CS4501: Introduction to Computer Vision
Feature Matching and Dense Stereo

Various slides from previous courses by:
D.A. Forsyth (Berkeley / UIUC), I. Kokkinos (Ecole Centrale / UCL), S. Lazebnik (UNC / UIUC), S. Seitz (MSR / Facebook), J. Hays (Brown / Georgia Tech), A. Berg (Stony Brook / UNC), D. Samaras (Stony Brook). J. M. Frahm (UNC), V. Ordonez (UVA), Steve Seitz (UW), Kristen Grauman (UT Austin) – Dense Stereo, Shree Nayar (Columbia) Sparse Feature Matching – Image Stitching.
Today’s Class and Assignment 3

• Feature Matching – Panorama Stitching
• Stereo Vision – Dense Stereo / Stereo Matching
Multiple views

Multi-view geometry, matching, invariant features, stereo vision

Hartley and Zisserman

Lowe
Why multiple views?

• Structure and depth are inherently ambiguous from single views.
Why multiple views?

- Structure and depth are inherently ambiguous from single views.

![Diagram showing why multiple views are necessary](image-url)
• What cues help us to perceive 3d shape and depth?
Shading

[Figure from Prados & Faugeras 2006]
Focus/defocus

Images from same point of view, different camera parameters

3d shape / depth estimates

[figs from H. Jin and P. Favaro, 2002]
Texture

Perspective effects

Image credit: S. Seitz
Motion

Figures from L. Zhang

http://www.brainconnection.com/teasers/?main=illusion/motion-shape
Estimating scene shape

• “Shape from X”: Shading, Texture, Focus, Motion...

• Stereo:
  • shape from “motion” between two views
  • infer 3d shape of scene from two (multiple) images from different viewpoints

Main idea:
Human eye

Rough analogy with human visual system:

Pupil/Iris – control amount of light passing through lens

Retina - contains sensor cells, where image is formed

Fovea – highest concentration of cones
Human stereopsis: disparity

Human eyes fixate on point in space – rotate so that corresponding images form in centers of fovea.

From Bruce and Green, Visual Perception, Physiology, Psychology and Ecology
Human stereopsis: disparity

**Disparity** occurs when eyes fixate on one object; others appear at different visual angles

From Bruce and Green, Visual Perception, Physiology, Psychology and Ecology

Adapted from David Forsyth, UC Berkeley
Stereo photography and stereo viewers

Take two pictures of the same subject from two slightly different viewpoints and display so that each eye sees only one of the images.

Invented by Sir Charles Wheatstone, 1838

Image from fisher-price.com
Public Library, Stereoscopic Looking Room, Chicago, by Phillips, 1923
Autostereograms

Exploit disparity as depth cue using single image.

(Single image random dot stereogram, Single image stereogram)

Images from magiceye.com
Estimating depth with stereo

• **Stereo**: shape from “motion” between two views

• We’ll need to consider:
  • Info on camera pose (“calibration”)
  • Image point correspondences
Key idea: Epipolar constraint

Potential matches for \( x \) have to lie on the corresponding line \( l' \).

Potential matches for \( x' \) have to lie on the corresponding line \( l \).
Epipolar geometry: notation

- **Baseline** – line connecting the two camera centers
- **Epipoles**
  = intersections of baseline with image planes
  = projections of the other camera center
- **Epipolar Plane** – plane containing baseline (1D family)
Epipolar geometry: notation

- **Baseline** – line connecting the two camera centers
- **Epipoles**
  = intersections of baseline with image planes
  = projections of the other camera center
- **Epipolar Plane** – plane containing baseline (1D family)
- **Epipolar Lines** - intersections of epipolar plane with image planes (always come in corresponding pairs)
Example: Converging cameras
Geometry for a simple stereo system

• First, assuming parallel optical axes, known camera parameters (i.e., calibrated cameras):
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
Focal length

World point

Depth of p

image point (left)

image point (right)

optical center (left)

baseline

optical center (right)

http://www.cse.psu.edu/~zyler/Chemo/Stereom20geometry.jpg
Geometry for a simple stereo system

- Assume parallel optical axes, known camera parameters (i.e., calibrated cameras). **What is expression for Z?**

Similar triangles \((p_l, P, p_r)\) and \((O_l, P, O_r)\):

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

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

disparity
Depth from disparity

So if we could find the corresponding points in two images, we could estimate relative depth...
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
Correspondence search
Correspondence search

Left

Right

scanline

Norm. corr
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 $\text{depth}(x) = B f / (x - x')$. 

*Image*:

- Hon. Abraham Lincoln, President of the United States.
Failures of correspondence search

Textureless surfaces

Occlusions, repetition

Non-Lambertian surfaces, specularities
Effect of window size

• Smaller window
  + More detail
  • More noise

• Larger window
  + Smoother disparity maps
  • Less detail

W = 3
W = 20
Results with window search

Data

Window-based matching

Ground truth
Better methods exist...


For the latest and greatest: [http://www.middlebury.edu/stereo/](http://www.middlebury.edu/stereo/)
When cameras are not aligned: Stereo image rectification

- Reproject image planes onto a common plane parallel to the line between optical centers
- Pixel motion is horizontal after this transformation
- Two homographies (3x3 transform), one for each input image reprojection

Rectification example
Panorama Stitching
Robust feature-based alignment

- Extract features
- Compute *putative matches*
- Loop:
  - *Hypothesize* transformation $T$ (small group of putative matches that are related by $T$)
  - *Verify* transformation (search for other matches consistent with $T$)

Source: L. Lazebnik
Robust feature-based alignment

- Extract features
- Compute putative matches
- Loop:
  - Hypothesize transformation $T$ (small group of putative matches that are related by $T$)
  - Verify transformation (search for other matches consistent with $T$)

Source: L. Lazebnik
The mosaic has a natural interpretation in 3D
- The images are reprojected onto a common plane
- The mosaic is formed on this plane
- Mosaic is a *synthetic wide-angle camera*

Source: Steve Seitz
• The mosaic has a natural interpretation in 3D
  • The images are reprojected onto a common plane
  • The mosaic is formed on this plane
  • Mosaic is a *synthetic wide-angle camera*
Image reprojection

• Basic question
  • How to relate two images from the same camera center?
    • how to map a pixel from PP1 to PP2

  Answer
  • Cast a ray through each pixel in PP1
  • Draw the pixel where that ray intersects PP2

Observation:
Rather than thinking of this as a 3D reprojection, think of it as a 2D **image warp** from one image to another.

Source: Alyosha Efros
Image reprojection: Homography

- A projective transform is a mapping between any two PPs with the same center of projection.
  - rectangle should map to arbitrary quadrilateral
  - parallel lines aren’t
  - but must preserve straight lines

- called **Homography**

\[
\begin{bmatrix}
wx' \\
w'x \\
w
\end{bmatrix} = \begin{bmatrix}
* & * & * \\
* & * & * \\
* & * & * \\
\end{bmatrix} \begin{bmatrix}
x \\
y \\
l
\end{bmatrix}
\]

\[
p' = \begin{bmatrix}
H & p
\end{bmatrix}
\]

Source: Alyosha Efros
Homography

To compute the homography given pairs of corresponding points in the images, we need to set up an equation where the parameters of $H$ are the unknowns...
Solving for homographies

\[ \mathbf{p}' = \mathbf{H} \mathbf{p} \]

\[
\begin{bmatrix}
wx' \\
wy' \\
w
\end{bmatrix} =
\begin{bmatrix}
a & b & c \\
d & e & f \\
g & h & i
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
1
\end{bmatrix}
\]

• Can set scale factor \( i = 1 \). So, there are 8 unknowns.

• Set up a system of linear equations:

\[ \mathbf{A}\mathbf{h} = \mathbf{b} \]

• Where vector of unknowns \( \mathbf{h} = [a,b,c,d,e,f,g,h]^T \)

• Need at least 8 eqs, but the more the better…

• Solve for \( \mathbf{h} \). If overconstrained, solve using least-squares:

\[
\min ||\mathbf{A}\mathbf{h} - \mathbf{b}||^2
\]

subject to \( ||\mathbf{h}||^2 = 1 \)
\[
\begin{bmatrix}
    x_d \\
    y_d \\
    1
\end{bmatrix}
\equiv
\begin{bmatrix}
    \tilde{x}_d \\
    \tilde{y}_d \\
    \tilde{z}_d
\end{bmatrix}
= \begin{bmatrix}
    h_{11} & h_{12} & h_{13} \\
    h_{21} & h_{22} & h_{23} \\
    h_{31} & h_{32} & h_{33}
\end{bmatrix}
\begin{bmatrix}
    x_s \\
    y_s \\
    1
\end{bmatrix}
\]

\[
x_{d(i)} = \frac{\tilde{x}_d^{(i)}}{\tilde{z}_d^{(i)}} = \frac{h_{11}x_s^{(i)} + h_{12}y_s^{(i)} + h_{13}}{h_{31}x_s^{(i)} + h_{32}y_s^{(i)} + h_{33}}
\]

\[
y_{d(i)} = \frac{\tilde{y}_d^{(i)}}{\tilde{z}_d^{(i)}} = \frac{h_{21}x_s^{(i)} + h_{22}y_s^{(i)} + h_{23}}{h_{31}x_s^{(i)} + h_{32}y_s^{(i)} + h_{33}}
\]

\[
x_{d(i)} \left( h_{31}x_s^{(i)} + h_{32}y_s^{(i)} + h_{33} \right) = h_{11}x_s^{(i)} + h_{12}y_s^{(i)} + h_{13}
\]

\[
y_{d(i)} \left( h_{31}x_s^{(i)} + h_{32}y_s^{(i)} + h_{33} \right) = h_{21}x_s^{(i)} + h_{22}y_s^{(i)} + h_{23}
\]
\[
\begin{align*}
x_d^{(i)} \left( h_{31} x_s^{(i)} + h_{32} y_s^{(i)} + h_{33} \right) &= h_{11} x_s^{(i)} + h_{12} y_s^{(i)} + h_{13} \\
y_d^{(i)} \left( h_{31} x_s^{(i)} + h_{32} y_s^{(i)} + h_{33} \right) &= h_{21} x_s^{(i)} + h_{22} y_s^{(i)} + h_{23}
\end{align*}
\]

\[
\begin{bmatrix}
x_s^{(i)} & y_s^{(i)} & 1 & 0 & 0 & 0 & -x_d^{(i)} & x_s^{(i)} & -x_d^{(i)} y_s^{(i)} & -x_d^{(i)} y_s^{(i)} & -y_d^{(i)} \\
0 & 0 & 0 & x_s^{(i)} & y_s^{(i)} & 1 & -y_d^{(i)} & x_s^{(i)} & -y_d^{(i)} y_s^{(i)} & -y_d^{(i)} y_s^{(i)} & -y_d^{(i)}
\end{bmatrix}
\begin{bmatrix}
h_{11} \\
h_{12} \\
h_{13} \\
h_{21} \\
h_{22} \\
h_{23} \\
h_{31} \\
h_{32} \\
h_{33}
\end{bmatrix}
= \begin{bmatrix} 0 \\ 0 \end{bmatrix}
\]
Combining the equations for all corresponding points:

\[
\begin{bmatrix}
    x_s^{(1)} & y_s^{(1)} & 1 & 0 & 0 & 0 & -x_d^{(1)} & x_s^{(1)} & -x_d^{(1)} & y_s^{(1)} & -x_d^{(1)} \\
    0 & 0 & 0 & x_s^{(1)} & y_s^{(1)} & 1 & -y_d^{(1)} & x_s^{(1)} & -y_d^{(1)} & y_s^{(1)} & -y_d^{(1)} \\
    \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
    x_s^{(i)} & y_s^{(i)} & 1 & 0 & 0 & 0 & -x_d^{(i)} & x_s^{(i)} & -x_d^{(i)} & y_s^{(i)} & -x_d^{(i)} \\
    0 & 0 & 0 & x_s^{(i)} & y_s^{(i)} & 1 & -y_d^{(i)} & x_s^{(i)} & -y_d^{(i)} & y_s^{(i)} & -y_d^{(i)} \\
    \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
    x_s^{(n)} & y_s^{(n)} & 1 & 0 & 0 & 0 & -x_d^{(n)} & x_s^{(n)} & -x_d^{(n)} & y_s^{(n)} & -x_d^{(n)} \\
    0 & 0 & 0 & x_s^{(n)} & y_s^{(n)} & 1 & -y_d^{(n)} & x_s^{(n)} & -y_d^{(n)} & y_s^{(n)} & -y_d^{(n)} \\
\end{bmatrix}
\begin{bmatrix}
h_{11} \\
h_{12} \\
h_{13} \\
h_{21} \\
h_{22} \\
h_{23} \\
h_{31} \\
h_{32} \\
h_{33}
\end{bmatrix}
= \begin{bmatrix}
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0
\end{bmatrix}
\]
Combining the equations for all corresponding points:

\[
\begin{bmatrix}
    x_s^{(1)} & y_s^{(1)} & 1 & 0 & 0 & 0 & -x_d^{(1)} & x_s^{(1)} & -x_d^{(1)} & y_s^{(1)} & -x_d^{(1)} \\
    0 & 0 & 0 & x_s^{(1)} & y_s^{(1)} & 1 & -y_d^{(1)} & x_s^{(1)} & -y_d^{(1)} & y_s^{(1)} & -y_d^{(1)} \\
    \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
    x_s^{(i)} & y_s^{(i)} & 1 & 0 & 0 & 0 & -x_d^{(i)} & x_s^{(i)} & -x_d^{(i)} & y_s^{(i)} & -x_d^{(i)} \\
    0 & 0 & 0 & x_s^{(i)} & y_s^{(i)} & 1 & -y_d^{(i)} & x_s^{(i)} & -y_d^{(i)} & y_s^{(i)} & -y_d^{(i)} \\
    \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
    x_s^{(n)} & y_s^{(n)} & 1 & 0 & 0 & 0 & -x_d^{(n)} & x_s^{(n)} & -x_d^{(n)} & y_s^{(n)} & -x_d^{(n)} \\
    0 & 0 & 0 & x_s^{(n)} & y_s^{(n)} & 1 & -y_d^{(n)} & x_s^{(n)} & -y_d^{(n)} & y_s^{(n)} & -y_d^{(n)} \\
\end{bmatrix}
\begin{bmatrix}
h_{11} \\
h_{12} \\
h_{13} \\
h_{21} \\
h_{22} \\
h_{23} \\
h_{31} \\
h_{32} \\
h_{33}
\end{bmatrix}
= \begin{bmatrix}
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0
\end{bmatrix}
\]

Solve for \( h \): \( A h = 0 \) such that \( \| h \|^2 = 1 \)

From Shree Nayar
Constrained Least Squares

Solve for $\mathbf{h}$: $A \mathbf{h} = 0$ such that $\|\mathbf{h}\|^2 = 1$

Define least squares problem:

$$\min_{\mathbf{h}} \|A\mathbf{h}\|^2 \text{ such that } \|\mathbf{h}\|^2 = 1$$

We know that:

$$\|A\mathbf{h}\|^2 = (A\mathbf{h})^T(A\mathbf{h}) = \mathbf{h}^T A^T A \mathbf{h} \quad \text{and} \quad \|\mathbf{h}\|^2 = \mathbf{h}^T \mathbf{h} = 1$$

$$\min_{\mathbf{h}} (\mathbf{h}^T A^T A \mathbf{h}) \text{ such that } \mathbf{h}^T \mathbf{h} = 1$$
Constrained Least Squares

\[
\min_{\mathbf{h}} (\mathbf{h}^T \mathbf{A}^T \mathbf{A} \mathbf{h}) \quad \text{such that} \quad \mathbf{h}^T \mathbf{h} = 1
\]

Define Loss function \( L(\mathbf{h}, \lambda) \):

\[
L(\mathbf{h}, \lambda) = \mathbf{h}^T \mathbf{A}^T \mathbf{A} \mathbf{h} - \lambda (\mathbf{h}^T \mathbf{h} - 1)
\]

Taking derivatives of \( L(\mathbf{h}, \lambda) \) w.r.t \( \mathbf{h} \):

\[
2 \mathbf{A}^T \mathbf{A} \mathbf{h} - 2\lambda \mathbf{h} = 0
\]

\[
A^T A \mathbf{h} = \lambda \mathbf{h}
\]

Eigenvector \( \mathbf{h} \) with smallest eigenvalue \( \lambda \) of matrix \( A^T A \) minimizes the loss function \( L(\mathbf{h}) \).
Strong recommend you watching Prof. Shree Nayar’s Lectures on Youtube

I created playlist of all the videos on Image Stitching – There is one on Warping and Blending that will help you with part 1.4 in your Assignment 3, and maybe help you get even better output results than mine.

- [https://www.youtube.com/playlist?list=PL-jil_nJgMKC9hZEGnjjnP-sqEb6Tixw5x](https://www.youtube.com/playlist?list=PL-jil_nJgMKC9hZEGnjjnP-sqEb6Tixw5x)
Stereo:
Epipolar geometry

Vicente Ordonez
University of Virginia
Questions?