ECE 418: Introduction to Image & Video Processing

Human Visual Perception

Shuai Huang

02/12/2019

University of Illinois at Urbana-Champaign
Overview

1. Introduction

2. Elements of visual perception
Introduction
Human visual perception

Do we perceive the world in the same way?

Figure 1: Is this dress “Black and Blue” or “White and Gold”?
Human visual perception

How do we study the Human Visual System (HVS)?

○ Psychometrics: Study what humans see based on collective statistics from many subjects.

○ Neurophysiology: Physical tests on individual HVS components, such as retina, visual model, etc.
Human visual perception

Human visual processing by different cognitive levels:

- Low-level vision: light receptors, formation of neural response, etc.

Figure 2: Human eye anatomy [1].
Human visual perception

Human visual processing by different cognitive levels:

- High-level vision: process of neural response in visual cortex, i.e. visual perception (pattern recognition and feature extraction).

Figure 3: Motor and sensory regions of the cerebral cortex [1].
Our understanding of HVS is very limited:

- Some attributes of the low-level vision are well understood.
- Research on High-level vision is just starting to take off.
  - Object recognition/classification via deep learning.
  - Neurology.
Human visual perception

Research on High-level vision is just starting to take off.

- Deep image reconstruction from human brain activity [2].

Figure 4: Deep image reconstruction framework [2].
Research on High-level vision is just starting to take off.

○ Deep image reconstruction from human brain activity [2].

Figure 5: Deep image reconstruction on real images [2].
Human visual perception

Research on High-level vision is just starting to take off.

- Deep image reconstruction from human brain activity [2].

Figure 6: Deep image reconstruction on symbols and letters [2].
Elements of visual perception
Image formation

The imaging system produces an image on the retina. Any point $P$ on the retina is specified by

- The angle $\theta$ between the line of sight and the visual axis.
- The angle $\phi$ between $FP$ and the vertical axis passing through the focal point $F$. 
Elements of visual perception

We are interested in studying the retina.

Figure 7: Different parts of the human eye [3].
Elements of visual perception

We are interested in studying the retina.

Structure of the Retina

Figure 8: Schematic diagram of the retina [4]
Elements of visual perception

- The light receptors (cones and rods) contain photosensitive pigments that can absorb light and initiate neural response.
- Horizontal and amacrine cells reduce the signals from adjacent cells: lateral inhibition
- Bipolar cells transmit signals.
- Ganglion cells fire electric pulses.
- Optical nerves transmit electric pulses to visual cortex.
There are two kinds of photoreceptors: 1) Cones

- 6~7 million cones in each eye.
- Concentrated in **fovea**: narrow, central portion of retina.
- Responsible for color vision.
- For medium-to-bright (daytime, photopic) vision

![Retina image: fovea is at the center](image-url)
Rod photoreceptor

There are two kinds of photoreceptors: 2) **Rods**

- 75~100 million rods in each eye, much more than cones.
- All over retina except fovea.
- Not involved in color vision.
- For dim-light (night, scotopic) vision

![Image of rod cells](image-url)

**Figure 10:** Cone and Rod cells [6].
Figure 11: Distribution of cones and rods in the retina [7].
What caused the blind spot?

Make sure your eyes are approximately at the same level with the cross and dot.

- Cover **Left** eye and stare at the cross with your **Right** eye.
- Place your eye at a distance from the screen approximately 3 times the distance between the cross and the dot.
- Move your eye towards or away from the screen until you notice the dot disappear.
What caused the blind spot?

All vertebrates have this blind spot. Cephalopod eyes do not.

Figure 12: Nerve fibers route before the retina in vertebrates eyes, blocking light [8]: 1 denotes the retina, 2 denotes the nerve fibers, 3 denote the optic nerve, 4 denotes the blind spot.
Human eye

Analogy between light receptors and pixels

- The angle subtended by cones/rods $\leftrightarrow$ image size
- Density of cones/rods $\leftrightarrow$ image sampling density.

The resolution of the human eye: approx $5.76 \times 10^8$ pixels.
Brightness adaption

Pixel intensities need to be quantized.

- Digital images usually quantized into $2^8 = 256$ levels (8 bits).
- Medical images quantized into $2^{12} = 4096$ levels (12 bits).

Human eye has a high, yet limited ability to discriminate between light intensity levels.

- Cannot operate over a wide range simultaneously.
- Adapt the overall sensitivity.
- Controlled by average light intensity often dominated by the background.
Brightness adaption

Perceived **brightness** is not the same as the actual **light intensity**.

Figure 13: Which block is brighter?
Threshold Vision

Weber’s Law

How well can we discriminate an intensity difference $\Delta I$ against a background intensity of $I$?

- A background with uniform intensity $I$
- Gradually increase the intensity in the circle by $\Delta I$.
- **Visibility threshold** $\Delta I$: change is detected 50% of the time.
Threshold vision

Weber ratio: \( \frac{\Delta I}{I} \)

- Weber ratio is approx. equal to 0.02 for a wide range of \( I \).
- When \( I \) is very small or very large, Weber ratio increases.
Threshold vision

If we work with $\ln I$

$$\frac{d \ln I}{dI} = \frac{1}{I} \quad \rightarrow \quad \Delta(\ln I) \approx \frac{\Delta I}{I} \approx 0.02$$  \hspace{1cm} (1)

- The sensitivity of human eyes is usually in ln-scale.
- Apply uniform quantizer on the contrast function:

$$c = \ln I$$  \hspace{1cm} (2)
Threshold vision

Example 1. If an imaging system has maximum contrast 100 : 1, a human observer can distinguish \( \frac{\ln 100}{0.02} = 230 \) grayscale levels.

- The grayscale levels uniformly spaced on the contrast scale \( c = \ln I \).
- 8-bit quantization is enough.

Why ln-scale?

- The firing rate of ganglion cells is approx. a logarithmic function of \( I \).
Visual masking

Real images are more complex.

- The thresholds are affected by presence of **structures** (edges, etc) close to the stimulus.
- The masking effect is stronger for high-contrast edges.
- The masking effect is local.
What is the just noticeable noise level? Frequency-dependent.

- Depend on the center frequency of noise.

Figure 14: The image corrupted by noise with the same SNR.
Lateral inhibition

Horizontal and amacrine cells reduce the signal from adjacent cells: **lateral inhibition**

- The spatial impulse response of the human eye $h(\vec{n})$.
- Equivalent to performing **spatial weighting** by the ganglion cells.
- Modulation Transfer function (MTF): $F(\omega) = FT(h(\vec{n}))$ is the frequency response of the human eye.
Modulation transfer function

Spatial frequency

Spatial frequency is a measure of the number of cycles subtended at the eye per degree.

Figure 15: (a) One cycle per degree. (b) Two cycles per degree. [9]
Modulation transfer function

The MTF typically exhibits a peak between 3 and 10 cycles per degree (cpd).

Figure 16: (a) The spatial impulse response of the eye. (b) The modulation transfer function, i.e. the frequency response of the eye. [10]
Modulation transfer function

Mach band effect can be explained by the MTF: the difference in the grayscale level is more obvious as we put two bands closer.
Mach band effect can be explained by the MTF: the difference in the grayscale level is more obvious as we put two bands closer.
Mach band effect can be explained by the MTF: the difference in the grayscale level is more obvious as we put two bands closer.
Modulation transfer function

Mach band effect can be explained by the MTF: the difference in the grayscale level is more obvious as we put two bands closer.
Example 2. A $500 \times 500$ image is shown on screen with height $H$, interpolated with zero-order-holder with intensities pattern $A, B, A, B, \cdots$. An observer sits at a distance $D = 8H$, is the pixelization artifact noticeable?

$$\arctan \frac{H}{250D} = \arctan \frac{1}{200} \approx \frac{1}{35} \text{ degree}$$

- The spatial frequency is then 35 cycles per degree.
- The MTF is low at this spatial frequency, hence the artifact should not be visible.
A simple vision model

Perceived *brightness* is not the same as the actual *light intensity*. 

\[
\begin{pmatrix}
\theta \\
\phi
\end{pmatrix}
\]
Acknowledgment: The slides are based on the lecture notes by Prof. Pierre Moulin, UIUC.

[1] B. staff,
“Medical gallery of blausen medical 2014,”

“Deep image reconstruction from human brain activity,”
*PLOS Computational Biology*, vol. 15, no. 1, pp. 1–23, 01 2019.

[3] Holly Fischer,
“ This image focuses on the three internal layers of the eye,”
https://commons.wikimedia.org/wiki/File:
Three_Internal_chambers_of_the_Eye.png,
Online; created 31 Jan 2013.


[7] Cmglee,
“Density of rod (dotted line) and cone (solid line) photoreceptors along a line passing through the fovea and the blind spot of a human eye vs the angle measured from the fovea, based on ’Foundations of Vision’ by Brian A. Wandell.,” https://commons.wikimedia.org/wiki/File:Human_photoreceptor_distribution.svg,
Online; created 29 November 2013.

[8] Caerbannog svg version, based on Jerry Crimson Mann png version.,
“Comparison of structures in vertebrate’s eye (left) with octopus’ eye (right).,”
https://commons.wikimedia.org/wiki/File:Evolution_eye.svg,
Online; created 4 April 2016.

[9] M. Kalloniatis and C. Luu,
“Visual acuity,“
[10] B. A. Wandell,

*Foundations of Vision*,