

Words and Pictures: more structure

D.A. Forsyth, UIUC

with

Okan Arikan (UT Austin), Nazli Ikizler (Boston U), Leslie Ikemoto (animate-me), Derek Hoiem (UIUC), Ali Farhadi (UIUC), Ian Endres (UIUC), Ryan White (Euclid media)

Computer vision

Obtain dataset

Build features

Light entertainment
(the way we do it)

Mess around with classifiers, probability, etc

Computer vision

Produce representation

Big questions

Computer vision

- What signal representation should we use for recognition?



Computer vision

- What should we say about what we see?

Structure

- Correlated words
 - waves go with beaches not cats
- Attributes
 - has nose
- Adjectives
 - green hat
- Relations
 - cat on mat
- Sentences
 - A dolphin holds a basketball as it swims on its back



A **dolphin** holds a **basketball** as it swims on its **back** .

A **dolphin** swimming upside down while holding a **basketball**

A **dolphin** swims upside down holding a **basketball** between it 's **flippers** .

A **seal** floats on it 's back in the **water** , holding a **basketball** .

The **dolphin** on his **back** holds the **orange basketball**

Correlated Words

- Simple method:
 - rack up some features, build a bunch of linear classifiers one per word
 - works poorly
 - few examples per word
 - many features, only some are stable

Learn this



$$\mathcal{D} \approx \mathcal{MX}$$

Word data (observed)

Image representation (observed)

Correlated words

- Idea
 - some features are not helpful
 - a low dimensional subspace is good at predicting most things (Ando +Zhang,)
 - We can find this space by penalizing rank in the matrix of linear classifiers

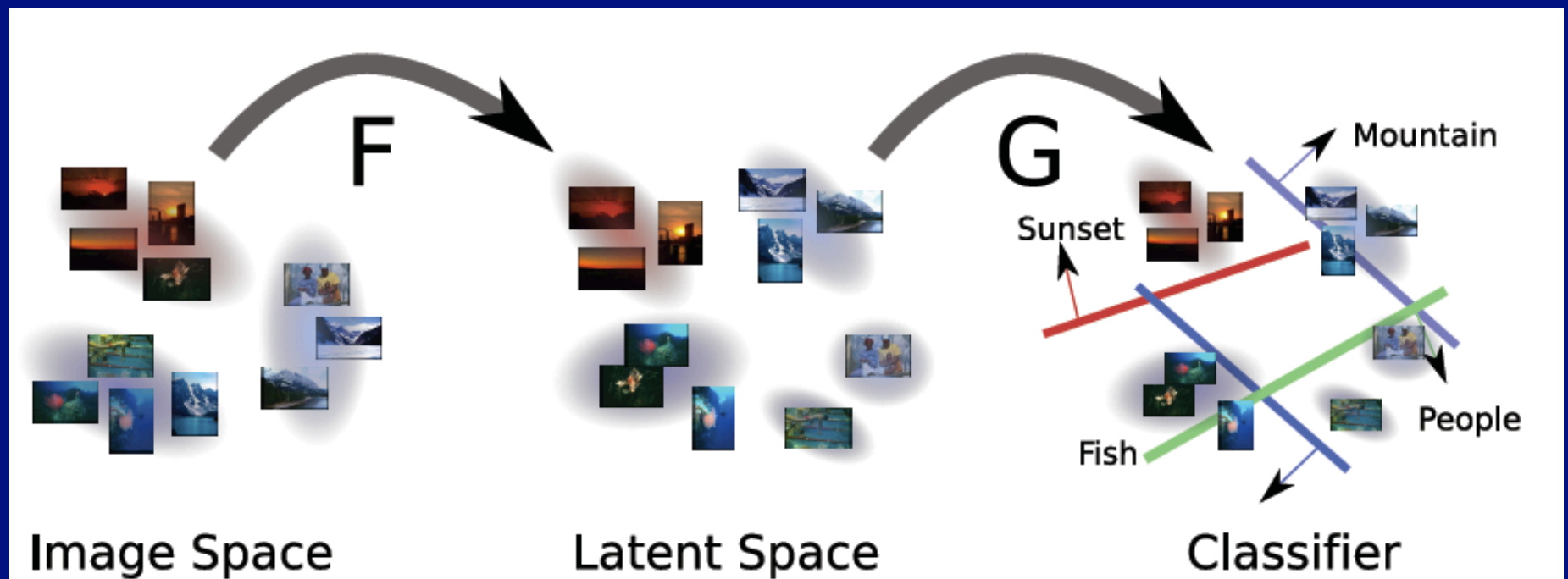
Learn this



$$\mathcal{D} \approx \mathcal{G}\mathcal{F}\mathcal{X}$$

Word data (observed)

Image representation (observed)



It was there and we didn't



sky, sun, clouds, sea,
waves, birds, water



tree, people, sand, road,
stone, statue, temple,
sculpture, pillar



tree, birds, snow, fly



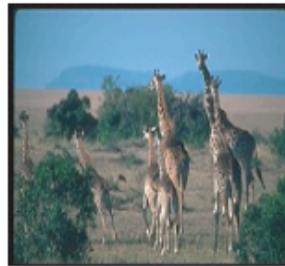
sky, water, tree, plane,
elephant, herd



mountain, sky, water,
clouds, tree



sky, sun, jet, plane



mountain, sky, water,
tree, grass, plane,
ground, giraffe



water, people, pool,
swimmers



tree, people, shadows,
road, stone, statue,
sculpture, pillar



people, buildings,
stone, temple, sculp-
ture, pillar, mosque

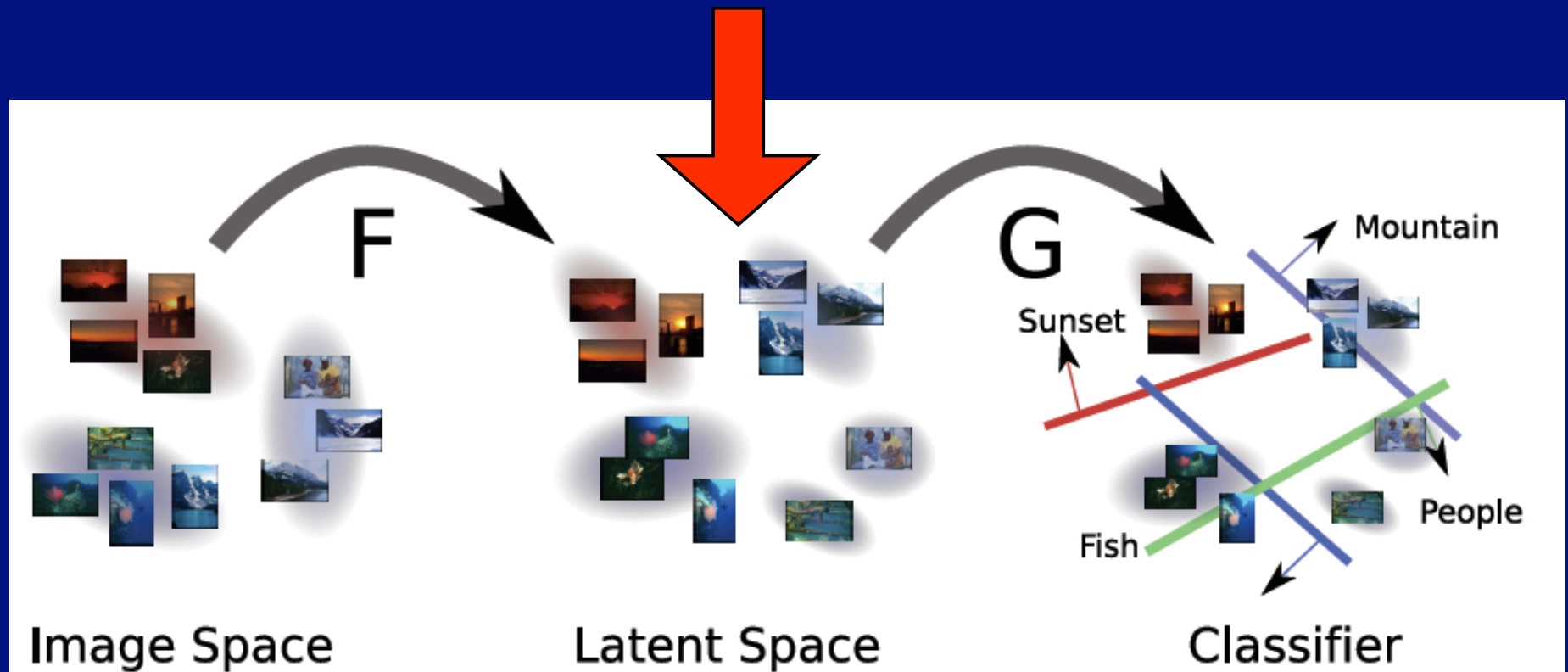
It was there and we predicted it

It wasn't and we did

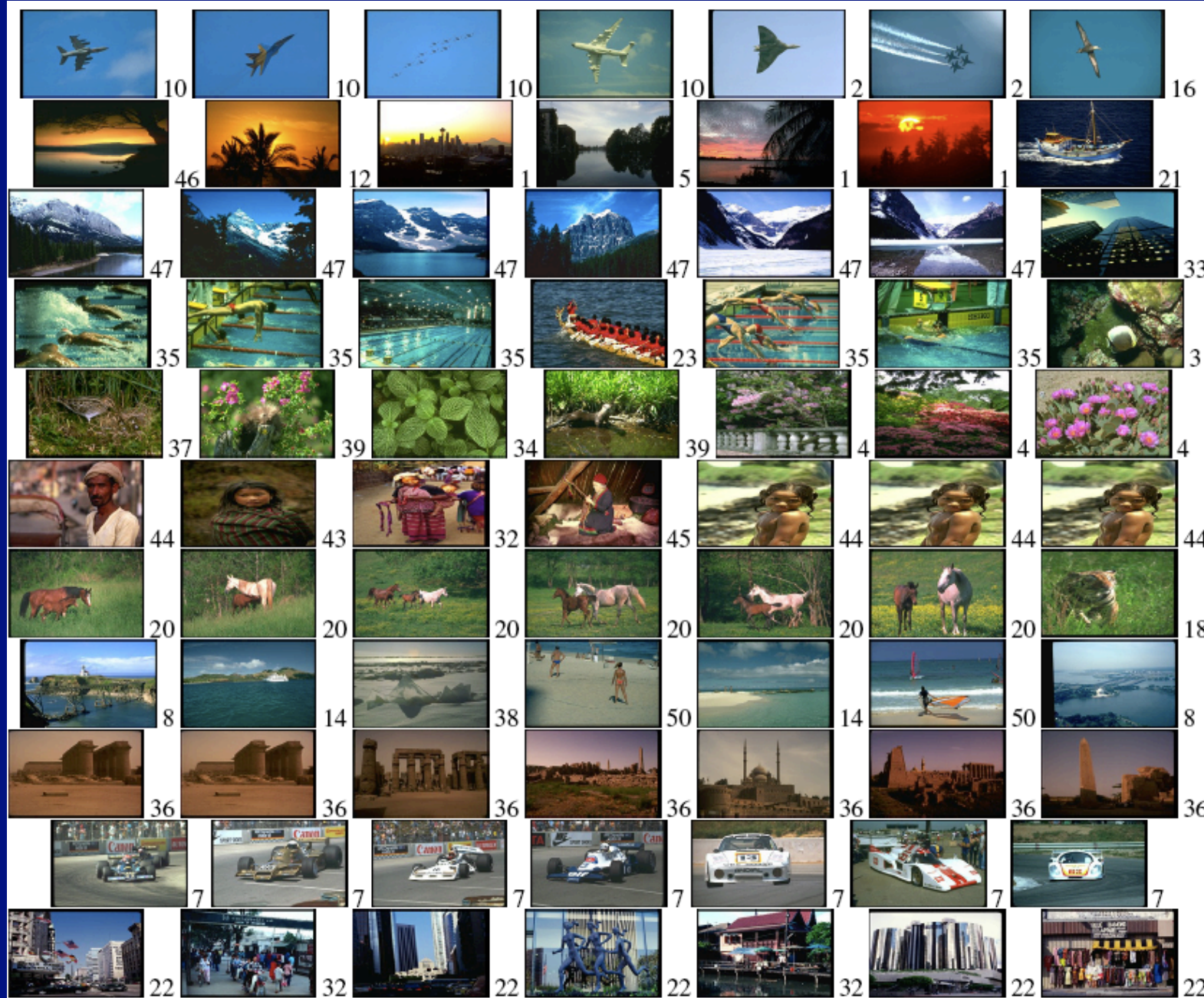
Scenes as object bags

- We could build collections of labelled scene images
 - useful, but..
 - kitchen, bathroom, outdoor, and then?
- We could collect images of similar appearance
 - but...
 - might not really have similar objects in them
- Unsupervised bag discovery
 - Pictures of the same scene tend to contain similar objects
 - i.e. tend to attract the same image annotations

In this space, images are “close” if they “look similar” AND if they predict “similar” words



Scene



CD #
(rough proxy)

Correlated word predictors are quite good

Method	P	R	F1	Ref
Co-occ	0.03	0.02	0.02	[53]
Trans	0.06	0.04	0.05	[27]
CMRM	0.10	0.09	0.10	[37]
TSIS	0.10	0.09	0.10	[19]
MaxEnt	0.09	0.12	0.10	[39]
CRM	0.16	0.19	0.17	[44]
CT-3×3	0.18	0.21	0.19	[82]
CRM-rect	0.22	0.23	0.23	[31]
InfNet	0.17	0.24	0.23	[50]
MBRM	0.24	0.25	0.25	[31]
MixHier	0.23	0.29	0.26	[17]
(section 2.2)	0.27	0.27	0.27	
(section 2.2, kernel)	0.29	0.29	0.29	
PicSOM	0.35*	0.35*	0.35*	[73]

Structure

- Correlated words
 - waves go with beaches not cats
- **Attributes**
 - **has nose**
- **Adjectives**
 - green hat
- **Relations**
 - cat on mat
- **Sentences**
 - A dolphin holds a basketball as it swims on its back



A **dolphin** holds a **basketball** as it swims on its **back** .

A **dolphin** swimming upside down while holding a **basketball**

A **dolphin** swims upside down holding a **basketball** between it 's **flippers** .

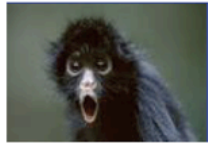
A **seal** floats on it 's back in the **water** , holding a **basketball** .

The **dolphin** on his **back** holds the **orange basketball**

Object recognition = k class classification

- current data sets ok,
 - improve coverage
- research agenda:
 - more features, better classifiers:
 - perhaps category hierarchies for statistical leverage (tying)

Are these monkeys?



Spider Monkey, Spider Monkey
Profile ...
470 x 324 - 29k - jpg
animals.nationalgeographic.com
[[More from](#)
animals.nationalgeographic.com]



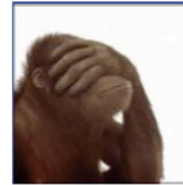
OMFG MONKEY
NIPS2.
444 x 398 - 40k - jpg
www.bestweekever.tv
[[More from](#)
www.bestweekever.tv]



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



... monkeys for ...
424 x 305 - 21k - jpg
thebitt.com



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



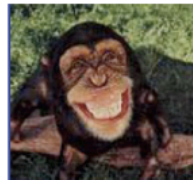
... be monkey ...
300 x 350 - 29k - jpg
my.opera.com



... 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



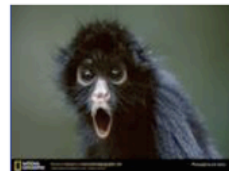
monkeys ...
400 x 310 - 85k - jpg
joaquinvargas.com



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.exzooberance.com

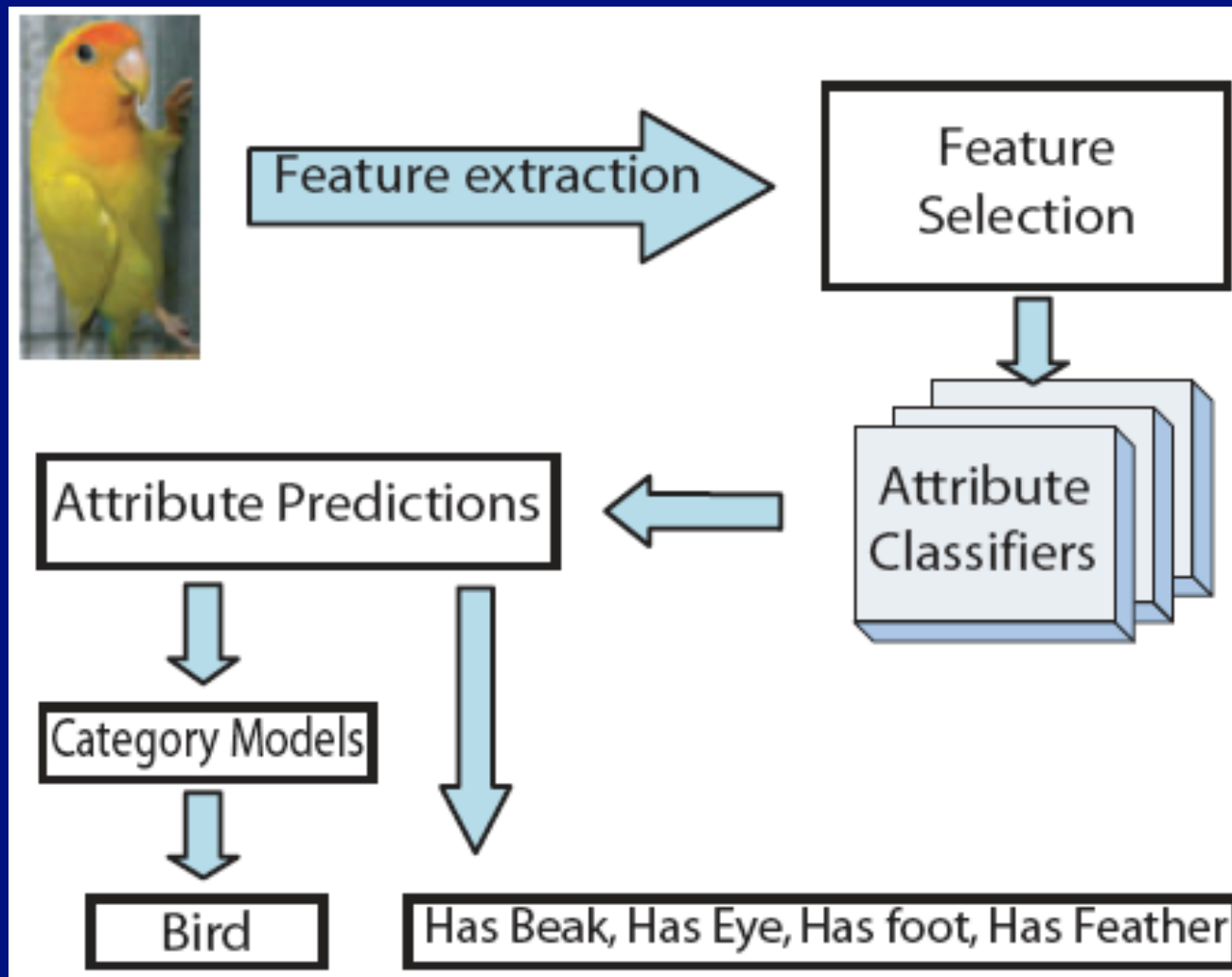


Monkeys ...
787 x 1024 - 131k - jpg
runrigging.blogspot.com

Object recognition = describing what objects are like

- current datasets
 - are largely of the wrong form
 - and no declarative data about objects
- research agenda
 - learning by reading
 - sensible responses to objects of unknown category
 - within class variance has semantics
 - architectures, representations, semantics

General architecture





'is 3D Boxy'

'is Vert Cylinder'

'has Window' ~~'has Screen'~~

'has Row Wind' ~~'has Headlight'~~



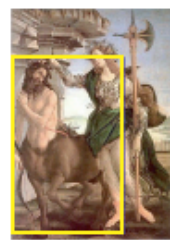
'has Hand'

'has Arm'

~~'has Screen'~~

'has Plastic'

'is Shiny'



'has Head'

'has Hair'

'has Face'

~~'has Saddle'~~

'has Skin'



'has Head'

'has Torso'

'has Arm'

'has Leg'

~~'has Wood'~~



'has Head'

'has Ear'

'has Snout'

'has Nose'

'has Mouth'



'has Head'

'has Ear'

'has Snout'

'has Mouth'

'has Leg'



~~'has Furniture Back'~~

~~'has Horn'~~

~~'s Screen'~~

'has Plastic'

'is Shiny'



'is 3D Boxy'

'has Wheel'

'has Window'

'is Round'

'has Torso'



'has Tail'

'has Snout'

'has Leg'

~~'has Text'~~

~~'has Plastic'~~



'has Head'

'has Ear'

'has Snout'

'has Leg'

'has Cloth'



'is Horizontal Cylinder'

~~'has Beak'~~

~~'has Wing'~~

~~'has Side mirror'~~

'has Metal'



'has Head'

'has Snout'

'has Horn'

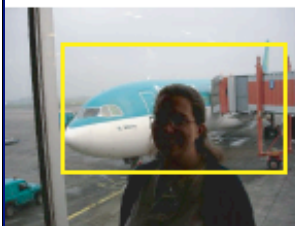
'has Torso'

~~'has Arm'~~

How is an object different from typical?

- Pragmatics suggests this is how adjectives are chosen
 - If we are sure it's a cat, and we know that
 - an attribute is different from normal
 - the detector is usually reliable
 - we should report the missing/extra attribute

Missing attributes



Aeroplane
No "wing"



Car
No "window"



Boat
No "sail"



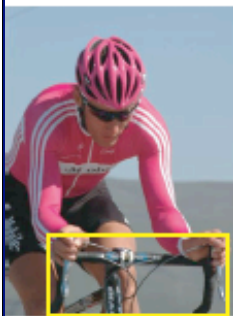
Aeroplane
No "jet engine"



Motorbike
No "side mirror"



Car
No "door"



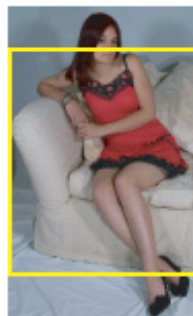
Bicycle
No "wheel"



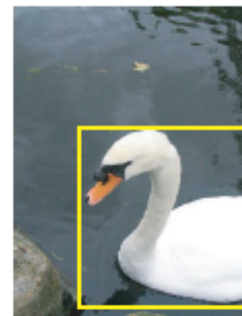
Sheep
No "wool"



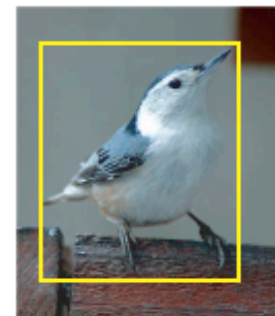
Train
No "window"



Sofa
No "wood"



Bird
No "tail"



Bird
No "leg"



Bus
No "door"

Extra attributes



Bird
"Leaf"



Bus
"face"



Motorbike
"cloth"



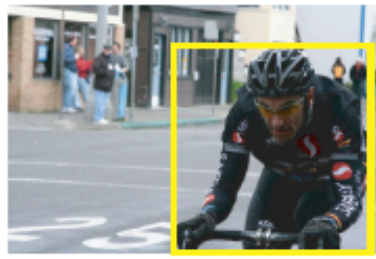
DiningTable
"skin"



People
"Furn.back"



Aeroplane
"beak"



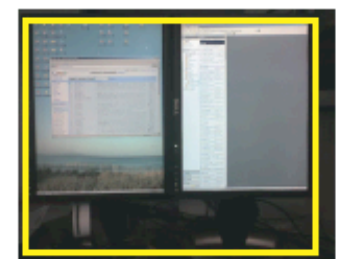
People
"label"



Sofa
"wheel"



Bike
"Horn"



Monitor
window"

Structure

- Correlated words
 - waves go with beaches not cats
- Attributes
 - has nose
- Adjectives
 - green hat
- Relations
 - cat on mat
- Sentences
 - A dolphin holds a basketball as it swims on its back



A **dolphin** holds a **basketball** as it swims on its **back** .

A **dolphin** swimming upside down while holding a **basketball**

A **dolphin** swims upside down holding a **basketball** between it 's **flippers** .

A **seal** floats on it 's back in the **water** , holding a **basketball** .

The **dolphin** on his **back** holds the **orange basketball**

“Pink” from Google





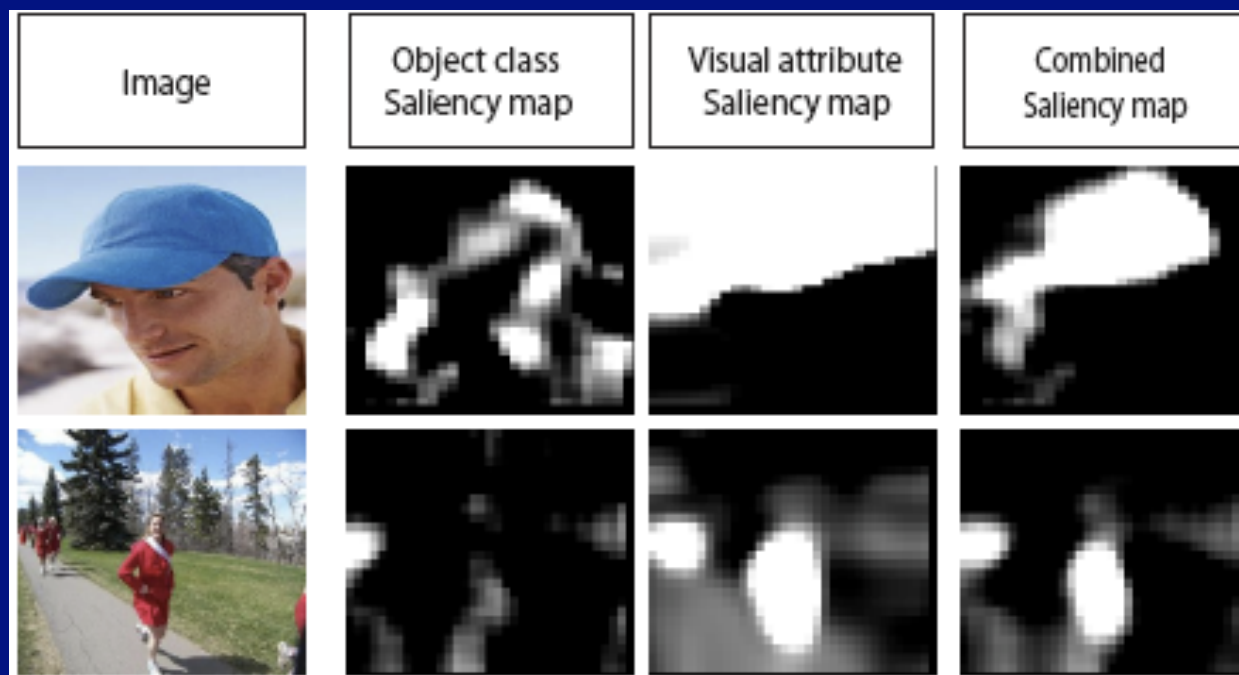
“Pink” after 10 EM iterations



Partially supervised recognition

- Training
 - we know what is in the image, but not where
- But if we have adjectives
 - we can improve location estimates
 - and so recognition

Training						
	Cap	Pants	Dress	Car	Flower	Umbrella
Blue						
Purple						
Red						
Yellow						



Wang et al 09



Structure

- Correlated words
 - waves go with beaches not cats
- Attributes
 - has nose
- Adjectives
 - green hat
- **Relations** Gupta and Davis 08,
 - **cat on mat** but there is still a lot here
- Sentences
 - A dolphin holds a basketball as it swims on its back



A **dolphin** holds a **basketball** as it swims on its **back** .

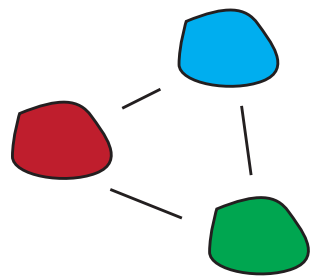
A **dolphin** swimming upside down while holding a **basketball**

A **dolphin** swims upside down holding a **basketball** between it 's **flippers** .

A **seal** floats on it 's back in the **water** , holding a **basketball** .

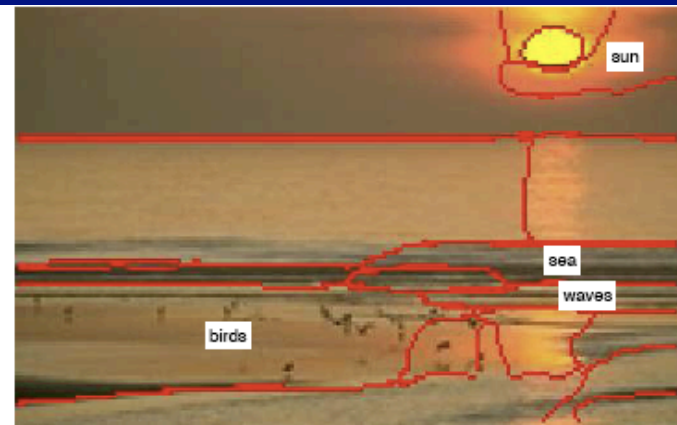
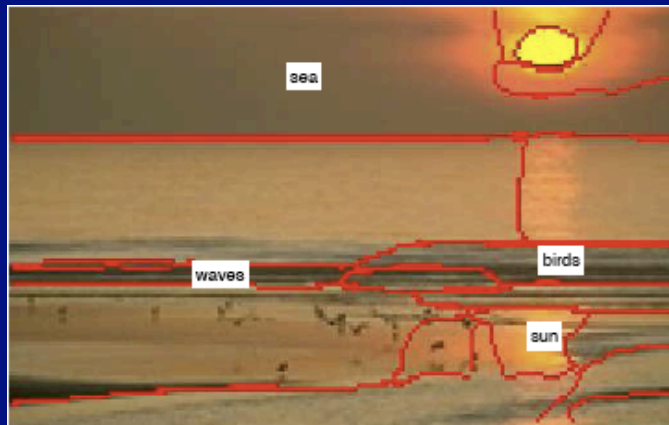
The **dolphin** on his **back** holds the **orange basketball**

Relations as an MRF



Sun IN Sky
Sky ABOVE Grass
Sun ABOVE Grass

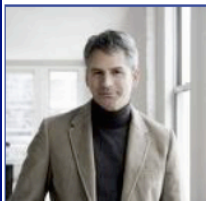
Gupta and Davis 08



Duygulu et al 02

Gupta and Davis 08

Relations distort participants



Man Smiling
482 x 473 - 28k - jpg
www.sap.com



Iron Man (Anthony Stark)
440 x 348 - 63k - jpg
marvel.com



Not what man ...
895 x 361 - 38k - jpg
www.mlahanas.de



**ONCE-HEAVIEST
MAN TIES THE KNOT**
531 x 411 - 38k - jpg
abcnews.go.com
[More from a.abcnews.com]



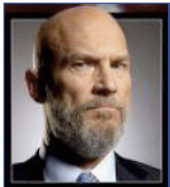
Australia's first pregnant man,
...
432 x 324 - 80k - jpg
www.powerhousemuseum.com
[More from www.powerhousemuseum.com]



The world's tallest man has put
his ...
700 x 465 - 74k - jpg
www.newscientist.com



'X-Men' stage reunion,
'Iron Man' ...
726 x 1080 - 79k - jpg
latimesblogs.latimes.com
[More from latimesblogs.latimes.com]



... Man," the ...
303 x 342 - 24k - jpg
www.oregonlive.com
[More from photobucket.com]



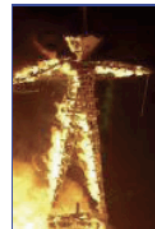
Spider-Man 3
460 x 300 - 28k - jpg
www.guardian.co.uk



**... to have a \$1500 Iron
Man phone**
318 x 500 - 46k - jpg
www.engadgetmobile.com



**Bao Xishun, the
World's Tallest Man ...**
611 x 404 - 89k - jpg
www.time.com



Greetings from Burning Man
2007.
607 x 924 - 85k - jpg
blogs.laweekly.com



... of Spider-Man 3.
357 x 458 - 41k - jpg
www.deadlinehollywooddaily.com



the-man.jpg (768 x 1024
- 102K)
768 x 1024 - 100k - jpg
pic.templetons.com

Relations distort participants



HerbWeb **Horses**: a **horse** on the beach
1024 x 768 - 88k - jpg
www.hedweb.com
[[More from](#)
www.hedweb.com]



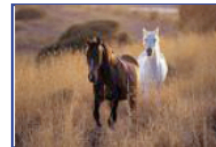
The Arabian **horses** at Smoky Mountain ...
550 x 681 - 254k - jpg
www.smokymountainparkarabians.com
[[More from](#)
www.smokymountainparkarabians.com]



This is the **horse** riding page of ...
657 x 430 - 58k - jpg
www.horseriding.gr



... **horse's** mouth.
613 x 525 - 63k - jpg
www.nurturalhorse.com



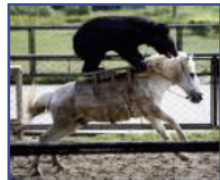
Welcome to the **Horse's** Maine!
600 x 400 - 52k - jpg
www.horsesmaine.com



horse
635 x 449 - 117k - jpg
www.historyforkids.org



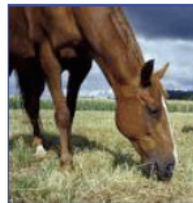
Przewalski's **Horse**, Przewa
...
470 x 324 - 30k - jpg
animals.nationalgeographic.com
[[More from](#)
animals.nationalgeographic.com]



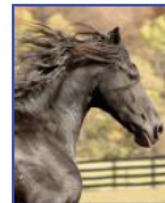
Bear **Horse**
550 x 458 - 160k - jpg
www.bestweekever.tv



... conditions of carriage **horses** ...
500 x 754 - 32k - jpg
advocacy.britannica.com



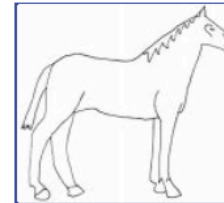
MRSA can cause infection in **horses**, ...
388 x 408 - 33k - jpg
www.wormsandgermsblog.com



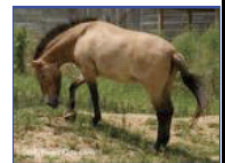
Friesian **Horses** - Wish Upon A Ster ...
340 x 420 - 37k - jpg
www.wishuponaster.com



Guide **Horse** Foundation - Miniature ...
359 x 300 - 32k - jpg
www.guidehorse.org



... jackets from books about **horses** ...
717 x 643 - 57k - jpg
www.tsl.state.tx.us



Animal Scoop - Przewalski's **Horse** ...
500 x 375 - 51k - jpg
www.billybear4kids.com

Relations distort participants



Puerto Rican **man riding horse**
400 x 286 - 17k - jpg
www.speedysigns.com
[[More from](http://www.speedysigns.com)]



Photo of **man riding a horse** near a ...
405 x 580 - 45k - jpg
www.americaslibrary.gov
[[More from](http://www.americaslibrary.gov)]



Image of **Man riding horse** on the way ...
525 x 350 - 103k - jpg
www.traveladventures.org



Man riding If you are a keen **horse** ...
194 x 281 - 12k - jpg
www.friendshipridingstables.com



... Mongolian **man riding horse** ...
400 x 267 - 16k - jpg
www.mongolia-travel-guide.com



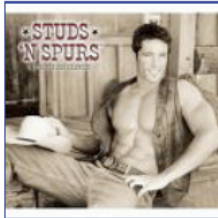
Stock Photo - **man riding a horse** ...
300 x 221 - 24k - jpg
www.fotosearch.com
[[More from](http://www.fotosearch.com)]



Stock Illustration - **man riding** ...
300 x 320 - 18k - jpg
www.fotosearch.com



Man riding a horse.
326 x 476 - 45k - jpeg
www.terragalleria.com
[[More from](http://www.terragalleria.com)]



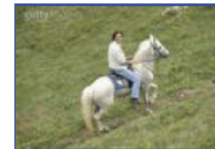
Seeing a **man riding a horse**, ...
500 x 500 - 57k - jpg
dixieugadawg.wordpress.com



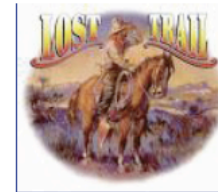
Man Riding A Horse
461 x 615 - 70k - jpg
www.publicdomainpictures.net



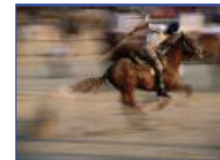
Photo of **man riding a horse** near a ...
200 x 200 - 11k - jpg
www.americaslibrary.gov



Man riding horse through meadow
504 x 339 - 45k
www.gettyimages.com



lost-trail-
soda-**man-riding-horse**.jpg
500 x 433 - 65k - jpg
losttrailsoda.com



Man Riding Horse At Annual Pushkar ...
400 x 300 - 29k - jpg
www.allposters.com.au
[[More from](http://www.allposters.com.au)]

Structure

- Correlated words
 - waves go with beaches not cats
- Attributes
 - has nose
- Adjectives
 - green hat
- Relations
 - cat on mat
- Sentences
 - A dolphin holds a basketball as it swims on its back



A **dolphin** holds a **basketball** as it swims on its **back** .

A **dolphin** swimming upside down while holding a **basketball**

A **dolphin** swims upside down holding a **basketball** between it 's **flippers** .

A **seal** floats on it 's back in the **water** , holding a **basketball** .

The **dolphin** on his **back** holds the **orange basketball**



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 .



A crowd of young adults in a dark room .

A girl in a brown shirt and a blue jean skirt is dancing with a young man dressed in a blue shirt wearing a black backpack .

A group of people standing in a dark building .

A large group of people dancing in a bar

Dancing at club and two guys bucking up