Bottom-up and Top-down processes in shape perception
Processing Framework Proposed by Marr

3D structure; motion characteristics; surface properties

- Shape From stereo
- Motion flow
- Shape From motion
- Color estimation
- Shape From contour
- Shape From shading
- Shape From texture

Edge extraction

Image
I'm sorry, Mr. Mitchell is only painted on the wall.
The perceptual task

3D shape recovery from a single 2D image
A constraint based system
An implemented constraint-based system

- Attneave & Frost, 1969
- Mackworth, 1973
- Perkins & Cooper, 1980
- Kanade, 1980
- Barrow & Tenenbaum, 1981
- Marill, 1989
- Fischler & Leclerc, 1992

Experimentation

Occam’s razor and complexity

Sinha & Adelson

A 3D shape recovery system with 3 general constraints
What are the three constraints?

1. Regularity

2. Planarity

3. Compactness
The system in action...

2D Inputs

Recovered 3D shapes
The system in action...
Where the system works...
And where it runs into trouble...

Henri Matisse
Is there a more general approach to shape recovery?
Learning based hypothesis

But, can learning really effect such a basic visual task?
Does learning influence perception?

Computational and psychophysical investigations
Does learning influence perception?
Survey says...

Empiricists vs nativists
Does learning influence perception?
Survey says...

1800s
- Yes: 50%
- No: 50%

2000s
- Yes: 100%
- No: 0%
Does learning influence perception?

Evidence - I

Early experience and neural development

Wiesel & Hubel, J. Neurophys., 1963
Hubel & Wiesel, J. Neurophys., 1965
Olson & Freeman, Exp. Brain Res., 1980

...
Does learning influence perception?

Evidence - II

Poggio, Fahle & Edelman, Science, 1992
Karni & Sagi, Nature, 1993

...
Does learning influence perception? Yes.
Common characteristics of perceptual learning phenomena

Perceptual learning is specific to:
- trained eye
- retinal location
- stimulus size and orientation

Change in a simple stimulus dimension
Implication

High-level ‘Object’ areas (jobs: recognition)

Low-level ‘Perception’ areas (jobs: 3D recovery, edge extraction)

Locus of learning
Open Question...

<table>
<thead>
<tr>
<th>Low-level perception</th>
<th>High-level perception</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low-level areas</td>
</tr>
<tr>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>High-level areas</td>
<td></td>
</tr>
<tr>
<td>?</td>
<td></td>
</tr>
</tbody>
</table>
Open question

Can learning in ‘higher’ object-specific areas influence early perception?
Does high-level learning influence early perception?

Why is this question interesting?

Marr’s hierarchy
Does high-level learning influence early perception?

Why is this question interesting?
Does high-level learning influence early perception?

The perceptual task:
Why this task?

1. Traditionally considered one of the most fundamental early/mid-level perceptual tasks.

2. Past work is overwhelmingly in favor of constraint based explanations devoid of high-level learning.
Road map

1. A constraint-based model devoid of high-level learning

“Here be dragons”

- Natural objects
- Arbitrary polyhedra
- Semi-regular polyhedra
- Polyhedra with rectangular facets
- Circular discs
Road map

1. A constraint based model devoid of high-level learning

2. Experimental evidence for high-level learning

3. A model for incorporating high-level learning in early perception
Experimental goal
Experimental Paradigm
Results
Results

(a) Population 1

(b) Population 2
Results
Illusions of non-rigidity
Illusions of non-rigidity
Illusions of non-rigidity
Results (contd.)

Learning is object-specific...

Percentage of trials over which subjects reported non-rigidity

Training set

X-Y positional noise added

Scale factor

0% 25% 50% 75% 100%

0% 25% 50% 75% 100%

1x 2x 3x
Results (contd.)

Learning is long lasting...

- No delay: 6 hours
- 24 hours

Percentage of trials over which subjects reported non-rigidity:

- Training set: 50%
- No delay: 40%
- 6 hours: 30%
- 24 hours: 20%
Results (contd.)

Learned shape may be represented implicitly...

Objects may be represented as collections of temporally associated views
Inferences

1. Arbitrary associations between 2D and 3D structures can be learned.

2. Learning is object-specific.

3. Recognition may, in some circumstances, precede 3D shape perception.
Road map

1. A constraint based model devoid of high-level learning

2. Experimental evidence for high-level learning

3. A model for incorporating high-level learning in early perception
Goal

Given an image, to estimate some attribute (e.g. 3D shape, edge-map, tactile feel etc.)
The model schematically

Attribute estimation via prototype combination

Learned knowledge about object-class 1
- Image1+attribute
- Image2+attribute
- ...
- Imagen+attribute

Learned knowledge about object-class n

Recognition (somehow…)

Result

Image
Combining learned instances

Step 1:
\[
\text{image}_{\text{new}} = \frac{C_0 \cdot \text{Image}_0}{+} \frac{C_1 \cdot \text{Image}_1}{+} \cdots \frac{C_n \cdot \text{Image}_n}{+}
\]

Step 2:
\[
\frac{C_0 \cdot \text{Attribute}_0}{+} \frac{C_1 \cdot \text{Attribute}_1}{+} \cdots \frac{C_n \cdot \text{Attribute}_n}{+} = \text{Attribute}_{\text{new}}
\]
The model in action:
Example 1 - implicit 3D shape recovery

<table>
<thead>
<tr>
<th>Image</th>
<th>45 degree face image</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attribute</td>
<td>Frontal face image</td>
</tr>
</tbody>
</table>
The model in action:
Example 1 - implicit 3D shape recovery

Image: 45 degree face image
Attribute: Frontal face image
The model in action:
Example 1 - implicit 3D shape recovery

<table>
<thead>
<tr>
<th>Image</th>
<th>Attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>45 degree face image</td>
<td>Frontal face image</td>
</tr>
</tbody>
</table>

(a) input  (b) computed  (c) real
The model in action:

Example 2 - Clean edge-map extraction

**Image**: frontal face image

**Attribute**: Clean edge-map

Input
The model in action:
Example 2 - Clean edge-map extraction

Input

Model output

Image

frontal face image

Attribute

Clean edge-map
The model in action:

Example 2 - Clean edge-map extraction

Image: frontal face image
Attribute: Clean edge-map

Input | Model output | Bottom-up output
Example 2 (contd.)

Strengths of high-level learning based model:
1. Ability to complete missing information
2. Ability to handle noise
Processing Framework Proposed by Marr

3D structure; motion characteristics; surface properties

Shape From stereo
Motion flow
Shape From motion
Color estimation
Shape From contour
Shape From shading
Shape From texture

Edge extraction

Image

Next!