Computer Vision

What is Computer Vision?

- Input: images or video
- Output: description of the world

What is Computer Vision?

- Input: images or video
- Output: description of the world
  - Many levels of description

Low-Level or “Early” Vision

- Considers local properties of an image

“There’s an edge!”

Mid-Level Vision

- Grouping and segmentation

"There's an object and a background!"

High-Level Vision

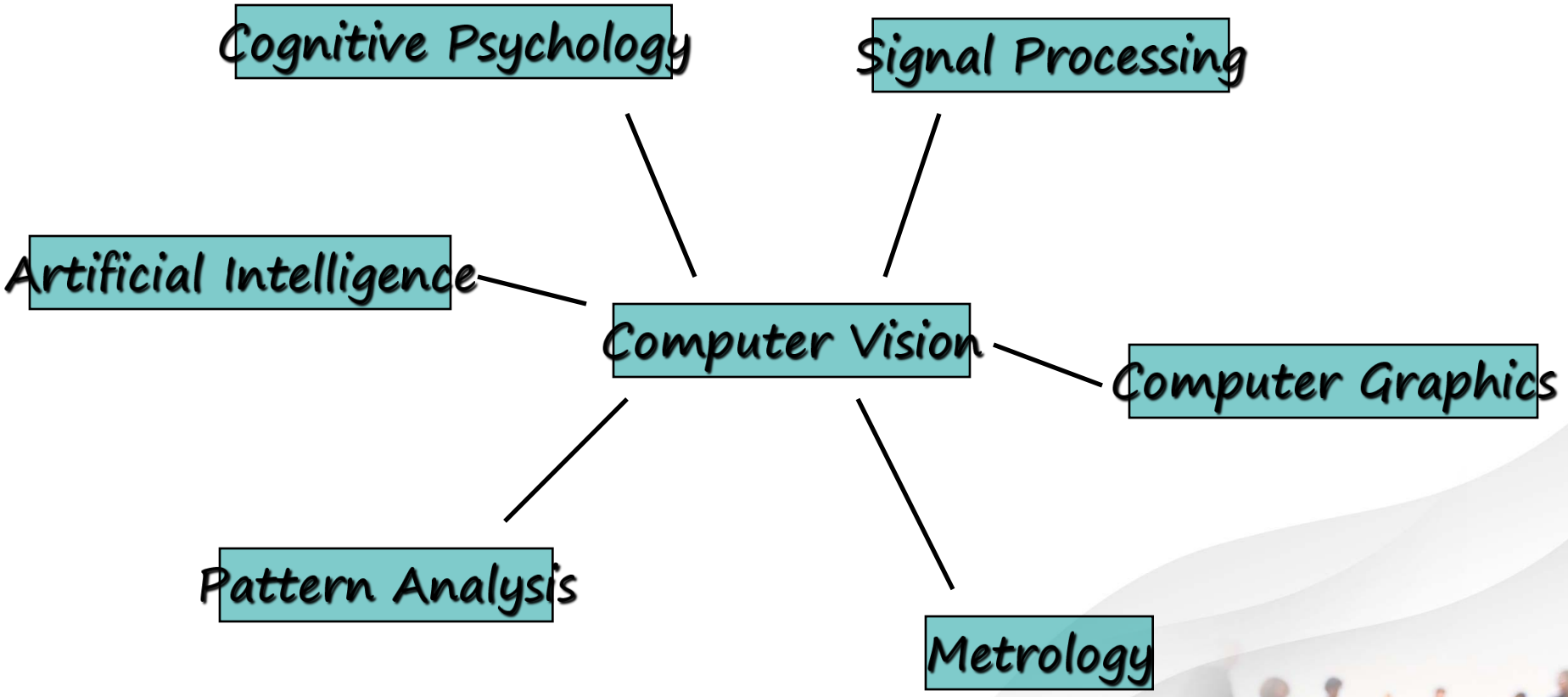
"It's a chair!"

- Recognition

Big Question #1: Who Cares?

- Applications of computer vision
  - In AI: vision serves as the “input stage”
  - In medicine: understanding human vision
  - In engineering: model extraction

Vision and Other Fields



Big Question #2: Does It Work?

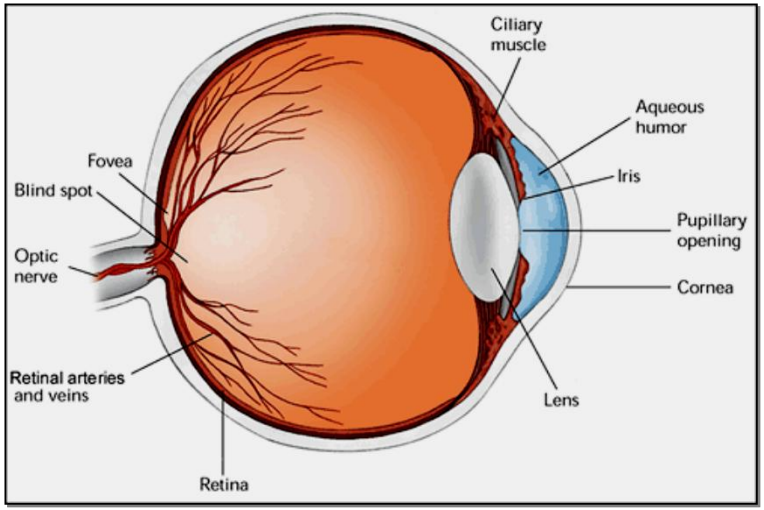
- Situation much the same as AI:
  - Some fundamental algorithms
  - Large collection of hacks / heuristics
- Vision is hard!
  - Especially at high level, physiology unknown
  - Requires integrating many different methods
  - Requires reasoning and understanding: "AI completeness"

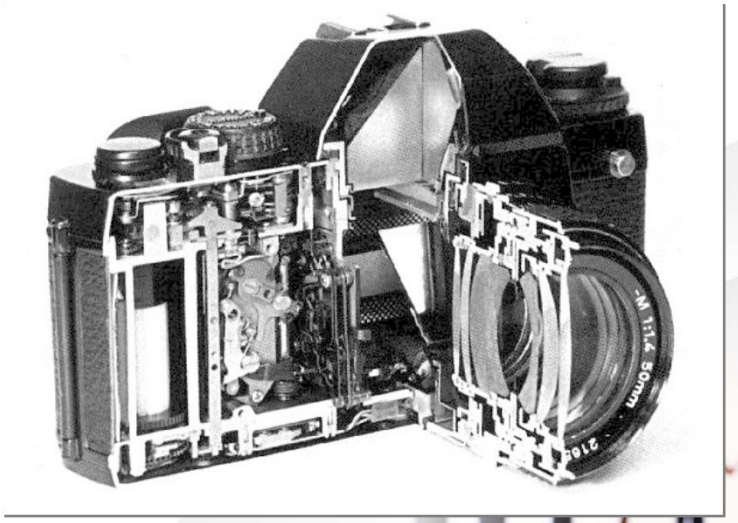
Computer and Human Vision

- Emulating effects of human vision
- Understanding physiology of human vision

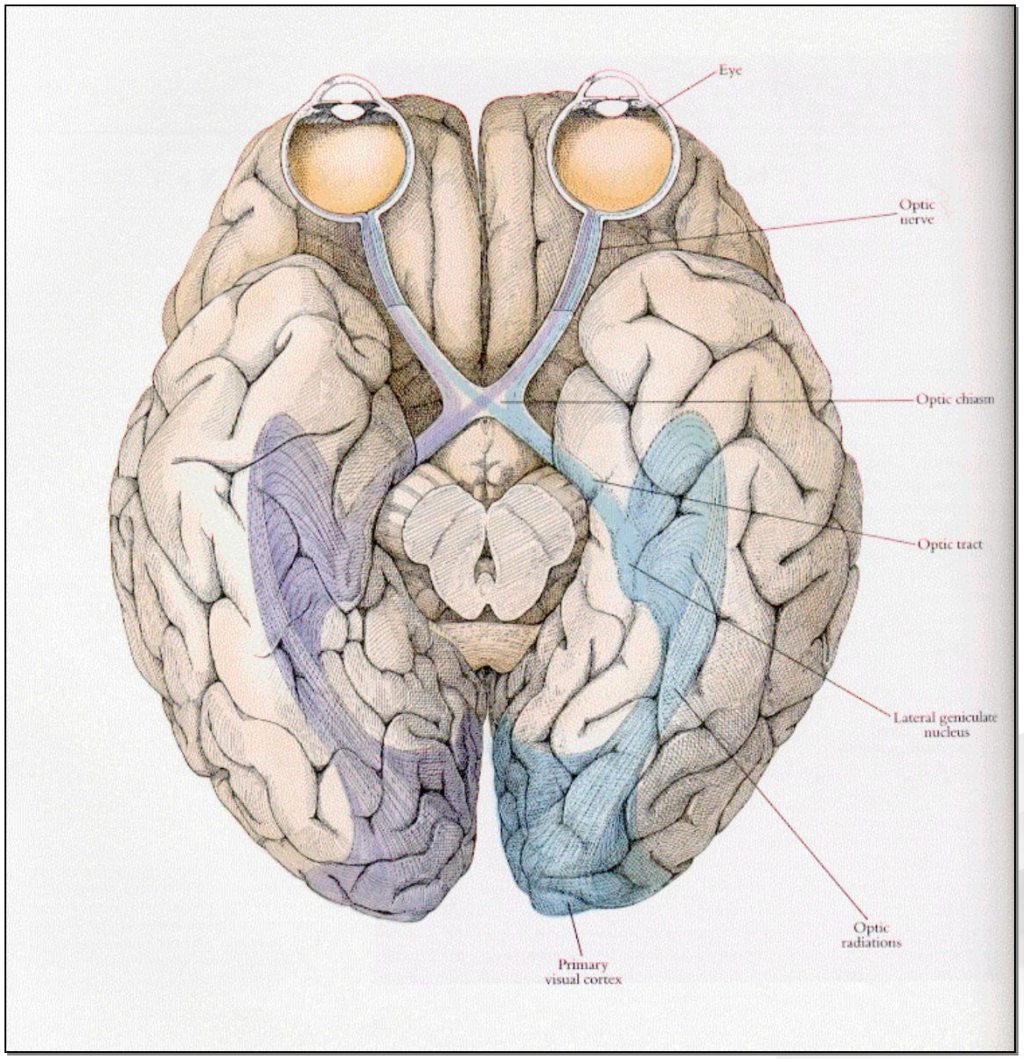
Image Formation

- Human: lens forms image on retina, sensors (rods and cones) respond to light
- Computer: lens system forms image, sensors (CCD, CMOS) respond to light





Low-Level Vision

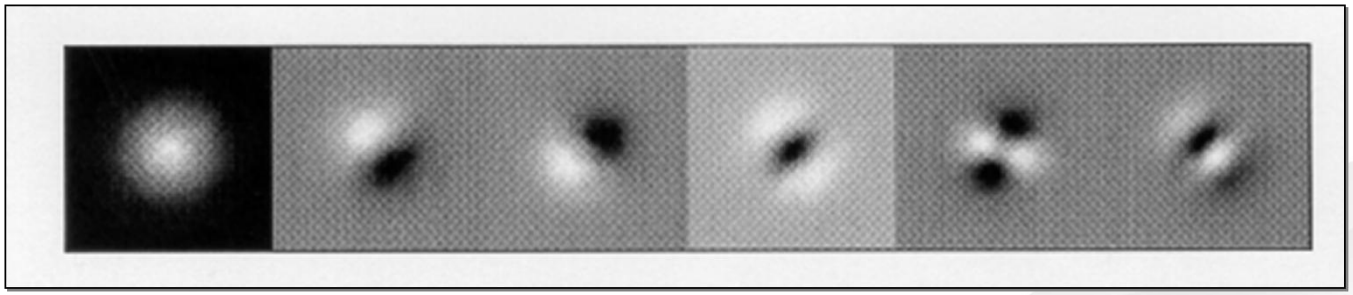


Low-Level Vision

- Retinal ganglion cells
- Lateral Geniculate Nucleus – function unknown (visual adaptation?)
- Primary Visual Cortex
  - Simple cells: orientational sensitivity
  - Complex cells: directional sensitivity
- Further processing
  - Temporal cortex: what is the object?
  - Parietal cortex: where is the object? How do I get it?

Low-Level Vision

- Net effect: low-level human vision can be (partially) modeled as a set of *multiresolution, oriented* filters



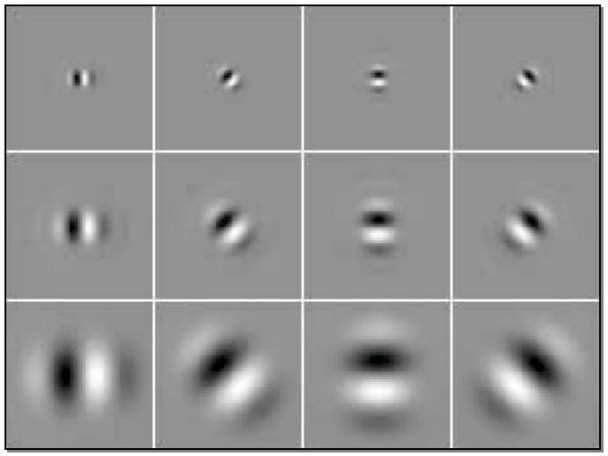
Low-Level Depth Cues

- Focus
- Vergence
- Stereo
- Not as important as popularly believed

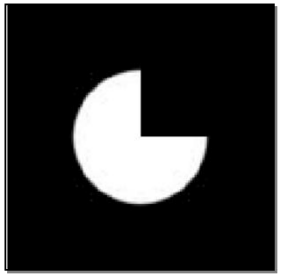
Low-Level Computer Vision

- Filters and filter banks
  - Implemented via convolution
  - Detection of edges, corners, and other local features
  - Can include multiple orientations
  - Can include multiple scales: “filter pyramids”
- Applications
  - First stage of segmentation
  - Texture recognition / classification
  - Texture synthesis

Texture Analysis / Synthesis



Multiresolution
Oriented
Filter Bank



Original
Image

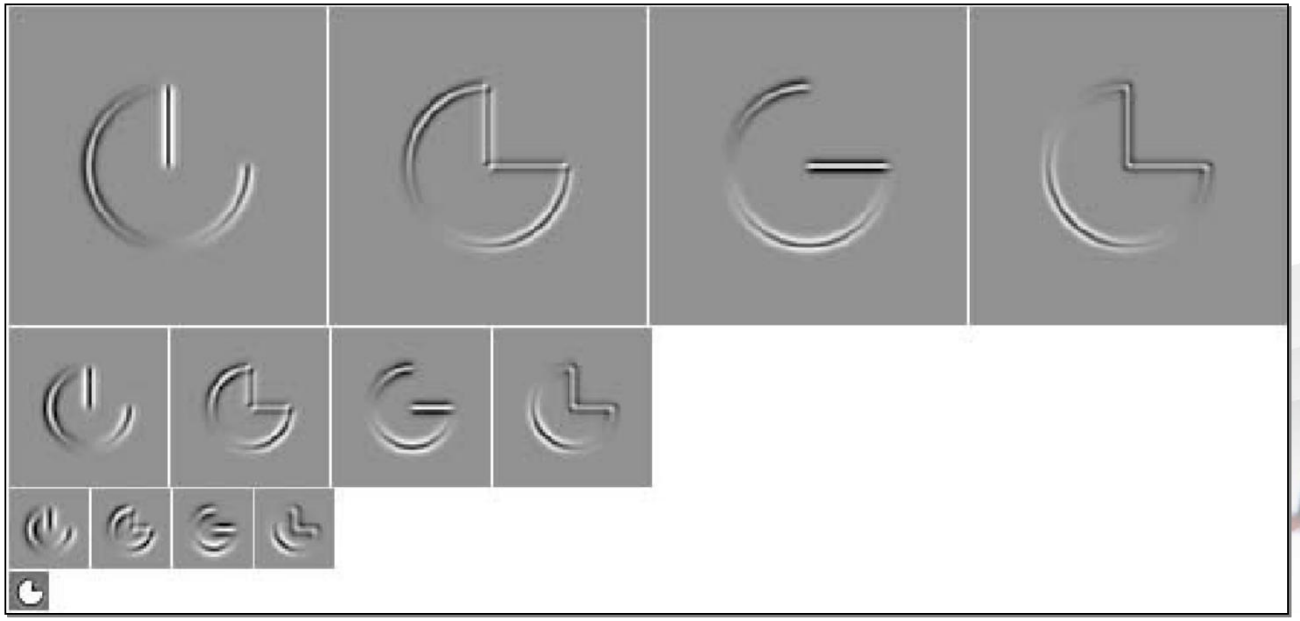
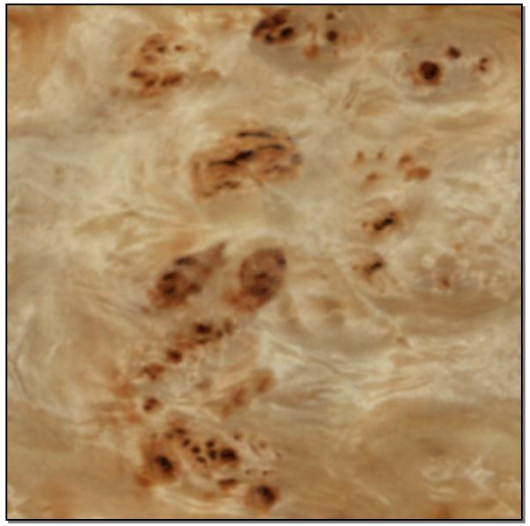
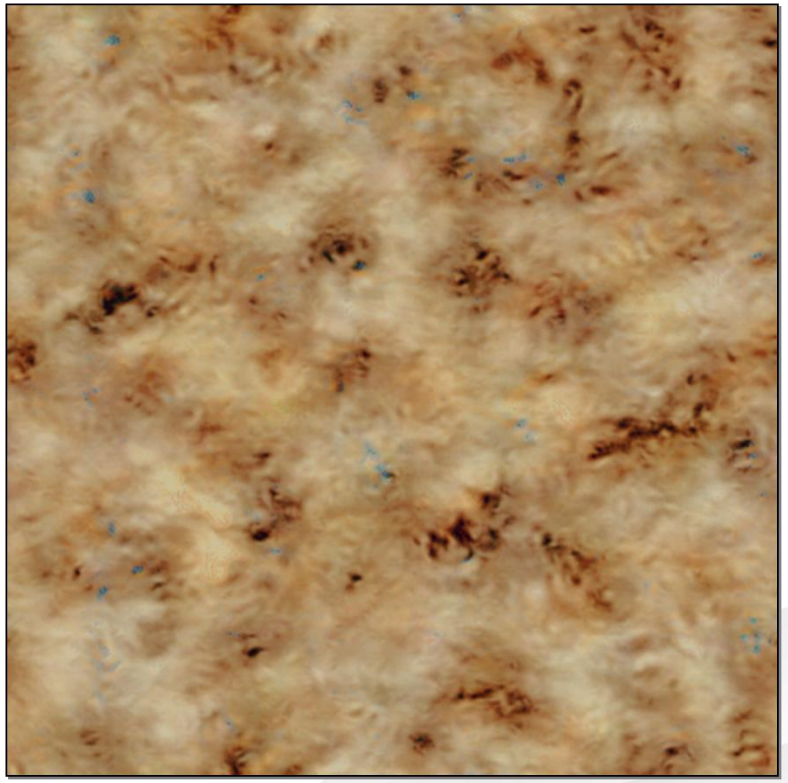


Image
Pyramid

Texture Analysis / Synthesis



Original Texture



Synthesized Texture

Low-Level Computer Vision

- Optical flow
  - Detecting frame-to-frame motion
  - Local operator: looking for gradients
- Applications
  - First stage of tracking

Optical Flow

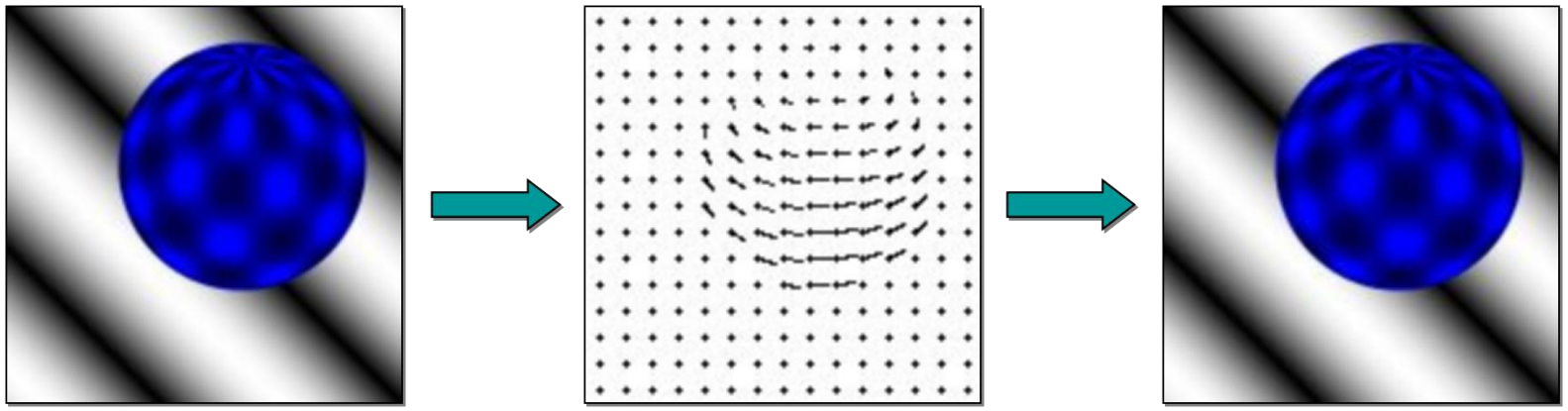


Image #1

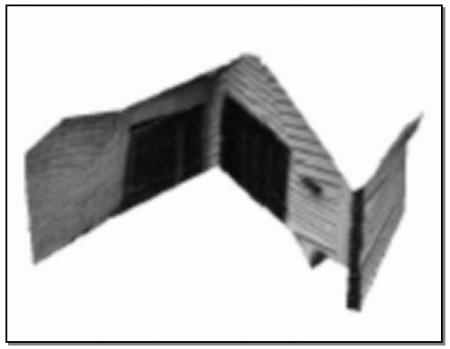
Optical Flow Field

Image #2

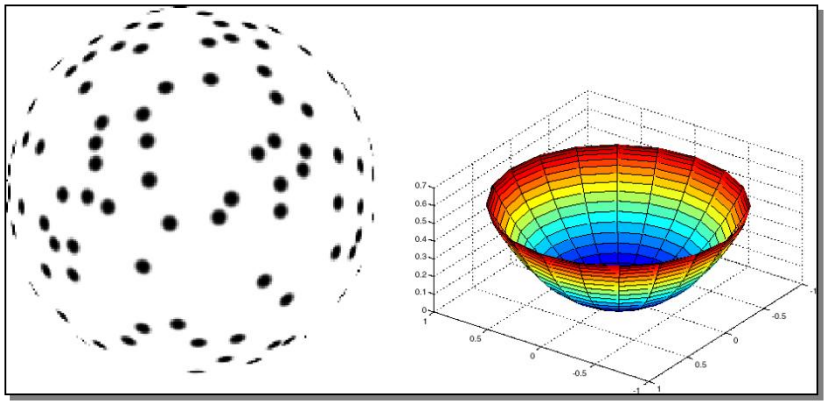
Low-Level Computer Vision

- Shape from X
  - Stereo
  - Motion
  - Shading
  - Texture foreshortening

3D Reconstruction

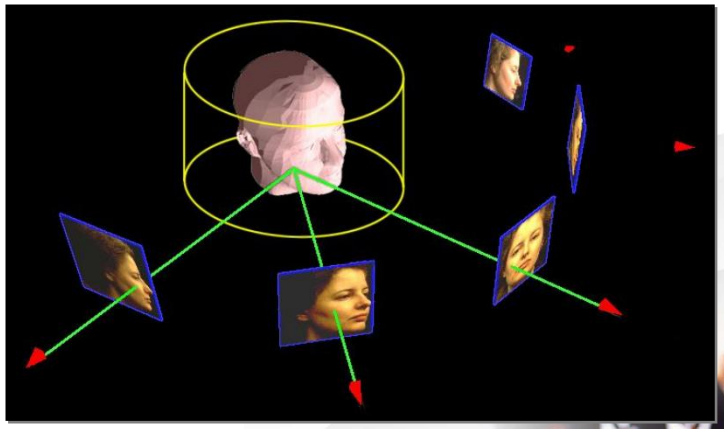


Tomasi+Kanade



Forsyth et al.

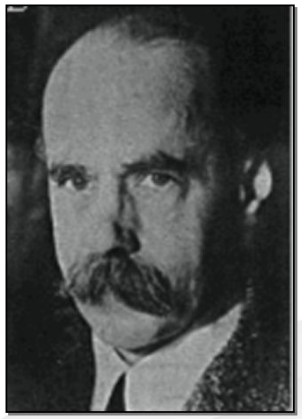
Debevec,Taylor,Malik



Phigin et al.

Mid-Level Vision

- Physiology unclear
- Observations by Gestalt psychologists
  - Proximity
  - Similarity
  - Common fate
  - Common region
  - Parallelism
  - Closure
  - Symmetry
  - Continuity
  - Familiar configuration



Wertheimer

Grouping Cues

Not grouped

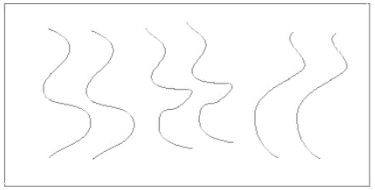
Proximity

Similarity

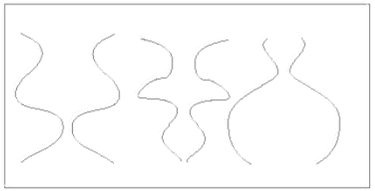
Similarity

Common Fate

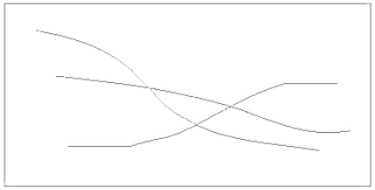
Common Region



Parallelism



Symmetry



Continuity

Closure

Mid-Level Computer Vision

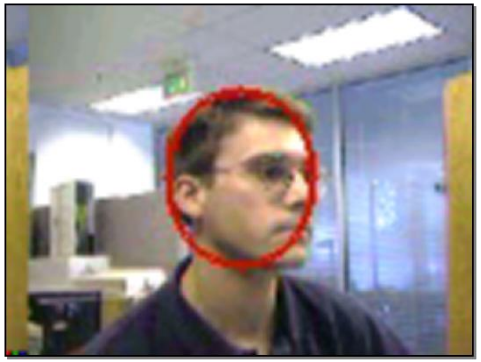
- Techniques
  - Clustering based on similarity
  - Limited work on other principles
- Applications
  - Segmentation / grouping
  - Tracking

Snakes: Active Contours

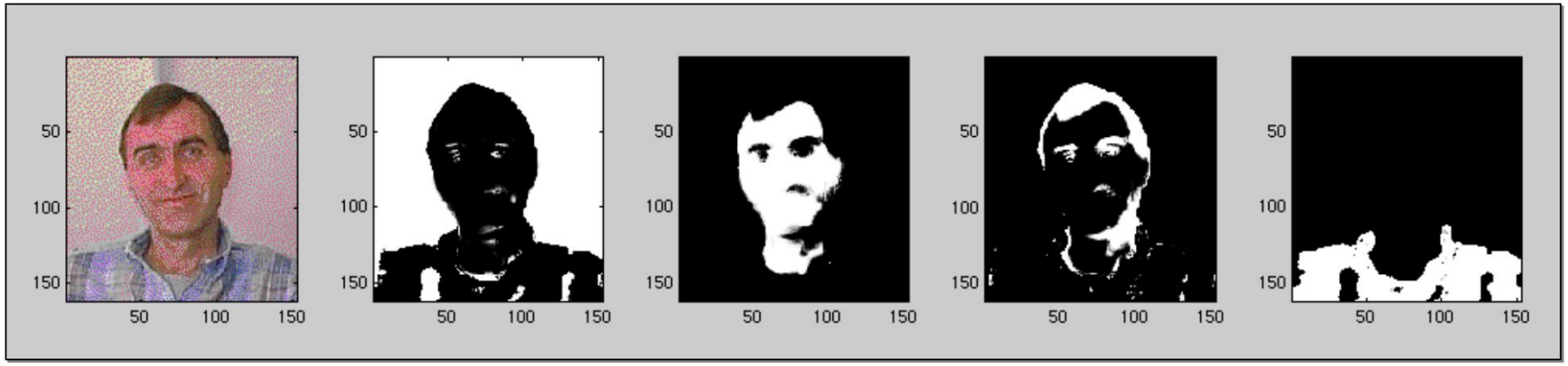
Contour Evolution for
Segmenting an Artery

Histograms





Expectation Maximization (EM)



Color Segmentation

Bayesian Methods

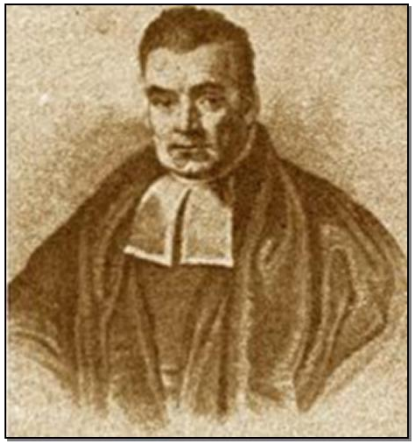
- Prior probability
  - Expected distribution of models
- Conditional probability P(A|B)
  - Probability of observation A given model B

Bayesian Methods

- Prior probability
  - Expected distribution of models
- Conditional probability P(A|B)
  - Probability of observation A given model B
- Bayes's Rule

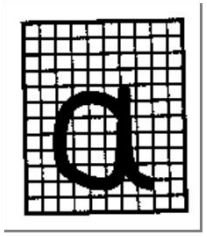
$$P(B|A) = P(A|B) \cdot P(B) / P(A)$$

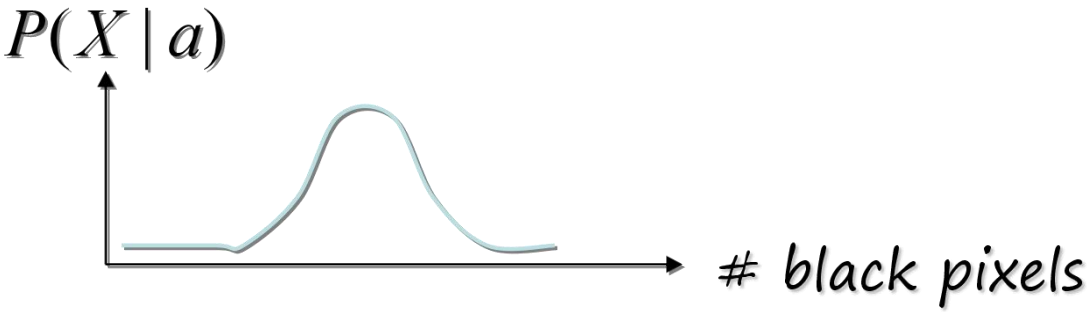
  - Probability of model B given observation A

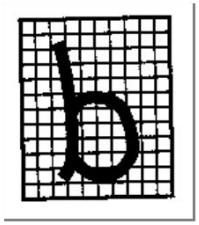


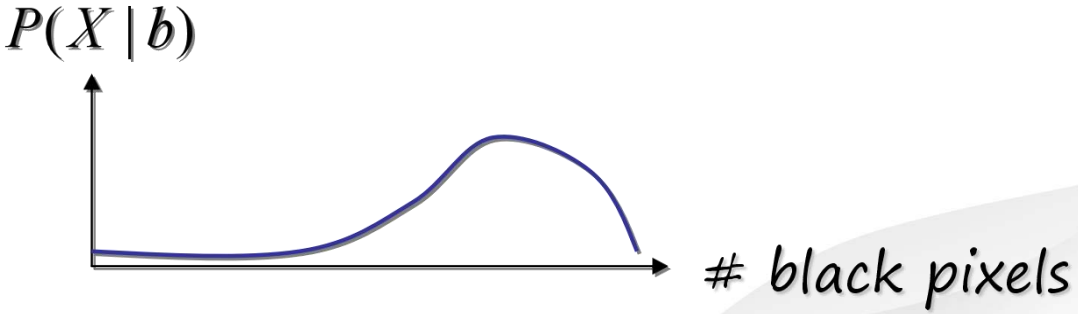
Thomas Bayes
(c. 1702-1761)

Bayesian Methods









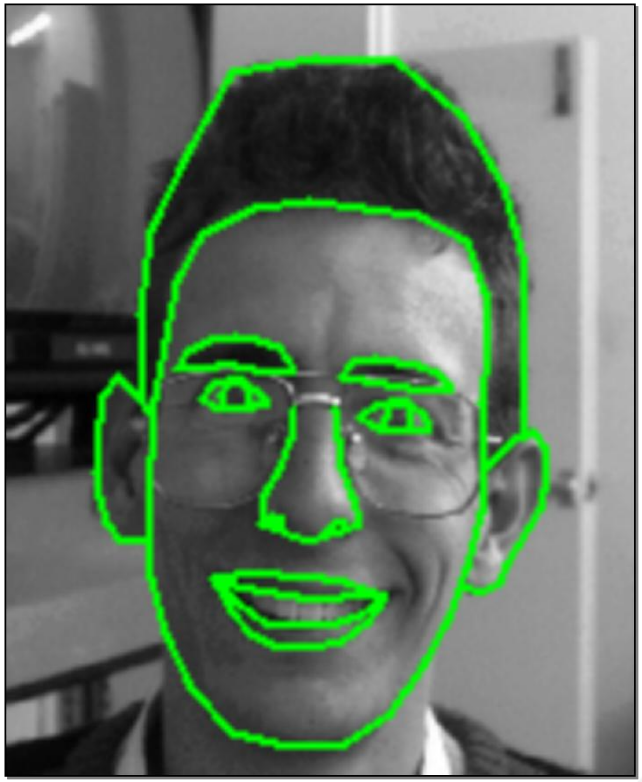
High-Level Vision

- Human mechanisms: ???

High-Level Vision

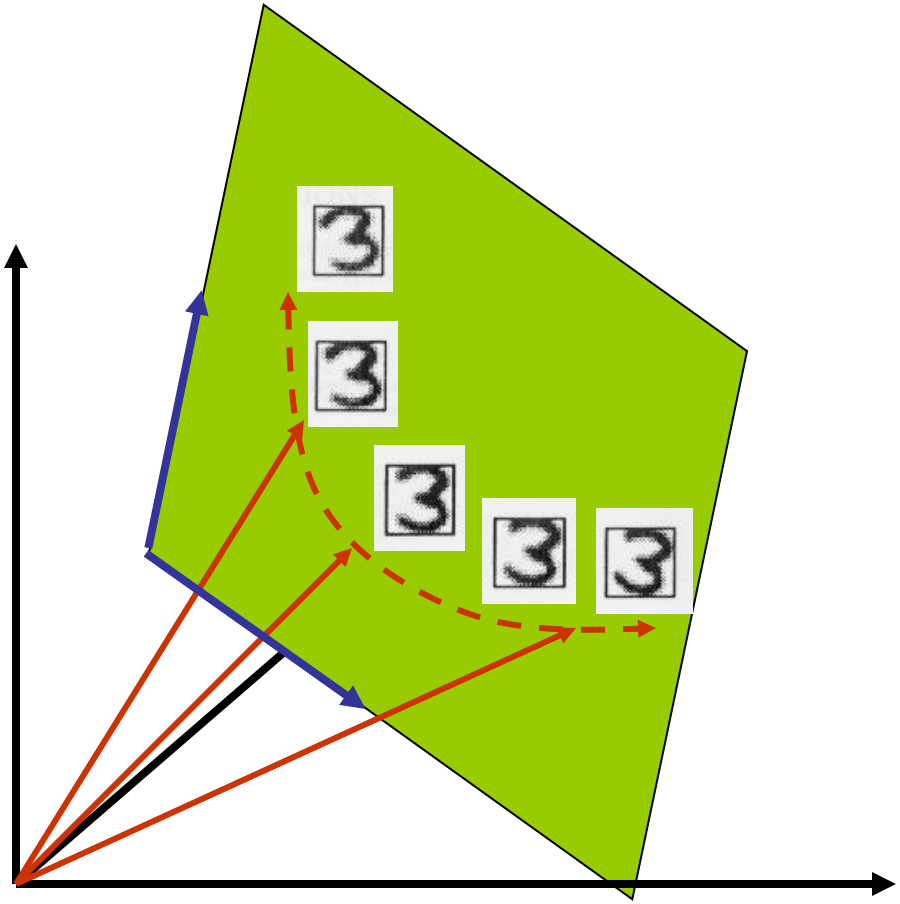
- Computational mechanisms
  - Bayesian networks
  - Templates
  - Linear subspace methods
  - Kinematic models

Template-Based Methods



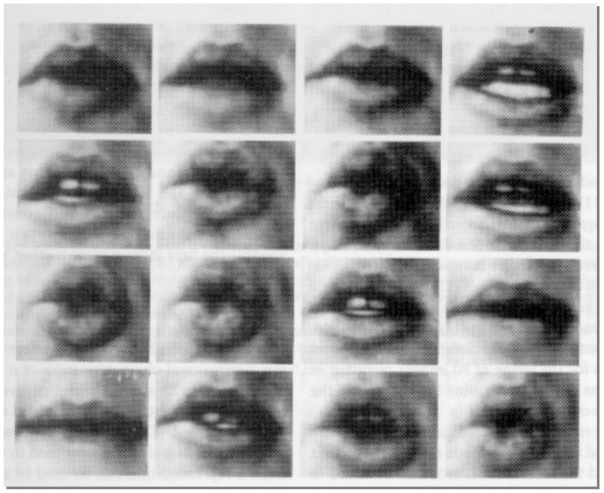
Cootes et al.

Linear Subspaces



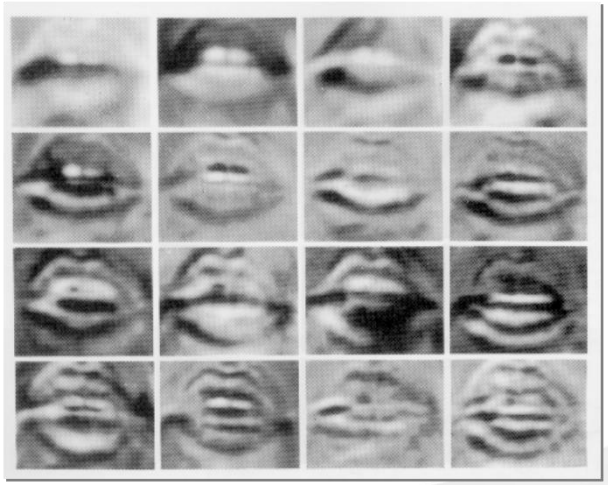
Principal Components Analysis (PCA)

Data



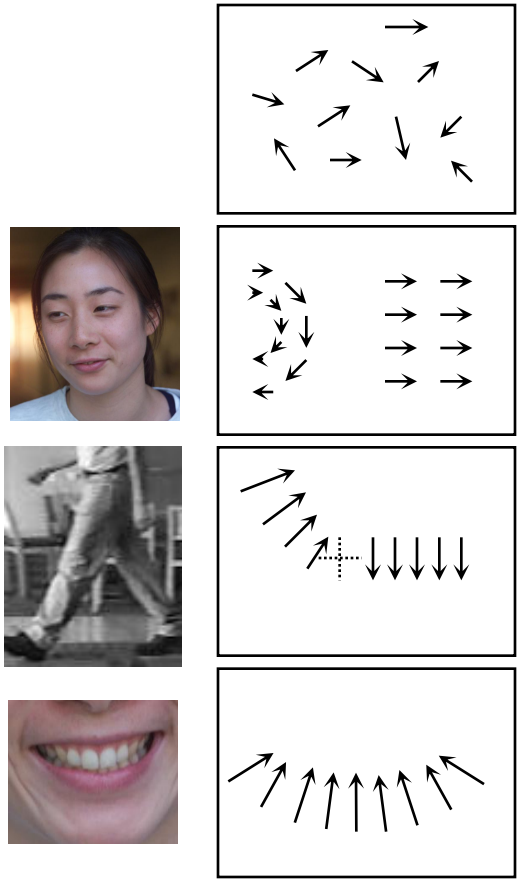
PCA

New Basis Vectors



Kirby et al.

Kinematic Models



- Optical Flow/Feature tracking: no constraints
- Layered Motion: rigid constraints
- Articulated: kinematic chain constraints
- Nonrigid: implicit / learned constraints

Real-world Applications

Osuna et al:

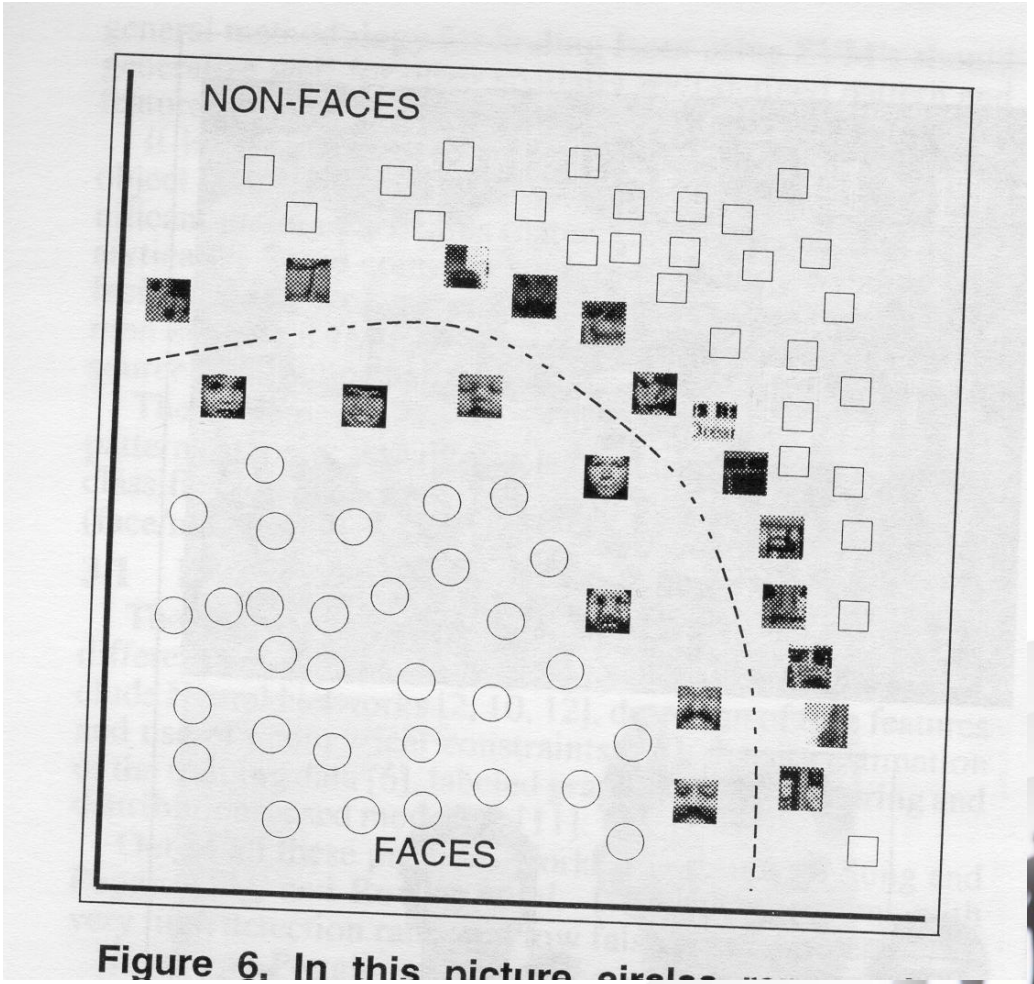


Figure 6. In this picture circles

Real-world Applications

Osuna et al:

Figure 5. Results from our Face Detection system

Course Outline

- Image formation and capture
- Filtering and feature detection
- Optical flow and tracking
- Projective geometry
- Shape from X
- Segmentation and clustering
- Recognition
- Applications: 3D scanning; image-based rendering

3D Scanning

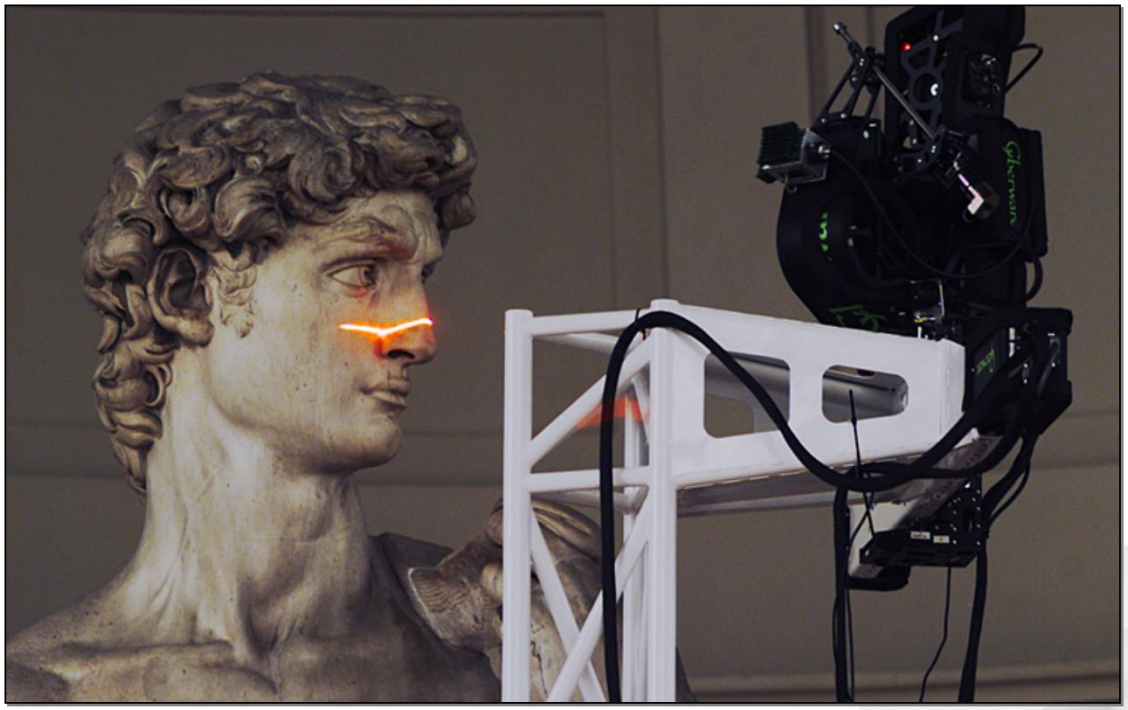
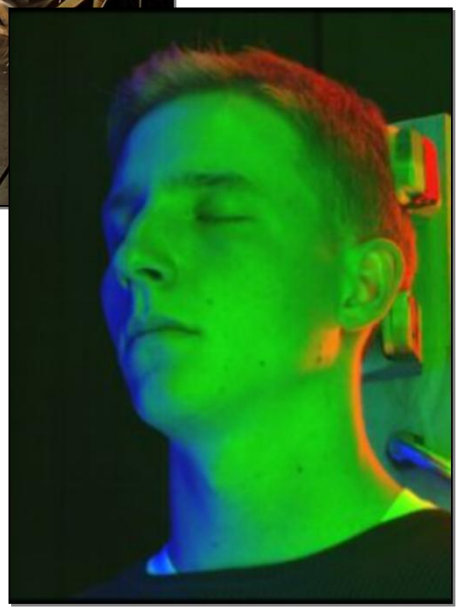


Image-Based Modeling and Rendering



Manex