DIRECT COMPUTATION OF DEPTH FROM SHADING FOR PERSPECTIVE PROJECTION

Kousuke Wakabayashi, Norio Tagawa and Kan Okubo
Graduate School of System Design, Tokyo Metropolitan University, Hino-shi, Tokyo, Japan

Keywords: Depth Recovery, Direct Computation, Shape from Shading, Perspective Projection.

Abstract: We present a method for recovering shape from shading in which the surface depth is directly computed. The already proposed method solving the same problem assumes that images are captured under the parallel projection, and hence, it can be correctly used only for the relative thin objects compared with the distance from the camera. If this method is formally extended for the perspective projection completely, the complicated calculations for differential are required. This gives rise to unstable recovery. In this study, we examine an extension of this method so as to treat the perspective projection approximately. In order to keep the simplicity of the original method, we propose the simple approximation of the derivative of the surface with respect to the image coordinate.

1 INTRODUCTION

Various algorithms for shape from shading have been enthusiastically studied, but most of them compute the surface orientation rather than surface depth (Brooks and Horn, 1985), (Szeliski, 1991), (Zhang et al., 1999). Computing surface orientation gives rise to two fundamental difficulties. First, the recovering problem is under-constrained, i.e. for the each point in an image, there is one observation but two unknown. To solve this problem, additional constraints, such as smoothness of the orientations, are required to obtain a unique solution. Secondly, arbitrary two functions $p(x,y)$ and $q(x,y)$ on an image will not generally correspond to the orientations of some continuous and differential surface.

Horn (Horn, 1990) developed a method which considered solving for three functions simultaneously: a surface function $Z(x,y)$ was recovered in addition to $p(x,y)$ and $q(x,y)$, which should represent the surface orientation. In this paper, we use the capital letter (X,Y,Z) for a three-dimensional point and the small case letter (x,y) for an image point. The objective function in (Horn, 1990) includes a term $(Z_X - p)^2 + (Z_Y - q)^2$ which makes these three functions to represent the same surface, but the actually recovered surface Z never exactly corresponds to the orientations (p,q).

Thereafter, Leclerc and Bobick (Leclerc and Bobick, 1991) developed a direct method for recovering shape from shading, which directly find a surface $Z(x,y)$ that minimizes the photometric error. In this method, the surface orientation is represented explicitly as the derivative of $Z(x,y)$, and the objective function is minimized with respect to $Z(x,y)$. By this method, additional constraints to ensure integrability of the surface orientation is not needed to be considered. However, this method assumes the parallel projection for imaging, and hence applicability of it is low. To recover shape collectively using various schemes including shape from stereo (Lazaros et al., 2008) and shape from motion (Simoncelli, 1999), (Bruhn and Weicke, 2005), the perspective projection has to be considered.

If this method is formally extended for the perspective projection, the objective function becomes complicated, and hence, the computation becomes unstable. To treat the perspective projection effectively with keeping the simplicity of the original method, we propose an approximation method for the objective function, and confirm the intrinsic performance of it numerically.

2 SHAPE FROM SHADING

Almost methods for shape from shading are based on the image irradiance equation:

$$I(x,y) = R(n(x,y)), \tag{1}$$

Wakabayashi K., Tagawa N. and Okubo K..
DIRECT COMPUTATION OF DEPTH FROM SHADING FOR PERSPECTIVE PROJECTION.
DOI: 10.5220/0003854804450448
In *Proceedings of the International Conference on Computer Vision Theory and Applications* (VISAPP-2012), pages 445-448
ISBN: 978-989-8565-03-7

which represents that image intensity I at a image point (x,y) is formulated as a function R of the surface normal n at the point (X,Y,Z) on a surface projecting to (x,y) in the image. Note that $x = X/Z$ and $y = Y/Z$ hold. General R contains other variables such as a view direction, a light source direction and albedo. These variables have to be determined in advance or simultaneously with the shape from images in general, and various methods have been studied.

Most formularizations of shape from shading problem have focused on determining surface orientation using the parameters (p,q) representing (Z_X, Z_Y), which is the first derivative of Z with respect to X and Y. Hence, we can express the shape from shading problem as solving for $p(x,y)$ and $q(x,y)$, with which the irradiance equation holds, by minimizing the following objective function.

$$J \equiv \int \{I(x,y) - R(p(x,y), q(x,y))\}^2 \, dxdy. \quad (2)$$

However, this problem is highly under-constrained, and additional constraints are required to determine a particular solution, for example a smoothness constraint. Additionally, the solutions $p(x,y)$ and $q(x,y)$ will not correspond to orientations of a continuous and differential surface $Z(x,y)$ in general. Therefore, the post processing is required, which generates a surface approximately satisfying the constraint $p_Y = q_X$.

3 DEPTH FROM SHADING

3.1 Algorithm for Parallel Projection

To avoid the problems mentioned in the previous section, we can represent $p(x,y)$ and $q(x,y)$ using $Z(x,y)$ explicitly in the discrete manner.

$$p_{i,j} = \frac{1}{2\Delta x}(Z_{i+1,j} - Z_{i-1,j}), \quad (3)$$

$$q_{i,j} = \frac{1}{2\Delta y}(Z_{i,j+1} - Z_{i,j-1}), \quad (4)$$

where Δx and Δy are the sampling intervals in an image along x and y directions respectively. By the same way, second finite differences of $Z(x,y)$ can be represented as follows:

$$u_{i,j} = \frac{1}{\Delta x^2}(Z_{i+1,j} - 2Z_{i,j} + Z_{i-1,j}), \quad (5)$$

$$v_{i,j} = \frac{1}{\Delta y^2}(Z_{i,j+1} - 2Z_{i,j} + Z_{i,j-1}). \quad (6)$$

Using these representations, Leclerc and Bobick (Leclerc and Bobick, 1991) defined the following objective function.

$$E \equiv \sum_{i,j} (1-\lambda) \left\{I_{i,j} - R(p_{i,j}, q_{i,j})\right\}^2 + \lambda \left(u_{i,j}^2 + v_{i,j}^2\right). \quad (7)$$

The parameter λ represents a degree of a smoothness constraint, that is initially set as 1 and is gradually decreased to near zero. In (Leclerc and Bobick, 1991) , λ is controlled using a hierarchical technique (Terzopoulos, 1983) which uses the multi-resolution image decomposition. This objective function is iteratively minimized by the standard conjugate gradient algorithm FRPRMN in conjunction with the line search algorithm DBRENT (Press et al., 1986).

3.2 Extension for Perspective Projection

In the parallel projection model, $x = X$ and $y = Y$ holds. However, in the perspective projection model, we have to consider the relations $x = X/Z$ and $y = Y/Z$. These relations cause the following formulations which is important for the perspective projection to define the objective function of Eq. 7.

$$\frac{\partial Z}{\partial X} = \frac{1}{Z}\frac{\partial Z}{\partial x}, \quad \frac{\partial Z}{\partial Y} = \frac{1}{Z}\frac{\partial Z}{\partial y}, \quad (8)$$

$$\frac{\partial^2 Z}{\partial X^2} = \frac{1}{Z^2}\frac{\partial^2 Z}{\partial x^2} - \frac{1}{Z^3}\left(\frac{\partial Z}{\partial x}\right)^2, \quad (9)$$

$$\frac{\partial^2 Z}{\partial Y^2} = \frac{1}{Z^2}\frac{\partial^2 Z}{\partial y^2} - \frac{1}{Z^3}\left(\frac{\partial Z}{\partial y}\right)^2. \quad (10)$$

s From Eq. 8, Eqs. 3 and 4 have to be altered as follows:

$$p_{i,j} = \frac{1}{2Z_{i,j}\Delta x}(Z_{i+1,j} - Z_{i-1,j}), \quad (11)$$

$$q_{i,j} = \frac{1}{2Z_{i,j}\Delta y}(Z_{i,j+1} - Z_{i,j-1}). \quad (12)$$

However, these definitions make the computation of the gradient of E complicated. Hence, we propose approximations of Eqs. 11 and 12 using a fixed value Z_0, which may be varied with i and j and is required to be close to an actual $Z_{i,j}$.

$$\tilde{p}_{i,j} = \frac{1}{2Z_0\Delta x}(Z_{i+1,j} - Z_{i-1,j}), \quad (13)$$

$$\tilde{q}_{i,j} = \frac{1}{2Z_0\Delta y}(Z_{i,j+1} - Z_{i,j-1}). \quad (14)$$

To approximate the second derivatives of the perspective projection, Z in the first term of the right-hand side of Eqs. 9 and 10 is replaced by Z_0 and the second term in the both equations is omitted.

$$\tilde{u}_{i,j} = \frac{1}{Z_0^2\Delta x^2}(Z_{i+1,j} - 2Z_{i,j} + Z_{i-1,j}), \quad (15)$$

$$\tilde{v}_{i,j} = \frac{1}{Z_0^2 \Delta y^2} (Z_{i,j+1} - 2Z_{i,j} + Z_{i,j-1}). \tag{16}$$

To make the derivation explicit, it is essential to specify the reflection model R. As the standard model, we can employ a Lambertian reflection model.

$$R_{i,j} = R(p_{i,j}, q_{i,j}) = n_{i,j} \cdot l = \frac{a p_{i,j} + b q_{i,j} - c}{\sqrt{1 + p_{i,j}^2 + q_{i,j}^2}}, \tag{17}$$

where n_{ij} is the unit vector indicating surface normal, and $l = (a,b,c)$ is the light source vector scaled by the albedo. Although various algorithms to estimate l have been studied, in this study we assume that l is known. The objective function of Eq. 7 is rewritten with this R and the proposed approximation as follows:

$$\tilde{E} \equiv \sum_{i,j} (1-\lambda) \left\{ I_{i,j} - R(\tilde{p}_{i,j}, \tilde{q}_{i,j}) \right\}^2 + \lambda \left(\tilde{u}_{i,j}^2 + \tilde{v}_{i,j}^2 \right), \tag{18}$$

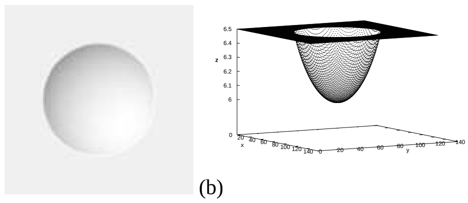
and the elements of the gradient of $\tilde{E}$ are derived as follows:

$$\begin{aligned} \frac{\partial \tilde{E}}{\partial Z_{i,j}} &= (1-\lambda) \times \\ &\left\{ \frac{I_{i-1,j} - R_{i-1,j}}{Z_0 \Delta x \sqrt{D_{i-1,j}}} \left(a - \frac{N_{i-1,j}}{D_{i-1,j}} \tilde{p}_{i-1,j} \right) \right. \\ &+ \frac{I_{i+1,j} - R_{i+1,j}}{Z_0 \Delta x \sqrt{D_{i+1,j}}} \left(-a + \frac{N_{i+1,j}}{D_{i+1,j}} \tilde{p}_{i+1,j} \right) \\ &+ \frac{I_{i,j-1} - R_{i,j-1}}{Z_0 \Delta y \sqrt{D_{i,j-1}}} \left(b - \frac{N_{i,j-1}}{D_{i,j-1}} \tilde{q}_{i,j-1} \right) \\ &\left. + \frac{I_{i,j+1} - R_{i,j+1}}{Z_0 \Delta y \sqrt{D_{i,j+1}}} \left(-b + \frac{N_{i,j+1}}{D_{i,j+1}} \tilde{q}_{i,j+1} \right) \right\} \\ &+ \frac{2\lambda}{Z_0^2} \left\{ \frac{\tilde{u}_{i+1,j} + \tilde{u}_{i-1,j} - 2\tilde{u}_{i,j}}{\Delta x^2} \right. \\ &\left. + \frac{\tilde{v}_{i,j+1} + \tilde{v}_{i,j-1} - 2\tilde{v}_{i,j}}{\Delta y^2} \right\}. \end{aligned} \tag{19}$$

3.3 Approximation Error of Depth

In this section, we assume that Z_0 is constant at the local region in the image plane. By minimizing Eq. 18, the surface $Z_{i,j}$, the orientation of which is close to the true value, is determined as a solution. Although the estimates of $\tilde{p}_{i,j}$ and $\tilde{q}_{i,j}$ corresponding to the determined surface can be considered as random variables according to the image noise, it is expected that these estimators have no bias, and the expectation values of them equals to the true values of $p_{i,j}$ and $q_{i,j}$, which are not the approximation values.

For qualitative analysis of the bias error of $\hat{Z}_{i,j}$,



(a) (b)

Figure 1: Example of the data used in the experiments: (a) artificial image; (b) true depth map.

which is the estimator of the surface by our method, we define the difference between the adjacent depths on the image plane.

$$\Delta Z_{i,j}^x = Z_{i+1,j} - Z_{i-1,j}, \tag{20}$$

$$\Delta Z_{i,j}^y = Z_{i,j+1} - Z_{i,j-1}. \tag{21}$$

Using Eqs. 11, 12, 13 and 14, and $\mathrm{E}[\hat{\tilde{p}}_{i,j}] = p_{i,j}$ and $\mathrm{E}[\hat{\tilde{q}}_{i,j}] = q_{i,j}$, the following relations can be derived.

$$\mathrm{E}\left[\hat{\tilde{\Delta Z}}{}_{i,j}^x \right] = \Delta Z_{i,j}^x + 2 p_{i,j} \Delta x (Z_0 - Z_{i,j}), \tag{22}$$

$$\mathrm{E}\left[\hat{\tilde{\Delta Z}}{}_{i,j}^y \right] = \Delta Z_{i,j}^y + 2 q_{i,j} \Delta y (Z_0 - Z_{i,j}). \tag{23}$$

In the above description, $\mathrm{E}[\cdot]$ indicates the expectation with respect to the image noise and $\hat{}$ indicates an estimator. From Eqs. 22 and 23, the gradient of the recovered depth has a statistical bias, and it is in proportion to the true value of the gradient and the difference between the true value of depth and the approximation of it.

4 NUMERICAL EVALUATIONS

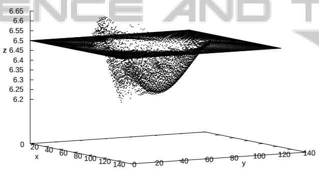
To confirm the effectiveness of the proposed method, we conducted numerical evaluations using artificial images. Figure 1(a) shows an example image generated by a computer graphics technique using the depth map shown in Fig. 1(b). The light source direction vector is set as $(0.25, 0.25, -1.0)$. The image size assumed in this study is 128×128 pixels, which corresponds to $-0.5 \leq x, y \leq 0.5$ measured using the focal length as a unit.

The proposed method can recover a surface up to scale, hence we assume that the background plane behind the hemisphere in Fig. 1(b) is known as a boundary condition. The steepest descent method was utilized for minimizing Eq. 18 with evaluating Eq. 19. Minimization of Eq. 18 can be repeatedly performed with updating Z_0 in Eq. 19 successively. If our approximation is effective, it is expected that after enough iteration of this minimization an accurate surface is recovered, and we confirmed it.

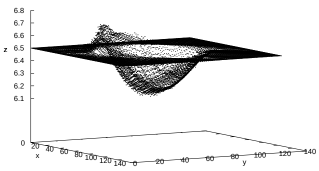
At the first minimization, a plane $Z = 6.5$ indicating the background plane shown in Fig. 1(b) was adopted as an initial value, and Z_0 in Eq. 19 is also set as this plane. We controlled the value of λ in Eq. 18 by starting with $\lambda = 1.0$ and decreasing it by 0.01 till $\lambda = 0.0$. For each value of λ, the steepest descent minimization is iterated until convergence. Differently from the method in (Leclerc and Bobick, 1991), the multi-resolution analysis was not used.

After the one minimization is finished, the obtained surface, which is not a plane any longer, are used as Z_0 in Eq. 19 and also as an initial value for the following minimization. This process is repeated to decrease the recovering error caused by the approximation of our method.

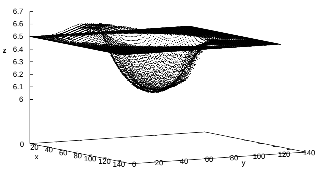
The recovered surfaces are shown in Fig. 2 for the each repetition of minimization. Figure 3 shows the RMSEs of the recovered surface with respect to the repetition number. From these results, the proposed method can be used repeatedly for recovering the accurate surface with stable computation.



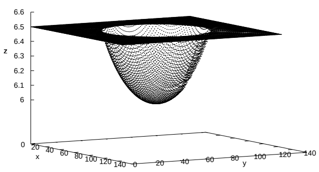
(a)



(b)



(c)



(d)

Figure 2: Recovered surfaces by proposed method: (a) repetition number is 1; (b) 2; (c) 5; (d) 10.

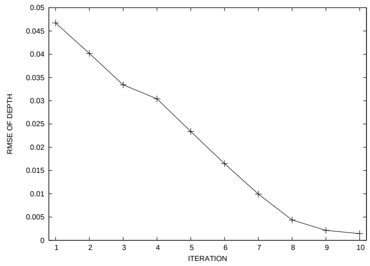


Figure 3: RMSE of recovered surface with repetition number of proposed minimization.

5 CONCLUSIONS

We propose a direct depth recovery method from shading information, which can be applied to the perspective projection. In this method, the representation of the surface gradient is approximated to avoid the complicated computation, which is caused by the straight-forward extension of the parallel projection method. Through numerical evaluations, we confirmed that the repeated application of our minimization method can recover a good surface.

REFERENCES

Brooks, M. J. and Horn, B. K. P. (1985). Shape and source from shading. In *proc. Int. Joint Conf. Art. Intell.*, pages 18–23.

Bruhn, A. and Weicke, J. (2005). Locas/kanade meets horn/schunk: combining local and global optic flow methods. *Int. J. Comput. Vision*, 61(3):211–231.

Horn, B. K. P. (1990). Height and gradient from shading. *Int. J. Computer Vision*, 5(1):37–75.

Lazaros, N., Sirakoulis, G. C., and Gasteratos, A. (2008). Review of stereo vision algorithm: from software to hardware. *Int. J. Optomechatronics*, 5(4):435–462.

Leclerc, Y. G. and Bobick, A. F. (1991). The direct computation of height from shading. In *proc. CVPR '91*, pages 552–558.

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1986). *Numerical methods, the art of scientific computing*. Cambridge U. Press, Cambridge, MA.

Simoncelli, E. P. (1999). Bayesian multi-scale differential optical flow. In *Handbook of Computer Vision and Applications*, pages 397–422. Academic Press.

Szeliski, R. (1991). Fast shape from shading. *CVGIP: Image Understanding*, 53(2):129–153.

Terzopoulos, D. (1983). Multilevel computational processes for visual surface reconstruction. *Computer Vision, Graphics, and Image Processing*, 24:52–96.

Zhang, R., Tsai, P.-S., Cryer, J. E., and Shah, M. (1999). Shape from shading: A survey. *IEEE Trans. PAMI*, 21(8):690–706.