Computer Vision - Lecture 17
Epipolar Geometry & Stereo Basics
13.01.2015

Bastian Leibe
RWTH Aachen
http://www.vision.rwth-aachen.de
leibe@vision.rwth-aachen.de

Announcements
- Seminar in the summer semester
  - “Current Topics in Computer Vision and Machine Learning”
  - Block seminar, presentations in 1st week of semester break
  - Registration period: 14.01.2015 - 28.01.2015
  - http://web-info8.informatik.rwth-aachen.de/apse
  - Quick poll: Who would be interested in that?

Course Outline
- Image Processing Basics
- Segmentation & Grouping
- Object Recognition
- Local Features & Matching
- Object Categorization
- 3D Reconstruction
  - Epipolar Geometry and Stereo Basics
  - Camera calibration & Uncalibrated Reconstruction
  - Multi-view Stereo
- Optical Flow

Topics of This Lecture
- Geometric vision
  - Visual cues
  - Stereo vision
- Epipolar geometry
  - Depth with stereo
  - Geometry for a simple stereo system
  - Case example with parallel optical axes
  - General case with calibrated cameras
- Stereopsis & 3D Reconstruction
  - Correspondence search
  - Additional correspondence constraints
  - Possible sources of error
  - Applications

Geometric vision
- Goal: Recovery of 3D structure
  - What cues in the image allow us to do this?

Visual Cues
- Shading

Slide credit: Svetlana Lazebnik

Slide credit: Steve Seitz

Merle Norman Cosmetics, Los Angeles
Our Goal: Recovery of 3D Structure

- We will focus on perspective and motion
- We need multi-view geometry because recovery of structure from one image is inherently ambiguous
Stereo Vision

What Is Stereo Vision?
- Generic problem formulation: given several images of the same object or scene, compute a representation of its 3D shape

What Is Stereo Vision?
- Narrower formulation: given a calibrated binocular stereo pair, fuse it to produce a depth image.
  - Humans can do it

What Is Stereo Vision?
- Narrower formulation: given a calibrated binocular stereo pair, fuse it to produce a depth image.
  - Humans can do it

http://www.well.com/~jimg/stereo/stereo_list.html

Stereograms: Invented by Sir Charles Wheatstone, 1838

Stereograms: Invented by Sir Charles Wheatstone, 1838

Dense depth map

Autostereograms: http://www.magiceye.com
What Is Stereo Vision?

- Narrower formulation: given a calibrated binocular stereo pair, fuse it to produce a depth image.
  - Humans can do it

Application of Stereo: Robotic Exploration

- Nomad robot searches for meteorites in Antarctica
- Real-time stereo on Mars

Topics of This Lecture

- Geometric vision
  - Visual cues
  - Stereo vision
- Epipolar geometry
  - Depth with stereo
  - Geometry for a simple stereo system
  - Case example with parallel optical axes
  - General case with calibrated cameras
- Stereopsis & 3D Reconstruction
  - Correspondence search
  - Additional correspondence constraints
  - Possible sources of error
  - Applications

Depth with Stereo: Basic Idea

- Basic Principle: Triangulation
  - Gives reconstruction as intersection of two rays
  - Requires
    - Camera pose (calibration)
    - Point correspondence

Camera Calibration

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

We'll assume for now that these parameters are given and fixed.
Geometry for a Simple Stereo System

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

$$\frac{T - (x_e - x_l)}{Z - f} = \frac{T}{Z}$$

$$Z = T \cdot \frac{z_l}{x_e - x_l}$$

Depth From Disparity

$$D(x, y) = (x', y') = (x + D(x, y), y)$$

Stereo Correspondence Constraints

- Given $p$ in the left image, where can the corresponding point $p'$ in the right image be?
Stereo Correspondence Constraints

Geometry of two views allows us to constrain where the corresponding pixel for some image point in the first view must occur in the second view.

Epipolar constraint: Why is this useful?
- Reduces correspondence problem to 1D search along conjugate epipolar lines.

Epipolar Geometry

- 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

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

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
Example: Converging Cameras

As position of 3D point varies, epipolar lines "rotate" about the baseline

Example: Motion Parallel With Image Plane

Let’s Formalize This!

- For a given stereo rig, how do we express the epipolar constraints algebraically?
- For this, we will need some linear algebra.
- But don’t worry! We’ll go through it step by step...

Stereo Geometry With Calibrated Cameras

- If the 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; translation: 3 vector.

Rotation Matrix

Express 3D rotation as series of rotations around coordinate axes by angles $\alpha$, $\beta$, $\gamma$

Overall rotation is product of these elementary rotations:

$R = R_z R_y R_x$
3D Rigid Transformation

\[
\begin{bmatrix}
X' \\
Y' \\
Z'
\end{bmatrix} =
\begin{bmatrix}
r_{11} & r_{12} & r_{13} \\
r_{21} & r_{22} & r_{23} \\
r_{31} & r_{32} & r_{33}
\end{bmatrix}
\begin{bmatrix}
X \\
Y \\
Z
\end{bmatrix} +
\begin{bmatrix}
T_x \\
T_y \\
T_z
\end{bmatrix}
\]

\[X' = RX + T\]

Stereo Geometry With Calibrated Cameras

- Camera-centered coordinate systems are related by known rotation \(R\) and translation \(T\):

\[X' = RX + T\]

Excursion: 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 is 0.

\[\vec{a} \times \vec{b} = \vec{c}\]

\[\vec{a} \cdot \vec{c} = 0\]

\[\vec{b} \cdot \vec{c} = 0\]

Matrix Form of Cross Product

\[
\begin{bmatrix}
a_x \\
-a_y \\
a_z
\end{bmatrix} = \begin{bmatrix}
0 & -a_z & a_y \\
a_z & 0 & -a_x \\
-a_y & a_x & 0
\end{bmatrix}
\]

\[\vec{a} \times \vec{b} = \begin{bmatrix}
a_x \\
-a_y \\
a_z
\end{bmatrix} \begin{bmatrix}
a_x \\
-a_y \\
a_z
\end{bmatrix}
\]

From Geometry to Algebra

\[X' = RX + T\]

\[X' \cdot (T \times X') = X' \cdot (T \times RX)\]

\[T \times X' = T \times RX + T \times T\]

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

\[= T \times RX\]
Essential Matrix

\[ X' \cdot (T \cdot RX) = 0 \]

Let \( E = T \cdot R \)

\[ X' \cdot (T \cdot RX) = 0 \]

- This holds for the rays \( p \) and \( p' \) that are parallel to the camera-centered position vectors \( X \) and \( X' \), so we have: \( p'^T E p = 0 \)

- \( E \) is called the essential matrix, which relates corresponding image points [Longuet-Higgins 1981]

Essential Matrix: Properties

- Relates image of corresponding points in both cameras, given rotation and translation.
- Assuming intrinsic parameters are known

\[ E = T \cdot R \]

Essential Matrix and Epipolar Lines

Epipolar constraint: if we observe point \( p \) in one image, then its position \( p' \) in second image must satisfy this equation.

\[ l' = E p \]

\( l' \) is the coordinate vector representing the epipolar line for point \( p' \)

\( l = E' p \) is the coordinate vector representing the epipolar line for point \( p' \)

Epipolar constraint: \( l \) and \( l' \) are parallel, so \( l' = t \cdot l \)

\[ l' = t \cdot l \]

Essential Matrix Example: Parallel Cameras

\[ R = \]

\[ T = \]

\[ E = [T] \cdot R = \]

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

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

More General Case

Image \( I(x,y) \)
Disparity map \( D(x,y) \)
Image \( I'(x',y') \)

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

What about when cameras' optical axes are not parallel?
Stereo Image Rectification

- In practice, it is convenient if image scanlines are the epipolar lines.

- Algorithm
  - Reproject image planes onto a common plane parallel to the line between optical centers
  - Pixel motion is horizontal after this transformation
  - Two homographies ($3 \times 3$ transforms), one for each input image reprojection

Topics of This Lecture

- Geometric vision
  - Visual cues
  - Stereo vision
- Epipolar geometry
  - Depth with stereo
  - Geometry for a simple stereo system
  - Case example with parallel optical axes
  - General case with calibrated cameras
- Stereopsis & 3D Reconstruction
  - Correspondence search
  - Additional correspondence constraints
  - Possible sources of error
  - Applications

Stereo Reconstruction

- Main Steps
  - Calibrate cameras
  - Rectify images
  - Compute disparity
  - Estimate depth

Correspondence Problem

- Multiple match hypotheses satisfy epipolar constraint, but which is correct?
**Dense Correspondence Search**

- For each pixel in the first image
  - Find corresponding epipolar line in the right image
  - Examine all pixels on the epipolar line and pick the best match (e.g., SSD, correlation)
  - Triangulate the matches to get depth information
- This is easiest when epipolar lines are scanlines ➞ Rectify images first

---

**Example: Window Search**

- Data from University of Tsukuba

---

**Effect of Window Size**

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

---

**Alternative: Sparse Correspondence Search**

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

What would make good features?

---

**Dense vs. Sparse**

- Sparse
  - Efficiency
  - Can have more reliable feature matches, less sensitive to illumination than raw pixels
  - But...
    - Have to know enough to pick good features
    - Sparse information
- Dense
  - Simple process
  - More depth estimates, can be useful for surface reconstruction
  - But...
    - Breaks down in textureless regions anyway
    - Raw pixel distances can be brittle
    - Not good with very different viewpoints
Difficulties in Similarity Constraint

Possible Sources of Error?
- Low-contrast / textureless image regions
- Occlusions
- Camera calibration errors
- Violations of brightness constancy (e.g., specular reflections)
- Large motions

Application: View Interpolation

Disparity
Application: Free-Viewpoint Video

http://www.liberovision.com

Summary: Stereo Reconstruction

- Main Steps
  - Calibrate cameras
  - Rectify images
  - Compute disparity
  - Estimate depth

- So far, we have only considered calibrated cameras...

- Next lecture
  - Uncalibrated cameras
  - Camera parameters
  - Revisiting epipolar geometry
  - Robust fitting

References and Further Reading

- Background information on epipolar geometry and stereopsis can be found in Chapters 10.1-10.2 and 11.1-11.3 of
  D. Forsyth, J. Ponce,
  Computer Vision - A Modern Approach.
  Prentice Hall, 2003

- More detailed information (if you really want to implement 3D reconstruction algorithms) can be found in Chapters 9 and 10 of
  R. Hartley, A. Zisserman
  Multiple View Geometry in Computer Vision
  2nd Ed., Cambridge Univ. Press, 2004