



**SC762**

**Device Simulation for Image Quality Evaluation**

Date: 17 January 2010

Time: 8:30 AM - 5:30 PM

Location: San Jose, United States

Course Level: Intermediate

Instructors:

Joyce E. Farrell, Peter B. Catrysse, Brian A. Wandell  
Stanford Center for Image Systems Engineering  
Stanford Univ. (United States)

Imageval, LLC  
Palo Alto, CA 94305

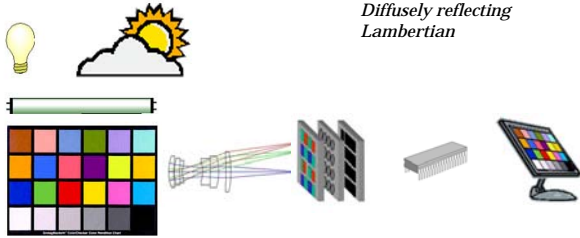


## Course schedule

- 8:30–8:45 Preview (Brian)
- 8:45–9:30 Scene radiance (Brian, Joyce)
- 9:45–10:30 Optics and irradiance (Peter)
- 10:45–11:30 Sensor model (Brian)
- 11:30–12:00 System design metrics (Joyce)
- 12:00 –1:00 Lunch
- 1:00–1:45 Sensor system components (Peter)
- 2:00–2:45 Human observer (Brian)
- 3:00–3:45 Image processing (Brian)
- 4:00–4:45 Metrics and preference (Joyce)
- 5:00–5:30 Post-view and discussion (All)

## Scene radiance

*Simplifying assumptions*  
 Planar surface  
 Diffusely reflecting  
 Lambertian

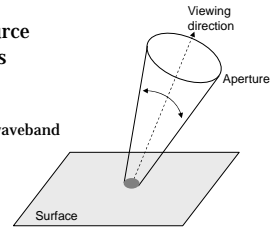


## Spectral radiance units

**Definition:** Light emitted or reflected from an extended source in a given direction. The light is specified in units of

energy/steradian/surface area/second/waveband

The surface area is foreshortened according to the viewing direction

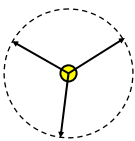


## Radiance geometry

*Functions of wavelength*

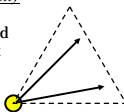
Radiant flux

Light emitted from a point



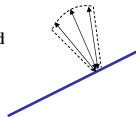
Radiant intensity

Light emitted from a point in a solid angle



Radiance

Light scattered from an extended source in a given direction

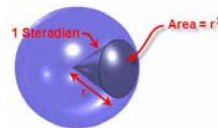


## Spectral radiance

$$L(\lambda) = \frac{E(\lambda)}{\omega A_s \cos(\theta) t \Delta\lambda}$$

• Solid angle of a cone (steradians): the surface area on a unit sphere intersected by the cone whose apex is at the center of the sphere

• A sphere is  $4\pi$  steradians



Solid angle

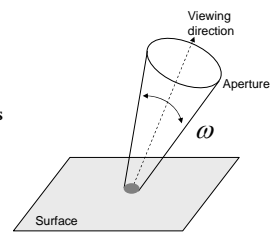
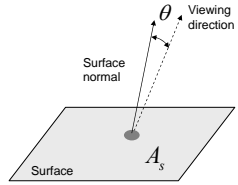


Image from <http://www.mathsisfun.com/geometry/steradian.html>

### Spectral radiance

$$L(\lambda) = \frac{E(\lambda)}{\omega A_s \cos(\theta) t \text{ nm}}$$

Foreshortened surface area



• The surface area is foreshortened and consistent with the area as seen from the viewing direction.

• As the angle approaches 90 deg, the area becomes zero and the energy becomes zero.

### Measuring spectral radiance

Spectroradiometer  
photons/(s sr nm m<sup>2</sup>)



Photoresearch PR-655



Spectral radiance  
[photons/(s sr nm m<sup>2</sup>)]

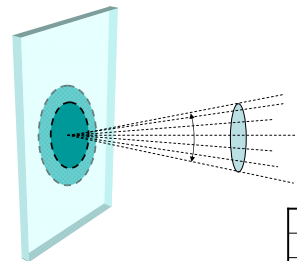
### Spectral radiance units

Term	Defining Equation	Application	SI Unit
Radiant Flux	$F = \frac{\Delta Q}{\Delta t}$	Light emitted from a point	watt
Radiant Intensity	$I = \frac{\Delta F}{\Delta \omega}$	Light emitted from a point in a given solid angle	watt / sr
Radiance	$L = \frac{\Delta I}{\Delta A_s \cos(\theta)}$	Light from an extended source in a given direction	watt /sr/m <sup>2</sup>

Q = energy (joules)  
F = Q/sec (watt)  
t = time (sec)  
ω = solid angle (steradian)  
A = area (meter<sup>2</sup>)  
θ = angle incident to plane



### Radiance geometry



Closer, acquires larger solid angle than farther

Parameter	Distance
Solid angle	
Surface area	

Spectroradiometer uses this principle to measure luminance without measuring distance

### Radiance geometry

*Closer, acquires  
larger solid angle  
than farther*

Parameter	Distance
Solid angle	Decreases
Surface area	

*Spectroradiometer uses this principle to measure luminance without measuring distance*

### Radiance geometry

*Closer, acquires  
smaller surface area  
than farther*

Parameter	Distance
Solid angle	Decreases
Surface area	Increases

*Spectroradiometer uses this principle to measure luminance without measuring distance*

### Radiance geometry

*Closer, acquires  
smaller surface area  
than farther*

Parameter	Distance
Solid angle	Decreases
Surface area	

*Spectroradiometer uses this principle to measure luminance without measuring distance*

### Luminance from spectral radiance

- Source Luminance  
– Symbol:  $L_v$   
– Units:  $\text{cd/m}^2$
- ←
- Source Radiance  
– Symbol:  $L$   
– Units:  $\text{W}/(\text{sr nm m}^2)$

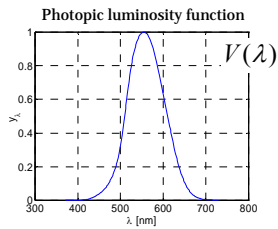
Photopic luminosity function

Wavelength λ [nm]	Relative Sensitivity V(λ)
380	0.0004
400	0.0012
420	0.0041
440	0.0171
460	0.0543
480	0.1375
500	0.2898
520	0.5033
540	0.7171
555	1.0000
570	0.7171
580	0.5033
600	0.2898
620	0.1375
640	0.0543
660	0.0171
680	0.0041
700	0.0012
720	0.0004
800	0.0000

### Luminance from spectral radiance

$$L_v = K_m \int V(\lambda) \cdot L(\lambda) d\lambda$$

- Conversion
  - Peak luminosity: unity at 555 nm
  - $K_m = 683 \text{ lm/W}$



### Luminance from spectral radiance

- Luminance
  - Symbol:  $L_v$
  - Units:  $\text{cd/m}^2$
- Spectral Radiance
  - Symbol:  $L$
  - Units:  $\text{W}/(\text{sr nm m}^2)$

$$L_v = K_m \int V(\lambda) \cdot L(\lambda) d\lambda$$

$$\frac{\text{cd}}{\text{m}^2} = \frac{\text{Lumen}}{\text{sr m}^2} = \left( \frac{\text{Lumen}}{\text{W}} \right) \int \frac{\text{W}}{\text{sr nm m}^2} d\lambda$$

### Luminance units

Term	Defining Equation	Application	SI Unit
Luminous Flux	$F_v = K_m \int F_e(\lambda) V(\lambda) d\lambda$	Light emitted from a point	Lumen (lm)
Luminous Intensity	$I_v = \frac{\Delta F_v}{\Delta \omega}$	Light emitted from a point in a given solid angle	candela (cd=lm/sr)
Luminance	$L_v = \frac{\Delta I_v}{\Delta A_v \cos(\theta)}$	Light emitted or reflected from an extended source in a given direction	$\text{cd/m}^2$

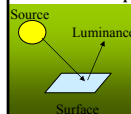
### Typical luminance levels

#### Outdoors ( $\text{cd/m}^2$ )

Sun	$6 \times 10^8$
Visual saturation	49,000
Just below saturation	25,000
Outdoor building façade	10,000
Blue sky (morning)	4,600
Concrete sidewalk	
in sun	3,200
in shadow	570
in deep shadow	290

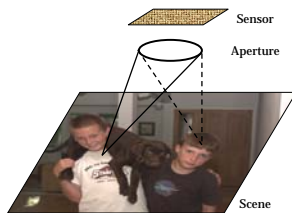
#### Indoors ( $\text{cd/m}^2$ )

Interior room (fluorescent lighting)	
floor/walls	90
in shadow	10
Interior room (no lighting)	
floor/walls	30
in shadow	5
in closet door	1



## Scenes and images

- Images are not extended sources (spectral radiance)
- Images are a collection of small light scattering (or emitting) regions
- We suppose the radiant intensity is equal in all directions\* (Lambertian)
- To simplify simulations, we store the radiance (not radiant intensity)



\* Since we view the scene from one angle, this assumption is not important for simulation

## Why are linear models useful?

- Compact representation of the data
- Work smoothly with conventional linear matrix computations
- Help make design decisions like the number of color types needed to measure a set of signals

## Data management: Low-dimensional linear models

- Introduction
- CIE Daylight model
- Surface reflectance models
- Application: Scene database

## Low dimensional linear models

- In most applications, the set of spectral functions of interest are a subset of all possible functions
- Consequently, we can represent nearly all of the variance in a typical function by the weighted ( $w_i$ ) sum of a small number ( $N$ ) of spectral basis functions ( $B_i$ )

$$s(\lambda) \approx \sum_{i=1}^N w_i B_i(\lambda)$$

### Linear models are efficient

For an M-dimensional linear model, we find the basis functions and weights that minimize

$$\sum_{\lambda} [s(\lambda) - \sum_{i=1}^M B_i(\lambda)w_i]^2$$

The basis functions are fixed – so each spectral function is summarized by only the M-weights per function plus the basis terms

$$s(\lambda) \approx \sum_{i=1}^M B_i(\lambda)w_i$$

### Linear model in matrix tableau

Spectral function	Basis functions	Weights
$\begin{pmatrix}   \\ s(\lambda) \\   \end{pmatrix}$	$\approx \begin{pmatrix}   &   \\ B_1(\lambda) & B_2(\lambda) \\   &   \end{pmatrix}$	$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$

### Linear Models: SVD

$$\begin{pmatrix} | \\ \text{Data in} \\ | \\ \text{columns} \end{pmatrix} \approx \begin{pmatrix} | \\ \text{Bases} \\ | \end{pmatrix} \begin{pmatrix} \text{Weights} \end{pmatrix}$$

D = USV<sup>t</sup>  
D ~ UW

The singular value decomposition is an excellent way to build linear models

$$D = USV^t$$

$$UU^t = U^tU = \text{Identity}$$

$$VV^t = V^tV = \text{Identity}$$

*S is diagonal with  $s_i \geq s_{i+1}$*

To compute basis functions: (typically, one removes the mean first)

$$D = USV^t$$

*U = Columns are basis functions*

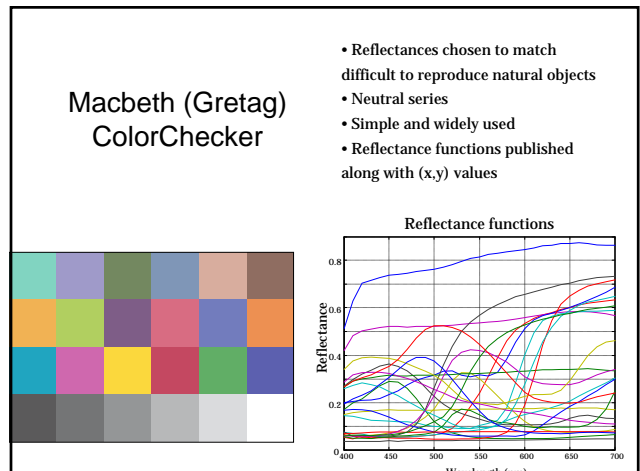
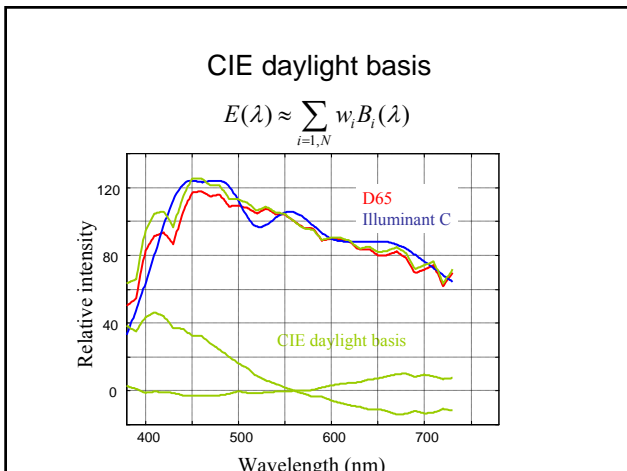
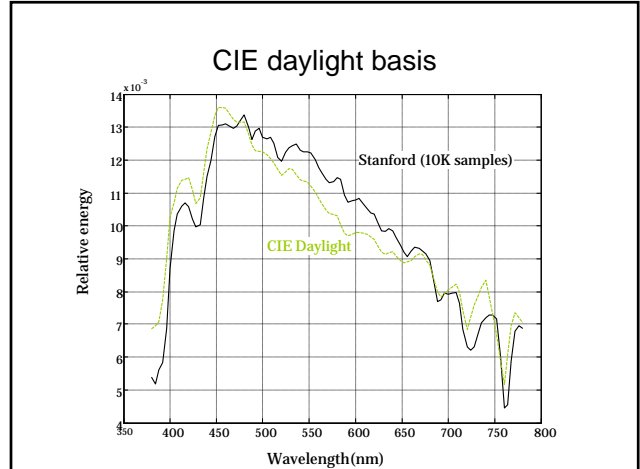
*SV<sup>t</sup> = Weights*

The singular value decomposition: an excellent way to build linear models

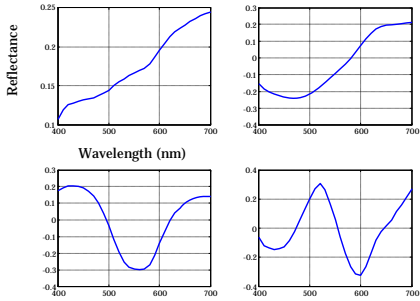
$$D = USV^t$$

$U$  = Columns are basis functions

$SV^t$  = Weights

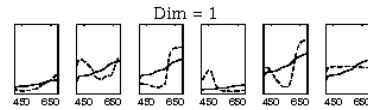


### Macbeth reflectance basis functions

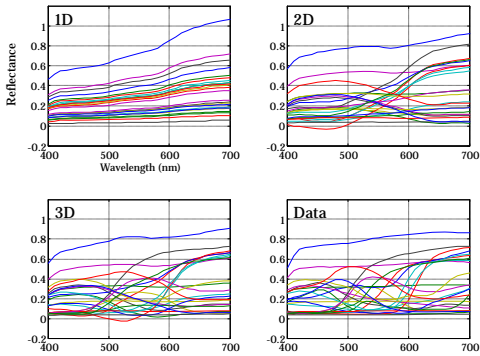


```
ref = macbethReadReflectance(wave);
[U S V]=svd(ref);
plot(wave,U(:,1:4))
```

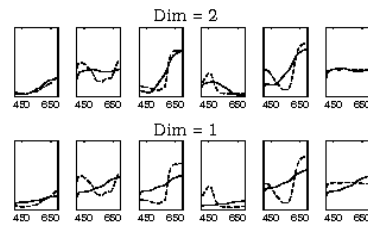
### Visualizing the number of bases



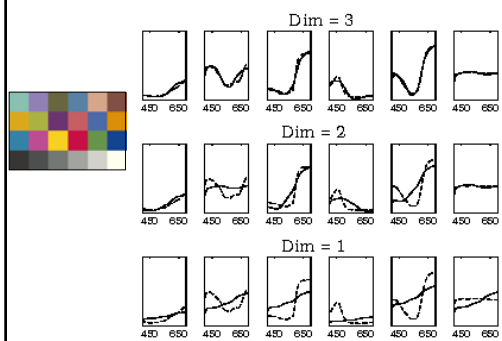
### MCC approximations



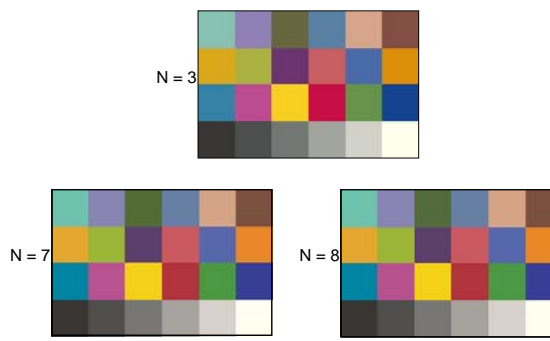
### Increasing the number of bases



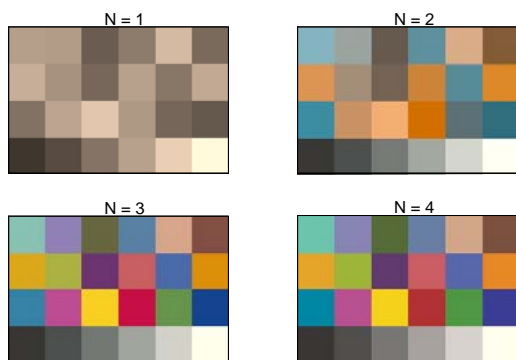
### Increasing the number of bases



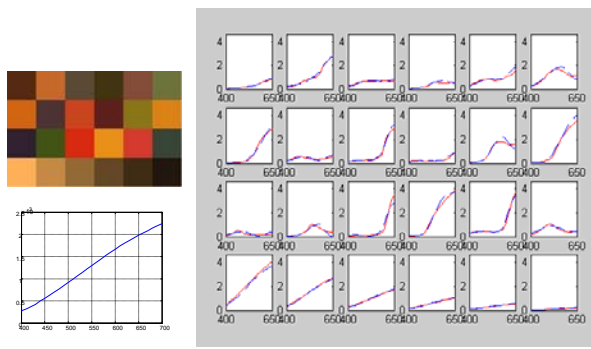
### Macbeth approximations: Appearance



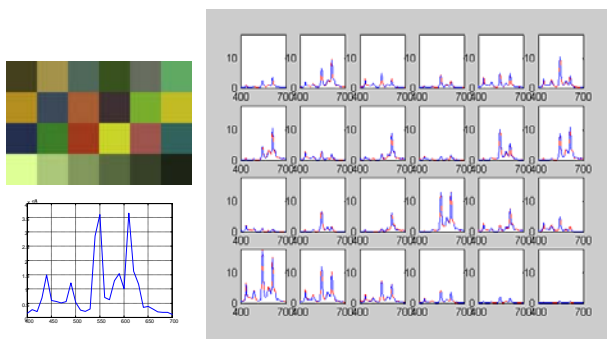
### Macbeth approximations: Appearance



### Example: MCC under tungsten illumination



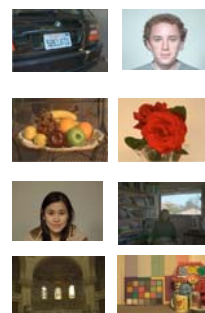
### Example: MCC under fluorescent illuminant



### Spectral scene database

Spectral radiance at each pixel in the sampled scene  
photons/(s sr nm m<sup>2</sup>)

- High-dynamic-range, High-resolution, Multi-spectral images
- Nikon D100, 2 additional color filters, exposure bracketing
- Linear models for surfaces and illuminants
- Natural scene dynamic range < 10,000:1

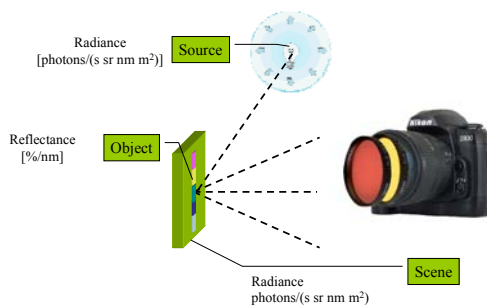


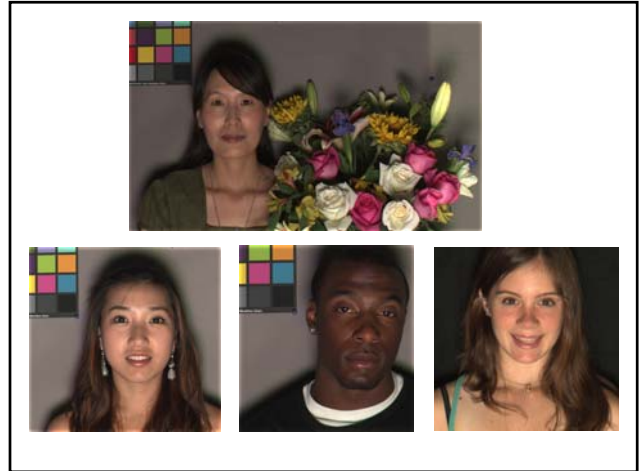
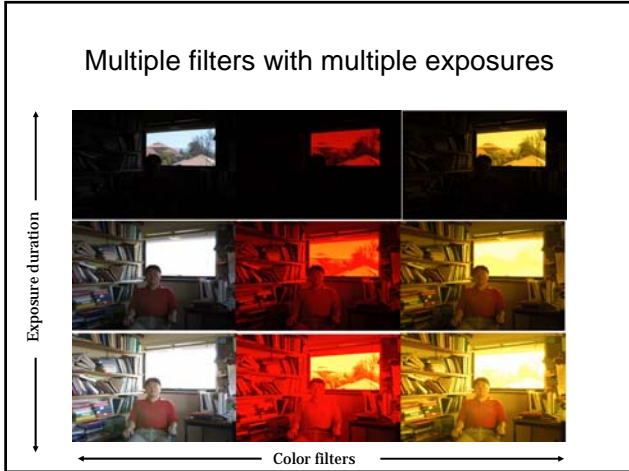
### Application: Spectral scene database

Using multi-capture methods and linear models to create calibrated, high dynamic range scene data

### High-dynamic-range multi-spectral Image capture

Multiple exposures, Multiple color sensors





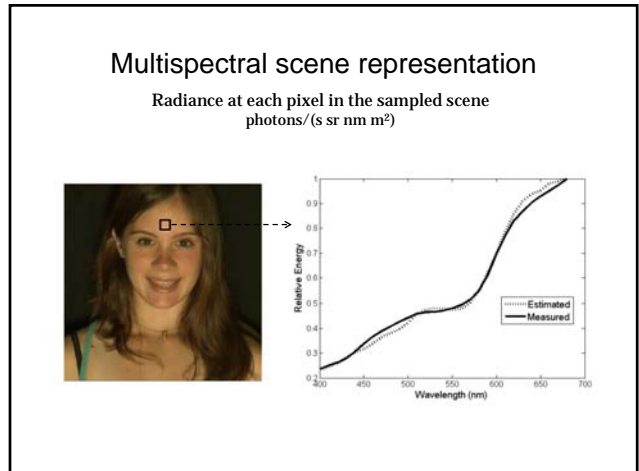
### Multispectral imaging system LED-based lighting system

- To address the need for shorter image acquisition times, we developed a multispectral imaging system that captures multiple acquisitions during a rapid sequence of differently colored LED lights.

Array of 216 LEDs with clusters of 8 different types of LEDs

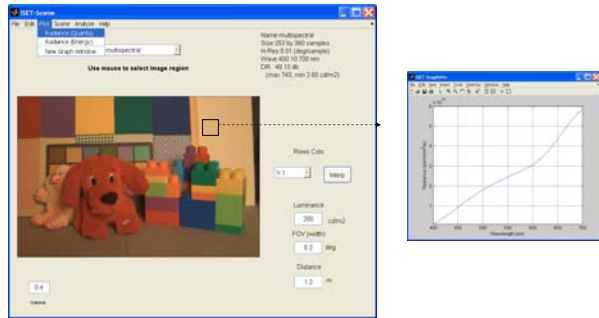
LEDs sample wavelengths between 380 and 1170 nm

We coupled the LED-based lighting system to a Nikon D2xs digital SLR camera which gives 12 MP images for obtaining high spatial resolution scenes.

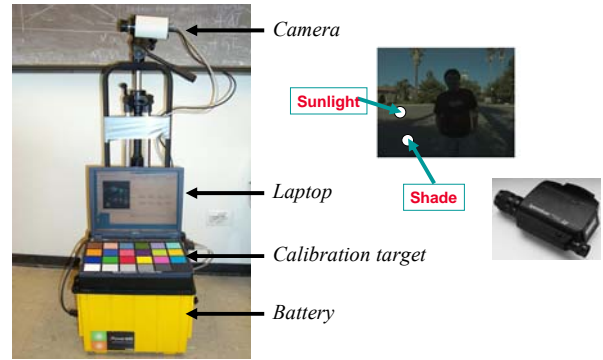


## Tools for managing spectral data

Radiance at each pixel in the sampled scene  
photons/(s sr nm m<sup>2</sup>)



## HDR: How high is high?



## High Dynamic Range Imaging of Natural Scenes

Feng Xiao, Jeffrey M. DiCarlo, Peter B. Catrysse and Brian A. Wandell  
Stanford University  
Stanford, California

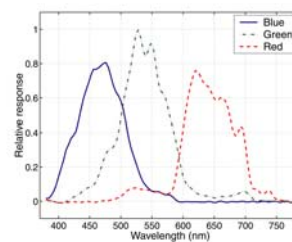
Volume 10/Nov. 2002/ Color Imaging Conference

### Rendering high dynamic range images

Jeffrey M. DiCarlo<sup>1</sup> and Brian A. Wandell<sup>1\*</sup>  
<sup>1</sup>Department of Electrical Engineering, Stanford University, CA 94305  
*Proceedings of the SPIE* (2000), V. 3965, pp.392-401

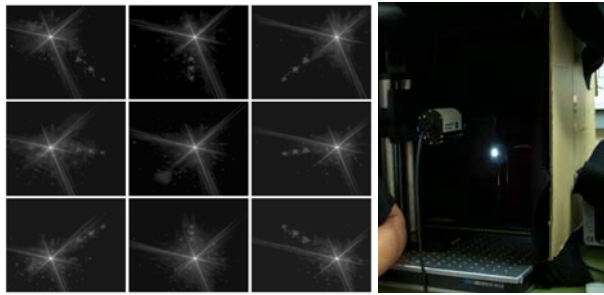


## System calibration



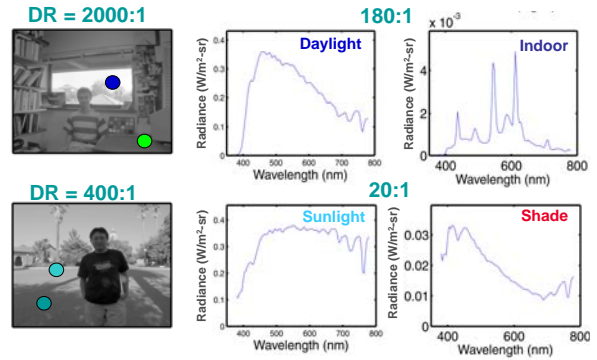
Dark current < 0.03 DV/sec  
Readout noise < 0.3 DV  
Exposure 40us ~ 15 minutes  
10-bit linear output  
1024x1280 Bayer pattern

### Lens calibration: Flare for HDR



Shift Variant (but linear)

### (1) HDR scenes have multiple illuminants



### Images captured using multiple exposures



Increases DR by factor:

$$\frac{T_{long}}{T_{short}} \quad (\text{Yang et al. 99})$$

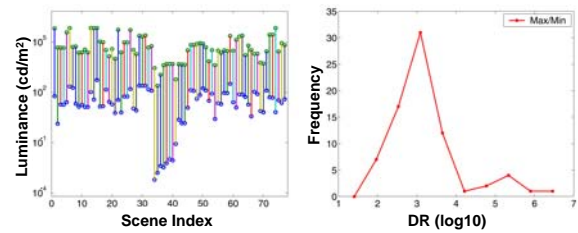
5-6 exposures at interval of 4 times

Measurable luminance range:

$[10^{-1} \sim 10^5 \text{ cd/m}^2]$

Dynamic range  $> 10^6:1$

### (2) Complete image DR is often 3-5 Log units



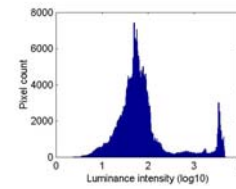
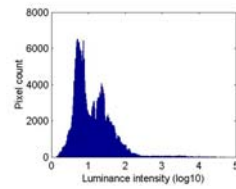
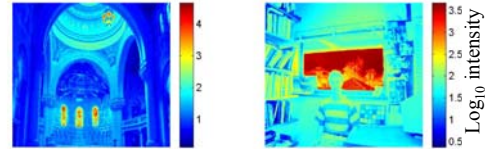
(3) The high range is often due to specular reflection or light

Original dynamic range = 6000:1

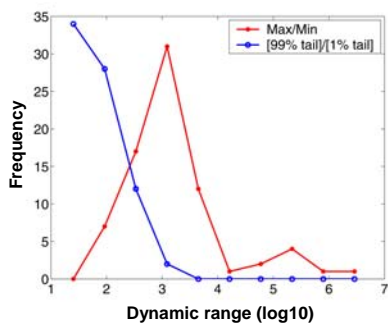


Effective dynamic range = [99% tail] / [1% tail]  
6000:1 reduced to 350:1

### HDR Database



Eliminating specular reflections or light sources  
brings DR down to 2-3 log units



Sensor calibration:  
Single exposure DR is 1000:1



10-bit linear output (w/o post-process)  
1024x1280 Bayer pattern  
Dark current < 0.03 DV/sec  
Readout noise < 0.3 DV

Image DR depends on the scene spatial structure

Scene DR > 2000:1

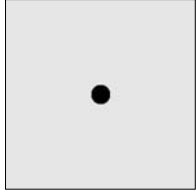


Image DR < 50:1

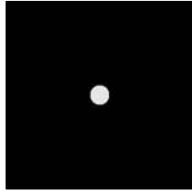


Image DR > 2000:1

HDR image database

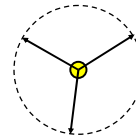


## Optics - Outline

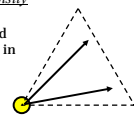
- Measuring light: Radiometry – Photometry
  - Irradiance and Illuminance
- Forming images: Imaging Optics
  - Image location and size
  - Image irradiance
  - Image resolution
- Modeling of imaging optics
  - Analytical models (diffraction-limited)
  - Numerical models (ray-tracing)

## Measuring Light - Radiometry

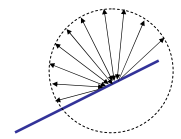
Radiant flux  
Light emitted from a point



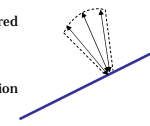
Radiant intensity  
Light emitted from a point in a solid angle



Irradiance  
Light density incident on a plane

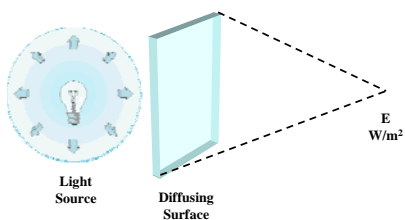


Radiance  
Light scattered from an extended source in a given direction



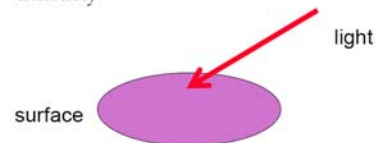
## Irradiance at the Image

- Irradiance (E)
  - incident radiant flux density:  $E = d\phi / dA$  [W/m<sup>2</sup>]



## Irradiance (E) properties

- ❖ Light power per unit area (watts per square meter) incident on a surface.
- ❖ If surface tilts away from light, same amount of light strikes bigger surface (less irradiance)(no foreshortening)
- ❖ E times pixel area times exposure time -> pixel intensity



Ramani (Maryland)

## Radiometry – Photometry Equivalents

- Source Radiance
  - Symbol:  $L$
  - Units:  $W/(sr\ m^2)$
- Source Luminance
  - Symbol:  $L_v$
  - Units:  $cd/m^2$
- Image Irradiance
  - Symbol:  $E$
  - Units:  $W/m^2$
- Image Illuminance
  - Symbol:  $E_v$
  - Units:  $lux = lm/m^2$



Courtesy P. Catrysse

## Radiance/Irradiance – Units

Q = energy (joules)  
 t = time (sec)  
 $\omega$  = solid angle (steradian)  
 $A$  = area (meter<sup>2</sup>)  
 $\theta$  = angle incident to plane  
 watt=joule/sec

Term	Definition	Application	SI Unit
Radiant Flux	$F = \frac{\Delta Q}{\Delta t}$	Total quantity of light emitted from a point	watt
Radiant Intensity	$I = \frac{\Delta F}{\Delta \omega}$	Total quantity of light emitted from a point in a given solid angle	watt /sr
Radiance	$L = \frac{\Delta I}{\Delta A_s \cos(\theta)}$	Light emitted or reflected from an extended source in a given direction	watt /sr/m <sup>2</sup>
Irradiance	$E = \frac{\Delta F}{\Delta A_r}$	Light density incident on a plane	watt /m <sup>2</sup>

Spectral quantities

## Luminance/Illuminance – Units

Q = energy (joules)  
 t = time (sec)  
 $\omega$  = solid angle (steradian)  
 $A$  = area (meter<sup>2</sup>)  
 $\theta$  = angle incident to plane  
 Watt = joule/sec

Term	Equation	Application	SI Unit
Luminous Flux	$F_v = K_m \int F_e(\lambda) V(\lambda) d\lambda$	Total quantity of light emitted from a point	lumen
Luminous Intensity	$I_v = \frac{\Delta F_v}{\Delta \omega}$	Total quantity of light emitted from a point in a given solid angle	candela (cd)
Luminance	$L_v = \frac{\Delta I_v}{\Delta A_s \cos(\theta)}$	Light emitted or reflected from an extended source in a given direction	lumens/m <sup>2</sup> (lux)
Illuminance	$E_v = \frac{\Delta L_v}{\Delta A_r}$	Light density incident on a plane	cd/m <sup>2</sup>

## Typical illuminance levels

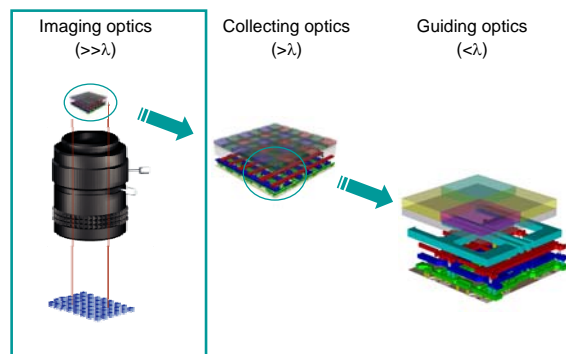
### Illuminance produced by various sources (lux = lm/m<sup>2</sup>)

Direct sunlight	110,000
Open shade	11,000
Overcast/dark day	110-1,100
Twilight	1.1-11
Full moon	0.11
Starlight	0.0011
Dark night	0.00011

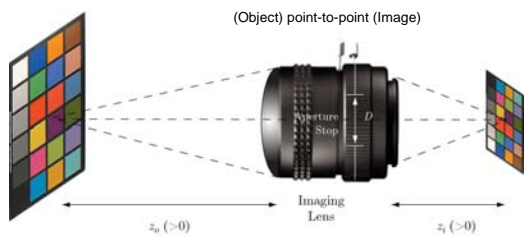
## Optics - Outline

- Measuring light: Radiometry – Photometry
  - Irradiance and Illuminance
- Forming images: Imaging Optics
  - Image location and size
  - Image irradiance
  - Image resolution
- Modeling of imaging optics
  - Analytical models (diffraction-limited)
  - Numerical models (ray-tracing)

## Optics of digital camera systems

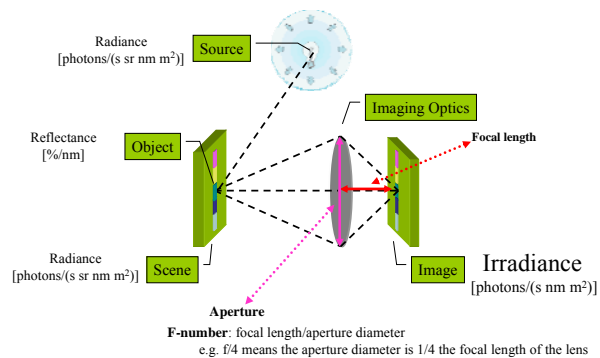


## Imaging optics: The basics



- Irradiance and exposure (# photons)
- On-axis vs. off-axis (relative illumination)
- PSF and MTF (resolution)

## Formation of the irradiance image



### Optics transform radiance to irradiance (photons/sec/nm/meter<sup>2</sup>)

Definition: Light density incident on a detector surface

Units: photons/(s nm m<sup>2</sup>)

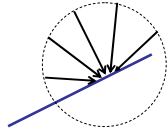


Image Irradiance

Units: photons/(s nm m<sup>2</sup>)  
W/(nm m<sup>2</sup>)

Photometric Conversion

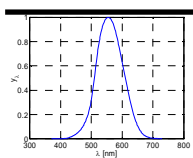
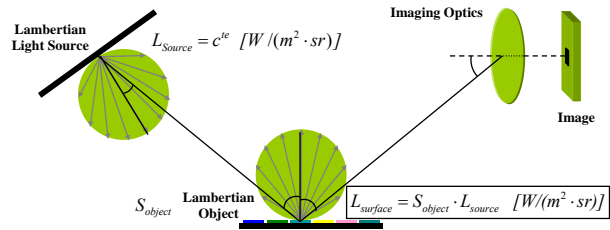


Image Illuminance

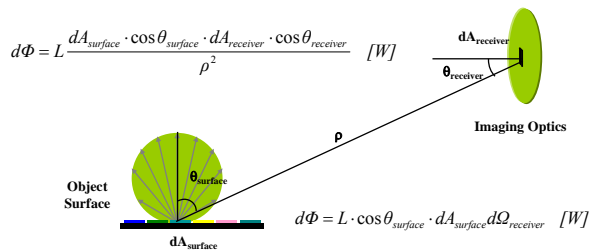
Units: lux = lm/m<sup>2</sup>

### Image irradiance calculation

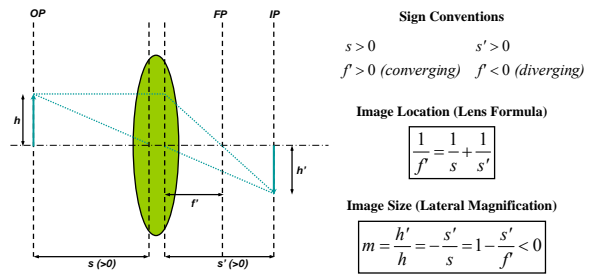


We will derive the number of photons incident at the imaging plane of an imaging system based on the number of photons in the imaged scene

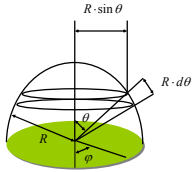
### From scene to imaging optics



### 1. Image formation



## 2. Integrating over solid angles



Annular Element of Solid Angle

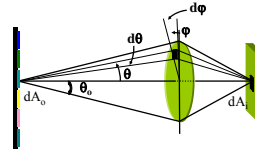
$$d\Omega = \frac{(2\pi \cdot R \cdot \sin \theta) \cdot (R \cdot d\theta)}{R^2}$$

$$\Downarrow$$

$$d\Omega = 2\pi \cdot \sin \theta \cdot d\theta$$

$$d^2\Phi = L \cdot \cos \theta \cdot dA_o d\Omega \quad [W] \quad \Rightarrow \quad d^2\Phi = 2\pi \cdot L \cdot dA_o \sin \theta \cdot \cos \theta \cdot d\theta \quad [W]$$

## 3. On-Axis image irradiance



Flux collected by Imaging Optics

$$d\Phi = \int_0^{2\pi} d\phi \int_0^{\theta_o} L \cdot dA_o \cos \theta \cdot \sin \theta \cdot d\theta \quad [W]$$

Image Irradiance

$$E = \frac{d\Phi}{dA_i} = \pi \cdot L \cdot (\sin \theta_o)^2 \frac{dA_o}{dA_i} \quad [W/m^2]$$

$\Downarrow$

$$E = \pi \cdot L \cdot (\sin \theta_o)^2 \quad [W/m^2]$$

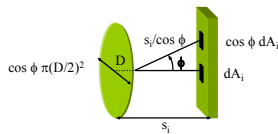
$\Downarrow$

$$E \cong \frac{\pi}{4} \frac{1}{(f/\#)^2 (1-m)^2} L \quad [W/m^2]$$

Sine Condition:  $dA_o/dA_i = (\sin \theta_i)^2 / (\sin \theta_o)^2$

$$s_i = f(1-m) \quad \sin \theta_i = \frac{D/2}{\sqrt{s_i^2 + (D/2)^2}}$$

## 4. Off-axis image irradiance



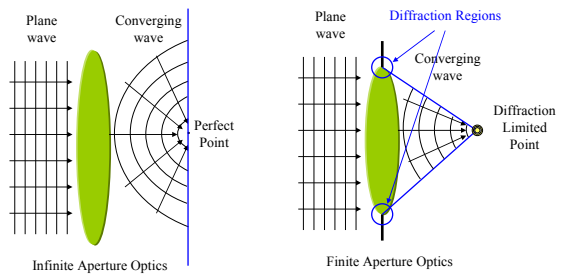
On-axis Image Irradiance

$$E \cong \frac{\pi}{4} \frac{L}{(f/\#)^2 (1-m)^2} \quad [W/m^2]$$

Off-axis Image Irradiance

$$E \cong \frac{\pi}{4} \frac{L}{(f/\#)^2 (1-m)^2} \cos^4 \phi \quad [W/m^2]$$

## Resolution limited by diffraction

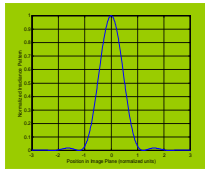


Diffraction (at aperture edges) is the fundamental reason why it is impossible to focus a plane wave into a perfect geometric point

## Diffraction in space domain

Diffraction is the fundamental reason why an object point-source spreads out to form a finite image spot

For an optical system with circular aperture the finite image spot forms an Airy Disk (Point Spread Function)



$$PSF(x) = \frac{[2 \cdot J_1(u)]^2}{u^2}$$

with  $u = \frac{\pi \cdot x}{\lambda \cdot f/\#}$   
and  $x_{cutoff} = 1.22 \cdot \lambda \cdot f/\#$



## Diffraction in frequency domain

- The PSF and the MTF of a diffraction limited optical imaging system are related

Point Spread Function (PSF)

Modulation Transfer Function (MTF)

$$PSF(x) = \frac{[2 \cdot J_1(u)]^2}{u^2}$$

$$\text{with } u = \frac{\pi \cdot x}{\lambda \cdot f/\#}$$

$$\text{and } x_{cutoff} = 1.22 \cdot \lambda \cdot f/\#$$

$\mathfrak{F}$

$\Rightarrow$

$\Leftarrow$

$\mathfrak{F}^{-1}$

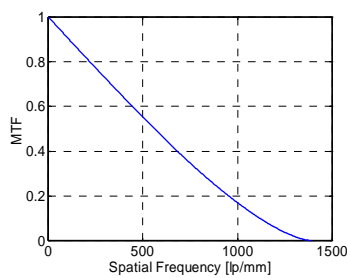
$$MTF(\omega) = \frac{2}{\pi} (\varphi - \cos \varphi \cdot \sin \varphi)$$

$$\text{with } \varphi = \cos^{-1}(\lambda \cdot \omega \cdot f/\#)$$

$$\text{and } \omega_{cutoff} = \frac{1}{\lambda \cdot f/\#}$$

- In general the optical MTF is difficult to find analytically

## Example: Diffraction-limited lens



$f/\# = 1.2$        $\lambda = 600 \text{ nm}$

$$MTF(\omega) = \frac{2}{\pi} (\varphi - \cos \varphi \cdot \sin \varphi)$$

$$\text{with } \varphi = \cos^{-1}(\lambda \cdot \omega \cdot f/\#)$$

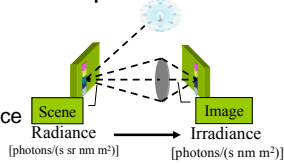
$$\text{and } \omega_{cutoff} = \frac{1}{\lambda \cdot f/\#}$$

## Optics - Outline

- Measuring light: Radiometry – Photometry
  - Irradiance and Illuminance
- Forming images: Imaging Optics
  - Image location and size
  - Image irradiance
  - Image resolution
- Modeling of imaging optics
  - Analytical models (diffraction-limited)
  - Numerical models (ray-tracing)

## Analytical models for optics

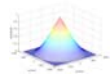
1. Convert radiance to irradiance



2. Apply relative illumination (cosine fall-off)



3. Calculate OTF given the f-number (diffraction limited)



1. Converting radiance to irradiance

### Irradiance formula

$$I = \frac{\pi}{1 + 4(f\#)^2(1 + |m|)} R$$

$R$  = radiance  
 $f\#$  = effective f-number  
 $m$  = magnification

## 2. Applying relative illumination (Cos-4th fall off)

For paraxial conditions, near the optical axis:

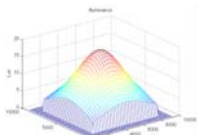


$$\text{Relative illumination} = (h/S)^4$$

$h$  = image field height

$$S = \sqrt{d^2 + h^2}$$

$d$  = distance from lens to image plane



## 3. Applying diffraction-limited OTF

$$OTF = \begin{cases} \frac{2}{\pi} [\arccos(\rho) - \rho\sqrt{1-\rho^2}] & (\rho < 1) \\ 0 & (\rho \geq 1) \end{cases}$$

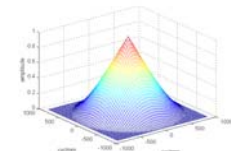
$\rho = f/(A/d\lambda)$  (normalized frequency)

$f$  = frequency in cycles/meter

$A$  = aperture diameter (m)

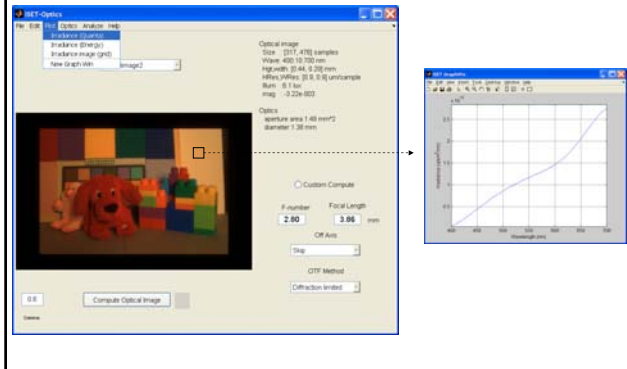
$d$  = distance from aperture to detector

$\lambda$  = wavelength

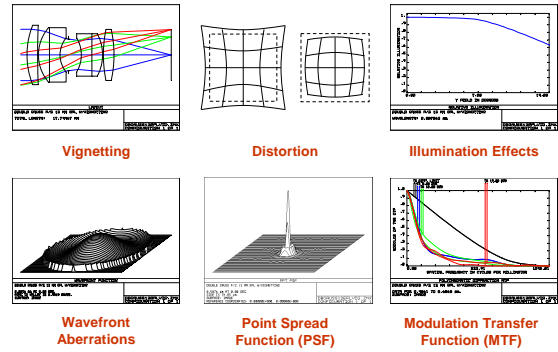


OTF at 550 nm  
 F-number: 2  
 Focal Length: 3.86 mm

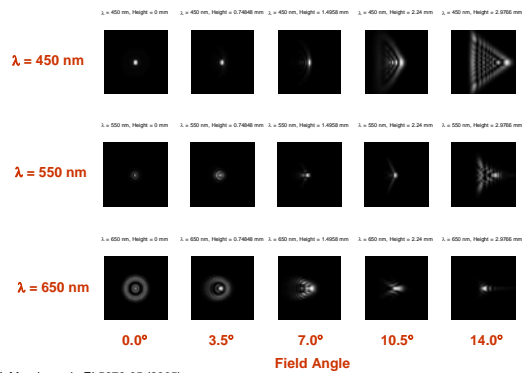
## Calculated irradiance image



## Numerical models for optics



## Variation with field and wavelength



P. Y. Maeda et al., EI 5678-05 (2005)

## Image formation modeling

- Must be shift-variant
- Computed over many wavelengths
- Development and verification of software needed to compute physical optics data is a large task
  - Integrate use of commercial lens design software like CodeV® and Zemax® to supply needed data
  - Use macro and programming languages in these programs to automate data generation process

## Linear shift-variant imaging system

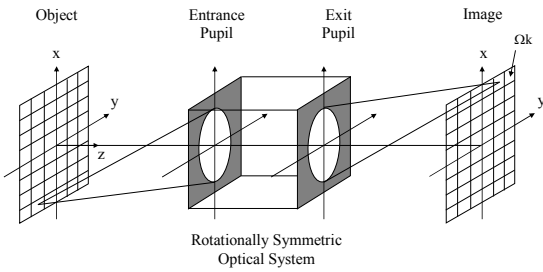


Image segmented into isoplanatic regions:  $\Omega_1, \Omega_2, \dots, \Omega_n$

P. Y. Maeda, P. B. Catrysse, and B. A. Wandell, Proc. SPIE Int. Soc. Opt. Eng. **5678**, 48 (2005)

## Linear shift-variant image formation

Let  $\Omega_k, k = 1, 2, \dots, n$ , be isoplanatic sections in the image plane, then:

$$I_{image}(x, y, \lambda) = \sum_k PSF_{\Omega_k}(x, y, \lambda) \otimes I_{ideal, \Omega_k}(x, y, \lambda)$$

or

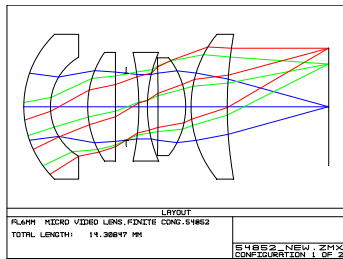
$$I_{image}(x, y, \lambda) = F^{-1} \left\{ \sum_k OTF_{\Omega_k}(f_x, f_y, \lambda) \cdot F \{ I_{ideal, \Omega_k}(x, y, \lambda) \} \right\}$$

To insure the radiometry is preserved, the OTFs are normalized.

$$MTF(f_x, f_y, \lambda) = |OTF(f_x, f_y, \lambda)| = \frac{F \{ PSF(x, y, \lambda) \}}{F \{ PSF(x, y, \lambda) \} \Big|_{f_x=0, f_y=0}}$$

P. Y. Maeda, P. B. Catrysse, and B. A. Wandell, Proc. SPIE Int. Soc. Opt. Eng. **5678**, 48 (2005)

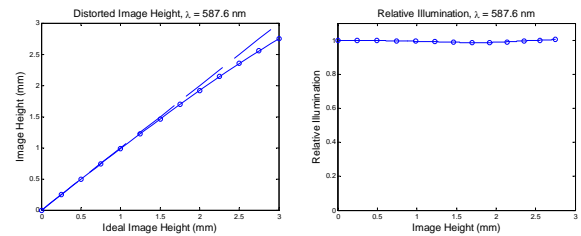
## Example: EO54852 lens



6.0 mm EFL, NA = 0.264, 250 mm object distance  
52.4 deg FOV. Vignetting at the larger off-axis image points  
Distortion < -10% (barrel type)  
The lens is not diffraction limited

P. Y. Maeda, P. B. Catrysse, and B. A. Wandell, Proc. SPIE Int. Soc. Opt. Eng. **5678**, 48 (2005)

## Distortion and relative illumination



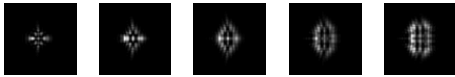
P. Y. Maeda, P. B. Catrysse, and B. A. Wandell, Proc. SPIE Int. Soc. Opt. Eng. **5678**, 48 (2005)

### EO54852 PSFs: $\lambda = 587.6$ nm

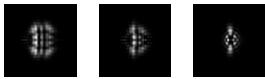
587.6 nm, 0 mm    587.6 nm, 0.24665 mm    587.6 nm, 0.49256 mm    587.6 nm, 0.73866 mm    587.6 nm, 0.97909 mm



587.6 nm, 1.2152 mm    587.6 nm, 1.4534 mm    587.6 nm, 1.6939 mm    587.6 nm, 1.9097 mm    587.6 nm, 2.1269 mm



587.6 nm, 2.3371 mm    587.6 nm, 2.538 mm    587.6 nm, 2.7277 mm

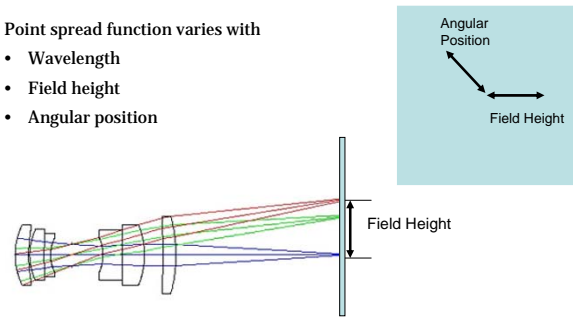


P. Y. Maeda, P. B. Catrysse, and B. A. Wandell, Proc. SPIE Int. Soc. Opt. Eng. 5678, 48 (2005)

### Ray trace modeling

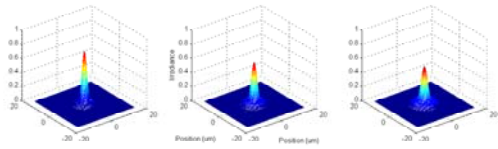
Point spread function varies with

- Wavelength
- Field height
- Angular position

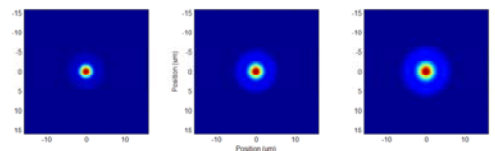


### Ray trace point spread: wavelength

475 nm    550 nm    625 nm

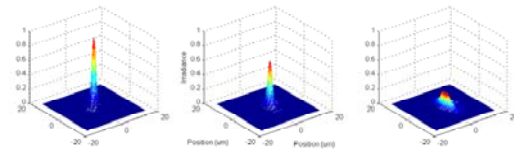


0.5mm field height

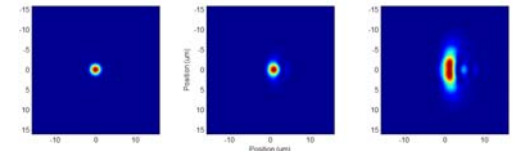


### Ray trace point spread: field heights

0 mm    0.57 mm    0.88 mm



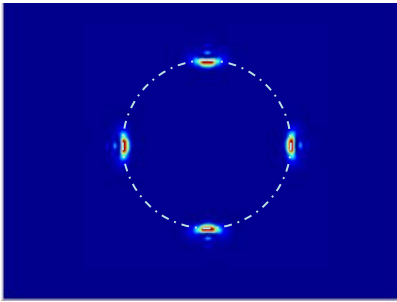
(400 nm)



### Point spreads are rotated at every pixel

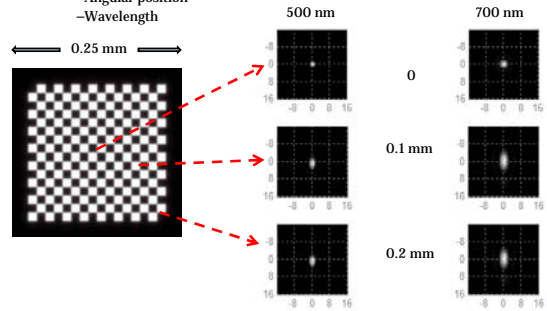
Four examples shown

Field height: 0.88 mm  
Wavelength: 400 nm

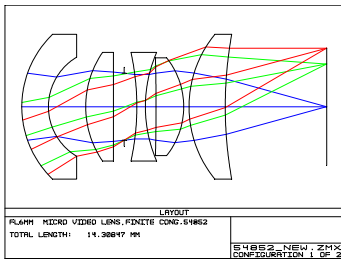


### Computation of irradiance Image

For each wavelength  
Convolve pixel with point spread function that varies with  
-Field height  
-Angular position  
-Wavelength

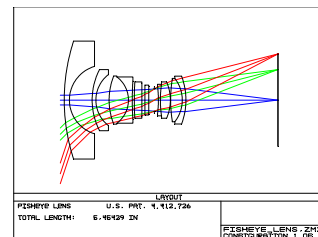


### EO54852 Micro Video Lens



6.0 mm EFL, NA = 0.264,  $m = -0.024$ , 250 mm object distance  
52.4 deg FOV, Distortion < -10% (barrel type distortion)  
The lens is not diffraction limited

### Fisheye lens



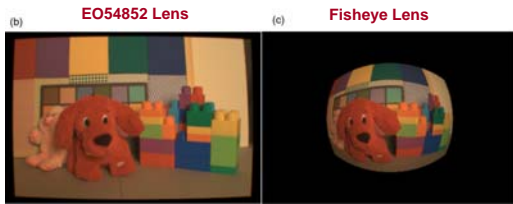
25.25 mm EFL,  $f/4$ , 140 deg FOV  
Distortion < -56.6%  
The lens is not diffraction limited

## Calculated irradiance image

For each wavelength

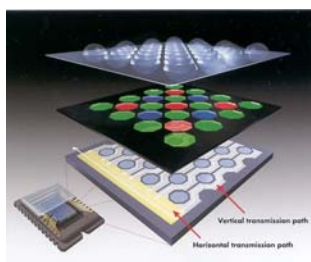
Convolve pixel with point spread function that varies with

- Field height
- Angular position
- Wavelength



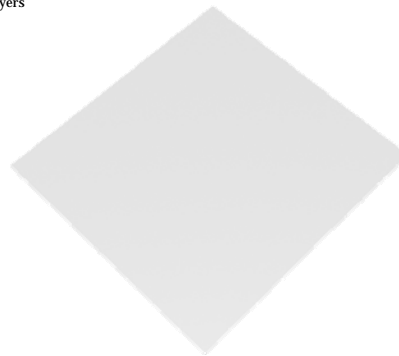
## Imagers

- CMOS imagers
- CCD imagers
- Imager-related components
- Figures of merit
- Triple-well (Foveon)
- HDR and high speed imaging (Pixim)



Building up the CMOS imager layers

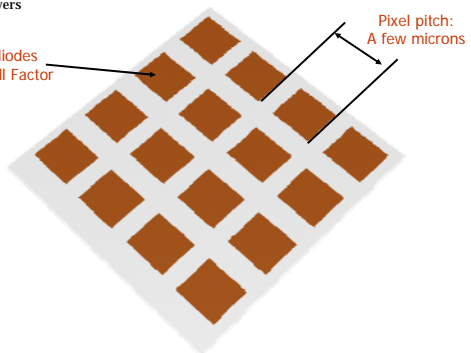
## CMOS



Courtesy R. Motta, Pixim

Building up the CMOS imager layers

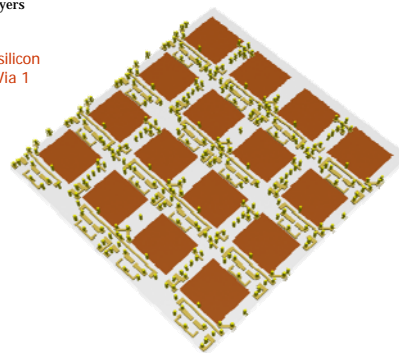
Photodiodes  
~50% Fill Factor



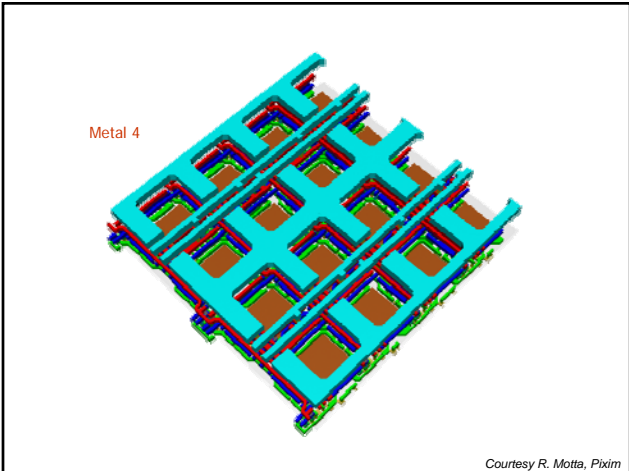
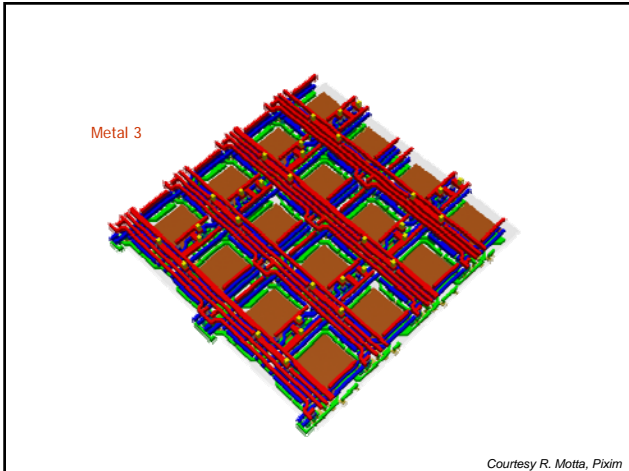
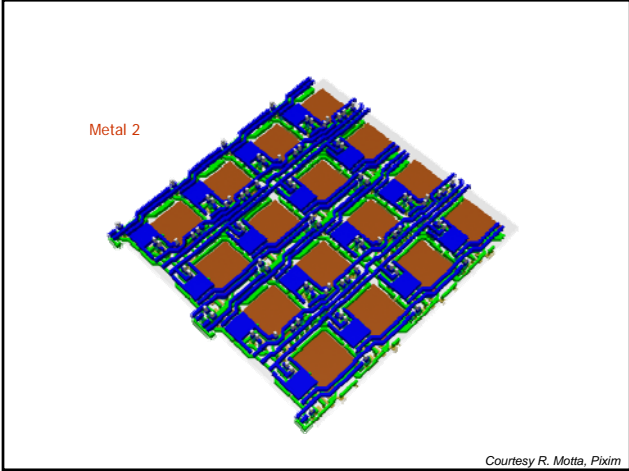
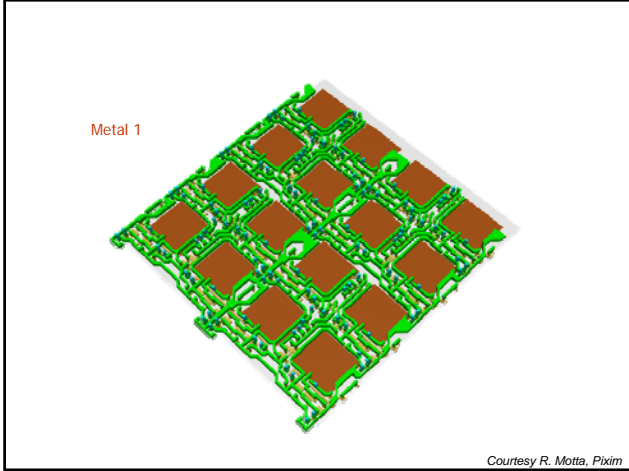
Courtesy R. Motta, Pixim

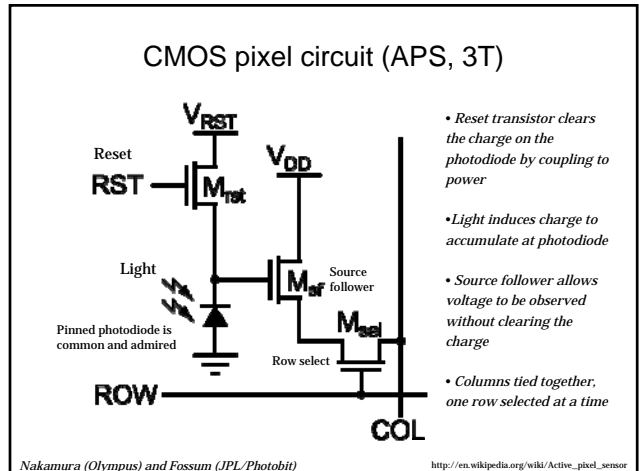
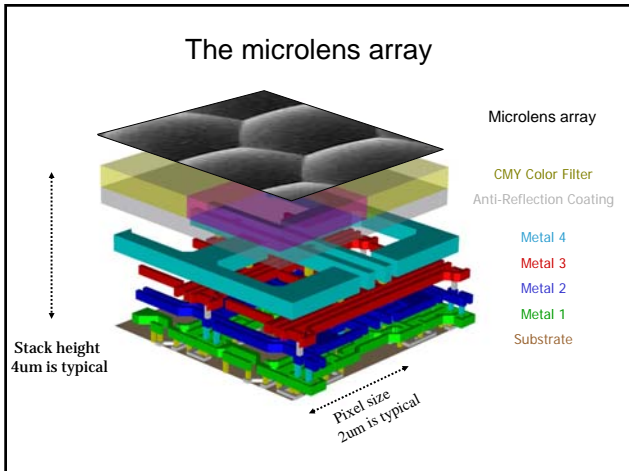
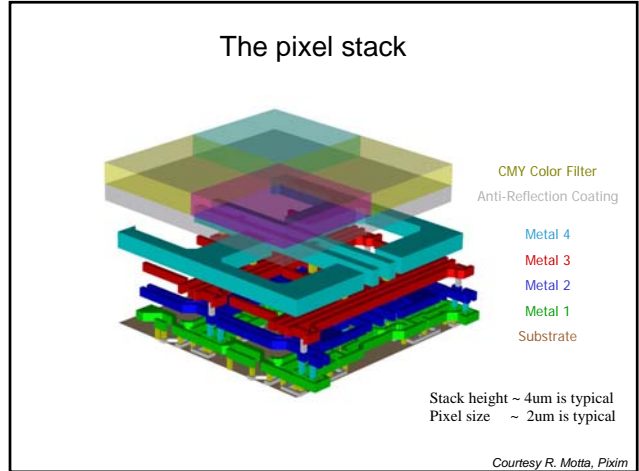
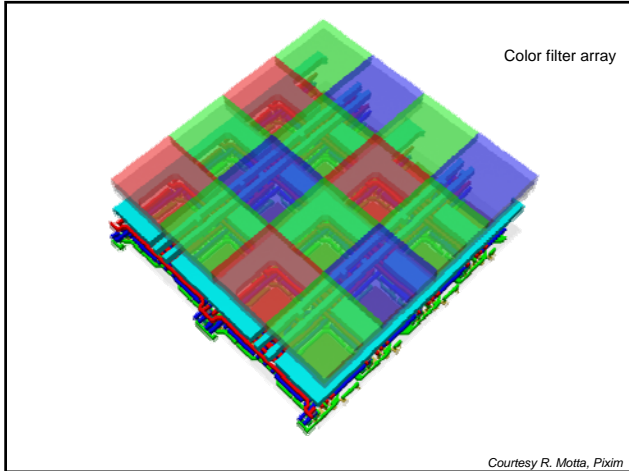
Building up the CMOS imager layers

Polysilicon  
& Via 1



Courtesy R. Motta, Pixim





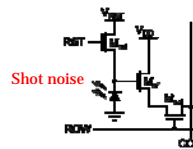
### Convert photons to electrons at the pixel

$$e = GT \int i(\lambda) s(\lambda) f(\lambda) d\lambda$$

- G = geometric factors (pixel optical cross-section, m<sup>2</sup>)
- T = integration time (s)
- i = spectral irradiance (photons/s/nm/m<sup>2</sup>)
- s = sensor spectral responsivity (electrons/photon)
- f = filter and media transmissivities
- e = electrons at the pixel
- Check units: e = m<sup>2</sup> s ∫ p/s/nm/m<sup>2</sup> e/p d(nm)

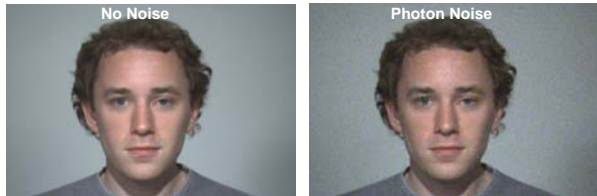
### Pixel circuit noise

- Electrons (Poisson random variable) are converted to volts by circuit properties (conversion gain)  $v_0 = c\tilde{e}$



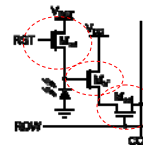
### Photon noise

No spatial structure  
Averaged away by repeated measures  
Visible at an SNR of 30 dB (1000 photons, Xiao et al.)



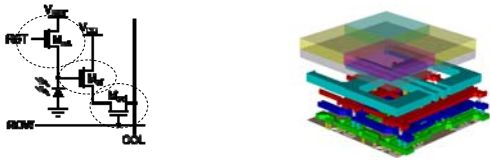
### Pixel noise

- Electrons (Poisson random variable) are converted to volts by circuit properties (conversion gain)  $v_0 = c\tilde{e}$
- Circuit imperfections introduce dark voltage that grows over time (T\*V/sec)  $v_1 = v_0 + T\bar{d}$
- The acts of reading and resetting the voltage are noisy (and can be grouped)  $v_2 = v_1 + \bar{r}_0 + \tilde{r}_1$



## Dark noise

Thermally generated electrons are indistinguishable from photo-generated electrons. They constitute a noise source known as 'Dark Current.' High end CCDs are designed (e.g., by cooling or in CMOS special circuitry) to reduce the amount of dark current.



## Read and reset noise

The act of reading data from a pixel and resetting the pixel level to an initialized state involve imperfect circuit events. These imperfections are sensor read and sensor reset noise. They can be separately identified, but it is convenient to group them.

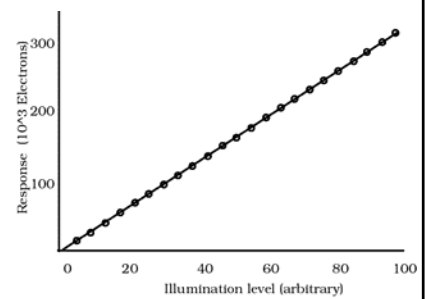


## Photon, read, reset, dark, prnu, dsnu

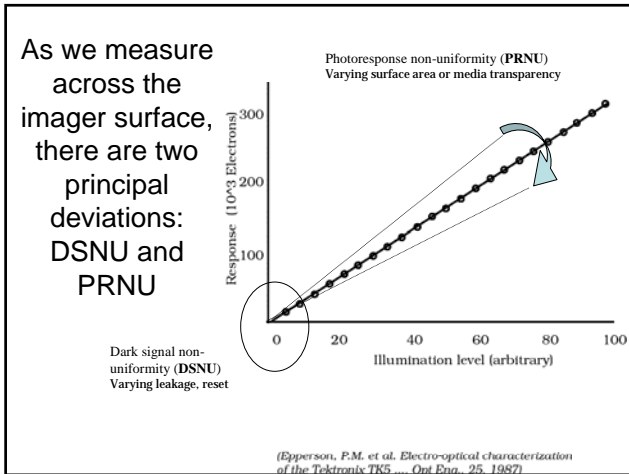
Adds to photon noise – so 1000 photons not enough for noise-free



CCD and CMOS transfer functions (photons to electrons) are linear  
(both are Si absorptions)



*(Epperson, P.M. et al. Electro-optical characterization of the Tektronix TK5 .... Opt Eng., 25, 1987)*



### Programmable gain/offset

- Electrons (Poisson random variable) are converted to volts by circuit properties (conversion gain)  $v_0 = c\tilde{e}$
- Circuit imperfections introduce dark voltage that grows over time ( $T^*V/sec$ )  $v_1 = v_0 + T\bar{d}$
- The acts of reading and resetting the voltage are noisy (and can be grouped)  $v_2 = v_1 + \bar{r}_0 + \tilde{r}_1$
- Variations in analog offset and gain; offset can be due to leakage and thus time-dependent  $v_3 = \tilde{a}(v_2 + T\tilde{o})$

Programmable gain amplifier and offset

### Column amplifier variation: A special place in imaging hell

Fixed column noise

Read Noise, Dark Voltage, PRNU, DSNU

Column FPN

In some systems, an entire column shares an amplifier. In this case, variations between the column amplifiers produce unwanted variations (as above). Consistent column variations are very noticeable and thus must be eliminated.

### Phenomonological model

$$v = \bar{Z}(\tilde{e} + \bar{X} + T\bar{Y})$$

Knowledge of the noise mechanisms is important for accurate simulation because the mechanisms have different temporal and spatial characteristics; in this form the equation does not make this important issue clear.

## Terminology: Temporal and fixed noise

- **Temporal noise:** Some noise terms differ across repeated reads of the pixel (e.g., read noise, reset noise, dark noise, photon noise); these can be averaged away

$$v_0 = c\tilde{e}$$

$$v_1 = v_0 + T\tilde{d}$$

- **Fixed pattern:** Some noise terms are simply variations of the device properties across the array; these are fixed over time (conversion gain, DSNU, PRNU, Column gain) and are not averaged away

$$v_2 = v_1 + \tilde{r}_0 + \tilde{r}_1$$

$$v_3 = \tilde{a}(v_2 + T\tilde{o})$$

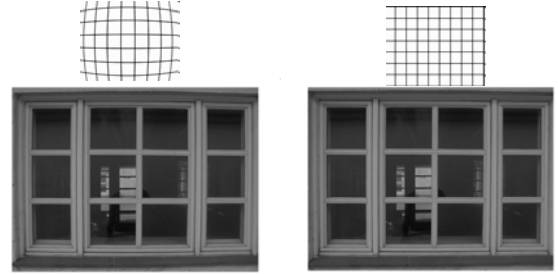
## System Design Metrics

- Geometric distortions
- Lens vignetting
- System MTF
  - Chromatic aberration
- Signal to noise Ratio
- Image quality tradeoffs
  - Resolution and sensitivity
  - Camera motion



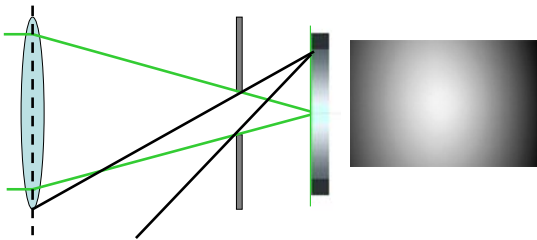
[www.imatest.com](http://www.imatest.com) and [www.dxo.com](http://www.dxo.com) sell software packages that include these measurements

## Measurement and characterization Geometric distortions

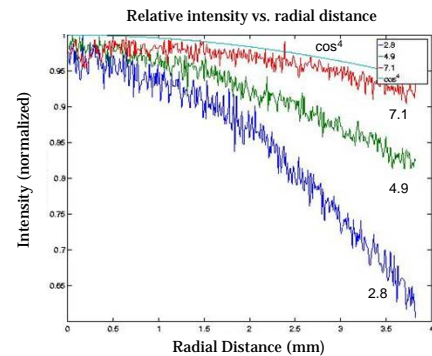


<http://scien.stanford.edu/class/psych221/projects/06/ddpatil/>

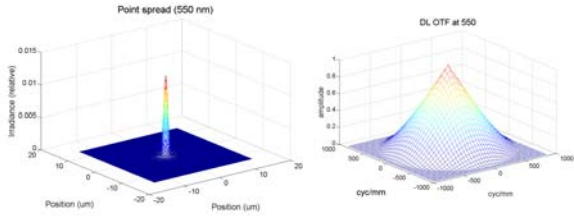
## Measurement and characterization Lens vignetting



## Measurement and characterization Lens vignetting

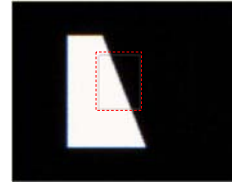


### Measurement and characterization Optical blur



### Measurement and characterization ISO 12233 Method for estimating the System Modulation Transfer Function (MTF)

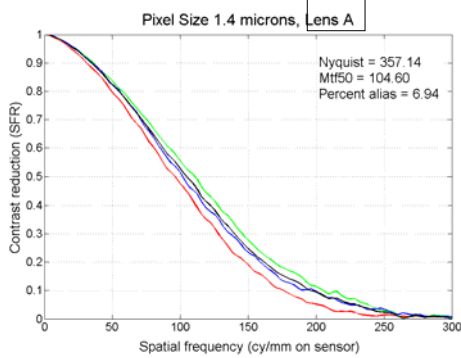
- The response in the vicinity of an edge is directly related to the line spread function (via a derivative).
- The line spread function is directly related to the MTF by a Fourier Transform.
- The ISO 12233 estimates the edge response to calculate the line spread, and then the MTF.



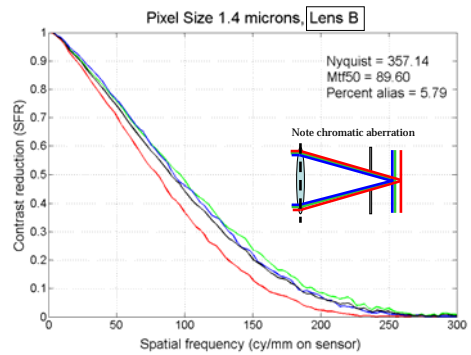
Reference  
Williams, D. and Burns, P. D. Low-Frequency MTF Estimation for Digital Imaging Devices using Slanted Edge Analysis. *Proceedings of the SPIE* Volume 5294, 2004.  
[http://www.isa.org/downloads\\_iso\\_tools.html](http://www.isa.org/downloads_iso_tools.html)

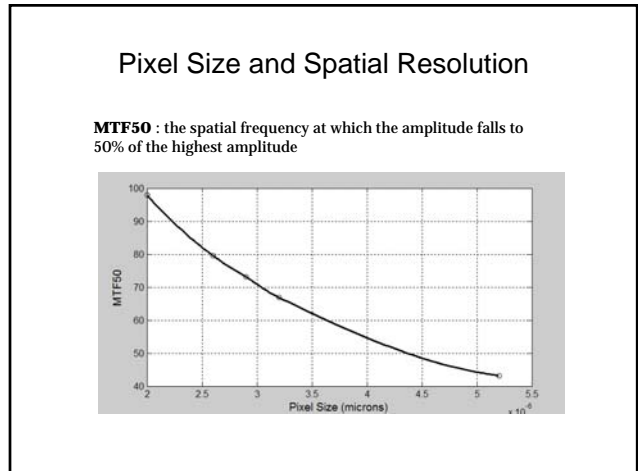
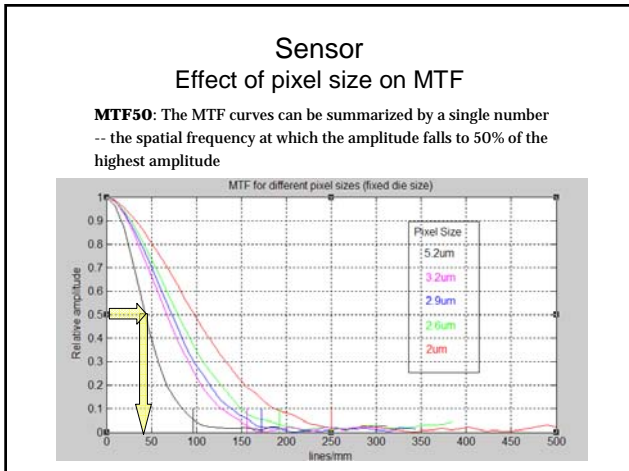
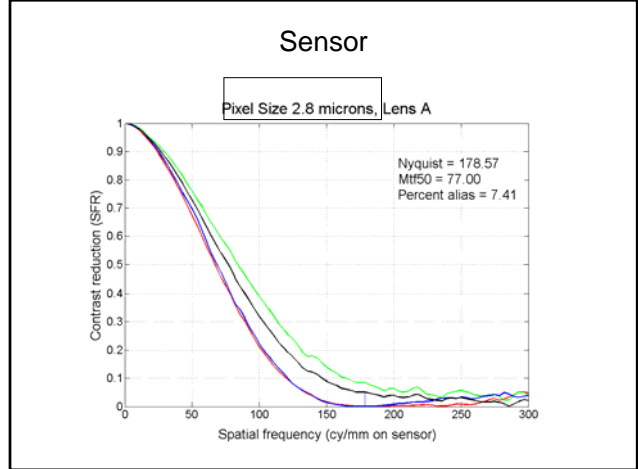
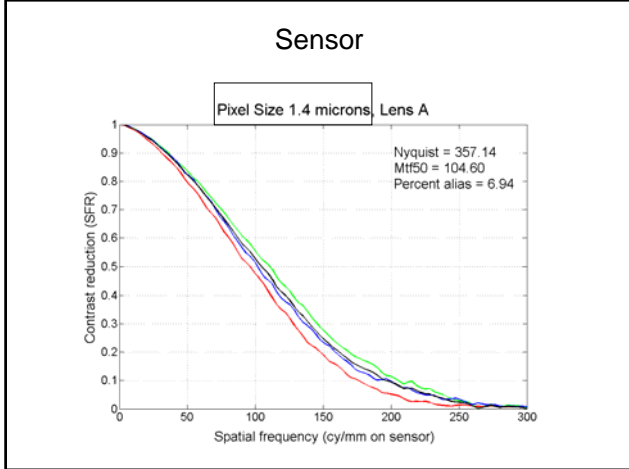
See also  
<http://www.imageval.com/public/Products/ISSET/ApplicationNotes/SlantedBarMTF.pdf>

### Measurement and characterization Optical blur



### Measurement and characterization Optical Blur





## Spatial Resolution (MTF) and Noise (SNR)

- MTF quantifies the effects of multiple system components on the spatial blur in the final image
- SNR quantifies the effects of multiple system components on noise in the final image

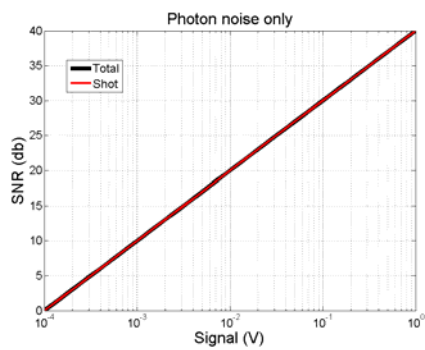
## Sensor Signal to noise ratio (SNR)

Signal level (units of electrons)

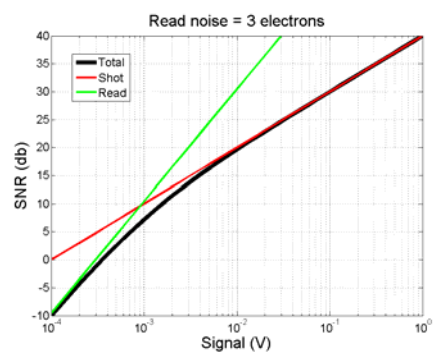
$$SNR(E_s) = 20 \log \left( \frac{E_s}{\sqrt{\sigma_s^2 + \sigma_R^2 + \sigma_{DSNU}^2 + \sigma_{PRNU}^2}} \right)$$

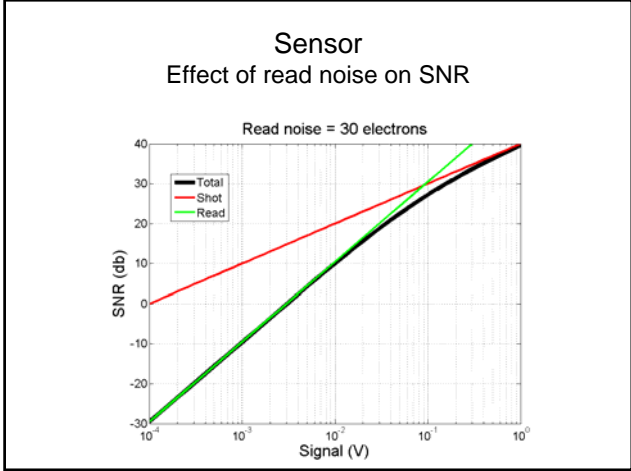
Variance in: Shot noise    Read and Reset noise    Dark signal non-uniformity    Photo-response non-uniformity

## Sensor Effect of photon noise on SNR



## Sensor Effect of read noise on SNR





- ### Image quality tradeoffs in sensor design
- Image quality is more than the number of pixels in the sensor
  - Resolution/sensitivity tradeoff
  - Camera motion

### Image quality tradeoffs

#### Resolution and sensitivity

- Resolution improves with decreasing pixel size
  - Make the pixel smaller, you will get more pixels/unit area
- Sensitivity suffers with decreasing pixel size
  - A smaller pixel captures fewer photons per unit time

### Photon noise

	Exposure Duration		
	1 ms	10 ms	100 ms
Pixel Size			
1.4 micron			
2.8 micron			

### Visibility of photon noise

Question: How many photons do you need to capture before you do not see the photon shot noise?



Answer: 1000 photons

- For all luminance values, 3 percent variation of luminance is detectable (Feng et al, 2005)
- For an ideal capture device with only photon shot noise, the noise threshold of 3 percent implies that the minimum number of photon absorptions per pixel should exceed 1000
- In terms of SNR, this is equivalent to requiring that image captured in the dark portion of a scene should exceed 33 db ( $-20 \log_{10}(1/0.03)$ ).

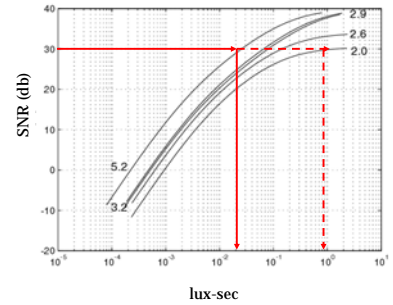
• F. Xiao, J. E. Farrell, B. Wandell, Psychophysical thresholds and digital camera sensitivity: The thousand photon limit. In Proceedings of the SPIE Electronic Imaging 2005 Conference, Santa Clara, CA, January 2005  
 • J. E. Farrell, F. Xiao, S. Kavusi, Resolution and light sensitivity tradeoff with pixel size In Proceedings of the SPIE Electronic Imaging 2006 Conference, Vol. 6069, 2006

### Sensor

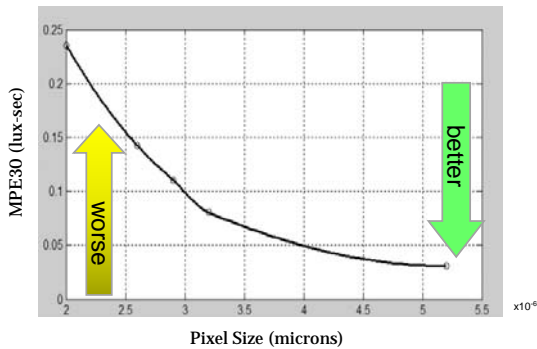
#### Minimum photometric exposure (MPE30)

MPE30: Minimum photometric exposure that will capture an image of a uniform field with

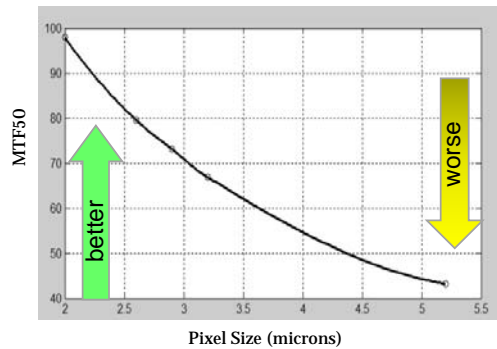
- less than 3% noise or, equivalently
- SNR > 30db

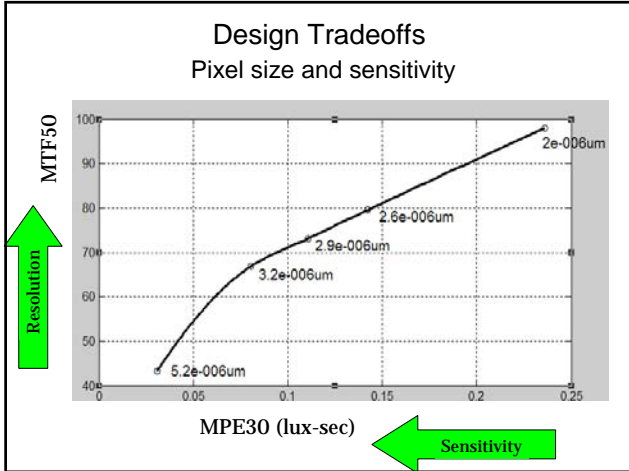


### Sensor MPE30 and pixel size



### Sensor MTF50 and pixel size





### Sensor Camera motion

Camera velocity from handshake is

- 4 deg/sec for typical camera phone (two-hand)
- 6 deg/sec (one-hand)

• F. Xiao, A. Silverstein, J. Farrell, Camera-motion and effective spatial resolution. *In Proceedings of the International Congress of Imaging Science, May 2006.*  
 • Xiao, F., Pincetti, J., John, G. and Johnson, K., Camera-motion and mobile imaging. *Proc. SPIE Vol. 6502, (2007).*

### Sensor Camera motion

Calculation of length of motion blur

1. Assume a uniformly illuminated surface with mean luminance (10 cd/m<sup>2</sup>)
2. T = Exposure duration (seconds) given MPE30 (expressed in lux-seconds) and illumination level (lux)
3. M<sub>l</sub> = Motion length (in degrees) = 4 degrees/sec x T
4. P<sub>deg</sub> = Pixels/degree = number of pixels/sensor field of view
5. M<sub>p</sub> = Motion length (in pixels) = M<sub>l</sub> x P<sub>deg</sub>

### Sensor Camera Motion

↓

Pixel Size (microns)	MPE30 (lux-secs)	Exposure Duration to achieve MPE30 (seconds)	Motion length (in degrees)	Motion length (in pixels)
2	0.24	0.2062	0.8249	32.9959
2.8	0.12	0.1031	0.4124	11.7857

2.8 micron pixel



2 micron pixel



### Sensor Effective pixel count

- The Effective Pixel Count (EPC) is a measure of the spatial resolution of a digital camera after optical blur, spatial blur and motion blur have been taken into account.

$$N_{EPC}(L) = S \times (2 \times MTF_{50}(L))^2$$

The number of effective pixels is a function of

- $L$ , illumination level
- $S$ , Sensor area (mm<sup>2</sup>)
- $MTF_{50}(L)$ , 50% cutoff frequency in line-pairs/mm

See F. Xiao, J. Farrell, P. Catrysse, B. Wandell, Mobile Imaging: The Big Challenge of the Small Pixel . In Proceedings of the SPIE Electronic Imaging 2006 Conference, 2009.

### Pixel optics

- Light inside imagers
  - # photons at the pixel is finite
  - Light collecting and guiding in pixels
  - Effects of pixel scaling

### 1,000 photon criterion (Xiao et al.)

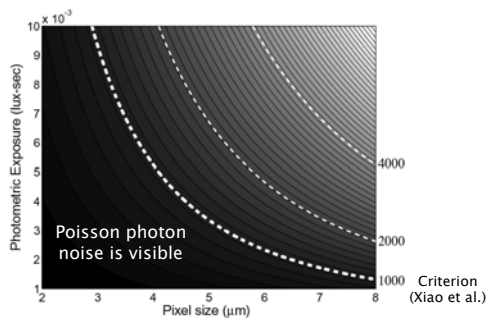
3% contrast-threshold for spatially uniform photon noise  
Photon noise becomes visible at 1,000 photons  
( $SNR = 1000/\sqrt{1000} = 33:1$ )



Figure 3: A pair of stimuli was displayed side by side: one was a uniform disk and the other one with noise superimposed.

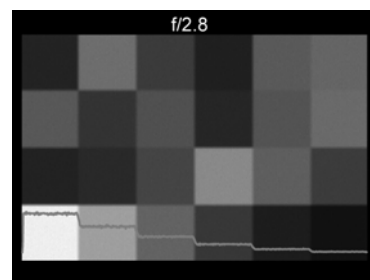
F. Xiao et al., EI 5678-07 (2005)

### Photon noise: Photons per pixel



P. B. Catrysse et al., EI 5678-01 (2005)

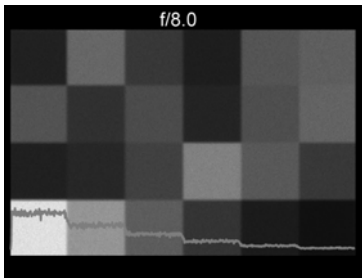
### Photon noise: Visibility (monochrome)



Luminance range: 12 – 1.25 cd/m<sup>2</sup>  
Exposure time: 10 ms

P. B. Catrysse et al., EI 5678-01 (2005)

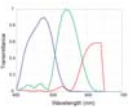
### Photon noise: Visibility (monochrome)



Luminance range: 12 - 1.25 cd/m<sup>2</sup>  
Exposure time: 10 ms

P. B. Catrysse et al., EI 5678-01 (2005)

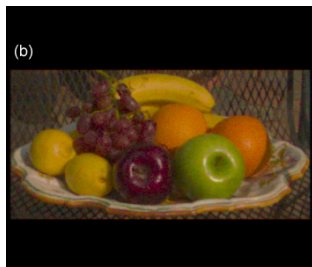
### Photon noise: Visibility (color)



Additional processing steps usually amplify noise in color images

Mean luminance: 100 cd/m<sup>2</sup>  
Exposure time: 10 ms  
Imaging lens: f/2.8

### Photon noise: Visibility (color)



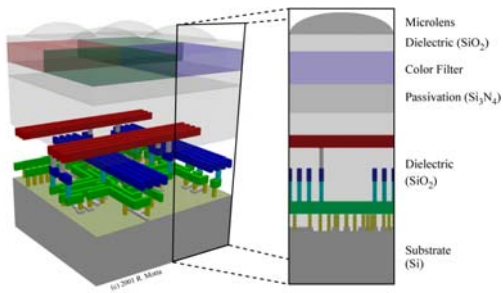
Mean luminance: 10 cd/m<sup>2</sup>  
Exposure time: 10 ms  
Imaging lens: f/2.8

### Photon noise: Visibility (color)

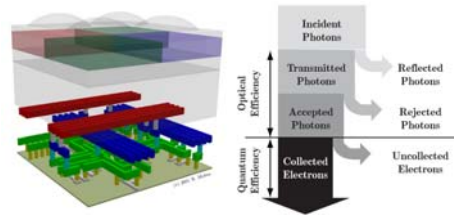


Mean luminance: 10 cd/m<sup>2</sup>  
Exposure time: 10 ms  
Imaging lens: f/8.0

### Typical image sensor pixel



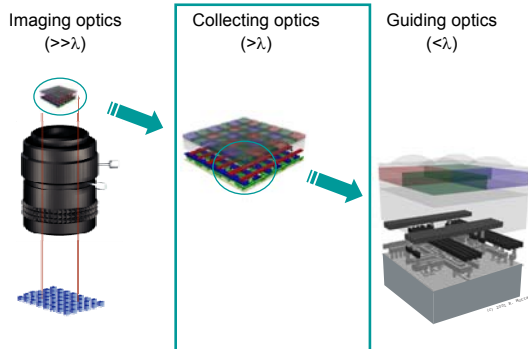
### Pixel optics: Capture every photon!



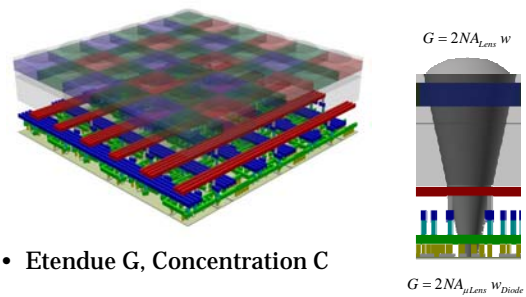
- Pixel QE = Optical Efficiency (OE) + internal QE
- Optical X-talk (OX) (in addition to electrical X-talk)
- Collecting and guiding: maximize OE & minimize OX

P. B. Catrysse *et al.*, JOSA A, Vol. 19, No. 8 (2002)

### Optics of digital camera systems



### Collecting optics: The basics



- Etendue  $G$ , Concentration  $C$

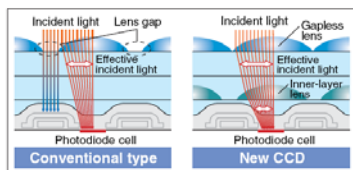
$$C = \frac{NA_{\mu\text{Lens}}}{NA_{\text{Lens}}} \approx \frac{n_{\mu\text{Lens}}}{n_{\text{Lens}}} \frac{f/\#\text{Lens}}{f/\#\mu\text{Lens}}$$

$$G = 2NA_{\text{Lens}} w$$

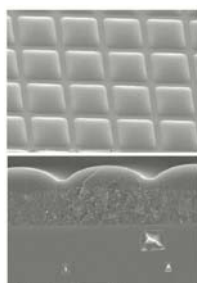
$$G = 2NA_{\mu\text{Lens}} w_{\text{Diode}}$$

## Collecting optics: Examples

- Micro-lens arrays
  - Single → double
  - Circular → gapless
  - Refractive → diffractive/Fresnel

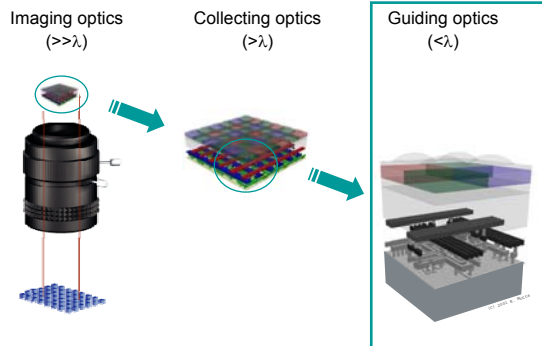


Panasonic CCD technology



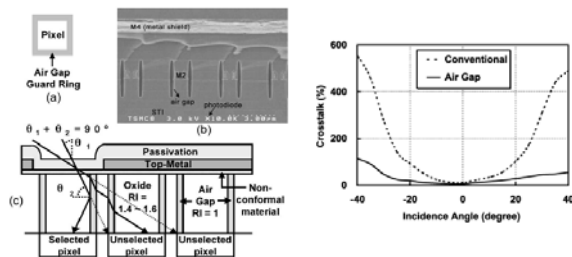
H. Rhodes *et al.*, IEEE (2004)

## Optics of digital camera systems



## Guiding optics: Examples

- Air gap “Guard ring”: TIR waveguide



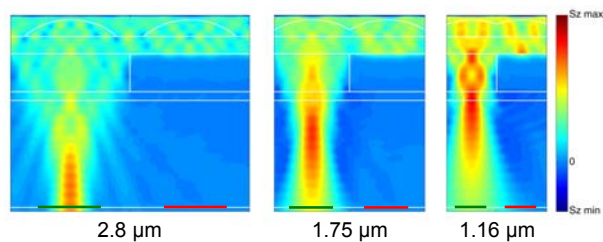
T. H. Hsu *et al.*, IEEE EDL 25, 6 (2004)

## Pixel scaling = Optical problems

- Motivations for reducing pixel size
  - Decrease cost per sensor
  - Increase number of pixels per area (spatial resolution)
- Reducing pixel size leads to:
  - Less light incident on each pixel
    - Noise visible when less than 1000 photons
  - More diffraction, spreading out light
    - Decrease in Optical Efficiency (OE)
    - Increase in Optical Crosstalk (OX)

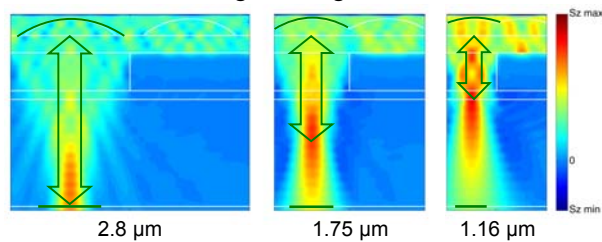
Catrysse, P.B. and Wandell, B.A., Proc. SPIE 5678, 1-13 (2005)  
Xiao, F. *et al.*, Proc. SPIE 5678, 75-84 (2005)

### Pixel shrinks, light escapes



- Electromagnetic field simulation for three-metal layer pixel of varying pixel size

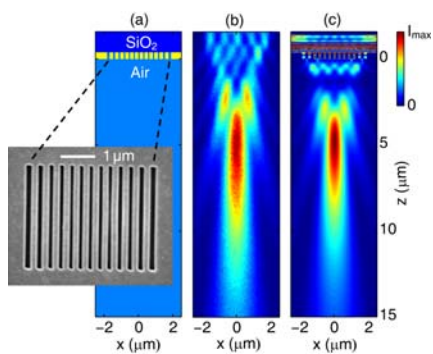
### Collecting: Failing microlens



- Optimized microlens no longer focuses onto photodetector

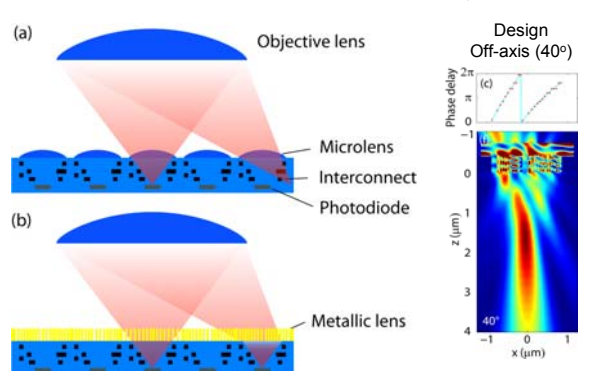
Fesenmaier, Huo, and Catrysse, Proc. SPIE 7250, 72500G (2009)

### Planar plasmonic lens

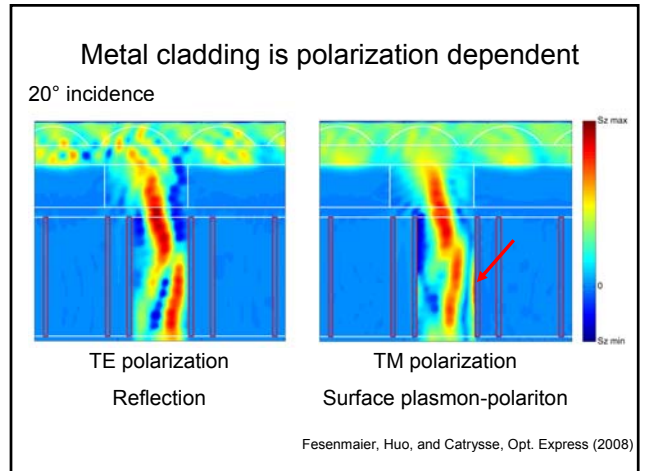
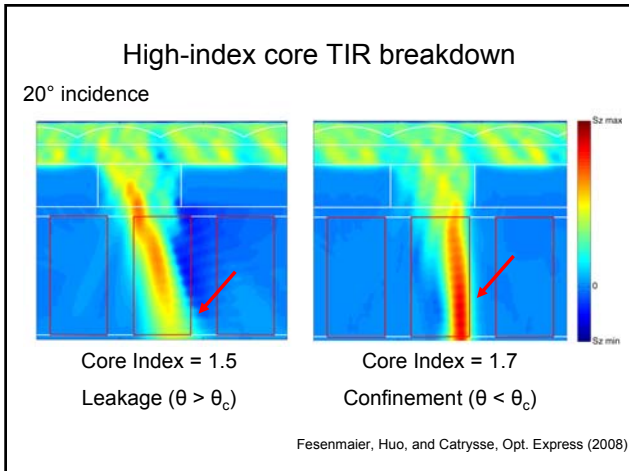
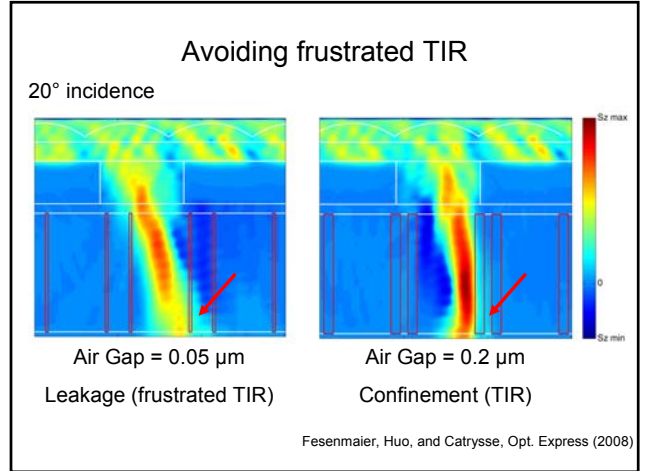
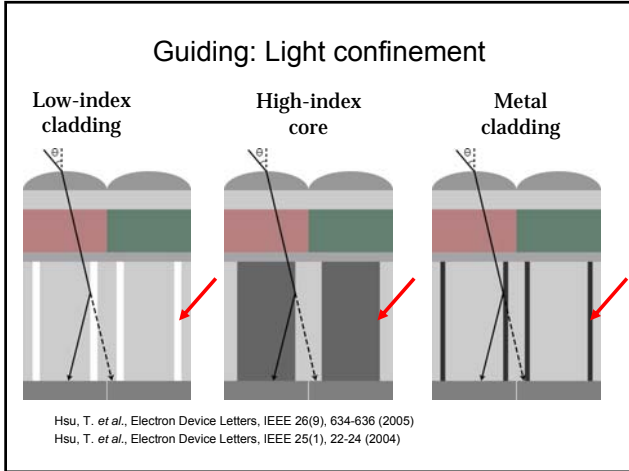


Verslegers, Catrysse et al., Nano Lett. 9, 238 (2009)

### Planar lenses offer flexibility

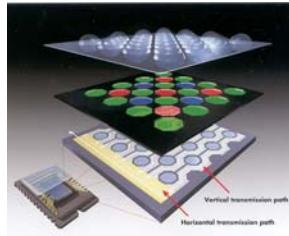


Verslegers, Catrysse et al., APL (2009)



## Sensor Properties: Overview

- Pixel Properties
  - Pixel height, width
  - Fill factor
  - Dark current
  - Read noise
  - Conversion gain
  - Voltage swing
- Array Properties
  - DSNU and PRNU
  - Pixel spacing
  - Number of pixels
- Color Properties
  - Color filter, transmittance
  - Photodetector QE
  - Infrared filter transmittance
  - Color filter array pattern
- Circuit Properties
  - Analog-to-digital conversion
  - Correlated Double Sampling
  - Column Amplifiers (Column FPN)
- Pixel Optics
  - Microlens array

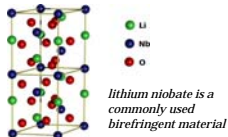


## Imager components

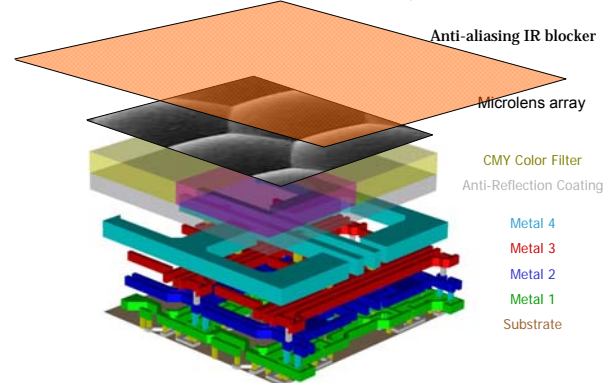
- Anti-aliasing and infrared filter
- Color filter array types

## Anti-aliasing and infrared blocking filter

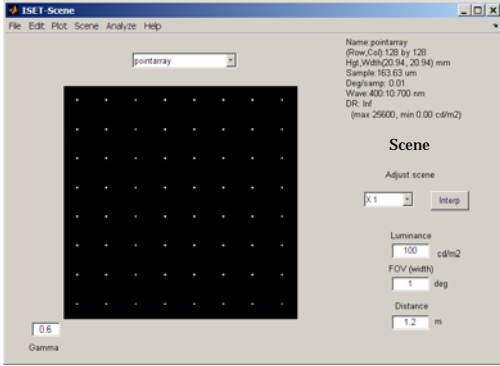
- Placed on the sensor surface
- Blocks IR – to discuss
- Blurs to prevent sampling artifacts –to discuss
- Two types – birefringent and diffusing



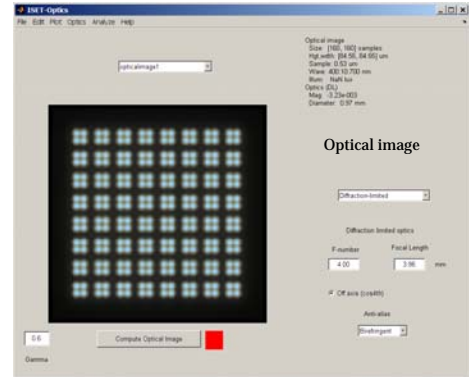
## The microlens array



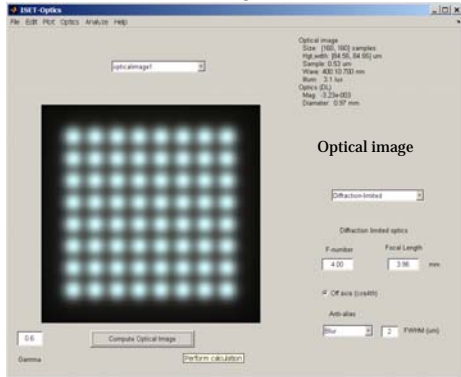
### Modeling the anti-aliasing filter: Scene



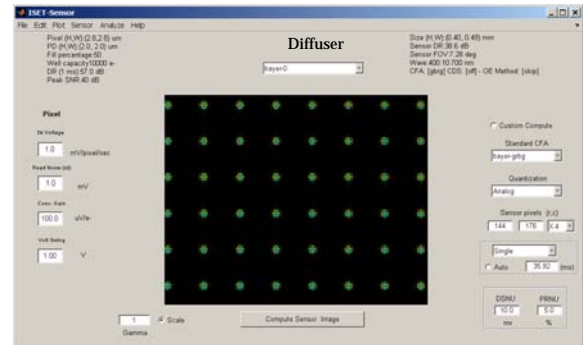
### Modeling the anti-aliasing filter: Birefringent filter



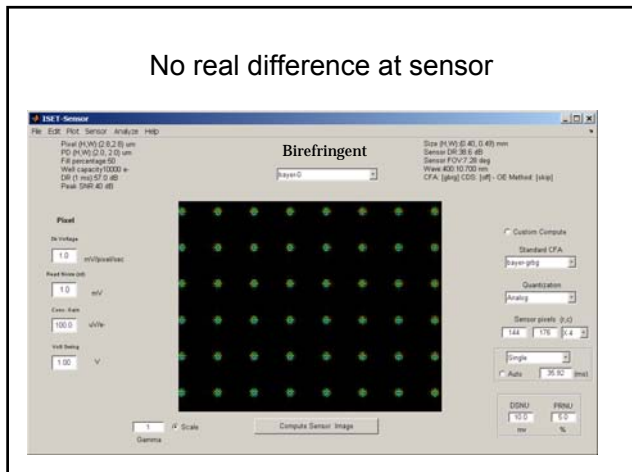
### Modeling the anti-aliasing filter: Diffusing filter



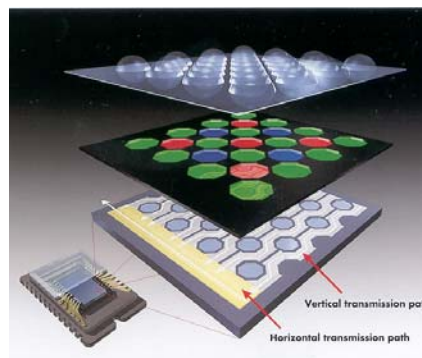
### No real difference at sensor



## No real difference at sensor



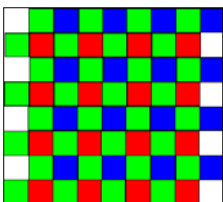
## Color architectures



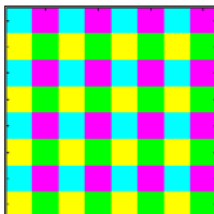
Fuji Corporation

## Camera CFA architectures

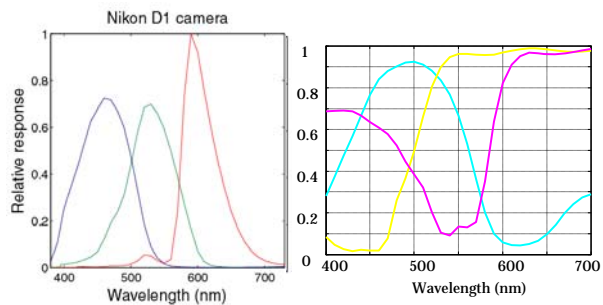
- Bayer array RGB
- Most common
- Green double the sampling rate
- Requires filling in missing pixel values – (demosaicking)



- CMYG – video cameras
- Fast demosaicking
- Broad filters, more photons
- Poor noise properties



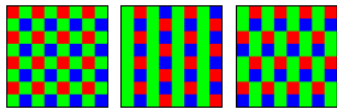
## Example sensor responsivity



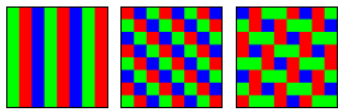
(Xiao and Wandell, personal comm.)

### Alternatives to Bayer RGB array

Most widely used  
⇒

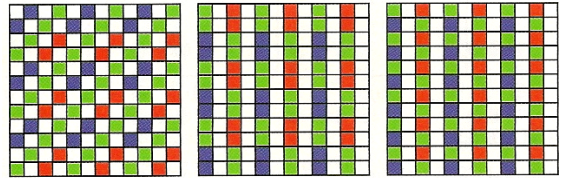


Bayer Yamanaka Lukac



Striped Diagonal striped Holladay

### Arrays with white pixels



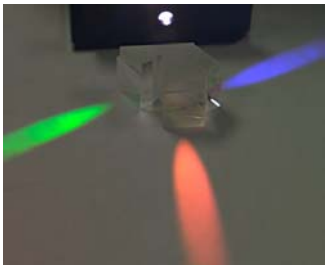
To discuss

### Camera color architectures: 3-Sensor

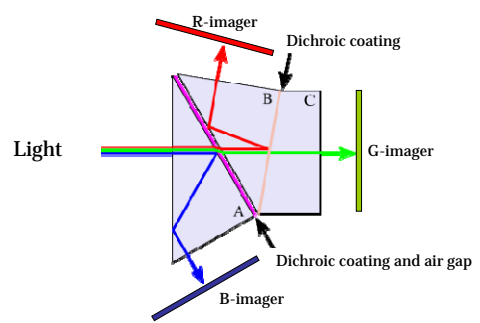
Prismatic optics

No demosaicking

Three (smaller) sensors and optical alignment



### Philips total internal reflection dichroic prism



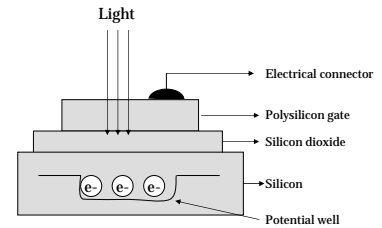
## Other imager architectures

- CCD – the classic
- Triple-well imagers (Foveon)
- Digital pixel sensors (Pixim)

## A few moments for the venerable CCD

Invented in the 1970s as memory devices; light sensitive properties were exploited for imaging applications; a major revolution in Astronomy.

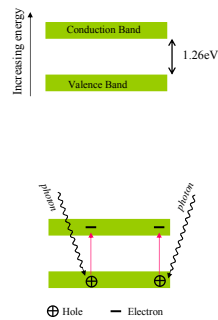
They improved the light gathering power of telescopes by almost two orders of magnitude. In 2001 an amateur astronomer with a CCD camera and a 15 cm telescope collects as much light as an astronomer of the 1960s equipped with a photographic plate and a 1m telescope.



## The photoelectric effect

Atoms in a silicon crystal have electrons arranged in discrete bands called the **Valence Band** (lower energy) and the **Conduction Band** (higher energy).

Most electrons occupy the **Valence band**. The absorption of a photon can excite electrons into the **conduction band**. The excitation requires the photon to provide at least 1.26 electron volts.

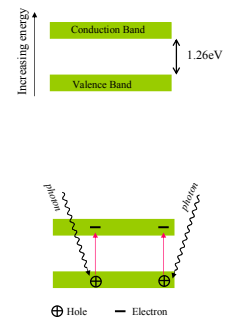


## The photoelectric effect

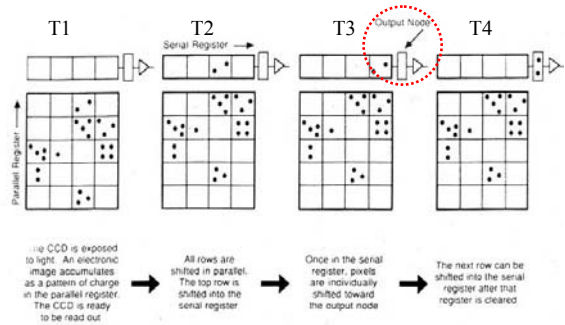
A 1100nm photon of light has 1.26eV. Beyond this wavelength silicon is transparent; hence silicon imagers are insensitive beyond this infra-red band.

In the conduction band electrons move freely about in the lattice of the silicon crystal; this leaves a 'hole' in the valence band which acts like a positively charged carrier. In the absence of an external electric field the hole and electron will recombine and be lost.

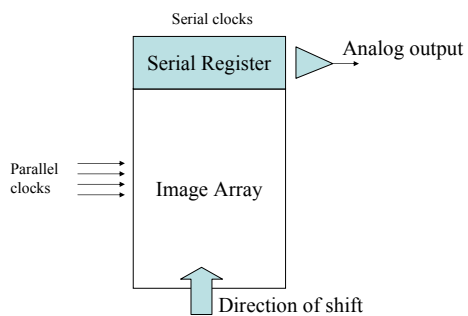
In a CCD an electric field is introduced to prevent recombination.



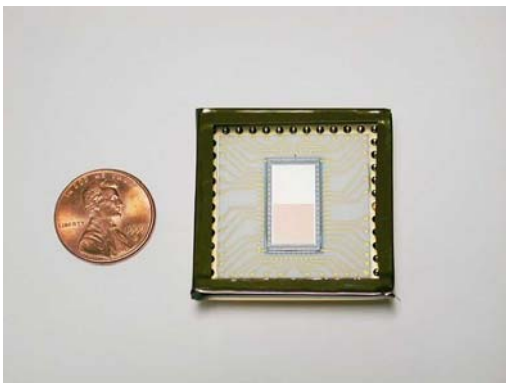
### CCD Charge transfer



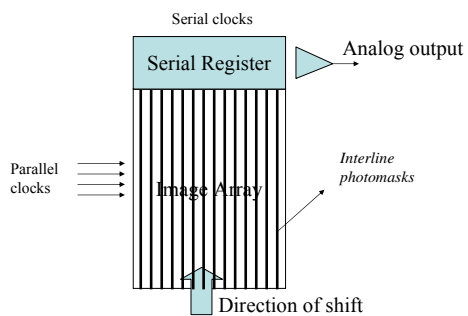
### Full Frame CCD



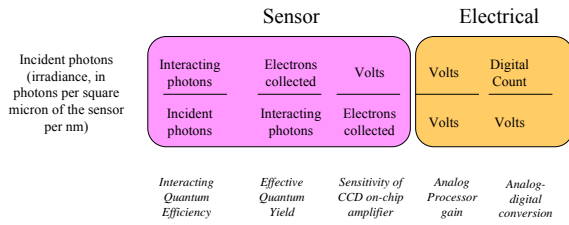
### Frame Transfer CCD



### Interline Transfer CCD



## Electrical Efficiency - Stages ( CCD, see Janesick's books, tutorials)



(After J. Janesick, Pixel Vision)

## Foveon and Sigma

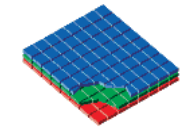
<http://www.foveon.com/article.php?a=67>



Figure 2. The Sigma SD9 is the first digital camera to use a full-color multi-layer sensor technology: The Foveon X3 sensor.

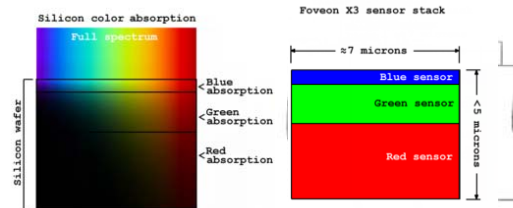
## Wavelengths penetrate to different depths

Foveon X3® Capture



Long-wavelength photons  
penetrate deeper than short.  
The spectral response of  
electrons at the surface differs  
from electrons deeper in the  
material

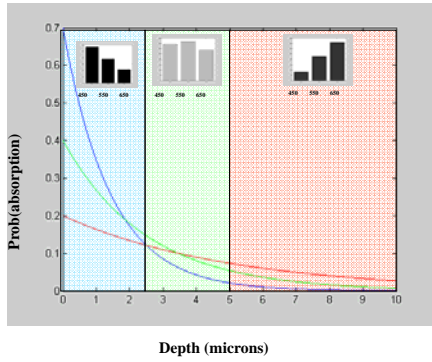
## Photon penetration into silicon



$$p(\lambda, d) = a(\lambda) e^{-a(\lambda)d}$$

Each photon goes somewhere

### Foveon X3 Pixel (Triple-Well)



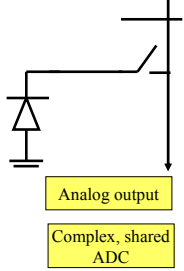
### High dynamic range image capture

- There are software and multi-capture approaches
- We attend to hardware ideas
- Limitation is pixel well-capacity

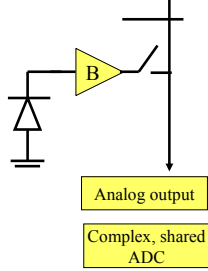


### Hardware Development: CMOS-PPS,APS,DPS

Passive pixel sensor (PPS)



Active pixel sensor (APS)



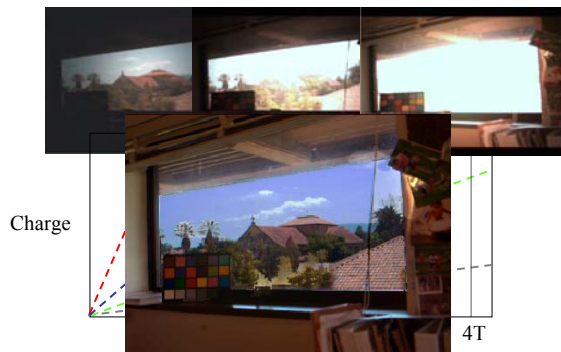
### 1<sup>st</sup> Generation Implementation

(D. Yang et al., IEEE JSSC, 1999)



- *Imager: 640 x 512*
- *Pixel: 10.5 micron*
- *Technology: 0.35 micron*
- *ADC shared x 4*
- *Control signals FPGA*
- *Fill factor ~ 29%*
- *QE ~10%*
- *Really cool, though*

### Multicapture HDR Illustration



### Pixim



## The Human Observer



## Eye and Brain



Optics blur the image that falls on the retina

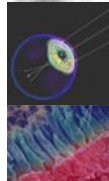
Chromatic aberration and point spread functions

Retinal image sampled by cone photoreceptors

Cone response within a linear transformation of XYZ functions

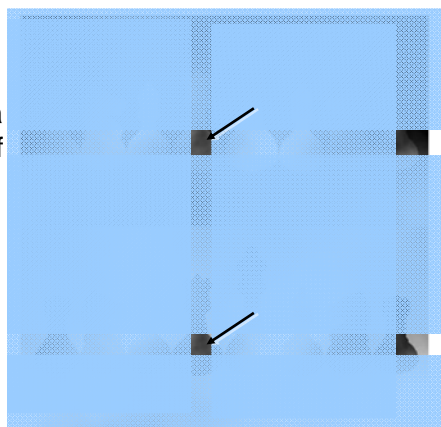
Human visual system separates luminance and chrominance

Neurons blur chrominance more than luminance



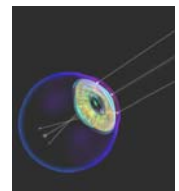
The visual brain does a great deal of processing:  
**Lightness Perception**  
(Anderson and Winawer)

*The chess pieces are the same shade of gray; really.*

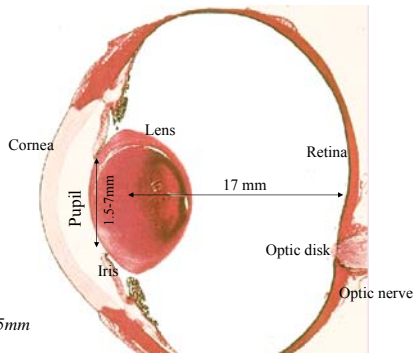


## Human image formation

- What is the quality of the optics of the human eye?
- How can we measure it and describe it empirically?

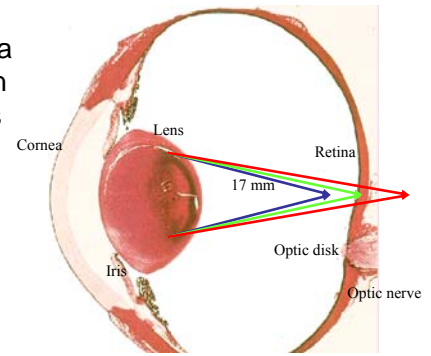


Human eye  
in cross-  
section

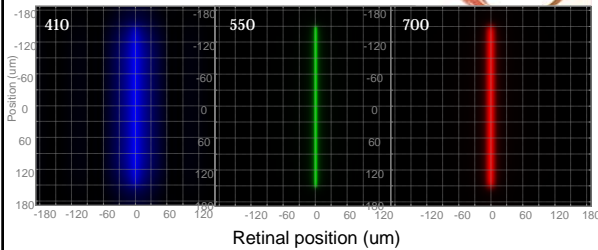
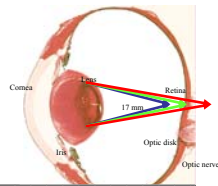


*F-number ~ 2.4-11*  
*Retinal thickness ~ 0.5mm*

Chromatic  
aberration is a  
differences in  
optical focus  
across  
wavelength

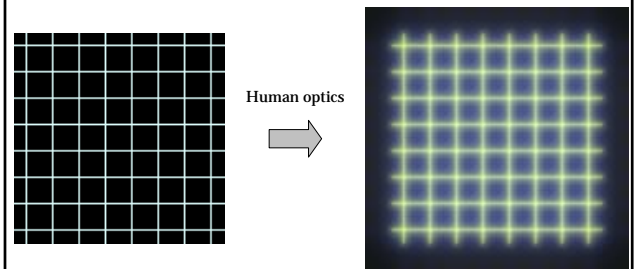


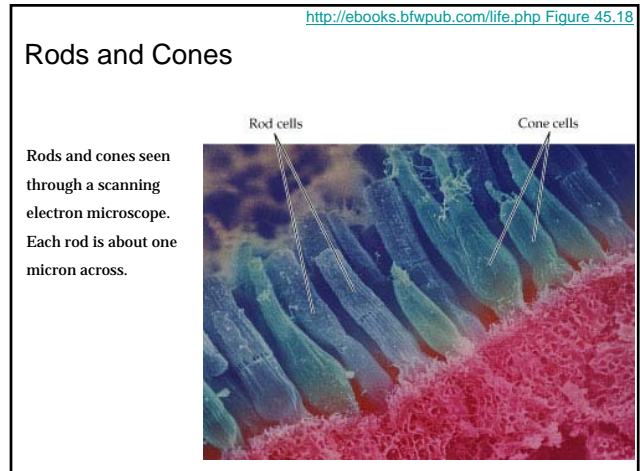
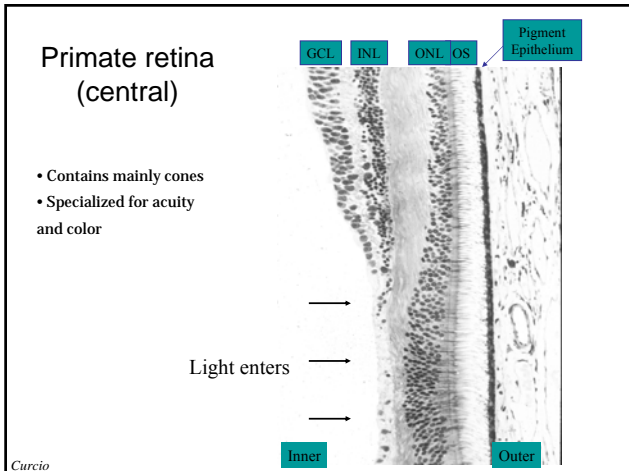
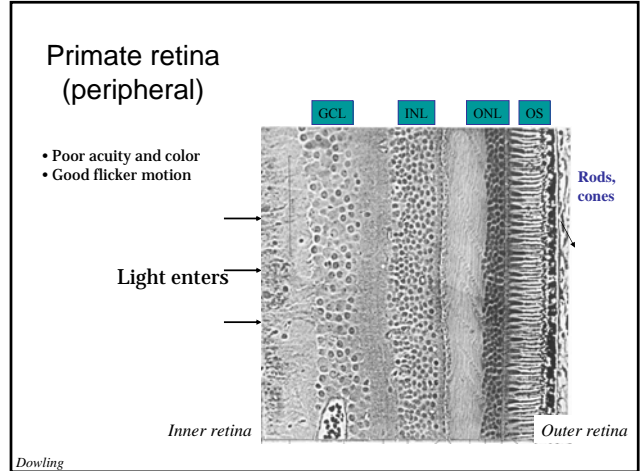
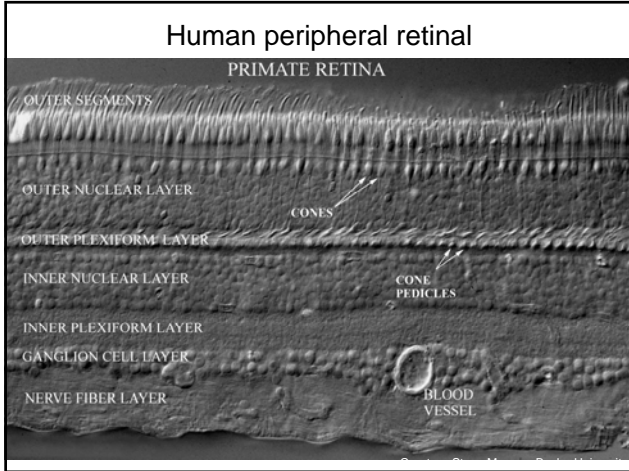
Retinal blur  
differs by  
wavelength



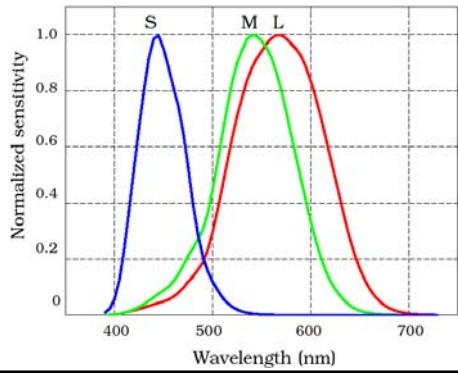
Short wavelength light spreads a lot

Broadband radiance produces chromatic irradiance



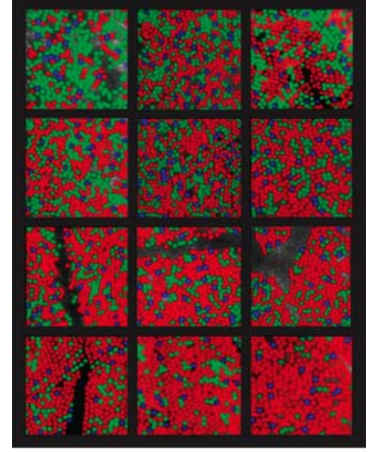


Three cone types: wavelength absorptions



Ratio of L/M cones differs greatly between normals

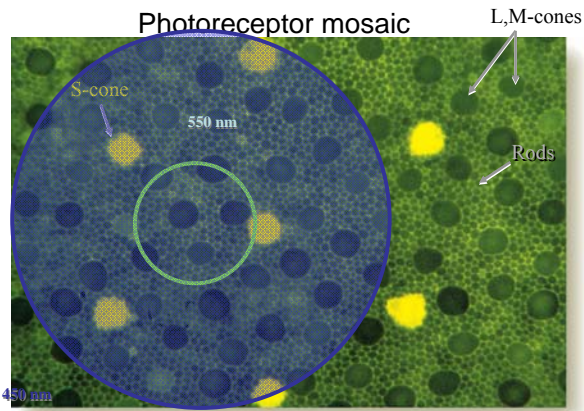
Hofer, H. et al. J. Neurosci. 2005;25:9669-9679



© 2007 Thomson Higher Education

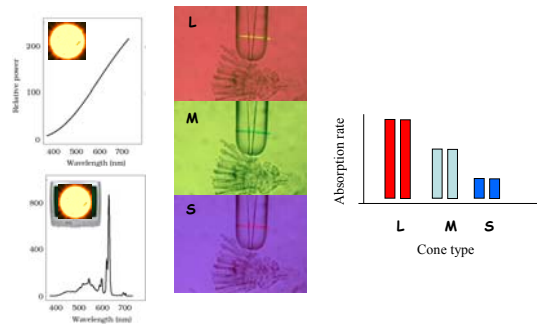
Fig. 7-CO, p. 140

Photoreceptor mosaic



© 2007 Thomson Higher Education

Stimuli causing equal cone absorptions match perceptually

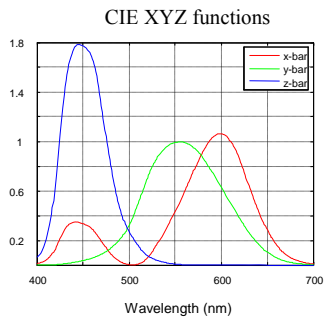


### The CIE-XYZ standard predicts color matches

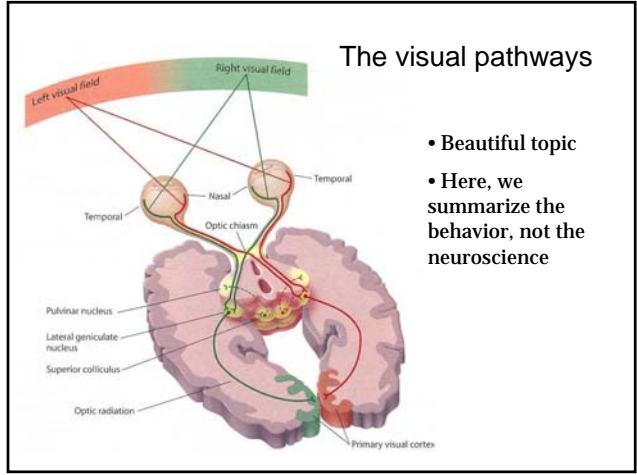
$$X = \int_{\lambda} \bar{x}(\lambda)t(\lambda)d\lambda$$

$$Y = \int_{\lambda} \bar{y}(\lambda)t(\lambda)d\lambda$$

$$Z = \int_{\lambda} \bar{z}(\lambda)t(\lambda)d\lambda$$



### The visual pathways



- Beautiful topic
- Here, we summarize the behavior, not the neuroscience

### Human color metric space: CIELAB ( $\Delta E$ )

- Industry has much experience with product color (paints, textiles)
  - Manufacturing consistency over time
  - Tolerance specifications for contracts
  - Early experience for uniform patches, not images
- Digital imaging required additional work to expand the metric

### Human color metric space: Why not use distance in XYZ (or cones) for metrics?

- Distance depends on coordinate frame
- Experiments show that distance in any frame that is linear with absorptions is wrong
- **Solution:** CIE defines a nonlinear transform on XYZ and distance metric (dE-2000)

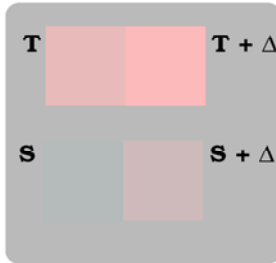
$\|x\| \neq \|Lx\|$

Discriminability  $x + \Delta$  from  $x$  depends on  $x$ ; distance does not

### Human color metric space:

Why not use distance in XYZ (or cones) for metrics?

- Experiments show that distance in any frame that is linear with absorptions is wrong



- This argument is invariant of linear transformation

### Human color metric space:

Why not use distance in XYZ (or cones) for metrics?

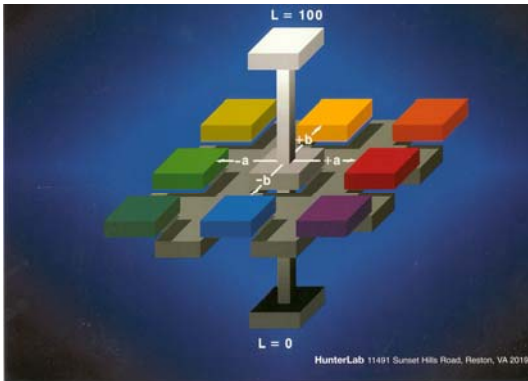
- **Solution:** CIE defines a nonlinear transform on XYZ and distance metric (dE-2000)

$$L^* = \begin{cases} 116 \left( \frac{Y}{Y_w} \right)^{1/3} - 16, & \text{if } \frac{Y}{Y_w} > 0.00856 \\ 903.3 \left( \frac{Y}{Y_w} \right) & \text{otherwise} \end{cases}$$

$$a^* = 500 \left\{ \left( \frac{X}{X_w} \right)^{1/3} - \left( \frac{Y}{Y_w} \right)^{1/3} \right\}$$

$$b^* = 200 \left\{ \left( \frac{Y}{Y_w} \right)^{1/3} - \left( \frac{Z}{Z_w} \right)^{1/3} \right\}$$

### CIELAB Color Solid



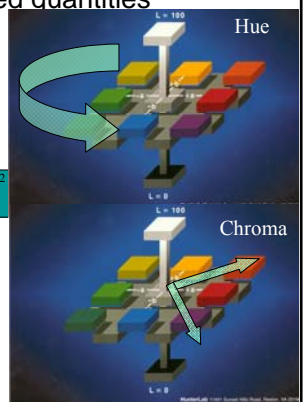
### CIELAB derived quantities

$$h_{ab} = \arctan(b^* / a^*)$$

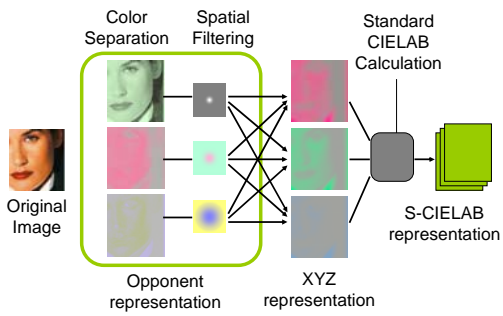
$$C_{ab}^* = (a^{*2} + b^{*2})^{1/2}$$

$$\Delta E_{ab} = \left( (\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2 \right)^{1/2}$$

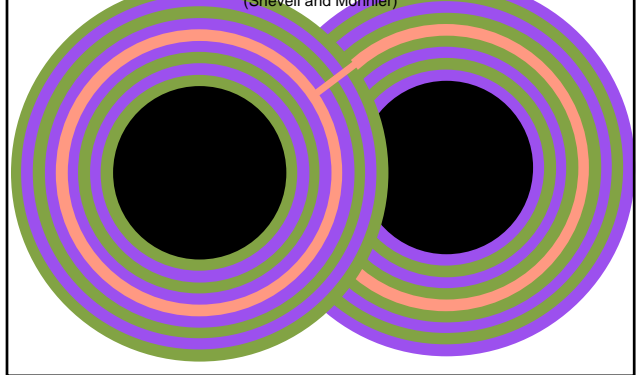
- Distance measure between two points
- Corresponds to perceptual similarity



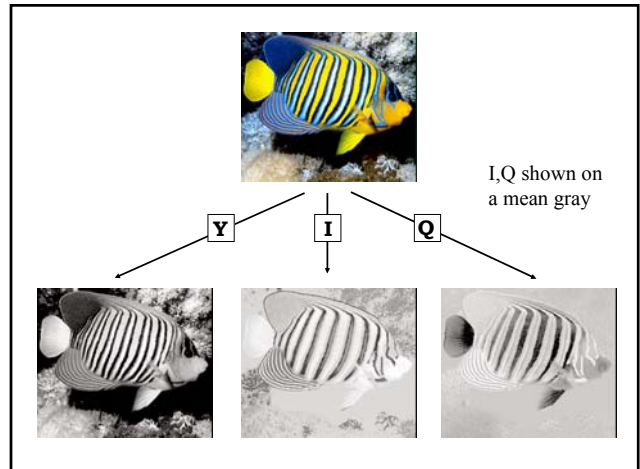
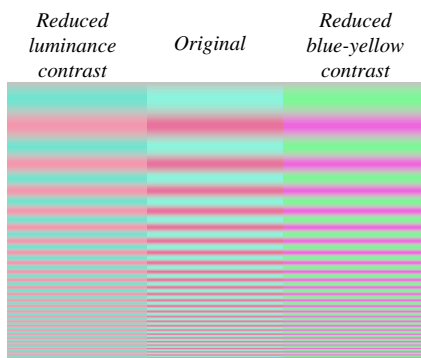
### Spatial CIELAB (S-CIELAB) Human Spatial Chromatic Sensitivity

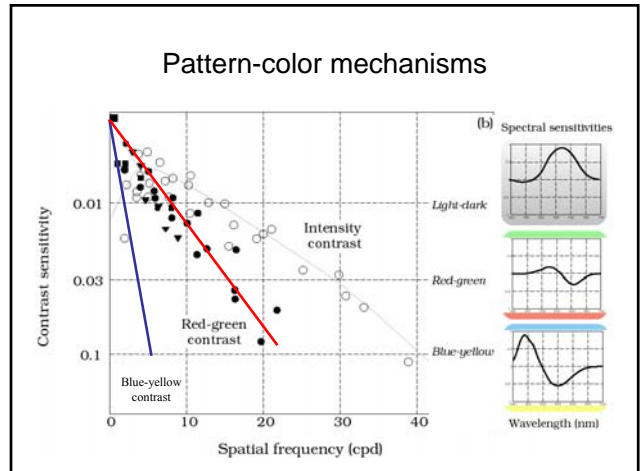
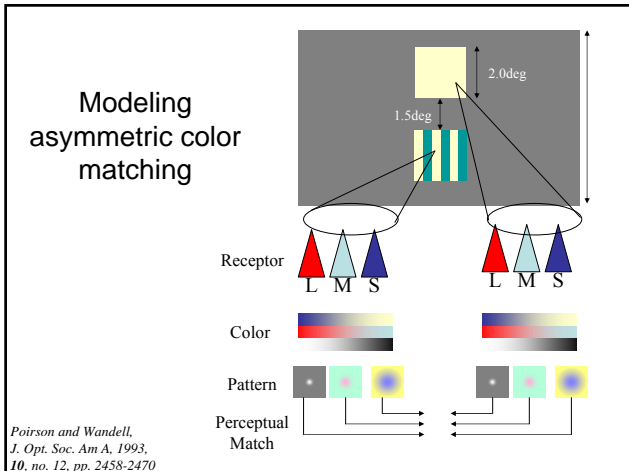
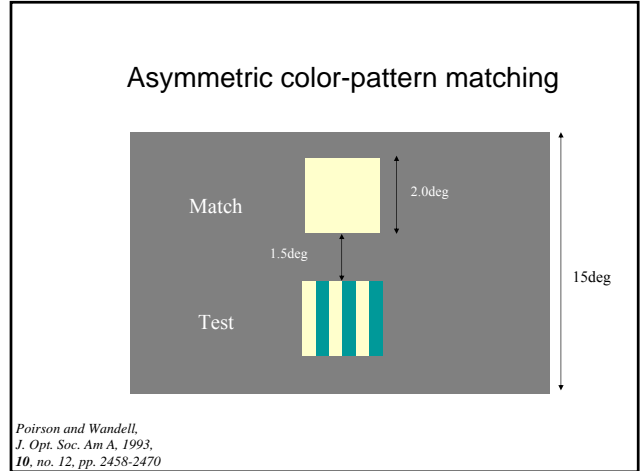
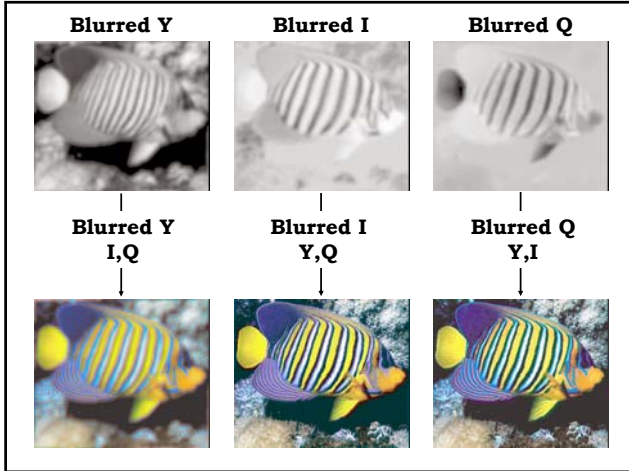


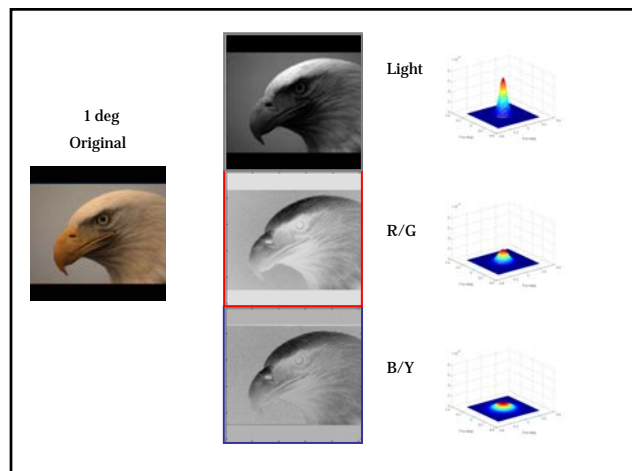
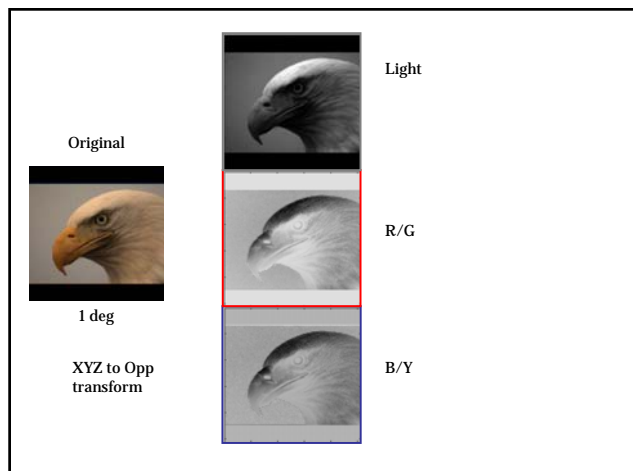
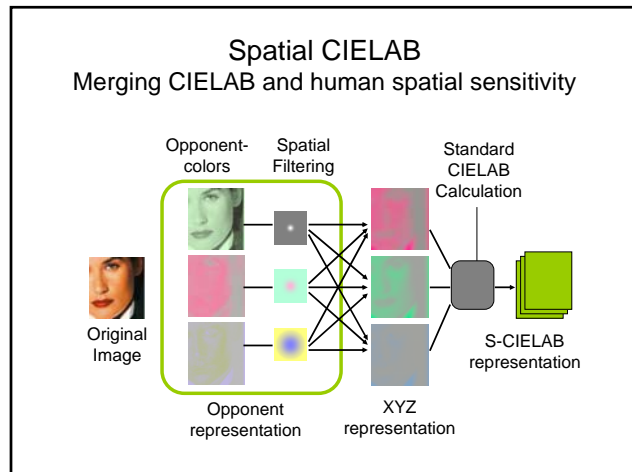
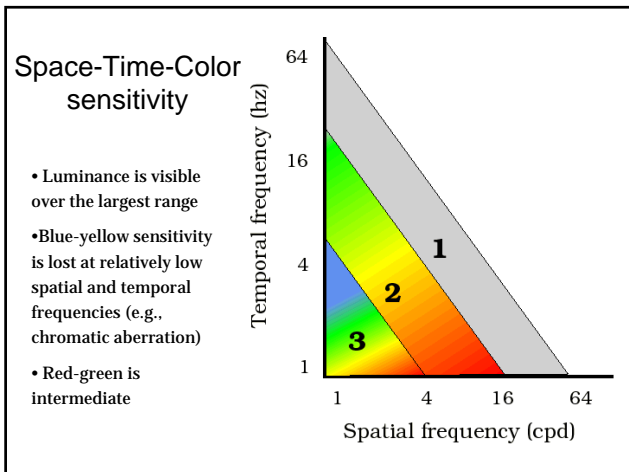
### Color Appearance Depends On The Spatial Pattern Across The Cone Mosaic (Shevell and Monnier)

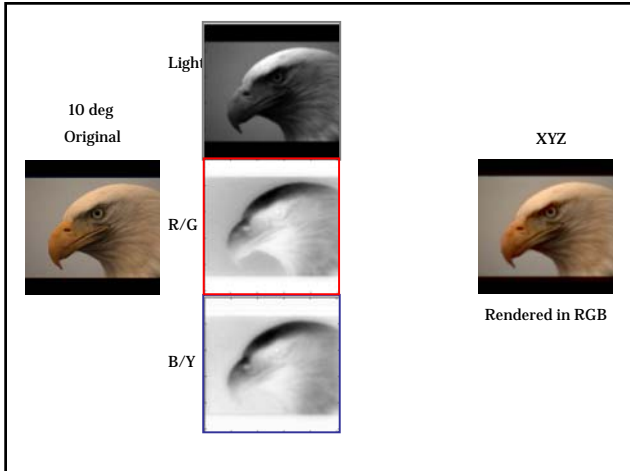
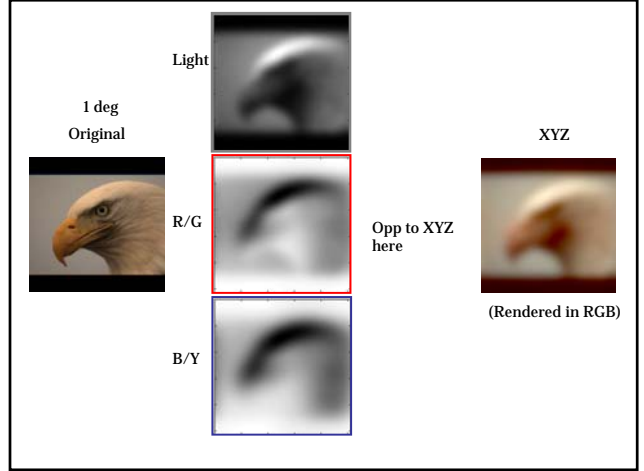
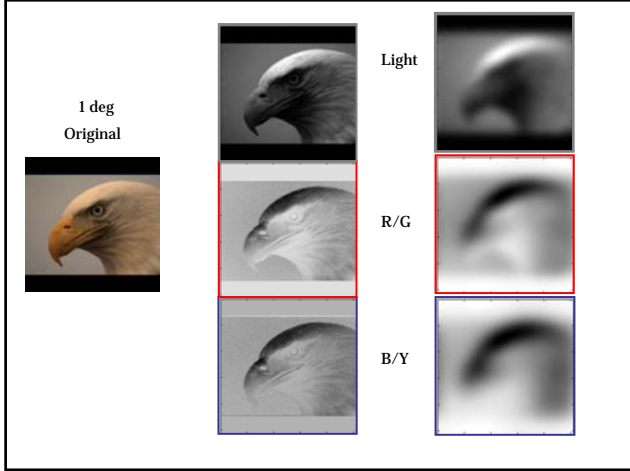


### Spatial and temporal variations in color sensitivity are significant

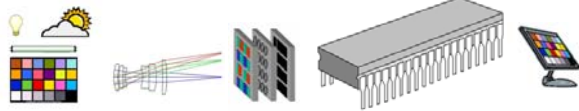








## Image processing pipeline



## Pipeline goals: A variety

- Reproduce an image whose appearance matches the original up to display intensity and gamut limitations, or
- Make a nice picture, guided by the original
- Preserve certain image metrics (edges, geometry)

## Pipeline goals

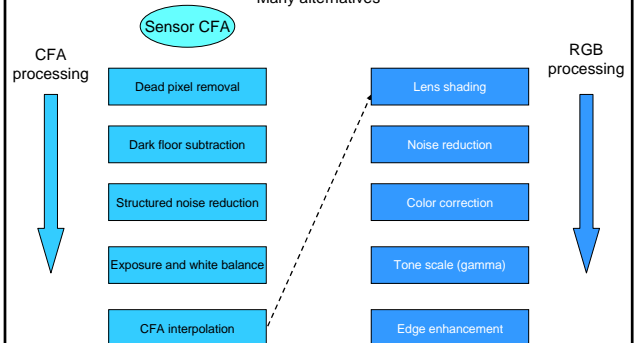
The processing pipeline contains a mixture of methods that

- Correct hardware imperfections
- Facilitate inter-device communication
- Address appearance (aesthetics) issues

## Reference pipeline

(after Adams and Hamilton, in Lukas, Fig. 3.1)

Many alternatives



## Reference pipeline II

### Design Considerations of Color Image Processing Pipeline for Digital Cameras

Wen-Chung Kao, Member, IEEE, Sheng-Hong Wang, Lien-Yang Chen, and Sheng-Yuan Lin

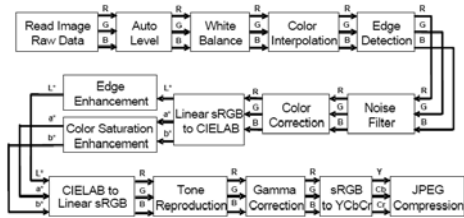


Fig. 1. The proposed image processing pipeline.

## Acquisition issues

- Exposure
- Demosaic
- Color management

## Exposure algorithms

Exposure value (EV) definition:

$$EV = \log_2\left(\frac{F^2}{T}\right) = 2\log_2(F) - \log(T)$$

$F$  : f-number

$T$  : exposure time

- EV accounts both for
  - F-number (focal length divided by aperture)
  - Integration time (T).
- EV becomes smaller as
  - Exposure duration increases
  - Aperture increases

## Exposure algorithms

Set the f-number (F) and exposure time (T) so that the image pixels accumulate charge that reaches, but does not exceed, the well-capacity of the pixels (sensor range).

The algorithm is based on a definition of Exposure Value (EV) and estimates of the current image brightness (B).

Definition:

$$EV = \log_2\left(\frac{F^2}{T}\right) = 2\log_2(F) - \log(T)$$

$F$  : f-number

$T$  : exposure time

EV accounts both for the f-number (focal length divided by aperture) and the integration time (T). The exposure value becomes smaller as the exposure duration increases, and it becomes larger as the f-number grows (aperture increases).

## Exposure algorithms: Principle

- For a particular image brightness,  $B_{opt}$  establish an optimal exposure value,  $EV_{opt}$
- Acquire an image with a pre-determined exposure value ( $EV_{pre}$ )
- Calculate the image brightness,  $B_{pre}$
- Adjust the  $EV_{pre}$  according to the formula

$$EV_{opt} = EV_{pre} + \log 2(B_{pre}) - \log 2(B_{opt})$$

Exposure algorithms differ in how they determine  $B$

## Exposure algorithms: Brightness

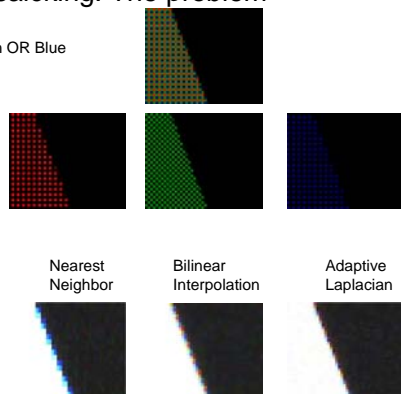
- *Mean*:  $B_{pre}$  is the mean image brightness
- *Center-weighted Mean*:  $B_{pre}$  is the weighted image brightness (more weight given to the image center)
- *Spot*:  $B_{pre}$  is the mean of the center (3%) area
- *Median*:  $B_{pre}$  is the median image brightness
- *Green*:  $B_{pre}$  is the mean of the green channel only

## Demosaicking: The problem

Sensor Array has Red, Green OR Blue at every pixel

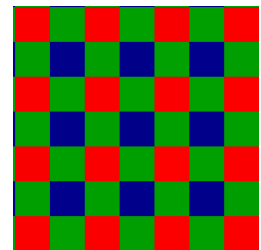
Interpolate the missing Red, Green, and Blue Pixel Values

Create an image that has a Red, Green AND Blue at every pixel

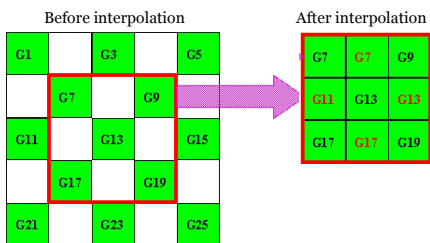


## Demosaicking depends on the mosaic The Bayer mosaic is most common

- G is present at twice the spatial resolution of the R or B
- G is associated with luminance; it is fully interpolated to the size of the sampling mosaic.



### Demosaicking: Nearest Neighbor



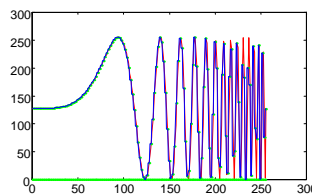
Similar – only more – for R and B channels

### Nearest Neighbor

Chromatic fringing on monochrome test patterns is a good visual test for demosaicking

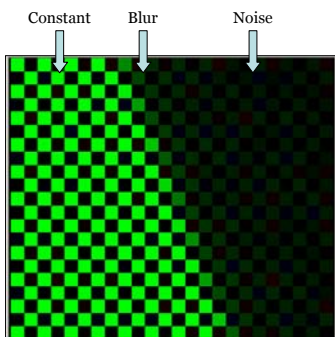


Very significant fringing for nearest neighbor



### Bilinear interpolation (within channel)

We leave the measured values in place  
 Demosaicking fills in the missing G values  
 We use nearby measured G to interpolate the missing G; similarly for R and B

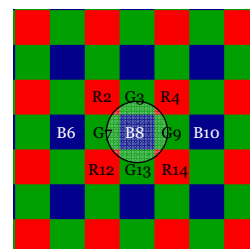


### Bilinear interpolation (within channel)

Existing values are left untouched. The average of adjacent green pixel values is interpolated.

For example:

$$G8 = (G3 + G7 + G9 + G13) / 4$$

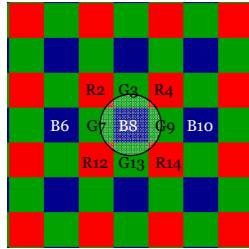


## Bilinear interpolation (across channel)

Across channel uses weighted sum from other colors, in particular those at the same location – these are usually very correlated.

For example:

$$G8 = \alpha B8 + \beta \frac{G3 + G7 + G9 + G13}{4} + (1 - \alpha - \beta) \frac{R2 + R4 + R12 + R14}{4}$$



## Chromatic demosaicing

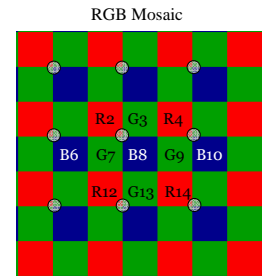
The RGB values are converted to two opponent-colors chromatic values (chrominance plane).

The chrominance plane has 1/4<sup>th</sup> the luminance resolution. This is consistent with common JPEG practice – and for the reasons we discussed about human vision.

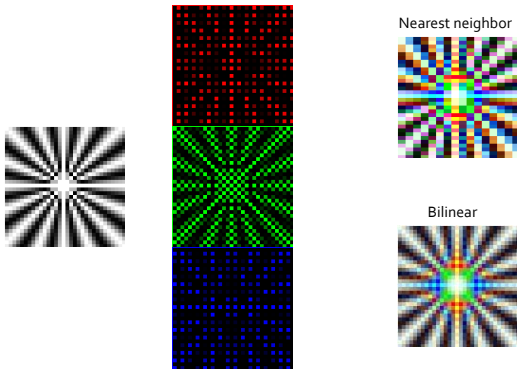
$$C_R = R - G$$

$$C_B = R - B$$

$$YCrCb$$



## Demosaicking comparison Nearest neighbor vs. bilinear



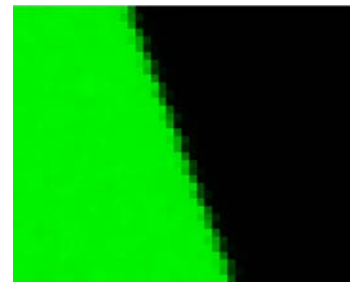
## Analysis: bilinear interpolation

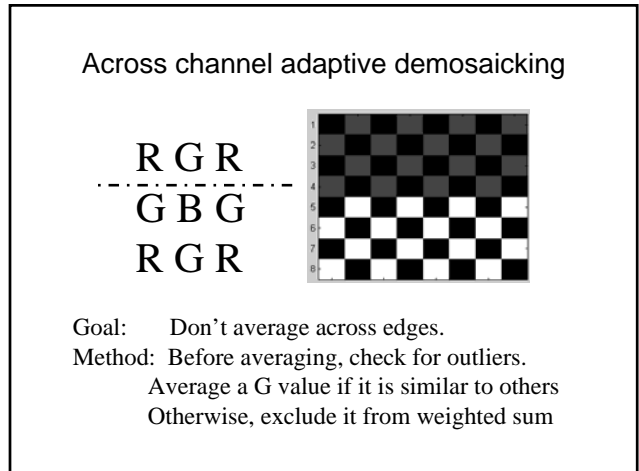
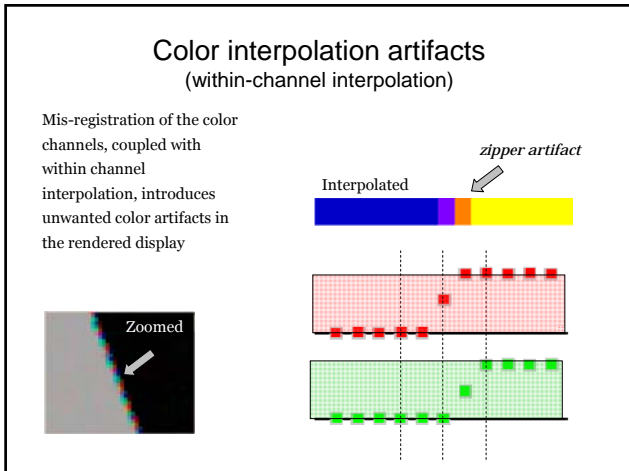
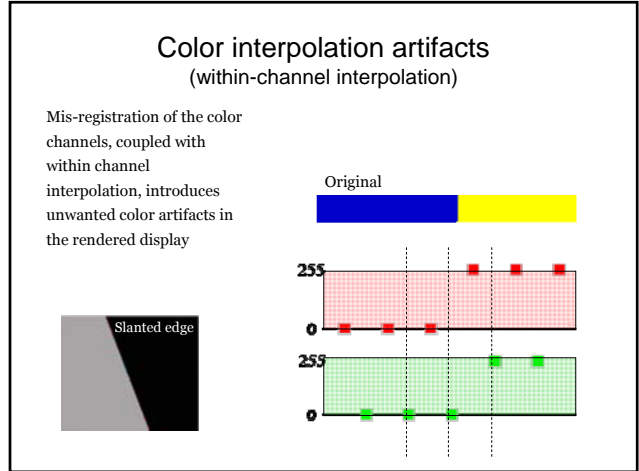
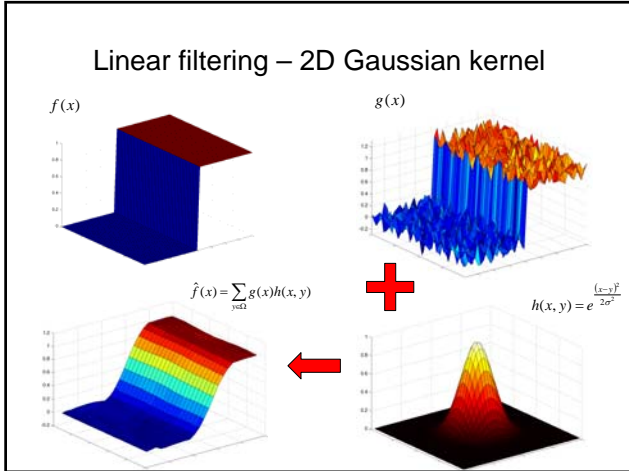
Results are OK, but not good

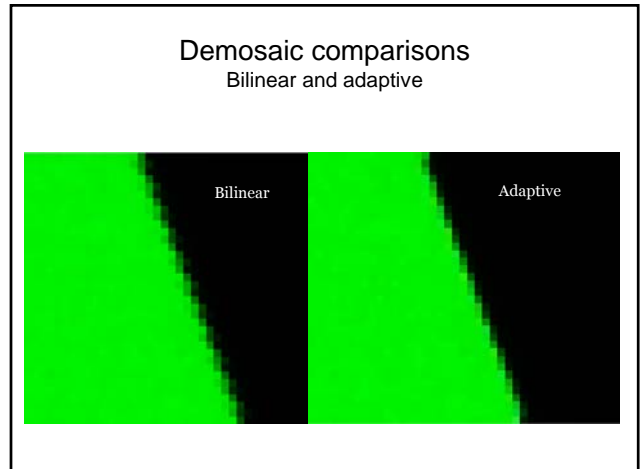
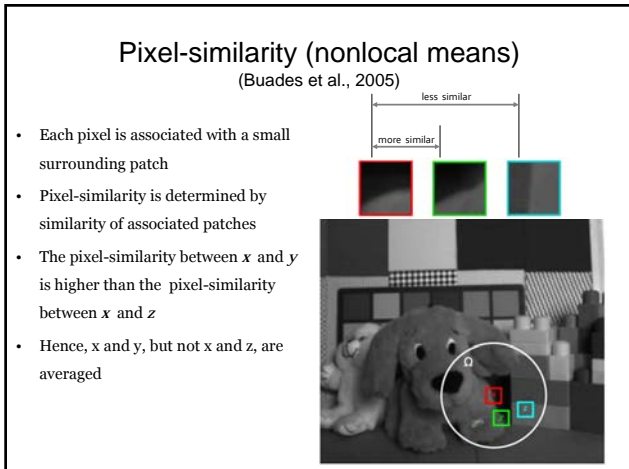
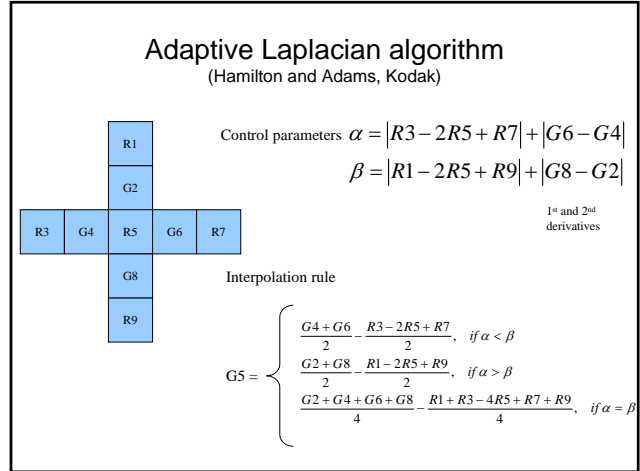
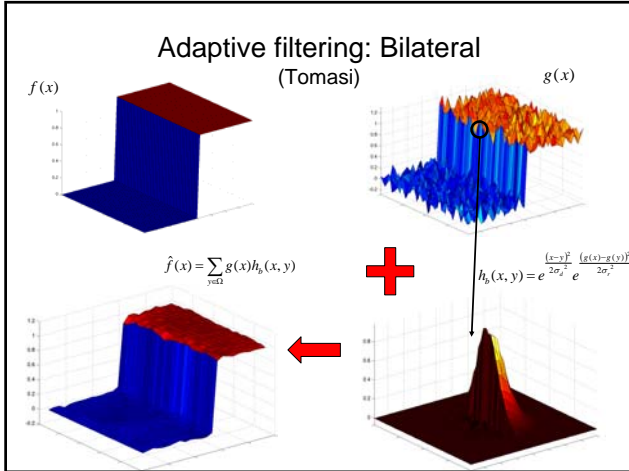
- The best information comes from the cross-color value at the same location
- The averaging blurs edges

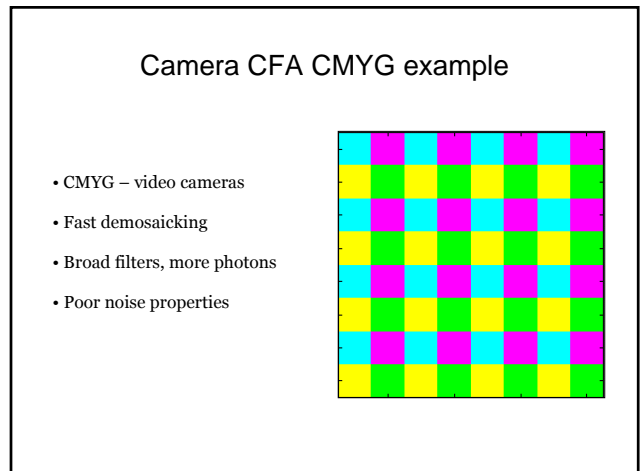
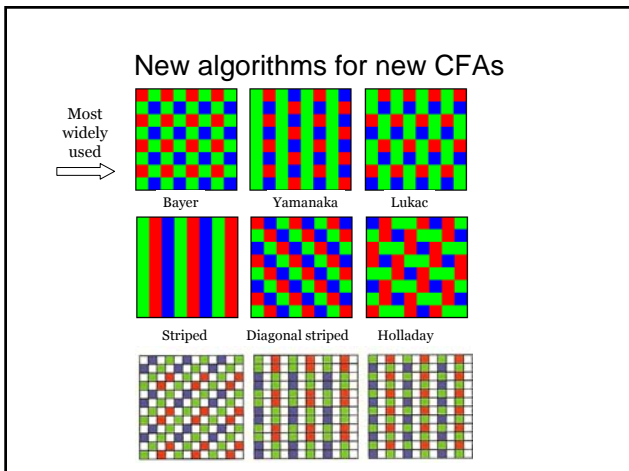
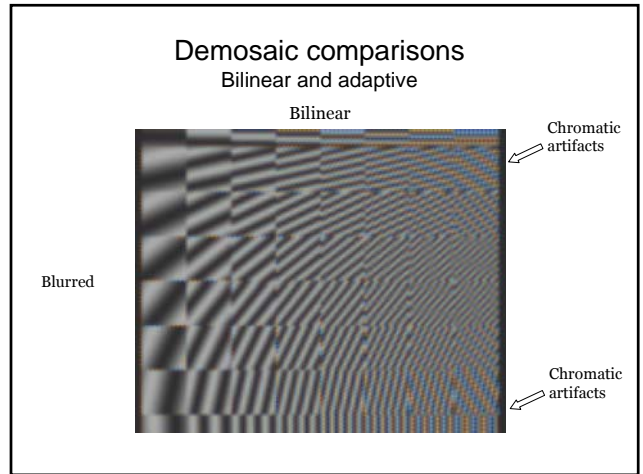
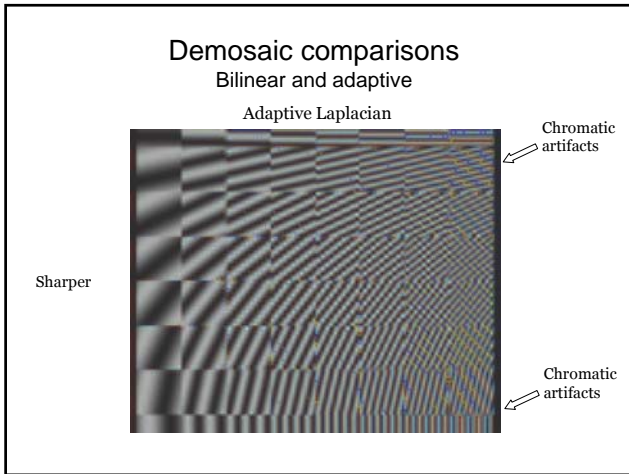
Most modern algorithms

- Combine information from color channels to make a decision
- Aim to preserve edges







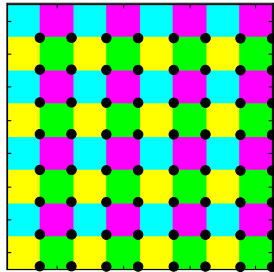


## Nice features of the CMYG mosaic

- The sum of the neighbors at each corner estimates luminance (black dots)

$$C+M+Y+G$$

- High spatial sampling, but spatially correlated



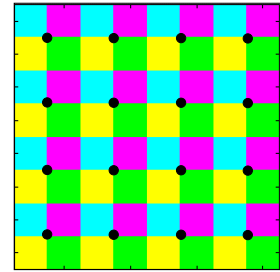
## Chromatic Demosaicing CMYG

- The chrominance plane can be calculated at half the spatial resolution using these formula:

$$C_R = (M + Y) - (G+C)$$

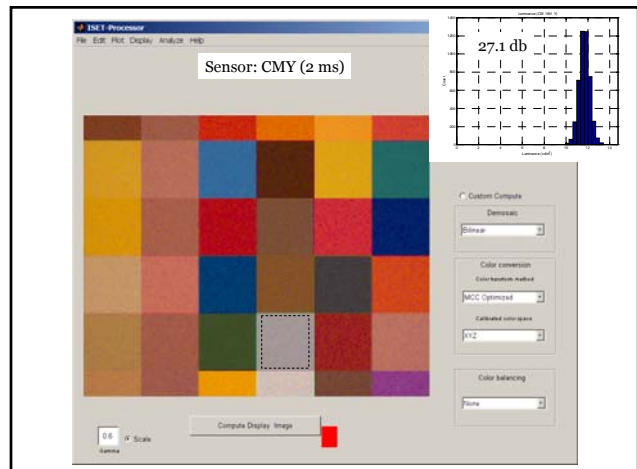
$$C_B = (M + C) - (G+Y)$$

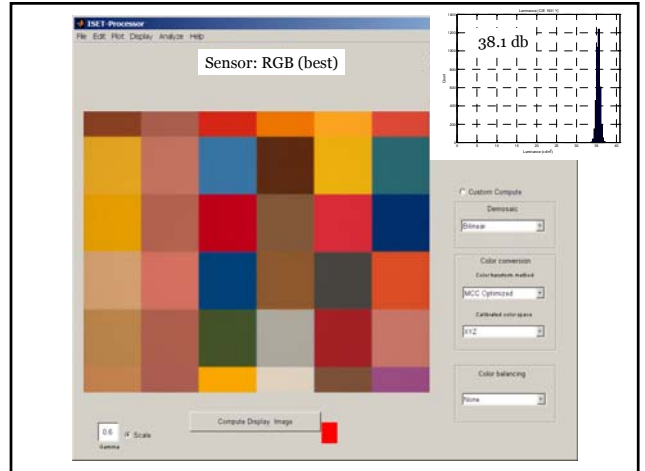
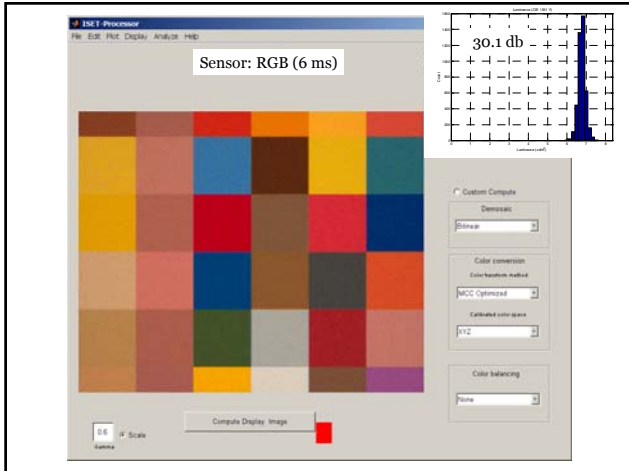
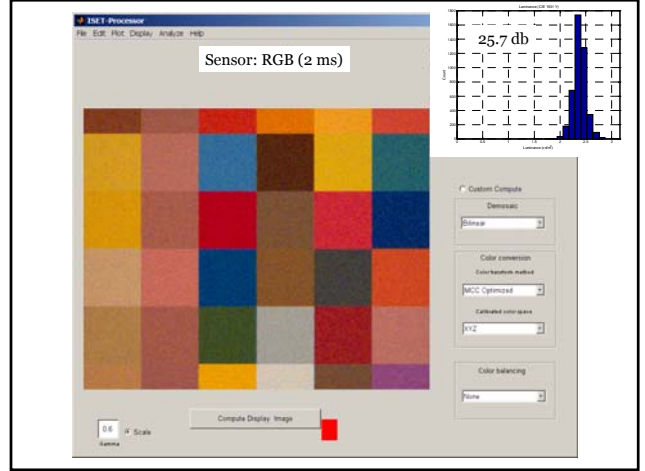
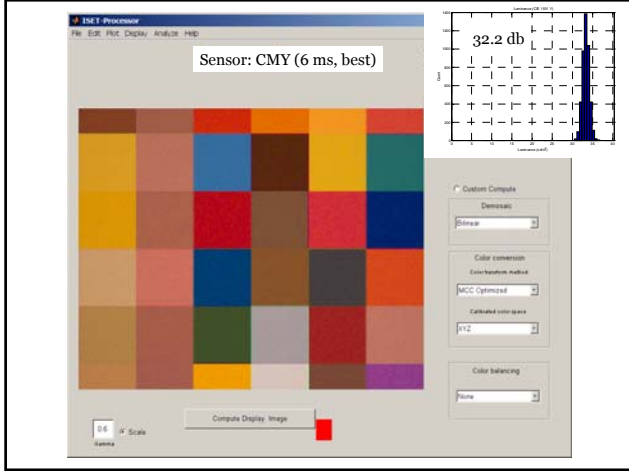
Note:  $R = 1 - \text{Cyan}$   
 $G = 1 - \text{Magenta}$   
 $B = 1 - \text{Yellow}$



## Comparing CMYG and RGB

- CMY captures more photons per unit time (that's good); at matched integration times CMY is about as good or better than RGB
- The CMY to display RGB transform amplifies noise (that's bad), limiting CMY peak performance
- At peak performance RGB is much better, though RGB peak requires a longer integration time





## Inter channel correlations – advanced

- The correlation of the mean and of the local differences has been used to motivate many demosaicing algorithms
  - Gunturk, B.K., Altunbasak, Y., and Mersereau, R.M., *Color plane interpolation using alternating projections*, IEEE Transactions on Image Processing, **11**(9), p.997-1013, 2002.
  - Adams, J.E. and Hamilton, J.F., *Design of practical color filter array interpolation algorithms for digital cameras*, Proceedings of the SPIE, **3028**, p.117-125, 1997.
  - Kimmel, R., *Demosaicing: image reconstruction from color CCD samples*, IEEE
- Practical implementations of demosaicking that exploit high inter-channel correlations rely on the higher rate of sampling in the green CFA channel.
- Most such algorithms derive the best estimate of high frequency image information using the green channel. They then use the inter-channel correlation to improve the interpolation in the red and blue channels

## On-camera denoising

- Only modest processing (e.g., median filter).
- <http://www.stanford.edu/~slansel/DenoiseLab/index.htm>
- Some discussions of denoising the CFA – ask me

## Color management

- Many informal terms describe color processing

- No formal definitions
- Used in different ways
- I won't try to explain them all

- This section describes
  - Why bad things happen to good colors, and
  - Several color processing approaches

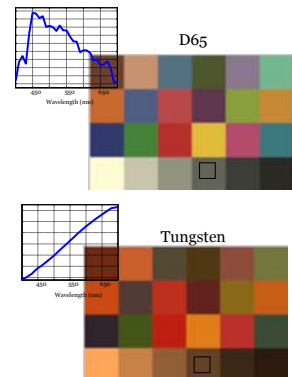


White balance  
 Color conversion  
 Color balance  
 Color correction  
 Color rendering  
 Illuminant transformation  
 Color constancy

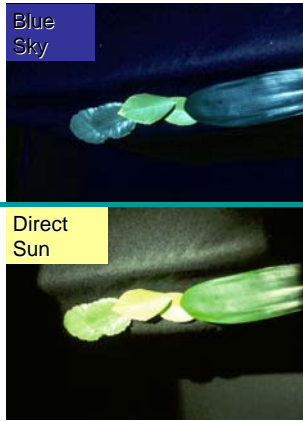
ICC profiles

## The illuminant and color management

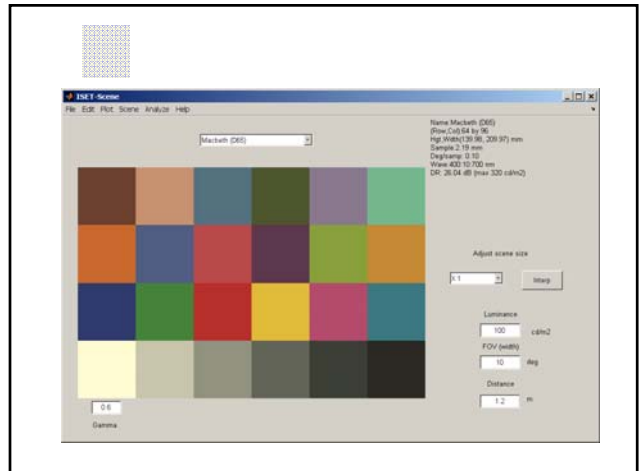
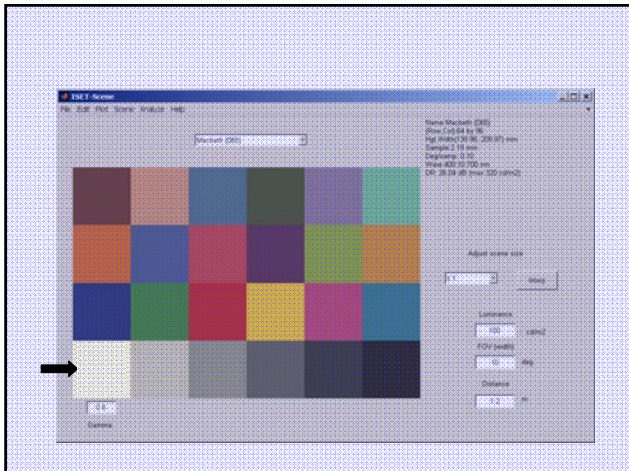
- Illuminant changes cause the scattered light to change
- The visual system adjusts to the illuminant (color constancy)
- If the pipeline doesn't adjust, the rendering has the wrong color balance
- Illuminant correction is a very important aspect of color balancing



Illumination and color appearance are coupled through color constancy



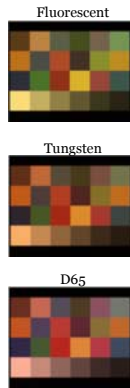
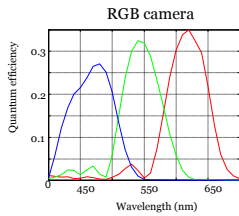
Color constancy: The brain adjusts



## Color management: Problem

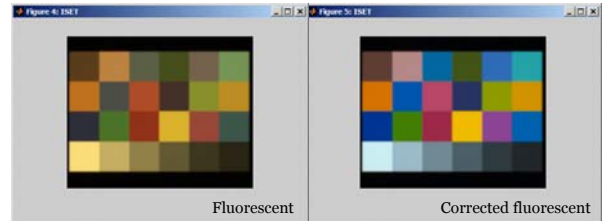
Copying sensor values directly to display graphics card or printer driver is a bad idea

- Devices differ (device-dependent, no calibration)
- Doesn't account for viewer
- Looks different for every light



## Color management: Principles

1. Color management
  - Guides the conversion of sensor data to display signals
  - Accounts for some human visual system properties
2. Color management computations may include
  - Linear transformations (global)
  - Look-up tables (locally linear)
  - Statistical characterizations of the image
  - Knowledge of the output medium



## Color adjustment methods

Device-dependent diagonal

$$\text{Display} \begin{pmatrix} R' \\ G' \\ B' \end{pmatrix} = \begin{pmatrix} s_r & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & s_b \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \text{ Camera}$$

Choose  $s_r$  and  $s_b$ , so that a neutral (gray) surface in the image is rendered as a neutral display output

You must make an educated guess about the camera RGB to a neutral surface. Example ideas:

- The average of the image is neutral
- The brightest elements of the image average to neutral
- Many other ideas

## Color adjustment methods

Device-independent diagonal

$$\text{Display} \begin{pmatrix} R' \\ G' \\ B' \end{pmatrix} = \begin{pmatrix} M & & \\ & \begin{pmatrix} s_r & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & s_b \end{pmatrix} & \\ & & L \end{pmatrix} \text{ Camera} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

- Transform the device data into a calibrated space (cones, XYZ), ( $L$ ).
- Perform the diagonal transformation, again choosing  $s_r$  and  $s_b$ , so that a neutral (gray) surface in the image is rendered as a neutral signal in that space
- Convert to the display representation from the calibrated space to display space ( $M$ )

### Color adjustment methods

Device-independent linear

$$\text{Display} \begin{pmatrix} R'_1 & R'_N \\ G'_1 & \dots & G'_N \\ B'_1 & B'_N \end{pmatrix} = \begin{pmatrix} C_E \end{pmatrix} \begin{pmatrix} R_1 & R_N \\ G_1 & \dots & G_N \\ B_1 & B_N \end{pmatrix} \text{Camera}$$

For each light condition, E, choose a matrix,  $C_E$ , that transforms multiple measurements to desirable display values.

Find these matrices for 30-50 likely lights.

The matrix  $C_E$  may be chosen by MSE or perceptual error minimization.

### Color adjustment methods

Device-independent locally linear

$$\text{Display} \begin{pmatrix} R'_1 & R'_N \\ G'_1 & \dots & G'_N \\ B'_1 & B'_N \end{pmatrix} = \begin{matrix} C_1 \\ C_2 \\ C_3 \\ C_4 \\ \dots \\ C_N \end{matrix} \mathbf{E} \begin{pmatrix} R_1 & R_N \\ G_1 & \dots & G_N \\ B_1 & B_N \end{pmatrix} \text{Camera}$$

Create lookup tables for each light condition, E, that map surfaces into a desired display for that light.

Such tables, which are essentially a collection of local linear transformations; they may be chosen by MSE or perceptual error minimization.

### Color adjustments summary

Estimate illuminant  $\mathbf{E}$

$\Rightarrow$  Apply transformation for illuminant E

Transform sensor data to calibrated color (XYZ)

Convert to display space, adjusting for gamma and gamut

Display

Sensor

L

$\Downarrow$

$C_E$

$\Downarrow$

M

$\Downarrow$

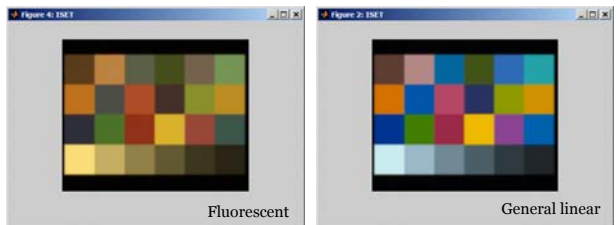
Display

### Device-dependent diagonal

Fluorescent

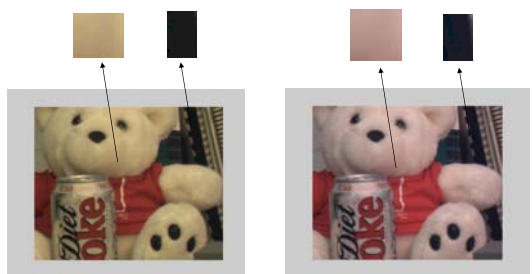
Device-diagonal

### Device-dependent diagonal



### Illuminant estimation

Best Information In Light Regions



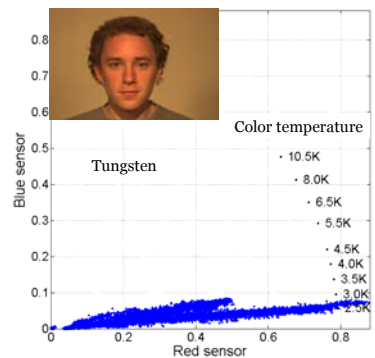
### Example of image statistics used in illuminant estimation

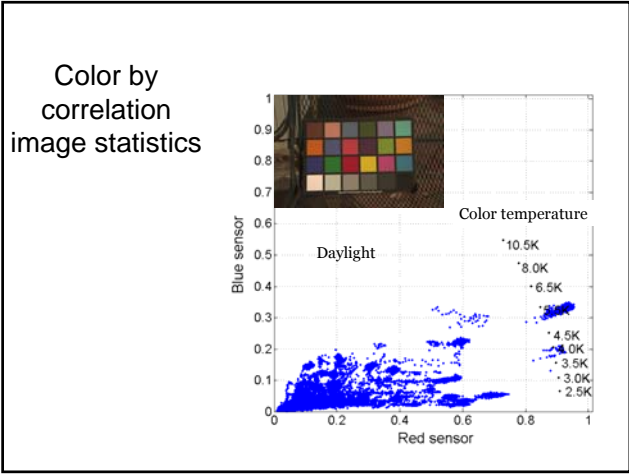
IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 23, NO. 11, NOVEMBER 2001 1209

### Color by Correlation: A Simple, Unifying Framework for Color Constancy

Graham D. Finlayson, Steven D. Hordley, and Paul M. Hubel

### Color by correlation image statistics





### Two-capture color balancing

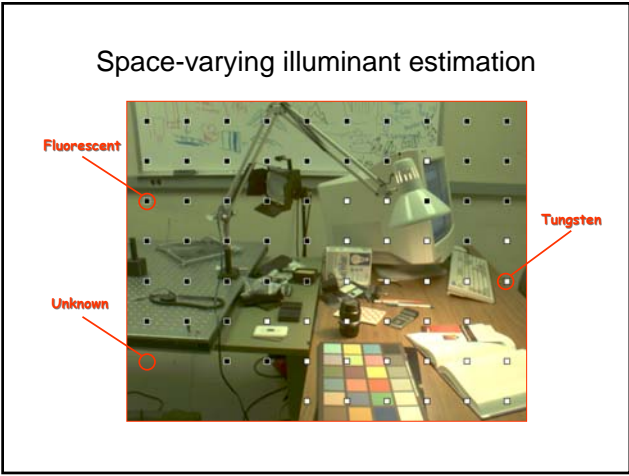
(DiCarlo et al., CIC, 2001)

**Multiple captures**

*Illuminant bracketing*

**Single image**

*Known illuminant image*



## Image Metrics and Preference

- Lots of things we can measure, but how do we know which are important?
- Metrics that incorporate properties of the human visual system
- In some cases, preference can be predicted by the visibility of annoying artifacts

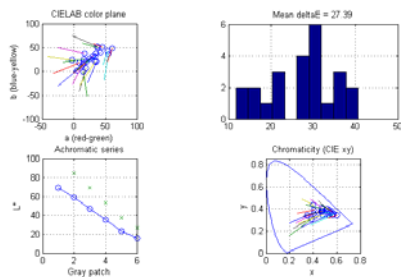
## Metrics that incorporate properties of the human visual system

- CIELAB
- S-CIELAB
- vSNR

### CIELAB $\Delta E$ Color Accuracy



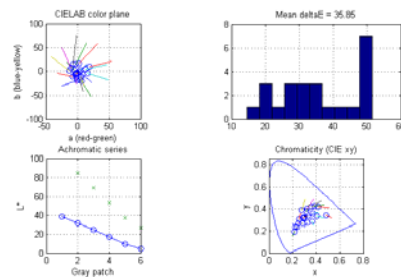
Not color-balanced



### CIELAB $\Delta E$ Color Accuracy



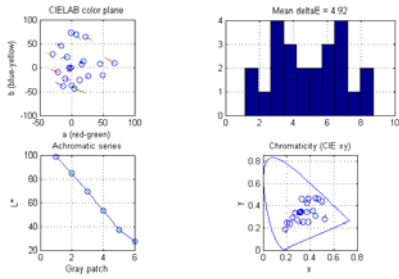
Color Balanced  
"Gray World"



## CIELAB $\Delta E$ Color Accuracy



Color Balanced  
Optimal Matrix



## Color Preference vs. Color Accuracy

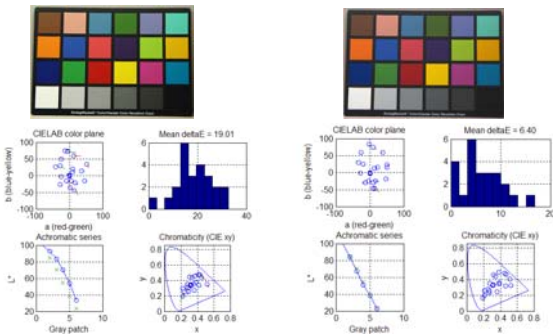
Which do you prefer?



## Color Preference vs Color Accuracy

Nikon Processed

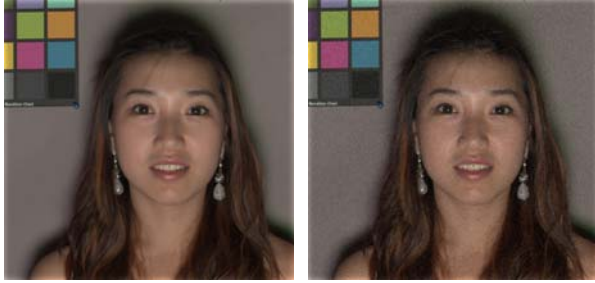
Optimized for Color Accuracy



## Preference Judgments

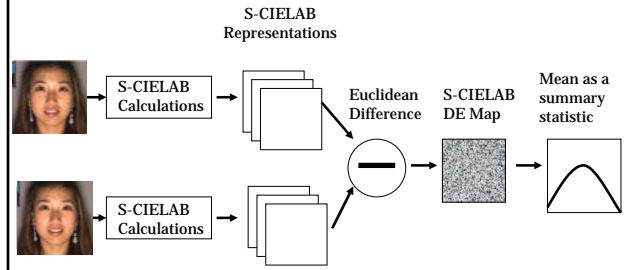
- **Discrimination:** I can see the difference between two images
- **Preference:** I like this image more than the other image

Which one do you prefer?

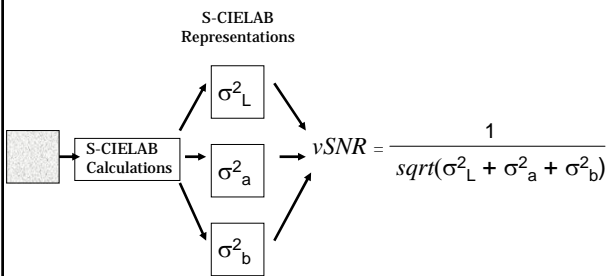


### S-CIELAB

Preference and Visibility of Uncorrelated Image Noise  
(e.g. read noise)

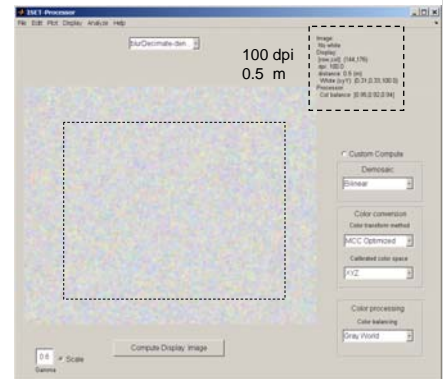


### Visible SNR (vSNR)



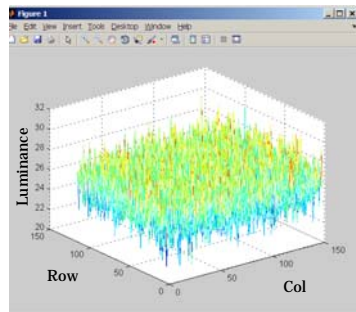
### vSNR: Overview

- Rendered uniform image
- Display and viewing distance are specified
- A region is selected for analysis



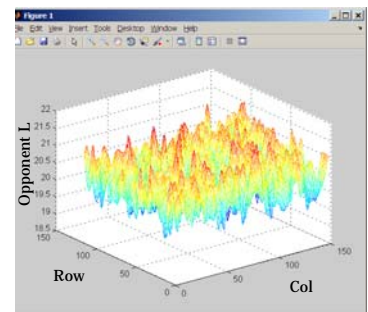
## vSNR: Uniform luminance

- The RGB data are expressed as radiance from the surface of the display
- The noise of the acquisition and rendering, plotted here in luminance (Y) units, is visible
- Is this noise visible?



## vSNR: Conversion to filtered Y

- Accounting for the viewing distance and pixel sampling, we apply the Spatial-CIELAB (S-CIELAB)
- The plot here shows the first opponent channel after the appropriate spatial blurring
- The blurring removes variance that the observer will not see

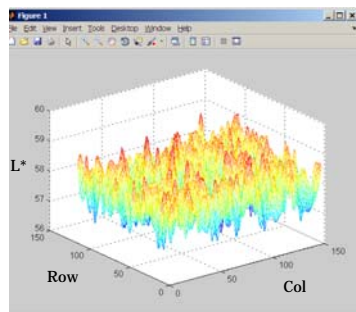


## vSNR: Conversion to L\*

- The opponent data are converted back to XYZ and then to CIELAB
- The L\* plot is shown here
- Similar operations are performed on the a\* and b\* terms and the inverse std. dev. of the LAB values is calculated (~ 1 in this case)

$$vSNR = \frac{1}{\text{sqrt}(\sigma_L^2 + \sigma_a^2 + \sigma_b^2)}$$

vSNR has units of 1/dE



## Correlated Image Noise Visual masking and orientation-specific effects

- S-CIELAB predicts the visibility of **uncorrelated image noise**
- S-CIELAB does not predict the visibility of **correlated image noise**
  - does not yet include masking and orientation selectivity
  - See Spatial Standard Observer (Watson and others)
- Example:
  - Bilinear demosaicking introduces uncorrelated image noise
  - Adaptive Laplacian demosaicking introduces correlated image noise



# Post-view

## Morning

- **Scene radiance (Brian, Joyce)**
  - Radiometry
  - Linear models
  - Scene database
- **Optics and irradiance (Peter)**
  - Measuring irradiance
  - Image formation
  - Optics modeling
- **Sensor model (Brian)**
  - CMOS circuit
  - Circuit modeling
- **System design metrics (Joyce)**
  - MTF, SNR
  - Resolution-sensitivity tradeoffs
  - Camera motion

## Afternoon

- **Sensor system components (Peter)**
  - Pixel optics
  - Anti-aliasing/IR filters
  - CFA, Prismatic optics
- **Human observer (Brian)**
  - Image formation and sampling
  - Wavelength encoding
  - Perceptual metrics
- **Image processing (Brian)**
  - Auto-exposure
  - Demosaicking
  - Color management
- **Metrics and preference (Joyce)**
  - Color accuracy
  - S-CIELAB, vSNR and preference