CPSC 532E — Week 8: Lecture

Shape Perception

0. Reference Frames

Intuition: Shape is an <u>objective</u> property of an object
(Just like colour)

<u>Objective Shape</u>: Set of spatial relations (distances) between points on the object

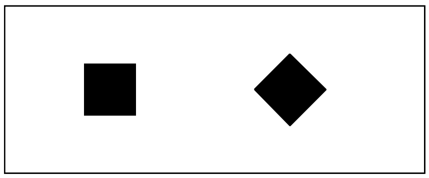
These are invariant under:
- 3D translation
- 3D rotation
- 3D reflection

Relative distances are invariant under:
- 3D scaling (dilation, expansion, shrinking)

But just like colour (or any other surface property), shape is also a subjective property

This isn't always the same as objective shape.

Eg.



Why do objectively similar shapes sometimes seem different?

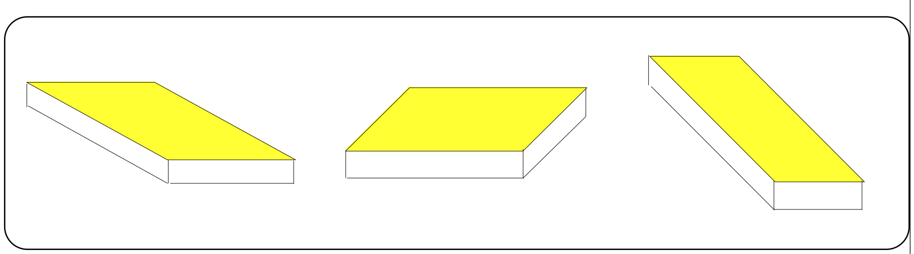
Why do objectively different shapes sometimes seem the same?

Answer (Part I): Shape Constancy

Just like colour, visual system attempts to compensate to undo effects of viewing angle, 3D, etc

- can suffer from similar illusions

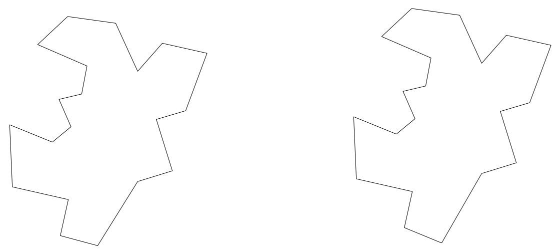
Question: Which shape is most like the center one?



Shepard Illusion

Answer (Part II): Representation Systems

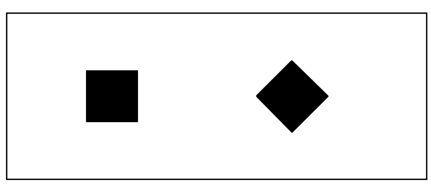
Why do objectively different shapes sometimes seem the same?



-> Different shapes given the same representation

- certain aspects ignored
- representations don't contain differences

Why do objectively similar shapes sometimes seem different?



-> Same shapes given different representation

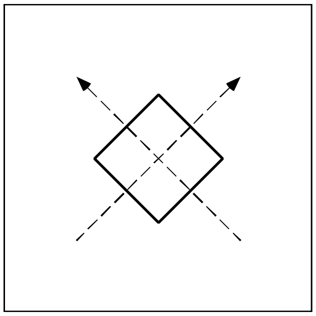
- certain aspects emphasized
- each representation reflects different aspects

One source of difference representations: <u>reference frame</u>

0. Reference Frames

1. Object-centered

- center point defined in terms of object
- orientation defined in terms of object
- [size defined in terms of object]

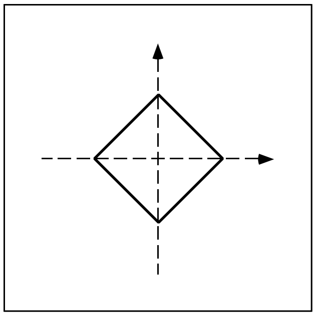


Insensitive to changes in

- location of object, viewer
- orientation of object, viewer
- [size of object]

2. Viewer-centered (egocentric)

- center point defined in terms of viewer (eye)
- orientation defined in terms of viewer (eye)
- [size defined in terms of viewer]

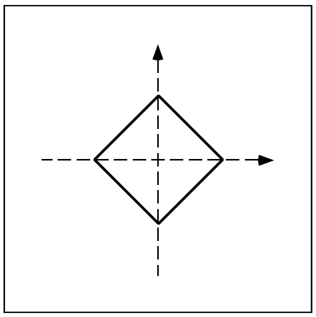


Sensitive to changes in

- location of object, viewer
- orientation of object, viewer
- [size of object]

3. Environment-centered (exocentric)

- center point defined in terms of environment (landmark)
- orientation defined in terms of environment (gravity)
- [size defined in terms of environment]



Sensitive to changes in

- location of object (not viewer)
- orientation of object (not viewer)
- [size of object]

Each type of reference frame has its advantages

Object-centered:

- facilitates recognition of object (shape) over many different kinds of transformation

Viewer-centered:

- facilitates interaction with object (that you're already holding)

Environment-centered:

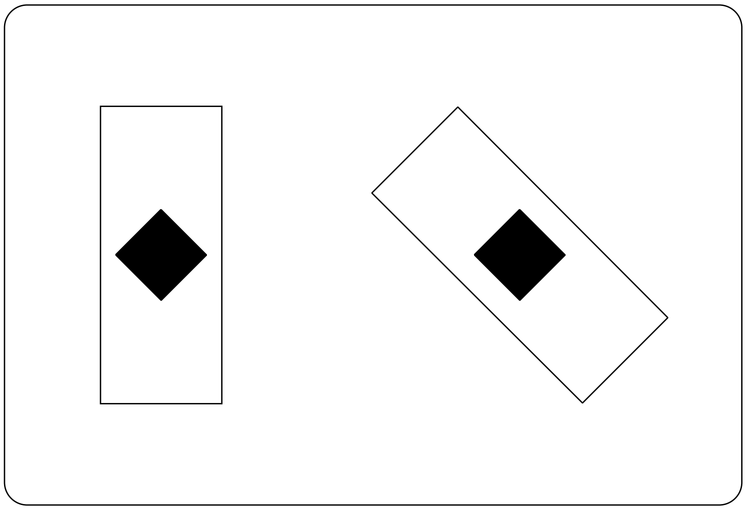
- facilitates interaction with object (that you're haven't yet contacted)

-> Relevant representation depends on task

- e.g. button could be pushed, or display information

-> Relevant representation also depends on context

- e.g.



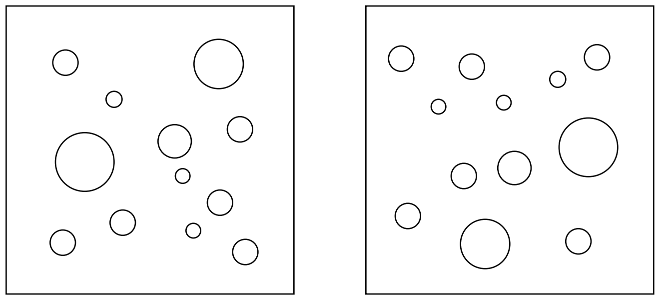
1. Invariant Features

Intuition: Describe object in terms of **features that don't change under various transformations**

E.g - relative measures
- relative distance (d1 - d2)
- relative orientation (o1 - o2)
- relative size (s1/s2)
  - aspect ratio (= length/width)

Also possible are:
- topological properties (e.g. number, closedness)
- statistical properties? (e.g. distribution of sizes)

Statistical properties of a group (Ariely, 2001)



500 ms 500 ms

Observers sensitive to mean + range of sizes

-> Also true for objects?

2. Shape Templates

Intuition: Describe object via **combinations of 2D features**
- "standard pieces" that are familiar
- describe parts of the entire shape

E.g - simple fixed geometric shapes
- disks, squares, lines
- letters, digits

- Comparison of object with template gives estimate of "fit"
  - same principle as edge detectors
  - more complex arrangement
    - detectors for squares, disks, rectangles
    - cf. "bug detectors"

Problem:

How to cope with different locations, orientations, sizes?

Answer:

- use a set of templates of different orientations, sizes
- response via the set of activities

(But this does require a lot of neural real estate)

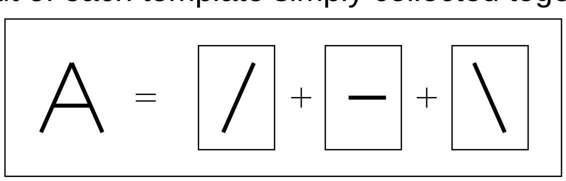
A More Serious Problem:

How to cope with many different kinds of shapes?

-***only one shape can be handled by each template***
-***how to combine outputs? - need structure?***

Possible strategy: **Feature lists**

Output of each template simply collected together



Proto-objects represented in this way?
(supports rapid perception of shape)

How to Determine Relevant Features?

Objects with different features can be compared by **multidimensional scaling (MDS)** (Shepard)

- **Input**: estimates of distances (similarities) between pairs of objects
- **Output**: a low-dimensional space that combines all relevant dimensions of the objects

Useful way to determine a set of shapes (or objects) that are maximally distinct from each other

Even More Serious Problems:

How to represent 3D structures?

How to cope with 3D rotations?

3. Structural Descriptions

Intuition: Describe shape via (3D) **structural relations** among a set of shape primitives

E.g - geon theory
- well-defined set of shape primitives (geons)
- well-defined set of structural relations
  (e.g. SIDE-CONNECTED, SMALLER-THAN)

Able to describe 3D shape of an object

But—does it describe what we see?
(cf. limits of attention)