

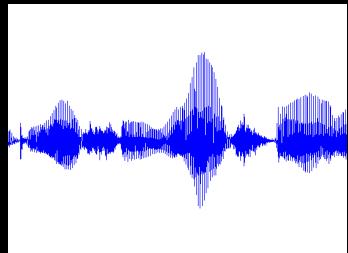
Things we want to do with data

Images



→ Label image

Audio



→ Speech recognition

Text



→ Web search

Andrew Ng

STanford AI Robot (STAIR)



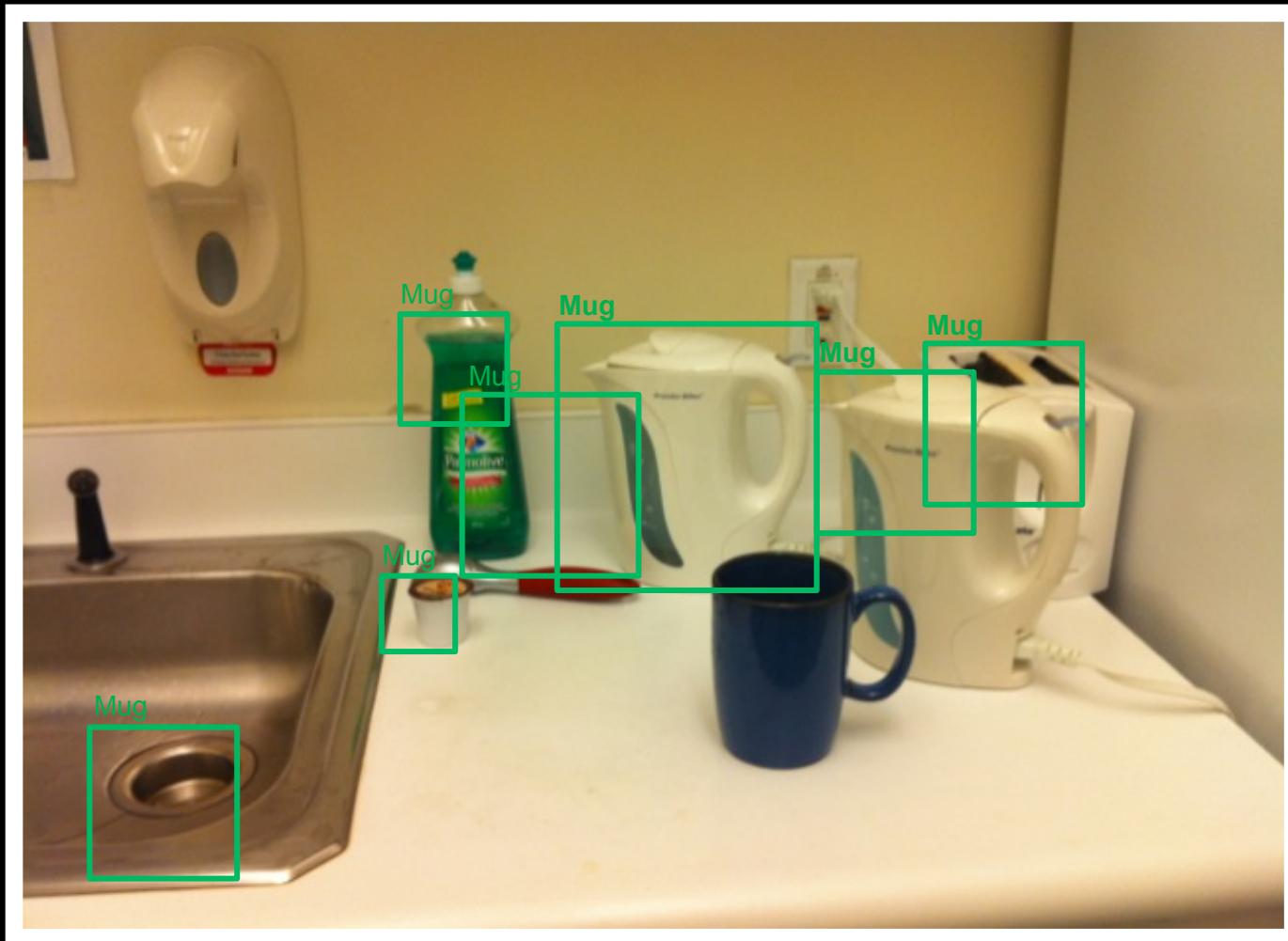
Andrew Ng

Computer vision: Identify coffee mug



Andrew Ng

Computer vision: Identify coffee mug



Andrew Ng

Why is computer vision hard?



The camera sees :

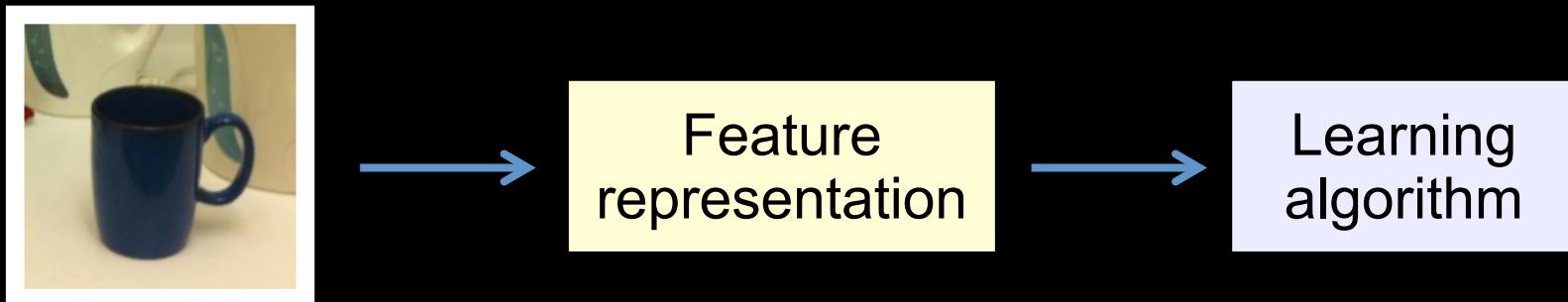
194	210	201	212	199	213	215	195	178	158	182	209
180	189	190	221	209	205	191	167	147	115	129	163
114	126	140	188	176	165	152	140	170	106	78	88
87	103	115	154	143	142	149	153	173	101	57	57
102	112	106	131	122	138	152	147	128	84	58	66
94	95	79	104	105	124	129	113	107	87	69	67
68	71	69	98	89	92	98	95	89	88	76	67
41	56	68	99	63	45	60	82	58	76	75	65
20	43	69	75	56	41	51	73	55	70	63	44
50	50	57	69	75	75	73	74	53	68	59	37
72	59	53	66	84	92	84	74	57	72	63	42
67	61	58	65	75	78	76	73	59	75	69	50

Computer vision

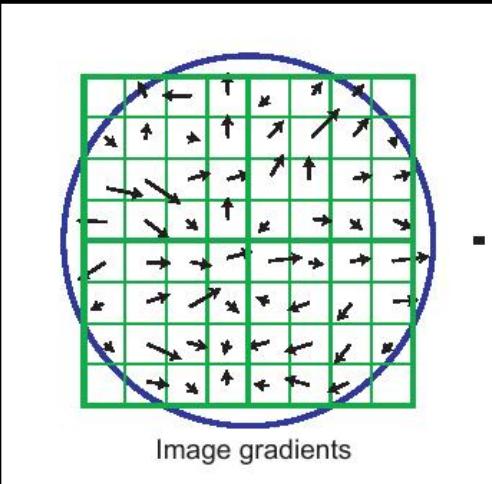


Learning
algorithm

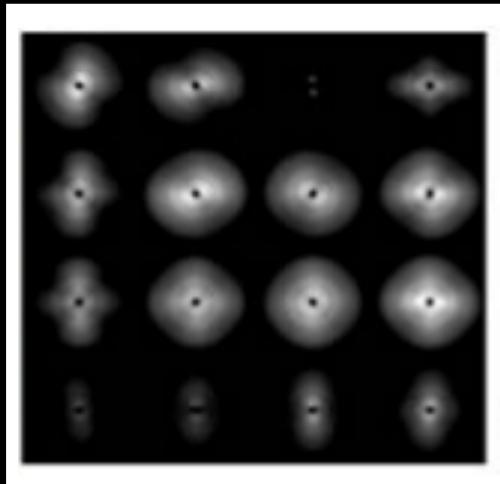
Computer vision



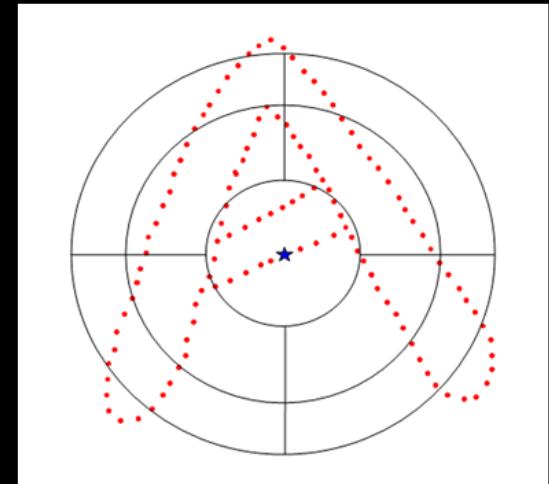
Features for vision



SIFT



GIST



Shape context

Features for machine learning

Images



Image

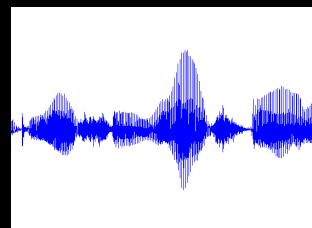


Vision features

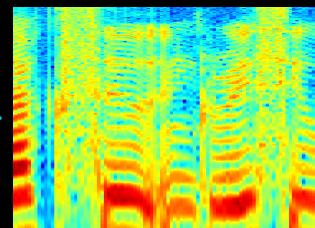


Detection

Audio



Audio



Audio features

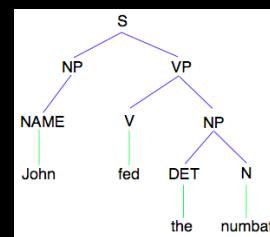


Speaker ID

Text



Text



Text features

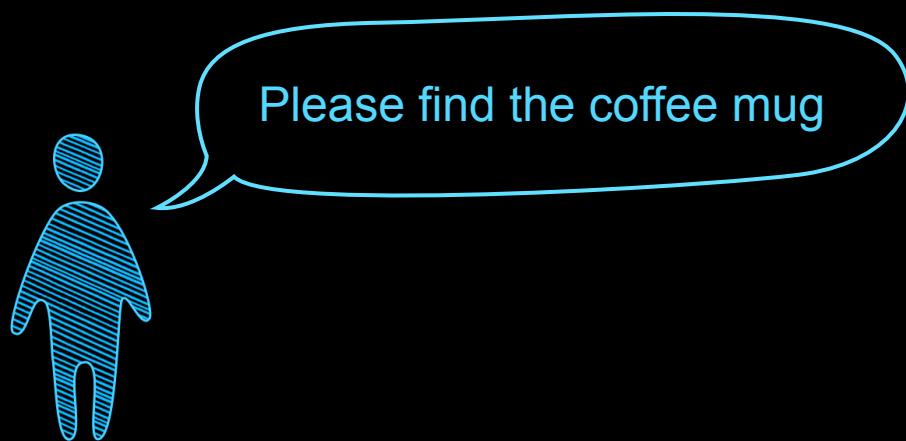
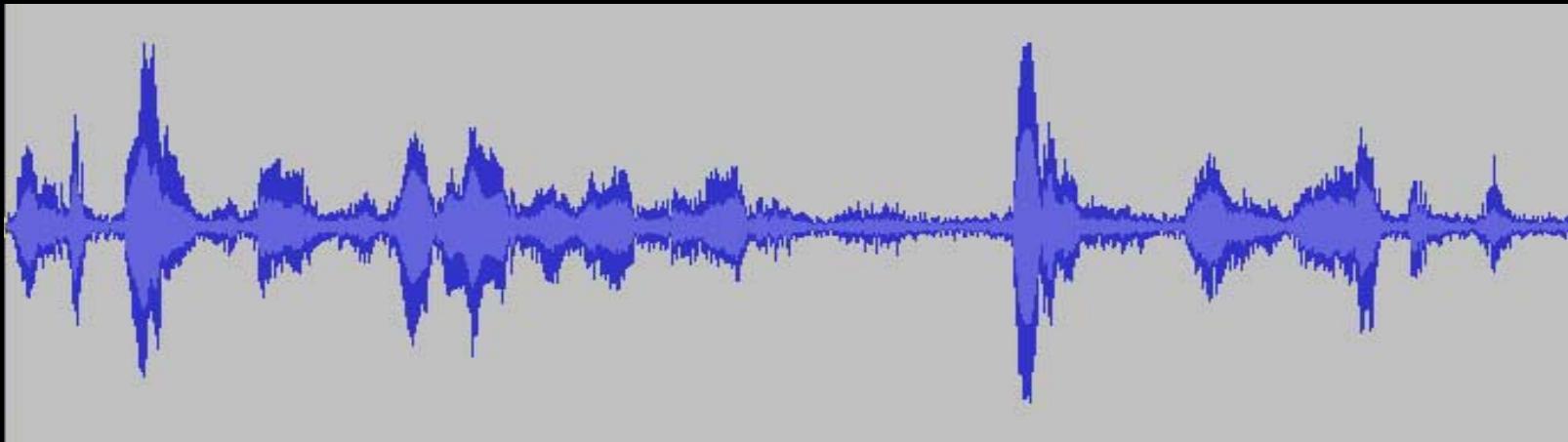
Web search

...

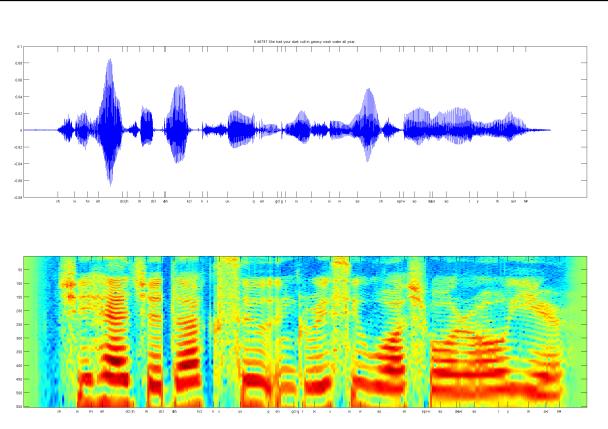
Andrew Ng

Why is speech recognition hard?

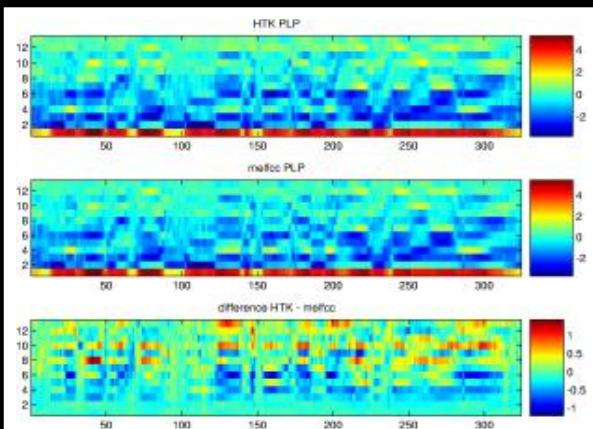
Microphone recording:



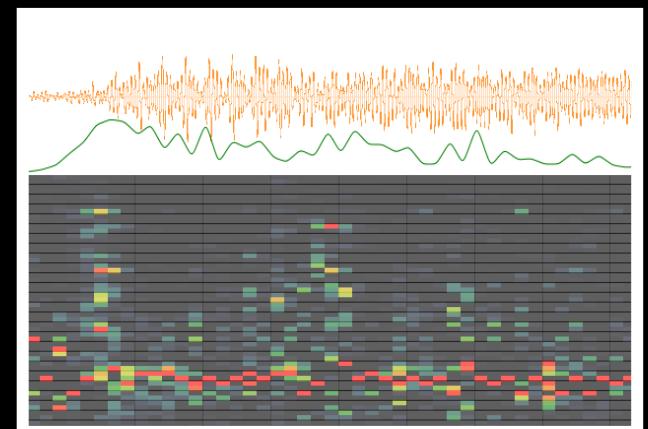
Features for audio



Spectrogram

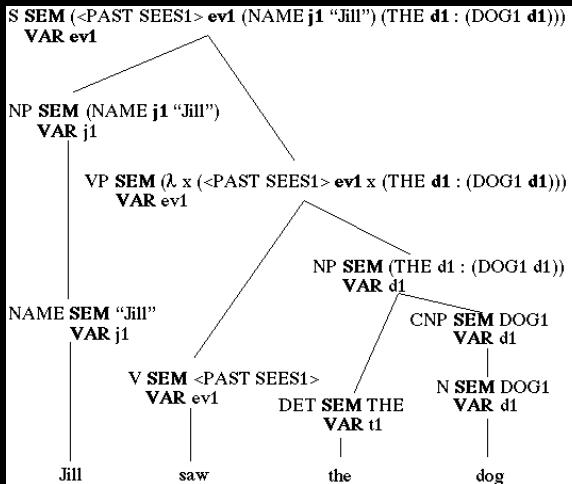


MFCC



Flux

Features for text



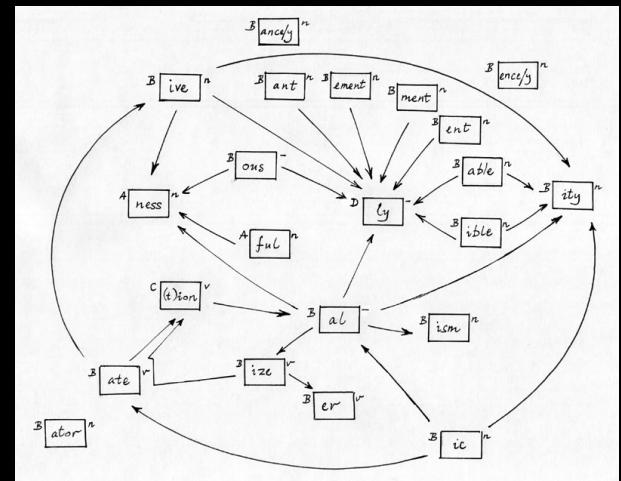
Parser

XML document structure:

```

<DOC>
<DOCID> wsj94_008.0212 </DOCID>
<DOCNO> 940413-0062. </DOCNO>
<HL> Who's News:
@ Burns Fry Ltd. </HL>
<DD> 04/13/94 </DD>
<SO> WALL STREET JOURNAL (J), PAGE B10 </SO>
<CO> MER </CO>
<IN> SECURITIES (SCR) </IN>
<TXT>
<p>
  BURNS FRY Ltd. (Toronto) -- Donald Wright, 4
  named executive vice president and director of
  brokerage firm. Mr. Wright resigned as president
  Canada Inc., a unit of Merrill Lynch & Co., to
  Kassirer, 48, who left Burns Fry last month. A
  spokeswoman said it hasn't named a successor to
  expected to begin his new position by the end o
</p>
</TXT>
</DOC>
  
```

Named entity



Stemming

The idea:

Most perception (input processing) in the brain may be due to one learning algorithm.



The idea:

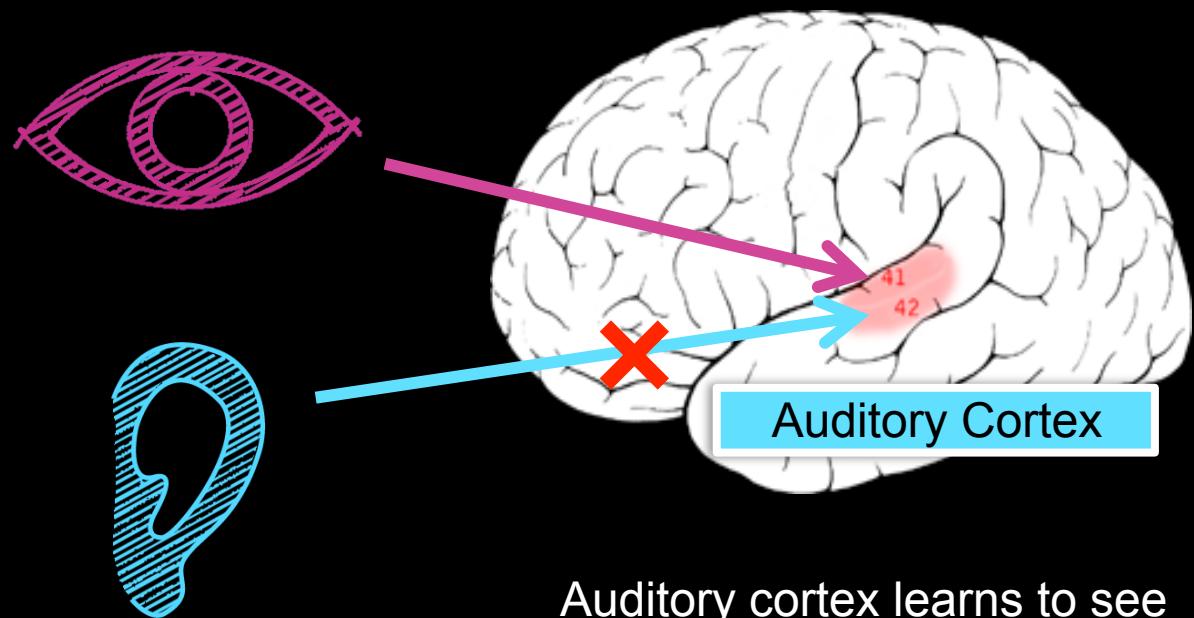
Build learning algorithms
that mimic the brain.

Most of human intelligence may
be due to one learning algorithm.



Andrew Ng

The “one learning algorithm” hypothesis

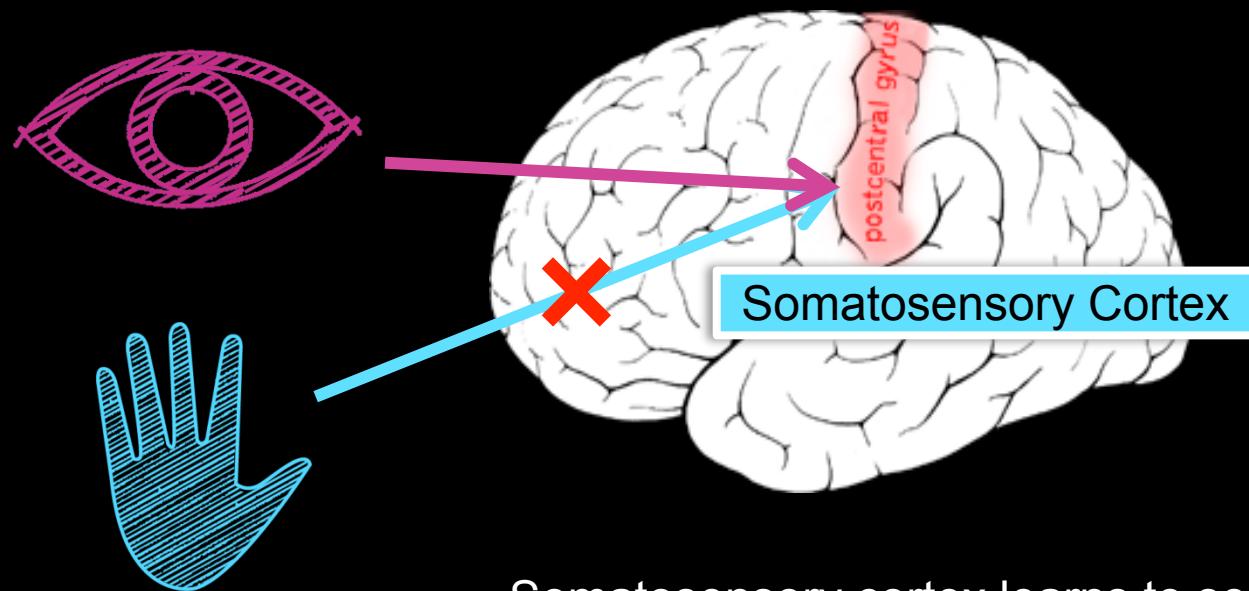


Auditory cortex learns to see

[Roe et al., 1992]

Andrew Ng

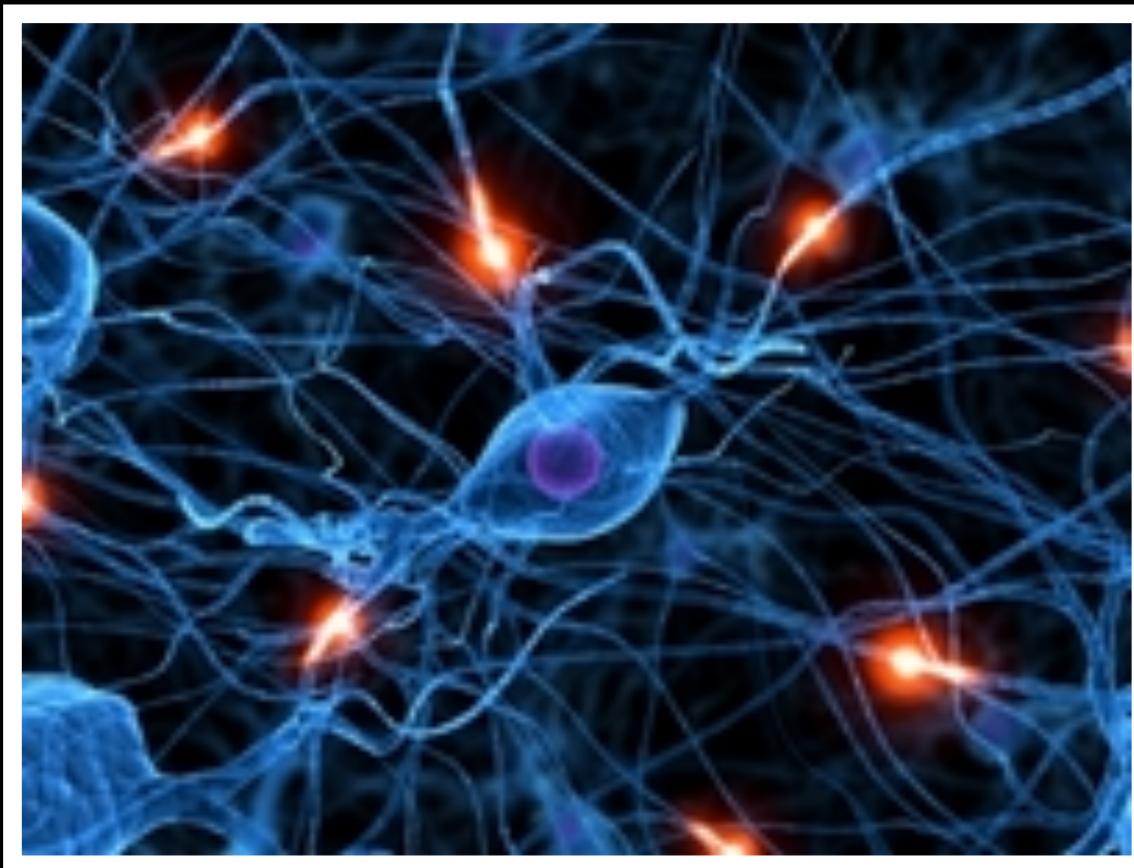
The “one learning algorithm” hypothesis



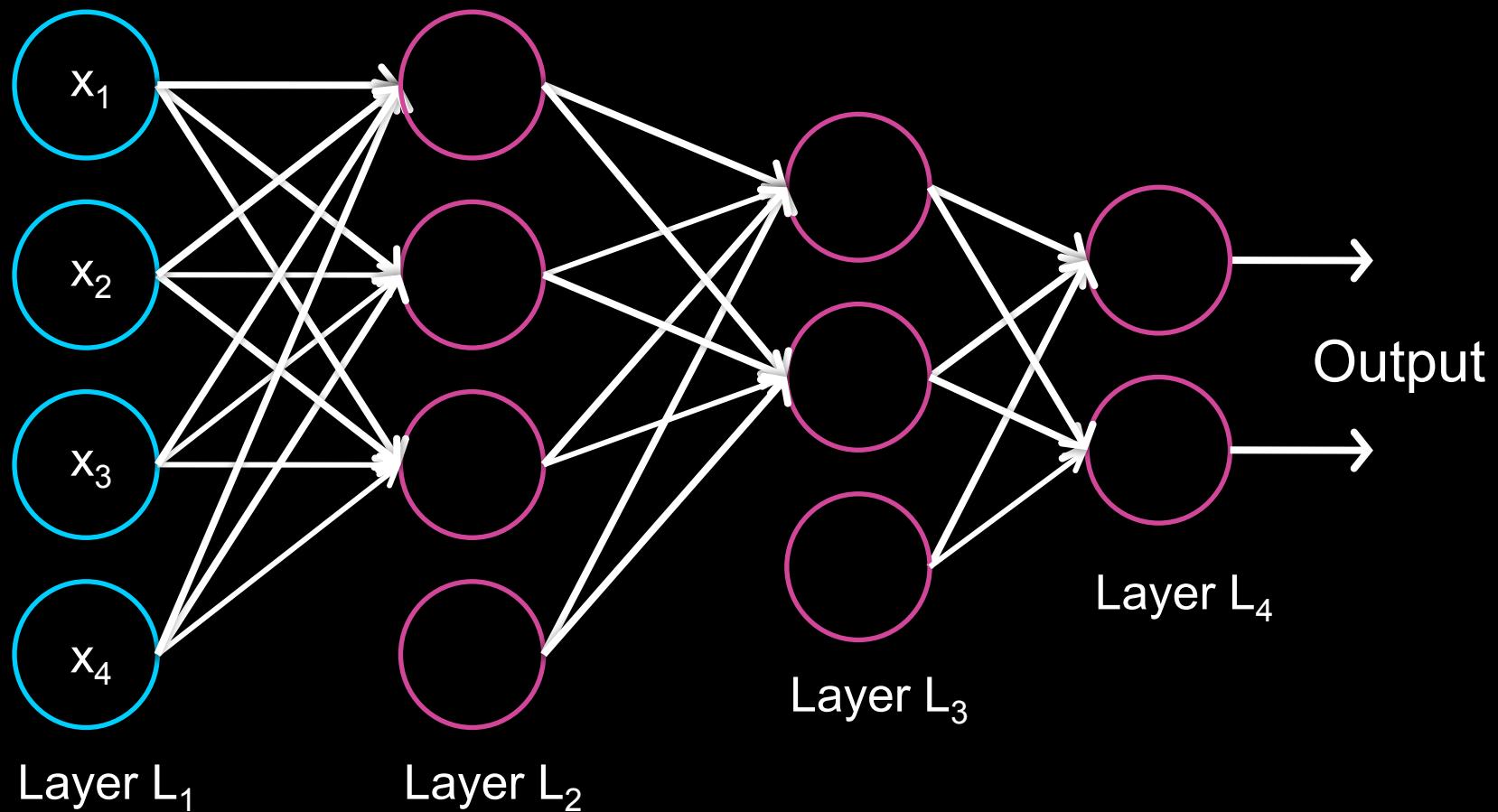
[Metin & Frost, 1989]

Andrew Ng

Neurons in the brain



Neural Network (Deep Learning)



Deep Learning trends

Now



0-2 years
Tagged data

3-5 years
Tagged & untagged data



Learning from tagged data (supervised)



Coffee mug



Coffee mug



Coffee mug



Coffee mug



Coffee mug

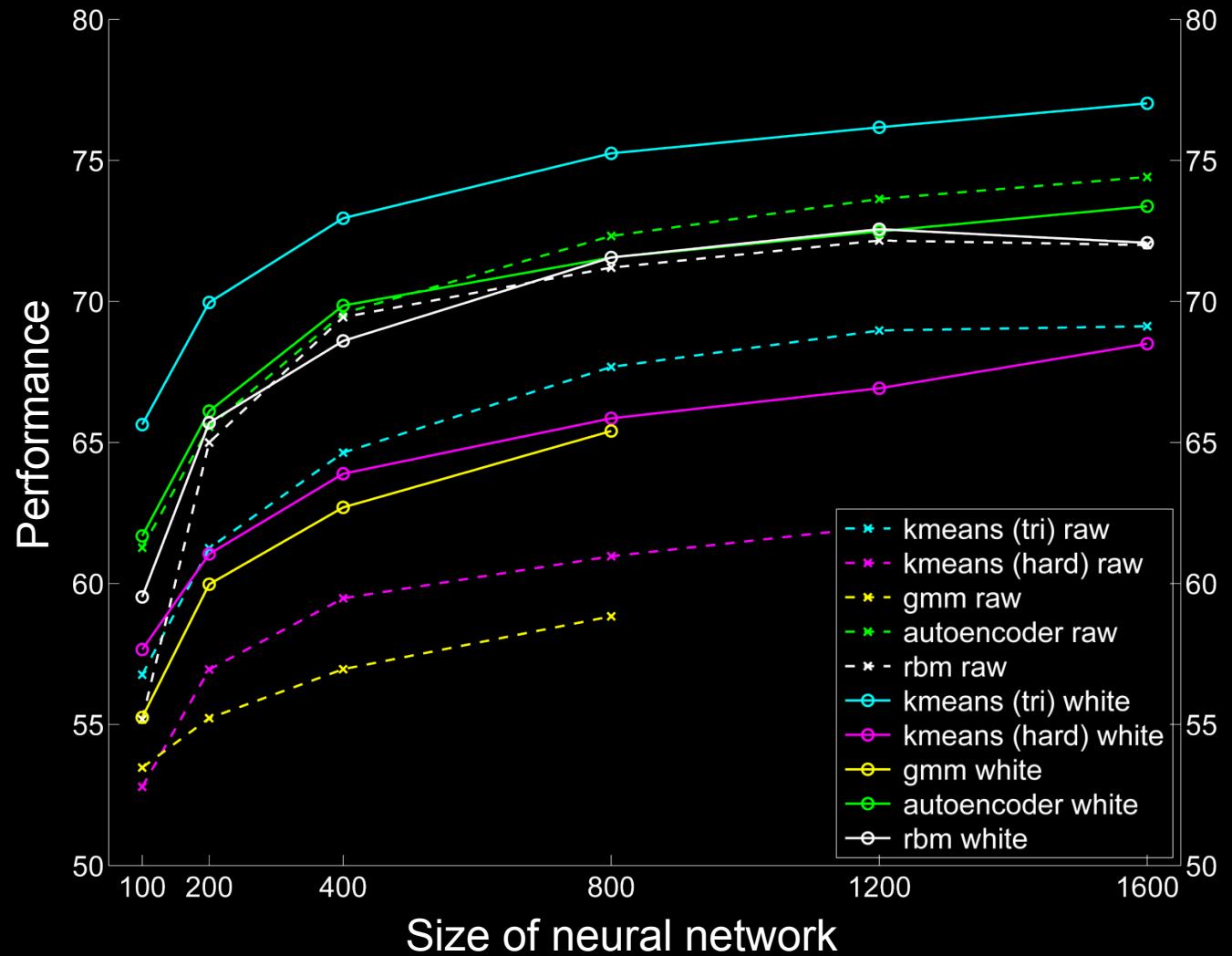


Coffee mug



Testing: What is this?

Bigger is better



[Adam Coates]

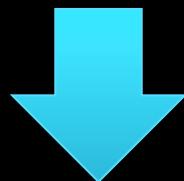
Google Brain



Andrew Ng

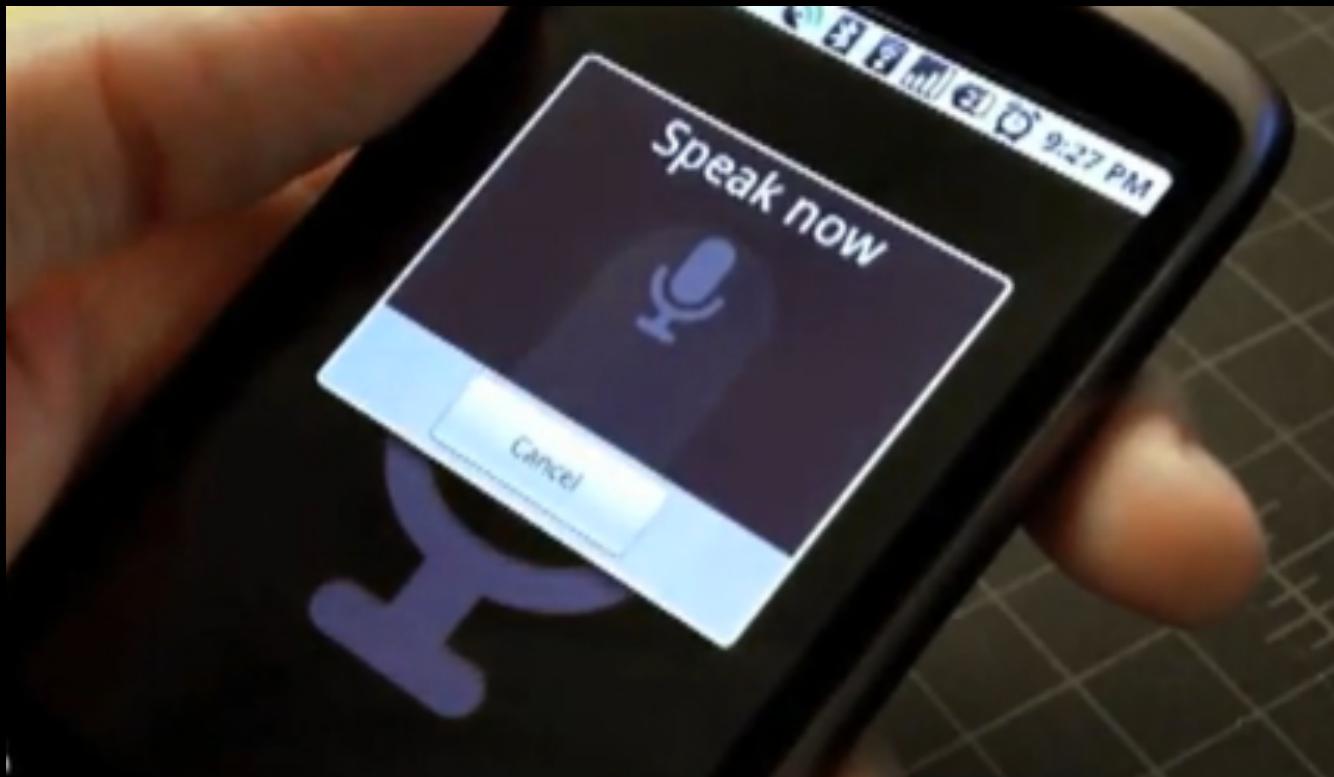
AI as a computer systems problem

10 million connections



1 billion connections

Speech recognition, and more....



[with Vincent Vanhoucke]

Andrew Ng

Tagged vs. untagged data



Coffee mug



Coffee mug



Coffee mug



Coffee mug



Coffee mug



Coffee mug

Untagged data (unsupervised learning)



Unknown



Unknown



Unknown



Unknown



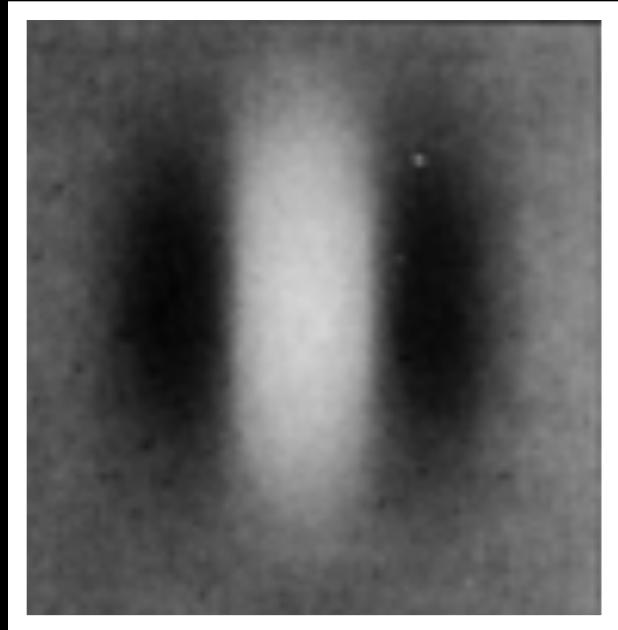
Unknown



Unknown

How does the brain process images?

Visual cortex looks for lines/edges.

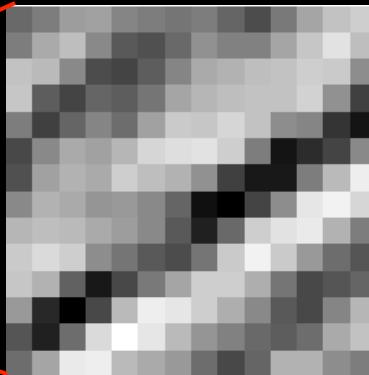


Neuron #1 of visual cortex
(model)



Neuron #2 of visual cortex
(model)

Start with Image patches

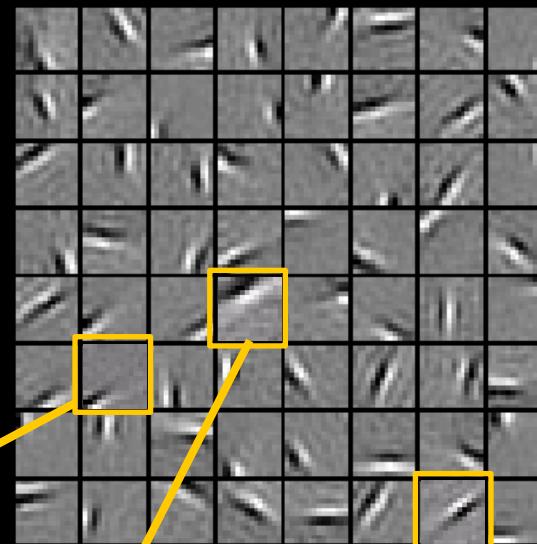


152	147	128	84
129	113	107	87
98	95	89	88
60	82	58	76
51	73	55	70

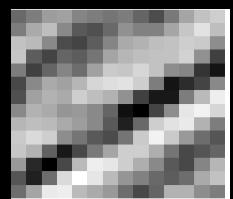
$R^{14 \times 14}$

Sparse Coding

$\phi_1, \phi_2, \dots, \phi_{64}$

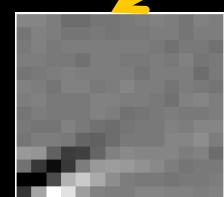


Test example

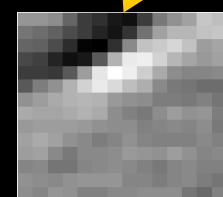


x

$$\approx 0.8 * \phi_{36}$$



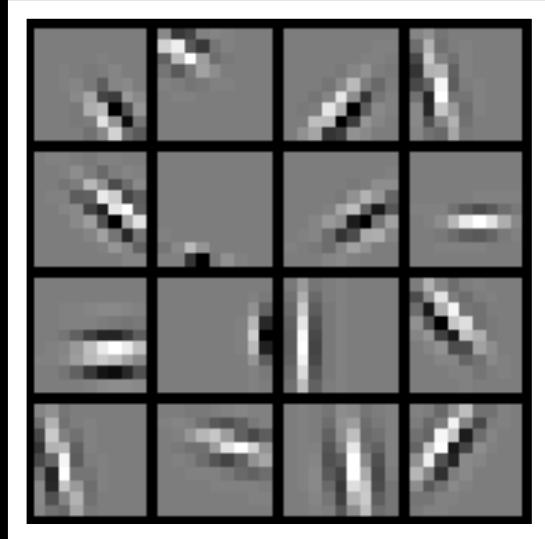
$$+ 0.3 * \phi_{42}$$



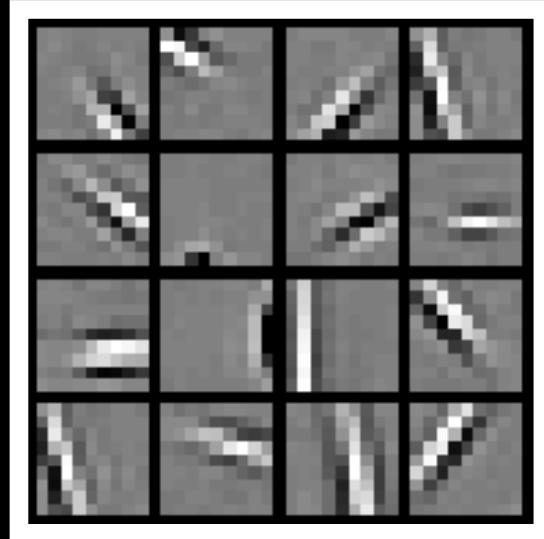
$$+ 0.5 * \phi_{63}$$

Comparing to Biology

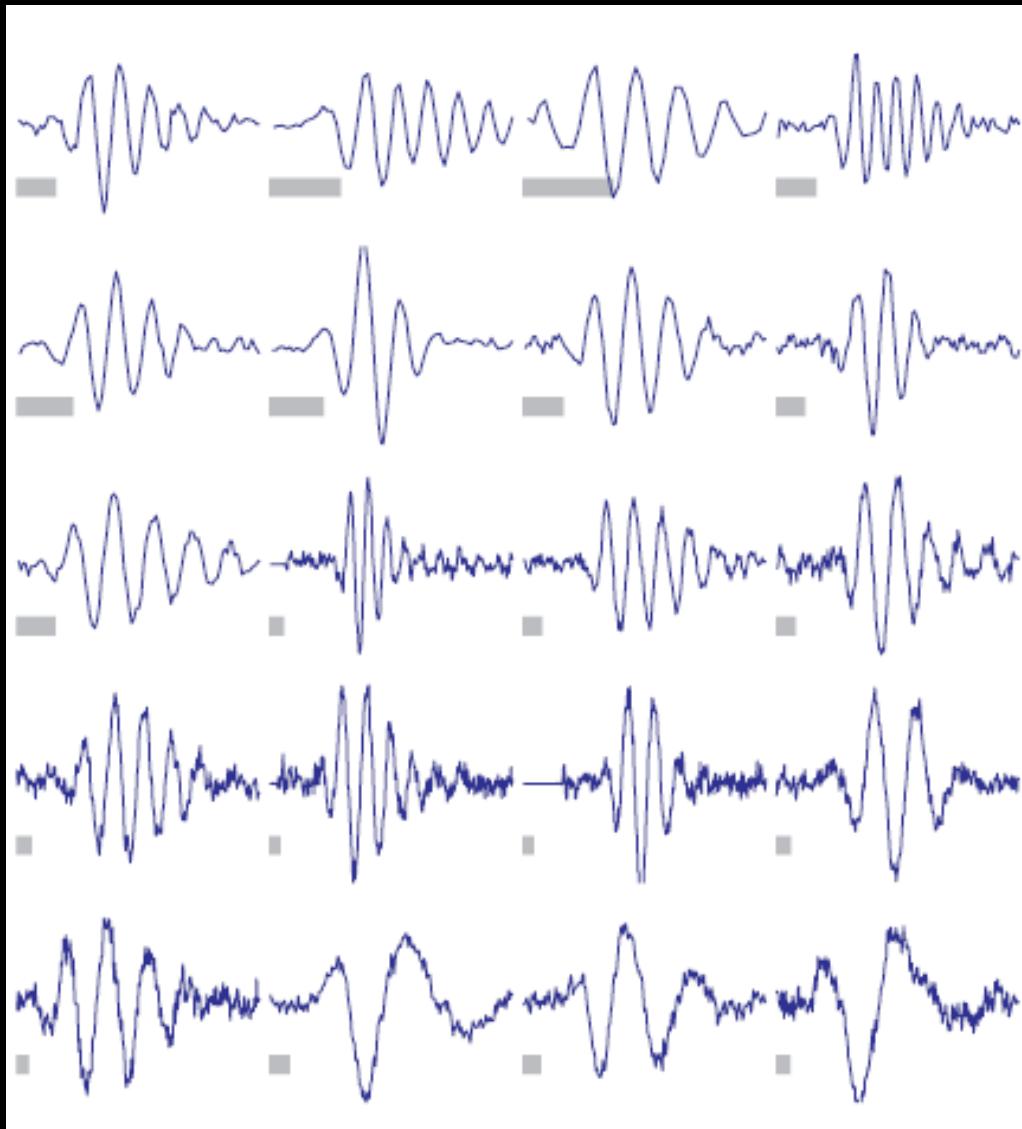
Brain (visual cortex)



Learning algorithm

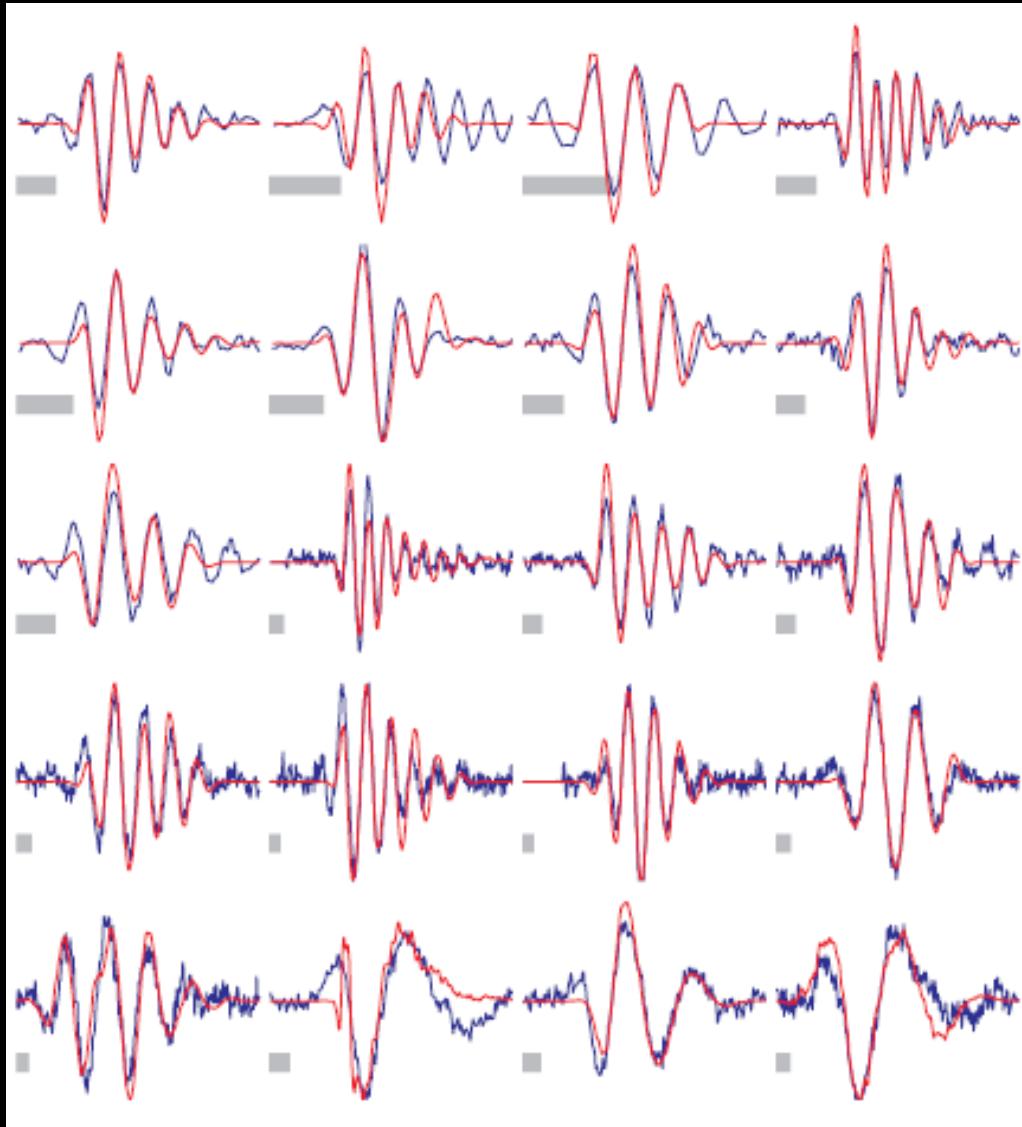


Comparing to Biology



[Evan Smith]

Comparing to Biology



[Evan Smith]

Learning from YouTube videos



Unknown



Unknown



Unknown



Unknown

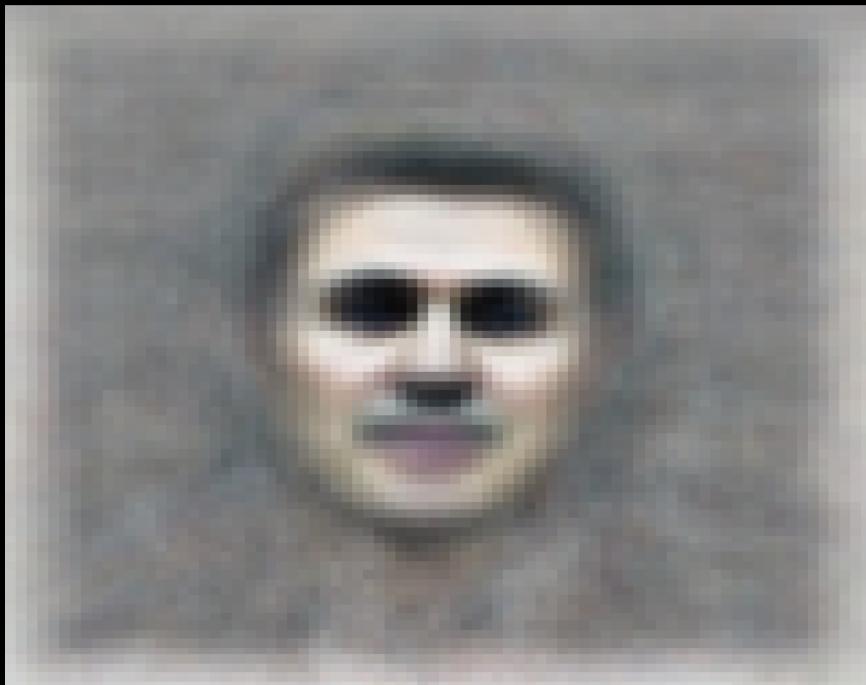


Unknown



Unknown

Face neuron



[Le et al., 2012]

Andrew Ng

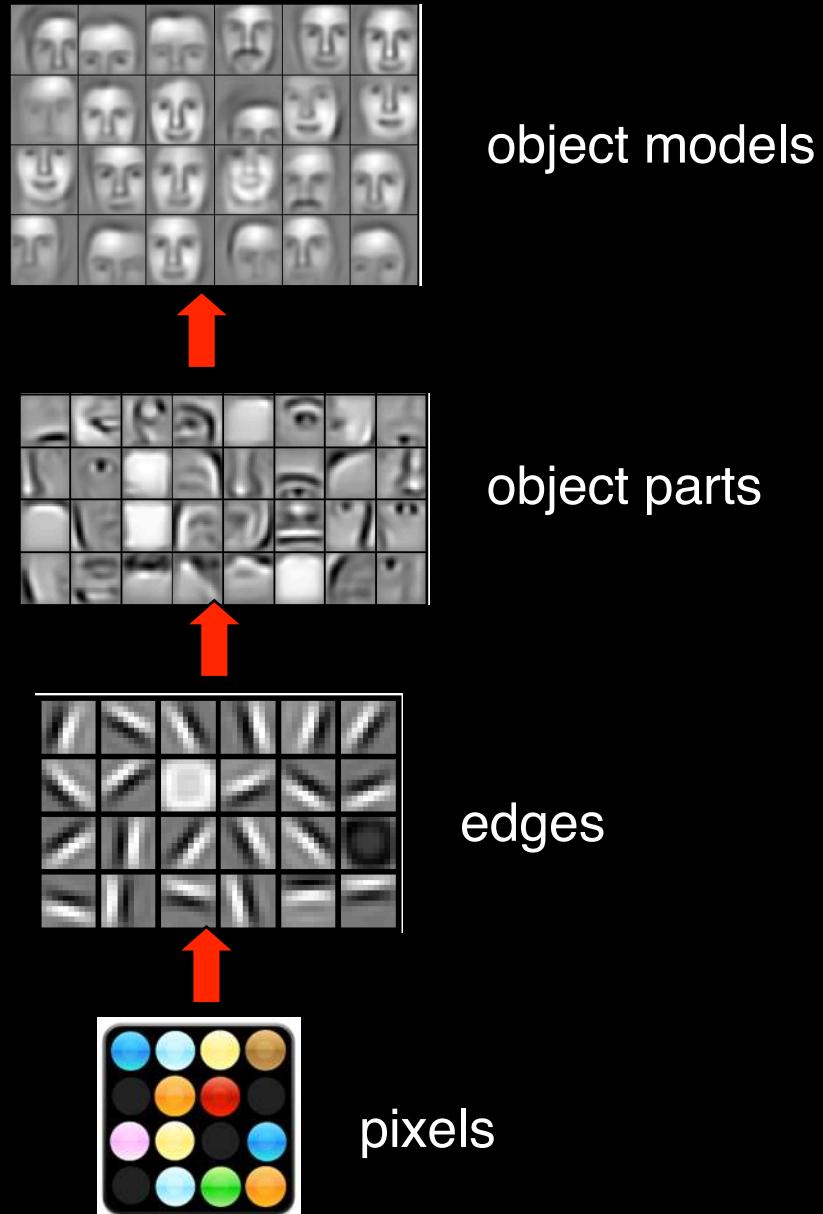
Cat neuron



[Le et al., 2012]

Andrew Ng

Deep Learning



Andrew Ng

16,000 CPUs is expensive



Andrew Ng

GPUs (Graphics Processor Unit)



[Adam Coates, Bryan Catanzaro, et al.]

Andrew Ng

Building huge neural networks

10 million connections

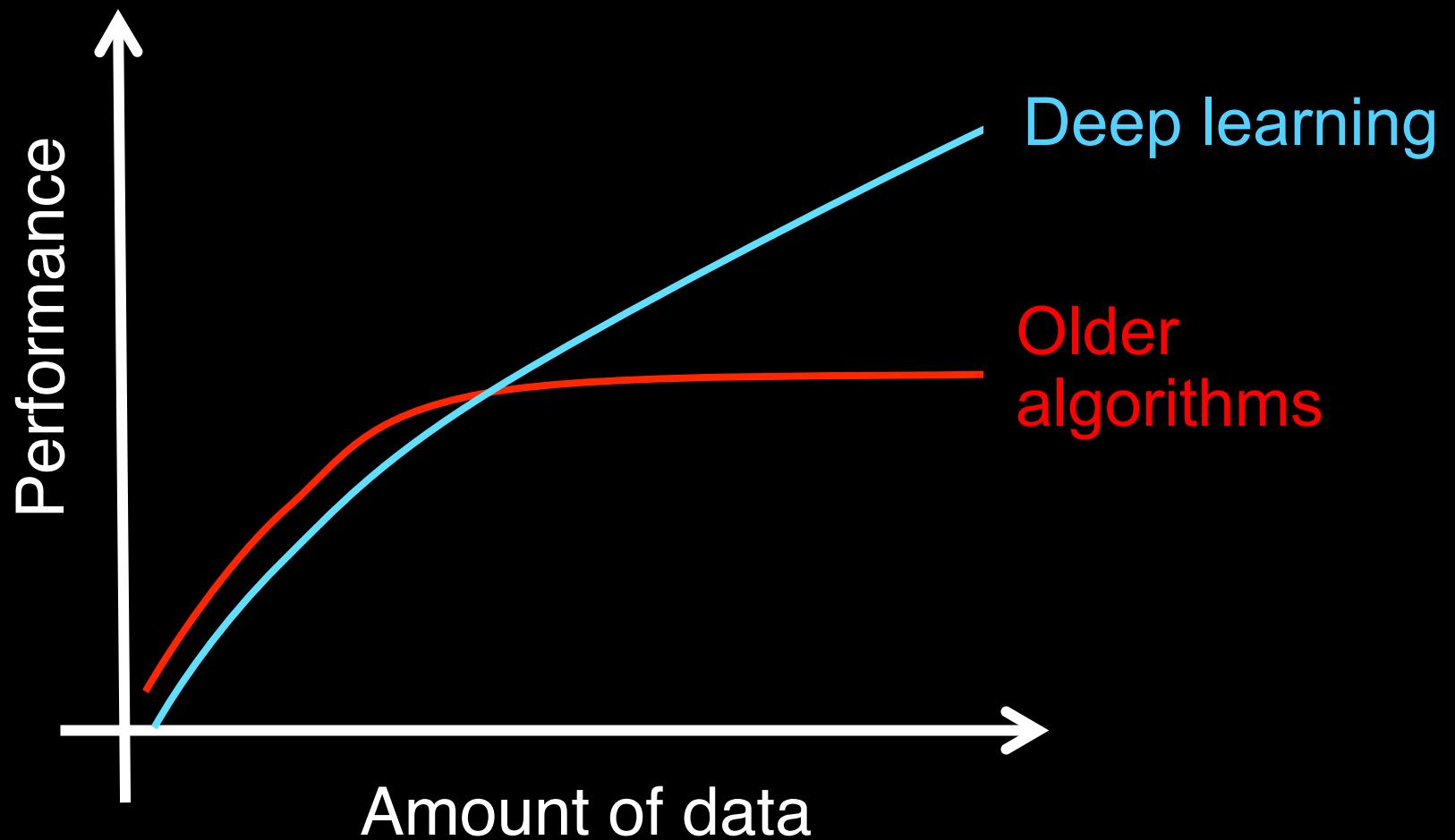


1 billion connections

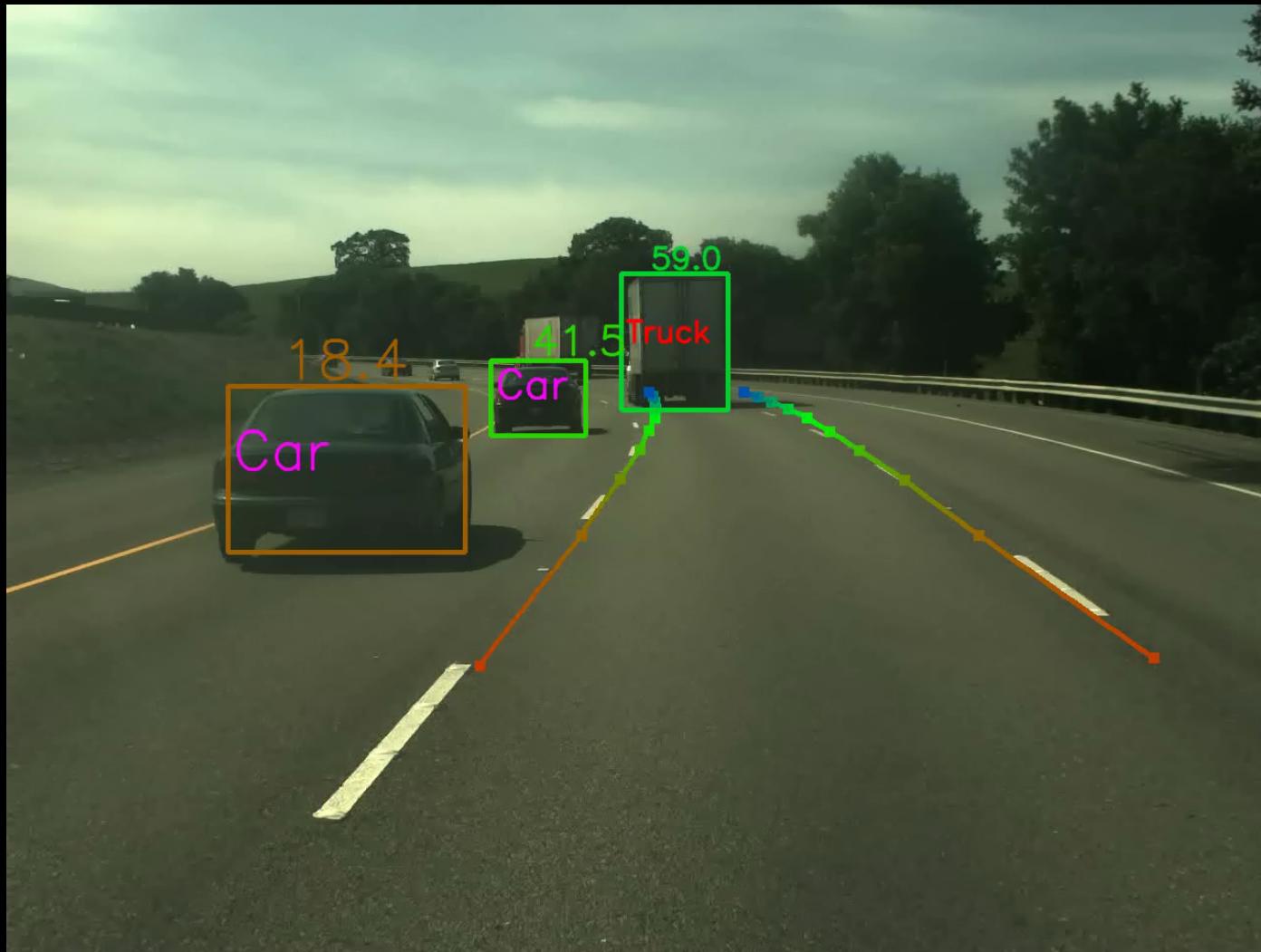


10 billion connections

Learning from tagged data



Highway perception



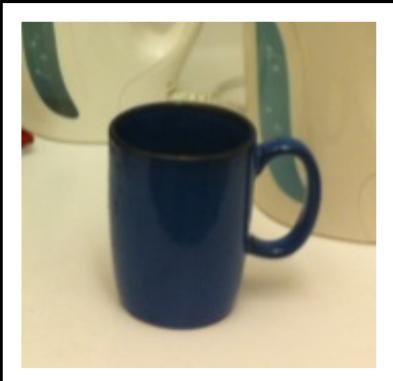
Andrew Ng

Deep Learning trends



0-2 years
Tagged data

3-5 years
Tagged & untagged data



Untagged data and AI (unsupervised learning)



Andrew Ng

AI will transform the internet



Speech



Images



Text

Andrew Ng



Speech



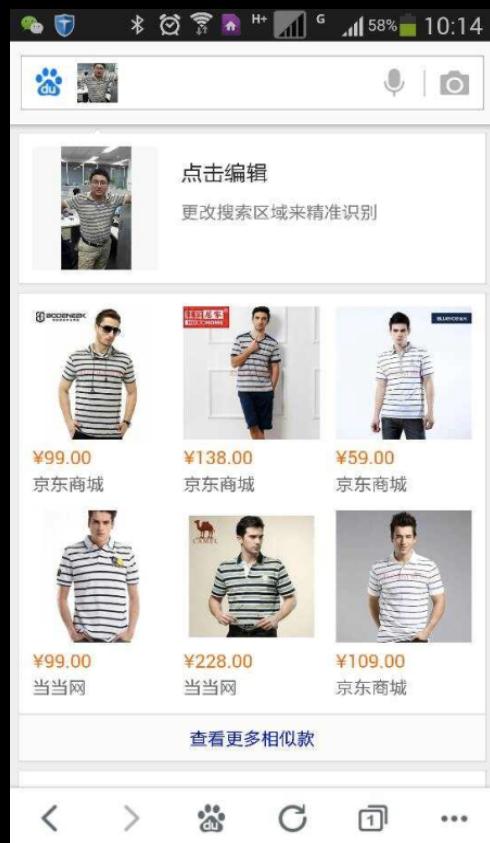
Images

Speech recognition

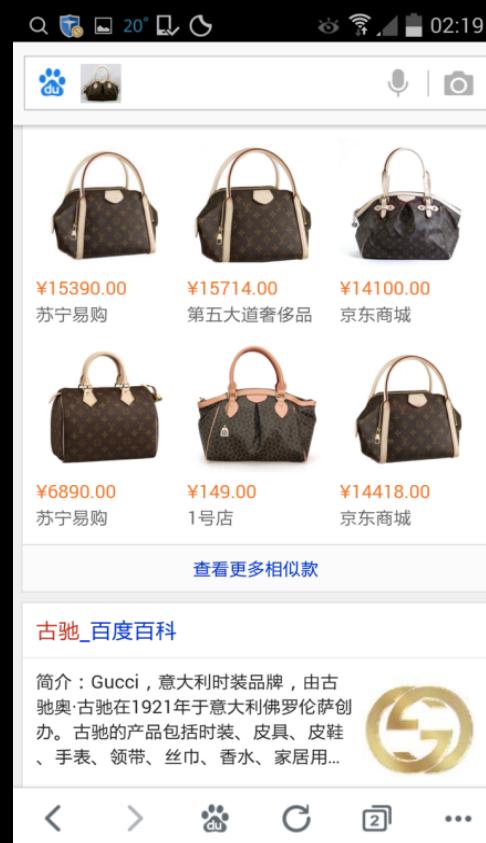


Andrew Ng

Image queries



Clothing

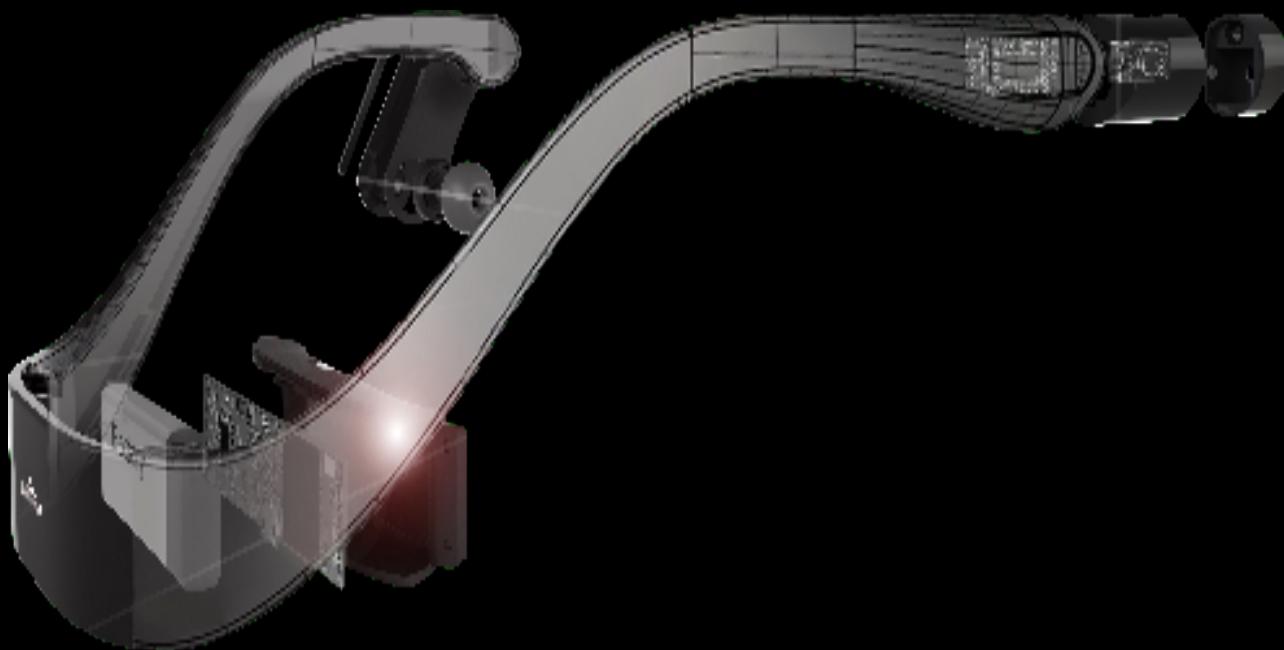


Bags



Fruits & Vegetables

Baidu Eye



Andrew Ng



Andrew Ng

“Smart glasses” designs



Andrew Ng

Extending human perception



Extending human perception



Comparison to “smart glasses” designs



Andrew Ng



Speech



Images



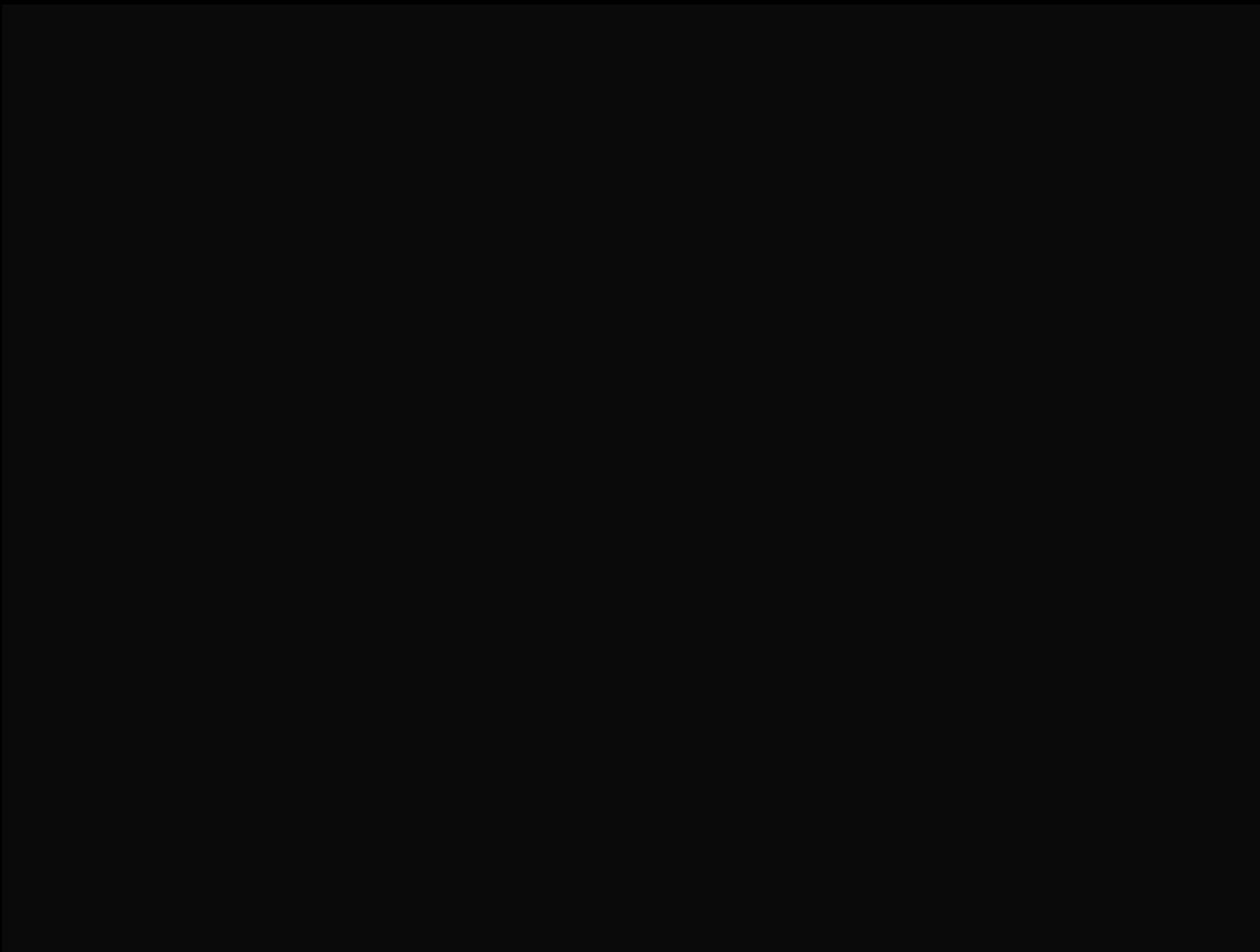
Text

From Control to Perception



Andrew Ng

Stanford's PR-1 robot



[Ken Salisbury]

AI will transform the internet

Technology areas with potential for paradigm shift:

- Computer vision
- Speech recognition & speech synthesis
- Language understanding: Machine translation;
Web search; Dialog systems;
- Advertising
- Personalization/recommendation systems
- Robotics

All this is hard: scalability, algorithms.