ISTA 352

Lecture 4

What is seeing?

Administrivia

Lecture slides and videos are now posted in one convenient table

HW1 start soon.

Next Friday

- We will continue our linear algebra tutorial
- Review expressing vectors with respect to an arbitrary coordinate system
- Focus on matrices for representing transformations
- I will assume that everyone is up to speed on
  - Basic representations
    - Coordinate systems, points in space represented by vectors
  - Basic algebraic operations
    - Adding, subtracting, and multiplying vectors and matrices
      - Associativity and non-commutativity

What is seeing?

- Do bats see? Blind people? Babies? How well?
- Do you need to learn to see? Learn what?
- Can a computer program see?
- If you think computers could see, how would we test one to find out if it sees?

More on seeing

- If we think of images as a spatial message, then seeing is interpreting that message.
- Most natural case is that the message is spatially indexed data (perhaps from eyes) about the world
  - Seeing is then inferring spatial and semantic representations about the world from such data
  - Computer vision (seeing by computers) is using images as evidence for the particulars of a world representation

According to Marr

" …, vision is the process of discovering from images what is present in the world, and where it is.

… our brains must be capable of representing this information … "

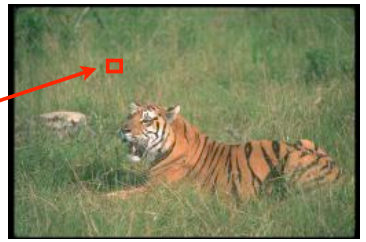
Marr 82, page 3.

More on seeing (2)

- Notice that the output of the process is a **different** kind of information
  - Images goes in, but the result is not an image
    - Contrast this with visualization, enhancement, and image processing
  - There is a notion of "translation" here
  - There is also an analogy with a scientific explanations of phenomena by reducing them to more fundamental constructs
    - e.g., surface color in terms of microscopic constructs
- Interpreting human constructed spatial messages requires learning the language
  - Art, maps, even television and photographs

Translation metaphor

(50,123,24)



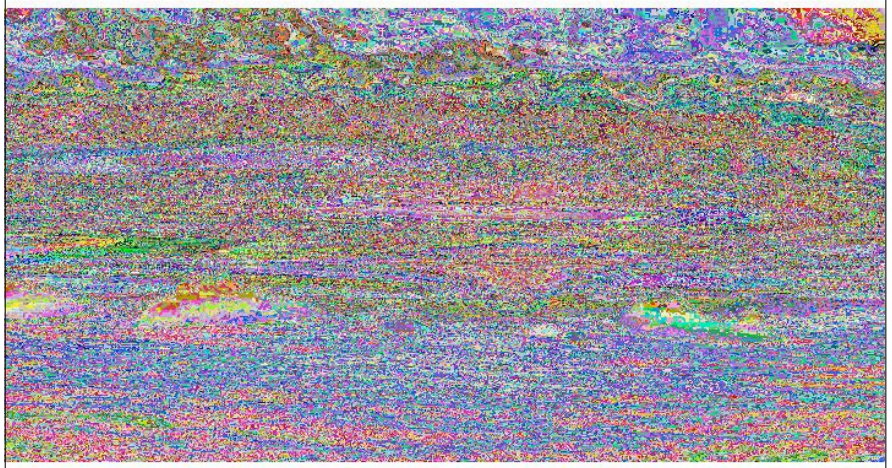
Entire image is a sequence of 256x384 triples of integers between 0 and 255.

Translation metaphor

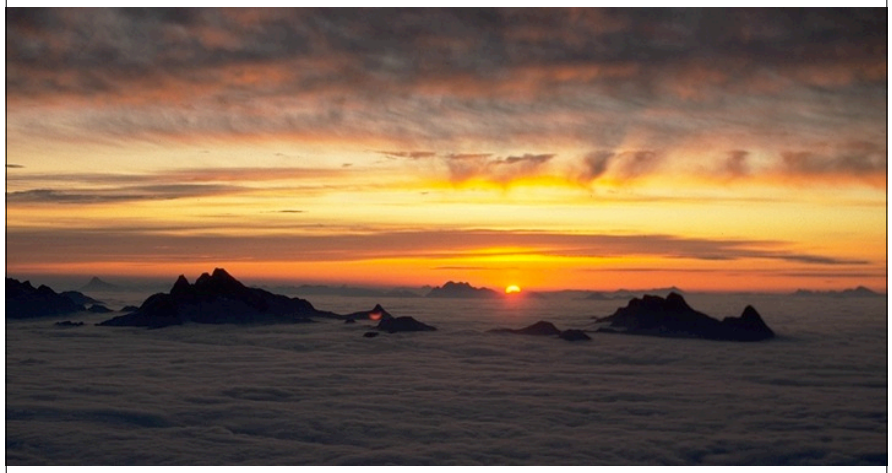
Your algorithm goes here

A tiger lying in the grass

Demo one (another mystery image)

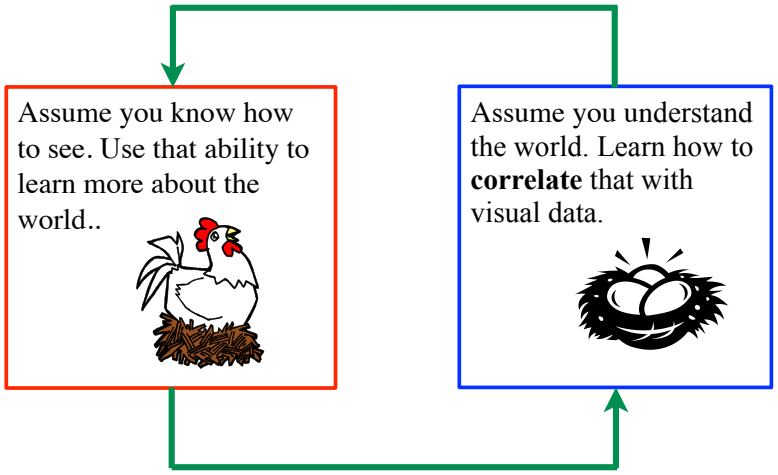


Mystery revealed



Learning to see?

- Need to learn about the world and its representation
  - Details about the representation in the brain are not well understood
- Need to learn how to connect the above with visual data
- But how do you learn about the world, if you cannot see?
  - We need to learn these together

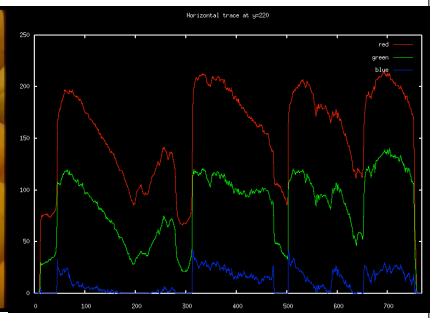


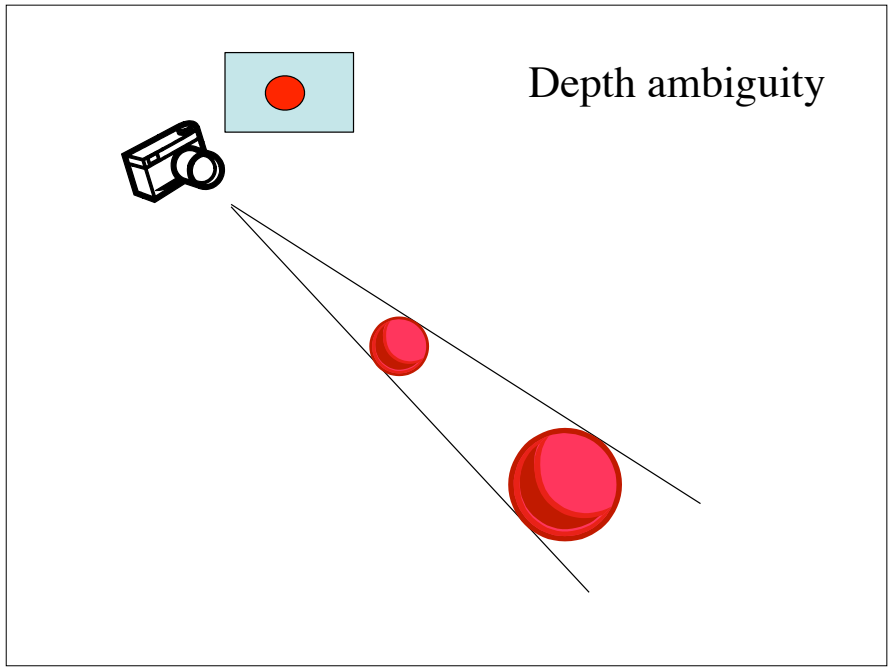
Ambiguity is everywhere

Ambiguity is everywhere

Do the apricots look spherical?

Ambiguity is everywhere




Depth ambiguity