Motion

Wed, April 20
Kristen Grauman
UT-Austin

Many slides adapted from S. Seitz, R. Szeliski, M. Pollefeys, and S. Lazebnik

Finally: Motion and tracking

Tracking objects, video analysis, low level motion

Video

• A video is a sequence of frames captured over time
• Now our image data is a function of space (x, y) and time (t)

Uses of motion

• Estimating 3D structure
• Segmenting objects based on motion cues
• Learning dynamical models
• Recognizing events and activities
• Improving video quality (motion stabilization)

Motion field

• The motion field is the projection of the 3D scene motion into the image

Motion parallax

http://psych.hanover.edu/KRANTZ/MotionParallax/MotionParallax.html
Motion field + camera motion

Length of flow vectors inversely proportional to depth $Z$ of 3d point.

Motion estimation techniques

- Direct methods
  - Directly recover image motion at each pixel from spatio-temporal image brightness variations
  - Dense motion fields, but sensitive to appearance variations
  - Suitable for video and when image motion is small

- Feature-based methods
  - Extract visual features (corners, textured areas) and track them over multiple frames
  - Sparse motion fields, but more robust tracking
  - Suitable when image motion is large (10s of pixels)

Optical flow

- Definition: optical flow is the apparent motion of brightness patterns in the image
- Ideally, optical flow would be the same as the motion field
- Have to be careful: apparent motion can be caused by lighting changes without any actual motion

Apparent motion $\neq$ motion field

Problem definition: optical flow

How to estimate pixel motion from image $H$ to image $I$?

- Solve pixel correspondence problem
  - given a pixel in $H$, look for nearby pixels of the same color in $I$

Key assumptions

- color constancy: a point in $H$ looks the same in $I$
- For grayscale images, this is brightness constancy
- small motion: points do not move very far
This is called the optical flow problem
Brightness constancy

Figure 6.5. Data conservation example. The highlighted region in the right image looks roughly the same as the region in the left image, despite the fact that it has moved.

Optical flow constraints (grayscale images)

\[
\begin{align*}
(x, y) & \quad \text{displacement} = (u, v) \\
H(x, y) & \quad I(x, y)
\end{align*}
\]

Let’s look at these constraints more closely

- brightness constancy: Q: what’s the equation?
  \[H(x, y) = I(x + u, y + v)\]
- small motion:
  \[I(x+u, y+v) = I(x, y) + \frac{\partial I}{\partial x} u + \frac{\partial I}{\partial y} v + \text{higher order terms}\]
  \[\approx I(x, y) + \frac{\partial I}{\partial x} u + \frac{\partial I}{\partial y} v\]

Optical flow equation

Combining these two equations

\[0 = I(x + u, y + v) - H(x, y)\]

shorthand: \[I_x = \frac{\partial I}{\partial x}\]

\[\approx I(x, y) + I_x u + I_y v - H(x, y)\]

\[\approx (I(x, y) - H(x, y)) + I_x u + I_y v\]

\[\approx I_x + I_y v\]

\[\approx I_x + \nabla I \cdot [u, v]\]

The aperture problem

Perceived motion

Optical flow equation

\[0 = I_t + \nabla I \cdot [u, v]\]

Q: how many unknowns and equations per pixel?

Intuitively, what does this ambiguity mean?

The aperture problem

Actual motion
The barber pole illusion

http://en.wikipedia.org/wiki/Barberpole_illusion

The barber pole illusion

http://www.sandlotscience.com/Ambiguous/Barberpole_Illusion.htm

Solving the aperture problem (grayscale image)

- How to get more equations for a pixel?
  - Spatial coherence constraint: pretend the pixel’s neighbors have the same \((u,v)\)

Figure 1.7: Spatial coherence assumption. Neighboring points in the image are assumed to belong to the same surface in the scene.

Figure by Michael Black

Solving the aperture problem

Prob: we have more equations than unknowns

\[
A \cdot d = b
\]

2x2 2x1 2x1

Solution: solve least squares problem

- minimum least squares solution given by solution (in \(d\)) of:

\[
(A^T A)^{-1} A^T b
\]

2x2 2x1 2x1

\[
\begin{bmatrix}
\sum I_x I_x \\
\sum I_x I_y
\end{bmatrix} \begin{bmatrix}
u \\
v
\end{bmatrix} = \begin{bmatrix}
\sum I_x I_x \\
\sum I_x I_y
\end{bmatrix} \begin{bmatrix}
u \\
v
\end{bmatrix}
\]

- The summations are over all pixels in the \(K \times K\) window
- This technique was first proposed by Lucas & Kanade (1981)

Conditions for solvability

\[
\begin{bmatrix}
\sum I_x I_x & \sum I_x I_y \\
\sum I_x I_y & \sum I_y I_y
\end{bmatrix} \begin{bmatrix}
u \\
v
\end{bmatrix} = \begin{bmatrix}
\sum I_x I_x \\
\sum I_x I_y
\end{bmatrix} \begin{bmatrix}
u \\
v
\end{bmatrix}
\]

A\(^T\)A

A\(^T\)b

When is this solvable?

- \(\mathbf{A}^\mathbf{A}\) should be invertible
- \(\mathbf{A}^\mathbf{A}\) should not be too small
  - eigenvalues \(\lambda_1\) and \(\lambda_2\) of \(\mathbf{A}^\mathbf{A}\) should not be too small
- \(\mathbf{A}^\mathbf{A}\) should be well-conditioned
  - \(\lambda_1,\lambda_2\) should not be too large (\(\lambda_1 = \) larger eigenvalue)
Edge
- gradients very large or very small
- large $\lambda_1$, small $\lambda_2$

Low-texture region
- gradients have small magnitude
- small $\lambda_1$, small $\lambda_2$

High-texture region
- gradients are different, large magnitudes
- large $\lambda_1$, large $\lambda_2$

Example use of optical flow: facial animation
http://www.fxguide.com/article333.html

Example use of optical flow: Motion Paint
Use optical flow to track brush strokes, in order to animate them to follow underlying scene motion.
http://www.fxguide.com/article333.html

Fun with flow
- http://www.youtube.com/watch?v=TbJrc6QCeU0&feature=related
- http://www.youtube.com/watch?v=pckFacsIWg4
Motion vs. Stereo: Similarities

- Both involve solving
  - Correspondence: disparities, motion vectors
  - Reconstruction

Motion vs. Stereo: Differences

- **Motion:**
  - Uses velocity: consecutive frames must be close to get good approximate time derivative
  - 3d movement between camera and scene not necessarily single 3d rigid transformation
- **Whereas with stereo:**
  - Could have any disparity value
  - View pair separated by a single 3d transformation

Coming up

Background subtraction, activity recognition, tracking