Depth from Gradient

Ying Xiong
School of Engineering and Applied Sciences
Harvard University
yxiong@seas.harvard.edu

Created: January 25th, 2014

1 Problem Statement

Given a (possibly noisy) gradient field $(p(x, y), q(x, y))$, we want to find a depth map $Z(x, y)$ such that

$$Z_x = \frac{\partial Z}{\partial x} = p, \quad Z_y = \frac{\partial Z}{\partial y} = q. \tag{1}$$

Define an error function on proposed depth Z with respect to given gradient (p, q), written as $E(Z; p, q)$. One simple and natural error function is

$$E(Z; p, q) = (Z_x - p)^2 + (Z_y - q)^2. \tag{2}$$

For more general error function, we refer readers to [1].

Note that the error function E is also a map defined on every point (x, y), and the problem of finding optimal $Z(x, y)$ can be formulated as a minimization of cost function

$$\text{cost}(Z) = \iint E(Z; p, q) dx dy. \tag{3}$$

2 Frankot Chellappa Algorithm [2]

2.1 General approach

Assuming the depth map can be written as a linear combination of basis function $\phi(x, y; \boldsymbol{\omega})$, where $\boldsymbol{\omega} = (\omega_x, \omega_y) = (u, v)$ is the a 2D index

$$Z(x, y) = \sum_{\boldsymbol{\omega}} C(\boldsymbol{\omega}) \phi(x, y; \boldsymbol{\omega}). \tag{4}$$

We write the partial derivatives of basis function as

$$\phi_x(x, y; \boldsymbol{\omega}) = \frac{\partial \phi}{\partial x}(x, y, \boldsymbol{\omega}), \quad \phi_y(x, y; \boldsymbol{\omega}) = \frac{\partial \phi}{\partial y}(x, y; \boldsymbol{\omega}), \tag{5}$$

and define

$$P_x(\boldsymbol{\omega}) = \iint |\phi_x(x, y; \boldsymbol{\omega})|^2 dx dy, \quad P_y(\boldsymbol{\omega}) = \iint |\phi_y(x, y; \boldsymbol{\omega})|^2 dx dy. \tag{6}$$

Theorem 1. *Given that members of* $\{\phi_x(x,y;\boldsymbol{\omega})\}_{\boldsymbol{\omega}}$ *as well as members of* $\{\phi_y(x,y;\boldsymbol{\omega})\}_{\boldsymbol{\omega}}$ *are mutually orthogonal, the best coefficients* $\widehat{C}(\boldsymbol{\omega})$ *in the (4) that minimizes the cost function (3) with square error (2) is*

$$\widehat{C}(\boldsymbol{\omega}) = \frac{P_x(\boldsymbol{\omega})\,\widehat{C}_1(\boldsymbol{\omega}) + P_y(\boldsymbol{\omega})\,\widehat{C}_2(\boldsymbol{\omega})}{P_x(\boldsymbol{\omega}) + P_y(\boldsymbol{\omega})}, \tag{7}$$

where $\widehat{C}_1(\boldsymbol{\omega})$ *and* $\widehat{C}_2(\boldsymbol{\omega})$ *comes from expansion*

$$p(x,y) = \sum_{\boldsymbol{\omega}} \widehat{C}_1(\boldsymbol{\omega})\,\phi_x(x,y;\boldsymbol{\omega}), \quad q(x,y) = \sum_{\boldsymbol{\omega}} \widehat{C}_2(\boldsymbol{\omega})\,\phi_y(x,y;\boldsymbol{\omega}). \tag{8}$$

2.2 Discrete Fourier basis

We use discrete Fourier basis for its computational efficiency

$$\phi(x,y;\boldsymbol{\omega}) = \exp\left(j2\pi\left(\frac{xu}{N} + \frac{yv}{M}\right)\right), \tag{9}$$

where M and N are the dimensions of the image. The partial derivatives of the basis is

$$\phi_x(x,y;\boldsymbol{\omega}) = \frac{j2\pi u}{N}\phi(x,y;\boldsymbol{\omega}), \quad \phi_y = \frac{j2\pi v}{M}\phi(x,y;\boldsymbol{\omega}), \tag{10}$$

and their powers are

$$P_x(\boldsymbol{\omega}) = \left(\frac{2\pi u}{N}\right)^2, \quad P_y(\boldsymbol{\omega}) = \left(\frac{2\pi v}{M}\right)^2. \tag{11}$$

The expansion coefficients $\widehat{C}_1(\boldsymbol{\omega})$ and $\widehat{C}_2(\boldsymbol{\omega})$ can be calculated from the Discrete Fourier Transform (DFT) of p and q, written as $\widehat{C}_p(\boldsymbol{\omega})$ and $\widehat{C}_q(\boldsymbol{\omega})$,

$$\widehat{C}_1(\boldsymbol{\omega}) = -\frac{jN}{2\pi u}\frac{\widehat{C}_p(\boldsymbol{\omega})}{MN}, \quad \widehat{C}_2(\boldsymbol{\omega}) = -\frac{jM}{2\pi v}\frac{\widehat{C}_q(\boldsymbol{\omega})}{MN}. \tag{12}$$

Putting everything together, the final output Z with respect to input p and q can be written as

$$Z = \mathcal{F}^{-1}\left\{-j\frac{\frac{2\pi u}{N}\mathcal{F}\{p\} + \frac{2\pi v}{M}\mathcal{F}\{q\}}{\left(\frac{2\pi u}{N}\right)^2 + \left(\frac{2\pi v}{M}\right)^2}\right\} = \mathcal{F}^{-1}\left\{-\frac{j}{2\pi}\cdot\frac{\frac{u}{N}\mathcal{F}\{p\} + \frac{v}{M}\mathcal{F}\{q\}}{\left(\frac{u}{N}\right)^2 + \left(\frac{v}{M}\right)^2}\right\}, \tag{13}$$

where $\mathcal{F}\{\cdot\}$ and $\mathcal{F}^{-1}\{\cdot\}$ are DFT and inverse DFT operations, respectively.

2.2.1 Implementation notes

1. The frequency indices (u,v) should range from $(-\lfloor N/2 \rfloor, \lfloor M/2 \rfloor)$ to $(\lceil N/2 \rceil, \lceil M/2 \rceil)$, not from $(0,0)$ to $(M-1, N-1)$. This will affect the calculation of derivatives in (10).

2. The DC compoent of depth can not be inferred from the gradient field. In the algorithm, when $\boldsymbol{\omega} = (0,0)$, the estimated $\widehat{C}(\boldsymbol{\omega})$ will be $\frac{0}{0}$ (or $\frac{\epsilon}{0}$), and should be reset to a given number, say 0.

3. The Fourier basis assumes periodic depth map

   $$Z(x,0) = Z(x,M), \quad Z(0,y) = Z(N,y). \tag{14}$$

   This will affect the calculation of gradient (1) at the boundary. To circumvent this restriction, we pad the surface as

   $$\widetilde{Z} = \begin{bmatrix} Z(x,y) & Z(N-x,y) \\ Z(x,M-y) & Z(N-x,M-y) \end{bmatrix}, \tag{15}$$

   with corresponding gradient fields

   $$\widetilde{p} = \begin{bmatrix} p(x,y) & -p(N-x,y) \\ p(x,M-y) & -p(N-x,M-y) \end{bmatrix}, \quad \widetilde{q} = \begin{bmatrix} q(x,y) & q(N-x,y) \\ -q(x,M-y) & -q(N-x,M-y) \end{bmatrix}. \tag{16}$$

References

[1] Amit Agrawal, Ramesh Raskar, and Rama Chellappa. What is the range of surface reconstructions from a gradient field? In *Computer Vision–ECCV 2006*, pages 578–591. Springer, 2006.

[2] Robert T. Frankot and Rama Chellappa. A method for enforcing integrability in shape from shading algorithms. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 10(4):439–451, 1988.