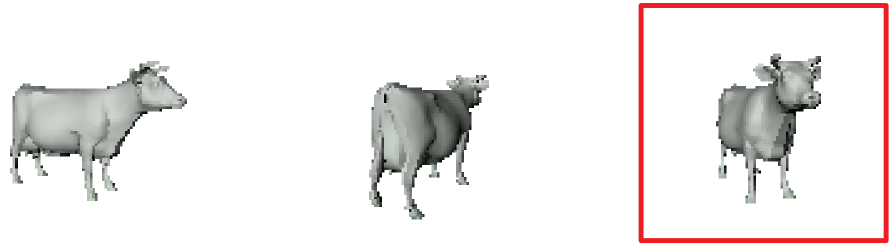
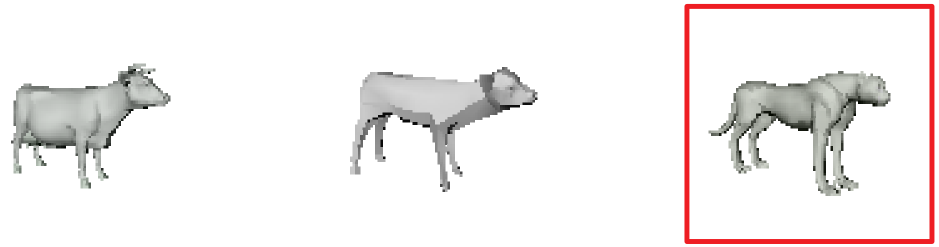
Representing Object Structure

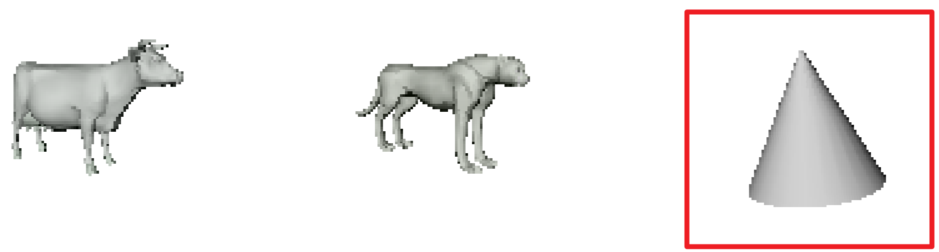
- **RECOGNITION:** how to deal with **novel views** of shapes?



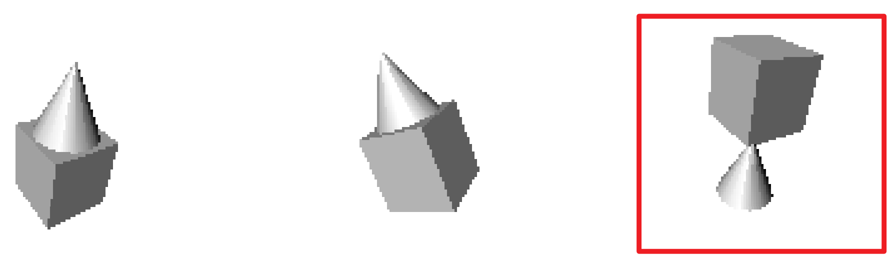
- **CATEGORIZATION:** how to deal with **novel instances** of shape categories?



- **META–CATEGORIZATION:** how to deal with **novel categories**?



- **REPRESENTATION of STRUCTURE:** how to deal with **novel arrangement**
  - of parts in an object?
  - of objects in a scene?



the **symbols + structure** idea:

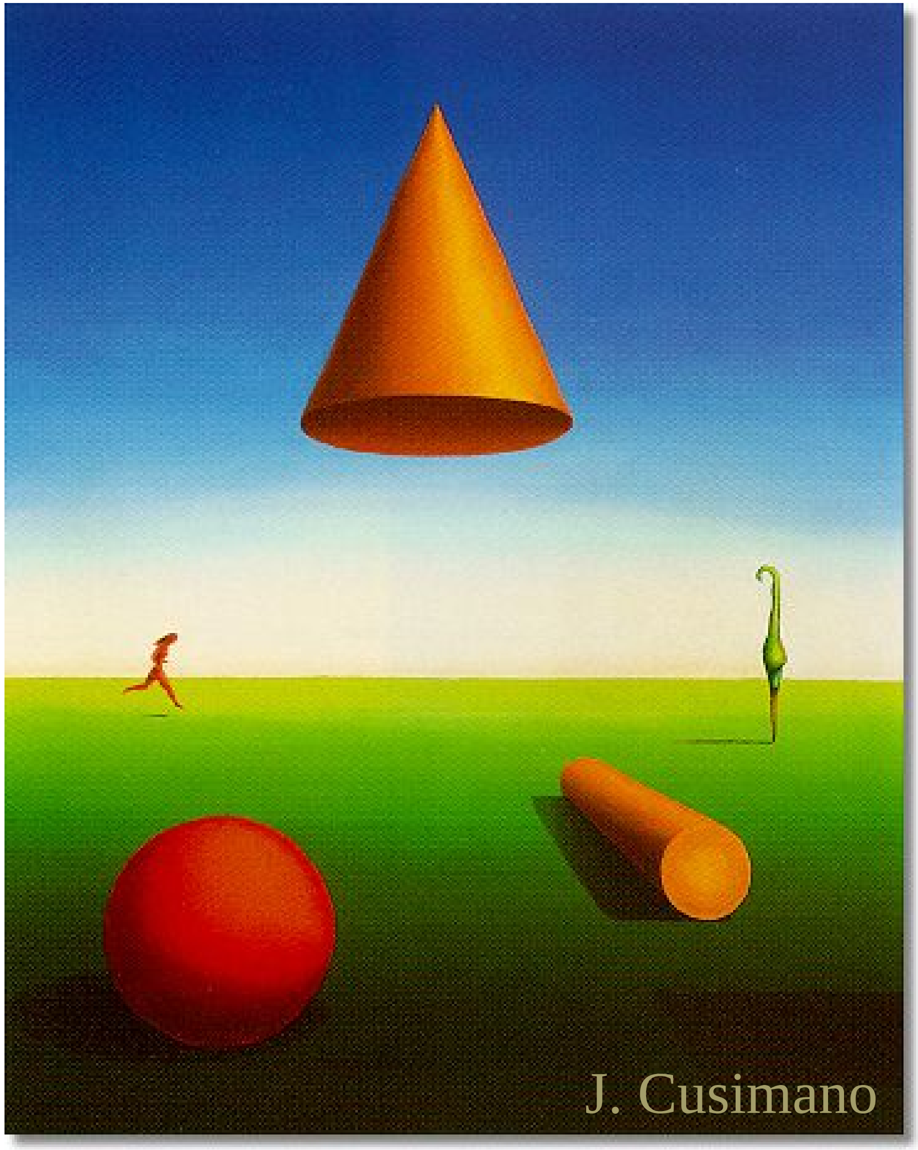
shapes = ***symbols*** standing for ***generic parts*** and ***categorical relationships*** which are ***bound*** together into structures

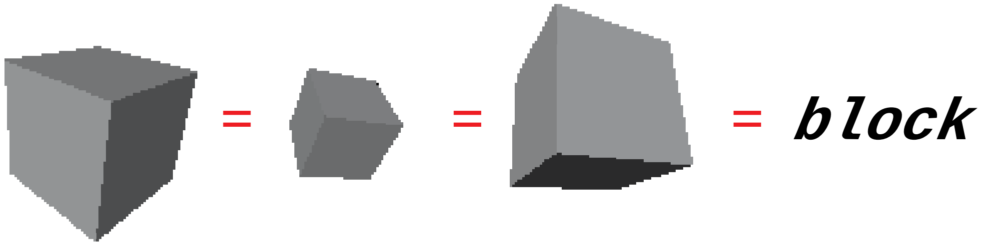
principles:

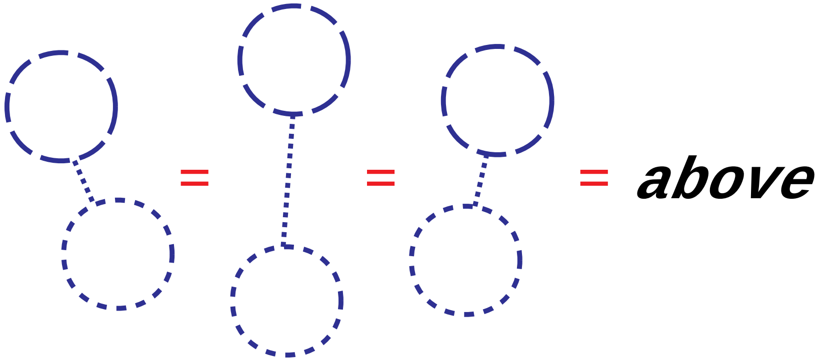
recognition: **invariance** to extraneous factors (pose etc.)

categorization: **invariance** to within−category differences

meta−cat, structure: **explicit coding** of parts and relationships

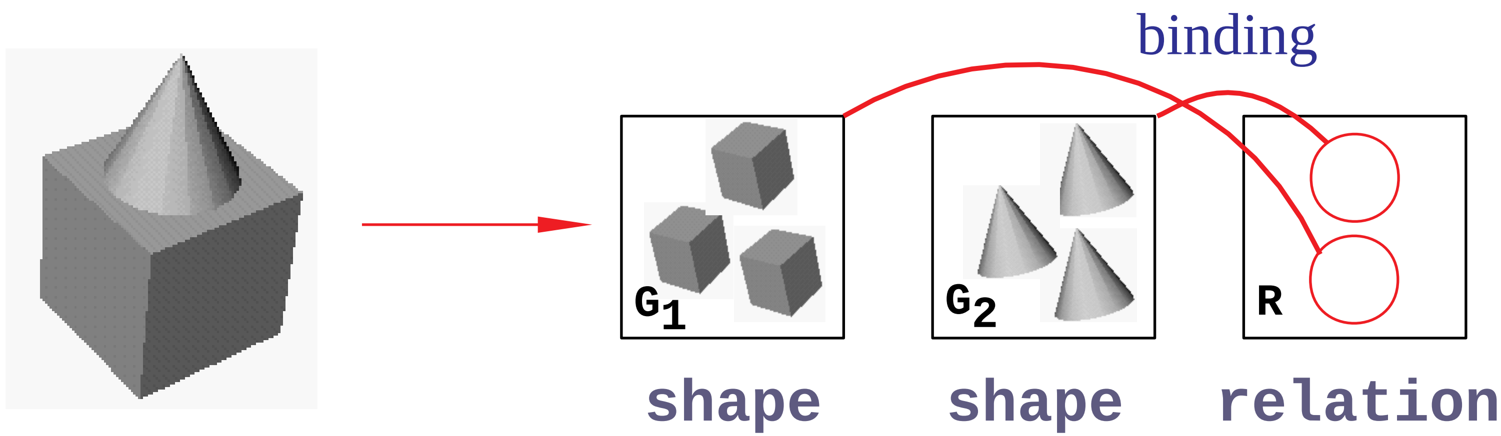






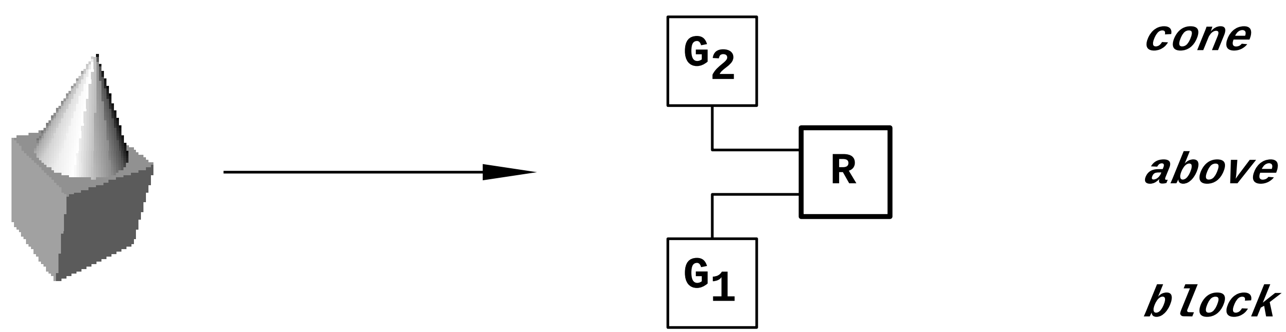
symbols + structure, applied

structural decomposition:

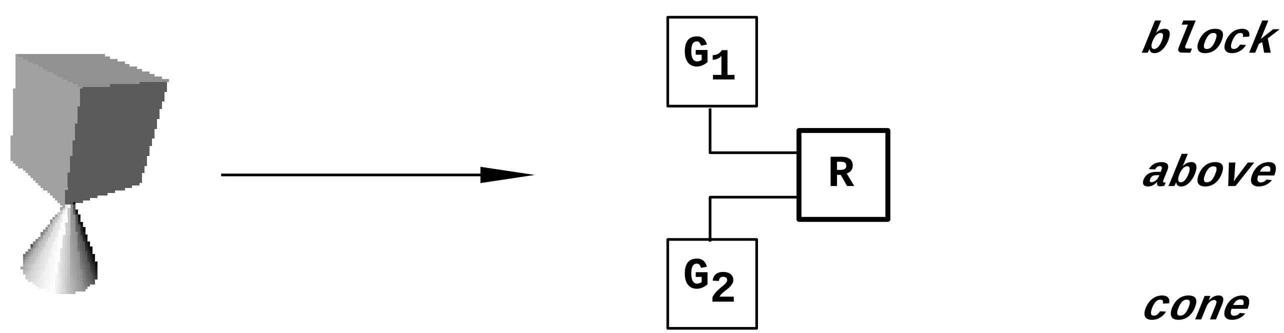


generalization:

old



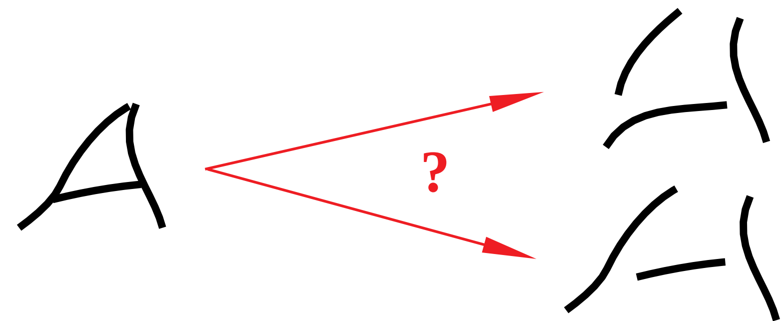
new



some **problems** with the
symbols+structure idea:

shapes = ***symbols***
standing for ***generic parts***
and ***categorical relationships***
which are ***bound*** into structures

* structural decomposition is not unique:



* metric (as in *metric vs. categorical*) issues are not resolved:

* structural decomposition defies computational implementation:

* contrary to the prediction of structural theories, recognition is generally not fully invariant (not even under object translation)

... a system of knowledge in which each constituent element is exactly measured, and in which the relations among the elements within the system are exactly measured.

But *definitio est negatio*. Boundaries which include, exclude.

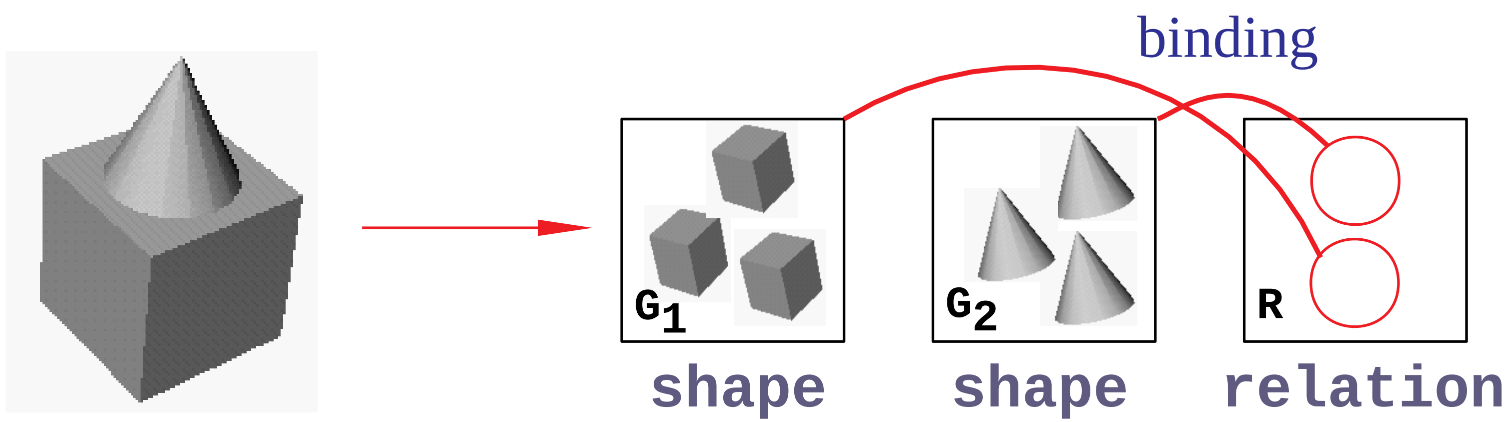
William Lowe Bryan
The Measured and the not−yet−Measured
Powell Lectures at Indiana University, 1940

Omnis determinatio negatio est.

B. Spinoza, *Epistolae* 50.41, 1674

D. Hofstadter
Variations on a theme as the crux of creativity, 1985

symbols + structure: **predictions** for psychophysics:



@ because absolute locations of parts do not figure at all in the structural decomposition, **translation invariance** is expected

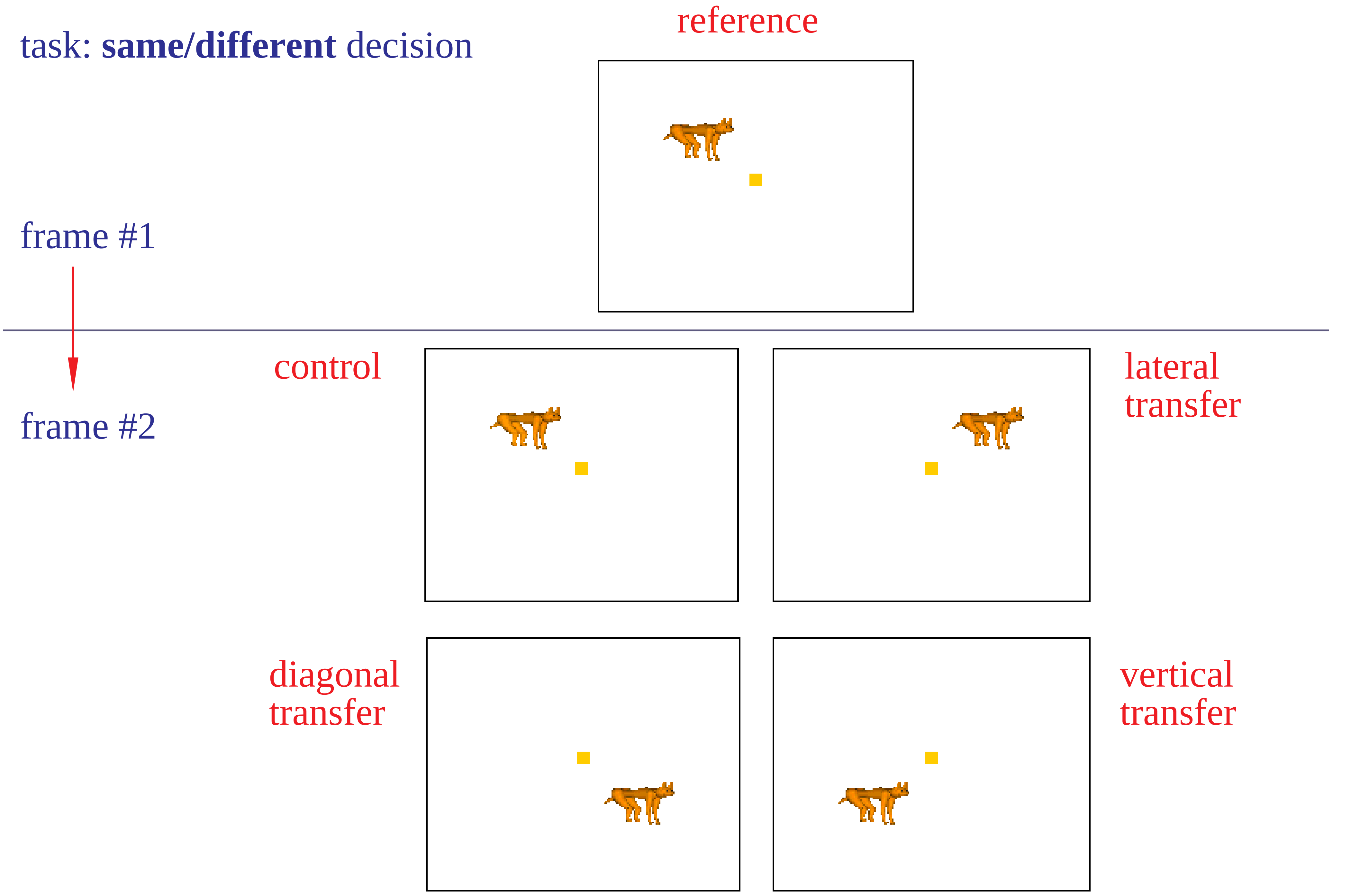
@ assuming that these "units" have real counterparts, they should be amenable to **priming:**

- **shape−based** priming, irrespective of location
- **relation−based** priming, irrespective of shape

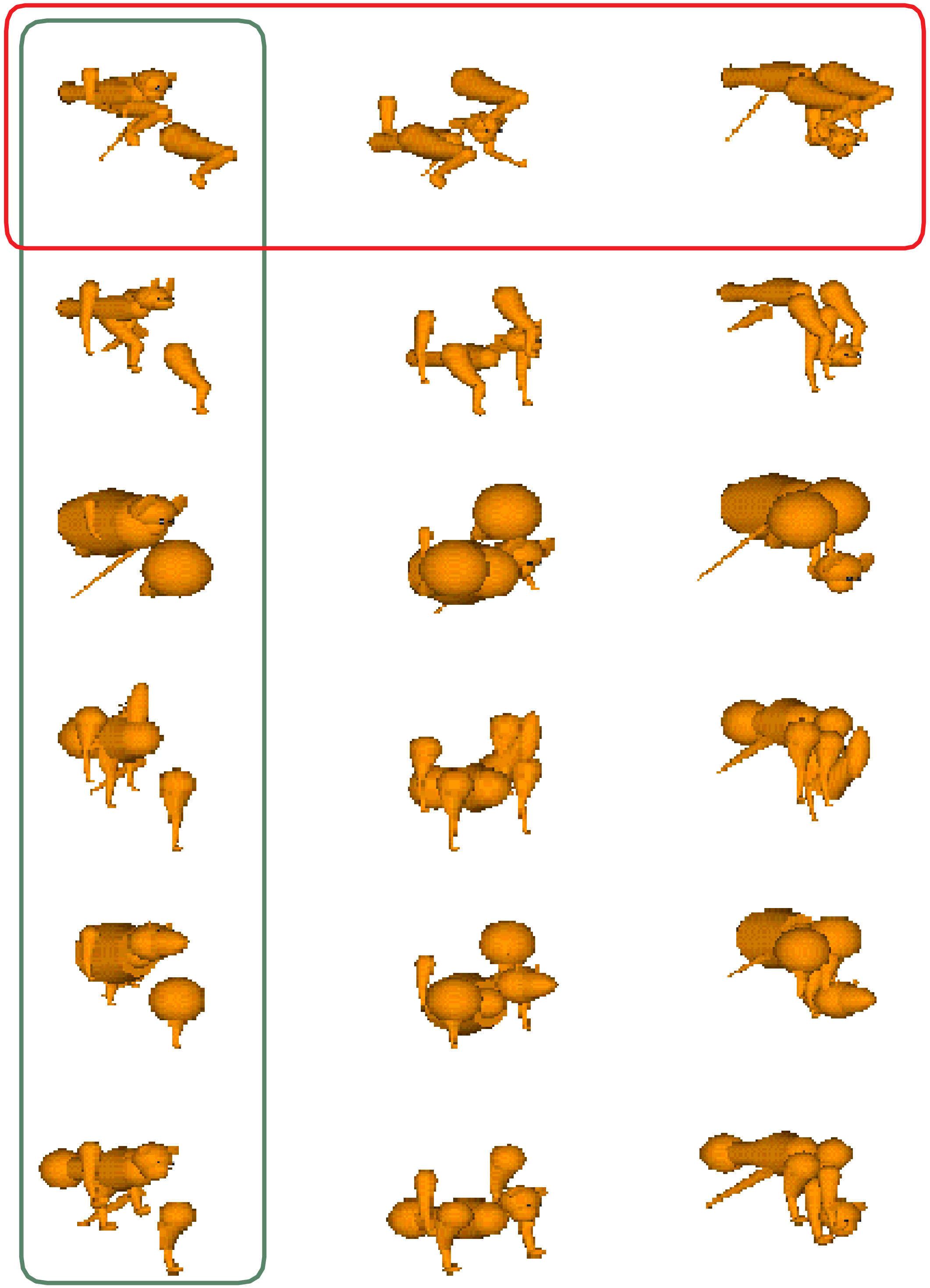
is there translation invariance?

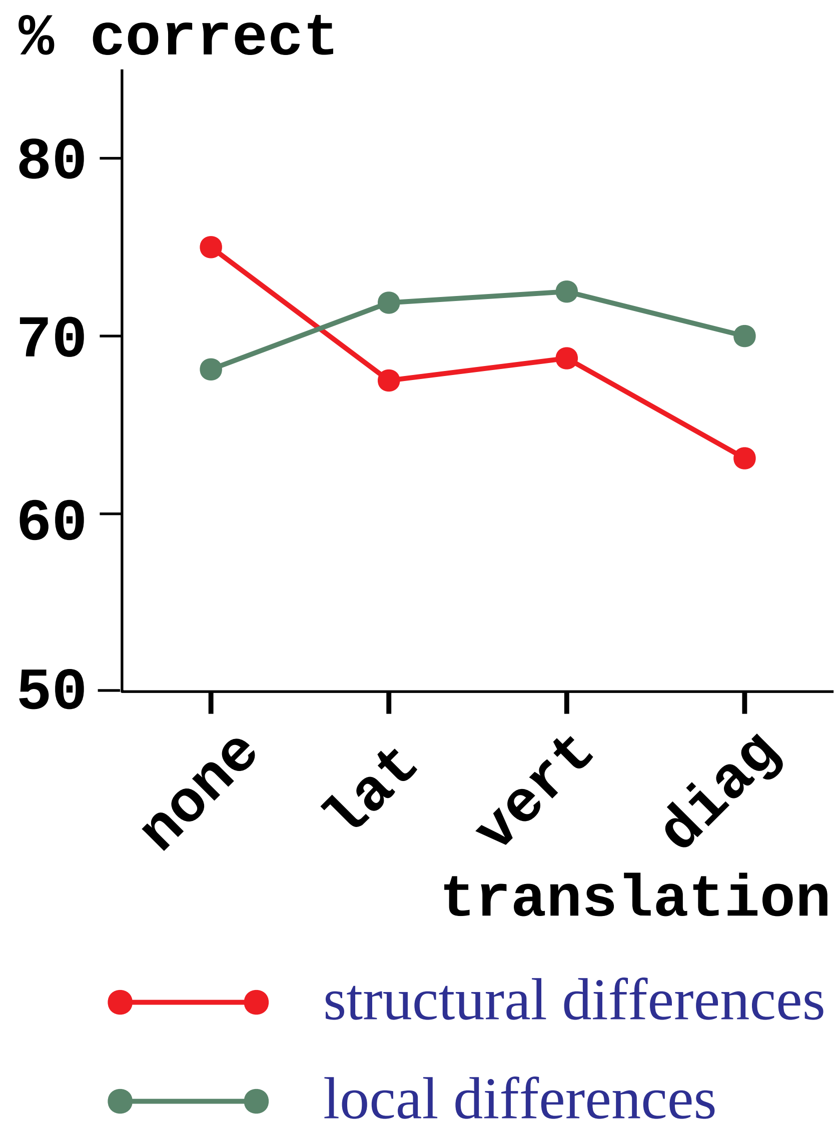
M. Dill & S. Edelman, 1997

task: **same/different** decision



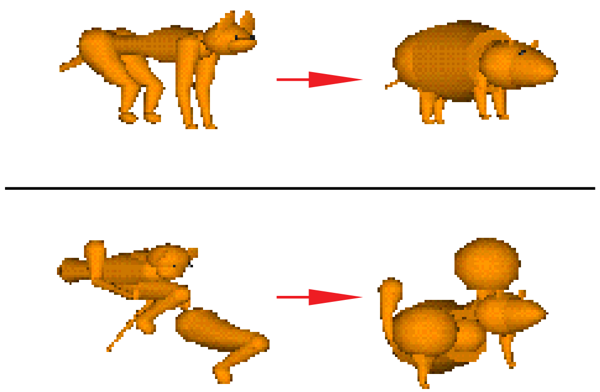
imperfect translation invariance



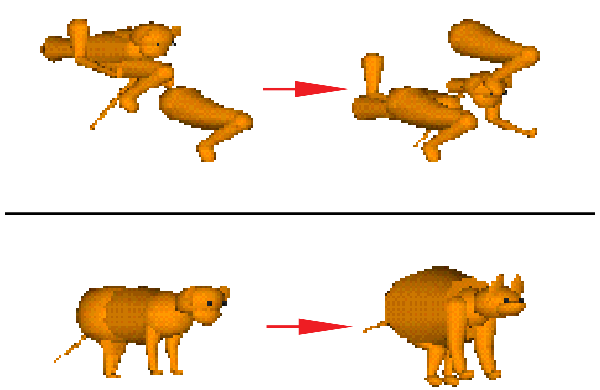


summary of invariance results

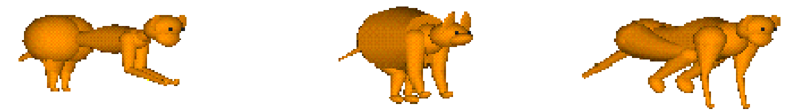
local features are diagnostic
→ full translation invariance

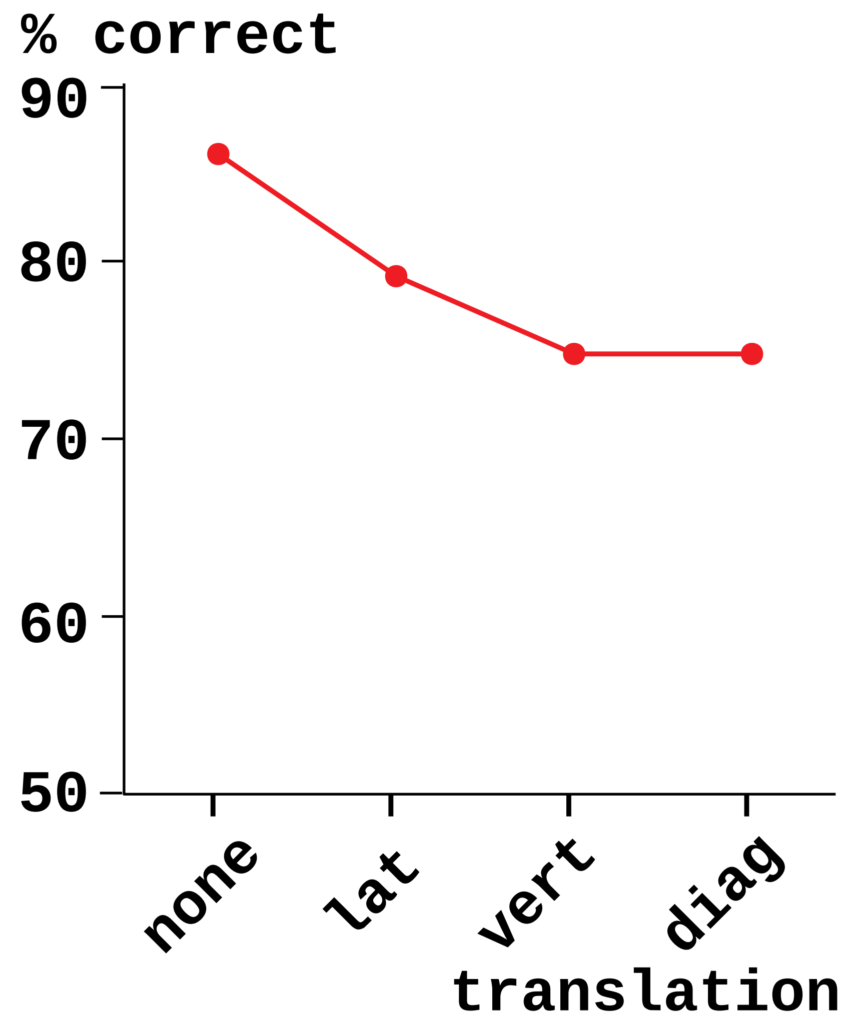


configurational features matter
→ **imperfect translation invariance**



imperfect translation invariance
with another class of objects

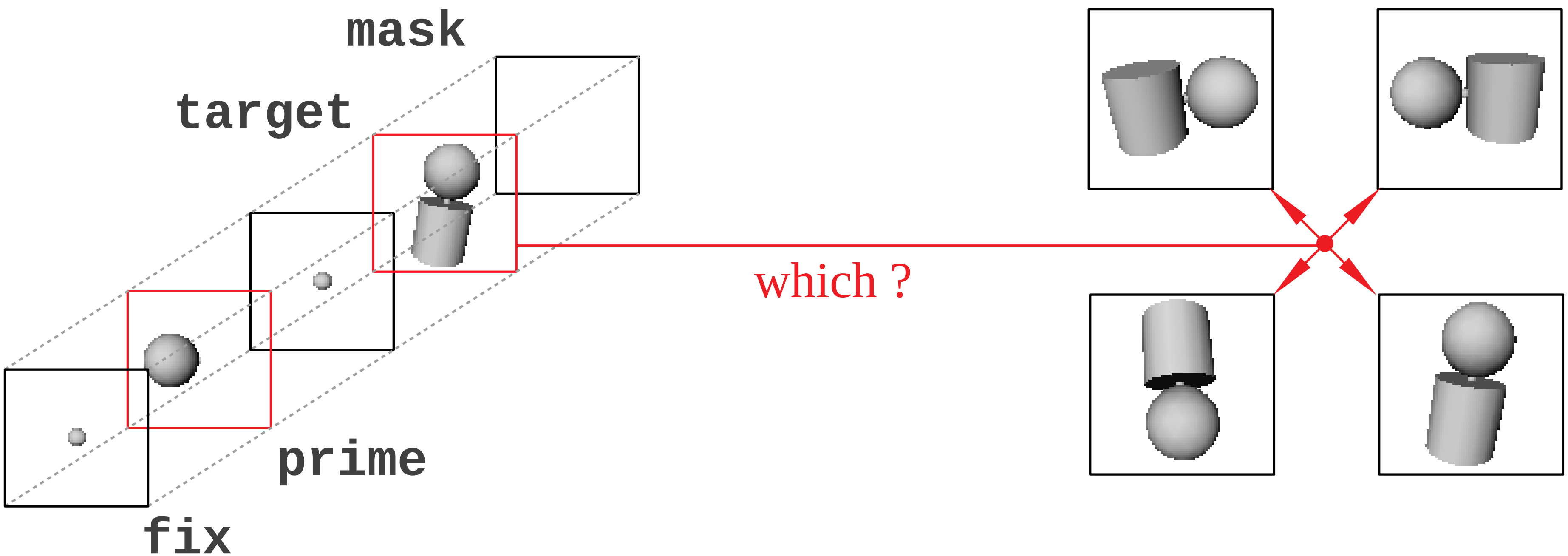




is there shape−based priming, irrespective of location?

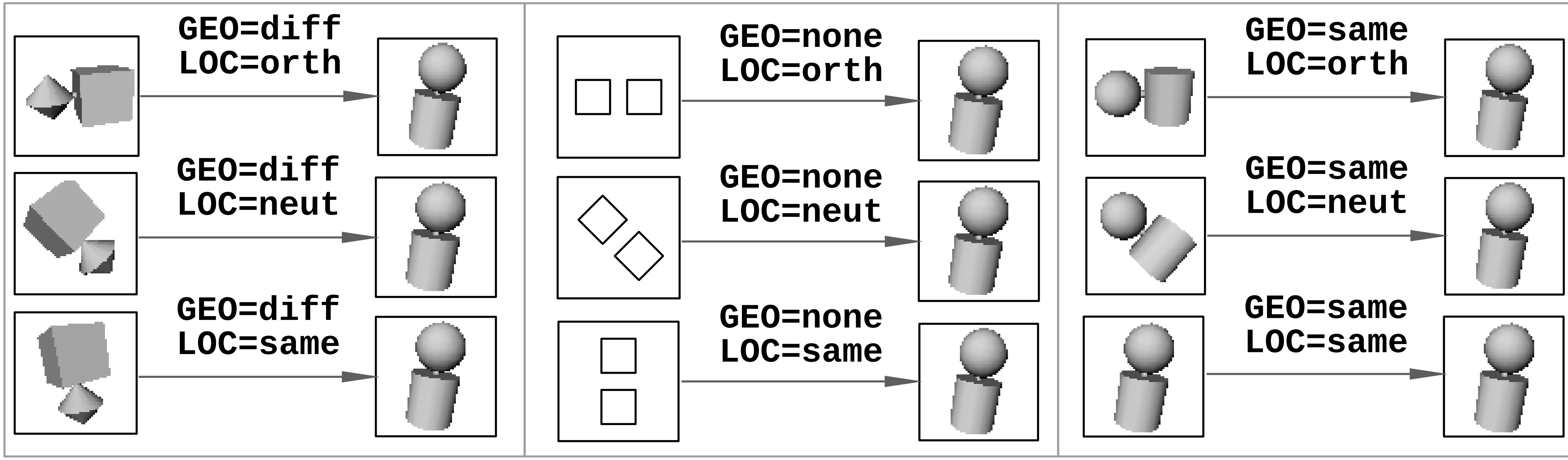
S. Edelman & F. Newell, 1998

task: four−alternative forced choice (**4AFC**)



manipulate separately:

- − prime/target shape
- − prime/target location

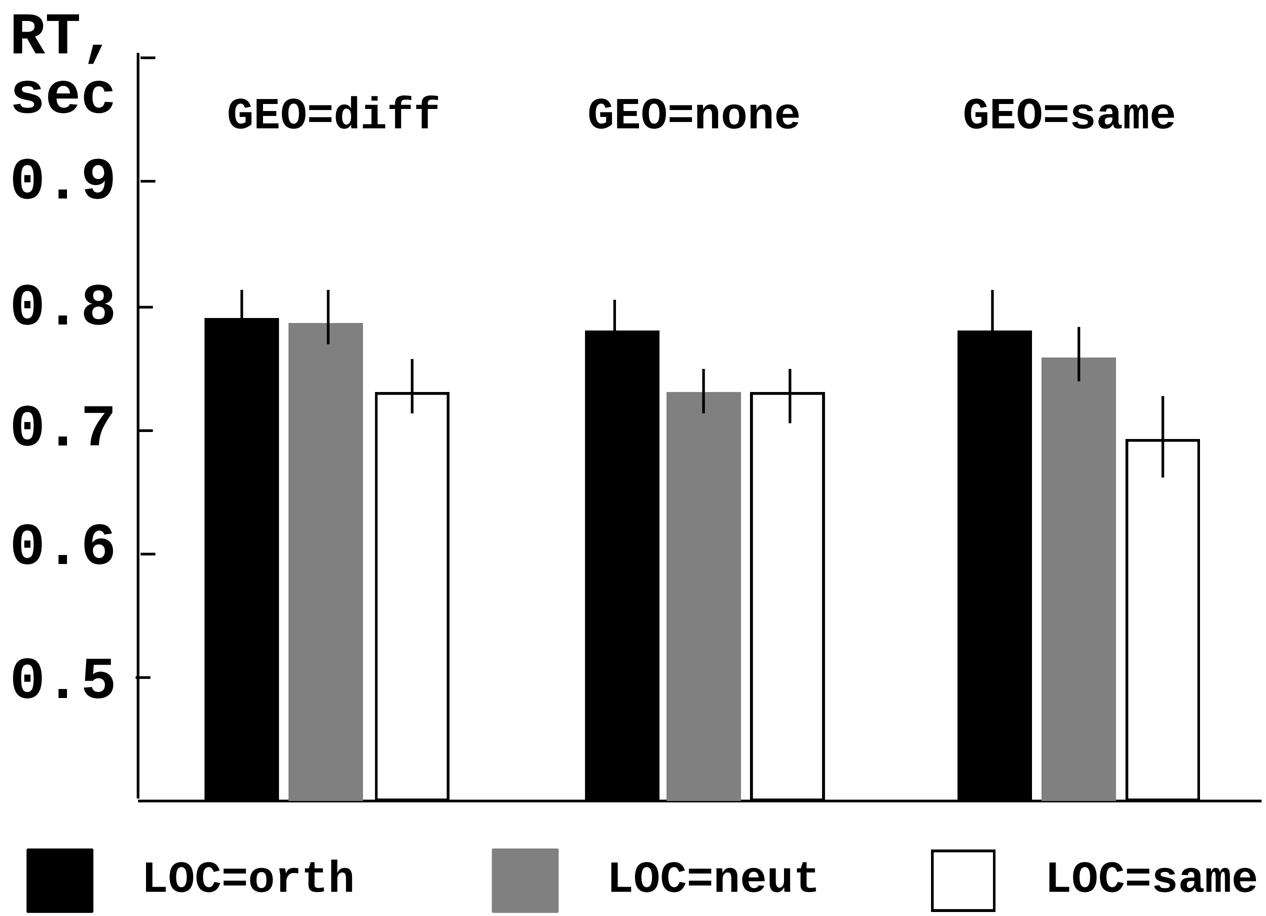


Response Times:

GEO gain: **n.s.**

LOC gain:
from **orth**: **66** *ms*
from **neut**: **42** *ms*

priming by shape (**GEO**)
only in conjunction
with location (**LOC**)



interim conclusion:

an alternative to the symbols + structure idea is needed

* **must fare better on the empirical front**
* **must support:**
  - recognition
  - categorization
  - meta−categorization
  - dealing with **structure**

---

consider: representation based on **similarity**

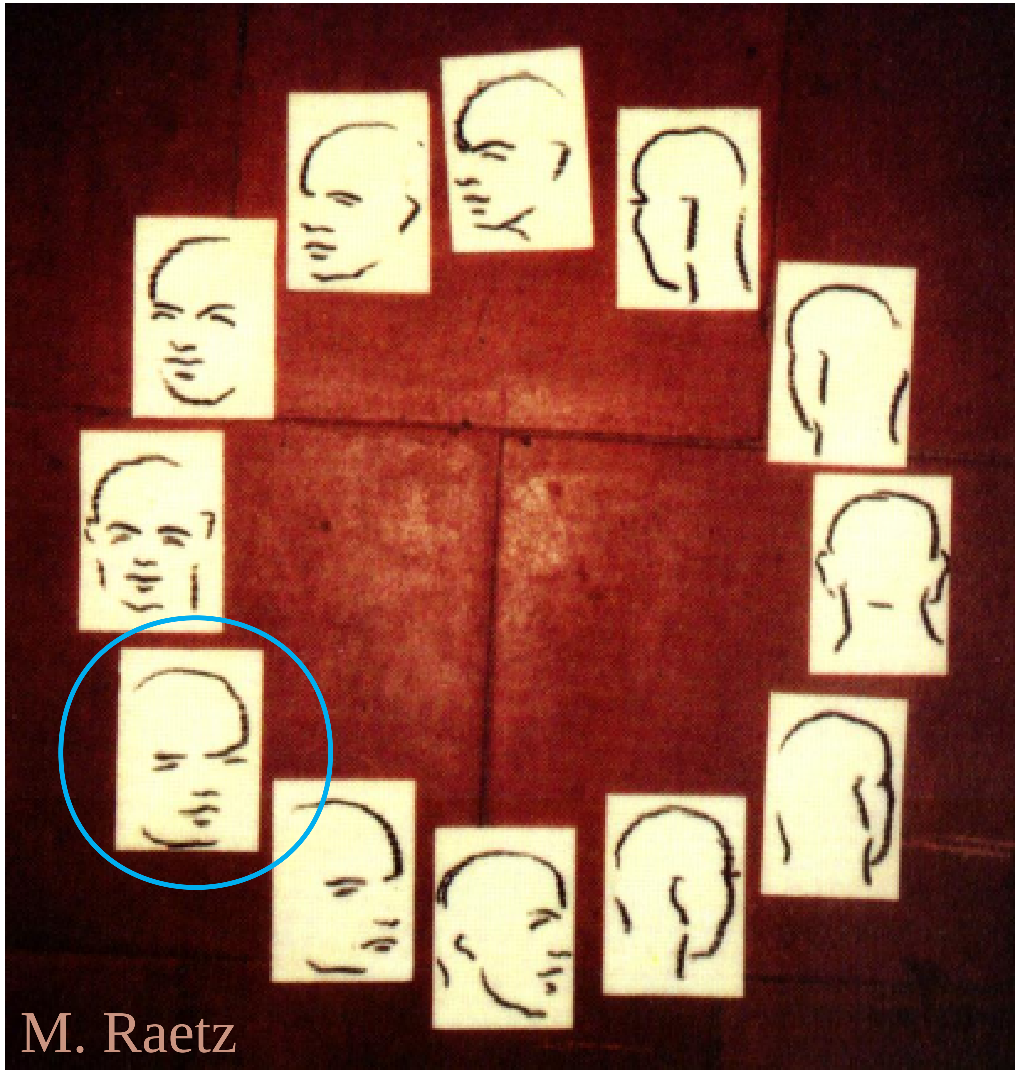
to spatially anchored reference−shape fragments

the similarity−based scheme; issue #1:

* RECOGNITION:

dealing with novel views
of shapes

principle: interpolation of viewspace

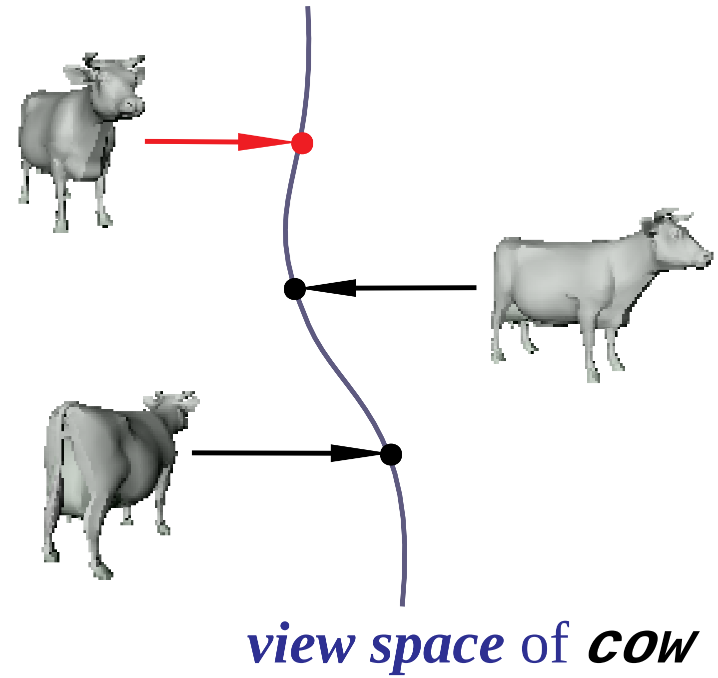


the similarity−based scheme; issue #1: T. Poggio & S. Edelman, 1990

⁕ RECOGNITION:

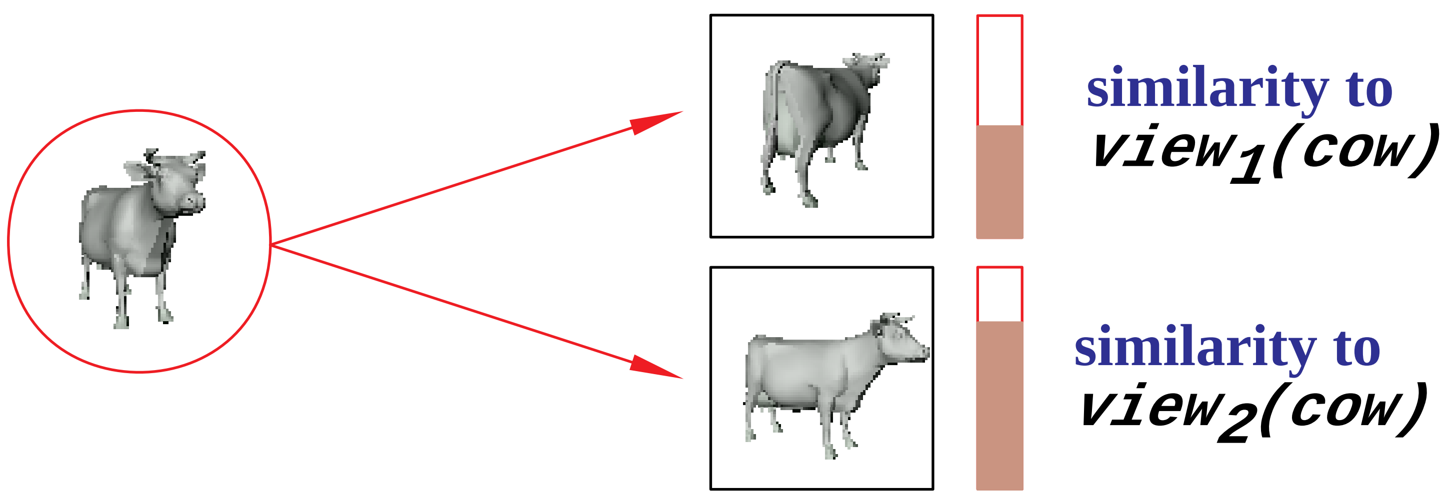
dealing with novel views of shapes

principle: interpolation of viewspace



implementation: similarities to sample views

similarities to multiple views

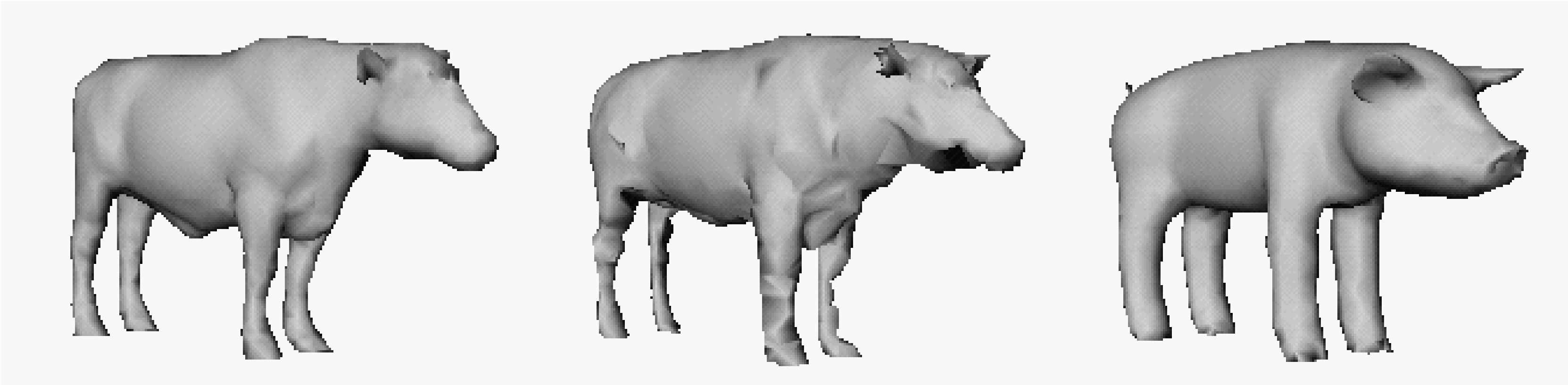


the similarity−based scheme; issue #2:

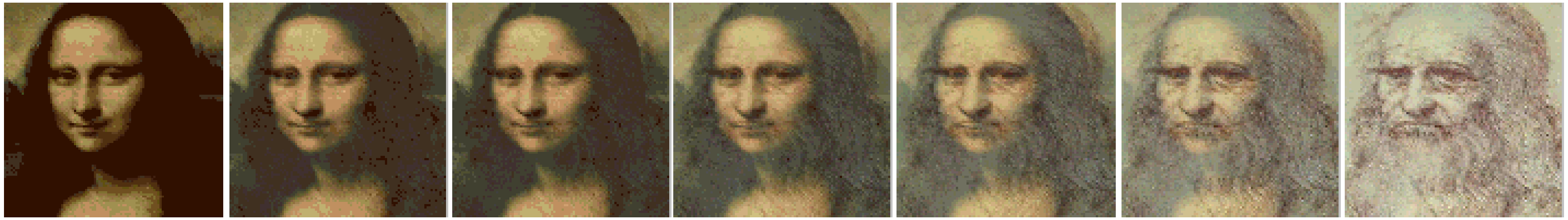
⁕ CATEGORIZATION:

dealing with novel instances
of shape categories

principle: interpolation of shape space

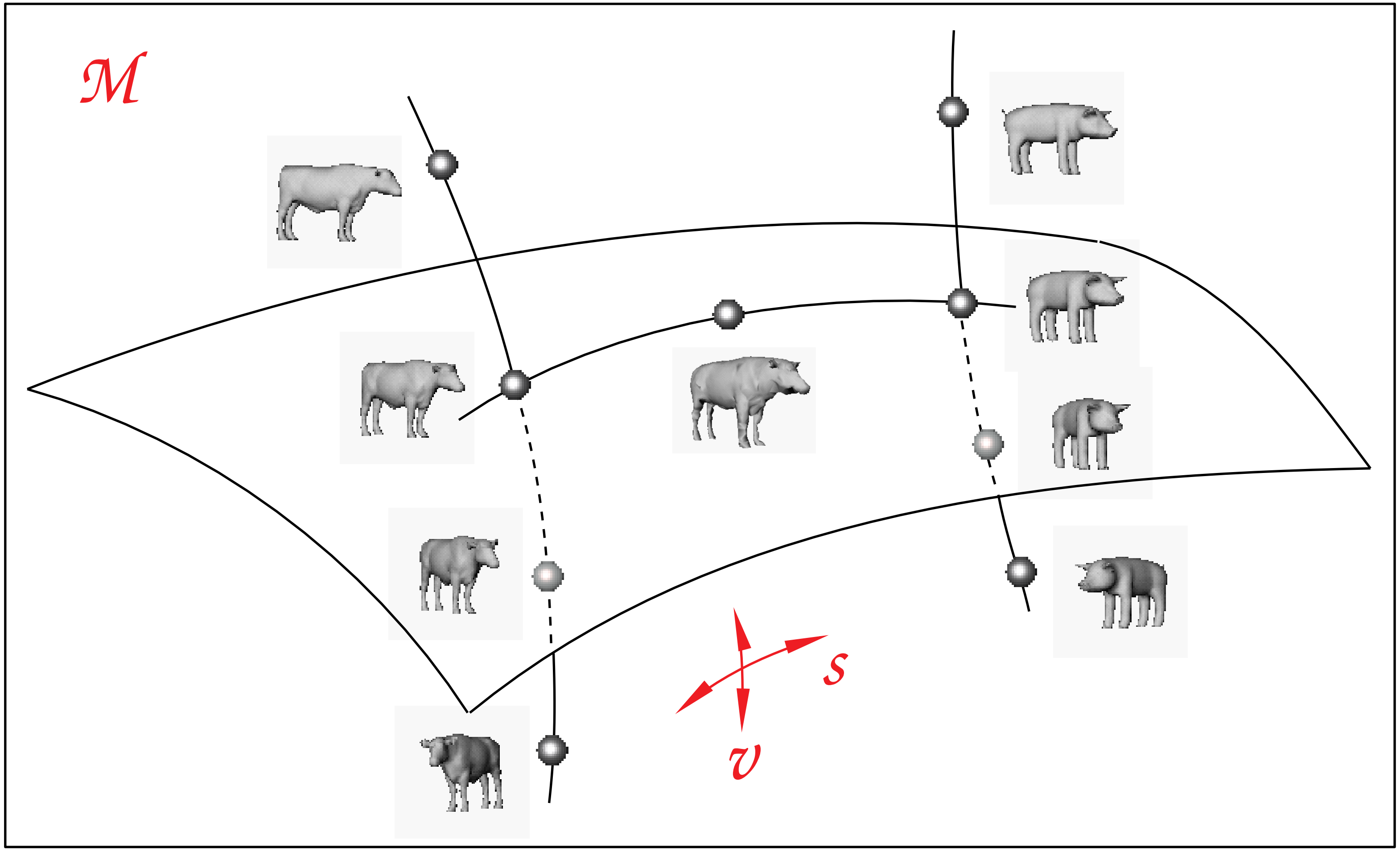


morphing



the similarity−based scheme; issue #2

an illustration of the relationship between view− and shape−space interpolation:



v **view** change, rotation (transformation)

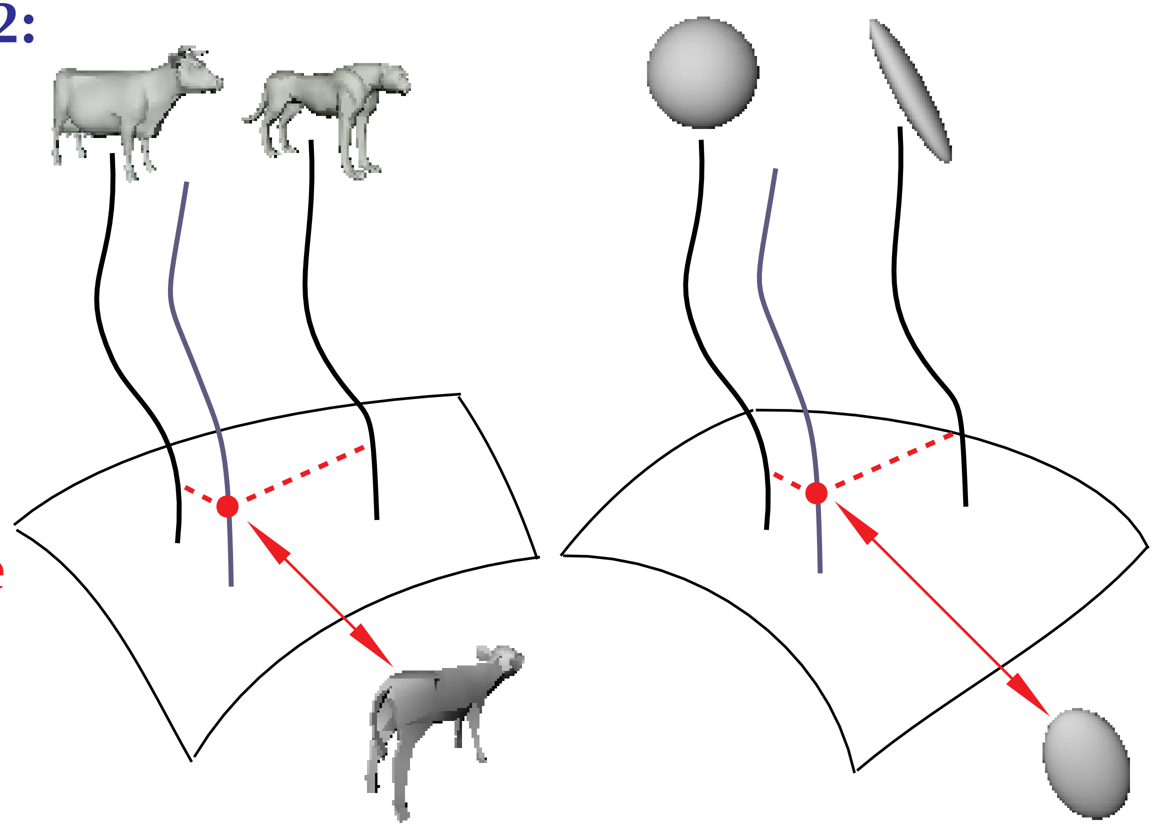
s **shape** change, morphing (deformation)

$\mathcal{M}$ **measurement** space (very high−dimensional; e.g., retina)

the similarity−based scheme; issue #2:

✱ CATEGORIZATION:

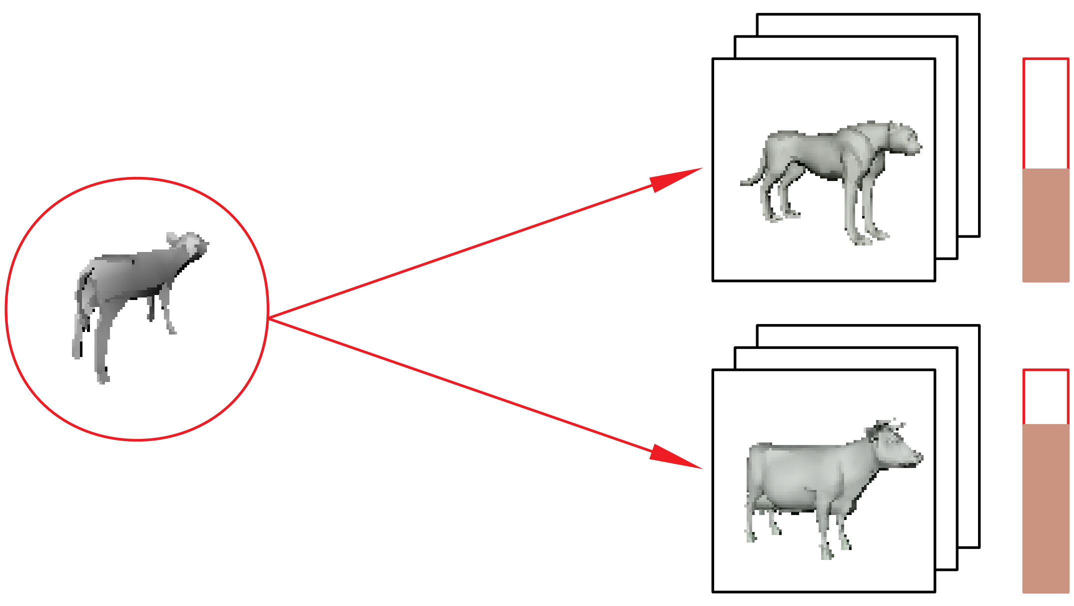
dealing with novel instances of shape categories



principle: interpolation of shape space

implementation: similarities to sample viewspaces

similarities to multiple class prototypes



similarity to ***cheetah***

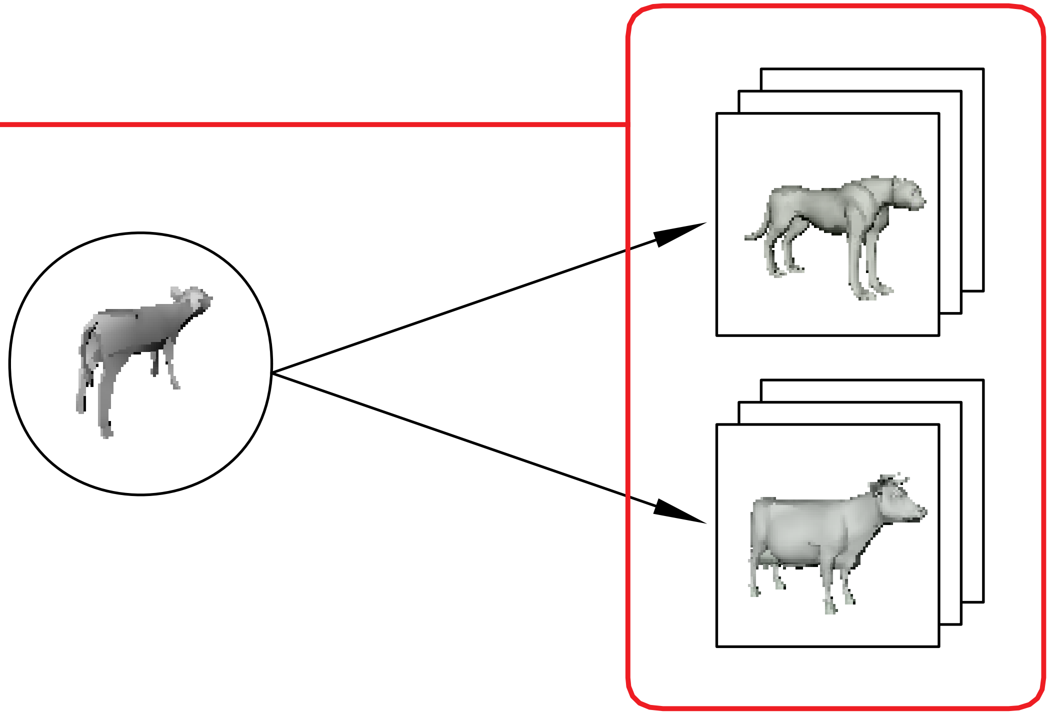
similarity to ***cow***

the similarity−based scheme; issues #1, 2:

⁕ RECOGNITION and CATEGORIZATION

implemented system:

− 10 reference shapes
− 70 test shapes



recognition: ~95%

categorization: ~85%

category−based processing:

− cluster by similarity

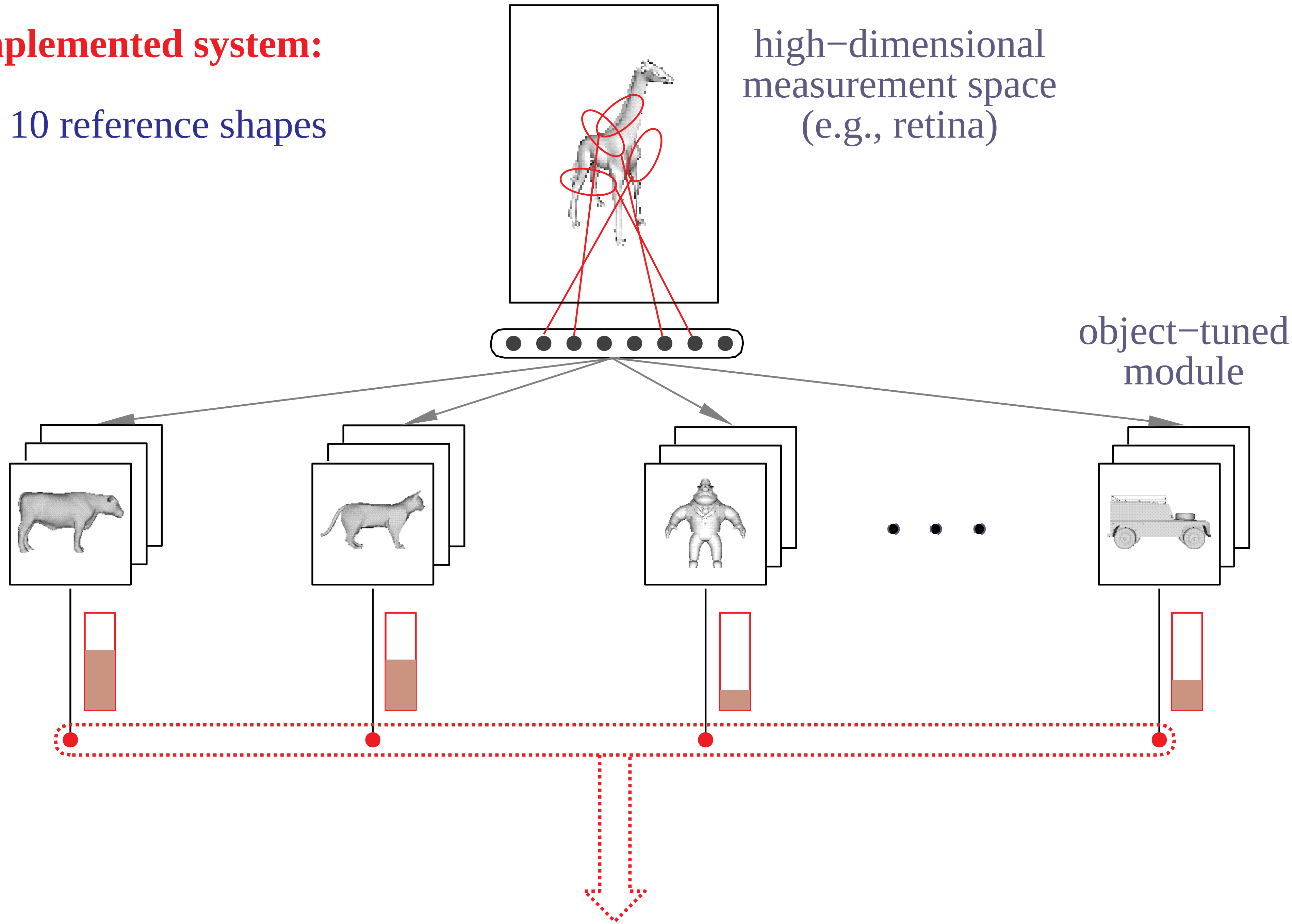
− estimate viewpoint

− imagine new view

S. Edelman & S. Duvdevani−Bar, 1997

implemented system:

– 10 reference shapes



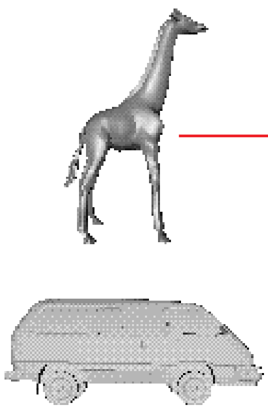
similarities to multiple class prototypes
("**Chorus of Prototypes**")

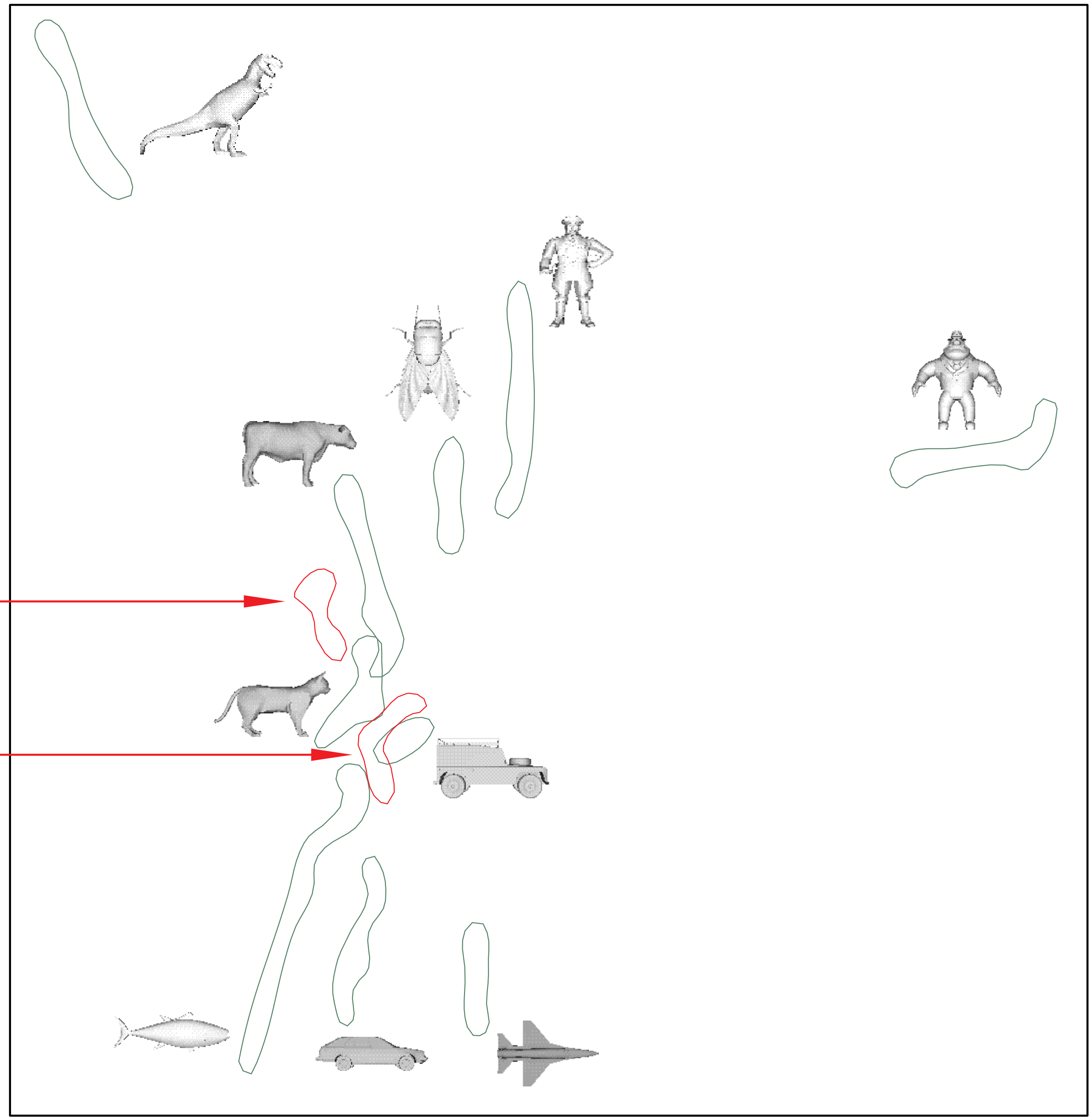
categorization

the **10−D** space spanned by similarities to the **10** reference objects

(embedded into **2−D** to facilitate visualization, using multidimensional scaling)

novel test objects



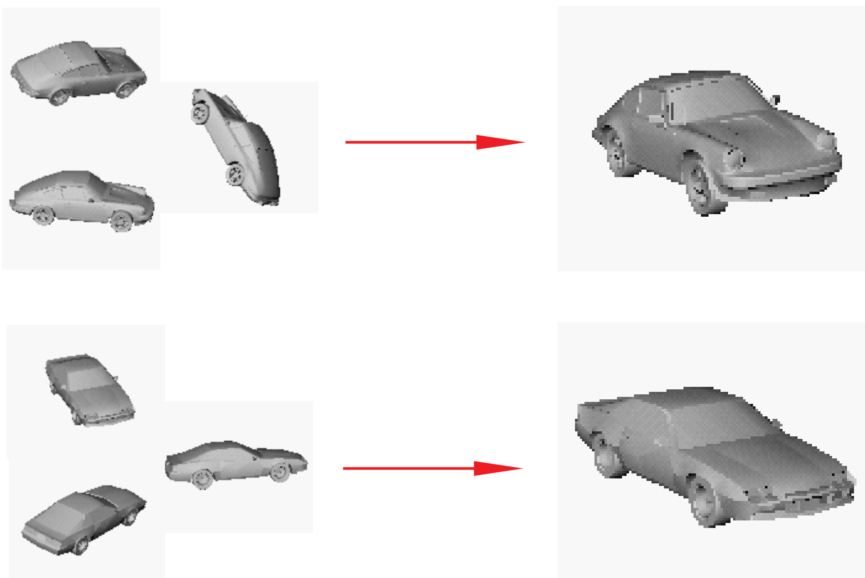


clustering by similarity

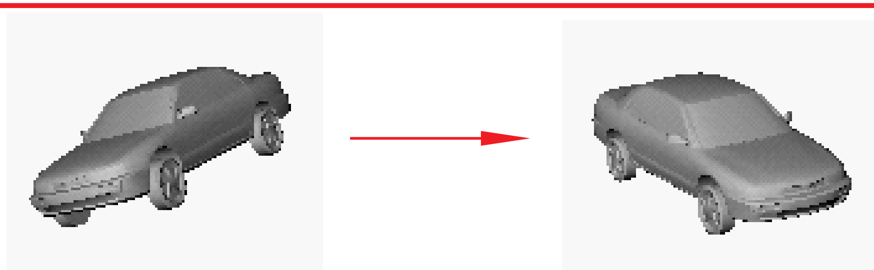
S. Duvdevani−Bar & S. Edelman, 1997

estimation of viewpoint

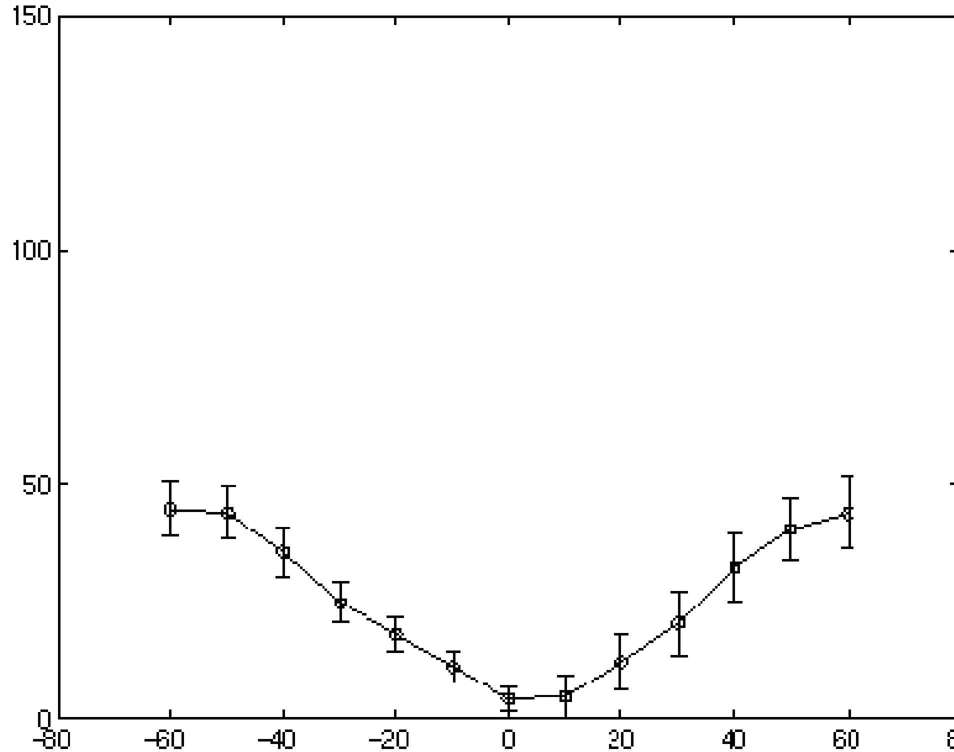
train:



test:



misorientation
between recovered
and true pose



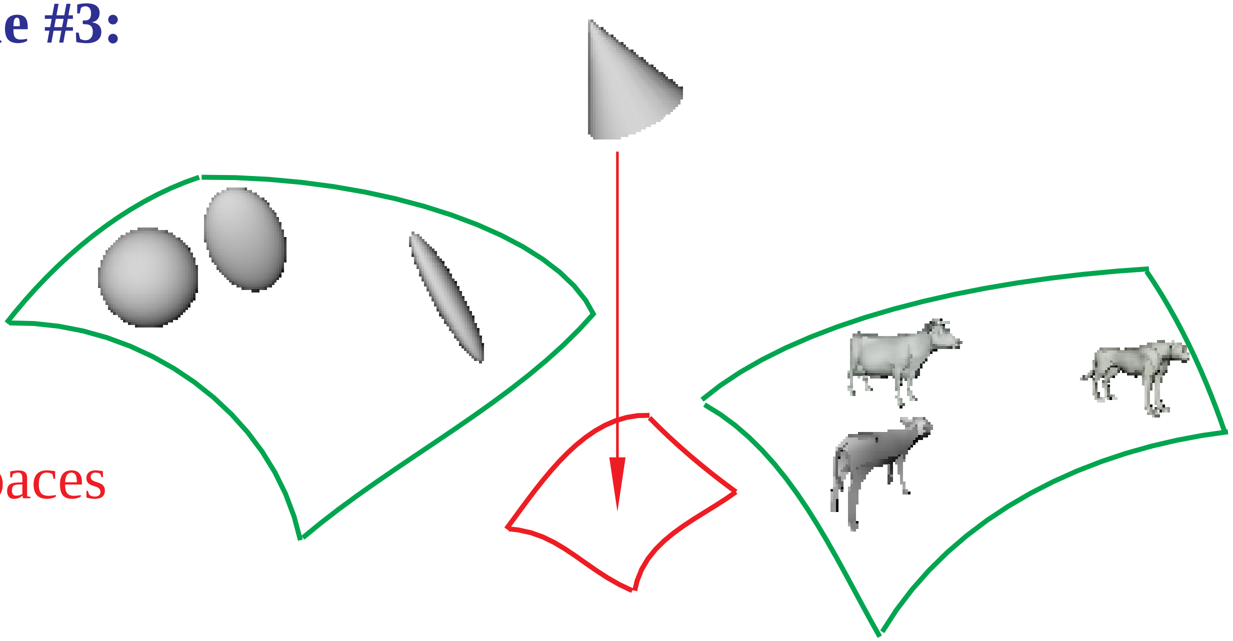
pose

the similarity−based scheme; issue #3:

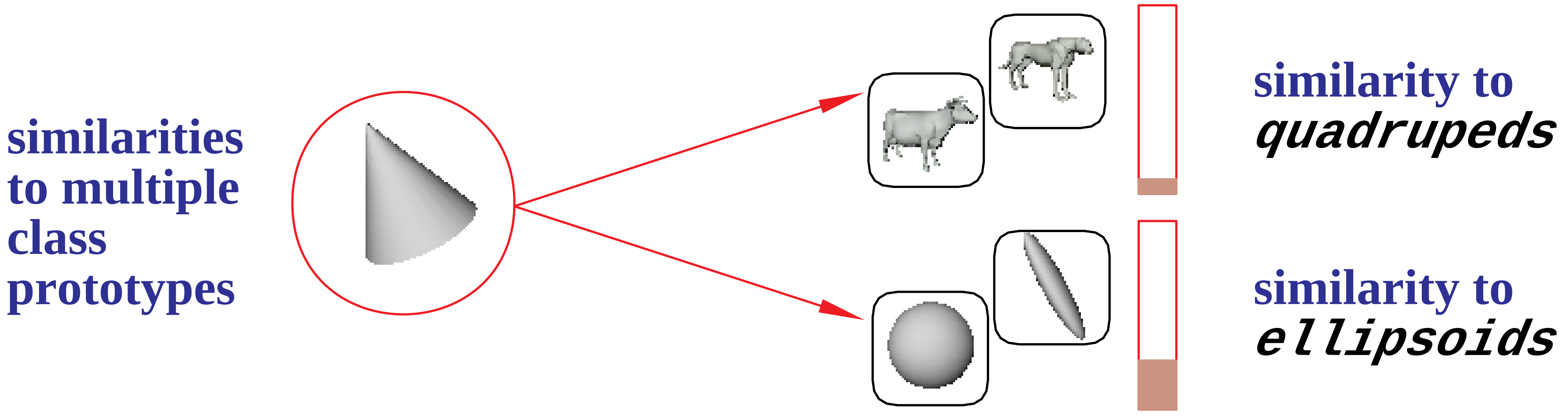
✱ META−CATEGORIZATION:

dealing with **novel categories**

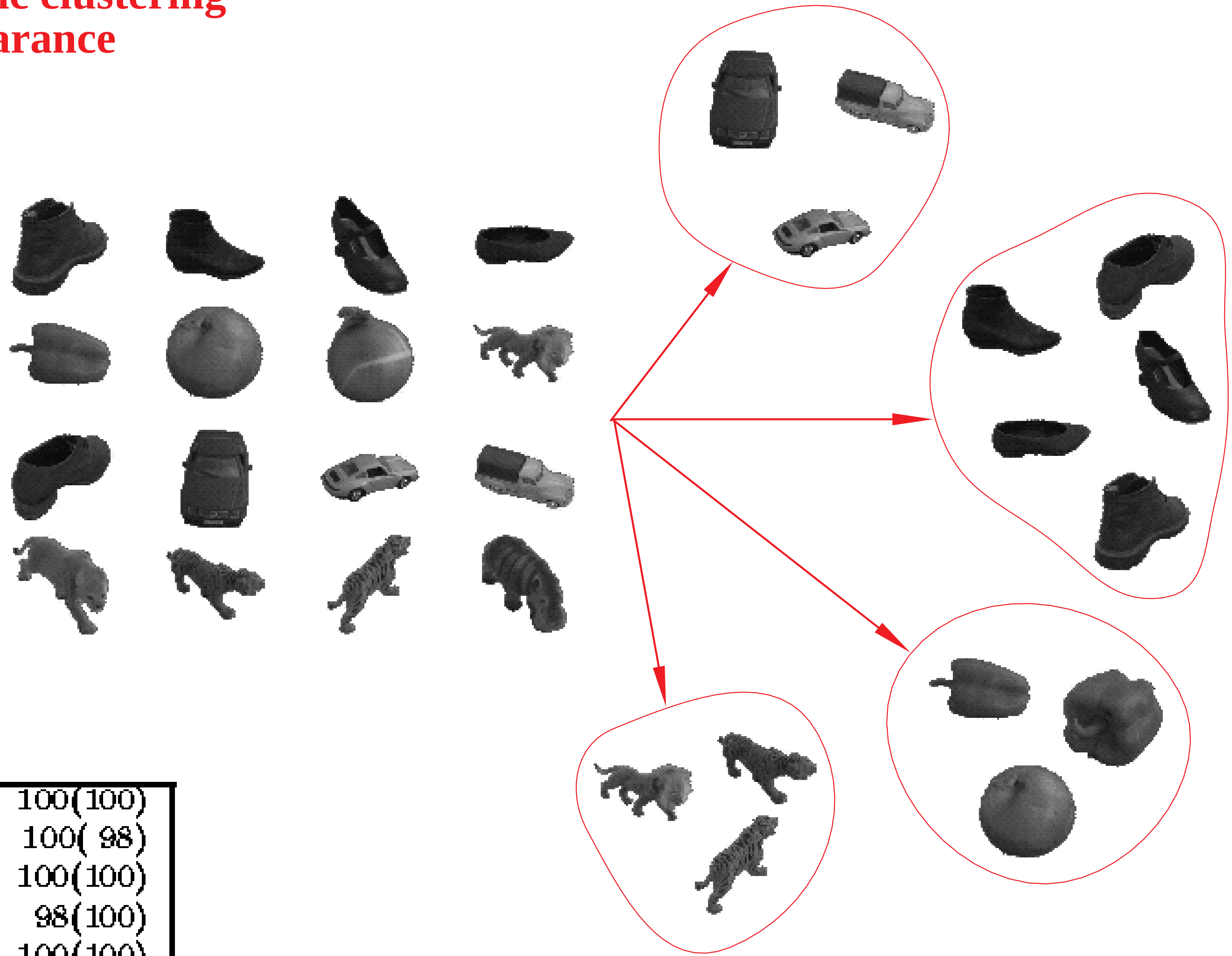
principle: creation of new shape spaces



implementation: comparing similarities to existing shape spaces



automatic clustering by appearance



Shoes	100(100)
Cars	100(98)
Veg.	100(100)
Cats	98(100)
Thick	100(100)
Mean	99(100)
+kNN	99(100)

R. Basri, D. Roth & D. Jacobs, 1998

the similarity−based scheme; issue #4:

⁕ REPRESENTATION of STRUCTURE:

dealing with novel arrangements

− of parts in an object...
− of objects in a scene

a possible solution:

representation based on **similarity**
to spatially anchored reference−shape
fragments **– `"what+where"`** units

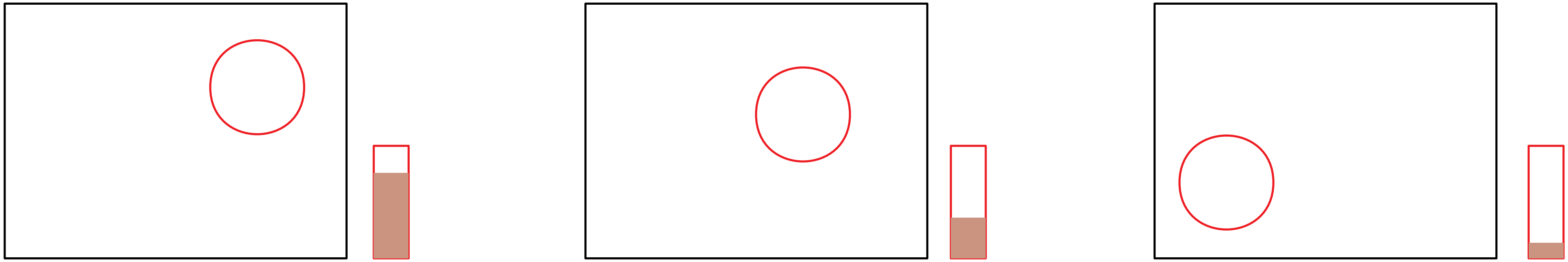
principle: simultaneous interpolation
in shape space and location ("space space")

implementation: similarities to localized shape fragments

a "what + where" unit

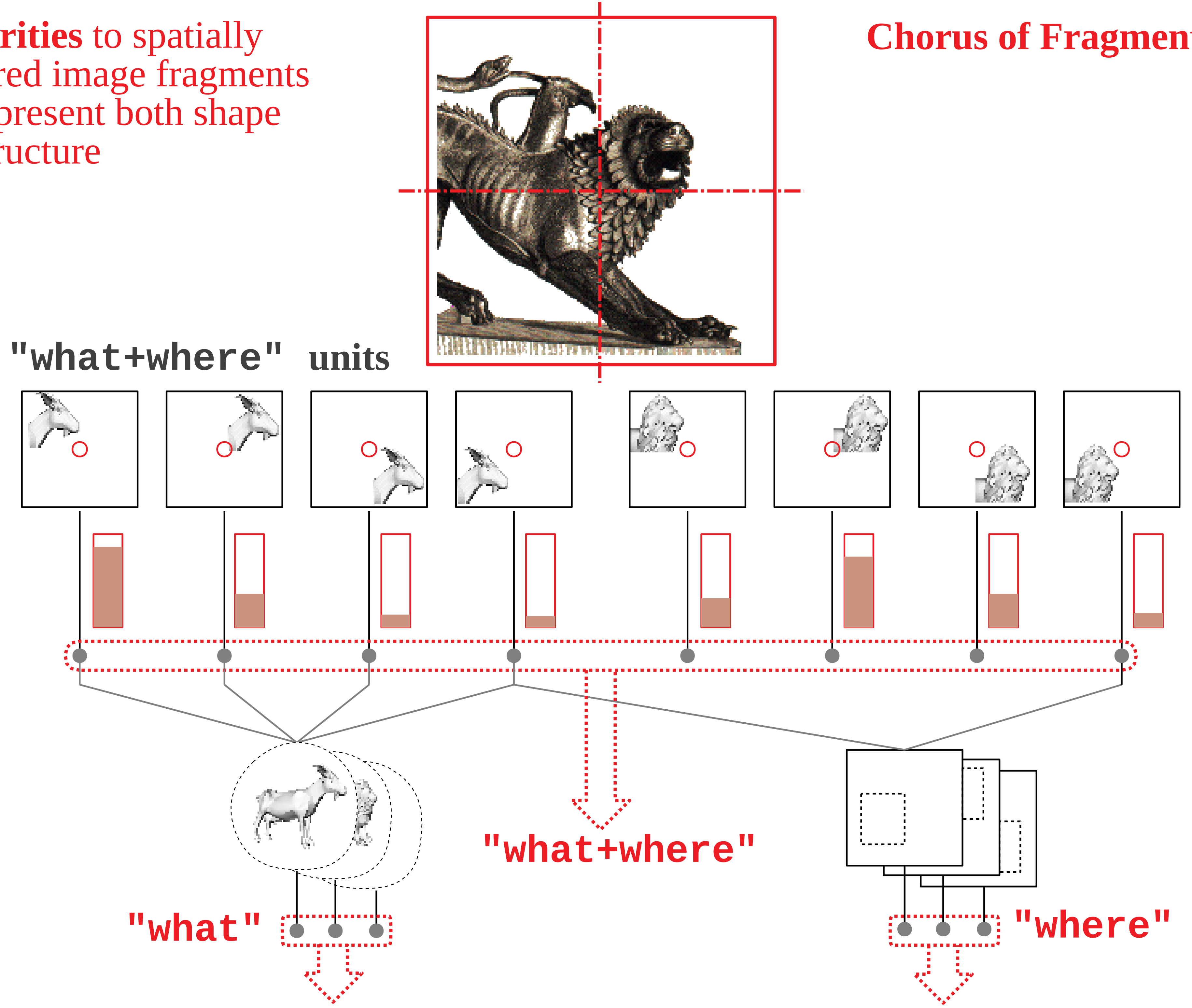
selectivity in shape space:

selectivity in space space:

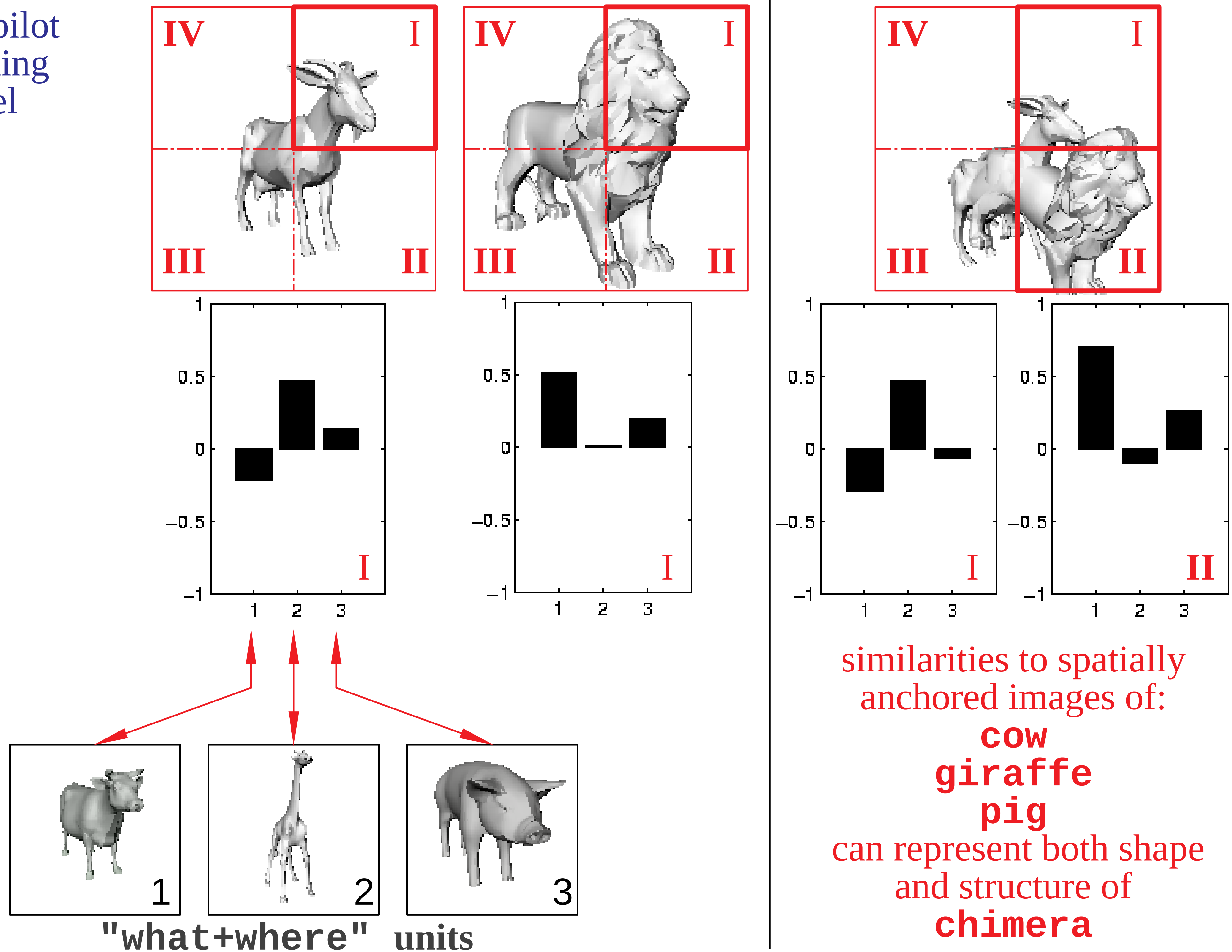


similarities to spatially anchored image fragments can represent both shape and structure

Chorus of Fragments



performance of a pilot working model

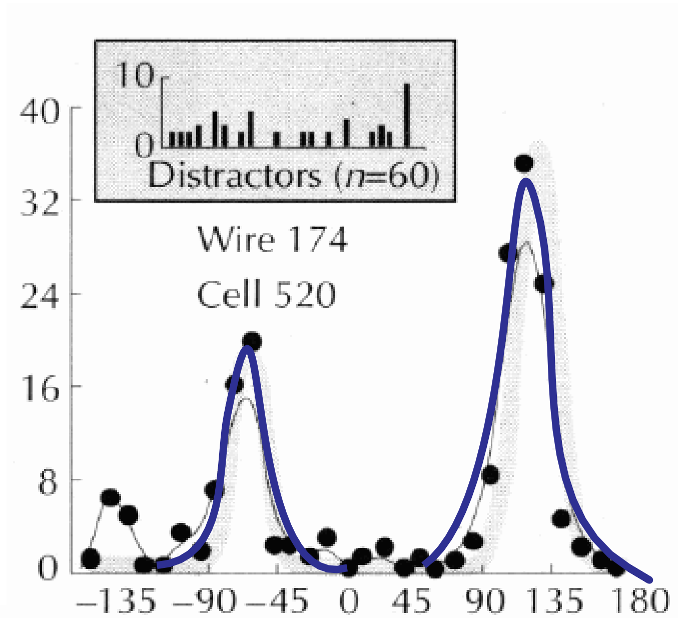


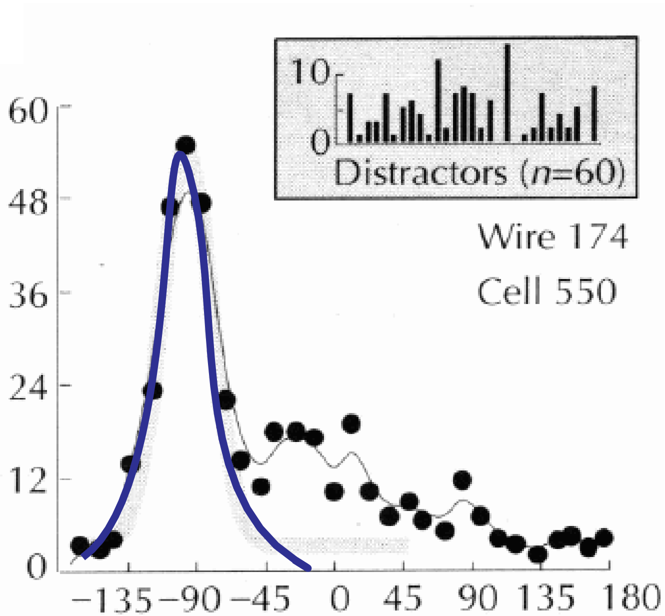
a neurobiological perspective:

N. Logothetis, J. Pauls, T. Poggio, Current Biology 5:552 (1995)

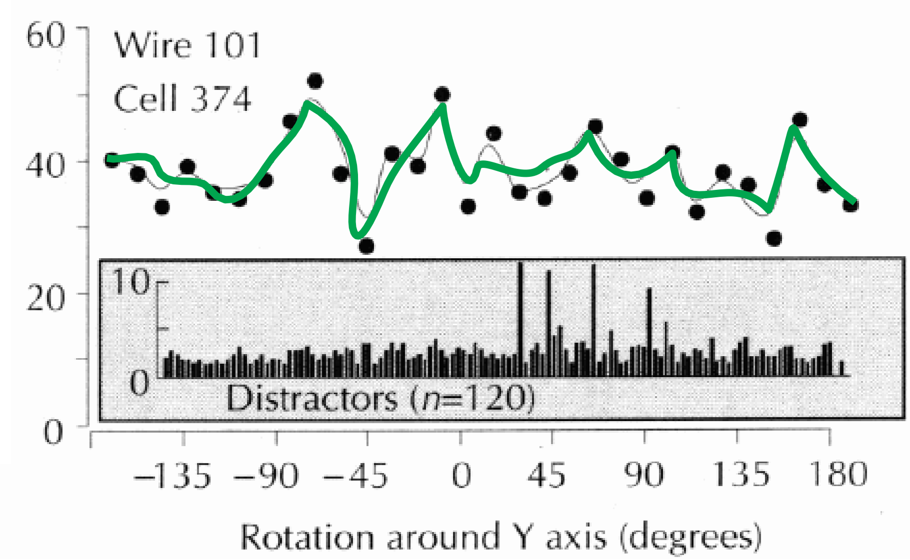
neurons in IT cortex tuned to:

– specific **views** of some objects





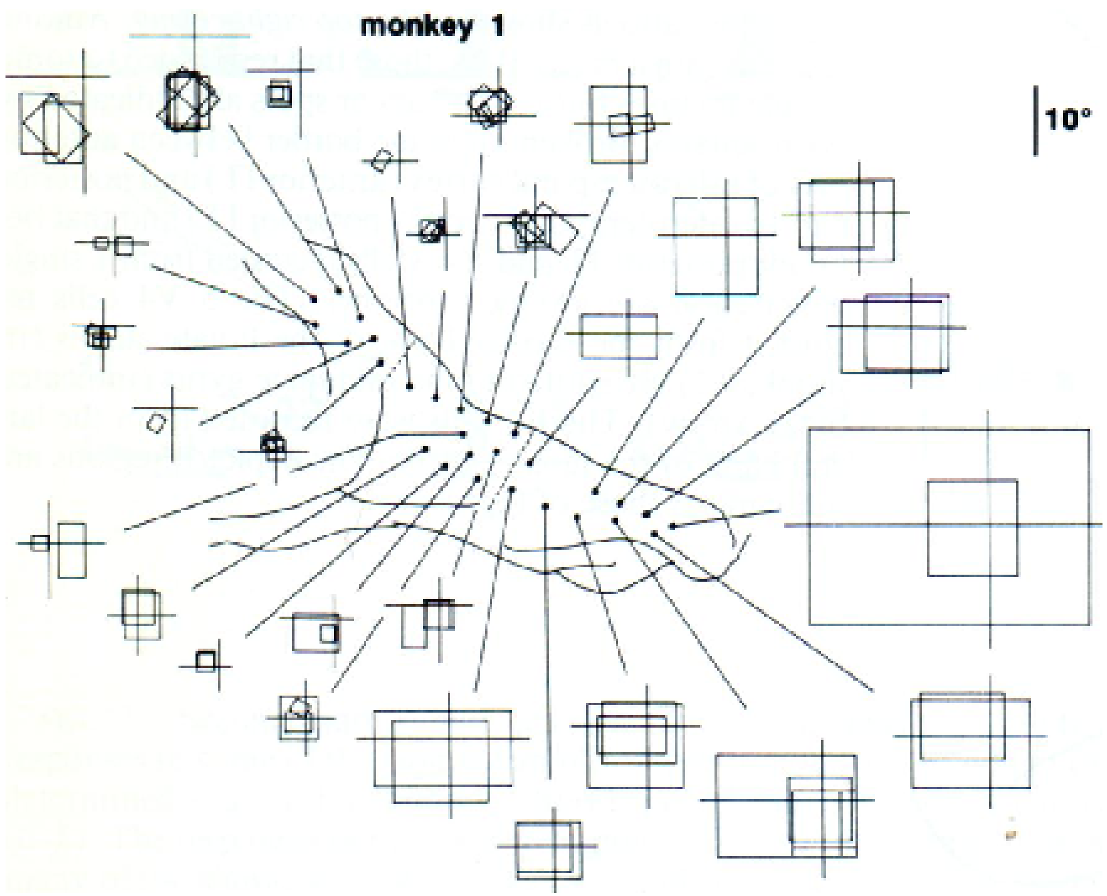
– entire **objects**

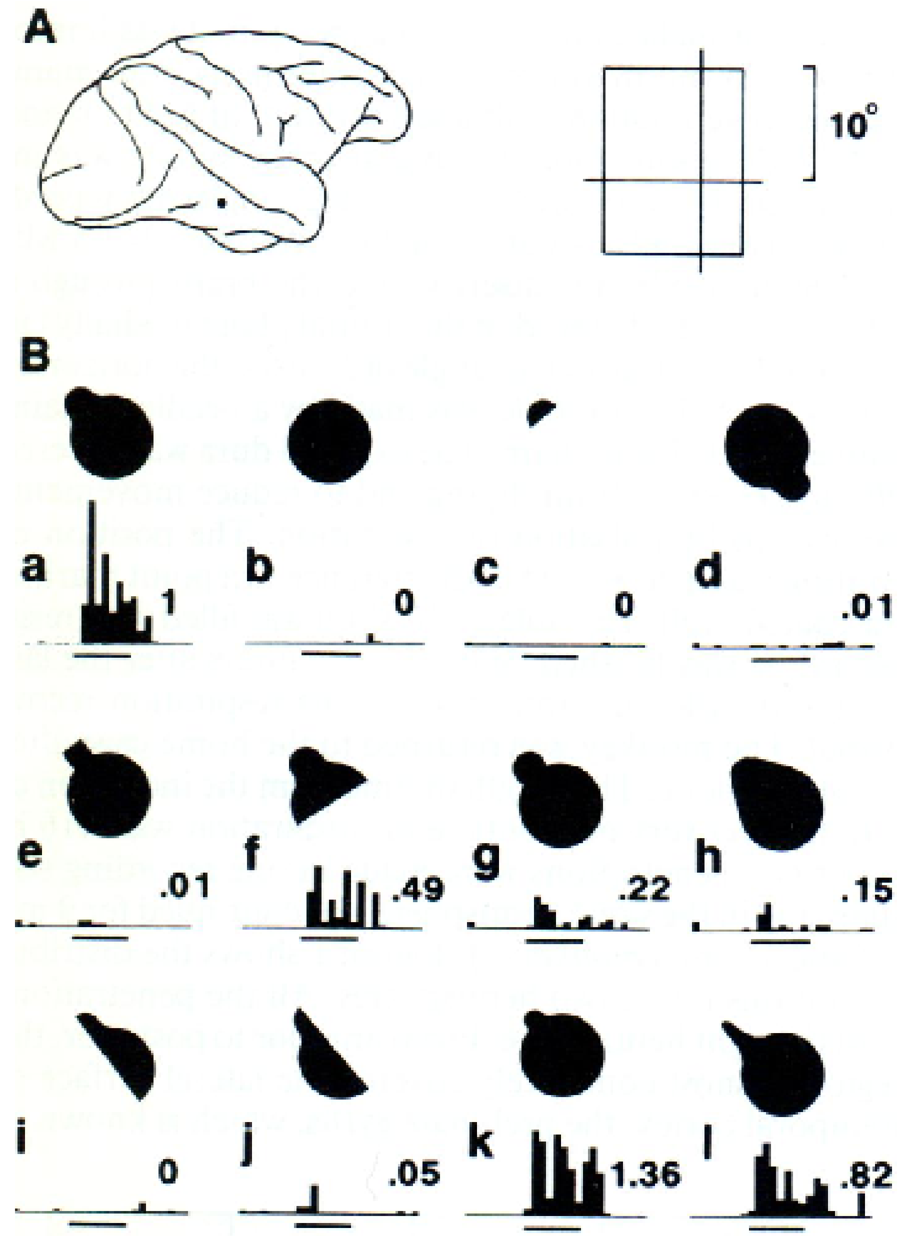


a neurobiological perspective:

E. Kobatake and K. Tanaka,
J. Neurophysiol. 71:856−867 (1994)

neurons in IT cortex
signal both "what"
and "where"

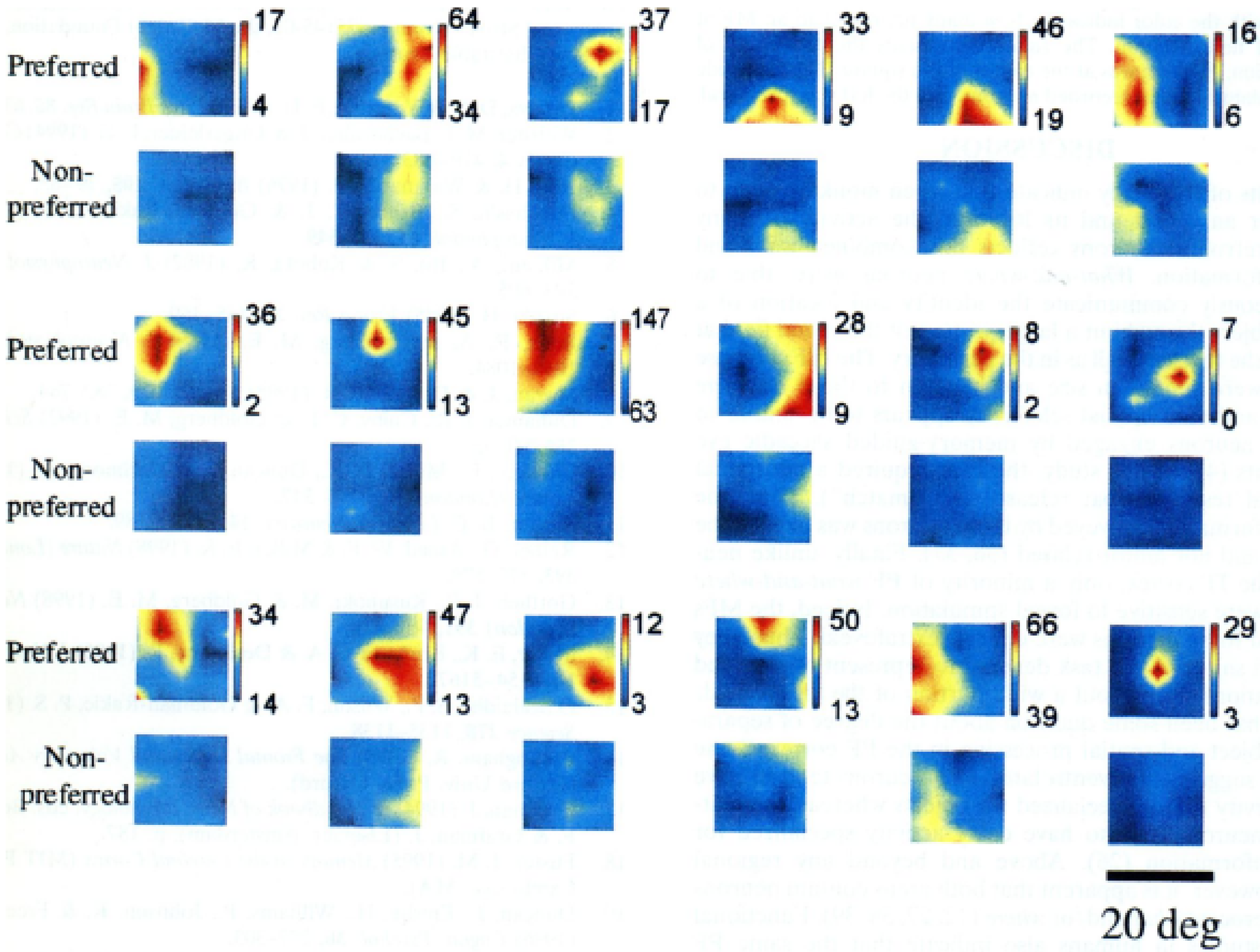




a neurobiological perspective:

G. Rainer, W. Asaad & E. Miller
PNAS 95:15008−15013 (1998)

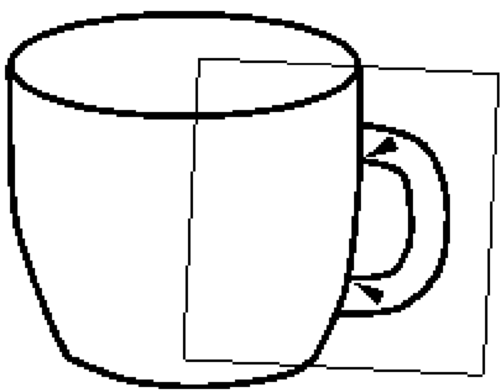
neurons in PF cortex signal both "`what`" and "`where`"

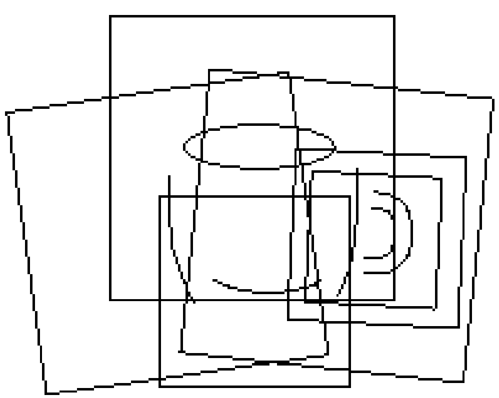


a computer vision perspective

successful systems use
`"what + where"` cues

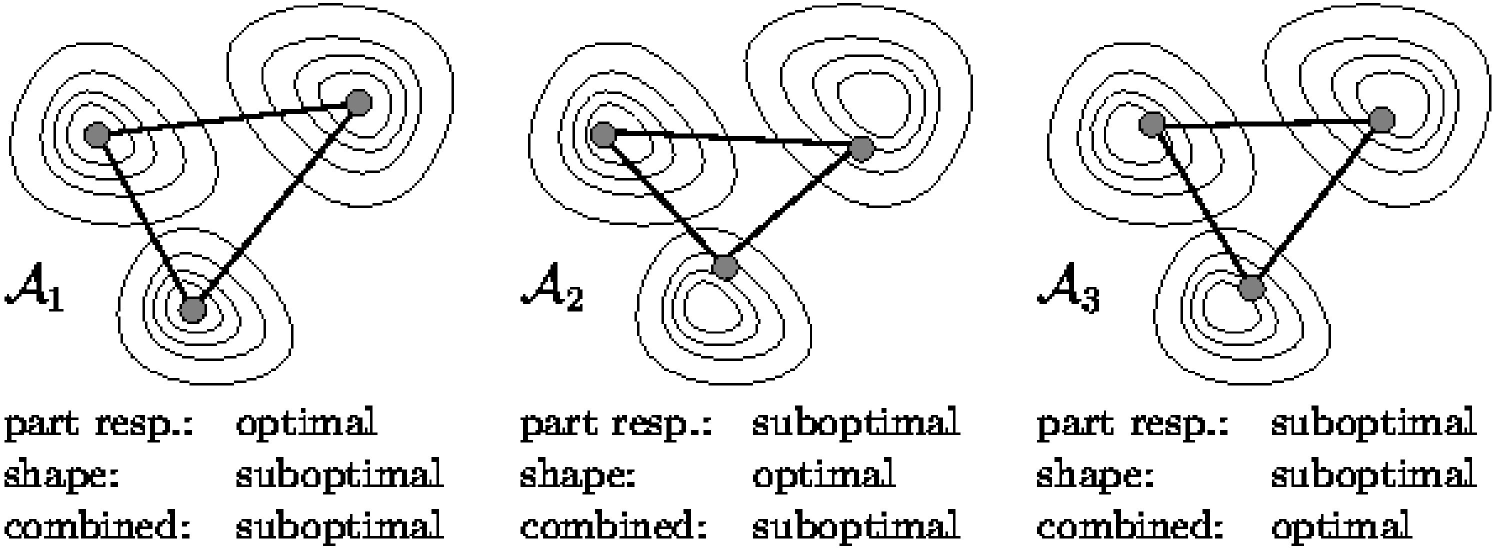
R. C. Nelson and A. Selinger,
Large−Scale Tests of a Keyed, Appearance−Based 3−D Object Recognition System,
Vision Research 38:2469−2488 (1998)





a computer vision perspective

successful systems use
"what + where" cues



M. C. Burl, M. Weber, and P. Perona,
A probabilistic approach to object recognition using local photometry and global geometry,
Proc. ECCV'98, 628–641 (1998)

Chorus of Fragments

???

a **lion**'s body with a **goat**'s head on the back, and a **snake**'s head at the end of the tail

– **Platonic**, categorical coding of shape

– **abstract**, categorical coding of structure

– **empirical** basis for the coding of shape

– **concrete** coding of structure

representation by similarities to spatially anchored shape fragments

extension
to **language**:
perceptual symbol systems

computational
implementation

exploring **implications**
for psychology,
neurobiology

a comprehensive theory
of shape/scene representation

binding:
the pegboard model

veridicality of
representation

mathematics of **compositionality**
and shape spaces

J. Cusimano