EECS 4422/5323 Computer Vision

Stereopsis 1

Calden Wloka

4 November, 2019

Announcements

- Site Visit this week
- Assignment 2 out
- Course swan song...

Final Weeks

- 5-6 topic lectures (including this one) left

Final Weeks

- 5-6 topic lectures (including this one) left
  - Stereopsis

Final Weeks

- 5-6 topic lectures (including this one) left
  - Stereopsis
  - Motion

Final Weeks

- 5-6 topic lectures (including this one) left
  - Stereopsis
  - Motion
- Assignment 2 due in a week and a half

Final Weeks

- 5-6 topic lectures (including this one) left
  - Stereopsis
  - Motion
- Assignment 2 due in a week and a half
- Projects need to be finished by the end of the month

Final Project Deliverables

- In-class demo

Final Project Deliverables

- In-class demo
  - Timing will be determined after the drop date

Final Project Deliverables

- In-class demo
  - Timing will be determined after the drop date
  - Organization will be similar to the site visit (randomized slot assignment)

Final Project Deliverables

- In-class demo
  - Timing will be determined after the drop date
  - Organization will be similar to the site visit (randomized slot assignment)
  - If you have any restrictions on timing, let me know by Nov. 11

Final Project Deliverables

- In-class demo
  - Timing will be determined after the drop date
  - Organization will be similar to the site visit (randomized slot assignment)
  - If you have any restrictions on timing, let me know by Nov. 11
- Final report is due by Dec. 2nd

For Those Paying Attention...

As of today, all material pertaining to marks has necessarily been presented.

Because of this, labs will be devoted to either Assignment 2 (next week) or project completion (after next week).

Outline

- Quick aside: Integral Images
- Non-stereo Depth Cues
- Parralax

Motivation

If we are computing the number of pixels with a certain property or average pixel values over subsets of pixels, we end up repeating many computations (any given candidate region must visit every pixel inside). When those computations are always the same, however, we can often operate much more efficiently using an *integral image*.

To compute the sum within the enclosed rectangle $ABCD$, it suffices to computer $D + A - B - C$

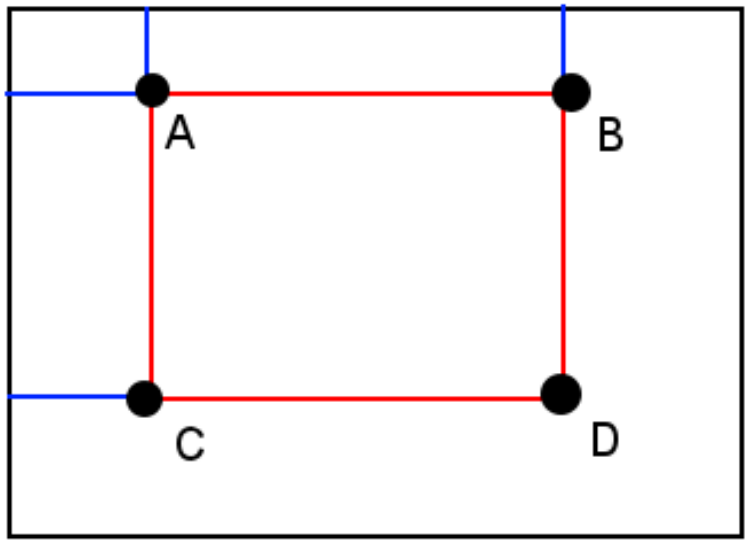


Image source: Wikipedia

Example Integral Image

Imagine a binary image representing some pixel quality we are interested in. We can represent the same information as an integral image, where the value of a pixel is equal to the sum of all pixels to the left and above it.

0	1	1	1	0
0	0	1	1	1
0	0	1	1	1
1	0	0	1	1
0	1	0	1	0

0	1	2	3	3
0	1	3	5	6
0	1	4	7	9
1	2	5	9	12
1	3	6	11	14

Recovering 3D from 2D

The general goal of stereopsis is to recover 3D information from 2D images.

Recovering 3D from 2D

The general goal of stereopsis is to recover 3D information from 2D images.

Stereo vision is not the only source of 3D structure, however. Typically, we refer to the recovery of 3D information from images as "Shape-from-X" or "Structure-from-X", where "X" is the type of cue involved.

Shape-from-X

- Single image cues:

Shape-from-X

- Single image cues:
  - Perspective

Shape-from-X

- Single image cues:
  - Perspective
  - Contour

Shape-from-X

- Single image cues:
  - Perspective
  - Contour
  - Texture

Shape-from-X

- Single image cues:
  - Perspective
  - Contour
  - Texture
  - Aerial Perspective

Shape-from-X

- Single image cues:
  - Perspective
  - Contour
  - Texture
  - Aerial Perspective
  - Lighting and Shading

Shape-from-X

- Single image cues:
  - Perspective
  - Contour
  - Texture
  - Aerial Perspective
  - Lighting and Shading
- Multiple image cues:

Shape-from-X

- Single image cues:
  - Perspective
  - Contour
  - Texture
  - Aerial Perspective
  - Lighting and Shading
- Multiple image cues:
  - Focus

Shape-from-X

- Single image cues:
  - Perspective
  - Contour
  - Texture
  - Aerial Perspective
  - Lighting and Shading
- Multiple image cues:
  - Focus
  - Motion (parallax, optical expansion)

Shape-from-X

- Single image cues:
  - Perspective
  - Contour
  - Texture
  - Aerial Perspective
  - Lighting and Shading
- Multiple image cues:
  - Focus
  - Motion (parallax, optical expansion)
  - Binocular Stereo

Perspective

Parallel lines appear to converge in the distance, which can provide a powerful cue for the relative distance of objects.

Image source: Photo by Julian Hochgesang

Contours and Occlusion

Object contours and occlusions provide very strong cues for relative distance (although little in the way of absolute distance).

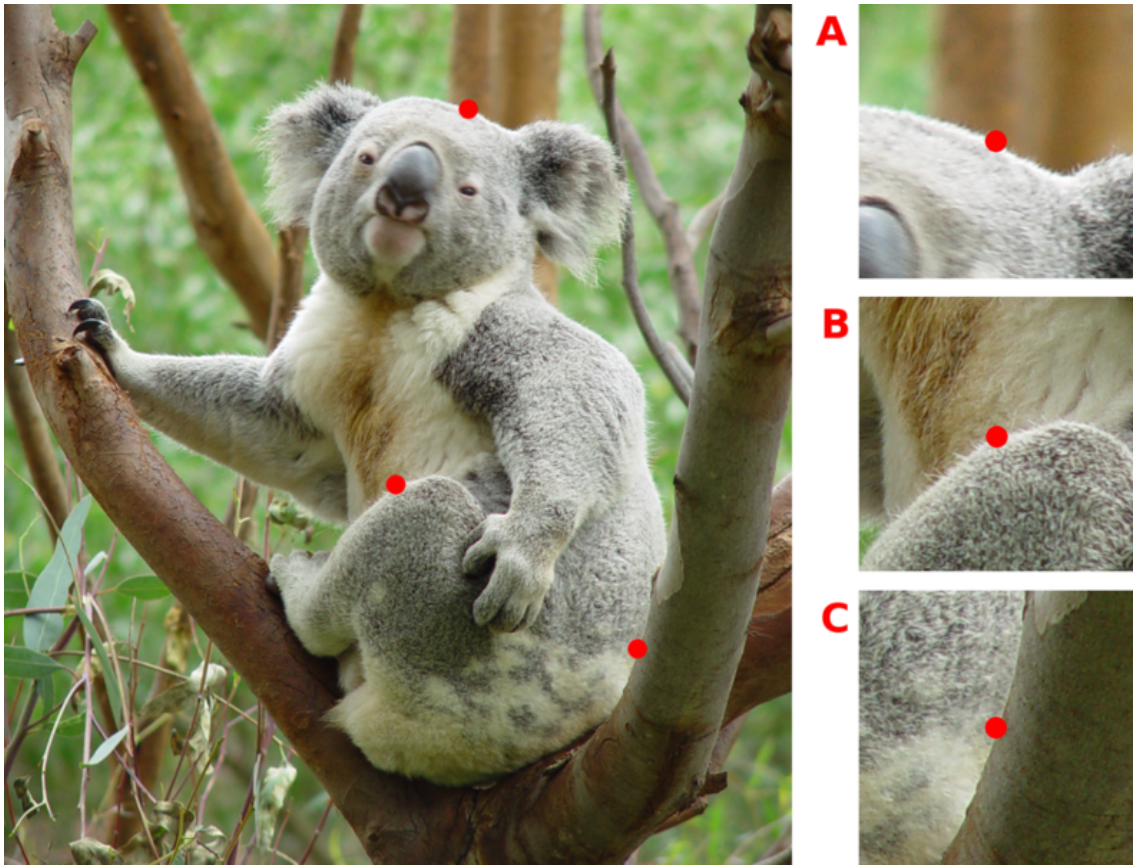


Image source: Williford & von der Heydt, 2013

Texture

For textured surfaces, the texture tends to be more clearly visible closer to the observer, and the texture elements grow more crowded and less distinct with distance.

Image source: Original image source unknown.

Aerial Perspective

The atmosphere scatters light, which tends to make objects further away seem hazy. This effect typically reduces colour saturation, contrast, and shifts colour values toward "skylight" (typically blue).

Image source: Wikipedia

Lighting and Shading

The pattern of interaction with light sources and resultant shadows gives a great deal of information about three dimensional form and structure.

Image source: Martin Spilker

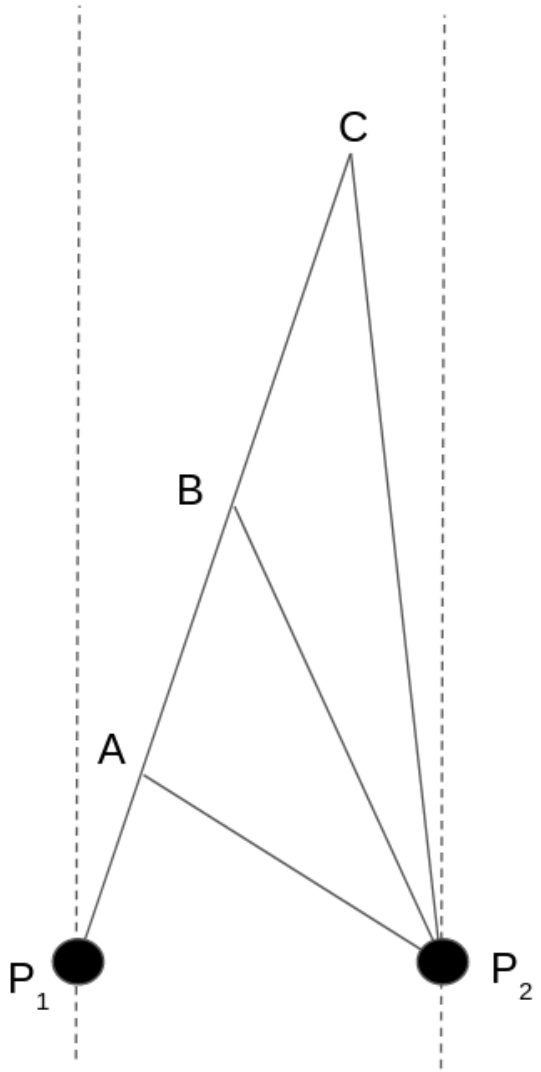
Focus

For a given depth of field and focus plane, the sharpness of objects provides a depth cue. However, out of focus typically doesn't provide a direction, which is why focus is typically viewed as a multi-image cue (taking multiple images with different planes of focus provides a directional gradient to the change of focus for a given object).

Image source: Abuolaim *et al.*, 2018

Basic Geometry of Parallax

Regardless of the source of displacement between positions P_1 and P_2 (motion or binocularity), the corresponding apparent shift in angle of an object depends on its depth from the observer.



Parallax Examples - Driving

- Side view
- Front view

What sets parallax apart?

As a depth cue, parallax is exceedingly powerful because it allows us to not only compute the relative distance between scene elements, but, so long as we know the displacement between viewpoints (the *baseline* in a stereo system), we can compute *actual* distance.

Calibration

- Calibration is conducted using a target pattern with known physical measurements

Calibration

- Calibration is conducted using a target pattern with known physical measurements
- Most common pattern is a checkerboard (why this shape?)

Calibration

- Calibration is conducted using a target pattern with known physical measurements
- Most common pattern is a checkerboard (why this shape?)
- Calibration video example

A Note About Stereo Vision in Humans

Human eyes are anisotropic; we typically want our foveas trained on a target of interest, but the required relative position of the foveas in both eyes changes with target depth, requiring compensatory movements (*vergence*)

A Note About Stereo Vision in Humans

Human eyes are anisotropic; we typically want our foveas trained on a target of interest, but the required relative position of the foveas in both eyes changes with target depth, requiring compensatory movements (*vergence*)

In augmented reality and virtual reality systems, there is typically a stereoscopic depth cue that is nevertheless always present in a single image plane (the screen or projection surface). Thus, the plane of focus remains constant (no need for the eye to *accomodate*), but the sterescopic need for vergence remains, which can lead to a disorienting or uncomfortable viewing experience.

The Curious Case of the Chameleon

Image source: First hit from Google search for "Face".

Image source: Original image source unknown.

Chameleon Depth Calculation

- Chameleons detect depth primarily from accomodation (changing the plane of focus) - Harkness, 1977, Ott *et al.*, 1998

Chameleon Depth Calculation

- Chameleons detect depth primarily from accomodation (changing the plane of focus) - Harkness, 1977, Ott *et al.*, 1998
- Chameleons still use two eyes to make this calculation, but these measures are independent (and are then combined to reduce measure noise)

Chameleon Depth Calculation

- Chameleons detect depth primarily from accomodation (changing the plane of focus) - Harkness, 1977, Ott *et al.*, 1998
- Chameleons still use two eyes to make this calculation, but these measures are independent (and are then combined to reduce measure noise)
- There is debate as to why chameleons do this

Chameleon Depth Calculation

- Chameleons detect depth primarily from accomodation (changing the plane of focus) - Harkness, 1977, Ott *et al.*, 1998
- Chameleons still use two eyes to make this calculation, but these measures are independent (and are then combined to reduce measure noise)
- There is debate as to why chameleons do this
  - Evolutionary holdover (monocular depth perception probably evolved first)

Chameleon Depth Calculation

- Chameleons detect depth primarily from accomodation (changing the plane of focus) - Harkness, 1977, Ott *et al.*, 1998
- Chameleons still use two eyes to make this calculation, but these measures are independent (and are then combined to reduce measure noise)
- There is debate as to why chameleons do this
  - Evolutionary holdover (monocular depth perception probably evolved first)
  - Independently moving eyes makes the baseline variable and noisier