Stereo: Epipolar geometry

Wednesday March 23
Kristen Grauman
UT-Austin

Announcements

• Reminder: Pset 3 due next Wed, March 30

Last time

• Image formation affected by geometry, photometry, and optics.
• Projection equations express how world points mapped to 2d image.
• Parameters (focal length, aperture, lens diameter,…) affect image obtained.

Review

• How do the perspective projection equations explain this effect?

Miniature faking

In close-up photo, the depth of field is limited.

Miniature faking


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

Shading

[Figure from Prados & Faugeras 2006]

Why multiple views?

- Structure and depth are inherently ambiguous from single views.

Why multiple views?

- What cues help us to perceive 3d shape and depth?

Images from Lana Lazebnik
Focus/defocus

Images from same point of view, different camera parameters

3d shape / depth estimates

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

Texture

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

Perspective effects

Image credit: S. Seitz

Motion

Figure from L. Zhang
http://www.brainconnection.com/sean/el/motion-shape.html

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 eye

Rough analogy with human visual system:

- Pupil/Iris – control amount of light passing through lens
- Retina - contains sensor cells, where image is formed
- Fovea – highest concentration of cones

Fig from Shapiro and Stockman

Human stereopsis: disparity

Disparity occurs when eyes fixate on one object; others appear at different visual angles

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

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

From 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 from fisher-price.com

http://www.johnsonshawmuseum.org

Public Library, Stereoscopic Looking Room, Chicago, by Phillips, 1923
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

![Camera parameters diagram](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.

Images from magiceye.com

Stereo vision

Two cameras, simultaneous views
Single moving camera and static scene
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):

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

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

Depth from disparity

So if we could find the corresponding points in two images, we could estimate relative depth...

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.

Stereo correspondence constraints

• Given \( p \) in left image, where can corresponding point \( p' \) be?

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.

Epipolar geometry

• Baseline: line joining the camera centers
• Epipole: point of intersection of baseline with 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

Why is the epipolar constraint useful?
Epipolar constraint

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

Example: converging cameras

Example: parallel cameras

What do the epipolar lines look like?

1.

2.

Example: converging cameras

Example: parallel cameras

• So far, we have the explanation in terms of geometry.
• Now, how to express the epipolar constraints algebraically?
Stereo geometry, with calibrated cameras

Main idea

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.

An aside: cross product

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

So here, c is perpendicular to both a and b, which means the dot product = 0.

From geometry to algebra

Another aside:
Matrix form of cross product

Can be expressed as a matrix multiplication.
From geometry to algebra

\[ X' = RX + T \]
\[ X'.(T \times RX) = 0 \]
\[ X'.((T \times RX) = 0 \]
\[ T \times X' = T \times RX + T \times T \]
\[ = T \times RX \]

Essential matrix

\[ E = [T \times RX] \]
\[ X'.(E)' = 0 \]

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 another image is constrained to lie on the line defined by above.

Note: these points are in **camera coordinate systems**.

Essential matrix example: parallel cameras

\[ R = \]
\[ p = [x, y, f] \]
\[ T = \]
\[ p' = [x', y', f] \]
\[ E = [T \times RX] \]

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

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

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

Stereo image rectification: example

\[ (x', y') = (x + \text{Disparity}(x, y), y) \]

What about when cameras’ optical axes are not parallel?

Source: Alyosha Efros
An audio camera & epipolar geometry

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

– Computing correspondences
– Non-geometric stereo constraints
– Weak calibration