ECE 634: Digital Video Systems
Frequency analysis: 1/17/17

Professor Amy Reibman
MSEE 356
reibman@purdue.edu

http://engineering.purdue.edu/~reibman/ece634/index.html

Outline 1/17/17

- Video Frequencies
  - Fourier Transforms
  - Spatial and temporal frequencies and the impact of motion
  - A bit about Human Visual Perception

One Dimensional Fourier Transform (Continuous time)

- $\psi(x)$ is a 1D continuous time sequence
- Forward Transform

$$\Psi(f) = \int_{-\infty}^{\infty} \psi(x) e^{-j2\pi fx}\, dx$$

- Inverse Transform

$$\psi(x) = \int_{-\infty}^{\infty} \Psi(f) e^{j2\pi fx}\, df$$

Two Dimensional Fourier Transform (Continuous space) (CSFT)

- ψ(x,y) is a 2D continuous space sequence
- Forward Transform

$$\Psi(f_x, f_y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi(x,y)\exp(-j2\pi(f_x x + f_y y))\,dx\,dy$$

- Inverse Transform

$$\psi(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Psi(f_x, f_y)\exp(j2\pi(f_x x + f_y y))\,df_x\,df_y$$

One dimensional Fourier Transform For Discrete Time Signal (DTFT)

- $\psi(n)$ is a 1D discrete time sequence
- Forward Transform

$$\Psi(f_x) = \sum_{n=-\infty}^{\infty} \psi(n) \exp(-j2\pi f_x n)$$

- Inverse Transform

$$\psi(n) = \int_{-1/2}^{1/2} \Psi(f_x) \exp(j2\pi f_x n) \, df_x$$

Two dimensional Fourier transform for Discrete Space signal (DSFT)

- ψ(m,n) is a 2D discrete space sequence
- Forward Transform

$$\Psi(f_x, f_y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \psi(m,n) \exp(-j2\pi(f_x m + f_y n))$$

- Inverse Transform

$$\psi(m,n) = \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} \Psi(f_x, f_y) \exp(j2\pi(f_x m + f_y n) df_x \, df_y$$

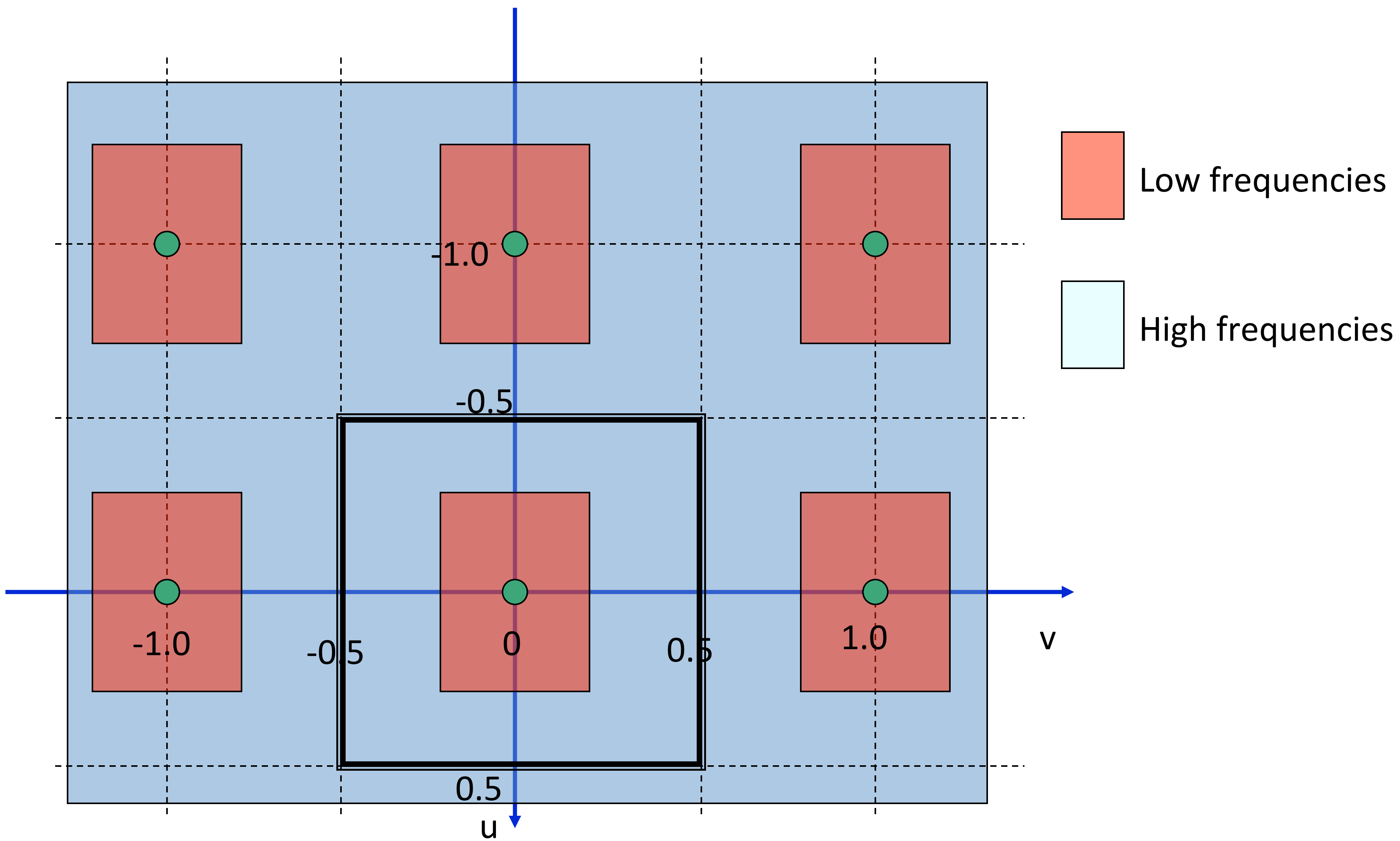
- Properties
  - Linearity, Periodicity, Energy Conservation

Periodicity of DSFT

$$\Psi(f_x, f_y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \psi(m,n) \exp(-j2\pi(f_x m + f_y n))$$

- $\Psi(f_x, f_y)$ is periodic in f_x, f_y with period 1, i.e., for all integers *k*, *l*:
  - $\Psi(f_x + k, f_y + l) = \Psi(f_x, f_y)$

Illustration of Periodicity



Fourier analysis and Human Visual system Frequency response: Outline

- Characterizing video signals in the frequency domain
- Frequency response of the HVS
- A bit about interlaced video

Frequency domain characterization of video signals

- Spatial frequency
- Temporal frequency
- Temporal frequency caused by motion

Spatial frequency (Matlab)

```
fx=5; fy=10;
[x,y]=meshgrid(0:.002:1,0:.002:1);
z=sin(2*fx*pi*x+2*fy*pi*y);
imtool(z,[])
imtool (abs(fftshift(fft2(z))),[])
```

Spatial Frequency

- Spatial frequency measures how fast the image intensity changes in the image plane
- Spatial frequency can be completely characterized by the variation frequencies in any two orthogonal directions (e.g horizontal and vertical)
  - f_x: cycles/horizontal unit distance
  - f_y : cycles/vertical unit distance
- It can also be specified by magnitude and angle of change $f_s = \sqrt{f_x^2 + f_y^2}, \quad \varphi = \arctan(f_y / f_x)$

Spatial frequency

- Sinusoidal signal can be defined by (f_x,f_y)
- Maximum spatial frequency present is defined by

$$f_s = \sqrt{f_x^2 + f_y^2}, \quad \varphi = \arctan(f_y \,/\, f_x)$$

Example of Spatial Frequency

$(f_x, f_y) = (5,0)$ $(f_x, f_y) = (5,10)$

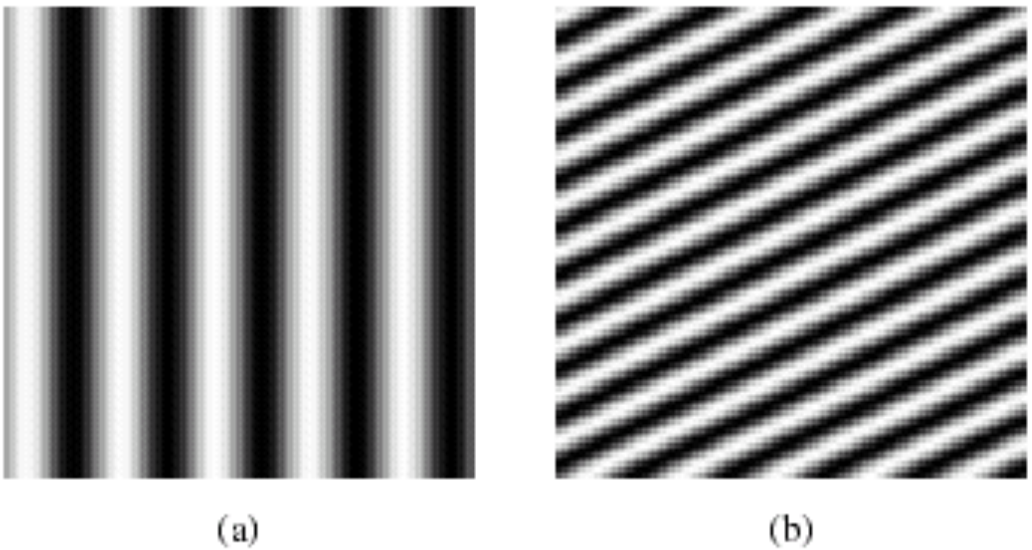


Figure 2.1 Two-dimensional sinusoidal signals: (a) $(f_x, f_y) = (5, 0)$; (b) $(f_x, f_y) = (5, 10)$. The horizontal and vertical units are the width and height of the image, respectively. Therefore, $f_x = 5$ means that there are five cycles along each row.

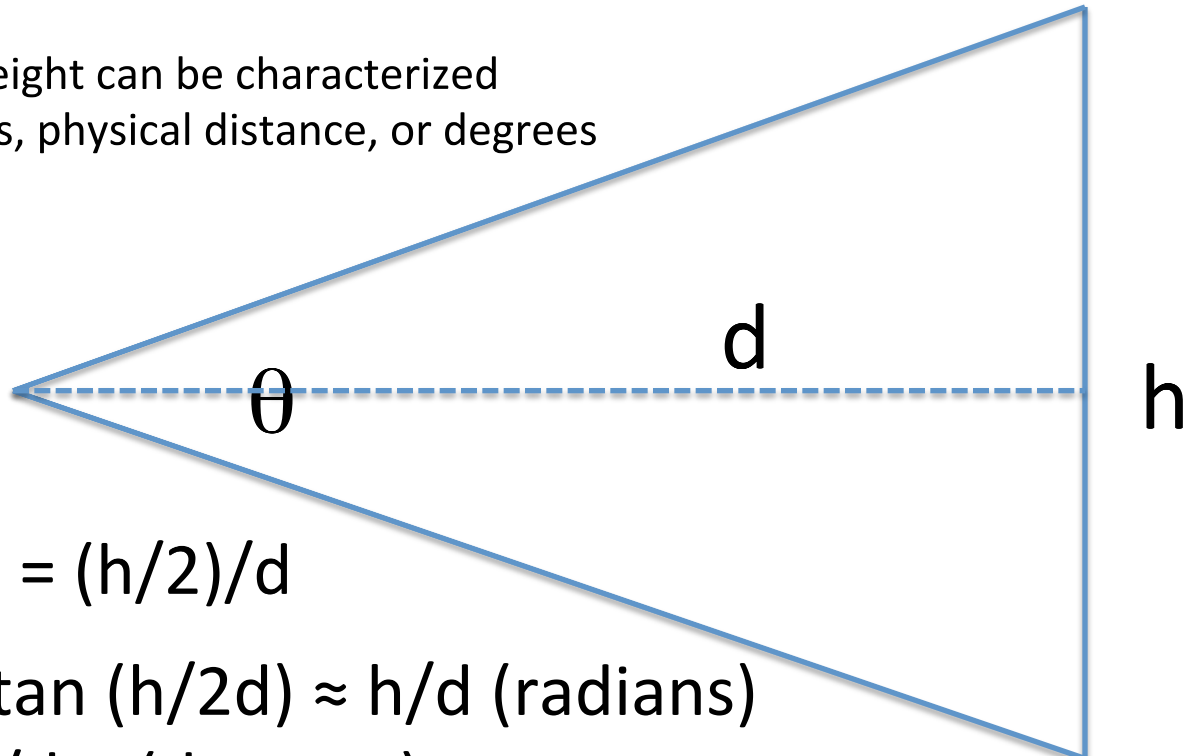
$$f_s = \sqrt{125}, \quad \varphi = \arctan(2)$$

Units of spatial frequency

- Cycles per image
- Cycles per pixel
- Cycles per degree

Angular frequency

The picture height can be characterized
in either pixels, physical distance, or degrees



$\tan \theta/2 = (h/2)/d$

$\theta = 2 \arctan (h/2d) \approx h/d$ (radians)

$\theta = 180h/d\pi$ (degrees)

f_s cycles per picture height into $f_\theta \approx f_s/\theta$ cycles per degree

$f_\theta = \pi d f_s/180h$

Why angular frequency?

Temporal Frequency

- Temporal frequency measures temporal variation (cycles/s or Hz)
- In a video, the *temporal frequency* is actually 2-dimensional; each point in space has its own temporal frequency
- Non-zero temporal frequency can be caused by camera or object motion

- Start simple: single object with constant velocity

Temporal frequency due to constant translational motion

- Still image $\psi_0(x,y)$ is translated from frame to frame
- Video $\psi(x,y,t)=\psi_0(x-v_xt, y-v_yt)$
- Fourier Transform
  $\Psi(f_x,f_y,f_t)=\delta(f_xv_x + f_yv_y + f_t)\ \Psi_0(f_x,f_y)$
- Non-zero ONLY on the plane
  $f_xv_x + f_yv_y + f_t=0$
- Relationship: $f_t= - f_xv_x - f_yv_y$
- Temporal frequency depends on motion as well as content (i.e. spatial frequency of object)

Relationship: temporal frequency and motion and spatial content

- Relationship: $f_t = - f_x v_x - f_y v_y$
- Temporal frequency depends on motion as well as content (i.e. spatial frequency of object)
- Temporal frequency is the projection of the velocity vector (v_x, v_y) onto the spatial gradient (defined by (f_x, f_y))
  - Moving sinusoid has $f_t=0$ if moving orthogonal to its spatial variations
  - If $f_x=f_y=0$, then $f_t=0$ even if there's motion

Illustration of the Relationship

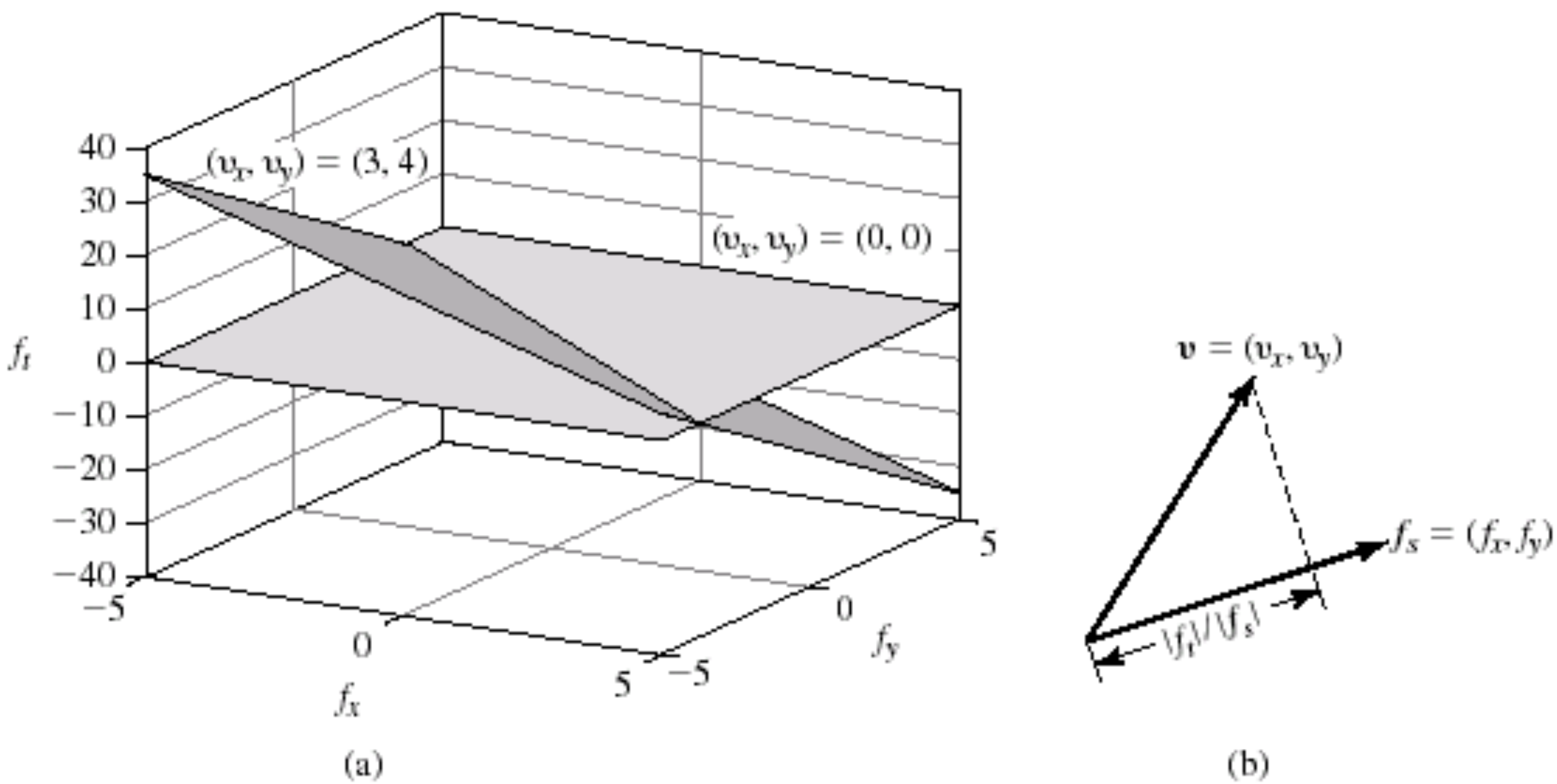


Figure 2.4 Relation between spatial and temporal frequencies under linear motions. (a) The spatiotemporal frequency plane in the (f_x, f_y, f_t) space, corresponding to two different velocity vectors; (b) the temporal frequencies is equal to the projection of the velocity onto the spatial gradient.

Interpretations

- If there's no spatial variation, temporal frequency is zero and the apparent motion is zero
- If the still image has spatial bandwidth $(f_{x,max},f_{y,max})$, and there's constant motion then the temporal bandwidth is
  $f_{t,max}= v_x f_{x,max}+ v_y f_{y,max}$

How do humans perceive frequency content?

Contrast sensitivity function

- Ability of an observer to see sine wave gratings at different frequencies
- Sensitivity is 1/(threshold)
- Many different definitions of contrast
  - Difference / average
  - Weber: $(I-I_b)/I_b$
  - Michelson: $(I_{max}-I_{min})/(I_{max}+I_{min})$
  - RMS: $$\sqrt{\frac{1}{MN}\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}(I_{ij}-\bar{I})^2},$$

Frequency response of the HVS

- Temporal frequency response and flicker

$$\psi(t) = B(1 + m\cos 2\pi f t)$$

- Spatial frequency response

$$\psi(x, y, t) = B(1 + m\cos 2\pi f x)$$

- Spatio-temporal response

$$\psi(x, y, t) = B(1 + m\cos 2\pi f_x x)\cos(2\pi f_t t)$$

- Smooth pursuit eye movement

Temporal frequency response

- Flat screen, brightness varies sinusoidally
- For fixed mean brightness B, frequency f, determine lowest modulation level m for which temporal variation is visible (Kelly '61)

$$\psi(t) = B(1 + m\cos 2\pi f t)$$

- Sensitivity is $1/m_{min}$
- Depends on: viewing distance, display brightness, ambient lighting

Temporal sensitivity

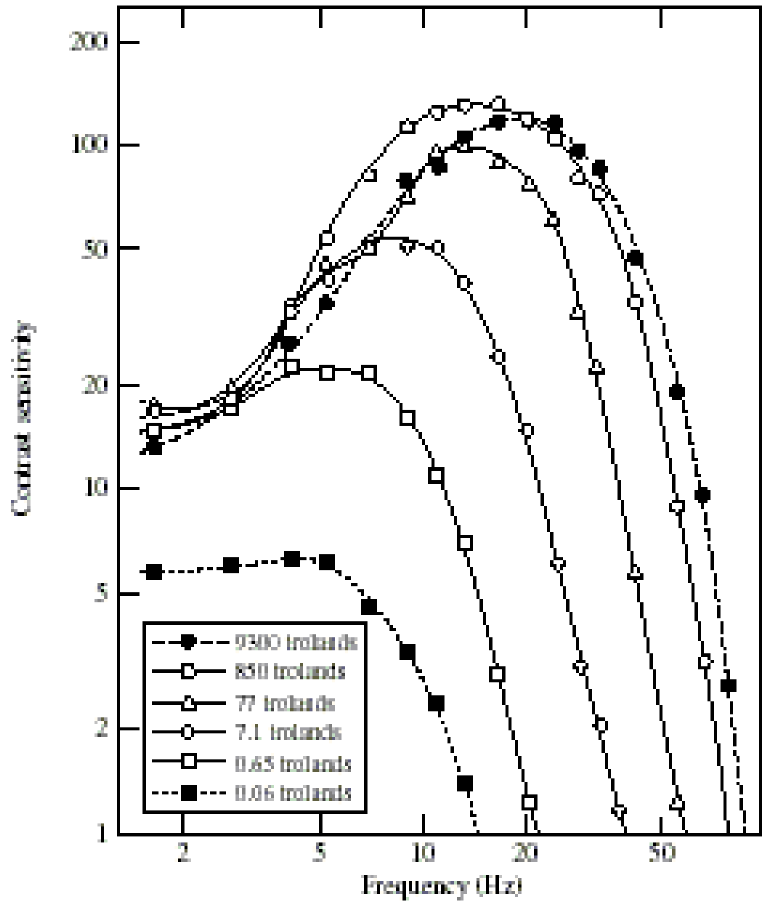


Figure 2.5 The temporal frequency response of the HVS obtained by a visual experiment. Different curves represent the responses obtained with different mean brightness levels, B, measured in trolands. The horizontal axis represents the flicker frequency f, measured in Hz. Reprinted from D. H. Kelly, Visual responses to time-dependent stimuli. I. Amplitude sensitivity measurements, *J. Opt. Soc. Am.* (1961) 51:422–29, by permission of the Optical Society of America.

$$\psi(t) = B(1 + m\cos 2\pi f t)$$

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.

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

60 Hz is typically sufficient for watching TV.

Watching a movie needs lower frame rate than TV

Computer needs higher frame rate

Spatial Response

$$\psi(t) = B(1 + m\cos 2\pi f x)$$

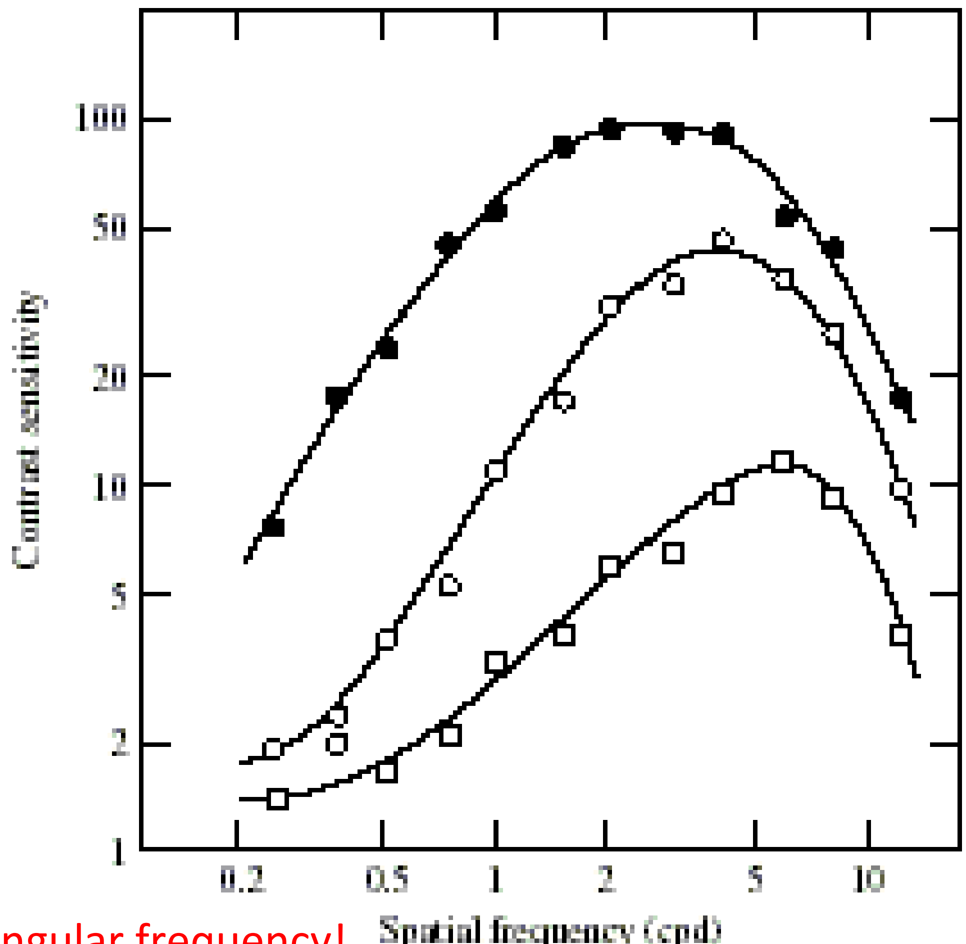
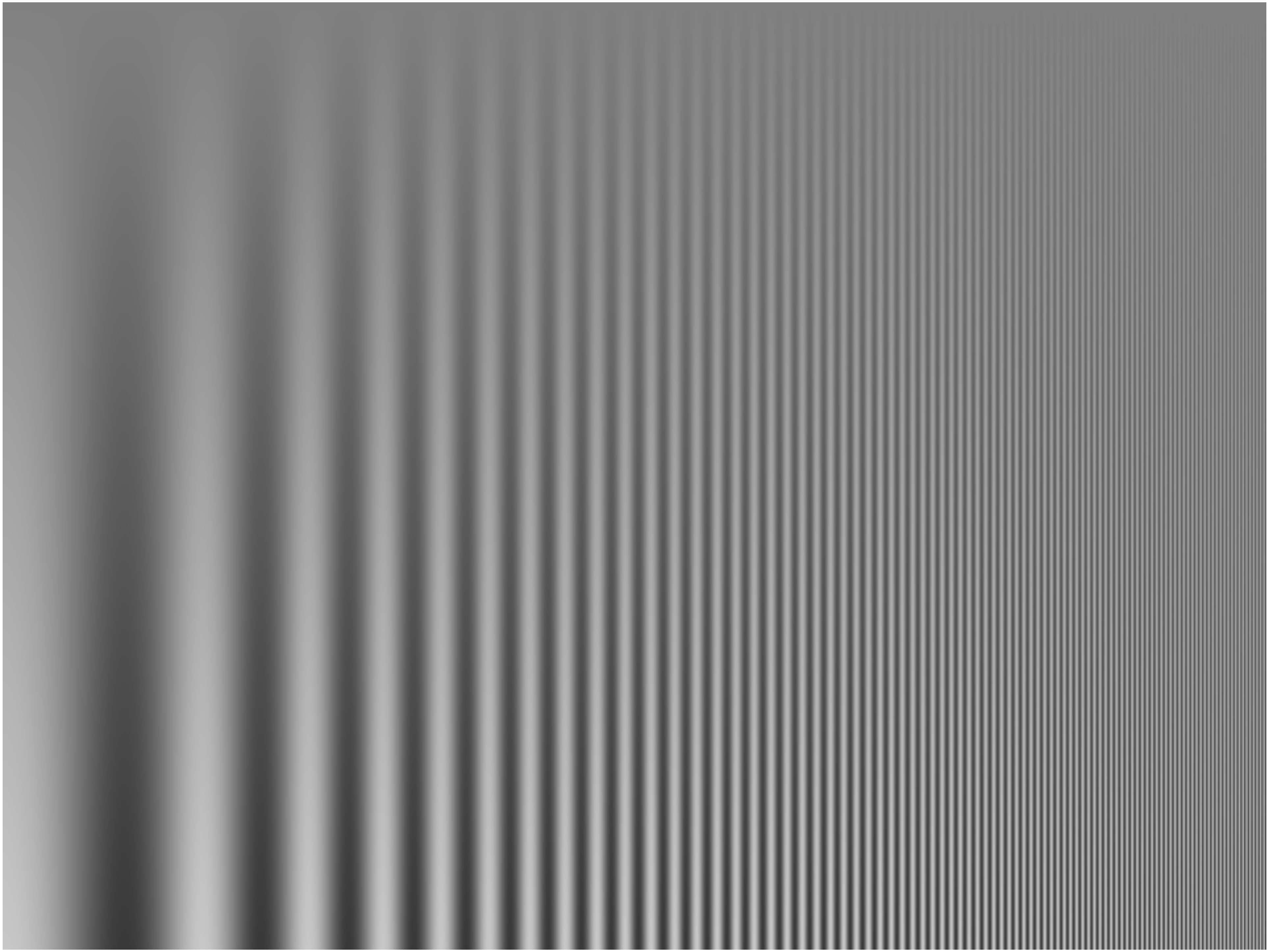
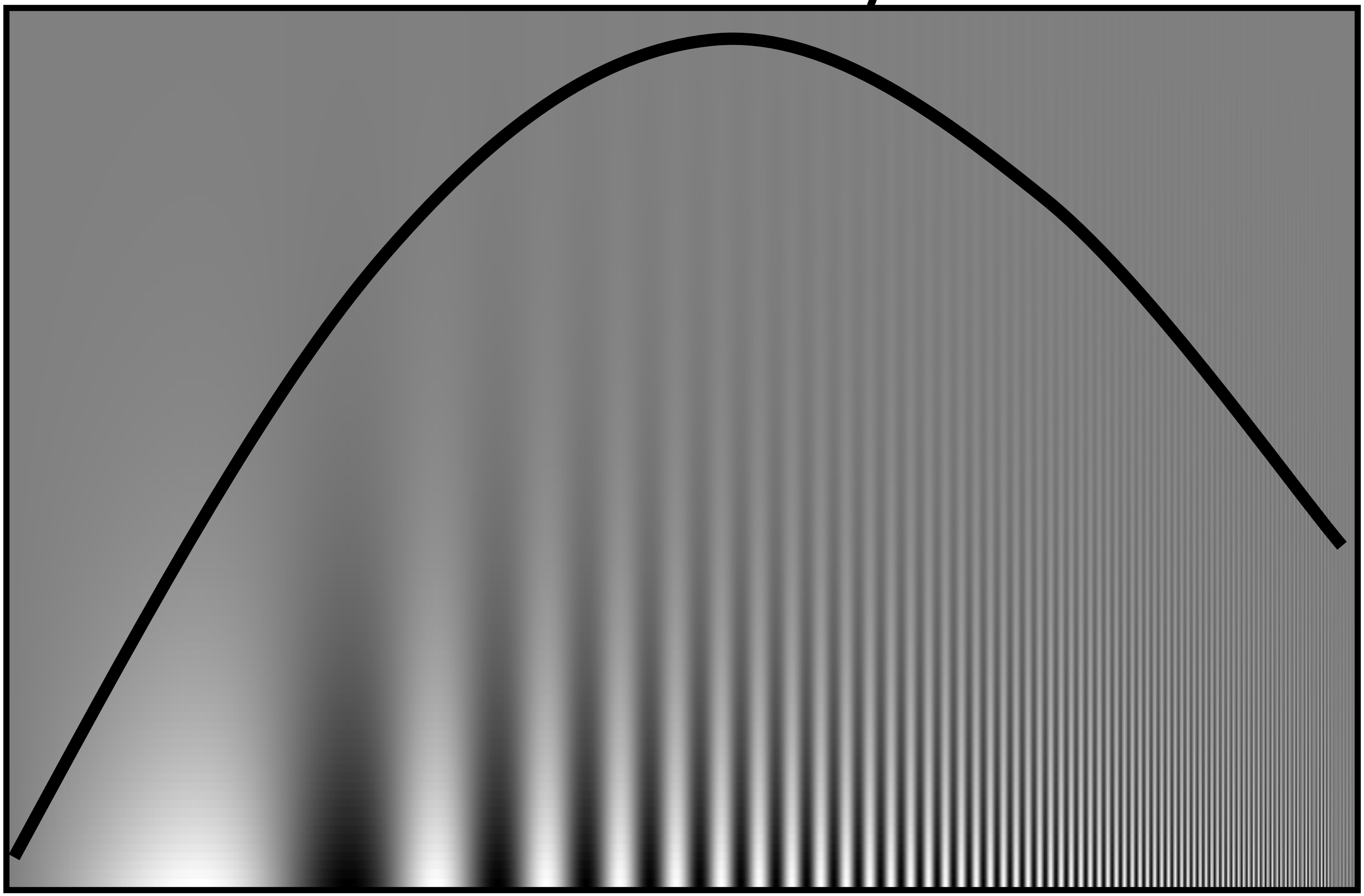


Figure 2.6 The spatial frequency response of the HVS, obtained by a visual experiment. The three curves result from different stabilization settings used to remove the effect of saccadic eye movements. Filled circles were obtained under normal, unstablized conditions; open squares, with optimal gain 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.

Angular frequency!



Contrast Sensitivity Curve



Matlab to create spatial CSF example

- [xx,mm]=meshgrid((0:.008:8)/2/pi,0:.1:100);
- ff=ones(1001,1)*(.5:.1/70:3);
- ff=ff(:,1:1001);
- z=128+mm.*cos(2*pi*xx.* 10.^(ff));
- imtool (z,[0 255],'InitialMagnification',100)
- imwrite(uint8(clip2(round(z), 0,255)),'csfillust.png')

A model for spatial CSF

- 2.6(0.0192+0.114f) exp(-(0.114f)^1.1)
- Due to Mannos and Sakrison
- J. L. Mannos, D. J. Sakrison, "The Effects of a Visual Fidelity Criterion on the Encoding of Images", *IEEE Transactions on Information Theory*, pp. 525-535, Vol. 20, No. 4, 1974

Frequency response of the HVS

- Temporal frequency response and flicker

$$\psi(t) = B(1 + m\cos 2\pi f t)$$

- Spatial frequency response

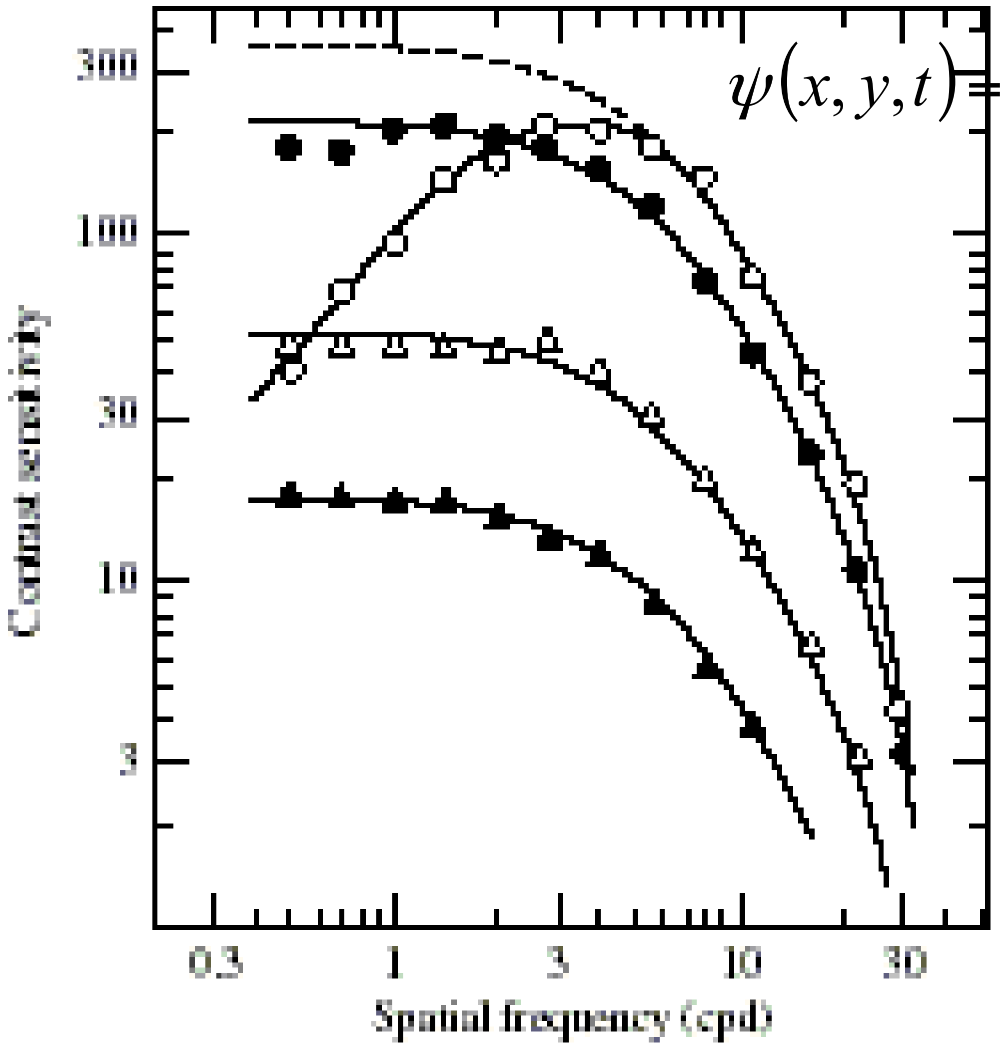
$$\psi(x, y, t) = B(1 + m\cos 2\pi f x)$$

- Spatio-temporal response

$$\psi(x, y, t) = B(1 + m\cos 2\pi f_x x)\cos(2\pi f_t t)$$

- Smooth pursuit eye movement

Spatiotemporal frequency response of HVS as function of spatial frequency

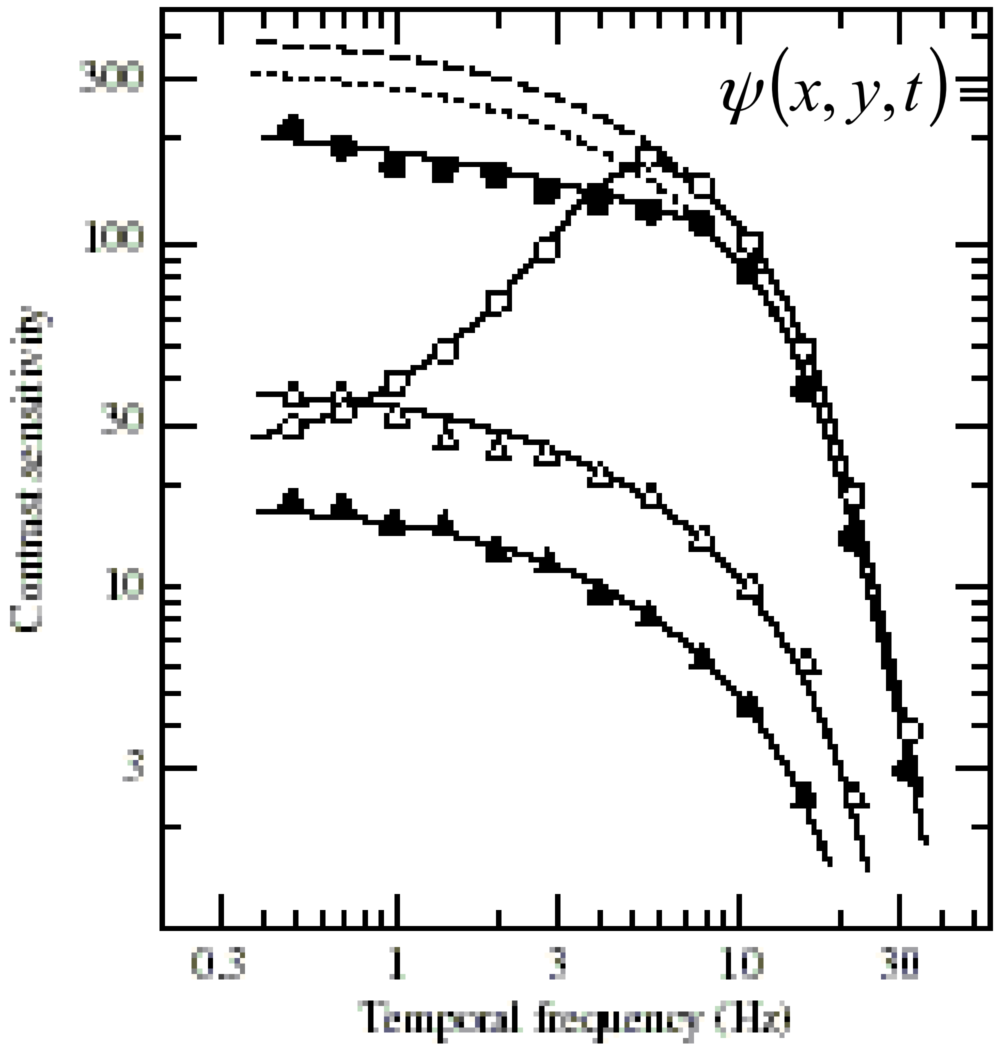


$$\psi(x,y,t) = B(1 + m\cos 2\pi f_x x)\cos(2\pi f_t t)$$

White circles: 1 Hz
Black circles: 6 Hz
White triangles: 16 Hz
Black triangles: 22 Hz

From J.G. Robson, Spatial and temporal contrast sensitivity functions of the visual system, J. Opt. Soc. Am. Vol. 56, pp 1141-42, 1966.

Spatiotemporal frequency response of HVS as function of temporal frequency



$$\psi(x,y,t) \equiv B(1+m\cos 2\pi f_x x)\cos(2\pi f_t t)$$

White circles: 0.5 cpd
Black circles: 4 cpd
White triangles: 16 cpd
Black triangles: 22 cpd

From J.G. Robson, Spatial and temporal contrast sensitivity functions of the visual system, J. Opt. Soc. Am. Vol. 56, pp 1141-42, 1966.

Observations about spatiotemporal CSF

- With nearly zero temporal (spatial) frequencies, the spatial (temporal) frequency response is band-pass (as seen before)
- At higher temporal frequencies, the spatial frequency response becomes more low-pass
  - High-frequency image pattern moving fast (on the retina) is difficult to see details
  - Same image pattern is easier to see when object is still (on the retina)

Observations about spatiotemporal CSF (2)

- At higher spatial frequencies, the temporal frequency response becomes more low-pass
- Line flicker: seen when viewing TV close
  - Lines further apart spatially creates lower angular frequencies
  - HVS is better at perceiving the temporal variations in signals with lower angular frequencies (relative to those with higher angular frequencies)

Observations about interlace

- A good trade-off between temporal and spatial resolution FOR A FIXED NUMBER OF SAMPLES
- Can display *still* objects with high spatial resolution
- Can display moving objects with “visually-necessary” resolution (HVS cannot see high spatial frequencies in fast moving scene EVEN IF rendered at high spatial resolution)

More observations about interlace

- For any temporal sampling, aliasing can be introduced when displaying objects with high spatial frequencies that are moving quickly
- To limit aliasing, spatial filtering is necessary
- Stronger spatial filtering is necessary for interlaced video than progressive video
- Interlace was an engineering solution that we no longer need

Smooth Pursuit Eye Movement

- The eye tracks moving objects
- Reduces velocity of moving objects on the retinal plane
- Eye can perceive much higher temporal frequencies, up to 1000 Hz

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 $(\tilde{v}_x, \tilde{v}_y)$:

$$\tilde{f}_t = f_t + (\tilde{v}_x f_x + \tilde{v}_y f_y)$$

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

Problems to practice (1)

- Consider a horizontal bar pattern on a TV screen with 100 cycles/picture-height. If the picture-height is 1 meter, and the viewer sits 3 meters from the screen, what is the equivalent angular frequency in cycles per degree? What if the viewer sits 1 meter or 5 meters away. In either case, would the viewer be able to perceive the vertical variation properly?

Problems to practice (2)

- Consider an object that has a flat homogeneously textured surface with maximum spatial frequency of $(f_x, f_y) = (3,4)$ cycles/meter, and that is moving at a constant speed of $(v_x, v_y) = (1,1)$ meters/second. What is the temporal frequency of the object surface at any point? What are the results for the following speeds (in meters per second): (4,-3), (4,0), (0,1)?

Problems to practice (3)

- Continuing, Suppose that the eye tracks the moving object with a speed that is equal to the object speed. What are the perceived temporal frequencies at the retina for the different speeds? What if the eye moves at a fixed speed of (2,2) meters/second?