Lecture-19

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.

Displacement Model

Point Correspondences

3-D Rigid Motion (displacement)

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = R\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + T = \begin{bmatrix} 1 & -\Delta\phi & \Delta\psi \\ \Delta\phi & 1 & -\Delta\theta \\ -\Delta\psi & \Delta\theta & 1 \end{bmatrix}\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix}$$

$$X' = X - \Delta\phi Y + \Delta\psi Z + T_X$$
$$Y' = \Delta\phi X + Y - \Delta\theta Z + T_Y$$
$$Z' = -\Delta\psi X + \Delta\theta Y + Z + T_Z$$

Orthographic Projection (displacement model)

$$X' = X - \Delta\phi Y + \Delta\psi Z + T_X$$
$$Y' = \Delta\phi X + Y - \Delta\theta Z + T_Y$$
$$Z' = -\Delta\psi X + \Delta\theta Y + Z + T_Z$$

$$x' = x - \Delta\phi y + \Delta\psi Z + T_X$$
$$y' = \Delta\phi x + y - \Delta\theta Z + T_Y$$

Perspective Projection (displacement)

$$X' = X - \gamma Y + \beta Z + T_X$$

$$Y' = \gamma X + Y - \alpha Z + T_Y$$

$$Z' = -\beta X + \alpha Y + Z + T_Z$$

$$x' = \frac{x - \gamma y + \beta + \frac{T_X}{Z}}{-\beta x + \alpha y + 1 + \frac{T_Z}{Z}}$$

$$y' = \frac{\gamma x + y - \alpha + \frac{T_Y}{Z}}{-\beta x + \alpha y + 1 + \frac{T_Z}{Z}}$$

Instantaneous Velocity Model

Optical Flow

3-D Rigid Motion

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} 1 & -\alpha & \beta \\ \alpha & 1 & -\gamma \\ -\beta & \gamma & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix}$$

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \left(\begin{bmatrix} 0 & -\alpha & \beta \\ \alpha & 0 & -\gamma \\ -\beta & \gamma & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix}$$

$$\begin{bmatrix} X' - X \\ Y' - Y \\ Z' - Z \end{bmatrix} = \begin{bmatrix} 0 & -\alpha & \beta \\ \alpha & 0 & -\gamma \\ -\beta & \gamma & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix}$$

$$\begin{bmatrix} \dot{X} \\ \dot{Y} \\ \dot{Z} \end{bmatrix} = \begin{bmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix}$$

$$\dot{X} = \Omega_2 Z - \Omega_3 Y + V_1$$
$$\dot{Y} = \Omega_3 X - \Omega_1 Z + V_2$$
$$\dot{Z} = \Omega_1 Y - \Omega_2 X + V_3$$

3-D Rigid Motion

$$\dot{X} = \Omega_2 Z - \Omega_3 Y + V_1$$
$$\dot{Y} = \Omega_3 X - \Omega_1 Z + V_2$$
$$\dot{Z} = \Omega_1 Y - \Omega_2 X + V_3$$

$$\mathbf{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}, \quad \Omega = \begin{bmatrix} \Omega_1 \\ \Omega_2 \\ \Omega_3 \end{bmatrix}$$

$$\dot{\mathbf{X}} = \Omega \times \mathbf{X} + \mathbf{V}$$

Cross Product

Orthographic Projection

$$\dot{X} = \square_2 Z \square\square_3 Y + V_1$$
$$\dot{Y} = \square_3 X \square\square_1 Z + V_2$$
$$\dot{Z} = \square_1 Y \square\square_2 X + V_3$$

$$y = Y$$
$$x = X$$

$$u = \dot{x} = \square_2 Z \square\square_3 y + V_1$$
$$v = \dot{y} = \square_3 x \square\square_1 Z + V_2$$

(u,v) is optical flow

Perspective Projection (arbitrary flow)

$$x = \frac{fX}{Z}$$
$$y = \frac{fY}{Z}$$

$$u = \dot{x} = \frac{fZ\dot{X} \;\square fX\dot{Z}}{Z^2} = f\frac{\dot{X}}{Z} \square x \frac{\dot{Z}}{Z}$$
$$v = \dot{y} = \frac{fZ\dot{Y} \square fY\dot{Z}}{Z^2} = f\frac{\dot{Y}}{Z} \square y \frac{\dot{Z}}{Z}$$

$$\dot{X} = \square_2 Z \square\square_3 Y + V_1$$
$$\dot{Y} = \square_3 X \square\square_1 Z + V_2$$
$$\dot{Z} = \square_1 Y \square\square_2 X + V_3$$

$$u = \dot{x} = f\frac{\dot{X}}{Z} \square x \frac{\dot{Z}}{Z} = f\frac{\square_2 Z \square\square_3 Y + V_1}{Z} \square x \frac{\square_1 Y \square\square_2 X + V_3}{Z}$$
$$v = \dot{y} = f\frac{\dot{Y}}{Z} \square y \frac{\dot{Z}}{Z} = f\frac{\square_3 X \square\square_1 Z + V_2}{Z} \square y \frac{\square_1 Y \square\square_2 X + V_3}{Z}$$

$$u = f(\frac{V_1}{Z} + \square_2) \square \frac{V_3}{Z} x \square\square_3 y \square \frac{\square_1}{f} xy + \frac{\square_2}{f} x^2$$
$$v = f(\frac{V_2}{Z} \square\square_1) + \square_3 x \square \frac{V_3}{Z} y + \frac{\square_2}{f} xy \square \frac{\square_1}{f} y^2$$

Perspective Projection (optical flow)

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

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

$$u = \frac{fV_1 - V_3 x}{Z} + f\Omega_2 - \Omega_3 y - \frac{\Omega_1}{f}xy + \frac{\Omega_2}{f}x^2$$

$$v = \frac{fV_2 - V_3 y}{Z} - f\Omega_1 + \Omega_3 x + \frac{\Omega_2}{f}xy - \frac{\Omega_1}{f}y^2$$

Pure Translation (FOE)

$$u^{(T)} = \frac{fV_1 - V_3 x}{Z}$$

$$v^{(T)} = \frac{fV_2 - V_3 y}{Z}$$

$$x_0 = f\frac{V_1}{V_3}, y_0 = f\frac{V_2}{V_3}$$

$$u^{(T)} = (x_0 - x)\frac{V_3}{Z}$$

$$v^{(T)} = (y_0 - y)\frac{V_3}{Z}$$

Pure Translation (FOE)

$$u^{(T)} = \frac{fV_1 - V_3 x}{Z}$$

$$v^{(T)} = \frac{fV_2 - V_3 y}{Z}$$

$$u^{(T)} = \frac{fV_1}{Z}$$

If $V_3=0$

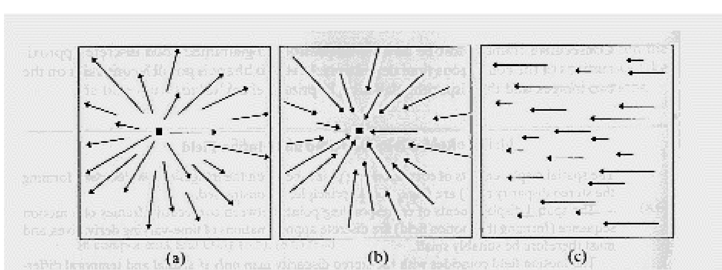
$$v^{(T)} = \frac{fV_2}{Z}$$

Pure Translation (FOE)

- If V3 is not zero, the flow field is radial, and all vectors point towards (or away from) a single point. If $V_3=0$, the flow field is parallel.
- The length of flow vectors is inversely proportional to the depth, if V_3 is not zero, then it is also proportional to the distance between p and p_0.

Pure Translation (FOE)

- p_0 is the vanishing point of the direction of translation.
- p_0 is the intersection of the ray parallel to the translation vector with the image plane.



Structure From Motion

ORTHOGRAPHIC PROJECTION

Orthographic Projection (displacement)

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = R\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + T = \begin{bmatrix} 1 & -\gamma & \beta \\ \gamma & 1 & -\alpha \\ -\beta & \alpha & 1 \end{bmatrix}\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix}$$

$$x' = x - \gamma y + \beta Z + T_X$$

$$y' = \gamma x + y - \alpha Z + T_Y$$

Simple Method

$$x' = x - \alpha y + \beta Z + T_X$$
$$y' = \alpha x + y - \gamma Z + T_Y$$

- **Two Steps Method**

  -Assume depth is known, compute motion

$$\begin{bmatrix} x' - x \\ y' - y \end{bmatrix} = \begin{bmatrix} -y & Z & 0 & 1 & 0 \\ x & 0 & -Z & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ T_x \\ T_y \end{bmatrix}$$

Simple Method

-Assume motion is known, refine depth

$$\begin{bmatrix} \beta \\ -\gamma \end{bmatrix} [Z] = \begin{bmatrix} x' - x + \alpha y - T_x \\ y' - y - \alpha x - T_y \end{bmatrix}$$

Structure from Motion:Perspective Projection

Heeger & Jepson sfm method

Three Step Algorithm

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

Heeger & Jepson sfm method

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

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

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

Heeger & Jepson sfm method

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

$$\Downarrow$$

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

One point (x,y)

Heeger & Jepson sfm method

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

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

n points

Heeger & Jepson sfm method

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

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

Heeger & Jepson sfm method

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

$$\Downarrow$$

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

$\mathbf{C}^{\square}(V)$ Orthogonal complemen to $\mathbf{C}(\mathbf{V})$

Find translation by search.

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

QR decomposition

Orthonormal & Upper triangular

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

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

$$\Downarrow$$ minimize

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

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

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

Null space

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

Translation

Unit vector translation can be represented
spherical coordinates:

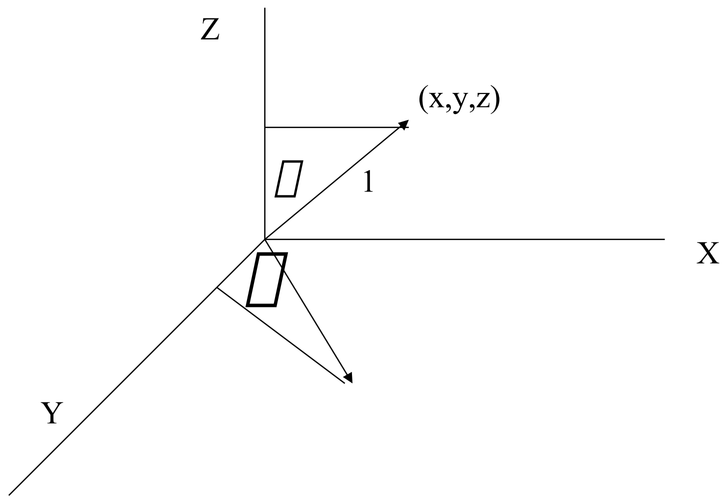
$$\mathbf{V} = (\sin\square\cos\square, \sin\square\sin\square, \cos\square)$$

$0 \square \square \square 90$ Slant

$0 \square \square \square 360$ Tilt

Only half of sphere can be considerec

Spherical Coordinates



Rotation

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

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

$d^T(x,y,V)$ is perpendicular to $\mathbf{A}(x,y)\mathbf{V}$

Rotation

Find rotation by least squares fit

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

$$\vdots$$

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

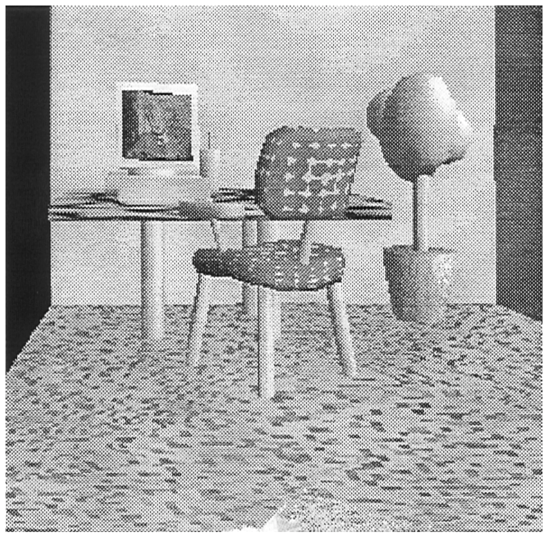
Depth

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

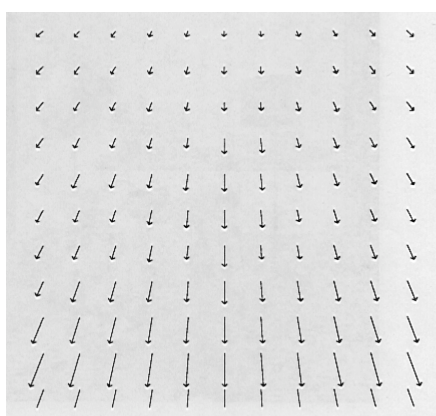
$$u = \square(\frac{V_1}{Z} + \square_2) + \frac{V_3}{Z}x + \square_3 y + \square_1 xy \square\square_2 x^2$$

$$v = \square(\frac{V_2}{Z} \square\square_1) \square\square_3 x + \frac{V_3}{Z}y \square\square_2 xy + \square_1 y^2$$

Synthetic Image

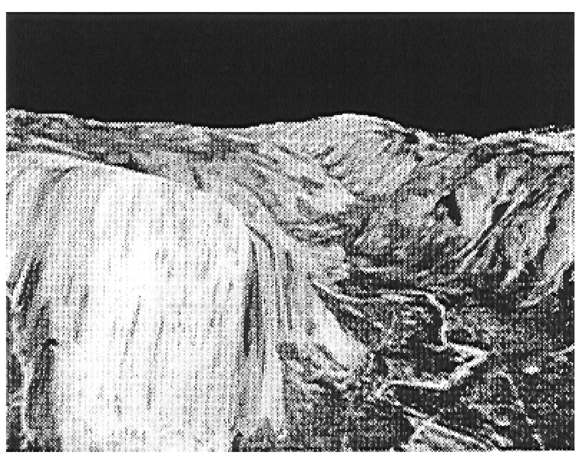


Optical Flow

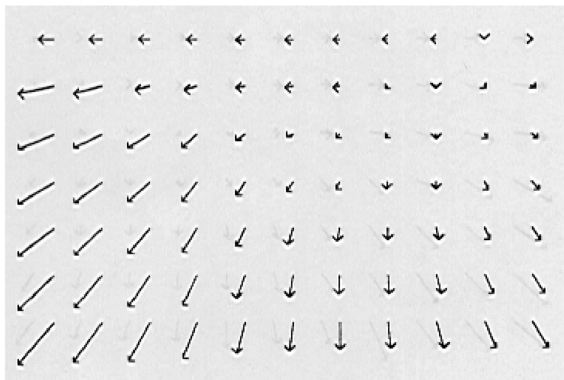


Computed Depth Map

Synthetic Image

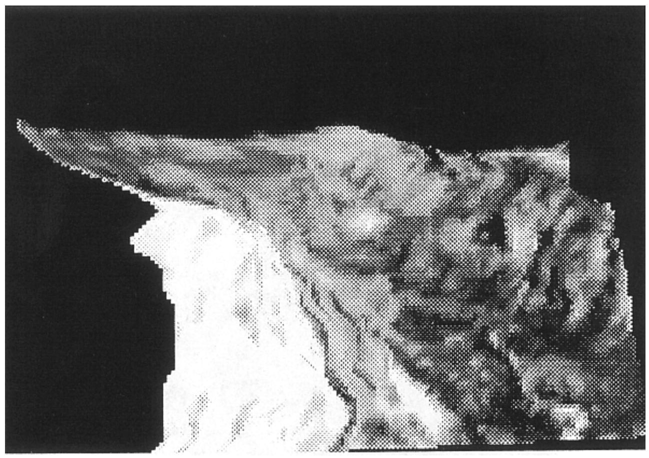


Optical Flow



Translation Search Space

Novel View Generated from Reconstructed Depth



Another Novel View Generated from Reconstructed Depth