Frequency Response of the HVS

M.A. Amer

Lecture Outline:
- The HVS as an optical system
- Retina
- Perceived brightness
- Spatial characteristics
- Spatial Response
- Temporal Response
- Spatio-temporal Response
- Smooth Pursuit Eye Movement

Figures courtesy of [Wang02]
Human Visual System

- We can model the HVS as a **non-linear, non-uniformly sampled, spatially varying, non-separable system**.

- The **human eye** is an organ which reacts to light allowing vision:
  - Convert the light into signals
  - Transmit signals to brain
The HVS as an optical system

The eye is made up of three transparent structures

- The **outermost** layer is composed of the cornea and sclera
- The **middle** layer consists of the choroid, ciliary body, and iris
- The **innermost** is the retina
  - The optics of the eye creates an image of the visual world on the retina
  - Consists of cons, rods, fovea
The HVS as an optical system

- Light striking the retina initiates a cascade of chemical and electrical events that ultimately trigger nerve impulses, which are sent to the brain through the fibers of the optic nerve.
The Fovea is located in the center of the retina
- Connected to optic nerve
- Responsible for **sharp central vision**, necessary where visual detail is of primary importance (e.g., in reading)

Retina covers inner plane of eyeball (except lens area)
- With 120 Mio photo-receptor cells (cones and rods)
HVS: Retina – Rods and cones

**Rods:** Sensitive only in low light levels – “scotopic vision”
- Serve for brightness perception at lower luminance (dark adapted eye)
- Responsible for brightness, higher sensitivity
  - No color perception

**Cones:** Sensitive only in high light levels – “photopic vision”
- Human being have difficulty to see color in the dark
- Responsible for color vision
- Serve for brightness and color perception at higher luminance (day light)
- 3 types of cones: S (Blue), M(Green), L(Red) in Fovea
- Not uniformly sensitive
  - 3 different responses: more sensitive to green than red (64%L, 32%M, 2%S)
Frequency Responses of Cones & the Luminous Efficiency Function

\[ C_i = \int C(\lambda) a_i(\lambda) d\lambda, \quad i = r, g, b, y \]

where

- \( C_i \): Response of cone \( i \)
- \( C(\lambda) \): Incoming light distribution
- \( a_i(\lambda) \): Frequency response or absorption function

\[ \lambda = \nu / f \]
HVS: Spectral Redundancy

- RGB channels are redundant: \( Y = 0.299 \ R + 0.587 \ G + 0.114 \ B \)
- The output of the 3 cone color is transformed into an achromatic channel (such as luminance) and two chromatic channels (opponent channels)

Retina represents color in the opponent color system using “matrix operation”:
- Luminance channel: \( Y \) (=G+R)
- 1st chromatic channel: \( Y-B \) (= G+R-B)
- 2nd chromatic channel: \( R-G \)

- High-level HVS is much more sensitive to the variation in the achromatic channel than in the chromatic ones
- The spectral redundancy occurs in the low-level vision system (first three layers of the retina)
- Visual perception takes place at higher levels of human brain
- HVS is very complex & only part of it is explored
  - “Do not always believe your eyes”: Halloween or Lady?
HVS: Perceived brightness

- Perceived brightness depends on:
  1. object luminance \( L \)
  2. surrounding luminance &
  3. adaptation state of the eye

- Perceived brightness (\( H \)) difference \( dH \): proportional to relative luminance \( dL/L \)

- Weber-Fechner’s law (general perception law):
  - \( dH = k \frac{dL}{L} \)
  - \( H = k \cdot \log L + \text{const} \)
  
  ➔ The eye senses brightness logarithmically
  - If \( L \) triples in strength: \( L: 3 \times 1 \rightarrow 3 \times 3 \)
    ➔ \( H = k \log 3 + k \log 3 \)

  ➔ The relationship between stimulus and perception is logarithmic
  ➔ If a stimulus is multiplied by a fixed factor, the corresponding perception is altered in additive constant amounts
  ➔ For multiplications in stimulus strength, the strength of perception only adds
Perceived brightness
HVS: Spatial characteristics

- Most lenses, including the human lens, are not perfect optical systems
  ➔ when “visual stimuli” are passed through lenses they undergo a certain degree of degradation

- Deviations are due to
  - Aperture of the eye
  - Focus errors (spherical aberration)
  - Chromatic aberration
  - Dispersion

- These deviations can be summarized by a 2D convolution with an “optical impulse response”
HVS: Spatial characteristics

- *This “optical impulse response” is called “point-spread function” PSF*
  - Easier to measure is the “Line-spread function” LSF

- *Modulation transfer function* (MTF) describes the response of an optical system to an image decomposed into sine waves
  - MTF describes the spatial (angular) variation as a function of spatial (angular) frequency
  - MTF describes how much an optical system **blurs** the image of an object
  - MTF can be interpreted as the Fourier transform of the optical LSF
**HVS: Spatial characteristics**

- **Lateral inhibition (Mach):** neural activity of a fiber of the optic nerve depends on activity of a “receptive field” → activity in any one area of the HVS will tend to produce a **dampening effect** on other areas of the same system.
  - For example, activity in a given photoreceptor will inhibit activation in nearby photoreceptors.

- **Lateral inhibition:** is to mathematically differentiate the signal being processed (or to spatially high-pass filter it).
  - Acts as edge enhancement due to high-pass filtering the visual stimulus in the retina (differentiation).
HVS: Spatial characteristics

- **Sensitivity to resolution** is a function of the distance: Objects appear smaller as they move further away.
- **Spatial frequency** components are visible up to 60 cycles per degree (cpd).
- HVS is less sensitive to high frequencies.
- **Eccentricity** (region of interest): to stimulate the fovea the stimulus must be centered on the screen and cover a visual angle of 2 degrees.
- **Spatial contrast sensitivity** of the HVS: band-pass.
  - High sensitivity to a band of spatial frequencies (approx. 1-10 cpd).
Spatial Response

- Saccadic eye movement: very rapid jumps from one fixation point to another
- Enhance contrast sensitivity BUT reduces the frequency at which peak response occurs
  - Fig. 2.6: HVS sensitivity is 10x higher with normal eye movement than without (i.e., with stabilization) BUT the peak responses occurs at 2 cpd (normal eye) compared to 4 cpd (stabilization)

Figure 2.6 The spatial frequency response of the HVS, obtained by a visual experiment. The three curves result from different stabilization settings used to remove the effect of saccadic eye movements. Filled circles were obtained under normal, unstabilized conditions; open squares, with optimal gain setting for stabilization; open circles, with the gain changed about 5 percent. Reprinted from D. H. Kelly, Motion and vision. I. Stabilized images of stationary gratings, J. Opt. Soc. Am. (1979), 69:1266–74, by permission of the Optical Society of America.
Spatial frequency response: sensitivity to a stationary pattern with different spatial frequency → Band-pass filter

- Spatial impulse response of the HVS: band-pass filter
  - Positive weighting to nearby cells and negative weighting to further cells

- Experiment: view a vertical bar that moves horizontally
HVS: Visibility & Masking

- Visibility thresholds often increased in the vicinity of edges but sometimes decreased ("masking")

- HVS is less sensitive to variations in high-frequency content: masking property
  - Noise is less visible to the HVS in high-frequency contents that in smooth areas
HVS: Temporal characteristics

- HVS is less sensitive to details moving cross retina
- Fast moving objects become “blurred”

- Spatio-temporal contrast sensitivity of the luminance channel has band pass characteristic
  - Contrast sensitivity function separable for high spatial and temporal frequencies only
Temporal Response

- Critical flicker frequency: The lowest frame rate at which the eye does not perceive flicker

- Provides guideline for determining the frame rate when designing a video system
  - Watching a movie needs lower frame rate than TV

- Critical flicker frequency depends on the mean brightness of the display

- 60 Hz is typically sufficient for watching TV
The reciprocal relation between spatial and temporal sensitivity was used in TV system design.

Interlaced scan provides tradeoff between spatial and temporal resolution.

Figure 2.7  Spatiotemporal frequency response of the HVS. (a) Spatial frequency responses for different temporal frequencies of 1 Hz (open circles), 6 Hz (filled circles), 16 Hz (open triangles), and 22 Hz (filled triangles). (b) Temporal frequency responses for different spatial frequencies of 0.5 cpd (open circles), 4 cpd (filled circles), 16 cpd (open triangles), and 22 cpd (filled triangles). Reprinted from J. G. Robson, Spatial and temporal contrast sensitivity functions of the visual systems, *J. Opt. Soc. Am.* (1966), 56:1141–42, by permission of the Optical Society of America.
Relation between Motion, Spatial and Temporal Frequency

Consider an object moving with speed \((v_x, v_y)\)

Assume the object image pattern at \(t = 0\) is \(\Psi_0(x, y)\),

the image pattern at time \(t\) is

\[
\Psi(x, y, t) = \Psi_0(x - v_x t, y - v_y t)
\]

\[\Leftrightarrow\]

\[
\Psi(f_x, f_y, f_t) = \Psi_0(f_x, f_y) \delta(f_t + v_x f_x + v_y f_y)
\]

Relation between motion, spatial, and temporal frequency:

\[
f_t = -(v_x f_x + v_y f_y)
\]

The temporal frequency of the image of a moving object depends on motion as well as on the spatial frequency of the object.

Example: A plane with vertical bar pattern, moving vertically, causes no temporal change; But moving horizontally, it causes fastest temporal change.
Smooth Pursuit Eye Movement

- Smooth Pursuit: the eye tracks moving objects
- Net effect: reduce the velocity of moving objects on the retinal plane, so that the eye can perceive much higher raw temporal frequencies than indicated by the temporal frequency response

Temporal frequency caused by object motion when the object is moving at \((v_x, v_y)\): 
\[ f_t = -(v_x f_x + v_y f_y) \]

Observed temporal frequency at the retina when the eye is moving at \((\tilde{v}_x, \tilde{v}_y)\): 
\[ \tilde{f}_t = f_t + (\tilde{v}_x f_x + \tilde{v}_y f_y) \]

\[ \tilde{f}_t = 0 \text{ if } \tilde{v}_x = v_x, \tilde{v}_y = v_y \]

- Smooth pursuit eye movement sheer window of perception in spatio-temporal frequency space
Smooth Pursuit Eye Movement

Figure 2.8 Spatiotemporal response of the HVS under smooth pursuit eye movements: (a) without smooth pursuit eye movement; (b) with eye velocity of 2 deg/s; (c) with eye velocity of 10 deg/s. Reprinted from Giroud, B. “Motion compensation: visual aspects, accuracy, and fundamental limits.” In Sezan, M. I., and R. L. Lagendijk, eds., Motion Analysis and Image Sequence Processing, Boston: Kluwer Academic Publishers, 1993, 126–52, by permission of Kluwer Academic Publishers.
Summary

- HVS is a set of filters
- Logarithmic relationship between luminance and subjective brightness perception
- Lateral inhibition $\implies$ spatial bandpass characteristic
- Chromaticity channels has lower bandwidth
- Visibility thresholds often increases in the vicinity of edges but sometimes decreased ("masking")
- Spatial frequency components visible up to 60 cpd
- Temporal frequency depends on motion & spatial frequencies of an object
- 60 Hz is typically sufficient for watching TV
- Smooth pursuit eye movement sheer window of perception