SHAPE FROM SHADING USING GRAPH CUTS

Ju Yong Chang, Kyoung Mu Lee**, and Sang Uk Lee**

*School of Electrical Engineering Seoul National University
Seoul, 151-742, Korea
**Dept. School of Electronics and Electrical Engineering,
Hong-Ik University

ABSTRACT

This paper describes a new semi-global method for SFS (Shape-from-Shading) using graph cuts. The new algorithm combines the local method proposed by Lee and Rosenfeld [1] and the global method using energy minimization technique. By employing a new global energy minimization formulation, the convex/concave ambiguity problem of the Lee and Rosenfeld method can be resolved efficiently. A new combinatorial optimization technique, graph cuts method is used for the minimization of the proposed energy functional. Experimental results on a variety of synthetic and real-world images show that the proposed algorithm reconstructs the 3-D shape of objects very efficiently.

1. INTRODUCTION

SFS has been a central problem in the field of computer vision since the early years. This problem is to compute the 3-D shape of an object from the brightness variations in an intensity image of that object. To solve this problem, it is very important to model how the images are formed. Although many reflectance models have been proposed, most of the SFS techniques assume a simple Lambertian reflection model. According to the Lambertian model the brightness information at a pixel in the image depends only on the albedo, surface normal vector, and the light source direction. Using this Lambertian model, the image formation process can be modeled by the following image irradiance equation.

$$I(x,y) = \rho\vec{S}\cdot\vec{N}(x,y)\,. \tag{1}$$

Thus, the SFS problem can be defined by finding the best way to reconstruct the geometric information of the object satisfying this image irradiance equation. The reconstructed geometric information can be either the surface depth or a set of surface normals, often described as a needle-map.

Many different approaches to solve the SFS problem have been proposed. These methods can be divided into two broad categories: global approaches [2], [3], [6], local approaches [1], [4]. Global approaches attempt to recover the entire surface by minimizing some energy functional associated with the surface to be estimated. However, since the problem is ill-posed, additional constraints are needed to solve it. Usually, the smoothness constraint such as a measure of "departure from smoothness" is added to the original energy functional. In general, this energy functional is minimized by a variational method. The global techniques have been shown to be more generally applicable to different types of input images, and robust to noise than local techniques. And also they generate more accurate results. However, they have demerits such as the tendency to oversmoothing.

Local approaches involve the use of local brightness information of an image, and recover surface patches which subsequently quilted together. These local techniques tend to be fast, and have better capacity for recovering the local features of a surface. But they are sensitive to noise, and require some restrictive assumption about the surface.

In this paper we propose a new semi-global shape-from-shading technique based on graph cuts in which the local and global methods are combined together. This new technique offers a number of advantages. First, local features are well reconstructed because new technique has the nature of local SFS approaches. Second, the global minimization and smoothing process make the new technique to be robust to noise. Third, computational demand is less than other global SFS approaches.

2. LOCAL SFS ALGORITHM

Lee and Rosenfeld proposed a local SFS algorithm that determines the surface normal locally using both the intensity and the intensity derivative information at a pixel in an image [1]. The surface normal can be represented as

$$\vec{N} = (N_x, N_y, N_z) = (\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi)\,, \tag{2}$$

0-7803-7750-8/03/$17.00 ©2003 IEEE

where ϕ and θ are the slant and tilt of the surface normal, respectively. If we assume the light source-centered coordinates, then the image irradiance equation can be rewritten as

$$\begin{aligned} I(x,y) &= \rho\vec{S}\cdot\vec{N} \\ &= \rho(0,0,1)\cdot(\sin\phi\cos\theta,\sin\phi\sin\theta,\cos\phi)^{\cdot} \\ &= \rho\cos\phi \end{aligned} \quad (3)$$

So, under the assumption that the albedo ρ is known, the slant ϕ of the surface normal can be determined by the intensity information at the pixel.

In order to determine the tilt θ of the surface normal, Lee and Rosenfeld assumed that a local surface patch could be approximated by a spherical patch. Under that assumption they proved that the tilt of the surface normal could be obtained from

$$\theta = \arctan\frac{I_y\cos\theta_S - I_x\sin\theta_S}{I_x\cos\theta_S\cos\phi_S + I_y\cos\phi_S\sin\theta_S}, \quad (4)$$

where I_x and I_y are intensity derivatives along the x and y directions, ϕ_S and θ_S are the slant and tilt of the light source direction, respectively. The slant and tilt of the surface normal computed by equation (3) and (4) is a value in the light source coordinate. Thus it should be transformed into the value in the viewer coordinate in order to determine the final surface normal vector.

In determining the tilt of the surface normal using Lee and Rosenfeld method, there exists an ambiguity between convex and concave cases. Determination of the tilt value by equation (4) results from the assumption that the surface is locally concave. Thus, the tilt in convex surface has an opposite direction to the tilt in concave case. Besides the convex/concave ambiguity, this method tends to be more sensitive to noise as the distance between the viewer and the light source get increased.

3. PROPOSED ENERGY FUNCTION

As stated above, Lee and Rosenfeld algorithm derives the surface normal vector by using the intensity and intensity gradient information. But under the local spherical assumption, there remains the ambiguity such that we cannot select the surface normal between two possible values. Therefore we must assume the surface type before applying the algorithm. So we cannot avoid large errors in regions where the assumption about the surface type is not appropriate. In this paper we aim to address this problem, proposing the algorithm that resolves the ambiguity globally using the energy minimization formulation. The energy minimization framework requires us to compose an energy functional. For that purpose we use the smoothness constraint.

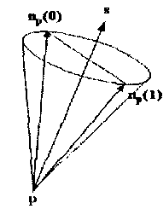

Fig. 1. Two allowable normal vectors.

Suppose that $\vec{n}_p(0)$ and $\vec{n}_p(1)$ are the two allowable surface normal vectors under the convex and concave assumptions at the pixel p, respectively, as in Figure 1. Then the smoothness constraint can be defined as

$$\begin{aligned} \varepsilon(X) &= \sum_{(p,q)\in N} V_{pq}(X_p, X_q) \\ &= \sum_{(p,q)\in N} \left\|\vec{n}_p(X_p) - \vec{n}_q(X_q)\right\|, X\in\{0,\ 1\} \end{aligned}, \quad (5)$$

where N is the set of all neighboring pixels, $V_{p,q}$ is a interaction function, and X is a function that is defined at each pixel which can take 0 or 1. By minimizing the above functional, we enforce the surface normal to change gradually, so that the convex/concave ambiguity is resolved. This assumption is suitable for general 3-D objects.

The proposed energy function has several advantages over those used in conventional global approaches. Firstly, the proposed energy function does not include the brightness constraint, while most conventional energy functions involve the brightness constraint as well as the smoothness constraint. The brightness constraint is derived directly from the image irradiance equation (1), and indicates the total brightness error of the reconstructed image compared with the input image. So, conventional global methods don't guarantee that the reconstructed pixel intensity is identical with the real pixel intensity. Therefore the reconstructed surface normal vector may have a slant value that differs from the slant value derived from the equation (3). Whereas in the case of the energy function proposed in this paper, allowable surface normal vectors at each pixel have the slant value determined by the image intensity. This results in imposing brightness constraint as a hard constraint. Similarly, Worthington has recently proposed a geometric framework for the same purpose [6]. Secondly, in the proposed energy function, the number of the allowable surface normal vectors at each pixel is limited to two, convex and concave cases. While, most of the conventional energy functions allow uncountable number of surface normal vectors as the candidates. Therefore those methods have used energy minimization technique in a continuous domain such as the variational method. On the other hand, since our energy minimization scheme is formulated in a discrete domain, so various efficient combinatorial energy minimization techniques can be applied.

4. ENERGY MINIMIZATION USING GRAPH CUTS

In order to minimize the proposed energy function, we use the graph cut method that was introduced as a combinatorial energy minimization technique [7]. This method is able to find a local minimum closer to the global minimum than other minimization techniques and the computational demand is relatively small.

In order to apply the graph cut method to the energy minimization problem, the interaction function defined between neighboring pixels, which eventually forms the energy functional, has to satisfy the metric conditions. The metric conditions for the interaction function $V(\alpha,\beta)$ is defined as follows

$$V(\alpha,\beta)=0 \Leftrightarrow \alpha=\beta, \quad V(\alpha,\beta)=V(\beta,\alpha)\geq 0, \quad V(\alpha,\beta)\leq V(\alpha,\gamma)+V(\gamma,\beta) \tag{6}$$

However, the interaction function,

$$V_{p,q}(X_p,X_q)=\left\|\vec{n}_p(X_p)-\vec{n}_q(X_q)\right\|,\ X\in\{0,1\}$$

defined in the equation (5) does not satisfy the metric conditions. Therefore, let us modify the energy function to make it satisfy the metric conditions. Let $\theta_p \in \{l_1,l_2,\ldots,l_k\}$ be the discrete tilt values of the surface normal vector at the pixel p. Then, equation (5) can be rewritten as

$$\varepsilon(\theta)=\sum_p D_p(\theta_p)+\sum_{(p,q)\in N} V_{p,q}(\phi_p,\phi_q,\theta_p,\theta_q)$$
$$D_p(\theta_p)=\begin{cases}0 & \text{if } \theta_p \text{ is a tilt in convex or concave case}\\ \infty & \text{otherwise}\end{cases} \tag{7}$$
$$V_{p,q}(\phi_p,\phi_q,\theta_p,\theta_q)=\left\|\vec{n}(\phi_p,\theta_p)-\vec{n}(\phi_q,\theta_q)\right\|.$$

Note that D_p enforces the surface normal vector at each pixel to have the tilt determined by the Lee and Rosenfeld method. And, the slant ϕ_p and ϕ_q are determined by equation (3). However, still the interaction function in equation (7) does not satisfy metric condition since $V_{p,q}(\phi_p,\phi_q,k,k)=\left\|\vec{n}(\phi_p,k)-\vec{n}(\phi_q,k)\right\|\neq 0$.

Therefore, let us define the interaction function by

$$V_{p,q}(\phi_p,\phi_q,\theta_p,\theta_q)=\left\|\vec{n}(\bar{\phi},\theta_p)-\vec{n}(\bar{\phi},\theta_q)\right\|, \quad \bar{\phi}=\min(\phi_p,\phi_q). \tag{8}$$

This interaction function is the magnitude of a new difference vector that is obtained by subtracting the original difference vector to the bias vector. This is illustrated in Figure 2. Note that this new interaction function satisfies the metric condition.

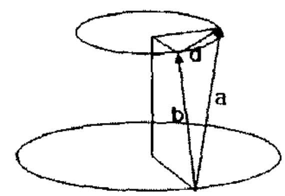

Fig.2. new difference vector **d**, original difference vector **a**, bias vector **b**

5. EXPERIMENTS

We have tested our algorithm on several synthetic and real world images and compared the results with those of conventional algorithms. Fig. 3 shows the Mozart and Lena images and the reconstruction results. The second row of Fig. 3 shows the reconstructed surface obtained by the Horn's algorithm, which is a representative global SFS method. The third row of Fig. 3 shows the result of the local method, Lee and Rosenfeld's algorithm. In the bottom row of Fig. 3, we show the result of the proposed algorithm. We note that the Horn's algorithm reconstructed surfaces without details due to oversmoothing. On the other hand, Lee and Rosenfeld's algorithm produced noisy surfaces despite its relatively good recovery of local features. Usually, this method becomes more sensitive to noise as the distance between the viewer and the light source get increased. However, the proposed algorithm captures sufficient surface details while performs proper smoothing process, thus recovers more suitable surfaces.

Finally, we have tested our algorithm on additional real world images. Figure 4 shows the reconstructed surfaces for the Pepper and Vase images. These results demonstrate that the proposed algorithm is able to reconstruct the 3D surface of various real world objects successfully.

6. CONCLUSIONS

In this paper, we have proposed a new semi-global SFS algorithm that combines a local and global energy minimization frameworks. The method uses the surface normal vectors obtained by the Lee and Rosenfeld's method as the initial allowable normals at each pixel. Then, by employing a new energy minimization formulation based on the smoothness constraint of local normal vectors, the method determines the final needle map. In order to minimize the proposed energy functional, we have applied the graph cut technique.

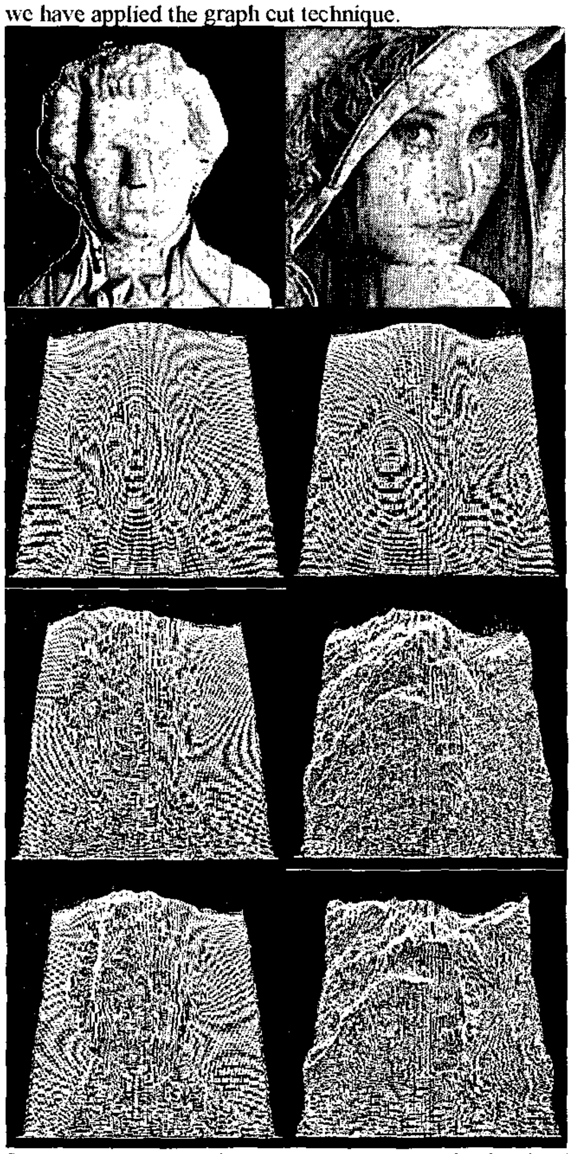

Fig. 3. Comparison of the reconstructed surfaces for the global method (Horn), the local method (Lee and Rosenfeld) and the proposed method.

Results on synthetic and real world objects reveal that the proposed algorithm produces more robust and enhanced results than the local methods, while exhibits local features better than other global methods.

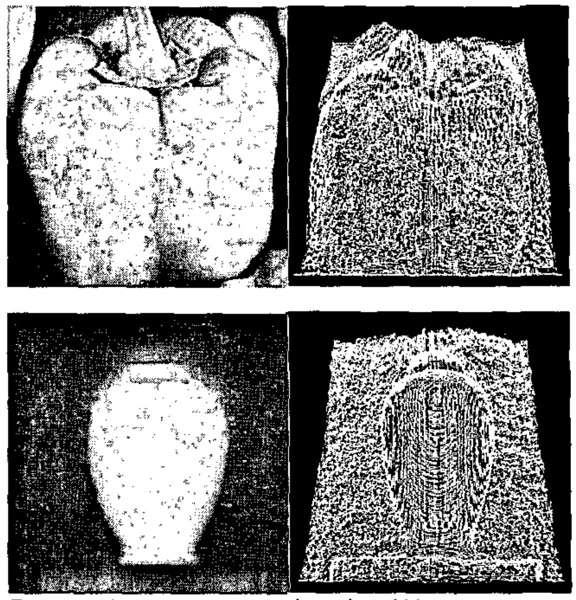

Fig. 4. Applying our method to the real world images.

7. REFERENCES

[1] C.H. Lee and A. Rosenfeld, "Improved methods of estimating shape from shading using the light source coordinate system," *Artificial Intelligence*, vol. 26, pp. 125-143, 1985.

[2] B.K.P. Horn, "Height and Gradient from Shading," *Int'l J. Computer Vision*, pp. 37-75, 1989.

[3] K.M. Lee and C.C.J. Kuo, "Shape from Shading with a Linear Triangular Element Surface Model," *IEEE Pattern Analysis and Machine Intelligence*, vol. 15, no. 8, pp. 815-822, 1993.

[4] P.S. Tsai and M. Shah, "Shape from Shading Using Linear Approximation," *Image and Vision Computing*, vol. 12, no. 8, pp. 487-498, 1994.

[5] R. Zhang, Ping-Sing Tsai, J.E. Cryer and M. Shah, "Shape from Shading: A survey," *IEEE Pattern Analysis and Machine Intelligence*, vol. 21, no. 8, pp. 690-706, 1999.

[6] P.L. Worthington and E.R. Hancock, "New Constraints on Data-closeness and Needle-map consistency for Shape-from-Shading," *IEEE Pattern Analysis and Machine Intelligence*, vol. 21, no. 12, pp. 1250-1267, 1999.

[7] Y. Boykov, O. Veksler and R. Zabih, "Fast Approximate Energy Minimization via Graph Cuts," *IEEE Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222-1239, 2001.