

Object Recognition:  
more than remembrance  
of things past?

Shimon Edelman



# Object Recognition:

mouse

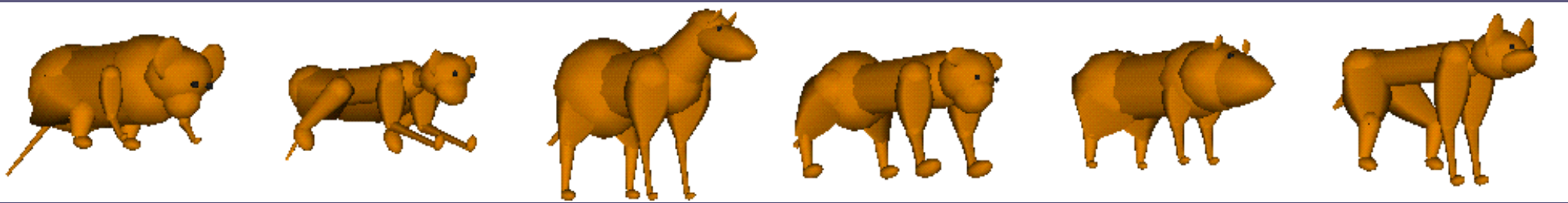
monkey

sheep

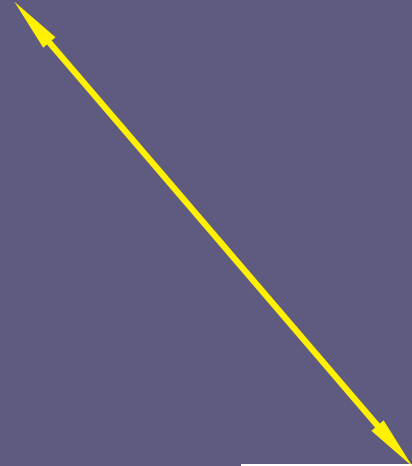
tiger

hippo

dog



things past



thing to be recognized

# Challenge #1: a novel view

mouse

monkey

sheep

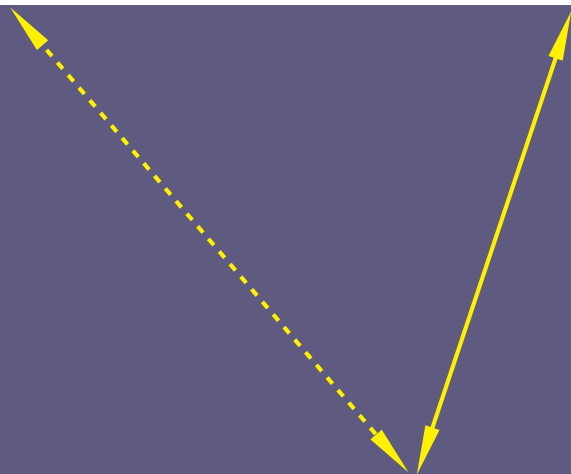
tiger

hippo

dog



remembered objects



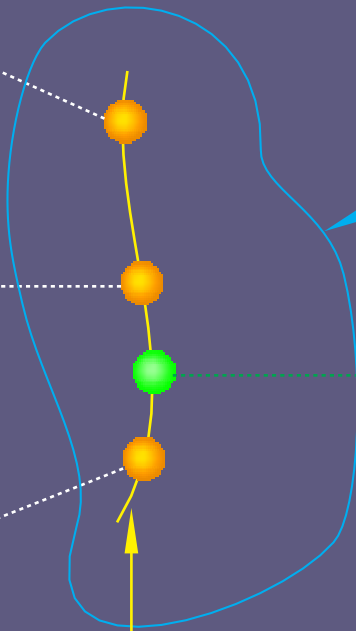
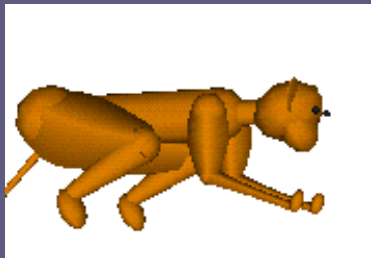
what if there is no exact match?

# View Space

smooth change in orientation

leads to

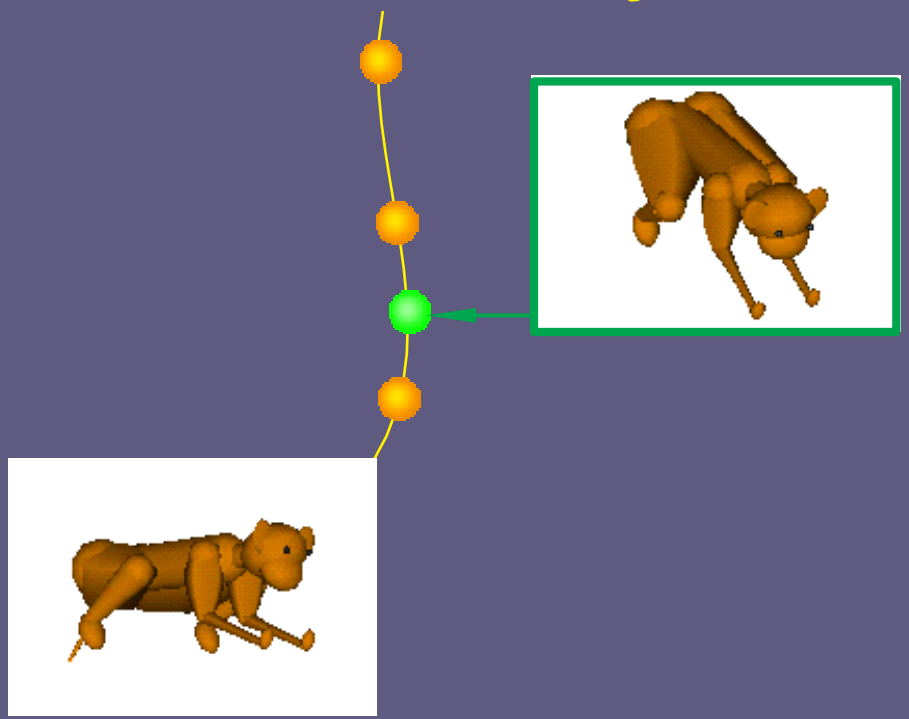
smooth change in **measurement space**



this suggests: **view space** can be INTERPOLATED.

# Recognition:

a familiar object

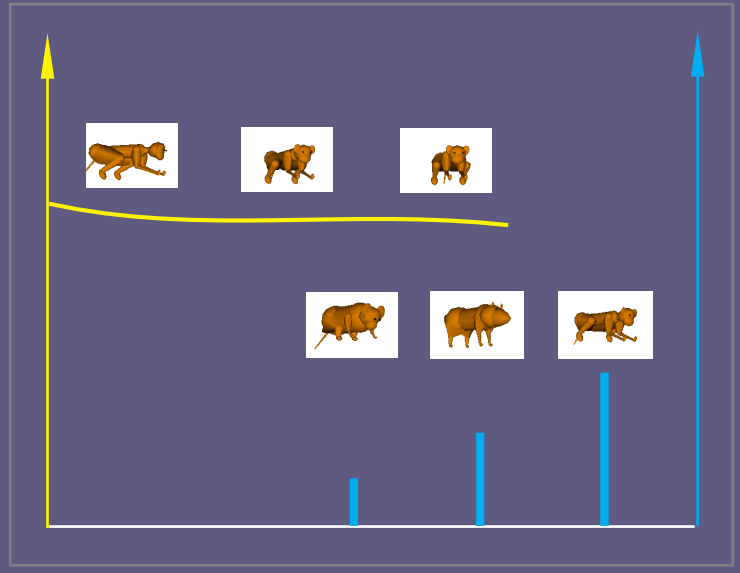


## strategy:

- \* interpolate view space;
- \* do Nearest Neighbor

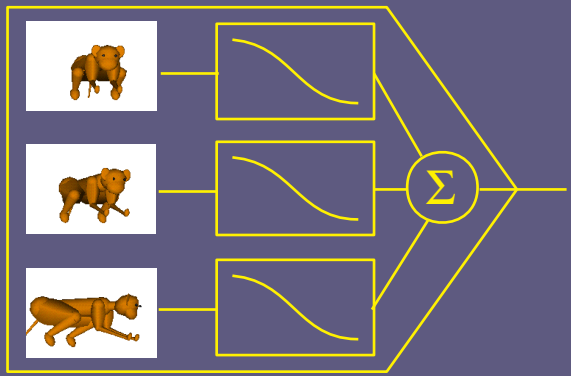
## functional requirements:

- \* near constant response
- \* rejection of other objects



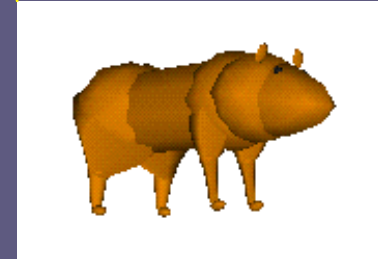
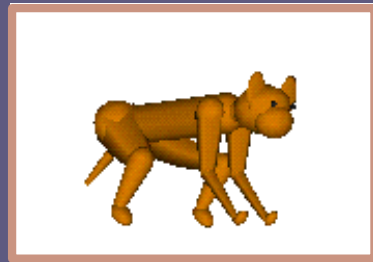
## possible implementation:

Radial Basis Function (RBF) interpolation.



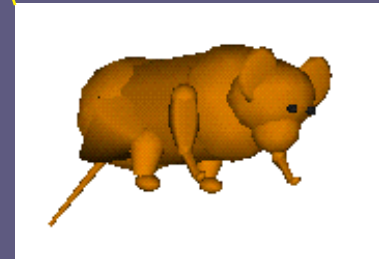
Categorization:

a "moderately" novel object



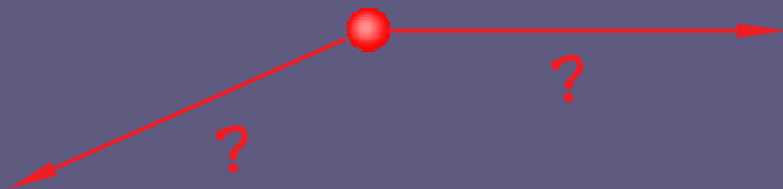
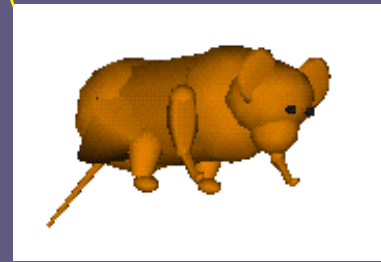
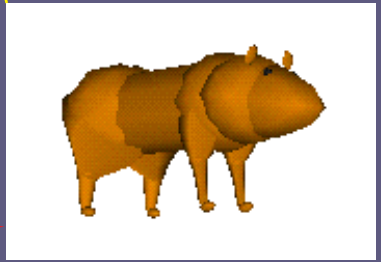
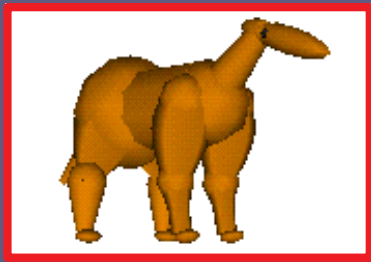
strategy:

- \* interpolate view space;
- \* do Nearest Neighbor



# Challenge #2:

a "radically" novel object



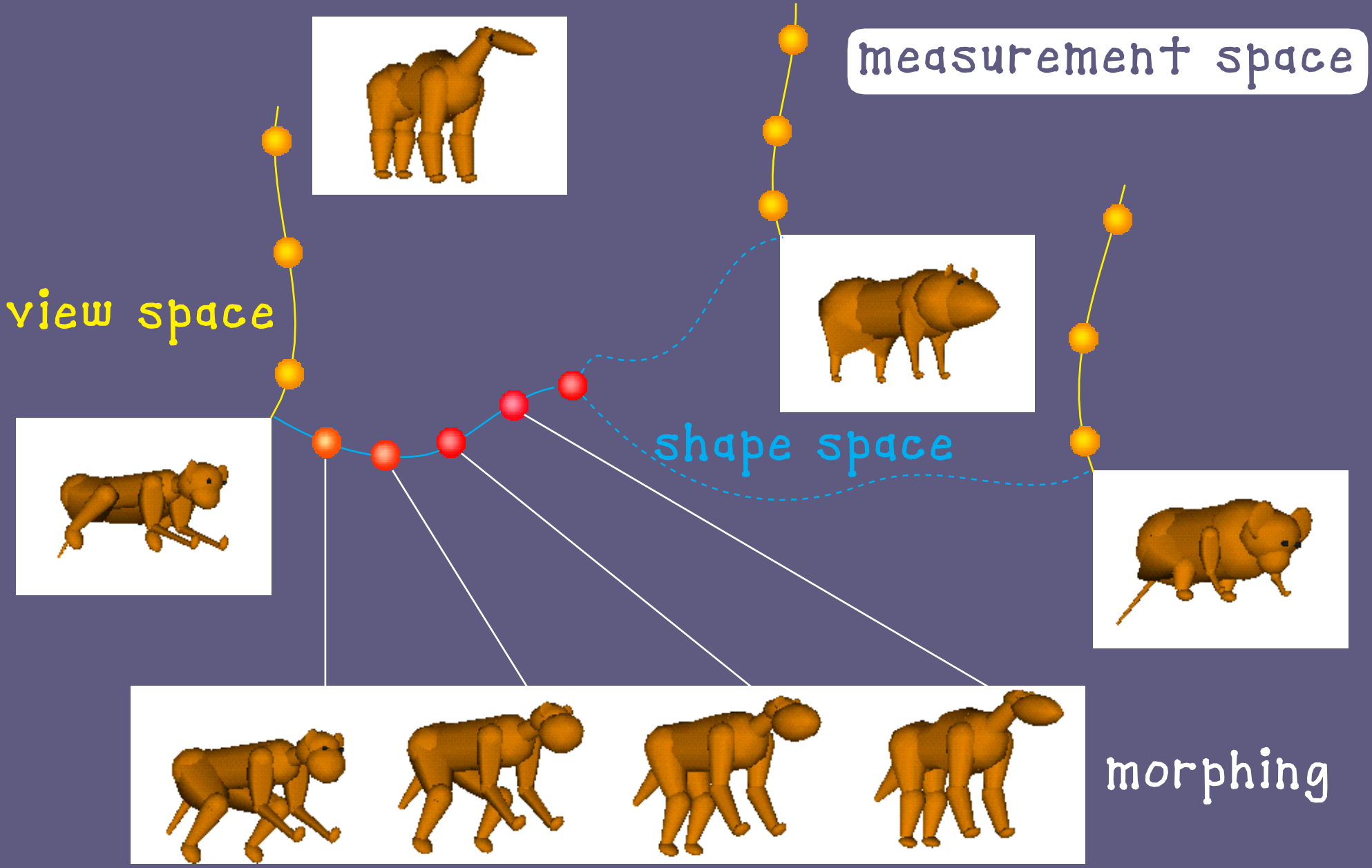
Nearest Neighbor strategy

a poor choice here;

Need: representation

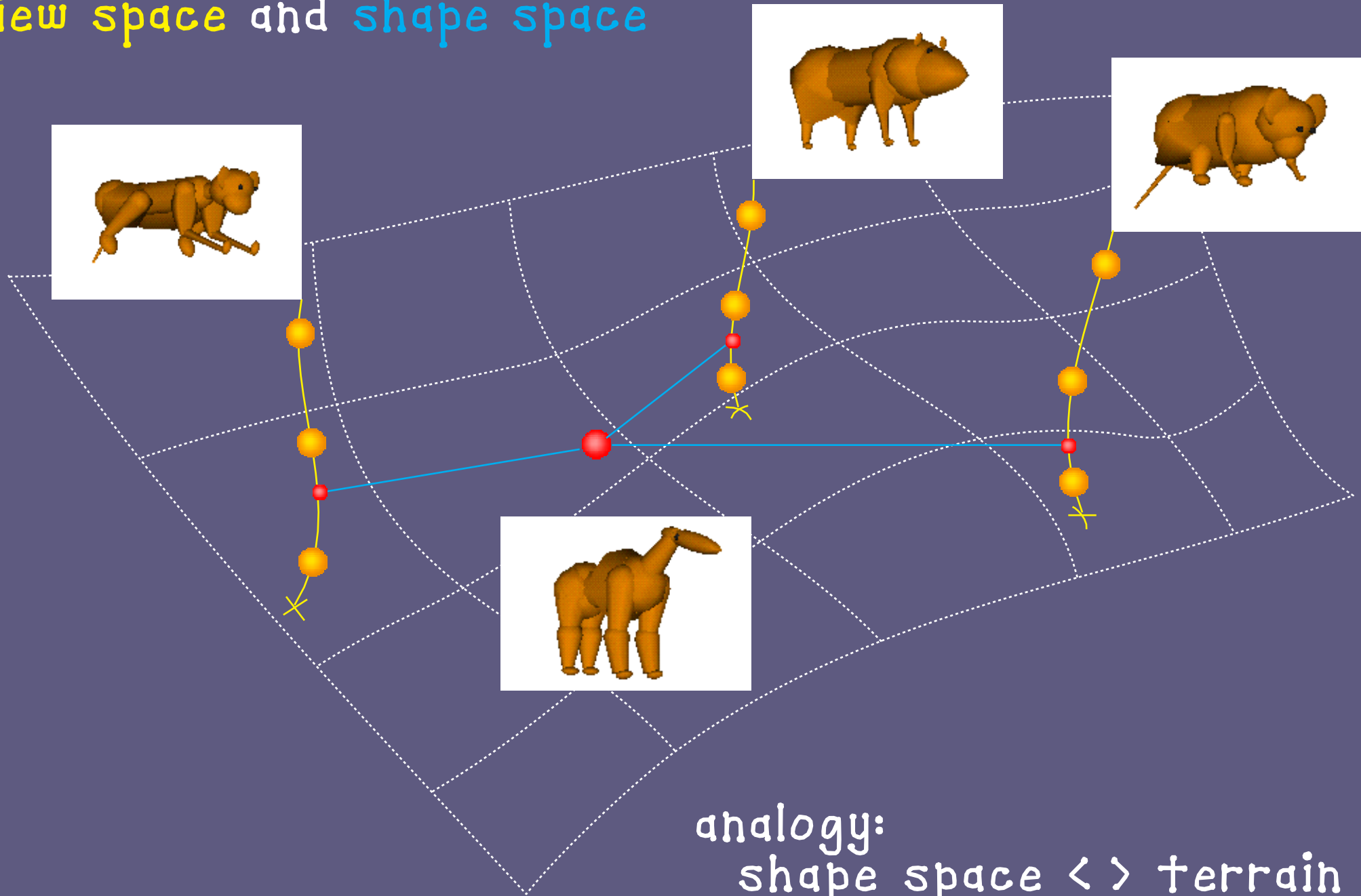
before decision

# A framework for the representation of shapes:



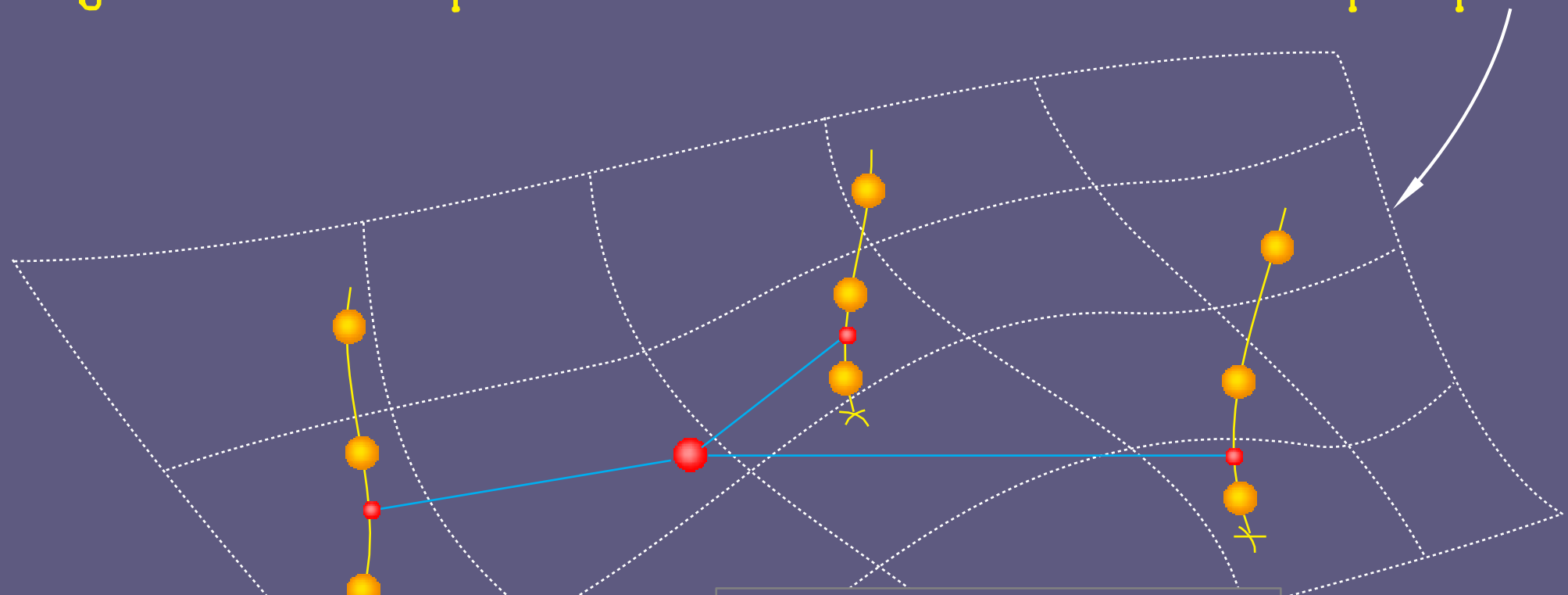


# view space and shape space



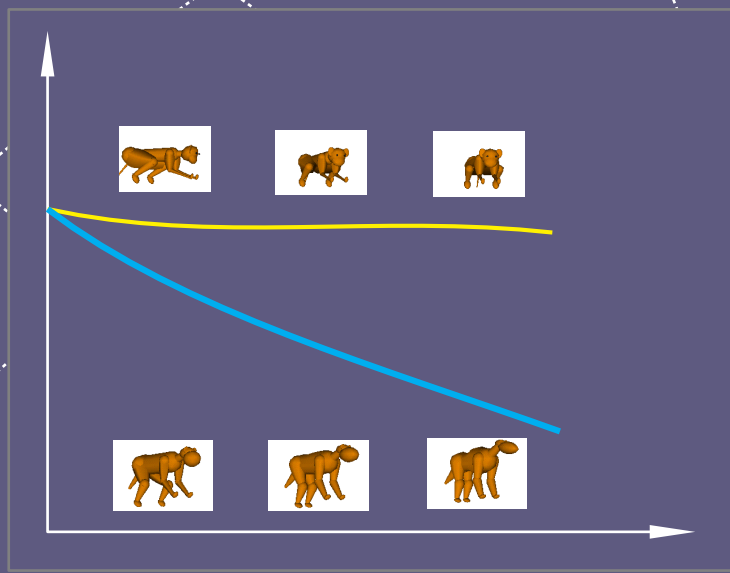
analogy:  
shape space < > terrain

# categorization, representation = localization in shape space



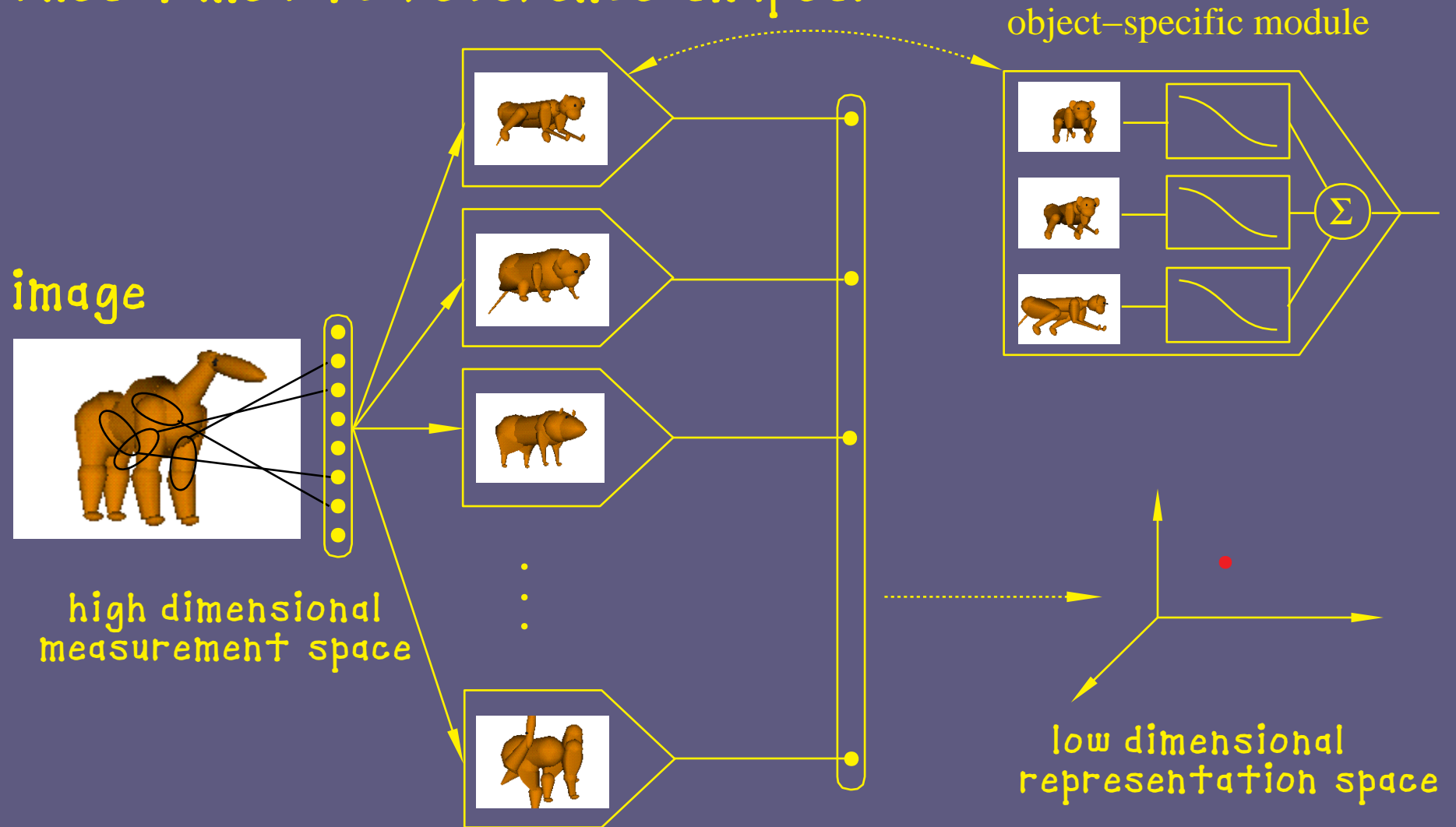
## REQUIRED:

- an active landmark
- mechanism to ignore distance along **view spaces**
- but respond to distance along the relevant **shape space** directions:



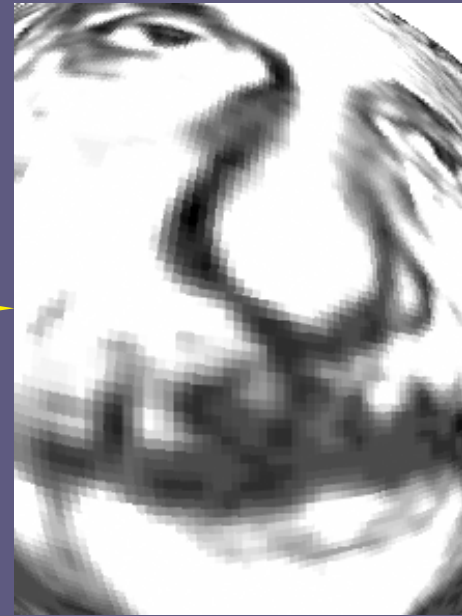
# A Chorus of Prototypes

(modules tuned to reference shapes)



High dimensional

measurement space

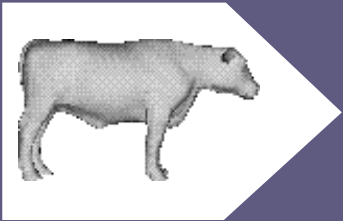


correlate with  
250 Gaussian  
filters  
("receptive fields")

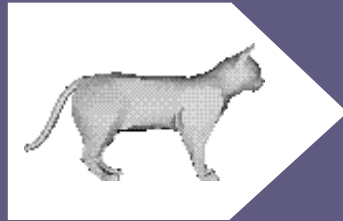
250 dimensional  
vector  
of measurements

10 training objects  
("reference shapes")

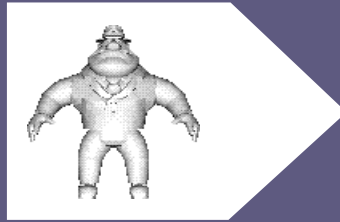
cow



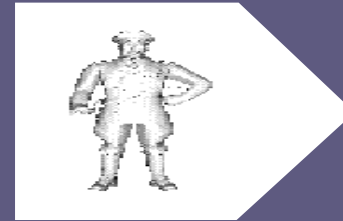
cat



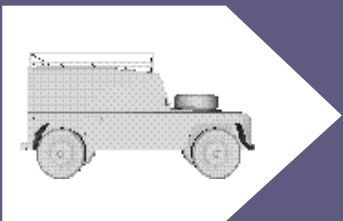
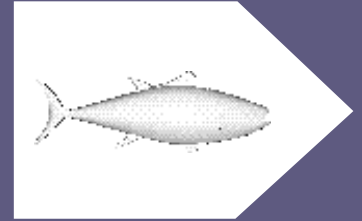
Al



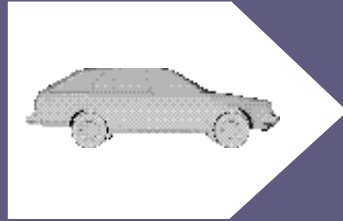
General



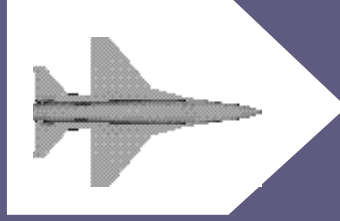
tuna



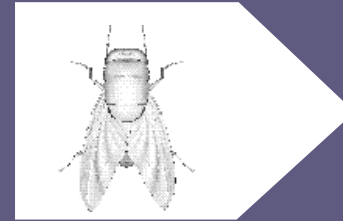
L\_rover



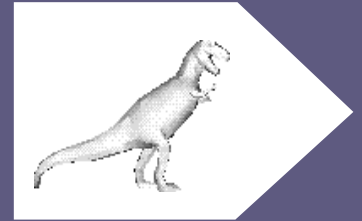
Nissan



F16



fly

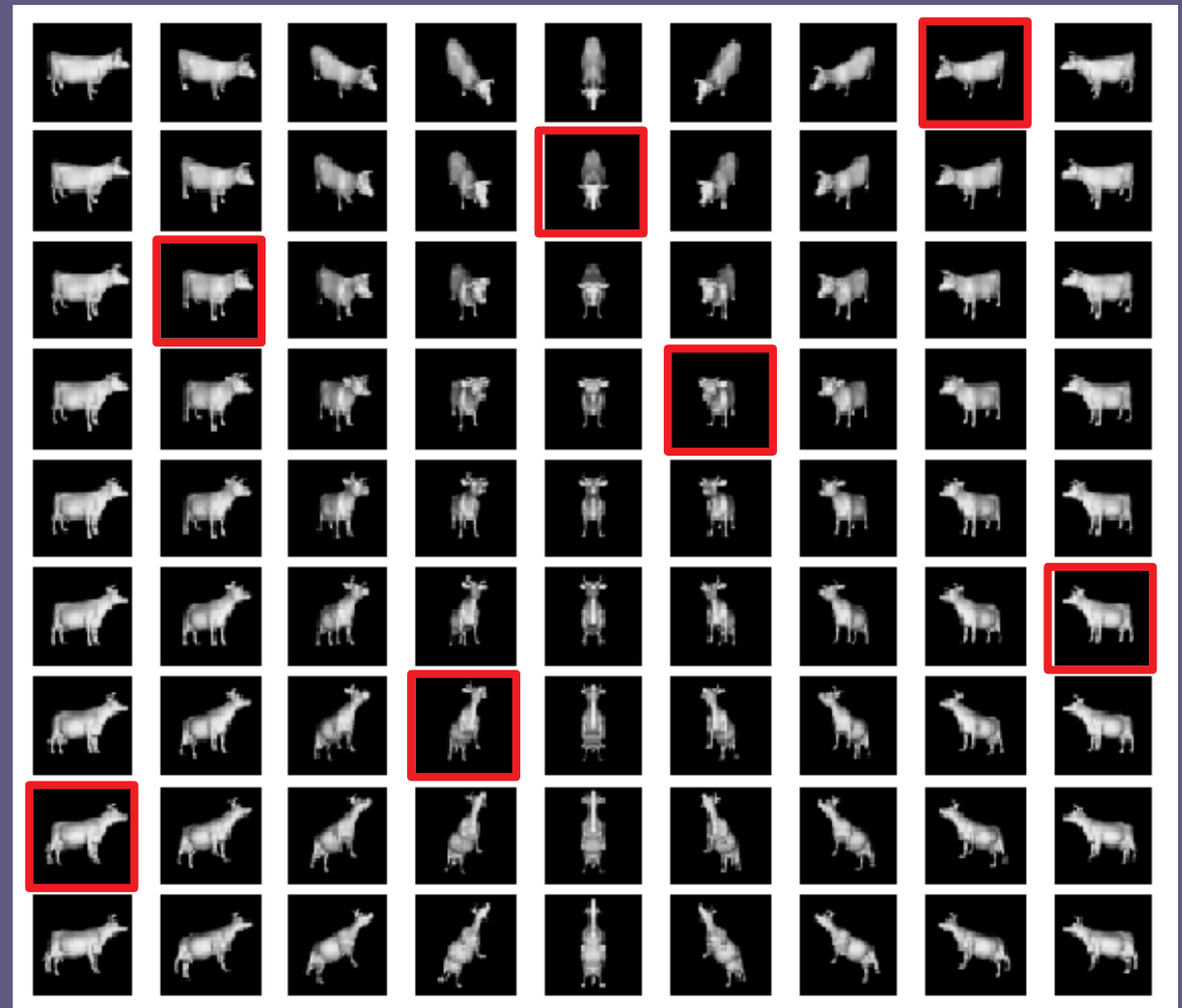


T.Rex

169 views/object

81 views of cow

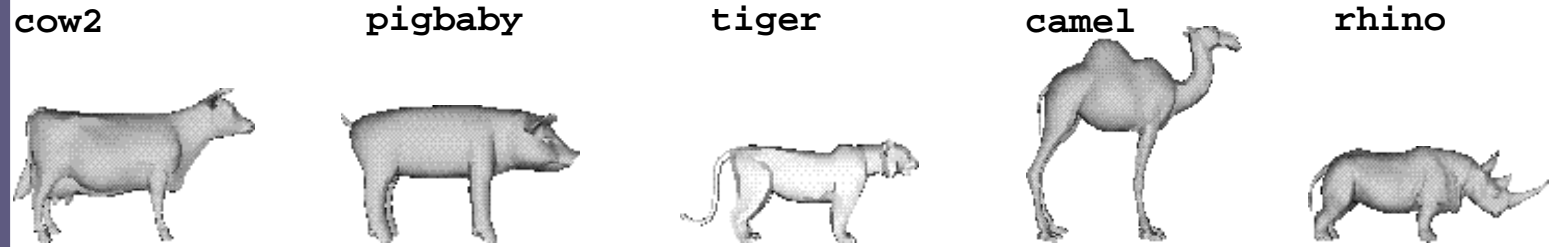
About 15 views/object  
were chosen  
(CVQ algorithm),  
to train the  
object specific  
modules.



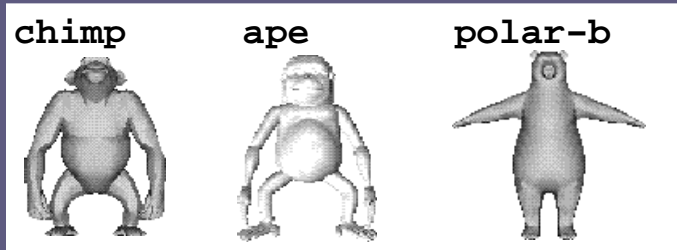
# Test set #1:

20 objects  
from same  
categories  
as  
training  
objects

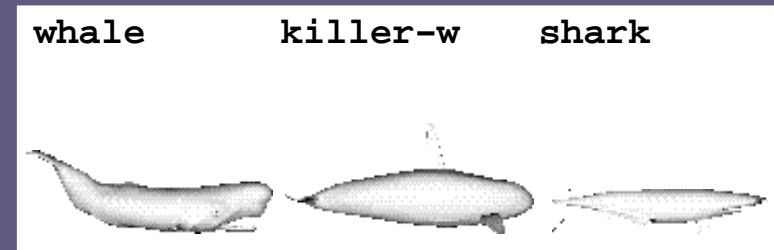
## QUADRUPEDS



## FIGURES



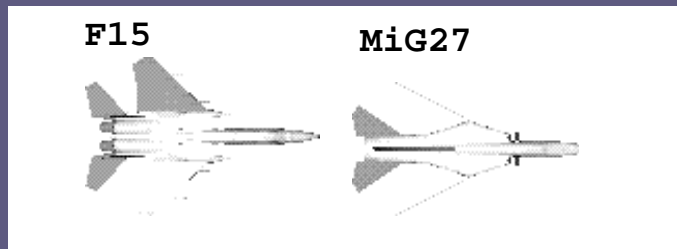
## "FISH"



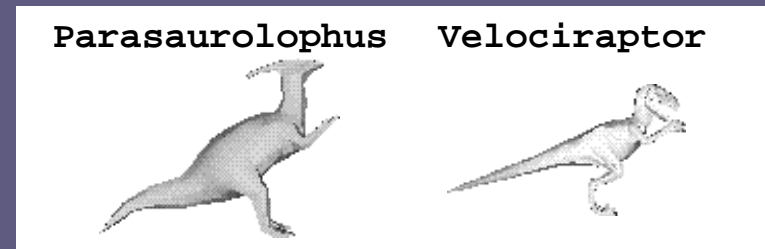
## "CARS"



## AIRCRAFT



## DINOSAURS



# Test set #1: results

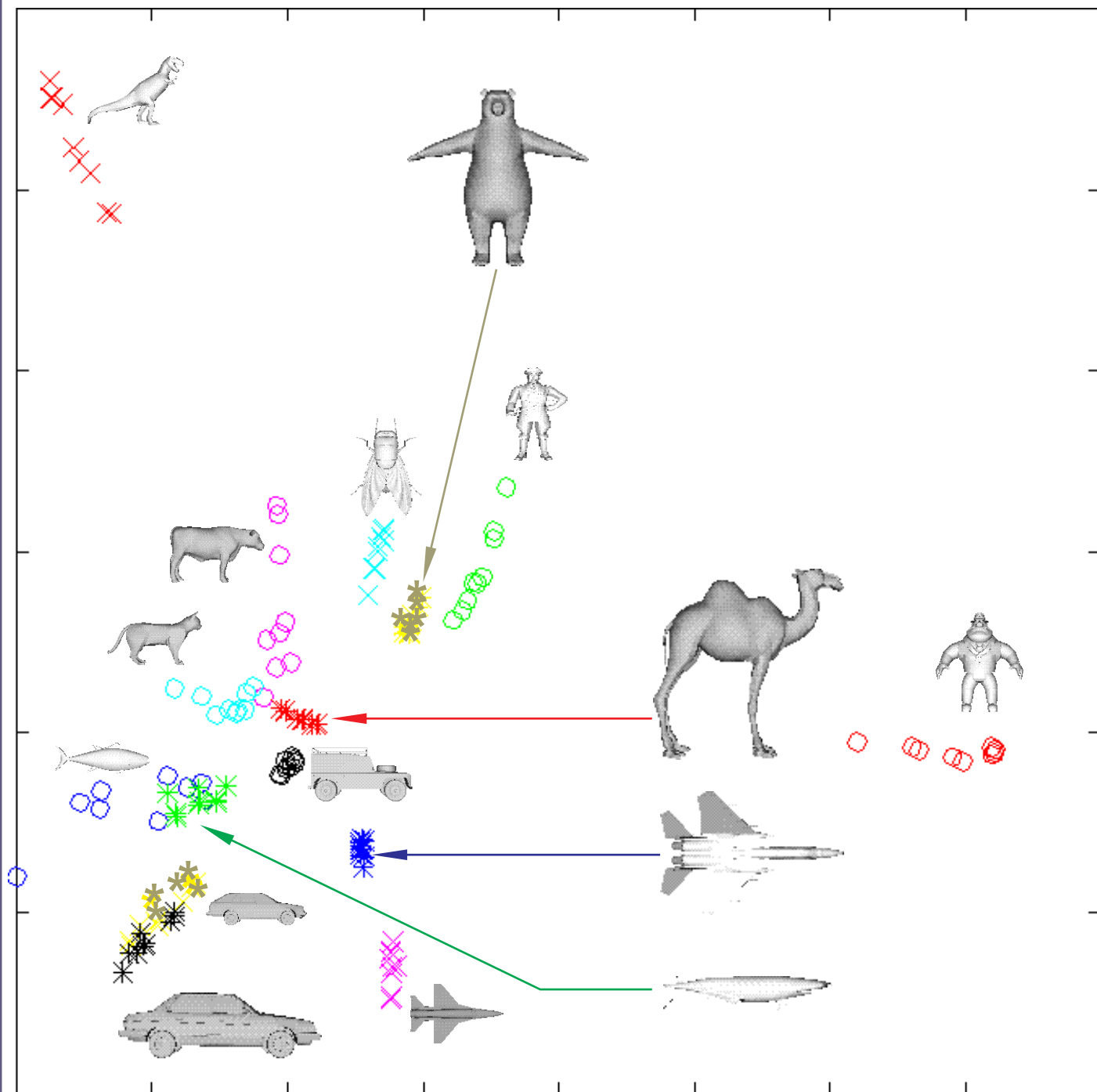
The 10-D space spanned  
by similarities to the  
10 training objects,  
embedded\* into 2-D  
for visualization.

Shown:  
10 training objects;  
5 test (novel) objects.

\* - by multidimensional  
scaling (MDS)

recognition rate: 83-98%

categorization: 79-85%





# Test set #2:

10 objects

randomly

chosen

from the

database

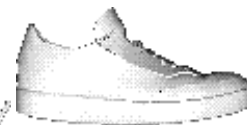
butterfly



frog



tennis



pump



Beethoven



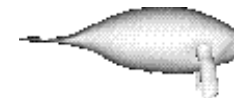
giraffe



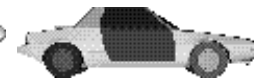
pawn



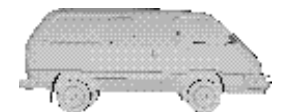
manatee



Fiat



Toyota



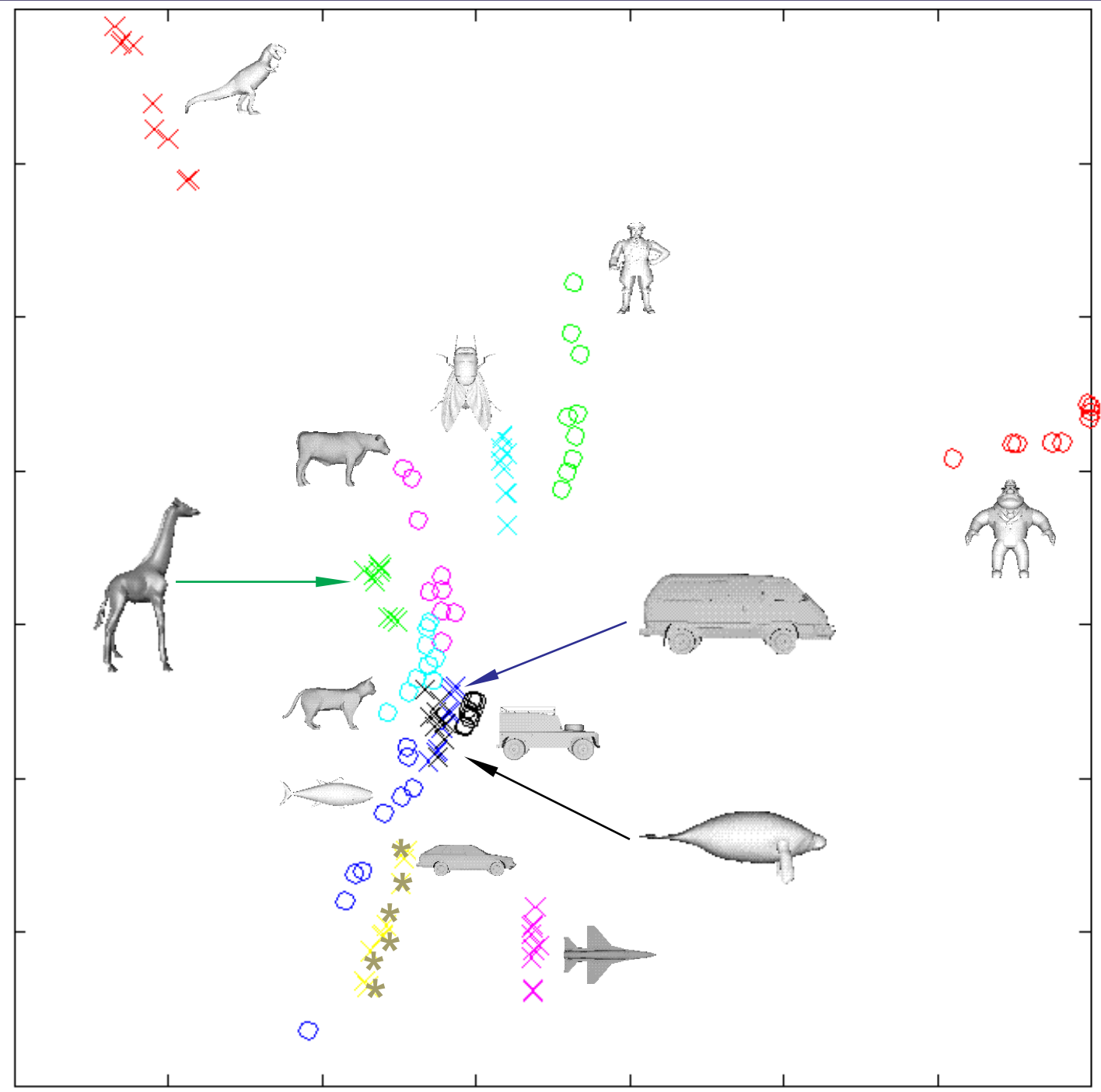
# Test set #2: results

The 10-D space spanned  
by similarities to the  
10 training objects,  
embedded\* into 2-D  
for visualization.

Shown:  
10 training objects;  
3 test (novel) objects.

\* - by multidimensional  
scaling (MDS)

recognition rate: 90-99%



# Test set #2: representation of novel objects

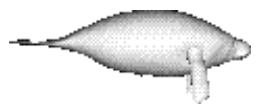
	cow	cat	Al	Gene	tuna	Lrov	Niss	F_16	fly	TRex
frog	0.19	0.12	0.29	0.09	0.20	0.08	0.08	0.08	<b>0.99</b>	0.10
shoe	0.25	0.31	0.05	0.06	<b>0.79</b>	0.15	<b>0.40</b>	0.27	<b>0.55</b>	0.09
pump	<b>0.77</b>	<b>0.58</b>	0.02	0.09	<b>1.12</b>	0.13	<b>0.75</b>	0.46	<b>0.65</b>	0.12
Beethoven	0.04	0.02	0.12	0.01	0.04	0.02	0.00	0.01	<b>0.39</b>	0.00
giraffe	<b>1.40</b>	<b>0.99</b>	0.02	0.28	<b>1.64</b>	0.07	0.68	0.78	<b>1.28</b>	<b>1.17</b>
<i>manatee</i>	<b>0.84</b>	<b>0.71</b>	<b>0.07</b>	<b>0.17</b>	<b>1.49</b>	<b>0.13</b>	<b>0.76</b>	<b>0.61</b>	<b>0.71</b>	<b>0.16</b>
Fiat	0.89	0.80	0.00	0.07	<b>1.98</b>	0.17	<b>1.61</b>	0.72	0.59	0.17
Toyota	1.17	1.06	0.08	0.12	<b>1.63</b>	0.87	<b>1.67</b>	0.66	0.71	0.1

*manatee*

tuna

cow

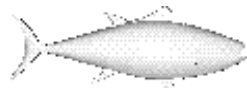
Nissan



=

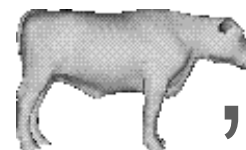
[

1.49



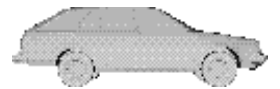
,

0.84



,

0.76

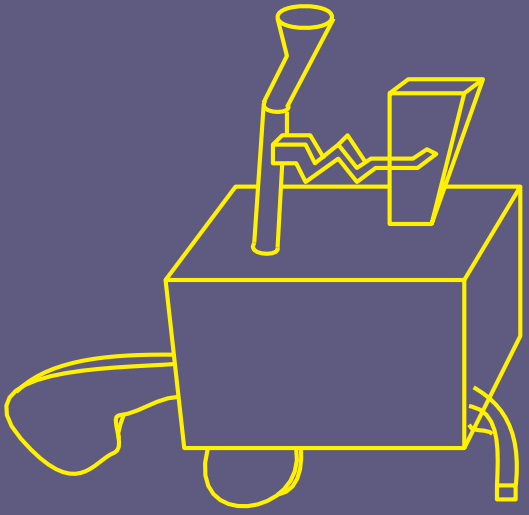


]

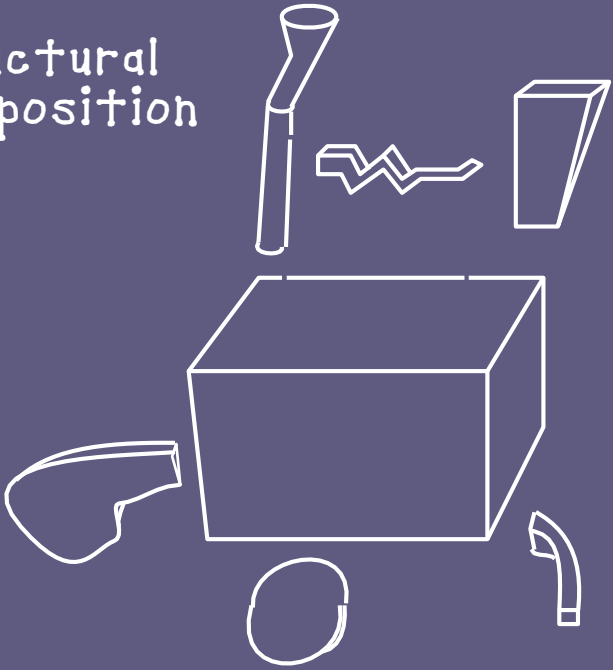
# A computer vision perspective

I. Biederman  
Psych. Review  
1987

ways to represent  
novel objects:

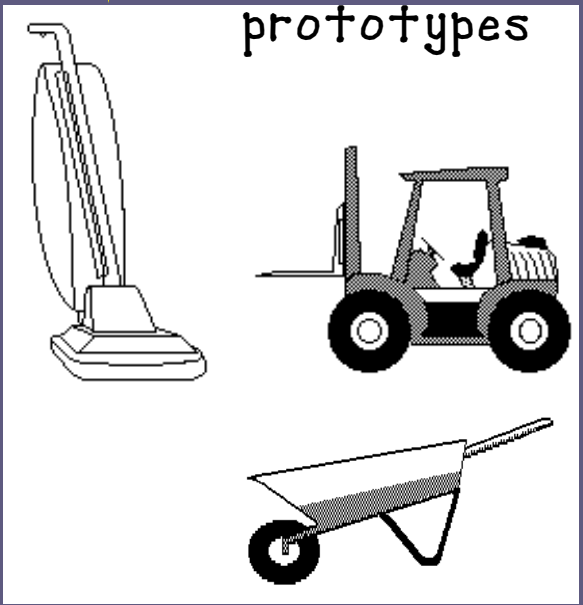


structural  
decomposition



the left way

similarities to  
prototypes



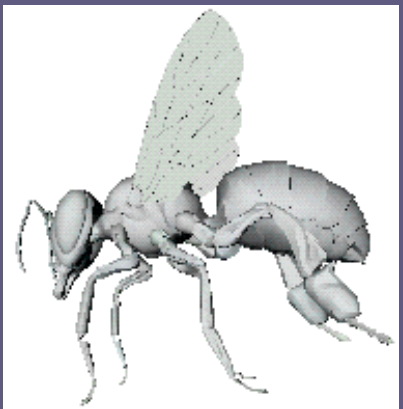
the right way

# some shortcomings of structural descriptions:

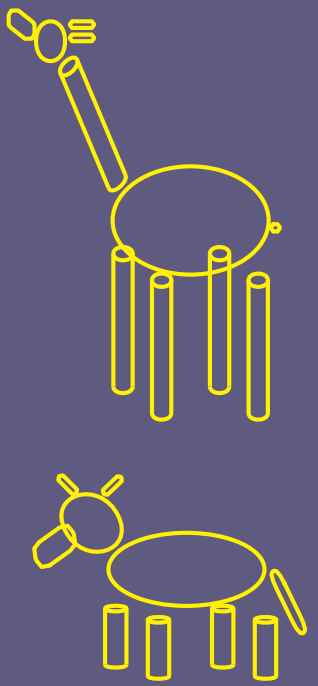
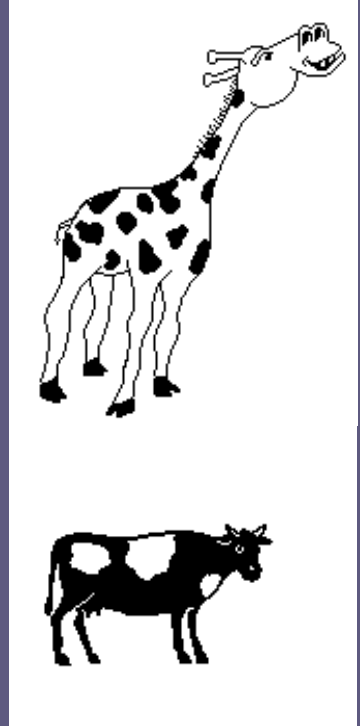
structural decomposition is not obvious for some simple common shapes...



...and is too cumbersome for some complex common ones...



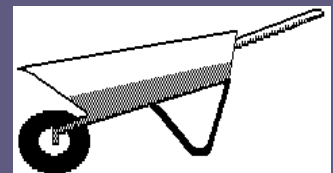
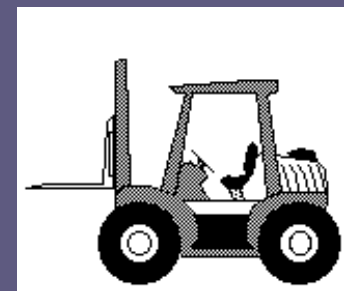
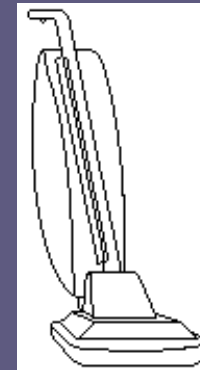
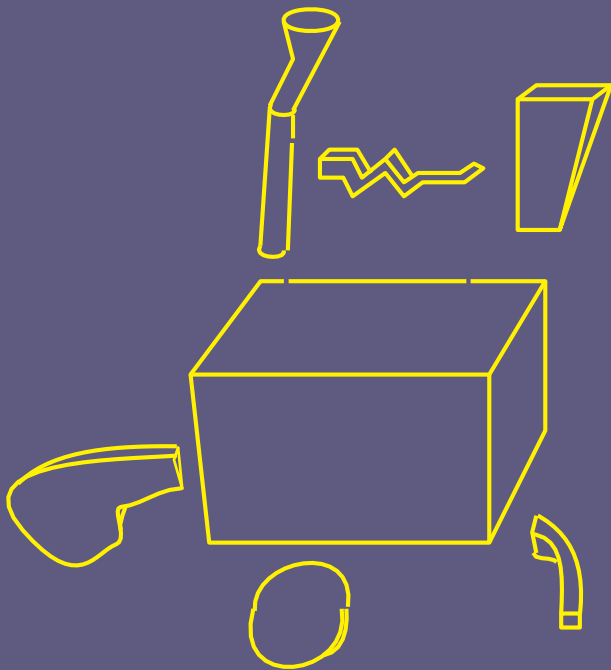
...metric information must be represented in any case...



structural decomposition  
is **very** difficult  
to compute automatically

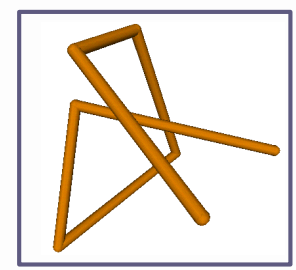
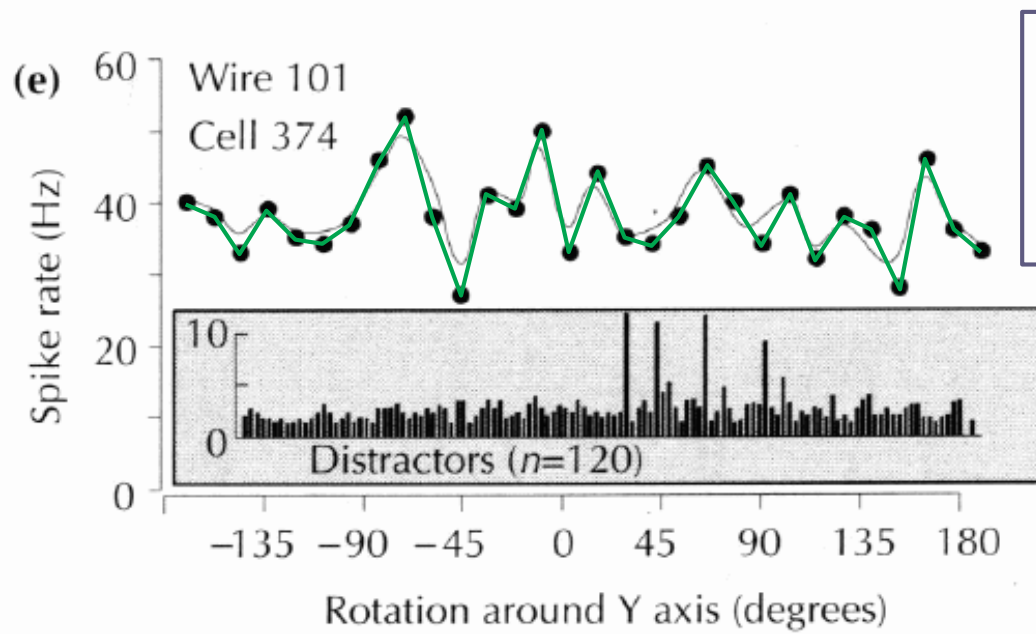
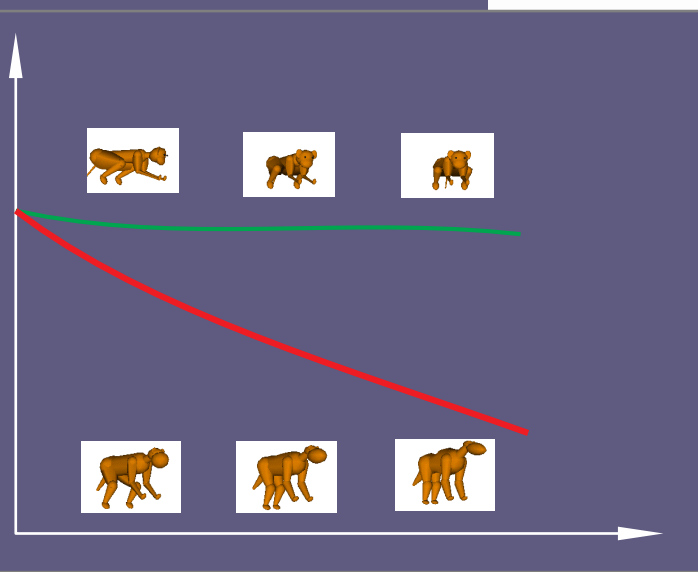
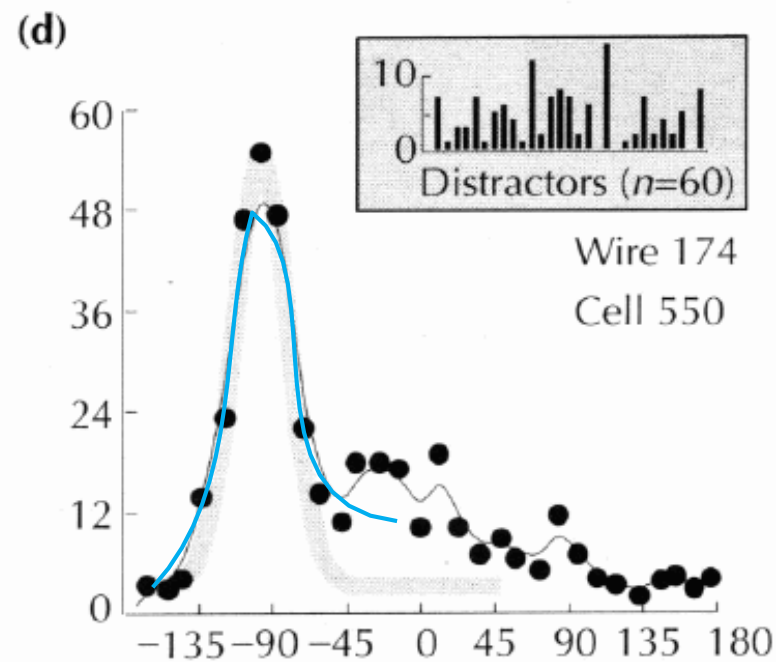
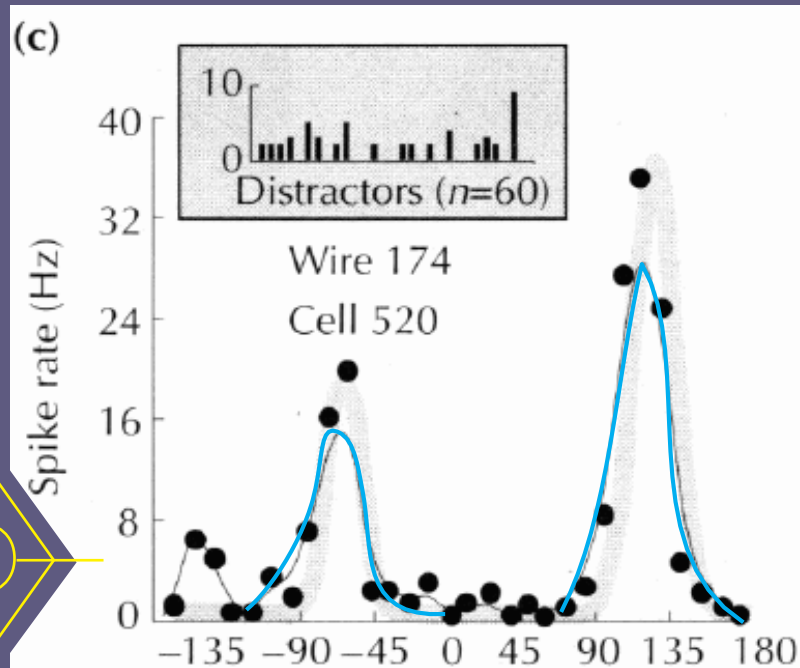
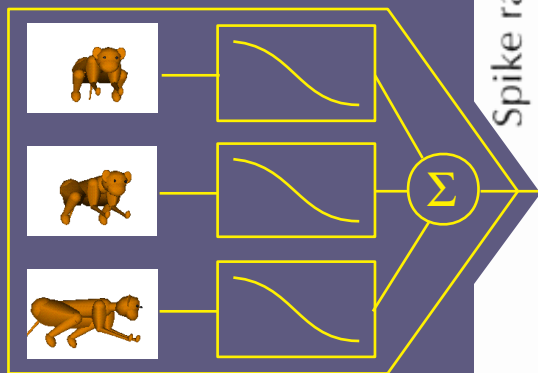


similarities to  
prototypes  
are easier



# A biological vision perspective

N. Logothetis, J. Pauls, T. Poggio, *Curr. Biol.* 5:552 (1995)

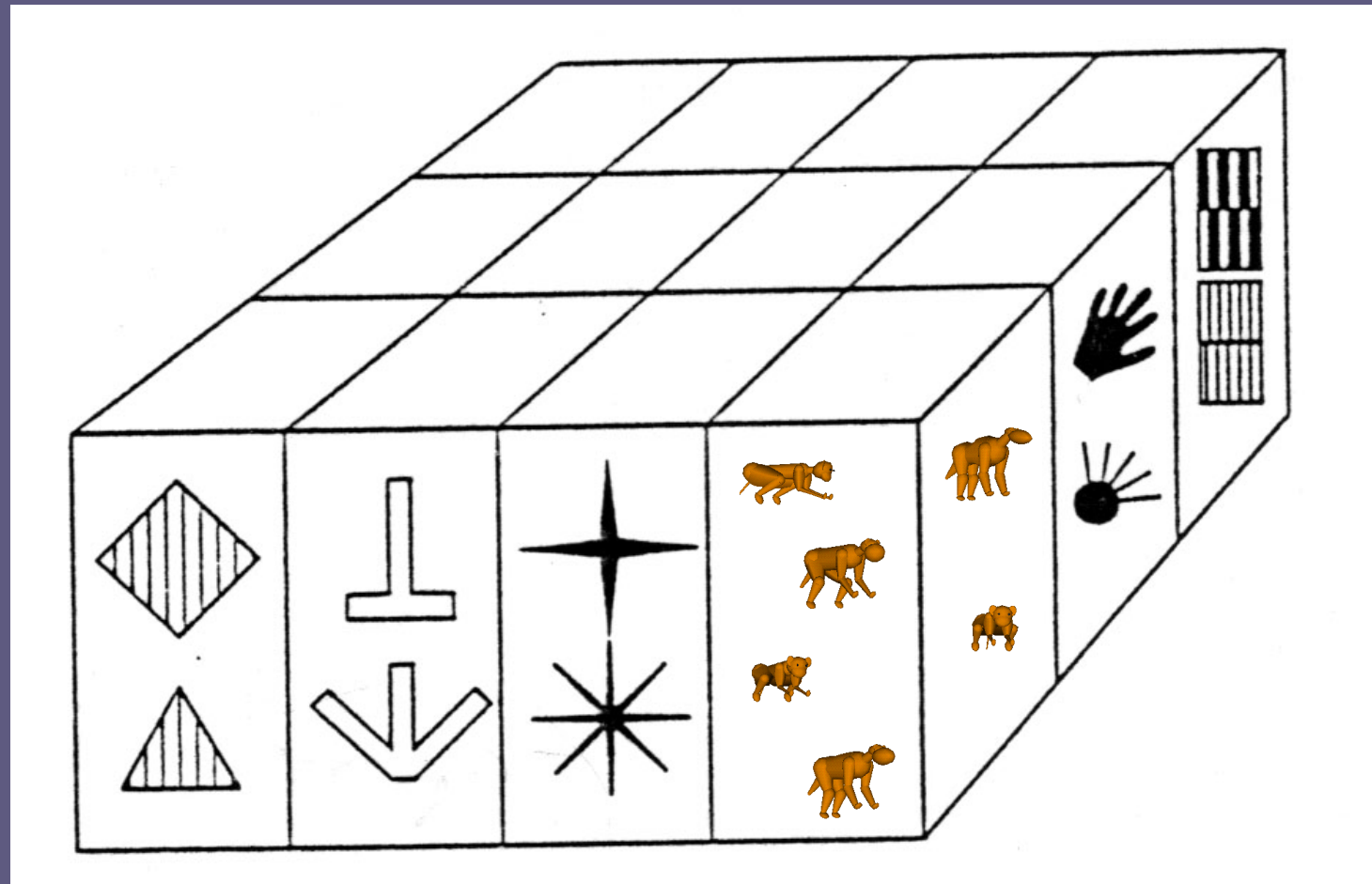




A tentative  
mapping of the  
model onto  
the functional  
architecture  
of the  
inferotemporal  
cortex:

adapted from:

K. Tanaka,  
Current Opinion in  
Neurobiology 2:502 (1992)



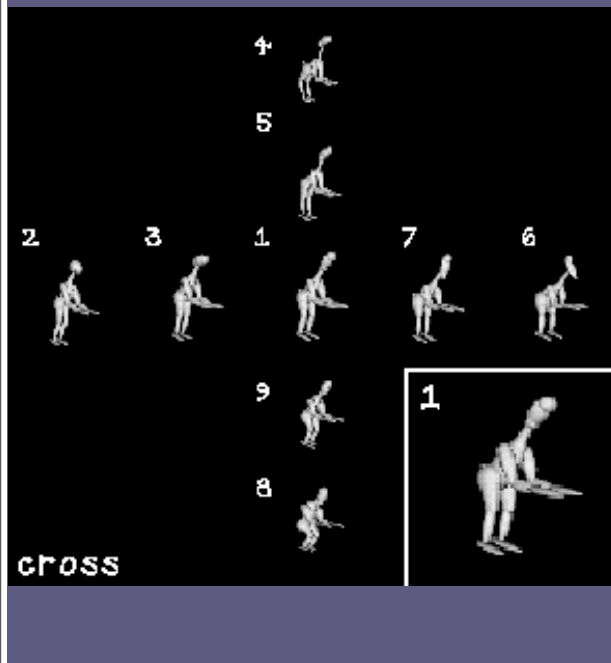
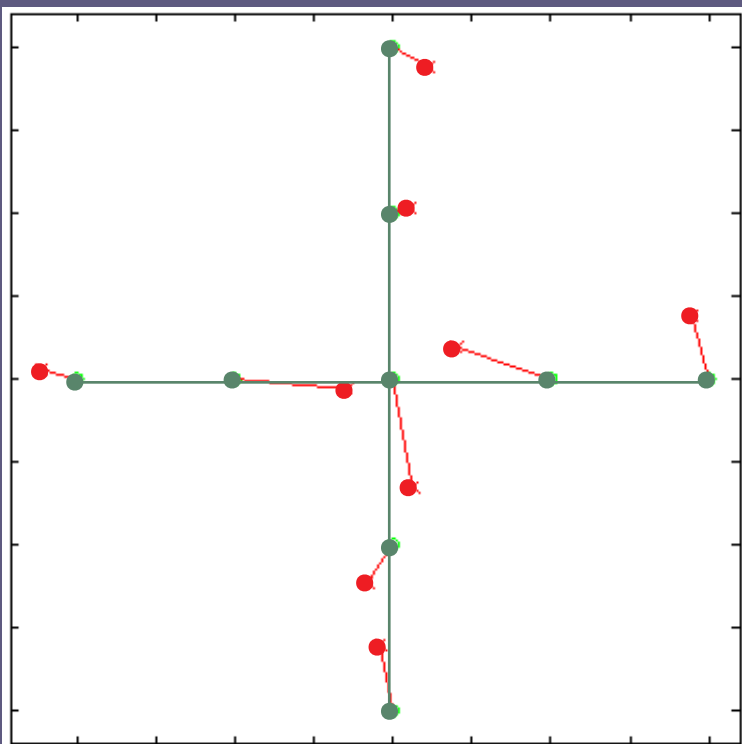


# human psychophysics: similarity perception

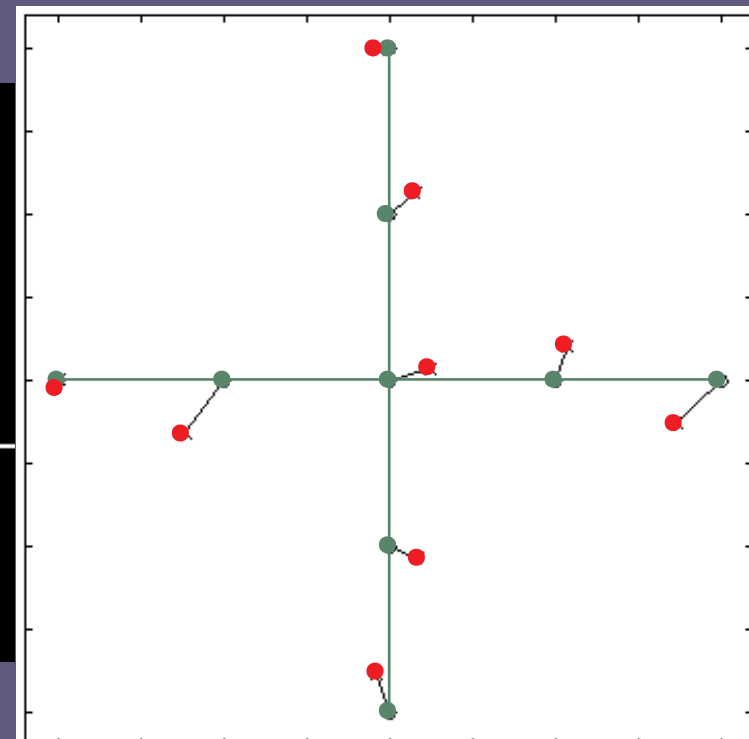
F. Cutzu & S. Edelman  
PNAS 93:12046 (1996)

1. parametrically manipulate stimulus shapes
2. use MDS to embed response data into 2D

computer model

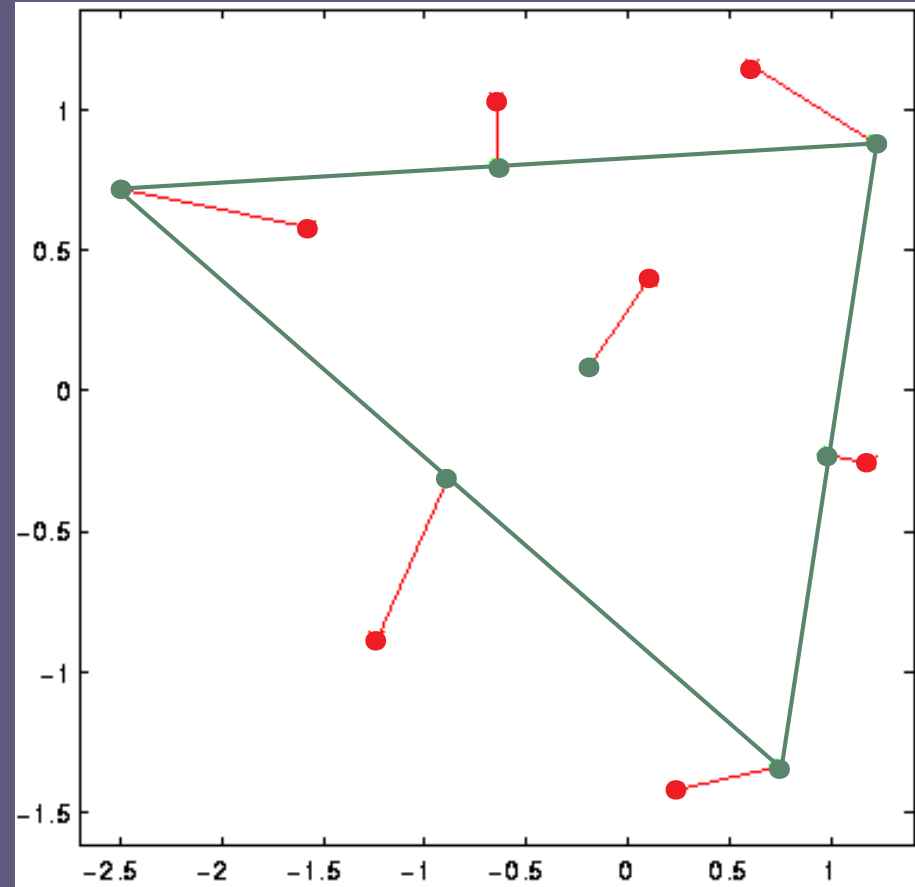
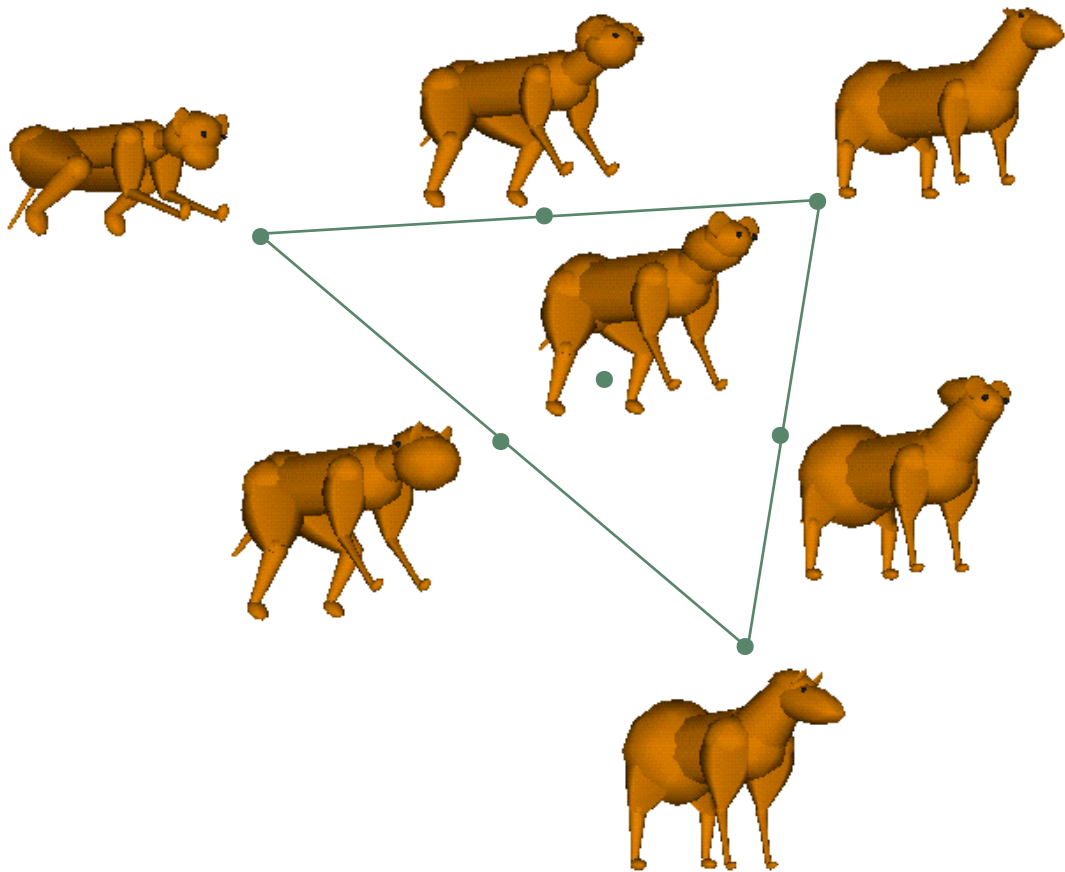


human subjects



# monkey psychophysics: similarity perception

T. Sugihara, S. Edelman, K. Tanaka  
Invest. Ophthalm. Vis. Sci., 1996



# Object Recognition: more than remembrance of things past?

Yes and No.

recognize

= remember a thing  
you saw before

categorize

= remember a thing  
you haven't seen before

represent

things current  
in terms of similarities  
to things past

Thanks to:  
Florin Cutzu, Sharon Duvdevani Bar

shantih shantih shantih