Recovering 3D observer motion & layout

FOE: focus of expansion

https://www.youtube.com/watch?v=m80b6esRpaQ

Application: Automated driving systems

DARPA Grand Challenge

https://www.wired.com/story/darpa-grand-urban-challenge-self-driving-car/

Observer motion problem

From image velocity field, compute:
- observer translation
  \((T_x, T_y, T_z)\)
- observer rotation
  \((R_x, R_y, R_z)\)
- depth at each location
  \(Z(x, y)\)

Human perception of heading

Warren & colleagues
Human accuracy:
1° - 2° visual arc

Observer heading to the left or right of target on horizon?

bird’s eye view
~ 2°
Observer just translates toward FOE

Directions of velocity vectors intersect at FOE

But... simple strategy doesn't work if observer also rotates

Observer Translation + Rotation

display simulates observer translation

observer rotates their eyes

display simulates translation + rotation

Still recover heading with high accuracy!

Observer motion problem, revisited

From image motion, compute:
- Observer translation
  \( (T_x, T_y, T_z) \)
- Observer rotation
  \( (R_x, R_y, R_z) \)
- Depth at each location
  \( Z(x,y) \)

Observer undergoes both translation + rotation

Equations of observer motion

<table>
<thead>
<tr>
<th>Translation ((T_x, T_y, T_z))</th>
<th>Rotation ((R_x, R_y, R_z))</th>
<th>Depth (Z(x,y))</th>
</tr>
</thead>
<tbody>
<tr>
<td>( V_x = \frac{-T_x + xT_z}{Z} + \frac{R_xxy - R_y(x^2 + 1) + R_yy}{Z} )</td>
<td>( V_y = \frac{-T_y + yT_z}{Z} + \frac{R_x(y^2 + 1) - R_yxy - R_yy}{Z} )</td>
<td></td>
</tr>
</tbody>
</table>

Translational Component
Rotational Component
Translational component of motion

\[ V_x(x, y) = \frac{-T_x + xT_z}{Z(x, y)} \]
\[ V_y(x, y) = \frac{-T_y + yT_z}{Z(x, y)} \]

- \( V_x, V_y, Z \) depend on position \((x, y)\)
- Note \( Z(x, y) \) in the denominator

- \( V_x, V_y \) depend on ratios: \( \frac{T_x}{Z}, \frac{T_y}{Z}, \frac{T_z}{Z} \)
  (e.g. doubling both observer speed & depth gives the same velocity field)
- Where is the FOE? \( x = \frac{T_x}{T_z}, y = \frac{T_y}{T_z} \)

Translational component of velocity

\[ V_x = (-T_x + xT_z)/Z \]
\[ V_y = (-T_y + yT_z)/Z \]

Where is the FOE?
\[ x = \quad y = \]

Example 1: \( T_x = T_y = 0 \) \( T_z = 1 \) \( Z = 10 \) everywhere
\[ V_x = \quad V_y = \]

Sketch the velocity field

Example 2: \( T_x = T_y = 2 \) \( T_z = 1 \) \( Z = 10 \) everywhere
\[ V_x = \quad V_y = \]

Longuet-Higgins & Prazdny

- Along a depth discontinuity, velocity differences depend only on observer translation
- Velocity differences point to the focus of expansion

Rieger & Lawton’s algorithm

(1) At each image location, compute distribution of velocity differences within neighborhood

Appearance of sample distributions:

(2) Find points with strongly oriented distribution, compute dominant direction

(3) Compute focus of expansion from intersection of dominant directions
Recovering the observer’s rotation

True FOE

Computed FOE

Velocity component due to observer’s translation
Velocity component due to observer’s rotation
Final velocity at each location

Velocity perpendicular to field lines must be due to observer’s rotation!

Find \((R_x, R_y, R_z)\) that best explains the motion perpendicular to the field lines

Finally, recovering 3D layout

Given \((R_x, R_y, R_z)\), compute image motions due to rotation...

... then subtract motions due to rotation, to obtain the image motions due to the observer’s translation alone.

Then, how can we compute the relative depth of surfaces?

What are we assuming about objects in the scene?
When is this assumption violated?