Two-View Stereo

Slides from S. Lazebnik, S. Seitz, Y. Furukawa
Stereo

- What cues tell us about scene depth?

Slide from L. Lazebnik.
Stereograms

- Humans can fuse pairs of images to get a sensation of depth

Stereograms: Invented by Sir Charles Wheatstone, 1838

Slide from L. Lazebnik.
Stereograms

Slide from L. Lazebnik.
Stereograms

- Humans can fuse pairs of images to get a sensation of depth

Autostereograms: www.magiceye.com
Stereograms

- Humans can fuse pairs of images to get a sensation of depth

Autostereograms: www.magiceye.com

Slide from L. Lazebnik.
Problem formulation

- Given a calibrated binocular stereo pair, fuse it to produce a depth image

image 1  image 2

Dense depth map

Slide from L. Lazebnik.
Basic stereo matching algorithm

• 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
  • Triangulate the matches to get depth information

• Simplest case: epipolar lines are corresponding scanlines
  • When does this happen?

Slide from L. Lazebnik.
Simplest Case: Parallel images

- Image planes of cameras are parallel to each other and to the baseline
- Camera centers are at same height
- Focal lengths are the same
Simplest Case: Parallel images

- Image planes of cameras are parallel to each other and to the baseline
- Camera centers are at same height
- Focal lengths are the same
- Then epipolar lines fall along the horizontal scan lines of the images

Slide from L. Lazebnik.
Epipolar constraint:

\[ x'^T E x = 0, \quad E = [t_x]R \]

\[ R = I \quad t = (T, 0, 0) \]

\[ E = [t_x]R = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -T \\ 0 & T & 0 \end{bmatrix} \]

\[
\begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -T \\ 0 & T & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = 0 \quad \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ -T \\ T v \end{bmatrix} = 0 \quad T v' = T v
\]

The y-coordinates of corresponding points are the same!
Stereo image rectification
Stereo image rectification

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

Rectification example
Another rectification example

Unrectified

Rectified

Slide from L. Lazebnik.
Basic stereo matching algorithm

- If necessary, rectify the two stereo images to transform epipolar lines into scanlines.
- 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.
Correspondence search

- Slide a window along the right scanline and compare contents of that window with the reference window in the left image
- Matching cost: SSD or normalized correlation

Slide from L. Lazebnik.
Correspondence search

Left

Right

scanline

SSD

Slide from L. Lazebnik.
Correspondence search

Scanline

Norm. corr

Slide from L. Lazebnik.
Basic stereo matching algorithm

- If necessary, rectify the two stereo images to transform epipolar lines into scanlines.
- For each pixel $x$ in the first image:
  - Find corresponding epipolar scanline in the right image.
  - Examine all pixels on the scanline and pick the best match $x'$.
  - Triangulate the matches to get depth information.

Slide from L. Lazebnik.
Depth from disparity

\[ \frac{x}{f} = \frac{B_1}{z} \quad \frac{-x'}{f} = \frac{B_2}{z} \]

\[ \frac{x - x'}{f} = \frac{B_1 + B_2}{z} \]

Disparity is inversely proportional to depth!
Depth from disparity

\[
\frac{x}{f} = \frac{B_1}{z} \quad \frac{x'}{f} = \frac{B_2}{z}
\]

\[
\frac{x - x'}{f} = \frac{B_1 - B_2}{z}
\]

\[
\text{disparity} = x - x' = \frac{B \cdot f}{z}
\]

Slide from L. Lazebnik.
Basic stereo matching algorithm

- If necessary, rectify the two stereo images to transform epipolar lines into scanlines.
- For each pixel $x$ in the first image:
  - Find corresponding epipolar scanline in the right image.
  - Examine all pixels on the scanline and pick the best match $x'$.
  - Compute disparity $x - x'$ and set $\text{depth}(x) = B^*f/(x - x')$.
Failures of correspondence search

- Textureless surfaces
- Occlusions, repetition
- Non-Lambertian surfaces, specularities

Slide from L. Lazebnik.
Effect of window size

- Smaller window
  + More detail
  - More noise

- Larger window
  + Smoother disparity maps
  - Less detail

Slide from L. Lazebnik.
Results with window search

Slide from L. Lazebnik.
Better methods exist...

Graph cuts

Ground truth


For the latest and greatest: [http://www.middlebury.edu/stereo/](http://www.middlebury.edu/stereo/)

Slide from L. Lazebnik.
How can we improve window-based matching?

- The similarity constraint is **local** (each reference window is matched independently)
- Need to enforce **non-local** correspondence constraints
Non-local constraints

- **Uniqueness**
  - For any point in one image, there should be at most one matching point in the other image.

Slide from L. Lazebnik.
Non-local constraints

- **Uniqueness**
  - For any point in one image, there should be at most one matching point in the other image

- **Ordering**
  - Corresponding points should be in the same order in both views
Non-local constraints

- **Uniqueness**
  - For any point in one image, there should be at most one matching point in the other image

- **Ordering**
  - Corresponding points should be in the same order in both views
Non-local constraints

- **Uniqueness**
  - For any point in one image, there should be at most one matching point in the other image

- **Ordering**
  - Corresponding points should be in the same order in both views

- **Smoothness**
  - We expect disparity values to change slowly (for the most part)
Scanline stereo

- Try to coherently match pixels on the entire scanline
- Different scanlines are still optimized independently

Slide from L. Lazebnik.
“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

- Can’t use dynamic programming to find spatially coherent disparities/ correspondences on a 2D grid
Stereo matching as global optimization

\[ E(D) = \sum_i \left( W_1(i) - W_2(i + D(i)) \right)^2 + \lambda \sum_{\text{neighbors } i,j} \rho \left( D(i) - D(j) \right) \]

- Energy functions of this form can be minimized using graph cuts
  

Slide from L. Lazebnik.
Stereo matching as a prediction problem

Review: Basic stereo matching algorithm

- For each pixel \( x \) in the reference image
  - Find corresponding epipolar scanline in the other image
  - Examine all pixels on the scanline and pick the best match \( x' \)
  - Compute disparity \( x - x' \) and set \( \text{depth}(x) = B \cdot f / (x - x') \)
Active sensing simplifies the problem of estimating point correspondences.
Kinect: Structured infrared light


Slide from L. Lazebnik.
Apple TrueDepth

https://www.cnet.com/news/apple-face-id-truedepth-how-it-works/

Slide from L. Lazebnik.
Laser scanning

Optical triangulation

- Project a single stripe of laser light
- Scan it across the surface of the object
- This is a very precise version of structured light scanning

Digital Michelangelo Project
Levoy et al.
http://graphics.stanford.edu/projects/mich/

Source: S. Seitz
Laser scanned models

The Digital Michelangelo Project, Levoy et al.

Source: S. Seitz
Laser scanned models

*The Digital Michelangelo Project, Levoy et al.*

Source: S. Seitz
Laser scanned models

*The Digital Michelangelo Project*, Levoy et al.

Source: S. Seitz
Laser scanned models

The Digital Michelangelo Project, Levoy et al.

Source: S. Seitz
Laser scanned models

1.0 mm resolution (56 million triangles)

*The Digital Michelangelo Project, Levoy et al.*
Stereo error (distance)

Error in distance estimate increases quadratically with the distance.

\[ Z = \text{distance} \]
\[ d = \text{disparity} \]
\[ Z = \frac{C}{d} \]
\[ \delta Z = -\frac{Z^2}{C} \delta d \]
\[ |\delta Z| = \frac{Z^2}{C} |\delta d| \]
\[ \text{error} \propto \text{distance}^2 \]
Multi-view stereo

Slide from L. Lazebnik.
Multi-view stereo: Basic idea

Source: Y. Furukawa
Multi-view stereo: Basic idea

Source: Y. Furukawa
Multi-view stereo: Basic idea

Source: Y. Furukawa
Multi-view stereo: Basic idea

Source: Y. Furukawa
Towards Internet-Scale Multi-View Stereo

YouTube video, CMVS software

Applications

Source: N. Snavely
Applications

Source: N. Snavely