STRUCTURE FROM MOTION

Problem

- Given optical flow or point correspondences, compute 3-D motion (translation and rotation) and depth.

Main Points

- What projection?
- How many points?
- How many frames?
- Single Vs multiple motions.
- Mostly non-linear problem.

Orthographic Projection (displacement)

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = R\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + T = \begin{bmatrix} 1 & -a & b \\ a & 1 & g \\ -b & g & 1 \end{bmatrix}\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix}$$

Perspective Projection (displacement)

$$x' = \frac{x - a y + b + \frac{T_X}{Z}}{-bx + gy + 1 + \frac{T_Z}{Z}}$$

$$y' = \frac{a x + y + g + \frac{T_Y}{Z}}{-bx + gy + 1 + \frac{T_Z}{Z}}$$

Orthographic Projection (optical flow)

Perspective Projection (optical flow)

$$u = f\left(\frac{V_1}{Z} + \Omega_2\right) - \frac{V_3}{Z}x - \Omega_3 y - \frac{\Omega_1}{f}xy + \frac{\Omega_2}{f}x^2$$

$$v = f\left(\frac{V_2}{Z} - \Omega_1\right) + \Omega_3 x - \frac{V_3}{Z}y + \frac{\Omega_2}{f}xy - \frac{\Omega_1}{f}y^2$$

Structure from Motion

Heeger & Jepson sfm method

Three Step Algorithm

- Find Translation using search.
- Find Rotation using least squares fit.
- Find Depth.

Heeger & Jepson sfm method

$$u = -(\frac{V_1}{Z} + \Omega_2) + \frac{V_3}{Z}x + \Omega_3 y + \Omega_1 xy - \Omega_2 x^2$$

$$v = -(\frac{V_2}{Z} - \Omega_1) - \Omega_3 x + \frac{V_3}{Z}y - \Omega_2 xy + \Omega_1 y^2$$

$$\begin{bmatrix} u(x,y) \\ v(x,y) \end{bmatrix} = \frac{1}{Z(x,y)} \begin{bmatrix} -1 & 0 & x \\ 0 & -1 & y \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix} + \begin{bmatrix} xy & -(1+x^2) & y \\ 1+y^2 & -xy & -x \end{bmatrix} \begin{bmatrix} \Omega_1 \\ \Omega_2 \\ \Omega_3 \end{bmatrix}$$

Heeger & Jepson sfm method

$$\begin{bmatrix} u(x,y) \\ v(x,y) \end{bmatrix} = \frac{1}{z(x,y)} \begin{bmatrix} -1 & 0 & x \\ 0 & -1 & y \end{bmatrix} \mathbf{V} + \begin{bmatrix} xy & -(1+x^2) & y \\ 1+y^2 & -xy & -x \end{bmatrix} \Omega$$

$$\Theta(\mathbf{x},\mathbf{y}) = p(x,y)\mathbf{A}(x,y)\mathbf{V} + \mathbf{B}(x,y)\Omega$$

One point (x,y)

Heeger & Jepson sfm method

$$\begin{bmatrix} \Theta(x_1,y_1) \\ \vdots \\ \vdots \\ \vdots \\ \Theta(x_n,y_n) \end{bmatrix} = \begin{bmatrix} \mathbf{A}(x_1,y_1)\mathbf{V} & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ \mathbf{0} & \cdots & \cdots & \cdots & \mathbf{A}(x_n,y_n)\mathbf{V} \end{bmatrix} \begin{bmatrix} p(x_1,y_1) \\ \vdots \\ \vdots \\ p(x_n,y_n) \\ \end{bmatrix} + \begin{bmatrix} \mathbf{B}(x_1,y_1) \\ \vdots \\ \vdots \\ \\ \mathbf{B}(x_n,y_n) \end{bmatrix} \Omega$$

n points

Heeger & Jepson sfm method

$$\begin{bmatrix} \Theta(x_1, y_1) \\ \vdots \\ \vdots \\ \vdots \\ \Theta(x_n, y_n) \end{bmatrix} = \begin{bmatrix} \mathbf{A}(x_1, y_1)\mathbf{V} & \cdots & \cdots & \cdots & 0 & \mathbf{B}(x_1, y_1) \\ \vdots & & & & & \vdots \\ \vdots & & & & & \vdots \\ \vdots & & & & & \vdots \\ \mathbf{0} & \cdots & \cdots & \cdots & \mathbf{A}(x_n, y_n)\mathbf{V} & \mathbf{B}(x_n, y_n) \end{bmatrix} \begin{bmatrix} p(x_1, y_1) \\ \vdots \\ \vdots \\ p(x_n, y_n) \\ \Omega \end{bmatrix}$$

$$\Theta = \mathbf{C}(\mathbf{V})\mathbf{q}$$

Heeger & Jepson sfm method

$$E(\mathbf{V},\mathbf{q}) = \left\| \Theta - \mathbf{C}(\mathbf{V})\mathbf{q} \right\|^2$$

$$\mathbf{E}(\mathbf{V}) = \left\| \Theta^T \mathbf{C}^{\perp}(V) \right\|^2$$

Orthogonal complement to

Find translation by search.

$$\mathbf{C}(\mathbf{V}) = \overline{\mathbf{C}}(\mathbf{V})\mathbf{U}(\mathbf{V})$$

QR decomposition

Orthonormal &
Upper triangular

$$E(\mathbf{V},\mathbf{q}) = \left\| \Theta - \overline{\mathbf{C}}(\mathbf{V})\mathbf{U}(\mathbf{V})\mathbf{q} \right\|^2$$

$$E(\mathbf{V},\mathbf{q}) = \left\| \Theta - \overline{\mathbf{C}}(\mathbf{V})\overline{\mathbf{q}} \right\|^2$$

$$\Downarrow \quad \text{minimize}$$

$$\hat{\mathbf{q}} = \overline{\mathbf{C}}^{\mathbf{T}}(\mathbf{V})\Theta$$

$$E(\mathbf{V}) = \left\| \Theta - \overline{\mathbf{C}}(\mathbf{V})\overline{\mathbf{C}}^{T}(\mathbf{V})\Theta \right\|^2$$

$$E(\mathbf{V}) = \left\| (I - \overline{\mathbf{C}}(\mathbf{V})\overline{\mathbf{C}}^{T}(\mathbf{V}))\Theta \right\|^2$$

Null space

$$\mathbf{E}(\mathbf{V}) = \left\| \Theta^{T}\mathbf{C}^{\perp}(V) \right\|^2$$

Translation

Unit vector translation can be represented by spherical coordinates:

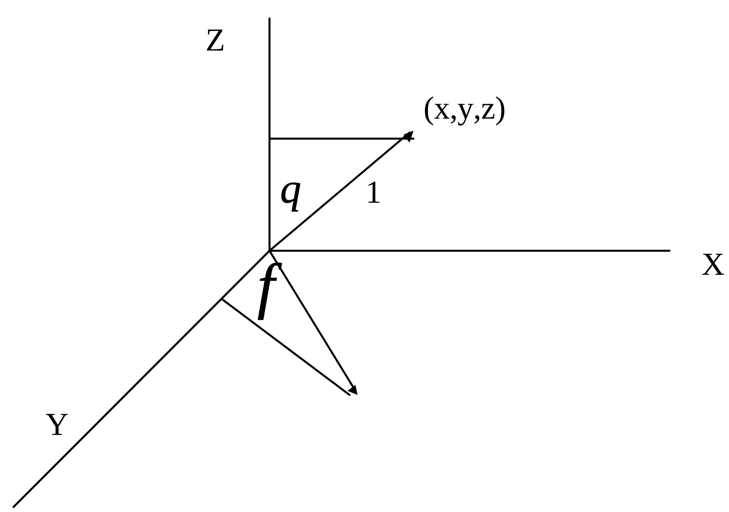
$$\mathbf{V} = (\sin q \cos f, \sin q \sin f, \cos q)$$

$0 \le q \le 90$ Slant

$0 \le f \le 360$ Tilt

Only half of sphere can be considered

Spherical Coordinates



Rotation

$$\Theta(\mathbf{x}, \mathbf{y}) = p(x,y)\mathbf{A}(x, y)\mathbf{V} + \mathbf{B}(x, y)\Omega$$

$$d^T(x, y, V)\Theta(\mathbf{x}, \mathbf{y}) = d^T(x, y, V)\mathbf{B}(x, y)\Omega$$

$d^T(x, y, V)$ is perpendicular to

Rotation

Find rotation by least squares fit

$$d^T(x_1, y_1, V)\Theta(\mathbf{x}_1, \mathbf{y}_1) = d^T(x_1, y_1, V)\mathbf{B}(x_1, y_1)\Omega$$

$$\vdots$$

$$d^T(x_n, y_n, V)\Theta(\mathbf{x}_n, \mathbf{y}_n) = d^T(x_n, y_n, V)\mathbf{B}(x_n, y_n)\Omega$$

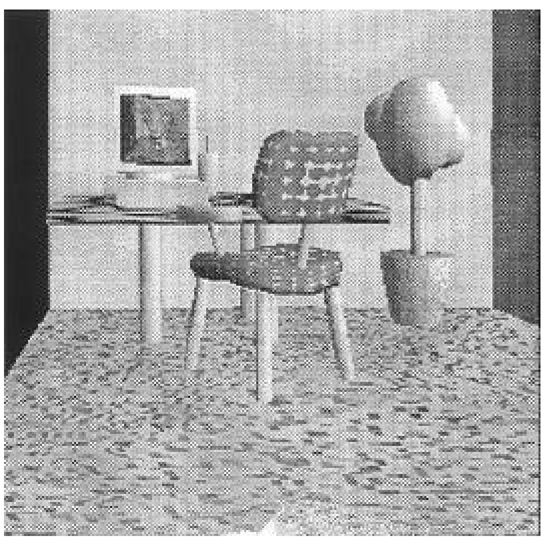
Depth

Find depth for each pixel (x,y) from following eqs

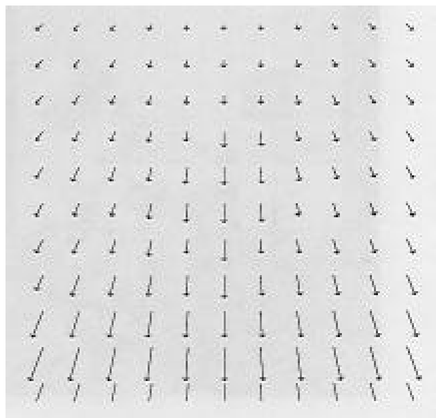
$$u = -(\frac{V_1}{Z} + \Omega_2) + \frac{V_3}{Z}x + \Omega_3 y + \Omega_1 xy - \Omega_2 x^2$$

$$v = -(\frac{V_2}{Z} - \Omega_1) - \Omega_3 x + \frac{V_3}{Z}y - \Omega_2 xy + \Omega_1 y^2$$

Synthetic Image

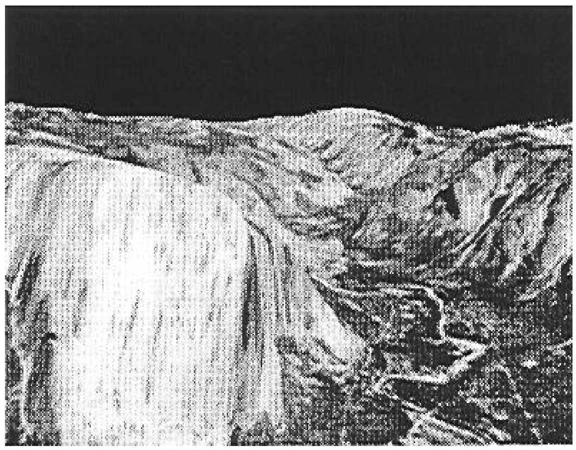


Optical Flow

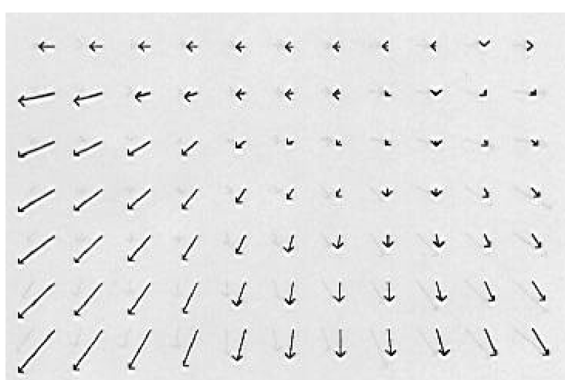


Computed Depth Map

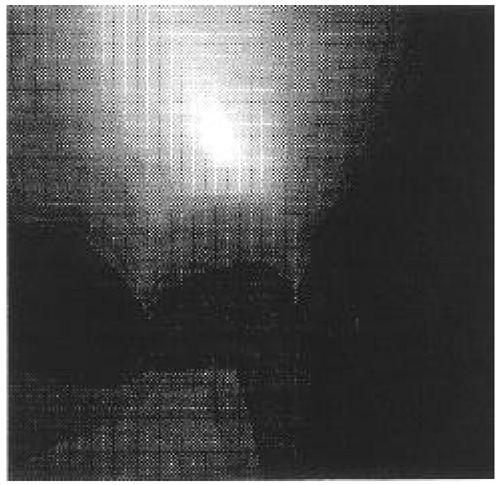
Synthetic Image



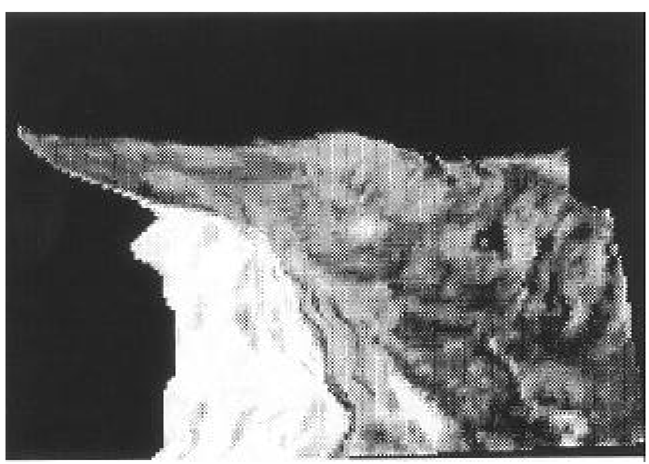
Optical Flow



Translation Search Space



Novel View Generated from Reconstructed Depth



Another Novel View Generated from Reconstructed Depth