Visual Object Recognition

Neurobiology 230 – Harvard / GSAS 78454

Gabriel Kreiman

**Email:** gabriel.kreiman@tch.harvard.edu

**Phone:** 617-919-2530

**Web site:** [http://tinyurl.com/vision-class](http://tinyurl.com/vision-class)

**Dates:** Mondays

**Time:** 3:30 – 5:30 PM

**Location:** Biolabs 1075
Gestalt laws — Basic phenomenological constraints

- **Law of Closure** — The mind may experience elements it does not perceive through sensation, in order to complete a regular figure (that is, to increase regularity).
- **Law of Similarity** — The mind groups similar elements into collective entities or totalities. This similarity might depend on relationships of form, color, size, or brightness.
- **Law of Proximity** — Spatial or temporal proximity of elements may induce the mind to perceive a collective or totality.
- **Law of Symmetry (Figure ground relationships)** — Symmetrical images are perceived collectively, even in spite of distance.
- **Law of Continuity** — The mind continues visual, auditory, and kinetic patterns.
- **Law of Common Fate** — Elements with the same moving direction are perceived as a collective or unit.
Law of closure – The Kanizsa triangle
Law of similarity
Law of proximity
Law of continuity
Law of common fate
Size tolerance
Position tolerance

bd    db    bd    bd    bd    db    db    bd    db    db    db    bd
Tolerance to viewpoint and illumination changes
Tolerance to viewpoint and illumination changes

(a) [Graph showing tolerance to viewpoint and illumination changes with various angles and responses]

(b) [Graph showing tolerance to viewpoint and illumination changes with various angles and responses]

Wire 526, cell 202

142 spikes sec⁻¹
One-shot learning for scale tolerance
Other transformations require example-based training
Beyond pixels – Context matters
Visual recognition depends on experience
Recognition from minimal features
Recognition of caricatures
Visual recognition can be extremely fast
Psychophysics: The study of the dependencies of psychological experiences upon the physical stimuli that generate them

If we are careful, we can learn about how the brain works by measuring it indirectly, using the behavioral responses it generates.

Basic measures:
• Reaction time — The time taken by subjects to perform a task or make a judgment can give an indication (or at least an upper bound) of how long the necessary psychological (and hence neural) processing takes.
• Performance — Often inversely related to reaction time. There are techniques for mitigating response biases.
• Threshold — Stimuli can be varied to determine the threshold for detection, discrimination, or some more complex psychological phenomenon.

Depending on experimental goals, one or more can be probed.
Visual recognition can be extremely fast.

Fig. 1. Choice saccade task. After a pseudo-random fixation period, a blank screen (gap period) for 200 ms preceded the simultaneous presentation of two natural scenes in the left and right hemifields (20 ms). The images were followed by two grey fixation crosses indicating the saccade landing positions.


<table>
<thead>
<tr>
<th>Subject</th>
<th>N</th>
<th>Accuracy (%)</th>
<th>Median RT (ms)</th>
<th>Min RT (ms)</th>
<th>Mean start (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>682</td>
<td>96.3</td>
<td>227</td>
<td>130</td>
<td>143</td>
</tr>
<tr>
<td>2</td>
<td>774</td>
<td>93.3</td>
<td>200</td>
<td>130</td>
<td>136</td>
</tr>
<tr>
<td>3</td>
<td>726</td>
<td>81.8</td>
<td>201</td>
<td>130</td>
<td>129</td>
</tr>
<tr>
<td>4</td>
<td>563</td>
<td>80.1</td>
<td>191</td>
<td>120</td>
<td>126</td>
</tr>
<tr>
<td>5</td>
<td>672</td>
<td>86.6</td>
<td>159</td>
<td>130</td>
<td>133</td>
</tr>
<tr>
<td>6</td>
<td>675</td>
<td>86.1</td>
<td>224</td>
<td>150</td>
<td>143</td>
</tr>
<tr>
<td>7</td>
<td>574</td>
<td>90.2</td>
<td>204</td>
<td>140</td>
<td>129</td>
</tr>
<tr>
<td>8</td>
<td>653</td>
<td>94.0</td>
<td>213</td>
<td>150</td>
<td>147</td>
</tr>
<tr>
<td>9</td>
<td>694</td>
<td>96.7</td>
<td>251</td>
<td>180</td>
<td>200</td>
</tr>
<tr>
<td>10</td>
<td>534</td>
<td>89.7</td>
<td>236</td>
<td>180</td>
<td>124</td>
</tr>
<tr>
<td>11</td>
<td>739</td>
<td>90.0</td>
<td>253</td>
<td>190</td>
<td>205</td>
</tr>
<tr>
<td>12</td>
<td>652</td>
<td>96.6</td>
<td>276</td>
<td>200</td>
<td>235</td>
</tr>
<tr>
<td>13</td>
<td>703</td>
<td>95.0</td>
<td>238</td>
<td>160</td>
<td>173</td>
</tr>
<tr>
<td>14</td>
<td>769</td>
<td>98.7</td>
<td>301</td>
<td>230</td>
<td>251</td>
</tr>
<tr>
<td>15</td>
<td>529</td>
<td>77.1</td>
<td>233</td>
<td>160</td>
<td>235</td>
</tr>
<tr>
<td>All</td>
<td>8998</td>
<td>90.1</td>
<td>228</td>
<td>120</td>
<td>140</td>
</tr>
</tbody>
</table>

The second column of this table indicates the total number of trials per participant (see Section 2 for details). Columns 3-5 give the mean accuracy, median and minimum reaction time values for each participant shown in Figs. 3B and C. The last column indicates the onset latency of the mean eye trace for each participant (see Fig. 5).
The visual system has a very large capacity
A massive recollection capacity

Gestalt laws apply to object recognition

Is information integrated over time?
Is information integrated over time?

Brief asynchronies disrupt object recognition, but some integration persists even beyond 100 ms.
Further reading


Original articles cited in class (see lecture notes for complete list)