Can computers see?

John MacCormick

computer vision

- WHAT is it?
- WHY is it interesting?
- HOW does it work?

what is computer vision?

Two points of view:

- “engineering view”:

  what useful tasks can be accomplished using cameras and computers?

- “AI (artificial intelligence) view”:

  to what extent can computers emulate the human visual system?

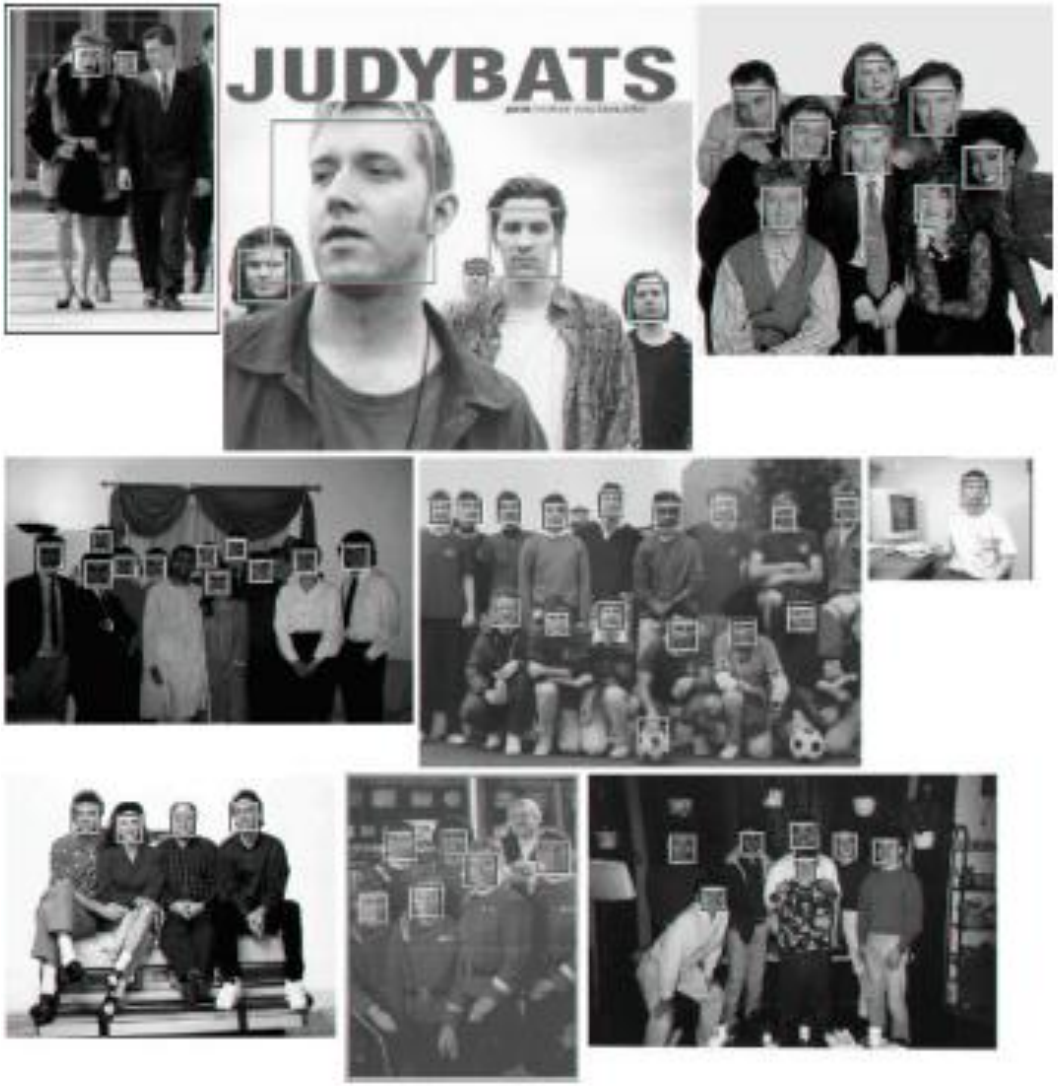
why is computer vision interesting?

two views again:

“engineering view”:

- LOTS of useful applications

face detection

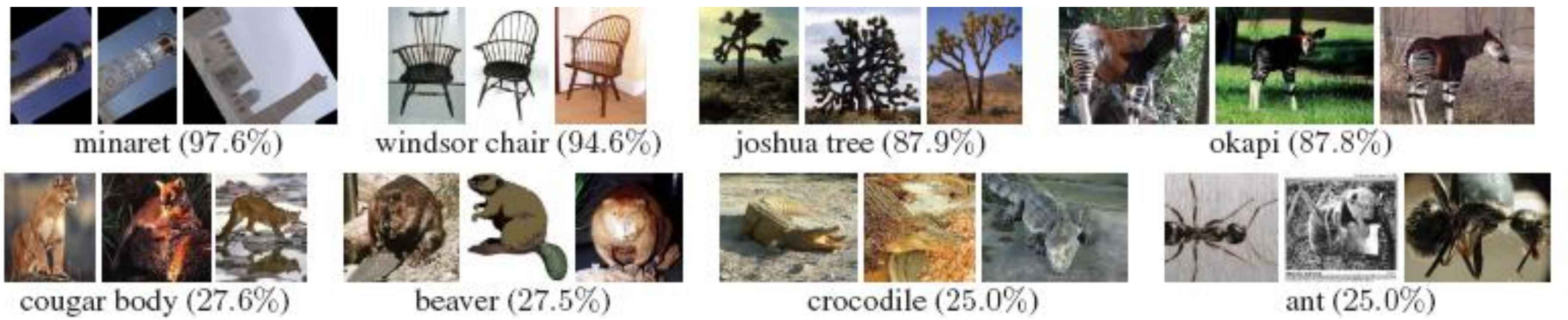


Viola and Jones 2001

face recognition

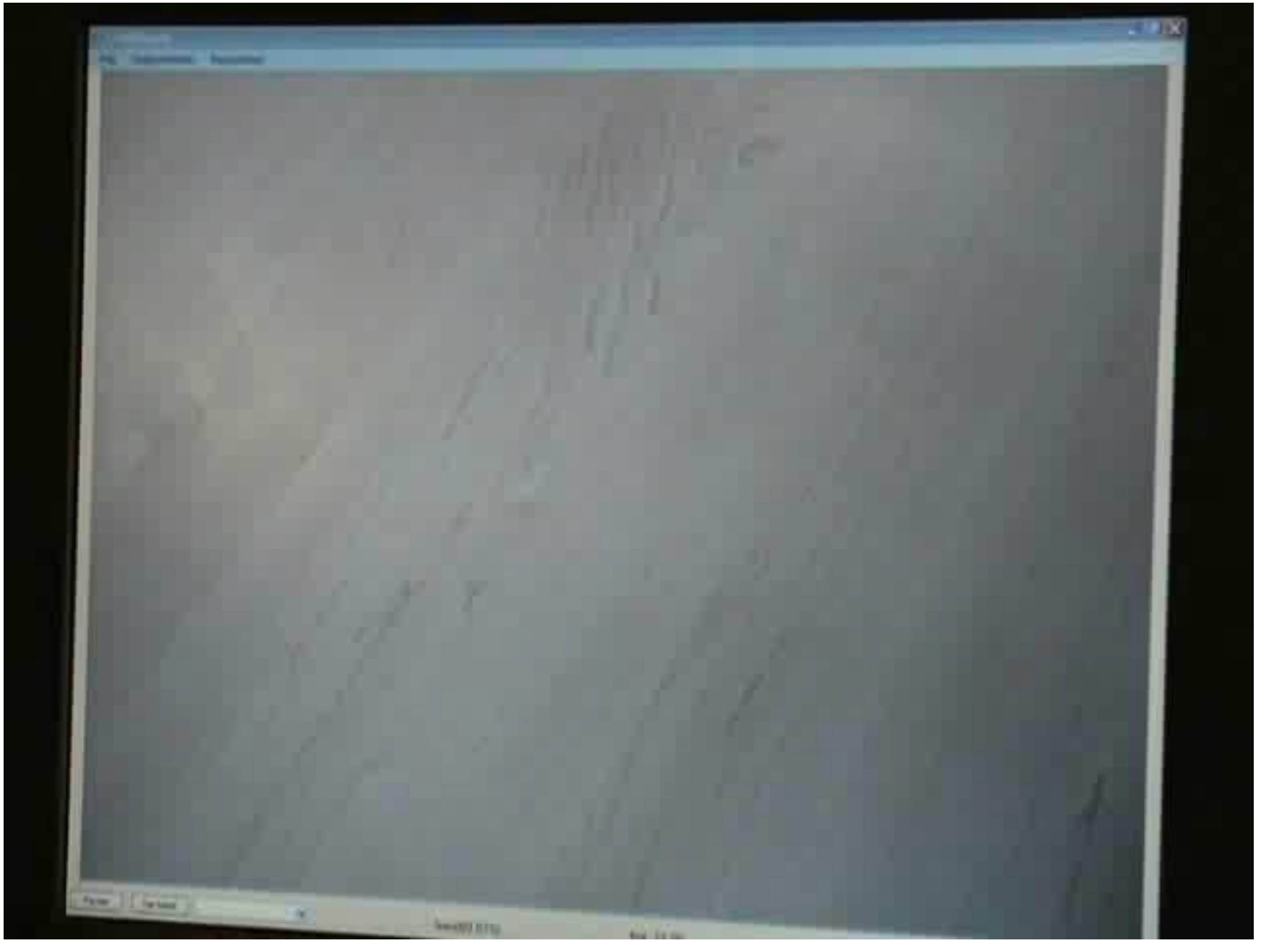
Decoro 2004

object recognition



Lazebnik et al 2006

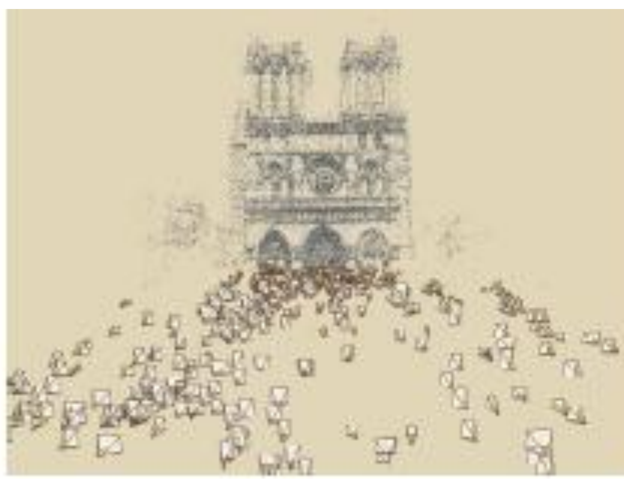
object recognition



Winn et al 2005

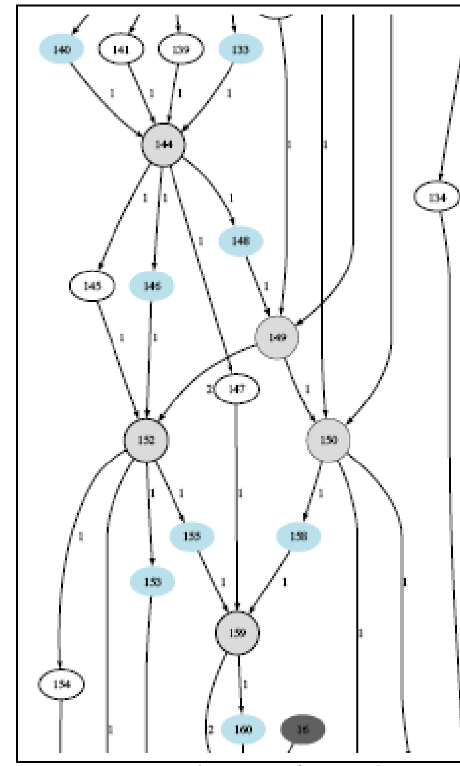
3D reconstruction

Seitz et al 2006



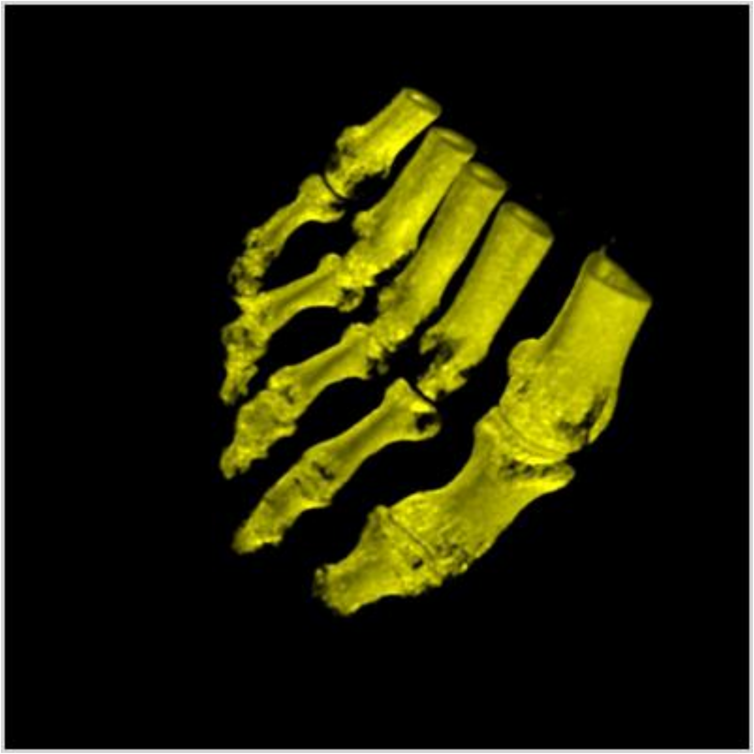
Snavely, Seitz, and Szeliski 2006

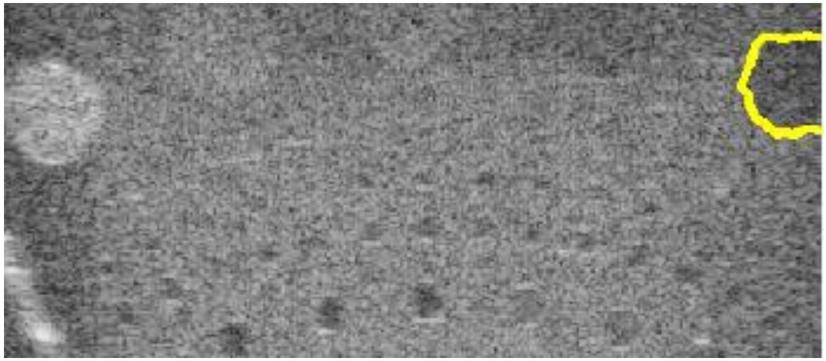
tracking moving objects



Nillius, Sullivan, and Carlsson 2006

medical imaging

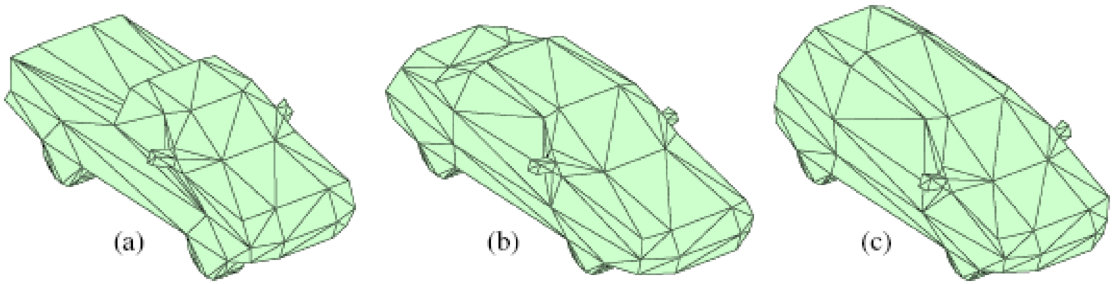


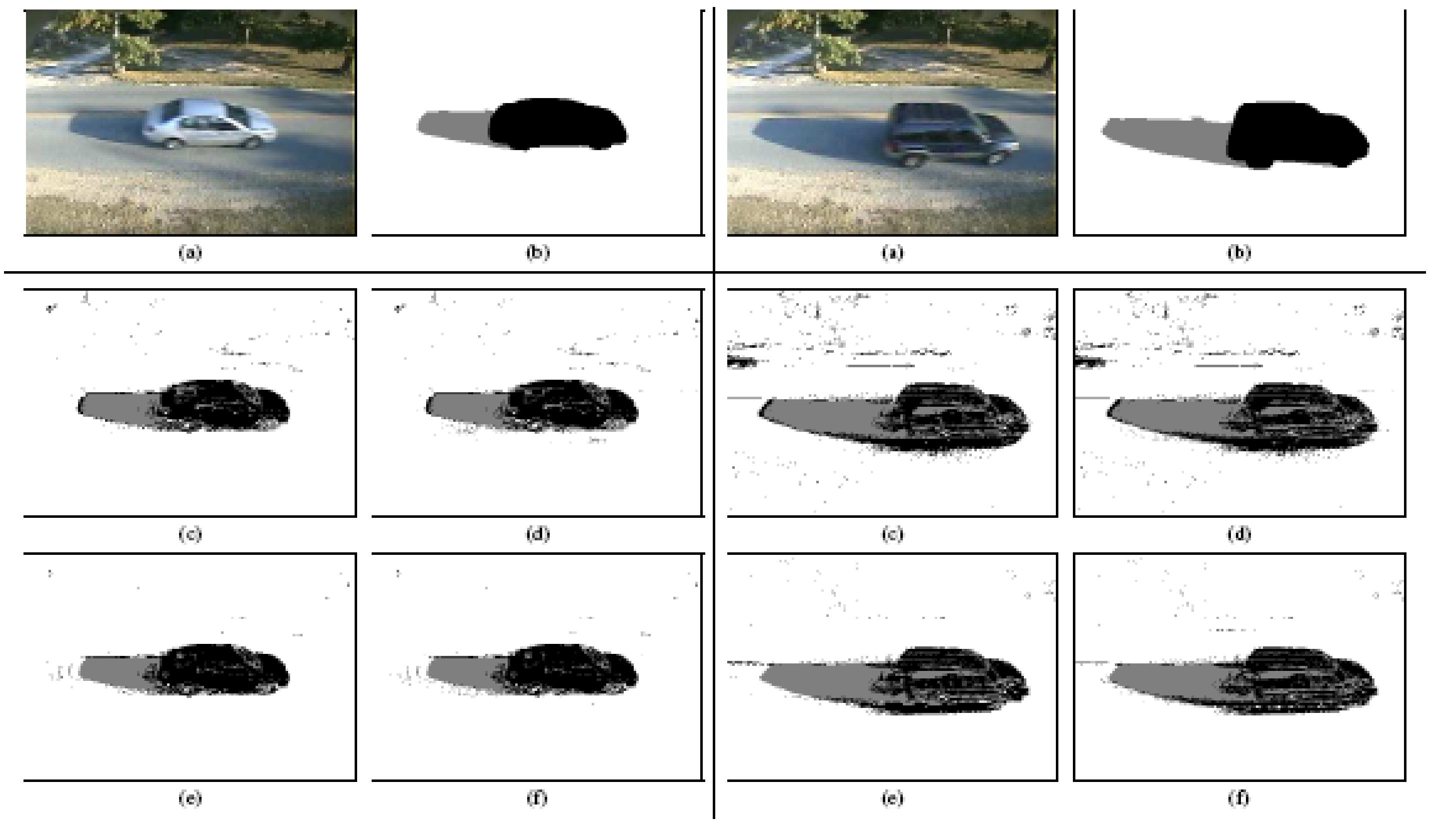


Yuan et al 2005

Slabaugh et al 2006

surveillance





Leotta and Mundy 2006

automotive applications

from a Mercedes advertisement:

"...Also offered is a rearview camera. Mercedes' optional Distronic Plus adaptive cruise control now operates in stop-and-go traffic as well as at higher speed. Newly optional is Mercedes' infrared Night View Assist, which projects onto an instrument panel screen a black-and-white image of objects beyond headlight range."

special effects in movies and games

why is computer vision interesting?

two views:

“engineering view”:

– LOTS of useful applications

“AI view”:

– computer vision demonstrates the HUGE gap between humans and computers
– also encourages us to close this gap in inventive ways

computer vision

- WHAT is it?
- WHY is it interesting?
- HOW does it work?

well-posed versus ill-posed

- I have $15 and spend $3. How much is left?
  - answer: $12

- I have $15 and spend some of it. How much is left?
  - answer: ???

well-posed versus ill-posed

- I have \$15 and spend \$3. How much is left?
  - answer: \$12

- I have \$15 and spend some of it. How much is left?
  - answer: ???

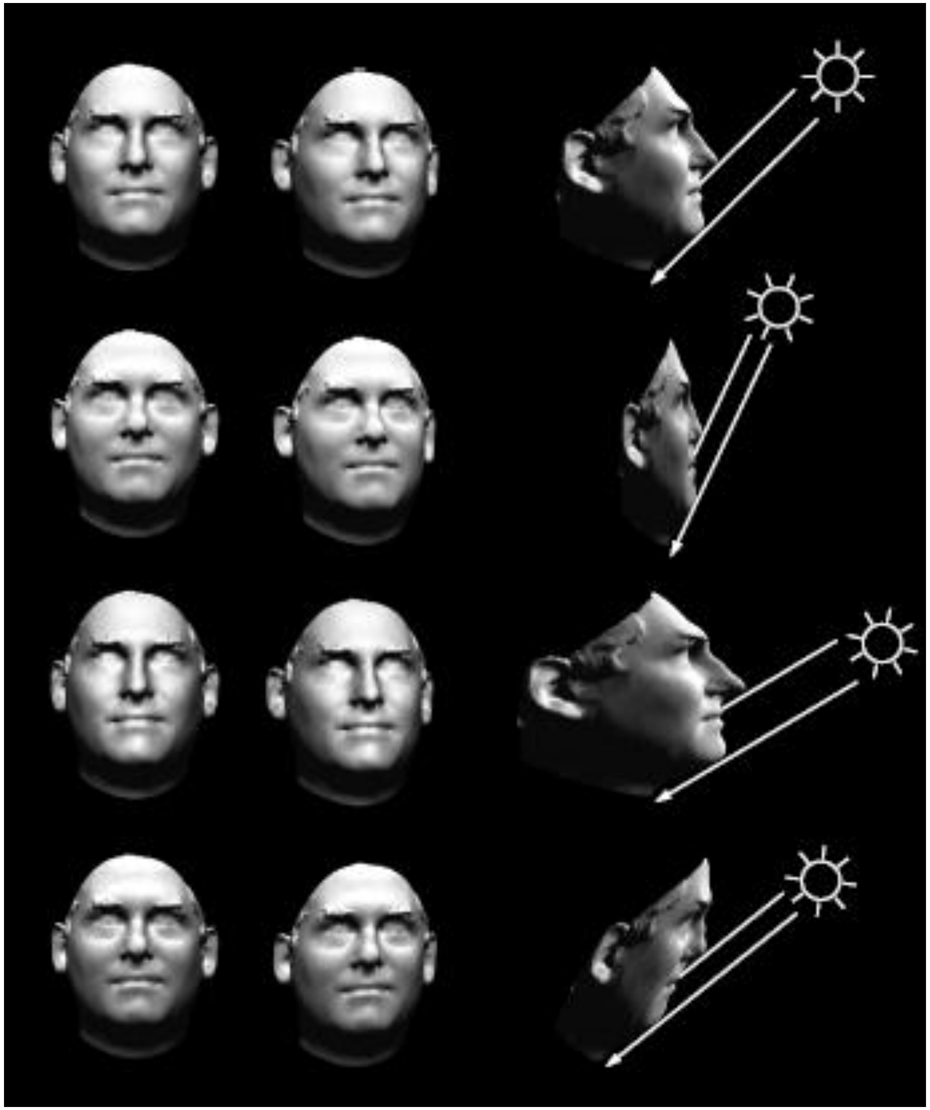
simultaneous equations

simultaneous equations

computer vision problems are typically ill-posed

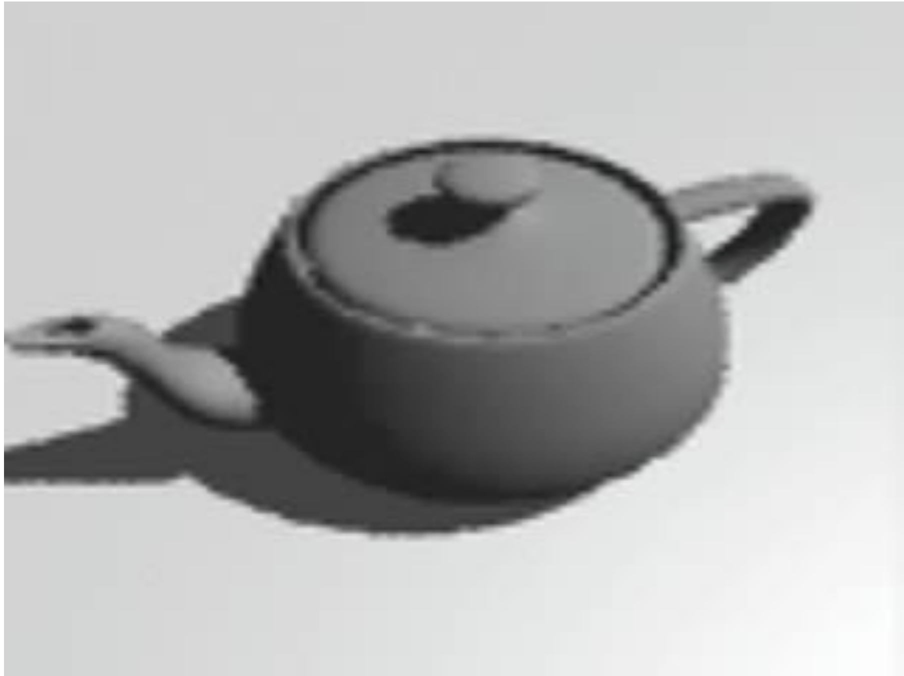
e.g. what 3-D object does this 2-D image represent?

solid objects are also ambiguous



Belhumeur,
Kriegman, Yuille
1997

solid objects are also ambiguous



Belhumeur,
Kriegman, Yuille
1997

the ancient Greeks and Romans exploited this “bas-relief” ambiguity

Belhumeur,
Kriegman, Yuille
1997

computer graphics: typically well-posed
computer vision: typically ill-posed

computer graphics

3D→2D

e.g. animation: given 3-D figure motion, produce 2-D rendering

computer vision

2D→3D

e.g. motion capture: given 2-D video, produce 3-D model

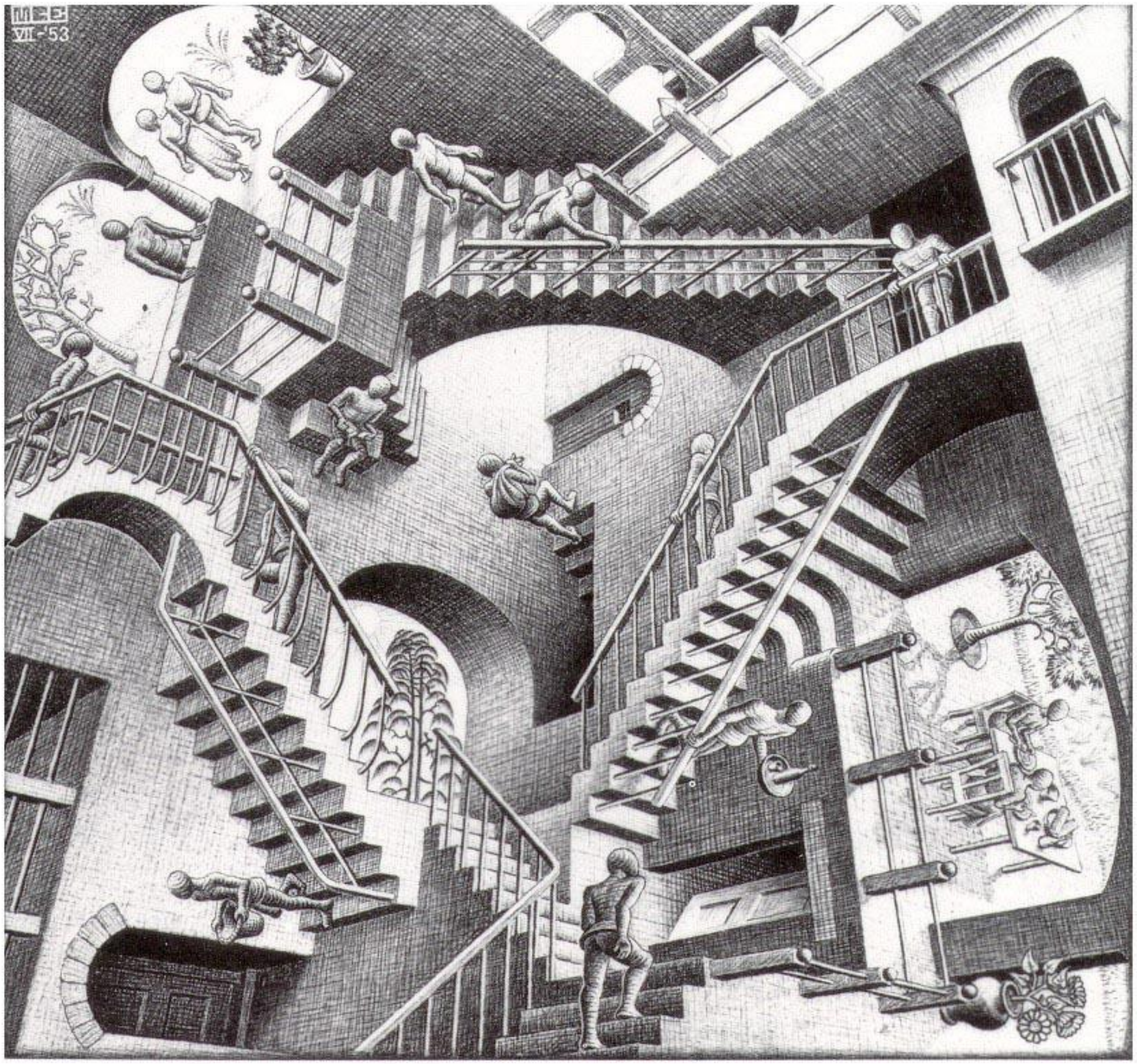
how does computer vision work?

- computer vision problems are typically *ill-posed*
- therefore, need to use *prior* information to choose “best” or “most likely” solution
- therefore, need to use *probability distributions*
- a common approach is to use *Bayes’ rule*

we need extra, or “prior” information to resolve these ambiguities

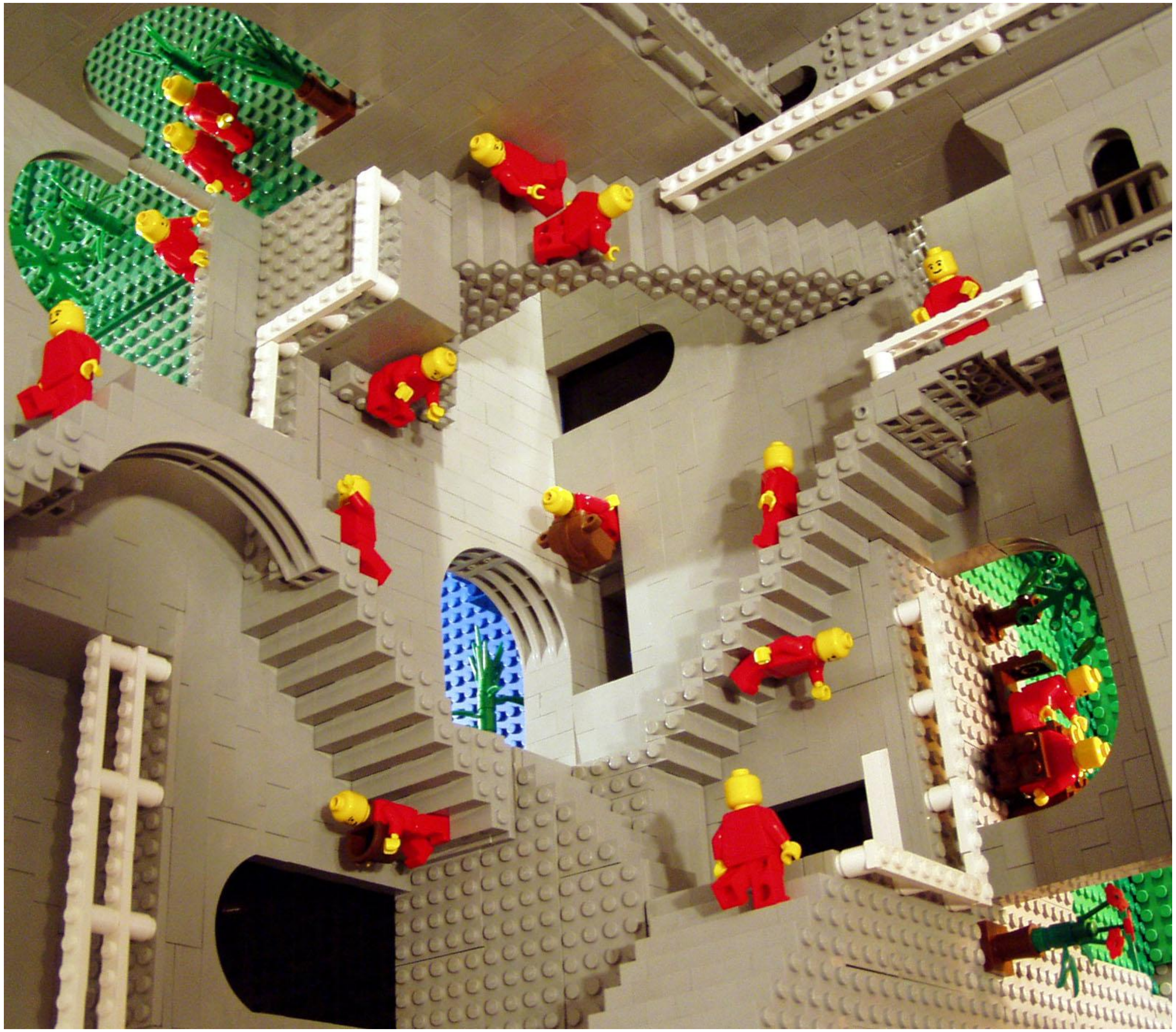
combine PRIOR information with OBSERVED information to draw CONCLUSIONS

aside: optical illusions exploit our visual priors



M C Escher

aside: optical illusions exploit our visual priors



a mathematically rigorous way to combine prior knowledge with a particular observation is to use

Bayes' rule

a mathematically rigorous way to combine prior knowledge with a particular observation is to use

Bayes' rule

consider a medical test for a particular disease

Prob(tests sick | actually sick) = 0.9
Prob(tests healthy | actually sick) = 0.1

question: what is Prob(actually sick | tests sick) ?

first thought experiment

- take 10 healthy people and test them
  - 9 test healthy, 1 tests sick
- take 10 sick people and test them
  - 1 tests healthy, 9 test sick

first thought experiment

- take 10 healthy people and test them
  - 9 test healthy, 1 tests sick
- take 10 sick people and test them
  - 1 tests healthy, 9 test sick

second thought experiment

- test 1000 people who are almost all healthy (only 10 are sick, other 990 are healthy)

second thought experiment

- test 1000 people who are almost all healthy (only 10 are sick, other 990 are healthy)

second thought experiment

- test 1000 people who are almost all healthy (only 10 are sick, other 990 are healthy)

second thought experiment

- test 1000 people who are almost all healthy (only 10 are sick, other 990 are healthy)

of the 108 people who TESTED sick, only 9 are ACTUALLY sick!

i.e. Prob(actually sick | tests sick) = 9/108 = 8%

Bayes’ rule

simple example of using Bayes' rule in computer vision

what is this letter?

simple example of using Bayes' rule in computer vision

what is this letter?

instead of a medical test, we have a character detector

Bayes' rule for optical character recognition

Bayes' rule for optical character recognition

[live demo of Bayes’ rule for handwriting recognition]

mathematical derivation of Bayes' rule

Bayes' rule works for continuous probability distributions too

the Rev. Thomas Bayes

there are many generalizations of Bayes' rule

- Bayes networks, graphical models
- Markov chains
- Markov random fields
- dynamic programming

how does computer vision work?

- computer vision problems are typically *ill-posed*
- therefore, need to use *prior* information to choose “best” or “most likely” solution
- therefore, need to use *probability distributions*
- a common approach is to use *Bayes’ rule*

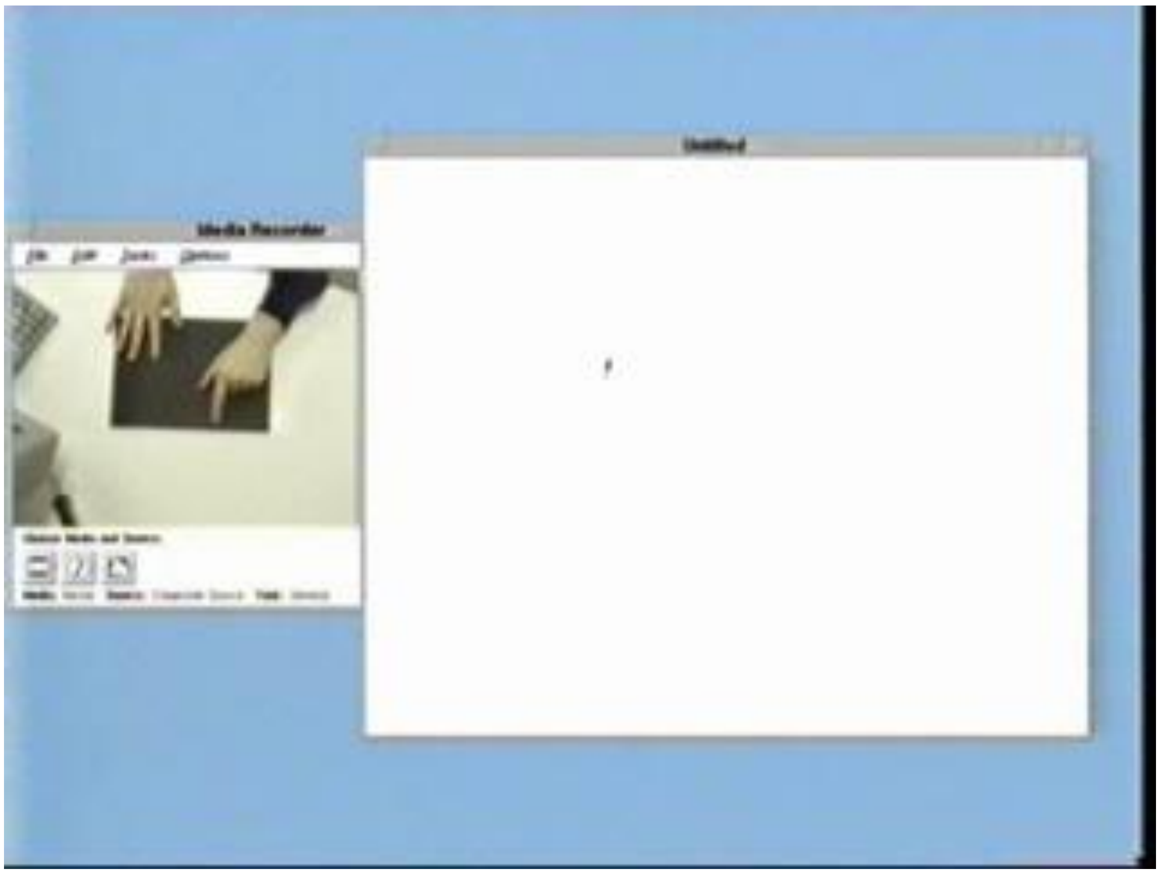
examples of Bayes' rule in action

examples of Bayes' rule in action

1. hand tracking for human-computer interaction

examples of Bayes' rule in action

1. hand tracking for human-computer interaction



examples of Bayes' rule in action

2. person tracking

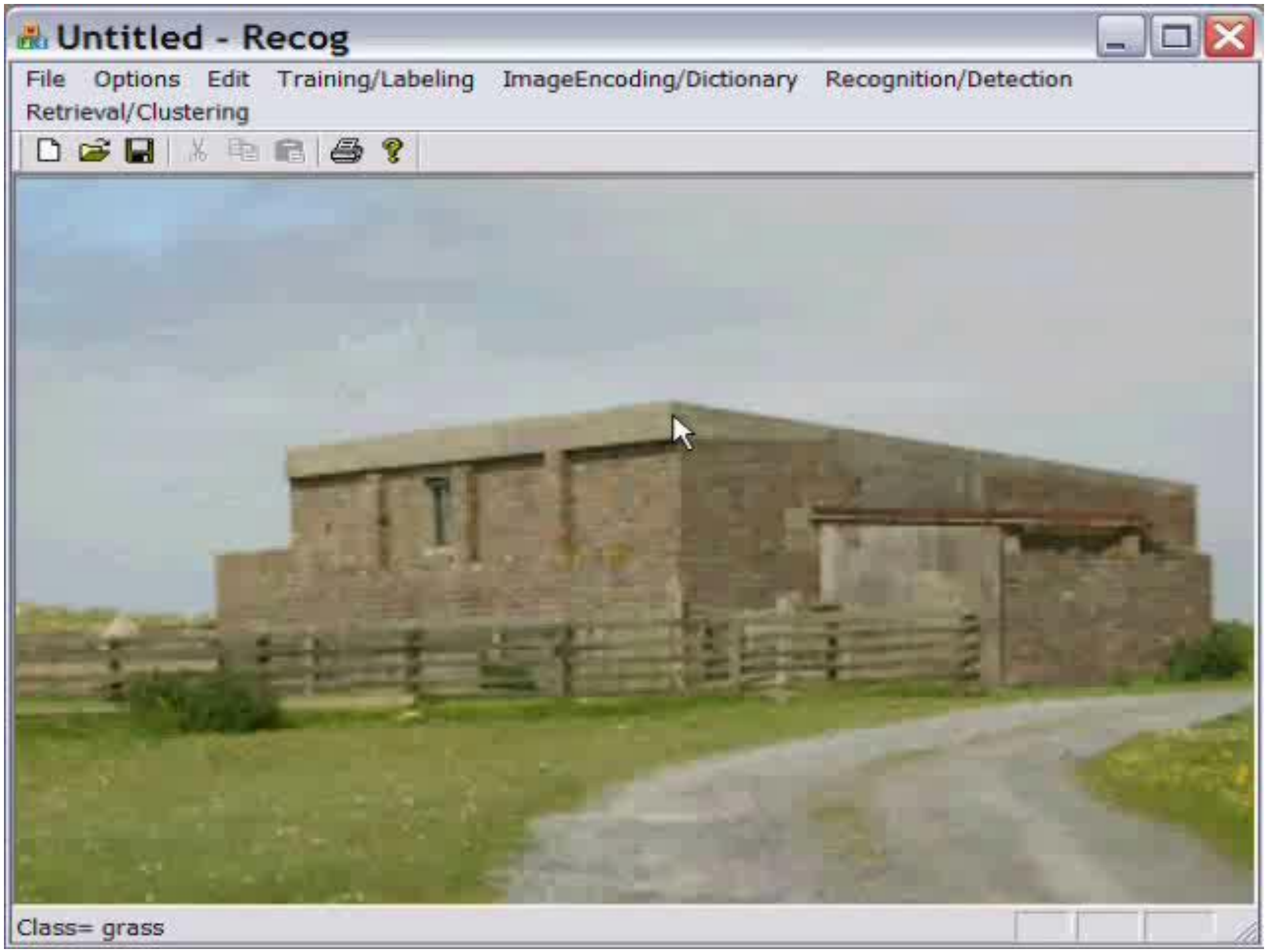
examples of Bayes' rule in action

3. layer segmentation

Kolmogorov et al, 2005

examples of Bayes' rule in action

4. object recognition



Winn et al 2006

examples of Bayes' rule in action

5. in-painting

Criminisi et al 2004

computer vision

- WHAT is it?
  - (engineering) do useful things with cameras and computers
  - (AI) try to emulate human visual system
- WHY is it interesting?
  - (engineering) many useful applications
  - (AI) huge gap between humans and computers
- HOW does it work?
  - computer vision problems usually ill-posed
  - need to incorporate prior information using probability distributions
  - Bayes' rule can achieve this

can computers see?

- not right now
  - e.g. object recognition:
    - computers get 20-90% success rate on database of tens of object classes
    - humans get virtually 100% success rate on thousands of object classes
  - web sites verify human-ness using object recognition

Computing Machinery and Intelligence (Alan Turing, 1950)

`I propose to consider the question, “Can machines think?”...’

Computing Machinery and Intelligence (Alan Turing, 1950)

`I propose to consider the question, “Can machines think?”...’

can computers see?

- not right now

- the future? Maybe computers *can* see.
  - the arguments in Turing's "Turing test" paper appear to apply equally well to a "visual Turing test"

thank you!