Epipolar geometry & stereo vision

Tuesday, Oct 20
Kristen Grauman
UT-Austin

Recap: Features and filters
Transforming and describing images; textures, colors, edges
Recap: Grouping & fitting

Clustering, segmentation, fitting; what parts belong together?

Multiple views

Multi-view geometry, matching, invariant features, stereo vision
Why multiple views?

- Structure and depth are inherently ambiguous from single views.

Images from Lana Lazebnik
• What cues help us to perceive 3d shape and depth?

Shading

[Figure from Prados & Faugeras 2006]
Focus/Defocus

[Figure from H. Jin and P. Favaro, 2002]

Texture

Perspective effects

Image credit: S. Seitz

Motion

Figures from L. Zhang
http://www.brainconnection.com/teasers/?main=illusion/motion-shape
Estimating scene shape

• “Shape from X”: Shading, Texture, Focus, Motion…

• **Stereo:**
  – shape from “motion” between two views
  – infer 3d shape of scene from two (multiple) images from different viewpoints

Main idea:

Outline

• Human stereopsis

• Stereograms

• Epipolar geometry and the epipolar constraint
  – Case example with parallel optical axes
  – General case with calibrated cameras
Human stereopsis: disparity

From Bruce and Green, Visual Perception, Physiology, Psychology and Ecology

Human eyes fixate on point in space – rotate so that corresponding images form in centers of fovea.

Disparity occurs when eyes fixate on one object; others appear at different visual angles.
Human stereopsis: disparity

Disparity: \( d = r - l = D - F \).

Random dot stereograms

- Julesz 1960: Do we identify local brightness patterns before fusion (monocular process) or after (binocular)?

- To test: pair of synthetic images obtained by randomly spraying black dots on white objects.
Random dot stereograms

Forsyth & Ponce
Random dot stereograms

• When viewed monocularly, they appear random; when viewed stereoscopically, see 3D structure.

• Conclusion: human binocular fusion not directly associated with the physical retinas; must involve the central nervous system

• Imaginary “cyclopean retina” that combines the left and right image stimuli as a single unit

Stereo photography and stereo viewers

Take two pictures of the same subject from two slightly different viewpoints and display so that each eye sees only one of the images.

Invented by Sir Charles Wheatstone, 1838

Image courtesy of fisher-price.com
Public Library, Stereoscopic Looking Room, Chicago, by Phillips, 1923

http://www.johnsonshawmuseum.org
Autostereograms

Exploit disparity as depth cue using single image.
(Single image random dot stereogram, Single image stereogram)

Images from magiceye.com
Estimating depth with stereo

- **Stereo**: shape from “motion” between two views
- We’ll need to consider:
  - Info on camera pose (“calibration”)
  - Image point correspondences

![Diagram of stereo vision](image)

Camera parameters

- **Extrinsic params**: rotation matrix and translation vector
- **Intrinsic params**: focal length, pixel sizes (mm), image center point, radial distortion parameters

*We’ll assume for now that these parameters are given and fixed.*
Outline

• Human stereopsis
• Stereograms
• Epipolar geometry and the epipolar constraint
  – Case example with parallel optical axes
  – General case with calibrated cameras

Geometry for a simple stereo system

• First, assuming parallel optical axes, known camera parameters (i.e., calibrated cameras):
Geometry for a simple stereo system

- Assume parallel optical axes, known camera parameters (i.e., calibrated cameras). We can triangulate via:

Similar triangles \((p_l, P, p_r)\) and \((O_l, P, O_r)\):

\[
\frac{T + x_l - x_r}{Z - f} = \frac{T}{Z}
\]

\[
Z = f \frac{T}{x_r - x_l}
\]
Depth from disparity

\[ (x', y') = (x + D(x, y), y) \]

Outline

- Human stereopsis
- Stereograms
- Epipolar geometry and the epipolar constraint
  - Case example with parallel optical axes
  - General case with calibrated cameras
General case, with calibrated cameras

- The two cameras need not have parallel optical axes.

Vs.

Stereo correspondence constraints

- Given p in left image, where can corresponding point p’ be?
Stereo correspondence constraints

Epipolar constraint

Geometry of two views constrains where the corresponding pixel for some image point in the first view must occur in the second view:

- It must be on the line carved out by a plane connecting the world point and optical centers.

Why is this useful?
Epipolar constraint

This is useful because it reduces the correspondence problem to a 1D search along an epipolar line.

Epipolar geometry

http://www.ai.sri.com/~luong/research/Meta3DViewer/EpipolarGeo.html
Epipolar geometry: terms

- **Baseline**: line joining the camera centers
- **Epipole**: point of intersection of baseline with the image plane
- **Epipolar plane**: plane containing baseline and world point
- **Epipolar line**: intersection of epipolar plane with the image plane

- All epipolar lines intersect at the epipole
- An epipolar plane intersects the left and right image planes in epipolar lines

Example
Example: converging cameras

Figure from Hartley & Zisserman

Example: parallel cameras

Where are the epipoles?

Figure from Hartley & Zisserman
• So far, we have the explanation in terms of geometry.
• Now, how to express the epipolar constraints algebraically?

Stereo geometry, with calibrated cameras

If the stereo rig is calibrated, we know:
how to **rotate** and **translate** camera reference frame 1 to get to camera reference frame 2.
Rotation: 3 x 3 matrix \( R \); translation: 3 vector \( T \).
Stereo geometry, with calibrated cameras

If the stereo rig is calibrated, we know:
how to rotate and translate camera reference frame 1 to get to camera reference frame 2. \( X'_c = RX_c + T \)

An aside: cross product

\( \vec{a} \times \vec{b} = \vec{c} \)
\( \vec{a} \cdot \vec{c} = 0 \)
\( \vec{b} \cdot \vec{c} = 0 \)

Vector cross product takes two vectors and returns a third vector that’s perpendicular to both inputs.

So here, \( \vec{c} \) is perpendicular to both \( \vec{a} \) and \( \vec{b} \), which means the dot product = 0.
From geometry to algebra

\[
\mathbf{X}' = \mathbf{RX} + \mathbf{T}
\]

\[
\mathbf{T} \times \mathbf{X}' = \mathbf{R} \times \mathbf{X'}
\]

Normal to the plane

Another aside:
Matrix form of cross product

\[
\mathbf{a} \times \mathbf{b} = \begin{bmatrix}
0 & -a_3 & a_2 \\
a_3 & 0 & -a_1 \\
-a_2 & a_1 & 0
\end{bmatrix} \begin{bmatrix}
b_1 \\
b_2 \\
b_3
\end{bmatrix} = \begin{bmatrix}
\mathbf{a} \cdot \mathbf{c} = 0 \\
\mathbf{b} \cdot \mathbf{c} = 0
\end{bmatrix}
\]

Can be expressed as a matrix multiplication.

\[
\mathbf{a} \times \mathbf{b} = [a_x] \mathbf{b}
\]
From geometry to algebra

\[ \begin{align*}
X' &= RX + T \\
T \times X' &= T \times RX + T \times T \\
\text{Normal to the plane} &\Rightarrow T \times RX
\end{align*} \]

Essential matrix

\[ \begin{align*}
X' \cdot (T \times RX) &= 0 \\
X' \cdot (T_x RX) &= 0
\end{align*} \]

Let \( E = T_x R \)

\[ X'^T E X = 0 \]

\( E \) is called the essential matrix, and it relates corresponding image points between both cameras, given the rotation and translation.

If we observe a point in one image, its position in other image is constrained to lie on line defined by above.

Note: these points are in camera coordinate systems.
Essential matrix example: parallel cameras

For the parallel cameras, image of any point must lie on same horizontal line in each image plane.

What about when cameras’ optical axes are not parallel?
Stereo image rectification

In practice, it is convenient if image scanlines (rows) are the epipolar lines.

reproject image planes onto a common plane parallel to the line between optical centers

pixel motion is horizontal after this transformation two homographies (3x3 transforms), one for each input image reprojection

Adapted from Li Zhang

Stereo image rectification: example

Source: Alyosha Efros
Summary

• Depth from stereo: main idea is to triangulate from corresponding image points.

• Epipolar geometry defined by two cameras
  – We’ve assumed known extrinsic parameters relating their poses

• Epipolar constraint limits where points from one view will be imaged in the other
  – Makes search for correspondences quicker

• Terms: epipole, epipolar plane / lines, disparity, rectification, intrinsic/extrinsic parameters, essential matrix, baseline

Coming up

• Thursday:
  – Computing correspondences
  – Non-geometric stereo constraints
  – Weak calibration