage of intensity gradient is its ability to force the solution surface as smooth as the actual surface and to establish an additional connection between image intensity and constructed surface radient values. Since deterministic methods cannot avoid local minima, an alternative stochastic approach will be used. Both methods will be combined in a multiresolutional framework in order to couple the efficiency and accuracy properties.



(a) original surface (b) global minimum (c) local minima

Figure 4: Solutions of the deterministic optimization methods may be stuck in local extremes. (b) and (c) are the solutions for the same object given in (a). Although they are computed using same algorithm, they differ since iterations start from different initial configurations.

References

[1] A. Crouzil, X. Descombes and J.-D. Durou, ”A Multiresolution Approach for Shape from Shading Coupling Deterministic and Stochastic Optimization,” *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 12, no. 6, pp. 529-540, November 2003.

[2] B.K.P. Horn, ”Understanding Image Intensities,” *Artificial Intelligence,* vol. 8, no. 2, pp. 201-231, April 1977.

[3] R. Szeliski, ”Fast Shape from Shading,” *Computer Vision, Graphics, and Image Processing: Image Understanding,* vol.53, no. 2, pp. 1250-1267, December 1999.

[4] J. Shah and H.H. Pien, ”Recovery of Surfaces with Discontinuities by Fusing Shading and Range Data Within a Variational Framework,” *IEEE Trans. Image Processing,* vol. 5, no. 8, August 1996.

[5] D.A. Kleffner and V.S Ramachandran, ”On the perception of shape from shading,” *Perception and Psychophysics,* vol. 52, no. 1, pp. 18-36, July 1992.

[6] Q. Zheng and R. Chellappa, ”Estimation of Illuminant Direction, Albedo, and Shape from Shading,” *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 13, no. 7, pp. 680-702, July 1991.

[7] A. Tankus, N. Sochen, and Y. Yeshurun, ”A New Perspective [on] Shape-from-Shading,” *Proc.* 9^{th} *IEEE International Conference on Computer Vision,* vol. II, pp. 862-869, October 2003.

[8] B.K.P. Horn and M.J. Brooks, ”The Variational Approach to Shape from Shading,” *Computer Vision, Graphics and Image Processing,* vol. 33, no.2, pp.174-208, 1986.

[9] B.K.P. Horn and R.W. Sjoberg, ”Calculating the Reflectance Map,” *Applied Optics,* vol. 18, no. 11, pp. 1770-1779, 1979.

[10] *Shape from Shading,* B.K.P. Horn and M.J. Brooks, eds. MIT press, 1989.

[11] R. Zang, P.S. Tsai, J.E.Cryer, M.Shah, ”Analysis of Shape from Shading Techniques” *IEEE CVPR-94,* pp. 377-384, June 1994.

[12] *Vision,* David Marr, Freeman, 1985