Stereo

EECS 442 – David Fouhey and Justin Johnson
Winter 2021, University of Michigan
https://web.eecs.umich.edu/~justincj/teaching/eecs442/WI2021/
Updates

• Project discussions this week. Not mandatory, but could be useful!
• Check your grades from earlier assignments. If you lost 50% of the grade because you submitted a .rar instead of a .zip, talk to us.
Two-View Stereo
Stereo

Slide credit: S. Lazebnik
How Two Photographers Unknowingly Shot the Same Millisecond in Time

https://petapixel.com/2018/03/07/two-photographers-unknowingly-shot-millisecond-time/
How Two Photographers Unknowingly Shot the Same Millisecond in Time

https://petapixel.com/2018/03/07/two-photographers-unknowingly-shot-millisecond-time/
How Two Photographers Unknowingly Shot the Same Millisecond in Time

https://petapixel.com/2018/03/07/two-photographers-unknowingly-shot-millisecond-time/
Stereograms

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

Stereograms: Invented by Sir Charles Wheatstone, 1838

Slide credit: S. Lazebnik
Stereograms

Slide credit: S. Lazebnik
Stereograms

What about this?

Slide Credit: S. Lazebnik, but idea of random dot stereogram is due to B. Julesz
Stereograms
Bela Julesz: Random Dot Stereogram
Shows that stereo can operate *without* recognition
Stereograms

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

Autostereograms: www.magiceye.com

Slide credit: S. Lazebnik
Stereograms

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

Autostereograms: www.magiceye.com
Problem formulation

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

image 1

image 2

Dense depth map

Slide credit: S. 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 = corresponding scanlines
  - When does this happen?

Slide credit: S. 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 the same

Slide credit: S. 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 the same
• Then epipolar lines fall along the horizontal scan lines of the images

Slide credit: S. Lazebnik
Essential matrix for parallel images

\[ \mathbf{pE} \mathbf{p}' = 0 \quad \mathbf{E} = [t_x]\mathbf{R} \]

What's \( \mathbf{R} \)?
What's \( \mathbf{t} \)?

\[ \mathbf{R} = \mathbf{I} \quad \mathbf{t} = [T, 0, 0] \]

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

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

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

The y-coordinates of corresponding points are the same!

Slide credit: S. Lazebnik
Stereo image rectification
Stereo image rectification

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


Slide credit: S. Lazebnik
Rectification example

Slide credit: S. Lazebnik
Another rectification example

Unrectified

Rectified

Slide credit: S. 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

Slide credit: S. Lazebnik
Correspondence Search

Slide window along the right scanline, compare contents of that window with reference window on left

Matching cost: SSD or normalized correlation

Slide credit: S. Lazebnik
Correspondence Search

Matching cost
Sum of squared differences

$\sum_i (l_i - r_i)^2$

Disparity

Slide credit: S. Lazebnik
Correspondence Search

Left

Right

Matching cost

Normalized correlation

\[ \hat{x}_i = \frac{x_i - \text{mean}(x)}{\text{std}(x)} \]

\[ \hat{l} \cdot \hat{r} \]

Disparity

Slide credit: S. 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 credit: S. Lazebnik
Triangulation: Older History

- From [Wikipedia](https://en.wikipedia.org): Gemma Frisius's 1533 diagram introducing the idea of triangulation into the science of surveying. Having established a baseline, e.g. the cities of Brussels and Antwerp, the location of other cities, e.g. Middelburg, Ghent etc., can be found by taking a compass direction from each end of the baseline, and plotting where the two directions cross. This was only a theoretical presentation of the concept — due to topographical restrictions, it is impossible to see Middelburg from either Brussels or Antwerp. Nevertheless, the figure soon became well known all across Europe.

Slide credit: S. Lazebnik
Triangulation: Modern History
Depth from disparity

\[ \frac{x}{f} = \frac{B_1}{z} \]

By similar triangles

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

Similarly by similar triangles

Diagram adapted from S. Lazebnik
Depth from disparity

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

Add them

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

Disparity

Diagram adapted from S. Lazebnik
Depth from disparity

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

Subtract them

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

Diagram adapted from S. 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 regions. Why?

Image credit: S. Lazebnik
Failures of Correspondence Search

Repeated Patterns. Why?

Image credit: S. Lazebnik
Failures of Correspondence Search

Specular Surfaces. Why?

Image credit: S. Lazebnik
Effect of window size

• Smaller window
  + More detail
  - More noise

• Larger window
  + Smoother disparity maps
  - Less detail

Image credit: S. Lazebnik
Results with window search

Data

Window-based matching

Ground truth

Image credit: S. Lazebnik
Better methods exist...


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

Image credit: S. Lazebnik
Improving Window-based Matching

- Similarity is **local** (each window independent)
- Need non-local correspondence constraints / cues.
Uniqueness

• Each point in one image should match at most one point in other image.

• **When might this not be true?**
Ordering

- Corresponding points should be in the same order.

![Diagram showing corresponding points in the same order.](image)
Ordering

- Not always true!
Smoothness

• We expect disparity values to change slowly (for the most part)

• When is this not true?
Scanline Stereo

- Try to coherently match pixels on the entire scanline
- Different scanlines are optimized (by dynamic programming) independently

Slide credit: S. Lazebnik
Coherent Stereo on 2D Grid

- Scanline stereo generates streaking artifacts
- Can’t use dynamic programming to find spatially coherent disparities on a 2D grid

Slide credit: S. Lazebnik
Stereo Matching as Optimization

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

Data term

Smoothness term

Solvable by graph cuts for certain smoothnesses \( \rho \)


Slide credit: S. Lazebnik
Is This Doable by Deep Network?

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

Data term

Smoothness term

Easy solution: replace the data term with a network
Deep Learning For Stereo

- Feed in two images to identical networks, concatenate outputs, learn multilayer perception
- Slow: why?

![Diagram showing the process of deep learning for stereo vision]
Deep Learning For Stereo

- Normalize outputs; treat dot product as prediction of match/no match
- Fast: why?

$$\frac{x}{\|x\|}, \quad \frac{x^T x'}{\|x'\|}$$

svm/hinge loss
Stereo datasets

- Middlebury stereo datasets
- KITTI
- Synthetic data?

Slide credit: S. Lazebnik
Active stereo with structured light

- Project “structured” light patterns onto the object
  - Simplifies the correspondence problem
  - Allows us to use only one camera

Active stereo with structured light


Slide credit: S. Lazebnik
Kinect: Structured infrared light

Apple TrueDepth

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

Slide credit: S. 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/
Laser scanned models

The Digital Michelangelo Project, Levoy et al.

Source: S. Seitz
Laser scanned models

Source: S. Seitz

The Digital Michelangelo Project, Levoy et al.
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.

Source: S. Seitz
Aligning range images

- One range scan not enough for complex surfaces
- Need techniques to register multiple range images

“Shortest paths” for scan-line stereo

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