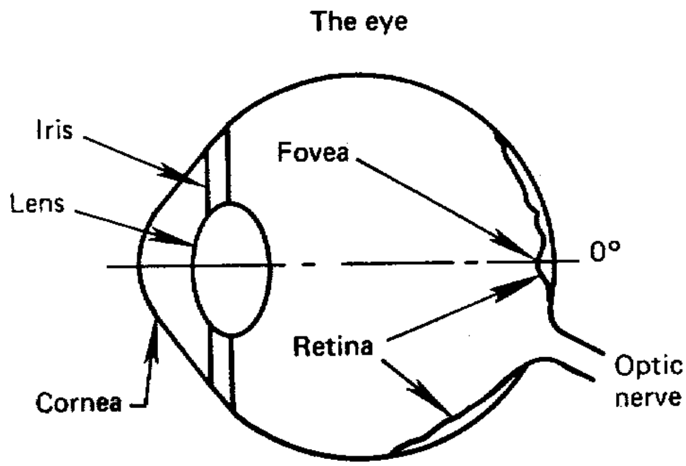
The Visual Perception of Images

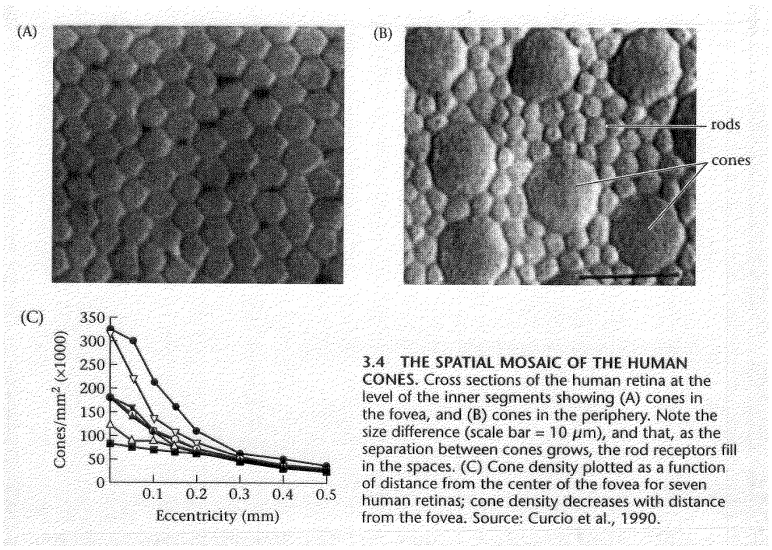
- In order to understand “images” you must understand how humans perceive visual stimulus.
- Objectives:
  - Understand contrast and how humans detect changes in images.
  - Understand photometric properties of the physical world.
  - Understand the percept of “color”.
  - Learn how to use this understanding to design imaging systems.

The Eye

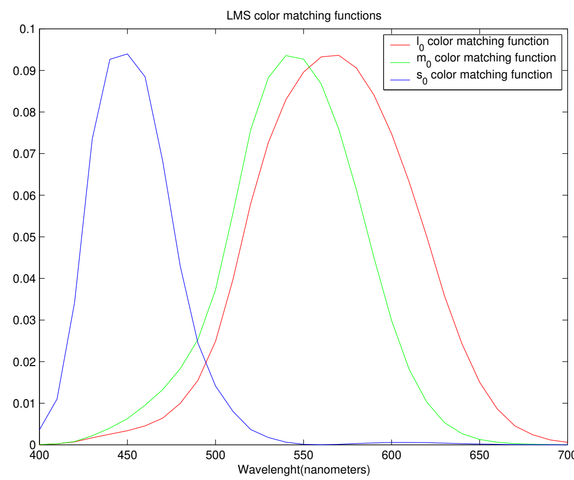


- Retina - the “focal plane array” on the back surface of the eye the detects and measures light.
- Photoreceptors - the nerves in the retina that detect light.
- Fovea - a small region in the retina ($\approx 1^\circ$) with high spatial resolution.
- Blind spot - a small region in the retina where the optic nerve is located that has no photoreceptors.

Visual System Basics



3.4 THE SPATIAL MOSAIC OF THE HUMAN CONES. Cross sections of the human retina at the level of the inner segments showing (A) cones in the fovea, and (B) cones in the periphery. Note the size difference (scale bar = 10 μm), and that, as the separation between cones grows, the rod receptors fill in the spaces. (C) Cone density plotted as a function of distance from the center of the fovea for seven human retinas; cone density decreases with distance from the fovea. Source: Curcio et al., 1990.



- Rods - a type of photoreceptor that is used for achromatic vision at very low light levels (scotopic vision).
- Cones - a type of photoreceptor that is used for color vision at high light levels (phototopic vision).
- Long, medium, and short cones - the three specific types of codes used to produce color vision. This cones are sensitive to long (L or red) wavelengths, medium (M or green) wavelengths, and short (S or blue) wavelengths.

Luminance

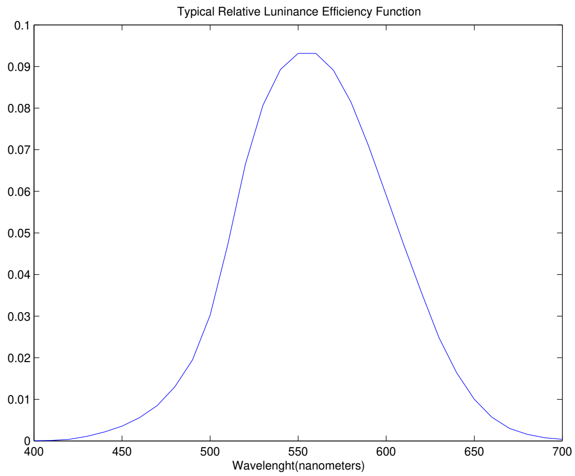
- Luminance describes the "achromatic" component of an image.
- λ - wavelength of light
- Most light contains a spectrum of energy at different wavelengths. So that

$$\text{Energy between } \lambda_1 \text{ and } \lambda_2 = \int_{\lambda_1}^{\lambda_2} I(\lambda) d\lambda$$

- Human visual system's (HVS) sensitivity is a function of wavelength. Most important region is from 400 nm to 700 nm.
- Informal definition: $y(\lambda)$ is the visual sensitivity as a function of wavelength.
- Luminance is defined as:

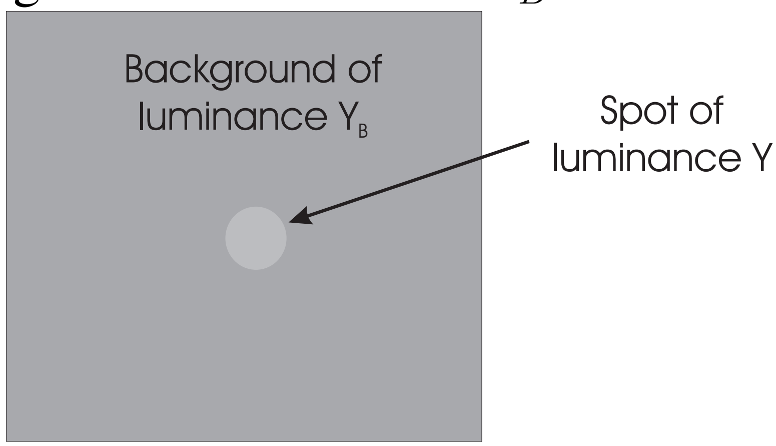
$$Y \triangleq \int_0^\infty I(\lambda) y(\lambda) d\lambda$$

- Note: Y is proportional to energy!!


Typical Relative Luninance Efficiency Function
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
400
450
500
550
600
650
700
Wavelenght(nanometers)

A Simple Visual Stimulus

- A single uniform dot of luminance Y ($\approx 10^\circ$) in a large uniform background of luminance Y_B.



Question: How much difference is necessary for a "standard observer" to notice the difference between Y and Y_B?

- Definitions:
  - The just noticeable difference (JND) is the difference that allows an observer to detect the center stimulus 50% of the time.
  - ΔY_{JND} is the difference in Y and Y_B required to achieve a just noticeable difference.

The Problem with Linear Luminance

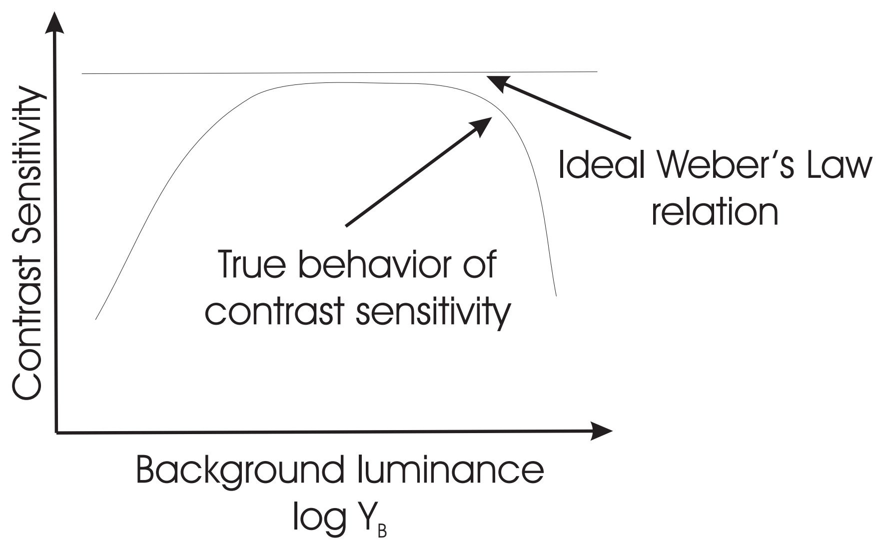
- Consider the following gedanken experiment:
  - **Experiment 1** - A visual experiment uses a background formed by a uniformly illuminated white board in an otherwise dark room. In this case, $Y_B = 1$ and $Y = 1.1$ achieves a JND. So, $\Delta Y_{JND} = 0.1$.
  - **Experiment 2** - A visual experiment uses a background formed by a uniformly illuminated white board in a bright outdoor environment. In this case, $Y_B = 1000$ and $Y = 1000.1$. Does $\Delta Y = 0.1$ still achieve a JND? **No!**
- Conclusion $\Rightarrow$ ΔY_{JND} is a strong function of the background luminance Y_B.

Weber's Contrast

- We need a quantity to measure JND changes in luminance which is independent (less dependent) on the background luminance Y_B.
- Definitions:
  - *Contrast* - $C \stackrel{\triangle}{=} \frac{Y - Y_B}{Y_B} = \frac{\Delta Y}{Y_B}$
  - *JND Contrast* - $C_{JND} \stackrel{\triangle}{=} \frac{\Delta Y_{JND}}{Y_B}$
  - *Contrast sensitivity* - $S \stackrel{\triangle}{=} \frac{1}{C_{JND}}$
- Comments:
  - Contrast is the *relative* change in luminance.
  - A small value of C_{JND} means that you are very sensitive to changes in luminance.
  - A large value of C_{JND} means that you are very insensitive to changes in luminance.

Weber's Law

- Weber's Law: The contrast sensitivity is approximately independent of the background luminance.
  - Relative changes in luminance are important.
  - Weber's law tends to break down for very dark and very bright luminance levels.



- At very low luminance, detector noise, and ambient light tend to reduce sensitivity, so the stimulus appears "black".
- At very high luminance, the very bright background tends to saturate detector sensitivity, thereby reducing sensitivity by "blinding" the subject.
- We are most concerned with the low and midrange luminance levels.

Perceptually Uniform Representations

- Problem:
  - Unit changes in Luminance Y **do not** correspond to unit changes in visual sensitivity.
  - When Y is large, changes luminance are less noticeable $\Rightarrow \Delta Y_{JND}$ is large
  - When Y is small, changes luminance are more noticeable $\Rightarrow \Delta Y_{JND}$ is small
- Define new quantity $L =$ "Lightness" so that...
  - $\Delta L_{JND} =$ constant
  - L is said to be perceptually uniform
  - Unit changes in L correspond to unit changes in visual sensitivity
  - $L = f(Y)$ for some function $f(\cdot)$
  - Quantization effects are much less noticeable in L than in Y

Derivation of Logarithmic Lightness Transformation

- If we believe Weber's Law, then ...

$$\Delta L_{JND} = \frac{\Delta Y_{JND}}{Y} = C_{JND}$$

  – In the differential limit

$$dL = \frac{dY}{Y}$$

$$\int dL = \int \frac{dY}{Y}$$

  – Integrating results in ...

$$\boxed{L = \log(Y)}$$

Log Luminance Transformations

$$L = \log Y$$

- Advantages:
  - Weber's Law says that fixed changes in L will correspond to equally visible changes in an image.
  - This makes L useful for problem such as image quantization, and compression.
- Problems:
  - $L = \log Y$ is not defined for $Y = 0$.
  - Weber's law is an approximation, particularly at low luminance levels were sensitivity is reduced.
  - We know that contrast sensitivity increases with Y
  - Over emphasizes sensitivity in dark regions

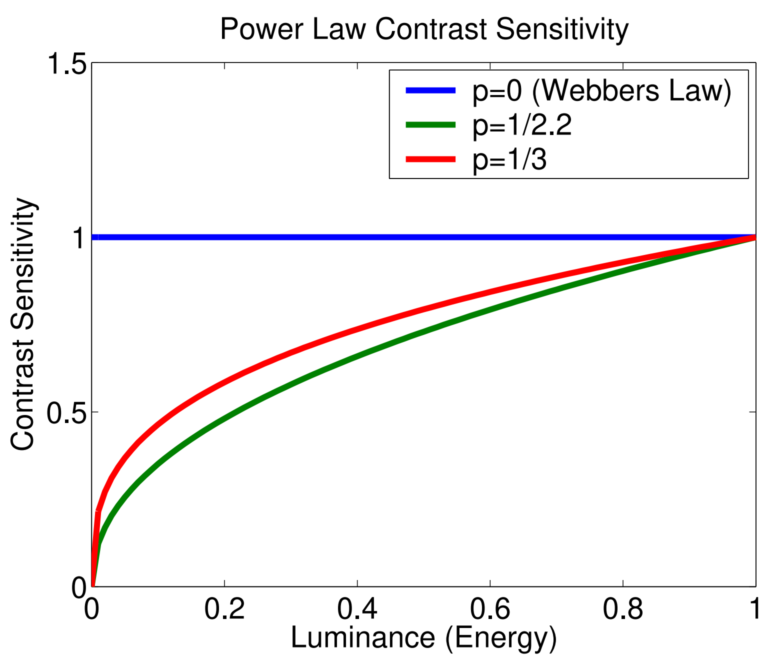
Power Law Correction to Weber's Law

- We can correct Weber's Law by weighting contrast with an increasing function of Y.

$$\begin{aligned} \Delta L_{JND} &= \frac{\Delta Y_{JND}}{Y} Y^p \\ &= C_{JND} Y^p \end{aligned}$$

- So the contrast sensitivity is given by ...

$$S = \frac{1}{C_{JND}} = \frac{1}{\Delta L_{JND}} Y^p$$



Derivation of Power Law Lightness Transformation

- If we believe

$$\Delta L_{JND} = \frac{\Delta Y_{JND}}{Y} Y^p = C_{JND} Y^p$$

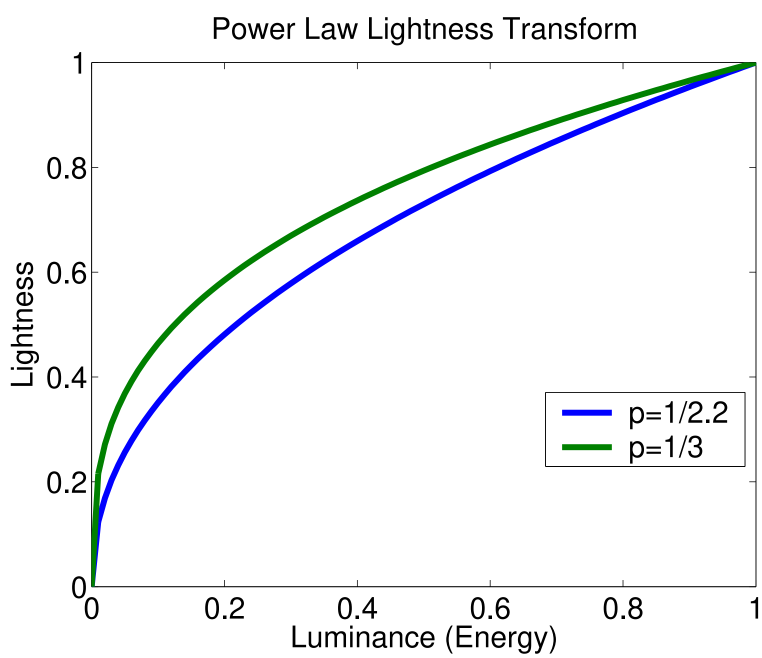
  - In the differential limit

$$dL = \frac{dY}{Y^{1-p}}$$

$$\int dL = \int \frac{dY}{Y^{1-p}}$$

  - Integrating and rescaling results in ...

$$\boxed{L = Y^p}$$

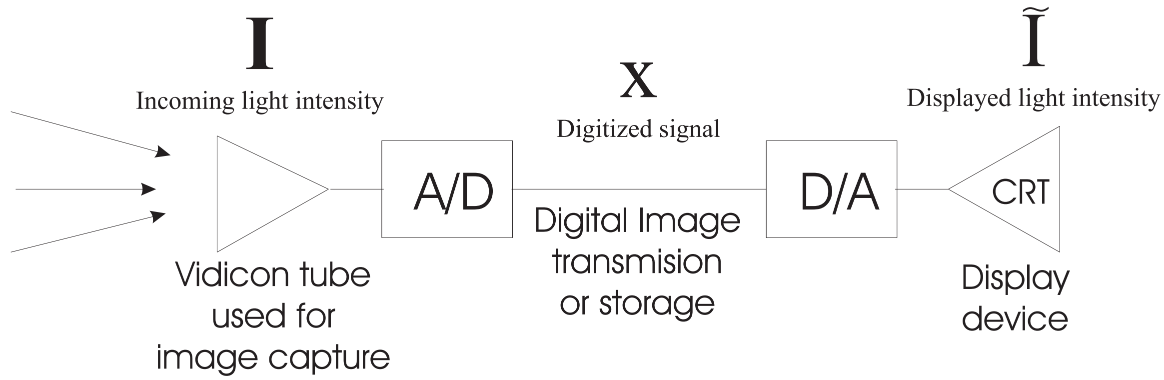


Power Law Luminance Transformations

$$L = Y^p$$

- L is well defined for $Y = 0$.
- L models the reduced sensitivity at low luminance levels.
- $p = 1/3$ is known to fit empirical data well.
- $p = 1/2.2$ is more robust and is widely used in applications.
- We will see that $p = \frac{1}{\gamma}$ where γ is the parameter used in "gamma correction."
- Typical values of p:
  - NTSC video $p = 1/2.2$
  - sRGB color standard $p = 1/2.2$
  - Standard PC and Unix displays $p = 1/2.2$
  - MacIntosh computers $p = 1/1.8$
  - $L^* a^* b^*$ visually uniform color space $p = 1/3$

Input and Output Nonlinearities in Imaging Systems



- $x(m,n)$ generally takes values from 0 to 255
- Videcon tubes are (were) nonlinear with input/output relationship.

$$x = 255 \left(\frac{I}{I_{in}} \right)^{1/\gamma_i}$$

  where I_{in} is the maximum input, and γ_i is a parameter of the input device.

- The cathode ray tube (CRT) has the inverse input/output relationship.

$$\tilde{I} = I_{out} \left(\frac{x}{255} \right)^{\gamma_o}$$

  where I_{out} is the maximum output and γ_o is a parameter of the output device.

Gamma Correction

- The input/output relationship for this imaging system is then

$$
\begin{aligned}
\tilde{I} &= I_{out}\left(\frac{x}{255}\right)^{\gamma_o} \\
&= I_{out}\left(\frac{255\left(\frac{I}{I_{in}}\right)^{1/\gamma_i}}{255}\right)^{\gamma_o} \\
&= I_{out}\left(\frac{I}{I_{in}}\right)^{\gamma_o/\gamma_i}
\end{aligned}
$$

  So we have that

$$
\frac{\tilde{I}}{I_{out}} = \left(\frac{I}{I_{in}}\right)^{\gamma_o/\gamma_i}
$$

- If $\gamma_i = \gamma_o$, then

$$
\tilde{I} = \frac{I_{out}}{I_{in}} I
$$

- Definition: The signal x is said to be *gamma corrected* because it is predistorted to display properly on the CRT.

Visual MTF

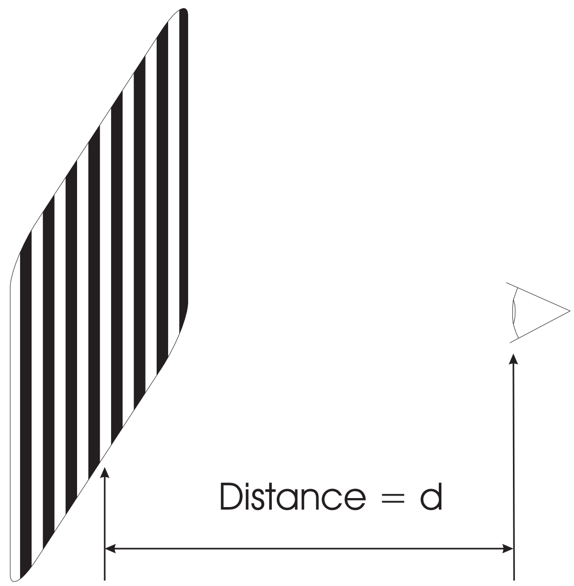
- How do we quantify the spacial frequency response of the visual system?
- Answer: Measure the contrast sensitivity as a function of spatial frequency.

Experiment for Measuring MTF

- Produce a horizontal sine wave pattern with the form

$$Y(x,y) = \frac{\Delta Y}{2} \cos\left(2\pi f_0 x\right) + Y_B$$

  and display the pattern to a viewer at a distance d.



- f_0 has units of cycles per inch.

$$\text{visual angle in degrees} \;=\; \sin^{-1}\left(\frac{\Delta x}{d}\right)\frac{180}{\pi}$$

$$\approx\; \frac{\Delta x}{d}\frac{180}{\pi}$$

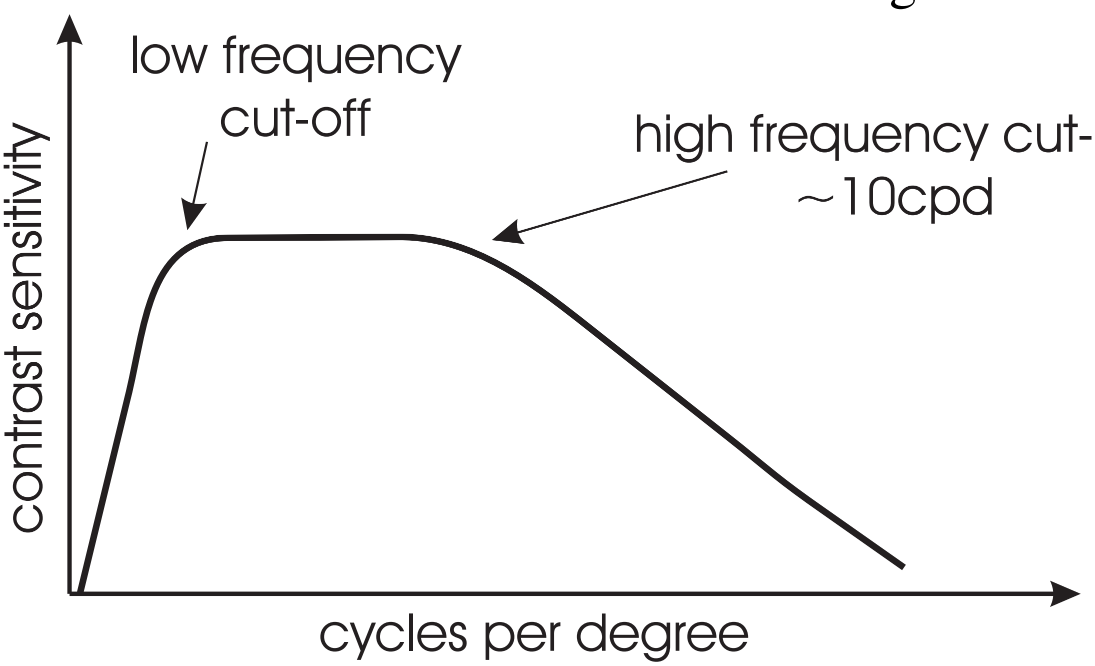
  So the spacial frequency in cycles per degree is given by

$$\tilde{f}_0 = \frac{\pi}{180} d\, f_0$$

- Larger distance $\Rightarrow$ higher frequency

Contrast Sensitivity Function (CSF)

- Let $S(\tilde{f})$ be the contrast sensitivity measured as a function of $\tilde{f}$ the spatial frequency in cycles per degree.
- $S(\tilde{f})$ is known as the contrast sensitivity function.
- Typical CSF function looks like the following.



  - Bandpass function
  - High frequency cut-off primarily due to optics of eye $\Rightarrow$ linear in energy.
  - Low frequency cut-off due to neural response.
  - Accurate measurement of CSF requires specialized techniques.

Image Fidelity and Quality Metrics

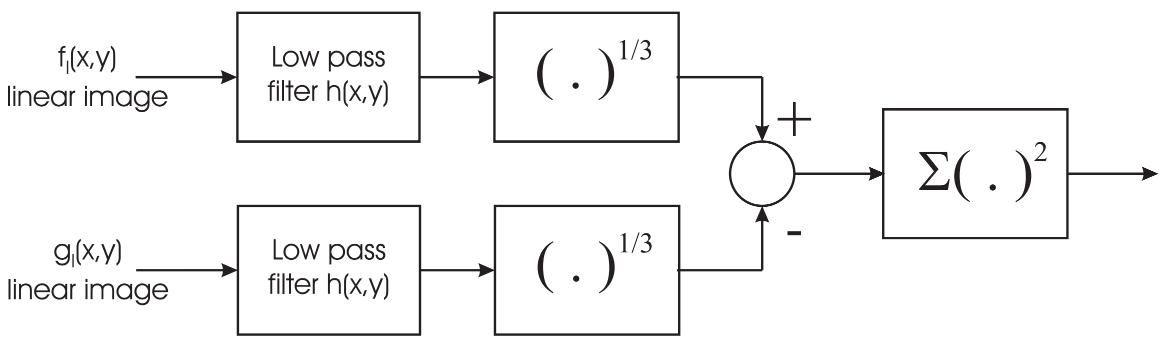
- Image Fidelity
  - Evaluates whether a processed image faithfully represents the original.
  - Usually differences are at or near JND levels.
  - Can be measured using well established psychophysical methods.
- Image Similarity
  - Attempts to quantify how similar two images are.
  - Usually differences are well above JND levels.
  - Very important in applications such as image database retrieval.
- Image Quality
  - Evaluates how pleasing an image is to the viewer.
  - Depends critically on factors such as sharpness, contrast, and color gamut.
  - Very important but more difficult to quantify.

A Simple Image Fidelity Model

- Let $f_g(x,y)$ and $g_g(x,y)$ be gamma corrected images.
- Let $f_l(x,y)$ and $g_l(x,y)$ be linear images.

$$
\begin{aligned}
f_l(x,y) &= \left(\frac{f_g(x,y)}{255}\right)^{\gamma} \\
g_l(x,y) &= \left(\frac{g_g(x,y)}{255}\right)^{\gamma}
\end{aligned}
$$

- Then a simple model for fidelity is:



  – Input is linear in energy.
  – $h(x,y)$ is low pass filter corresponding to CSF.
  – Usually, low frequency cut-off in CSF is ignored.
  – Total squared error is computed after cube-root is computed.