



GAMES 204



Computational Imaging

Lecture 05: Image Signal Processing (ISP)



Qilin Sun (孙启霖)

香港中文大学（深圳）

点昀技术 (Point Spread Technology)



Today's Topic

- Bayer Domain Processing
 - Dead Pixel Correction
 - Black Level Compensation
 - Denoising
 - 3As
 - Lens Shading Correction
 - Anti-aliasing
 - Demosaicing
- Color Correction
- Edge Enhancement
- False Color Suppression
- Brightness/Contrast Control
- Gamma and Tone Mapping



Assignment 01

- Realize a basic image signal pipeline
 - Bayer Domain Processing (50 pts)
 - Dead Pixel Correction (10pts)
 - Black Level Compensation (10pts)
 - Denoising (Optional, 20pts extra for RAW NLM)
 - Anti-aliasing (10pts)
 - Demosaicing (20pts)
 - Color Correction (Optional, 10pts extra)
 - Edge Enhancement (10pts)
 - False Color Supression (10pts)
 - Brightness/Contrast Control (10pts)
 - Gamma and Tone Mapping (20pts)
- The requirements, code framework and data will be released through **CES-alpha** system in the weekend
- Strict Deadline: 11.59pm, Aug 15
- Programming Language: Python

Slide credits

Many of these slides were adapted from:

OpenISP

Wolfgang Heidrich(KAUST)

Ren Ng(Uc Bakery)

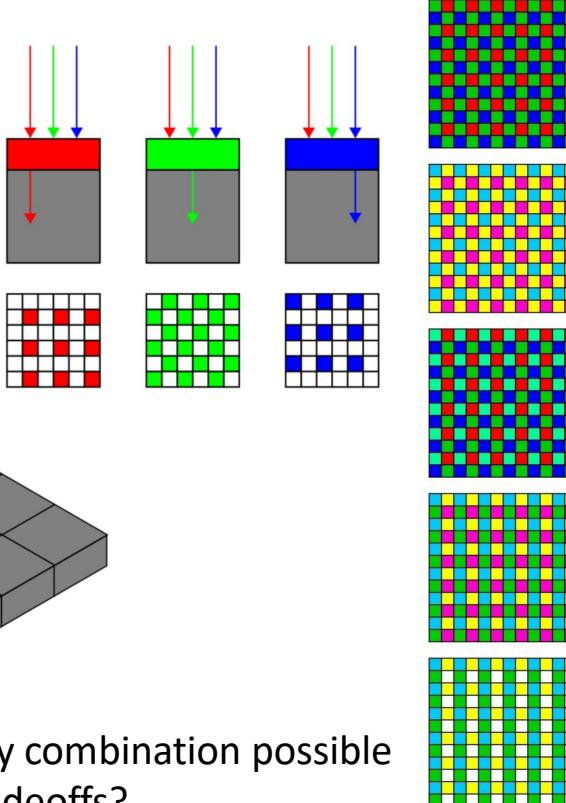
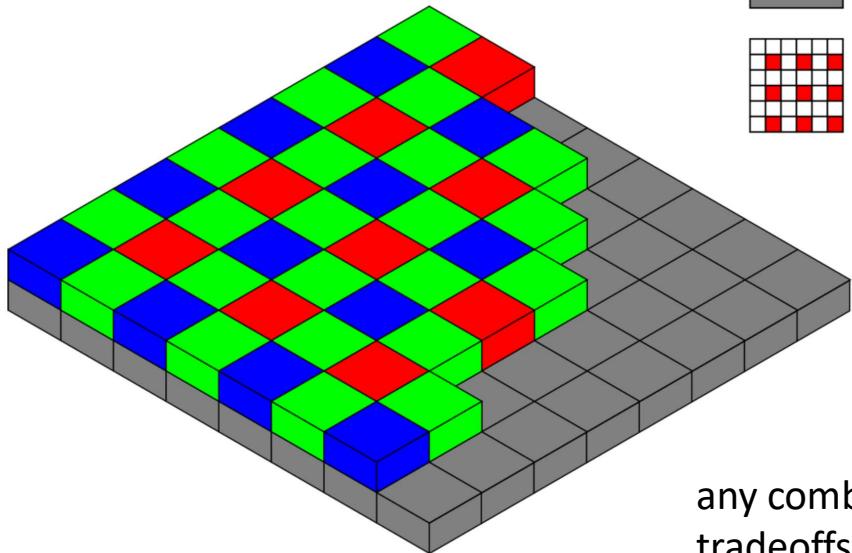
Marc Levoy (Stanford)

Kris Kitani, Ioannis Gkioulekas (15-463, 15-663, 15-862).



Review Color Filter Arrays

Bayer pattern



any combination possible
tradeoffs?

wikipedia



Image Formation

- High-dimensional integration over angle, wavelength, time

Plenoptic Function

$$i(x) \approx \iiint_{\Omega_{\theta,\lambda,t}} l(x, \theta, \lambda, t) d\theta d\lambda dt$$

- Plenoptic Function of light field

$$L(x, y, z, \theta, \Phi, \lambda, t, P<DoP, AoP>)$$

plenoptic function: [Adelson 1991]



Recall: Photons to RAW Image

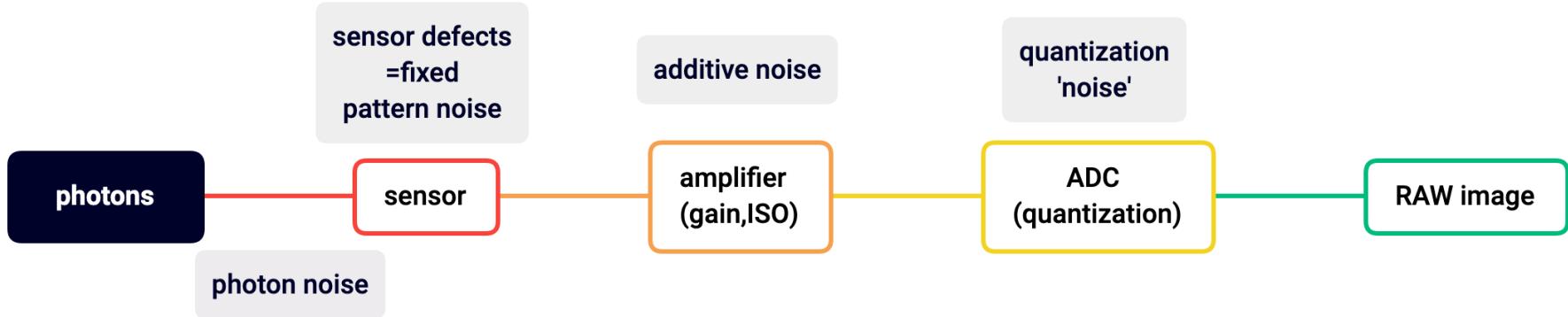




Image Processing Target

RAW image
(dcraw -D)



JPEG image

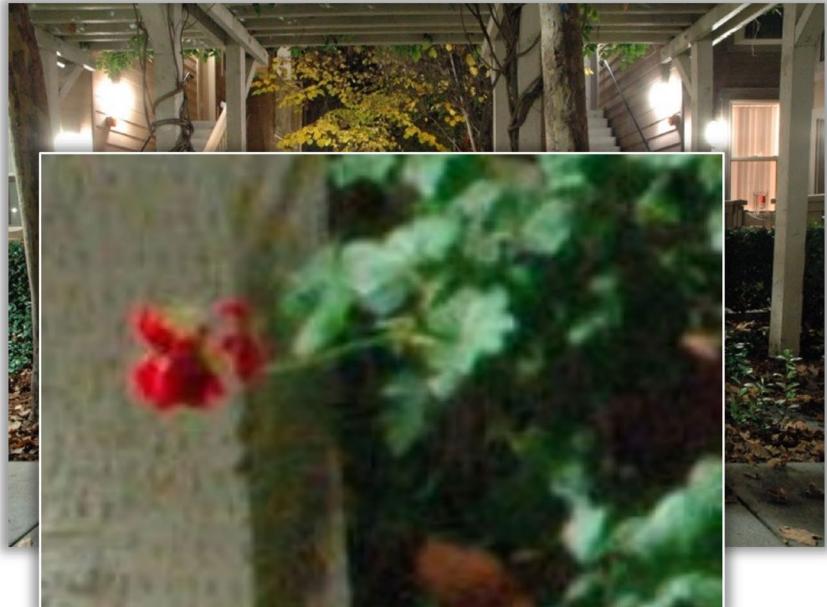




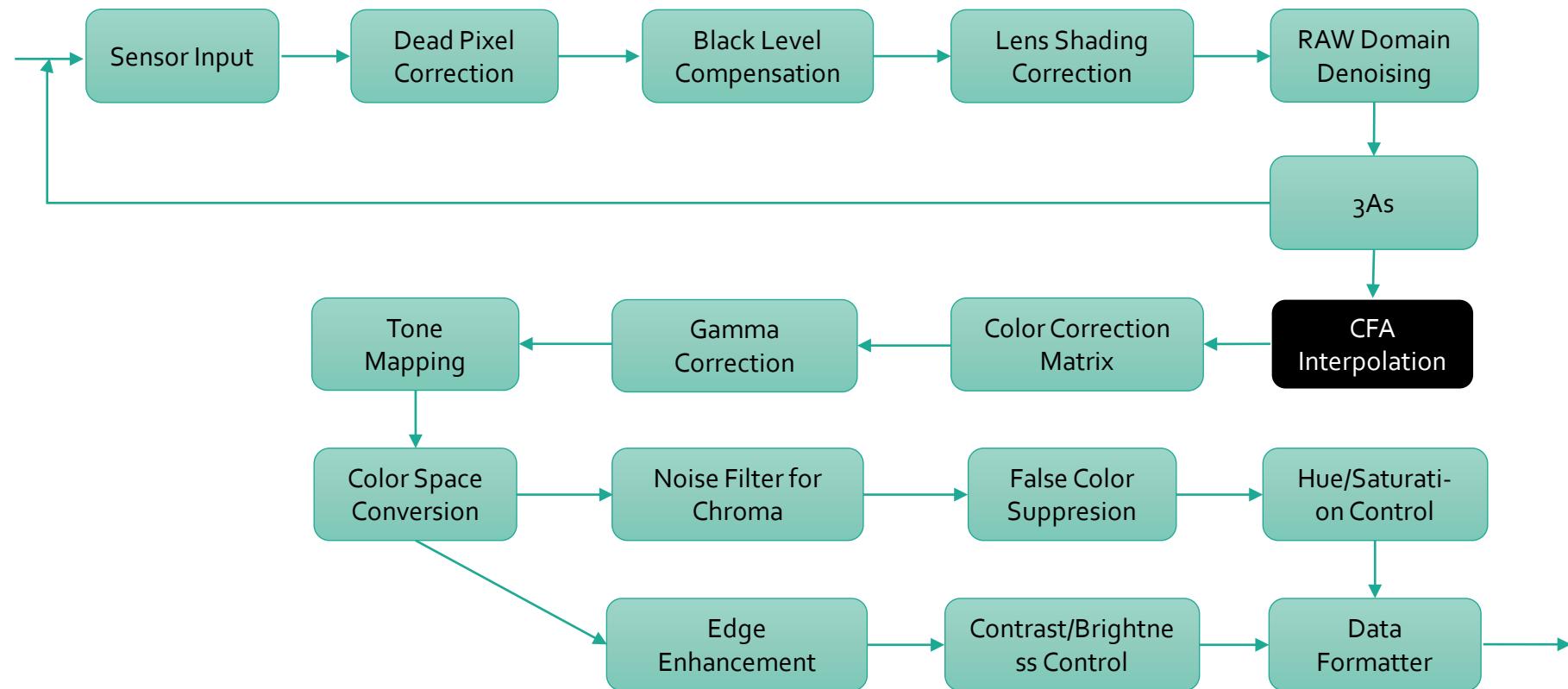
Image Processing Pipeline

- Darkcount Correction
- Hot/Dead Pixels Removal
- Vignetting Correction
- White Balance
- Auto Exposure
- Distortion Correction
- Anti-alising
- Demosaicking
- Edge Enhancement
- Color Correction
- Tone Mapping
- Compression





Image Processing Pipeline





Exif Meta Data (exchangeable)

Filename - night_nikon.JPG
Make - NIKON CORPORATION
Model - NIKON D70S
Orientation - Top left
XResolution - 300
YResolution - 300
ResolutionUnit - Inch
Software - Ver.1.00
DateTimeOriginal - 2005:09:01 12:16:43
VColorPositioning - Co-Sited
ExifOffset - 216
ExposureTime - 10 seconds
FNumber - 13.00
ExposureProgram - Manual control
ExifVersion - 0221
DateTimeOriginal - 2005:09:01 12:16:43
DateTimeDigitized - 2005:09:01 12:16:43
ComponentsConfiguration - YCbCr
CompressedBitsPerPixel - 1 (bits/pixel)
ExposureBiasValue - 0.50
MaxApertureValue - F 3.48
MeteringMode - Center weighted average
LightSource - Auto
Flash - Not fired
FocalLength - 18.00 mm
UserComment - (c) Gordon Wetzstein
SubsecTime - 00
SubsecTimeOriginal - 00
SubsecTimeDigitized - 00
FlashPixVersion - 0100
ColorSpace - sRGB
Thumbnail:
Compression - 6 (JPG)
XResolution - 300
YResolution - 300
ResolutionUnit - Inch
JpegIFOffset - 29368
JpegIFByteCount - 8393
YCbCrPositioning - Co-Sited

Filename - night_nikon.JPG
Make - NIKON CORPORATION

Model - NIKON D70S
Orientation - Top left
XResolution - 300
YResolution - 300
ResolutionUnit - Inch
Software - Ver.1.00
DateTime - 2005:09:01 12:16:43
YCbCrPositioning - Co-Sited

ExifOffset - 216
ExposureTime - 10 seconds

FNumber - 13.00

ExposureProgram - Manual control

ExifVersion - 0221

DateTimeOriginal - 2005:09:01 12:16:43

DateTimeDigitized - 2005:09:01 12:16:43

ComponentsConfiguration - YCbCr

CompressedBitsPerPixel - 1 (bits/pixel)

ExposureBiasValue - 0.50

MaxApertureValue - F 3.48

MeteringMode - Center weighted average

LightSource - Auto

Flash - Not fired

FocalLength - 18.00 mm

UserComment - (c) Gordon Wetzstein

SubsecTime - 00

SubsecTimeOriginal - 00

SubsecTimeDigitized - 00

FlashPixVersion - 0100

ColorSpace - sRGB

Maker Note (Vendor): -
Data version - 0210 (808595760)

ISO Setting - 1600

Image Quality - BASIC

White Balance - AUTO

Image Sharpening - MED.L

Focus Mode - MANUAL

Flash Setting - NORMAL

Flash Mode -

White Balance Adjustment - 0

Exposure Adjustment - 1.7

Thumbnail IFD offset - 1430

Flash Compensation - 67072

ISO 2 - 1600

Tone Compensation - AUTO

Lens type - AF-D G

Lens - 618

Exposure Mode - Auto

AF Focus Position - Center

Bracketing - 131072

Color Mode - MODE1a

Light Type - NORMAL

Hue Adjustment - 0

Noise Reduction - FPNR

Total pictures - 22346

Optimization - PORTRAIT

Thumbnail -

Compression - 6 (JPG)

Interlaced -

YResolution - 300

ResolutionUnit - Inch

JpegIFOffset - 29368

JpegIFByteCount - 8393

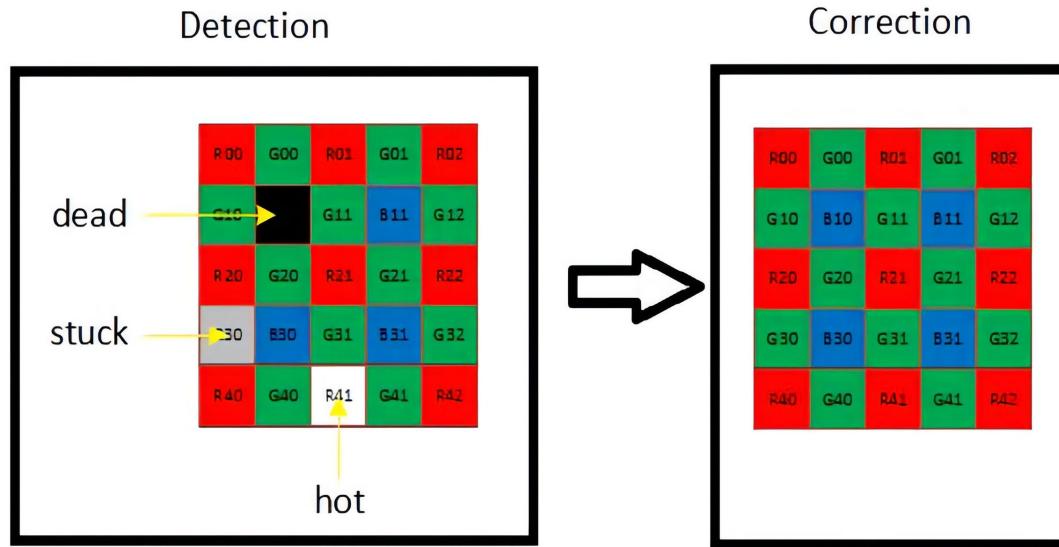
YCbCrPositioning - Co-Sited

Bayer Domain Processing



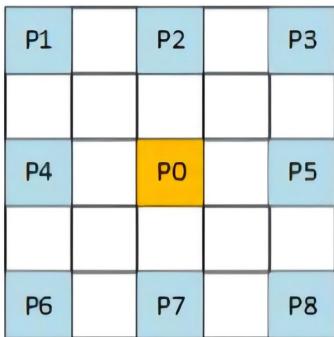
Dead Pixel Correction

- Source: sensor manufacturing
- Types of defective pixels:
 - dead (always low)
 - hot (always high)
 - stuck (always a value between low and high)





Dead Pixel Correction



- The differences between center pixel and neighbored 8 pixels are calculated.

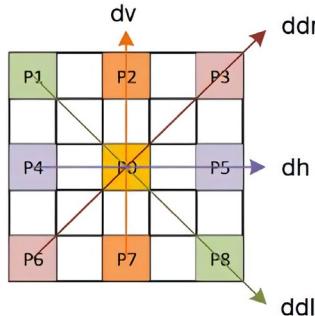
$$diff(x) = abs(P_{ij} - P_{mn})$$

$$P_{ij} = \begin{cases} dead & \text{if } \&(diff(x) > \text{thres}) \text{ where } x \text{ is 0-8} \\ not\ dead & \text{others} \end{cases}$$



Dead Pixel Correction

- Interpolate with gradient-based filter



- Get directions gradient

$$dv = |2P_0 - P_2 - P_7|$$

$$dh = |2P_0 - P_4 - P_5|$$

$$ddr = |2P_0 - P_1 - P_8|$$

$$ddl = |2P_0 - P_3 - P_6|$$

- The output

$$P_{ij} = \begin{cases} (P_{i,j-2} + P_{i,j+2})/2 & \text{if } \min(dv,dh,ddr,ddl) == dv \\ (P_{i-2,j} + P_{i+2,j})/2 & \text{if } \min(dv,dh,ddr,ddl) == dh \\ (P_{i-2,j-2} + P_{i+2,j+2})/2 & \text{if } \min(dv,dh,ddr,ddl) == ddl \\ (P_{i+2,j-2} + P_{i-2,j+2})/2 & \text{others} \end{cases}$$



Black Level Compensation

- Source: circuit design (not pure dark)
- Correction:

$$R' = R + R_{offset}$$

$$Gr' = Gr + Gr_{offset} + \alpha R$$

$$Gb' = Gb + Gb_{offset} + \beta B$$

$$B' = B + B_{offset}$$

where R_{offset} , Gr_{offset} , Gb_{offset} , B_{offset} , α , β are configurable parameters.



Anti-aliasing

Source: Sampling high freq image with low ratio



No anti-aliasing filtered



Anti-aliasing filtered



Lens Shading Correction

Source: Irradiance is proportional to

- Projected area of aperture as seen from pixel
- Projected area of pixel as seen from aperture
- Distance² from aperture to pixel

Combining all these: light drops as $\cos^4 \theta$

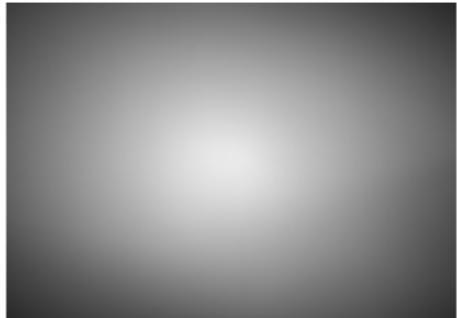
Types of Lens Shading (Vignetting):

- Luma shading
- Color shading.

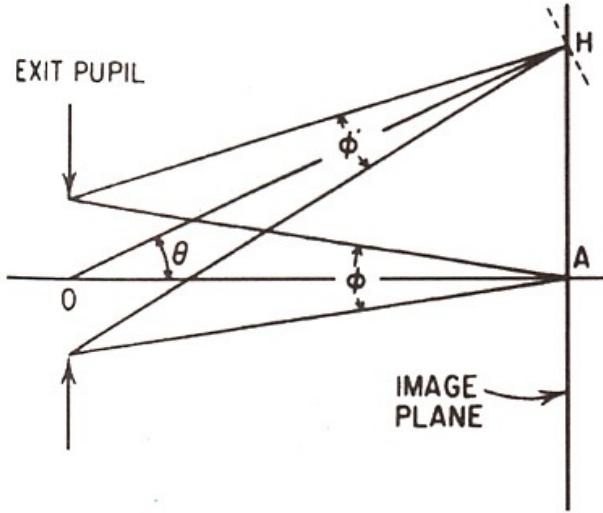
Color shading



Luma shading



Ideal





Lens Shading Correction

Multiply a gain on each pixel at different locations

- Take a image with uniform light condition. And divide the image into $m \times n$ blocks. Four points of each block have a correction coefficient. Store the coefficients into look-up table.
- Based on the pixel location, calculating the pixel falls into which block. Then get four coefficient of the block from LUT. (Or use surface fitting)
- Calculating the correction gain of the pixel by interpolation.
- Multiplying the correction gain by the pixel.



Noise Reduction

Non-local Means

- Most image details occur repeatedly
- Each color indicates a group of squares in the image which are almost indistinguishable
- Image self-similarity can be used to eliminate noise
 - It suffices to average the squares which resemble each other

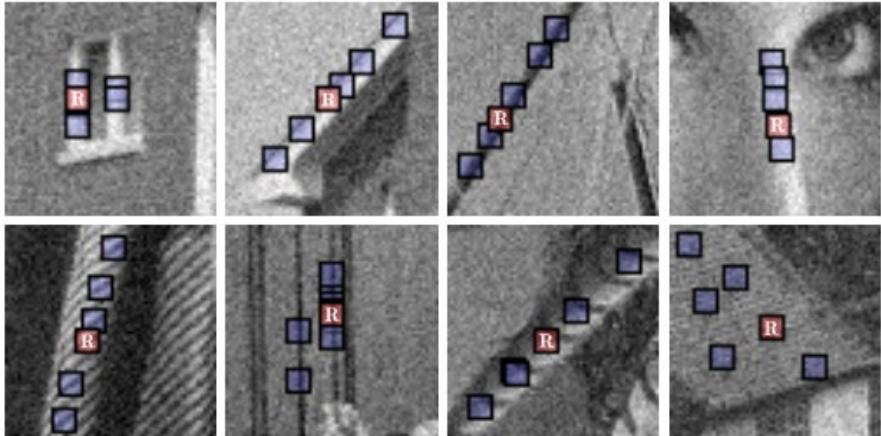


Image and movie denoising by nonlocal means
Buades, Coll, Morel, IJCV 2006

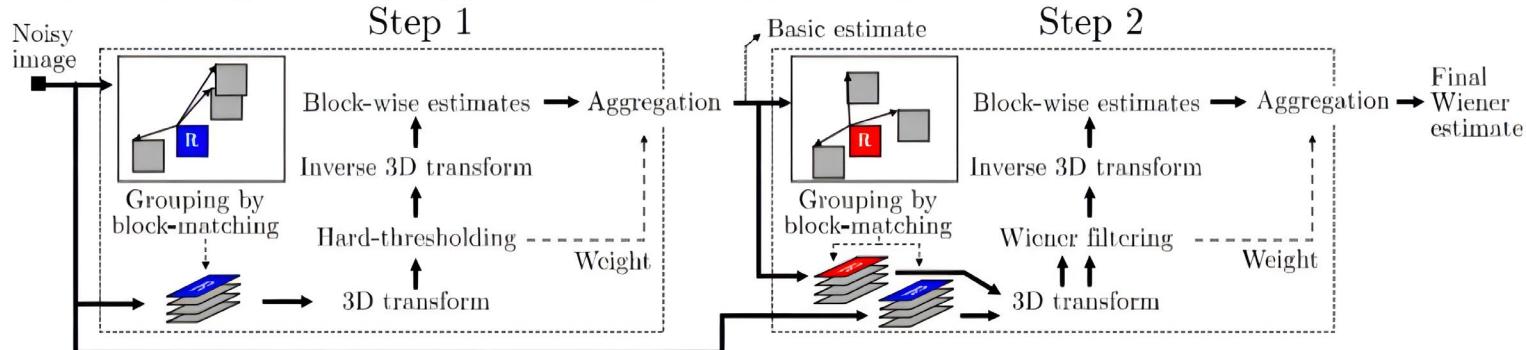


Noise Reduction

BM₃D (Block Matching 3D)



2022.07, Point Spread Technology using deep optimized ASIC realizes RAW BM₃D for 4K video in **Real Time!**



3As

Auto-exposure
Auto-focus
Auto-white-balance



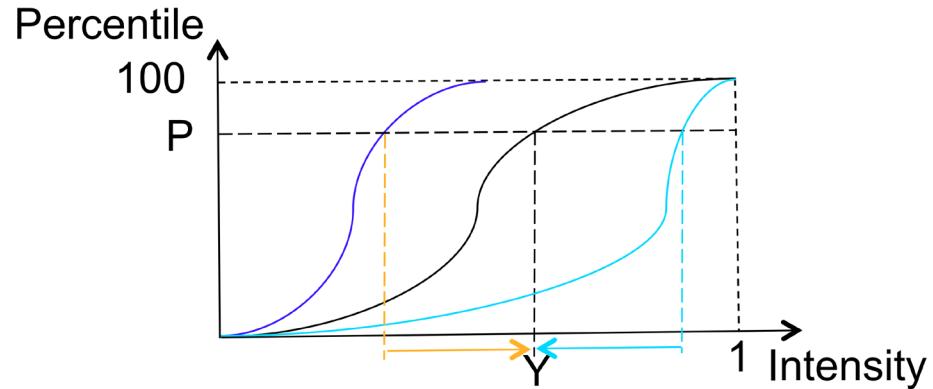
Auto-exposure

- Goal: well-exposed image (not a very well defined goal!)
- Possible parameters to adjust
 - Exposure time
 - Longer exposure time leads to brighter image, but also motion blur a
 - Aperture (f-number)
 - Larger aperture (smaller f-number) lets more light in causing the image to be brighter, also makes depth of field shallower
 - Phone cameras often have fixed aperture
- Analog and digital gain
 - Higher gain makes image brighter but amplifies noise as well
- ND filters on some cameras



Exposure metering

- Cumulative Density Function of image intensity values
 - P percent of image pixels have an intensity





Exposure metering examples

- Adjustment examples
 - $P = 0.995, Y = 0.9$
 - max 0.5% of pixels are saturated (highlights)
 - $P = 0.1, Y = 0.1$
 - max 10% of pixels are under-exposed (shadows)
 - Auto-exposure somewhere in between, e.g., $P = 0.9, Y = 0.4$



Highlights



Auto-exposure

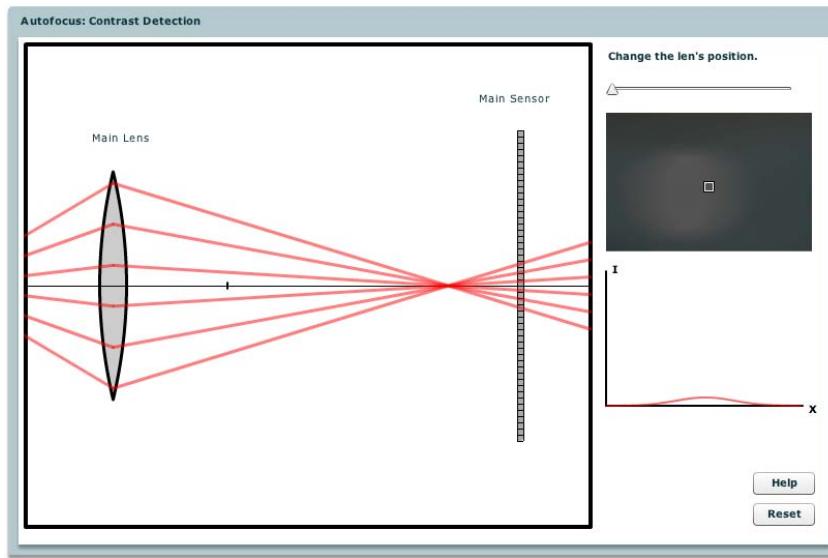


Shadows



Auto Focus

- Passive autofocus: contrast detection

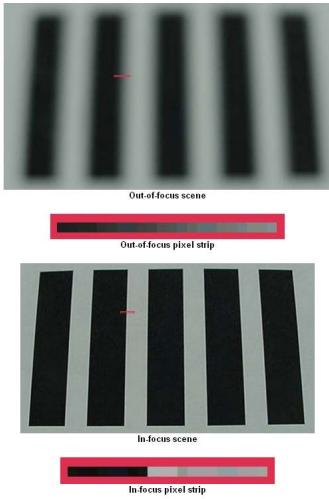


- An image sensor will see an object as contrasty if it's in focus, or of low contrast if it's not move the lens until the image falling on the sensor is contrasty compute contrasty-ness using local gradient of pixel values.

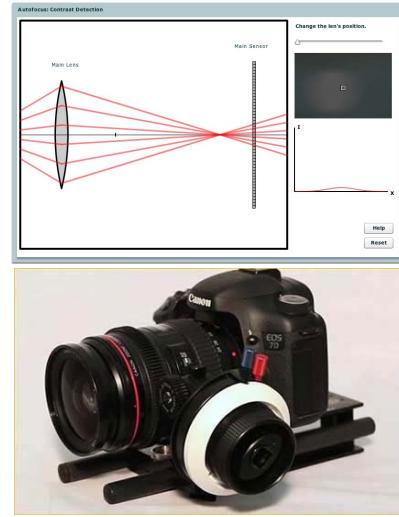


Auto Focus

Passive autofocus: contrast detection



cinema lenses
do not autofocus



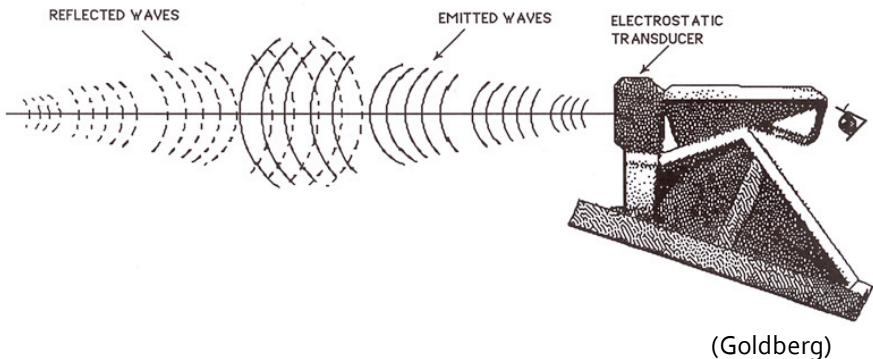
- Requires repeated measurements as lens moves, measurements are captured using the main sensor
 - Equivalent to depth-from-focus in computer vision
- Slow, requires hunting, suffers from overshooting
 - It's ok if still cameras overshoot, but video cameras shouldn't



Auto Focus Modes

- AI servo (Canon) / Continuous servo (Nikon)
 - predictive tracking so focus doesn't lag axially moving objects
 - continues as long as shutter is pressed halfway
- Focusing versus metering
 - autofocus first, then meter on those points "trap focus"
 - trigger a shot if an object comes into focus (Nikon)
- Depth of field focusing
 - find closest and furthest object; set focus and N accordingly
- Overriding autofocus
 - manually triggered autofocus (AF-ON in Canon)
- All autofocus methods fail if object is textureless!

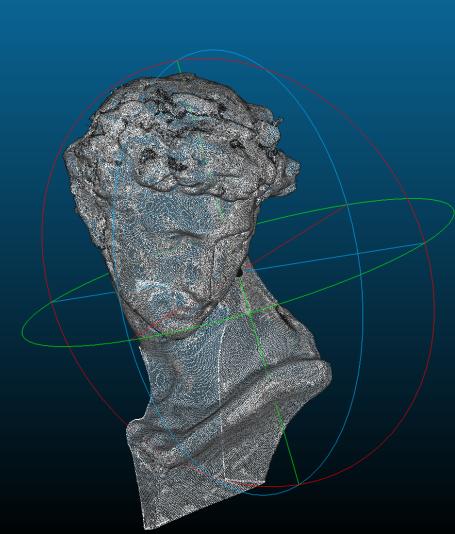
Active Auto Focus: Time-of-Flight



- SONAR = Sound Navigation and Ranging
- Polaroid system used ultrasound (50KHz)
 - well outside human hearing (20Hz - 20KHz)
- Limited range, stopped by glass
- Hardware salvaged and re-used in amateur robotics



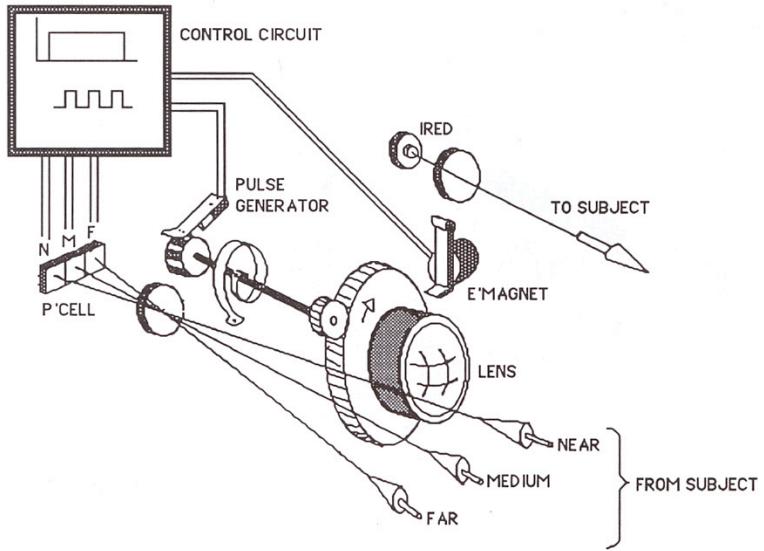
Active Auto Focus: Time-of-Flight



- LIDAR = Light Detection and Ranging
- PointSpread Tech OKulo uses 940nm infrared light
- Light travels 1 foot per nanosecond, so accuracy requires fast circuitry ($\pm 1\text{mm}$ is achieved)
- Work in bright/dark environment



Active Auto Focus: Triangulation



The technology depicted here is not used on any current consumer camera. Don't confuse it with the LED that turns on to illuminate a dark scene and thereby assist many cameras with phase or contrast based passive autofocus. That LED is sometimes called the "autofocus assist light".

- Infrared (IR) LED flash reflects from subject
- Angle of returned reflection depends on distance
- Fails on dark or shiny objects



White Balance Gain Control

Target: removing unrealistic color casts





Auto White Balance Gain Control

- Estimate the color temperature first

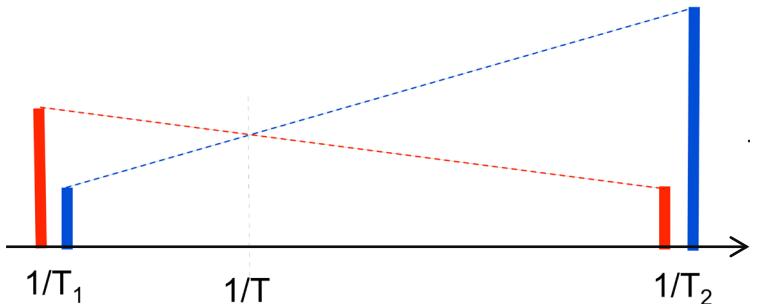
$$\begin{bmatrix} R' \\ Gr' \\ Gb' \\ B' \end{bmatrix} = \begin{bmatrix} R & 0 & 0 & 0 \\ 0 & Gr & 0 & 0 \\ 0 & 0 & Gb & 0 \\ 0 & 0 & 0 & B \end{bmatrix} * \begin{bmatrix} \frac{R_{avg}}{G_{avg}} * Gain \\ Gain \\ Gain \\ \frac{B_{avg}}{G_{avg}} * Gain \end{bmatrix}$$

- Where R_{avg} , G_{avg} , B_{avg} is the average of different channels on bayer domain.
- Gain applied to all channels to keep the image brightness constant.

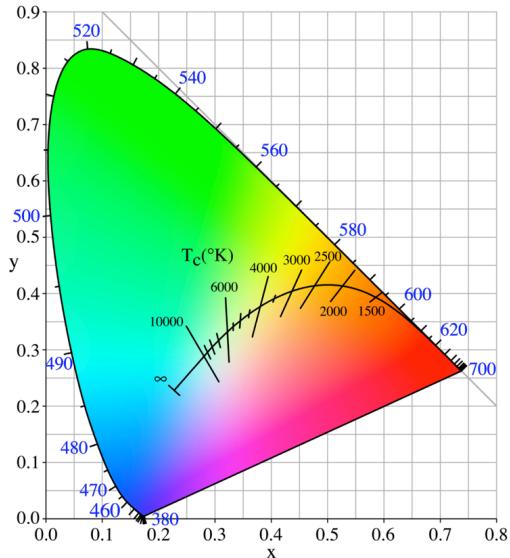


Estimating the Color Temperature

- Use scene mode
- Use gray world assumption ($R = G = B$) in sRGB space
 - really, just $R = B$, ignore G
- Estimate color temperature in a given image
 - apply pre-computed matrix to get sRGB for T_1 and T_2
 - calculate the average values R, B
 - solve α , use to interpolate matrices (or $1/T$)



$$R = (1 - \alpha)R_1 + \alpha R_2, B = (1 - \alpha)B_1 + \alpha B_2$$



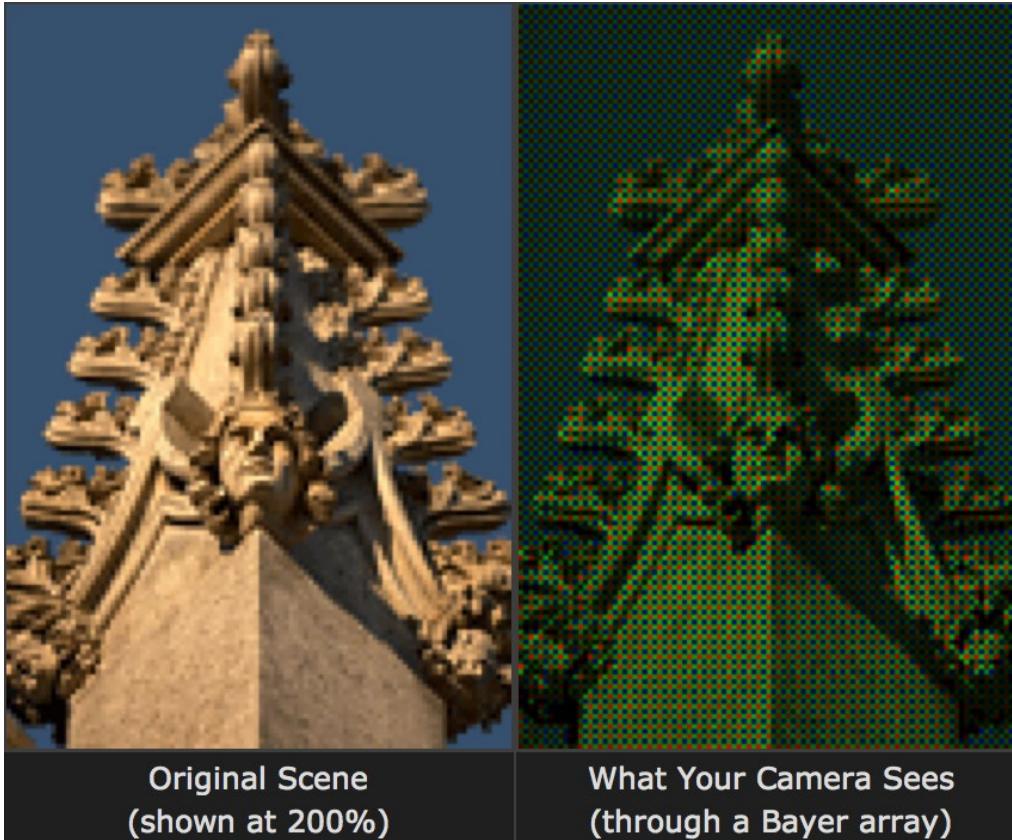
$$\frac{1}{T} = (1 - \alpha) \frac{1}{T_1} + \alpha \frac{1}{T_2}$$

Demosacing: CFA Interpolation



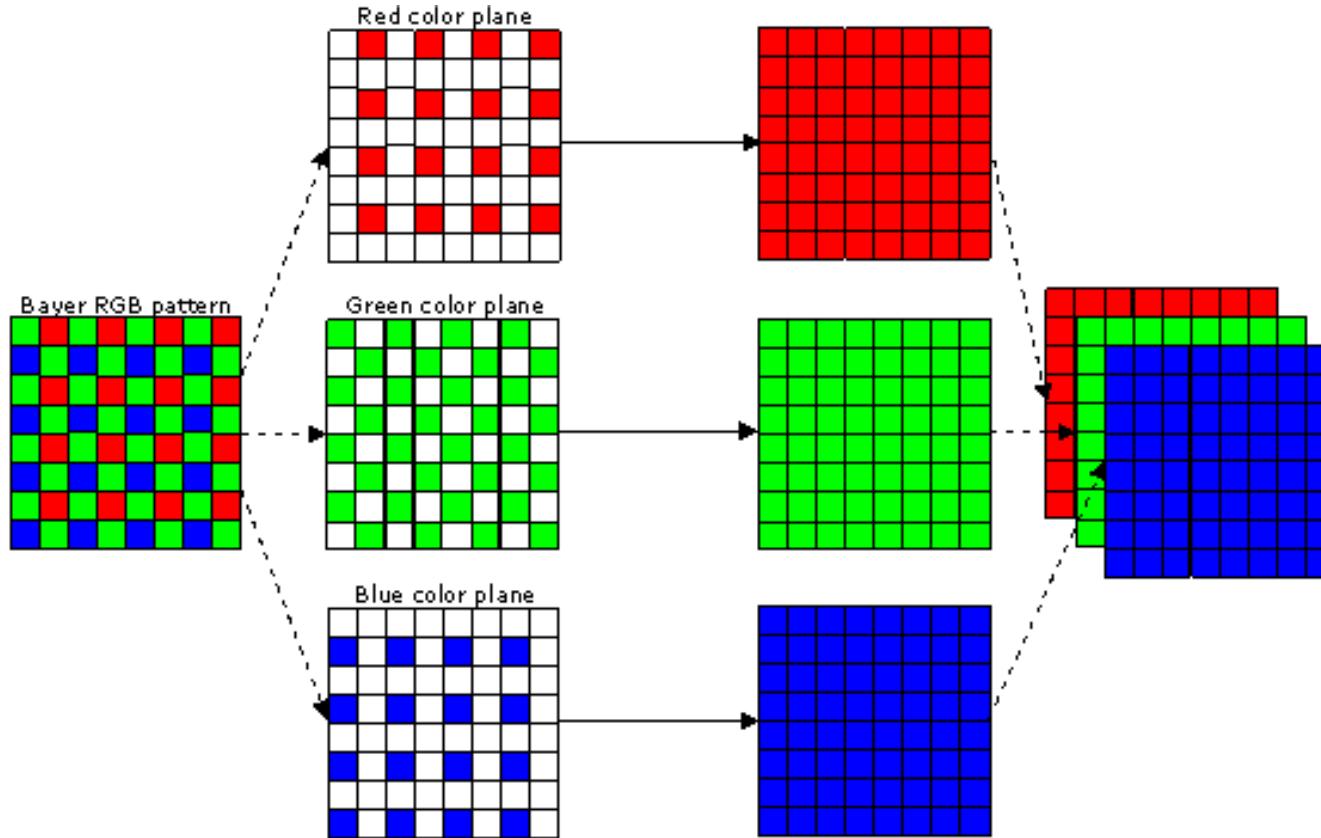
Demosacing: CFA Interpolation

Source: sensor color filter array





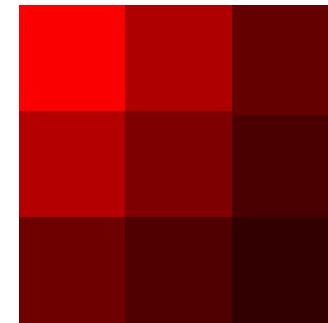
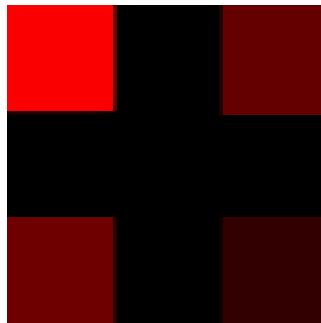
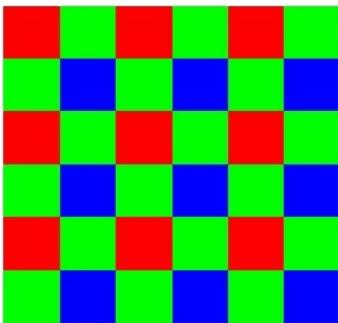
Demosacing: CFA Interpolation



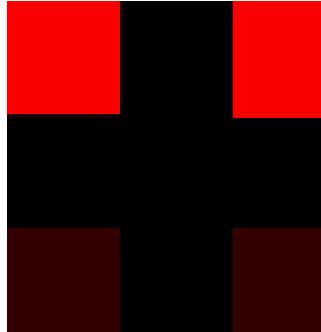


Demosacing: Bilinear Interpolation

- Advantage: Easy to implement



- Disadvantage: Fails at sharp edges





Demosacing: Take Edges into Account

- Use bilateral filtering
 - Avoid interpolating across edges

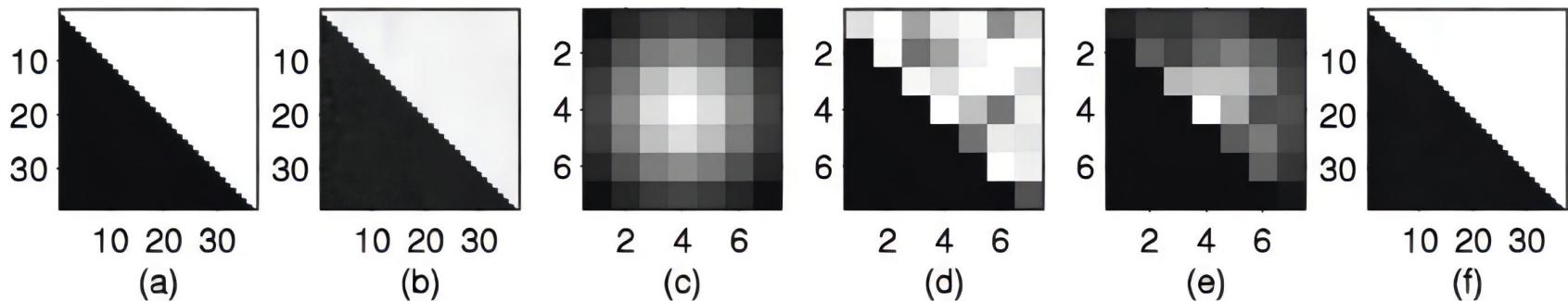


Fig. 3 Bilateral filtering: (a) original image, (b) image corrupted by Gaussian noise, (c) 7×7 blur kernel, (d) 7×7 similarity kernel at row=18, col=18, (e) 7×7 bilateral filter kernel, and (f) resulting image (denoised and sharpened).

ADAPTIVE DEMOSAICKING Ramanath, Snyder, JEI 2003

Demosacing: Take Edges into Account

- Predict edges and adjust
 - assumptions
 - luminance correlates with RGB
 - edges = luminance change
- When estimating G at R
 - if the R differs from bilinearly estimated R
 - →luminance changes
- Correct the bilinear estimate
 - by the difference between the estimate and real value

$$\hat{G}(i, j) = \hat{G}(i, j) + \alpha \Delta_{\hat{R}}(i, j)$$

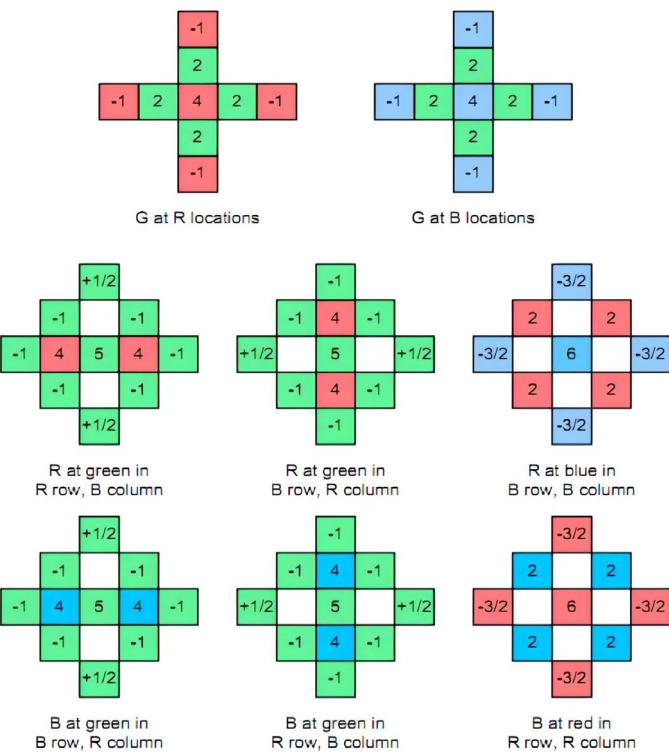


Figure 2. Filter coefficients for our proposed linear

HIGH-QUALITY LINEAR INTERPOLATION FOR
DEMOASICING OF BAYER-PATTERNED COLOR IMAGES
Malvar, He, Cutler, ICASSP 2004

Color Correction



Color Correction

Source:

- Spectral characteristics of the optics (lens, filters)
- Lighting source variations like daylight, fluorescent, or tungsten
- Characteristics of the color filters of the sensor

Color Correction Matrix (CCM) is dedicated to transfer the sensor RGB color space to sRGB/other color space.

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} RR & RG & RB \\ GR & GG & GB \\ BR & BG & BB \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} R_{offset} \\ G_{offset} \\ B_{offset} \end{bmatrix}$$

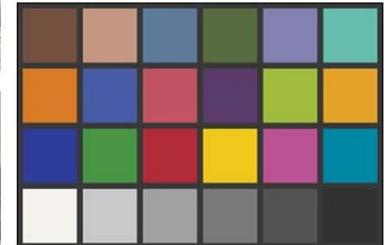
Before Color Correction



After Color Correction



Ideal Target Color Chart



Edge Enhancement



Edge Enhancement

Unsharp masking algorithm:

- **Gain.** This controls the extent to which contrast in the edge detected area is enhanced.
- **Radius or aperture.** This affects the size of the edges to be detected or enhanced, and the size of the area surrounding the edge that will be altered by the enhancement. A smaller radius will result in enhancement being applied only to sharper, finer edges, and the enhancement being confined to a smaller area around the edge.
- **Threshold.** Where available, this adjusts the sensitivity of the edge detection mechanism. A lower threshold results in more subtle boundaries of color being identified as edges. A threshold that is too low may result in some small parts of surface textures, film grain or noise being incorrectly identified as being an edge.



Edge Enhancement

- Extract the edge map from Y channel (YCrCb)

$$edge_filter = \begin{bmatrix} -1 & 0 & -1 & 0 & -1 \\ -1 & 0 & 8 & 0 & -1 \\ -1 & 0 & -1 & 0 & -1 \end{bmatrix} * \frac{1}{8}$$

- The filter is applied on 3x5 region.

n-2	n-1	n	n+1	n+2
m-1				
m				
m+1				

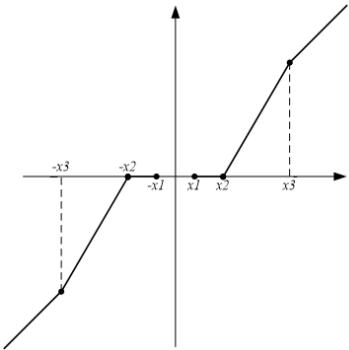
- The extracted edge map is given by

$$EM = \frac{1}{8} \sum_{i,j=-2}^{i,j=2} Y_{m+i,n+j} * edge_filter_{i,j}$$

Where i, j ranges from -2 to 2 with interval of 1.

Edge Enhancement

- The edge map (EM) is further modified through a lookup table (EMLUT) as



$$EMLUT(x) = \begin{cases} m_1 x & x \leq -x_3 \\ m_1 x_3 \times \left(\frac{x+x_2}{x_3-x_2}\right) & -x_3 < x \leq -x_2 \\ 0 & -x_2 < x \leq -x_1 \\ -m_2 x \text{ or } Avg - Y & -x_1 < x \leq x_1 \\ 0 & x_1 < x \leq x_2 \\ m_1 x_3 \times \left(\frac{x-x_2}{x_3-x_2}\right) & x_2 < x \leq x_3 \\ m_1 x & x_3 \leq x \end{cases}$$

Where m_1, m_2 is the gain for different thresholds. When $-x_1 < x \leq x_1$ the pixel is noise not the edge.

- There are two methods to calculate the Y value:

$$Y = Y - m_2 x$$

$$Y = Avg$$



Edge Enhancement

- The **Avg** is calculated by the gradient of neighbor pixels

$$DDH = \text{abs}(2Y_{m,n} - Y_{m,n-1} - Y_{m,n+1})$$

$$DDV = \text{abs}(2Y_{m,n} - Y_{m-1,n} - Y_{m+1,n})$$

$$DDL = \text{abs}(2Y_{m,n} - Y_{m-1,n-1} - Y_{m+1,n+1})$$

$$DDR = \text{abs}(2Y_{m,n} - Y_{m-1,n+1} - Y_{m+1,n-1})$$

- Find the minimum of , then **Avg** takes the value of that direction.

$$AvgH = \frac{Y_{m,n} + Y_{m,n-1} + Y_{m,n+1}}{3}$$

$$AvgV = \frac{Y_{m,n} + Y_{m-1,n} + Y_{m+1,n}}{3}$$

$$AvgL = \frac{Y_{m,n} + Y_{m-1,n-1} + Y_{m+1,n+1}}{3}$$

$$AvgR = \frac{Y_{m,n} + Y_{m-1,n+1} + Y_{m+1,n-1}}{3}$$

- The EMULUT(x) is further clipped before adding to the original Y channel

$$Y' = Y + \text{clip}(EMULT(x))$$

- To avoid over-shoot on edges, the Y' is further clipped for output.

$$Y' = \text{clip}(Y', Y_{max}, Y_{min})$$

False Color Suppression



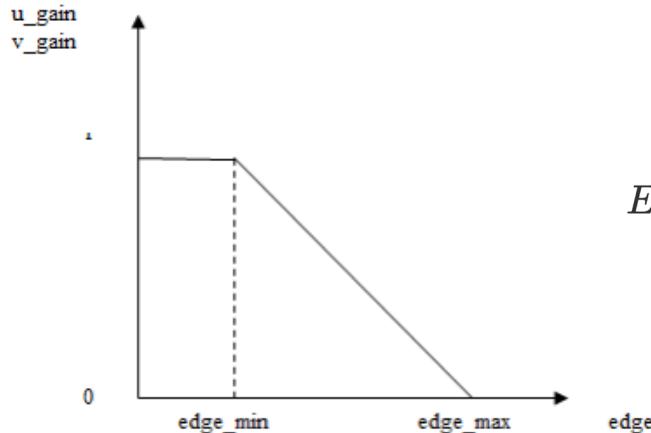
False Color Suppression

Source:

- Demosaicing phase where very fine details resolved
- Lens, optical filter, dynamic compression etc

Conducted on chroma (Cb, Cr) channel. Luma (Y) is not influenced

- The edge map acquired from Edge Enhancement is taken by absolute value: $EM' = \text{abs}(EM)$. Then clipping the absolute edge map into $\{\text{edge_min}, \text{edge_max}\}$.



$$EM'' = \text{clip}(\text{edge_min}, \text{edge_max})$$



False Color Suppression

- The chroma gain is calculated by

$$chroma_gain = K_edge * (edge_max - EM'')$$

Where

$$K_edge = \frac{65536}{(edge_max - edge_min)}.$$

- Do false color suppression on chroma channel when edge is larger than $edge_min$.

$$(Cb', Cr') = \begin{cases} (Cb, Cr) & edge \leq edge_min \\ \frac{chroma_gain * (EM'' - 128))}{65536} + 128 & edge_min < edge \leq edge_max \\ (0, 0) & edge_max < edge \end{cases}$$

Brightness/Contrast Control



Brightness/Contrast Control

- Brightness and contrast are both applied on luma channel.

$$Y' = Y + \text{brightness}$$

Where *brightness* ranges from [-128, 127]

$$Y' = Y + (Y - 127) * \text{contrast}$$

Where *contrast* is the adjustment ratio, ranges from [0, 1.992x].

Gamma and Tone Mapping



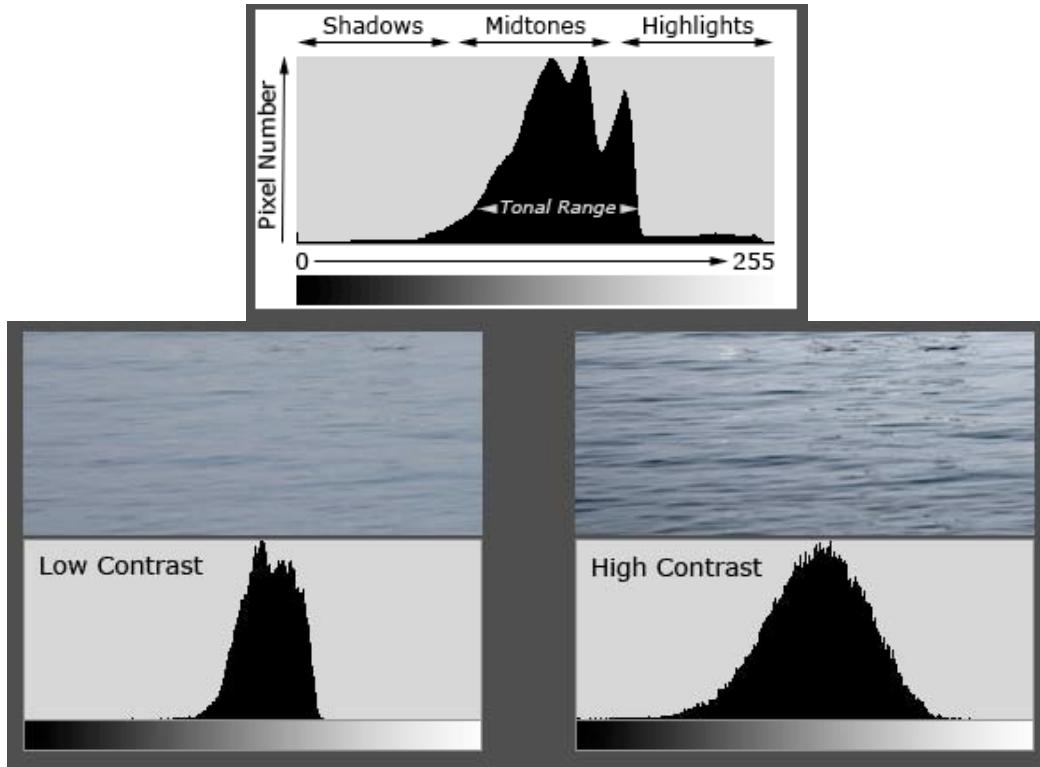
Gamma Correction

- Goal: accurately reproduce relative scene luminances on a display screen
 - Absolute luminance is impossible to reproduce
 - Humans are sensitive to relative luminance anyway
- In some workflows, pixel value is made proportional to scene luminance, in other systems to perceived brightness
- In CRTs luminance was proportional to voltage γ with $\gamma \approx 2.5$, so TV cameras were designed to output voltage \propto scene luminance $^{1/\gamma}$
- Pixel value \propto luminance $^{1/2.5}$ is roughly perceptually uniform, so it's a good space for quantization, for example in JPEG files



Contrast Correction (a.k.a. tone mapping)

- Manual editing
 - Capture image in RAW mode, then fiddle with histogram in Photoshop, dcraw, Canon Digital Photo Professional, etc.
 - To expand contrast, apply an S-curve to pixel values



(cambridgeincoulour.com)

Contrast Correction (a.k.a. tone mapping)

- Manual editing
 - Capture image in RAW mode, then fiddle with histogram in Photoshop, dcraw, Canon Digital Photo Professional, etc.
 - To expand contrast, apply an S-curve to pixel values
- Gamma transform (in addition to RAW→JPEG gamma)
 - Output = input γ (for $0 \leq li \leq 1$)
 - Simple but crude
- Global versus local transformations



original



$\gamma = 0.5$

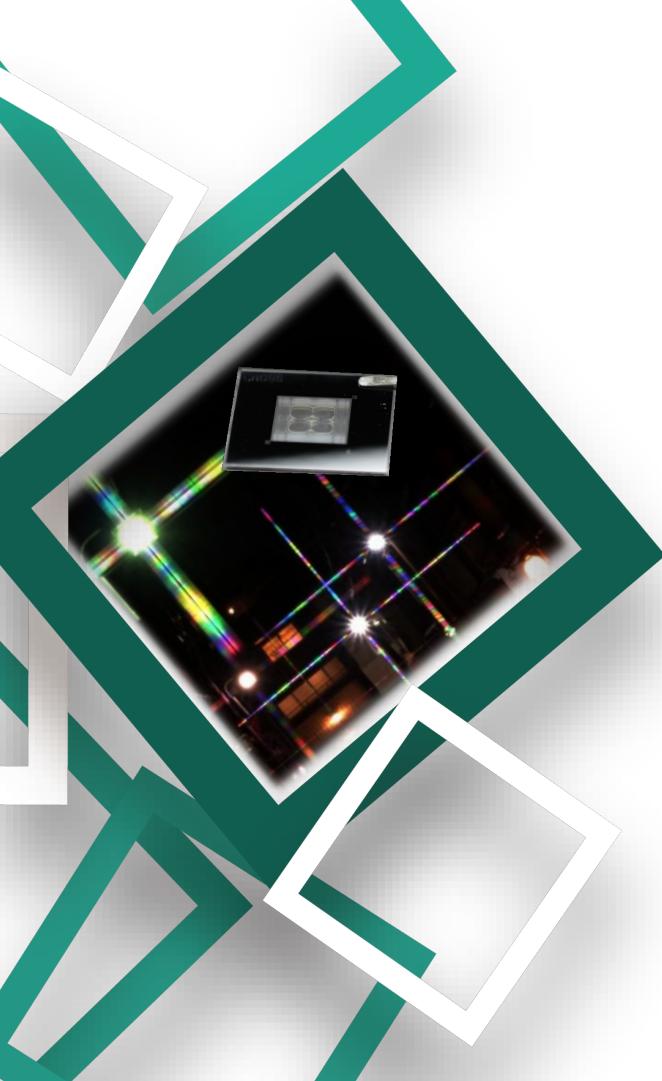


$\gamma = 2.0$



Today's Topic

- Bayer Domain Processing
 - Dead Pixel Correction
 - Black Level Compensation
 - Denoising
 - 3As
 - Lens Shading Correction
 - Anti-aliasing
 - Demosaicing
- Color Correction
- Edge Enhancement
- False Color Suppression
- Brightness/Contrast Control
- Gamma and Tone Mapping



GAMES 204



Thank You!



Qilin Sun (孙启霖)

香港中文大学（深圳）
点昀技术（Point Spread Technology）