What is (computer) vision?

“ …, 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.

Visual Representation → Semantic Representation

→ A tiger lying in the grass

Is vision computational?

Why study Computer Vision?

- Images and movies are everywhere
- Fast-growing collection of useful applications
  - building representations of the 3D world from pictures
  - automated surveillance (who’s doing what)
  - movie post-processing
  - face finding
- Various deep and attractive scientific mysteries
  - how does object recognition work?
- Greater understanding of human vision

More Applications

Image and video retrieval and data mining

Robotics

- Defect spotting
- Driving aids, autonomous flight
- Surveillance, identification

Graphics, Virtual Reality, Printing

Properties of Vision

- One can "see the future"
  - Cricketers avoid being hit in the head
    - There's a reflex --- when the right eye sees something going left, and the left eye sees something going right, move your head fast.
  - Gannets pull their wings back at the last moment
    - Gannets are diving birds; they must steer with their wings, but wings break unless pulled back at the moment of contact.
    - Area of target over rate of change of area gives time to contact.

Properties of Vision

- 3D representations are easily constructed
  - There are many different cues.
  - Useful
    - to humans (avoid bumping into things; planning a grasp; etc.)
    - in computer vision (build models for movies).
  - Cues include
    - multiple views (motion, stereopsis)
    - texture
    - shading

Properties of Vision

- People draw distinctions between what is seen
  - "Object recognition"
  - This could mean "is this a fish or a bicycle?"
  - It could mean "is this George Washington?"
  - It could mean "is this poisonous or not?"
  - It could mean "is this slippery or not?"
  - It could mean "will this support my weight?"
  - Great mystery
    - How to build programs that can draw useful distinctions based on image properties.

Vision versus Graphics

Vision

images --> model of the world

Graphics

model of the world --> images

Vision Systems

Biological

eye + brain

Man made

camera + computer

Computer Vision Research

Part of Artificial Intelligence

Connected to
- cognitive psychology
- perceptual psychology
- robotics
- databases
- imaging science

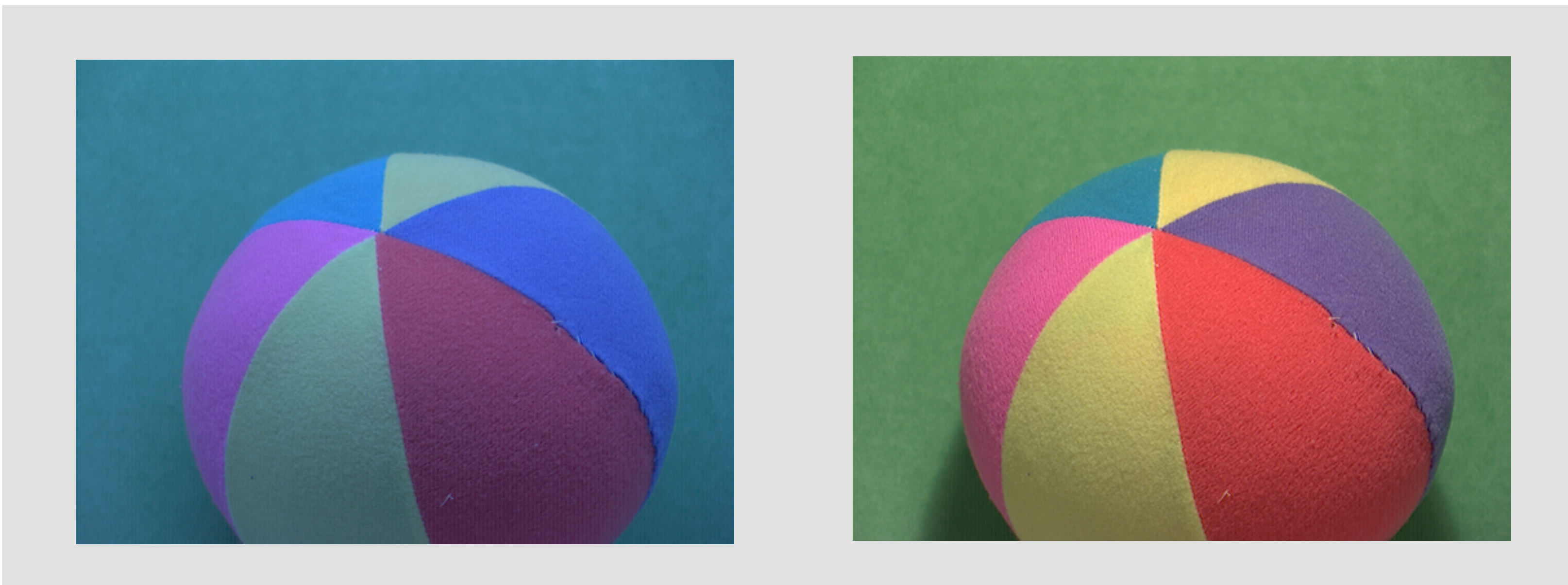
Computer Vision History

(Nope, it does not work yet)

Sample Problems

Low level vision
- Shape from shading
- Stereo vision
- Structure from motion
- Segmentation
- Tracking
- Colour constancy

The Computational Colour Constancy Problem

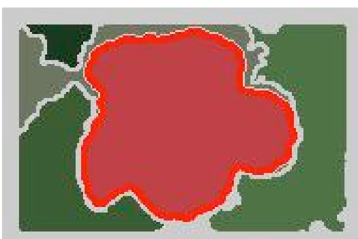


(Same scene, but different illuminant)

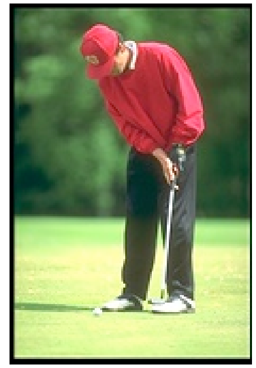
Sample Problems

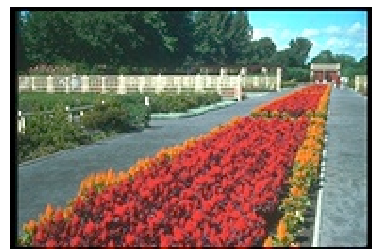
Mid level vision
- Grouping of image elements
- Face detection and identification
- Gesture identification
- Fingerprint identification
- Retrieving images by content

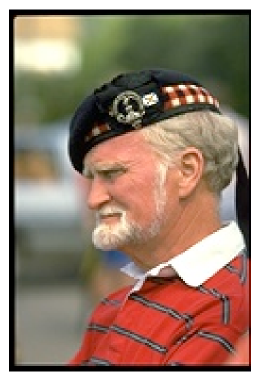
Query on

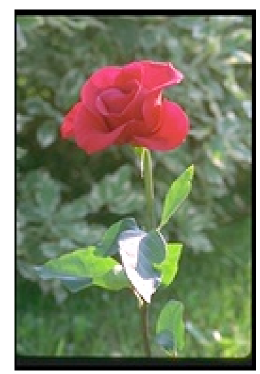


Example from Berkeley Blobworld system









Methods

Physics of image formation

Studying biological vision systems

Machine learning

Computational experimentation

Brief Course Intro
(Subject to change)

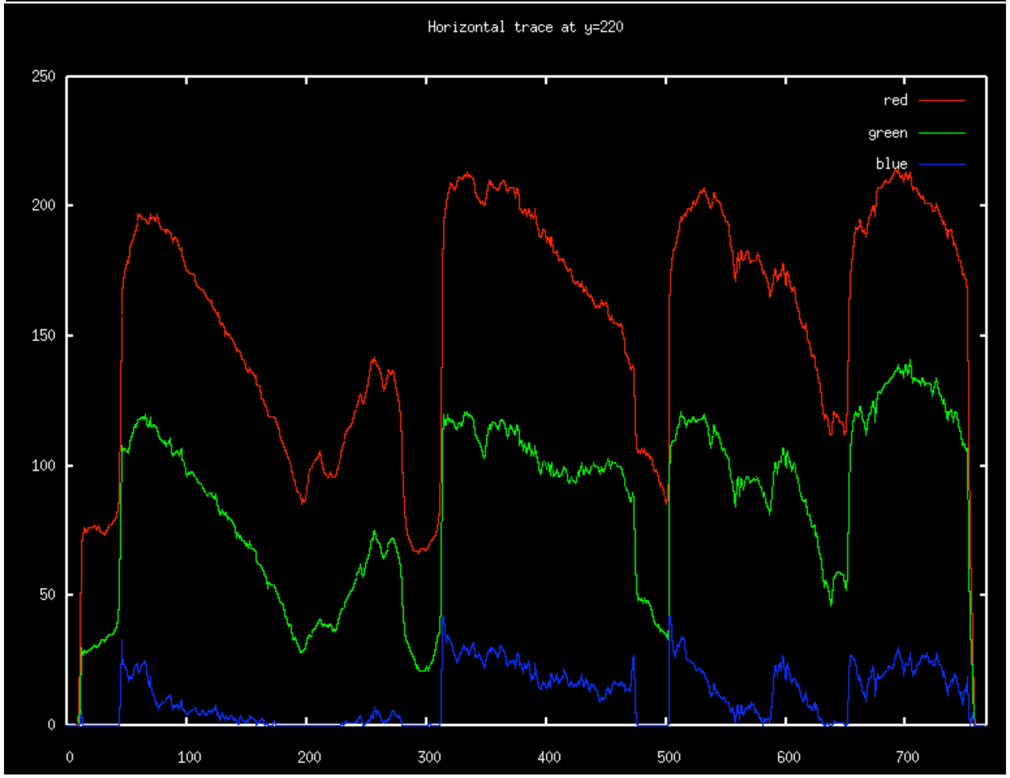
Part I: The Physics of Imaging

- How images are formed
  - Cameras
    - What a camera does
    - How to tell where the camera was
  - Light
    - How to measure light
    - What light does at surfaces
    - How the brightness values we see in cameras are determined
  - Color
    - The underlying mechanisms of color
    - How to describe it and measure it

Part II: Early Vision in One Image

- Representing small patches of image
  - We wish to establish correspondence between (say) points in different images, so we need to describe the neighborhood of the points
  - Sharp changes are important in practice --- known as “edges”
  - Representing texture by giving some statistics of the different kinds of small patch present in the texture.
    - Tigers have lots of bars, few spots
    - Leopards are the other way

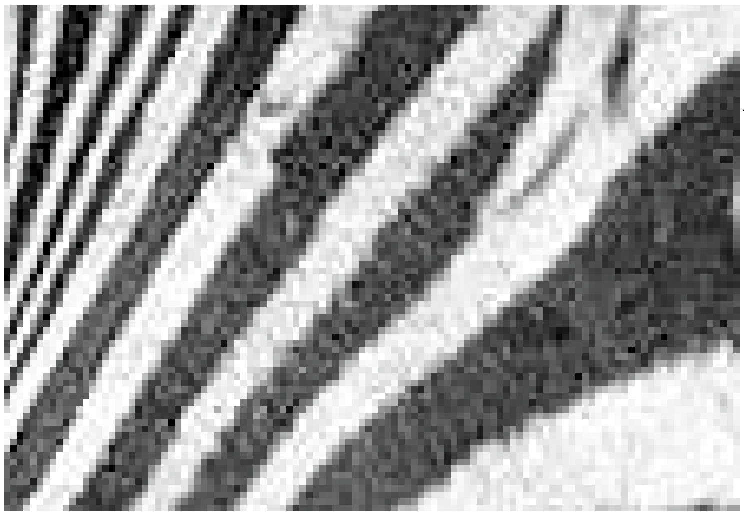
Shading Cues



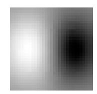
Representing an image patch

- Filter outputs
  - essentially form a dot-product between a pattern and an image, while shifting the pattern across the image
  - strong response -> image locally looks like the pattern
  - e.g. derivatives measured by filtering with a kernel that looks like a big derivative (bright bar next to dark bar)

Convolve this image

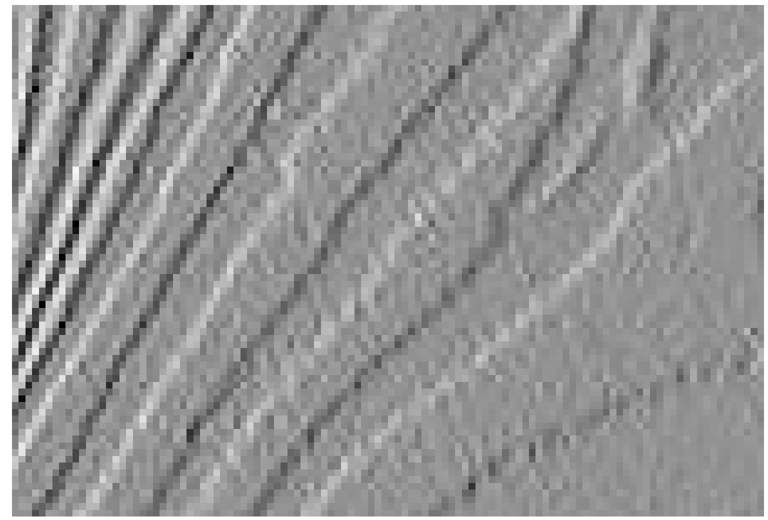


With this kernel



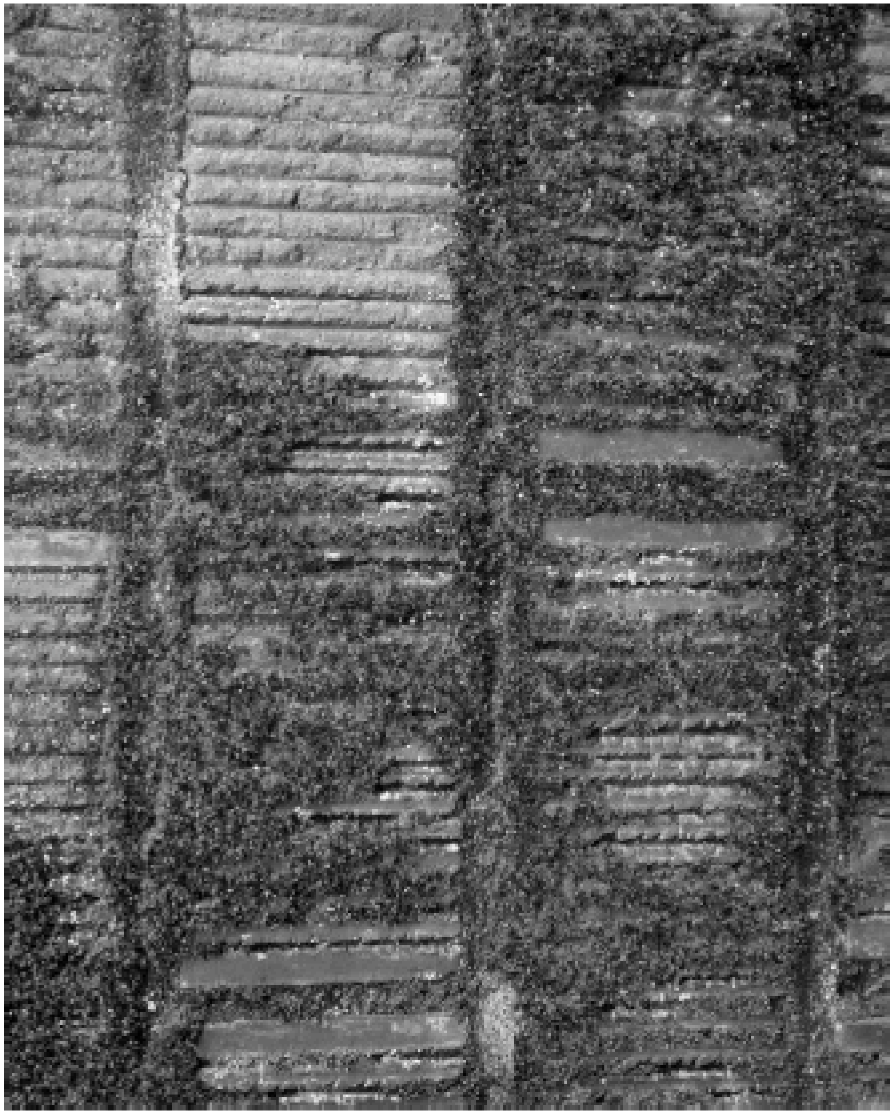
→

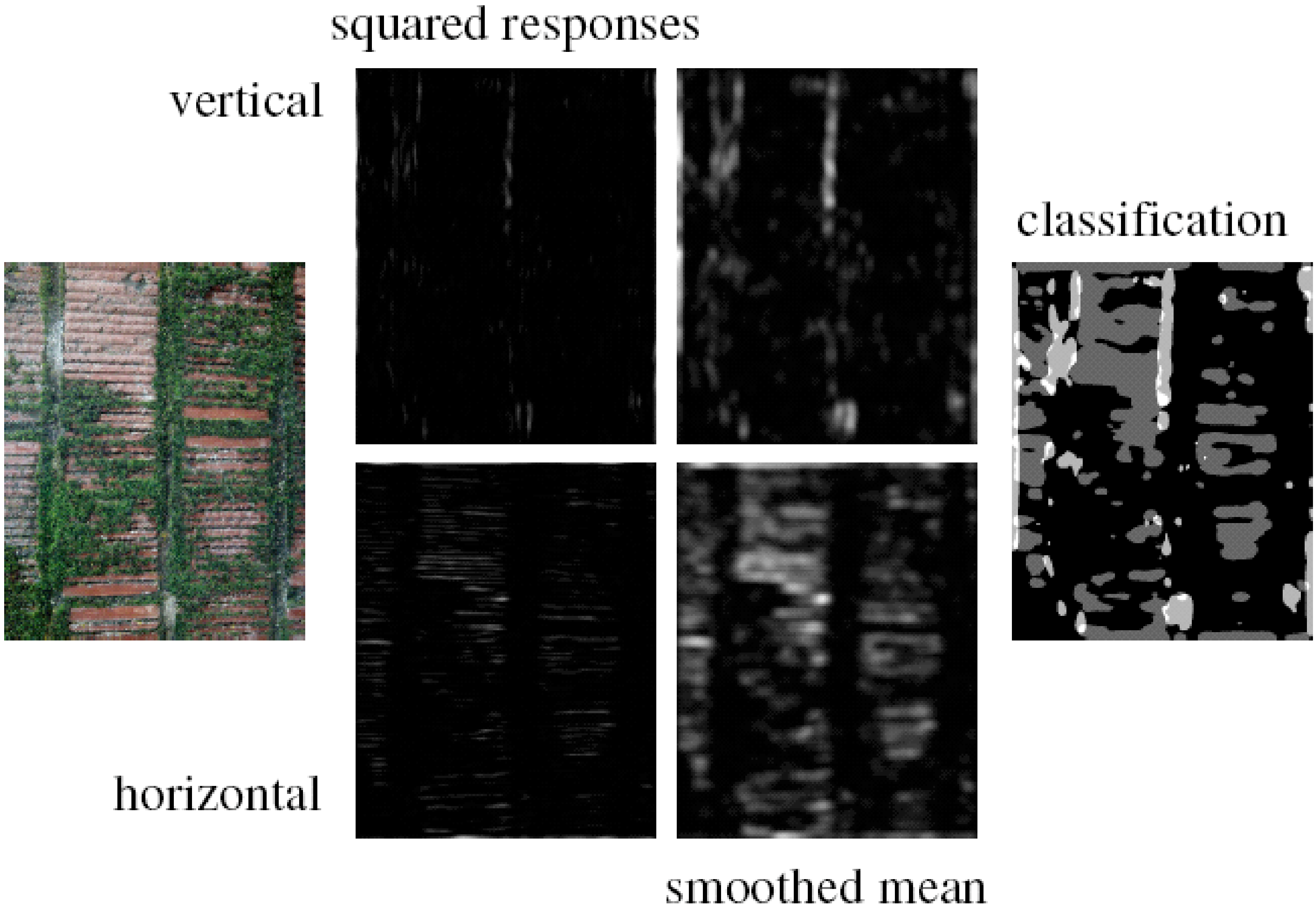
To get this



Texture

- Many objects are distinguished by their texture
  - Tigers, cheetahs, grass, trees
- We represent texture with statistics of filter outputs
  - For tigers, bar filters at a coarse scale respond strongly
  - For cheetahs, spots at the same scale
  - For grass, long narrow bars
  - For the leaves of trees, extended spots
- Objects with different textures can be segmented
- The variation in textures is a cue to shape





squared responses
vertical
classification
horizontal
smoothed mean

Shape from texture