# Shape From Shading

Shape from shading aims to recover the $shape$ of an object from a measurement of the intensity of light reflected off of the object's surface.

Physical modeling of the imaging process leads to the following model:

$$I(x, y) = \rho(x, y)R(p(x, y), q(x, y))$$

$I(x, y)$ is the sensed image.

$\rho(x, y)$ is the surface albedo (ratio of surface brightness to that of a reference Lambertian surface)

$p(x, y)$ and $q(x, y)$ are the slopes of the object surface in the $x$ and $y$ directions:

$$p(x, y) = \frac{\partial Z(x, y)}{dx}, \qquad q(x, y) = \frac{\partial Z(x, y)}{dy}$$

$Z(x, y)$ is the surface "height" function.

$R(p, q)$ is the $Reflectance\ Map$ of the surface.

For a Lambertian surface with a point light source located at infinity, and at a direction $\hat{s}$ relative to the surface,

$$R_l(p,q) = \frac{(1 + p_s p + q_s q)}{\sqrt{1+p^2+q^2}\sqrt{1+p_s^2+q_s^2}}$$

$$\hat{s} = (-p_s, -q_s, 1)/\sqrt{1+p_s^2+q_s^2}$$

The $(p,q)$ values are related to the unit vector, $\hat{n}$, normal to the surface:

$$\hat{n} = \frac{(-p, -q, 1)}{\sqrt{1+p^2+q^2}}$$

$$R_l(\hat{n}) = \hat{n} \cdot \hat{s}$$

For a Specular (mirror-like) surface with a point light source located at infinity, and at a direction $\hat{s}$ relative to the surface,

$$R_s(x, y) = (\hat{k} \cdot \hat{h}(x, y))^m$$

$\hat{k}$ is the unit vector in the viewer direction

$\hat{h}$ is a unit vector that depends on $\hat{s}$ and $\hat{n}$ (or $(p, q)$)

$$\hat{h}(x, y) = 2(\hat{n}(x, y) \cdot \hat{s})\hat{n}(x, y) - \hat{s}$$

The parameter $m$ is a number (assumed to be known) corresponding to the sharpness of the specularity.

# Regularization of Shape from Shading

The shading equations involve two unknowns ($(p, q)$ or any two independent components of $\hat{n}$) and one known $I(x, y)$ at each image point $x, y$.

Thus, there can be an infinity of solutions.

This ambiguity can be resolved by adding a very small amount of information, such as the $(p, q)$ values on some 1-D contour in the image (as in Horn's characteristics approach).

Solution stability remains as a problem, however. Thus, regularization has been applied to the shape from shading problem.

# Horn-Brooks Method

(Horn, B.K.P., and Brooks, M.J., "The variational approach to shape from shading.", MIT AI Lab memo No. 813, March 1985)

Minimize the following functional:

$$\int\int \left((I - R(\hat{n}))^2 + \lambda(|\frac{\partial \hat{n}}{\partial x}|^2 + |\frac{\partial \hat{n}}{\partial y}|^2) + \mu(x,y)(|\hat{n}|^2 - 1)\right) dxdy$$

$\mu$ is a Lagrange multiplier which must be chosen so as to enforce the hard constraint that $|\hat{n}| = 1$.

The term multiplied by $\lambda$ is a "smoothness" constraint (Tikhonov stabilizer)

The Euler equations are, for Labertian shading:

$$(I - \hat{n} \cdot \hat{s})\hat{s} + \lambda \nabla^2 \hat{n} - \mu \hat{n} = 0$$

These can be discretized as:

$$\hat{n}_{i,j}^{k+1} = \frac{4\lambda^2}{4\lambda^2 + \mu_{i,j}\epsilon^2}\left(\bar{n}_{i,j}^k + \frac{\epsilon^2}{4\lambda}(I_{i,j} - \hat{n}_{i,j}^k \cdot \hat{s})\hat{s}\right)$$

$$\bar{n}_{i,j} = \frac{1}{4}(\hat{n}_{i,j+1} + \hat{n}_{i,j-1} + \hat{n}_{i+1,j} + \hat{n}_{i-1,j})$$

# Data Fusion

Some problems are ill-posed due to lack of information.

Data Fusion aims to combine different sources of information in order to obtain unique and, hopefully, stable solutions.

Usually, however, data fusion algorithm designers concentrate on the uniqueness problem and pay little attention to the stability problem.

One should *always*, if possible, do a stability or sensitivity analysis of their vision algorithms.

If an analytical closed form solution is available then one can usually derive expressions for the derivatives of the solution elements with respect to the data elements.

It often turns out that Data Fusion algorithms need to be regularized to achieve stability.

# Example of Data Fusion

(Clark, J.J. and Yuille, A.L., "Shape from shading via the fusion of specular and Lambertian image components", Proceedings of the International Conference on Pattern Recognition, June 1990.
Clark, J.J., and Yuille, A.L., **Data Fusion for Sensory Information Processing Systems,** Kluwer Academic Publishers, Boston, MA, 1990)

This technique assumes that we can decompose the image intensity into specular and Lambertian components.

Two ways that this can be done are:

- Use colour images. Specular reflections take the colour of the light source, while Lambertian reflections take the colour of the surface.

  (Klinker, G.J., Shafer, S.A., and Kanade, T., "Color image analysis with an intrinsic reflection model", Proceedings of the 2nd International Conference on Computer Vision, Tampa, Florida, Dec. 1988, pp 292-296)

- Polarization. Specular and Lambertian surfaces affect the polarization of incident light in different ways.

  Wolff, L.B., "Using polarization to separate reflectance components", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, San Diego, 1989, pp 361-369

We can combine the two shading equations for the specular and Lambertian image components and solve for the surface normal:
(left as an exercise!)

$$\hat{n}(\hat{x}) = \alpha(\hat{x})\hat{s} + \beta(\hat{x})\hat{k} + \gamma(\hat{x})\hat{s} \times \hat{k}$$

where

$$\alpha = \frac{1}{s^2\Omega}(E_l(\vec{x}) - \frac{E_s^{\frac{1}{m}}(\hat{x})c\Omega}{2E_l(\hat{x})} - \frac{c^2\Omega}{2E_l(\hat{x})})$$

$$\beta = \frac{1}{s^2\Omega}(\frac{E_s^{\frac{1}{m}}(\vec{x})}{2E_l(\hat{x})} + \frac{c\Omega}{2E_l(\hat{x})} - c\Omega E_l(\vec{x}))$$

$$\gamma = \pm\frac{(4s^2\Omega E_l^2 - 4E_l^4 - (E_s^{\frac{1}{m}} + c\Omega)^2 + 4E_l^2(E_s^{\frac{1}{m}} + c\Omega)c\Omega)^{1/2}}{2E_l s\Omega}$$

with $c\Omega = \cos\Omega = \hat{k} \cdot \hat{s}$ being the cosine of the angle between the viewer direction and the light source, and $s\Omega = \sin\Omega$.

# Stability of the Data Fusion

Suppose the data is perturbed by $\delta I_l$ and $\delta I_s$

The perturbation in the computed $\hat{n}$ is then:

$$\delta\hat{n} = (\delta\alpha)\hat{s} + (\delta\beta)\hat{k} + (\delta\gamma)\hat{s}\times\hat{k}$$

where

$$\delta\alpha = \frac{1}{s^2\Omega}\Big(\left[1 + \frac{E_s^{1/m}c\Omega}{2E_l^2} + \frac{c^2\Omega}{2E_l^2}\right]\delta E_l - \left[\frac{c\Omega}{2mE_lE_s^{(m-1)/m}}\right]\delta E_s\Big)$$

$$\delta\beta = \frac{1}{s^2\Omega}\Big(-\left[c\Omega + \frac{c\Omega}{2E_l^2} + \frac{E_s^{1/m}}{2E_l^2}\right]\delta E_l + \left[\frac{1}{2mE_lE_s^{(m-1)/m}}\right]\delta E_s\Big)$$

$$\delta\gamma = \frac{-1}{\tau s^2\Omega}\left[\frac{(E_s^{1/m}+c\Omega)^2 - 4E_l^4}{2E_l^2}\right]\delta E_l + \frac{-1}{\tau s^2\Omega}\left[\frac{2c\Omega E_l^2 - E_s^{1/m} - c\Omega}{2mE_lE_s^{(m-1)/m}}\right]\delta E_s$$

where

$$\tau = \pm(4s^2\Omega E_l^2 - 4E_l^4 - (E_s^{1/m}+c\Omega)^2 + 4E_l^2(E_s^{1/m}+c\Omega)c\Omega)^{1/2}$$

Instabilities occur as $\Omega \mapsto 0$, $E_l \mapsto 0$, $E_s \mapsto 0$ and as $\tau \mapsto 0$.

The instability as $\tau \mapsto 0$ occurs when $\hat{s}$, $\hat{n}$ and $\hat{k}$ are coplanar and $\gamma = 0$, it is non-generic in the sense that any small perturbation of $\hat{n}$ will move one away from the locus of instability.

The instability as $\Omega \mapsto 0$ can only be avoided when one has control over the light source position or the viewer position as this case occurs when $\hat{k}$ and $\hat{s}$ are parallel. The instability is generic in that it occurs independently of the value of $\hat{n}$, and is therefore unaffected by perturbations in $\hat{n}$.

In this situation information about only one component of $\hat{n}$ is available. Similarly the instabilities as $E_l \mapsto 0$ and $E_s \mapsto 0$ cannot be avoided.

## Estimates of Error Variances

Assume that $\delta I_l$ and $\delta I_s$ are Gaussian, with zero mean and variances $\sigma_l^2$ and $\sigma_s^2$ respectively.

Linearize the variation of $\delta\hat{n}$ with respect to $\delta I_s$ and $\delta I_l$ to give:

$$\delta\hat{n} = \left(\frac{\partial\hat{n}(I_s, I_l, \hat{s}, \hat{k})}{\partial I_s}\right)\delta I_s + \left(\frac{\partial\hat{n}(I_s, I_l, \hat{s}, \hat{k})}{\partial I_l}\right)\delta I_l$$

where

$$\frac{\partial\hat{n}}{\partial I_s} = \frac{\left[\hat{k} - \hat{s}c\Omega + \frac{1}{\tau}(\hat{s}\times\vec{k})(2c\Omega I_l^2 - I_s^{\frac{1}{m}} - c\Omega)\right]}{2mI_l I_s^{\frac{m-1}{m}} s\Omega}$$

$$\begin{aligned}\frac{\partial\hat{n}}{\partial I_l} &= \frac{\vec{s}(2I_l^2 + I_s^{\frac{1}{m}}c\Omega + c^2\Omega) - \hat{k}(2I_l^2 c\Omega + c\Omega + I_s^{\frac{1}{m}})}{2I_l^2 s\Omega} \\ &\quad - \frac{(\hat{s}\times\hat{k})((I_s^{\frac{1}{m}} + c\Omega)^2 - 4I_l^4)}{2\tau I_l^2 s\Omega} \qquad (1)\end{aligned}$$

The linear approximation to $\delta\hat{n}$ given above will be zero mean gaussian:

$$\hat{\sigma}_n^2 = E\{\delta\hat{n}^2\} = E\{(\delta n_x^2, \delta n_y^2)\} = \left(\frac{\partial\hat{n}}{\partial I_s}\right)^2\sigma_s^2 + \left(\frac{\partial\hat{n}}{\partial I_l}\right)^2\sigma_l^2$$

# Regularization of the Data Fusion

We can use regularization to stabilize the data fusion process.

Minimize the following functional, with respect to $\hat{n}$:

$$\int[w_l(I_l - \hat{n}\cdot\hat{s})^2 + w_s(I_s - (\hat{k}\cdot\hat{h})^m)^2 + \mu(||\hat{n}|| - 1)^2 + \lambda(||\nabla\hat{n}||^2)]dA$$

The weight $\lambda$ determines the amount of smoothing done on the resulting surface normal field.

The choice of $\lambda$ affects the stability of the numerical minimization process. It cannot be too small.

$w_l$ and $w_s$ are weights related to the reliability associated with the input data sources $I_l$ and $I_s$.

They are chosen to be the inverse of the maximum update in the discretized solution process.

$$w_s = \frac{1}{\max|(I_s - R_s)\frac{\partial R_s}{\partial\hat{n}}|}$$

and

$$w_l = \frac{1}{\max|(I_l - R_l)\frac{\partial R_l}{\partial\hat{n}}|}$$

# Adaptive Weighting

The previous weighting scheme only uses information about the noise in the input data.

We also have information about the sensitivity of the shape from shading algorithm to noise in the inputs.

We should make use of this information!
Reduce the weights when $\partial\hat{n}/\partial I_{l,s}$ is high (i.e. where $\hat{n}$ is sensitive to noise in $I$)

$$w_s(x,y) = w_{s_{max}} \left[\frac{\min(\sigma_l^2, \sigma_s^2)}{\sigma_s^2}\right] \frac{1}{1 + \log\left(1 + |\frac{\partial\hat{n}(x,y)}{\partial I_s}|\right)}$$

$$w_l(x,y) = w_{l_{max}} \left[\frac{\min(\sigma_l^2, \sigma_s^2)}{\sigma_l^2}\right] \frac{1}{1 + \log\left(1 + |\frac{\partial\hat{n}(x,y)}{\partial I_l}|\right)}$$