

# IST605: Human Information Processing — Perception Summary

---

## Document Outline

- The Problem of Perception
  - Visual Perception (Low-Level and High-Level Vision)
  - Feature Detection Theories
  - Structural Theories
  - Template Matching and Alignment
  - Levels and Integration of Information
  - Perceptual Context Effects
  - The Word Superiority Effect
- 

## 1. Sensation and Perception

### Sensation

- **Definition:** The process of receiving, converting, and transmitting information from the outside world.
- Sensory organs contain **receptors** that **transduce** sensory energy into nerve impulses carried to the brain.
- **Three types of processing:**
  - **Transduction:** Converting sensory stimuli into neural impulses.
  - **Sensory reduction:** Filtering and analyzing incoming sensations before sending to the brain.
  - **Coding:** Converting particular sensory input into a specific sensation sent to differing parts of the brain.

### Sensory Thresholds (Psychophysics)

- **Psychophysics:** Subfield studying physical stimuli and their interaction with sensory systems.
- **Absolute threshold:** Smallest amount of a stimulus we can detect.
- **Difference threshold (JND):** Minimal difference needed to detect a stimulus change.

### Perceptual Processes

- **Selection:** Choosing which of many stimuli will be processed.
  - **Organization:** Collecting information into some pattern.
  - **Interpretation:** Understanding the pattern.
  - Perceptions can be in error (e.g., illusions—visual stimuli that are misinterpreted).
- 

## 2. The Problem of Perception

- We perceive the world through eyes, ears, touch, etc., and end up with a coherent experience.
- **Perception is very hard for computers** because of **ambiguities**:

1. **Same objects, different views:** Objects seen from different vantage points and under different lighting produce different retinal images.
  2. **2-D vs 3-D:** Images on the retina are 2-D (from light bouncing off 3-D objects). An unlimited number of 3-D arrangements can produce the same 2-D retinal image.
- The perceptual system resolves these ambiguities before we consciously experience the visual world.
- 

### 3. Signal Detection Theory

#### Purpose

- Measures the ability to differentiate between **information-bearing patterns (signal)** and **noise** (random, distracting patterns).
- Explains how we make decisions about stimuli in **uncertain** situations.
- Goal: Tease out the signal from background noise.

#### Key Concepts

- **Signal:** The stimulus being detected (e.g., sound, image).
- **Noise:** Distracting information or random background interference.
- **Sensitivity ( $d'$ ):** Ability to discriminate signal from noise; independent of the decision criterion.
- **Criterion ( $\beta$ ):** The point at which we decide the signal is present. Can be biased (cautious vs liberal).

#### Outcomes

- **Hit:** Signal present, responded "yes."
- **Miss:** Signal present, responded "no."
- **False alarm:** Signal absent, responded "yes."
- **Correct rejection:** Signal absent, responded "no."

#### Applications

- Sensory perception (e.g., faint sound in a noisy room).
  - Medical imaging (distinguishing tumour from tissue).
  - Telecommunications (detecting signals in noise).
  - UX research (how often users notice important information).
- 

### 4. Visual Perception

- **An active process** in which different levels of analysis interact.
  - **Low-level vision:** Extracting preliminary information from the pattern of light on the retina.
  - **High-level vision:** Perception of larger-scale elements—meaning and identity of objects and scenes.
- 

### 5. Low-Level Vision

#### Goals

- Form internal representations of the 3-D visual world.

- Extract features, determine location, and track motion to help people interact with the environment.

## Edge Detection

- **Object boundary:** Coarse edges; **parts of object:** Fine details.
- **Low spatial frequency:** Gradual changes, coarse details, overall shape and structure.
- **High spatial frequency:** Abrupt changes, fine details, sharp edges and textures.
- Processing may begin with **global (coarse)** information and gradually incorporate **local (fine)** information.
- Edge detection uses **differences in brightness**; the visual system enhances dark/light contrasts (e.g., illusory squares).

## Localization: Segregation

- **Segregation:** Organizing objects—separating visual information for each object from others.
- **Gestalt Laws of Organization:** Principles for perceiving visual elements as unified wholes.
  - **Similarity:** Grouping like items.
  - **Proximity:** Grouping nearby items.
  - **Closure:** Filling in gaps to see complete shapes.
  - **Continuity:** Perceiving smooth, continuous lines.
  - **Figure-ground:** Separating foreground from background.
  - **Prägnanz (Simplicity):** Perceiving the simplest possible arrangement.
  - **Connectedness:** Connected elements (by colour, lines, etc.) are perceived as grouped.

## Figure-Ground Cues

- **Blurriness:** Foreground crisp, background blurry.
- **Contrast:** High contrast leads to figure-ground perception (e.g., Rubin vase).
- **Size:** Larger images perceived as closer (figure); smaller as background.
- **Separation:** Isolated objects more likely seen as figure.

## Localization: Distance (Depth Cues)

### Monocular cues (one eye):

- **Interposition (occlusion):** One object interrupting another is seen as closer.
- **Linear perspective:** Converging parallel lines imply distance.
- **Relative size:** Similar-sized objects—smaller retinal image implies farther away.
- **Shadows:** Position and shape from cast shadows.
- **Size constancy:** Object perceived as constant size despite changing retinal size with distance.
- **Shape constancy:** Familiar object's shape perceived as constant from different angles.
- **Texture gradients:** Distant objects have smoother texture.
- **Relative clarity:** Distant objects less clear (outdoors).
- **Motion parallax:** As head moves, nearby objects move faster and in opposite direction; distant objects move slower and in same direction.

### Binocular cues (both eyes):

- **Stereopsis:** Interpreting information from both visual fields.

- **Retinal disparity:** Slightly different images from left and right retina; brain merges them into 3-D. More effective when objects are close; disparity decreases with distance.
- 

## 6. High-Level Vision

### Focus

- Meaning and identity of objects and scenes.
- Object recognition, scene understanding, property estimation, integration with cognition.

### How We Identify Objects

- **Feature detection theories**
  - **Structural theories**
  - **Template matching and alignment**
- 

## 7. Feature Detection Theories

### Core Idea

- Objects are composed of **separable features**.
- Recognition involves decomposing a stimulus into features and matching against stored representations.
- A small set of features can describe many objects (like letters forming words).
- **Visual search:** Locating a target among distractors; a departure from the standard (e.g., tilted vs vertical line) is detected as an extra feature.

### Limitation

- Ignores **spatial relationships** among features.
  - Example: A horizontal and vertical line describe both T and +; feature list alone cannot distinguish them.
- 

## 8. Structural Theories

### Core Idea

- Include **features** and **spatial relations** (how parts are connected).
- Example: T = t-junction(Line1, Line1); + = cross(Line1, Line1, 90°).
- Enables recognition of many objects with few features.

### Biederman's Recognition-by-Components (RBC)

- Objects recognized by breaking them into **geons** (geometric icons)—simple 3-D shapes (cylinders, blocks, cones, wedges; ~36 geons).
- **Viewpoint-invariance:** Geon properties (e.g., straight vs curved edges, parallelism, symmetry) are view-invariant; objects recognized from many viewpoints.

- **Structural descriptions:** Built from geons and spatial relations (e.g., "cylinder and handle" for a mug).

**Strengths:** Limited elements, implemented in computer models.

**Limitations:** May not distinguish objects with similar parts (e.g., wolf vs cat); some objects (e.g., loaf of bread) are hard to decompose into geons.

---

## 9. Template Matching and Alignment

### Core Idea

- Representations are **2-D arrays** (pixels).
- A **template** is a copy of the image; matching involves comparing current image to stored templates.
- Problem: Changing size or rotation breaks matching; humans do not have this trouble.

### Two-Stage Solution

1. **Alignment:** Use transformations to align the object with stored models/templates.
  2. **Matching:** Search for the best match.
- 

## 10. Levels and Integration of Information

### Bottom-Up vs Top-Down

- **Bottom-up:** Processing from sensory receptors upward; relies on raw sensory data.
- **Top-down:** Uses prior knowledge, expectations, and context to interpret sensory information.

### Perceptual Context Effects

- Context influences perception.
  - Same middle character read as "13" (top-to-bottom) or "B" (left-to-right).
  - Same ambiguous character read as "H" in one word and "A" in another.
  - Partial letter (e.g., partly blotted E) still perceived when context supports it.
- 

## 11. The Word Superiority Effect

- **Definition:** Letters are recognized **more easily in words** than in isolation or in non-word strings.
- Top-down processing of the word aids letter identification.
- Meaningfulness helps identify component letters.
- Pronounceability: Non-words (e.g., "TVXC") are harder than real words (e.g., "TREE").

### Sentence-Level Effect

- Background knowledge guides reading and prediction of the next word.
  - Unfamiliar or random text is harder to predict and read more slowly.
  - Unexpected words (e.g., "I would like to drink a microwave") slow processing and cause re-checking.
- 

## Quick Reference

Topic	Key idea
Sensation	Transduction, sensory reduction, coding
Absolute threshold	Smallest detectable stimulus
JND	Minimal difference to detect change
Signal detection	Signal vs noise; hits, misses, false alarms, correct rejections; sensitivity vs criterion
Low spatial frequency	Coarse, global, gradual
High spatial frequency	Fine, local, abrupt
Gestalt laws	Similarity, proximity, closure, continuity, figure-ground, simplicity, connectedness
Monocular depth cues	Interposition, linear perspective, relative size, shadow, texture, motion parallax
Binocular depth cues	Stereopsis, retinal disparity
Feature theory	Objects as separable features; ignores spatial relations
Structural theory	Features + spatial relations; geons in RBC
Template matching	Pixel arrays; alignment + matching
Bottom-up	Data-driven
Top-down	Knowledge- and context-driven
Word superiority	Letters easier in words than in isolation/non-words