

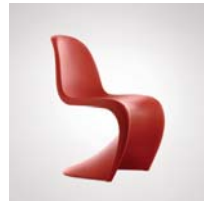
High-Level Vision

Object Recognition II



CS332 Visual Processing
Department of Computer Science
Wellesley College

What is a chair?



1-2

Approaches to recognition...

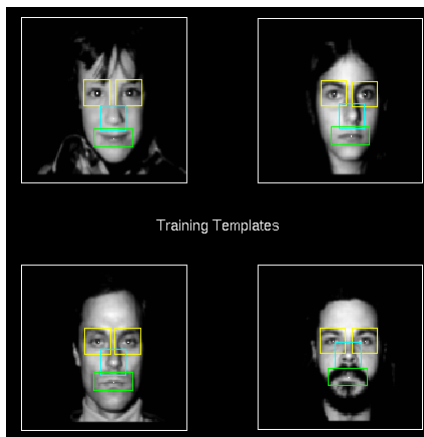
... differ in how *regularities* are used to constrain the interpretation of the viewed object

Three main approaches:

- invariant properties
- parts decomposition
- alignment

1-3

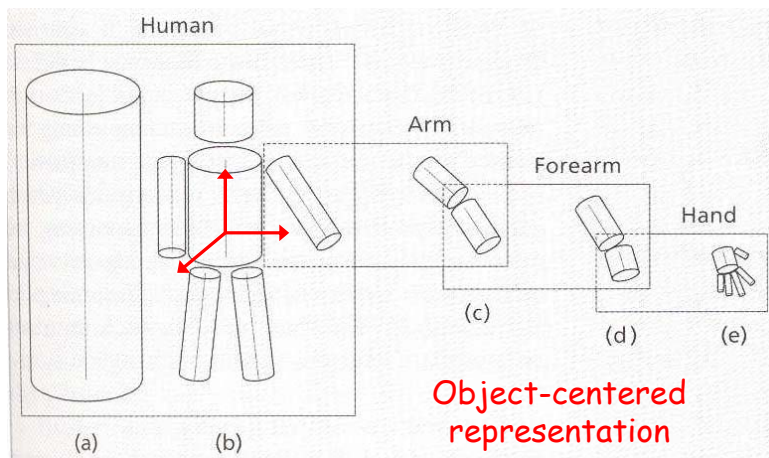
Face recognition by parts decomposition



MIT Media Lab Vision & Modeling Group

1-4

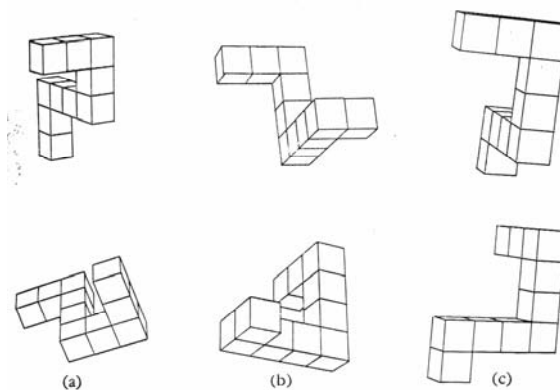
Feature hierarchies



Marr & Nishihara

1-5

Mental rotation

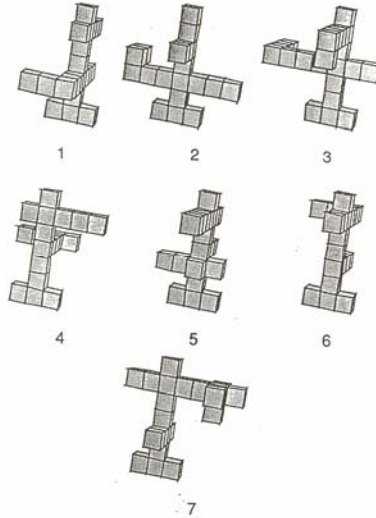


Time needed to determine whether pair of objects are *the same* is proportional to angle of rotation between pair

1-6

Viewer-centered object representation?

Tarr, '95: After learning to recognize a set of 3-D objects from a small set of specific 2-D views of these objects, the time needed to recognize a novel view is proportional to the 3-D angle between the new view and closest learned view



1-7

The debate continues...

Viewpoint invariant
object representations



Viewpoint dependent
object representations



Biederman

Object-centered!

Viewer-centered!



Bülhoff

1-8

Alignment methods

Find an object model and geometric transformation that *best match* the viewed image

- V viewed object (image)
 M_i object models
 T_{ij} allowable transformations between viewed object and models
 F measure of fit between V and the expected appearance of model M_i under the transformation T_{ij}

GOAL: Find a combination of M_i and T_{ij} that maximizes the fit F

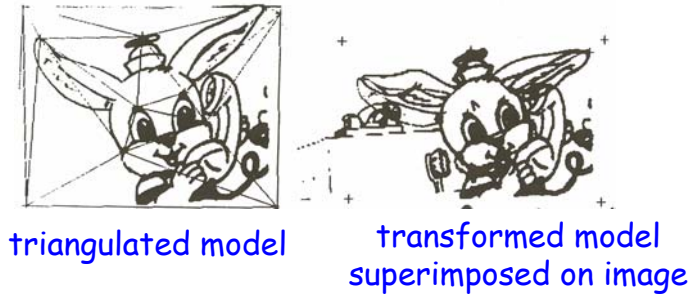
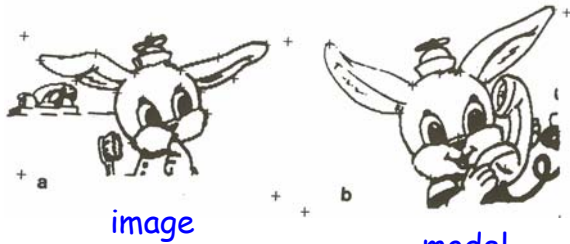
1-9

Alignment method: recognition process

- (1) Find best transformation T_{ij} for each model M_i (optimizing over possible views)
- (2) Find M_i whose best T_{ij} gives the best match to image V

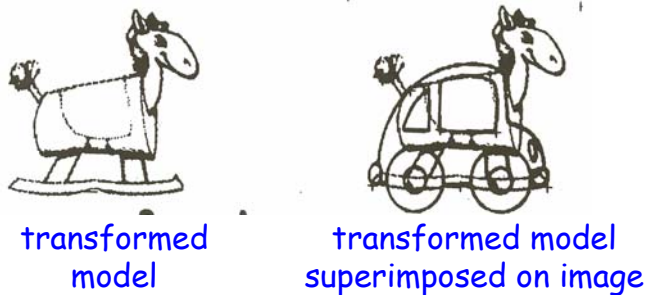
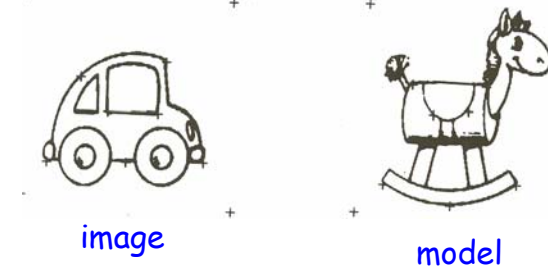
1-10

Aligning pictorial models



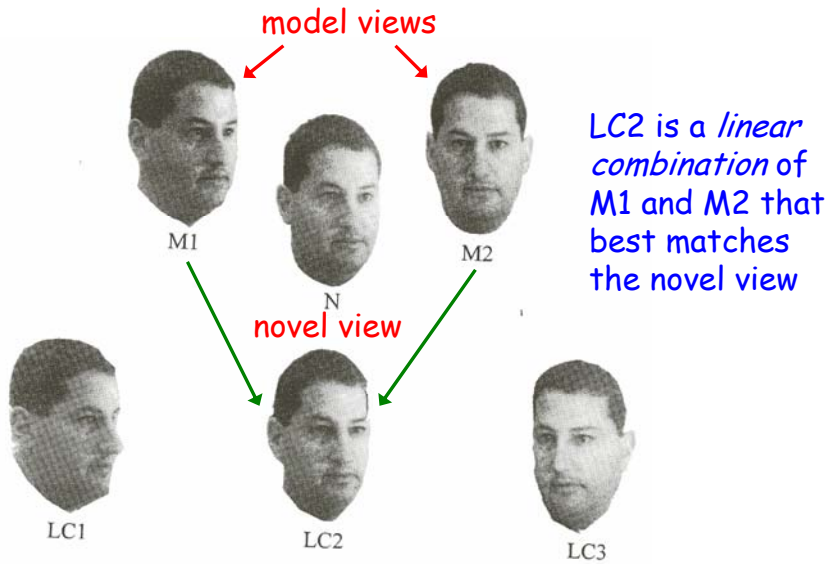
1-11

When the model doesn't fit...



1-12

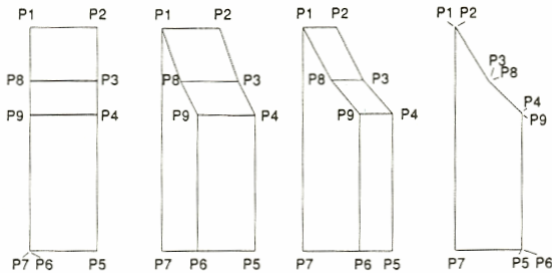
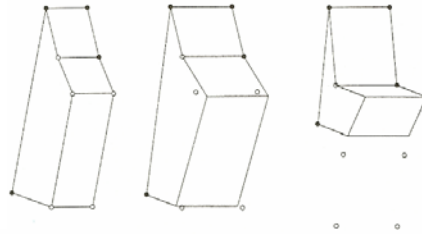
Recognition by linear combination of views



1-13

Obelisk, jukebox or seat?

Each object model consists of multiple 2-D views

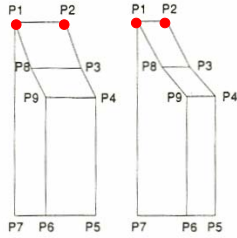


obelisk model

Goal: recognize novel views of these objects

1-14

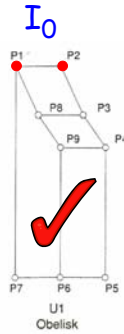
Predicting object appearance



I_1

I_2

two known views
of obelisk



Recognition process:

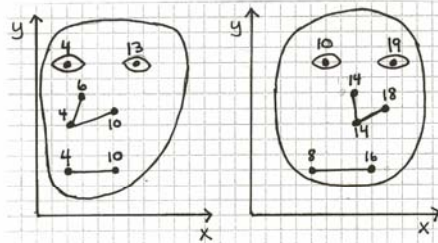
- (1) compute α, β that predict P1 and P2
- (2) use α, β to predict other points
- (3) evaluate fit of model to image

$$X_{P_1 I_0} = \alpha X_{P_1 I_1} + \beta X_{P_1 I_2}$$

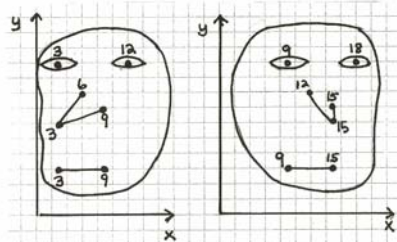
$$X_{P_2 I_0} = \alpha X_{P_2 I_1} + \beta X_{P_2 I_2}$$

1-15

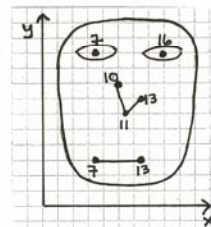
Face recognition by linear combination of views



Harry model



Henry model



Who am I?

1-16