



Computer Vision: The Modern Approach

David Forsyth
Professor
University of Illinois, Urbana Champaign

February 20-22, 2013 | Orlando World Marriott Center | Orlando, Florida USA

What do we do in Computer Vision?

- Two major intertwining themes
 - Reconstruction
 - Build me a model of it
 - Recognition
 - What is this like
- Wildly successful field
 - 20 years ago:
 - eccentric preoccupation of few
 - Now:
 - massive impact, including numerous applications



Reconstruction - SOA

- Multiple cameras
 - Incredible results at huge geometric scales
 - centimeters of error from multiple building quadrangles
 - essentially automatic
 - video is better than scattered single images, but...
- Single views
 - much more difficult, but some progress
 - cues include: symmetry, stylized shapes, contour, texture, shading





M. Pollefeys, L. Van Gool, M. Vergauwen, F. Verbiest, K. Cornelis, J. Tops, R. Koch, Visual modeling with a hand-held camera, International Journal of Computer Vision 59(3), 207-232, 2004



ROBOTIC INDUSTRIES ASSOCIATION



Why is visual object recognition useful?

- If you want to act, you must draw distinctions
- For robotics
 - recognition can predict the future
 - is the ground soggy?
 - is that person doing something dangerous?
 - does it matter if I run that over?
 - which end is dangerous?
- For information systems
 - recognition can unlock value in pictures
 - for search, clustering, ordering, inference, ...
- General engineering
 - recognition can tell what people are doing
- If you have vision, you have some recognition system



Words and Pictures Affect One Another



Marc by Marc Jacobs
Adorable peep-toe pumps, great for any occasion.
Available in an array of uppers. Metallic fabric trim
and bow detail. Metallic leather lined footbed.
Lined printed design.
Leather sole. 3 3/4" heel.

Zappos.com



soft and glassy patent calfskin trimmed with
natural vachetta cowhide, open top satchel for
daytime and weekends, interior double slide
pockets and zip pocket, seersucker stripe cotton
twill lining, kate spade leather license plate logo,
imported
2.8" drop length
14"h x 14.2"w x 6.9"d

Katespade.com



It's the perfect party dress. With distinctly feminine
details such as a wide sash bow around an empire
waist and a deep scoopneck, this linen dress will
keep you comfortable and feeling elegant all
evening long. Measures 38" from center back, hits
at the knee.

- * Scoopneck, full skirt.
- * Hidden side zip, fully lined.
- * 100% Linen. Dry clean.

bananarepublic.com

E-commerce transactions
and \$198 billion (Forrester Research).



Conclusion

- Recognition is subtle
 - strong basic methods based on classifiers
- Important recognition technologies coming
 - the unfamiliar
 - phrases
 - geometry
 - selection
- Crucial open questions
 - dataset bias
 - links to utility

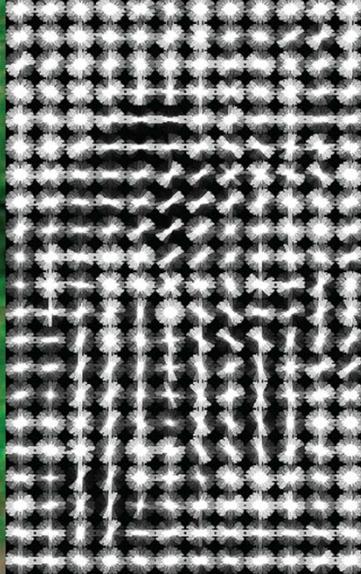


A Belief Space About Recognition

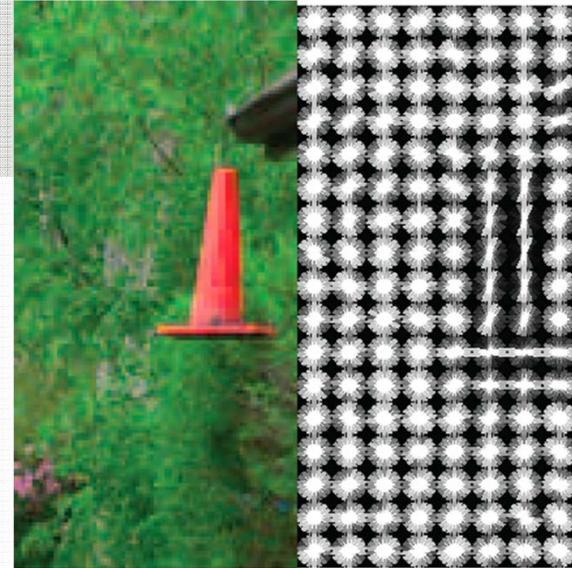
Platonism?

- Object categories are fixed and known
 - Each instance belongs to one category of k
- Good training data for categories is available
- Object recognition=k-way classification
- Detection = lots of classification

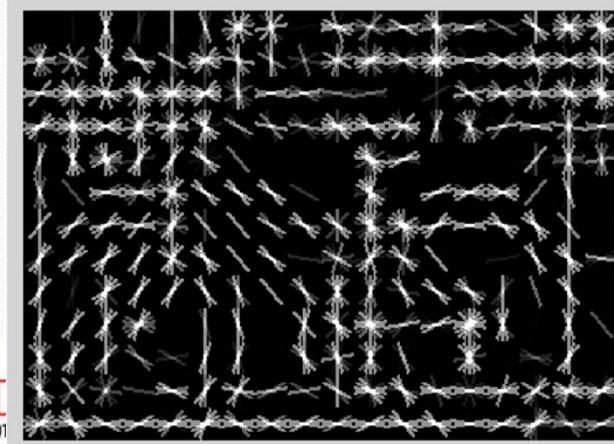




Features

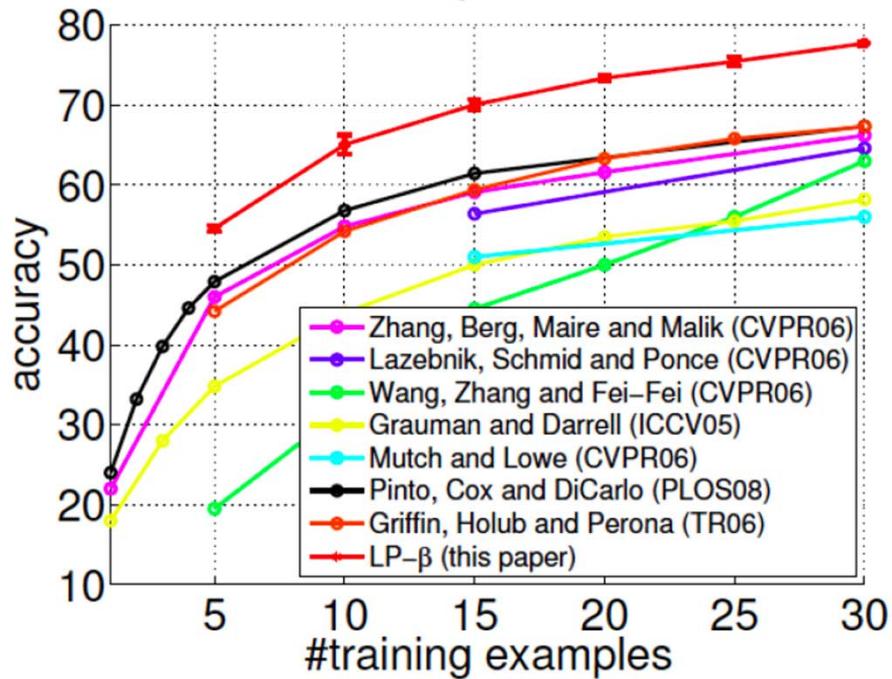


- Principles
 - illumination invariant (robust) -> gradient orientation features
 - windows always slightly misaligned -> local histograms
- HOG, SIFT features (Lowe, 04; Dalal+Triggs 05)

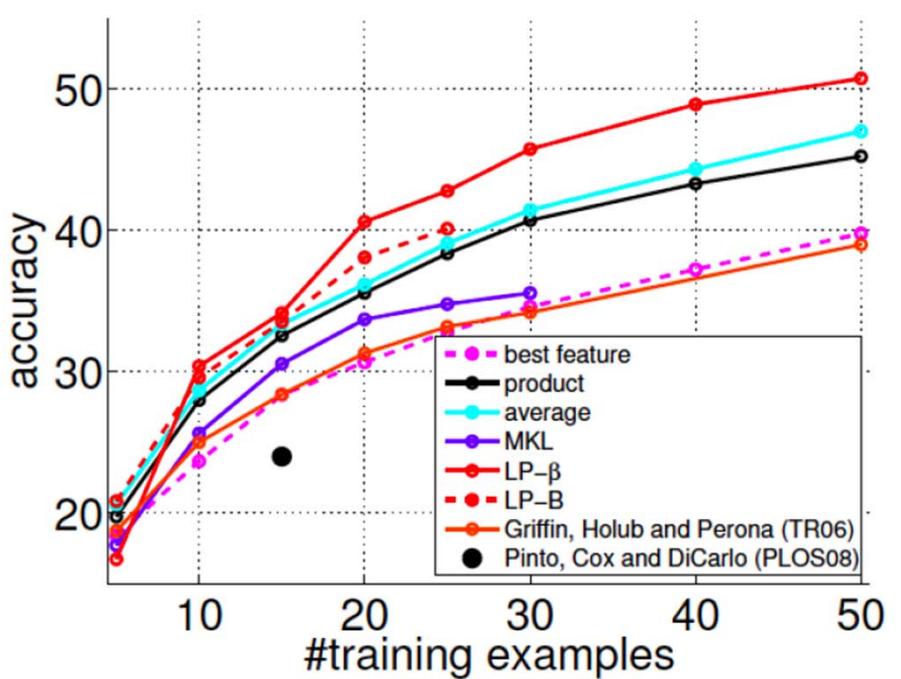


Classification Works Well

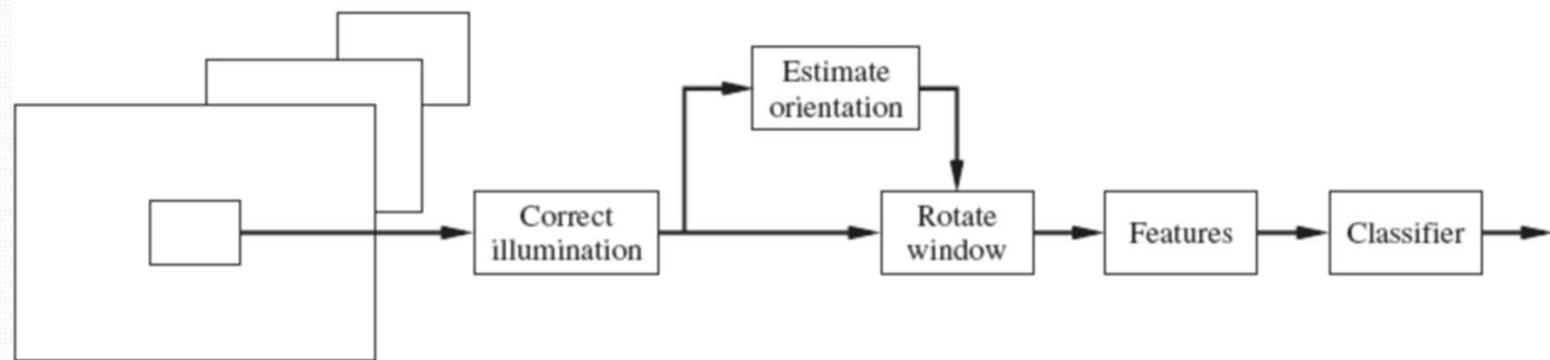
Caltech101 comparison to literature

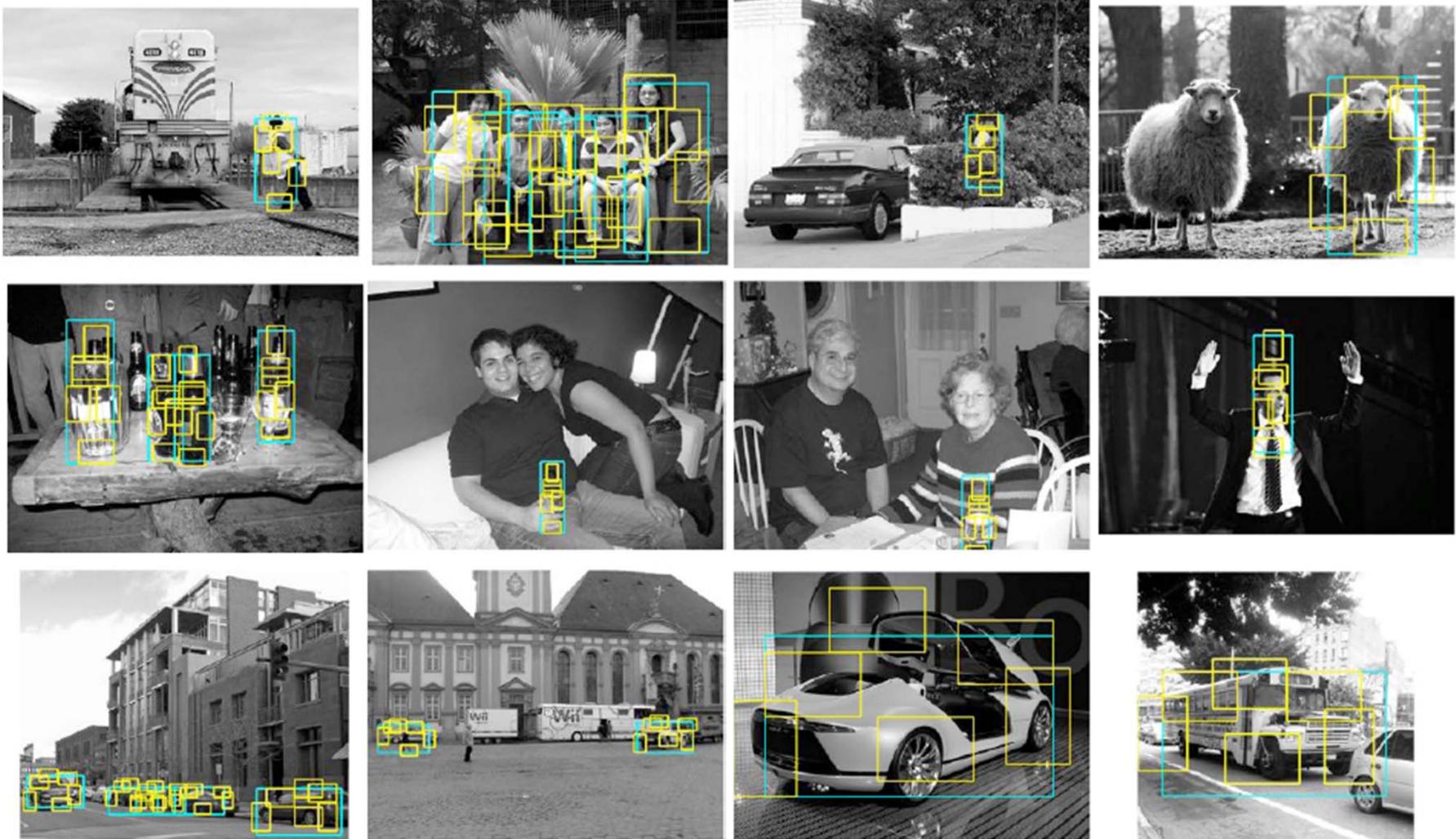


Caltech-256 (39 kernels)



Detection with a Classifier





J. Felzenszwalb, D. McAllester, D. Ramanan. "A Discriminatively Trained, Multiscale, Deformable Part Model" CVPR 2008.

A Belief Space About Eecognition

Platonism?

- Object categories are fixed and known
 - Each instance belongs to one category of k
 - Good training data for categories is available
 - Object recognition=k-way classification
 - Detection = lots of classification
- Obvious nonsense Obvious nonsense Obvious nonsense



Are these monkeys?



Spider **Monkey**, Spider **Monkey**
Profile ...
470 x 324 - 29k - jpg
animals.nationalgeographic.com
[More from animals.nationalgeographic.com] www.bestweekever.tv]



OMFG **MONKEY**
NIPSY2.
444 x 398 - 40k - jpg
www.bestweekever.tv
[More from www.bestweekever.tv]



Vampire **Monkey**
350 x 500 - 32k - jpg
paranormal.about.com



... monkeys for ...
424 x 305 - 21k - jpg
thebit.t.com



The **Monkey Cage**
300 x 306 - 35k - jpg
www.themonkeycage.org



... be **monkey** ...
300 x 350 - 29k - jpg
my.opera.com



... **monkey's** interests ...
378 x 470 - 85k - jpg
www.schwimmerlegal.com



"You will be a **monkey**.
358 x 480 - 38k - jpg
kulxp.blogspot.com



... **monkey** and I am ...
342 x 324 - 17k - jpg
www.azcazandco.com



Monkey
353 x 408 - 423k - bmp
www.graphicshunt.com



The **Monkey Park**
400 x 402 - 24k - jpg
www.lysator.liu.se



Monkey cloning follow
up ...
450 x 316 - 17k - jpg
blog.bioethics.net



So here's one of my
monkeys.
400 x 300 - 13k - jpg
www.gamespot.com



monkeys ...
400 x 310 - 85k - jpg
joaquinvargas.com



MONKEY TEETH
308 x 311 - 18k - jpg
repairstemcell.wordpress.com



The **Blow Monkey** is ...
500 x 500 - 30k - jpg
www.uberreview.com



Spider **Monkey** Picture, Spider
Monkey ...
800 x 600 - 75k - jpg
animals.nationalgeographic.com



a..... **monkey!**
mammal **monkey**
525 x 525 - 99k - jpg
www.sodahead.com



WTF **Monkey**
374 x 300 - 23k - jpg
www.myspace.com



Monkey
512 x 768 - 344k - jpg
www.exzoobearance.com



Monkeys ...
787 x 1024 - 131k - jpg
runrigging.blogspot.com

Conclusion

- Recognition is subtle
 - strong basic methods based on classifiers
- Important recognition technologies coming
 - the unfamiliar
 - phrases
 - geometry
 - sentences
- Crucial open questions
 - dataset bias
 - links to utility



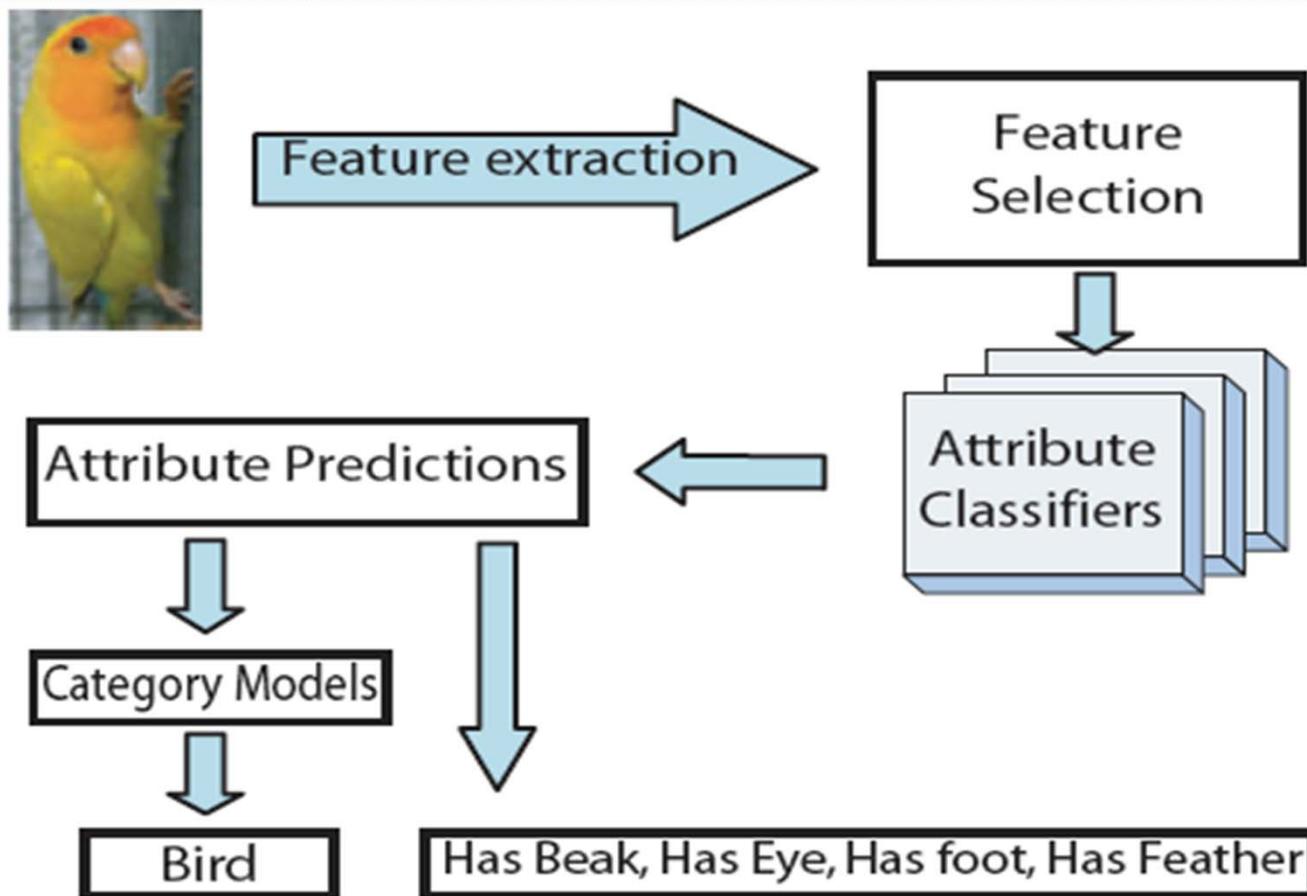
The Unfamiliar



The Unfamiliar



General Architecture



Farhadi et al 09; cf Lampert et al 09

Attribute Predictions for Unknown Objects

						
'is 3D Boxy' 'is Vert Cylinder' 'has Window' 'has Row Wind' X'has Headlight'	'has Hand' 'has Arm' X'has Screen' 'has Plastic' X'has Headlight'	'has Head' 'has Hair' 'has Face' X'has Saddle' 'is Shiny'	'has Head' 'has Torso' 'has Arm' 'has Leg'	'has Head' 'has Ear' 'has Snout' 'has Nose' 'has Mouth'	'has Head' 'has Ear' 'has Snout' 'has Mouth' 'has Leg'	X'has Furniture Back' X'has Horn' X'has Screen' 'has Plastic' 'is Shiny'
						
'is 3D Boxy' 'has Wheel' 'has Window' 'is Round' 'has Torso'	'has Tail' 'has Snout' 'has Leg' X'has Text' X'has Plastic'	'has Head' 'has Ear' 'has Snout' 'has Leg' 'has Cloth'	'is Horizontal Cylinder' X'has Beak' X'has Wing' X'has Side mirror' 'has Metal'			

Farhadi et al 09; cf Lampert et al 09

Conclusion

- Recognition is subtle
 - strong basic methods based on classifiers
 - many meanings, useful in different contexts
- Important recognition technologies coming
 - attributes
 - **phrases**
 - geometry
 - sentences
- Crucial open questions
 - dataset bias
 - links to utility



Meaning Comes in Clumps



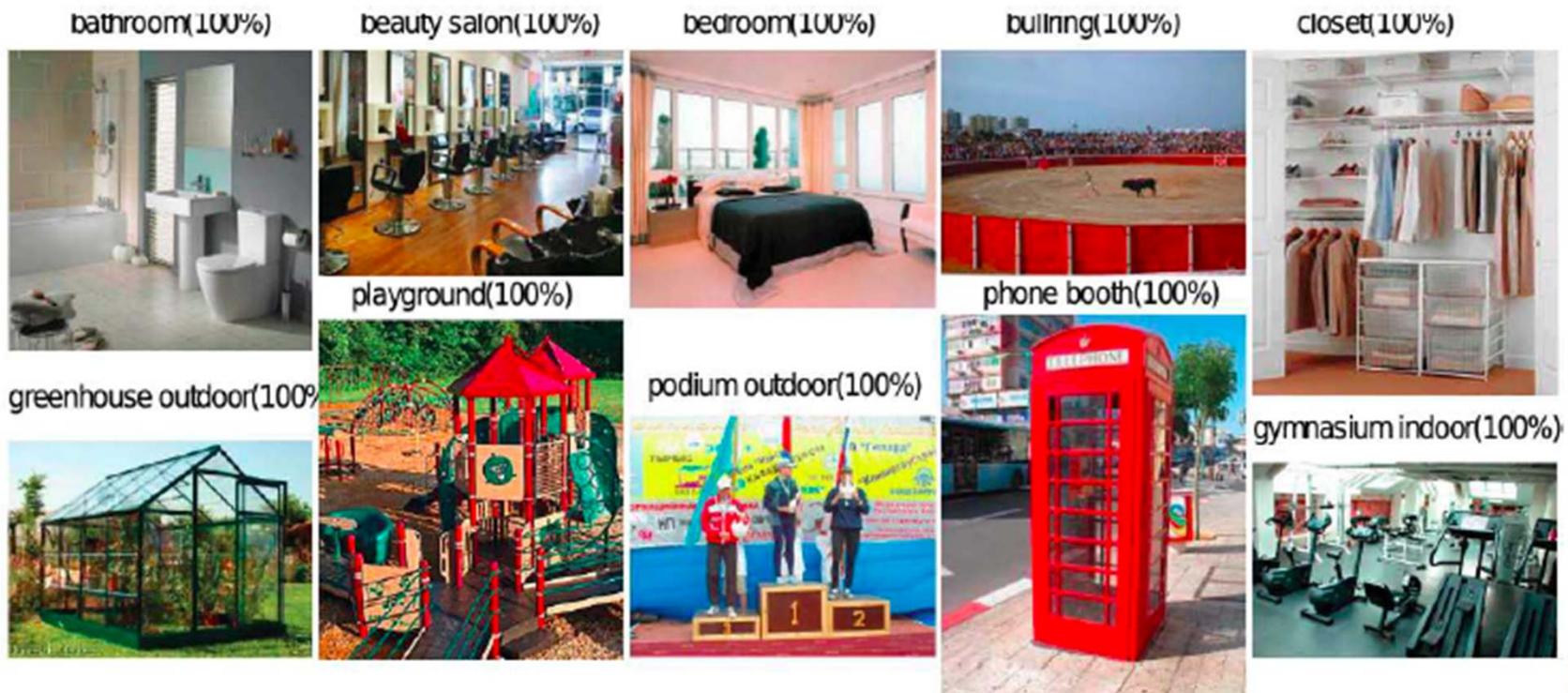
“Sledder”
Is this one thing?
Should we cut her off her sled?



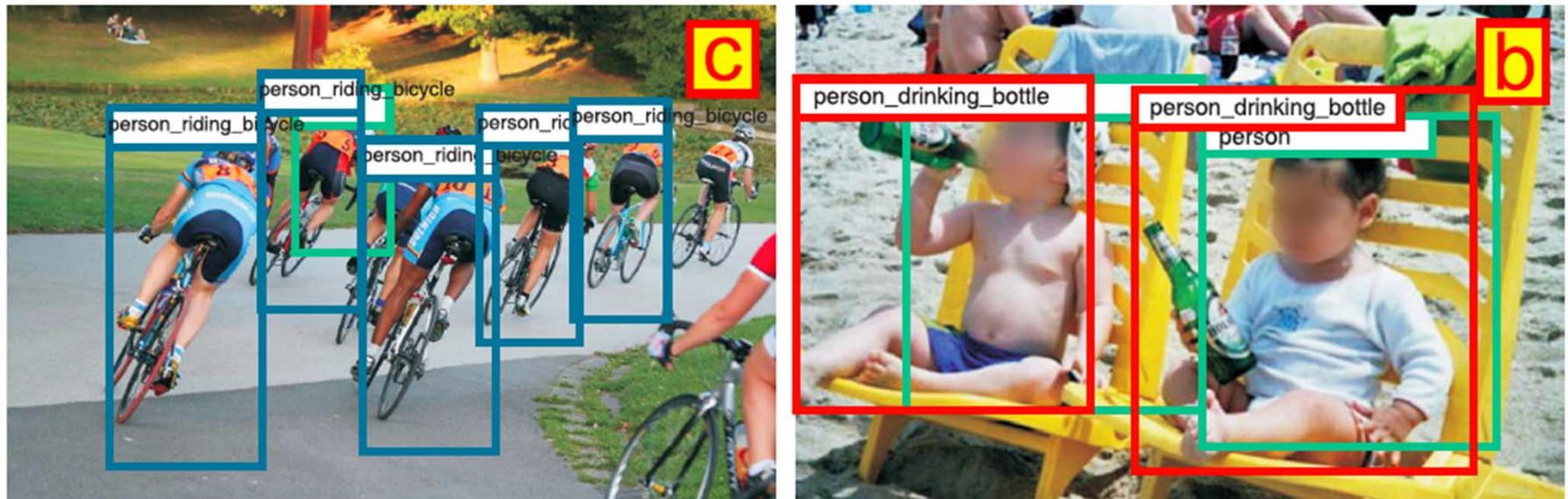
Scenes

- Likely stages for
 - Particular types of object
 - Particular types of activity

Xiao et al 10



Scenes > Visual phrases > Objects



- Composites
 - easier to recognize than their components
 - because appearance is simpler

Farhadi + Sadeghi 11



Decoding



Farhadi Sadeghi 11



Decoding Helps

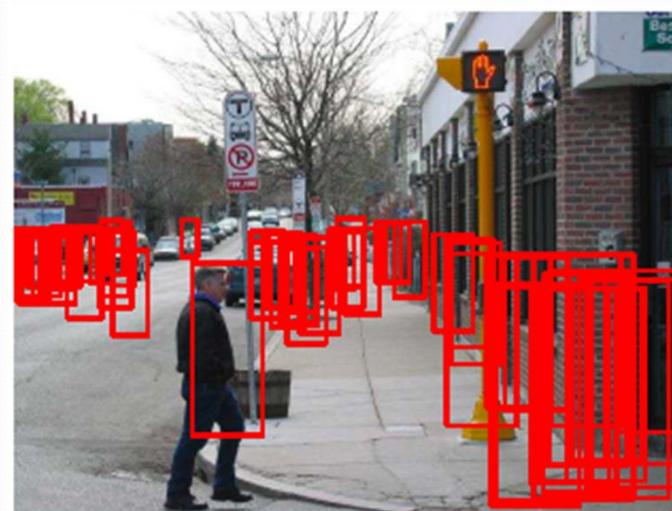


Conclusion

- Recognition is subtle
 - strong basic methods based on classifiers
- Important recognition technologies coming
 - the unfamiliar
 - phrases
 - **geometry**
 - selection
- Crucial open questions
 - dataset bias
 - links to utility

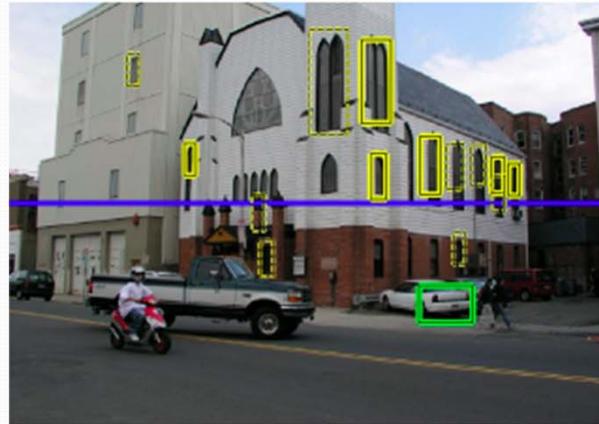


Environmental Knowledge

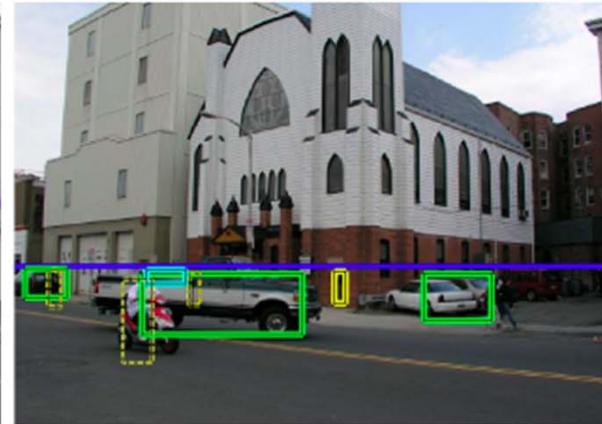


Hoiem et al 06

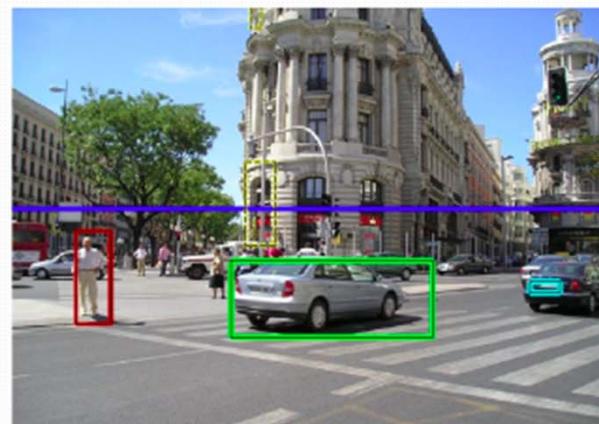
Environmental Knowledge is Powerful



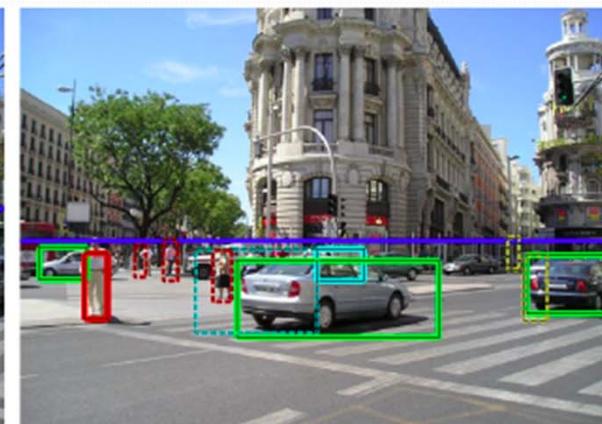
(b) Local Detection



(b) Full Model Detection



(a) Local Detection



(a) Full Model Detection

Conclusion

- Recognition is subtle
 - strong basic methods based on classifiers
 - many meanings, useful in different contexts
- Important recognition technologies coming
 - attributes
 - phrases
 - geometry
 - **selection**
- Crucial open questions
 - dataset bias
 - links to utility



Selection: What is worth saying?



Two girls take a break to sit and talk .

Two women are sitting , and one of them is holding something .

Two women chatting while sitting outside

Two women sitting on a bench talking .

Two women wearing jeans , one with a blue scarf around her head , sit and talk .

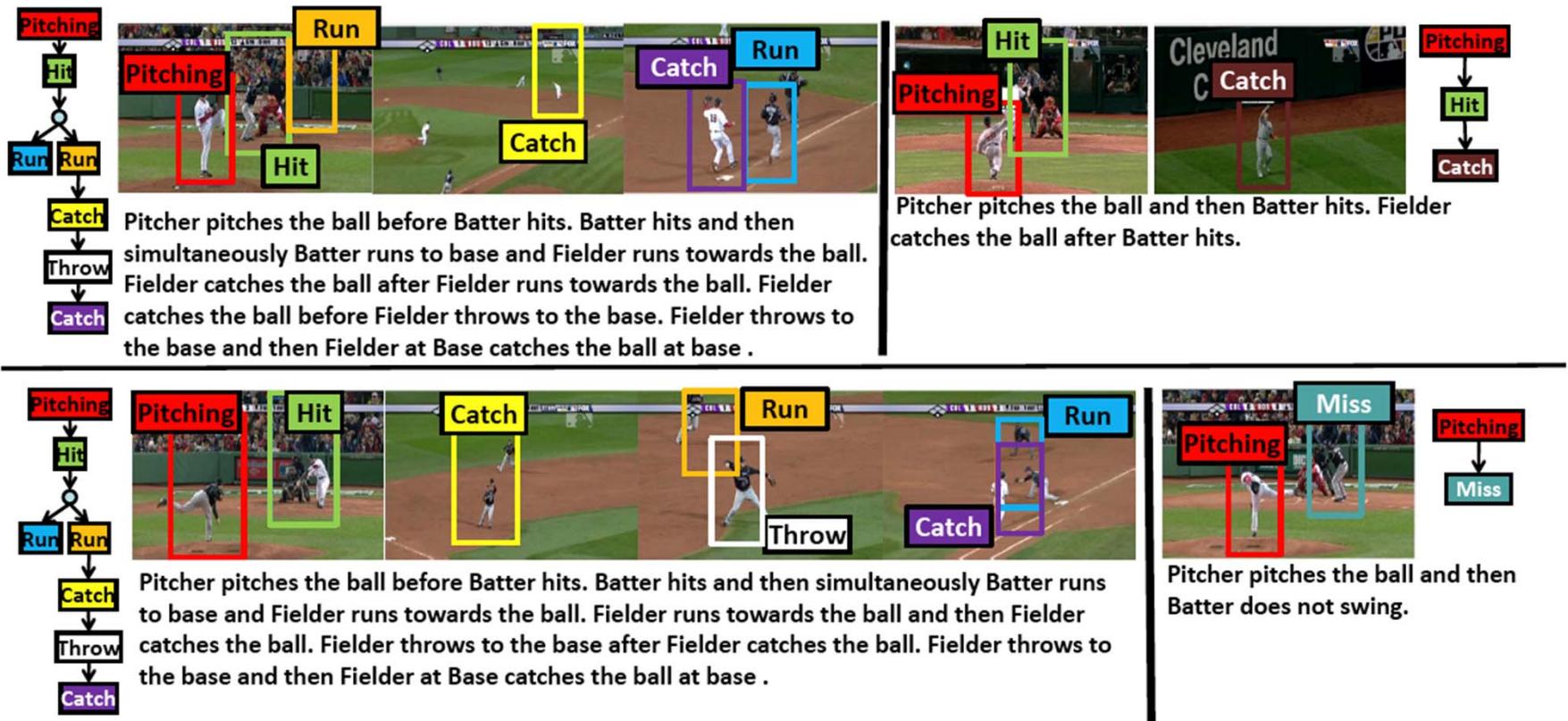
Sentences from Julia Hockenmaier's work

For language people: Pragmatics - what is worth saying?

Rashtchian ea 10



Describing Structured Events



Gupta ea 09

Adding Attributes and Prepositions



This is a photograph of one sky, one road and one bus. The blue sky is above the gray road. The gray road is near the shiny bus. The shiny bus is near the blue sky.



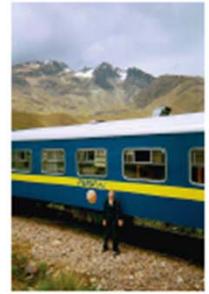
There are two aeroplanes. The first shiny aeroplane is near the second shiny aeroplane.



There are one cow and one sky. The golden cow is by the blue sky.



There are one dining table, one chair and two windows. The wooden dining table is by the wooden chair, and against the first window, and against the second white window. The wooden chair is by the first window, and by the second white window. The first window is by the second white window.



Here we see one person and one train. The black person is by the train.



This is a picture of one sky, one road and one sheep. The gray sky is over the gray road. The gray sheep is by the gray road.



Here we see one road, one sky and one bicycle. The road is near the blue sky, and near the colorful bicycle. The colorful bicycle is within the blue sky.



Here we see two persons, one sky and one aeroplane. The first black person is by the blue sky. The blue sky is near the shiny aeroplane. The second black person is by the blue sky. The shiny aeroplane is by the first black person, and by the second black person.



This is a picture of two dogs. The first dog is near the second furry dog.



This is a photograph of two buses. The first rectangular bus is near the second rectangular bus.

Conclusion

- Recognition is subtle
 - strong basic methods based on classifiers
 - many meanings, useful in different contexts
- Important recognition technologies coming
 - attributes
 - phrases
 - geometry
 - sentences
- Crucial open questions
 - dataset bias
 - links to utility



Bias is Pervasive

lion - Google Search

6/8/10 12:25

Web Images Videos Maps News Shopping Gmail more ▾

Search settings | Sign in

SafeSearch off ▾

Google

lion

About 23,100,000 results (0.05 seconds)

Advanced search

Related searches: lion roaring, lioness, lion drawing, lion tattoo

Everything

Images

Videos

More

Any size

Medium

Large

Icon

Larger than...

Exactly...

Any type

Face

Photo

Clip art

Line drawing

Any color

Full color

Black and white

□□□□

□□□□

Lions Kill Giraffe
479 x 450 - 48k - jpg
abolitionist.com
Find similar images

Lion on Horseback
468 x 393 - 39k - jpg
raincoaster.com
Find similar images

3, Lion
434 x 341 - 41k - jpg
bluepyramid.org
Find similar images

Interestingly, the
470 x 324 - 30k - jpg
bostonherald.com
Find similar images

Description : Asian
792 x 768 - 99k - jpg
photocase.org
Find similar images

I was doing research
on
400 x 300 - 27k - jpg
lowkayhwa.com
Find similar images

Lion Tiger Size
500 x 553 - 65k - jpg
indravit.wordpress.com
Find similar images

Lion Park, South
450 x 300 - 30k - jpg
africa-nature-photog...
Find similar images

Lion Limited
500 x 500 - 76k - jpg
onlineartdemos.co.uk
Find similar images

Lion
395 x 480 - 47k - jpg
ibexcinc.wordpress.com
Find similar images

lions
1200 x 800 - 243k - jpg
lifeasastudentnurse...
Find similar images

African Lion
500 x 333 - 57k - jpg
itsnature.org
Find similar images

LIONS:
604 x 800 - 225k - jpg
edge.org
Find similar images

Lion. Panthera leo
459 x 480 - 35k - jpg
shoarns.com
Find similar images

lions, cuddle
620 x 400 - 70k - jpg
telegraph.co.uk
Find similar images

lion
350 x 504 - 28k - jpg
sodahead.com
Find similar images

LION!
500 x 385 - 74k - jpg
firermice.wordpress.com
Find similar images

Starring horse-riding
800 x 626 - 53k - jpg
dailymail.co.uk
Find similar images

Picture: 17 stone
466 x 602 - 93k - jpg
dailymail.co.uk
Find similar images

human-lion
470 x 324 - 31k - jpg
seesdifferent...
Find similar images

Lion at Sunset
400 x 316 - 25k - jpg
art.com
Find similar images

http://www.google.com/images?hl=en&source=imghp&q=lion&gbv=2&aq=f&aql=g10&aql=&oq=&gs_rfai=

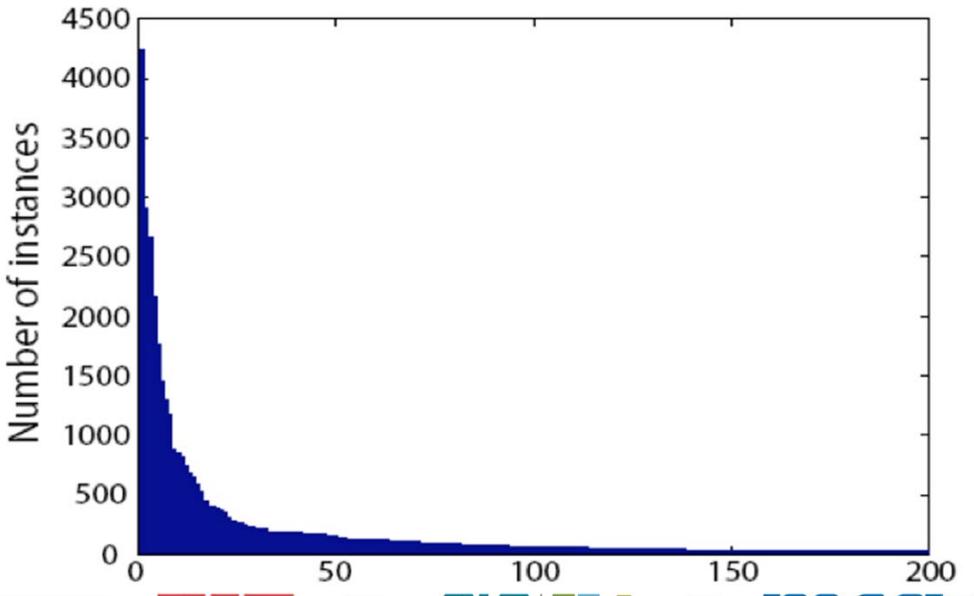
Page 1 of 2

Object Recognition is Biased

- The world can't be like the dataset because
 - many things are rare
 - this exaggerates bias



Instance number of the top 200 object categories



Defenses Against Bias

- Appropriate feature representations
 - e.g. illumination invariance
- Appropriate intermediate representations
 - which could have less biased behavior
 - perhaps attributes? scenes? visual phrases?
- Appropriate representations of knowledge
 - e.g. geometry --- pedestrian example



What should we say about visual data?

- Most important question in vision
 - What does the output of a recognition system consist of?
- A useful representation of reasonable size
 - dubious answer
 - Useful in what way?
 - How do we make the size reasonable?



Object Categories Depend on Utility



Monkey or Plastic toy or both or irrelevant

Some of this depends on what you're trying to do, in ways we don't understand



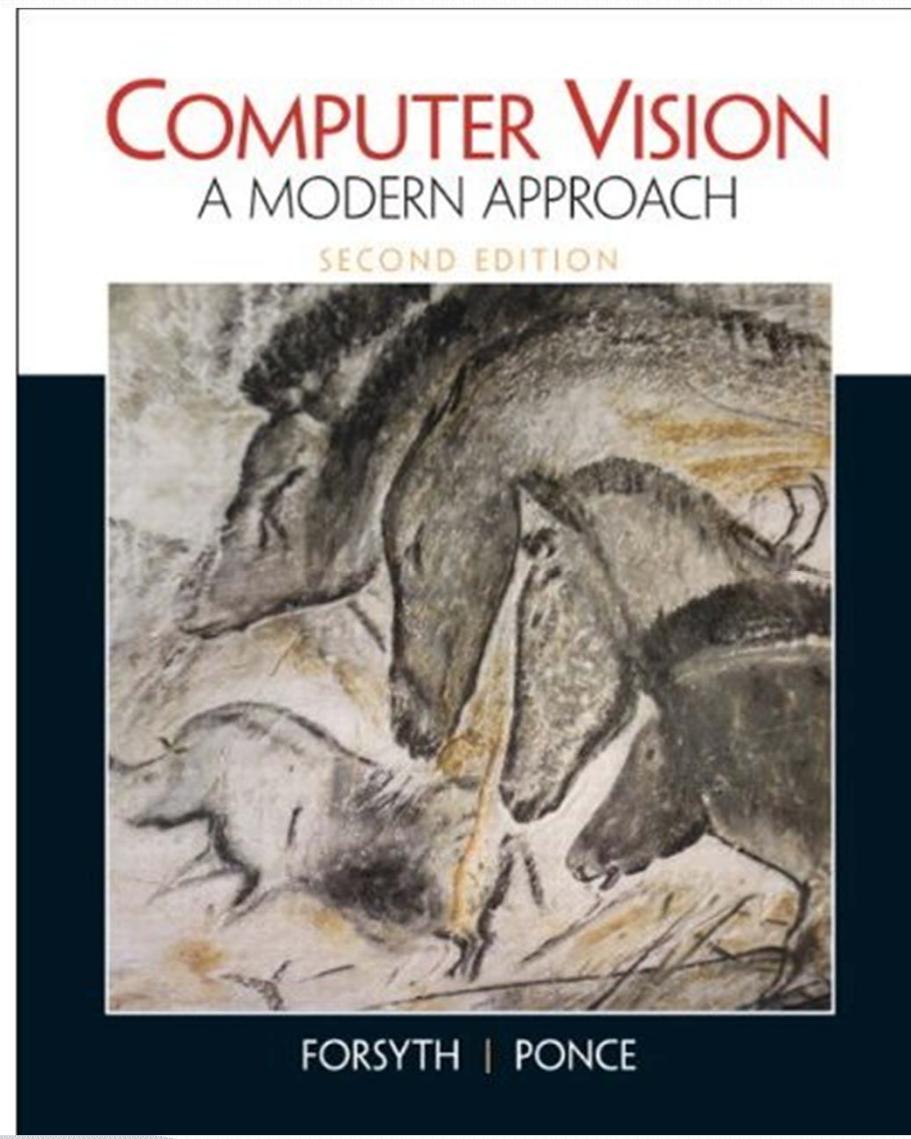
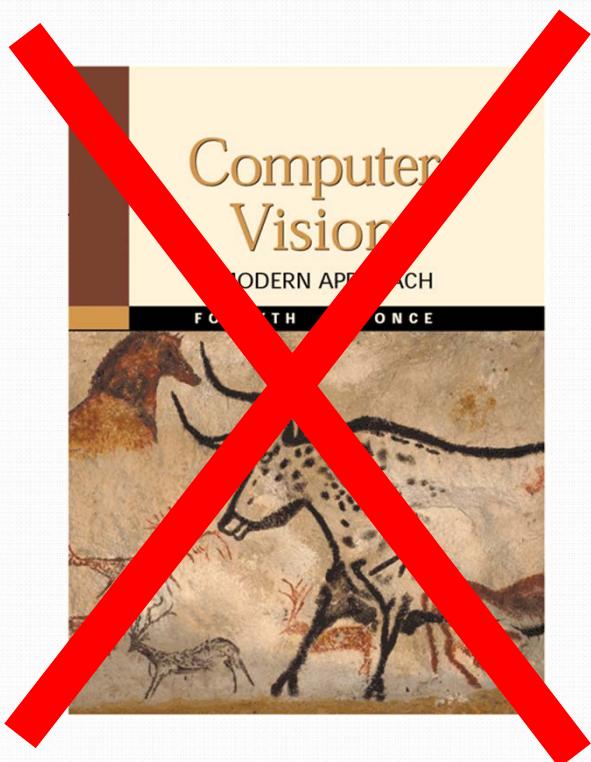
Person or child or beer drinker or beer-drinking child or tourist or holidaymaker or obstacle or potential arrest or irrelevant or...

Conclusion

- Recognition is subtle
 - strong basic methods based on classifiers
- Important recognition technologies coming
 - the unfamiliar
 - phrases
 - geometry
 - selection
- Crucial open questions
 - dataset bias
 - links to utility



More Information



Contact Information

David Forsyth

Professor

University of Illinois

Computer Science Department

Urbana-Champaign, Illinois 61801

USA

Email: daf@uiuc.edu

Web: <http://luthuli.cs.uiuc.edu/~daf>

