Stereo: Correspondence and Calibration
Mon, March 28
Prof. Kristen Grauman
UT-Austin

Today
- Recap: epipolar constraint
- Stereo image rectification
- Stereo solutions
  • Computing correspondences
  • Non-geometric stereo constraints
- Calibration
- Example stereo applications

Last time:
Estimating depth with stereo
• Stereo: shape from “motion” between two views
• We need to consider:
  • Info on camera pose ("calibration")
  • Image point correspondences

Last time:
Epipolar constraint
- Potential matches for \( p \) have to lie on the corresponding epipolar line \( l' \).
- Potential matches for \( p' \) have to lie on the corresponding epipolar line \( l \).
Example: parallel cameras

An audio camera & epipolar geometry

An audio camera & epipolar geometry

Last time:
Essential matrix

Essential matrix example: parallel cameras

What about when cameras' optical axes are not parallel?
Today

- Recap: epipolar constraint
- Stereo image rectification
- Stereo solutions
  - Computing correspondences
  - Non-geometric stereo constraints
- Calibration
- Example stereo applications

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.

Slide credit: Li Zhang

Correspondence problem

Multiple match hypotheses satisfy epipolar constraint, but which is correct?

Beyond the hard constraint of epipolar geometry, there are "soft" constraints to help identify corresponding points:

- Similarity
- Uniqueness
- Ordering
- Disparity gradient

To find matches in the image pair, we will assume:
- Most scene points visible from both views
- Image regions for the matches are similar in appearance

Correspondence problem

Today

- Recap: epipolar constraint
- Stereo image rectification
- Stereo solutions
  - Computing correspondences
  - Non-geometric stereo constraints
- Calibration
- Example stereo applications
Dense correspondence search

For each epipolar line
For each pixel / window in the left image
- compare with every pixel / window on same epipolar line in right image
- pick position with minimum match cost (e.g., SSD, correlation)

Adapted from Li Zhang

Correspondence problem

Parallel camera example: epipolar lines are corresponding image scanlines

Correspondence problem

Intensity profiles
- Clear correspondence between intensities, but also noise and ambiguity

Source: Andrew Zisserman

Correspondence problem

Neighborhoods of corresponding points are similar in intensity patterns.

Source: Andrew Zisserman

Correlation-based window matching

Source: Andrew Zisserman

Textureless regions
Effect of window size

W = 3
W = 20

Want window large enough to have sufficient intensity variation, yet small enough to contain only pixels with about the same disparity.

Source: Li Zhang

Foreshortening effects

Occlusion

Sparse correspondence search

- Restrict search to sparse set of detected features (e.g., corners)
- Rather than pixel values (or lists of pixel values) use feature descriptor and an associated feature distance
- Still narrow search further by epipolar geometry

Tradeoffs between dense and sparse search?

Correspondence problem

- Beyond the hard constraint of epipolar geometry, there are “soft” constraints to help identify corresponding points
  - Similarity
  - Uniqueness
  - Disparity gradient
  - Ordering
**Uniqueness constraint**
- Up to one match in right image for every point in left image

**Disparity gradient constraint**
- Assume piecewise continuous surface, so want disparity estimates to be locally smooth

**Ordering constraint**
- Points on *same surface* (opaque object) will be in same order in both views

**Ordering constraint**
- Won’t always hold, e.g. consider transparent object, or an occluding surface

**Scanline stereo**
- Beyond individual correspondences to estimate disparities:
  - Optimize correspondence assignments jointly
    - Scanline at a time (DP)
    - Full 2D grid (graph cuts)
“Shortest paths” for scan-line stereo

Can be implemented with dynamic programming
Ohta & Kanade ’85, Cox et al. ’96

Coherent stereo on 2D grid

• Scanline stereo generates streaking artifacts

Recap: stereo with calibrated cameras

• Given image pair, \( R, T \)
• Detect some features
• Compute essential matrix \( E \)
• Match features using the epipolar and other constraints
• Triangulate for 3d structure

Error sources

• Low-contrast ; textureless image regions
• Occlusions
• Camera calibration errors
• Violations of brightness constancy (e.g., specular reflections)
• Large motions

Today

– Recap: epipolar constraint
– Stereo image rectification
– Stereo solutions
  • Computing correspondences
  • Non-geometric stereo constraints
– Calibration

– Example stereo applications
Stereo in machine vision systems

Left: The Stanford cart sports a single camera moving in discrete increments along a straight line and providing multiple snapshots of outdoor scenes.

Right: The INRIA mobile robot uses three cameras to map its environment.

Depth for segmentation

Edges in disparity in conjunction with image edges enhances contours found.

Model-based body tracking, stereo input

David Demirdjian, MIT Vision Interface Group
http://people.csail.mit.edu/demirdji/movie/artic-tracker/hum-around.m1v

Virtual viewpoint video

Virtual viewpoint video

Uncalibrated case

• What if we don’t know the camera parameters?
  
  Two possibilities:
  1. Calibrate with a calibration object
  2. Weak calibration

Calibrating a camera

• Compute intrinsic and extrinsic parameters using observed camera data

Main idea

• Place “calibration object” with known geometry in the scene
• Get correspondences
• Solve for mapping from scene to image

Perspective projection

Thus far, in camera’s reference frame only.

Camera parameters

• Extrinsic: location and orientation of camera frame with respect to reference frame
• Intrinsic: how to map pixel coordinates to image plane coordinates

Extrinsic camera parameters

\[ \mathbf{P}_c = \mathbf{R}(\mathbf{P}_w - \mathbf{T}) \]

\[ \mathbf{P}_c = (X, Y, Z)^T \]
Camera parameters

- Extrinsic: location and orientation of camera frame with respect to reference frame
- Intrinsic: how to map pixel coordinates to image plane coordinates

Intrinsic camera parameters

- Ignoring any geometric distortions from optics, we can describe them by:
  \[
  x = -(x_{im} - o_x) s_x \\
  y = -(y_{im} - o_y) s_y
  \]

Camera parameters

- We know that in terms of camera reference frame:
  \[
  x = f \frac{X}{Z} \quad y = f \frac{Y}{Z} \quad P_c = R(p_w - T) \\
  p_c = (X, Y, Z)^T
  \]

- Substituting previous eqns describing intrinsic and extrinsic parameters, can relate pixels coordinates to world points:
  \[
  -(x_{im} - o_x) s_x = f \frac{X - (p_w - T)}{R_i - (p_w - T)} \\
  -(y_{im} - o_y) s_y = f \frac{Y - (p_w - T)}{R_j - (p_w - T)}
  \]

Projection matrix

- This can be rewritten as a matrix product using homogeneous coordinates:
  \[
  \begin{bmatrix}
  wX_{im} \\
  wY_{im} \\
  w
  \end{bmatrix} = M P_w
  \]

where:
  \[
  M_w = \begin{bmatrix}
  -f/s_x & 0 & o_x \\
  0 & -f/s_y & o_y \\
  0 & 0 & 1
  \end{bmatrix}
  \]

  \[
  M_{ext} = \begin{bmatrix}
  r_{11} & r_{12} & r_{13} & -R_i^T \\
  r_{21} & r_{22} & r_{23} & -R_j^T \\
  r_{31} & r_{32} & r_{33} & -R_k^T
  \end{bmatrix}
  \]

Calibrating a camera

- Compute intrinsic and extrinsic parameters using observed camera data

  Main idea
  - Place “calibration object” with known geometry in the scene
  - Get correspondences
  - Solve for mapping from scene to image: estimate \( M = M_{ext} M_{int} \)

When would we calibrate this way?

- Makes sense when geometry of system is not going to change over time

  …when would it change?
Weak calibration

• Want to estimate world geometry without requiring calibrated cameras
  – Archival videos
  – Photos from multiple unrelated users
  – Dynamic camera system

• Main idea:
  – Estimate epipolar geometry from a (redundant) set of point correspondences between two uncalibrated cameras

From before: Projection matrix

• This can be rewritten as a matrix product using homogeneous coordinates:

\[
\begin{bmatrix}
wx_m \\
w'y_m \\
w \\
1
\end{bmatrix} = M_{\text{int}} M_{\text{ext}} \begin{bmatrix}
x_w \\
y_w \\
z_w \\
1
\end{bmatrix}
\]

\[
p_{\text{im}} = M_{\text{int}} M_{\text{ext}} p_w
\]

\[
p_{\text{im}} = M_{\text{ext}} p_c
\]

Uncalibrated case

For a given camera:

\[
p_{\text{im}} = M_{\text{int}} p_c
\]

So, for two cameras (left and right):

\[
p_{\text{c, left}} = M_{\text{int, left}}^{-1} p_{\text{im, left}}
\]

\[
p_{\text{c, right}} = M_{\text{int, right}}^{-1} p_{\text{im, right}}
\]

Internal calibration matrices, one per camera

Computing F from correspondences

Each point correspondence generates one constraint on F

\[
[p_{\text{im, right}}^T F p_{\text{im, left}}] = 0
\]

Collect n of these constraints

\[
\begin{bmatrix}
u' & v' & 1 & f_{11} & f_{12} & f_{13} & f_{21} & f_{22} & f_{23} & f_{31} & f_{32} & f_{33} & u & v
\end{bmatrix} = 0
\]

Solve for f, vector of parameters.
Fundamental matrix

- Relates pixel coordinates in the two views
- More general form than essential matrix: we remove need to know intrinsic parameters
- If we estimate fundamental matrix from correspondences in pixel coordinates, can reconstruct epipolar geometry without intrinsic or extrinsic parameters.

Stereo pipeline with weak calibration

- So, where to start with uncalibrated cameras?
  - Need to find fundamental matrix $F$ and the correspondences (pairs of points $(u',v') \leftrightarrow (u,v)$).

1) Find interest points

2) Match points within proximity to get putative matches

3) Compute epipolar geometry -- robustly with RANSAC
   - Select random sample of putative correspondences
   - Compute $F$ using them
     - determines epipolar constraint
   - Evaluate amount of support
     - inliers within threshold distance of epipolar line
   - Choose $F$ with most support (inliers)

4) Refine

Example from Andrew Zisserman
Summary

- **Rectification**: make epipolar lines align with scanlines

Stereo solutions:
- **Correspondence**: dense, or at interest points
- **Non-geometric stereo constraints** (e.g., similarity, order, smoothness)

Calibration
- **With calibration object** in scene: relate world coordinates to image coordinates
- **Weak calibration**: solve for fundamental matrix, relate image coordinates to image coordinates