Image Artifacts in the Digital Era:

The Psychophysics of Artifact Visibility

SMPTE 2008:Annual Technology Conference & Expo
Hollywood, CA

Wednesday October 29, 2008

James Larimer, Ph.D.
ImageMetrics, LLC
www.imagemetrics.com
jim@imagemetrics.com
(650) 678-0658

For the sake of persons of different types of mind, scientific truth should be presented in different forms and should be regarded as equally scientific whether it appears in the robust form and vivid colouring of a physical illustration or in the tenuity and paleness of a symbolic expression.

James Clerk Maxwell, from The Scientific Papers of J. Clerk Maxwell

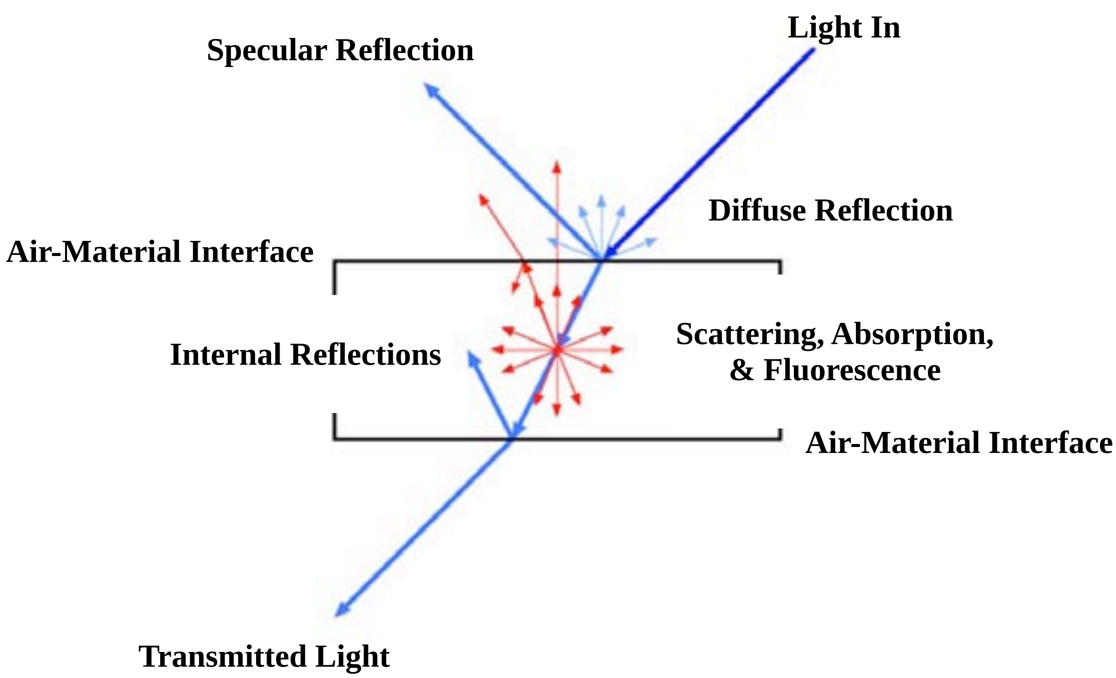
Image Origins

or

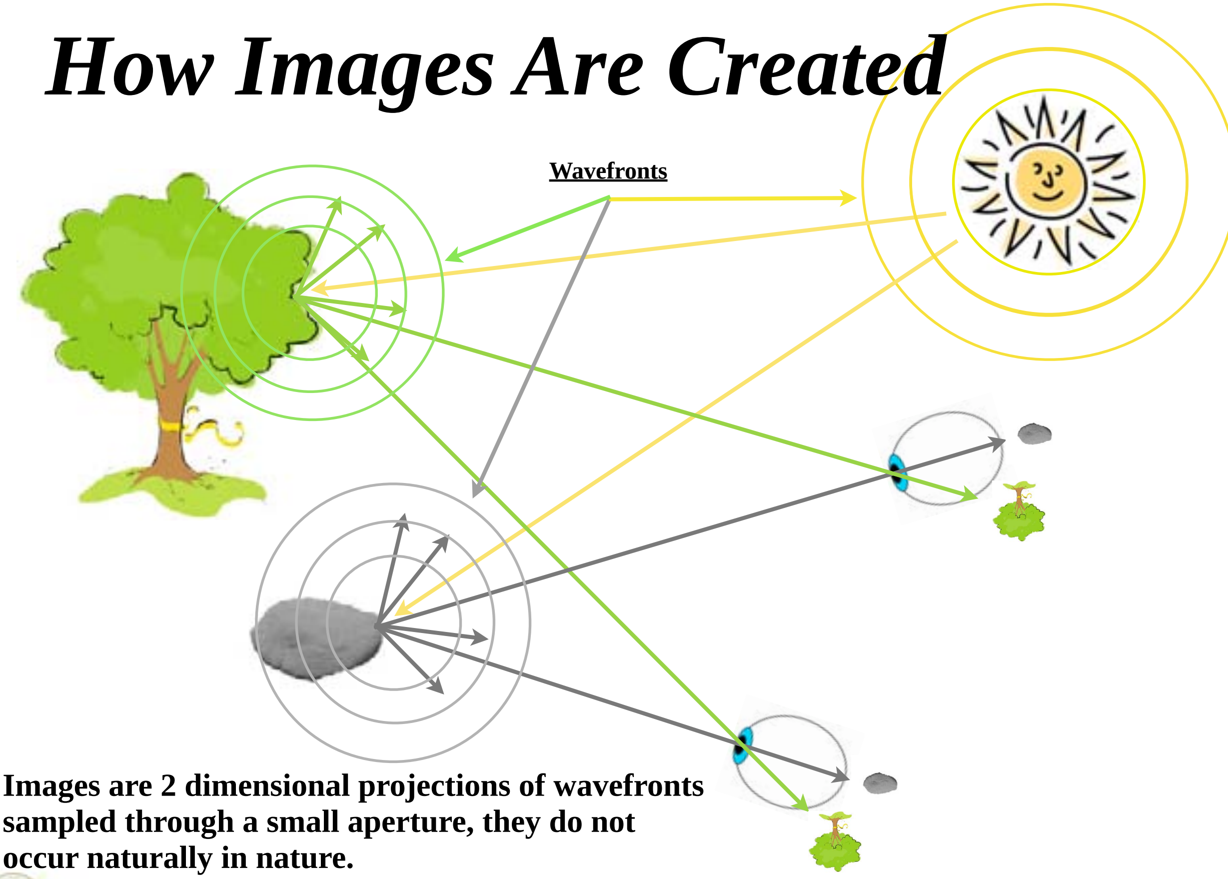
What Is An Image: Natural Phenomena or Evolutionary Invention?

www.imagemetrics.com
(650) 678-0658

Light's Interaction with Materials Creates The Signals for Brightness and Contrast



How Images Are Created



Images are 2 dimensional projections of wavefronts sampled through a small aperture, they do not occur naturally in nature.

SMPTE 2008: Annual Tech Conference & Expo
Oct. 28-30, 2008 • Symposium Oct. 27, 2008
Hollywood, CA

© J. Larimer
www.imagemetrics.com
(650) 678-0658

Images do not naturally in nature only wavefronts exist. Images are formed by a lossy sampling process either through an aperture or via an array of directional detectors. A great deal of information is embedded in the wavefront and only a small fraction is sampled by the eye.

Information Embedded in Multiple Wavefronts

Object Localization:

- **Object relative location**
- **Object heading**
- **Object distance - time of flight, parallax**

Object Surface, Material, & General Properties:

- **Uniformity, Shape, & Size (object attributes)**
- **Albedo (reflectivity)**
  - **surface orientation - polarization**
  - **differential wavelength reflectivity - color**
  - **directional albedo (BDRF), different intensity in different reflection directions**

www.imagemetrics.com
(650) 678-0658

Bi-directional reflectance function.

The Four Domains of Imagery

1. **Intensity (amplitude)**
   - (a) **Hue - color**
   - (b) **Saturation - purity of color**
   - (c) **Brightness - gradient**
2. **Spatial**
3. **Temporal**
4. **Direction & Distance (Holographic/3D)**

In this lecture the fourth domain will not be discussed, but it is the most primitive and earliest of visual sensitivities to develop within the Darwinian evolution of life on earth. All known forms of eyes and visual systems evolved approximately 580 to 530 million years ago, man evolved 1 million years ago.

Artifact Categories by Domain

- ***Intensity domain:***
  - **Metamerism & gamut clipping - color changes due to inadequate color capture and/or reconstruction. (The original colors cannot be reconstructed.)**
  - **Discrete color replacement (possibly for visual effect)**
  - **Banding artifacts - under-sampled tone scale; insufficient levels**
  - **Contrast clipping, noise limits (sensor dead-band, gray blacks, Weber fraction)**
- ***Temporal domain:***
  - **Flicker**
  - **Judder**
  - **Judder induced edge flicker (smooth pursuit eye movements)**
  - **Motion Blur (camera motion and/or smooth pursuit eye movements)**
  - **Color breakup (saccadic eye movements)**
  - **Tone scale flicker & aliasing (a feature of PWM, i.e., time division)**
- ***Spatial Domain:***
  - **Jaggies**
  - **Dither**
  - **Blur - halation**
  - **Depth of field - blur in z direction**
  - **Spatial resampling - aliasing**
  - **Codex - Compression artifacts**

© J. Larimer
www.imagemetrics.com
(650) 678-0658

The topic of color & metamerism is a whole lecture, so we are only going to mention it briefly. There are 5 meanings to metamerism: (1) physcially different lights can appear to have the same color, (2) lighting changes can make reflected surfaces that appear identical in one light appear different in another illuminant, (3) on screen metamerism where two lights that are identical according to the CIE standard observer look different on different parts of the screen, this is a problem with multi-primary displays where p>3, and (4) off screen matemerism where a color match according to the CIE standard observer between a on-screen and off-screen object do not match, and (5) individual differences between observers where a flesh tone may look bluish to one observer and not to another.

The movie Pleasantville is an example of non-linear (discontinuous) color manipulation.

The dark values and bright values in the scene or signal can be lost at capture or reconstruction, ie. a 500;1 signal rendered on an 80:1 display.

Tone scale flicker is often due to coarse dithering at the bottom of the tonescale where the weber fraction can be large. It looks like scintillation in the dark areas on a plasma TV.

Display Reconstruction Methods & Techniques

www.imagemetrics.com
(650) 678-0658

Methods:

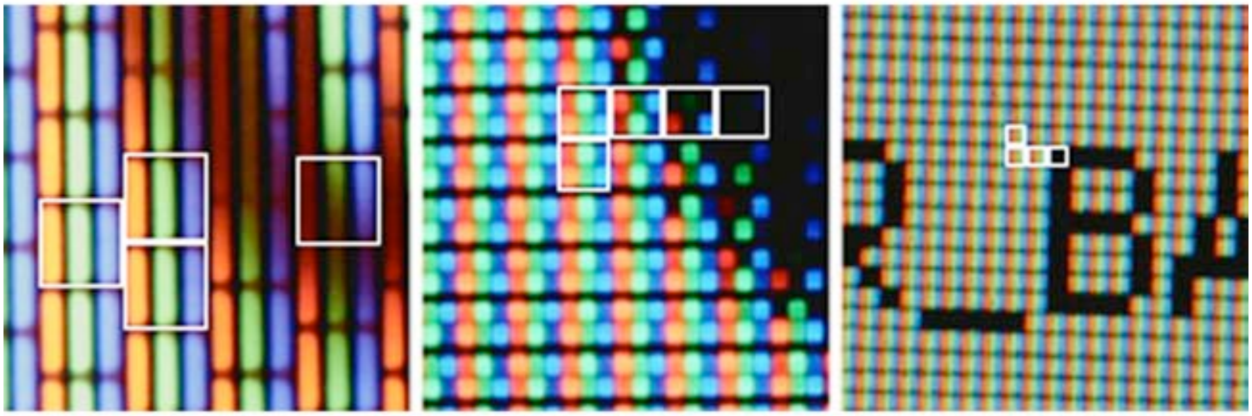
- ***Intensity domain (levels & directions, i.e., intensity is a vector):***
  - **Current Modulation (CRT, OLED, LED) - analog &/or discrete steps (digital)**
  - **Light valve transmissivity (LCD) - analog &/or discrete steps (digital)**
  - **Pulse Width Modulation (DLP, Plasma, OLED, LED) - time division**
  - **Frame sequential Color (multiple primaries, layered time division)**
  - **Dither - in time and/or space**
- ***Temporal domain (frames per second):***
  - **Always discrete and quantized (inherently digital from the beginning of film/ video imagery, i.e., the Lumiere Brothers & Edison)**
  - **Time expansion/contraction by varying the relationship between capture and reconstruction frame rates. For example 96 fps capture reconstructed at 24 fps is 4x time expansion.**
  - **Various reconstruction up-sampling methods (time domain scaling or re-sampling)**
    - **replication - double or triple shuttering, etc.**
    - **fading (convex linear operator, t•F1+(1-t)•F2)**
    - **3/2 pull down, etc.**
    - **motion vector interpolation (non-linear operator)**
- ***Spatial Domain (pixels per picture height):***
  - **Rectangular pixel lattices - discrete (there are soft pixel architectures)**
  - **Stripes, triads, quads, n-tuples - discrete/digital**
  - **Scaling: RxC to R'xC' where R≠R' and/or C≠C' (anamorphic as well as aspect ratio preserving resampling) - analog or digital**

One reconstruction domain may be used to reconstruct the signal in another domain. Spatial-temporal dither is an example of spatial and temporal trade-offs being used to produce an intensity level. Laptop computer screens use both types of dither to produce 24 bit images, essentially giving up some spatial and temporal resolution to enhance the tonescale resolution. PWM is also called “time division”.

Three Pixel Architectures

Delta Triad CRT
defocused e-beam raster painting shadow mask grill

Rectangular Lattice AMLCD
High resolution RGB stripe Laptop TN screen

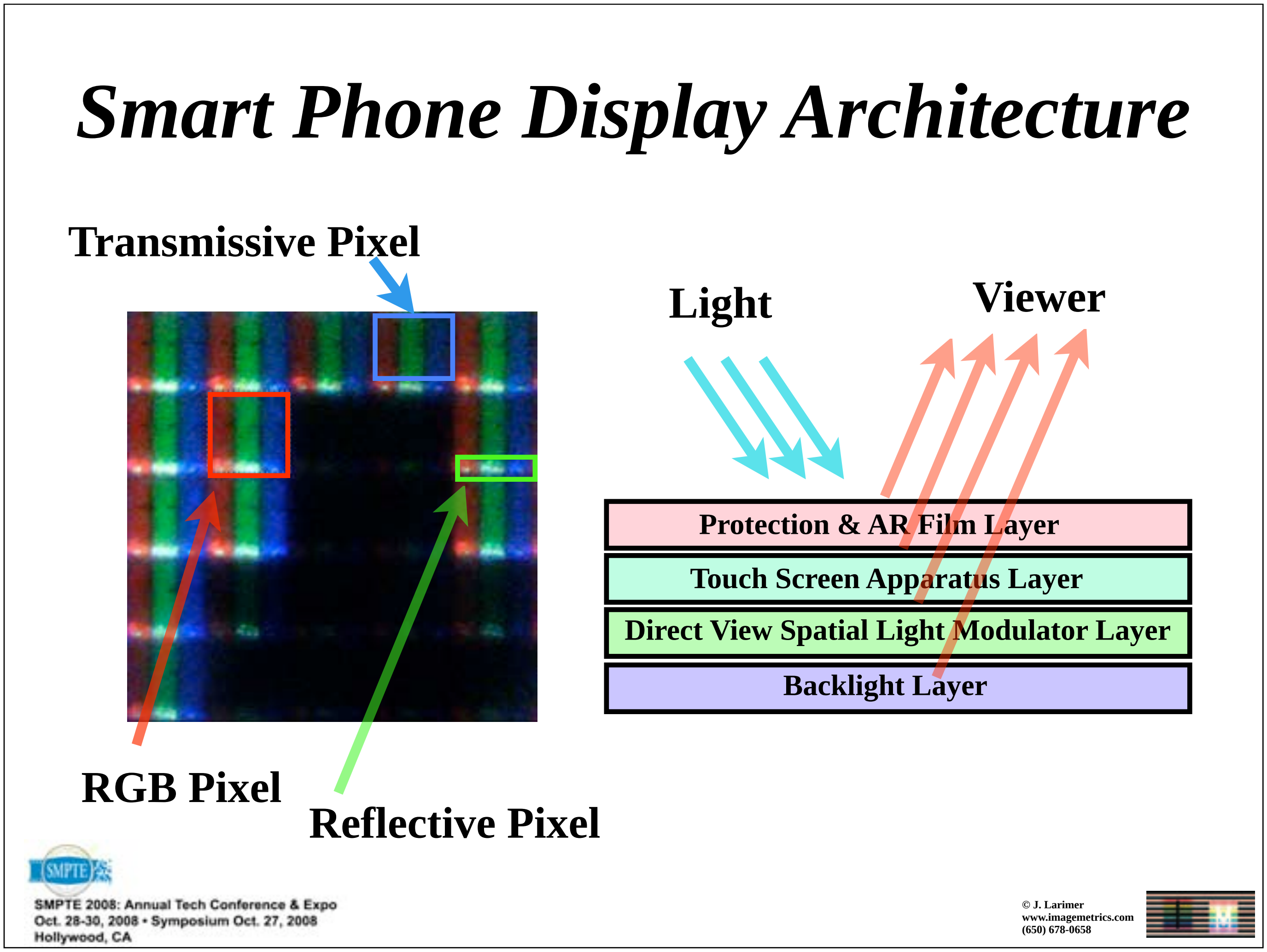


Rectangular Lattice AMLCD
Dithered 6 sub-pixel for anti-aliasing

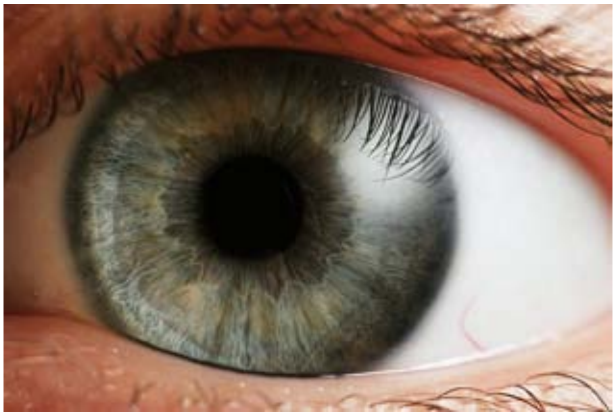
SMPTE 2008: Annual Tech Conference & Expo
Oct. 28-30, 2008 • Symposium Oct. 27, 2008
Hollywood, CA

www.imagemetrics.com
(650) 678-0658

Three standard types of pixel architectures.



This is a hybrid architecture where the goal is performance out of doors in bright ambient and indoors using a back light. The backlight is replaced with a diffuser mirror for 1/4th the display surface area. Internally the liquid crystal layer must be thinned by a factor of 2 to correlate the optical performance in the mirror area and backlight transmissive areas. This requires a multigap design with unique LC modes.



Vision & The Eye Evolved to Analyze Wavefront Information For Darwinian Goals

© J. Larimer
www.imagemetrics.com
(650) 678-0658

Evolution of Vision

- **Vision & sensory physiology emerged during the *Cambrian Explosion* 530 million years ago. Complex life forms evolved during this period.**
- **The Darwinian drivers for sensory capabilities <u>are</u> *Predation & Foraging*.**
- **The image formed on the retina supports extraction of physical object properties and geometries from the reflected light signal, i.e., wavefronts, sampled through the pupil.**
- **The sampling process is 2-dimensional with the 3rd & 4th dimensions extracted by temporal processing. Temporal information includes relative rate of change, i.e., parallax, and closing rates, i.e., object size growth (towards) or shrinking (away), in the retinal image. Vision must separate self from object motion. Motion sickness occurs when vestibular and visual information are uncoupled.**
- ***Seeing* is primarily a cognitive-sensory process - we see objects not images. We remember scenic attributes not scene images. Cameras record patterns of projected light, images, people do not despite claims of eidetic imagers.**
- **A great deal of information is lost by the biological process of seeing.**

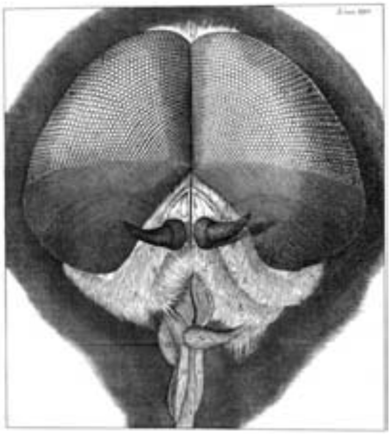
SMPTE 2008: Annual Tech Conference & Expo
Oct. 28-30, 2008 • Symposium Oct. 27, 2008
Hollywood, CA

Prior to the Cambrian, -542 Myrs to -490 Myrs, all life forms were either single celled like the the cyanobacteria. The Ediacaran period proceeded the Cambrian. During this period the Earth’s atmosphere did not contain sufficient oxygen to sustain complex life forms. Eidetic imagery is a controversial topic in the vis. research literature.

Eye Evolution

Two Variants:

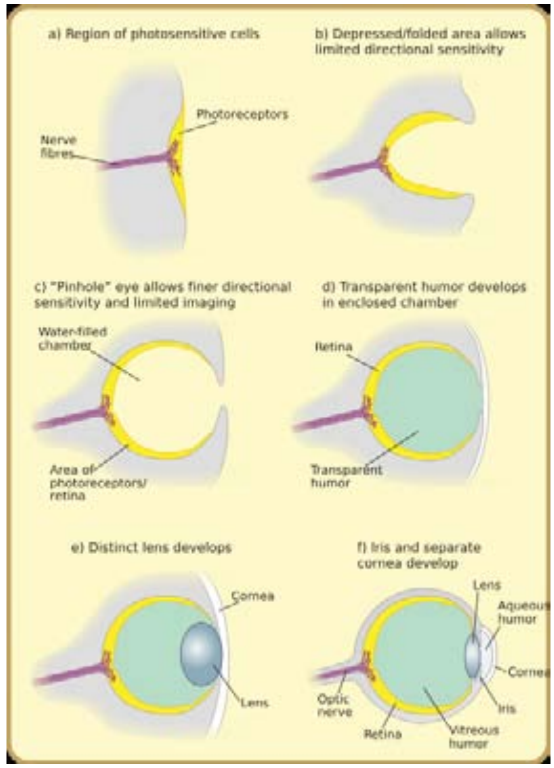
1. Compound Eye



Robert Hooke's 1664 drawing of a grey drone fly compound eye.

- **Directional Sensitivity**
- **Shadow detection**
- **Good motion detection**
- **Poor image quality**

2. Complex Chambered Eye



- **Directional Sensitivity**
- **Shadow detection**
- **Good motion detection**
- **Good image quality**

The two eye types: compound and chambered.

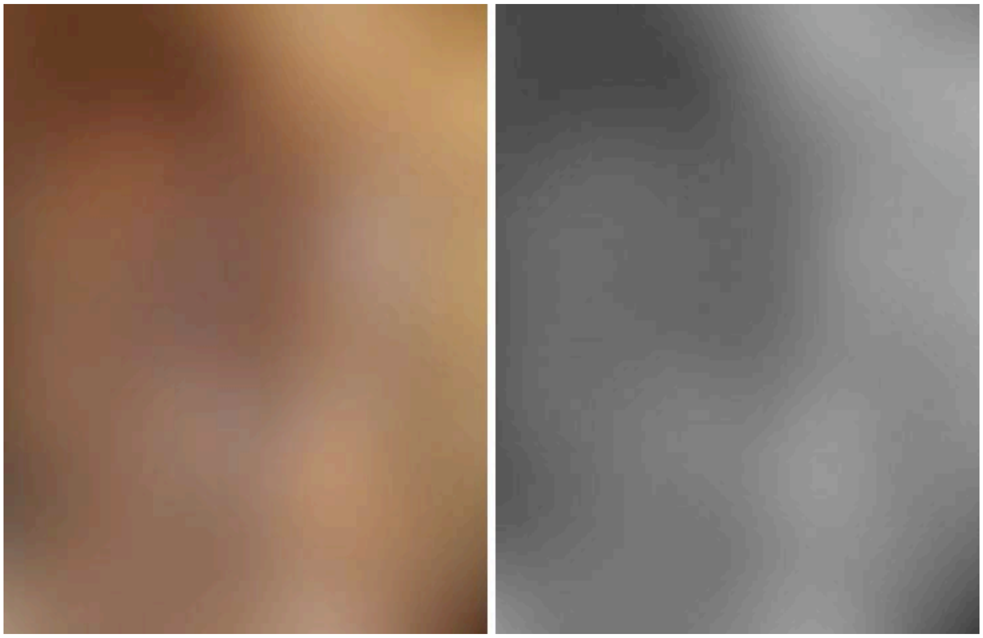
Vision's Darwinian Goals

- Find prey and avoid being prey.
- Navigation - avoid collisions, finding a safe path.
- Detecting danger, breaking camouflage.
- Detecting threats, for example spoiled food, a sick person who might make you sick. This is why flesh tones are so critical to "image quality".
- Spatial detail required for object recognition - samples on target.

www.imagemetrics.com
(650) 678-0658

This is also why perfume, cosmetics, beauty aids has been a part of human culture for all of recorded history.

An Illustration

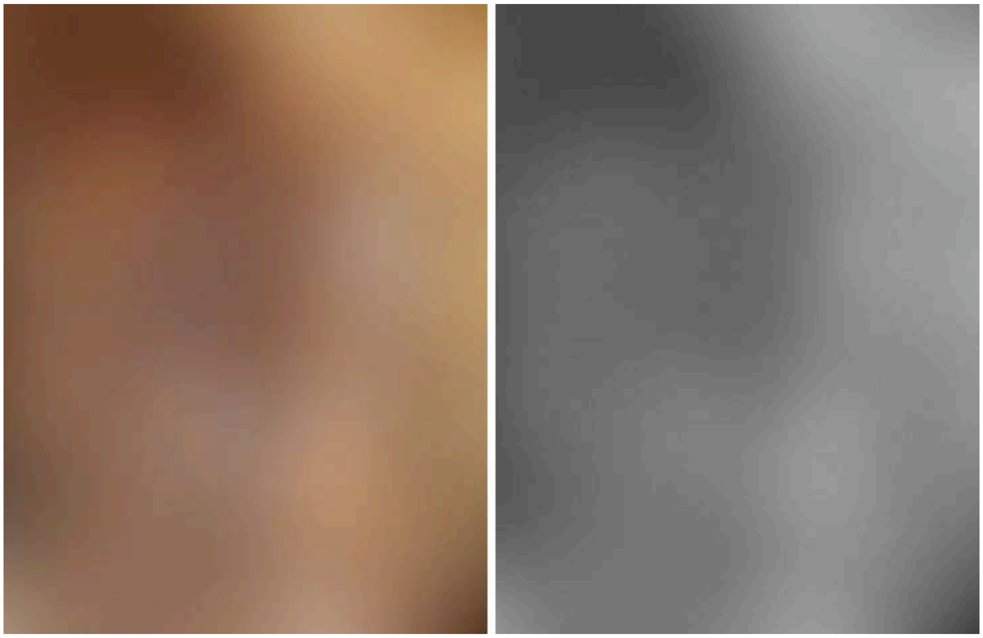


SMPTE 2008: Annual Tech Conference & Expo
Oct. 28-30, 2008 • Symposium Oct. 27, 2008
Hollywood, CA

www.imagemetrics.com
(650) 678-0658

This video demonstrates the value of spatial resolution is apparent in this example.

Notice the Judder & Extra Salience Produced by Color



SMPTE 2008: Annual Tech Conference & Expo
Oct. 28-30, 2008 • Symposium Oct. 27, 2008
Hollywood, CA

© J. Larimer
www.imagemetrics.com
(650) 678-0658

The differential salience of the colored vs black and white imagery is demonstrated in this video. The jerky motion in this slow motion video is called judder.
In the real world the spectral albedo (reflectivity) and motion of the grass and tiger would be additional cues to breaking the camouflage. Our eyes evolved for these tasks.

The Darwinian Value of Color & Edge Definition

The edge contrast in luminance units is the same for both the colored and monochrome squares; color adds salience.

SMPTE 2008: Annual Tech Conference & Expo
Oct. 28-30, 2008 • Symposium Oct. 27, 2008
Hollywood, CA

If you stare into the center of the dark spot on the maximum blur frame of this video demonstration (i.e., a static image not the video) with your gaze held steady, you may experience the “filling-in” phenomena where the green center vanishes. This is due to the effort made to make the green center and red surround equi-luminant (of equal luminance - not brightness). Holding the gaze steady means that any residual luminance gradient resulting from individual differences or the monitor are shifted to a lower temporal frequency. When the edge loses salience the center vanishes and the left side appears 100% reddish without a green center.

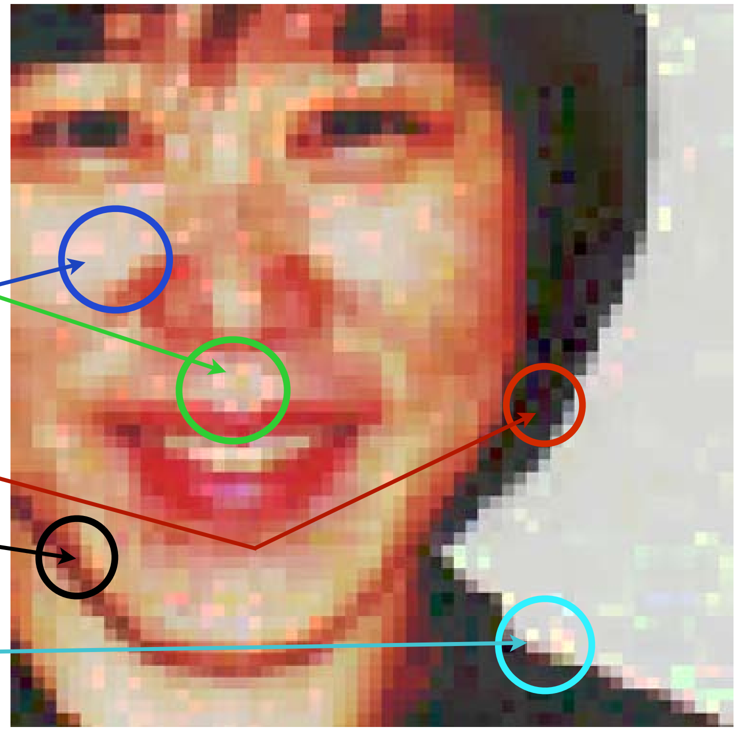
What is an Image Artifact?

- **Sampling is a process that loses information; it can distort the signal and add new signals, artifacts, to the resulting record. Color experience is a biological artifact.**
- **Imaging systems sample at capture, coding, & reconstruction; this process is lossy. Artifacts are signals added by sampling and processing. Color is an artifact produced by the visual nervous system. Stair-stepping, jaggies, screen door, film grain, flicker, and judder are signals added by imaging technologies.**
- **Artifacts are only important when we can see them. Without visual salience they have no consequence. In the artist's hands artifacts may be attributes, properties of a desired "look" in the image product.**
- **Distortion is not lossy, it is <u>not</u> an artifact, distortion includes:**
  - **Blur**
  - **Geometric distortion: keystone, pin cushion, etc.**
  - **Tone scale distortion - gamma (typically monotonic)**
- **Distortion can be reversed or altered by signal processing - up to the noise limit of the system.**

For example if you blur an image to the point where it is a uniform field, obviously all of the distinctions in the image have been lost and cannot be recovered. This happens if the window is finite or the system's amplitude resolution is noise limited. Blurred high spatial frequency information is often lost in the system noise!

Familiar Artifacts

Magnified 10x



Artifacts:

- **Tonescale Under-sampling**
  - **Bayer pattern chroma errors**
  - **Quantization Noise or level binning**
  - **Dynamic Range Compression**
- **Sensor Noise**
  - **Scintillation**
- **Spatial Under-sampling**
  - **jaggies (stair-stepping)**
  - **Blocky appearance**
  - **diagonal edge thickening**
- **JPEG Compression**
  - **Block noise**
  - **Chroma errors**

www.imagemetrics.com
(650) 678-0658

Measuring Artifact Salience

The Human Condition, 1934 Rene Magritte

“In front of a window seen from inside a room, I placed a painting representing exactly that portion of the landscape covered by the painting. Thus, the tree in the picture hid the tree behind it, outside the room. For the spectator, it was both inside the room and within the painting and outside the real landscape. This is how we see the world. We see it outside ourselves, and at the same time we only have a representation of it in ourselves.”

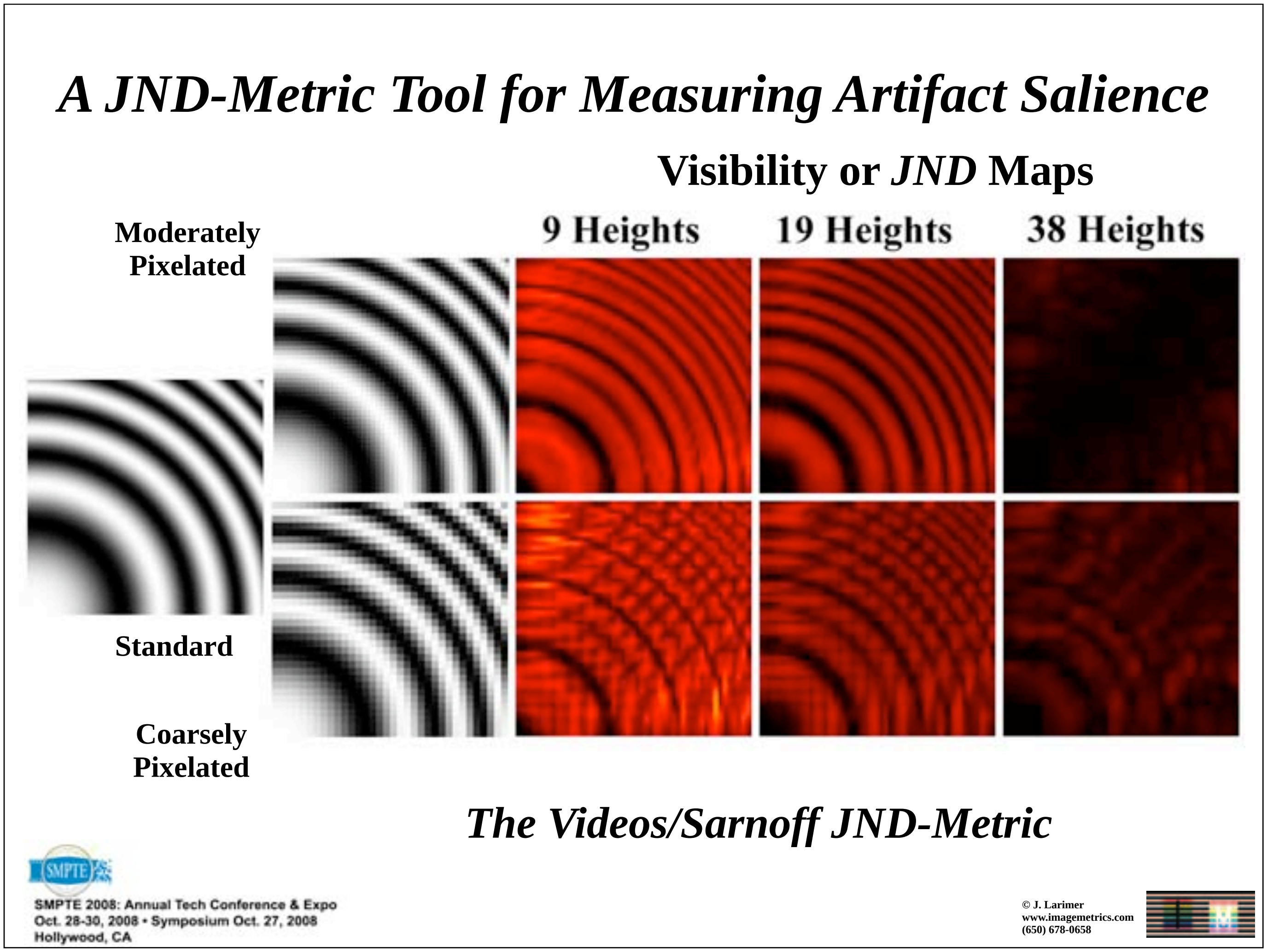
Image quality has at least two meanings: aesthetic and technical. The later must be measured in units of visibility.

Measuring Image Quality

In 1978, Carlson and Cohen [3] proposed a definition for image quality and the perfect image. They suggested that '. . . a "perfect" image is one that looks like a piece of the world viewed through a picture frame'. They went on to propose that the discriminability of the reconstructed image *relative to the actual scene* should be the measure of image quality. We should measure image quality in terms of the human visual system's ability to see differences. A *Just Noticeable Differences* (JND) is defined to be the physical signal difference required to see a difference at or near the threshold of discrimination.

We amend this to read *" . . . relative to the desired image look . . . "*

This definition would make image quality completely objective and quantifiable but requires a reference image as a rendering target. The reference image embodies the aesthetic values of image quality, judging rendering against a standard makes it objective and quantifiable.



The six images on the right are the output of the human vision model and predict the visibility of differences noticeable between either the left most image and the image to the right of it. The bottom row compares the left most image to the lower B&W quarter zoneplate and similarily the top row compares the left most image to the upper B&W quarter zoneplate. These two images are approximately 2 octaves apart in spatial resolution.

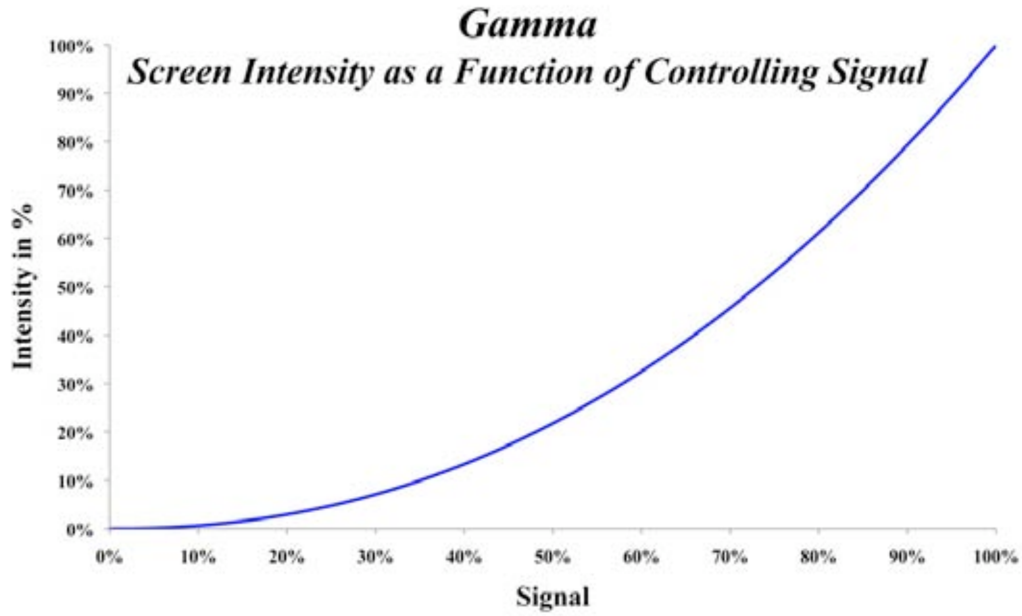
The Goal of the Artist is to Control the Outcome

SMPTE 2008: Annual Tech Conference & Expo
Oct. 28-30, 2008 • Symposium Oct. 27, 2008
Hollywood, CA

Image quality is an artistic aesthetic quantity, it is outside the realm of purely sensory analysis of the image or its bandwidth and information characteristics. Fixing these images to “improve” or optimize the visibility of signal features would destroy the artist’s goals in creating the image.

Sampling Intensity

The Continuous Tone-scale of Film & Analog Television

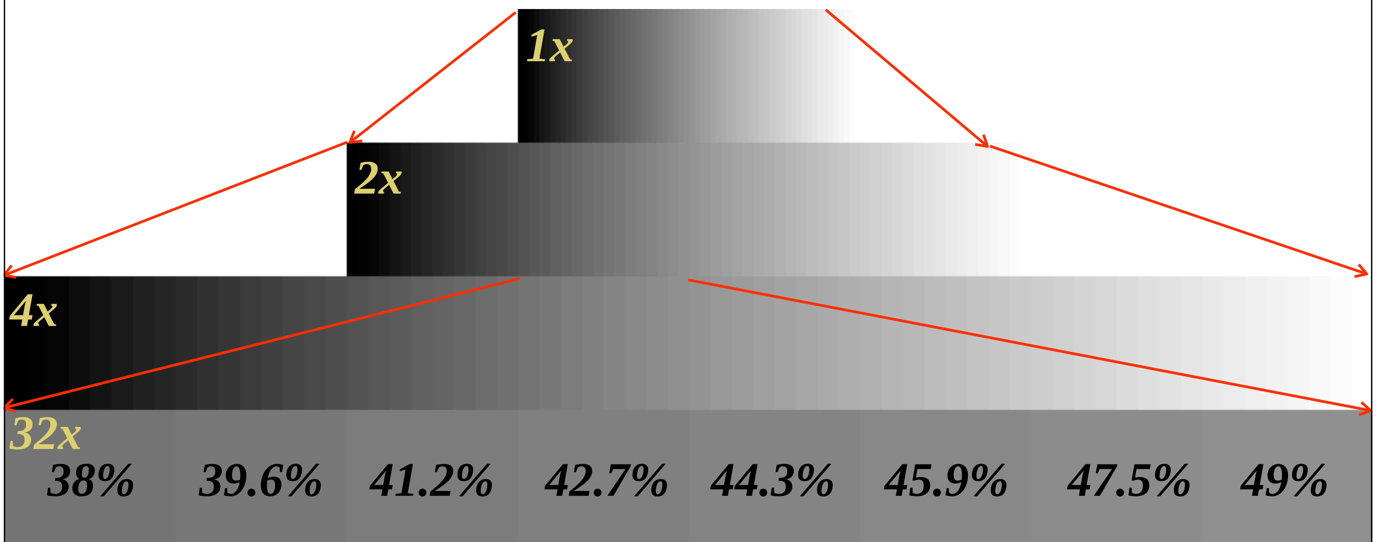


SMPTE 2008: Annual Tech Conference & Expo
Oct. 28-30, 2008 • Symposium Oct. 27, 2008
Hollywood, CA

www.imagemetrics.com
(650) 678-0658

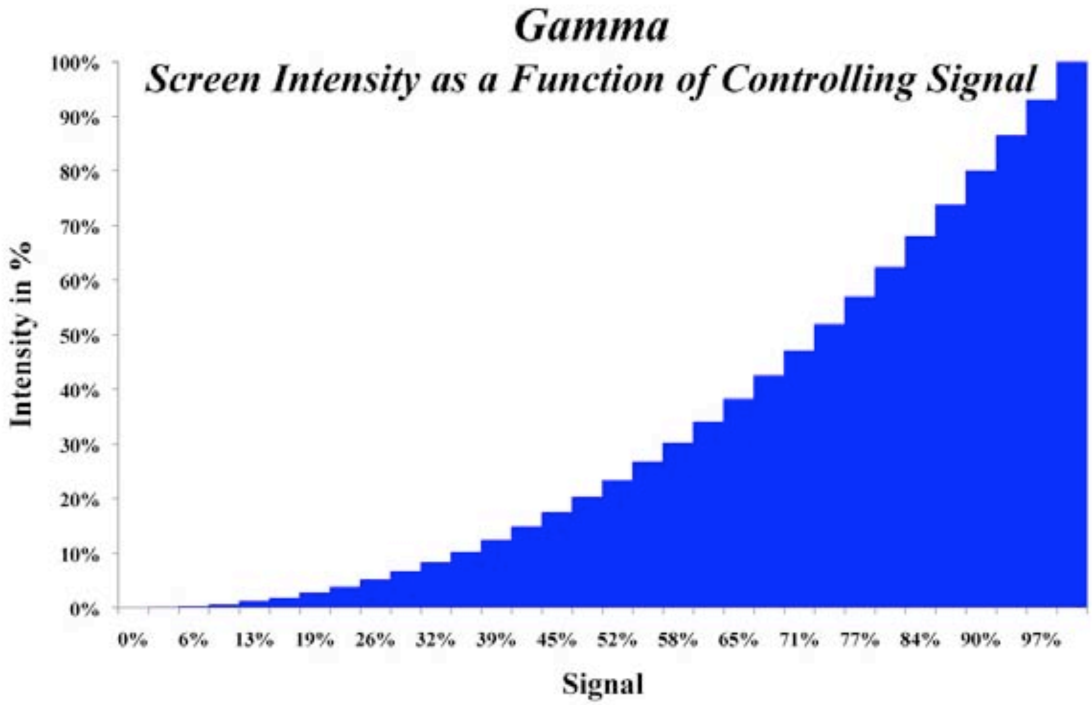
The intensity range in the original scene may have had a peak of 8,000 cd/m^2 whereas the reconstruction yields a range with a peak of only 27.4 cd/m^2 (8 ftL). If the black is not black in the reconstruction, then the contrast can also be “clipped”.

Unlike Film & CRTs Digital Video Is Discrete



www.imagemetrics.com
(650) 678-0658

The Discrete Tone-scale of Digital Cinema & HD-Television



Notice that risers or step sizes at the bottom are smaller than those at the top. Although this is shown as going going to zero, there is a noise floor in the signal and there is always some light even in the black state.

The Continuous Dynamic Range of Wavefront Signals

www.imagemetrics.com
(650) 678-0658

Dynamic Range in Nature

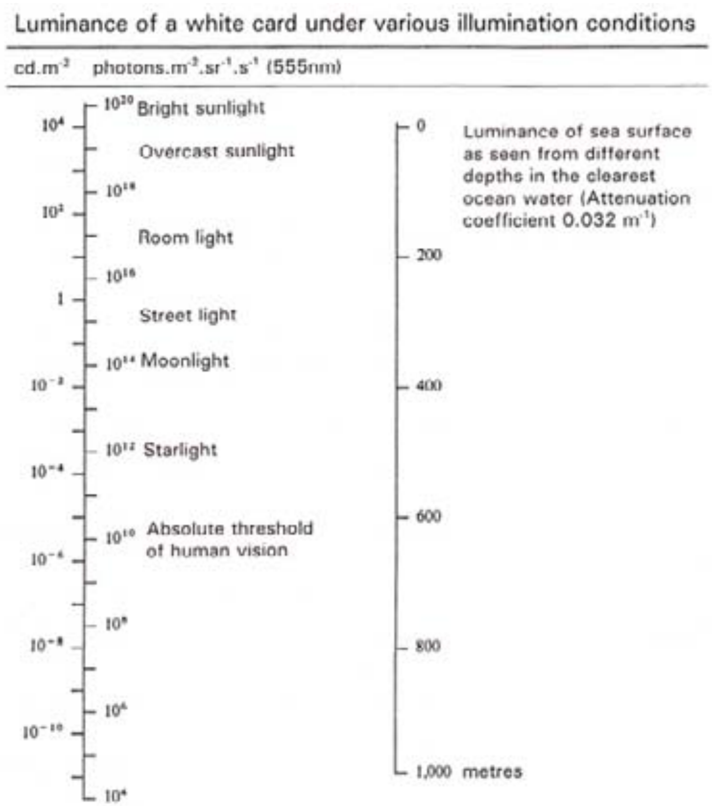


Figure from Animal Eyes by Land & Nilsson

SMPTE 2008: Annual Tech Conference & Expo
Oct. 28-30, 2008 • Symposium Oct. 27, 2008
Hollywood, CA

www.imagemetrics.com
(650) 678-0658

The dynamic range of wavefront amplitudes is several orders of magnitude in natural setting and throughout the diurnal cycle.

Retinal Local Dynamic Range Control

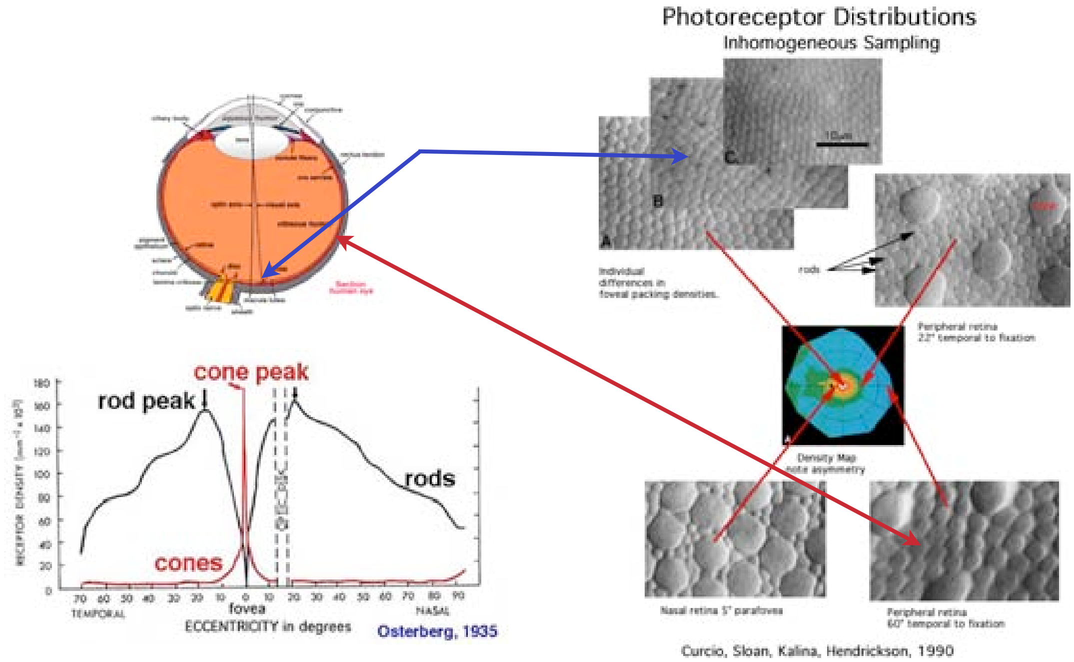
What a Camera Would Record

What a Visual System Would See

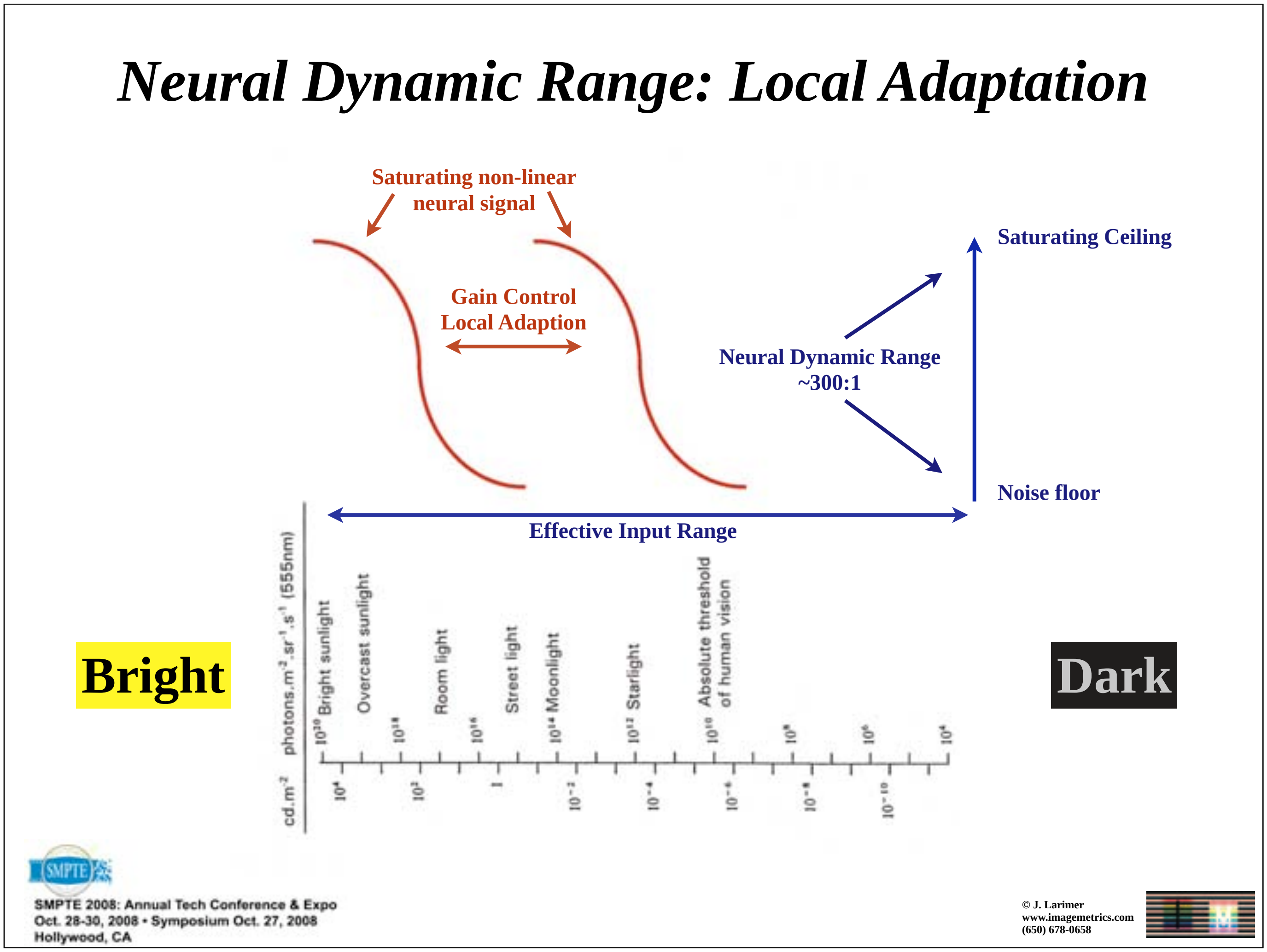
An outdoor scene on a bright day can have 10,000:1 dynamic range.

The image on the right looks more like what we experience. The global dynamic range control in a camera means that signal is lost at the top and bottom of the dynamic range,

Cone Sampling on the Retina

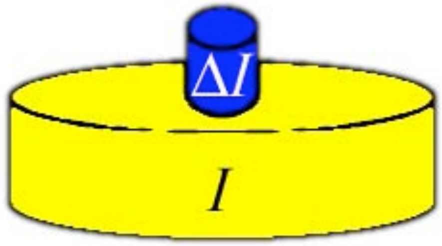


On the right are electron micrographs of human retinas. The sampling mosaic is randomly jittered hexagonal scheme with variable aperture size increasing away from the optical axis on the fovea. Peripheral cone sampling is also sparse.



The visual system's neural pathways have a 300:1 dynamic range, so range compression is part of the sampling story in the eye. A local gain control has evolved in biological vision so two of these functions can operate simultaneously in different regions of the retina, this corresponds to ***dodging and burning*** in photography.

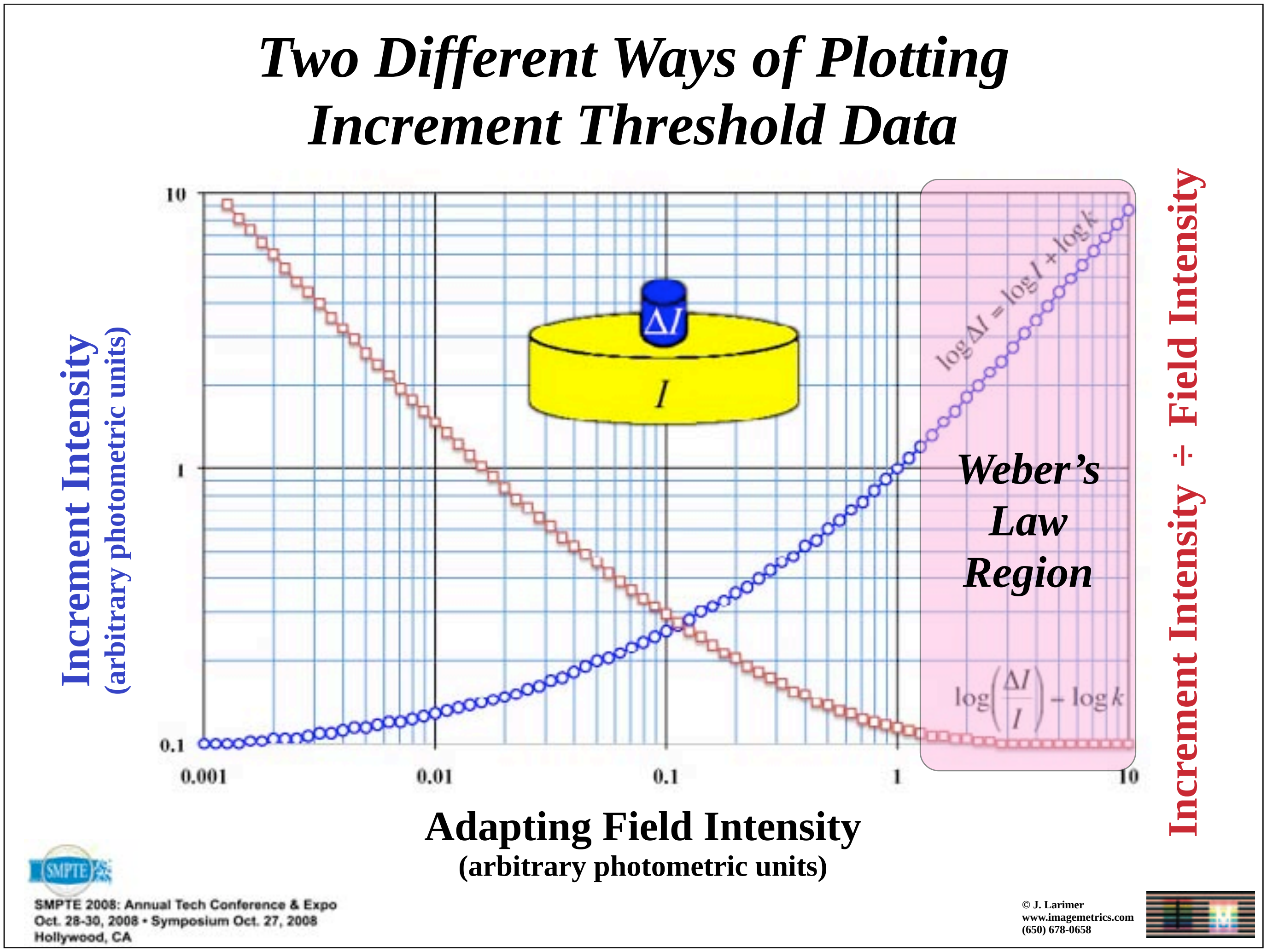
Weber's Law



$$\frac{\Delta I}{I} = k$$

The intensity of an increment seen on a background is a constant fraction of the background intensity. This is why we could plot gain as a shape invariant function translating along a log intensity axis.

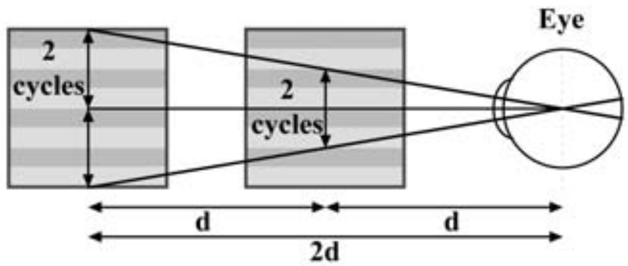
Although Weber's Law is generally the rule, it does not apply uniformly everywhere. There are transition regions where sensitivity is changing, where we are becoming more sensitive. Note that Weber's Law is a measure of sensitivity and the smaller the value of k the more sensitive the system.



The area to the left denoted by the pink region of the graph is called the Weber-Fechner region because Weber's law applies here independently of level. As the DC or average area brightness increases the sensitivity measured as a Weber fraction decreases until it achieves a minimum value. From this DC level and above the Weber fraction does not improve, i.e., reduce in size.

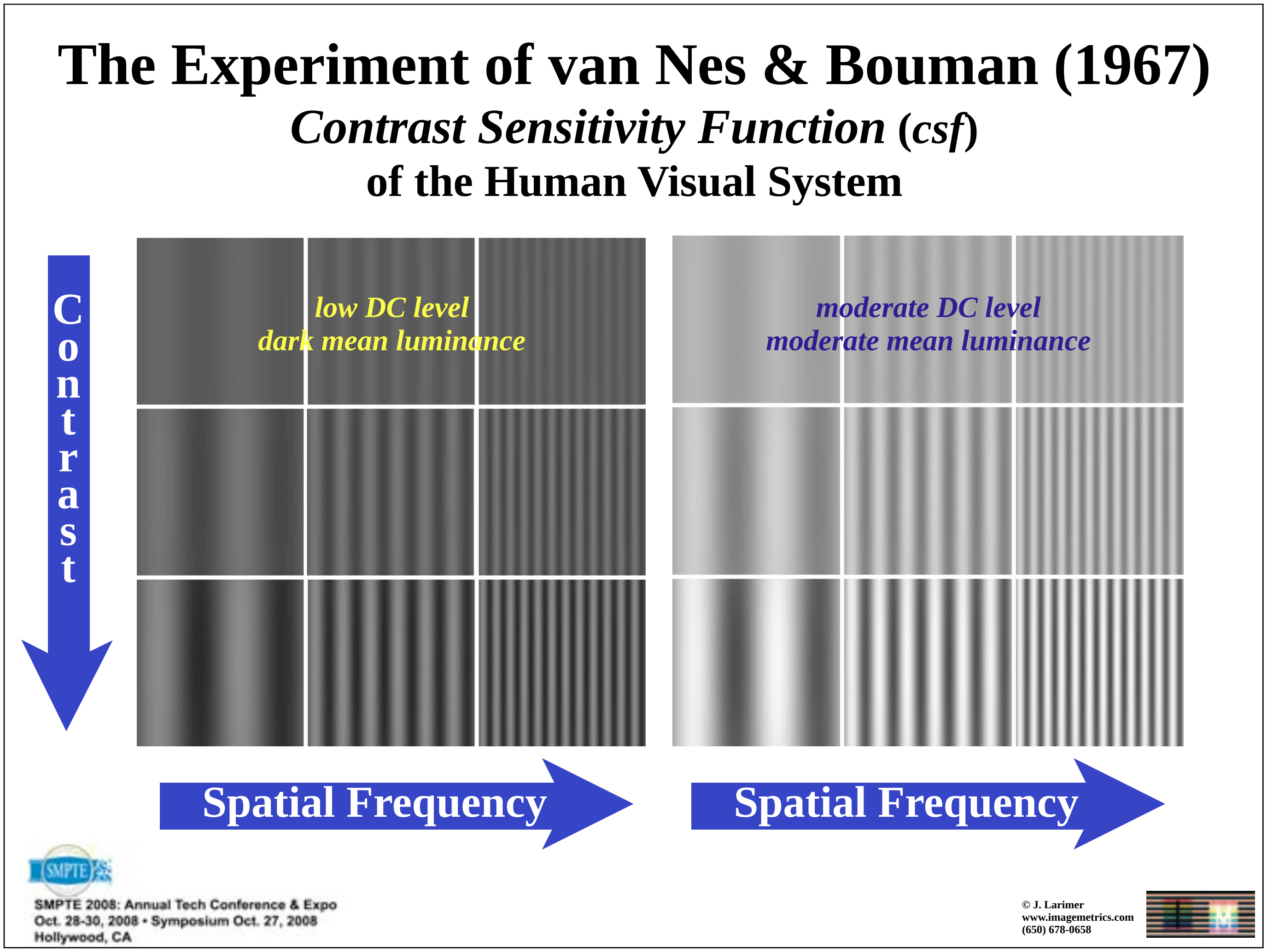
Spatial Tuning of the Visual System: Understanding Intensity Requires Understanding Spatial Sensitivity

The Metrology of Pattern Vision: Measuring Spatial Sensitivity



The eye measures spatial detail in cycles (or line pairs) per visual angle. Resolution measured this way is invariant with distance; doubling the distance to an object, however, reduces its linear size proportional to the change in distance so as shown here 2 cycles per angle becomes 4 cycles per angle at twice the distance.

If I move a surface away from the viewer, the spatial frequency will increase proportionally with the viewing distance and retinal images will shrink proportionally with increasing viewing distance. Hence everything looks the same at some distance. Any visual attribute can be hidden by increasing the viewing distance.



39

Van Nes and Bouman used sinewave gratings of various spatial frequencies and average luminance levels. For each frequency and level they determined the amplitude modulation around the mean luminance required to just detect the bars or grating. This extensive data set was one of the first systematic efforts to characterize the spatial tuning characteristic of the eye. It is one of the cornerstone data sets used to day in computational models like the ViDEOS-Sarnoff human vision model. Models of this sort, called channel models, all rely on spatial-temporal filters designed to replicate standard data sets such as van Nes and Bouman or Kelly. An additional “secret sauce” in this class of models is how they handle masking typically modeled as a saturating non-linearity within the spatial-temporal channels.

Two & One Half Kinds of Contrast

1. Simple Contrast:

$$C_{Simple} = \frac{L_{\max}}{L_{\min}}$$

2. Michelson Contrast:

$$C_{Michelson} = \frac{L_{\max} - L_{\min}}{L_{\max} + L_{\min}}$$

2.5 Image Feature Contrast:

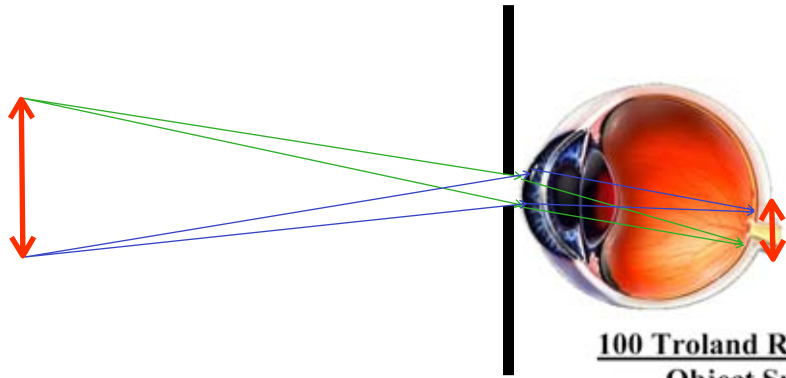
$$C_{ImageFeature} = \frac{EdgeDifference}{LocalAverage}$$

SMPTE 2008: Annual Tech Conference & Expo
Oct. 28-30, 2008 • Symposium Oct. 27, 2008
Hollywood, CA

www.imagemetrics.com
(650) 678-0658

Notice that Michelson Contrast and what I am calling Image Feature Contrast is very similar in functional form to the Weber fraction k.

The Troland: A Unit of Retinal Image Intensity Independent of the Pupil



Td=L•p

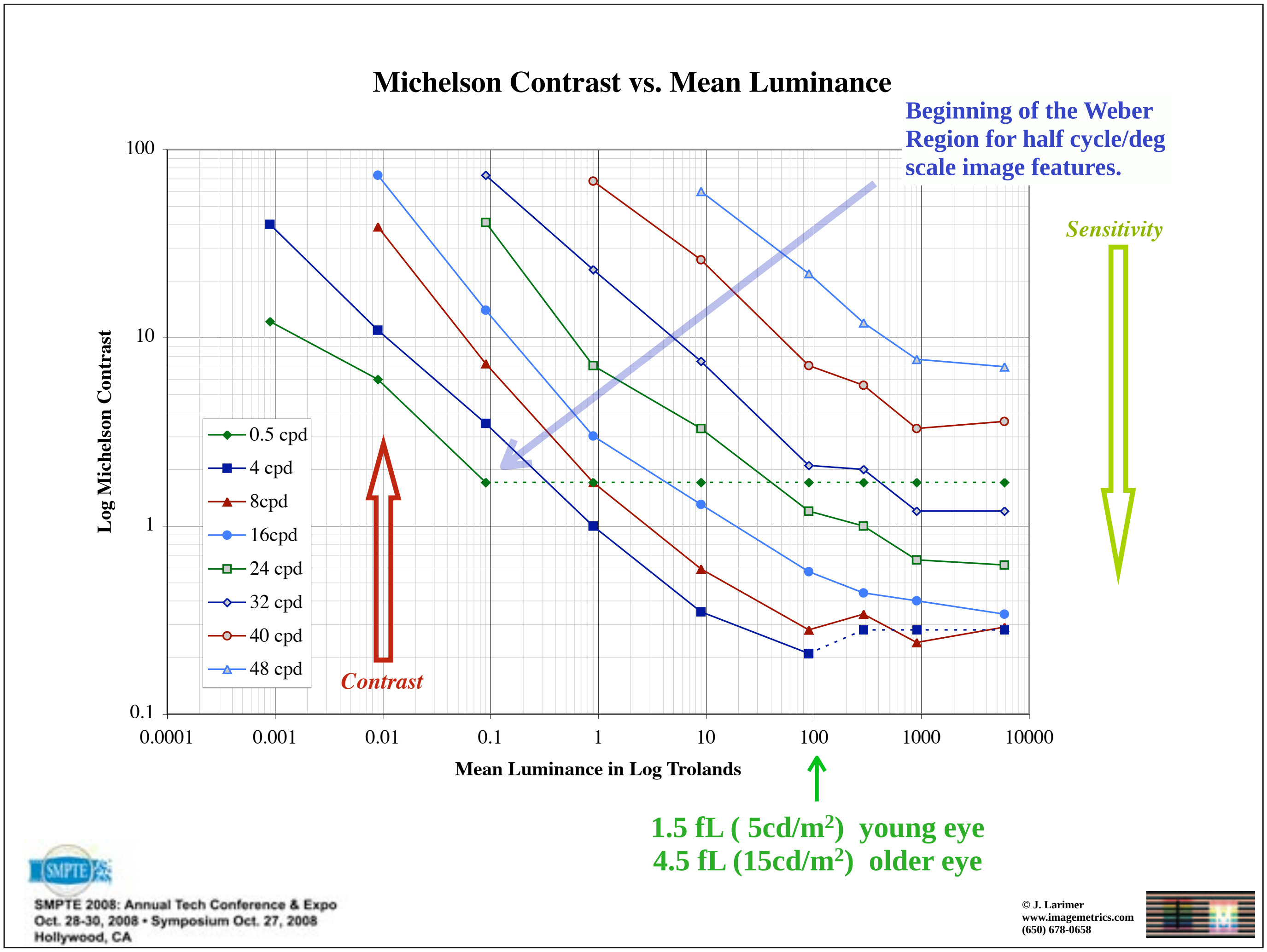
Td = Trolands
L= luminance in cd/m^2
p = pupil area in mm^2

An *<u>artificial pupil</u>* restricts amount of light entering the eye and if smaller than the pupil eliminates the pupil's ability to modulate retinal image intensity.

<u>100 Troland Retinal Luminance</u> in Equivalent Object Surface Luminance in cd/m^2

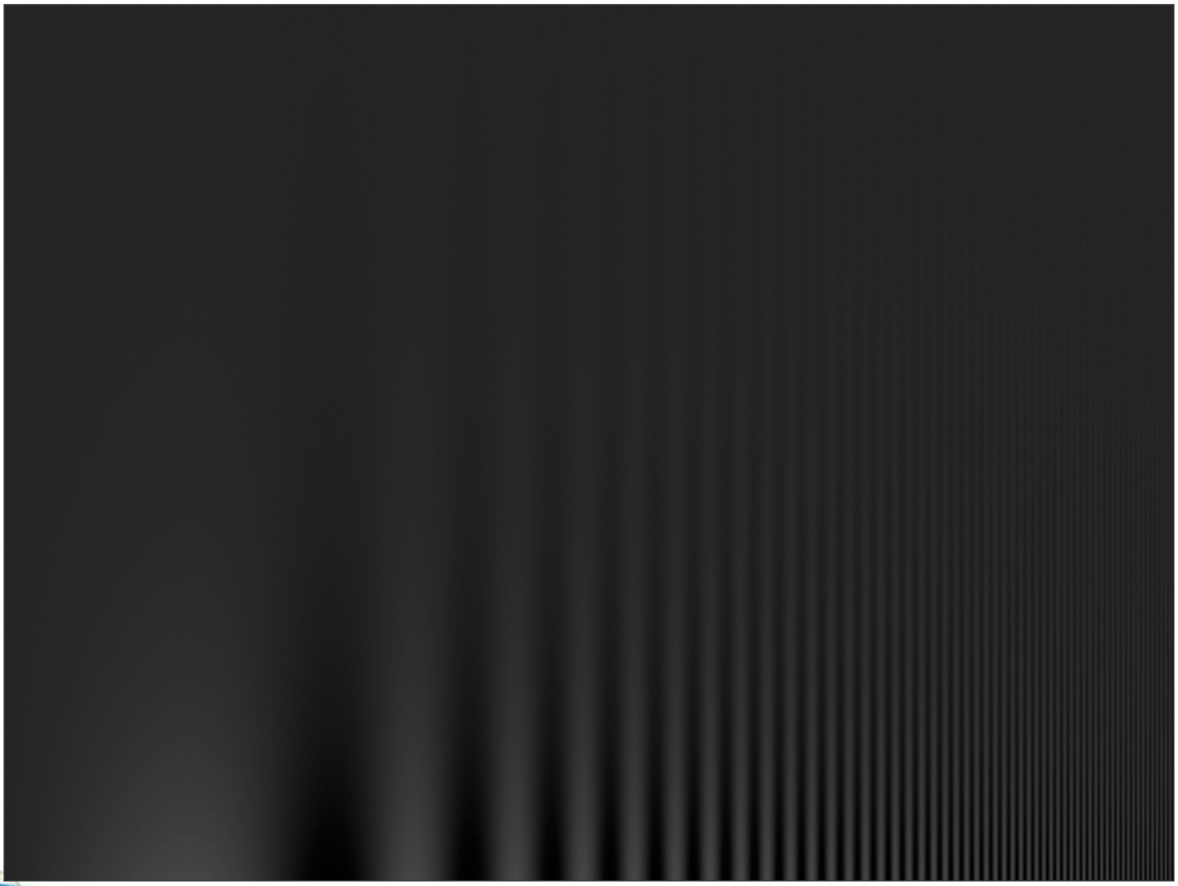
Note: To obtain eqivalent retinal intensities (Tr) the object surface intensity (cd/m^2) is inversely proportional to pupil area.

		Older Person	
Young Person			
Indoors			Outdoors
5 mm pupil	4 mm pupil	3 mm pupil	2 mm pupil
5 cd/m^2	8 cd/m^2	14 cd/m^2	32 cd/m^2

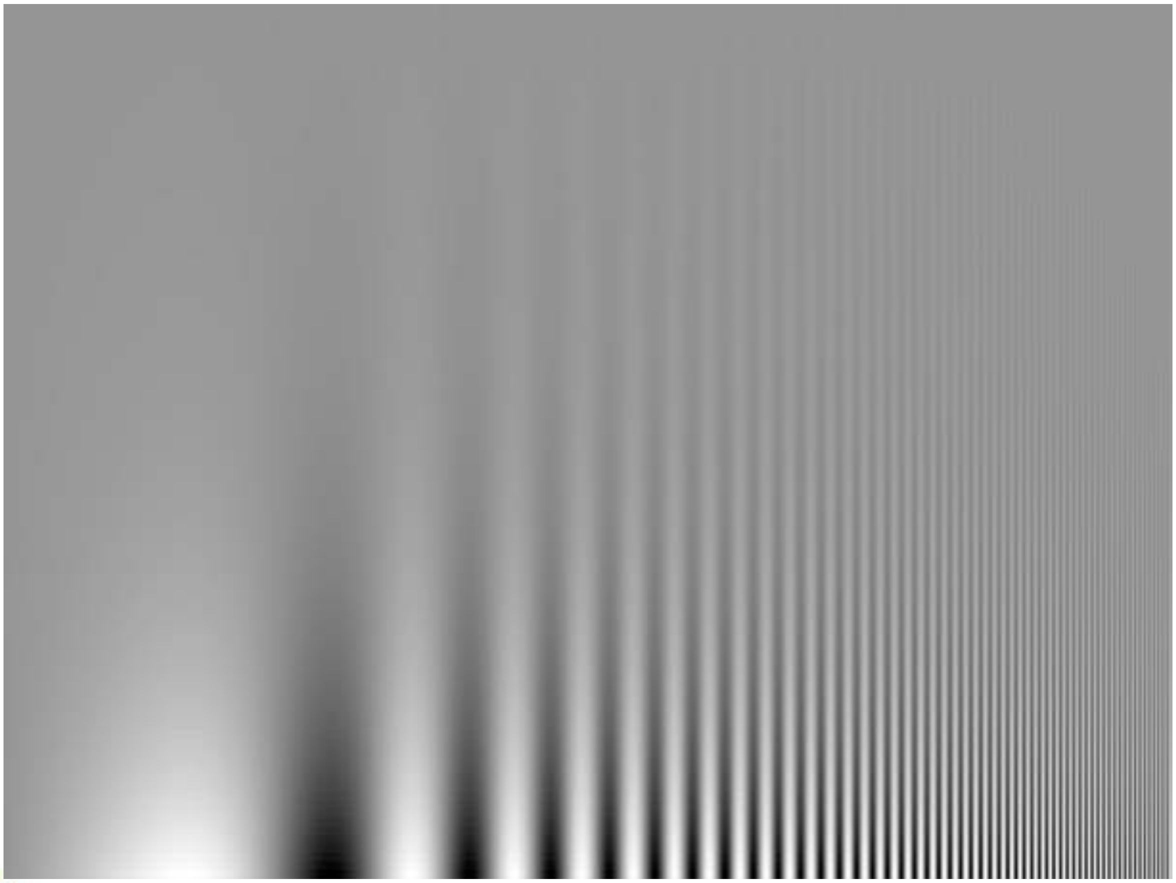


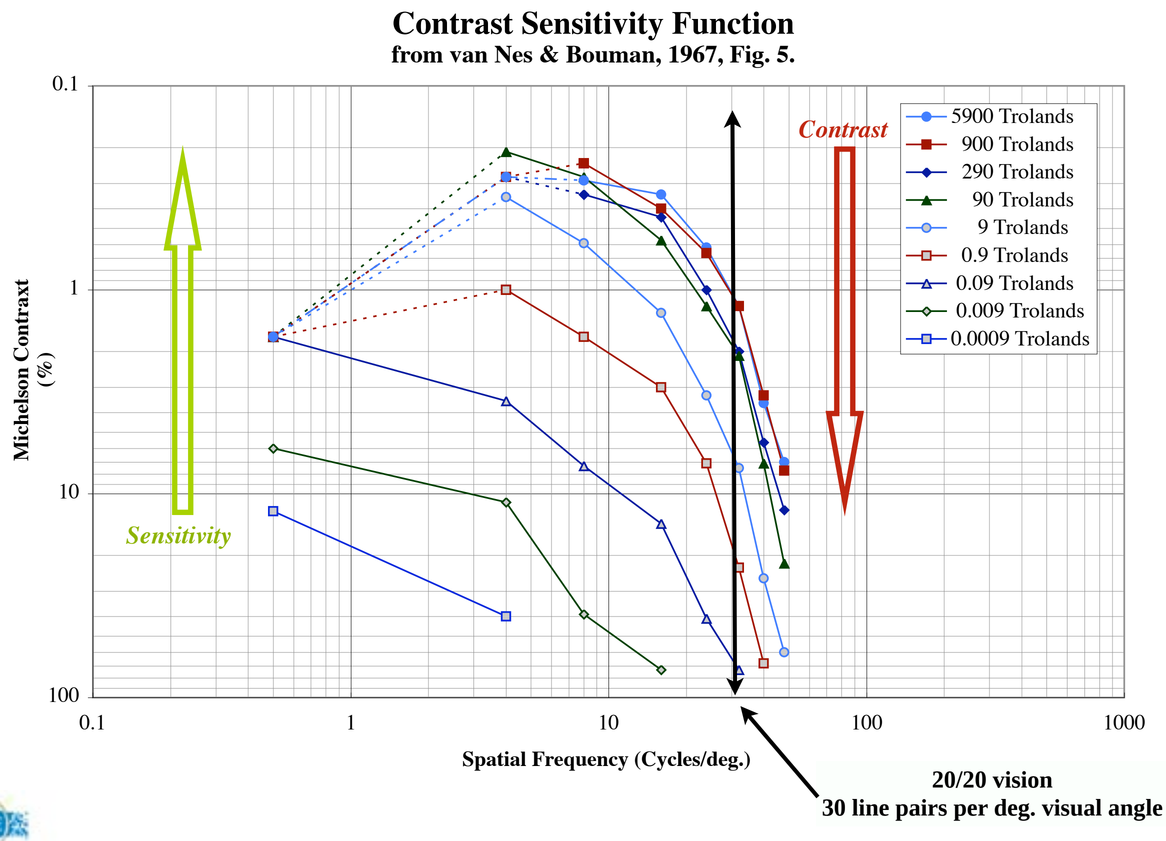
Most desktop monitors used as a computer terminal render at a max of around 15 cycles/deg when viewed from 0.5m or 19 inches.

For a young eye, 100 Tr is about 1.5 ftL (5cd/m^2) & 3x brighter than the 4.5 ftL (14cd/m^2) for an older person. Young people will see problems in the dark areas as more salient than an older member of the audience due to the smaller pupil aperture in an older individual. This effect is somewhat offset by increasing optical defects with a larger pupil, i.e., more optical distortion due to lens imperfections which are more significant with a large pupil.



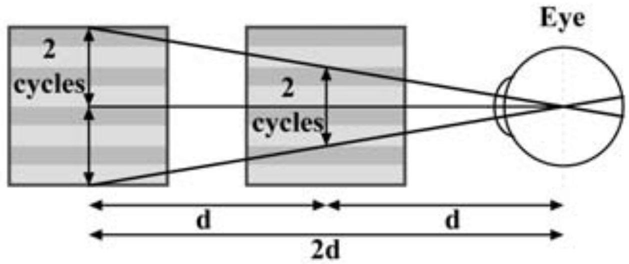
www.imagemetrics.com
(650) 678-0658





SMPTE 2008: Annual Tech Conference & Expo
Oct. 28-30, 2008 • Symposium Oct. 27, 2008
Hollywood, CA

Using the CSF to Find Optimal Viewing Distance



60 lines centered on the line of sight on the screen should subtend 1° visual angle to achieve a *eye-matched*, i.r., 20/20 resolution, viewing distance to a display screen.

		Picture Height in lines	Vertical Screen Angle Ø	*Eye-matched* Viewing Distance (in Screen Heights)
	NTSC/VGA	480	8.0	7.2
	720 line HD	720	12.0	4.8
	1080 line HD	1080	18.0	3.2
Quad HDTV =>	2160 line HD	2160	36.0	1.5
	Cinema	3000	50.0	1.1

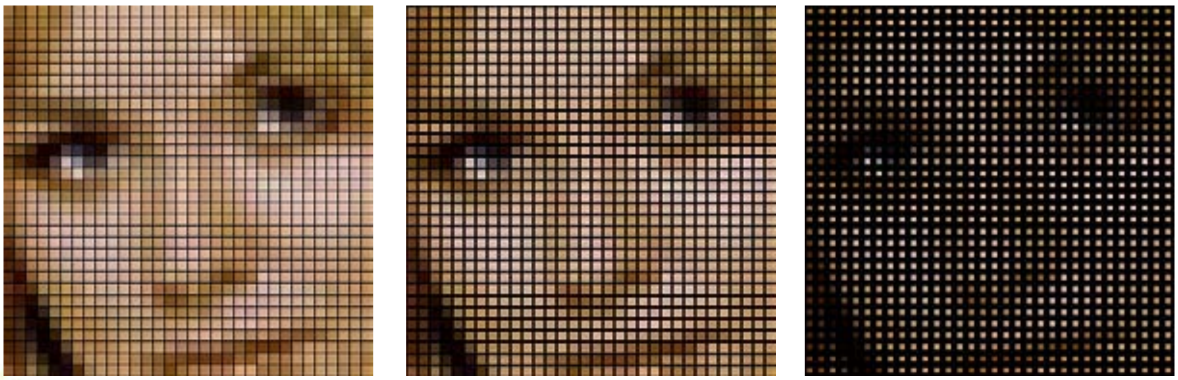
This chart gives the design eye viewing distances for each of the screen resolution formats. Viewing in front of this distance lowers the spatial frequencies in the imagery and viewing farther away increases them. The design distance matches 20/20 Snellen Eye Chart Acuity specs and indicates that a 20/20 observer would be able to see image features that subtend 1 minarc.

Jaggies Are More Apparent With Square Matrix Hard Edge Pixels

3 Octave Spatial Resolution Difference

© J. Larimer
www.imagemetrics.com
(650) 678-0658

The Fill Factor Resulting From the Black Matrix Reduces Image Brightness

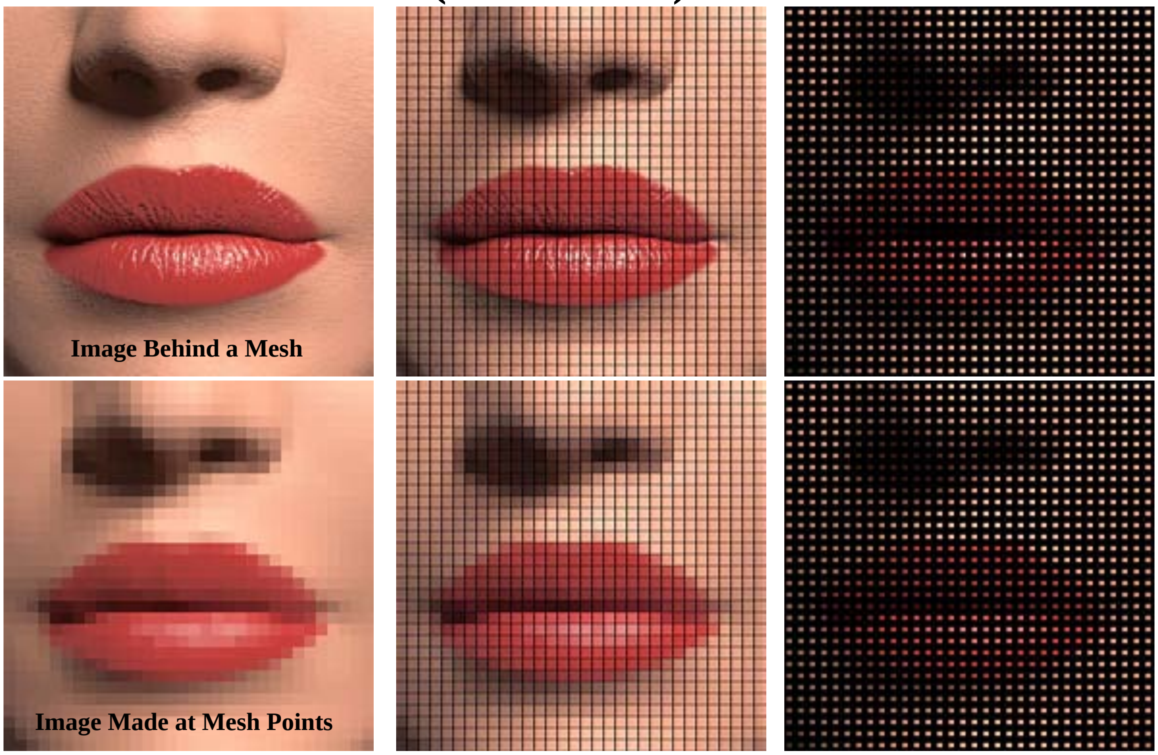


SMPTE 2008: Annual Tech Conference & Expo
Oct. 28-30, 2008 • Symposium Oct. 27, 2008
Hollywood, CA

www.imagemetrics.com
(650) 678-0658

The screen door or black matrix effect darkens the image but it also masks the jaggies.

The Black Matrix (*Fill Factor*) Can Be Your Friend



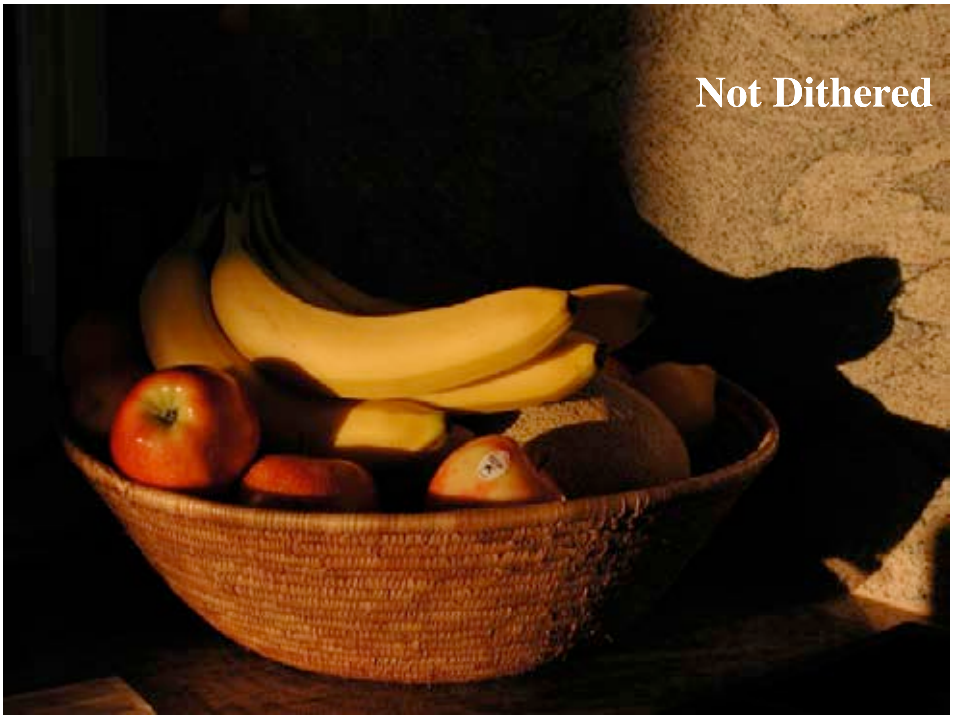
The Black Matrix Will Mask Image Detail When Pixel Structure Is Within the Window Of Visibility

The masking effect of the black matrix is illustrated here by placing either a high resolution or low resolution image behind a mesh of different aperture sizes.

Why Dither & Anti-aliasing Work

- **At high spatial frequencies the Weber Fraction increases. We are less sensitive to contrast at high spatial frequencies, so high spatial freq. details are less salient.**
- **Intensity steps at a spatial frequency that are below the contrast threshold will not be seen as steps but will contribute to lower spatial frequencies.**
- **The same rules that apply to printing (hard copy) apply to electronic displays (soft copy), so variable tonescale gives soft copy devices greater flexibility for dithering and anti-aliasing schemes like sub-pixel addressing.**

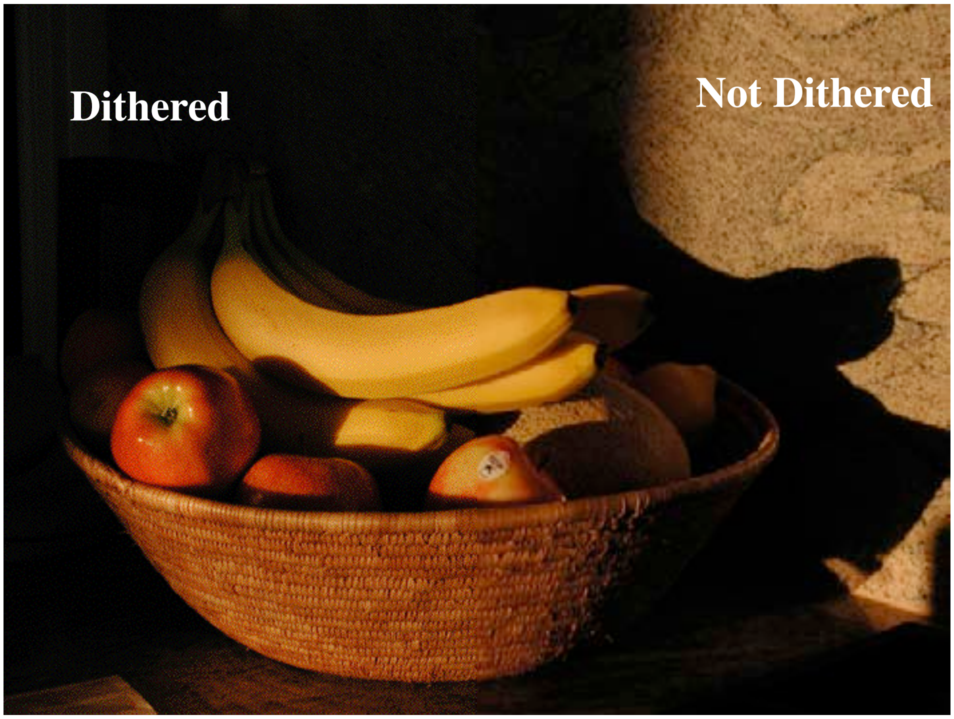
A Web-Safe Dithered Example



www.imagemetrics.com
(650) 678-0658

Here we have a full 24 bit resolution image.

A Web-Safe Dithered Example



The same image rendered out of 216 web-safe colors. The right side is as before, the left side has been down-sampled and dithered.

A Web-Safe Dithered Example

www.imagemetrics.com
(650) 678-0658

Dither Artifacts Are Perceived as Very Small Edges or Dots

Another Tone Scale Artifact is Banding When the Number of Levels is Insufficient

www.imagemetrics.com
(650) 678-0658

Dither Artifacts Are Perceived as Very Small Edges or Dots

Another Tone Scale Artifact is Banding When the Number of Levels is Insufficient

www.imagemetrics.com
(650) 678-0658

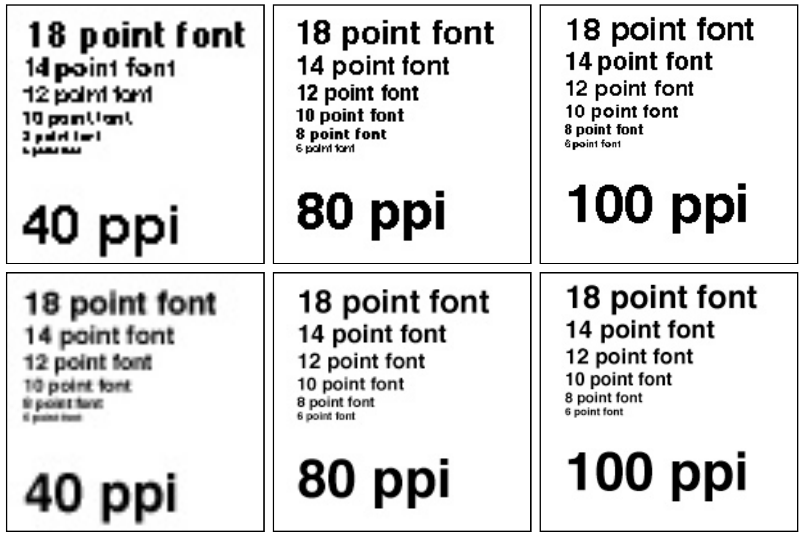
Dither Artifacts Are Perceived as Very Small Edges or Dots

Another Tone Scale Artifact is Banding When the Number of Levels is Insufficient

www.imagemetrics.com
(650) 678-0658

Text Legibility
The Legge[†] Rule of Thumb: 2.5-3 lp/letter Width

- **Helvetica font at 3 simulated resolutions with & without smoothing or anti-aliasing. Anti-aliased fonts on the bottom row.**
- **Font smoothing & other techniques such as sub-pixel addressing lose their value as pixels/inch (ppi) increases & pixelation becomes less visible.**



[†]Legge, G.E., Pelli, D.G ., Rubin, G.S ., Schleske, M.M., Psychophysics Of Reading: I. Normal Vision. *Vision Research*, Vol. 25, No. 2, pp. 239-252, 1985.

Using this rule of thumb, 5 pixels per character width, you can determine the resolution required to make text legible in a scene.

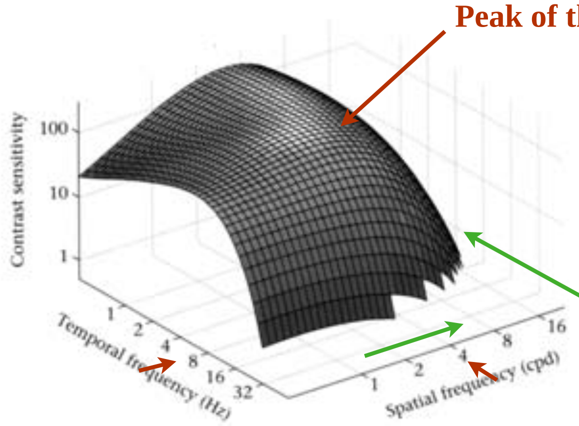
Temporal Tuning of the Visual System

www.imagemetrics.com
(650) 678-0658

The Spatial-Temporal CSF Measured by Kelly

(Figure from Wandell, Foundations of Vision, 1995)

Peak of the Spatial-Temporal CSF



As Spatial frequency increases our ability to see temporal changed decreases!

7.22 HUMAN SPATIOTEMPORAL CONTRAST-SENSITIVITY FUNCTION. The two lower axes represent the spatial and temporal frequencies of a contrast-reversing pattern; the vertical axis represents the observer's contrast sensitivity to each of the contrast-reversing patterns. The data used to estimate this surface were made on a mean background luminance of 1,000 Trolands. Curves running parallel to the spatial-frequency axis define a set of spatial contrast-sensitivity functions measured at different temporal frequencies (cf. Figure 7.4). Curves running parallel to the temporal-frequency axis represent the temporal contrast sensitivity measured at different spatial frequencies. Human spatiotemporal contrast sensitivity is not space–time separable. Source: Kelly, 1966, 1979a, b.

We do not see speeding bullets & tweed on marathoners is wasted on bystanders viewing the race.

Margin Views of the Spatial-Temporal CSF

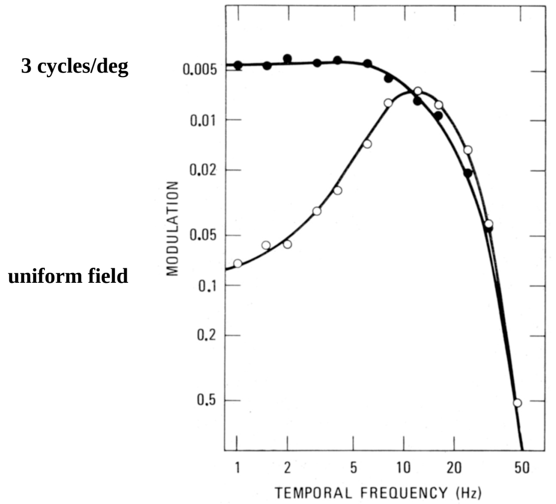


FIG. 1. Maximum and minimum achromatic flicker-sensitivity curves, obtained at spatial frequencies of 3 cycles/deg (filled circles) and zero (open circles); artificial pupil, 2.3 mm; retinal illuminance, 10^3 td. These may be regarded as two parallel sections through the spatio-temporal threshold surface described in the text.

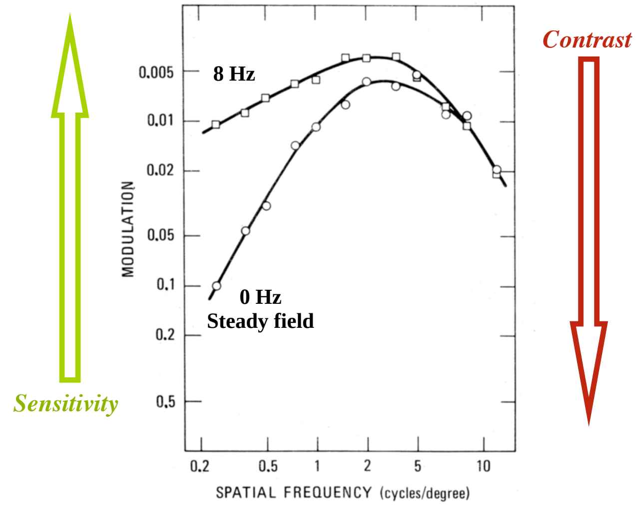


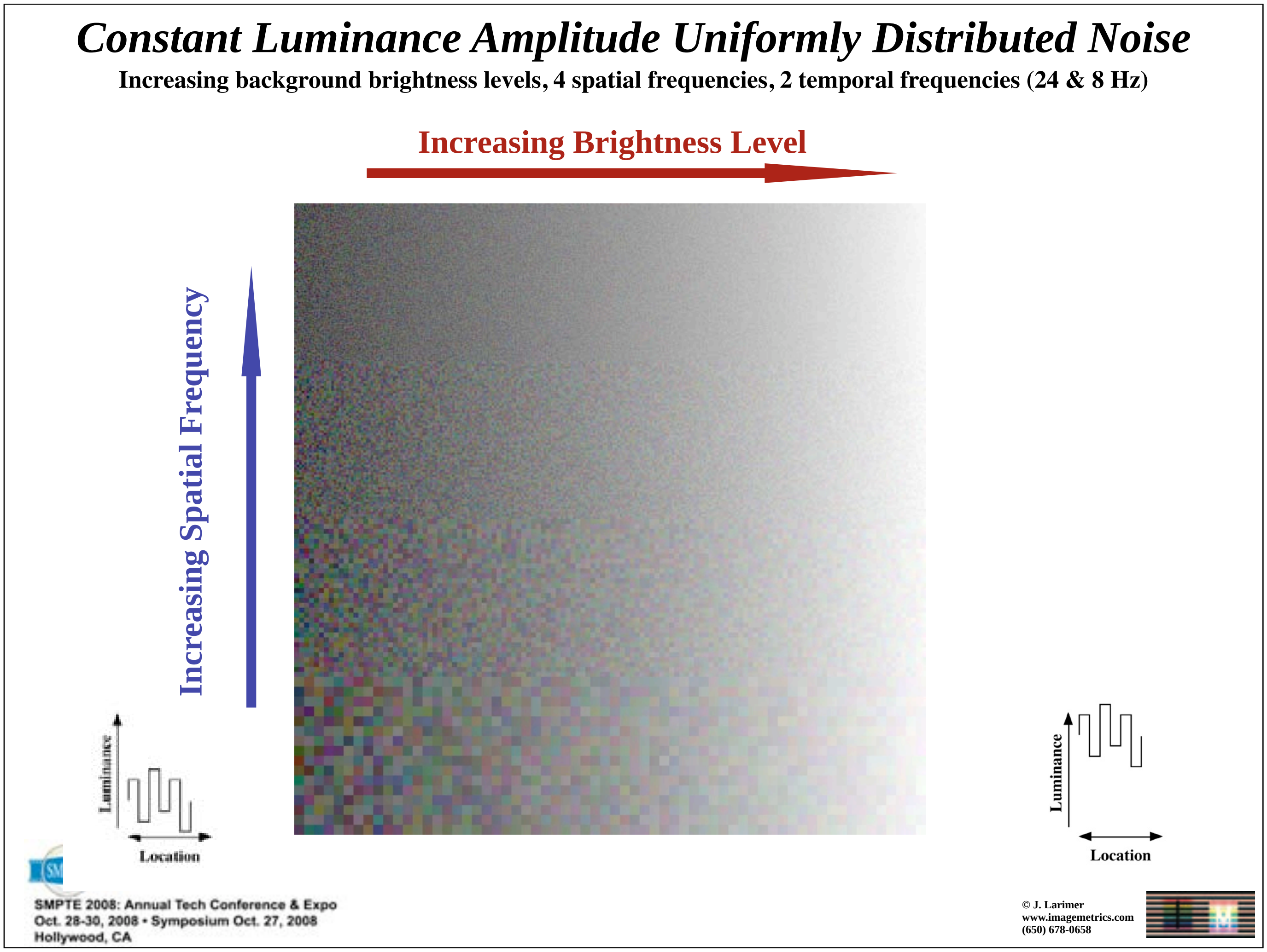
FIG. 2. Maximum and minimum achromatic contrast sensitivity curves, obtained at temporal frequencies of 8 Hz (squares) and zero (circles); other conditions, same as Fig. 1. These may be regarded as two parallel sections through the spatio-temporal threshold surface, forming a rectangle with those of Fig. 1.

From D.H. Kelly, JOSA,64,1974,983-990.

SMPTE 2008: Annual Tech Conference & Expo
Oct. 28-30, 2008 • Symposium Oct. 27, 2008
Hollywood, CA

© J. Larimer
www.imagemetrics.com
(650) 678-0658

This data from Kelly shows two cross sections from the previous slide showing the combined spatial-temporal tuning surface for the eye at a fixed DC level well into the photopic range and where Weber-Fechner performance dominates. Recall that Delta(I)/I=k, i.e., sensitivity has peaked, typifies the performance data at these DC (brightness) levels. When the Weber fraction saturates, the region is called the Weber-Fechner region. Above these levels when all spatial temporal channels have achieved optimal sensitivities, the performance characteristic or tuning surface does not change. Below these levels the surface topography of the tuning characteristic surface will depend upon the DC level at which the characterization was measured. **LEFT:** On the left are two slices through the surface perpendicular to the spatial frequency axis; a uniform (i.e., zero spatial frequency) and 3 cycles/deg (near the spatial peak) are shown. **RIGHT:** On the right are two slices through the surface perpendicular to the temporal frequency axis; a steady temporally invariant field and a field varying in counter-phase close to the temporal sensitivity peak. Kelly additionally found that if the grating images could be completely stabilized on the retina so that the target temporal frequency exactly matched the temporal frequency in the retinal image (note that eye motion will change spatial-temporal rates) that the steady field without a visible edge would completely vanish independent of the grating pattern's spatial frequency. This means that temporal change is required to see, or colloquially that the eye is AC coupled, it computes a time derivative while coding the signal on the retina.

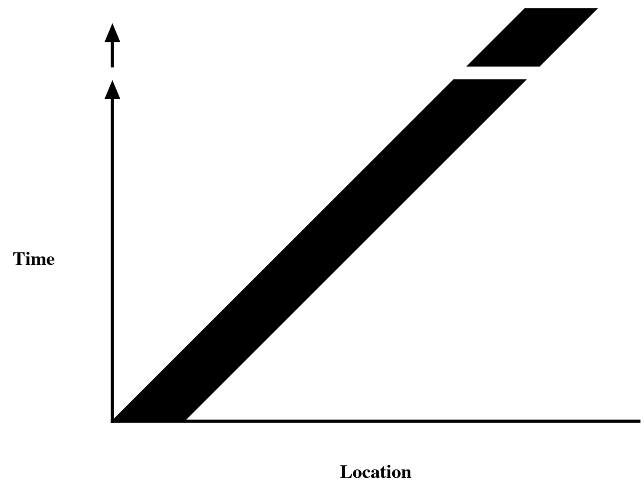


Spatial-temporal artifact visibility depends upon temporal rates and spatial size.

Judder

Judder is the temporal domain analog to jaggies. It is a discrete sampling artifact.

The Motion of a Moving Bar

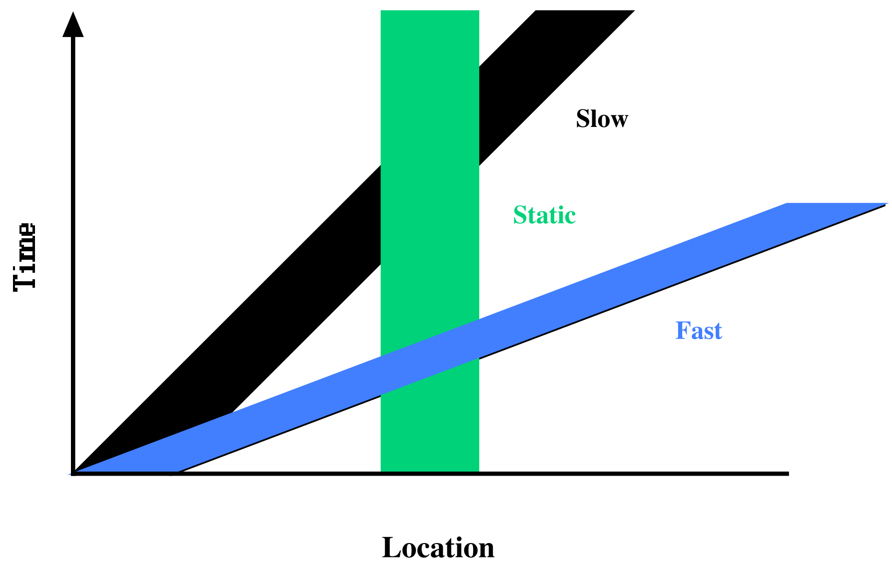


Space-Time Diagram of Moving Line
Brightness represents line intensity

A space time diagram shows the location of a bar in space (x-axis) at an instant in time (y-axis).

Space-time Diagram of 3 Motions:

Fast, Slow, Stationary



Space-Time Diagram of Moving Line

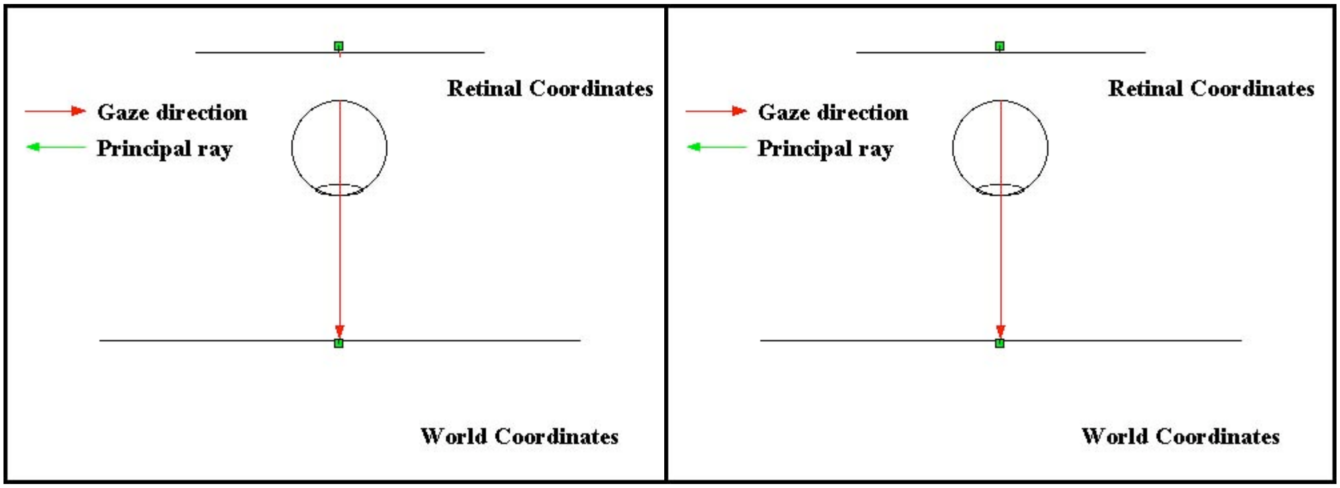
Three different rates.

SMPTE 2008: Annual Tech Conference & Expo
Oct. 28-30, 2008 • Symposium Oct. 27, 2008
Hollywood, CA

Three different rates of translation.

How the Eye Looks at Moving Bars

Staring Eye: Fixed Gaze



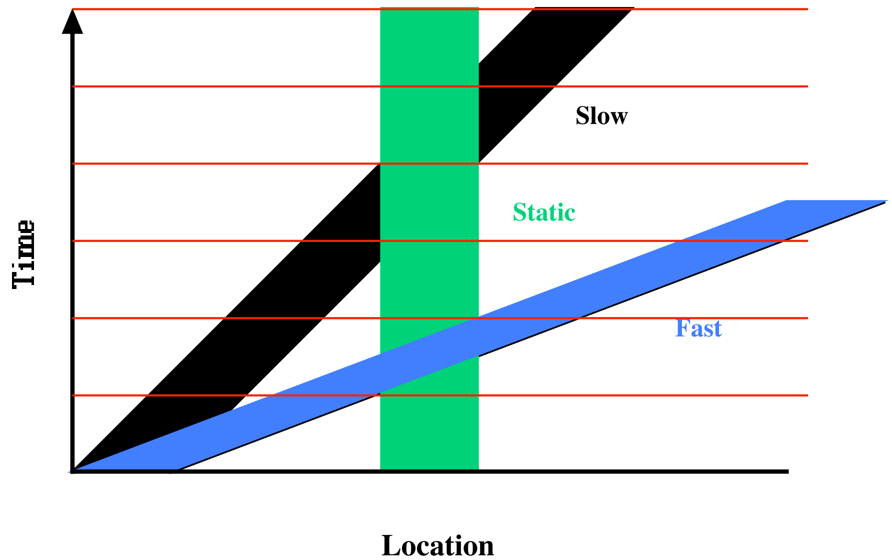
Smooth Pursuit Eye Movements

www.imagemetrics.com
(650) 678-0658

A staring eye on the left and a smooth pursuit eye motion on the right.

Sampling in World Coordinates

Time is divided into bins & each bin is temporally sampled.

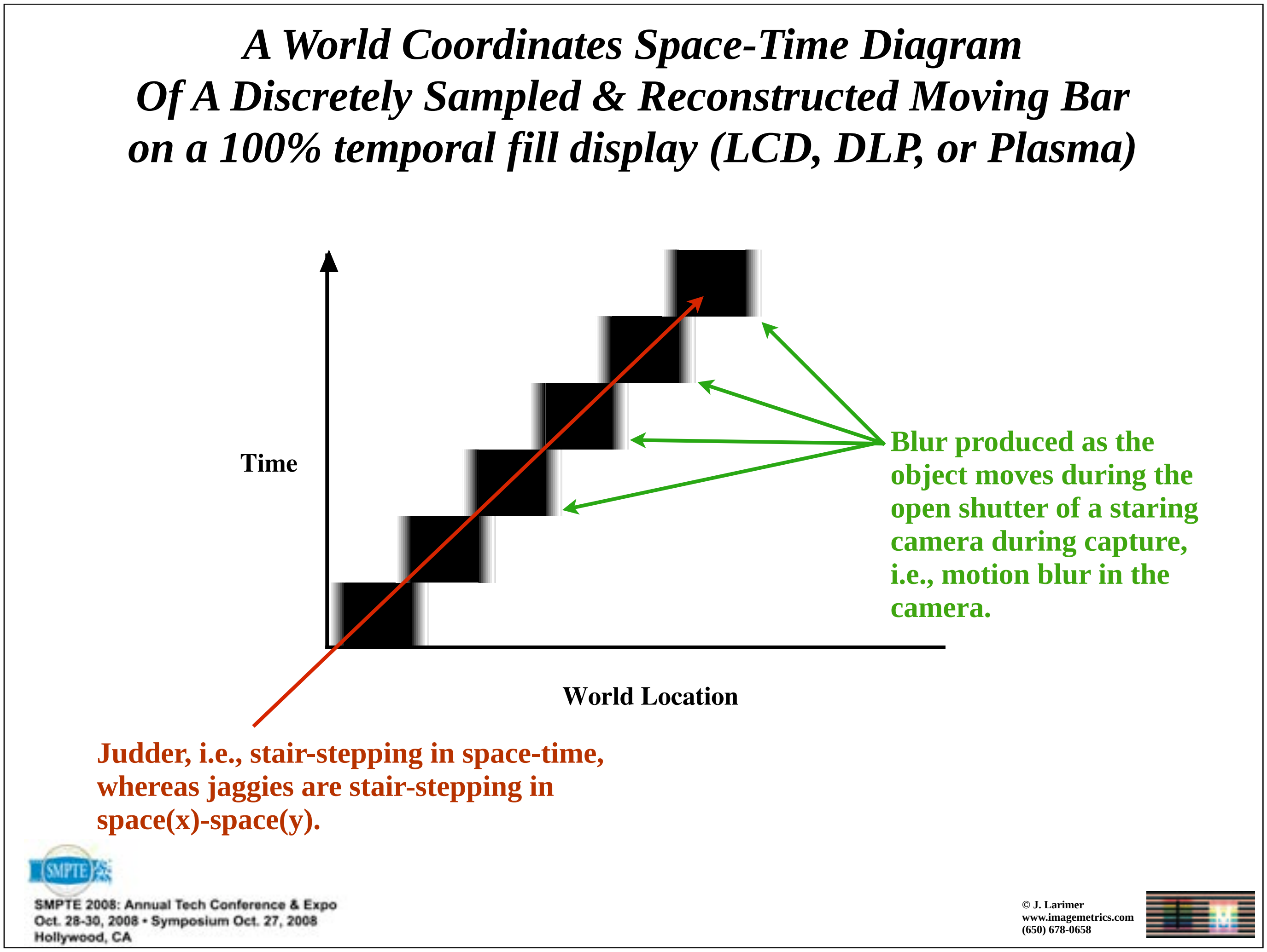


Space-Time Diagram of Moving Line
Three different rates.
Red lines represent sampling periods.

SMPTE 2008: Annual Tech Conference & Expo
Oct. 28-30, 2008 • Symposium Oct. 27, 2008
Hollywood, CA

www.imagemetrics.com
(650) 678-0658

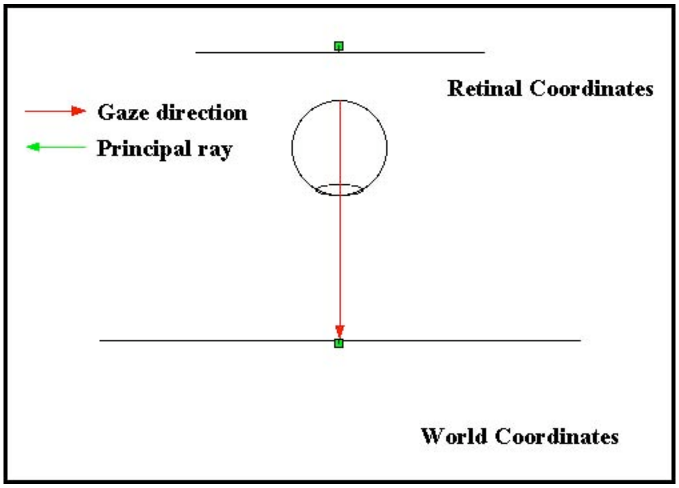
The time axis is sampled discretely. The dwell time of the sample can be constant, variable, or random variable and the camera can be staring or pursuing. The period between samples can be variable, fixed or random. Each of these factors determines the temporal and spatial power spectra of the resulting captured image signal. In this example, the camera is staring and the dwell time is 100% with uniform with constant temporal frequency sampling.



The reconstruction space-time diagram based upon the constant rate sampling shown in the previous slide is shown here. The edges are blurred due to motion blur for a moving bar and staring camera. Depending upon how the focal plane image is sampled and read out on the camera sensor plane, a skew can also be added to the image if the camera is operating as a push-broom sampling device. The capture sampling can be complex in all domains, space, time and intensity. CMOS sensors are creating new possibilities for new and varied wavefront sampling strategies so the space time components within the captured signal can be designed to meet specific goals for information capture.

Continuous Pursuit of a Moving Object on a Display with Discrete Temporal Reconstruction,

for example, an LCD, DLP, or Plasma Display

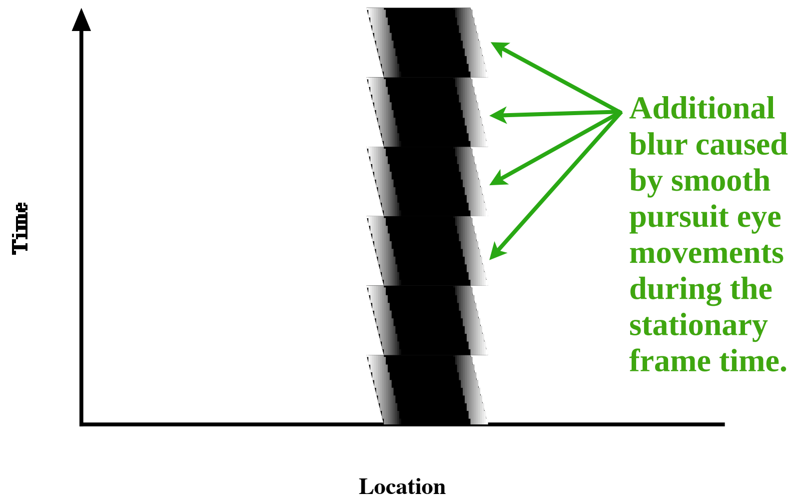


A smooth pursuit eyeball viewing a stationary discrete image reconstruction on a fast display.

Retinal Coordinate Space-Time Diagram when the Eye Smoothly Pursues the Moving Object:

Depending upon motion rate this produces judder induced edge flicker.

- **The image shearing is caused by the eye continuously pursuing the motion.**
- **Pursuit eye movements tend to follow the blob energy in the image.**
- **Along the edges is a simple flicker signal (a sawtooth) at the frame rate frequency.**
- **No surprise, we see flicker!!**
- **It will take a few thousand centuries of viewing digitally reconstructed images for evolution to "fix" this problem, so we better not wait.**



Space-Time Diagram of Perfectly Pursued Moving Line
Brightness represents line intensity

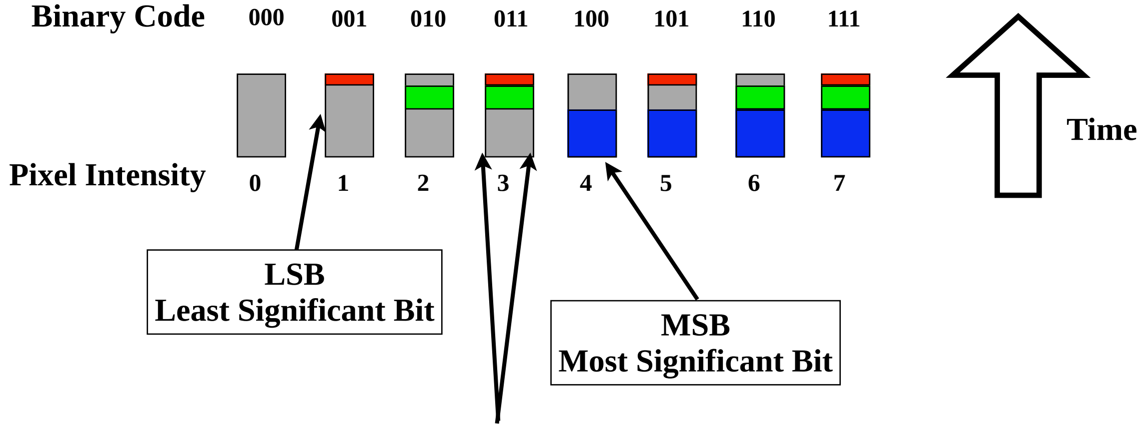
The space time diagram for the image signal reconstruction on the retina.

(PWM)
Pulse Width Modulation Artifacts

www.imagemetrics.com
(650) 678-0658

Time division to control intensity is widely used in modern displays. Typically the frame time is divided into binary weighted intervals or primarily binary weighted intervals. It is also possible to use equal temporal intervals and control the light intensity during these brief subframe periods.

Example of a 3 Bit PWM System



Pixel Aperture or Footprint

f = frame rate (fps) in Hz
c = clock rate = f x (2^n-1)

This requires fast optical switching, for example at 60 fps and 8 bits the clock frequency is 15.3 kHz. The pixel must switch in less than 65 µsec.

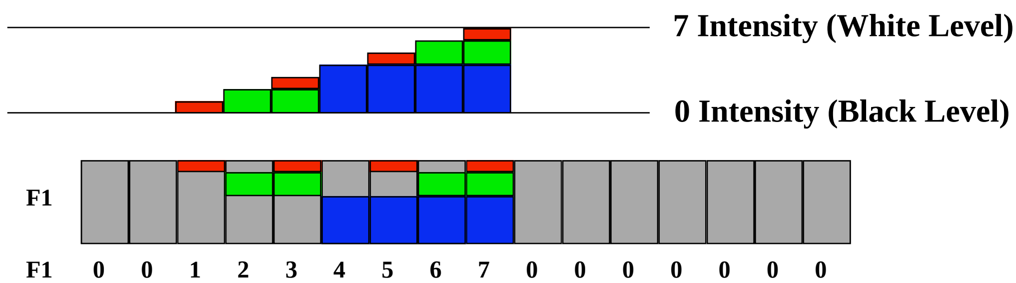
SMPTE 2008: Annual Tech Conference & Expo
Oct. 28-30, 2008 • Symposium Oct. 27, 2008
Hollywood, CA

© J. Larimer
www.imagemetrics.com
(650) 678-0658

Binary weighted intervals for a 3 bit solution are 1:2:4 respectively for 0 to 7 levels.

Single Frame Image of a Ramp Viewed with a Staring Eye

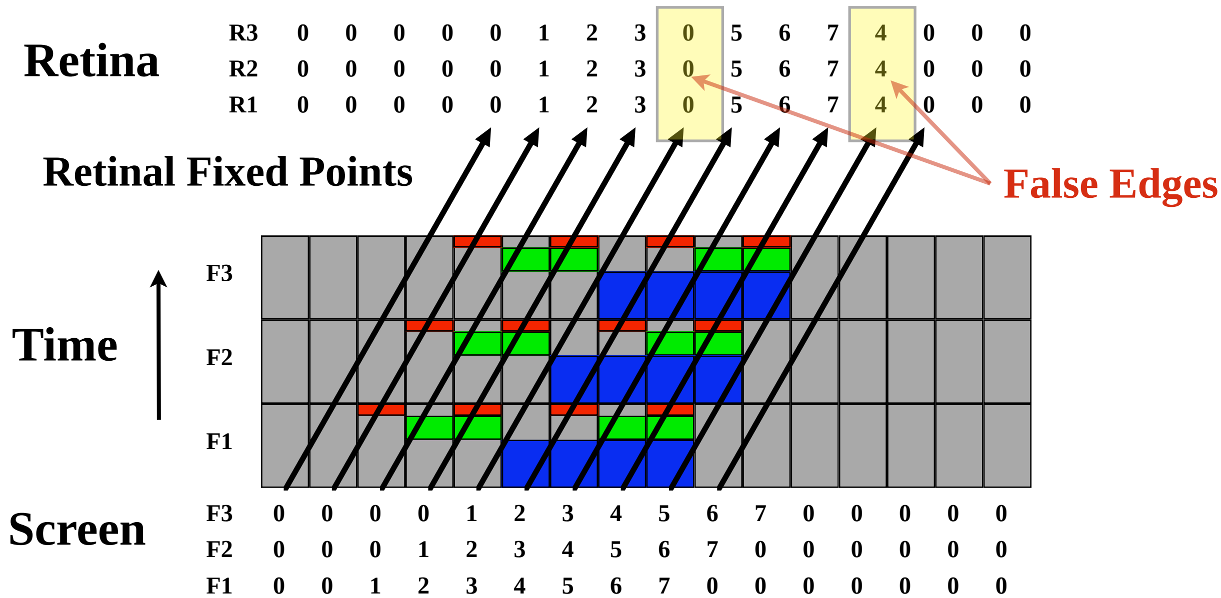
(fixed gaze)



SMPTE 2008: Annual Tech Conference & Expo
Oct. 28-30, 2008 • Symposium Oct. 27, 2008
Hollywood, CA

A spatial ramp reconstructed with a binary weighted PWM intensity system.

Moving Ramp Viewed With Smooth Pursuit Eye Movements



SMPTE 2008: Annual Tech Conference & Expo
Oct. 28-30, 2008 • Symposium Oct. 27, 2008
Hollywood, CA

www.imagemetrics.com
(650) 678-0658

Smooth pursuit eye movements produce false edges or contours in the ramp, due to the pulse width modulation scheme used to produce the tone scale. The smooth pursuit eye will literally “look” through the PWM schema producing false edges on the retinal focal plane.

Additionally the most significant bit creates a signal with potential energy at 2x the frame rate. Aliasing due the to eye’s auto-phase locking behavior can also produce indiscernible tone scale levels. For example a temporal variation that flips the MSB in a fixed bit order reconstruction will look exactly like a “all-off” “all-on” temporal pattern. To avoid this requires breaking up the MSB and permutation ordering of the bit stream.

Distortion and artifacts are image features. Distortion does not lose information whereas artifacts add information to the image signal. Both are neither good or bad from an aesthetic perspective. Both must be controlled so that the video content producer can create the image look they desire. The engineering task therefore is to provide interfaces and mechanisms to control these visual affects.

This lecture has just touched on the more evident aspects of artifact generation. Controlling artifacts is possible, but requires a mating of the technology for sampling to the image signal and technology reconstructing it. All of this is within the realm of the possible with modern digital capture, coding, editing, and display technologies. Making these signals available to the artist is a remaining challenge for the digital imaging engineer.

Thank You for Your Attention

www.imagemetrics.com
(650) 678-0658