Motion and Flow II

Structure from Motion

\[ \mathbf{u} = \mathbf{u}_t + \mathbf{u}_{rot} \]

\[ \mathbf{u}_t = \frac{1}{Z} \left( \mathbf{z} \times (\mathbf{t} \times \mathbf{r}) \right) \]

\[ \mathbf{u}_{rot} = \frac{1}{F} \left( \mathbf{z} \times (\mathbf{r} \times (\mathbf{r}) \right) \]

Passive Navigation and Structure

The system move with a rigid motion with translational velocity \( \mathbf{t} = (U, V, W)^T \) and rotational velocity \( \mathbf{y} = (x, y, z)^T \).

Scene points \( \mathbf{X} = (X, Y, Z)^T \) projected to image points \( \mathbf{r} = (x, y, f) \) and the 3D velocity \( \mathbf{v} = (U, V, W) \) of a scene point is observed in the image as velocity \( \mathbf{v} = (u, v, 0) \).
Image Flow due to Rigid Motion

The velocity of a point with respect to the XYZ coordinate system is

\[ \mathbf{v} = -U \mathbf{w} + \mathbf{a} \]

where \( \mathbf{a} \) is the acceleration.

\[ \mathbf{w} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \]

Scaling ambiguity: We can compute the translation only up to a scale factor \((K, K)\) gives the same flow as \((t, Z)\).

Translational flow field:

\[ \mathbf{v} = \begin{pmatrix} U \\ V \\ W \end{pmatrix} \]

Rotational flow field:

\[ \mathbf{v} = \begin{pmatrix} U \\ V \\ W \end{pmatrix} \]

Classical Structure from Motion

- Established approach is the epipolar minimization. The “derotated flow” should be parallel to the translational flow.

Uniqueness

Let there be two translations \( t_1, t_2 \) and two surfaces \( Z_1, Z_2 \),

\[ t_1 = (U_1, V_1, W_1) \]

\[ t_2 = (U_2, V_2, W_2) \]

must hold for all \( x \) and \( y \)

\[ U_1 V_1 = U_2 V_1 \]

\[ V_1 W_1 = V_2 W_1 \]

A translational flow field determines the motions of the camera uniquely up to a scaling factor.

Step 2: Differentiate with respect to \( U, V, W \), set expression to zero.

Let \( K = \frac{ab - va + vb}{(a^2 + b^2)} \)

\[ \int \left[ (a^2 + b^2)dx + (a^2 + b^2)dy \right] = 0 \]

3 linearly dependent equations \((U, U \| W, W = W)\)
The Rotational Case

\[ \int (u - u_{\text{ref}})^2 + (v - v_{\text{ref}})^2 \rightarrow \min \]
\[ u = \alpha x + \beta (x^2 + 1) - \gamma y = 0 \]
\[ v = \alpha y + \beta (y^2 + 1) - \gamma x = 0 \]
\[ \begin{pmatrix}
  xy \\
  (x^2 + 1) \\
  -xy \\
  (y^2 + 1)
\end{pmatrix}
= \begin{pmatrix}
  u \\
  v
\end{pmatrix} \]

In matrix form:

\[ A\mathbf{x} = \mathbf{u} \]
\[ \mathbf{a} = (A^T A)^{-1} A^T \mathbf{u} \]

The General Case

Minimization of epipolar distance

\[ \int (u - u_{\text{ref}})^2 + (v - v_{\text{ref}})^2 \rightarrow \min \]

or, in vector notation

\[ \int (\| (u - u_{\text{ref}}) \|)^2 dx dy \rightarrow \min \]

Motion Estimation Techniques

- Linearization (Tsai Huang 1984, Longuet-Higgins 1981) in the discrete case

Optical flow difficulties

- The aperture problem
- Depth discontinuities

Translational Normal Flow

- In the case of translation each normal flow vector constrains the location of the FOE to a half-plane.
- Intersection of half-planes provides FOE.

Egoestimation from normal flow

- patterns defined on the sign of normal flow along particular orientation fields
- positive depth constraint
- 2 classes of orientation fields: copoint vectors and coaxis vectors
Copoint vector fields

Copoint vectors: \( v_{\{r\}} \) perpendicular to translational flow defined by \( t_r \)
\[ v_{\{r\}} = \hat{z} \times u_{\{r\}} \]

The components of flow along \( v_{\{r\}} \) amount to
\[ r \times \frac{1}{2} (\hat{z} \times t_r) r + (u \times \hat{z}) (t_r, r) \]

Thus the translational component is separated by a line into positive and negative values

The rotational component is separated by a second-order curve \( (u \times \hat{z}) (t_r, r) = 0 \) into positive and negative values

Pattern with positive areas, negative areas, and some undefined areas

Coaxis vector fields

Coaxis vectors: \( v_{\{w\}} \) perpendicular to rotation:
\[ v_{\{w\}} = \hat{z} \times u_{\{w\}} = \hat{z} \times (r \times (u \times r)) \]

The components of flow along \( v_{\{w\}} \) amount to
\[ r \times \frac{1}{2} (u \times \hat{z}) r + \frac{1}{2} (r \times r) (u \times r) \]

Thus the translational component is separated by a second-order curve \( (r \times r) (u \times r) = 0 \)
and the rotational component is separated by a line \( (u \times \hat{z}) r = 0 \)

Intersection of patterns provides the FOE.

Three coaxis vector fields

(a) (b) (c)
Depth variability constraint
- Errors in motion estimates lead to distortion of the scene estimates.
- The distortion is such that the correct motion gives the “smoothest” (least varying) scene structure.

Scene depth can be estimated from normal flow measurements:

\[ u_n = u \cdot n = \frac{1}{Z} u_n \cdot n + u_{rot} \cdot n \]

\[ \frac{1}{Z} = \frac{u_n - u_{rot}(i) \cdot n}{u_n(t) \cdot n} \]

Visual Space Distortion
- Wrong 3D motion gives rise to a rugged (unsmooth) depth function (surface).
- The correct 3D motion leads to the “smoothest” estimated depth.

\[ \hat{Z} = Z : D, \quad D = \frac{u_n(i) \cdot n}{|u_n(t) - u_{rot}(d) \cdot n|} \]

The error function
- A normal flow measurement:

\[ u_n = \frac{1}{Z} u_n \cdot n + u_{rot} \cdot n \]

The error function to be minimized:

\[ \Theta = \sum_{i} W(u_{i} - u_n)^2 \]

Global parameters: \[ i, j \]

Local parameter: \[ \hat{Z} \]

Error function evaluation
- Given a translation candidate \[ \hat{i} \], each local depth can be computed as a linear function of the rotation \[ \hat{j} \].
- We obtain a second order function of the rotation; its minimization provides both the rotation and the value of the error function.
Handling depth discontinuities

• Given a candidate motion, the scene depth can be estimated and further processed to find depth discontinuities.
• Split a region if it corresponds to two depth values separated in space.

The algorithm

• Compute spatio-temporal image derivatives and normal flow.
• Find the direction of translation that minimizes the depth-variability criterion.
  – Hierarchical search of the 2D space.
  – Iterative minimization.
  – Utilize continuity of the solution in time; usually the motion changes slowly over time.

Sources:

• Horn (1986)
• http://www.cfar.umd.edu/ftp/TRs/CVL-Reports-1993/TR3064-Fermuller.ps.gz
• http://www.cfar.umd.edu/ftp/TRs/CVL-Reports-1999/TR4000-brodsky.ps.gz (depth variability constraint)