Models of early visual perception

Part 1/6 – perceived brightness of light

Rafal Mantiuk

Computer Laboratory, University of Cambridge
Many graphics/display solutions are motivated by visual perception.

Image & video compression

Display spectral emission - metamerism

Halftoning

Display’s subpixels

Camera’s Bayer pattern

Color wheel in DLPs
Luminance (again)

- Luminance – measure of light weighted by the response of the achromatic mechanism. Units: cd/m²

\[ L_V = \int_{350}^{700} kL(\lambda)V(\lambda)d\lambda \]

\[ k = 683.002 \]
Steven’s power law for brightness

- Stevens (1906-1973) measured the perceived magnitude of physical stimuli
  - Loudness of sound, tastes, smell, warmth, electric shock and brightness
  - Using the magnitude estimation methods
    - Ask to rate loudness on a scale with a known reference

- All measured stimuli followed the power law:
  \[ \phi(I) = kI^a \]
  - Perceived magnitude
  - Exponent
  - Constant
  - Physical stimulus

- For brightness (5 deg target in dark), \( a = 0.3 \)
Steven’s law for brightness
Steven’s law vs. Gamma correction

Stevens’ law
a=0.3

Gamma function
Gamma = 2.2
Models of early visual perception

Part 2/6 – contrast detection

Rafal Mantiuk
Computer Laboratory, University of Cambridge
Detection thresholds

- The smallest detectable difference between
  - the luminance of the object and
  - the luminance of the background
Threshold versus intensity (t.v.i.) function

- The smallest detectable difference in luminance for a given background luminance

![Graph showing the relationship between log10 detection threshold \(\Delta L\) and log10 background luminance in cd/m².](image)
t.v.i. measurements – Blackwell 1946
Psychophysics
Threshold experiments

L + ΔL

Luminance difference ΔL

Detection threshold

Psychometric function

P = 0.75

Probability

Luminance difference ΔL

1
0.8
0.6
0.4
0.2
0
10^{-2}
10^{-1}
10^0
10^1
10^2
t.v.i. function / c.v.i. function / Sensitivity

- The same data, different representation

Threshold vs. intensity

Contrast vs. intensity

Sensitivity

\[ \Delta L = L_{\text{disk}} - L_{\text{background}} \]

\[ C = \frac{\Delta L}{L_{\text{background}}} \]

\[ S = \frac{1}{C} = \frac{L_{\text{background}}}{\Delta L} \]
Sensitivity to luminance

- Weber-law – the just-noticeable difference is proportional to the magnitude of a stimulus

\[ \frac{\Delta L}{L} = k \]

- The smallest detectable luminance difference
- Background (adapting) luminance

Typical stimuli:

Ernst Heinrich Weber
[From wikipedia]
Consequence of the Weber-law

- Smallest detectable difference in luminance
  \[ \frac{\Delta L}{L} = k \]
  For \( k = 1\% \)

<table>
<thead>
<tr>
<th>L</th>
<th>( \Delta L )</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 cd/m(^2)</td>
<td>1 cd/m(^2)</td>
</tr>
<tr>
<td>1 cd/m(^2)</td>
<td>0.01 cd/m(^2)</td>
</tr>
</tbody>
</table>

- Adding or subtracting luminance will have different visual impact depending on the background luminance
- Unlike LDR luma values, luminance values are **not** perceptually uniform!
How to make luminance (more) perceptually uniform?

- Using “Fechnerian” integration

\[
\frac{dR}{dl}(L) = \frac{1}{\Delta L(L)}
\]

Detection threshold

Luminance transducer:

\[
R(L) = \int_{L_{\text{min}}}^{L} \frac{1}{\Delta L(l)} \, dl
\]
Assuming the Weber law

\[ \frac{\Delta L}{L} = k \]

- and given the luminance transducer

\[ R(L) = \int \frac{1}{\Delta L(l)} dl \]

- the response of the visual system to light is:

\[ R(L) = \int \frac{1}{kL} dL = \frac{1}{k} \ln(L) + k_1 \]
Fechner law

\[ R(L) = a \ln(L) \]

- Response of the visual system to luminance is \textit{approximately} logarithmic
But...the Fechner law does not hold for the full luminance range

- Because the Weber law does not hold either
- Threshold vs. intensity function:
If we allow detection threshold to vary with luminance according to the t.v.i. function:

\[ R(L) = \int_0^L \frac{1}{tvi(l)} dl \]
Fechnerian integration and Stevens’ law

\[ R(L) = \int_0^L \frac{1}{tvi(l)} dl \]

- **R(L)** - function derived from the t.v.i. function
Applications of JND encoding – R(L)

- **DICOM grayscale function**
  - Function used to encode signal for medical monitors
  - 10-bit JND-scaled (just noticeable difference)
  - Equal visibility of gray levels

- **HDMI 2.0a (HDR10)**
  - PQ (Perceptual Quantizer) encoding
  - Dolby Vision
  - To encode pixels for high dynamic range images and video
Models of early visual perception

Part 3/6 – spatial contrast sensitivity and contrast constancy

Rafal Mantiuk

Computer Laboratory, University of Cambridge
Resolution and sampling rate

- Pixels per inch [ppi]
  - Does not account for vision
- The visual resolution depends on
  - screen size
  - screen resolution
  - viewing distance
- The right measure
  - Pixels per visual degree [ppd]
  - In frequency space
    - Cycles per visual degree [cpd]
Fourier analysis

- Every N-dimensional function (including images) can be represented as a sum of sinusoidal waves of different frequency and phase

\[ \text{Pixel value} = \sum \]

- Think of “equalizer” in audio software, which manipulates each frequency
Spatial frequency in images

- Image space units: cycles per sample (or cycles per pixel)

- What are the screen-space frequencies of the red and green sinusoid?

- The visual system units: cycles per degree

  - If the angular resolution of the viewed image is 55 pixels per degree, what is the frequency of the sinusoids in cycles per degree?
Nyquist frequency

- Sampling density restricts the highest spatial frequency signal that can be (uniquely) reconstructed
  - Sampling density – how many pixels per image/visual angle/…

- Any number of sinusoids can be fitted to this set of samples
- It is possible to fit an infinite number of sinusoids if we allow infinitely high frequency
Nyquist frequency

- Sampling density restricts the highest spatial frequency signal that can be (uniquely) reconstructed
  - Sampling density – how many pixels per image/visual angle/…

- Any number of sinusoids can be fitted to this set of samples
  - It is possible to fit an infinite number of sinusoids if we allow infinitely high frequency
Nyquist frequency

- Sampling density restricts the highest spatial frequency signal that can be (uniquely) reconstructed
  - Sampling density – how many pixels per image/visual angle/…

Any number of sinusoids can be fitted to this set of samples
- It is possible to fit an infinite number of sinusoids if we allow infinitely high frequency
Nyquist frequency

- Sampling density restricts the highest spatial frequency signal that can be (uniquely) reconstructed
  - Sampling density – how many pixels per image/visual angle/…

- Any number of sinusoids can be fitted to this set of samples
- It is possible to fit an infinite number of sinusoids if we allow infinitely high frequency
Nyquist frequency / aliasing

- Nyquist frequency is the highest frequency that can be represented by a discrete set of uniform samples (pixels)
- Nyquist frequency = 0.5 sampling rate
  - For audio
    - If the sampling rate is 44100 samples per second (audio CD), then the Nyquist frequency is 22050 Hz
  - For images (visual degrees)
    - If the sampling rate is 60 pixels per degree, then the Nyquist frequency is 30 cycles per degree
- When resampling an image to lower resolution, the frequency content above the Nyquist frequency needs to be removed (reduced in practice)
  - Otherwise aliasing is visible
Modeling contrast detection

- Photoreceptors
- Lens
- Cornea
- Defocus & Aberrations
- Glare
- Retinal ganglion cells
- Adaptation
- Colour opponency
- Luminance masking
- Spectral sensitivity
- P & M visual pathways
- LGN
- Visual Cortex
- Detection
- Integration
- Contrast masking
- Spatial- / orientation- / temporal- Selective channels
- Contrast Sensitivity Function
Campbell & Robson contrast sensitivity chart

Spatial frequency [cycles per degree]
Contrast sensitivity function

\[ CSF = S(\rho, \theta, \omega, l, i^2, d, e) \]
CSF as a function of spatial frequency

Sensitivity ($L/\Delta L$) vs. Spatial frequency [cpd]

$L_\text{b}$ = 0.001 cd/m²
CSF as a function of background luminance
CSF as a function of spatial frequency and background luminance
Contrast constancy

Experiment: Adjust the amplitude of one sinusoidal grating until it matches the perceived magnitude of another sinusoidal grating.

Contrast constancy
No CSF above the detection threshold
CSF and the resolution

- CSF plotted as the detection contrast
  \[ \frac{\Delta L}{L_b} = S^{-1} \]

- The contrast below each line is invisible

- Maximum perceivable resolution depends on luminance

Expected contrast in natural images

CSF models:
https://doi.org/10.1117/12.537476
Spatio-chromatic CSF
Spatio-chromatic contrast sensitivity

- CSF as a function of **luminance** and **frequency**

**Black-White**

**Red-Green**

**Violet-Yellow**

http://dx.doi.org/10.2352/issn.2169-2629.2020.28.1
CSF and colour ellipses

- Colour discrimination as a function of
  - Background colour and luminance [LMS]
  - Spatial frequency [cpd]
  - Size [deg]
Visibility of blur

- The same amount of blur was introduced into light-dark, red-green and blue-yellow colour opponent channels
- The blur is only visible in light-dark channel
- This property is used in image and video compression
  - Sub-sampling of colour channels (4:2:1)
Models of early visual perception

Part 4/6 – lateral inhibition and multi-resolution models

Rafal Mantiuk

Computer Laboratory, University of Cambridge
Mach Bands – evidence for band-pass visual processing

- “Overshooting” along edges
  - Extra-bright rims on bright sides
  - Extra-dark rims on dark sides

- Due to “Lateral Inhibition“
Centre-surround (Lateral Inhibition)

- “Pre-processing” step within the retina
- Surrounding brightness level weighted negatively
  - A: high stimulus, maximal bright inhibition
  - B: high stimulus, reduced inhibition & stronger response
  - D: low stimulus, maximal inhibition
  - C: low stimulus, increased inhibition & weaker response
Centre-surround: Hermann Grid

- Dark dots at crossings
- Explanation
  - Crossings (A)
    - More surround stimulation (more bright area)
    - Less inhibition
    - Weaker response
  - Streets (B)
    - Less surround stimulation
    - More inhibition
    - Greater response
- Simulation
  - Darker at crossings, brighter in streets
  - Appears more steady
  - What if reversed?
some further weirdness
Spatial-frequency selective channels

- The visual information is decomposed in the visual cortex into multiple channels
  - The channels are selective to spatial frequency, temporal frequency, and orientation
  - Each channel is affected by different "noise" level
  - The CSF is the net result of information being passed in noise-affected visual channels

From: Wandell, 1995
Multi-scale decomposition

Steerable pyramid decomposition
Multi-resolution visual model

- Convolution kernels are band-pass, orientation selective filters

- The filters have the shape of an oriented Gabor function

From: Wandell, 1995
Applications of multi-scale models

- JPEG2000
  - Wavelet decomposition

- JPEG / MPEG
  - Frequency transforms

- Image pyramids
  - Blending & stitching
  - Hybrid images

Hybrid Images by Aude Oliva
http://cvcl.mit.edu/hybrid_gallery
Models of early visual perception
Part 5/6 – light and dark adaptation

Rafal Mantiuk
Computer Laboratory, University of Cambridge
Light and dark adaptation

- Light adaptation: from dark to bright
- Dark adaptation: from bright to dark (much slower)
Time-course of adaptation

- Bright -> Dark
- Dark -> Bright
Temporal adaptation mechanisms

- Bleaching & recovery of photopigment
  - Slow asymmetric (light -> dark, dark -> light)
  - Reaction times (1-1000 sec)
  - Separate time-course for rods and cones

- Neural adaptation
  - Fast
  - Approx. symmetric reaction times (10-3000 ms)

- Pupil
  - Diameter varies between 3 and 8 mm
  - About 1:7 variation in retinal illumination
Night and daylight vision

Vision mode:

<table>
<thead>
<tr>
<th>Vision mode</th>
<th>SCOTOPIC</th>
<th>MESOPIC</th>
<th>PHOTOPIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>rod activity</td>
<td>night light</td>
<td>office light</td>
<td>daylight</td>
</tr>
<tr>
<td>cone activity</td>
<td>Luminance [log cd/m²]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Mode properties:

- monochromatic vision
- limited visual acuity
- good color perception
- good visual acuity

Luminous efficiency

![Graph showing rod and cone sensitivity](image)
Advanced Graphics and Image Processing

Models of early visual perception

Part 6/6 – high(er) level vision

Rafal Mantiuk

Computer Laboratory, University of Cambridge
Simultaneous contrast
High-Level Contrast Processing
High-Level Contrast Processing

Checker-shadow illusion:
The squares marked A and B are the same shade of gray.

Edward H. Adelson
Shape Perception

- Depends on surrounding primitives
  - Directional emphasis
  - Size emphasis

http://www.panoptikum.net/optischetaeuscheinungen/index.html
Shape Processing: Geometrical Clues

- Automatic geometrical interpretation
  - 3D perspective
  - Implicit scene depth

http://www.panoptikum.net/optischetaeuschen/index.html
Impossible Scenes

- Escher et.al.
  - Confuse HVS by presenting contradicting visual clues
  - Local vs. global processing

http://www.panoptikum.net/optischetaeuschen/index.html
caused by saccades, motion from dark to bright areas
Law of closure
References

  
  Available online: [https://foundationsofvision.stanford.edu/](https://foundationsofvision.stanford.edu/)

  
  Section 2.4