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

Visual Cues

- Shading
  
Merle Norman Cosmetics, Los Angeles

- Texture
  
The Visual Cliff, by William Vandivert, 1960
Visual Cues
- Shading
- Texture
- Focus
- Perspective
- Motion

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

To Illustrate This Point…
- Structure and depth are inherently ambiguous from single views.

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

Autostereograms: [http://www.magiceye.com](http://www.magiceye.com)
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

Extrinsic parameters: Camera frame ↔ Reference frame

Intrinsic parameters: Image coordinates relative to camera ↔ Pixel coordinates

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

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:

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

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

  *(disparity)*

Depth From Disparity

\[\text{Image } I(x, y) \quad \text{Disparity map } D(x, y) \quad \text{Image } I'(x', y')\]

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

General Case With Calibrated Cameras

- The two cameras need not have parallel optical axes.

Stereo Correspondence Constraints

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

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

- Epipolar Plane
- Epipoles
- Baseline
- Epipolar Lines

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

- 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: Converging Cameras

As position of 3D point varies, epipolar lines "rotate" about the baseline.
Example: Motion Parallel With Image Plane

Example: Forward Motion

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:

\[
R = R_x R_y R_z
\]

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 $0$.

From Geometry to Algebra

- This holds for the rays $p$ and $p'$ that are parallel to the camera-centered position vectors $X$ and $X'$, so we have:

$$X' \cdot (T \times RX) = 0$$
$$X' \cdot (T \times RX') = 0$$

Let $E = T \cdot R$

$$X' \cdot E \cdot X = 0$$

- $E$ is called the essential matrix, which relates corresponding image points [Longuet-Higgins 1981].
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.

$$l' = Ep$$

$l' = E^T p'$ is the coordinate vector representing the epipolar line for point $p'$.

Essential Matrix: Properties

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

$$E = TR$$

Essential Matrix Example: Parallel Cameras

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

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

More General Case

$$\langle x', y' \rangle = \langle x + D(x,y), y \rangle$$

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

Stereo Image Rectification: Example

Source: Alyosha Efros
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?

Example: Window Search

- Data from University of Tsukuba

Example: Window Search

- Data from University of Tsukuba

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

- Window-based matching
  - (best window size)
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

- Untextured surfaces
- Occlusions

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

Right Image
Application: View Interpolation

Left Image

Application: View Interpolation

Disparity

Application: View Interpolation

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

- 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